Heliyon 10 (2024) e27388

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification and analysis of prognostic metabolic characteristics in colon adenocarcinoma

Yang Yang [a,b,1], Xinyu Yang [a,c,1], Shiqi Ren [d,1], Yang Cao [e], Ziheng Wang [f,g,*], Zhouyang Cheng [h,**]

[a] Department of Trauma Center, Affiliated Hospital of Nantong University, Medicial School of Nantong University, No.20 Xisi Road, Chongchuan District, 226001, Nantong City, Jiangsu Province, China
[b] Department of Emergency Medicine, Affiliated Hospital of Nantong University, No.20 Xisi Road, Chongchuan District, Nantong City, Jiangsu Province, 226001, China
[c] Medical School of Nantong University, Qixiu Road, 226001, Nantong City, Jiangsu Province, China
[d] Department of Hand Surgery, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong, 226001, China
[e] Department of Operation Room, Affiliated Hospital of Nantong University, No.20 Xisi Road, Chongchuan District, 226001, Nantong City, Jiangsu Province, China
[f] MOE Frontier Science Centre for Precision Oncology, University of Macau, Macau SAR, China
[g] Department of Clinical-Biobank, Affiliated Hospital of Nantong University, No.20 Xisi Road, Chongchuan District, 226001, Nantong City, Jiangsu Province, China
[h] Department of General Surgery, Affiliated Hospital of Nantong University, No.20 Xisi Road, Chongchuan District, 226001, Nantong City, Jiangsu Province, China

ARTICLE INFO

Keywords:
Colon
Oncology
Proliferator-activated receptors
Biomarker

ABSTRACT

Colon adenocarcinoma (COAD) is a highly lethal gastrointestinal malignancy. The five-year survival rate of metastatic colorectal cancer remains low, at 14 percent. Numerous publications have suggested a role for peroxisome proliferator-activated receptors (PPARs) in malignancy. Recent studies have shown that PPARs, as nuclear transcription factors, may serve as potential targets for the treatment of metabolic syndrome tumors and their associated complications. However, the molecular mechanism has not been thoroughly investigated. Hence, in order to enhance the prediction of personalized medicine for PPAR-associated modulators in malignancy treatment, a timely review becomes essential. Utilizing TCGA-COAD expression profile data and patient overall survival (OS) information, this study systematically conducted investigations to identify and develop Hub stem cell-related diagnostic and prognostic identification models, aiming to enhance the multi-gene markers for COAD. Utilizing the differential expression profiles of stem cell-related genes, an 11-gene (SLC27A4, CPT1C, CPT1B, CPT2, CYP4A11, FABP3, FABP7, AQP7, MMP1, ACOX1, ANGPTL4) diagnostic and prognostic model was developed. This model demonstrated precise diagnostic and prognostic capabilities and holds the potential to characterize the clinicopathologic features of COAD. Univariate and multivariate Cox proportional hazards regression analyses were conducted to ascertain the independent factors influencing OS outcomes in COAD. The results revealed that CPT1B, SLC27A4, and FABP3 were identified as independent risk prognostic factors for OS in COAD, whereas ACOX1 and CPT2 served as independent protective prognostic factors. The hub genes associated with PPARs were

* Corresponding author. Affliated Hospital of Nantong University, Nantong City, Jiangsu, China.
** Corresponding author.
E-mail addresses: wang.ziheng@connect.um.edu.mo (Z. Wang), chengzhouyang1210@hotmail.com (Z. Cheng).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e27388
Received 3 September 2023; Received in revised form 28 February 2024; Accepted 28 February 2024
Available online 5 March 2024

identified through the differential expression of contrast agent COAD and normal tissues. Finally, the investigation of variations in immune infiltration and the analysis of relevant biological pathways validate the prognostic significance of the independent post-factors within this molecular model. This research aims to provide references for comprehending the mechanism of post-transcriptional regulation of COAD and molecular therapy.

1. Introduction

Colon adenocarcinoma (COAD) is a prevalent cancer and continues to be the third leading contributor to cancer-related mortality on a global scale [1,2]. The occurrence of colorectal cancer in China has risen over the last twenty years; however, the prognosis has not improved significantly [3,4]. Therefore, new biomarkers for prognostic evaluation should be further explored to investigate the molecular mechanism of COAD development [5]. Furthermore, there is growing evidence that, in addition to genetic mutations, the tumor microenvironment (TME) [6] and immune cells are critically involved in the progression of COAD. Anti-tumor immune patterns through cytokine interactions are closely associated with the maintenance of intestinal homeostasis and the intrinsic mechanisms of intestinal cancer progression [7].

Peroxisome proliferators-activated receptors (PPARs) were first peroxisome in 1990. PPARs belong to the steroid hormone receptor family, a group of ligands that activate transcription factors that help regulate the expression of Hub genes involved in lipid metabolism, lipogenesis, and glucose control [8]. PPARs are also transcription factors activated by ligands [9]. Upon binding to their fatty acid ligand, PPAR forms a heterodimer complex with retinoic acid X receptor (RXR) to regulate transcription [10,11]. Recent evidence suggests that PPARs are implicated in inflammation and belong to the processes related to the regulation of gene expression, They function as connectors between lipid metabolic disorders and innate immunity, and their activation by fatty acids and their derivatives facilitates signal transduction between the cell nucleus and the cell surface [9]. PPARs have become an exciting therapeutic target for several diseases. Peroxisome proliferator-activated receptor α (PPARα) is a nuclear receptor that acts as an exogenous and lipid sensor to regulate energy burning, lipid homeostasis, and inflammation [12]. A study by Yuhong Luo et al. reported that deletion of PPARα in the intestine promotes colon carcinogenesis by increasing DNMT1-mediated P21 methylation and PRMT6-mediated p27 methylation in mice [13]. Peroxisome proliferator-activated receptor δ (PPAR δ) is a member of the ligand-activated PPAR nuclear receptor family [14]. A study by Yuhong Luo et al. reported that downregulation of PPAR δ promoted colon cancer growth by inducing less differentiation, accelerated proliferation of tumor cells and VEGF expression in vivo, and reduced tumor sensitivity to bevacizumab [15]. PPARs are expressed in various types of tumors, including tumor cells, and their role is intimately linked to the onset and progression of COAD [16]. Therefore, further research is being carried out in this field to discover new PPARs-related therapeutic targets, offering potential avenues for safe and effective treatment of COAD.

This study analyzed genomic information from 461 CRC samples to comprehensively assess and construct the PPARS prognostic risk score model. Univariate and multivariate Cox proportional hazards regression analyses were further employed to ascertain the independent variables influencing overall survival (OS) outcomes in COAD. CPT1B, SLC27A4, and FABP3 were found to be independent risk prognostic factors for OS outcomes in COAD, whereas aCOX1 and CPT2 served as independent protective prognostic factors. Furthermore, an assessment was executed to examine the link between the prognostic risk score model and the features of TME cell infiltration. These outcomes introduce a fresh angle for investigating the metabolic mechanism and therapeutic approaches for COAD.

2. Material and methods

2.1. Data collection

Tertiary data for mRNA sequencing in individuals with COAD were obtained from TCGA (https://TCGA-data.nci.nih.gov/TCGA/). The RNAseq data comprised 521 cases, consisting of 41 paracancerous tissues and 480 colon cancer tissues. A total of 461 patients with available clinical information were selected for additional analysis The obtained gene expression data were initially recorded as “One million fragments per kilobase” (FPKM), and the raw data were subsequently transformed into “Transcription number in millionths” (TPM).

2.2. Identification of differentially expressed genes

For the identification of differentially expressed genes (DEGs) between COAD and normal tissues, the “limma” R-package was used [17]. The criteria for DEG screening were set at p-values <0.05 and $|\log FC| > 1.5$. Additionally, 67 PPAR pathway genes were downloaded from the “WP” section of the MsigDB online database (https://www.gsea-MsigDB.org/gsea/MsigDB/index.jsp) [18].

2.3. Enrichment analysis

Gene Ontology (GO) [19] and Kyoto Encyclopedia of genes and genomes (KEGG) analyses were executed utilizing pertinent software packages with the ‘clusterProfiler’ R package [20,21]. Gene set enrichment analysis (GSEA) was conducted to ascertain

differential functions or gene sets that could be preferentially defined between two biological states. Meaningful potential functional enrichment pathways within the "c2.cp.v7.2″ category were scrutinized in the risk groups determined by the prognostic model. Furthermore, the analysis produced the Normalized Enrichment Score (NES, nom *P*-value, and FDR q-value using GSEA to divide the enrichment pathway into two phenotypes. $|NES| > 1$, nom P-value < 0.05 and FDR q-value < 0.25 were used as thresholds.

2.4. Machine learning selection of biomarkers

Machine learning models were utilized to screen genes significantly associated with prognosis and construct a model with excellent predictive performance. The gene expression values underwent normalization utilizing "Log2ðx + 1ði" and "Minmax" normalization methods. The least absolute shrinkage and selection operator (LASSO) regression algorithm was applied to screen genes closely associated with OS in one step, and 10-fold cross-validation was used to screen candidate signature genes and determine the penalty parameter (λ), which corresponds to the minimum value of the biased likelihood deviation. The coefficients obtained by LASSO regression were applied to each normalized gene expression value, leading to the formulation of the risk scoring formula [22,23].

$$\text{risk}Score = \sum_{i} Coefficient\ (hub\ gene_i) * mRNA\ Expression\ (hub\ gene_i)$$

2.5. Development and validation of the risk model

The important values of genes in the two models were quantified, and the 11 genes with the most important values were chosen as pivotal genes for further study. Prognostic analysis of clinical data was acquired from TCGA. Survival analysis showed differences in OS. Furthermore, univariate and multivariate Cox regression analyses were conducted to determine independent predictive variables. Receiver Operating Characteristic (ROC) curves and calibration curves were employed for discrimination and correction [24].

2.6. Evaluation of tumor microenvironment (TME)

The CIBERSORT algorithm in R was utilized to determine the TME status for each COAD sample [25]. CIBERSORT was employed to predict the relative levels of 22 tumor-infiltrating lymphocytes (TILs) in COAD samples. A selection criterion of a $P \leq 0.05$ was applied to ensure the reliability of the prediction results [26].

2.7. qRT-PCR

The total RNA was extracted from GES1, MKN-45, and SUN-16 cell lines using the TRIzol reagent (Thermo Fisher, USA). From 500 ng of RNA, cDNA was generated utilizing HiScript II SuperMix (Vazyme, China). qRT-PCR was conducted in the ABI 7500 System (Thermo Fisher, USA) employing SYBR Green Master Mix. PCR amplification conditions involved 46 cycles of 94 °C for 10 min, 94 °C for 10 s, and 60 °C for 45 s. GAPDH was utilized as the internal reference. Below is a list of primer pair sequences for the targeted genes.

Gene	Forward primer sequence (5′-3′)	Reverse primer sequence (5′-3′)
CPT1B	TGTATCGCCGTAAACTGGACCG	TGTCTGAGAGGTGCTGTAGCAC
CYP4A11	CATGGCAGACTCTGTACGAGTG	CTGATGGCTGAAGGCACACTTC
CPT2	GCAGATGATGGTTGAGTGCTCC	AGATGCCGCAGAGCAAACAAGTG
CPT1C	TGCCATGTCGTTCCATTCTCCC	GCCGACTCATAAGTCAGGCAGA
SLC27A4	GGACCAACTTTTCCAGCCGCTT	TGCGGCTATTGAAACCACAGGC
FABP3	GTGGAGTTCGATGAGACAACAGC	TGGTCTCTTGCCCGTCCCATTT
FABP7	CTGTTGTTAGCCTGGATGGAGAC	CTCATAGTGGCGAACAGCAACC
AQP7	GAGGAAGATGGTGCGAGAGTTC	CAAGTTGACACCAAGGTAGCTCC
MMP1	ATGAAGCAGCCCAGATGTGGAG	TGGTCCACATCTGCTCTTGGCA
ACOX1	GGCGCATACATGAAGGAGACCT	AGGTGAAAGCCTTCAGTCCAGC
ANGPTL4	GATGGCTCAGTGGACTTCAACC	TGCTATGCACCTTCTCCAGACC
GAPDH	CTGGGCTACACTGAGCACC	AAGTGGTCGTTGAGGGCAATG

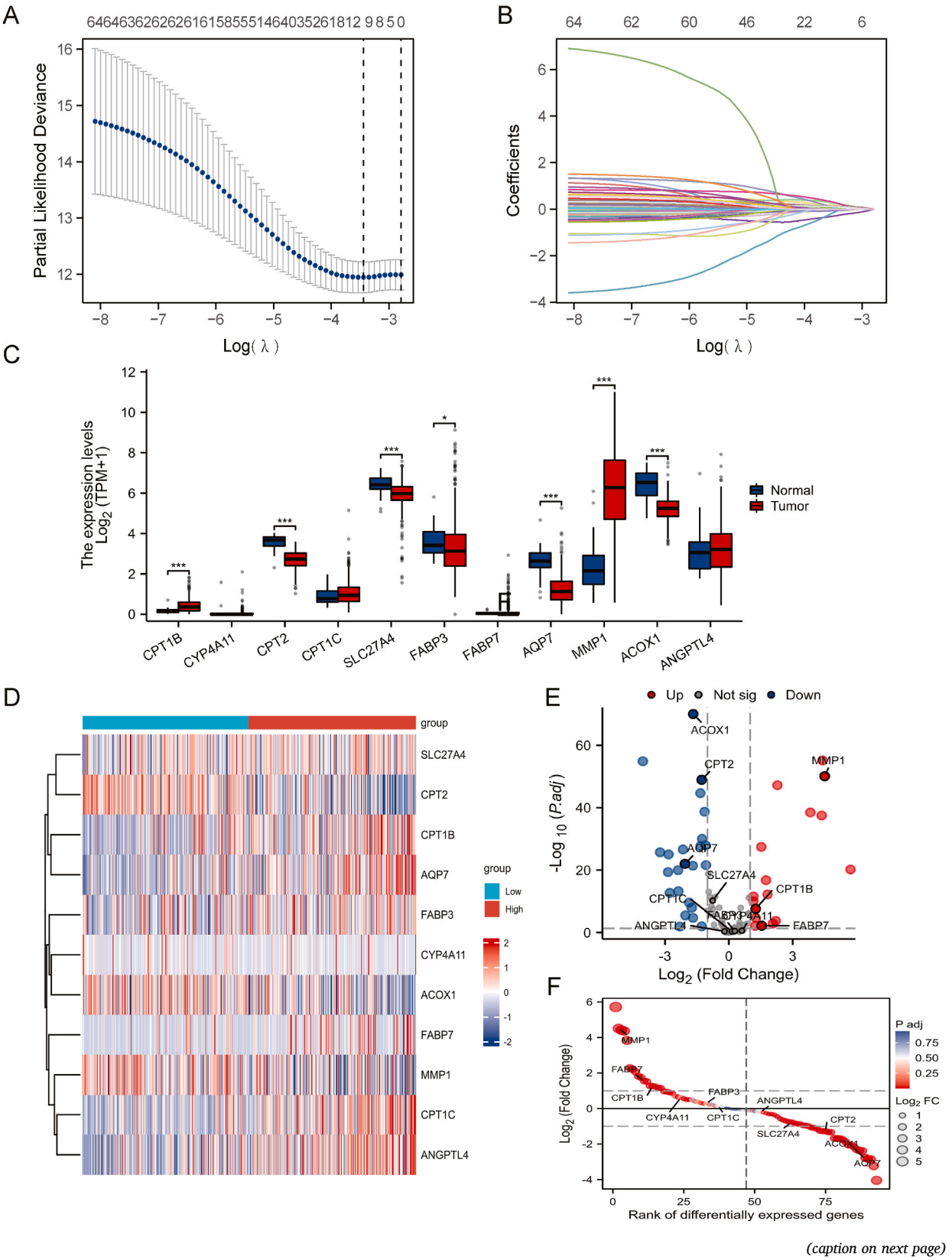
2.8. Statistical analyses

Data are expressed as means ± standard error (SD). Differences between groups were assessed through a Student's t-test. Statistical analyses were conducted utilizing R 4.0.2. $P < 0.05$ (two-tailed) was deemed as a statistically significant value.

3. Results

3.1. Construction of lasso model based on PPRAs of OS in COAD

Owing to the biological importance of PPARs in tumor processes, a diagnostic and OS prognostic model for COAD was constructed.



(caption on next page)

Fig. 1. Establishment of prognostic risk scoring model. (A) LASSO coefficients of 67 PPARs-associated genes; (B) Genetic identification of prognostic risk score model development; (C) Expression difference box plots of prognostic model genes in TCGA-COAD database; (D) Heatmap showing remarkable variations in the expression of hub PPRAs prognostic genes between high and low-risk groups, with red indicating high risk and blue indicating low risk; (E–F) Differential volcano and differential expression sequencing maps of prognostic PPARs-related genes in COAD.

This model was based on 67 genes associated with the PPARs pathway from the WP database. It can be utilized in regression analysis for diagnostic and prognostic purposes. Initially, lasso regression analysis was conducted for regression screening, aiming to ascertain the optimal lambda values for the identification of 11 statistically significant genes. A 11-gene diagnostic model consisting of SLC27A4, CPT1C, CPT1B, CPT2, CYP4A11, FABP3, FABP7, AQP7, MMP1, ACOX1 and ANGPTL4 was obtained. The Lambda and Min values are visualized through the LASSO Logit model algorithm, and their stability is ensured by specifying the loss function of LASSO (Fig. 1A and B). The expression differences among these prognostic genes were examined within the TCGA-COAD database. Notably, CPT2, SLC27A4, FABP3, AQP7, and ACOX1 exhibited significantly lower expression levels in tumor tissues. Conversely, CPT1B and MMP1 exhibited elevated expression in tumor tissues (Fig. 1C). The LASSO model of PPRAS outcomes based on OS classifies individuals with COAD into high- and low-risk groups. The heatmap demonstrated remarkable variations in the expression of hub PPRA prognostic genes between both risk groups. For the risk group, red denotes high risk, and blue denotes low risk (Fig. 1D). Subsequently, the differential expression of these prognostic PPARs-related genes in COAD tissues was demonstrated by volcano plots and differential expression sequencing maps (Fig. 1E and F).

Risk Factor Association maps show differences in population proportion, OS time and survival status, and gene expression between high-risk and low-risk COAD patients as distinguished by the construction of the LASSO model (Fig. 2A). To depict the correlation between the prognostic genes, a correlation heatmap was utilized. The findings indicated considerable variations between the two molecules with positive correlation. However, most of the negative correlations were not significant (Fig. 2B). Subsequently, an analysis was conducted to ascertain if the prognostic risk model held independent prognostic significance. This was achieved through Cox regression, considering both the survival status and OS time of COAD patients. Univariate and multivariate Cox regression analysis showed that only CPT1B, SLC27A4, ACOX1, FABP3, and CPT2 were independent prognostic factors for OS in individuals with COAD (Fig. 2C and D).

3.2. Verification of prognostic molecular efficacy

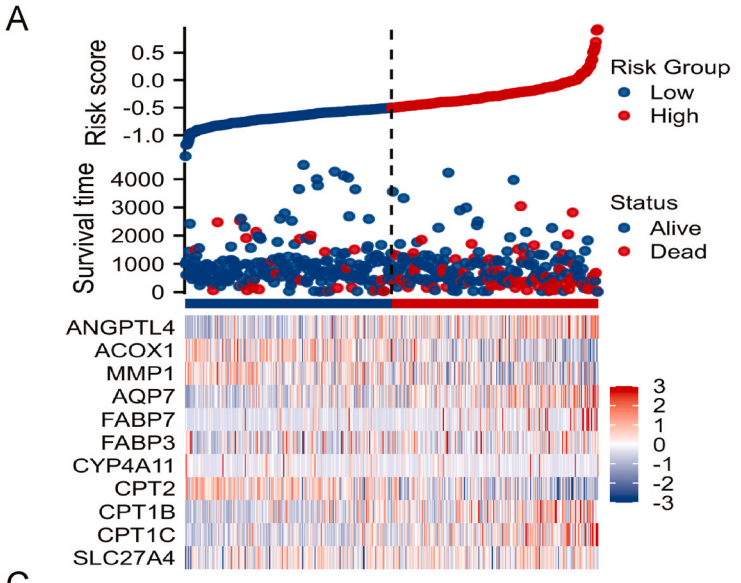
To confirm the predictive capability of the diagnostic model, the R software package pROC was employed to generate individual gene ROC curves and compute the corresponding areas under the curve (AUC). Next, to reveal the relationship between the 11 PPARs-related markers in COAD and the diagnostic value of the disease and the OS prognosis, R package RMS was employed. This facilitated the construction of time-dependent ROC curves for the 11 hub genes that held significant weights in a prior lasso model. Only FABP3 [1−year (AUC = 0.530); 3−year (AUC = 0.545); 5−year (AUC = 0.577)], ANGPTL4[1−year (AUC = 0.530); 3−year (AUC = 0.545); 5−year (AUC = 0.577)], CPT1B[1−year (AUC = 0.592); 3−year (AUC = 0.600); 5−year (AUC = 0.598], and CPT1C[1−year (AUC = 0.581); 3−year (AUC = 0.556); 5−year (AUC = 0.558)] were found to have predictive value for OS at 1,3, and 5 years in individuals with COAD (Fig. 3A–D). Meanwhile, the survival curve on OS time found that only patients with elevated expression levels of CPT1B ($P = 0.046$) and SLC27A4 ($P = 0.046$) among hub PPARs-related genes in the model had significantly worse prognostic significance (Fig. 3E and F).

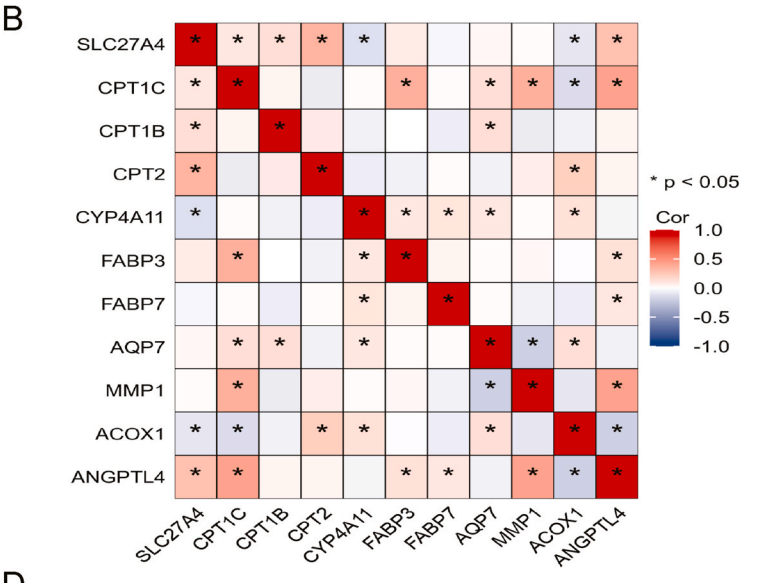
3.3. Gene enrichment analysis

To delve deeper into the biological implications of the six stem cell-related marker models in COAD, comprehensive functional enrichment investigations were conducted. These investigations involved genes with varying dimensions and levels and were executed through GO enrichment analysis from three levels: biological process (BP), molecular function (MF), and cellular component (CC). Kyoto Encyclopedia of Genes and Genomes (KEGG) is an extensively utilized database that houses data concerning genomes, diseases, biological pathways, and medicines. All significantly differentially expressed genes underwent annotation with GO functions through the utilization of the R-package clustering analyzers. This approach aimed to determine remarkable enrichment in BP. The outcomes of the enrichment analysis were visualized utilizing the R package GO plot. The threshold for significance in the enrichment analysis was established at p-values <0.05. The outcomes of the functional enrichment analysis conducted on the PPARs prognostic genes indicated a predominant enrichment of CC in the outer mitochondrial membrane and the outer membrane of organelles. Furthermore, significant functions related to palmitoyltransferase activity, fatty acid binding, o-acyltransferase activity, and carboxylic acid-binding were enriched. Bioengineering is centered on the study of fatty acid transport and catabolism, as well as the catabolism of monocarboxylic acids, carboxylic acids, and organic acids (Fig. 4A, C, E). The prognostic model was concentrated on fatty acid degradation, fatty acid metabolism, alcoholic liver disease, and thermogenesis (Fig. 4B, D, F).

3.4. Gene set enrichment analysis of prognostic risk model

The prognostic model associated with high-risk groups exhibited enrichment in various pathways, including fatty acid metabolism, folate biosynthesis, nitrogen metabolism, oxidative phosphorylation, terpenoid backbone biosynthesis, and valine leucine and isoleucine degradation, as identified in the KEGG database (Fig. 5A). The prognostic low-risk group exhibited enriched entries





C

Characteristics	Total(N)	HR(95% CI) Univariate analysis	P value Univariate analysis
CPT1C	477		0.145
Low	238	Reference	
High	239	1.336 (0.904 - 1.973)	0.146
SLC27A4	477		0.046
Low	238	Reference	
High	239	1.486 (1.004 - 2.200)	0.048
CPT1B	477		0.013
Low	238	Reference	
High	239	1.643 (1.107 - 2.439)	0.014
ANGPTL4	477		0.241
Low	238	Reference	
High	239	1.261 (0.856 - 1.858)	0.241
ACOX1	477		0.055
Low	239	Reference	
High	238	0.682 (0.459 - 1.012)	0.058
MMP1	477		0.135
Low	239	Reference	
High	238	0.744 (0.503 - 1.099)	0.137
AQP7	477		0.287
Low	238	Reference	
High	239	1.237 (0.835 - 1.831)	0.289
FABP3	477		0.066
Low	238	Reference	
High	239	1.441 (0.974 - 2.133)	0.067
CPT2	477		0.009
Low	238	Reference	
High	239	0.591 (0.396 - 0.882)	0.01

D

Characteristics	Total(N)	HR(95% CI) Multivariate analysis	P value Multivariate analysis
CPT1C	477		
Low	238		
High	239		
SLC27A4	477		
Low	238	Reference	
High	239	1.675 (1.105 - 2.538)	0.015
CPT1B	477		
Low	238	Reference	
High	239	1.693 (1.139 - 2.517)	0.009
ANGPTL4	477		
Low	238		
High	239		
ACOX1	477		
Low	239	Reference	
High	238	0.718 (0.482 - 1.069)	0.102
MMP1	477		
Low	239		
High	238		
AQP7	477		
Low	238		
High	239		
FABP3	477		
Low	238	Reference	
High	239	1.341 (0.900 - 2.000)	0.149
CPT2	477		
Low	238	Reference	
High	239	0.528 (0.347 - 0.804)	0.003

Fig. 2. Prognostic model risk analysis. Risk curves for (A) risk genome (top), survival status (middle), risk heatmap (bottom), and (B) relevant heatmap for prognostic model genes; (C–D) Univariate and multivariate Cox regression analysis of forest plots for risk model genes based on overall survival of COAD.

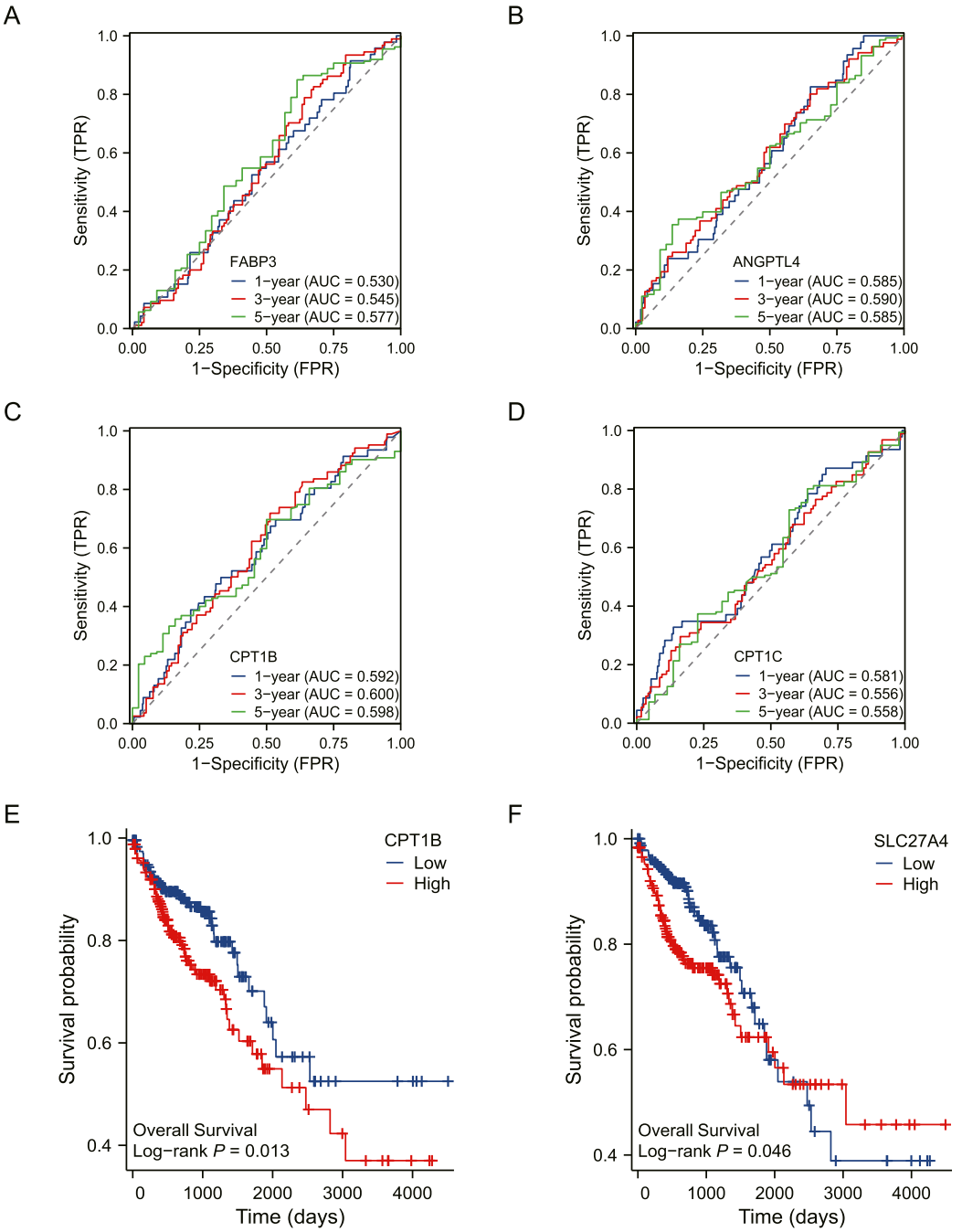
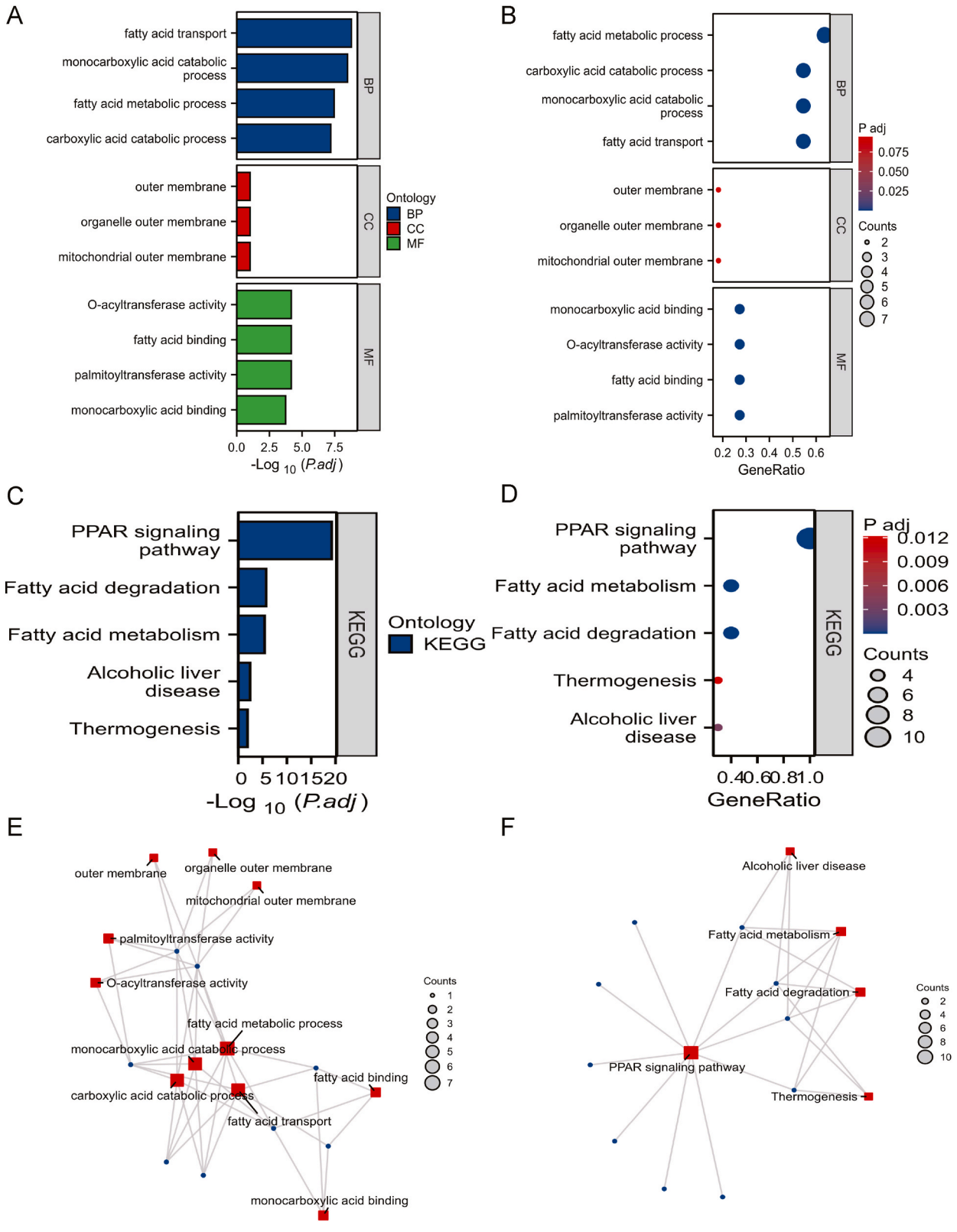


Fig. 3. Validation of diagnostic models for predicting the diagnosis and prognosis of COAD. (A–D) Single-gene time-dependent ROC curves were calculated based on OS at 1,3, and 5 years in COAD patients and the area under the curves. (E–F) KM survival curves showed that elevated expression of CPT1B ($P = 0.046$) and SLC27A4 ($P = 0.046$) had remarkably worse OS outcomes, and the variation between high and low-expression groups was statistically significant.

including acute myeloid leukemia, arrhythmogenic right ventricular cardiomyopathy (ARVC), hypertrophic cardiomyopathy (HCM), and small cell lung cancer (SCLC) (Fig. 5B). Within the REACTOME database, high-risk samples exhibited enrichment in pathways related to cholesterol biosynthesis, L-Cysteine and homocysteine degradation, chemiosmosis ATP formation, mitochondrial fatty acid β oxidation, and other molecular functional pathways (Fig. 5C). The low-risk group exhibited remarkable enrichment in the REACTOME database, specifically in pathways involving receptor-dependent activation of caspase, PTEN translation-related processes, DCC-mediated signaling, apoptotic soma formation, and the binding of GLI protein to the HH response gene promoter, facilitating transcription, among other factors (Fig. 5D). The high-risk group exhibited significant enrichment in various entries, including pathways



(caption on next page)

Fig. 4. Enrichment analysis of PPARs-related prognostic genes. (A) GO enrichment analysis bar plot with column length representing the number of gene enrichment; (B) GO enrichment analysis bubble plot with bubble size representing the number of gene enrichment and color representing significance, with significance increasing gradually from yellow to blue; (C) KEGG enrichment analysis bar plot with column length representing the number of gene enrichment; (D) KEGG enrichment analysis bubble plot showing bubble size representing the number of gene enrichment and color representing significance, with gradual increase from yellow to blue. (E–F) GO and KEGG Enrichment Analysis Enrichment Network diagram.

such as benzoapyrene metabolism, electron transport chain oxphos system in mitochondria, iron metabolism in placenta, as well as mitochondrial CII and CIII assembly within the WP database (Fig. 5E). IL17, IL3 and mapk/nfkb signalling pathways inhibited by yersinia yopj, mir5093p alteration of yap1ecm axis, mirna biogenesis and targets in ecm and membrane receptors were enriched entries for low-risk groups in WP database (Fig. 5F).

3.5. Immune infiltration analysis of predictive risk model

To investigate how the prognostic marker model affects the immune cell infiltration within the COAD microenvironment, the transcriptome expression matrix was deconvolved using the CIBERSORT algorithm as per the linear support vector regression principle. Subsequently, the composition of immune cells within the mix was assessed. Samples with a significance level of $P < 0.05$ were selected and the matrix infiltrated by the immune cells was obtained (Fig. 6A). Subsequent analysis of immune cell infiltration levels indicated variations in the impact of the prognostic model on multiple immune cells. Notably, resting memory CD4 T cells, activated memory CD4T cells, resting dendritic cells, activated dendritic cells, and eosinophils exhibited remarkably reduced infiltration levels in the high-risk group. Macrophages M0 and regulatory T cells (Tregs) exhibited remarkably elevated infiltration levels in the high-risk group (Fig. 6B). Additionally, the degree of infiltration between immune cells was significantly correlated (Fig. 6C). Subsequent correlation analysis indicated a negative correlation between PPAR-related markers and the level of immune cell infiltration, implying that inadequate immune cell infiltration might play a role in accelerated tumor advancement and unfavorable prognosis (Fig. 6D). The prognostic risk score exhibited negative correlation with resting dendritic cells ($r = -0.28$, $P = 6.7e-05$) and resting memory CD4 T cells ($R = -0.32$, $P = 5.5E-06$), while it exhibited a positive correlation with Tregs ($R = 0.28$, $P = 6.9e-05$). These results are consistent with the fact that dendritic cells and memory CD4 T cells play a beneficial role in enhancing immune cell-mediated tumor killing (Fig. 6E–G).

3.6. qRT-PCR for detecting the reliability of bioinformatics models

To test the reliability of bioinformatics models, the expression of CPT1B, CYP4A11, CPT2, CPT1C, SLC27A4, FABP3, FABP7, AQP7, MMP1, ACOX1, and ANGPTL4 were examined in normal colonic epithelial cells (NCM-460), as well as in colonic adenocarcinoma cell lines (RKO and SW48) using qRT-PCR. It was observed that CPT1B and MMP1 exhibited higher expression levels in colonic adenocarcinoma cell lines, whereas the expression levels of CPT2, SLC27A4, FABP3, AQP7, and ACOX1 were reduced in colonic adenocarcinoma cell lines. This alignment between the observed results and the predictions made by the bioinformatics model lends support to the validity of the model (see Fig. 7).

4. Discussion

COAD ranks as the second primary cause of cancer-related mortality worldwide [3], with metastasis standing out as the leading contributor to cancer-related fatalities [27–29]. The 5-year OS rate of individuals with metastatic colorectal cancer (advanced stage (IV)) is only 4–12% [30,31].

Lasso regression analysis was conducted to identify molecular models linked to OS outcomes in individuals with COAD, encompassing 67 PPARs pathway-related genes. The 11-gene prognostic model, comprised of SLC27A4, CPT1C, CPT1B, CPT2, CYP4A11, FABP3, FABP7, AQP7, MMP1, ACOX1 and ANGPTL4, was consequently established. The diagnosis of COAD and its association with OS were analyzed as per the transcriptome expression data of the prognostic genes. Univariate and multivariate Cox regression analyses confirmed that CPT1B, SLC27A4, and FABP3 were independent risk prognostic factors, while ACOX1 and CPT2 were independent protective prognostic factors (Table 1). Subsequent enrichment analysis also suggested a close association with metabolic pathways, metabolic functions, and associated metabolic syndrome, such as fatty acids.

To assess the impact of prognostic risk models on immune cell infiltration in the COAD microenvironment, the transcriptome expression matrix was deconvoluted using the CIBERSORT algorithm. Notably, differences in immune cell infiltration between the prognostic molecular models were evident. CD4 T cells, Tregs, dendritic cells, and eosinophil granulocytes exhibited reduced infiltration in the high-risk group. Conversely, the high-risk group manifested notably heightened infiltration of M0 macrophages and Tregs. Concurrently, PPAR-related markers predominantly demonstrated inverse associations with the level of immune cell infiltration. This coherence suggests the potential implication that insufficient immune cell infiltration could potentially contribute to accelerated tumor progression and unfavorable prognosis outcomes. Additionally, there was also a remarkable correlation between the degree of infiltration among immune cells. Further correlation analysis found that PPARs-related prognostic markers were inversely correlated with the extent of immune cell infiltration, suggesting that insufficient immune cell infiltration may contribute to rapid tumor progression and poor prognosis. The prognostic risk score displayed a noteworthy correlation with the infiltration of immune cells, including dendritic cells, CD4 T cells, and Tregs. This alignment is congruent with the understanding that dendritic cells and

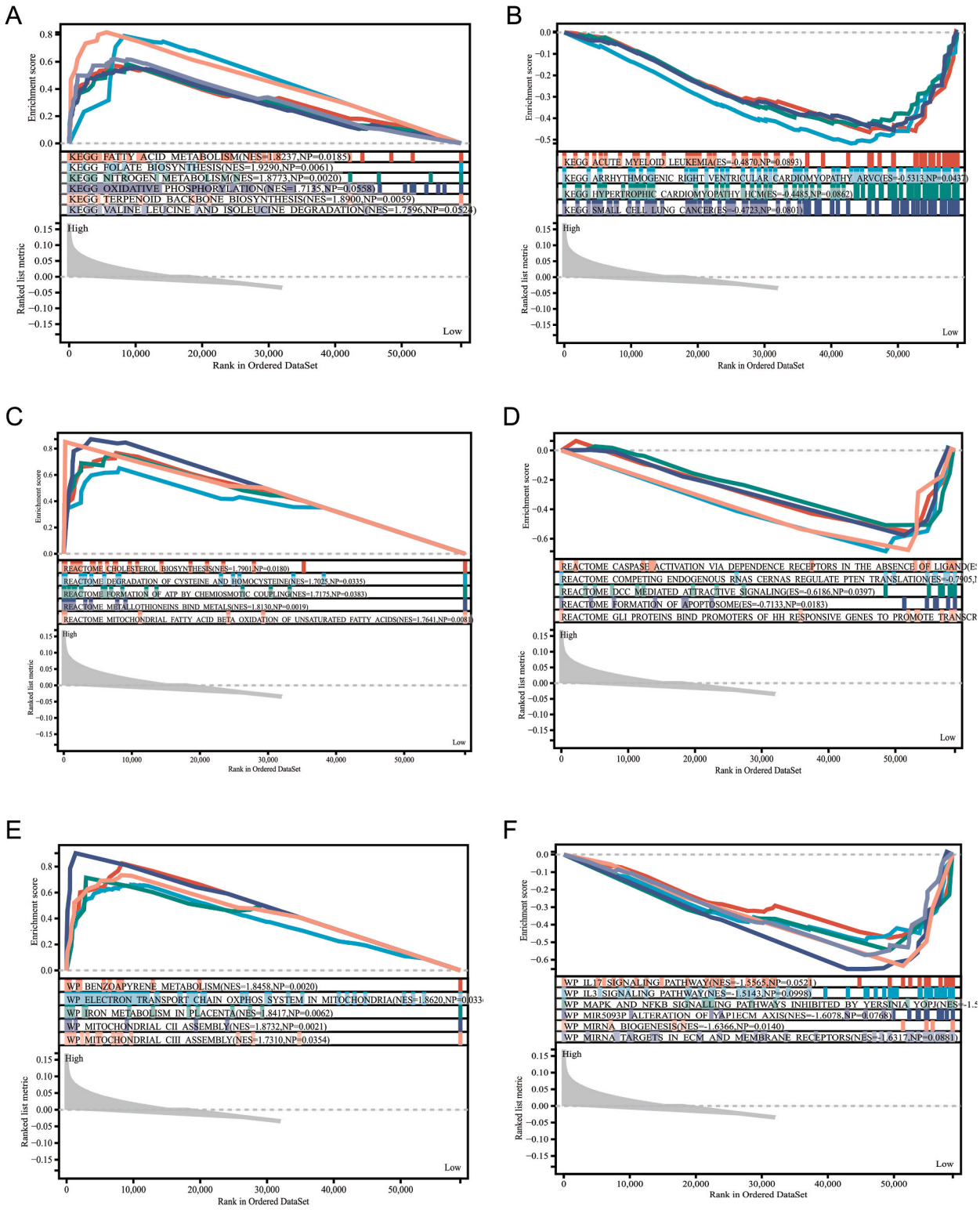


Fig. 5. Prognostic model of GSEA. GSEA outcomes of differential expression of TMEM131L. Elevated (A, C, E) and reduced (B, D, F) expression levels of TMEM131L exhibited enrichment in BIOCARTA, REACTOME, KEGG, and WP.

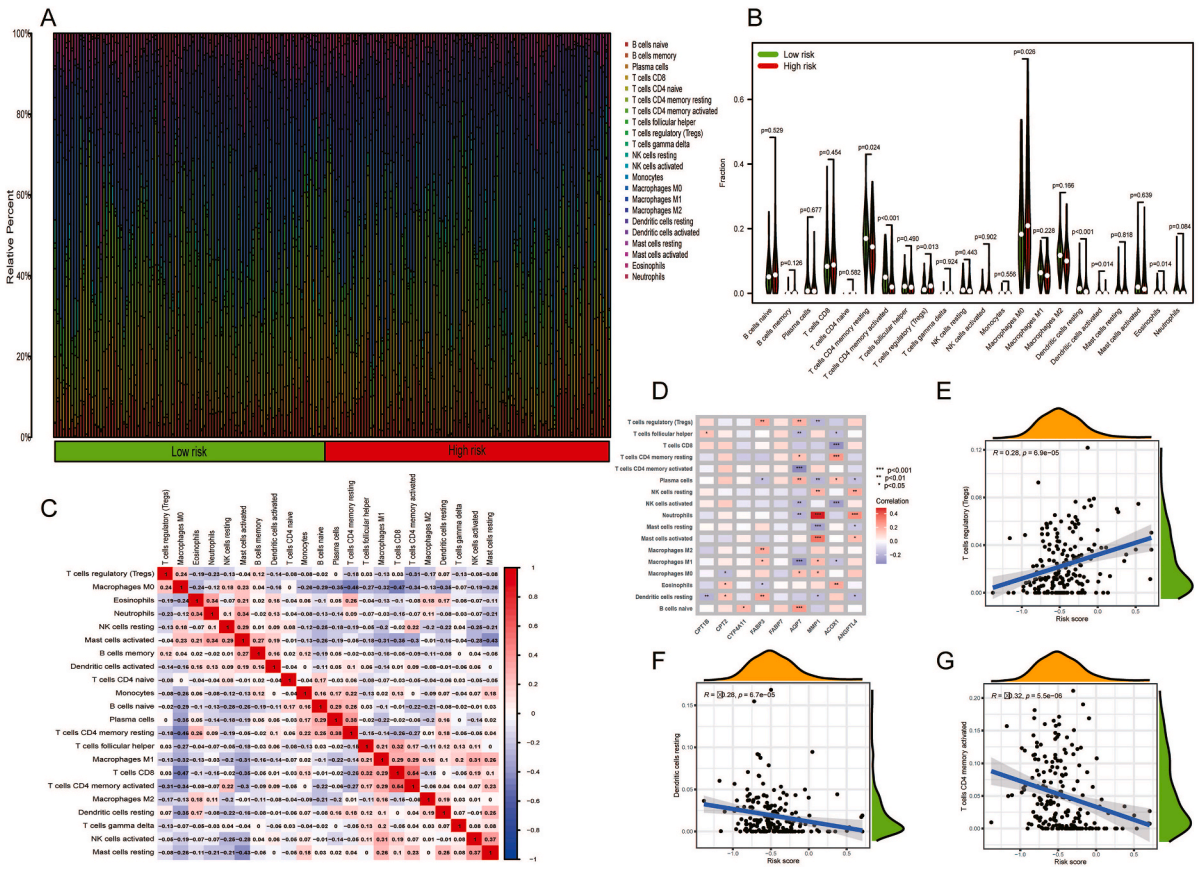


Fig. 6. Evaluation of immune infiltration. (A) Stacked bar graphs showing the proportion of infiltrating immune cells between the high-and low-risk groups; (B) Boxplot of differences in immune cell infiltration between high-and low-risk groups (Wilcoxon's test *< 0.05; * * * < 0.001). The horizontal axis indicates the type of infiltrating immune cells; the vertical axis indicates the proportion of infiltrating immune cells; (C) Correlation heatmap of infiltrating immune cell distribution, (D) Correlation heatmap of PPARs-related markers with the degree of infiltration of immune cells, (E) Positive correlation scatter plot of PPARS prognostic model risk score with the degree of infiltration of T-cell regulatory (Tregs); (F–G) Negative correlation scatter plot of PPARS prognostic model risk score with resting dendritic cell (r = −0.28, P = 6.7e-05) and resting memory CD4 T-cell (R = −0.32, P = 5.5e-06).

memory CD4 T cells are advantageous for immune cell killing of tumors [32].

Of course, there are some shortcomings in our study. First, we only used TCGA data for the analysis. Secondly, the ROC of the pivotal gene was only slightly higher than 0.05, which may be due to the small sample size. In addition, the clinicopathologic characteristics of the COAD patients included in the study were limited, so we need more practical and valuable factors to predict the survival outcome of COAD patients. Finally, we only performed qRT-PCR analysis to verify the expression of the genes used to construct the model. Therefore, we still need to verify the clinical value of the model and the disease molecules through more prospective studies in the future.

5. Conclusion

In summary, this study constructed a COAD-related clinical diagnostic model that demonstrates accurate diagnostic and prognostic capabilities and has the potential to characterize the clinicopathological features of COAD. The model can accurately stratify COAD and thus improve individualized patient treatment. More, CPT1B, SLC27A4 and FABP3 were identified as independent risk prognostic factors for OS in COAD, while ACOX1 and CPT2 were independent protective prognostic factors. We hope that this study will inform the understanding of post-transcriptional regulatory mechanisms and molecular therapy in COAD.

Funding

This study was supported by grants from Jiangsu Provincial Research Hospital (YJXYY202204-YSB20), the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX23 3433), the Science and Technology Project of Nantong Municipal Health Commission (NO. MS2023016).

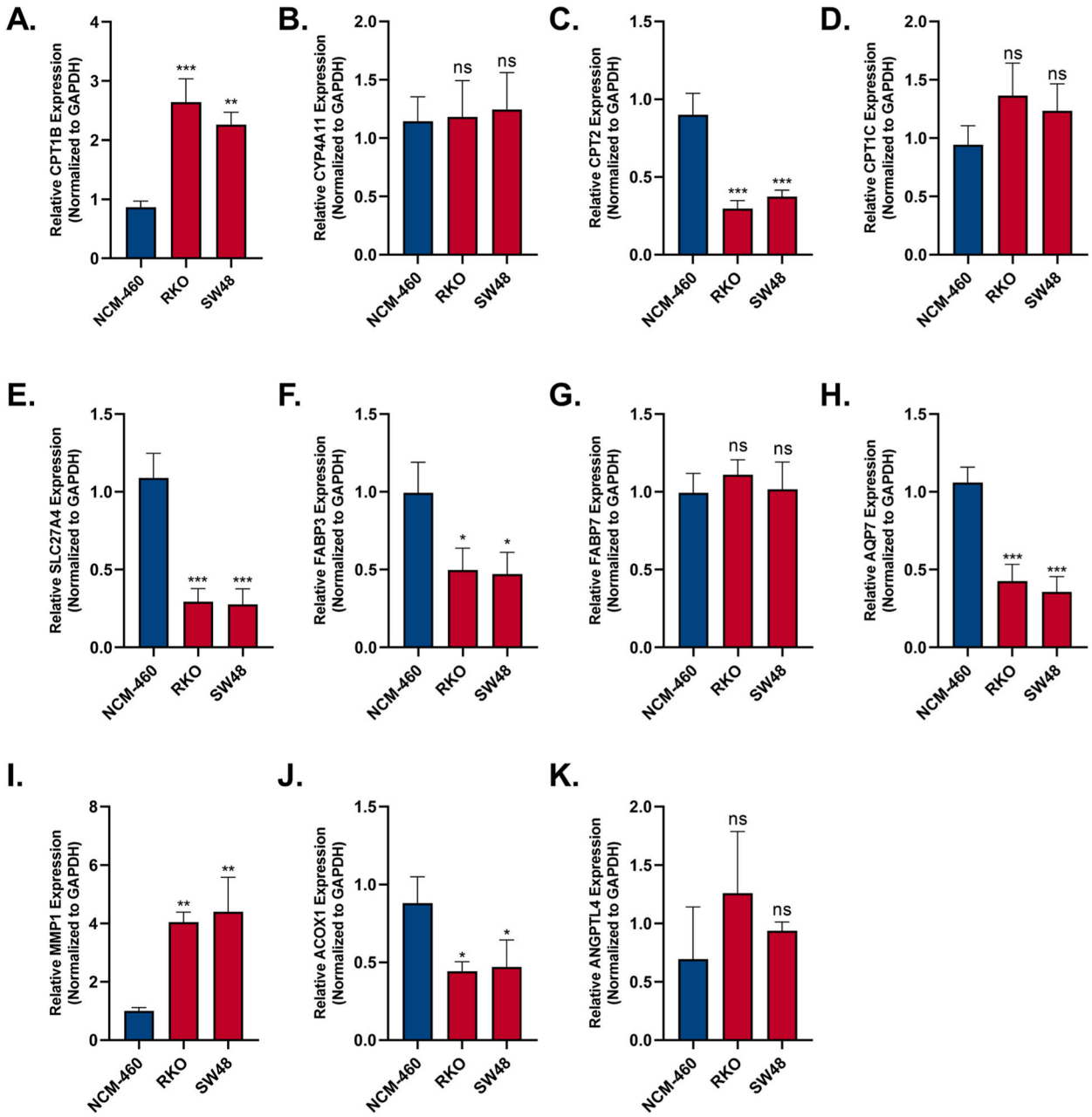


Fig. 7. qRT-PCR for detecting the reliability of bioinformatics models. (A–K) qRT-PCR assay for the expression of CPT1B, CYP4A11, CPT2, CPT1C, SLC27A4, FABP3, FABP7, AQP7, MMP1, ACOX1, and ANGPTL4 in normal colonic epithelial cells, NCM-460, as well as in colonic adenocarcinoma cell lines, RKO and SW48. N = 3, *≤0.05, **≤0.01 , ***≤0.001 , ****≤0.0001. The findings are expressed as mean ± SD.

Data availability statement

All raw data can be access from correspondence authors.

CRediT authorship contribution statement

Yang Yang: Writing – original draft, Software, Investigation, Formal analysis, Data curation, Conceptualization. **Xinyu Yang:** Writing – original draft, Formal analysis, Data curation, Conceptualization. **Shiqi Ren:** Writing – review & editing, Software, Investigation, Formal analysis, Data curation, Conceptualization. **Yang Cao:** Writing – original draft, Methodology. **Ziheng Wang:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Data curation,

Table 1
Univariate and multivariate Cox regression analysis of risk model genes based on OS of COAD.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		HR(95% CI)	*P* value	HR(95% CI)	P value
CPT1C	477		0.145		
Low	238	Reference			
High	239	1.336 (0.904–1.973)	0.146		
SLC27A4	477		0.046		
Low	238	Reference		Reference	
High	239	1.486 (1.004–2.200)	0.048	1.675 (1.105–2.538)	0.015
CPT1B	477		0.013		
Low	238	Reference		Reference	
High	239	1.643 (1.107–2.439)	0.014	1.693 (1.139–2.517)	0.009
ANGPTL4	477		0.241		
Low	238	Reference			
High	239	1.261 (0.856–1.858)	0.241		
ACOX1	477		0.055		
Low	239	Reference		Reference	
High	238	0.682 (0.459–1.012)	0.058	0.718 (0.482–1.069)	0.102
MMP1	477		0.135		
Low	239	Reference			
High	238	0.744 (0.503–1.099)	0.137		
AQP7	477		0.287		
Low	238	Reference			
High	239	1.237 (0.835–1.831)	0.289		
FABP3	477		0.066		
Low	238	Reference		Reference	
High	239	1.441 (0.974–2.133)	0.067	1.341 (0.900–2.000)	0.149
CPT2	477		0.009		
Low	238	Reference		Reference	
High	239	0.591 (0.396–0.882)	0.010	0.528 (0.347–0.804)	0.003

Conceptualization. **Zhouyang Cheng:** Writing – review & editing, Writing – original draft, Visualization, Resources, Project administration, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Ziheng Wang is the associated editor of Heliyon If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] O. Hoffstad, et al., Diabetes, lower-extremity amputation, and death, Diabetes Care 38 (10) (2015) 1852–1857.
[2] S. Haotong, et al., Increased MAD2L2 expression predicts poor clinical outcome in Colon Adenocarcinoma, Biocell 47 (3) (2023) 607–618.
[3] S. Dhall, et al., Release of insulin from PLGA-alginate dressing stimulates regenerative healing of burn wounds in rats, Clin. Sci. (Lond.) 129 (12) (2015) 1115–1129.
[4] A. Loboda, et al., Role of Nrf2/HO-1 system in development, oxidative stress response and diseases: an evolutionarily conserved mechanism, Cell. Mol. Life Sci. 73 (17) (2016) 3221–3247.
[5] C. P, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships, Cell Rep. 18 (1) (2017 Jan 3) 248–262, https://doi.org/10.1016/j.celrep.2016.12.019 (- 2211-1247 (Electronic)): p. - 248-262.
[6] D.C. Chambers, et al., Transcriptomics and single-cell RNA-sequencing, Respirology 24 (1) (2019) 29–36.
[7] B. Ma, et al., Tumor-associated macrophage (TAM) and angiogenesis in human colon carcinoma, Open Access Maced. J. Med. Sci. 3 (2) (2015 Jun 15) 209–214 (- 1857-9655 (Print)): p. - 209-214.
[8] A. Romero, et al., The angiotensin-(1-7)/Mas receptor axis protects from endothelial cell senescence via klotho and Nrf2 activation, Aging Cell 18 (3) (2019) e12913.
[9] R. Ah, et al., - Peroxisome proliferator-activated receptors: a therapeutic target in COPD? Eur. Respir. J. 31 (3) (2008 Mar) 502–508, https://doi.org/10.1183/09031936.00068207 (- 1399-3003 (Electronic)): p. - 502-508.
[10] K.E. Chatwin, et al., The role of foot pressure measurement in the prediction and prevention of diabetic foot ulceration-A comprehensive review, Diabetes Metab Res Rev 36 (4) (2020) e3258.
[11] M. D, et al., A retinoid X receptor partial agonist attenuates pulmonary emphysema and airway, Respir. Res. 20 (1) (2019 Jan 3) 2, https://doi.org/10.1186/s12931-018-0963-0 (- 1465-993X (Electronic)): p. - 2.
[12] S.R. Pyper, et al., PPARalpha: energy combustion, hypolipidemia, inflammation and cancer, Nucl. Recept. Signal. 8 (2010) e002.
[13] Y. Luo, et al., Intestinal PPARα protects against colon carcinogenesis via regulation of methyltransferases DNMT1 and PRMT6, Gastroenterology 157 (3) (2019) 744–759.e4.
[14] S.A. Kliewer, et al., Differential expression and activation of a family of murine peroxisome proliferator-activated receptors, Proc. Natl. Acad. Sci. U. S. A. 91 (15) (1994) 7355–7359.
[15] L. Yang, et al., Knockdown of PPAR δ gene promotes the growth of colon cancer and reduces the sensitivity to bevacizumab in nude mice model, PLoS One 8 (4) (2013) e60715.
[16] A.A. Kolodziejczyk, et al., The technology and biology of single-cell RNA sequencing, Mol. Cell 58 (4) (2015) 610–620.

[17] R. Me, et al., Limma powers differential expression analyses for RNA-sequencing and microarray, Nucleic Acids Res. 43 (7) (2015 Apr 20) e47, https://doi.org/10.1093/nar/gkv007. Epub 2015 Jan, (- 1362-4962 (Electronic)): p. - e47.
[18] M.B. Kursa, Robustness of Random Forest-based gene selection methods, BMC Bioinf. 15 (2014) 8.
[19] R.E. Mirza, et al., Blocking interleukin-1β induces a healing-associated wound macrophage phenotype and improves healing in type 2 diabetes, Diabetes 62 (7) (2013) 2579–2587.
[20] N.R. Garge, G. Bobashev, B. Eggleston, Random forest methodology for model-based recursive partitioning: the mobForest package for R, BMC Bioinf. 14 (2013) 125.
[21] T. Wu, et al., clusterProfiler 4.0: a universal enrichment tool for interpreting omics data, Innovation 2 (3) (2021) 100141.
[22] Z. Deng, et al., LncRNA SNHG14 promotes OGD/R-induced neuron injury by inducing excessive mitophagy via miR-182-5p/BINP3 axis in HT22 mouse hippocampal neuronal cells, Biol. Res. 53 (1) (2020) 38.
[23] S.Q. Ren, et al., Development and validation of a clinical prognostic model based on immune-related genes expressed in clear cell renal cell carcinoma, Front. Oncol. 10 (2020).
[24] H. Pj, Z. Y, Survival model predictive accuracy and ROC curves, Biometrics 61 (1) (2005 Mar) 92–105, https://doi.org/10.1111/j.0006-341X.2005.030814.x (-0006-341X (Print)): p. - 92-105.
[25] S.J. Dixon, et al., Ferroptosis: an iron-dependent form of nonapoptotic cell death, Cell 149 (5) (2012) 1060–1072.
[26] N. Am, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015 May) 453–457, https://doi.org/10.1038/nmeth.3337. Epub 2015 Mar 30., (-1548-7105 (Electronic)): p. - 453-457.
[27] B. Npm, et al., An overview of 25 years of incidence, treatment and outcome of colorectal cancer, Int. J. Cancer 143 (11) (2018 Dec 1) 2758–2766, https://doi.org/10.1002/ijc.31785. Epub 2018 Sep, (- 1097-0215 (Electronic)): p. - 2758-2766.
[28] B. Zhang, et al., Efficacy and safety of CTLA-4 inhibitors combined with PD-1 inhibitors or chemotherapy in patients with advanced melanoma, Int. Immunopharm. 68 (2019) 131–136.
[29] N. Lianli, et al., HOXB8 contributed to oxaliplatin chemo-resistance in colon cancer cells by activating STAT3, Biocell 47 (10) (2023) 2245–2254.
[30] R. M, et al., Patterns of metastasis in colon and rectal cancer, Sci. Rep. 6 (2016 Jul 15) 29765, https://doi.org/10.1038/srep29765 (- 2045-2322 (Electronic)): p. - 29765.
[31] Z. Lr, et al., Colorectal cancer liver metastasis: evolving paradigms and future directions, Cell Mol. Gastroenterol. Hepatol. 3 (2) (2017 Jan 20) 163–173 (- 2352-345X (Print)): p. - 163-173.
[32] C. Liming, et al., Expression of eIF6 and its relationship with cell proliferation in colorectal adenocarcinoma, Oncologie 0 (0) (2023).