



OPEN ACCESS

EDITED BY

Meng Zhou,
Wenzhou Medical University, China

REVIEWED BY

Fangfang Duan,
Shenzhen Campus of Sun Yat-sen
University, China
Ping Zheng,
The University of Melbourne, Australia
Zhenhuan Zhao,
University of Virginia, United States

*CORRESPONDENCE

Changhong Miao
15111230032@fudan.edu.cn
Di Zhou
judy612542@163.com
Minli Sun
sunminli@sina.com

[†]These authors have contributed
equally to this work and share
first authorship

SPECIALTY SECTION

This article was submitted to
Cancer Immunity
and Immunotherapy,
a section of the journal
Frontiers in Immunology

RECEIVED 23 May 2022

ACCEPTED 01 August 2022

PUBLISHED 23 August 2022

CITATION

Weng M, Li T, Zhao J, Guo M, Zhao W,
Gu W, Sun C, Yue Y, Zhong Z, Nan K,
Liao Q, Sun M, Zhou D and Miao C
(2022) mRNAsi-related metabolic risk
score model identifies poor prognosis,
immuno-evasive contexture, and low
chemotherapy response in
colorectal cancer patients
through machine learning.
Front. Immunol. 13:950782.
doi: 10.3389/fimmu.2022.950782

COPYRIGHT

© 2022 Weng, Li, Zhao, Guo, Zhao, Gu,
Sun, Yue, Zhong, Nan, Liao, Sun, Zhou
and Miao. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

mRNAsi-related metabolic risk score model identifies poor prognosis, immuno-evasive contexture, and low chemotherapy response in colorectal cancer patients through machine learning

Meilin Weng^{1,2†}, Ting Li^{1,2†}, Jing Zhao^{3†}, Miaomiao Guo^{1,2},
Wenling Zhao^{1,2}, Wenchao Gu^{4,5}, Caihong Sun^{1,2}, Ying Yue^{1,2},
Ziwen Zhong^{1,2}, Ke Nan^{1,2}, Qingwu Liao^{1,2}, Minli Sun^{1,2*},
Di Zhou^{1,2*} and Changhong Miao^{1,2*}

¹Department of Anesthesiology, Zhongshan hospital, Fudan University, Shanghai, China, ²Shanghai Key Laboratory of Perioperative Stress and Protection, Zhongshan hospital, Fudan University, Shanghai, China, ³Department of Pathology, Obstetrics and Gynecology Hospital, Fudan University, Shanghai, China, ⁴Department of Diagnostic and Interventional Radiology, University of Tsukuba, Ibaraki, Japan, ⁵Department of Diagnostic Radiology and Nuclear Medicine, Gunma University Graduate School of Medicine, Maebashi, Japan

Colorectal cancer (CRC) is one of the most fatal cancers of the digestive system. Although cancer stem cells and metabolic reprogramming have an important effect on tumor progression and drug resistance, their combined effect on CRC prognosis remains unclear. Therefore, we generated a 21-gene mRNA stemness index-related metabolic risk score model, which was examined in The Cancer Genome Atlas and Gene Expression Omnibus databases (1323 patients) and validated using the Zhongshan Hospital cohort (200 patients). The high-risk group showed more immune infiltrations; higher levels of immunosuppressive checkpoints, such as CD274, tumor mutation burden, and resistance to chemotherapeutics; potentially better response to immune therapy; worse prognosis; and advanced stage of tumor node metastasis than the low-risk group. The combination of risk score and clinical characteristics was effective in predicting overall survival. Zhongshan cohort validated that high-risk score group correlated with malignant progression, worse prognosis, inferior adjuvant chemotherapy responsiveness of CRC, and shaped an immuno-evasive contexture. This tool may provide a more accurate risk stratification in CRC and screening of patients with CRC responsive to immunotherapy.

KEYWORDS

colorectal cancer, mRNAsi, stemness, risk score model, immunotherapy, metabolism, immune evasion, Machine learning

Introduction

Colorectal cancer (CRC) is one of the deadliest cancers of the digestive system (1, 2). Although there is an increasing number of potential therapeutic approaches for CRC, such as surgery, chemotherapy, radiotherapy, and molecular targeted therapy, the clinical prognosis remains unsatisfactory, especially for patients with distant metastasis of CRC (3, 4). Therefore, accurate medical treatment is essential for the eradication of malignancy. At the same time, due to the high molecular heterogeneity of CRC, most existing biomarkers lack strong predictive accuracy (5). Hence, it has become an urgent problem to find a powerful index to predict and evaluate the clinical prognosis and therapeutic effect to achieve accurate clinical intervention.

Cancer stem cells (CSCs) play a crucial part in the progression, recurrence, and drug resistance of solid malignant tumors (6). Furthermore, CSCs promoted immunosuppression, immune escape, tumor metastasis, and therapeutic resistance by interacting with immune cells (7). For example, in a co-transplantation environment, CSCs can promote the polarization of CD14⁺ peripheral monocytes into immunosuppressive M2 macrophages and the generation of tumorigenic myeloid cells, followed by the acceleration of tumor growth in immunocompromised mice (8). CSCs also drive the recruitment and polarization of TH17 cells and Treg cells by secreting CCL1, CCL2, CCL5, and TGF- β , resulting in an immunosuppressive environment (7). In recent years, the mRNA expression-based stemness index (mRNAsi) developed by a machine learning algorithm has been used to quantify the stemness characteristics of tumors (9) such as esophageal cancer (10), gastric cancer (11), hepatocellular carcinoma (12), and glioma (13). However, the risk score model for stemness features associated with immunological properties in CRC remains uninvestigated.

Metabolic reprogramming, a hallmark of cancer, is another important factor leading to antitumor immunity and immune escape. For example, excessive glycolysis in tumor cells produces a large amount of lactate, which leads to acidification of the microenvironment and, consequently, inhibits the proliferation and function of cytotoxic T cells. Other studies also showed that inhibition of mTOR or the glycolysis pathway regulated T-cell differentiation into naive and memory phenotypes (14). Furthermore, when CAR-T cells were expanded *in vitro*, inhibition of AKT improved their metabolism and promoted their differentiation to the memory phenotype, thus improving the progression of acute lymphoblastic leukemia (15). Therefore, further elucidation of the effects of tumor stemness and metabolic characteristics on the immune microenvironment may provide significant clinical benefits.

In this study, we generated a new risk prediction model using mRNAsi and metabolism-related genes using CRC expression data

retrieved from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases (n = 1323). We also determined the association between the risk score and several functional and clinical features of patients with CRC. Clinical prognosis, tumor microenvironment and immunophenotype, response to chemotherapy and immunotherapy, and genomic variation between two risk score groups were evaluated comprehensively. We then validated the mRNAsi-related metabolic risk score model using the Zhongshan Hospital cohort (n = 200). Our data aimed to provide new insights into the screening of patients more likely to benefit from immunotherapy, and to improve individualized treatments for CRC patients.

Methods

Data collection and processing

The expression profile data of colon adenocarcinoma (COAD) and rectal adenocarcinoma (READ) of 591 patients, and their clinicopathological annotation were retrieved from the TCGA GDC website (<https://portal.gdc.cancer.gov/>). TPM values were converted from FPKM. Furthermore, tumor mutation burden (TMB) was obtained by analyzing the copy number variation (CNV) and somatic mutation data using the *maftools* package of R.

In addition, CRC gene expression data of GSE17536 (16–18) and GSE39582 (19) and the clinicopathological features of the patients were also downloaded from the GEO database. GSE17536 included 177 CRC tissue samples, and GSE39582 included 555 CRC tissue samples. Subsequently, the TCGA and GEO data were merged (n = 1323), and the *limma* (20) R package and *sva* (21) R package were used to combine and eliminate any batch effect.

Analysis of mRNAsi and differentially expressed genes

Based on the relative expression data provided by Zheng et al. (22), the mRNAsi of each sample was determined conforming to the gene expression matrix by the single-sample gene-set enrichment analysis (ssGSEA) method using the R-package GSVA. According to the mRNAsi of each sample obtained, combined with the survival status of the patients, the best cut-off value of mRNAsi was set, and the patients with CRC were distributed into the high-mRNAsi group and low-mRNAsi group.

DEGs between the high-mRNAsi groups and low-mRNAsi groups in patients with CRC were analyzed using the “*limma*” R package. DEGs were defined as genes with Log₂ (fold change) > 1.0 and *P* < 0.05. Metabolism-related gene sets were copied from

the Molecular Signature Database (MSigDB) V7.0 (22). Finally, the overlap between DEGs and metabolism-related genes resulted in the identification of metabolism-related DEGs.

Weighted gene co-expression network analysis

WGCNA was achieved using the WGCNA package in R, which aims to determine the correlation between genes by building important modules. First, a scale-free gene co-expression network was constructed according to the weight of the correlation coefficient, and a hierarchical clustering tree was established depending on the adjacency matrix of the network. The module significance (MS) was then calculated to judge the correlation between the mRNAsi value and different modules. The genes in each module were recorded and defined as module characteristic genes. Modules with maximum and minimum MS values were regarded as positive and negative modules, respectively. After selecting the modules of interest according to the MS values, all gene expressions in the modules were identified as genes highly correlated with mRNAsi.

Construction of mRNAsi-related metabolic risk score model

By integrating the results of metabolism-related DEG analysis and WGCNA, mRNAsi-related metabolic genes were finally obtained. Significantly differentially expressed mRNAsi-related metabolic genes were included in the model, dimensionality reduction analysis was performed using the minimum absolute contraction and selection operator (least absolute shrinkage and selection operator, LASSO) algorithm, and the characteristic genes related to prognosis were obtained. Using the normalized gene expression value weighted by the penalty coefficient obtained by LASSO Cox analysis, a risk score formula was established, and patients were divided into high-risk group and low-risk group according to the median risk score.

$$\text{Risk score} = \sum_i \text{Coefficient}(\text{hub gene}_i) \times \text{mRNA Expression}(\text{hub gene}_i)$$

Functional and pathway enrichment analyses

Gene Ontology (GO) analysis is a widely used method for functional enrichment studies and generates data related to biological processes (BP), molecular functions (MF), and cellular components (CC). Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database for systematic analysis of gene

function, linking genome information with more orderly biological function information. The clusterProfiler package of R (23) was used for GO analysis and KEGG pathway enrichment in the mRNAsi-related metabolic risk score model. FDR < 0.05 was regarded as significant.

To investigate differences in BP between different groups, we employed gene-set enrichment analysis (GSEA) (22). The “h.all.v7.2.symbols.gmt” gene set was copied from the MSigDB for the GSEA. $P < 0.05$ was considered significant.

Molecular network analysis

The STRING database (<https://cn.string-db.org>) (24) was used to construct a protein-protein interaction (PPI) network. Genes with scores greater than 0.4 were chosen to build a network model, which was visualized using Cytoscape (v3.7.2) (25). Then, eight hub genes were selected using the CytoHubba plug-in (26) in the Cytoscape software. Furthermore, we use the GOSemSim package in R (27) to judge the GO semantic similarity of the eight genes (28).

Information regarding miRNA-mRNA interactions from the miRTarBase database was downloaded before analyzing the basic statistics. Based on the core mRNA obtained by PPI analysis, the miRTarBase database was used to predict the miRNAs that may be regulated and to further predict the related lncRNAs. Cytoscape software was used to visually display the results of ceRNA analysis.

Analysis of tumor immune infiltrating cells

An ssGSEA algorithm was deployed to measure the relative number of tumor-infiltrating immune cells in patients with CRC (29). The enrichment score calculated by ssGSEA using the GSVA R package (30) indicates the level of each immune cell type in each sample. In addition, depending on the gene expression profile, the ESTIMATE R package (31) was used to quantify the level of immune infiltration of tumor samples, and the immune score of each tumor sample was obtained. The differences in the immune infiltration characteristics of CRC patients between the high-risk score group and low-risk score groups were evaluated.

Analysis of drug sensitivity and immunotherapy response

The Genomics of Drug Sensitivity in Cancer (GDSC) (<https://www.cancerrxgene.org/>) is an open database for molecular therapy and mutation exploration in cancer. The pRRophetic package of R (32) was used to download the cell

line gene mutation data and the IC50 values of different anticancer drugs from GDSC (33) and to analyze the correlation between patients with high and low-risk scores and different anticancer drug sensitivities.

In addition, we used online tumor immune dysfunction and exclusion (TIDE) scores (34) to examine immunotherapy sensitivity and compare the scores of tumor immunotherapy markers, such as CD8 and CD274, between the high-risk score groups and the low-risk scoring groups. The response of immune-checkpoint blockade was predicted.

CNV analysis

To analyze the changes in copy number in different risk score groups of patients with TCGA-CRC, we used the TCGAbiolinks package of R to obtain the masked copy number segment data of the patients. The downloaded CNV fragments were analyzed using GISTIC 2.0, with default settings in GenePattern5. Finally, the analysis results of GISTIC 2.0 were visualized through the maftools package of R.

Establishment of a prognostic model

Univariate and multivariate Cox analyses were used to predict the overall survival (OS) of patients with CRC. The clinicopathological features were then incorporated into the risk score model to construct a clinical predictive nomogram. To quantify the differential performance of the nomogram, Harrell's consistency index (C-index) was estimated. A calibration curve was produced and the capability of the nomogram was evaluated.

Patients and CRC tissue samples

The Zhongshan Hospital cohort included 200 patients who underwent CRC surgery between January 2008 and December 2014. The patients' baseline characteristics included sex, age, adjuvant chemotherapy, tumor location, tumor histology, tumor differentiation, nerve invasion, surgical margin positivity, and stage of tumor node metastasis (TNM). Tumor staging was performed according to the 7th edition of the American Joint Commission on Cancer (AJCC) TNM Classification (35). Conforming to the National Comprehensive Cancer Network guidelines and patient wishes, patients with stage III-IV TNM were treated with ACT after surgery. OS was described as the time from the date of diagnosis to death or last follow-up. Disease-free survival (DFS) was described as the time from the date of diagnosis to relapse or last follow-up. The follow-up period ended on December 31, 2020. Clinical data validation was approved by the ethics committee of the Zhongshan Hospital (B2022-068R2).

RNA separation and quantitative reverse transcription PCR

The mRNA expression of 21 mRNAs-related metabolic genes was measured by qRT-PCR in Zhongshan cohort. Total RNA was obtained using the TRIzol reagent (Invitrogen, Waltham, MA, USA). cDNA was obtained by reverse transcription using the PrimeScript RT kit (Takara). The expression of candidate genes and the housekeeping gene GAPDH was evaluated by quantitative reverse transcription PCR using the ABI 7900HT real-time PCR system (Applied Biosystems, Carlsbad, CA, USA). Relative transcription levels were calculated using the $\Delta\Delta C_t$ method (36). The primer sequences used are listed in [Supplementary Table 7](#).

Immunohistochemical staining

We randomly selected 20 cases from 200 Zhongshan patients for IHC, including 10 cases in high risk group and 10 cases in low risk group. Paraffin-embedded tissues were stained with antibodies. The staining score was decided by two experienced pathologists at the Zhongshan Hospital. Six high-power fields (HPFs, $\times 200$ magnification) were randomly counted by two independent pathologists (each with three fields), and the densities of CD8⁺T cells, Foxp3⁺Tregs, CD19⁺B cells, CD11c dendritic cells, immunosuppressive checkpoints (PD-1, PD-L1) and effector molecules (GZMB, PRF1) were recorded. Immunohistochemistry antibodies are listed in [Supplementary Table 8](#).

Statistical analysis

All data processing and analyses were accomplished using the R software (version 3.6.2) and SPSS (version 25; IBM, Armonk, USA). For the comparison of two groups of continuous variables, the statistical significance of normally distributed variables was calculated using an independent *t*-test, and the difference between non-normally distributed variables was measured using the Mann-Whitney U test. Chi-square test or Fisher's exact test was used to analyze the significant differences between the two groups of classified variables. The survival package in R was conducted for survival analysis. The receiver-operating characteristic (ROC) curve was drawn by the pROC package of R (37) and the area under the curve (AUC) was calculated to evaluate the performance of the risk score model. Univariate and multivariate Cox analyses were used to determine independent prognostic factors. All statistical P values were bilateral, and $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ were regarded as statistically significant.

Results

Relationship between colorectal cancer stemness characteristics and clinical features

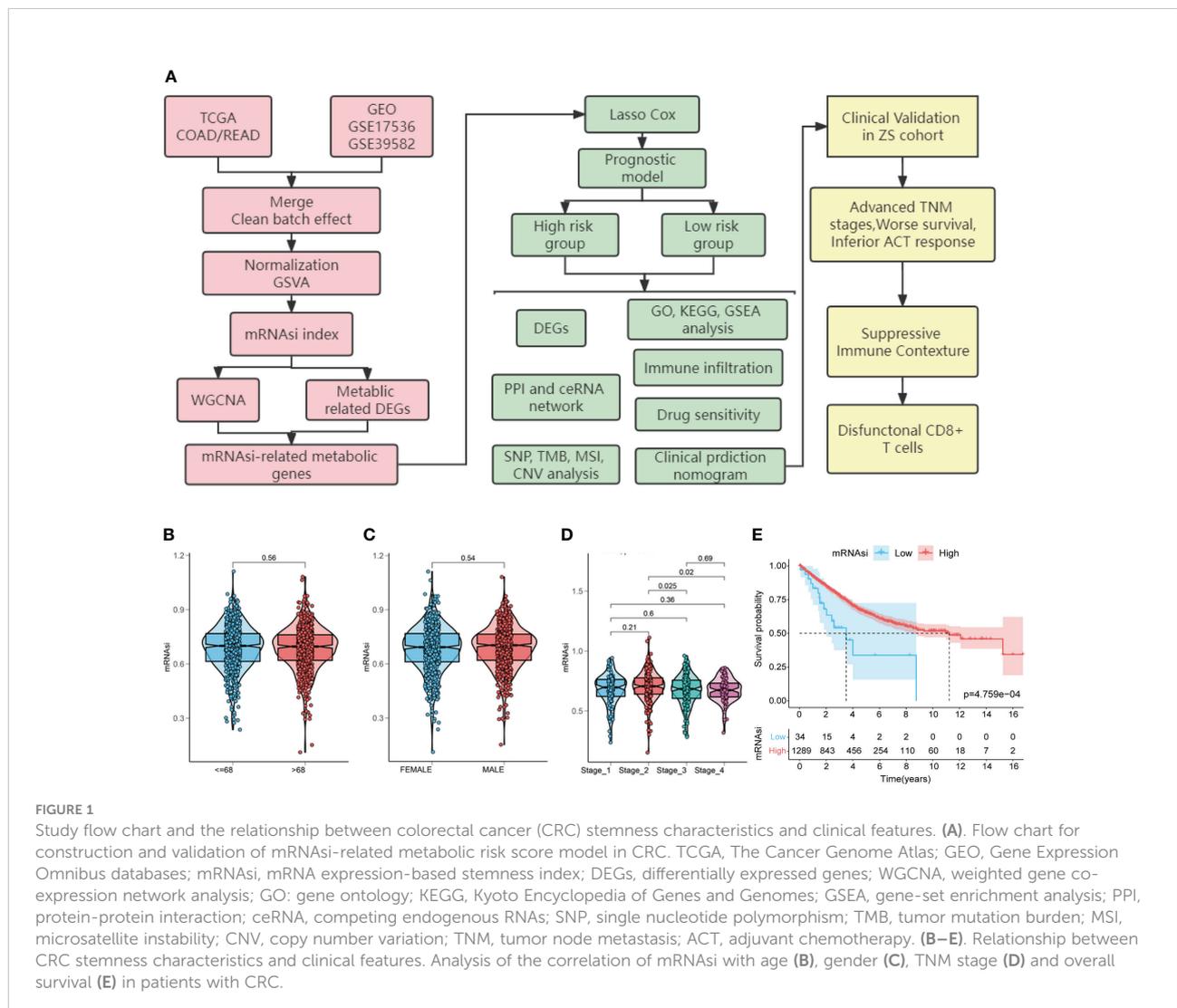
A flowchart of this study is shown in **Figure 1A**. To explore the role of mRNasi on the progression of CRC, including COAD and READ, the gene expression matrices of GSE17536 and GSE39582 datasets and TCGA database were downloaded (**Supplementary Figures 1A, B**). The data from the two databases were then merged ($n = 1323$) and cleaned from any batch effect (**Supplementary Figures 1C, D**).

First, to explore the correlation between mRNasi and clinical characteristics, we determined CRC mRNasi using the ssGSEA algorithm. Then, according to the optimal mRNasi cut-off value, the patients with CRC were separated into high-

mRNasi and low-mRNasi groups. The relationship between CRC stemness characteristics and clinical characteristics is shown in **Figures 1B–E**. No significant correlation between mRNasi and age ($P = 0.56$) or gender ($P = 0.54$) was observed (**Figures 1B, C**). However, higher mRNasi were associated with staging of TNM (stage 2 vs. stage 3, $P = 0.025$; stage 2 vs. stage 4, $P = 0.02$; **Figure 1D**). Furthermore, patients with high mRNasi showed a significant increase in OS compared to those with low mRNasi (log-rank $P < 0.001$, **Figure 1E**).

Identification of mRNasi-related metabolic genes in patients with CRC

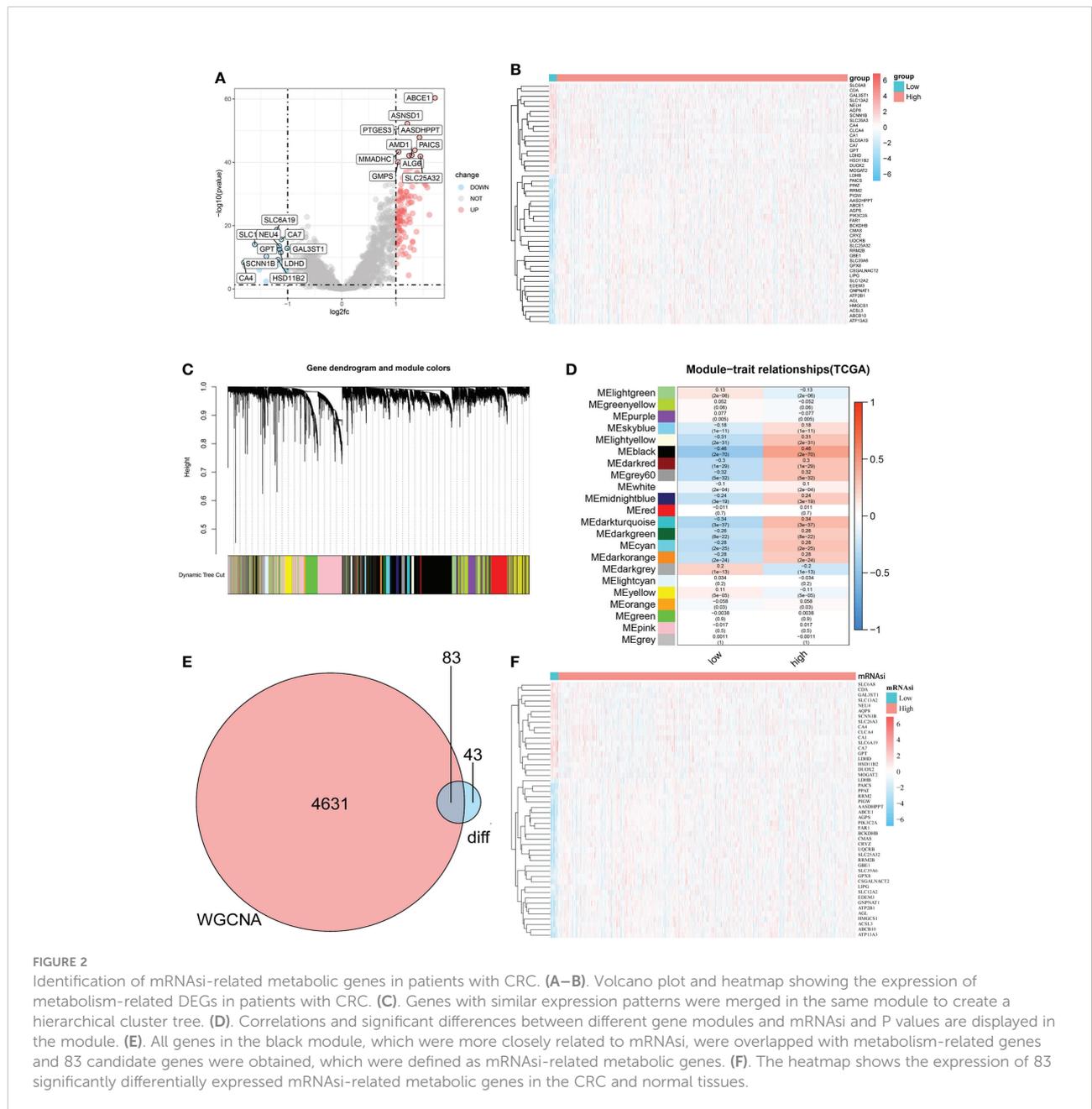
To determine the role of the mRNasi in metabolic processes in CRC, DEGs between the high-mRNasi and low-mRNasi groups were identified and intersected with a metabolic gene set



(2752 genes). One hundred and twenty-six genes were obtained and labeled as metabolism-related DEGs, of which 108 genes were significantly upregulated, and 18 genes were significantly downregulated (Figures 2A, B and Supplementary Figure 3A). WGCNA was used to identify modules closely related to mRNAsi-related genes. A total of 22 co-expression modules were identified, with the black module showing the strongest correlation with mRNAsi in CRC (Figures 2C, D). All genes in the black module were intersected with the metabolism-related DEGs, and 83 mRNAsi-related metabolic genes were obtained for further analysis, as shown in the Venn diagram (Figures 2E, F and Supplementary Figure 3B).

Construction of mRNAsi-related metabolic risk score model

To quantitatively evaluate the predictive value of identified mRNAsi-related metabolic genes in the clinical prognosis of CRC, we constructed a risk score model based on these genes. First, the expression characteristics of the 83 mRNAsi-related metabolic genes were included in the LASSO Cox analysis and 21 genes with the optimal predictive value were selected (Figures 3A, B). Simultaneously, a risk score formula was established based on the normalized expression of important



characteristic genes weighted by the penalty coefficient calculated by LASSO Cox analysis, and a risk score for each sample was calculated. An example of the formula used to calculate the risk score is given below.

$$\begin{aligned} \text{Risk score} = & (-0.1053) \times \text{PTGES3} + (-0.1874) \times \text{PAICS} + \\ & (-0.0133) \times \text{GNPNAT1} + \\ & (0.02893) \times \text{PGM3} + (0.08862) \times \text{MTHFD2} + (-0.0043) \times \\ & \text{DCK} + (0.06428) \times \text{MTAP} + \\ & (0.22468) \times \text{SLC25A36} + (0.00308) \times \text{GBE1} + (-0.0679) \times \\ & \text{RRM2} + (0.0029) \times \text{KCTD3} \\ & + (-0.0436) \times \text{ACADSB} + (-0.0339) \times \text{ABCD3} + (-0.0137) \times \\ & \text{BCKDHB} + (-0.0525) \times \\ & \text{PHOSPHO2} + (-0.1185) \times \text{FUT4} + (0.00089) \times \text{EDEM3} + \\ & (-0.0684) \times \text{NEU4} + (0.6165) \\ & \times \text{SLC16A1} + (-0.0162) \times \text{ELOVL7} + (0.04312) \times \text{SLC6A8} \end{aligned}$$

Then we performed the time-dependent ROC curve analysis and found that the model had appropriate accuracy in predicting OS in patients with CRC, and the AUC of 1-year, 2-year and 3-year OS was 0.647, 0.644, and 0.672, respectively (Figure 3C). Kaplan–Meier analysis showed a reduction in OS in patients with high-risk scores (log-rank $P < 0.001$; Figure 3D). In addition, there was a significant negative correlation between mRNasi and risk scores ($Rho = -0.2$, $P < 0.001$, Figure 3E). The distribution of the risk score, survival status, and expression pattern of characteristic genes is shown in Figure 3F.

GSEA, GO, KEGG analyses of DEGs between high-risk and low-risk patients in mRNasi-related metabolic risk score model

To analyze the impact of mRNasi-related metabolic risk score models on the occurrence and development of CRC, we

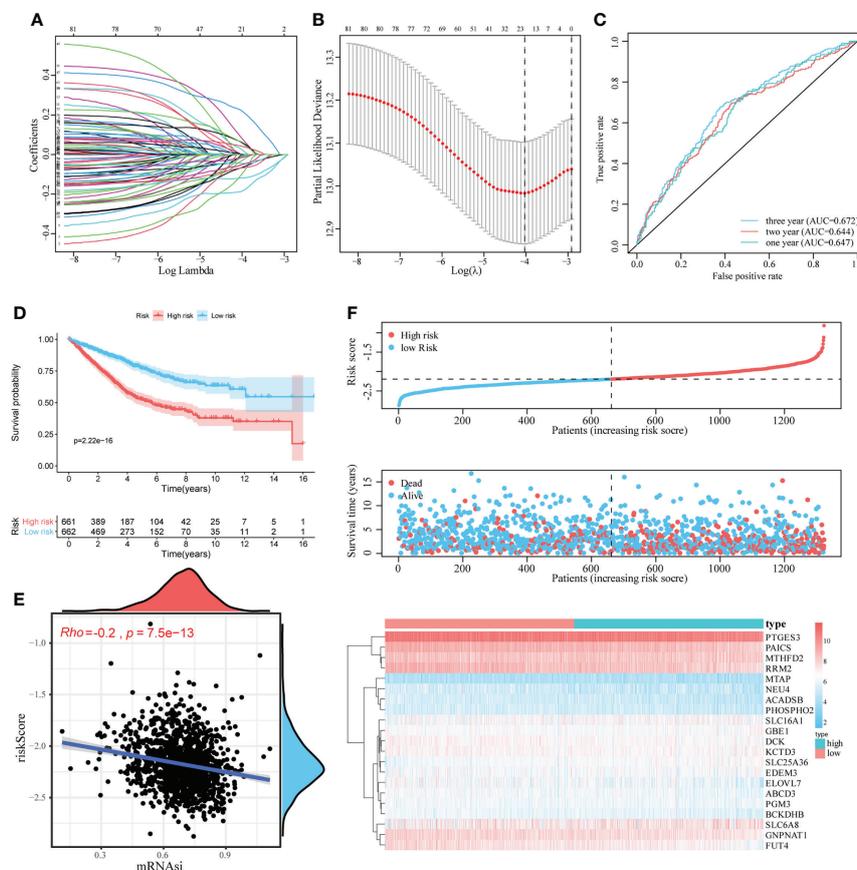


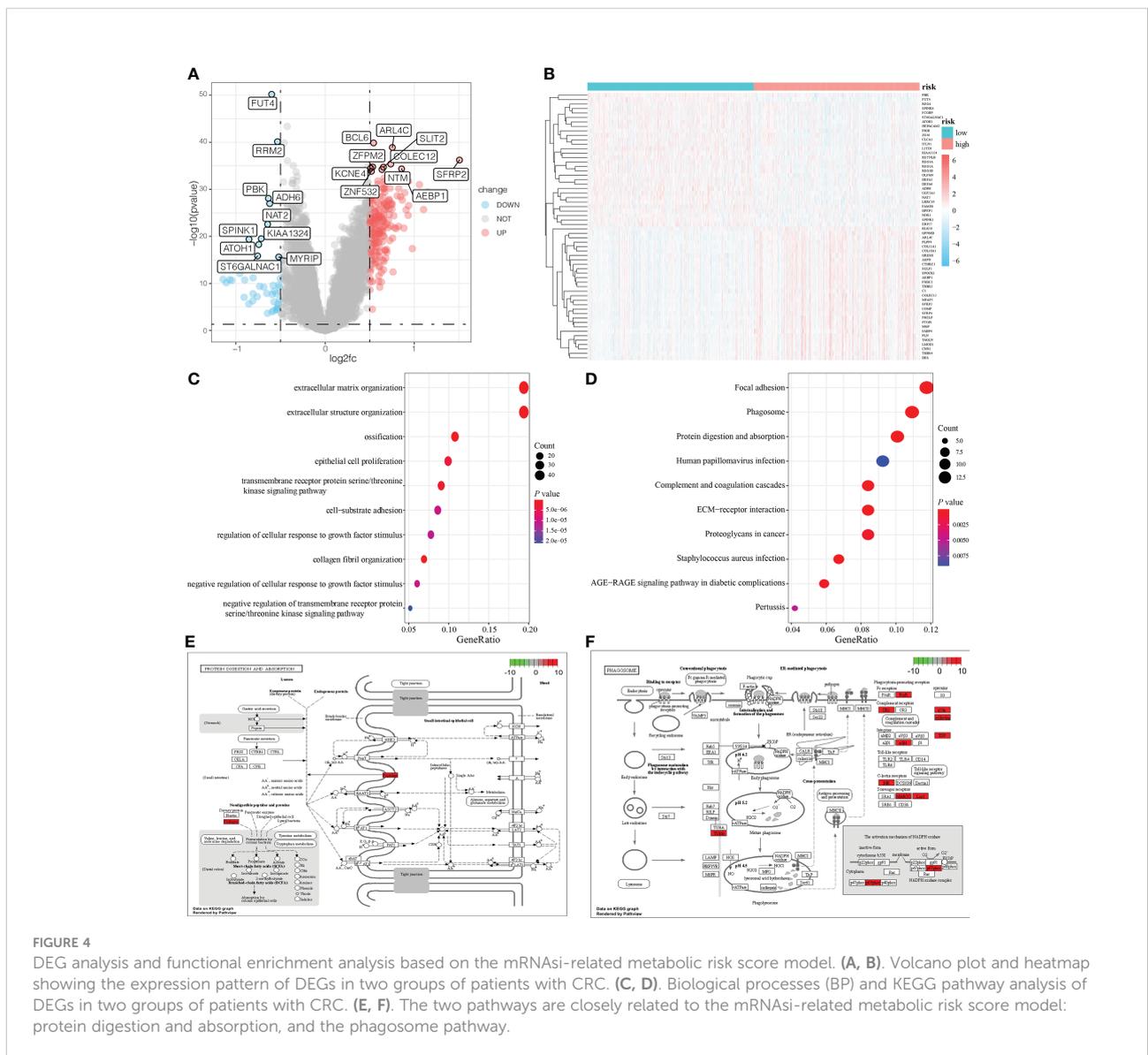
FIGURE 3

Construction and evaluation of the mRNasi-related metabolic risk score model. (A, B). LASSO Cox analysis identified 21 genes most associated with OS in the TCGA dataset. (C). Time-dependent ROC curve analysis of risk score. (D). The effect of the risk score assessed by Kaplan–Meier curve on the overall survival rate of patients with CRC. (E). Spearman rank correlation analysis was used to analyze the relationship between mRNasi and risk score. (F). The risk score distribution, survival status, and heatmap of characteristic gene expression in patients with CRC.

used the median LASSO Cox risk score of CRC cases from TCGA dataset and divided the CRC cases into high-risk and low-risk score groups. There were 242 DEGs between high-risk and low-risk patients ($\text{Log}_2(\text{fold change}) > 1.0$ and $P < 0.05$), of which 195 were significantly upregulated, and 47 were significantly downregulated (Figures 4A, B). The correlation between the risk score with the clinical characteristics of CRC patients in the TCGA and GEO database is shown in Supplementary Table 1. The functional annotation of the GO showed the DEGs were closely related to several BP, including the organization of the extracellular matrix, the organization of the extracellular structure and ossification, as well as several MF such as extracellular matrix structural constituent, glycosaminoglycan binding, and extracellular matrix structural constituent conferring tensile strength (Figure 4C,

Supplementary Table 2). KEGG analysis indicated that DEGs were particularly involved in focal adhesion, phagosome, protein digestion and absorption, complement and coagulation cascades, and ECM-receptor interaction pathways (Figure 4D and Supplementary Table 3). Two pathways, protein digestion and absorption ($P = 2.82E^{-08}$) and phagosome pathways ($P = 2.96E^{-07}$), which were highly related to the mRNAi-related metabolic risk score model, are shown in Figures 4E, F.

Furthermore, GSEA showed that ascorbate and aldarate metabolism ($\text{NES} = -1.89$, $P = 0.002$), citrate cycle (TCA cycle) ($\text{NES} = -2.28$, $P = 0.002$), glyoxylate and dicarboxylate metabolism pathways ($\text{NES} = -2.03$, $P = 0.002$), propanoate metabolism ($\text{NES} = -2.07$, $P = 0.002$), arginine and proline metabolism ($\text{NES} = -1.94$, $P = 0.002$), pyruvate metabolism ($\text{NES} = -2.18$, $P = 0.002$), hallmark oxidative



phosphorylation (NES = -2.86, $P < 0.001$), and hallmark fatty acid metabolism (NES = -1.93, $P < 0.001$), were abundant in low-risk patients, whereas hallmark hypoxia was significantly enriched in high-risk patients (NES = 2.04, $P < 0.001$), (Supplementary Figures 2A–I and Supplementary Table 4).

Construction of PPI network and related regulation network

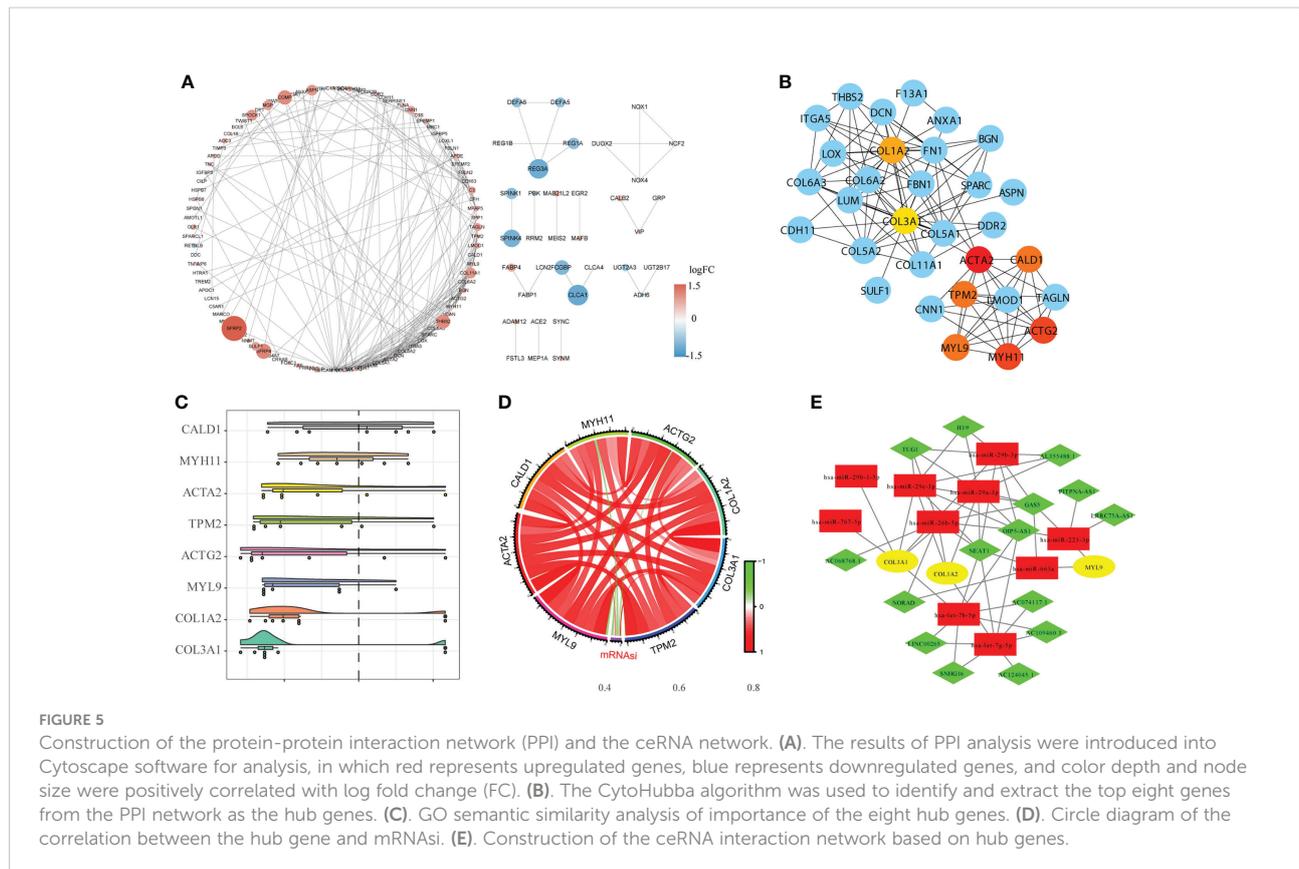
We used the STRING database to establish the PPI network between DEGs, and imported the interaction between genes into Cytoscape software to obtain Figure 5A, in which the upregulated genes were represented in red and the downregulated genes were represented in blue.

The hub genes were analyzed using Cytoscape software (Figure 5B). GO semantic similarity analysis showed that the *CALD1* gene played an important role in the hub genes (Figure 5C). Subsequently, we conducted correlation analysis between hub genes and mRNAsi, and found a significant co-expression pattern between hub genes, whereas the relationship of each hub gene and mRNAsi was not consistent (Figure 5D). Finally, based on information about miRNA-mRNA interaction downloaded from the miRTarBase database; the hub genes

obtained *via* the PPI network were used to construct the ceRNA network of miRNA-mRNA-lncRNA interaction (Figure 5E).

Immune contexture difference between high-risk and low-risk patients

We then evaluated the immune contexture heterogeneity between the high and low-risk score groups. As shown in Figure 6, the immune and stromal scores of the high-risk score group were significantly higher than those in the low-risk score group (both $P < 0.001$, Figures 6A, B). In addition, to evaluate the degree of immune cell infiltration in tumor tissue, we used the ssGSEA algorithm and obtained the relative enrichment scores of 28 subtypes of immune cells between the two groups, as shown in the heatmap in Figure 6C. The correlation analysis showed that the infiltration levels of most immune cells were positively correlated (Figure 6D). Further analysis revealed that infiltration of CD4⁺T cells, CD8⁺T cells, B cells, dendritic cells, eosinophils, mast cells, macrophages, myeloid-derived suppressor cells (MDSCs), natural killer cells, regulatory T cells, and T helper cells was higher in the high-risk score group (Figure 6E). In addition, in this study the expression of HLA family members and several immunotherapy-related target genes, such as *CD274* (*PD-L1*), *CTLA-4*, and *LAG-3*, was



elevated in the tumor environment of the high-risk group compared with the low-risk group (Figures 6F, G).

Sensitivity to chemotherapy and immunotherapy in high-risk and low-risk patients with CRC

To analyze the differences in the sensitivity of patients with CRC to different drugs and small-molecule drugs based on the risk score, we downloaded the CRC cell line gene mutation data and the half-maximal inhibitory concentration (IC50) values of several anticancer drugs from the GDSC database. In GDSC, IC50 values for patients with CRC were predicted based on the responses of cell lines to 138 chemotherapeutic agents and small-molecule anticancer agents. This suggested that patients in the

high-risk score group were less susceptible to multiple chemotherapeutic and small-molecule anticancer drugs, including Metformin, PF.4708671, Sorafenib, Mitomycin, Methotrexate, and gemcitabine (Figure 7A, all $P < 0.05$).

Because of the important role of immune-checkpoint inhibitor (ICI) therapy in tumors, we examined the sensitivity of two groups of patients with CRC to ICI therapy using the TIDE algorithm, which models two mechanisms of immune-evasion: T-cell dysfunction and reduced T-cell infiltration, to predict the immunotherapy response. As shown in Figure 7B, although no significant differences in the scores for two immune markers CD8 and CD274 between the high-risk and low-risk score groups was observed, the TIDE score in the high-risk score group was lower than that in the low-risk score group, suggesting a better response to the ICI therapy in the high-risk score group than in the low-risk score group.

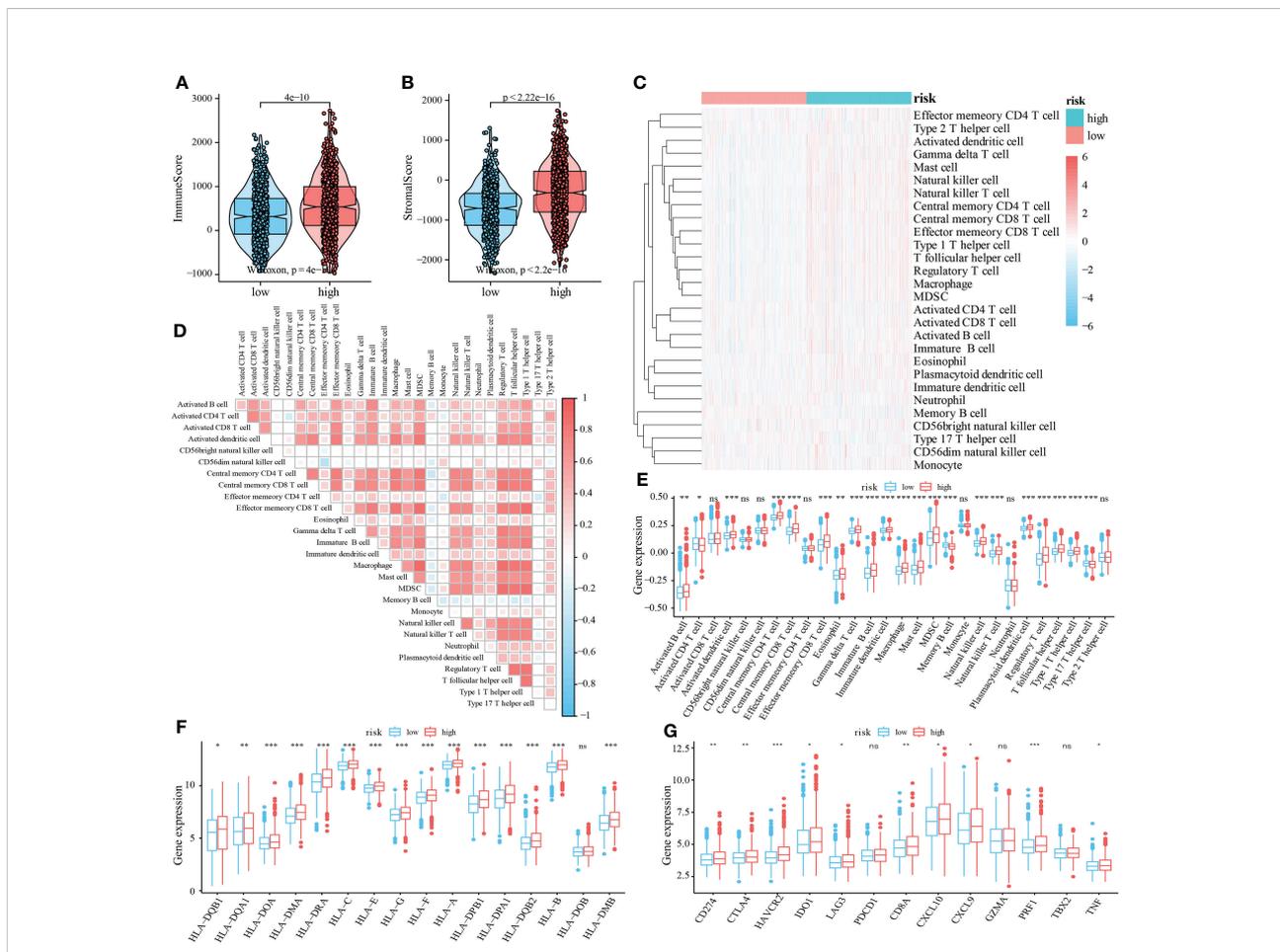


FIGURE 6 Relationship between mRNAsi-related metabolic risk score groups and infiltration of different immune cell subtypes. (A, B) Differential analysis of immune scores and stromal scores between the high- and low-risk score group of patients with CRC. (C) The heatmap showed the infiltration levels of 28 immune cell subtypes in CRC samples from TCGA and GEO datasets. (D) Correlation heatmap showed the correlation between different levels of immune cell infiltration. (E) Analysis of the difference of 28 levels of immune cell infiltration between two groups (F). Multiple HLA family members, and (G) immunotherapy-related targets in high and low-risk score groups of patients with CRC. Differences were considered significant at $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, compared to the low-risk group. ns, not significant.

Analysis of genomic variation between high and low-risk score patients

Research has suggested that genomic variation affects tumor response to immunotherapy (38). Furthermore, we evaluated the differences in genomic variation in patients with CRC in the high and low-risk groups, including single nucleotide polymorphism (SNP), TMB, microsatellite instability (MSI), and CNV.

Difference in the level of the top SNP between the two groups was detected (Figure 8A). Furthermore, the TMB in the high-risk score group was higher than in the low-risk score group; however, no significant differences in MSI between the two groups were detected (Figures 8B, C). In addition, compared with the high-risk score group, the low-risk score group showed a significant increase in CNV, mainly characterized by deletion events (Figures 8D, E).

Construction and validation of clinical prediction nomogram based on mRNAsi-related metabolic risk score model

Next, we evaluated the association between mRNAsi-associated metabolic risk scores and clinicopathological characteristics in patients with CRC. The results showed no significant correlation between the risk scores and the age and gender of the patients (Figures 9A, B). However, high-risk scores were associated with lower mRNAsi and an advanced TNM state (Figures 9C, D).

In addition, univariate and multivariate Cox analyses showed that a high mRNAsi-related metabolic risk score was an independent predictor of prognosis in patients with CRC (Figure 9E and Supplementary Table 5). The mRNAsi-related metabolic risk score was then combined with different

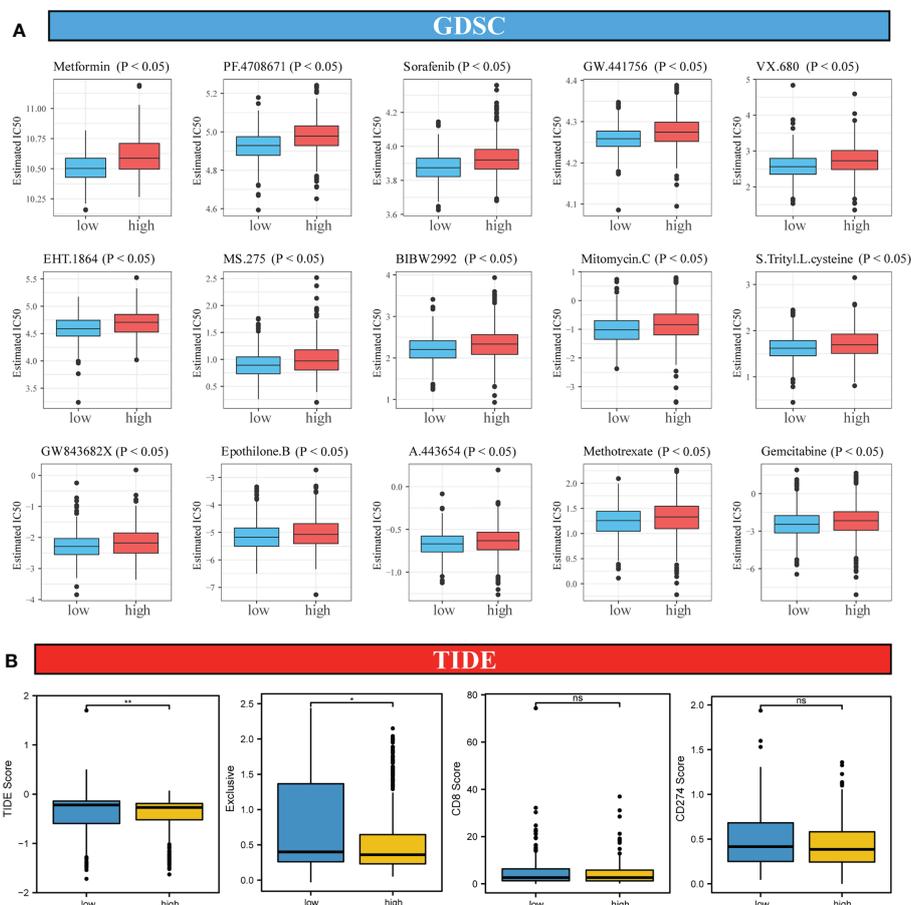


FIGURE 7

Analysis of sensitivity differences between high-risk and low-risk patients to different chemotherapeutic agents, small-molecule anticancer agents, and immunotherapy. (A). Difference in sensitivity between high-risk and low-risk patients to 138 small-molecule anticancer agents and chemotherapeutic agents. (B). Differences in TIDE score, immune exclusive, scores of immunotherapy targets, CD8 and CD274 in high-risk and low-risk score groups. Differences were considered significant at $*P < 0.05$, $**P < 0.01$, compared to the low-risk group. ns, not significant.

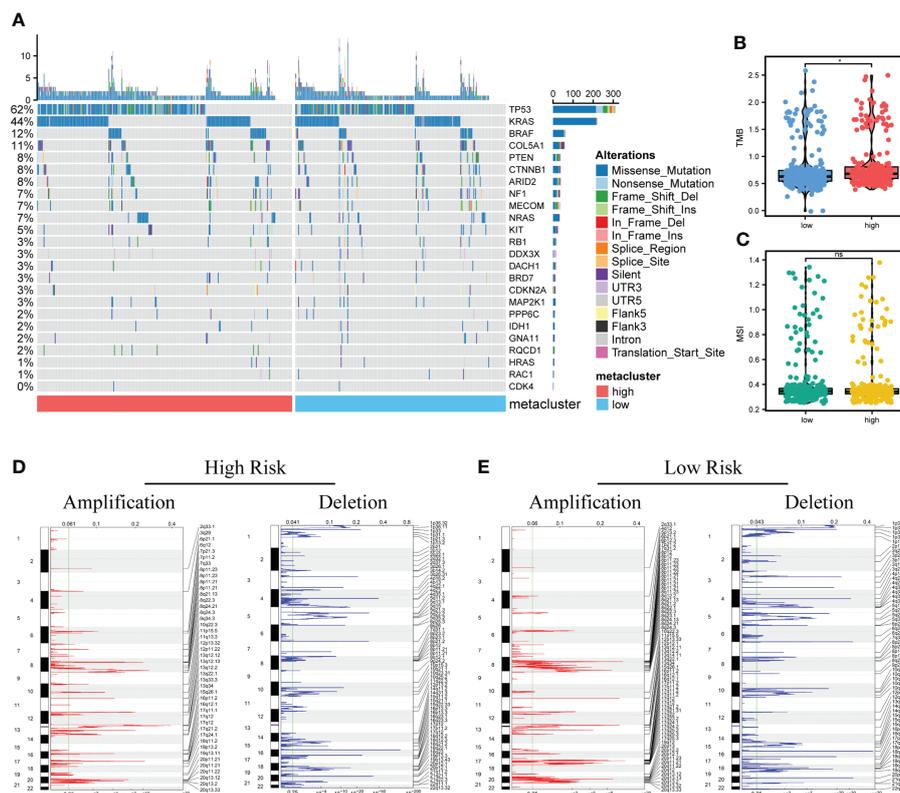


FIGURE 8

Analysis of genomic variation between high-risk and low-risk patients. (A) Mutation profiles of common tumor-related genes in patients in high and low-risk score groups (Left: high-risk score group, right: low-risk score group). (B, C) The difference of microsatellite instability (MSI) and tumor mutation burden (TMB) between two groups. (D, E) Copy number variation in patients between two groups. Red indicates the amplified genes, and blue indicates the deleted genes. Differences were considered significant at $*P < 0.05$, compared with low-risk group. ns, not significant.

clinicopathological features to create a nomogram to predict OS in patients with CRC (Figure 9F). Then we reformed C-index to evaluate the differentiation of nomogram and found it has high discriminative ability (mean: 0.751 [range: 0.700–0.802]). In addition, the calibration curve showed that the 1-, 2-, and 3-year OS estimated by the nomogram matched with the actual OS values of the patients (Figure 9G).

High-risk score group correlated with malignant progression, worse prognosis, inferior adjuvant chemotherapy responsiveness of CRC

To further determine the clinical significance of the risk score model in CRC, we evaluated the correlation between the high and low-risk score groups and the clinicopathological characteristics of patients with CRC in the Zhongshan Hospital cohort. The mRNA expression of 21 mRNAsi-related metabolic genes was measured by qRT-PCR. The median

expression level of risk score was used as the cutoff value. Patients were divided into high and low-risk score groups. The high-risk score group was positively correlated with right-sided colon, poorer differentiation, uroid adenocarcinoma and signet-ring cell carcinoma, nerve invasion, surgical margin positivity, and higher TNM stage (all $P < 0.001$, Supplementary Table 6). These findings suggest the gene set defining the high-risk score group is potentially involved in tumor progression.

To investigate the association between the risk score model and long-term outcomes of patients with CRC, Kaplan–Meier analysis was performed. The high-risk score group predicted worse survival of patients with CRC in the Zhongshan Hospital cohort (OS: $P < 0.001$, log-rank = 13.102; DFS: $P < 0.001$, log-rank = 26.309; Figures 10A, B). These results indicate that the high-risk score was related to a worse outcome for patients with CRC.

In addition, we evaluated the interaction between the risk score model and therapeutic responsiveness to adjuvant chemotherapy (ACT) for TNM stage III–IV patients with CRC. In this study, ACT could improve patient survival in the

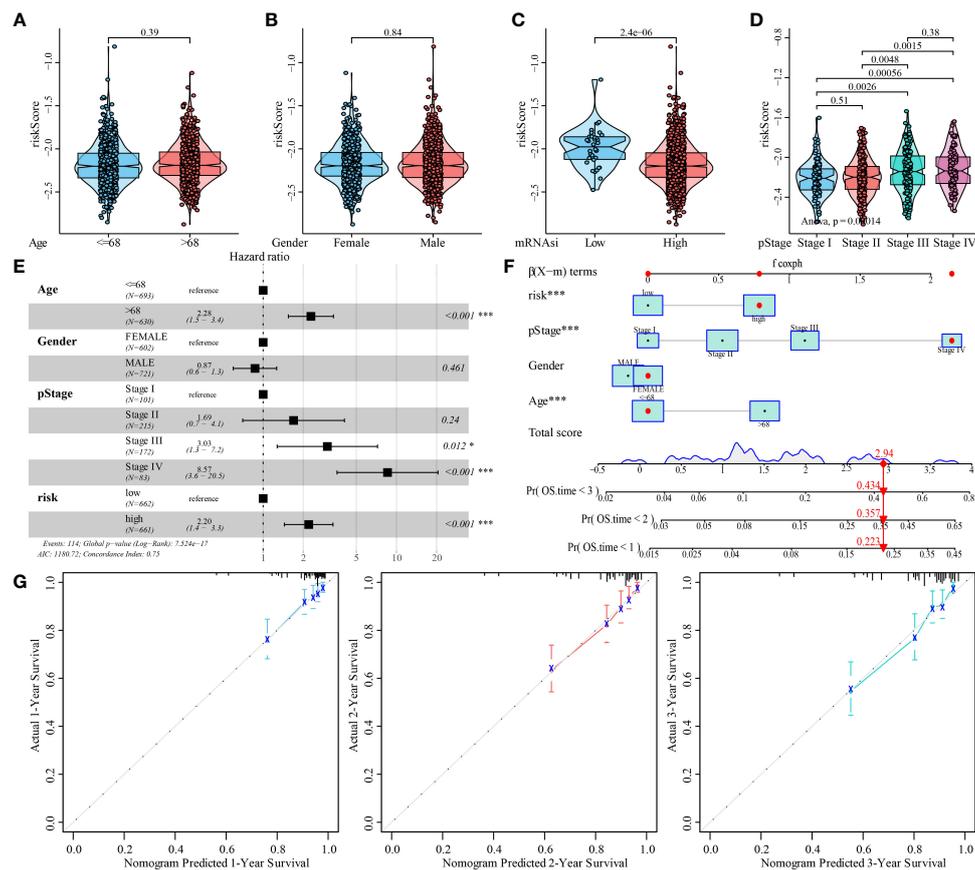


FIGURE 9

Analysis of the predictive ability of mRNA-related metabolic risk score model for the prognosis of patients with CRC. (A–D). Analysis of correlation between mRNA-related metabolic risk scores and clinicopathological features of patients with CRC. (E). Multivariate Cox regression analysis of HR and P values of risk score, combined with clinicopathological features. (F). mRNA-related metabolic risk score combined with clinicopathological features to construct a clinical predictive model. (G). The calibration curve of the nomogram showed that the risk score model had a good predictive ability for the overall survival rate of 1-, 2- and 3-year OS in patients. Differences were considered significant at $*P < 0.05$, $***P < 0.001$, compared to the reference.

low-risk score group (OS: $P = 0.043$, log rank = 4.094, Figure 10C; DFS: $P = 0.005$, log rank = 7.860, Figure 10D) but had no significant benefit in the high-risk score group (OS: $P = 0.209$, log rank = 1.579, Figure 10E; DFS: $P = 0.413$, log rank = 0.670, Figure 10F). Therefore, these results suggest that the high-risk score group might have impaired therapeutic responsiveness to ACT in TNM stage III-IV CRC.

High-risk score group shaped immunoevasive contexture

To explore the underlying mechanism, we performed IHC staining of tumor-infiltrating immune cells in CRC tissues obtained from the Zhongshan Hospital cohort. The number of CD8⁺T cells ($P = 0.0080$), CD19⁺B cells ($P = 0.0013$), Foxp3⁺Tregs ($P < 0.001$), and CD11c dendritic cells ($P = 0.0028$) was more

abundant in the high-risk group (Figures 11A–H). But the ratio of Foxp3⁺Treg cells to CD8⁺T cells also increased markedly in the high-risk score group ($P = 0.029$) (Figure 11I), suggesting a more immunosuppressive tumor microenvironment with increased Treg cell infiltration. We further investigated whether the high-risk score group could affect CD8⁺T-cell function. The results indicated that CD8⁺T cells in the high-risk score group showed an exhausted T-cell phenotype with increased expression of immunosuppressive checkpoints, programmed cell death protein 1 (PD-1) ($P = 0.0027$) and programmed cell death-ligand 1 (PD-L1) ($P = 0.0013$), and decreased expression of CD8⁺T-cell effector molecules, granzyme B (GZMB) ($P = 0.0028$) and perforin (PRF1) ($P = 0.0020$), compared to the low-risk score group (Figures 11J–Q). Taken together, these data suggest that the high-risk score group may orchestrate an immunoevasive contexture and direct CD8⁺T-cell dysfunction in CRC.

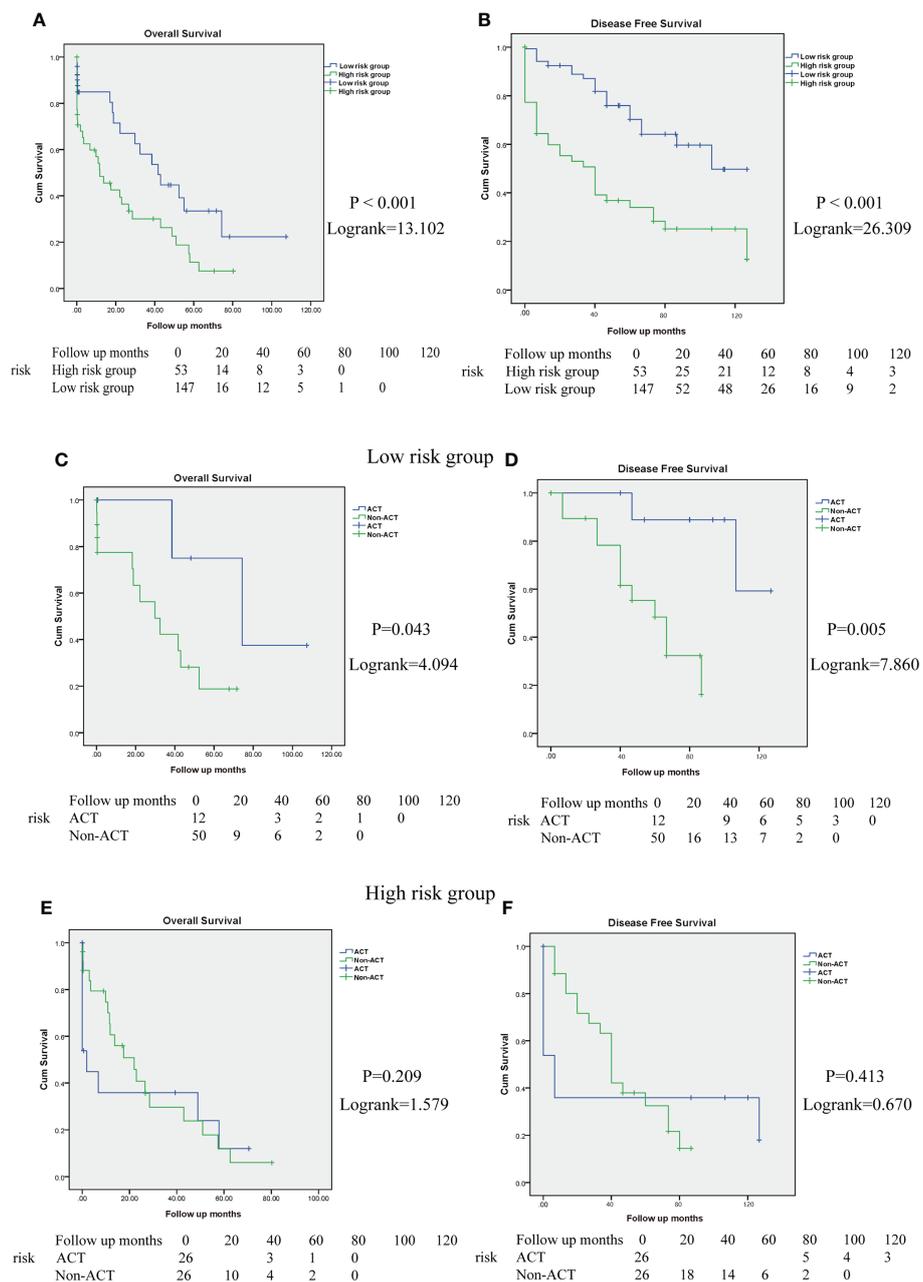


FIGURE 10

High-risk score group determines poor prognosis and impairs the ACT responsiveness of patients with CRC in Zhongshan cohort. (A, B). Overall survival (OS) and disease-free survival (DFS) curves between high and low-risk score group in Zhongshan Hospital cohort. (C, D). The OS and DFS curves for TNM stage III-IV patients with CRC in low-risk score group with or without ACT treatment. (E, F). The OS and DFS curves for TNM stage III-IV patients with CRC in high-risk score group with or without ACT treatment.

Discussion

Tumor recurrence and drug resistance have always been obstacles to the treatment of CRC. Studies have shown that CSCs and metabolic reprogramming promote immunosuppression, immune escape, and therapeutic resistance by interacting with

immune cells (39, 40). Therefore, by integrating differential expression analysis between high and low mRNasi, metabolism-related genes, co-expression network analysis, and LASSO Cox regression analysis, 21 mRNasi-related metabolic genes with the highest prognostic value were identified and used to construct the risk score model. After validation first in the

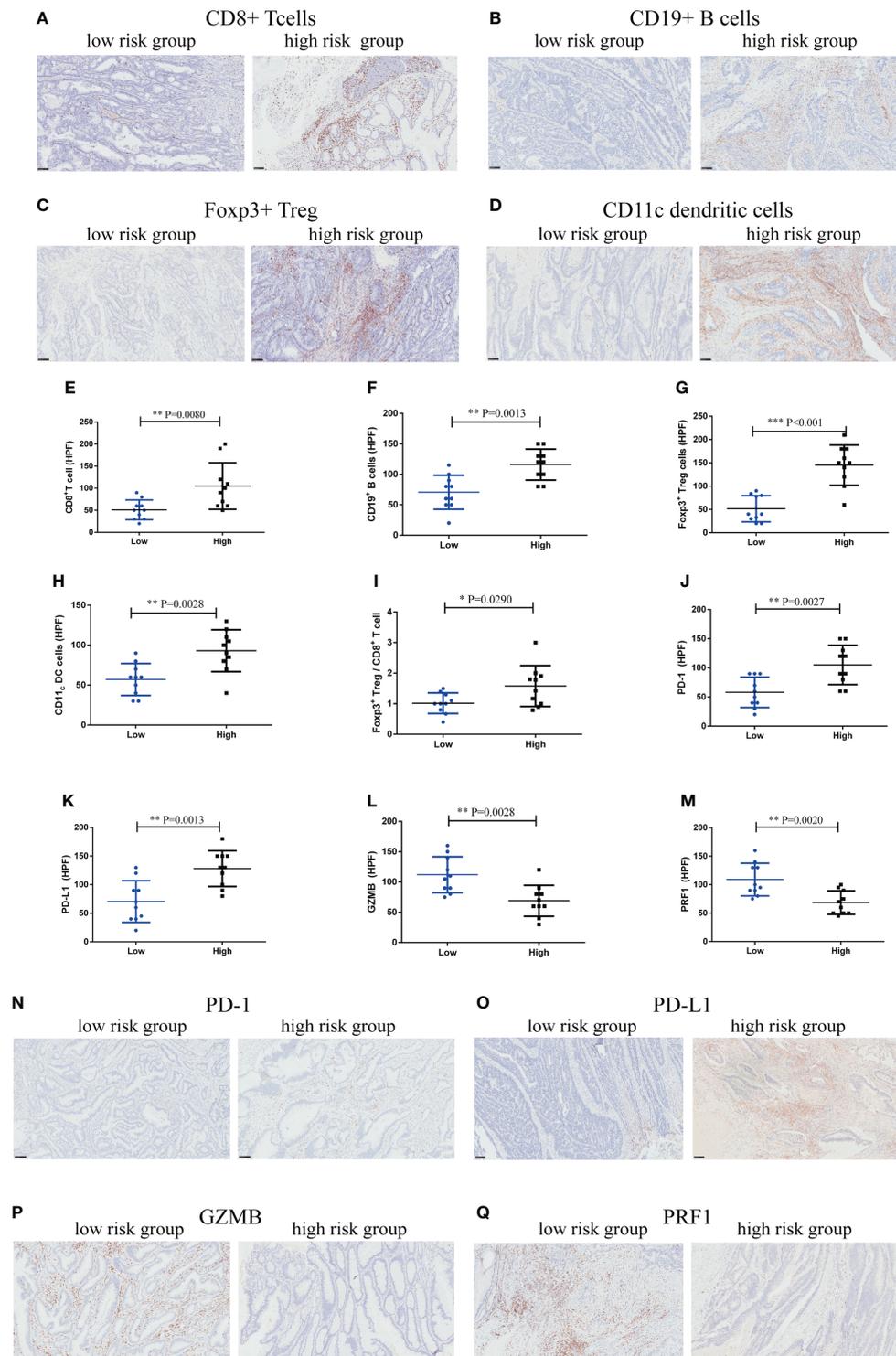


FIGURE 11

High-risk score group drives immunoevasive contexture and damages CD8⁺ T-cell function in CRC in Zhongshan cohort. (A–D). Representative immunohistochemical (IHC) staining of four significant tumor-infiltrating immune cell subtypes between high and low-risk score groups, including CD8⁺ T cells, CD19⁺ B cells, Foxp3⁺ Tregs, CD11c dendritic cells. (E–I). Comparison of CD8⁺ T cells, CD19⁺ B cells, Foxp3⁺ Tregs, CD11c dendritic cells and the ratio of Foxp3⁺ Tregs to CD8⁺ T cells between two groups. (J–M). Expression of immunosuppressive checkpoints (PD-1, PD-L1) and effector molecules (GZMB, PRF1) between two groups. (N–Q). Representative IHC staining of immunosuppressive checkpoints (PD-1, PD-L1) and effector molecules (GZMB, PRF1) between two groups. n = 10 in each group, scale bar: 250um. Differences were considered significant at *P < 0.05, **P < 0.01, ***P < 0.001, compared to the low-risk group.

data retrieved from the TCGA and GEO databases, and then in the Zhongshan Hospital cohort, the CRC samples in the high-risk score group exhibited poor clinical outcome, increased immune-evasion, reduced sensitivity to chemotherapy, whereas potentially better response to immunotherapy, and higher genomic variation. This risk score model could be a tool to screen for patients with worse prognosis and inferior chemotherapy response, optimizing targeted treatment for CRC patients.

According to the GO functional enrichment results, most of these genes were clustered in functional groups related to the extracellular matrix (ECM). The ECM, as the main component of the tumor microenvironment, is considered to play a leading role in the progression of various cancers, including CRC, and promotes the invasion and metastasis of cancer (41–43). In addition, Ortensi et al. have demonstrated that the stemness characteristic of cancer tissue contributes to glioma invasiveness, which is closely related to the ECM (44). KEGG analysis also showed these genes were significantly abundant in inflammation, immunity, adhesion, invasion, and other processes. These results suggest that mRNAi-related metabolic genes are involved in CRC metastasis and invasion.

Furthermore, GSEA revealed that hallmark hypoxia was significantly enriched in the high-risk score group. Our study and other studies have shown that hypoxia may promote glycolysis in CRC cells by activating the HIF-1 α signaling pathway, thus promoting the proliferation and metastasis of CRC cells (45, 46). Hypoxia reduced the sensitivity of CRC to 5-fluorouracil chemotherapy (47). This suggests that mRNAi-related metabolic genes participate in hypoxia-related pathways, leading to a poor prognosis and chemotherapy resistance in CRC.

Hub genes were obtained by PPI network construction and Cytoscape software analysis. GO semantic similarity analysis showed a key role for the *CALD1* gene. Li et al. found that *CALD1* upregulated the expression of PD-L1 through the JAK/STAT signaling pathway and promoted malignant progression of bladder cancer (48). Several bioinformatics analyses and cellular studies have shown that *CALD* promoted the proliferation, metastasis, and invasion of CRC cells and is related to a reduction in OS (49). However, the exact mechanism of *CALD* involvement in CRC remains to be clarified.

Considering that the risk score model was derived from the stemness index and metabolism-related genes, which were significantly associated with antitumor immunity (13, 50), we further investigated the immune contexture heterogeneity between the high-risk and low-risk score groups. Tumor-infiltrating immune cell analysis showed that the high-risk score group had greater infiltration of CD8+ T cells, CD4+ T cells, B cells, Treg cells, dendritic cells, macrophages, MDSCs, neutrophils, regulatory T cells, and T helper cells, which is consistent with infiltrating immune cells from colorectal

cancer in a highly inflammatory state (51). Meanwhile, Zhongshan cohort validated that CD8⁺T cells, CD19⁺B cells, Foxp3⁺Tregs, and CD11c dendritic cells was more abundant in the high-risk group. But an increase in the ratio of Foxp3⁺Treg cells to CD8⁺T cells in high-risk group, which leads to an immunosuppressive microenvironment. Treg cells promote tumorigenesis and development by inhibiting adaptive anti-tumor immunity, which is the key mechanism of tumor immune escape (52). The ratio of Foxp3⁺Treg cells to CD8⁺T cells as is a better variable because they are more representative of the biological characteristics of infiltrating immune cells (53). The immune microenvironment of tumor is closely related to clinical outcome and drug resistance (54). Further immunostaining experiments confirmed CD8⁺T-cell dysfunction with decreased levels of cytotoxic molecules (GZMB and PRF1) and increased the expression of immune checkpoints (PD-1 and PD L1) in the high-risk score group, resulting in a highly exhausted state and impaired immune function. The effect of tumor on immune cells can lead to T cell anergy or dysfunction, which promotes tumor escape and therapeutic drug resistance (55). Our results suggested that high-risk group induced an immunoevasive contexture and impaired antitumor immunity, which explains the poor clinical prognosis.

Immunotherapy is a novel cancer treatment approach. Although the effect is significant, only a fraction of the patients responds to the treatment (56, 57). In the TCGA and GEO databases, patients with CRC with high-risk scores had higher expression of immunotherapeutic molecules (such as PD-L1, CTLA4, HAVCR2 and LAG3), suggesting that CRC with high-risk scores may be more likely to be affected by the immune-checkpoint pathway, inhibit the antitumor immune response, and lead to deterioration of prognosis. Similarly, the TIDE algorithm predicted that patients with high-risk scores were more sensitive to ICI therapy. In the Zhongshan Hospital cohort, we confirmed that CD8⁺T cells in high-risk score patients showed an exhausted T-cell phenotype with increased expression of the immunosuppressive checkpoint, PD-1 and PD-L1, compared with low-risk score patients. These findings also suggest that CRC patients with high-risk scores may clinically benefit from immunotherapy. The TMB score of patients with the high-risk score group is higher, suggesting that PD-1 blocking therapy has a certain curative effect on these patients (13). Low TMB is an important reason for patient resistance to immunotherapy (58). This provides a new approach for stratifying this subgroup of patients with CRC to identify those who may achieve a superior response to immunotherapy.

Because of the close relationship of the stemness index and chemotherapeutic drug resistance in cancer, we analyzed the predictive ability of the risk score to chemotherapeutic drug sensitivity and found that patients with high-risk scores were less susceptible to a variety of small molecular anticancer drugs and chemotherapeutic medicines in TCGA and GEO databases. In

the Zhongshan Hospital cohort, our findings suggest that the high-risk score group might have impaired therapeutic responsiveness to ACT in TNM stage III-IV CRC. Some studies have shown that in CRC, tumor stem cells lead to chemotherapy resistance by inhibiting antiapoptotic gene expression and reducing mitochondrial transcription initiation (59, 60). Due to the heterogeneity of CRC, the response of patients to chemotherapy is different, even at the same stage (61). This further suggests that patients with high-risk scores are potentially more suitable for immunotherapy than for traditional chemotherapy. However, the exact correlation between the risk score and the response to anticancer treatment needs to be further explored in a larger CRC cohort.

Finally, to improve clinical application, mRNAsi-related metabolic risk scores were combined with different clinicopathological characteristics to construct a prognostic nomogram and verify the predictive ability of the nomogram in TCGA and GEO datasets. In the Zhongshan Hospital cohort, a high-risk score was related to malignant progression and worse clinical outcomes in patients with CRC. This risk score model contains 21 important prognostic genes and has never been reported to identify the immunoevasive subgroup of patients with CRC in previous publications related to the CRC stemness index (62). Furthermore, this risk score model could help molecular typing and screening of differential subgroups to optimize personalized treatment and facilitate clinical translation.

Our study bears limitations. First, the mechanism of crosstalk between mRNAsi and metabolic reprogramming remains unclear, and further experimental studies are needed. Second, the clinical data of patients receiving immunotherapy in this study were limited, and the robust ability of the risk score model to predict immunotherapy responsiveness needs to be verified in a future larger immunotherapy cohort. Finally, although the new model and nomogram could accurately forecast the survival of patients with CRC in the TCGA and GEO databases and the Zhongshan Hospital cohort, more cell experiments, animal models, and clinical samples are needed to verify the value of this mRNAsi-related metabolic risk score model before developing immunotherapy strategies for subgroups of patients with CRC.

In this study, we proposed and validated a new risk score model according to 21 mRNAsi-related metabolic genes. The high-risk score group had a poorer clinical prognosis, inferior sensitivity to chemotherapy, a potentially better response to immunotherapy, and an immunoevasive environment, which sheds light on more accurate risk stratification and divides subgroups of patients with CRC for immunotherapy.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The studies involving human participants were reviewed and approved by This study was approved by the ethics committee of Zhongshan Hospital, Fudan University (B2022-068R2). The patients/participants provided their written informed consent to participate in this study.

Author contributions

MLW and CHM designed the study. MLW, TL, JZ, MMG, and WLZ, performed the study. MLW, WCG, CHS, QWL, YY, KN, and ZWZ analyzed the data. MLW and TL wrote the paper. MLW, MLS and DZ revised the paper. All authors contributed to the article and approved the submitted version.

Funding

Our study was supported by the National Natural Science Foundation of China (No.82002538, 82072213); Shanghai Pujiang Talent Plan (No. 2020PJD013); Clinical Research Plan of SHDC (No. SHDC2020CR1005A); and National Key Research and Development Program of China (No. 2020YFC2008400).

Acknowledgments

We thank Rongkui Luo (Department of Pathology, Zhongshan hospital Fudan University) for providing technical consultation.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.950782/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Data collection and processing. Box plot of CRC samples in the GSE17536 (A) and GSE39582 (B) datasets. Box plots of TCGA, GSE17536, and GSE39582 datasets before (C) and after (D) merging and de-batching.

SUPPLEMENTARY FIGURE 2

GSEA of DEGs in high and low-risk score groups of patients with CRC. (A–I). GSEA analysis of DEGs in high and low-risk score groups of patients with CRC.

SUPPLEMENTARY FIGURE 3

(A). The heatmap showing the expression of metabolism-related DEGs in patients with CRC. (B). The heatmap shows the expression of 83 significantly differentially expressed mRNAsi-related metabolic genes in the CRC and normal tissues.

SUPPLEMENTARY TABLE 1

Association between risk score and clinical characteristics of patients with CRC.

SUPPLEMENTARY TABLE 2

GO analysis of DEGs in high and low-risk score groups of patients with CRC.

SUPPLEMENTARY TABLE 3

KEGG analysis of DEGs in high and low-risk score groups of patients with CRC.

SUPPLEMENTARY TABLE 4

GSEA of DEGs in high and low-risk score groups of patients with CRC.

SUPPLEMENTARY TABLE 5

Univariate and multivariate Cox analyses of OS prediction in patients with CRC based on mRNAsi-related metabolic risk score model.

SUPPLEMENTARY TABLE 6

Relationship between high and low-risk score groups and clinicopathological features of patients with CRC in Zhongshan Hospital cohort.

SUPPLEMENTARY TABLE 7

Primers used in reverse transcription PCR analysis.

SUPPLEMENTARY TABLE 8

Immunohistochemistry antibodies.

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
- Kyrochristos ID, Roukos DH. Comprehensive intra-individual genomic and transcriptional heterogeneity: Evidence-based colorectal cancer precision medicine. *Cancer Treat Rev* (2019) 80:101894. doi: 10.1016/j.ctrv.2019.101894
- Zhai Z, Yu X, Yang B, Zhang Y, Zhang L, Li X, et al. Colorectal cancer heterogeneity and targeted therapy: Clinical implications, challenges and solutions for treatment resistance. *Semin Cell Dev Biol* (2017) 64:107–15. doi: 10.1016/j.semcdb.2016.08.033
- Linnekamp JF, Wang X, Medema JP, Vermeulen L. Colorectal cancer heterogeneity and targeted therapy: a case for molecular disease subtypes. *Cancer Res* (2015) 75(2):245–49. doi: 10.1158/0008-5472.CAN-14-2240
- Molinari C, Marisi G, Passardi A, Matteucci L, De Maio G, Ulivi P. Heterogeneity in colorectal cancer: A challenge for personalized medicine? *Int J Mol Sci* (2018) 19(12):3733. doi: 10.3390/ijms19123733
- Yi Y, Hsieh IY, Huang X, Li J, Zhao W. Glioblastoma stem-like cells: Characteristics, microenvironment, and therapy. *Front Pharmacol* (2016) 7:477. doi: 10.3389/fphar.2016.00477
- Bayik D, Lathia JD. Cancer stem cell-immune cell crosstalk in tumour progression. *Nat Rev Cancer* (2021) 21(8):526–36. doi: 10.1038/s41568-021-00366-w
- Yamashina T, Baghdadi M, Yoneda A, Kinoshita I, Suzu S, Dosaka-Akita H, et al. Cancer stem-like cells derived from chemoresistant tumors have a unique capacity to prime tumorigenic myeloid cells. *Cancer Res* (2014) 74(10):2698–709. doi: 10.1158/0008-5472.CAN-13-2169
- Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173(2):338–54. doi: 10.1016/j.cell.2018.03.034
- Yi L, Huang P, Zou X, Guo L, Gu Y, Wen C, et al. Integrative stemness characteristics associated with prognosis and the immune microenvironment in esophageal cancer. *Pharmacol Res* (2020) 161:105144. doi: 10.1016/j.phrs.2020.105144
- Chen X, Zhang D, Jiang F, Shen Y, Li X, Hu X, et al. Prognostic prediction using a stemness index-related signature in a cohort of gastric cancer. *Front Mol Biosci* (2020) 7:570702. doi: 10.3389/fmolb.2020.570702
- Xu Q, Xu H, Chen S, Huang W. Immunological value of prognostic signature based on cancer stem cell characteristics in hepatocellular carcinoma. *Front Cell Dev Biol* (2021) 9:710207. doi: 10.3389/fcell.2021.710207
- Wang Z, Wang Y, Yang T, Xing H, Wang Y, Gao L, et al. Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients. *Brief Bioinform* (2021) 22(5):bbab032. doi: 10.1093/bib/bbab032
- Li X, Wenes M, Romero P, Huang SC, Fendt SM, Ho PC. Navigating metabolic pathways to enhance antitumour immunity and immunotherapy. *Nat Rev Clin Oncol* (2019) 16(7):425–41. doi: 10.1038/s41571-019-0203-7
- Klebanoff CA, Crompton JG, Leonard AJ, Yamamoto TN, Chandran SS, Eil RL, et al. Inhibition of AKT signaling uncouples T cell differentiation from expansion for receptor-engineered adoptive immunotherapy. *JCI Insight* (2017) 2(23):e95103. doi: 10.1172/jci.insight.95103
- Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* (2010) 138(3):958–68. doi: 10.1053/j.gastro.2009.11.005
- Freeman TJ, Smith JJ, Chen X, Washington MK, Roland JT, Means AL, et al. Smad4-mediated signaling inhibits intestinal neoplasia by inhibiting expression of beta-catenin. *Gastroenterology* (2012) 142(3):562–71. doi: 10.1053/j.gastro.2011.11.026
- Williams CS, Bernard JK, Demory BM, Almohazey D, Washington MK, Smith JJ, et al. ERBB4 is over-expressed in human colon cancer and enhances cellular transformation. *Carcinogenesis* (2015) 36(7):710–18. doi: 10.1093/carcin/bgv049
- Marisa L, de Reynies A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* (2013) 10(5):e1001453. doi: 10.1371/journal.pmed.1001453
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007
- Zeng L, Fan X, Wang X, Deng H, Zhang K, Zhang X, et al. Bioinformatics analysis based on multiple databases identifies hub genes associated with hepatocellular carcinoma. *Curr Genomics* (2019) 20(5):349–61. doi: 10.2174/1389202920666191011092410
- 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. *Proc Natl Acad Sci U.S.A.* (2005) 102(43):15545–50. doi: 10.1073/pnas.0506580102

23. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OmicS* (2012) 16(5):284–87. doi: 10.1089/omi.2011.0118
24. 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. doi: 10.1093/nar/gky1131
25. 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. doi: 10.1101/gr.1239303
26. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* (2014) 8 Suppl 4:S11. doi: 10.1186/1752-0509-8-S4-S11
27. Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* (2010) 26(7):976–78. doi: 10.1093/bioinformatics/btq064
28. Wang Y, Zhao M, Zhang Y. Identification of fibronectin 1 (FN1) and complement component 3 (C3) as immune infiltration-related biomarkers for diabetic nephropathy using integrated bioinformatic analysis. *Bioengineered* (2021) 12(1):5386–401. doi: 10.1080/21655979.2021.1960766
29. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* (2013) 39(4):782–95. doi: 10.1016/j.immuni.2013.10.003
30. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* (2009) 462(7269):108–12. doi: 10.1038/nature08460
31. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
32. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* (2014) 9(9):e107468. doi: 10.1371/journal.pone.0107468
33. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2013) 41(Database issue):D955–61. doi: 10.1093/nar/gks1111
34. Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, et al. Large-Scale public data reuse to model immunotherapy response and resistance. *Genome Med* (2020) 12(1):21. doi: 10.1186/s13073-020-0721-z
35. Edge SB, Compton CC. The American joint committee on cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* (2010) 17(6):1471–74. doi: 10.1245/s10434-010-0985-4
36. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(delta delta C(T)) method. *Methods* (2001) 25(4):402–08. doi: 10.1006/meth.2001.1262
37. 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 Bioinf* (2011) 12:77. doi: 10.1186/1471-2105-12-77
38. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* (2015) 160(1–2):48–61. doi: 10.1016/j.cell.2014.12.033
39. Allavena P, Digifico E, Belgiovine C. Macrophages and cancer stem cells: a malevolent alliance. *Mol Med* (2021) 27(1):121. doi: 10.1186/s10020-021-00383-3
40. Bottcher M, Baur R, Stoll A, Mackensen A, Mougiakakos D. Linking immunoevasion and metabolic reprogramming in b-Cell-Derived lymphomas. *Front Oncol* (2020) 10:594782. doi: 10.3389/fonc.2020.594782
41. Kim MS, Ha SE, Wu M, Zogg H, Ronkon CF, Lee MY, et al. Extracellular matrix biomarkers in colorectal cancer. *Int J Mol Sci* (2021) 22(17):9185. doi: 10.3390/ijms22179185
42. Yuzhalin AE, Lim SY, Kutikhin AG, Gordon-Weeks AN. Dynamic matrisome: ECM remodeling factors licensing cancer progression and metastasis. *Biochim Biophys Acta Rev Cancer* (2018) 1870(2):207–28. doi: 10.1016/j.bbcan.2018.09.002
43. Crotti S, Piccoli M, Rizzolio F, Giordano A, Nitti D, Agostini M. Extracellular matrix and colorectal cancer: How surrounding microenvironment affects cancer cell behavior? *J Cell Physiol* (2017) 232(5):967–75. doi: 10.1002/jcp.25658
44. Ortensi B, Setti M, Osti D, Pelicci G. Cancer stem cell contribution to glioblastoma invasiveness. *Stem Cell Res Ther* (2013) 4(1):18. doi: 10.1186/scrt166
45. Weng ML, Chen WK, Chen XY, Lu H, Sun ZR, Yu Q, et al. Fasting inhibits aerobic glycolysis and proliferation in colorectal cancer via the Fdft1-mediated AKT/mTOR/HIF1alpha pathway suppression. *Nat Commun* (2020) 11(1):1869. doi: 10.1038/s41467-020-15795-8
46. Wu Z, Zuo M, Zeng L, Cui K, Liu B, Yan C, et al. OMA1 reprograms metabolism under hypoxia to promote colorectal cancer development. *EMBO Rep* (2021) 22(1):e50827. doi: 10.15252/embr.202050827
47. Nijhuis A, Thompson H, Adam J, Parker A, Gammon L, Lewis A, et al. Remodelling of microRNAs in colorectal cancer by hypoxia alters metabolism profiles and 5-fluorouracil resistance. *Hum Mol Genet* (2017) 26(8):1552–64. doi: 10.1093/hmg/ddx059
48. Li C, Yang F, Wang R, Li W, Maskey N, Zhang W, et al. CALD1 promotes the expression of PD-L1 in bladder cancer via the JAK/STAT signaling pathway. *Ann Transl Med* (2021) 9(18):1441. doi: 10.21037/atm-21-4192
49. Zheng H, Bai Y, Wang J, Chen S, Zhang J, Zhu J, et al. Weighted gene Co-expression network analysis identifies CALD1 as a biomarker related to M2 macrophages infiltration in stage III and IV mismatch repair-proficient colorectal carcinoma. *Front Mol Biosci* (2021) 8:649363. doi: 10.3389/fmobi.2021.649363
50. Zhang C, Chen T, Li Z, Liu A, Xu Y, Gao Y, et al. Depiction of tumor stemlike features and underlying relationships with hazard immune infiltrations based on large prostate cancer cohorts. *Brief Bioinform* (2021) 22(3). doi: 10.1093/bib/bbaa211
51. Vayrynen JP, Tuomisto A, Klintrup K, Makela J, Karttunen TJ, Makinen MJ. Detailed analysis of inflammatory cell infiltration in colorectal cancer. *Br J Cancer* (2013) 109(7):1839–47. doi: 10.1038/bjc.2013.508
52. Li C, Jiang P, Wei S, Xu X, Wang J. Regulatory T cells in tumor microenvironment: new mechanisms, potential therapeutic strategies and future prospects. *Mol Cancer* (2020) 19(1):116. doi: 10.1186/s12943-020-01234-1
53. Ino Y, Yamazaki-Itoh R, Shimada K, Iwasaki M, Kosuge T, Kanai Y, et al. Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. *Br J Cancer* (2013) 108(4):914–23. doi: 10.1038/bjc.2013.32
54. Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol* (2013) 14(10):1014–22. doi: 10.1038/ni.2703
55. Dersh D, Holly J, Yewdell JW. A few good peptides: MHC class I-based cancer immunosurveillance and immunoevasion. *Nat Rev Immunol* (2021) 21(2):116–28. doi: 10.1038/s41577-020-0390-6
56. Billan S, Kaidar-Person O, Gil Z. Treatment after progression in the era of immunotherapy. *Lancet Oncol* (2020) 21(10):e463–76. doi: 10.1016/S1470-2045(20)30328-4
57. Schoenfeld AJ, Hellmann MD. Acquired resistance to immune checkpoint inhibitors. *Cancer Cell* (2020) 37(4):443–55. doi: 10.1016/j.ccell.2020.03.017
58. Touat M, Li YY, Boynton AN, Spurr LF, Iorgulescu JB, Bohrsen CL, et al. Mechanisms and therapeutic implications of hypermutation in gliomas. *Nature* (2020) 580(7804):517–23. doi: 10.1038/s41586-020-2209-9
59. Colak S, Medema JP. Human colonic fibroblasts regulate stemness and chemotherapy resistance of colon cancer stem cells. *Cell Cycle* (2016) 15(12):1531–37. doi: 10.4161/15384101.2014.973321
60. Colak S, Zimmerlin CD, Fessler E, Hogdal L, Prasetyanti PR, Grandela CM, et al. Decreased mitochondrial priming determines chemoresistance of colon cancer stem cells. *Cell Death Diff* (2014) 21(7):1170–77. doi: 10.1038/cdd.2014.37
61. Kishore C, Bhadra P. Current advancements and future perspectives of immunotherapy in colorectal cancer research. *Eur J Pharmacol* (2021) 893:173819. doi: 10.1016/j.ejphar.2020.173819
62. Wei R, Quan J, Li S, Liu H, Guan X, Jiang Z, et al. Integrative analysis of biomarkers through machine learning identifies stemness features in colorectal cancer. *Front Cell Dev Biol* (2021) 9:724860. doi: 10.3389/fcell.2021.724860