

A Risk Score Model Associated with Tumor Microenvironment Predicts Disease-Free Survival After Radical Prostatectomy

Hao Zhao

Zhengzhou University

Xuening Zhang

Zhengzhou University

Zhan Shi

Zhengzhou First People Hospital

Songhe Shi (✉ zzussh@163.com)

Zhengzhou University <https://orcid.org/0000-0002-2214-1440>

Primary research

Keywords: prostate cancer, radical prostatectomy, microenvironment, stromal, immune, prognosis, prediction

Posted Date: July 15th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-41768/v1>

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

[Read Full License](#)

Abstract

Background Tumor microenvironment (TME) and immune checkpoint inhibitors has been shown to promote active immune responses through different mechanisms. We aimed to identify the important prognostic genes and prognostic characteristics related to TME in prostate cancer (PCa).

Methods The gene transcriptome profiles and clinical information of PCa patients were obtained from the TCGA database, and the immune, stromal and estimate scores were calculated by the ESTIMATE algorithm. We evaluated the prognostic value of risk score (RS) model based on univariate Cox and LASSO Cox regression models analysis, and established a nomogram to predict disease-free survival (DFS) in PCa patients. The GSE70768 data set was used for external validation. Finally, 22 subsets of tumor-infiltrating immune cells (Tiics) were analyzed using the Cibersort algorithm.

Results In this study, the patients with higher immune, stromal, and estimate scores were associated with poorer DFS, higher Gleason score, and higher AJCC T stage. Based on the immune and stromal scores, the Venny diagram screened out 515 cross DEGs. The univariate COX and Lasso Cox regression models were used to select 18 DEGs from 515 DEGs, and constructed a RS model. The DFS of the high-RS group was significantly lower than that of the low-RS group ($P < 0.001$). The AUC of 1-year, 3-year and 5-year DFS rates in RS model were 0.778, 0.754 and 0.750, respectively. In addition, the RS model constructed from 18 genes was found to be more sensitive than Gleason score (1, 3, 5 year AUC= 0.704, 0.677 and 0.682). The nomograms of DFS were established based on RS and Gleason scores. The AUC of the nomograms in the first, third, and fifth years were 0.802, 0.808, and 0.796, respectively. These results have been further validated in GEO. In addition, the proportion of Tregs was higher in high-RS patients ($P < 0.05$), and the expression of five immune checkpoints (CTLA-4, PD-1, LAG-3, TIM-3 and TIGIT) was higher in high-RS patients ($P < 0.05$).

Conclusion We identified 18 TME-related genes from the TCGA database, which were significantly related to DFS in PCa patients.

Introduction

Prostate cancer (PCa) is the most common malignant tumor among men in Western developed countries [1]. According to the 2018 Annual Report of Cancer in China, the incidence of prostate cancer among men is second only to lung cancer, and the incidence is gradually increasing, with an annual growth rate of even 8.92% [2, 3]. With the progress of cancer treatment technology, localized PCa could be cured by radical prostatectomy (RP), and mortality is also significantly reduced. However, the recurrence rate of PCa after RP is still high, resulting in treatment failure. Previous studies have reported approximately 40% of patients relapse within five years after RP, and about 27%-53% of patients eventually develop local recurrence or distant metastasis within 10 years after RP [4, 5].

In recent years, immunotherapy with cytokines and immune checkpoint inhibitors has been shown to promote active immune responses through different mechanisms [6]. Therefore, a new classification of

PCa in combination with immunotherapy is needed to more accurately predict postoperative recurrence, thereby contributing to personalized treatment of clinical decision-making and reducing the recurrence rate of PCa patients.

Tumor microenvironment (TME) refers to the surrounding microenvironment of tumor cells, including immune cells, stromal cells, endothelial cells, inflammatory cells, and fibroblasts [7]. Among them, tumor infiltrating immune cells (Tiics) and stromal cells are two major non-tumor cell components, which have been considered important for the diagnosis and prognostic evaluation of cancer patients [8]. Therefore, understanding the cell composition and function of TME has great potential in effectively preventing cancer recurrence and immune response.

Bioinformatics analysis uses a combination of biological, statistical, computer science, and informatics techniques to process and analyze large amounts of complex biological data [9]. And the establishment of public databases, such as The Cancer Genome Atlas (TCGA) database and Gene Expression Omnibus (GEO) database, provide new data resources and technical means for TME research [10, 11]. Yoshihara et al [12] first proposed the ESTIMATE algorithm in 2013. This algorithm uses the unique properties of the transcription profile of cancer samples to infer infiltrating stromal/immune cells. According to reports, the researchers have explored the tumor characteristics and prognosis assessment of lung cancer [13], breast cancer [14], clear cell renal cell carcinoma [15] based on the ESTIMATE algorithm. However, the value of immune and stromal scores for PCa has not been verified.

In this study, we examined the TME in PCa, calculated the immune and stromal scores for each cancer sample, and established a PCa risk score prognostic model based on TCGA database. In addition, the model was validated using the GEO database. Finally, we also explored the relationship between high-RS and low-RS PCa patients and immune cell infiltration and immune checkpoints based on the Cibersort method, so as to provide some references for immunotherapy and postoperative management of PCa patients.

Methods

Data collection and processing

We obtained the Fragments per kilobase million (FPKM) data of RNA-Seq from TCGA-PRAD cohort (<https://portal.gdc.cancer.gov/>), including 499 PCa patients and 52 normal samples, and then the FPKM data was transferred to transcripts per million (TPM) expression data. The gene expression levels of duplicate samples were averaged and normal samples were deleted for subsequent analysis.

We used the GDC tool and cBioPortal website (<http://www.cbioportal.org/>) to download the corresponding clinical information, including age, pathological T stage, lymph node status, Gleason score, surgical margin status, tumor laterality, and prognostic information. We excluded samples with incomplete key clinical information, and finally included 480 PCa patients with complete transcriptome data and corresponding clinical information for follow-up analysis. We utilized the “limma” package for

normalization processing, and then immune, stromal and estimate scores were calculated using ESTIMATE algorithm. For further verification, PCa patients with gene expression information and clinical information were retrieved from Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). The inclusion criteria were as follows: (1) patients diagnosed with PCa; (2) RP was performed; (3) Gene levels in tissue samples were detected. The exclusion criteria are as follows: (1) clinical data without prognostic information, (2) data set with small sample size ($n < 50$). Finally, the eligible data set, GEO70768 ($n = 111$), was selected and the normalized expression matrix was used for subsequent analysis.

Correlation Analysis and Survival Analysis

For data satisfying the parameter test, T test was used for comparison between two groups, and ANOVA analysis was used for comparison of three groups and above. For the unsatisfactory parameter test data, Wilcoxon test was used for the comparison between the two groups, and Kruskal-Wallis analysis was used for the comparison of three groups and above. The relationship between immune scores or stromal scores and important clinical phenotypes was explored by comparing the differences of immune, stromal, and estimated scores in different clinical subgroups. Disease-free survival (DFS) was defined as the time from the date of diagnosis to the date of recurrence or death and the last follow-up, with DFS as the main prognostic endpoint.

According to the stromal and immune score of each PCa patient, the best cut-off value based on the R package "maxstat" (i.e., the maximum selective rank statistic method) [16] was used to divide the patients into high and low groups. Based on "survival" packages, the Kaplan-Meier (K-M) method and log rank test were used to compare the prognosis of the two groups.

Differentially expressed genes (DEGs) screening

The "limma" package in R software was used to screen for DEGs between high and low groups of immune scores and stromal scores. In this study, the adjusted P value < 0.05 and fold change ≥ 1.5 were regarded as the critical value for screening DEGs. The immune-related DEGs and stromal-related DEGs showing the same expression trend were selected for further analysis using Venny diagram. We used the "ggplot2" and "pheatmap" packages to generate volcano maps and heat maps.

DEGs functional enrichment analysis

The David online database (<http://david.ncifcrf.gov>) was used to explore the potential functions of DEGs. Gene ontology (GO) analysis included biological processes, molecular functions and cellular components, which demonstrated by barplot. Kyoto Encyclopedia of Genes and Genomes (KEGG) were performed to conduct the pathway analysis, which was illustrated by dotplot. With false discovery rate (FDR) < 0.05 as cut-off value, all enrichment results were visualized with "ggplot2" package.

Establishment of risk score (RS) and Survival Analysis

The univariate COX model was used to determine the relationship between TME-related DEGS expression and DFS. Then, the least absolute shrinkage and selection operator (LASSO) regression analysis was used to build an optimal risk signature with the minimum number of genes [17]. A set of genes and their coefficients were determined by minimum criteria, which involves selecting the best penalty parameter λ associated with the 10-fold cross validation [18]. The risk score (RS) was calculated as follows: $RS = \sum (\beta_i * Exp_i)$ ('i' = the number of prognostic hub genes, ' β_i ' represented the coefficient of each gene, and ' Exp_i ' represented gene expression.) In addition, PCa patients were divided into high-RS and low-RS groups according to the optimal cut-off value of the risk score. The receiver operating characteristic curve (ROC) and the consistency index (C-index) were then used to assess the predictive ability of the risk score. The K-M method and log rank test were used to analyze the difference in survival between the high-RS group and the low-RS group.

Validation of Risk Score model in the testing dataset

GSE70768 independent data set was used for verification. According to the RS calculation formula of the training data set, the samples in the test data set were divided into the high-RS group and the low-RS group. K-M survival analysis and ROC curve were used to evaluate the predictive ability of this model, and the predictive ability of this model was compared with Gleason score model.

Identification of Independent Prognostic Factors

The univariate and multivariate Cox analysis were used to study the independent prognostic value of RS and other clinical characteristics, and a clinical prediction model was established using a nomogram, and then the performance of the nomogram was evaluated by ROC and calibration curves.

Estimating the composition of immune cells

Cibersort is a deconvolution algorithm based on the principle of linear support vector regression to describe the infiltration of immune cells in the sample. LM22 is composed of 547 genes that accurately distinguish 22 human hematopoietic cell phenotypes, including seven T cell types, naïve and memory B cells, plasma cells, NK cells, and myeloid subsets [19]. We used Cibersort and LM22 to jointly estimate the scores of 22 human immune cell types in PCa samples from the TCGA cohort. For each sample, the sum of all estimated immune cell type scores was equal to 1. We compared differences in the composition of immune cell types between high and low RS groups.

Statistical analysis

Statistical analysis was performed using R software (version 3.6.1). All statistical tests were two sided and the P value < 0.05 represented as statistically significant.

Results

Immune scores and stromal scores were correlated with clinical features of PCa

The work flow chart of this study was shown in Supplementary material 1. A total of 480 male PCa patients were included in the TCGA database. Elderly PCa patients (≥ 65 years) accounted for 33.54%. There were 186 cases (38.75) of \leq pT2c stage patients, 153 cases (31.88%) of pT3a stage patients, 141 cases (29.38%) of \geq pT3b stage patients. Gleason score: 44 cases (9.17%) in the < 7 score group, 240 cases (50%) in the 7 score group, and 196 cases (40.83%) in the > 7 score group. For other detailed information, see Table 1 for clinical information.

Table 1
Clinical characteristics of 480 PCa patients included in study from TCGA cohort

Variables	Whole cohort (N = 480)	Relapse or death		Log-rank <i>P</i>
		No(N = 386)	Yes(N = 94)	
Age				0.462
< 65	319(66.46)	257(80.56)	62(19.44)	
≥ 65	161(33.54)	129(80.12)	32(19.88)	
Tumor laterality				0.212
Left/Right	56(11.67)	48(85.71)	8(14.29)	
Bilateral	424(88.33)	338(79.72)	86(20.28)	
Pathological T stage				< 0.001
\leq pT2c	186(38.75)	170(91.40)	16(8.60)	
pT3a	153(31.88)	120(78.43)	33(21.57)	
\geq pT3b	141(29.38)	96(68.09)	45(31.91)	
Lymph node status				0.007
pN0/pNx	402(83.75)	331(82.34)	71(17.66)	
pN1	78(16.25)	55(70.51)	23(29.49)	
Gleason score				< 0.001
<7	44(9.17)	43(97.73)	1(2.27)	
7	240(50.00)	214(89.17)	26(10.83)	
>7	196(40.83)	129(65.82)	67(34.18)	
Surgical margin status				< 0.001
Negative	333(69.38)	283(84.98)	50(15.02)	
Positive	147(30.63)	103(70.07)	44(29.93)	

The Immune, stromal and estimate scores of each sample were respectively calculated using an ESTIMATE algorithm. The immune score ranged from - 1404.50 to 2963.33, the stromal score ranged from - 1897.04 to 1762.53, and the estimate score ranged from - 3237.41 to 3584.35. The relationship between immune and stromal scores and clinical characteristics showed that higher immune scores were significantly associated with higher T stage ($P= 0.015$), positive surgical marginal status ($P= 0.038$) and lymph node metastasis ($P= 0.047$) (Fig. 1A). Higher stromal scores were significantly associated with higher T stage ($P< 0.001$), positive surgical margin status ($P= 0.006$) and higher Gleason score ($P= 0.002$) (Fig. 1B). The relationship between the estimate score and clinical characteristics was similar to the stromal score (Supplementary Fig. 1).

Immune score and stromal score were significantly related to PCa prognosis

The K-M survival curve of the relationship between immune and stromal scores and PCa patient prognosis showed that patients with lower immune and stromal scores had higher DFS rates ($P= 0.011$, $P= 0.037$) (Fig. 1D and 1E). Consistently, PCa patients with low-estimate scores also had higher DFS rates than patients with high-estimate scores ($P= 0.02$) (Fig. 1F). These observations consistently suggested that patients with low immune or stromal scores had a more favorable outcome.

Identification of DEGs based on immune score and stromal score in PCa

In order to explore the DEGs closely related to the tumor microenvironment, the limma package was used to process the Affymetrix microarray data from 480 PCa patients. Figure 2A showed a heat map of 804 DEGs between high and low immune scores, and Fig. 2B showed a heat map of 1098 DEGs between high and low stromal scores.

In addition, the volcano plot showed DEGs based on immune score and stromal score (Supplementary Fig. 2). For the immune score, there were 68 up-regulated DEGs and 736 down-regulated DEGs in the high group compared with the low group. For stromal score, compared with the low score group, there were 104 up-regulated DEGs and 994 down-regulated DEGs in the high score group. Venny diagram showed 41 cross-up-regulated DEGs and 474 cross-down-regulated DEGs between the immune and stroma groups (Fig. 2C).

Function and pathway enrichment analysis of DEGs

Functional enrichment analyses for DEGs, including biological processes (BP), cellular component (CC), molecular function (MF) and KEGG pathways were conducted using the David gene annotation tool. BP indicated that these genes may be associated with immune response, inflammatory response, cell adhesion, extracellular matrix organization, and leukocyte migration. CC indicated that these genes may be associated with plasma membrane and extracellular exosome, region, and space. MF indicated that these genes may be associated with calcium ion binding, receptor activity, and serine-type endopeptidase

activity. The results of KEGG enrichment were related to immune response, including phagosome, infection, cytokine receptor interaction, and cell inflammatory molecules (CAMs) (Fig. 2D). Overall, our results confirmed that TME-related DEGs activated the immune response in patients with PCa and promoted tumor progression.

Risk Score (RS) calculation and Survival Analysis

In order to explore the potential role of DEGs in DFS, a univariate COX proportional hazards regression model was first conducted, and the results showed that 172 prognostic genes were selected by univariate analysis. Then, the 10-fold cross-validation random sampling method was used, according to the -2 log-likelihood test, through repeated calculation and verification, the model was optimized at the penalty parameter $\log \lambda = -1.59$, and constructed a risk score (RS) model of 18 genes (Fig. 3A, 3B and Supplementary Fig. 3A). $RS = (-0.11366 * APOF) + (0.24979 * C1QC) + (-0.00804 * CHRNA2) + (-0.42763 * CLIC6) + (-0.01610 * COL1A1) + (-0.11550 * EGR1) + (-0.07356 * FEV) + (0.01102 * FOS) + (-0.05767 * GJB1) + (-0.70028 * GNG2) + (-0.02392 * HOPX) + (-0.08264 * HSD11B1) + (0.03916 * ITGAX) + (-0.05485 * OLFML3) + (-0.23884 * PLTP) + (0.57912 * STAB1) + (0.60611 * TGFB1) + (-0.09685 * TGM3)$

In addition, survival curves of 18 DEGs were constructed to explore the prognostic value of each gene (Fig. 4). The results of K-M curve showed that, the prognosis was better when the expression levels of C1QC, COL1A1, HOPX, ITGAX, STAB1, and TGFB1 were lower. In contrast, we found that low expression of other hub genes was associated with poor prognosis in PCa patients.

A total of 353 patients with risk scores below the critical value (-0.17539) were classified as "low-RS group", and the remaining 127 patients were classified as "high-RS group" according to the optimal cut-off value of the RS (Supplementary Fig. 3B). Kaplan-Meier curve showed that the DFS of high-RS patients was significantly worse ($P < 0.001$) (Fig. 3C). At the same time, for sensitivity verification, RS was grouped by the median, and the result was consistent (Supplementary Fig. 3C).

To test whether RS can predict DFS without considering tumor residual disease, we performed a risk-stratified analysis. The results showed that in the absence of tumor residuals, patients with low-RS had significantly longer DFS than patients with high-RS ($P < 0.001$) (Supplementary Fig. 3D)

In order to evaluate the predictive ability of the RS model, we drew the ROC curve based on the RS and calculated the AUC of the area under the curve. The AUC of the first, third and fifth years DFS prognostic models were 0.790, 0.806 and 0.795 respectively (Fig. 3D), indicating that the RS model had good predictive accuracy.

To further compare the accuracy of RS model, a prediction model based on Gleason score was established. The AUC of the first, third and fifth years of the Gleason score model were 0.704, 0.677 and 0.682 respectively (Supplementary Fig. 3E), and we found that the RS model was more accurate in predicting prognosis than Gleason score model.

Validation of the RS model

In order to verify the generalization value of the RS model based on the TCGA cohort, we calculated the risk score of each sample for the 111 PCa patients in GSE70768 using the above RS formula. The K-M survival curve indicated that the low-RS group had higher DFS ($P=0.007$) (Fig. 3E). In addition, the ROC curves based on RS model showed that the AUC for the first and third year of DFS prognostic models were 0.820 and 0.604, respectively (Fig. 3F). The ROC curves based on Gleason score model showed that the AUC of the first and third DFS prognostic models were 0.677 and 0.599, respectively (Supplementary Fig. 3F). The result of validation set also indicated that the RS model had better robustness and was superior to Gleason score model.

Identification of Independent Prognostic Factors

The univariate COX model analysis showed that higher pathological T stage, Gleason score and RS, lymph node metastasis, and positive surgical margin status were risk factors affecting prognosis. In the multivariate Cox analysis, the meaningful variables of the univariate COX model analysis were included, and the results showed that after the adjustment of variables including variables like lymph T stage, Gleason score, surgical margin Status, RS was independent predictor (HR: 4.04, 95%CI: 0.25–2.44), similar to and independent of Gleason score (Table 2).

Table 2
Univariate and multivariate Cox analysis of clinical information and RS

Variables	Univariate analysis		Multivariate analysis	
	HR(95%CI)	P	HR(95%CI)	P
Age				
< 65	1			
≥ 65	0.46(0.77–1.8)	0.462		
Tumor laterality				
Left/Right	1			
Bilateral	1.58(0.77–3.27)	0.215		
Pathological T stage				
≤pT2c	1		1	
pT3a	1.8(1.43–4.77)	0.002	1.49(0.79–2.8)	0.222
≥pT3b	4.59(2.57–8.2)	< 0.001	1.16(0.56–2.37)	0.691
Lymph node status				
pN0/pNx	1		1	
pN1	1.89(1.18–3.03)	0.009	0.65(1.55 – 0.39)	0.098
Gleason score				
<7	1		1	
7	4.73(0.64–34.89)	0.127	3.88(0.26–0.52)	0.185
>7	17.34(2.41–125)	0.005	8.5(0.12–1.13)	0.038
Surgical margin status				
Negative	1		1	
Positive	2.34(1.56–3.52)	< 0.001	1.51(0.66–0.97)	0.066
Risk score				
Low	1		1	
High	5.9(3.88–8.97)	< 0.001	4.04(0.25–2.44)	< 0.001

The nomogram of the predicted DFS was further established based on RS and Gleason score (Fig. 5A). The calibration curve of the nomogram showed good consistency between actual observations and predictions (Fig. 5C-5E). The AUC of the first, third and fifth years of the nomogram were 0.802, 0.808 and

0.796, respectively (Fig. 5B). Further validation of an independent cohort of 111 PCa patients also showed good predictive power (Fig. 5F).

Estimating the composition of immune cells.

We used Cibersort to estimate the immune cell composition of 480 samples and to quantify the relative levels of different cell types in the mixed cell population. As shown in the Fig. 6, we compared different cell types of patients in the low-RS group with those in the high-RS group. The results indicated that the expression levels of T cells CD4 memory resting, T cells CD8, Macrophage M0, Macrophage M1, Plasma cells, eosinophils, monocytes, B cells naïve, and Mast cells resting in the low-risk group were significantly higher than those in the high-risk group ($p < 0.05$). In contrast, the expression levels of Macrophage M2, T cells regulatory (Tregs), Dendritic cells resting and Neutrophils in the high-RS group were significantly higher than those in the low-RS group ($P < 0.05$).

We also explored the expression of immune checkpoints between high-RS and low-RS PCa patients. The expression levels of CTLA-4, PD-1, LAG-3, TIM-3 and TIGIT in high-RS patients were significantly higher than those in low-RS patients ($P < 0.05$) Fig. 7.

Discussion

Prostate cancer (PCa) is the most common cancer among men in developed countries. Although with the development of treatment methods, including chemotherapy, radiotherapy, hormones and surgical resection, the recurrence rate of PCa patients is still high [20]. Previous studies have shown that TME plays a vital role in the development, progression and recurrence of cancer [21]. Therefore, it is important to study the TME of PCa in this study to determine biomarkers that can predict DFS of PCa patients after RP.

In order to study the TME of PCa, we calculated the immune score, stromal score and estimate score of each PCa sample extracted from the TCGA database by applying an ESTIMATE algorithm. The results showed that higher immune, stromal, and estimate scores were associated with poorer DFS, higher Gleason score, and higher AJCC T stage in PCa patients with malignant tumors. Subsequently, we divided PCa patients into high/low immune score groups and high/low stromal score groups, and identified 515 cross-sectional DEGs. The GO and KEGG analyses of DEGs showed that DEGs mainly participated in TME, such as immune response, inflammatory response, cell adhesion, extracellular matrix organization, and leukocyte migration. These processes may inhibit tumor progression and metastasis, thereby improving DFS. We found that these DEGs have a strong correlation with the immune response and tumor immune microenvironment. In addition, we applied univariate COX and Lasso-Cox regression model to construct a RS model based on 18 DEGs which screened from 515 DEGs. In this model, DFS in the high-RS group was significantly lower than that in the low-risk group, so recurrence in PCa patients could be well predicted. The AUC of 1-year, 3-year and 5-year survival rates in the RS model were 0.778, 0.754 and 0.75, respectively. In addition, we found that the RS model constructed with 18 genes was

more sensitive to prognosis than Gleason score. The stratified analysis showed that the RS model also had strong prognostic capability for PCa patients with negative surgical margins (R0).

Among them, the expression levels of C1QC, COL1A1, HOPX, ITGAX, STAB1, TGFB1 were low, and the prognosis was good. On the contrary, we found that the low expression levels of other genes were related to the poor prognosis of PCa patients.

C1QC belongs to C1Q and plays an important role in adaptive and innate immune responses. Studies have shown that C1QC can promote the adhesion, migration and proliferation of malignant pleural mesothelioma [22], and the increase in C1QC levels in patients with sarcoma is associated with poor prognosis [23]. The high expression of COL1A1 gene will cause unrestricted growth factors, which in turn will benefit tumor proliferation [24]. The HOPX gene may be involved in the malignant transformation of cancer cells. Studies have shown that higher HOPX expression is an independent adverse prognostic factor for acute myeloid leukemia [25]. ITGAX expression level is positively correlated with aggressive prostate cancer [26]. STAB1 is an identified oncogene whose increased expression promotes tumorigenesis and tumor progression [27], and it is associated with poor prognosis in many cancers. TGFB1 is often up-regulated in tumor cells and highly secreted into the prostate environment, partially mediating the immunosuppressive effect on NK cells and promoting the invasion and metastasis of prostate cancer [28].

APOF can act as a tumor suppressor for hepatocellular carcinoma, and the decreased expression of APOF is associated with poor prognosis [29]. The genetic variation of the nicotinic cholinergic receptor gene (CHRNS) may affect the risk of lung cancer [30]. The low expression of CLIC6 in breast cancer is related to high histological grade [31]. The tumor suppressor gene EGR-1 can directly mediate the apoptotic function through the transcriptional upregulation of Bax-mRNA and protein and the increase of oligomerization and activation [32]. FEV is rich in alanine c-terminal, indicating that it may act as a transcriptional inhibitor [33]. FOS is considered as a regulator of cell proliferation, differentiation, and transformation and participates in MAPK signaling pathway [34, 35]. GJB1 is abundantly expressed in other well-differentiated cell types such as prostate and pancreas. In prostate tumors, the ability to assemble GJS from GJB1 and GJA1 is impaired [36]. GNG2 is involved in the signal transduction of GPCR and CCR5 pathway in macrophages, and the expression level of GNG2 in malignant melanoma is decreases [37, 38]. The protein encoded by the HSD11b1 gene is a microsomal enzyme involved in the synthesis and regulation of prostaglandins. Up-regulation of OLFML3 enhances self-renewal of glioma stem cells and triggers primary tumor immunity PLTP plays a crucial role in mediating the association between triacyl lipid A and lipoprotein, which is beneficial to the anticancer properties[39]. As a candidate cancer suppressor, low TGM3 expression is associated with poor overall survival rate of neck cancer [40].

In conclusion, the RS prognostic model constructed by these 18 genes has not been reported and may be a new prognostic factor for PCa. Furthermore, multivariate COX model showed that RS and Gleason scores were two independent prognostic indicators. Based on RS and Gleason score, the nomogram for the prediction of DFS rate was established. The calibration diagram and AUC for the probability of

recurrence at 1, 3 and 5 years showed that the prediction was consistent with the actual observation, and the prediction ability was strong. The main purpose of the nomogram is to quantify the risk of clinical events based on various predictive factors, provide personalized scores for each patient, and provide new ideas for personalized treatment of PCa patients.

Differences in the composition of T cells subsets and immune checkpoints between high-RS and low-RS PCa patients were also explored. The concentration of M2 macrophages and Tregs was higher in the high-RS group. In contrast, the low-RS group had a higher proportion of CD8 T cells, resting memory CD4 T cells, and monocytes. The expressions levels of CTLA-4, PD-1, LAG-3, TIM-3 and TIGIT in high-RS patients were significantly higher than those in low-RS patients ($P < 0.05$). Previous studies have shown that resting memory CD4 T cells can further differentiate and have multiple functions, including restoring immune tolerance to autoantigens or heteroantigens and promoting CD8 T cells against tumors [41, 42]. Tregs expressing CTLA-4 play a crucial role in the maintenance of immunological self-tolerance and homeostasis and suppressing the anti-tumor immune response[43]. CTLA-4 is expressed in activated CD + 4 and CD + 8 T cells, which can terminate the response of activated T cells and mediate the inhibitory function of Tregs [44]. Overexpression of PD-1 on CD8 + T cells is an indicator of T cell depletion [45]. Inhibiting or knocking out LAG-3 will release the inhibitory function of Tregs on T cells. TIM-3 suppresses anti-tumor immunity by mediating T cell depletion. TIGIT can suppress immune cells in multiple steps of the tumor immune cycle [46]. In our study, the proportion of Tregs in high-RS patients was higher, the expression of immune checkpoint CTLA-4, PD-1, LAG-3, TIM-3 and TIGIT was higher, and the prognosis was poor, suggesting that the immunosuppressive environment and the high expression of immune checkpoint may be the reason for the poor prognosis of PCa. In addition, these results suggested that anti-CTLA4 drugs blocking immune checkpoint leads to T-cell activation, which is an ideal strategy for treating cancer. Anti-immune checkpoint antibody treatment will be more beneficial to high-risk PCa patients than low-risk patients, resulting in a better prognosis.

However, this study also has certain limitations. First, this study only conducted bioinformatics research on public databases, and the follow-up time of the validation set has not exceeded five years. Next, we should verify the results of this study through clinical patients in the prospective design. Secondly, 18 hub genes related to immune cell infiltration should be further studied to clarify the regulatory mechanism of PCa immune infiltration.

Conclusion

In summary, our study established and validated a model of risk score based on 18 immune-related genes, which provided a theoretical basis for predicting DFS of PCa, and further revealed the TME-related features associated with tumor immunity infiltration abundance of potential, these may be of great significance for the individualization and immunotherapy of PCa patients and postoperative rehabilitation management.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

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

Competing interests

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.

Funding

This research was funded by the National Key Research and Development Programme of China (Grant Nos 2017YFC1307705 and 2016YFC0106907), the Science and Technology Development Programme of Henan (Grant Number: 201403007). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contribution

H.Z., and X.Z., designed the study. H.Z., X.Z. and Z.S. participated in the data collection and analysis. H.Z. and X.Z. drafted this manuscript. H.Z., X.Z. and S.S. interpreted the data. S.S. reviewed and revised this manuscript. All authors have approved the final manuscript.

ACKNOWLEDGMENTS

We are grateful for the joint efforts of the members of the research group. We also extend our thanks all participants of this research.

References

1. Litwin MS, Tan HJ: The Diagnosis and Treatment of Prostate Cancer: A Review. JAMA. 2017; 317(24):2532-2542.

2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018; 68(6):394-424.
3. Ye DW, Zhu Y: Prostate cancer and prostatic diseases Best of China, 2018. *Prostate Cancer Prostatic Dis.* 2019; 22(1):1-2.
4. Bott SR: Management of recurrent disease after radical prostatectomy. *Prostate Cancer Prostatic Dis.* 2004; 7(3):211-216.
5. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, Davis M, Peters TJ, Turner EL, Martin RM *et al*: 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. *N Engl J Med.* 2016; 375(15):1415-1424.
6. Lu X, Horner JW, Paul E, Shang X, Troncoso P, Deng P, Jiang S, Chang Q, Spring DJ, Sharma P *et al*: Effective combinatorial immunotherapy for castration-resistant prostate cancer. *Nature.* 2017; 543(7647):728-732.
7. Neal JT, Li X, Zhu J, Giangarra V, Grzeskowiak CL, Ju J, Liu IH, Chiou SH, Salahudeen AA, Smith AR *et al*: Organoid Modeling of the Tumor Immune Microenvironment. *Cell.* 2018; 175(7):1972-1988 e1916.
8. Ni J, Wu Y, Qi F, Li X, Yu S, Liu S, Feng J, Zheng Y: Screening the Cancer Genome Atlas Database for Genes of Prognostic Value in Acute Myeloid Leukemia. *Front Oncol.* 2019; 9:1509.
9. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD *et al*: The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med.* 2015; 21(8):938-945.
10. Cancer Genome Atlas Research N: The Molecular Taxonomy of Primary Prostate Cancer. *Cell.* 2015; 163(4):1011-1025.
11. Coordinators NR: Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2016; 44(D1):D7-19.
12. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA *et al*: Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013; 4:2612.
13. Li J, Li X, Zhang C, Zhang C, Wang H: A signature of tumor immune microenvironment genes associated with the prognosis of nonsmall cell lung cancer. *Oncol Rep.* 2020; 43(3):795-806.
14. Li B, Geng R, Wu Q, Yang Q, Sun S, Zhu S, Xu Z, Sun S: Alterations in Immune-Related Genes as Potential Marker of Prognosis in Breast Cancer. *Front Oncol.* 2020; 10:333.
15. Luo J, Xie Y, Zheng Y, Wang C, Qi F, Hu J, Xu Y: Comprehensive insights on pivotal prognostic signature involved in clear cell renal cell carcinoma microenvironment using the ESTIMATE algorithm. *Cancer Med.* 2020.
16. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015; 43(7):e47.

17. Bovelstad HM, Nygard S, Storvold HL, Aldrin M, Borgan O, Frigessi A, Lingjaerde OC: Predicting survival from microarray data—a comparative study. *Bioinformatics*. 2007; 23(16):2080-2087.
18. Chai RC, Li YM, Zhang KN, Chang YZ, Liu YQ, Zhao Z, Wang ZL, Chang YH, Li GZ, Wang KY *et al*: RNA processing genes characterize RNA splicing and further stratify lower-grade glioma. *JCI Insight*. 2019; 5.
19. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12(5):453-457.
20. Nguyen PL, Ma J, Chavarro JE, Freedman ML, Lis R, Fedele G, Fiore C, Qiu W, Fiorentino M, Finn S *et al*: Fatty acid synthase polymorphisms, tumor expression, body mass index, prostate cancer risk, and survival. *J Clin Oncol*. 2010; 28(25):3958-3964.
21. Wang H, Wu X, Chen Y: Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer. *Front Oncol*. 2019; 9:1212.
22. Zhang J, Chen M, Zhao Y, Xiong H, Sneh T, Fan Y, Wang J, Zhou X, Gong C: Complement and coagulation cascades pathway correlates with chemosensitivity and overall survival in patients with soft tissue sarcoma. *Eur J Pharmacol*. 2020; 879:173121.
23. Mangogna A, Belmonte B, Agostinis C, Zacchi P, Iacopino DG, Martorana A, Rodolico V, Bonazza D, Zanconati F, Kishore U *et al*: Prognostic Implications of the Complement Protein C1q in Gliomas. *Front Immunol*. 2019; 10:2366.
24. Lichner Z, Ding Q, Samaan S, Saleh C, Nasser A, Al-Haddad S, Samuel JN, Fleshner NE, Stephan C, Jung K *et al*: miRNAs dysregulated in association with Gleason grade regulate extracellular matrix, cytoskeleton and androgen receptor pathways. *J Pathol*. 2015; 237(2):226-237.
25. Lin CC, Hsu YC, Li YH, Kuo YY, Hou HA, Lan KH, Chen TC, Tzeng YS, Kuo YY, Kao CJ *et al*: Higher HOPX expression is associated with distinct clinical and biological features and predicts poor prognosis in de novo acute myeloid leukemia. *Haematologica*. 2017; 102(6):1044-1053.
26. Williams KA, Lee M, Hu Y, Andreas J, Patel SJ, Zhang S, Chines P, Elkahloun A, Chandrasekharappa S, Gutkind JS *et al*: A systems genetics approach identifies CXCL14, ITGAX, and LPCAT2 as novel aggressive prostate cancer susceptibility genes. *PLoS Genet*. 2014; 10(11):e1004809.
27. Lin SY, Hu FF, Miao YR, Hu H, Lei Q, Zhang Q, Li Q, Wang H, Chen Z, Guo AY: Identification of STAB1 in Multiple Datasets as a Prognostic Factor for Cytogenetically Normal AML: Mechanism and Drug Indications. *Mol Ther Nucleic Acids*. 2019; 18:476-484.
28. Gao F, Alwhaibi A, Sabbineni H, Verma A, Eldahshan W, Somanath PR: Suppression of Akt1-beta-catenin pathway in advanced prostate cancer promotes TGFbeta1-mediated epithelial to mesenchymal transition and metastasis. *Cancer Lett*. 2017; 402:177-189.
29. Wang YB, Zhou BX, Ling YB, Xiong ZY, Li RX, Zhong YS, Xu MX, Lu Y, Liang H, Chen GH *et al*: Decreased expression of ApoF associates with poor prognosis in human hepatocellular carcinoma. *Gastroenterol Rep (Oxf)*. 2019; 7(5):354-360.
30. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, Caporaso NE, Johansson M, Xiao X, Li Y *et al*: Large-scale association analysis identifies new lung cancer susceptibility loci and

- heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet.* 2017; 49(7):1126-1132.
31. Ko JH, Ko EA, Gu W, Lim I, Bang H, Zhou T: Expression profiling of ion channel genes predicts clinical outcome in breast cancer. *Mol Cancer.* 2013; 12(1):106.
 32. Zagurovskaya M, Shareef MM, Das A, Reeves A, Gupta S, Sudol M, Bedford MT, Prichard J, Mohiuddin M, Ahmed MM: EGR-1 forms a complex with YAP-1 and upregulates Bax expression in irradiated prostate carcinoma cells. *Oncogene.* 2009; 28(8):1121-1131.
 33. Renzi S, Anderson ND, Light N, Gupta A: Ewing-like sarcoma: An emerging family of round cell sarcomas. *J Cell Physiol.* 2019; 234(6):7999-8007.
 34. Sharma NV, Pellegrini KL, Ouellet V, Giuste FO, Ramalingam S, Watanabe K, Adam-Granger E, Fossouo L, You S, Freeman MR *et al*: Identification of the Transcription Factor Relationships Associated with Androgen Deprivation Therapy Response and Metastatic Progression in Prostate Cancer. *Cancers (Basel).* 2018; 10(10).
 35. Schlomm T, Hellwinkel OJ, Bunes A, Ruschhaupt M, Lubke AM, Chun FK, Simon R, Budaus L, Erbersdobler A, Graefen M *et al*: Molecular cancer phenotype in normal prostate tissue. *Eur Urol.* 2009; 55(4):885-890.
 36. Ray A, Katoch P, Jain N, Mehta PP: Dileucine-like motifs in the C-terminal tail of connexin32 control its endocytosis and assembly into gap junctions. *J Cell Sci.* 2018; 131(5).
 37. Yajima I, Kumasaka MY, Yamanoshita O, Zou C, Li X, Ohgami N, Kato M: GNG2 inhibits invasion of human malignant melanoma cells with decreased FAK activity. *Am J Cancer Res.* 2014; 4(2):182-188.
 38. Yajima I, Kumasaka MY, Naito Y, Yoshikawa T, Takahashi H, Funasaka Y, Suzuki T, Kato M: Reduced GNG2 expression levels in mouse malignant melanomas and human melanoma cell lines. *Am J Cancer Res.* 2012; 2(3):322-329.
 39. Chen P, Hsu WH, Chang A, Tan Z, Lan Z, Zhou A, Spring DJ, Lang FF, Wang YA, DePinho RA: Circadian Regulator CLOCK Recruits Immune-Suppressive Microglia into the GBM Tumor Microenvironment. *Cancer Discov.* 2020; 10(3):371-381.
 40. Wu X, Cao W, Wang X, Zhang J, Lv Z, Qin X, Wu Y, Chen W: TGM3, a candidate tumor suppressor gene, contributes to human head and neck cancer. *Mol Cancer.* 2013; 12(1):151.
 41. Rosenberg J, Huang J: CD8(+) T Cells and NK Cells: Parallel and Complementary Soldiers of Immunotherapy. *Curr Opin Chem Eng.* 2018; 19:9-20.
 42. Crouse J, Xu HC, Lang PA, Oxenius A: NK cells regulating T cell responses: mechanisms and outcome. *Trends Immunol.* 2015; 36(1):49-58.
 43. Sakaguchi S, Mikami N, Wing JB, Tanaka A, Ichiyama K, Ohkura N: Regulatory T Cells and Human Disease. *Annu Rev Immunol.* 2020; 38:541-566.
 44. Rowshanravan B, Halliday N, Sansom DM: CTLA-4: a moving target in immunotherapy. *Blood.* 2018; 131(1):58-67.

45. Wherry EJ, Kurachi M: Molecular and cellular insights into T cell exhaustion. Nat Rev Immunol. 2015; 15(8):486-499.
46. Anderson AC, Joller N, Kuchroo VK: Lag-3, Tim-3, and TIGIT: Co-inhibitory Receptors with Specialized Functions in Immune Regulation. Immunity. 2016; 44(5):989-1004.

Figures

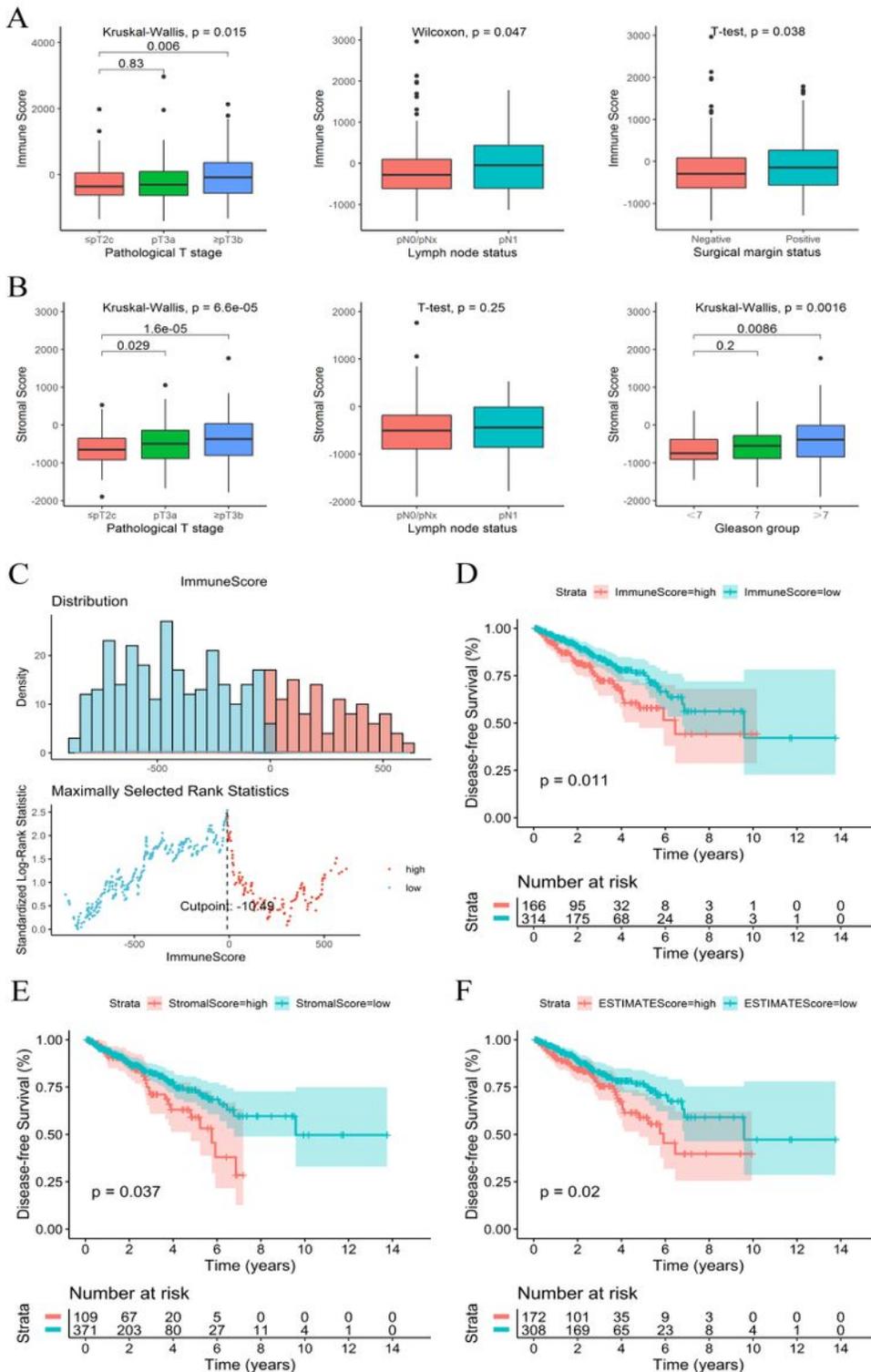


Figure 1

Immune scores and stromal scores were correlated with clinical features and prognosis of PCa. (A) Immune scores correlated with clinical features. (B) Stromal scores correlated with clinical features. (C) The best cutoff value for immunity score. (D-F) The K-M survival curve of the relationship between immune, stromal and estimate scores and PCa patient prognosis.

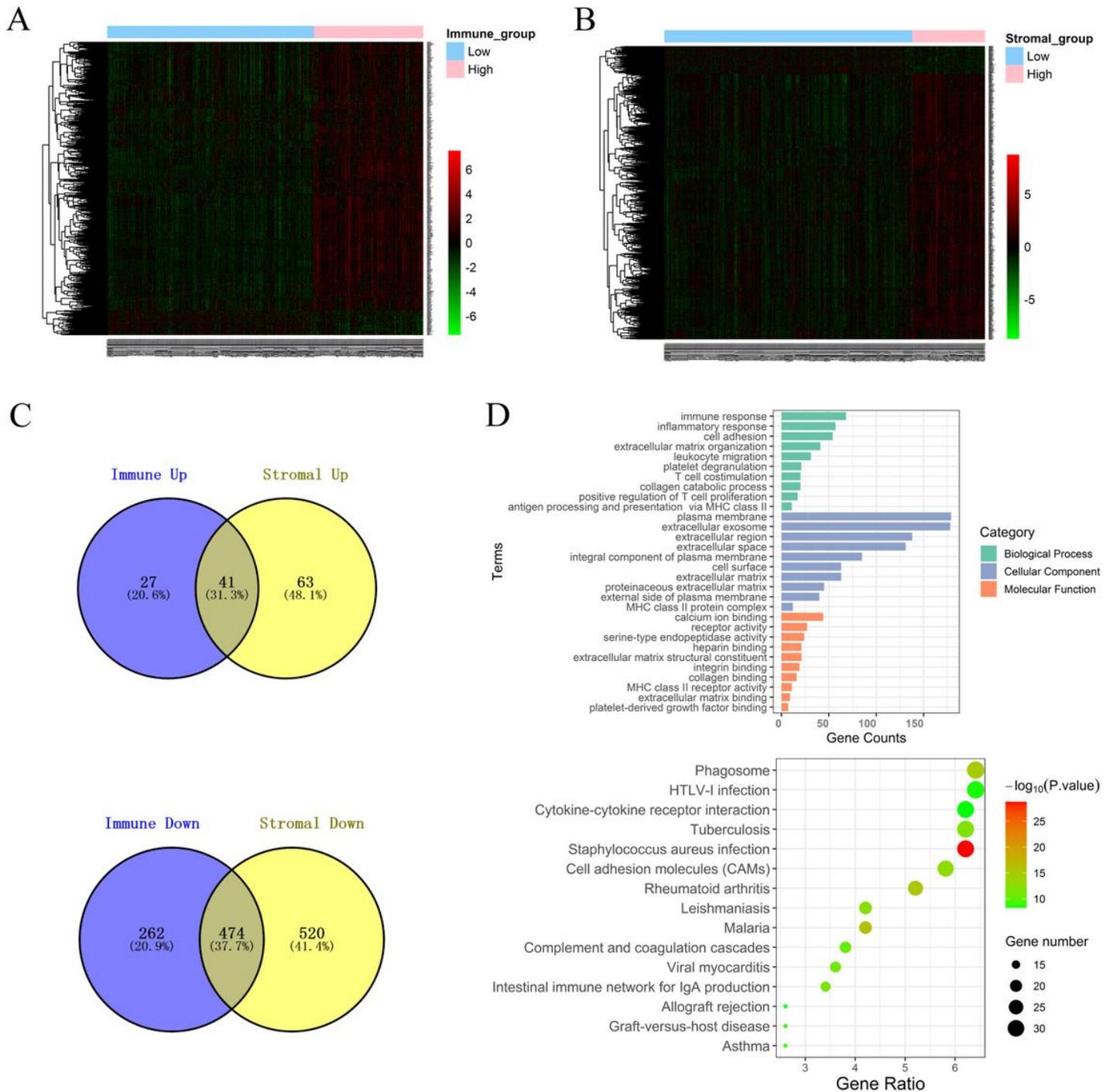


Figure 2

Identification of DEGs and Function and pathway enrichment analysis (A) A heat map of 804 DEGs between patients with high or low immune scores. (B) A heat map of 1098 DEGs between patients with high or low stromal scores. (C) Cross-up-regulated and Cross-down-regulated DEGs between the immune and stroma groups. (D) Function and pathway enrichment analysis of DEGs by GO and KEGG.

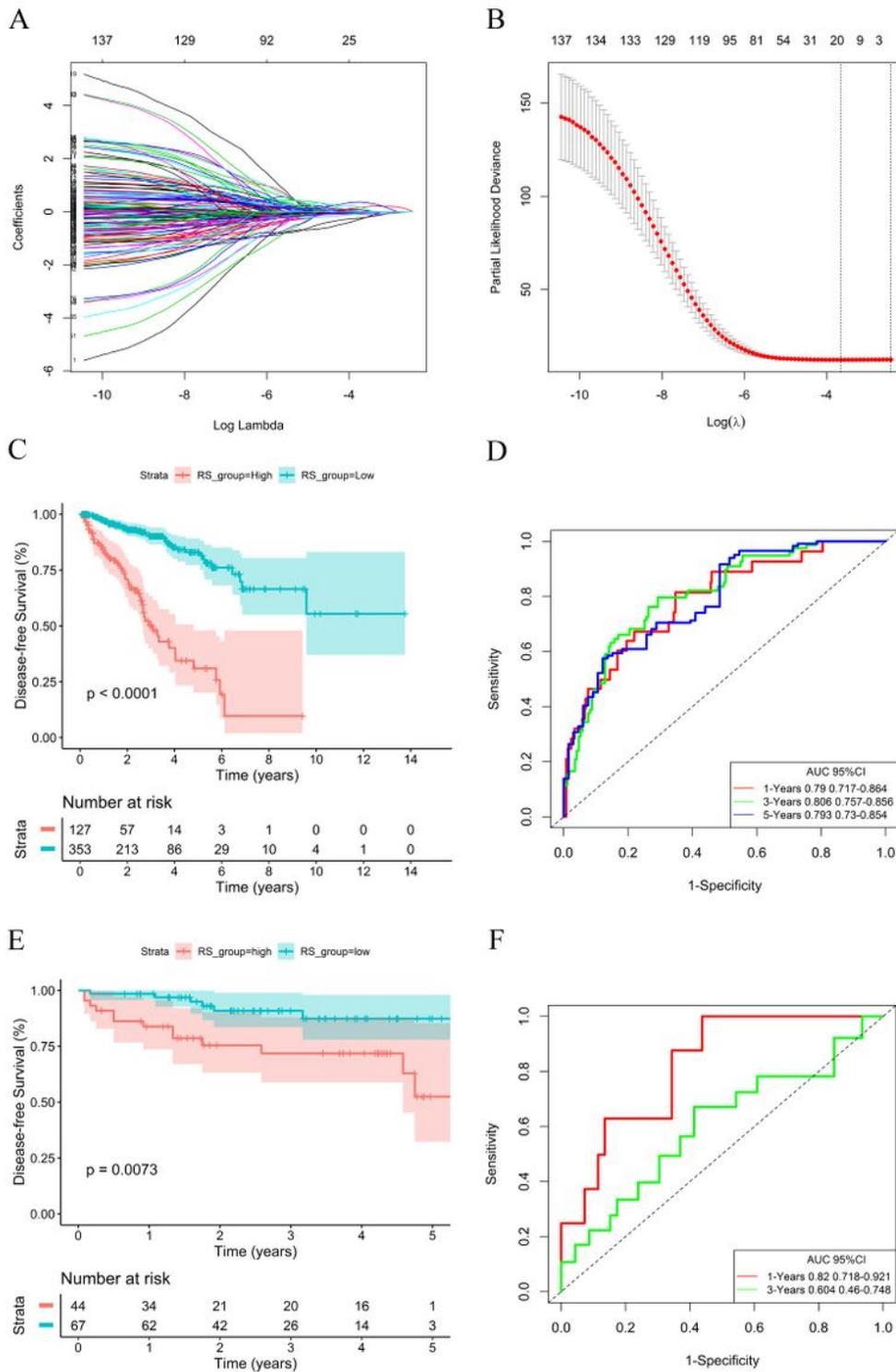


Figure 3

The RS model calculation and survival analysis. (A) LASSO coefficient profiles. (B) 10-fold cross-validations result which identified optimal values of the penalty parameter λ . (C) Kaplan-Meier curve of high and low RS groups in TCGA database (D) The ROC curve based on the RS model in TCGA database. (E) Kaplan-Meier curve of high and low RS groups in GEO database (F) The ROC curve based on the RS model in GEO database.

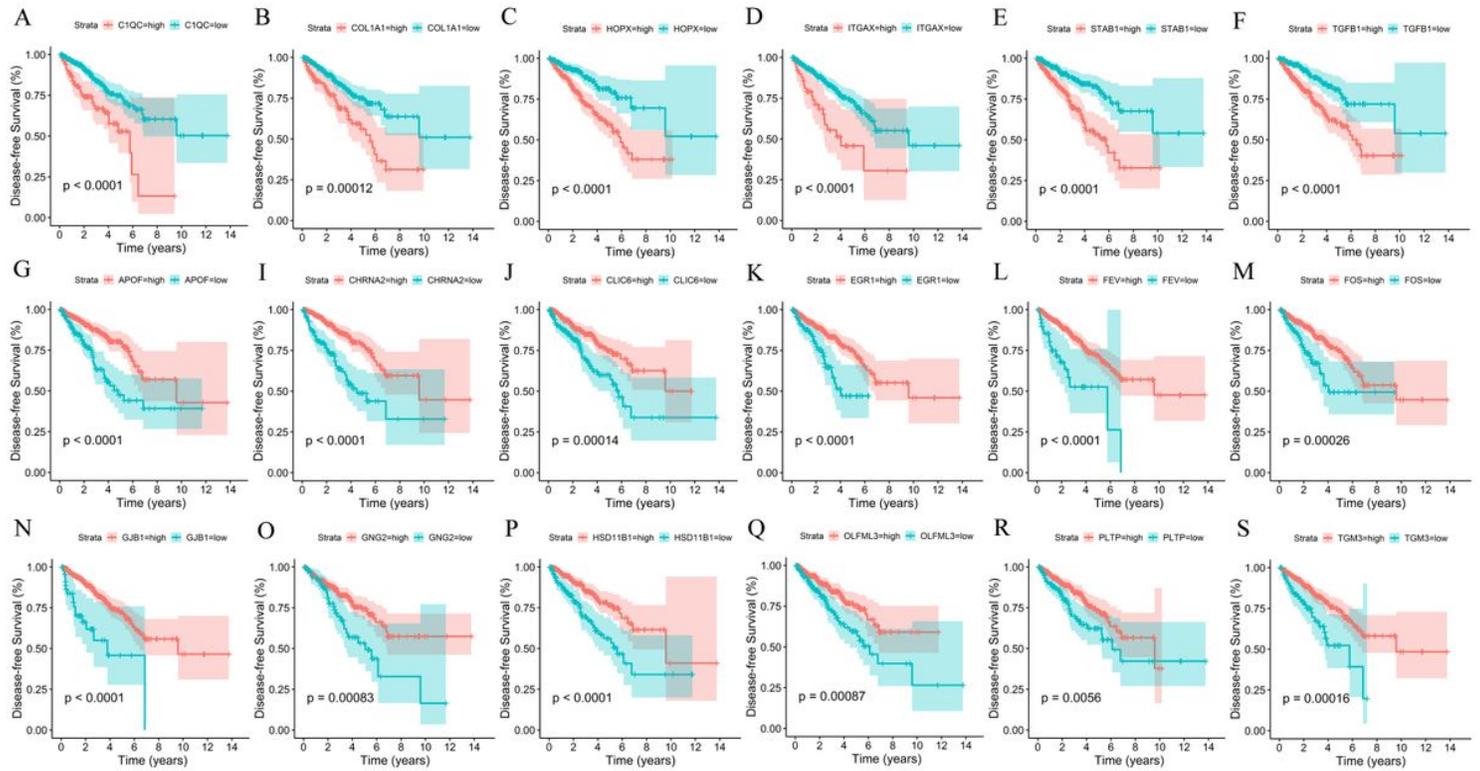


Figure 4

Survival curves of 18 prognostic hub genes.

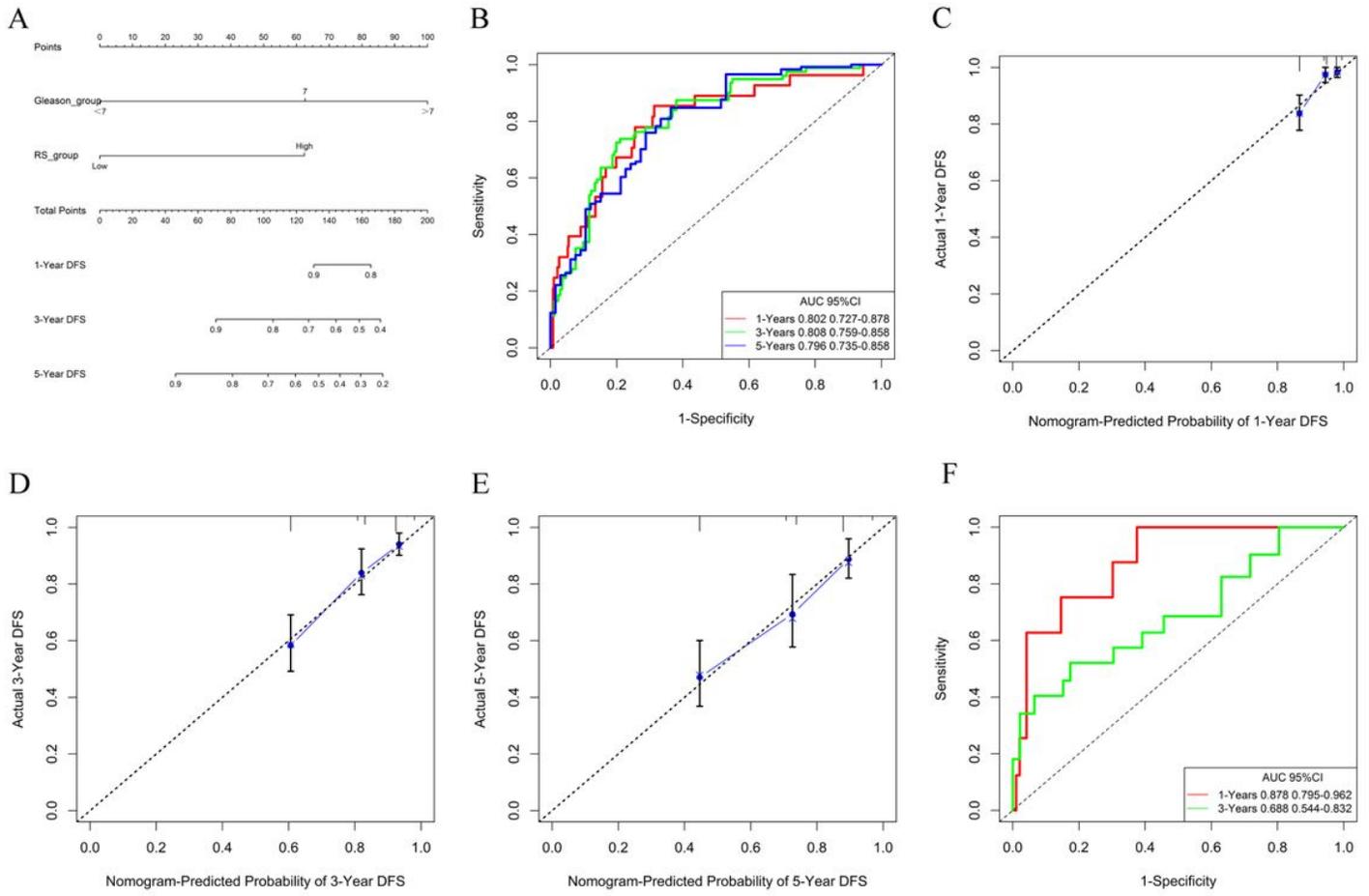


Figure 5

Identification of Independent Prognostic Factors (A) The nomogram of the predicted DFS was established based on RS and Gleason score. (B) The AUC of the first, third and fifth years of the nomogram in TCGA. (C-E) The calibration curve of the nomogram in TCGA. (F) The AUC of the first and third years of the nomogram in GEO.

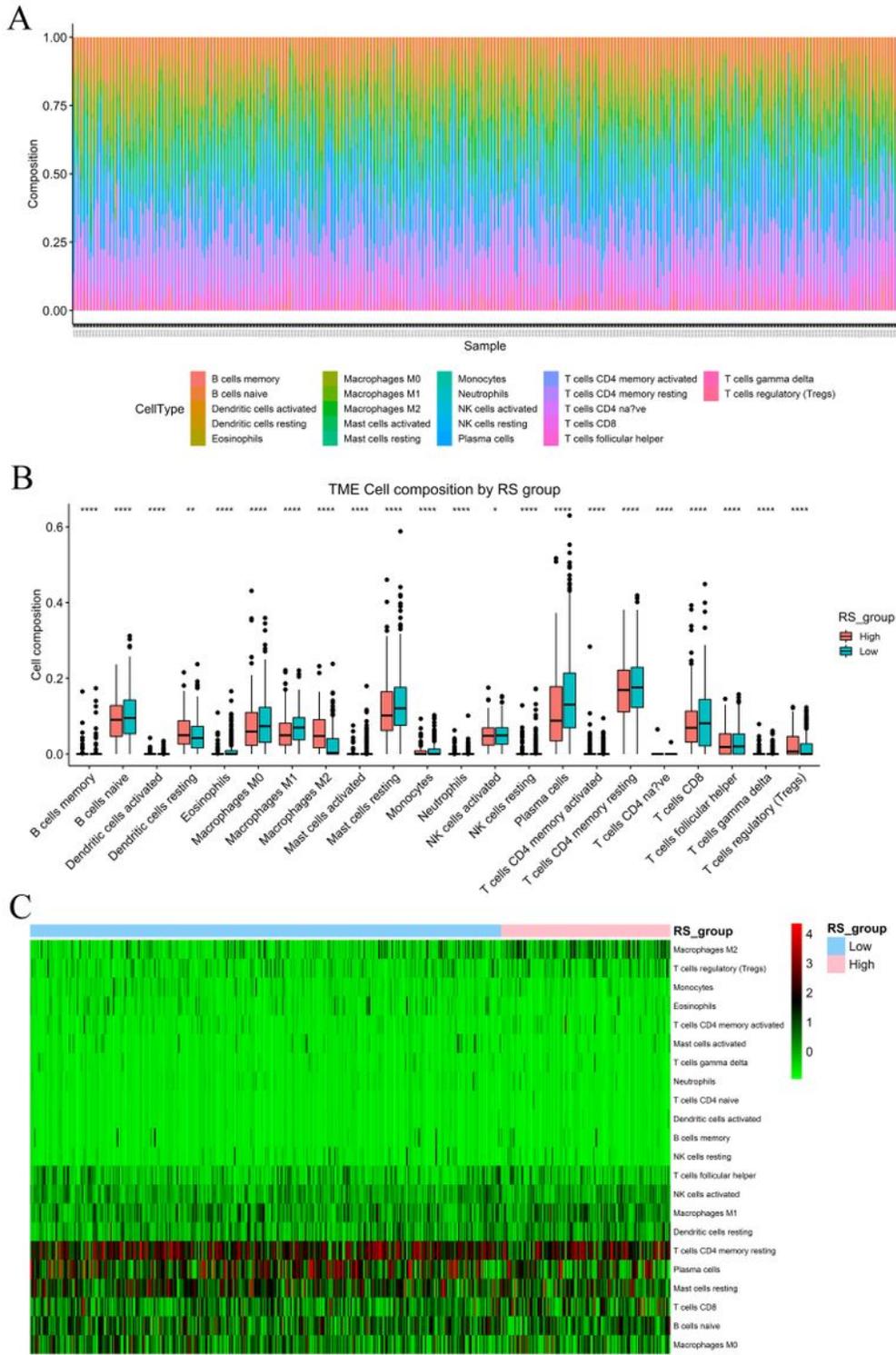


Figure 6

Estimating the composition of immune cells

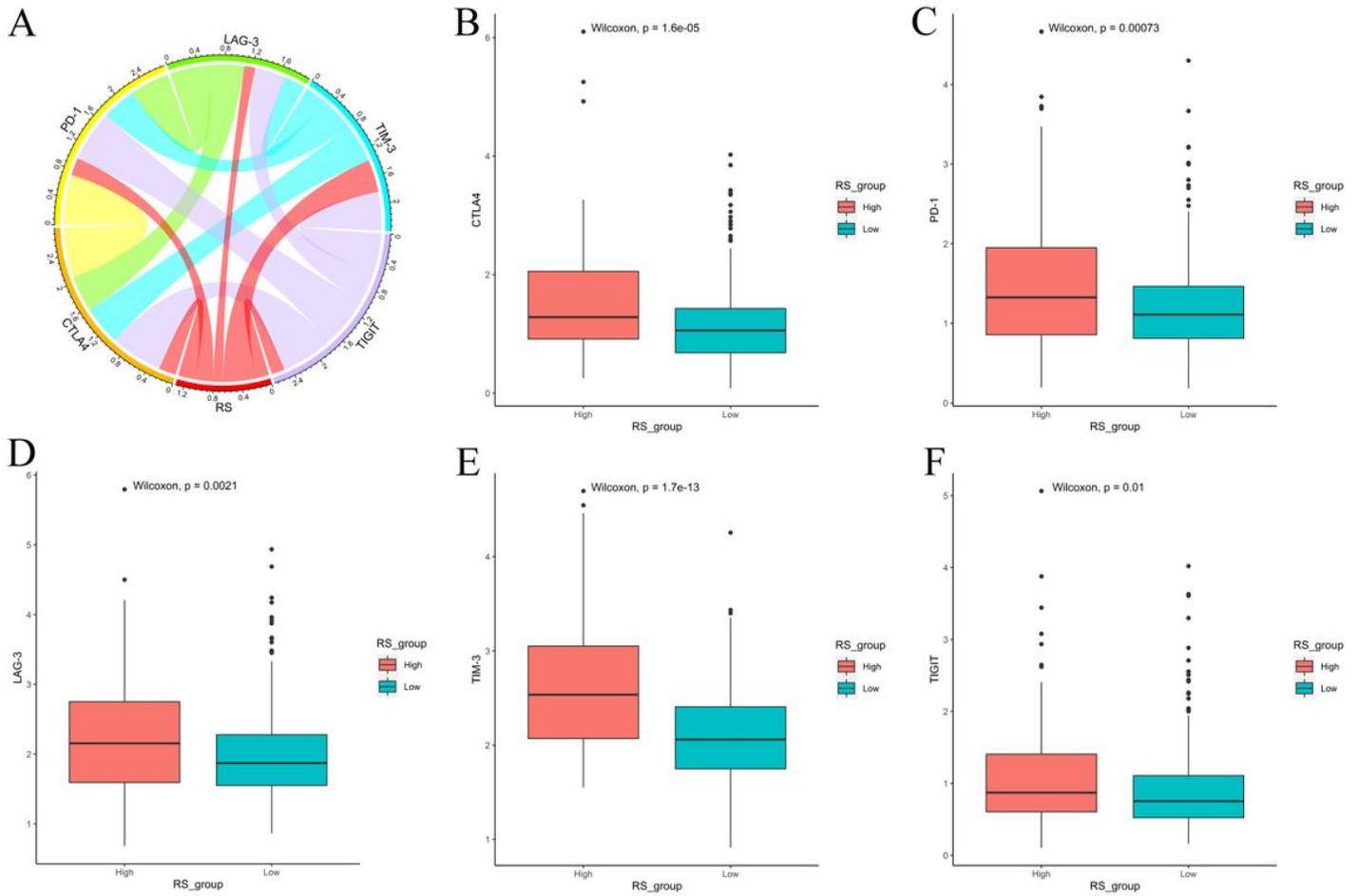


Figure 7

The expression of immune checkpoints between high-RS and low-RS PCa patients.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial1.docx](#)
- [SupplementaryFigureLegend.docx](#)
- [SupplementaryFigure3.tif](#)
- [SupplementaryFigure2.tif](#)
- [SupplementaryFigure1.tif](#)