



# Construction of an Immunogenomic Risk Score for Prognostication in Colon Cancer

Han Zhang<sup>1,2†</sup>, Chuan Qin<sup>3†</sup>, Hua Gan<sup>1\*</sup>, Xiong Guo<sup>1</sup> and Li Zhang<sup>2</sup>

<sup>1</sup> First Clinical Medical College, Chongqing Medical University, Chongqing, China, <sup>2</sup> Department of Digestive Oncology, Three Gorges Hospital, Chongqing University, Chongqing, China, <sup>3</sup> Department of Gastrointestinal Surgery, Three Gorges Hospital, Chongqing University, Chongqing, China

## OPEN ACCESS

### Edited by:

Robert Friedman,  
University of South Carolina,  
United States

### Reviewed by:

Xinyan Zhang,  
Georgia Southern University,  
United States  
Soichiro Ishihara,  
The University of Tokyo, Japan

Hao Zhang,  
Jilin University, China

### \*Correspondence:

Hua Gan  
ganhua113@sohu.com

<sup>†</sup>Members of the Collaborative  
Working Group

### Specialty section:

This article was submitted to  
Bioinformatics and Computational  
Biology,  
a section of the journal  
Frontiers in Genetics

**Received:** 03 February 2020

**Accepted:** 22 April 2020

**Published:** 21 May 2020

### Citation:

Zhang H, Qin C, Gan H, Guo X  
and Zhang L (2020) Construction  
of an Immunogenomic Risk Score  
for Prognostication in Colon Cancer.  
*Front. Genet.* 11:499.  
doi: 10.3389/fgene.2020.00499

Immune-related genes (IRGs) play regulatory roles in the immune system and are involved in the initiation and progression of colon cancer. This study aimed to develop an immunogenomic risk score for predicting survival outcomes among colon cancer patients. We analyzed the expressions of IRGs in colon specimens and discovered 484 differentially expressed IRGs when we compared specimens from colon cancer and adjacent normal tissue. Univariate Cox regression analyses were performed to identify 26 IRGs that were associated with survival. A Cox proportional hazards model with a lasso penalty identified five optimal IRGs for constructing the immunogenomic risk score (*CD1B*, *XCL1*, *PLCG2*, *NGF*, and *OXTR*). The risk score had good performance in predicting overall survival among patients with colon cancer and was correlated with the amount of tumor-infiltrating immune cells. Our findings suggest that the immunogenomic risk score may be useful for prognostication in colon cancer cases. Furthermore, the five IRGs included in the risk score might be useful targets for investigating the initiation of colon cancer and designing personalized treatments.

**Keywords:** immune-related genes, immune system, colon cancer, prognostication, biomarkers

## INTRODUCTION

Colorectal cancer (CRC) is the third most common malignancy and the second leading cause of cancer-related death, with an estimated 1,800,000 new diagnoses and 881,000 deaths in 2018 (Bray et al., 2018). Despite advances in diagnosis and treatment, the 5-year survival rate remains approximately 57% (Holleczek et al., 2015), which highlights the need for strategies to provide better outcomes. The tumor-node-metastasis (TNM) system is an essential prognostic tool for guiding treatment selection (Hu et al., 2011; Kim et al., 2011; Park et al., 2012), although some patients with the same disease stage experience different survival outcomes, which is related to molecular heterogeneity. Therefore, it is critical to identify biomarkers that can help predict the risks of recurrence and death, which can facilitate early interventions and improve outcomes among patients with colon cancer.

There is increasing evidence that dysregulation of the immune system is involved in the initiation and progression of cancer (Fridman et al., 2012; Janssen et al., 2017). For example, cancer cells can escape immune system recognition and elimination by upregulating and downregulating immune-related genes (IRGs), which help promote tumor growth (de Vries et al., 2016). Many reports have also confirmed that IRGs are attractive targets for regulating cancer progression

(Weiß and Schnieders, 2017; Cabrero-de Las Heras and Martínez-Balibrea, 2018). Therefore, IRG-based features may be useful for prognostication in CRC cases.

Several studies have evaluated the relationships between IRG-based genetic signatures and the prognosis of CRC (Ge et al., 2019; Wu et al., 2019). However, there are significant differences between colon and rectal cancers in terms of their embryological origin, anatomy, and functional implications (Tamas et al., 2015; Riihimäki et al., 2016; Paschke et al., 2018). Tumor-based heterogeneity is also apparent between individuals and between tumor sites (Imperial et al., 2018), and we suspect that IRG expression profiles also vary between colon and rectal cancers. However, we are not aware of any studies that have evaluated the characteristics and regulatory mechanisms of IRGs in colon cancer. Therefore, this study aimed to evaluate IRGs in colon cancer and to develop an immunogenomic risk score to predict the prognosis of patients with colon cancer.

## MATERIALS AND METHODS

### Data Retrieval

Transcriptome expression profiles for colon cancer samples were downloaded from The Cancer Genome Atlas data portal (TCGA)<sup>1</sup> (Liu et al., 2018). Clinical data for the corresponding patients were also retrieved from the database, which included sex, age, tumor stage, and survival information. Patients without survival data or with <30 days of data were excluded because they might have died because of lethal complications (e.g., digestive tract infection or hemorrhage), rather than colon cancer. The expression of protein-coding genes were annotated in the TCGA data portal based on fragments per kilobase of transcript per million mapped reads (FPKM).

Data regarding 2,498 IRGs were downloaded from the Immunology Database and Analysis Portal database<sup>2</sup> (Bhattacharya et al., 2018). Data regarding cancer-associated transcription factors (TFs) were also obtained from the Cistrome project<sup>3</sup> (Mei et al., 2017). Immune infiltrate data were collected from the Tumor Immune Estimation Resource<sup>4</sup> (Li et al., 2017), which contains information regarding the relative proportions of six types of tumor-infiltrating immune cells (B-cells, CD4<sup>+</sup> T-cells, CD8<sup>+</sup> T-cells, neutrophils, macrophages, and dendritic cells).

### Differential Gene Analysis

Genes and TFs that were differentially expressed between colon cancer specimens and adjacent normal colon specimens were identified using the “limma” package for R software, with a |log<sub>2</sub> fold-change [log<sub>2</sub>FC]| of >1 and an adjusted false-discovery rate (FDR) of <0.05. Differentially expressed IRGs were identified among the differentially expressed genes, with expression patterns of significant differentially expressed genes and IRGs

visualized using heatmaps and volcano plots. These elements were created using the “pheatmap” package for R software.

### Functional Enrichment Analysis of the Differentially Expressed IRGs

Biological processes, molecular functions, and cellular components that were potentially associated with the differentially expressed IRGs were evaluated using Gene Ontology data (Ashburner et al., 2000) from the Database for Annotation, Visualization and Integrated Discovery (version 6.8)<sup>5</sup>. These data were analyzed using the “gplot” package for R software. Enrichment analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (Kanehisa et al., 2017) and data from the KOBAS database (version 3.0)<sup>6</sup> (Xie et al., 2011). These data were analyzed using the “clusterProfiler” package for R software.

### Construction of a TF Regulatory Network

Regulatory mechanisms were evaluated by screening TFs that were differentially expressed between colon cancer and normal colon tissues (*P*-values of <0.05). We identified clinically relevant TFs based on a correlation assessment (correlation coefficient >0.4), and constructed a regulatory network of the relevant IRGs and potential TFs using Cytoscape software (version 3.7.2) (Su et al., 2014).

### Development and Validation of the Risk Score

The patients were randomly assigned to a training dataset and a testing dataset. The training dataset was used to develop the risk score. Survival-associated IRGs were identified using univariate Cox analyses and the “survival” package for R software (*P*-value <0.01). Next, a Cox proportional hazards model with a lasso penalty was used to identify the genetic model with the best prognostic value, which was performed using the “glmnet” and “survival” packages for R software. The risk score was then created based on the gene model, using the Gene Expression Profiling Interactive Analysis website<sup>7</sup> (Tang et al., 2017) to analyze the expression of each IRG that was included in the risk score. The risk score was calculated as:

$$\text{Risk score} = \sum_{i=1}^N (\text{Exp}_i \times \text{Coef}_i)$$

with *N* representing the number of signature genes, *Exp<sub>i</sub>* representing the gene expression levels, and *Coef<sub>i</sub>* representing the estimated regression coefficient value from the Cox proportional hazards analysis.

The risk score’s predictive value was evaluated using receiver operating characteristic (ROC) curve analysis in the training and testing datasets. For that analysis, an area under the ROC curve (AUC) of >0.75 was judged to have excellent predictive value.

<sup>1</sup><https://cancergenome.nih.gov/>

<sup>2</sup><https://immport.niaid.nih.gov>

<sup>3</sup><http://cistrome.org/>

<sup>4</sup><https://cistrome.shinyapps.io/timer/>

<sup>5</sup><http://david.ncifcrf.gov/home.jsp>

<sup>6</sup><http://kobas.cbi.pku.edu.cn/kobas3>

<sup>7</sup><http://gepia.cancer-pku.cn/>

The predictive value was also evaluated by grouping patients into high-risk and low-risk groups (based on the median risk score), with univariate and multivariate analyses of survival then performed for the risk score and clinical factors. We also evaluated the relationships of the risk score with clinical data and tumor-infiltrating immune cells using a correlation assessment (significant at a  $P$ -value of  $<0.01$ ).

## Statistical Analysis

The data were analyzed using R software (version 3.6.1)<sup>8</sup>. Differentially expressed genes were identified using the Wilcoxon test. Survival analyses were performed using the Kaplan-Meier method and the log-rank test.

## RESULTS

### Identification of Differentially Expressed IRGs

We screened the expression levels of 60,483 genes in colon cancer specimens ( $n = 398$ ) and normal colon tissues ( $n = 39$ ). This screening identified 6,501 differentially expressed genes, including 4,478 genes that were upregulated and 2,023 genes that were downregulated, relative to the levels in normal colon tissues (log<sub>2</sub> fold-change [logFC] of  $>1.0$ , FDR of  $<0.05$ ) (Figures 1A,C). We searched these differentially expressed genes and identified 484 differentially expressed IRGs, including 173 upregulated IRGs and 311 downregulated IRGs (logFC of  $>1.0$ , FDR of  $<0.05$ ) (Figures 1B,D). All of the IRGs are protein-coding genes. The Gene Ontology analyses of differentially expressed IRGs revealed that “immune response” was the most common biological process, “extracellular region” was the most common cellular component, and “antigen-binding” was the most common molecular function (Figure 2A). The Kyoto Encyclopedia of Genes and Genomes pathway analyses revealed that most of the IRGs played roles in cytokine-cytokine receptor interactions (Figure 2B).

### Identification of Survival-Associated IRGs

Univariate Cox regression analyses were performed for the differentially expressed IRGs, which revealed that 26 IRGs were significantly associated with overall survival among colon cancer patients ( $P < 0.01$ ) (Table 1). Their regulatory mechanisms were evaluated using mRNA levels of TFs in colon cancer and normal colon tissue, which revealed 71 differentially expressed TFs (logFC of  $>1.0$ , FDR of  $<0.05$ ) (Figures 3A,B). Nine of these TFs were associated with overall survival among colon cancer patients. Based on these findings, we created a regulatory network using the 9 TFs and 26 IRGs that were associated with survival (Figure 3C).

### Construction and Validation of the Prognostic Risk Score

After excluding patients without survival data or with  $<30$  days of data, we separated the remaining patients into a training dataset ( $n = 251$ , 70%) and a testing dataset ( $n = 104$ , 30%). The risk score was developed using the training dataset and validated using the testing dataset.

Screening of the survival-associated IRGs identified five relevant IRGs (*CD1B*, *XCL1*, *PLCG2*, *NGF*, and *OXTR*) via a Cox proportional hazards model, which were used to develop the risk score. The *CD1B* gene was considered a protective gene (coefficient of  $-3.62$ ) and the other four IRGs were considered risk genes. The risk score for each patient was calculated as: risk score =  $(-3.62 \times \text{expression of } CD1B) + (0.492 \times \text{expression of } XCL1) + (0.52 \times \text{expression of } PLCG2) + (0.876 \times \text{expression of } NGF) + (0.203 \times \text{expression of } OXTR)$ . The expression level of each IRG was evaluated using the samples from the TCGA (Figure 4).

The prognostic value of the risk score was evaluated using ROC curves for the training dataset. The AUC values were 0.788 for predicting 1-year survival and 0.787 for predicting 3-year survival (Figures 5A,C). The risk score appeared to provide excellent prognostic value relative to the AUC values for predicting survival based on tumor stage (1-year survival: 0.813, 3-year survival: 0.771). Validation using the testing dataset revealed AUC values of 0.807 for predicting 1-year survival and 0.710 for predicting 3-year survival (Figures 5B,D).

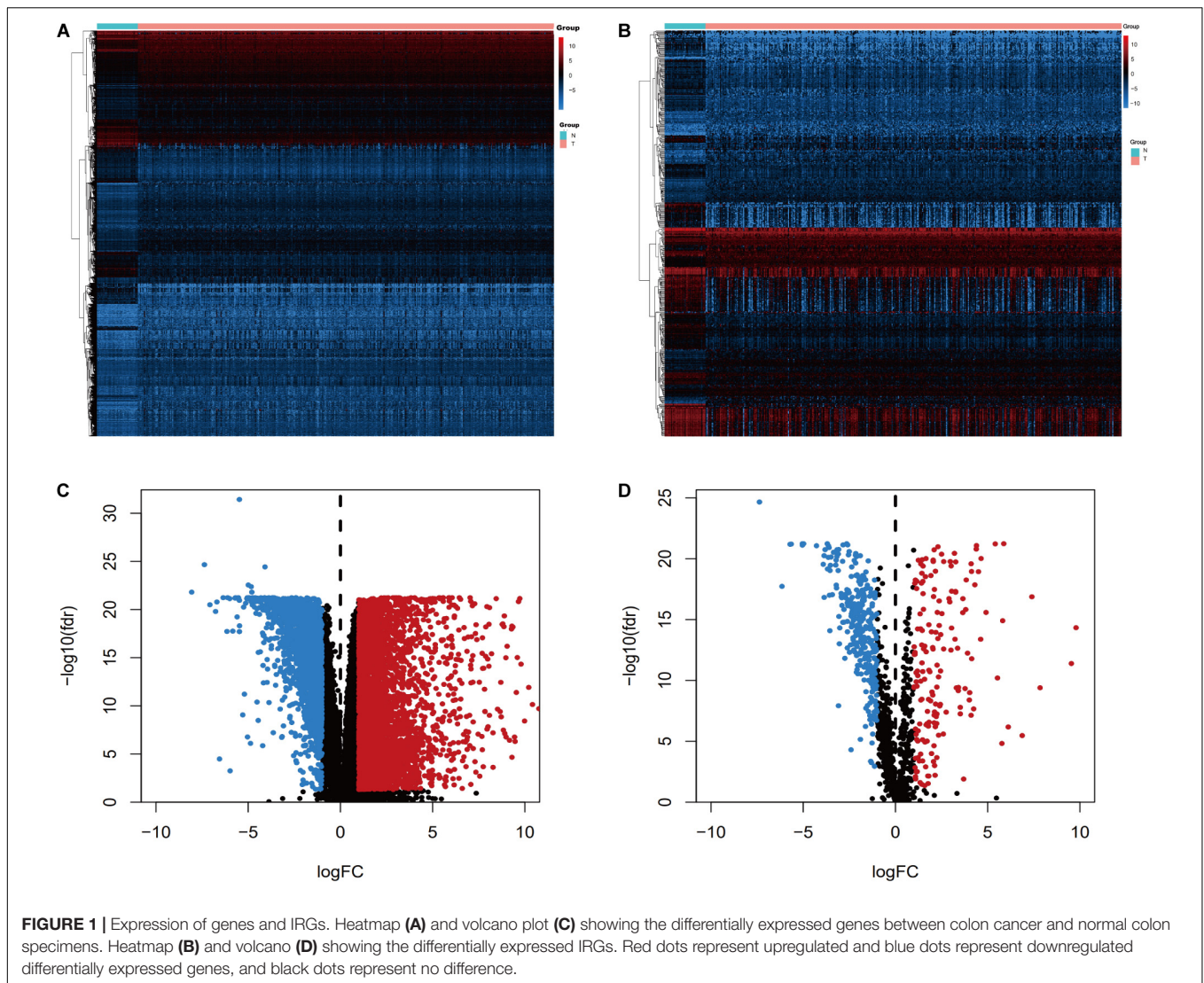
Patients from the training set were assigned to a high-risk group ( $n = 125$ ) and a low-risk group ( $n = 126$ ) based on the median risk score. The Kaplan-Meier survival curves revealed significantly better survival in the low-risk group than in the high-risk group ( $P = 1.97e-04$ ) (Figure 6A). Patients from the testing set were also divided according to risk score (high-risk: 46 patients vs. low-risk: 58 patients), and the survival analyses also revealed significantly better survival in the low-risk group ( $P = 9.483e-03$ ) (Figure 6B).

The risk score distributions, survival statuses, and risk gene expressions in the training and testing datasets are shown in Figures 7A–F. The high-risk group had clearly higher values for the risk score and the mortality rate. In contrast, the low-risk group had significantly higher expression of the protective gene (*CD1B*) and lower expression of the four risk genes. These results indicated that our risk model was capable of accurately predicting the prognosis of colon cancer patients. Multivariate Cox regression analyses also revealed that the risk score was an independent predictor of survival in the training and testing datasets, after adjusting for age, sex, and tumor stage (Table 2).

### Clinical Utility of the Prognostic Risk Score

We also evaluated whether the risk score could predict progression of colon cancer by evaluating the IRGs' relationships with clinical variables (age, sex, and tumor stage) (Table 3). The expression of *CD1B* was significantly lower in cases involving advanced-stage disease, lymph node metastasis, and distant

<sup>8</sup><http://www.Rproject.org>



metastasis (Figures 8A–C). The expression of *OXTR* was notably higher in cases with an advanced T classification (Figure 8D). The expression of *NGF* was markedly higher in cases involving elderly patients, advanced-stage disease, and lymph node metastasis (Figures 8E–G).

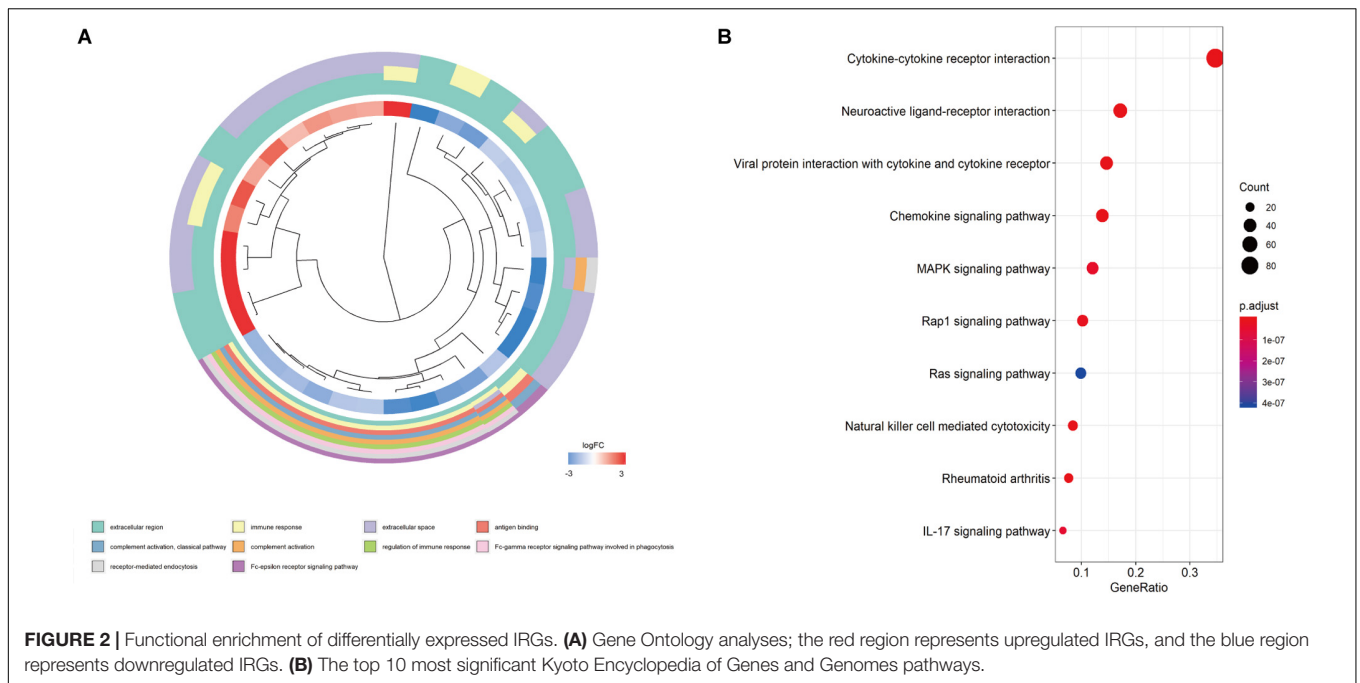
Furthermore, we evaluated whether the risk score could reflect the tumor microenvironment, based on the correlation between the risk score and immune cell infiltration in the colon cancer patients. Higher risk scores were correlated with increasing values for tumor-infiltrating immune cells, which included B-cells, CD4<sup>+</sup> T-cells, CD8<sup>+</sup> T-cells, neutrophils, macrophages, and dendritic cells (all  $P < 0.05$ ) (Figure 9).

## DISCUSSION

This study revealed that 5 IRGs were associated with survival among patients with colon cancer. To the best of our knowledge, this is the first study to construct a reliable

immunogenomic model for predicting overall survival among patients with colon cancer.

Immune evasion is an indispensable step in cancer development and this process is mediated by various IRGs (Concha-Benavente and Ferris, 2017), which prompted us to consider the role of IRGs in prognostication. We screened differentially expressed IRGs in colon cancer and identified 173 upregulated IRGs and 311 downregulated IRGs. These results indicated that IRGs were closely related to the development of colon cancer, which agrees with previously reported data. As expected, the Gene Ontology analyses revealed that the most common biological processes for the IRGs involved “immune response”, while the Kyoto Encyclopedia of Genes and Genomes analyses revealed that the differentially expressed IRGs were related to cytokine-cytokine receptor interactions. In this context, cytokines are mainly expressed by immune cells and tumor cells, and can alter anti-tumor immunity and tumor progression (Zhang et al., 2018; Weinstein et al., 2019). Some cytokines have also been used for cancer therapy, such as IL-8



**TABLE 1 |** General characteristics of survival-associated IRGs in patients with colon cancer.

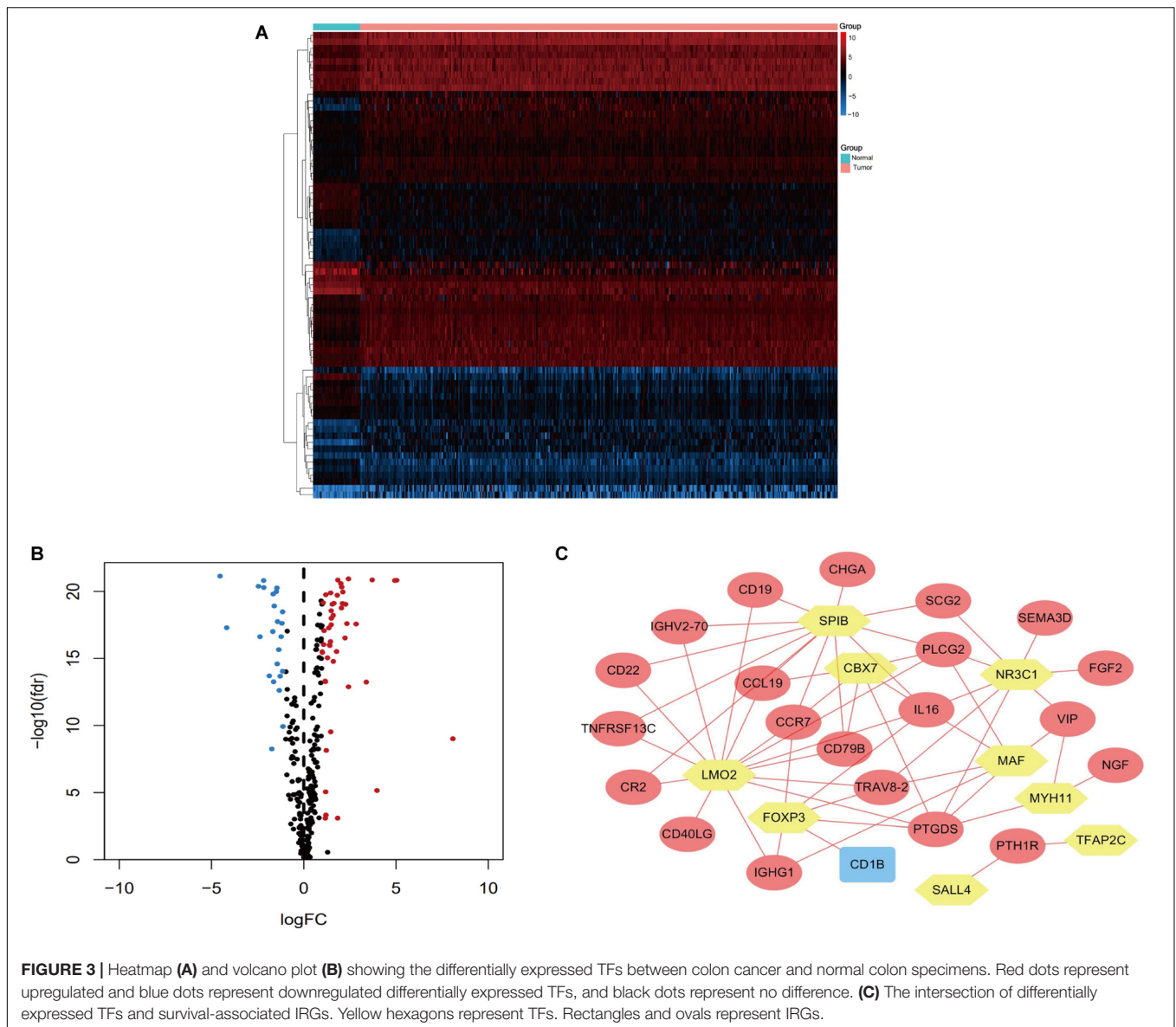
Gene symbol	HR	HR.95L	HR.95H	P-value
PLCG2	1.78771957	1.329491	2.403883	0.000121
CHGA	1.01049537	1.005125	1.015895	0.000123
OXTR	1.40633083	1.181286	1.674249	0.000127
CD19	1.30216573	1.134585	1.494498	0.000172
CCL19	1.0235064	1.011097	1.036068	0.000189
CD79B	1.12218511	1.05124	1.197919	0.000541
TNFRSF13C	1.46244744	1.178929	1.814149	0.000546
CD22	1.21130031	1.086407	1.350552	0.000555
XCL1	2.15293999	1.381475	3.35522	0.000705
VIP	1.05870861	1.023354	1.095285	0.000994
FGF2	1.47456065	1.166011	1.864759	0.001186
TRAV8-2	3.32850327	1.583458	6.996669	0.001511
CR2	1.06879614	1.025544	1.113872	0.001596
IGHG1	1.00060367	1.000221	1.000986	0.00198
INHBE	3.61792417	1.593117	8.216207	0.002121
FGF9	2.42824902	1.358618	4.339995	0.00275
PTH1R	1.53935823	1.142105	2.074787	0.00462
IGHV2-70	1.01935188	1.005701	1.033188	0.005331
CGR7	1.17300899	1.046284	1.315083	0.006226
CD1B	0.04002157	0.003874	0.41346	0.006908
IL16	1.62583077	1.140763	2.317156	0.007178
SCG2	1.11004978	1.027623	1.199089	0.007999
SEMA3D	2.07808955	1.204535	3.585164	0.008569
PTGDS	1.03559807	1.008817	1.06309	0.00888
NGF	2.86414111	1.290971	6.354366	0.00965
CD40LG	1.7639811	1.14685	2.713197	0.009775

HR, hazard ratio.

(Ning and Lenz, 2012) and IL-6 (Wang et al., 2019). Thus, our findings may provide a basis for future studies regarding ways to target these IRGs.

Previous studies have indicated that TF dysregulation, which significantly modifies gene expressions, was related to the progression of colon cancer (Laissue, 2019). For example, the FOXE1 TF is a critical tumor inhibitor that regulates tumor growth and glycolysis by suppressing HK2 in CRC (Dai et al., 2020). The ZBP-89 TF can also drive a feed-forward loop involving  $\beta$ -catenin expression in CRC (Essien et al., 2016). Thus, to explore the regulatory mechanisms for the IRGs, we identified differentially expressed TFs in colon cancer and established a network involving the differentially expressed TFs and IRGs. We identified nine important TFs, with CBX7, FOXP3, LMO2, MAF, MYH11, NR3C1, and SPIB dominating the network. These results suggested that TFs could influence the effects of IRGs on survival outcomes.

Among the differentially expressed IRGs, we identified 5 IRGs (*CD1B*, *XCL1*, *PLCG2*, *NGF*, and *OXTR*) for inclusion in the risk score, and previous studies have indicated that these genes were related to immune processes and cancer progression. The stability of the risk score was evaluated based on the expression level of each IRG, which were stably expressed in the colon specimens. The *CD1B* gene encodes a member of the Group 1 CD1 family of transmembrane glycoproteins, which present a wide array of self and foreign lipid antigens to T-cell receptors (Shahine, 2018). Previous reports have indicated that *CD1B*-restricted self-lipid-reactive T-cells responded more potently to lipid from tumor cells than to an equivalent amount of lipids from normal cells, and the adoptive transfer of these T-cells into mice resulted in tumor control (Bagchi et al., 2016). Furthermore, *CD1B* expression was related to the prognosis for localized prostate cancer (Lee et al., 2019), and *CD1B* was detectable within mononuclear cells from liver tumor specimens, but was not expressed in healthy livers (Kenna et al., 2007).

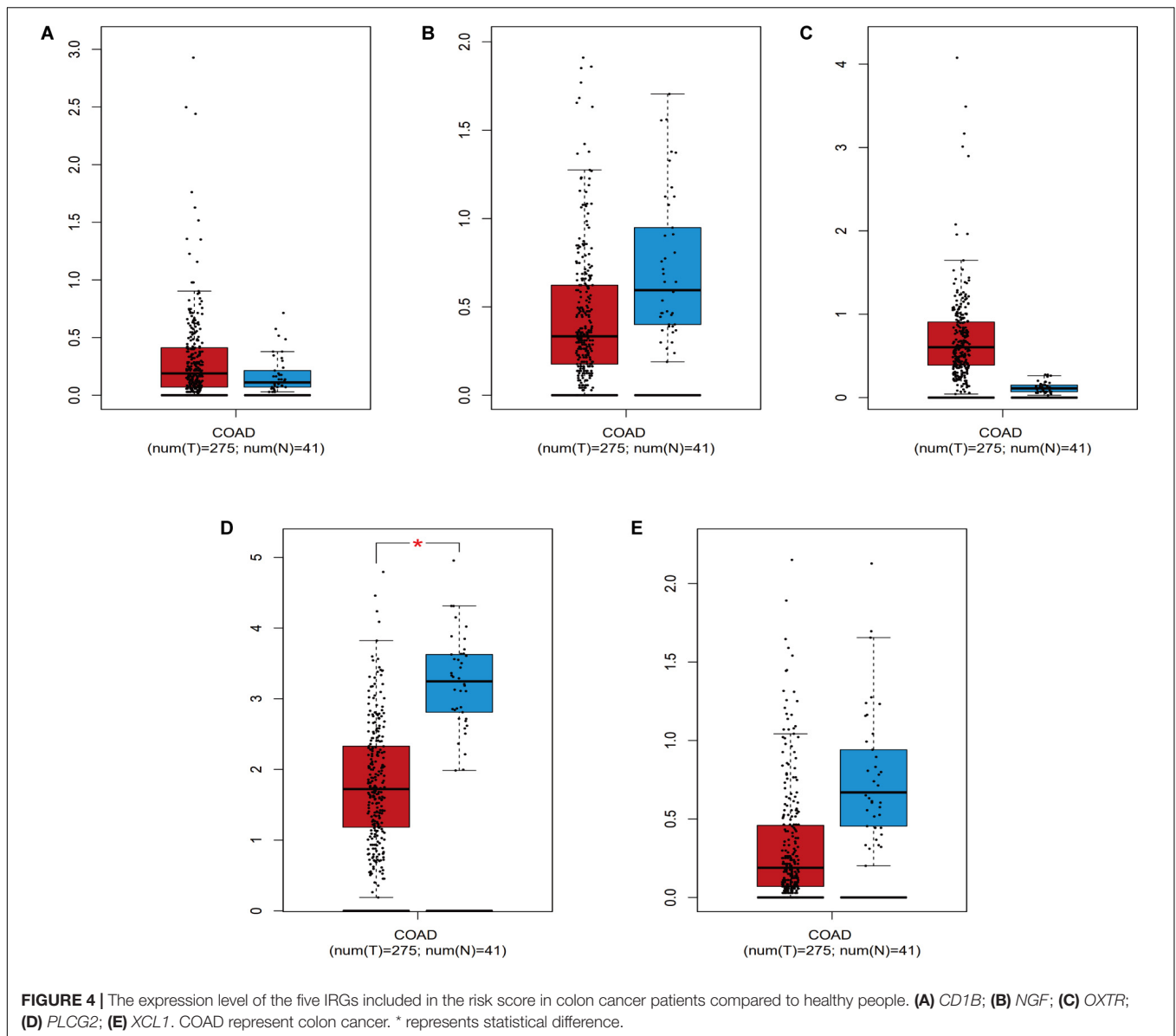


The *XCL1* gene encodes a C class chemokine that is also known as lymphotactin (Lei and Takahama, 2012). The ability of *XCL1* to contribute to the anti-tumor activity via attracting DC1s has been demonstrated in a mouse tumor model (Böttcher et al., 2018). In addition, *XCL1* expression is related to the number of tumor-infiltrating CD8<sup>+</sup> T-cells and PD-L1 expression on ovarian tumor cells, which indicated that *XCL1* might be a biomarker for anti-PD1/PD-L1 immunotherapy (Tamura et al., 2020).

The *PLCG1* gene encodes an intracellular signaling molecule that is positioned at the convergence of various signaling pathways for cell proliferation, migration, and invasion (Wells and Grandis, 2003). Dysfunction of *PLCG1* is closely associated with inflammation, immune disorders, and cancer (Koss et al., 2014), and *in vitro* testing revealed that *PLCG1* mediated high glucose levels and insulin-induced cell proliferation and migration in SW480 colon cancer cells (Tomas et al., 2012). The

expression of *PLCG1* also promoted hepatoma cell carcinogenesis *in vitro* and *in vivo* (Tang W. et al., 2019).

The *NGF* gene encodes a growth factor that can be released into or produced by the tumor microenvironment (Bradshaw et al., 2015). Expression of *NGF* influences the activities of various immune cells, including macrophages, granulocytes, T-cells, B-cells, NK cells, and eosinophils. These cells can also synthesize, store, and release consistent amounts of NGF, which suggests that NGF may influence the anti-tumor immune response (Aloe et al., 2016). Moreover, NGF activates breast cancer stem cells through the promotion of epithelial-mesenchymal transition and by increasing the number of symmetric divisions, which indicates that NGF is involved in the self-renewal and plasticity of cancer stem cells (Tomellini et al., 2015). Expression of *NGF* also increases angiogenesis via the COX-2/PGE2 signaling axis in epithelial ovarian cancer (Garrido et al., 2019).

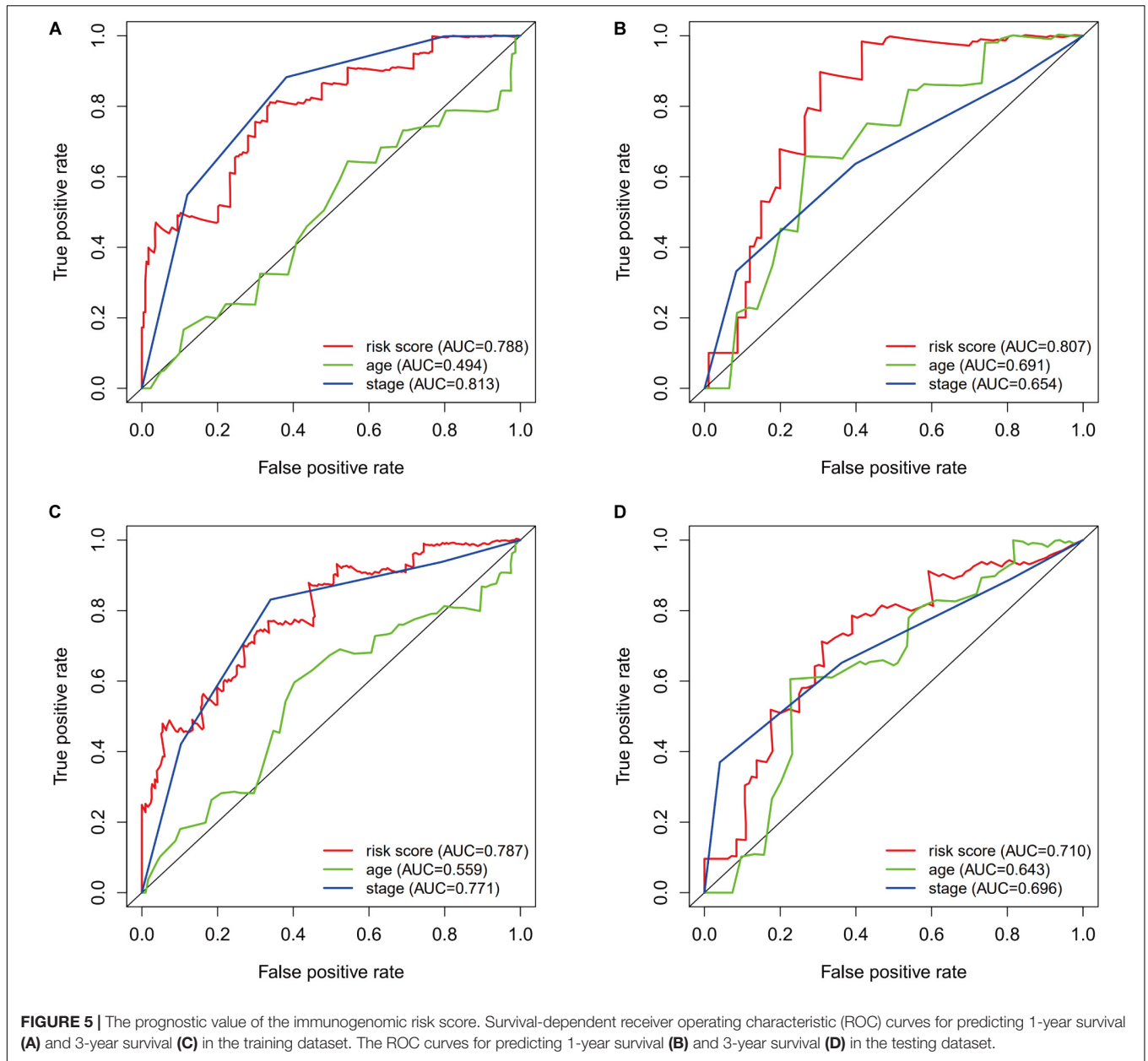


The *OXTR* gene encodes the oxytocin receptor, and the oxytocin/*OXTR* axis is known to decrease the sensitivity of macrophages to lipopolysaccharides, with lower expression of inflammatory cytokines (e.g., IL-1 $\beta$ , IL-6, and TNF- $\alpha$ ) and to increase the sensitivity to IL-4 stimulation in the intestinal microenvironment (Tang Y. et al., 2019). Thus, the oxytocin/*OXTR* axis may be useful for inhibiting CRC development via down-regulating immunosuppressive proteins (FAP $\alpha$  and CCL-2) (Ma et al., 2019).

The risk score provided AUC values of 0.788 for predicting 1-year survival and 0.787 for predicting 3-year survival, which indicated the score had good prognostic ability. We also divided the patients according to the median risk score, and observed remarkable differences in their survival curves, which confirmed that the risk score was useful for identifying patients with a high risk of death. Nevertheless, medical improvements have

allowed even some advanced cancer patients to achieve long-term survival (Li et al., 2019), and frequent follow-up with active management may still be useful for patients with stage IV colon cancer (Karoui et al., 2011; Becerra et al., 2015). In clinical practice, tumor staging can be used to guide patient classification and more personalized treatment (Lu et al., 2015; Babaei et al., 2018), although patients with the same disease stage can still experience different outcomes (Mayanagi et al., 2018). Thus, for patients with the same colon cancer stage, our risk score may be useful as an auxiliary tool to identify patients with a high risk of death.

We evaluated the relationship between the risk score and tumor-infiltrating immune cells, which revealed a positive correlation with the number of infiltrating immune cells. In this context, tumor immune evasion is mediated by B-cells, CD4<sup>+</sup> T-cells, CD8<sup>+</sup> T-cells, neutrophils, macrophages, and dendritic

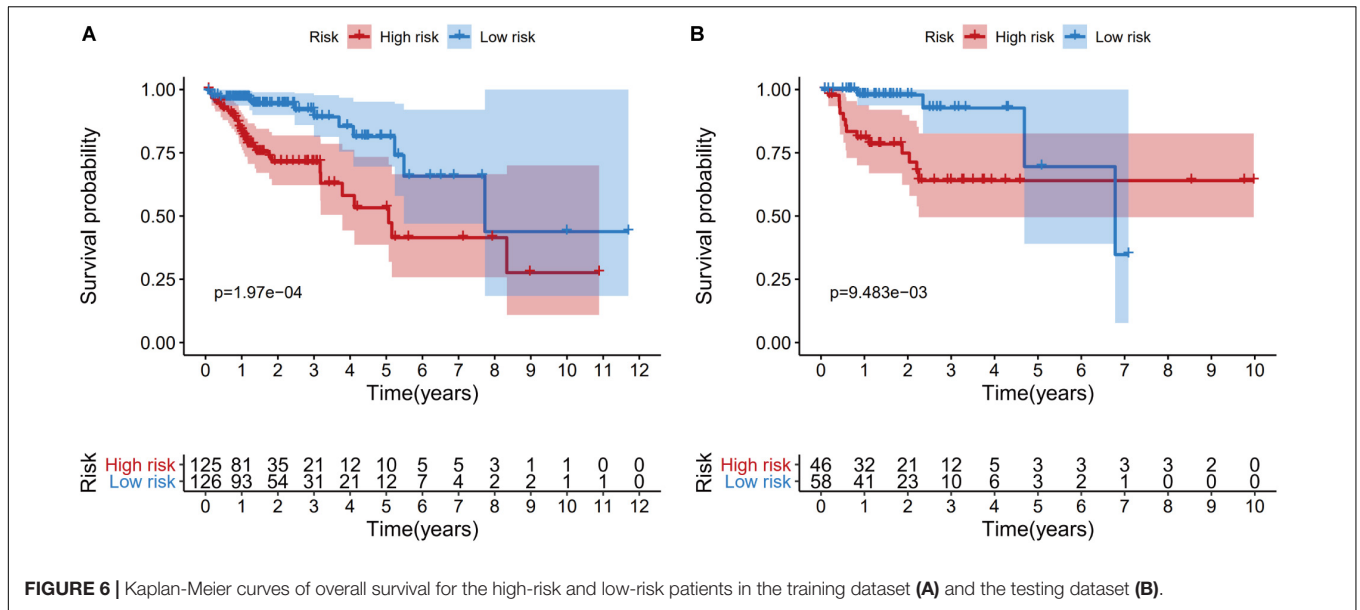


**TABLE 2 |** Univariate and multivariate Cox regression analysis of clinicopathologic factors and risk score for OS in training and testing sets.

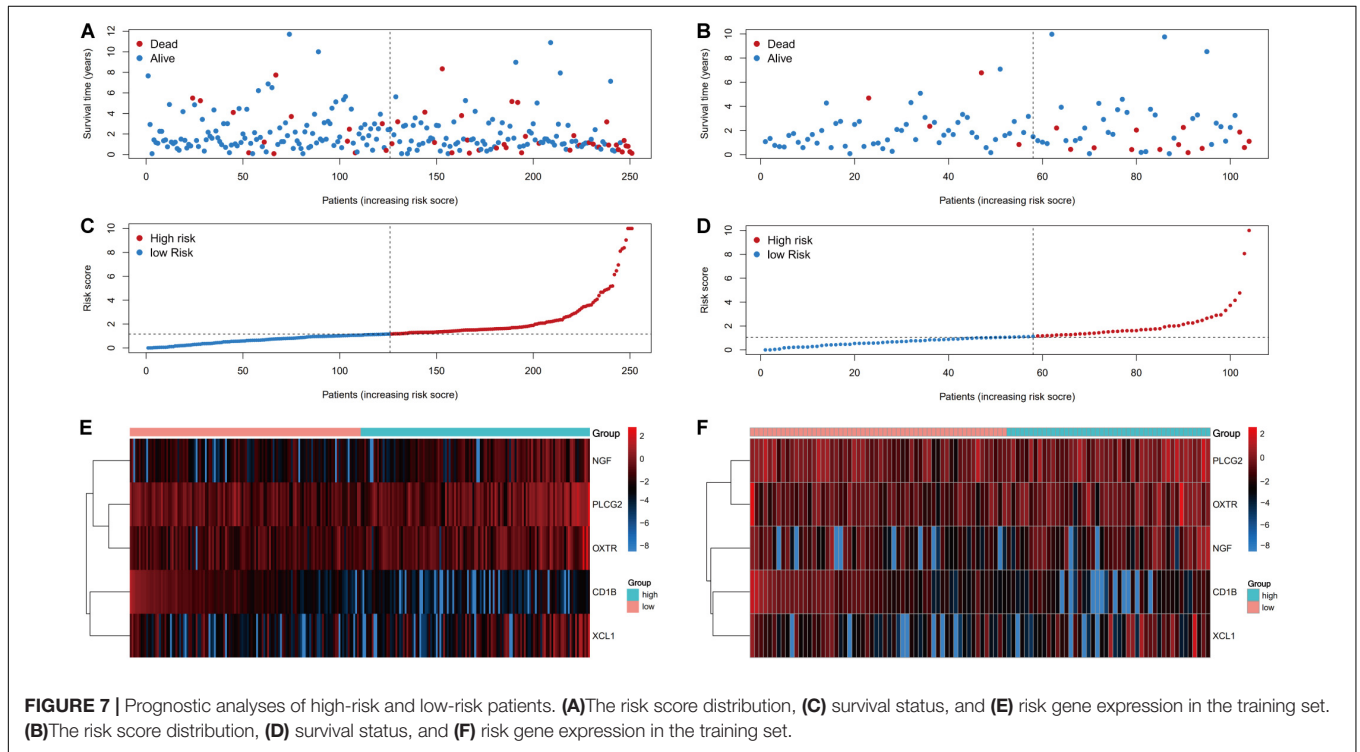
	Variable	Univariate		Multivariate	
		HR	P-value	HR	P-value
Training	Risk score (low risk vs. high risk)	1.020	3.82E-05	1.019	<0.001
	Age ( $\leq 65$ vs. $> 65$ )	1.009	0.503	1.030	0.036
	Gender (male vs. female)	0.706	0.264	0.828	0.561
	Stage (I and II vs. III and IV)	2.628	6.46E-08	2.768	5.49E-08
Testing	Risk score (low risk vs. high risk)	1.348	0.001	1.390	0.005
	Age ( $\leq 65$ vs. $> 65$ )	1.041	0.102	1.063	0.033
	Gender (male vs. female)	0.798	0.676	0.629	0.445
	Stage (I and II vs. III and IV)	2.620	0.003	2.380	0.009

HR, hazard ratio; OS, overall survival.





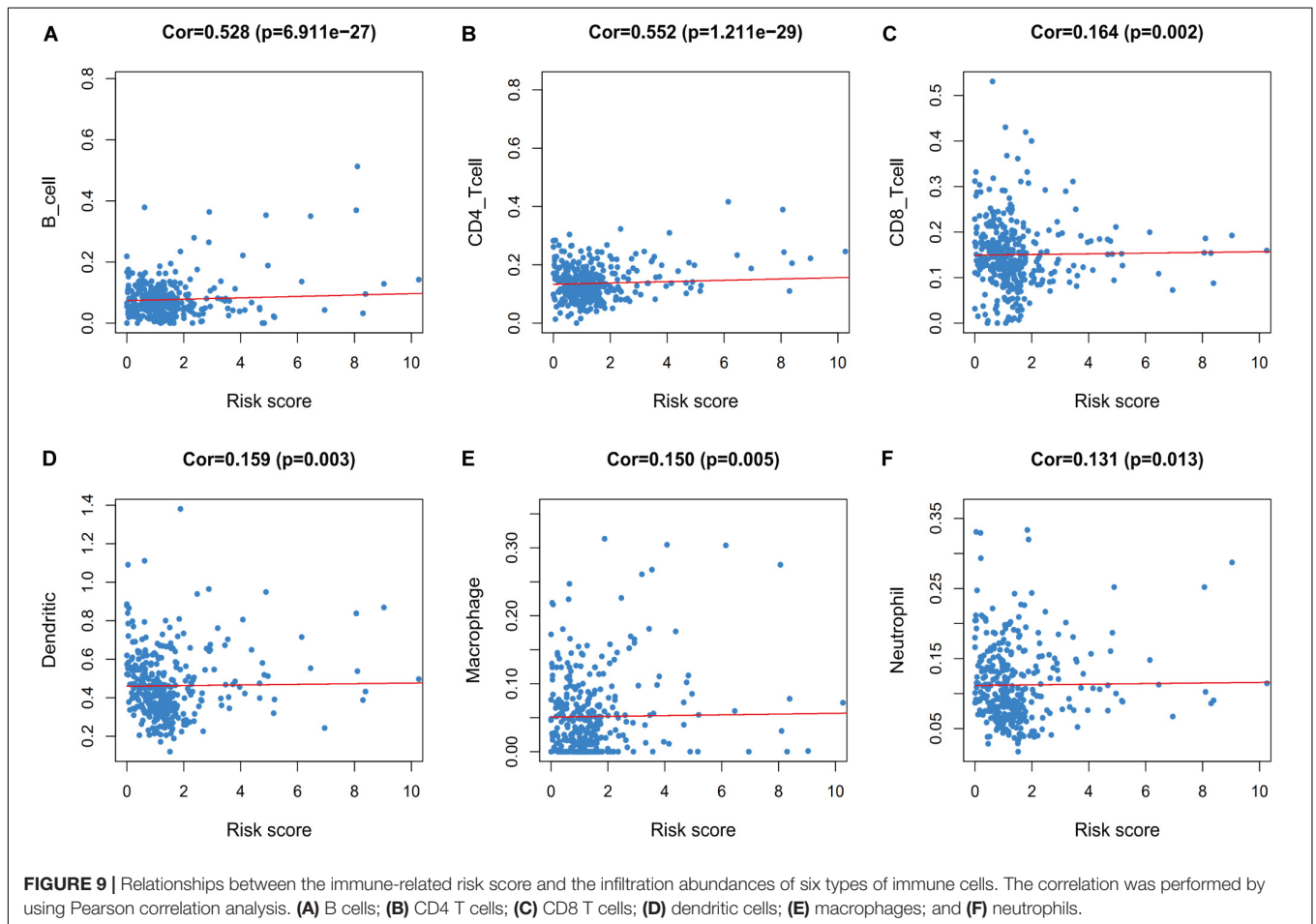
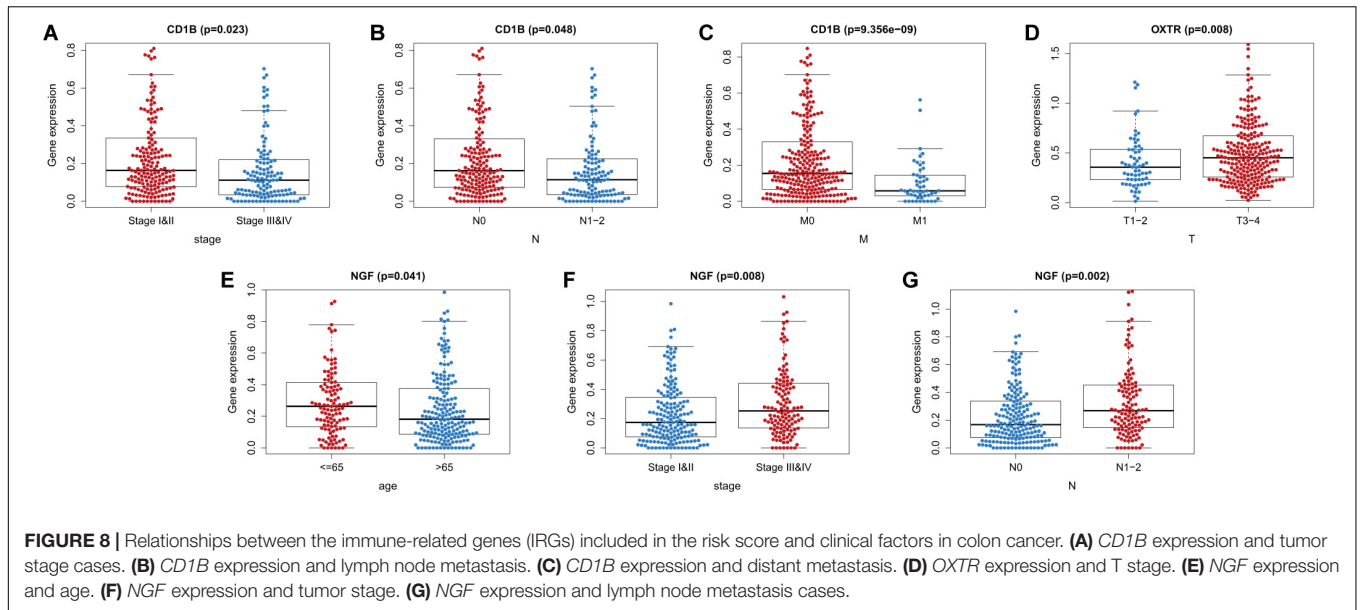
**FIGURE 6 |** Kaplan-Meier curves of overall survival for the high-risk and low-risk patients in the training dataset (A) and the testing dataset (B).



**FIGURE 7 |** Prognostic analyses of high-risk and low-risk patients. (A)The risk score distribution, (C) survival status, and (E) risk gene expression in the training set. (B)The risk score distribution, (D) survival status, and (F) risk gene expression in the testing set.

**TABLE 3 |** Relationships between risk factors (the risk score and the risk genes) and clinical variables in colon cancer.

Gene symbol	Age ( $\leq 65 / > 65$ )	Sex (male/female)	Tumor stage (I and II/III and IV)	T stage (T1-2/T3-4)	M stage (M0/M1)	N stage (N0/N1)
CD1B	-1.051 (0.294)	1.49 (0.137)	2.277 (0.023)	1.63 (0.106)	5.928 (9.36e-09)	1.987 (0.048)
XCL1	0.518 (0.605)	1.238 (0.217)	-1.122 (0.263)	-0.338 (0.736)	-1.261 (0.213)	-1.02 (0.309)
PLCG2	0.209 (0.835)	0.368 (0.713)	-1.423 (0.156)	-0.871 (0.385)	-0.871 (0.387)	-1.822 (0.070)
NGF	2.063 (0.041)	0.868 (0.386)	-2.697 (0.008)	-1.911 (0.059)	-1.84 (0.071)	-3.061 (0.002)
OXTR	0.128 (0.898)	-1.203 (0.230)	-1.529 (0.128)	-2.654 (0.008)	-0.918 (0.363)	-1.522 (0.130)
Risk Score	1.062 (0.291)	-0.975 (0.331)	-1.453 (0.149)	-1.624 (0.106)	-1.323 (0.192)	-1.451 (0.149)



cells (Liu and Cao, 2016; Spranger, 2016), which are closely related to tumorigenesis, progression, and metastasis (Berntsson et al., 2016; Liu and Cao, 2016; Väyrynen et al., 2016; Prizment

et al., 2017; Van den Eynde et al., 2018). Thus, our risk score might be useful for evaluating the patient's immune status and guiding treatment.

Previous immunogenomic models have shown favorable predictive value in this setting. Wu et al. developed a prognostic model using 19 IRG pairs for colon cancer (Wu et al., 2019), while Ge et al. developed prognostic models for stage I–II colon cancer (using 2 IRGs) and stage III–IV colon cancer (using 3 IRGs) (Ge et al., 2019). However, the IRGs identified in those studies were different from the IRGs in our risk score, which may be related to differences in the studies designs and patient populations. Moreover, we performed comprehensive analyses to identify the IRGs for the model, based on multiple algorithms, univariate Cox analyses, lasso regression analysis, and a Cox proportional hazards model.

Our study has several limitations. First, the risk score could not be validated based on clinical practice because of a lack of clinical specimens, and our risk score needs to be validated in other well-powered studies. Second, the predictive value of the risk scores was not discussed in the left-sided or right-sided colon cancer, respectively. We will collect colon cancer patients with tumor sites to verify our results in future study. Furthermore, the effects of the IRGs on colon cancer development remain unclear, although cancer progression has been linked to *CD1B* (Lee et al., 2019), *XCL1* (Chou et al., 2020; Tamura et al., 2020), *PLCG1* (Wells and Grandis, 2003), *NGF* (Retamales-Ortega et al., 2017; Garrido et al., 2019), and *OXTR* (Ma et al., 2019). Therefore, the underlying mechanisms require further investigation using *in vivo* and *in vitro* experiments.

## CONCLUSION

The present study developed an immunogenomic risk score that appears to have promising value for prognostication in cases of colon cancer. While it will not replace traditional cancer staging,

it may help refine prognostication among patients who had the same disease stage. However, further prospective studies are needed to evaluate the score's prognostic accuracy and clinical utility in the management of colon cancer.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the datasets (Transcriptome expression profiles and Clinical data for the corresponding patients), for this study can be found in the [The Cancer Genome Atlas (TCGA) data portal] (<https://cancergenome.nih.gov/>), the datasets (IMMUNE-RELATED GENE DATA) for this study can be found in the [Immunology Database and Analysis Portal (ImmPort) database] (<https://immport.niaid.nih.gov/>), the datasets (Immune infiltrate data) for this study can be found in the (Tumor Immune Estimation Resource) (<https://cistrome.shinyapps.io/timer/>).

## AUTHOR CONTRIBUTIONS

HZ and CQ: study conception, design, and performance, as well as writing of the report. XG: provided critical suggestions regarding the figures and tables. LZ: data collection. HG: led the research team. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

We would like to thank Editage ([www.editage.cn](http://www.editage.cn)) for English language editing.

## REFERENCES

- Aloe, L., Rocco, M. L., Balzamino, B. O., and Micera, A. (2016). Nerve growth factor: role in growth, differentiation and controlling cancer cell development. *J. Exp. Clin. Cancer Res.* 35:116. doi: 10.1186/s13046-016-0395-y
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Babaei, M., Balavarca, Y., Jansen, L., Lemmens, V., van Erning, F. N., van Eycken, L., et al. (2018). Administration of adjuvant chemotherapy for stage II-III colon cancer patients: an European population-based study. *Int. J. Cancer* 142, 1480–1489. doi: 10.1002/ijc.31168
- Bagchi, S., Li, S., and Wang, C. R. (2016). CD1b-autoreactive T cells recognize phospholipid antigens and contribute to antitumor immunity against a CD1b T cell lymphoma. *Oncoimmunology* 5:e1213932. doi: 10.1080/2162402x.2016.1213932
- Becerra, A. Z., Probst, C. P., Tejani, M. A., Aquina, C. T., González, M. G., Hensley, B. J., et al. (2015). Opportunity lost: adjuvant chemotherapy in patients with stage III colon cancer remains underused. *Surgery* 158, 692–699. doi: 10.1016/j.surg.2015.03.057
- Berntsson, J., Nodin, B., Eberhard, J., Micke, P., and Jirstrom, K. (2016). Prognostic impact of tumour-infiltrating B cells and plasma cells in colorectal cancer. *Int. J. Cancer* 139, 1129–1139. doi: 10.1002/ijc.30138
- Bhattacharya, S., Dunn, P., Thomas, C. G., Smith, B., Schaefer, H., Chen, J., et al. (2018). ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sc. Data* 5:180015. doi: 10.1038/sdata.2018.15
- Böttcher, J. P., Bonavita, E., Chakravarty, P., Bles, H., Cabeza-Cabrero, M., Sammicheli, S., et al. (2018). NK cells stimulate recruitment of cDC1 into the tumor microenvironment promoting cancer immune control. *Cell* 172, 1022–1037.e14. doi: 10.1016/j.cell.2018.01.004
- Bradshaw, R. A., Pundavela, J., Biarc, J., Chalkley, R. J., Burlingame, A. L., and Hondermarck, H. (2015). NGF and ProNGF: regulation of neuronal and neoplastic responses through receptor signaling. *Adv. Biol. Regul.* 58, 16–27. doi: 10.1016/j.jbior.2014.11.003
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Cabrero-de Las Heras, S., and Martínez-Balibrea, E. (2018). CXC family of chemokines as prognostic or predictive biomarkers and possible drug targets in colorectal cancer. *World J. Gastroenterol.* 24, 4738–4749. doi: 10.3748/wjg.v24.i42.4738
- Chou, W. C., Hsiung, C. N., Chen, W. T., Tseng, L. M., Wang, H. C., Chu, H. W., et al. (2020). A functional variant near *XCL1* gene improves breast cancer survival via promoting cancer immunity. *Int. J. Cancer* 146, 2182–2193. doi: 10.1002/ijc.32855
- Concha-Benavente, F., and Ferris, R. L. (2017). Oncogenic growth factor signaling mediating tumor escape from cellular immunity. *Curr. Opin. Immunol.* 45, 52–59. doi: 10.1016/j.coi.2017.01.004
- Dai, W., Meng, X., Mo, S., Xiang, W., Xu, Y., Zhang, L., et al. (2020). FOXE1 represses cell proliferation and Warburg effect by inhibiting HK2 in colorectal cancer. *Cell Commun. Signal.* 18:7. doi: 10.1186/s12964-019-0502-8

- de Vries, N. L., Swets, M., Vahrmeijer, A. L., Hokland, M., and Kuppen, P. J. (2016). The immunogenicity of colorectal cancer in relation to tumor development and treatment. *Int. J. Mol. Sci.* 17:1030. doi: 10.3390/ijms17071030
- Essien, B. E., Sundaresan, S., Ocadiz-Ruiz, R., Chavis, A., Tsao, A. C., Tessier, A. J., et al. (2016). Transcription factor ZBP-89 drives a feedforward loop of  $\beta$ -catenin expression in colorectal cancer. *Cancer Res.* 76, 6877–6887. doi: 10.1158/0008-5472.can-15-3150
- Fridman, W. H., Pagès, F., Sautès-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* 12, 298–306. doi: 10.1038/nrc3245
- Garrido, M. P., Hurtado, I., Valenzuela-Valderrama, M., Salvatierra, R., Hernández, A., Vega, M., et al. (2019). NGF-enhanced vasculogenic properties of epithelial ovarian cancer cells is reduced by inhibition of the COX-2/PGE signaling axis. *Cancers* 11:1970. doi: 10.3390/cancers11121970
- Ge, P., Wang, W., Li, L., Zhang, G., Gao, Z., Tang, Z., et al. (2019). Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of colorectal cancer. *Biomed. Pharmacother.* 118:109228. doi: 10.1016/j.biopha.2019.109228
- Holleczeck, B., Rossi, S., Domenic, A., Innos, K., Minicozzi, P., Francisci, S., et al. (2015). On-going improvement and persistent differences in the survival for patients with colon and rectum cancer across Europe 1999–2007 - Results from the EURO-CARE-5 study. *Eur. J. Cancer* 51, 2158–2168. doi: 10.1016/j.ejca.2015.07.024
- Hu, H., Krasinskas, A., and Willis, J. (2011). Perspectives on current tumor-node-metastasis (TNM) staging of cancers of the colon and rectum. *Semin. Oncol.* 38, 500–510. doi: 10.1053/j.seminoncol.2011.05.004
- Imperial, R., Ahmed, Z., Toor, O. M., Erdoğan, C., Khaliq, A., Case, P., et al. (2018). Comparative proteogenomic analysis of right-sided colon cancer, left-sided colon cancer and rectal cancer reveals distinct mutational profiles. *Mol. Cancer* 17:177. doi: 10.1186/s12943-018-0923-9
- Janssen, L. M. E., Ramsay, E. E., Logsdon, C. D., and Overwijk, W. W. (2017). The immune system in cancer metastasis: friend or foe? *J. Immunother. Cancer* 5:79. doi: 10.1186/s40425-017-0283-9
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkx1092
- Karoui, M., Roudot-Thoraval, F., Mesli, F., Mitry, E., Aparicio, T., Des Guetz, G., et al. (2011). Primary colectomy in patients with stage IV colon cancer and unresectable distant metastases improves overall survival: results of a multicentric study. *Dis. Colon Rectum* 54, 930–938. doi: 10.1097/DCR.0b013e31821cccd0
- Kenna, T., O'Brien, M., Hogan, A. E., Exley, M. A., Porcelli, S. A., Hegarty, J. E., et al. (2007). CD1 expression and CD1-restricted T cell activity in normal and tumour-bearing human liver. *Cancer Immunol. Immunother.* 56, 563–572. doi: 10.1007/s00262-006-0215-x
- Kim, K. H., Yang, S. S., Yoon, Y. S., Lim, S. B., Yu, C. S., and Kim, J. C. (2011). Validation of the seventh edition of the American Joint Committee on Cancer tumor-node-metastasis (AJCC TNM) staging in patients with stage II and stage III colorectal carcinoma: analysis of 2511 cases from a medical centre in Korea. *Colorectal Dis.* 13, e220–e226. doi: 10.1111/j.1463-1318.2011.02625.x
- Koss, H., Bunney, T. D., Behjati, S., and Katan, M. (2014). Dysfunction of phospholipase C $\gamma$  in immune disorders and cancer. *Trends Biochem. Sci.* 39, 603–611. doi: 10.1016/j.tibs.2014.09.004
- Laissue, P. (2019). The forkhead-box family of transcription factors: key molecular players in colorectal cancer pathogenesis. *Mol. Cancer* 18:5. doi: 10.1186/s12943-019-0938-x
- Lee, C. H., Chen, L. C., Yu, C. C., Lin, W. H., Lin, V. C., Huang, C. Y., et al. (2019). Prognostic value of CD1B in localised prostate cancer. *Int. J. Environ. Res. Public Health* 16:4723. doi: 10.3390/ijerph16234723
- Lei, Y., and Takahama, Y. (2012). XCL1 and XCR1 in the immune system. *Microbes Infect.* 14, 262–267. doi: 10.1016/j.micinf.2011.10.003
- Li, S., Dong, D., Geng, J., Zhu, X., Shi, C., Zhang, Y., et al. (2019). Prognostic factors and optimal response interval for stereotactic body radiotherapy in patients with lung oligometastases or oligoprogression from colorectal cancer. *Front. Oncol.* 9:1080. doi: 10.3389/fonc.2019.01080
- Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77, e108–e110. doi: 10.1158/0008-5472.CAN-17-0307
- Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 173, 400–416.e11. doi: 10.1016/j.cell.2018.02.052
- Liu, Y., and Cao, X. (2016). Immunosuppressive cells in tumor immune escape and metastasis. *J. Mol. Med.* 94, 509–522. doi: 10.1007/s00109-015-1376-x
- Lu, C. S., Chang, P. Y., Chen, Y. G., Chen, J. H., Wu, Y. Y., and Ho, C. L. (2015). Stage III colon cancer: the individualized strategy of adjuvant chemotherapy for aged under and over 70. *PLoS One* 10:e0138632. doi: 10.1371/journal.pone.0138632
- Ma, M., Li, L., Chen, H., and Feng, Y. (2019). Oxytocin inhibition of metastatic colorectal cancer by suppressing the expression of fibroblast activation protein- $\alpha$ . *Front. Neurosci.* 13:1317. doi: 10.3389/fnins.2019.01317
- Mayanagi, S., Kashiwabara, K., Honda, M., Oba, K., Aoyama, T., Kanda, M., et al. (2018). Risk factors for peritoneal recurrence in stage II to III colon cancer. *Dis. Colon Rectum* 61, 803–808. doi: 10.1097/dcr.0000000000001002
- Mei, S., Meyer, C. A., Zheng, R., Qin, Q., Wu, Q., Jiang, P., et al. (2017). Cistrome cancer: a web resource for integrative gene regulation modeling in cancer. *Cancer Res.* 77, e19–e22. doi: 10.1158/0008-5472.CAN-17-0327
- Ning, Y., and Lenz, H. J. (2012). Targeting IL-8 in colorectal cancer. *Expert Opin. Ther. Targets* 16, 491–497. doi: 10.1517/14728222.2012.677440
- Park, J. S., Choi, G. S., Hasegawa, S., Sakai, Y., Huh, J. W., Kim, H. R., et al. (2012). Validation of the seventh edition of the American Joint Committee on cancer tumor node-staging system in patients with colorectal carcinoma in comparison with sixth classification. *J. Surg. Oncol.* 106, 674–679. doi: 10.1002/jso.23117
- Paschke, S., Jafarov, S., Staib, L., Kreuser, E.-D., Maulbecker-Armstrong, C., Roitman, M., et al. (2018). Are colon and rectal cancer two different tumor entities? A proposal to abandon the term colorectal cancer. *Int. J. Mol. Sci.* 19:2577. doi: 10.3390/ijms19092577
- Prizment, A. E., Vierkant, R. A., Smyrk, T. C., Tillmans, L. S., Nelson, H. H., Lynch, C. F., et al. (2017). Cytotoxic T cells and granzyme B associated with improved colorectal cancer survival in a prospective cohort of older women. *Cancer Epidemiol. Biomarkers Prev.* 26, 622–631. doi: 10.1158/1055-9965.EPI-16-0641
- Retamales-Ortega, R., Oróstica, L., Vera, C., Cuevas, P., Hernández, A., Hurtado, I., et al. (2017). Role of nerve growth factor (NGF) and miRNAs in epithelial ovarian cancer. *Int. J. Mol. Sci.* 18:507. doi: 10.3390/ijms18030507
- Riihimäki, M., Hemminki, A., Sundquist, J., and Hemminki, K. (2016). Patterns of metastasis in colon and rectal cancer. *Sci. Rep.* 6:29765. doi: 10.1038/srep29765
- Shahine, A. (2018). The intricacies of self-lipid antigen presentation by CD1b. *Mol. Immunol.* 104, 27–36. doi: 10.1016/j.molimm.2018.09.022
- Spranger, S. (2016). Mechanisms of tumor escape in the context of the T-cell-inflamed and the non-T-cell-inflamed tumor microenvironment. *Int. Immunol.* 28, 383–391. doi: 10.1093/intimm/dxw014
- Su, G., Morris, J. H., Demchak, B., and Bader, G. D. (2014). Biological network exploration with cytoscape 3. *Curr. Protoc. Bioinformatics* 47, 8.13.11–8.13.24.
- Tamas, K., Walenkamp, A. M., de Vries, E. G., van Vugt, M. A., Beets-Tan, R. G., van Etten, B., et al. (2015). Rectal and colon cancer: not just a different anatomic site. *Cancer Treat. Rev.* 41, 671–679. doi: 10.1016/j.ctrv.2015.06.007
- Tamura, R., Yoshihara, K., Nakaoka, H., Yachida, N., Yamaguchi, M., Suda, K., et al. (2020). XCL1 expression correlates with CD8-positive T cells infiltration and PD-L1 expression in squamous cell carcinoma arising from mature cystic teratoma of the ovary. *Oncogene* 39, 3541–3554. doi: 10.1038/s41388-020-1237-0
- Tang, W., Zhou, Y., Sun, D., Dong, L., Xia, J., and Yang, B. (2019). Oncogenic role of phospholipase C- $\gamma$ 1 in progression of hepatocellular carcinoma. *Hepatol. Res.* 49, 559–569. doi: 10.1111/hepr.13309
- Tang, Y., Shi, Y., Gao, Y., Xu, X., Han, T., Li, J., et al. (2019). Oxytocin system alleviates intestinal inflammation by regulating macrophages polarization in experimental colitis. *Clin. Sci.* 133, 1977–1992. doi: 10.1042/cs20190756
- Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102. doi: 10.1093/nar/gkx247

- Tomas, N. M., Masur, K., Piecha, J. C., Niggemann, B., and Zänker, K. S. (2012). Akt and phospholipase C $\gamma$  are involved in the regulation of growth and migration of MDA-MB-468 breast cancer and SW480 colon cancer cells when cultured with diabetogenic levels of glucose and insulin. *BMC Res. Notes* 5:214. doi: 10.1186/1756-0500-5-214
- Tomellini, E., Touil, Y., Lagadec, C., Julien, S., Ostyn, P., Ziental-Gelus, N., et al. (2015). Nerve growth factor and proNGF simultaneously promote symmetric self-renewal, quiescence, and epithelial to mesenchymal transition to enlarge the breast cancer stem cell compartment. *Stem Cells* 33, 342–353. doi: 10.1002/stem.1849
- Van den Eynde, M., Mlecnik, B., Bindea, G., Fredriksen, T., Church, S. E., Lafontaine, L., et al. (2018). The link between the multiverse of immune microenvironments in metastases and the survival of colorectal cancer patients. *Cancer Cell* 34, 1012–1026.e3. doi: 10.1016/j.ccell.2018.11.003
- Väyrynen, J. P., Kantola, T., Väyrynen, S. A., Klintrup, K., Bloigu, R., Karhu, T., et al. (2016). The relationships between serum cytokine levels and tumor infiltrating immune cells and their clinical significance in colorectal cancer. *Int. J. Cancer* 139, 112–121. doi: 10.1002/ijc.30040
- Wang, S., Li, L., Zhou, Y., He, Y., Wei, Y., and Tao, A. (2019). Heterotypic cell-in-cell structures in colon cancer can be regulated by IL-6 and lead to tumor immune escape. *Exp. Cell Res.* 382:111447. doi: 10.1016/j.yexcr.2019.05.028
- Weinstein, A. M., Giraldo, N. A., Petitprez, F., Julie, C., Lacroix, L., Peschard, F., et al. (2019). Association of IL-36 $\gamma$  with tertiary lymphoid structures and inflammatory immune infiltrates in human colorectal cancer. *Cancer Immunol. Immunother.* 68, 109–120. doi: 10.1007/s00262-018-2259-0
- Wells, A., and Grandis, J. R. (2003). Phospholipase C-gamma1 in tumor progression. *Clin. Exp. Metastasis* 20, 285–290. doi: 10.1023/a:1024088922957
- Weß, L., and Schnieders, F. (2017). Immuno-oncology—the translational runway for gene therapy: gene therapeutics to address multiple immune targets. *Hum. Gene Ther.* 28, 1130–1137. doi: 10.1089/hum.2017.145
- Wu, J., Zhao, Y., Zhang, J., Wu, Q., and Wang, W. (2019). Development and validation of an immune-related gene pairs signature in colorectal cancer. *Oncoimmunology* 8:1596715. doi: 10.1080/2162402x.2019.1596715
- Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39, W316–W322. doi: 10.1093/nar/gkr483
- Zhang, X., Hu, F., Li, G., Li, G., Yang, X., Liu, L., et al. (2018). Human colorectal cancer-derived mesenchymal stem cells promote colorectal cancer progression through IL-6/JAK2/STAT3 signaling. *Cell Death Dis.* 9:25. doi: 10.1038/s41419-017-0176-3

**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.

Copyright © 2020 Zhang, Qin, Gan, Guo and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.