



Identifying Potential Biomarkers of Prognostic Value in Colorectal Cancer via Tumor Microenvironment Data Mining

Lei Li^{1,2}, Xiao Du^{2,3*} and Guangyi Fan^{2,3*}

¹College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, ²BGI-Qingdao, BGI-Shenzhen, Qingdao, China, ³BGI-Shenzhen, Shenzhen, China

OPEN ACCESS

Edited by:

Peng Wang,
Harbin Medical University, China

Reviewed by:

Feng Gao,
The Sixth Affiliated Hospital of Sun
Yat-sen University, China
Xin Wang,
The Chinese University of Hong Kong,
China

*Correspondence:

Guangyi Fan
fanguangyi@genomics.cn
Xiao Du
duxiao@genomics.cn

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 30 September 2021

Accepted: 16 December 2021

Published: 03 February 2022

Citation:

Li L, Du X and Fan G (2022) Identifying
Potential Biomarkers of Prognostic
Value in Colorectal Cancer via Tumor
Microenvironment Data Mining.
Front. Genet. 12:787208.
doi: 10.3389/fgene.2021.787208

Colorectal cancer (CRC) is a common cancer that has increased rapidly worldwide in the past decades with a relatively high mortality rate. An increasing body of evidence has highlighted the importance of infiltrating immune and stromal cells in CRC. In this study, based on gene expression data of CRC patients in TCGA database we evaluated immune and stromal scores in tumor microenvironment using ESTIMATE method. Results showed there was potential correlation between these scores and the prognosis, and that patients with higher immune score and lower stromal score had longer survival time. We found that immune score was correlated with clinical characteristics including tumor location, tumor stage, and survival time. Specifically, the right-sided colon cancer had markedly elevated immune score, compared to left-sided colon cancer and rectal cancer. These results might be useful for understanding tumor microenvironment in colorectal cancer. Through the differential analysis we got a list of genes significantly associated with immune and stromal scores. Gene Set Enrichment and protein-protein interaction network analysis were used to further illustrate these differentially expressed genes. Finally, 15 hub genes were identified, and three (CXCL9, CXCL10 and SELL) of them were validated with favorable outcomes in CRC patients. Our result suggested that these tumor microenvironment related genes might be potential biomarkers for the prognosis of CRC.

Keywords: immune, stromal, hub genes, colorectal cancer, survival analysis, tumor location

INTRODUCTION

Colorectal cancer (CRC) is one of the most commonly occurring cancers, whose incidence occupies 10% of all cancer diagnoses (Sung et al., 2021; Wong et al., 2021). As the second most common cause of cancer death, Colorectal cancer has been increasing rapidly in the past decades with over 1.9 million new cases reported in 2020 (Arnold et al., 2017; Sawicki et al., 2021). Colorectal cancer may develop on either the proximal colon (right side), the distal colon (left side) or the rectum. Right-sided colon cancer (RCC) differs from the left-sided colon cancer (LCC) and rectal cancer (RC) in pathogenesis and prognosis, exhibiting distinct molecular characteristics and histology (Baran et al., 2018; Imperial et al., 2018; Siegel et al., 2020). Presently, CRC screening is not common and the diagnosis is usually made after the onset of symptoms. Because the tumor status and TNM stage at diagnosis have a fundamental role in CRC prognosis, early symptom investigation and diagnosis are of high importance (Bosch et al., 2011; Kawakami et al., 2015). However, although CRC prevalence is

high, the awareness of colorectal cancer and its symptoms is relatively low. Due to wide variation in colorectal cancer and complexity in treatment outcome prediction, investigation for new strategies and new biomarkers is necessary in CRC for improving prognosis.

It has been documented that tumor microenvironment (TME) has a great impact on tumor cells and clinical outcomes (Turley et al., 2015; Lim et al., 2018). Apart from tumor cells, TME also comprises a variety of nontumor components including endothelial cells, immune cells, inflammatory mediators, and extracellular matrix (ECM) molecules (Lorusso and Rüegg, 2008; Bolouri, 2015). The cells and molecules in the TME are in a dynamic process, jointly promoting tumor immune escape, tumor growth and metastasis (Quail and Joyce, 2013). Accumulating evidence suggests that the stromal and immune cells, which constitute two main nontumor components in the TME, are valuable in investigating tumor diagnosis and clinical outcome (Kalluri and Zeisberg, 2006; Hanahan and Weinberg, 2011; Fridman et al., 2012). Recent evidence has indicated that tumor microenvironment plays a significant role in colorectal carcinogenesis, metastasis and the choosing of therapy strategies (Peddareddigari et al., 2010; Pedrosa et al., 2019). T cells, a major part of the immune system, were described to be of major importance for tumor growth, invasion, early metastasis and prognosis in colorectal cancer (Pagès et al., 2005; Mlecnik et al., 2011). Calon et al. suggested that high expression of mesenchymal genes associated with poor outcomes in CRC patients is primarily caused by stromal cells instead of epithelial cancer cells (Calon et al., 2015). To promote the understanding of cancer prognosis, efforts have been made in studying tumor microenvironment components and developing novel immunotherapeutic strategies in recent years. Algorithms such as ESTIMATE (Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data) (Yoshihara et al., 2013) have been developed to predict tumor purity and levels of infiltrating stromal and immune cells in tumor, such as gastric cancer, hepatocellular carcinoma and colorectal cancer (Mao et al., 2018; Deng et al., 2019; Wang et al., 2019).

To promote the understanding of CRC microenvironment and prognosis, in this study we took use of ESTIMATE algorithm and public database to evaluate the tumor-infiltrating immune and stromal cells of TME. By performing survival analysis and correlation analysis, we explored the relationship between immune/stromal score and clinical factors in CRC. Moreover, we aim to extract a list of tumor microenvironment associated genes of prognostic value, through the differential analysis, network construction and survival analysis. We hope to provide insights to investigate stromal and immune cells of CRC and offer evidence to potential prognostic markers.

MATERIALS AND METHODS

Data Collection and Preprocessing

In this study, gene expression profiles of colorectal cancer were downloaded and collected from The Cancer Genome Atlas

(TCGA) data portal (<https://portal.gdc.cancer.gov/>) using TCGAbiolinks (Colaprico et al., 2016) R package. Relevant clinical information including age, gender, survival time, pathologic stage, and tissue or organ of origin were also obtained from TCGA database. Patients with primary tumor expression and survival information were retained in this study. Before further analysis, TCGA gene expression profiles were normalized using R package DESeq2 (Love et al., 2014).

GSE41258 expression and clinical data were also downloaded from the Gene Expression Omnibus (GEO) database as the validation set. The GSE41258 dataset was processed *via* the Affymetrix MAS5 background correction algorithm using affy package (Gautier et al., 2004) in R and log₂ transformation. Probe sets were transformed into gene sets by retaining only the probes with the highest expression levels if one gene corresponds to multiple probes. When multiple genes per probe, this probe would be discarded.

Estimation of Immune and Stromal Scores

The normalized expression data was analyzed by the ESTIMATE algorithm for calculating the Immune and Stromal Score. We used ESTIMATE to calculate the fraction of immune and stromal cells in tumor using the gene expression data. In our study v.1.0.13 estimate R package (Yoshihara et al., 2013) was used to predict the level of infiltrating immune cells (immune score) and the level of infiltrating stromal cells (stromal score).

Survival Analysis Based on Immune and Stromal Scores

Survival analysis was performed by R package survival (Therneau, 2019) and survminer (Kassambara et al., 2019) to assess the association of immune and stromal score with prognosis. The best cut-off value of immune/stromal score was inferred using R program `surv_cutpoint`. Subsequently, patients were divided into two groups (high vs. low) based on the cut-off value. The Kaplan-Meier (KM) method was used to estimate the likelihood of survival based on the observed overall survival time. Overall Survival differences between high and low score groups were compared by log-rank test.

Differential Gene Expression Analysis

We analyzed differentially expressed genes (DEGs) between high score and low score groups using R package DESeq2 (Love et al., 2014), which based on the negative binomial distribution algorithm. And $|\log_2 \text{fold change (FC)}| > 2$ and p value < 0.01 were selected as the criteria to select the significantly different genes. R package pheatmap (Kolde, 2019) was used to visualize the DEGs.

Function Analysis

To explore the potential function of DEGs, function analysis was carried out by using the Gene Set Enrichment Analysis (GSEA) web server (Mootha et al., 2003; Subramanian et al., 2005). Enrichment analyses of hallmark gene sets, ontology gene

TABLE 1 | Summary and Cox Regression Analysis of overall survival for TCGA CRC study dataset.

Characteristics	Count	Univariate Cox		Multivariate Cox	
		Hazard ratio (95% CI)	P-value	Hazard ratio (95% CI)	P-value
Age	613	1.03 (1.015-1.049)	<0.001	1.04 (1.021-1.067)	<0.001
Gender					
Female	286	1	-	1	-
Male	327	1.02 (0.710-1.454)	0.93	0.88 (0.554-1.400)	0.59
Location					
Right	189	1	-	1	-
Left	132	0.70 (0.435-1.134)	0.15	0.58 (0.349-0.974)	0.04
Rectum	85	0.72 (0.386-1.326)	0.29	0.55 (0.279-1.094)	0.09
Stage					
Stage I	103	1	-	1	-
Stage II	227	1.72 (0.712-4.150)	0.23	1.03 (0.384-2.775)	0.95
Stage III	177	3.19 (1.345-7.580)	0.01	2.31 (0.870-6.151)	0.09
Stage IV	86	8.62 (3.647-20.370)	<0.001	7.36 (2.803-19.327)	<0.001
Stromal score					
High	230	1	-	1	-
Low	383	0.69 (0.483-0.998)	0.05	0.66 (0.332-1.312)	0.24
Immune score					
High	425	1	-	1	-
Low	188	1.44 (1.001-2.071)	0.05	2.07 (1.060-4.043)	0.03

terms (cellular component, molecular function, and biological process), and KEGG gene sets were selected to extract biological insight in different risk groups. The top 20 biological functional terms with False discovery rate (FDR) *q*-value below 0.01 were selected.

PPI Network Construction and Hub Gene Selection

To further investigate the relationship between different genes, the protein-protein interaction (PPI) network analysis was performed *via* the version 11.5 STRING (Search Tool for the Retrieval of Interacting Genes, <https://string-db.org/>) (Szklarczyk et al., 2015), an online tool and database of protein-protein interaction. A minimum required interaction score > 0.7 were selected and reconstructed in the Cytoscape (Shannon et al., 2003) software. In a gene candidate module, one gene with high correlation with other genes is called a hub gene. In this study, We used CytoHubba plugin (Chin et al., 2014) in Cytoscape v3.7.1 to find hub genes in PPI network. The top 15 genes with the highest prediction scores calculated by the Maximal Clique Centrality (MCC) algorithm were defined as the hub genes.

Statistical Analysis

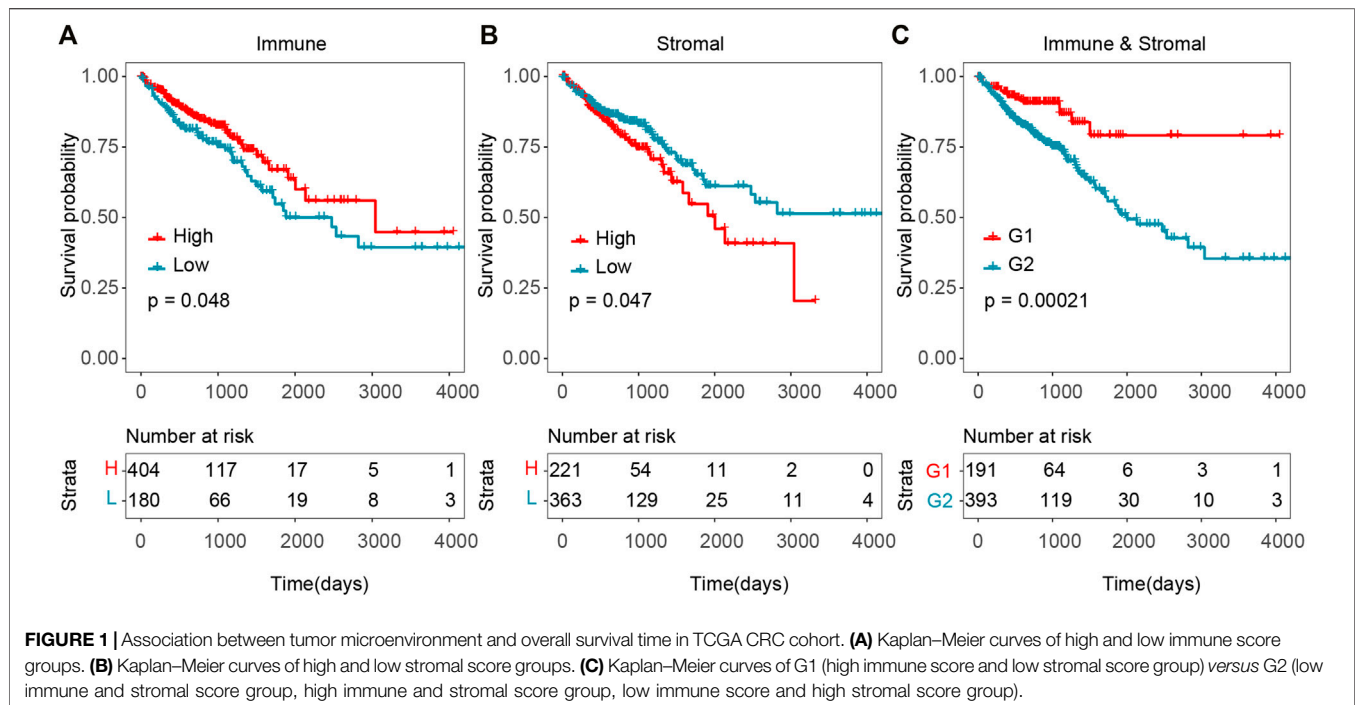
All statistical analyses were performed in R statistical environment version \geq 3.5.0. Cox proportional hazard regression survival analysis was applied to overall survival time with different clinical features including age, gender, tumor location, tumor stage, immune score and stromal score. Correlations between the clinical factors and immune/stromal score were also calculated in this study. Kruskal-Wallis Test for three or more groups and Wilcoxon Test for two groups were used to estimate the P value.

RESULTS

Tumor Immune and Stromal Scores Significantly Associated With Prognosis in CRC

HTSeq-Counts based gene expression profiles and clinical information of 613 CRC patients were downloaded from TCGA database. In this cohort, patients were diagnosed with colorectal cancer between 1998 and 2013 and their sequencing and clinical information were collected into the TCGA database. Among them, 286 (46.7%) patients were female and 327 (53.3%) patients were male. The ages ranged from 31 to 90. Clinical diagnosis included 189 (30.8%) cases with right-sided colon cancer, 132 (21.5%) cases with left-side colon cancer, and 85 (13.9%) cases with rectal cancer. The Pathologic stage I, stage II, stage III and stage IV accounted for 16.8% ($n = 103$), 37.0% ($n = 227$), 28.9% ($n = 177$) and 14.0% ($n = 86$) of the total number (**Table 1**). In addition, based on ESTIMATE algorithm immune and stromal scores were obtained. Stromal scores for the analyzed CRC cohort ranged from -2,531.36 to 1,481.74, and immune scores were distributed between -1,724.23 and 1,856.93, respectively. The average immune score was -600.92 and the median was -658.63. The average stromal score was -966.83 and the median was -1,026.69.

To explore the potential correlation of overall survival time with stromal and immune scores, 613 CRC cases were divided into high- and low-score groups according to the cut-off of stromal/immune scores. As shown in **Figure 1**, survival analysis indicated that both the immune and stromal scores were significantly correlated with overall survival time, and that patients with high immune score or low stromal score significantly correlated with better overall survival time (**Figures 1A,B**, p -value = 0.048 for immune



score and p -value = 0.047 for stromal score, log-rank test). Patients with high immune score had a median overall survival time of 101.4 months, while patients with low immune score had a median survival of 62.7 months. Patients with lower stromal score also had a longer median overall survival compared to those with high stromal score. Especially, patients with combined high immune score and low stromal score have a significantly better overall survival time than others (**Figure 1C**, p -value = 0.00021, log-rank test).

In order to validate these results which were obtained from the TCGA database, we downloaded and analyzed another independent cohort in GEO database. We retrieved 182 CRC patients' gene expression data and clinical information from GSE41258 cohort as validation set. Although the difference was not statistically significant, Patients with high immune score displayed a longer median survival (**Supplementary Figure S1A**, high- vs. low-score = 91 vs. 86 months). And patients with lower stromal score showed a longer median survival (**Supplementary Figure S1B**, high- vs. low-score = 72 vs. 113 months). Consistently, patients with high immune score and low stromal score in the validation cohort had a longer survival time (**Supplementary Figure S1C**, p -value = 0.021, log-rank test). These results indicated that higher level of immune score and lower level of stromal score in CRC might mean the favorable survival outcome, which might provide potential prognosis stratification factors for clinical predictions.

Immune Scores Correlated With Tumor Location and Tumor Stage in CRC

To determine the clinical significance of immune and stromal scores, we investigated the association between immune/stromal score and clinical features, and the results suggested

that the right-sided colon cancer have a significantly higher immune score. Immune score significantly correlated with tumor stage and tumor location (**Figures 2A,B**, p -value < 0.01). The median immune score of the RCC patients ranked the highest of all three tumor location subtypes, and the LCC subtype cases had the lowest immune scores (RCC vs RC = -571.65 vs -838.1 , p -value = 0.019, RCC vs LCC = -571.65 vs -860.61 , p -value = 0.00012, LCC vs RC = -860.61 vs -838.1 , p -value = 0.23, Wilcoxon Test) (**Figure 2B**). Similarly, the rank order of immune scores across tumor stage from highest to lowest was Stage I > Stage II > Stage III > Stage IV (**Figure 2A**). What's more, we found immune score was also significantly associated with tumor location and the RCC also had the highest immune score in GSE41258 dataset (**Supplementary Figure S2B**, p -value = 0.032), which indicated that immune score might be predictive in the classification of CRC tumor location. However, we found no significant differences between stromal scores with CRC tumor stage or location (**Figures 2C,D**, p -value > 0.05). Consequently, immunotherapy is likely to be more effective for right-sided colon cancer with more immune infiltration and activation in CRC.

Differential Expressed Genes Revealed by Immune and Stromal Scores in CRC

To reveal the correlation of gene expression profiles with immune and stromal scores, we performed differential expression genes analysis using DESeq2, and 318 DEGs were screened out in total. By comparing immune scores between high- and low-score groups, 188 genes were identified to be differentially expressed genes. A total of 150

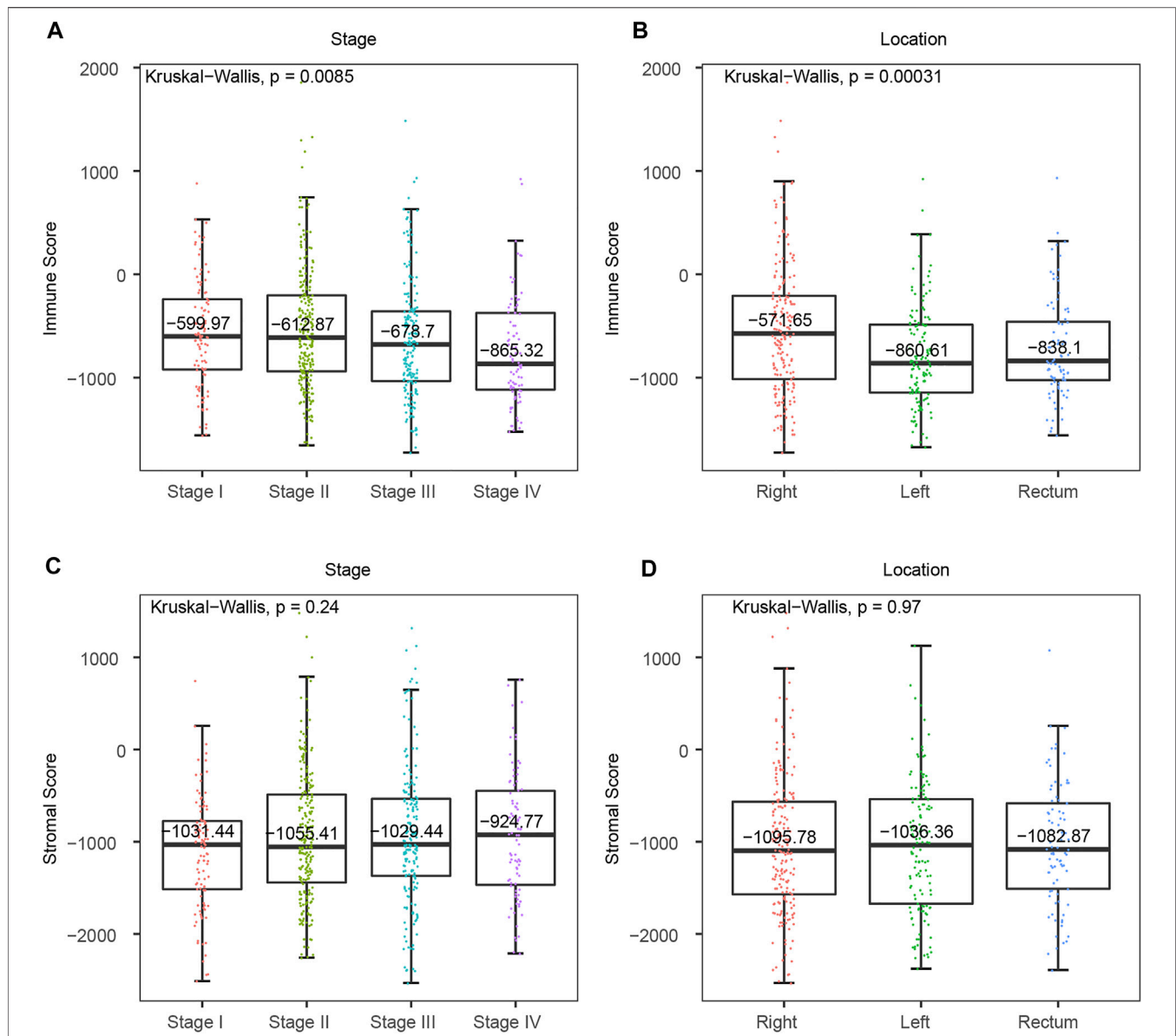
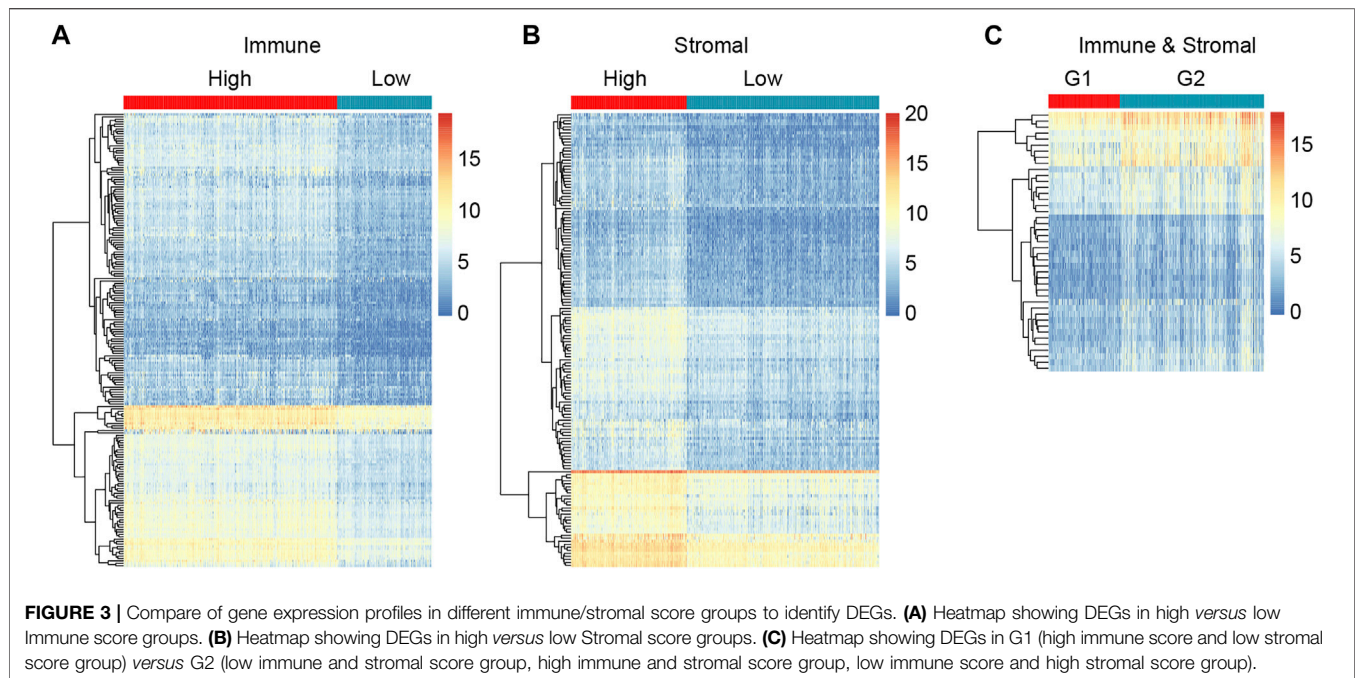


FIGURE 2 | Association between tumor microenvironment scores and clinical features in TCGA CRC cohort. **(A)** Distribution of immune scores in consecutive CRC tumor stages. **(B)** Distribution of immune scores from distinct CRC primary tumor locations. **(C)** Distribution of stromal scores in consecutive CRC tumor stages. **(D)** Distribution of stromal scores from different CRC primary tumor locations.

DEGs were found for high stromal score as compared to low stromal score. What's more, we got 43 DEGs when high immune and low stromal score patients were compared to the rest. The expression level of the DEGs in each group was displayed in heatmap (**Figure 3**). The subsequent analysis in our study were based on these DEGs.

To better understand the potential biological functions and mechanisms of DEGs in different immune and stromal score groups, Gene Set Enrichment Analysis was used to annotate the biological roles of these DEGs. GO: BP, GO: CC, GO: MF, KEGG pathways and hallmark gene sets were included in the functional enrichment analysis. The top 20 functional terms of DEGs in each

group were shown in **Figure 4**. For the immune score group, the DEGs were mostly enriched in the regulation of immune system process and defense response. For the stromal score group and combined group, the top biological terms were external encapsulating structure and muscle system process. Moreover, circulatory system development, collagen containing extracellular matrix, external encapsulating structure, intrinsic component of plasma membrane, and skeletal system development were enriched in at least 2 groups. According to the result of GSEA, it could be concluded that these 318 DEGs were mostly involved in the immune regulation biological process that modulates the



frequency, rate, extent of an immune system process, and cytokine-cytokine receptor interaction pathway.

Hub Gene Selection Based on PPI Network

In order to evaluate the protein interactive relationships among DEGs, PPI network was constructed based on STRING database and nodes that reported high scores in the network were screened as hub genes. A total of 318 differential expressed genes comprised 318 nodes and 372 edges based on STRING database, and result was visualized in **Figure 5** after hidden disconnected nodes in the network. Following STRING analysis, the network was reconstructed in the Cytoscape. According to the calculation of CytoHubba plugin module, we identified a list of important genes, from which the top fifteen genes identified by the MCC algorithm were used for further analysis. Finally, 15 genes were selected as hub genes (CD86, ITGAM, PTPRC, FCGR3A, FCGR3B, MRC1, CD163, CCR2, SELL, CD69, CXCL10, CXCL8, CXCL9, CCL19 and CCL4), which were marked with red color in the PPI network (**Figure 5**). And we found that these genes were significantly enriched in the external side of plasma membrane, cell surface and chemokine receptor binding according to Gene Set Enrichment Analysis (**Supplementary Table S1**).

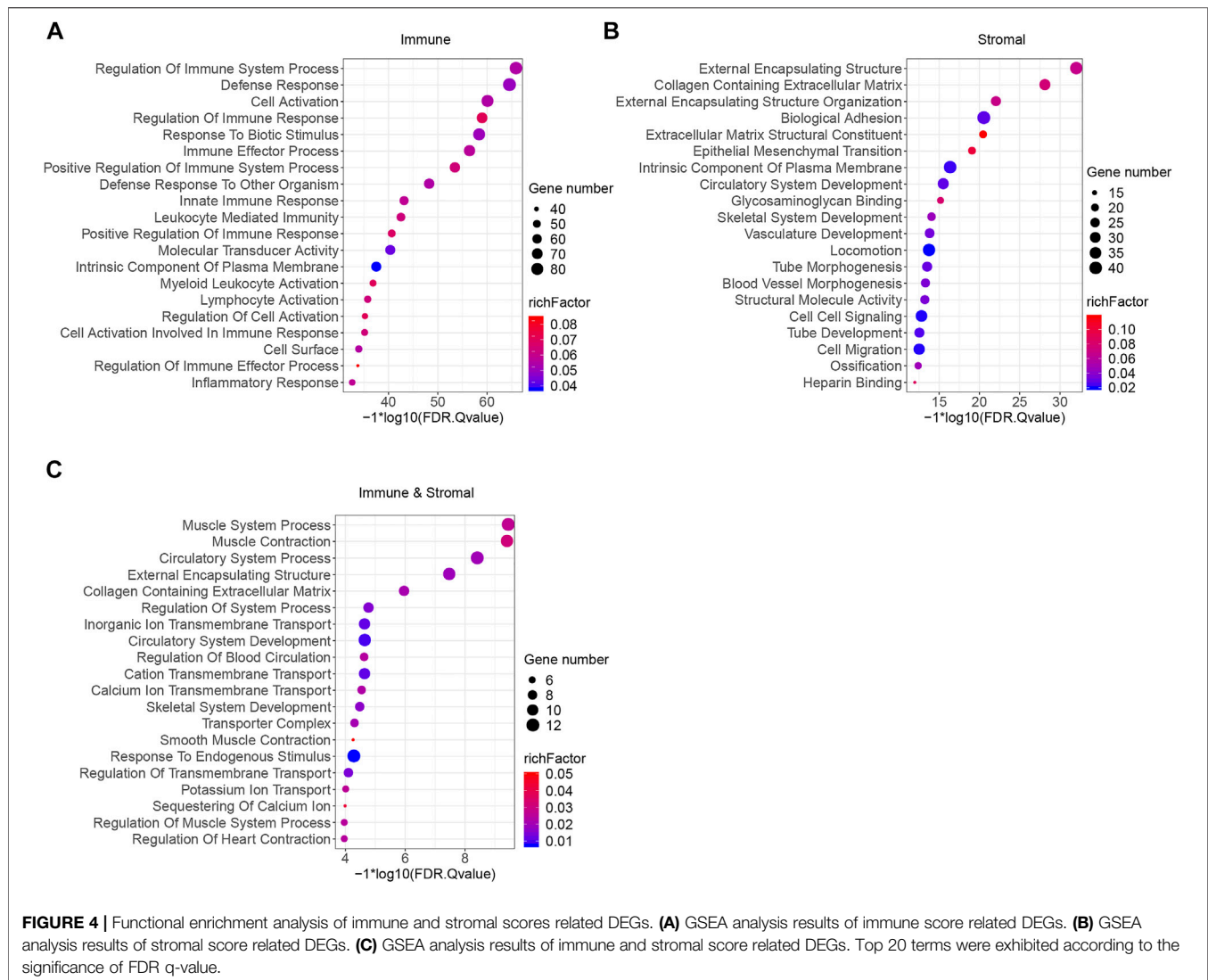
Identification and Validation of Survival Related Hub Genes

We performed survival analysis between the 15 hub genes and the overall survival time to identify potential prognostic or predictive markers for CRC. Colorectal cancer samples were split into high- and low- expression groups according to the optimal survival cut-off. We found that 11 hub genes were correlated with survival in TCGA dataset (**Figures 6A–C** and

Supplementary Figure S3, p -value < 0.05, log-rank test). As shown in **Figure 6**, CXCL9 and CXCL10 were significantly correlated with the overall survival time in TCGA dataset (**Figures 6A,B**, p -value < 0.05, log-rank test), and a higher expression of them might correspond to better survival. Importantly, similar result was observed in the validation set GSE41258 (**Figures 6D,E**, p -value < 0.05, log-rank test). Moreover, high expression of SELL also showed longer overall survival in TCGA dataset (**Figure 6C**, p -value < 0.05, log-rank test), even though this pattern was not statistically significant in GSE41258 cohort (**Figure 6F**, p -value = 0.053, log-rank test). Higher expression of PTPRC and CCL4 had a better survival time in TCGA dataset (**Supplementary Figure S3**, p -value < 0.05, log-rank test) and showed a longer median survival time in GSE41258, but this correlation was not statistically significant (**Supplementary Figure S4**, $0.05 < p$ -value < 0.1, log-rank test).

DISCUSSION

Colorectal cancer is one of the most common pathological types of cancer. Previous research have demonstrated that tumor microenvironment play an important role in the occurrence and development of CRC (Peddareddigari et al., 2010; Kamal et al., 2020). Data from previous studies indicated that the infiltration of immune cells into the tumor bed may be a valuable prognostic factor in the treatment of colorectal tumor (Pagès et al., 2005; Galon et al., 2006; Galon et al., 2007; Ganesh et al., 2019). Research showed that the high density of infiltrating memory CD45RO+ T cells, one type of immune cell, was associated with the absence of signs of early tumor lymphovascular and perineural invasion, a less advanced tumor stage, and a good clinical outcome (Pagès et al., 2005).



Cancer-associated fibroblasts (CAFs) are one of the most abundant and key components of the tumor mesenchyme among all the stromal cells (Liu et al., 2019). According to the study of Isella et al., the presence of high levels of CAFs was associated with poor prognosis in untreated CRC (Isella et al., 2015). Understanding the relationship between tumor microenvironment and patients' clinical features is vital in figuring out cancer recurrence and metastasis mechanisms. However, this mechanism is not well-understood yet.

In this study, we used the ESTIMATE algorithm to evaluate the infiltration degree of immune and stromal cells in colorectal cancer. A total of 613 CRC patients were divided into two groups based on the immune and stromal scores calculated by the R function ESTIMATE. As a result, we found high immune score was related with prolonged survival time. This observation was in general agreement with the study of Mlecnik et al. (Mlecnik et al., 2016). Besides, we found lower stromal score indicated a longer overall survival time, which further confirmed previous work by Calon et al. (Calon et al., 2015). More importantly, when patients

had high immune and low stromal scores, they displayed a significantly better clinical outcome. The similar trends were also observed in another independent dataset GSE41258. These results from our study may help elucidate the underlying mechanisms in colorectal cancer microenvironment and prognosis.

Apart from that, we found clinical factors including primary tumor location and tumor stage were significantly correlated with immune score in CRC. It is worth noting that right-sided colon cancer had significantly higher immune score, as compared to left-sided colon cancer or rectum cancer. These findings might explain why right-sided colon cancer, presenting a high level of neoantigens, responded well to immunotherapies rather than adjuvant chemotherapies (Ribic et al., 2003; Wang et al., 2015; Passardi et al., 2017; Baran et al., 2018). To the best of our knowledge, previous researches mainly focused on the difference between right-sided and left-sided colon cancer (Petrelli et al., 2017; Mao et al., 2018; Zhang et al., 2018). Our study provides a more comprehensive analysis about right-sided, left-sided, and

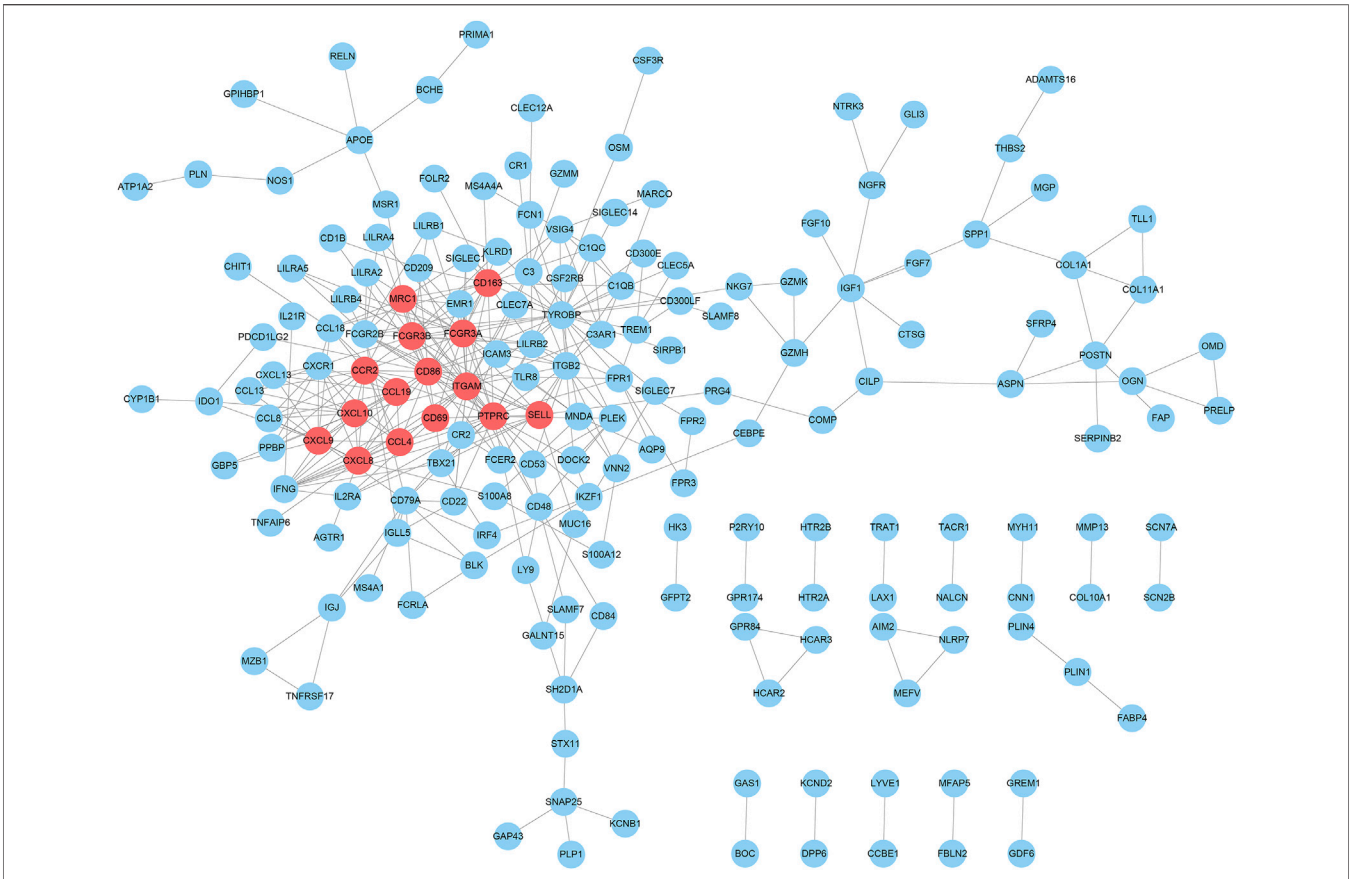


FIGURE 5 | PPI network analysis of DEGs and their hub genes screen. The hub gene nodes were highlighted in red.

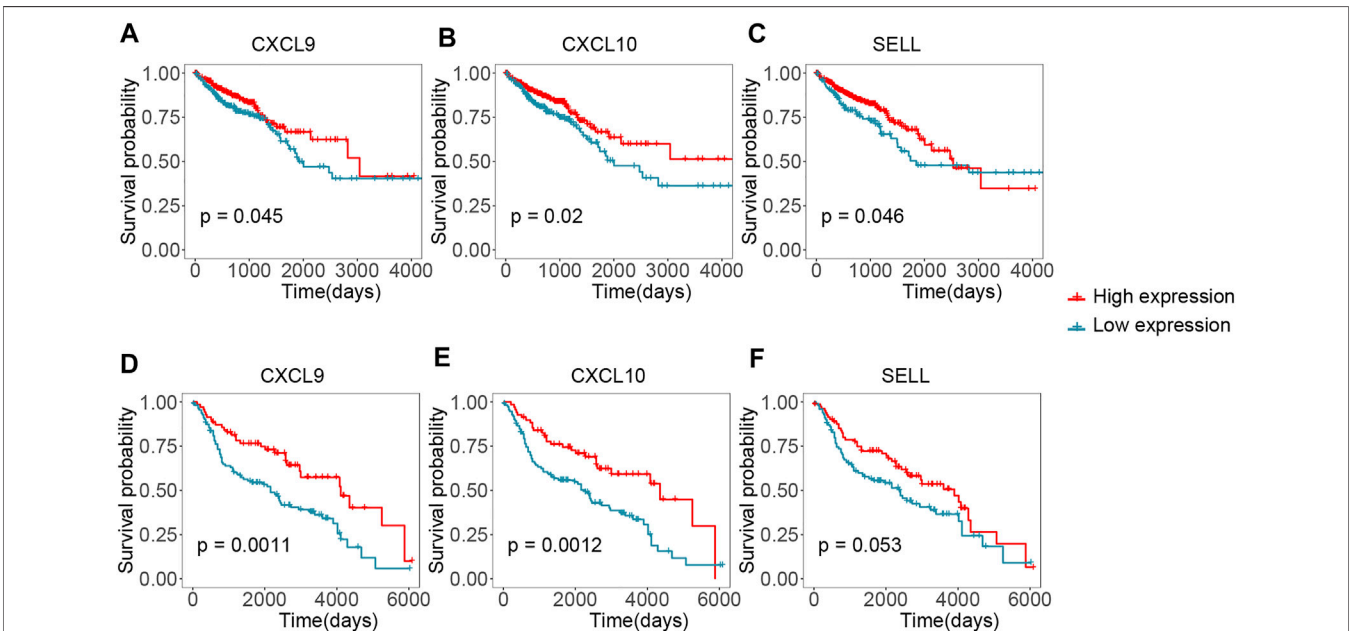


FIGURE 6 | Validating the hub genes by survival time in TCGA and GEO cohorts. **(A-C)** Kaplan-Meier plots reflecting the overall survival status of CRC patients in TCGA cohort. **(D-F)** Kaplan-Meier plots showing the overall survival status of CRC patients in GSE41258 cohort.

rectum in CRC patients. Our results further indicated that immune infiltration was different among right, left, and rectal CRCs. These immune infiltrating differences might contribute to the different survival time of CRC patients and providing a potential explanation for prognostic survival associated with primary tumor location (Petrelli et al., 2017).

Through the immune and stromal scores related DEGs analysis, a total of 318 DEGs were screened out and many of them were involved in tumor microenvironment related biological processes and pathways. Specifically, based on the DEGs analysis and GSEA annotation results, 188 DEGs were significantly correlated with immune score and most of them were involved in function that modulates the frequency or extent of an immune system process. Based on the analysis of DEGs and annotation of GSEA, 150 genes were significantly correlated with stromal score and mainly enriched in a structure that lies outside the plasma membrane and surrounds the entire cells.

Via PPI network construction, 15 genes (CD86, ITGAM, PTPRC, FCGR3A, FCGR3B, MRC1, CD163, CCR2, SELL, CD69, CXCL10, CXCL8, CXCL9, CCL19 and CCL4) were selected as hub genes. Especially, three genes (CXCL9, CXCL10, and SELL) were detected to be correlated with overall survival time both in the TCGA dataset and the validation GEO dataset. As shown in **Figure 6**, their higher expression was associated with an increased survival rate, indicating that they might be potential prognostic targets of CRC.

C-X-C motif chemokine ligand 9 (CXCL9, also known as CMK and MIG) and C-X-C Motif Chemokine Ligand 10 (CXCL10, also known INP10 and SCYB10) are mainly involved in selective and non-covalent interaction with the CXCR3 chemokine receptor and cytokine activity according to the Gene Ontology annotation. The protein encoded by CXCL9 is a member of CXC chemokine family that participates in T cell trafficking. Previous study suggested that CXCL9 plays an important role in different types of tumors (Ding et al., 2016). CXCL9 can be a tumor suppressor in breast cancer, non-small cell lung carcinoma, and colorectal cancer (Addison et al., 2000; Denkert et al., 2010; Wu et al., 2016). Conversely, it acts as tumor promoter in various types of cancer such as hepatocellular carcinoma, oral cavity squamous cell carcinoma, squamous cell cervical cancer, and chronic lymphocytic leukemia (Yan et al., 2011; Chang et al., 2013; Zhi et al., 2014; Liu et al., 2015). CXCL10 which is an important paralog of CXCL9, binds CXCR3 receptor to induce a variety of processes including chemotaxis, regulation of cell growth and apoptosis, regulation of angiostasis, and activation of immune cells (Liu et al., 2011; Sidahmed et al., 2012). The study of Chen et al. revealed that lower expression of CXCL10 was significantly associated with unsatisfied survival time (Chen et al., 2020). Our result showed that high expression of CXCL9 and CXCL10 were correlated with a better prognosis, which is consistent with studies of colorectal cancer in recent years (Wu et al., 2016; Chen et al., 2020).

SELL, also known as CD62L and L-selectin, belongs to the selectin family of glycoprotein adhesion molecules (Lefer, 2000), which is expressed on multiple tumor-infiltrating immune cells and abundant in the surface of neutrophils (Lefer, 2000; Kumari et al., 2021). Recent study suggest that L-selectin might be a favorable prognosis factor in breast cancer (Kumari et al., 2021). To the best of our knowledge, there are limited studies about SELL expression and overall survival time in colorectal cancer. In this study, the high level of SELL was found correlated with better survival of CRC patients, indicating that SELL might be a new potential prognostic biomarker in CRC.

In Summary, based on the tumor immune and stromal analysis, we found that tumor microenvironment was related to CRC survival outcome and clinical characteristics such as tumor stage and location. And we identified a series of candidate genes which might serve as prognostic biomarkers in CRC. However, there were some limitations in our study. All analysis was based on public data mining instead of experiments. More experiments need to be carried out in order to further verify our conclusion and have a comprehensive insight on the potential link between the tumor microenvironment and colorectal cancer. Our current findings might provide insights into understanding the potential role of tumor microenvironment in CRC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found in TCGA (<https://portal.gdc.cancer.gov/>) and GEO (<https://www.ncbi.nlm.nih.gov/geo/>).

AUTHOR CONTRIBUTIONS

LL, GF, and XD designed the study. LL collected and analyzed data. LL and XD wrote the manuscript. LL, XD, and GF contributed to writing this study. All authors had full access to the final version of the manuscript and agreed to its submission.

ACKNOWLEDGMENTS

We are grateful to the group of TCGA and GEO databases for the availability of the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.787208/full#supplementary-material>

REFERENCES

- Addison, C. L., Arenberg, D. A., Morris, S. B., Xue, Y.-Y., Burdick, M. D., Mulligan, M. S., et al. (2000). The CXCL Chemokine, Monokine Induced by Interferon-Gamma, Inhibits Non-small Cell Lung Carcinoma Tumor Growth and Metastasis. *Hum. Gene Ther.* 11 (2), 247–261. doi:10.1089/10430340050015996
- Arnold, M., Sierra, M. S., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2017). Global Patterns and Trends in Colorectal Cancer Incidence and Mortality. *Gut* 66 (4), 683–691. doi:10.1136/gutjnl-2015-310912
- Baran, B., Mert Ozupek, N., Yerli Tetik, N., Acar, E., Bekcioglu, O., and Baskin, Y. (2018). Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. *Gastroenterol. Res.* 11 (4), 264–273. doi:10.14740/gr1062w
- Bolouri, H. (2015). Network Dynamics in the Tumor Microenvironment. *Semin. Cancer Biol.* 30, 52–59. doi:10.1016/j.semcancer.2014.02.007
- Bosch, L. J. W., Carvalho, B., Fijneman, R. J. A., Jimenez, C. R., Pinedo, H. M., van Engeland, M., et al. (2011). Molecular Tests for Colorectal Cancer Screening. *Clin. Colorectal Cancer* 10 (1), 8–23. doi:10.3816/cc.2011.n.002
- Calon, A., Lonardo, E., Berenguer-Llgero, A., Espinet, E., Hernando-Mombona, X., Iglesias, M., et al. (2015). Stromal Gene Expression Defines Poor-Prognosis Subtypes in Colorectal Cancer. *Nat. Genet.* 47 (4), 320–329. doi:10.1038/ng.3225
- Chang, K.-P., Wu, C.-C., Fang, K.-H., Tsai, C.-Y., Chang, Y.-L., Liu, S.-C., et al. (2013). Serum Levels of Chemokine (C-X-C Motif) Ligand 9 (CXCL9) Are Associated with Tumor Progression and Treatment Outcome in Patients with Oral Cavity Squamous Cell Carcinoma. *Oral Oncol.* 49 (8), 802–807. doi:10.1016/j.oraloncology.2013.05.006
- Chen, J., Chen, Q.-L., Wang, W.-H., Chen, X.-L., Hu, X.-Q., Liang, Z.-Q., et al. (2020). Prognostic and Predictive Values of CXCL10 in Colorectal Cancer. *Clin. Transl. Oncol.* 22 (9), 1548–1564. doi:10.1007/s12094-020-02299-6
- Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: Identifying Hub Objects and Sub-Networks from Complex Interactome. *BMC Syst. Biol.* 8 (4), S11–S17. doi:10.1186/1752-0509-8-S4-S11
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiLinks: An R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507
- Deng, Z., Wang, J., Xu, B., Jin, Z., Wu, G., Zeng, J., et al. (2019). Mining TCGA Database for Tumor Microenvironment-Related Genes of Prognostic Value in Hepatocellular Carcinoma. *Biomed. Res. Int.* 2019 (3), 1–12. doi:10.1155/2019/2408348
- Denkert, C., Loibl, S., Noske, A., Roller, M., Müller, B. M., Komor, M., et al. (2010). Tumor-Associated Lymphocytes as an Independent Predictor of Response to Neoadjuvant Chemotherapy in Breast Cancer. *J. Clin. Oncol.* 28 (1), 105–113. doi:10.1200/jco.2009.23.7370
- Ding, Q., Lu, P., Xia, Y., Ding, S., Fan, Y., Li, X., et al. (2016). CXCL9: Evidence and Contradictions for its Role in Tumor Progression. *Cancer Med.* 5 (11), 3246–3259. doi:10.1002/cam4.934
- 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 (4), 298–306. doi:10.1038/nrc3245
- Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., et al. (2006). Type, Density, and Location of Immune Cells within Human Colorectal Tumors Predict Clinical Outcome. *Science* 313 (5795), 1960–1964. doi:10.1126/science.1129139
- Galon, J., Fridman, W.-H., and Pagès, F. (2007). The Adaptive Immunologic Microenvironment in Colorectal Cancer: A Novel Perspective: Figure 1. *Cancer Res.* 67 (5), 1883–1886. doi:10.1158/0008-5472.can-06-4806
- Ganesh, K., Stadler, Z. K., Cercek, A., Mendelsohn, R. B., Shia, J., Segal, N. H., et al. (2019). Immunotherapy in Colorectal Cancer: Rationale, Challenges and Potential. *Nat. Rev. Gastroenterol. Hepatol.* 16 (6), 361–375. doi:10.1038/s41575-019-0126-x
- Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). affy--Analysis of Affymetrix GeneChip Data at the Probe Level. *Bioinformatics* 20 (3), 307–315. doi:10.1093/bioinformatics/btg405
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013
- 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 (1), 177. doi:10.1186/s12943-018-0923-9
- Isella, C., Terrasi, A., Bellomo, S. E., Petti, C., Galatola, G., Muratore, A., et al. (2015). Stromal Contribution to the Colorectal Cancer Transcriptome. *Nat. Genet.* 47 (4), 312–319. doi:10.1038/ng.3224
- Kalluri, R., and Zeisberg, M. (2006). Fibroblasts in Cancer. *Nat. Rev. Cancer* 6 (5), 392–401. doi:10.1038/nrc1877
- Kamal, Y., Schmit, S. L., Frost, H. R., and Amos, C. I. (2020). The Tumor Microenvironment of Colorectal Cancer Metastases: Opportunities in Cancer Immunotherapy. *Immunotherapy* 12 (14), 1083–1100. doi:10.2217/imt-2020-0026
- Kassambara, A., Kosinski, M., Bieчек, P., and Fabian, S. (2019). Survminer: Drawing Survival Curves Using 'ggplot2'. R package version 0.4.6. Available at: <https://CRAN.R-project.org/package=survminer> (Accessed September 4, 2019).
- Kawakami, H., Zaanani, A., and Sinicrope, F. A. (2015). Microsatellite Instability Testing and its Role in the Management of Colorectal Cancer. *Curr. Treat. Options. Oncol.* 16 (7), 30. doi:10.1007/s11864-015-0348-2
- Kolde, R. (2019). Pheatmap: Pretty Heatmaps. R package version 1.0.12. Available at: <https://CRAN.R-project.org/package=pheatmap> (Accessed January 4, 2019).
- Kumari, S., Arora, M., Singh, J., Chauhan, S. S., Kumar, S., and Chopra, A. (2021). L-Selectin Expression Is Associated with Inflammatory Microenvironment and Favourable Prognosis in Breast Cancer. *3 Biotech.* 11 (2), 1–13. doi:10.1007/s13205-020-02549-y
- Lefer, D. J. (2000). Pharmacology of Selectin Inhibitors in Ischemia/reperfusion States. *Annu. Rev. Pharmacol. Toxicol.* 40, 283–294. doi:10.1146/annurev.pharmtox.40.1.283
- Lim, C. J., Lee, Y. H., Pan, L., Lai, L., Chua, C., Wasser, M., et al. (2018). Multidimensional Analyses Reveal Distinct Immune Microenvironment in Hepatitis B Virus-Related Hepatocellular Carcinoma. *Gut* 68 (5), 916–927. doi:10.1136/gutjnl-2018-316510
- Liu, M., Guo, S., Hibbert, J. M., Jain, V., Singh, N., Wilson, N. O., et al. (2011). CXCL10/IP-10 in Infectious Diseases Pathogenesis and Potential Therapeutic Implications. *Cytokine Growth Factor. Rev.* 22 (3), 121–130. doi:10.1016/j.cytogfr.2011.06.001
- Liu, R.-X., Wei, Y., Zeng, Q.-H., Chan, K.-W., Xiao, X., Zhao, X.-Y., et al. (2015). Chemokine (C-X-C Motif) Receptor 3-Positive B Cells Link Interleukin-17 Inflammation to Protumorigenic Macrophage Polarization in Human Hepatocellular Carcinoma. *Hepatology* 62 (6), 1779–1790. doi:10.1002/hep.28020
- Liu, T., Han, C., Wang, S., Fang, P., Ma, Z., Xu, L., et al. (2019). Cancer-Associated Fibroblasts: an Emerging Target of Anti-Cancer Immunotherapy. *J. Hematol. Oncol.* 12 (1), 86–15. doi:10.1186/s13045-019-0770-1
- Lorusso, G., and Rüegg, C. (2008). The Tumor Microenvironment and its Contribution to Tumor Evolution toward Metastasis. *Histochem. Cel. Biol.* 130 (6), 1091–1103. doi:10.1007/s00418-008-0530-8
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8
- Mao, Y., Feng, Q., Zheng, P., Yang, L., Liu, T., Xu, Y., et al. (2018). Low Tumor Purity Is Associated with Poor Prognosis, Heavy Mutation burden, and Intense Immune Phenotype in colon Cancer. *Cancer Manag. Res.* 10, 3569–3577. doi:10.2147/cmcr.s171855
- Mlecnik, B., Bindea, G., Angell, H. K., Maby, P., Angelova, M., Tougeron, D., et al. (2016). Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. *Immunity* 44 (3), 698–711. doi:10.1016/j.immuni.2016.02.025
- Mlecnik, B., Tosolini, M., Kirilovsky, A., Berger, A., Bindea, G., Meatchi, T., et al. (2011). Histopathologic-Based Prognostic Factors of Colorectal Cancers Are Associated with the State of the Local Immune Reaction. *J. Clin. Oncol.* 29 (6), 610–618. doi:10.1200/jco.2010.30.5425
- Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1 α -Responsive Genes Involved in Oxidative Phosphorylation Are Coordinately Downregulated in Human Diabetes. *Nat. Genet.* 34 (3), 267–273. doi:10.1038/ng1180

- Pagès, F., Berger, A., Camus, M., Sanchez-Cabo, F., Costes, A., Molitor, R., et al. (2005). Effector Memory T Cells, Early Metastasis, and Survival in Colorectal Cancer. *N. Engl. J. Med.* 353 (25), 2654–2666. doi:10.1056/NEJMoa051424
- Passardi, A., Canale, M., Valgiusti, M., and Ulivi, P. (2017). Immune Checkpoints as a Target for Colorectal Cancer Treatment. *Int. J. Mol. Sci.* 18 (6), 1324. doi:10.3390/ijms18061324
- Peddareddigari, V. G., Wang, D., and DuBois, R. N. (2010). The Tumor Microenvironment in Colorectal Carcinogenesis. *Cancer Microenvironment* 3 (1), 149–166. doi:10.1007/s12307-010-0038-3
- Pedrosa, L., Esposito, F., Thomson, T. M., and Maurel, J. (2019). The Tumor Microenvironment in Colorectal Cancer Therapy. *Cancers (Basel)* 11 (8), 1172. doi:10.3390/cancers11081172
- Petrelli, F., Tomasello, G., Borgonovo, K., Ghidini, M., Turati, L., Dallera, P., et al. (2017). Prognostic Survival Associated with Left-Sided vs Right-Sided Colon Cancer. *JAMA Oncol.* 3 (2), 211–219. doi:10.1001/jamaoncol.2016.4227
- Quail, D. F., and Joyce, J. A. (2013). Microenvironmental Regulation of Tumor Progression and Metastasis. *Nat. Med.* 19 (11), 1423–1437. doi:10.1038/nm.3394
- Ribic, C. M., Sargent, D. J., Moore, M. J., Thibodeau, S. N., French, A. J., Goldberg, R. M., et al. (2003). Tumor Microsatellite-Instability Status as a Predictor of Benefit from Fluorouracil-Based Adjuvant Chemotherapy for colon Cancer. *N. Engl. J. Med.* 349 (3), 247–257. doi:10.1056/nejmoa022289
- Sawicki, T., Ruzskowska, M., Danielewicz, A., Niedźwiedzka, E., Arlukowicz, T., and Przybyłowicz, K. E. (2021). A Review of Colorectal Cancer in Terms of Epidemiology, Risk Factors, Development, Symptoms and Diagnosis. *Cancers* 13 (9), 2025. doi:10.3390/cancers13092025
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303
- Sidahmed, A. M. E., León, A. J., Bosinger, S. E., Banner, D., Danesh, A., Cameron, M. J., et al. (2012). CXCL10 Contributes to P38-Mediated Apoptosis in Primary T Lymphocytes. *In Vitro. Cytokine* 59 (2), 433–441. doi:10.1016/j.cyto.2012.05.002
- Siegel, R. L., Miller, K. D., Goding Sauer, A., Fedewa, S. A., Butterly, L. F., Anderson, J. C., et al. (2020). Colorectal Cancer Statistics, 2020. *CA A. Cancer J. Clin.* 70 (3), 145–164. doi:10.3322/caac.21601
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING V10: Protein-Protein Interaction Networks, Integrated over the Tree of Life. *Nucleic Acids Res.* 43 (D1), D447–D452. doi:10.1093/nar/gku1003
- Therneau, T. M. (2019). Survival: A Package for Survival Analysis in R. R Package Version 3.1-8. Available at: <https://CRAN.R-project.org/package=survival> (Accessed December 3, 2019).
- Turley, S. J., Cremasco, V., and Astarita, J. L. (2015). Immunological Hallmarks of Stromal Cells in the Tumour Microenvironment. *Nat. Rev. Immunol.* 15 (11), 669–682. doi:10.1038/nri3902
- Wang, F., Bai, L., Liu, T. S., Yu, Y. Y., He, M. M., Liu, K. Y., et al. (2015). Right-Sided colon Cancer and Left-Sided Colorectal Cancers Respond Differently to Cetuximab. *Chin. J. Cancer* 34 (3), 384–393. doi:10.1186/s40880-015-0022-x
- Wang, H., Wu, X., and Chen, Y. (2019). Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer. *Front. Oncol.* 9, 1212. doi:10.3389/fonc.2019.01212
- Wong, M. C. S., Huang, J., Lok, V., Wang, J., Fung, F., Ding, H., et al. (2021). Differences in Incidence and Mortality Trends of Colorectal Cancer Worldwide Based on Sex, Age, and Anatomic Location. *Clin. Gastroenterol. Hepatol.* 19 (5), 955–966. e61. doi:10.1016/j.cgh.2020.02.026
- Wu, Z., Huang, X., Han, X., Li, Z., Zhu, Q., Yan, J., et al. (2016). The Chemokine CXCL9 Expression Is Associated with Better Prognosis for Colorectal Carcinoma Patients. *Biomed. Pharmacother.* 78, 8–13. doi:10.1016/j.biopha.2015.12.021
- Yan, X.-J., Dozmorov, I., Li, W., Yancopoulos, S., Sison, C., Centola, M., et al. (2011). Identification of Outcome-Correlated Cytokine Clusters in Chronic Lymphocytic Leukemia. *Blood* 118 (19), 5201–5210. doi:10.1182/blood-2011-03-342436
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4 (1), 2612. doi:10.1038/ncomms3612
- Zhang, L., Zhao, Y., Dai, Y., Cheng, J.-N., Gong, Z., Feng, Y., et al. (2018). Immune Landscape of Colorectal Cancer Tumor Microenvironment from Different Primary Tumor Location. *Front. Immunol.* 9, 1578. doi:10.3389/fimmu.2018.01578
- Zhi, W., Ferris, D., Sharma, A., Purohit, S., Santos, C., He, M., et al. (2014). Twelve Serum Proteins Progressively Increase with Disease Stage in Squamous Cell Cervical Cancer Patients. *Int. J. Gynecol. Cancer* 24 (6), 1085–1092. doi:10.1097/IGC.0000000000000153

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.

Copyright © 2022 Li, Du and Fan. 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.