Prediction and identification of immune genes related to the prognosis of patients with colon adenocarcinoma and its mechanisms

Sihan Chen

Anhui Medical University

Cao GD

Anhui Medical University

Wu wei

Anhui Medical University

Lu Yida

Anhui Medical University

He Xiaobo

Anhui Medical University

Yang Lei

Anhui Medical University

Chen Ke

Anhui Medical University

Bo Chen

Anhui Medical University

MaoMing Xiong (✉ ayfyxmm@163.com)

Research

Keywords: Bioinformatics analysis, prognosis, colon adenocarcinoma, immune genes

Posted Date: May 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-26039/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on June 29th, 2020. See the published version at <https://doi.org/10.1186/s12957-020-01921-9>.

Abstract

Background

Colon adenocarcinoma(COAD) is a type of gastrointestinal tumor with a high degree of malignancy, and its immunotherapy method has been stagnant. Recently, with the development of network databases and the development of bioinformatics tools, we can analyze the data in the current database to find out which genes may be the target of immunotherapy for COAD.

Methods

We use various codes and packages in R language to analyze the downloaded data, construct a riskScore model of immune genes and clinical data, and use Cox regression analysis to explore the clinical relationship between riskScore and COAD.

Results

We found that seven immune genes associated with the survival prognosis of COAD were related to the construction of a riskScore model. And Cox regression analysis found that there was clinical significance and statistical significance between the riskScore model and clinical data of COAD. Some traditional immune microenvironment cells also increase their cell content with the increase of riskScore.

Conclusions

We found 7 immune genes (SLC10A2, CXCL3, IGHV5-51, INHBA, STC1, UCN, and OXTR) that affect the clinical prognosis of COAD patients and can construct a riskScore model. It may be a target for future COAD immunotherapy.

1. Background

COAD is a type of malignant tumor of the digestive tract, which can be subdivided into right COAD and left COAD according to location. According to a WHO report in 2018, COAD is the third most common adenocarcinoma worldwide, and 1.8 million COAD cases were diagnosed in 2018 (10% of all tumors). Adenocarcinoma of the colon is relatively common in both men and women. There were 881,1,000 patients who died of COAD in 2018 [1]. There are many ways to treat COAD. Traditional treatment methods include surgery and chemotherapy. However, immunotherapy has been stagnant in the treatment of COAD. We look for new targets for COAD immunotherapy, which may help us with COAD immunotherapy [2].

With the development of various network databases, such as: clinical database TCGA, immune gene database IMMPort, TIMER, and transcription factor database Cistrome. We can find the immune genes related to the prognosis of COAD through data analysis methods such as R language. Cox regression analysis was used to make a riskScore model of immune genes related to prognosis and further correlate clinical data. Finally, seven immune genes that constituted a riskScore model and related to the prognosis of COAD were obtained.

In this study, our goal was to find immune genes that not only constitute a risk score model, but also correlate with the prognosis of COAD.

To facilitate us to provide more diagnosis and treatment ideas in clinical function and COAD immunotherapy.

2. Methods

2.1 Gets relevant data from the network database. The COAD gene expression data and clinical data of Asian yellow people were obtained from the TCGA database [3]. Immune gene names were obtained from the IMMPort database, and immune genes were screened from the downloaded data [4]. Transcription factor data from CISTROME database [5], Tumor microenvironment-related gene infiltration data were obtained from the TIMER database [6].

2.2 Acquisition of differentially expressed genes of interest. Use the R language, BiocManager package and Limmma package to obtain differentially expressed genes of interest. The screening conditions for differential genes and differential immune genes are $\text{LogFC} \geq 2$, $\text{FDR} < 0.05$. The screening conditions for differential transcription factors were: $\text{LogFC} > 1$; $P < 0.05$.

2.3 Calculate RiskScore. The calculation formula of risk score is: $\sum (\text{Exp}^{\text{mRNA1-n}} \times \text{coef}^{\text{mRNA1-n}})$. "exp" indicates the expression level of the gene, "coe" indicates the correlation coefficient of the gene.

2.4 Cox regression analysis and assessment of risk scores. Cox regression analysis Survival package using R language. The filter coefficient of univariate Cox regression of immune genes was $P = 0.01$. Multivariate Cox regression models were used to assess the riskScore of immune genes and quantify them.

2.5 Immune genes will draw the interaction network of transcription factors. Mapping the interaction network of immune genes and transcription factors using R and Cytoscape, the screening conditions are $\text{Cor} = 0.4$, $P = 0.01$. [7]

2.6 Correlation analysis of immune genes and clinical data. Correlation analysis between selected immune genes and clinical data was performed using the beeswarm package in R language.

2.7 Related picture drawing. Drawing of heat maps, use the pheatmap package in R to draw a heat map. Use the survival package and survminer package of R language to draw the Kaplan-Meier survival curve of

riskScore(Screening conditions for survival prognosis data: follow-up data less than 90 days were excluded). Use the survivalROC package to draw the ROC curve. Use R-related code to draw and draw survival status charts risk curves Volcano map and forest map, etc.

3. Result

3.1 The related process is shown in Fig. 1. Acquisition of immune DEGs and differentially expressed transcription factors.

The research team downloaded clinical data and gene expression data of 385 COAD patients from the TCGA database and obtained differentially expressed genes (DEGs) through screening (screening conditions: $|\text{LogFC}| \geq 2$, $\text{FDR} < 0.05$) (Fig. 2A,B). The IMMPORT database contains the names of a large number of immune genes. We obtain the differentially expressed immune genes through the intersection of the immune gene names and DEGs (Fig. 2C,D). The transcription factor names were obtained from the Cistrome database, and the differentially expressed transcription factors were obtained by intersecting with the DEGs. the screening conditions for transcription factors are: $\text{LogFC} > |1|$; $P < 0.05$. The volcano and heat maps were drawn (Fig. 2E,F).

3.2 Construction of the prognostic related immune gene model and the interaction network of prognostic related immune genes and transcription factors.

Univariate Cox regression analysis was used to study the differentially expressed immune genes related to survival prognosis. The results showed that there are 12 immune genes that are closely related to the prognosis of gastric adenocarcinoma in Asian yellow people. (Fig. 3A). $\text{Cor} = 0.4$, $P = 0.01$ screening criteria were used to establish the interaction between immune genes and transcription factors, and network diagrams were made (Fig. 3B). Details of the regulatory relationship between transcription factors and the immune genes associated with COAD prognosis are shown in Table 1. The results showed that 8 immune genes are closely related to the regulation of transcription factors and belong to positive regulation.

Table 1
Details of the regulatory relationship between transcription factors
and immune genes.

TF	immune Gene	cor	p-value	Regulation
E2F1	NOX4	-5.63E-01	1.14E-06	negative
	INHBA	-5.26E-01	3.04E-05	negative
	STC1	-4.67E-01	2.17E-03	negative
	VIP	-4.71E-01	1.75E-03	negative
FOSL1	INHBA	4.67E-01	2.16E-03	positive
	STC1	4.67E-01	2.18E-03	positive
	VIP	4.57E-01	2.01E-06	positive
	OXTR	4.85E-01	7.00E-04	positive
HOXC11	SLC10A2	9.34E-01	3.98E-10	positive
KLF4	CXCL3	5.15E-01	7.63E-05	positive
	NOX4	-4.96E-01	3.09E-04	negative
LEF1	SLC10A2	4.87E-01	5.95E-04	positive
	CXCL3	-4.56E-01	4.22E-03	negative
	NOX4	4.82E-01	8.60E-04	positive
	INHBA	4.84E-01	7.51E-04	positive
	VIP	4.65E-01	2.56E-03	positive
MYBL2	NOX4	-5.45E-01	5.71E-06	negative
	INHBA	-5.31E-01	2.01E-05	negative
	STC1	-4.53E + 01	4.94E-03	negative
	VIP	-4.58E-01	3.75E-03	negative
MYH11	CXCL3	-4.58E-01	3.80E-03	negative
	NOX4	4.55E-01	4.44E-03	positive
	CCL19	6.40E-01	1.84E-10	positive
	INHBA	4.80E-01	9.71E-04	positive
	STC1	4.68E-01	2.04E-03	positive
	VIP	7.22E-01	5.91E-20	positive

TF	immune Gene	cor	p-value	Regulation
SALL4	SLC10A2	6.22E-01	1.66E-09	positive
	CXCL3	-4.84E-01	7.37E-04	negative
	INHBA	4.70E-01	1.81E-03	positive
SPIB	CCL19	4.77E-01	1.17E-03	positive
TFAP2A	STC1	4.49E-01	6.26E-03	positive

3.3 Calculation of immune gene riskScore and construction of survival prognosis model.

We calculated the Coef coefficient of immune genes related to survival prognosis by multivariate Cox regression, and then calculated the riskScore of immune genes by the product of Coef coefficient and gene expression. Among the immune genes related to survival prognosis, 7 immune genes are closely related to the composition of risk scores, which are SLC10A2, CXCL3, IGHV5-51, INHBA, STC1, UCN and OXTR (Table 2), this is also the key immune gene that we will study later. According to the median riskScore, the riskScore is divided into two groups. Survival and surmiser packages in R were used to correlate risk scores with survival prognosis and draw the Kaplan-Meier survival curves. The results showed that $P = 8.876e-04$. The survival prognosis of the high-risk group was significantly worse than that of the low-expression group (Fig. 4A), The survivalROC package of R language is used to draw the ROC curve. The results show that the AUC of the ROC curve = 0.749 (Fig. 4B). Detailed data on the survival rates of high- and low-risk patients are shown in Table 3 and Table 4.

Table 2
Details of the seven immune genes used to construct the riskScore model.

id	coef	HR	HR.95L	HR.95H	p-value
SLC10A2	0.650	1.916	1.203	3.050	0.006
CXCL3	-0.019	0.981	0.964	0.998	0.033
IGHV5-51	0.002	1.002	1.000	1.003	0.005
INHBA	0.046	1.047	1.003	1.093	0.038
STC1	0.058	1.059	0.991	1.133	0.092
UCN	0.405	1.499	1.198	1.876	0.000
OXTR	0.229	1.258	0.996	1.588	0.054

Table 3
Detailed data for high risk survival analysis.

Time(year)	n.Risk	n.Event	Survival(%)	Std.err	Lower 95% CI	Upper 95% CI
0.266	167	1	0.994	0.00597	0.9824	1
0.419	163	1	0.988	0.0085	0.9714	1
0.427	160	1	0.982	0.01045	0.9615	1
0.436	159	1	0.976	0.01207	0.9522	1
0.471	158	1	0.969	0.01348	0.9433	0.996
0.515	154	1	0.963	0.01479	0.9345	0.993
0.586	150	1	0.957	0.01602	0.9258	0.989
0.625	149	1	0.95	0.01715	0.9172	0.984
0.718	142	1	0.944	0.01829	0.9084	0.98
0.795	136	1	0.937	0.01943	0.8993	0.975
0.827	134	1	0.93	0.0205	0.8903	0.971
0.838	131	2	0.915	0.02251	0.8724	0.961
0.907	127	1	0.908	0.02346	0.8634	0.955
0.926	123	1	0.901	0.0244	0.8543	0.95
1.008	117	1	0.893	0.02538	0.8448	0.944
1.049	112	1	0.885	0.02638	0.835	0.938
1.085	108	1	0.877	0.02738	0.8249	0.932
1.104	106	1	0.869	0.02834	0.8149	0.926
1.162	101	1	0.86	0.02934	0.8045	0.92
1.167	99	1	0.851	0.0303	0.7941	0.913
1.293	88	1	0.842	0.03146	0.7823	0.906
1.359	81	1	0.831	0.03275	0.7696	0.898
1.4	77	1	0.821	0.03405	0.7565	0.89
1.762	57	1	0.806	0.03637	0.7379	0.881
1.836	53	1	0.791	0.03874	0.7186	0.871
1.868	52	1	0.776	0.04087	0.6996	0.86
2.036	49	1	0.76	0.04299	0.6801	0.849

Time(year)	n.Risk	n.Event	Survival(%)	Std.err	Lower 95% CI	Upper 95% CI
2.205	46	1	0.743	0.04512	0.66	0.837
3.693	16	1	0.697	0.06175	0.5858	0.829
3.784	15	1	0.65	0.07305	0.5219	0.811
4.688	10	1	0.585	0.09017	0.4329	0.792
5.066	9	1	0.52	0.10092	0.3558	0.761
5.233	8	1	0.455	0.10724	0.287	0.722
5.488	7	1	0.39	0.10989	0.2248	0.678
8.334	3	1	0.26	0.12903	0.0984	0.688

Table 4
Detailed data for low risk survival analysis.

Time(year)	n.Risk	n.Event	Survival(%)	Std.err	Lower 95% CI	Upper 95% CI
0.247	167	1	0.994	0.00597	0.982	1
0.4	161	1	0.988	0.00855	0.971	1
0.419	159	2	0.975	0.01214	0.952	1
0.564	154	1	0.969	0.01362	0.943	0.996
0.663	148	1	0.963	0.01502	0.934	0.992
0.918	138	1	0.956	0.01645	0.924	0.988
0.978	133	1	0.948	0.01783	0.914	0.984
1.211	114	1	0.94	0.01951	0.903	0.979
2.252	71	1	0.927	0.0233	0.882	0.974
2.351	69	1	0.913	0.02655	0.863	0.967
2.463	66	1	0.9	0.02954	0.843	0.959
2.997	52	1	0.882	0.03366	0.819	0.951
3.173	43	1	0.862	0.03863	0.789	0.941
3.184	42	1	0.841	0.04281	0.761	0.929
4.09	31	1	0.814	0.04928	0.723	0.917
4.118	30	1	0.787	0.0546	0.687	0.902
5.153	16	1	0.738	0.06992	0.613	0.888
6.781	10	1	0.664	0.09412	0.503	0.877
7.729	7	1	0.569	0.11925	0.377	0.858

3.4 Immune gene risk curve mapping and independent prognostic analysis.

Use R language related codes to draw related pictures of the risk curve. The results showed that with the gradual increase of the immune gene risk value, the survival time of patients gradually decreased (Fig. 5A,B). Heat map of related immune gene expression(Fig. 5C). Univariate independent prognostic analysis showed that the Hazard ratio of riskScore was 1.033 (1.018–1.049), and $P < 0.001$. Multivariate independent prognostic analysis showed that Hazard ration of riskScore was 1.026 (1.011–1.042), $P \leq 0.001$ (Fig. 6A,B). RiskSore are clinically and statistically significant.

3.5 Correlation analysis of immune genes and clinical data.

We analyzed the correlation between the 7 immune genes that make up the immune score and clinical data, using the R language beeswarm package. The results showed that there were statistically significant correlations between seven immune genes and clinical data, namely, CXCL3, OXTR and STC1 (Fig. 7). Among them, the expression of CXCL3 were statistically significant in correlation with stage, while the expressions of OXTR and STC1 were statistically significant in correlation with T. CXCL3 also has significant differences in N and M.

3.6 Correlation Analysis of Risk Score and Tumor Microenvironment Cells.

Correlation analysis was performed between the risk scores assessed by our research and immune microenvironment genes, and the results showed that in CD4, CD8, Dendritic, Macrophage, and Neutrophil cells, as the risk scores increased, the expression levels of these genes became upward. And It has statistical significance $P < 0.05$. Correlation analysis of our risk value model with some widely recognized genes that constitute the immune microenvironment showed that CD4, CD8, Dendritic, Macrophage, and Neutrophil cells were positively correlated with the riskScore model. As the riskScore increases, so does the expression of these genes (Fig. 8).

3.7 Correlation between RiskScore model and tumor microenvironment.

In order to evaluate the difference in immune cell content between samples with high RiskScore and samples with low RiskScore, we selected 20 sets of samples, which were selected from the 10 samples with the highest RiskScore and the 10 samples with the lowest RiskScore, through the EPIC database. Calculate the difference in the amount of their direct immune cells. The results showed that in 10 samples with high RiskScore, the content of CAFs cells was significantly higher than that with 10 samples with low RiskScore. From this we can know that CAFs cells play an important role in RiskScore (Fig. 9).

4. Discussion

The process is shown in Fig. 1. We searched the TCGA database for 385 cases of COAD and downloaded them. The clinical data and gene expression data in the download data were integrated, and the Limme package of R language was used to extract differentially expressed genes (DEGs). Using the immune gene names provided by IMMPORT, we can easily screen out the differential gene immune genes from DEGs. In the same way, we download the names of transcription factors from the CISTROME database and screen out the differentially expressed transcription factors from DEGs (screening conditions: $|\text{LogFC}| \geq 2$, $\text{FDR} < 0.05$). We then conduct further analysis of differentially expressed immune genes (screening conditions: $|\text{LogFC}| \geq 2$, $\text{FDR} < 0.05$) and differentially expressed transcription factors (screening conditions: $|\text{LogFC}| \geq 1$, $\text{FDR} < 0.05$). The differentially expressed immune genes were analyzed by univariate Cox regression using the survival package of R language and clinical data to obtain immune genes related to prognosis. Prognostic related immune genes are as follows: SLC10A2, CXCL3, NOX4, CCL19, IGHG1, IGHV5-51, IGKV1-33, INHBA, STC1, UCN, VIP and OXTR (Fig. 3A). We

performed an interaction network analysis of prognostic-associated immune genes and differentially expressed transcription factors, and the results are shown in (Fig. 3B).

We excluded samples with a survival time of less than 90 days from the downloaded clinical data, and assessed the survival prognosis by the level of riskScore. The results showed that patients with high risk scores had significantly worse survival prognosis than patients with low risk scores, $P = 8.876e-04$ (Fig. 4A). The ROC curve showed that $AUC = 0.749$, and the riskScore and prognosis model were more reliable (Fig. 4B). This allows us to group patients according to risk scores in clinical work to predict their prognosis. According to Fig. 5A and Fig. 5B, we can find that as the riskScore increases, the survival time of the patient decreases. The heat map of Fig. 5C also shows that the genes that construct risk scores have higher expression levels in the high-risk score array. We included clinical data on COAD and the riskScore evaluated in this study into the Cox regression analysis. The results showed that stage, T, M, N, and riskScore were statistically significant and clinically significant in the survival prognosis of the patients in the univariate Cox regression analysis. However, in the results of multivariate Cox regression analysis, age, stage, T, and riskScore have statistical significance and clinical significance. Based on the analysis of the seven genes and clinical data used to construct the riskScore model, the results show that the expression of CXCL3 gene in M, N, and stage is higher than that in late stage [8]. This is likely to be related to the mechanism of the immune microenvironment. Studies on the immune microenvironment have shown that some genes that construct the immune microenvironment can promote tumor progression (Fig. 7). Some genes that make up the tumor microenvironment, such as B, CD4, CD8, Dendritic, Macrophage, and Neutrophil cells, have been shown in research to be correlated with the survival prognosis of many types of tumor patients [8]. We downloaded the data of these cells through the TIMER database and performed correlation analysis with the riskScore model we constructed. The results showed that the expression of CD4, CD8, Dendritic, Macrophage, and Neutrophil cells increased with the increase of the riskScore. This also confirms on the side that the risk score model we constructed has a certain predictive ability for the clinical prognosis of patients.

The formation of the tumor microenvironment is closely related to the occurrence and development of tumors [9]. By studying the cells that constitute the tumor microenvironment, we can effectively find many cells or genes that are closely related to the clinical prognosis of patients. In order to evaluate the difference in immune cell content between samples with high RiskScore and samples with low RiskScore, we further evaluated them in the EPIC database. We selected 20 sets of samples, which were selected from the 10 samples with the highest RiskScore and the 10 samples with the lowest RiskScore, which passed the EPIC database [24]. Calculate the difference in the amount of their direct immune cells. The results showed that in 10 samples with high RiskScore, the content of CAFs cells was significantly higher than that with 10 samples with low RiskScore. From this we can know that CAFs cells play an important role in RiskScore. The related literature reports that CAFs cells can promote tumor deterioration in multiple tumors. We judge that the poor prognosis of patients is related to the excessive expression of CAFs cells [25].

In this study, we constructed an immune gene risk score model for 385 COAD patients through correlation analysis. Through a series of analyses of the disease, it was found that the riskScore is closely related to the survival prognosis of patients. In future clinical treatments, we can use the risk score model to effectively predict the survival prognosis of patients with COAD, and we can do targeted immunotherapy for 7 immune genes (SLC10A2, CXCL3, IGHV5-51, INHBA, STC1, UCN, and OXTR) that constitute the riskScore to improve the prognosis of patients and improve the treatment effect.

SLC10A2 is mainly used to mediate the bile in the intestinal circulation and assist colonization of the intestinal flora. Recently, some research reports have reported that slc10a2 / PPAR γ / PTEN / mTOR Signaling Pathway is related to the development of lung cancer [10, 11]. CXCL3 is related to the occurrence and development of prostate cancer, colon cancer, and breast cancer. There are also reports in the literature that the effect of suppressing the development of colon cancer can be achieved by immunosuppression of CXCL3 [12–16]. INHBA has a significant relationship with the occurrence and development of gastric, esophageal, and ovarian cancers, and studies have reported that the immunosuppressive treatment of INHBA can reduce the rate of deterioration of gastric and ovarian cancers [17–19]. STC1, OXTR, and UCN can promote the metastasis of colon cancer [20–23].

5. Conclusion

We download data for COAD, immune genes and transcription factors through a series of bioinformatics databases. A riskScore model of COAD immune genes was constructed. Through a series of clinical correlation analysis, it was found that 7 immune genes (SLC10A2, CXCL3, IGHV5-51, INHBA, STC1, UCN and OXTR) were correlated with clinical prognosis and riskScore of patients with COAD. These seven immune genes combined with their research background are expected to become target genes for immunotherapy.

Abbreviations

colon adenocarcinoma(COAD); Hazard ratio(HR); Transcription factor(TF)

Declarations

Ethics approval and consent to participate: There were no cell, tissue, or animal studies. No ethical requirements are involved.

Consent for publication: All authors agree to publish the paper.

Availability of data and material: The data used to support the findings of this study are included within the article.

Competing interests: The authors declare that they have no competing interests.

Funding: The Key Research and Development Plan Projects of Anhui Province (No.201904a07020045)

Authors' contributions: Sihan Chen is responsible for writing and submitting the papers; Cao GD, Wu wei, LU Yida are responsible for data analysis and collation; He Xiaobo, Yang Lei, Chen Ke are responsible for the production of pictures; Bo Chen, MM Xiong, are responsible for the manuscript fees and ideas guidance.

Acknowledgements: Thanks to the Fund (The Key Research and Development Plan Projects of Anhui Province) for supporting this research.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin*. 2018;68(6):394–424.
2. Villarreal DO, L'Huillier A, Armington S, Mottershead C, Filippova EV, Coder BD, Petit RG, Princiotta MF. Targeting CCR8 Induces Protective Antitumor Immunity and Enhances Vaccine-Induced Responses in Colon Cancer. *Cancer research*. 2018;78(18):5340–8.
3. Sharma S, Ciufu S, Starchenko E, Darji D, Chlumsky L, Karsch-Mizrachi I, Schoch CL: The NCBI BioCollections Database. *Database (Oxford)* 2018, 2018: bay006.
4. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, Hu Z, Zalocusky KA, Shankar RD, Shen-Orr SS, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci Data*. 2018;5:180015–5.
5. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, Chen C-H, Brown M, Zhang X, Meyer CA, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res*. 2019;47(D1):D729–35.
6. Li B, Severson E, Pignon J-C, Zhao H, Li T, Novak J, Jiang P, Shen H, Aster JC, Rodig S, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol*. 2016;17(1):174–4.
7. Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol*. 2019;20(1):185–5.
8. Wu T, Dai Y. Tumor microenvironment and therapeutic response. *Cancer letters*. 2017;387:61–8.
9. Frankel T, Lanfranca MP, Zou W. The Role of Tumor Microenvironment in Cancer Immunotherapy. *Adv Exp Med Biol*. 2017;1036:51–64.
10. Ai X, Mao F, Shen S, Shentu Y, Wang J, Lu S. Bexarotene inhibits the viability of non-small cell lung cancer cells via slc10a2/PPAR γ /PTEN/mTOR signaling pathway. *BMC Cancer*. 2018;18(1):407.
11. Aymeric L, Donnadieu F, Mulet C, du Merle L, Nigro G, Saffarian A, Bérard M, Poyart C, Robine S, Regnault B, et al: Colorectal cancer specific conditions promote *Streptococcus gallolyticus* gut colonization. 2018, 115(2):E283-e291.

12. Gui SL, Teng LC, Wang SQ, Liu S, Lin YL, Zhao XL, Liu L, Sui HY, Yang Y, Liang LC, et al: Overexpression of CXCL3 can enhance the oncogenic potential of prostate cancer. 2016, 48(5):701–709.
13. Hänggi K, Ruffell B. Oncogenic KRAS Drives Immune Suppression in Colorectal Cancer. *Int Urol Nephrol.* 2019;35(4):535–7.
14. Liao W, Overman MJ, Boutin AT, Shang X, Zhao D, Dey P, Li J, Wang G, Lan Z, Li J, et al. KRAS-IRF2 Axis Drives Immune Suppression and Immune Therapy Resistance in Colorectal Cancer. *Cancer cell.* 2019;35(4):559–72.e557.
15. See AL, Chong PK, Lu SY, Lim YP. CXCL3 is a potential target for breast cancer metastasis. *Curr Cancer Drug Targets.* 2014;14(3):294–309.
16. Xin H, Cao Y, Shao ML, Zhang W, Zhang CB, Wang JT, Liang LC, Shao WW, Qi YL, Li Y, et al: Chemokine CXCL3 mediates prostate cancer cells proliferation, migration and gene expression changes in an autocrine/paracrine fashion. 2018, 50(5):861–868.
17. Katayama Y, Oshima T, Sakamaki K, Aoyama T, Sato T, Masudo K, Shiozawa M, Yoshikawa T, Rino Y, Imada T, et al. Clinical Significance of INHBA Gene Expression in Patients with Gastric Cancer who Receive Curative Resection Followed by Adjuvant S-1 Chemotherapy. *Disease markers.* 2017;31(4):565–71.
18. Li X, Yang Z, Xu S, Wang Z, Jin P, Yang X, Zhang Z, Wang Y, Wei X, Fang T, et al: Targeting INHBA in Ovarian Cancer Cells Suppresses Cancer Xenograft Growth by Attenuating Stromal Fibroblast Activation. 2019, 2019:7275289.
19. Lyu S, Jiang C, Xu R, Huang Y, Yan S. INHBA upregulation correlates with poorer prognosis in patients with esophageal squamous cell carcinoma. *Cancer management research.* 2018;10:1585–96.
20. Peña C, Céspedes MV, Lindh MB, Kiflemariam S, Mezheyeuski A, Edqvist PH, Hägglöf C, Birgisson H, Bojmar L, Jirström K, et al. STC1 expression by cancer-associated fibroblasts drives metastasis of colorectal cancer. *Cancer research.* 2013;73(4):1287–97.
21. Rezapour S, Bahrami T, Hashemzadeh S, Estiar MA, Nemati M, Ravanbakhsh R, Feizi MA, Kafil HS, Pouladi N, Ghojzadeh M, et al. STC1 and NF- κ B p65 (Rel A) is Constitutively Activated in Colorectal Cancer. *Clinical laboratory.* 2016;62(3):463–9.
22. Signore M, Buccarelli M, Pillozzi E, De Luca G, Cappellari M, Fanciulli M, Goeman F, Melucci E, Biffoni M, Ricci-Vitiani L. UCN-01 enhances cytotoxicity of irinotecan in colorectal cancer stem-like cells by impairing DNA damage response. *Oncotarget.* 2016;7(28):44113–28.
23. Ma M, Li L, Chen H, Feng Y. Oxytocin Inhibition of Metastatic Colorectal Cancer by Suppressing the Expression of Fibroblast Activation Protein- α . *Front NeuroSci.* 2019;13:1317.

Figures

Figure 1

Flow chart of this study.

Figure 2

A. Heat map of genes expression; B. Volcano map of genes expression; C. Heat map of immune genes expression; D. Volcano map of immune genes expression; E. Heat map of transcription factors expression; F. Volcano map of transcription factors expression.

Figure 3

A. Forest plot of 12 immune genes associated with survival prognosis in patients with COAD. B. Regulatory network of immune genes and transcription factors related to the prognosis of COAD. The circles represent immune genes (Green represents a downward adjustment and red represents an upward adjustment.) and the triangles represent transcription factors.

Figure 4

A. Kaplan-Meier survival curve of high-risk group and low-risk group. B. ROC curve of survival prognosis model.

Figure 5

A. Risk curve of riskScore growth trend. B. Diagram of the relationship between riskScore and patient survival time. C. Construction of a heat map of immune genes in the riskScore model.

Figure 6

A. Forest plot of univariate Cox regression analysis between immune genes and clinical data constituting the riskScore model. B. Forest plot of multivariate Cox regression analysis between immune genes and clinical data that constitute the riskScore model.

Figure 7

Correlation analysis of genes (building a riskScore model) and clinical data. A. Correlation analysis between CXCL3 and M; B. Correlation analysis between CXCL3 and N; C. Correlation analysis between CXCL3 and COAD stage; D. Correlation analysis between OXTR and T; E. Correlation analysis between STC1 and T.

Figure 8

Correlation analysis between the expression of immune microenvironment cells and riskScore. A. Correlation analysis between B cells and riskScore; B. Correlation analysis between CD4 cells and riskScore; C. Correlation analysis between CD8 cells and riskScore; D. Correlation analysis between Dendritic cells and riskScore; E. Correlation analysis between Macrophage cells correlation and riskScore; F. Correlation analysis between Neutrophil cells and riskScore.

Figure 9

A. Comparison of immune cell differences between 10 groups of high RiskScore samples and 10 groups of low RiskScore samples, each color represents a different cell type. B. The expression of immune cells differs between 10 groups of high RiskScore samples and 10 groups of low RiskScore samples. C. Immune cell correlation matrix, the positive correlation is red, and the negative correlation is blue.