



## OPEN ACCESS

## EDITED BY

David A. Gewirtz,  
Virginia Commonwealth University,  
United States

## REVIEWED BY

Elena De Mattia,  
Aviano Oncology Reference Center,  
(IRCCS), Italy  
Donatella Lucchetti,  
Catholic University of the Sacred  
Heart, Italy

## \*CORRESPONDENCE

Bang-li Hu  
hubangli@gxmu.edu.cn  
Rong Lin  
selinalin@hotmail.com

## SPECIALTY SECTION

This article was submitted to  
Pharmacology of Anti-Cancer Drugs,  
a section of the journal  
Frontiers in Oncology

RECEIVED 26 May 2022

ACCEPTED 05 August 2022

PUBLISHED 30 August 2022

## CITATION

Guo X-W, Li S-Q, Lei R-E, Ding Z,  
Hu B-L and Lin R (2022) Tumor-  
infiltrating immune cells based  
TMEScore and related gene signature  
is associated with the survival of CRC  
patients and response to  
fluoropyrimidine-based  
chemotherapy.  
*Front. Oncol.* 12:953321.  
doi: 10.3389/fonc.2022.953321

## COPYRIGHT

© 2022 Guo, Li, Lei, Ding, Hu and Lin.  
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.

# Tumor-infiltrating immune cells based TMEScore and related gene signature is associated with the survival of CRC patients and response to fluoropyrimidine-based chemotherapy

Xian-Wen Guo<sup>1,2</sup>, Si-Qi Li<sup>3</sup>, Rong-E Lei<sup>4</sup>, Zhen Ding<sup>1</sup>,  
Bang-li Hu<sup>3\*</sup> and Rong Lin<sup>1\*</sup>

<sup>1</sup>Department of Gastroenterology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, <sup>2</sup>Department of Gastroenterology, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China, <sup>3</sup>Department of Research, Guangxi Medical University Cancer Hospital, Nanning, China, <sup>4</sup>Department of Gastroenterology, The First Affiliated Hospital of Guangxi Medical University, Nanning, China

**Background:** Tumor-infiltrating immune cells (TIICs) are associated with chemotherapy response. This study aimed to explore the prognostic value of a TIIC-related tumor microenvironment score (TMEScore) in patients with colorectal cancer (CRC) who underwent chemotherapy and construct a TMEScore-related gene signature to determine its predictive value.

**Methods:** Gene profiles of patients who underwent fluoropyrimidine-based chemotherapy were collected, and their TIIC fractions were calculated and clustered. Differentially expressed genes (DEGs) between clusters were used to calculate the TMEScore. The association between the TMEScore, chemotherapy response, and survival rate was analyzed. Machine learning methods were used to identify key TMEScore-related genes, and a gene signature was constructed to verify the predictive value.

**Results:** Two clusters based on the TIIC fraction were identified, and the TMEScore was calculated based on the DEGs of the two clusters. The TMEScore was higher in patients who responded to chemotherapy than in those who did not, and was associated with the survival rate of patients who underwent chemotherapy. Three machine learning methods, support vector machine (SVM), decision tree (DT), and Extreme Gradient Boosting (XGBoost), identified three TMEScore-related genes (ADH1C, SLC26A2, and NANS) associated with the response to chemotherapy. A TMEScore-related gene signature was constructed, and three external cohorts validated that the gene signature could predict the response to chemotherapy. Five datasets

and clinical samples showed that the expression of the three TMEscore-related genes was increased in tumor tissues compared to those in control tissues.

**Conclusions:** The TIIC-based TMEscore was associated with the survival of CRC patients who underwent fluoropyrimidine-based chemotherapy, and predicted the response to chemotherapy. The TMEscore-related gene signature had a better predictive value for response to chemotherapy than for survival.

#### KEYWORDS

colorectal cancer, tumor-infiltrating immune cells, TMEscore, chemotherapy, machine learning, gene signature

## Introduction

The prognosis of patients with colorectal cancer (CRC) is strongly dependent on the clinical stage, as the 5-year survival rate of patients at an advanced stage is much lower than that of patients at an early stage (1). Unfortunately, most patients with CRC are initially diagnosed at an advanced stage. Chemotherapy remains an important and effective approach for the treatment of patients with CRC, particularly those at advanced stages. Fluoropyrimidine-based chemotherapy, including FOLFOX, FOLFIRI, or CapeOX, is recommended as the first-line treatment for patients with unresectable and/or recurrent CRC (2, 3). However, chemotherapy resistance, which occurs in a significant proportion of patients and results in tumor progression and ultimately death, is a major obstacle to the successful treatment of CRC. Many mechanisms have been found to underlie the occurrence of chemotherapy resistance, including alterations in drug metabolism, aberrant DNA repair and proliferation of cancer cells, activation of detoxifying enzymes, and cancer cell death inhibition (4, 5). In recent decades, growing evidence has shown that alteration of the tumor microenvironment (TME) is an important factor that significantly contributes to chemotherapy resistance (6).

The TME is composed of tumor cells, stromal cells, immune cells, and other cell types. Changes in the cell diversity and cytokines in the TME contribute to the pathogenesis, progression, therapy response, and prognosis of cancers (7, 8). Alterations in TME cells, especially tumor-infiltrating immune cells (TIICs), and the protein molecules and cytokines secreted by these cells, mediate the occurrence of chemotherapy resistance (9). Recently, the association between TIICs and chemotherapy resistance in CRC has been reported in several studies. For example, a study reported that the presence of T follicular helper cells and M0 macrophages in the TME of

patients with CRC who underwent fluoropyrimidine-based chemotherapy were associated with better survival rates, whereas eosinophils in the TME were associated with worse survival rates (10). Another study used a TME-specific gene signature to identify CRC subtypes. In this study, a “signature associated with FOLFIRI resistance and the microenvironment” (SFM) was constructed to identify both TIICs and drug sensitivity in CRC patients (11). A single-cell atlas of liver metastases of CRC reveals tumors treated with preoperative chemotherapy show activation of B cells, lower diversity of tumor-associated macrophages with immature and less activated phenotype, lower abundance of both dysfunctional T cells and ECM-remodeling cancer-associated fibroblasts, and accumulation of myofibroblasts (12). These results demonstrate that TIICs in the TME could be used to identify patients with chemoresistance and predict the prognosis in CRC patients undergoing chemotherapy.

The TMEscore is an index that reflects the fraction of immune cells and is calculated from a gene expression matrix using principal component analysis (PCA) (13). Compared with the immune cell fraction, the TMEscore could be more reliable in representing the composition of the TME. The TMEscore has recently been used to predict the survival rate of patients with various cancer types, such as gastric cancer (13), glioma (14), and ovarian cancer (15). However, an association between the TMEscore and chemotherapy resistance has not been previously reported. Therefore, in this study, we calculated the TMEscore of patients with CRC who underwent fluoropyrimidine-based chemotherapy based on TIICs, and constructed a TMEscore-related gene signature to predict the survival rate of these patients. Our results provide novel insights into the role of TIIC-mediated chemotherapy resistance and precise treatment options for patients who undergo fluoropyrimidine-based chemotherapy.

## Materials and methods

### CRC dataset acquisition and preprocessing

The CRC datasets were retrieved and freely downloaded from the gene expression omnibus (GEO) database ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)). Datasets containing information from 2061 patients with CRC, including GSE72970 (n=124), GSE39582 (n=585), GSE87211 (n=363), GSE104645 (n=193), and GSE28702 (n=83) were included in the analysis. The cancer genome atlas (TCGA)-CORDREAD (n=635) dataset (transcripts per million normalized) with corresponding clinical information was downloaded freely from the Xena database (GDC hub: <https://gdc.xenahubs.net>). Four GEO datasets, GSE106582 (n=194), GSE31737 (n=80), GSE117606 (n=208), and GSE74602 (n=60), were downloaded to validate gene expression differences between tumor and control tissues. The GEO datasets were preprocessed by performing background adjustment using the RMA algorithm.

### Quantification of TIICs in the TME

The TIICs in the TME of CRC tissues were quantified using the CIBERSORT algorithm (16), which generates a fraction of 22 immune cell phenotypes by calculating gene expression. The CRC samples with CIBERSORT P of <0.05 were retained for subsequent analysis because this inferred a highly reliable cell composition.

### Consensus clustering analysis and differentially expressed gene analysis

The “ConsensusClusterPlus” package (17) was used to perform unsupervised consensus clustering of the samples based on the TIIC fraction. The cumulative distribution function (CDF) curves determined the optimal number of clusters, which was indexed by k values from 2 to 6. The DEG analyses between different clusters in each dataset were performed using the “limma” package. This package provides an integrated solution for analyzing data from gene expression experiments. It contains rich features for handling complex experimental designs. To date, the “limma” package is a classic method for the differentially expressed analysis of microarray datasets (18).

### Functional enrichment analysis for the DEGs

Gene set variation analysis (GSVA) was applied to screen significantly enriched pathways between the two clusters using the Molecular Signatures Database (MSigDB, version 7.4) (19).

Those with a P value of <0.05 were considered significant pathway terms. Kyoto Encyclopedia of Genes and Genomes (KEGG) terms were identified with a cutoff P value of <0.01 and a false discovery rate of <0.05. These analyses were conducted using the clusterProfiler package. The clusterProfiler package provides a universal interface for functional enrichment analysis in thousands of organisms based on internally supported ontologies and pathways, as well as annotation data provided by users or derived from online databases (20).

### TMEscore calculation and cluster analysis

The TMEscore of the gene dataset was calculated using the “TMEscore” package (13, 21). This package provides functionality for calculating the TMEscore using PCA or z-score methods. For the dataset with survival data, this package divides the samples into high or low clusters based on the TMEscore and survival data.

### Chemotherapeutic drugs and immunotherapeutic response prediction

The potential chemotherapeutic drugs that were associated with the TMEscore clusters were screened using the “pRRophetic” package (22), which is used to calculate the half-maximal inhibitory concentration (IC50) based on the gene expression profile for 251 chemotherapeutic drugs. Furthermore, the response of the TMEscore clusters to immunotherapy was predicted by comparing the expression of immune checkpoint inhibitors.

### Screening and validation of key genes using multiple machine learning algorithms

The datasets were randomly classified into training and testing sets in a ratio of 7:3. First, support vector machine (SVM), decision tree (DT), and Extreme Gradient Boosting (XGBoost) algorithms were used to screen the most important genes *via* the e1071, rpart, and XGBoost packages in the R language. The predictive value of the three machine learning algorithms for important gene selection in the training set was estimated by the receiver operating characteristic (ROC) curves and areas under the curve (AUC) using the “pROC” package. Then, the intersecting genes among the three algorithms were considered as the key genes and visualized using a Venn diagram. Finally, multivariate logistic regression analysis was conducted to construct the predictive signature using the key genes, which evaluated the predictive value of the signature *via* AUC indices.

## Reverse transcription-polymerase chain reaction assay for clinical samples

The expression of key genes in the CRC clinical samples was determined using RT-PCR. Forty tumor tissues and their corresponding adjacent tissues were collected between February 2020 and August 2021. This study was approved by the hospital's ethics committee. Total RNA extraction, RT-PCR, and analysis were performed as previously described (23). TaqMan probe-based RT-PCR was performed using a commercial kit (Thermo Fisher Scientific, Inc.) according to the manufacturer's instruction. The primers used for RT-PCR were as follows: ADH1C forward: TGA TAA AGT CAT CCC GCT CTT T, reverse: CAT TCT CAT CCA CCA CTG TGT A; SLC26A2 forward: AAG AGC AAC ATA ACG TTT CAC C, reverse: GTC TGC ATT GAT CAT TGG TCT C; NANS forward: TCA GAA GCT CTT TCC TGA CAT T; reverse: GTC CAA AGT TAT GTG ACG TTC C. The reaction conditions for the RT-PCR were as follows: 1 cycle at 95°C for 5 min; 15 cycles at 95°C for 25 sec, 64°C for 20 sec, 72°C for 20 sec; and a final 31 cycles at 93°C for 25 sec, 64°C for 20 sec, 72°C for 20 sec. Amplicons were detected using capillary electrophoresis on an ABI 3130xl Genetic Analyzer (Life Technologies, Grand Island, NY).

## Statistical analysis

Data with normally distributed variables were tested using the unpaired Student's *t* test for two-group comparisons; otherwise, the Mann-Whitney U-test was used. The Kruskal-Wallis test and one-way analysis of variance were used for multiple group comparisons. The chi-squared test was used to compare categorical variables. Kaplan-Meier and log-rank (Mantel-Cox) tests were employed to compare patient survival rates using the "survminer" package. All statistical analyses were conducted using the R language. P values of <0.05 were considered statistically significant. P values were two-sided.

## Results

### Establishment of clusters based on the TIIC fraction

The GSE72970 dataset included 124 samples from CRC patients who underwent fluoropyrimidine-based chemotherapy, including FOLFOX, FOLFIRI, FOLFIRINOX, combined with bevacizumab. The dataset only included patients with metastatic CRC. Sixty-three patients had shown a response (complete response + partial response) to chemotherapy, whereas 61 patients had not (stable disease +

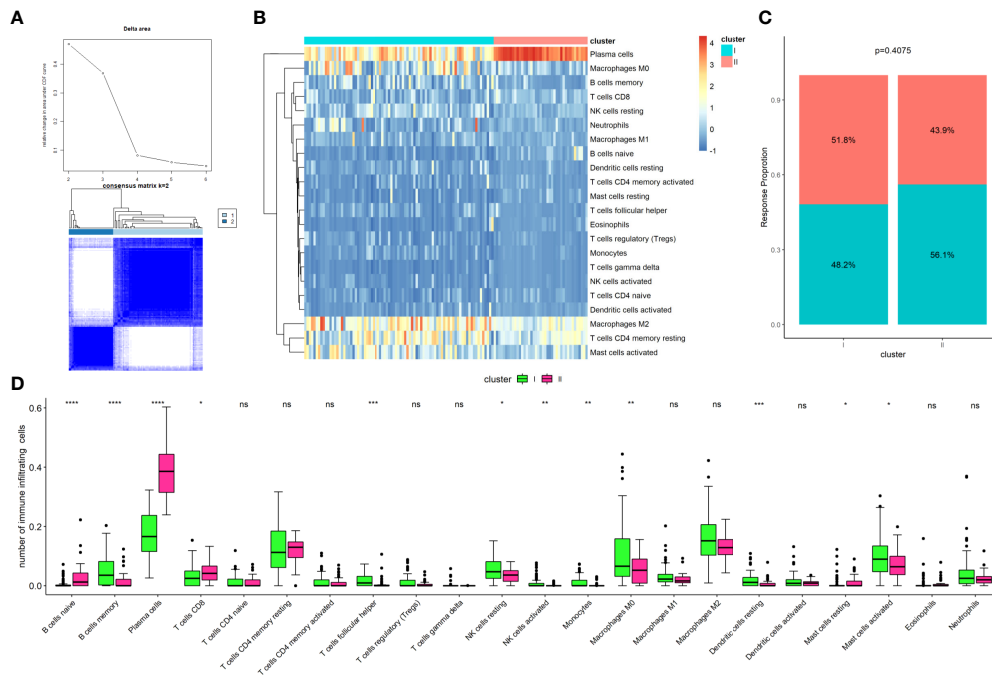
progressive disease). The TIIC fraction of the dataset was calculated using the CIBERSORT algorithm. Then, the 22 TIIC fraction was clustered using unsupervised clustering methods *via* the ConsensusClusterPlus package. The results suggested that TIICs can be optimally divided into two clusters (cluster I, n=83; cluster II, n=41; Figure 1A). Figure 1B shows a heatmap of the clusters for the TIIC fraction. The frequency of the clusters was not significantly different to that of the response to chemotherapy (Figure 1C, P=0.407). Figure 1D shows a comparison of the TIIC fraction between clusters I and II. These results indicated that the TIICs of patients who underwent fluoropyrimidine-based chemotherapy could be divided into two clusters, and that the results were associated with the response to chemotherapy.

### GSVA and DEG analyses for the TIIC clusters

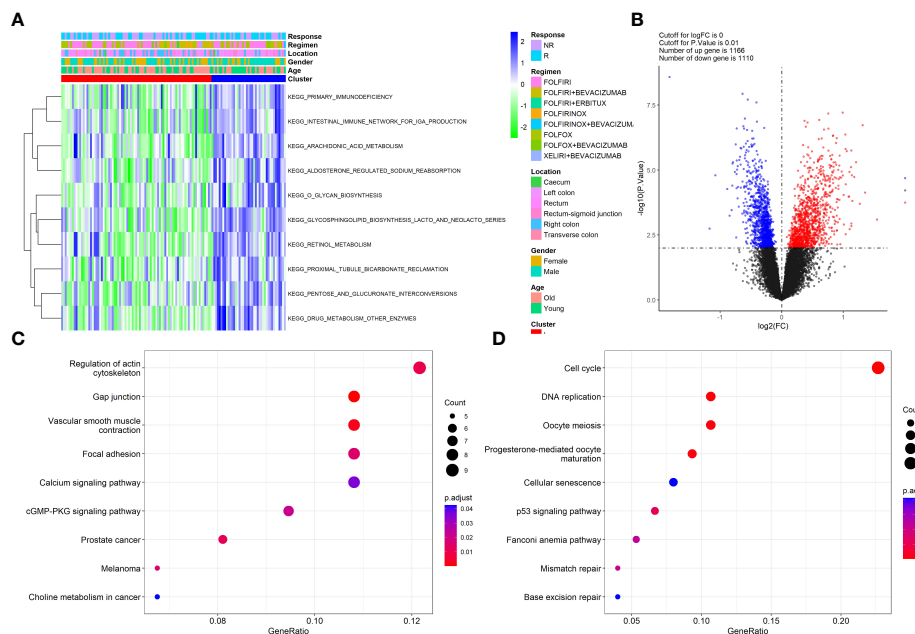
To reveal the potential pathways involved in the TIIC clusters of patients with CRC, the GSVA algorithm was applied to clusters I and II. Generally, GSVA is used to estimate the variation in pathway activity over a sample population in an unsupervised manner. As illustrated in Figure 2A, the top three enriched pathways by GSVA between the two clusters were primary immunodeficiency, the intestinal immune network for IGA production, and arachidonic acid metabolism, suggesting that these two clusters were involved in the immune system processes. Next, DEGs between clusters I and II were screened using the "limma" package in the GSE72970 dataset with the criterion of a P value of <0.05. A total of 2267 immune cell-related DEGs were identified between the two clusters (Figure 2B). Afterwards, the pathways that the DEGs were involved in were determined using the top 200 upregulated and downregulated DEGs *via* the "clusterProfiler" package. The results revealed that the upregulated DEGs were mainly involved in the regulation of the actin cytoskeleton, calcium signaling pathway, and cGMP-PKG signaling pathway (Figure 2C); the downregulated DEGs were mainly involved in the cell cycle, DNA replication, and p53 signaling pathway (Figure 2D). Altogether, these results provided information on the role of TIIC clusters and DEGs in biological processes and the pathophysiology of diseases.

### TMEscore calculation and the association with chemotherapy response and survival

Next, the TMEscore was calculated by incorporating the 2267 DEGs and survival data of the GSE72970 dataset using the "TMEscore" package. The TMEscore was lower in those who did not respond to chemotherapy compared to those who did



**FIGURE 1** Establishment of clusters based on immune cells infiltrating fraction. **(A)** Unsupervised consensus clustering for the samples based on the TIICs infiltration, with the k value as 2; **(B)** Clustering of the tumor-infiltrating immune cells; **(C)** Comparison of cluster frequency with the response status; **(D)** Comparison of tumor-infiltrating immune cells between Cluster I and Cluster II. \*P < 0.05; \*\*P < 0.01; \*\*\*P < 0.001; ns, not significant.

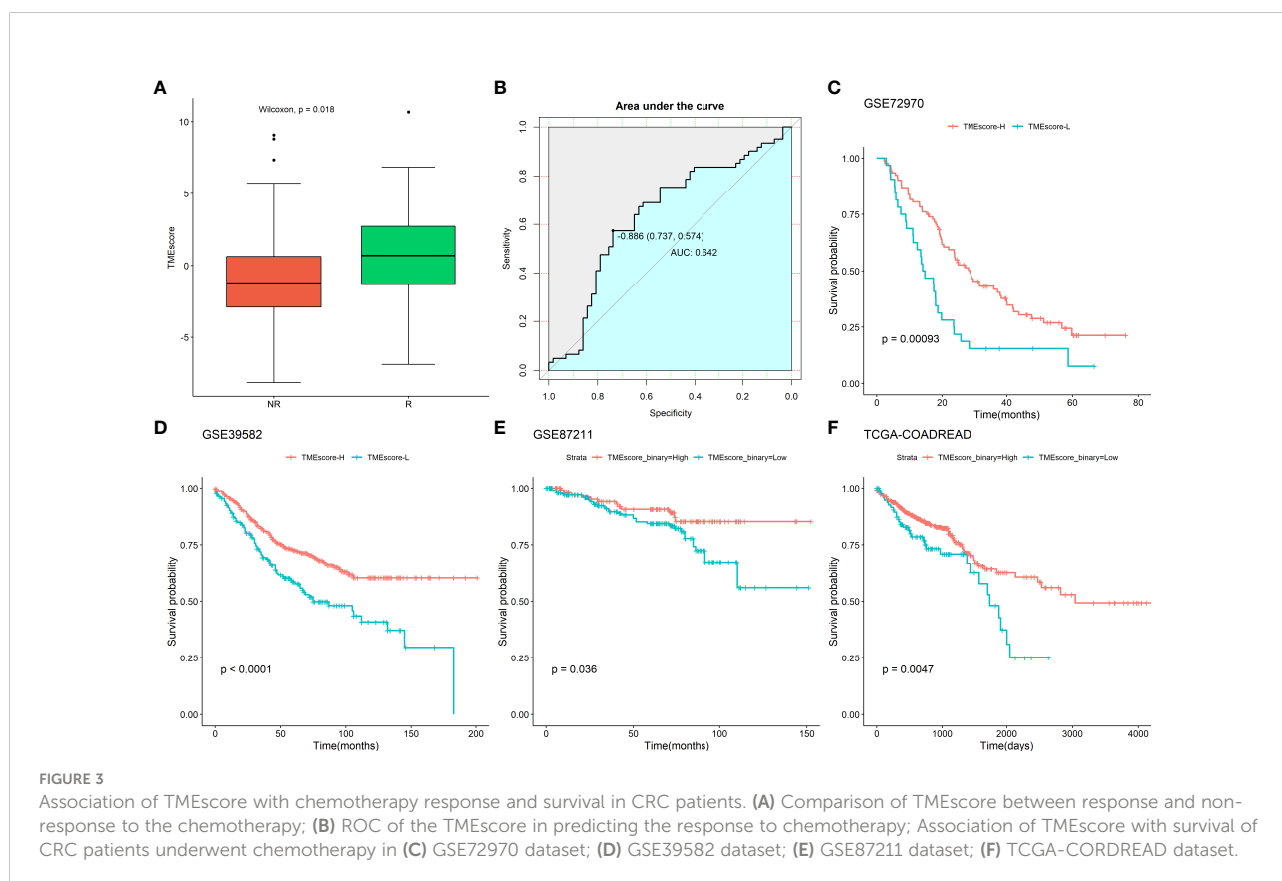


**FIGURE 2** Functional enrichment of clusters and the DEGs. **(A)** Heatmap of GSEA analysis revealed the significant pathways that Clusters involved; **(B)** Volcano plot showed the DEGs between the two clusters; **(C)** KEGG pathways that up-regulated DEGs involved; **(D)** KEGG pathways that down-regulated DEGs involved.

respond to chemotherapy (Wilcoxon test,  $P=0.018$ ; Figure 3A). The ROC curve revealed that the TMEscore had a moderate predictive value for identifying patients who responded to chemotherapy (AUC:0.642, Figure 3B). Next, the samples were divided into two groups, TMEscore-High (TMEscore-H) and TMEscore-Low (TMEscore-L) based on the optimal cutoff point of the TMEscore. The Kaplan-Meier plot showed that there was a significantly different survival time between TMEscore-H and TMEscore-L patients (Figure 3C). Subsequently, the association of the TMEscore group with the survival of CRC patients who underwent fluoropyrimidine-based chemotherapy was validated in three datasets (GSE39582,  $n=585$ ; GSE87211,  $n=363$ ; TCGA-CORDREAD,  $n=635$ ; Figures 3C–F). The results revealed that patients with a high TMEscore showed better survival time compared to those with a low TMEscore. In addition, we conducted subgroup analyses by dividing the samples into two groups according to the administration of irinotecan or oxaliplatin in combination with fluoropyrimidine. We then compared the TMEscore and other clinical parameters between the two groups, and the results failed to show any significance regarding the TMEscore, patients' age, T stage, N stage, and response status (Supplementary Table S1). Altogether, these results indicate that the TMEscore based on the TIIC-related DEGs could predict the prognosis of CRC patients who underwent chemotherapy.

## Clinical therapeutic strategies using the TMEscore

To explore the clinical therapeutic strategies using the TMEscore, we applied the “pRRophetic” package (22) to predict the potential therapeutic effects of CRC first-line chemotherapy drugs, including 5-fluorouracil, cisplatin, docetaxel, gemcitabine, and paclitaxel. As shown in Figures 4A, B, the IC50 of the TMEscore-H group was not significantly different to that of the TMEscore-L group for cisplatin and docetaxel ( $P>0.05$ ), but was significantly different for 5-fluorouracil, gemcitabine, and paclitaxel (Figures 4C–E). Furthermore, five chemotherapy drugs (temozolomide (24), pyrimethamine (25), lapatinib (26), doxorubicin (27), and ruxolitinib (28)), which are reported to be associated with the treatment of CRC, showed significant differences between the TMEscore-H and TMEscore-L groups (Figures 4F–J). Thereafter, the expression levels of six known immune checkpoint inhibitors, namely PD1 (PDCD1), PD-L1, PD-L2 (PDCD1LG2), CD80, CD86, CTLA4, and CD274, were compared based on the TMEscore clusters. Only the expression of CD86 and PD-L2 was significantly different between the TMEscore-H and TMEscore-L groups (Figure 4K). In summary, these results suggest that CRC patients in the TMEscore-H group tend to benefit from



specific chemotherapy drugs but are unlikely to benefit from immunotherapy.

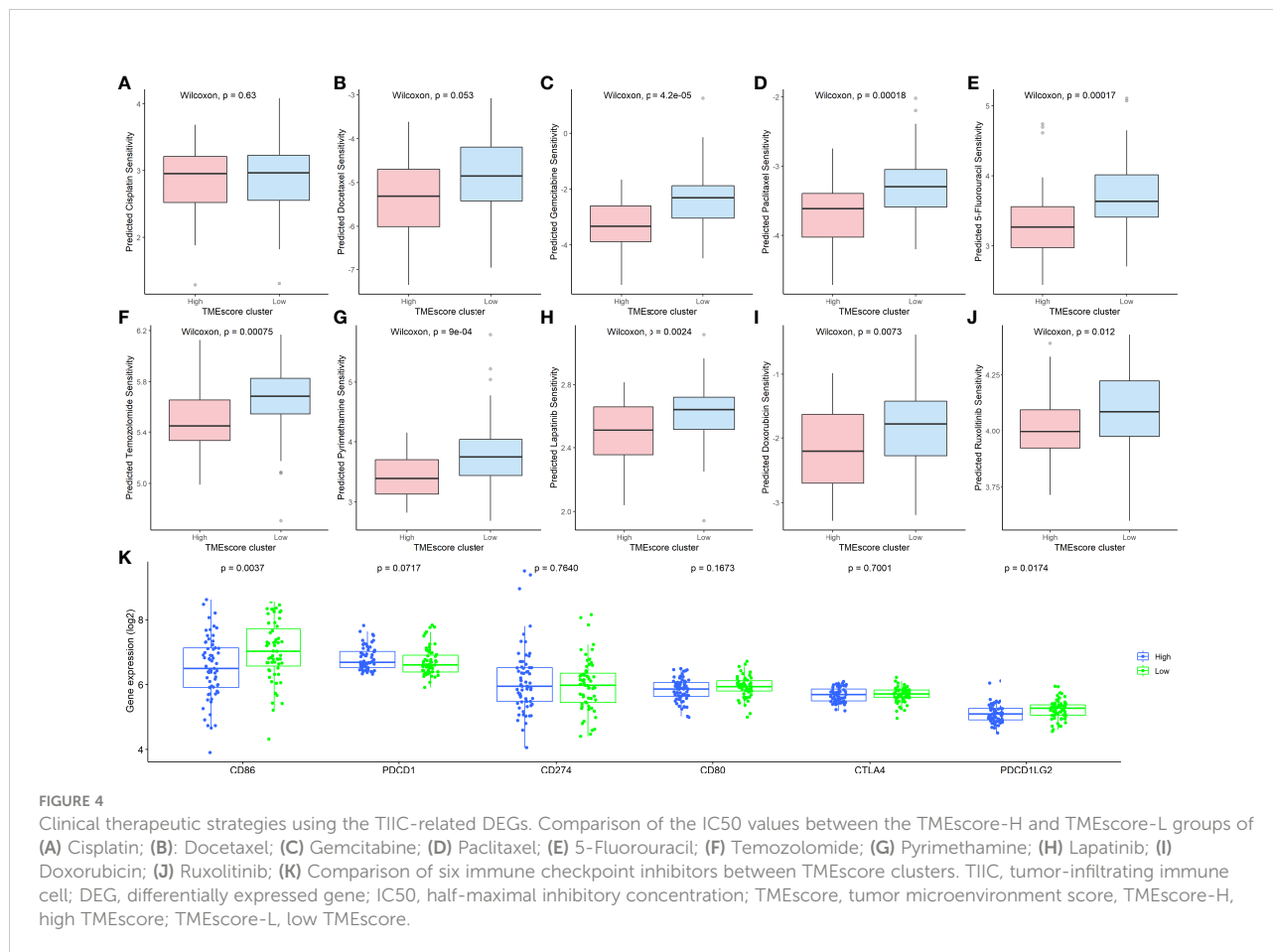
## Machine learning methods screened key genes for chemotherapy response

To screen key genes associated with the response to chemotherapy of CRC patients, three machine learning methods (XGBoost, SVM, and DT) were employed to analyze the 2267 TIIC-related DEGs in the datasets. The samples were divided into training and test sets at a ratio of 7:3, with 768 and 321 samples in the training and test sets, respectively. The number of key genes screened by each method was 18, 30, and 8 in the training sets, which suggested that these genes are important in identifying the patients who respond to chemotherapy. Thereafter, we determined the predictive value of the above key genes using each machine learning method in the tested set. The results showed that the AUC values for the tested set were 0.690, 0.663, and 0.729, respectively (Figures 5A–C), suggesting that the key genes from each machine learning method achieved a moderate predictive value. Finally, three key genes (ADH1C, SLC26A2, and NANS) were obtained by

overlapping the genes found using the three machine learning methods (Figure 5D). A logistic regression model was employed to determine the predictive value using these three key genes, and the results indicated that the signature constructed by the three key genes had good predictive performance, with an AUC of 0.867 (Figure 5E), indicating that the three key genes derived from the three machine learning methods improved the predictive performance of the model.

## Validation of the predictive value of the three-gene signature for chemotherapy response

To validate the predictive value of the three-gene (ADH1C, SLC26A2, and NANS) signature in patients with CRC who underwent fluoropyrimidine-based chemotherapy, two GEO datasets (GSE104645,  $n=193$ ; GSE28702,  $n=83$ ) were screened, which provided the response status data of CRC patients who underwent FOLFOX or FOLFOX-based regimens. Logistic regression was then employed to determine the predictive value of the three-gene signature. As Figures 6A, B shows, the predictive value of the signature constructed using the three



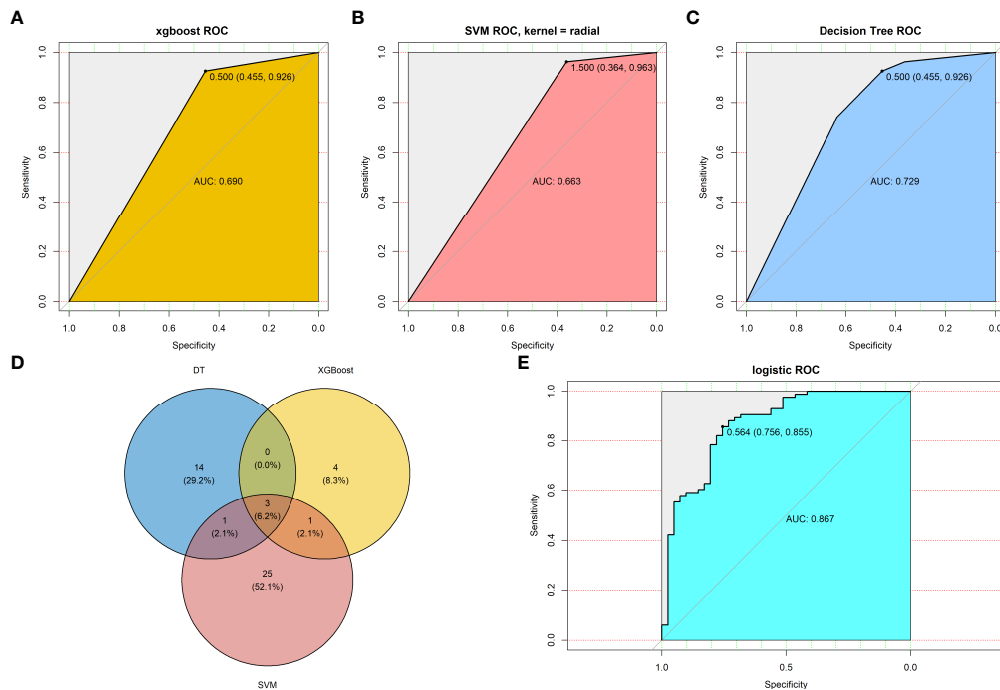


FIGURE 5

Machine learning methods screen key genes response to chemotherapy. Predictive value of TMEScore-relate genes in testing set by (A) XGBoost algorithms; (B) SVM algorithms; (C) DT algorithms; (D) Venn plot revealed three overlapped among the three machine learning methods; (E) Predictive value of the three TMEScore-relate genes signature in the response to chemotherapy by multivariate logistic regression.

genes was moderate, with AUC values of 0.616 and 0.719, respectively. Next, samples from patients who only underwent 5-fluorouracil treatment ( $n=36$ ) were extracted from the TCGA-COADREAD dataset, and the predictive value increased, with an AUC value of 0.764 (Figure 6C). These results indicate that the signature constructed by ADH1C, SLC26A2, and NANS could help to identify patients who will probably benefit from fluoropyrimidine-based chemotherapy.

## Validation of the expression of key genes in external cohorts and clinical samples

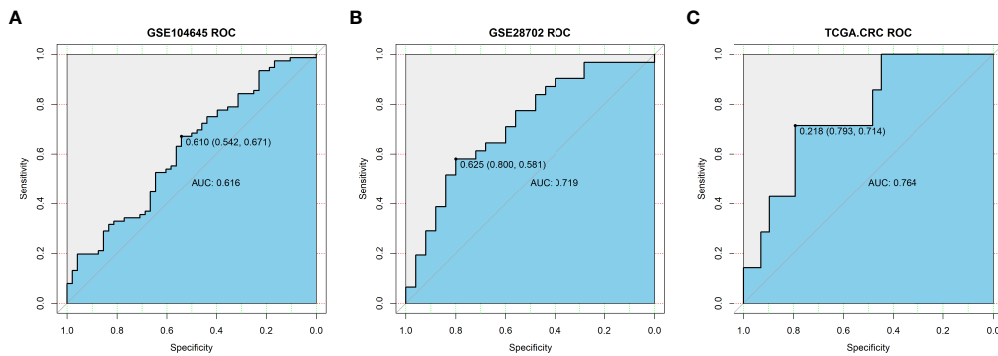
The expression of ADH1C, SLC26A2, and NANS was determined in the TCGA-COADREAD dataset ( $n=689$ ) and four large external cohorts (GSE106582,  $n=194$ ; GSE31737,  $n=80$ ; GSE117606,  $n=208$ ; GSE74602,  $n=60$ ), and the expression of the three genes were compared between tumor and control tissues. As shown in Figures 7A–E, the results from the above five datasets revealed that the levels of all three genes were decreased in tumor tissues compared with those in control tissues in CRC samples ( $P<0.01$ ). Finally, clinical samples of CRC were collected and gene expression was tested using RT-PCR. Consistent with the results from the datasets, the

expression levels of ADH1C, SLC26A2, and NANS were substantially decreased in tumor tissues compared with those in the corresponding adjacent tissues ( $P<0.01$ ; Figure 7F). These results demonstrated that the levels of ADH1C, SLC26A2, and NANS were decreased in CRC tissues compared to those in control tissues.

## Discussion

Currently, fluoropyrimidine-based chemotherapy regimens, mainly 5-fluorouracil and its oral prodrug capecitabine combined with irinotecan or oxaliplatin, remain the mainstay of treatment for many types of solid tumors (29), such as CRC, gastric cancer, and pancreatic cancer. However, only half of the patients respond to this regimen, which substantially hampers the treatment effect (30). Due to the relatively slow advancement of new agents to substitute traditional chemotherapeutics, predicting whether the patient will develop fluoropyrimidine-based chemotherapy resistance is critical in achieving effective clinical management of patients with CRC. Many biomarkers have been shown to be associated with fluoropyrimidine-based chemotherapy resistance, such as complement component 3 (31), ZEB2 (32), and PKM2 (33). These biomarkers may help



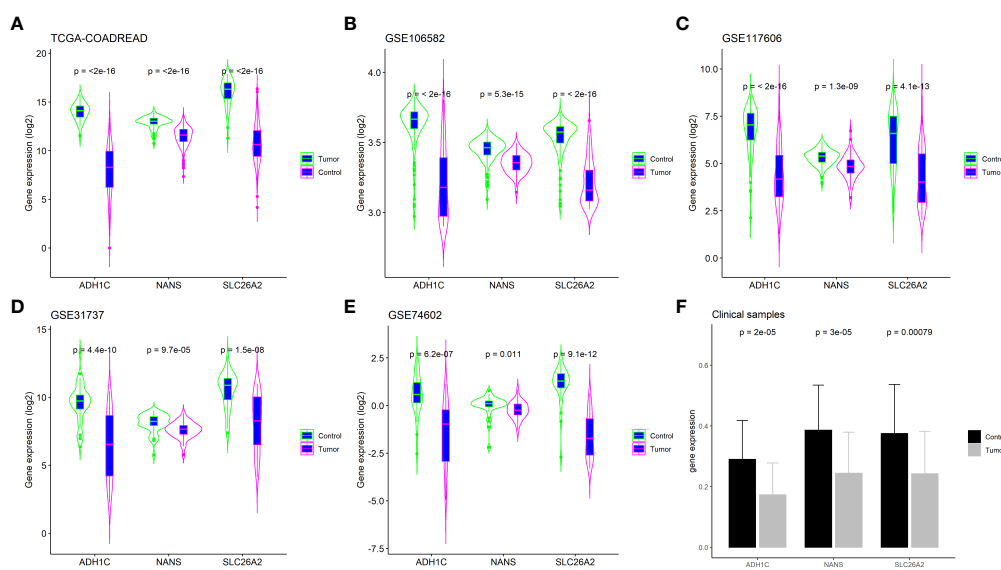


**FIGURE 6** Validation of predictive value of TMEScore-relate genes signature in the response to chemotherapy in (A) GSE104645 dataset; (B) GSE28702 dataset; (C) TCGA-COADREAD dataset (36 patients underwent 5-Fluorouracil treatment).

identify patients who are suitable for fluoropyrimidine-based chemotherapy.

In the present study, the TIIC fraction was analyzed in a dataset with information on fluoropyrimidine-based chemotherapy. It was found that TIICs could be divided into two clusters, which were associated with the response to chemotherapy. The DEGs between the two clusters of TIICs were screened and then used to calculate the TMEScore and construct the TMEScore signature. These results indicated that the TMEScore was significantly associated with the survival of patients with CRC who underwent chemotherapy, suggesting

that it could serve as an important indicator for predicting the prognosis of patients with CRC. Given that chemotherapy resistance is induced not only by a variety of cells, but also by protein molecules and cytokines in the cells, the genes that represent the TMEScore signature were also screened using three machine learning methods; three genes with good performance in predicting the response to chemotherapy in CRC patients were identified. Finally, the expression of these three genes was validated in larger cohorts and clinical samples. These results demonstrated that TIICs were associated with the response to fluoropyrimidine-based chemotherapy in patients with CRC,



**FIGURE 7** Validation key genes expression in external cohorts and clinical samples. (A) TCGA-COADREAD dataset; (B) GSE106582 dataset; (C) GSE117606 dataset; (D) GSE31737 dataset; (E) GSE74602 dataset; (F) Clinical samples by RT-PCR assay.

and that the TMEscore signature derived from TIICs could predict patient survival rate. In addition, the genes representative of the TMEscore signature could better predict the response to chemotherapy, rather than patient survival rates.

Previous studies have shown that TIICs are associated with chemotherapy resistance. The cytokine IL-22, which is produced by T and natural killer cells, protects CRC cells from chemotherapy by activating the STAT3 pathway and inducing the autocrine expression of IL-8 (34). A study described that 5-fluorouracil-based chemotherapy regimens increase the expression of CXCR2 and ligand CXCL7 in liver metastasis of CRC, thus explaining the aggressiveness of relapsed drug-resistant tumors (35). The TMEscore, a novel index that characterizes TIICs, is associated with the survival of patients with several cancer types. Since the TMEscore is calculated by the DEGs from different clusters, the TMEscore-based signature generally has a more reliable predictive value compared to that of individual TIIC fractions. Studies have also reported that the TMEscore could predict the responses to several immunotherapies in neuroblastoma (36) and breast cancer (37).

Although the TMEscore is closely associated with the response to chemotherapy or immunotherapy, the clinical application value of the TMEscore remains to be determined. In this study, by overlapping the results from the three machine learning methods, three genes (ADH1C, SLC26A2, and NANS) were identified that showed good predictive value regarding the response to chemotherapy. ADH1C has been shown to inhibit the progression of CRC through the ADH1C/PHGDH/PSAT1/serine metabolic pathway (38), and downregulation of ADH1C has been associated with poor prognosis in patients with CRC (39). Higher SLC26A2 expression in tumor tissues indicates a longer survival for patients with CRC (40). NANS, together with seven other genes, could better predict the survival rate of patients with CRC (41). The results also suggested that these three genes participate in the development of CRC, and can be used to predict the prognosis of CRC patients. However, the role of these three genes in chemotherapy resistance have previously not been investigated. Thus, future relevant studies may further uncover the mechanism of chemotherapy resistance of each gene, and their association with the immune cells in CRC is worth further exploration.

Compared to previous studies (42–44), this study identified the prognostic value of the TMEscore in CRC and the predictive value of TMEscore-related genes in the response to fluoropyrimidine-based chemotherapy, which has not previously been reported. Moreover, our results were derived from a gene dataset with large cohorts of public databases, and

validated in several independent cohorts, thereby guaranteeing the robustness of the results. Furthermore, the results from the individual analysis of each gene were consistent with the overall analysis, which makes the utility of our findings in clinical decision-making easier, and doctors can estimate the prognosis of CRC patients based on fewer gene expressions, and administer appropriate treatment.

However, there were several limitations in this study. Firstly, the GSE72970 dataset used in the analysis only includes patients with metastatic CRC who receiving fluoropyrimidine-based chemotherapy as first-line therapy; therefore, the results might not be representative of non-metastatic CRC. Secondly, the fraction of TIICs obtained were from microarray datasets, and the exact number of TIICs will still need to be determined by the immunohistochemistry method or flow cytometry. Thirdly, the predictive value of the TMEscore in this study was verified using microarray datasets; thus, the exact predictive value will still need to be validated in a larger clinical cohort. Finally, more experiments are needed to confirm the association of the key genes with resistance to chemotherapy agents.

## Conclusion

In the present study, the TIIC-based TMEscore was associated with the survival rate of patients with CRC who underwent fluoropyrimidine-based chemotherapy and could predict the response to chemotherapy. Furthermore, the TMEscore-related gene signature had a better predictive value for the response to chemotherapy compared to that of survival rates.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#). Further inquiries can be directed to the corresponding authors.

## Ethics statement

This study was reviewed and approved by Ethics Committee of Guangxi Medical University Cancer Hospital. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

Study concept and design: X-WG, ZD and RL. Collection and assembly of data: X-WG and B-LH. Performed the experiment: S-QL, B-LH and X-WG. Data analysis and interpretation: S-QL, B-LH and X-WG. Manuscript writing and review: All authors. All authors contributed to the article and approved the submitted version.

## Funding

This study was partially supported by research funding from the National Natural Science Foundation of China (No. 82060104; No. 81860417; No. 82160446).

## Acknowledgments

We thank Dr. Jianming Zeng (University of Macau), and all the members of his bioinformatics team, biotrainee, for generously sharing their experience and codes.

## References

- Virk GS, Jafri M, Mehdi S, Ashley C. Staging and survival of colorectal cancer (CRC) in octogenarians: Nation-wide study of US veterans. *J Gastrointest Oncol* (2019) 10(1):12–8. doi: 10.21037/jgo.2018.09.01
- Juez I, Rubio C, Figueras J. Multidisciplinary approach of colorectal liver metastases. *Clin Transl Oncol* (2011) 13(10):721–7. doi: 10.1007/s12094-011-0722-x
- Kanemitsu Y, Shitara K, Mizusawa J, Hamaguchi T, Shida D, Komori K, et al. Primary tumor resection plus chemotherapy versus chemotherapy alone for colorectal cancer patients with asymptomatic, synchronous unresectable metastases (JCOG1007; iPACS): A randomized clinical trial. *J Clin Oncol* (2021) 39(10):1098–107. doi: 10.1200/JCO.20.02447
- Mansoori B, Mohammadi A, Davudian S, Shirjang S, Baradaran B. The different mechanisms of cancer drug resistance: A brief review. *Adv Pharm Bull* (2017) 7(3):339–48. doi: 10.15171/apb.2017.041
- Poulikakos PI, Sullivan RJ, Yaeger R. Molecular pathways and mechanisms of BRAF in cancer therapy. *Clin Cancer Res* (2022). doi: 10.1158/1078-0432.CCR-21-2138
- Ge R, Wang Z, Cheng L. Tumor microenvironment heterogeneity an important mediator of prostate cancer progression and therapeutic resistance. *NPJ Precis Oncol* (2022) 6(1):31. doi: 10.1038/s41698-022-00272-w
- Anderson NM, Simon MC. The tumor microenvironment. *Curr Biol* (2020) 30(16):R921–5. doi: 10.1016/j.cub.2020.06.081
- Malla R, Padmaraju V, Kundrapu DB. Tumor-associated macrophages: Potential target of natural compounds for management of breast cancer. *Life Sci* (2022) 8:120572. doi: 10.1016/j.lfs.2022.120572
- Cathcart AM, Smith H, Labrie M, Mills GB. Characterization of anticancer drug resistance by reverse-phase protein array: new targets and strategies. *Expert Rev Proteomics* (2022) 19(2):115–29. doi: 10.1080/14789450.2022.2070065
- Mo X, Huang X, Feng Y, Wei C, Liu H, Ru H, et al. Immune infiltration and immune gene signature predict the response to fluoropyrimidine-based chemotherapy in colorectal cancer patients. *Oncoimmunology* (2020) 9(1):1832347. doi: 10.1080/2162402X.2020.1832347
- Zhu X, Tian X, Ji L, Zhang X, Cao Y, Shen C, et al. A tumor microenvironment-specific gene expression signature predicts chemotherapy resistance in colorectal cancer patients. *NPJ Precis Oncol* (2021) 5(1):7. doi: 10.1038/s41698-021-00142-x

## 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/fonc.2022.953321/full#supplementary-material>

- Che LH, Liu JW, Huo JP, Luo R, Xu RM, He C, et al. A single-cell atlas of liver metastases of colorectal cancer reveals reprogramming of the tumor microenvironment in response to preoperative chemotherapy. *Cell Discovery* (2021) 7(1):80. doi: 10.1038/s41421-021-00312-y
- Zeng D, Wu J, Luo H, Li Y, Xiao J, Peng J, et al. Tumor microenvironment evaluation promotes precise checkpoint immunotherapy of advanced gastric cancer. *J Immunother Cancer* (2021) 9(8):e002467. doi: 10.1136/jitc-2021-002467
- Zhang N, Zhang H, Wang Z, Dai Z, Zhang X, Cheng Q, et al. Immune infiltrating cells-derived risk signature based on Large-scale analysis defines immune landscape and predicts immunotherapy responses in glioma tumor microenvironment. *Front Immunol* (2021) 12:691811. doi: 10.3389/fimmu.2021.691811
- Shen J, Liu T, Lv J, Xu S. Identification of an immune-related prognostic gene CLEC5A based on immune microenvironment and risk modeling of ovarian cancer. *Front Cell Dev Biol* (2021) 9:746932. doi: 10.3389/fcell.2021.746932
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337
- Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007
- Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst* (2015) 1(6):417–25. doi: 10.1016/j.cels.2015.12.004
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* (2021) 2(3):100141. doi: 10.1016/j.xinn.2021.100141
- Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol Res* (2019) 7(5):737–50. doi: 10.1158/2326-6066.CIR-18-0436

22. 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
23. Tang YP, Yin YX, Xie MZ, Liang XQ, Li JL, Li KZ, et al. Systematic analysis of the clinical significance of hyaluronan-mediated motility receptor in colorectal cancer. *Front Mol Biosci* (2021) 8:733271. doi: 10.3389/fmolb.2021.733271
24. Crisafulli G, Sartore-Bianchi A, Lazzari L, Pietrantonio F, Amatu A, Macagno M, et al. Temozolomide treatment alters mismatch repair and boosts mutational burden in tumor and blood of colorectal cancer patients. *Cancer Discovery* (2022) 12(7):1656–75. doi: 10.1158/2159-8290.CD-21-1434
25. Dong H, Hu L, Li W, Shi M, He L, Wang C, et al. Pyrimethamine inhibits cell growth by inducing cell senescence and boosting CD8(+) T-cell mediated cytotoxicity in colorectal cancer. *Mol Biol Rep* (2022) 49(6):4281–92. doi: 10.1007/s11033-022-07262-y
26. Wang C, Fakhri M. Response to trastuzumab and lapatinib in a metastatic colorectal cancer harboring HER2 amplification and HER2 S310F mutation. *J Natl Compr Canc Netw* (2021) 19(6):670–4. doi: 10.6004/jncn.2021.7023
27. Rembialkowska N, Novickij V, Baczynska D, Dubinska-Magiera M, Saczko J, Rudno-Rudzinska J, et al. Micro- and nanosecond pulses used in doxorubicin electrochemotherapy in human breast and colon cancer cells with drug resistance. *Molecules* (2022) 27(7):2052. doi: 10.3390/molecules27072052
28. Fogelman D, Cubillo A, Garcia-Alfonso P, Miron MLL, Nemunaitis J, Flora D, et al. Randomized, double-blind, phase two study of ruxolitinib plus regorafenib in patients with relapsed/refractory metastatic colorectal cancer. *Cancer Med* (2018) 7(11):5382–93. doi: 10.1002/cam4.1703
29. Hamzic S, Aebi S, Joergler M, Montemurro M, Ansari M, Amstutz U, et al. Fluoropyrimidine chemotherapy: recommendations for DPYD genotyping and therapeutic drug monitoring of the Swiss group of pharmacogenomics and personalised therapy. *Swiss Med Wkly* (2020) 150:w20375. doi: 10.4414/smw.2020.20375
30. Spada F, Antonuzzo L, Marconcini R, Radice D, Antonuzzo A, Ricci S, et al. Oxaliplatin-based chemotherapy in advanced neuroendocrine tumors: Clinical outcomes and preliminary correlation with biological factors. *Neuroendocrinology* (2016) 103(6):806–14. doi: 10.1159/000444087
31. He XS, Zou SY, Yao JL, Yu W, Deng ZY, Wang JR, et al. Transcriptomic analysis identifies complement component 3 as a potential predictive biomarker for chemotherapy resistance in colorectal cancer. *Front Mol Biosci* (2021) 8:763652. doi: 10.3389/fmolb.2021.763652
32. Sreekumar R, Al-Saihati H, Emaduddin M, Moutasim K, Mellone M, Patel A, et al. The ZEB2-dependent EMT transcriptional programme drives therapy resistance by activating nucleotide excision repair genes ERCC1 and ERCC4 in colorectal cancer. *Mol Oncol* (2021) 15(8):2065–83. doi: 10.1002/1878-0261.12965
33. Sfakianaki M, Papadaki C, Tzardi M, Trypaki M, Manolakou S, Messaritakis I, et al. PKM2 expression as biomarker for resistance to oxaliplatin-based chemotherapy in colorectal cancer. *Cancers (Basel)* (2020) 12(8):2058. doi: 10.3390/cancers12082058
34. Wu T, Wang Z, Liu Y, Mei Z, Wang G, Liang Z, et al. Interleukin 22 protects colorectal cancer cells from chemotherapy by activating the STAT3 pathway and inducing autocrine expression of interleukin 8. *Clin Immunol* (2014) 154(2):116–26. doi: 10.1016/j.clim.2014.07.005
35. Desurmont T, Skrypek N, Duhamel A, Jonckheere N, Millet G, Leteurtre E, et al. Overexpression of chemokine receptor CXCR2 and ligand CXCL7 in liver metastases from colon cancer is correlated to shorter disease-free and overall survival. *Cancer Sci* (2015) 106(3):262–9. doi: 10.1111/cas.12603
36. Feng C, Li T, Xiao J, Wang J, Meng X, Niu H, et al. Tumor microenvironment profiling identifies prognostic signatures and suggests immunotherapeutic benefits in neuroblastoma. *Front Cell Dev Biol* (2022) 10:814836. doi: 10.3389/fcell.2022.814836
37. Xu L, Hu Y, Liu W. Tumor microenvironment-mediated immune profiles characterized by distinct survival outcome and immunotherapeutic efficacy in breast cancer. *Front Genet* (2022) 13:840348. doi: 10.3389/fgene.2022.840348
38. Li S, Yang H, Li W, Liu JY, Ren LW, Yang YH, et al. ADH1C inhibits progression of colorectal cancer through the ADH1C/PHGDH/PSAT1/serine metabolic pathway. *Acta Pharmacol Sin* (2022). doi: 10.1038/s41401-022-00894-7
39. Li M, Liu Z, Song J, Wang T, Wang H, Wang Y, et al. Identification of down-regulated ADH1C is associated with poor prognosis in colorectal cancer using bioinformatics analysis. *Front Mol Biosci* (2022) 9:791249. doi: 10.3389/fmolb.2022.791249
40. Pitule P, Vycital O, Bruha J, Novak P, Hosek P, Treska V, et al. Differential expression and prognostic role of selected genes in colorectal cancer patients. *Anticancer Res* (2013) 33(11):4855–65.
41. Ren J, Feng J, Song W, Wang C, Ge Y, Fu T. Development and validation of a metabolic gene signature for predicting overall survival in patients with colon cancer. *Clin Exp Med* (2020) 20(4):535–44. doi: 10.1007/s10238-020-00652-1
42. Wang X, Xu Z, Liu Z, Lin W, Cao Z, Feng X, et al. Characterization of the immune cell infiltration landscape uncovers prognostic and immunogenic characteristics in lung adenocarcinoma. *Front Genet* (2022) 13:902577. doi: 10.3389/fgene.2022.902577
43. Zheng H, Liu H, Li H, Dou W, Wang J, Zhang J, et al. Characterization of stem cell landscape and identification of stemness-relevant prognostic gene signature to aid immunotherapy in colorectal cancer. *Stem Cell Res Ther* (2022) 13(1):244. doi: 10.1186/s13287-022-02913-0
44. Shi J, Bao M, Wang W, Wu X, Li Y, Zhao C, et al. Integrated profiling identifies PLOD3 as a potential prognostic and immunotherapy relevant biomarker in colorectal cancer. *Front Immunol* (2021) 12:722807. doi: 10.3389/fimmu.2021.722807