

# The Correlation Between Novel Survival-related Alternative Splicing Signature and Prognostic Prediction and Immune Microenvironment in Clear Cell Renal Cell Carcinoma

**Wengang Jian**

The First Affiliated Hospital of Harbin Medical University

**Gang Wang**

The First Affiliated Hospital of School of Medicine, Zhejiang University

**Yipeng Yu**

The First Affiliated Hospital of Harbin Medical University

**Licheng Cai**

The First Affiliated Hospital of Harbin Medical University

**Yongchun Yu**

The First Affiliated Hospital of Harbin Medical University

**Wei Xue**

The First Affiliated Hospital of Harbin Medical University

**Tengda Wang**

The First Affiliated Hospital of Harbin Medical University

**Yuyang Meng**

The First Affiliated Hospital of Harbin Medical University

**Zhinan Xia**

The First Affiliated Hospital of Harbin Medical University

**Wenhao Yao**

The First Affiliated Hospital of Harbin Medical University

**Cheng Zhang** (✉ [zhangcheng13836182568@zju.edu.cn](mailto:zhangcheng13836182568@zju.edu.cn))

School of Medicine, Zhejiang University

---

## Research Article

**Keywords:** Clear Cell Renal Cell Carcinoma, Alternative Splicing, Prognosis, Immune Cell Infiltration, Immune Checkpoint

**Posted Date:** December 30th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1209772/v1>

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

[Read Full License](#)

---

# Abstract

## Background

Although extensive researches related alternative splicing (AS) as the prognostic markers of patients in cancers, which remains unknown in clear cell renal cell carcinoma (ccRCC), especially in immunotherapy. Therefore, a novel survival-associated AS signature was established to predict prognosis of patients and explored its correlation with immune cell infiltration or immune checkpoint expression to explain the phenomenon of resistance to immunotherapy in ccRCC.

## Methods

According to AS data, clinical information and gene expression data in ccRCC, overall survival-related AS events was identified and further the AS-related prognostic risk model established by LASSO regression and multi-Cox regression analysis was evaluated by Kaplan-Meier survival analysis, the ROC curves and a nomogram model. Then we clarified the biological processes and pathways by GSEA, and further measured the immune cell infiltration by ESTIMATE, CIBERSORT and ssGSEA. Finally we analysed the clinical features and immune features of different parental genes, and queried the splicing factors regulating riskScore-related AS events by Spearman correlation analysis.

## Results

We obtained the most significant 5 AS events, including C4orf19|69001|AT, UACA|31438|AP, FAM120C|89237|AT, TRIM16L|39629|AP and SEC31A|100881|ES, to establish the prognostic risk model, and further illustrated the stability and importance of the riskScore prognostic signatures. Then we found that in high risk group, most of the top 10 GO enrichments and the KEGG pathway were closely related to the immunity, and the higher immune cell infiltration, and higher expression of classic immune checkpoints such as PD1 and CTLA4. In addition, 6 different parental genes were obtained, including C4orf19, ARHGAP24 DNASE1L3, P4HA1, SLC39A14 and TAF1D. These 6 genes could not be the independent prognostic signatures, but the expression of these genes was closely related to immune cells infiltration and the expression of immune checkpoints. Finally, we got aberrant 52 splicing factors regulating riskScore-related AS events.

## Conclusion

Our study discovered that overall survival-related AS events mediated by aberrant splicing factors can be constructed a prognostic risk model to predict prognosis of patients and utilized to index the situation of immune cell infiltration and immune checkpoint expression that impact tumor immune microenvironment in ccRCC.

# Introduction

Renal cell carcinoma (RCC) is a malignant tumor originated from the renal tubular epithelium, as accounting for about 90% of all malignant kidney tumors, of which is the most common form [1]. The carcinoma encompasses ten subtypes in histopathology, of which clear cell RCC (ccRCC) is the most prevalent and accounts for about 80% [2]. After partial or radical nephrectomy for patients with stage I or II in ccRCC, of which the 5-year survival rate is more than 90%, but less than 10% in those with stage IV due to lack of targeted drugs and the poor sensitivity to radiotherapy and chemotherapy [3]. Even if the combination of immunotherapy and/or targeted therapy has prolonged reasonably the effective survival time of ccRCC patients, the beneficiary group remains extremely limited and even develops into drug resistance [4]. Hence, many researchers have devoted themselves to exploring the prognostic markers of ccRCC [5, 6], especially in immunotherapy, but unsatisfactorily, those results remain completely unascertained.

The prominent biodiversity and functional complexity of human body are fundamentally caused by protein diversity, of which one of the molecular mechanism is alternative splicing (AS) and processing modification for pre-mRNA [7]. As a post-transcriptional process, AS enables the organized generation of multiple mRNAs from pre-mRNA of a single gene, which as molecular templates are translated into diverse proteins in mitochondria [8], indicating that it plays an important role in regulating the development and metabolism of cells and even provides potential opportunities for tumorigenesis or other disease when it expresses abnormally [9]. The discovery of mutations or aberrant expression in genes encoding splicing factors (SFs), a protein binding and splicing to the pre-mRNA, provides genetic evidence of a direct link between splicing misregulation and cancer pathogenesis [10]. According to recent reports, the phenomenon of tumorigenesis caused by mutations in SFs has been discovered and confirmed in multiple cancers, especially the most common in hematologic malignancies [11]. For instance, SF3B1, SRSF2 and U2AF1, as the most extensively studied SFs, their mutations are closely related to a variety of tumors, such as chronic lymphocytic leukemia [12], uveal melanoma [13] and many solid tumors including breast carcinoma and cutaneous melanoma [14]. Aberrant expression of SFs, by reason of copy number variation or modification by transcription factors such as MYC amplified in cancers, possess oncogenic properties by bringing about abundant AS events and engendering the generation of different protein isoforms [15, 16]. So far, AS events regulated by SFs mainly involve the following seven types: alternate acceptor site (AA), alternate donor site (AD), alternate promoter (AP), alternate terminator (AT), exon skip (ES), mutually exclusive exons (ME) and retained intron (RI) [17].

Recently, increasing number of researchers pay attention to and lucubrate the AS events, as an essential regulatory mechanism, for which even have developed the effective targeted drugs in the field of tumors, such as ovarian cancer [18] and prostate cancer [19]. In RCC, it has been reported that some AS events, such as PKM and HnRNPL, can promote the occurrence or development of tumors [20, 21] and even resistance to chemotherapy drugs [22], however, mainly are focused on as the novel prognostic factors [23-25]. Besides, AS events are closely related to the immune microenvironment, meanwhile the feasibility and safety of splicing isoforms produced by AS as potential tumor immunotherapy targets are in-depth discussed, of which the clinical challenges are in-detail analysed [26]. Due to the limited research

conditions, in the field of tumors, it is still more researched and reported as an indicator or a marker of tumor immune microenvironment (TIME) or immune cell infiltration [27-29].

In this study, of which details were shown in the workflow (**Figure 1A**), we constructed the prognostic risk models based on survival-related AS events in ccRCC samples from the TCGA database, then found that in high-risk patient groups, more abundance of immune cell infiltration such as regulatory T cells (Treg cells), M0 macrophages, etc., but no significant difference in infiltrative CD8<sup>+</sup> T cells, and the higher expression of immune checkpoints such as PD-1, CTLA-4, etc. These comprehensive results indicate the heterogeneity of TIME and the limited effect of immune checkpoint inhibitors (ICIs) in ccRCC.

## Methods And Materials

### 1: Data Acquisition and Processing

We downloaded gene expression data and corresponding clinical information of KIRC (kidney renal clear cell carcinoma, same as ccRCC) samples from The Cancer Genome Atlas (TCGA) data portal website (<https://portal.gdc.cancer.gov/>) [30]. Alternative splicing (AS) data from KIRC samples were acquired from SpliceSeq data website ([https://bioinformatics.mdanderson.org/TCGA\\_SpliceSeq/PSIdownload](https://bioinformatics.mdanderson.org/TCGA_SpliceSeq/PSIdownload)), which provides a resource for investigation of tumor-normal alterations in mRNA splicing patterns of TCGA RNASeq data [31]. The following screening criteria were used: (1) patients pathologically diagnosed with KIRC; (2) samples with complete clinical characteristics, including age, gender, survival time, survival state, histological grade and pathological stage, including T, N, and M; (3) samples with integral mRNA expression data and corresponding AS data; (4) deletions of AS events where the percentage of samples with percent-spliced-in (PSI) value is greater than 75%; (5) deletions of samples where the percentage of AS events with PSI value is less than 70%; (6) deletions of AS events of which the mean of PSI value is greater than 0.95 or less than 0.05; (7) deletions of AS events of which the standard deviation (SD) of PSI value is less than 0.05. For minimizing the possible bias caused by the NA values of PSI, the “Knn” algorithm of “impute” (version: 3.13) package in R (version: 3.4.2) was used to perform multiple interpolations of the missing values. The “UpSetR” (version: 1.3.3) package was used to create the UpSet diagram, which shows the gene interaction sets in 7 types of all AS events.

### 2: Identification of survival-associated AS events in KIRC

To combine the clinical information of samples and the PSI value of AS events of the corresponding samples by Perl (version: 5.30.0), which was used to estimate the correlation ( $P < 0.05$ ) between AS events and overall survival (OS) by implementing the “survival” (version: 3.2-3) package and univariate Cox regression analysis in R. The “UpSetR” package was used to create the UpSet diagram of the AS events relevant to OS. The volcano plot and the bubble graphs were drawn by the “ggplot2” (version: 3.2.1) package.

### 3: Establishment and Evaluation of AS-related prognostic risk models

Based out of the OS-related AS events identified by univariate Cox regression analysis, we selected the top 50 of these AS events to employ the LASSO (least absolute shrinkage and selection operator) regression to identify the final elimination of potential predictors with nonzero coefficients by the “glmnet” (version: 1.0.6) package, which avoids model overfitting to obtain the smallest parameter model. Then the multivariate Cox regression analysis was adopted to further comprehensively evaluate the contribution of each OS-related AS events filtrated by the LASSO regression. Finally, the riskScores were calculated using the following equations:

$$riskScore = \sum_{i=0}^n (PSI_i * Coef_i)$$

where  $PSI_i$  was the PSI value of OS-related AS event,  $Coef_i$  was the correspondent regression coefficient. Based on the riskScore, ccRCC patients were divided into two groups (high/low risk), and the riskScore and survival state of these patients were severally showed through the diagrams. The “pheatmap” (version: 1.0.12) package was used to display the OS-related AS events in the low-risk and high-risk groups. The “survival” and “survminer” (version: 0.4.9) packages were employed to perform Kaplan-Meier survival analyse with log-rank tests. Besides the univariate and multivariate Cox regression analysis by the “survival” packages were further used to analyze whether the riskScore can be an independent prognostic indicator. The 1-, 3- and 5-year ROC analysis were performed by the “survivalROC” (version: 1.0.3) package. The nomogram was created by the “rms” (version: 6.2-0) package to comprehensively evaluate or predict the 1-, 2- and 3-year survival probability of ccRCC patients.

Meanwhile, we further respectively analyzed the top 20 of other 7 types of OS-related AS events in the same way in turn.

#### **4 Gene Set Enrichment Analysis (GSEA)**

Differential gene sets (DGSs) in the high/low risk subgroups were further analysed by GSEA, including GO (Gene Ontology) analysis (c5.go.v7.4.symbols.gmt) which composed of BP (biological process), MF (molecular function) and CC (cellular component), and KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis (c2.cp.kegg.v7.4.symbols.gmt) which was KEGG pathway, which were conducted by the “limma” (version: 3.48.3), “org.Hs.eg.db” (version: 3.13.0), “clusterProfiler” (version: 4.0.5), and “enrichplot” (version: 1.12.3) packages.  $P < 0.05$  was considered statistically significant.

#### **5: Measurement of the Tumor Immune Microenvironment (TIME)**

We extracted the expression profile data of each ccRCC sample, and calculated the tumor microenvironment (TME) score by the “estimate” package, which included the stromal cell score, immune cell score and tumor purity. Then the “ggpubr” (version: 0.4.0) package was used to plot between the risk clusters and TME scores. CIBERSORT was used to evaluate the fractions of 22 types of immune cell infiltration in each ccRCC sample based on expression profile data. Then the “ggpubr” and “vioplot” (version: 0.3.7) packages was used to plot between the risk clusters and the fractions of 22

types of immune cell infiltration. On the other hand, the scores of immune cells and their functions in each ccRCC sample were obtained by ssGSEA which need to use the “GSVA” (version: 1.40.1) and “GSEABase” (version: 0.3.7) packages. The “ggpubr”, “reshape2” (version: 1.4.4) and “pheatmap” packages was used to plot between the risk clusters and the scores of immune cells and their functions. The “corrplot” package was used to display the correlation between the riskScore and significant immune checkpoint which included PD-1 (PDCD1), PD-L1 (CD274), CTLA4, HAVCR2, IDO1 and PDCD1LG2, meanwhile the “ggpubr” and “reshape2” packages was used to display the expression of immune checkpoint related genes in the risk clusters.

## 6: Identification of different parental genes of riskScore-related AS events in ccRCC

The “limma” package was implemented to obtained different parental genes of riskScore-related AS events in ccRCC,  $|\log_{2}FC| > 0.585$  and  $P < 0.05$  as the cutoffs. The clinical features and immune features of these genes were analysed by the preceding methods.

## 7: The regulatory network between riskScore-related AS events and splicing factors

A total of 404 human splicing factors (SFs) were downloaded from the reference [32]. A total of 212 SFs related to OS were testified by univariate Cox regression analysis. The correlation coefficient between the PSI values of riskScore-related AS events and the SFs expression values which were extracted from the RNA sequencing data of ccRCC samples in TCGA database, were calculated by Spearman correlation ( $r > 0.6$ ,  $P < 0.05$ ). The regulatory network of SFs and AS events was constructed and plotted by Cytoscape (version 3.9.0).

## 8: Statistical analysis

All statistical analyses were conducted in R software (version: 3.4.2). All statistical tests with  $p < 0.05$  (two-sided test) were considered statistically significant.

# Results

## 1: Overview of AS events and identification of the OS-related AS events in ccRCC

According to the screening and filtering criteria described in the methods, finally, a total of 525 ccRCC samples and 72 matched normal samples were enrolled in our study cohort, and the clinical characteristics of 525 patients were summarized in **Supplementary Table 1**. Meanwhile, this research retained a total of 31407 AS events and 9734 parental genes, which were used to crate the UpSet diagram (**Figure 2A**), showing the interaction sets in 7 types (AA, AD, AP, AS, ES, ME, RI) of all AS events, and respectively displayed the type and number of AS events and parental genes through a histogram (**Figure 2B**). In addition to this, it could be clearly seen that ES (AS=12142, Gene=5592) was the most common AS events, and ME (AS=151, Gene=145) was the rarest (**Figure 2B**). Proverbially, one type of AS events, such as ES, can exist in multiple genes, but interestingly, at the most five different AS events occurred simultaneously in one single gene, for instance, shows of 39 genes of this circumstance in the

UpSet which just included the top 50 gene intersection sets in number such as the top one was the ES (n=1779) (**Figure 2A**).

To identify the OS-related AS events in ccRCC, we conducted a univariate Cox proportional hazards regression analysis ( $P < 0.05$ ), and then detected 9339 survival-associated AS events in 4480 parental genes, of which the types and number were respectively recorded in a histogram (**Figure 2D**). The interaction between these OS-related AS events and parental genes were shown in an UpSet plot (**Figure 2C**). At this point, the AP (AS=2838, Gene=1461) and ES (AS=2517, Gene=1723) events were the highest in number among them, and the ME (AS=54, Gene=54) events remained the least (**Figure 2D**). The volcano plot (**Figure S1A**) intuitively showed the all AS events, including the significant prognostic AS events ( $P$  value  $< 0.05$ ) and no significant AS events. The top 20 OS-related AS events of each of the 7 types, respectively, were clearly shown in the bubble graphs (**Figure S1B-H**).

## **2: Establishment and evaluation of the riskScore prognostic signatures about OS-related AS events in ccRCC**

In order to find the most significant prognostic signatures, we then conducted a LASSO regression and a multivariate Cox regression according to the OS-related AS events. We obtained the most significant 17 events from the top 50 OS-related AS events through the LASSO regression (**Figure 3A and 3B**), and further filtrated a list of prognostic signatures (**Supplementary Excel 1**) by the multivariate Cox regression and calculated the riskScore of each ccRCC samples based on the weight of each gene (**Supplementary Excel 2**). Now, the 525 KIRC patients were classified into the low-risk (n=263) cohort and the high-risk (n=262) cohort (**Figure 3C**) depended on the median of those riskScores, and then renewedly verified the survival state of each ccRCC patients with riskScore (**Figure 3D**). The most significant OS-related AS events after the multivariate Cox regression were also shown in the heat map (**Figure 3E**). It was demonstrated that ccRCC patients with high riskScore were associated with poor OS ( $P < 0.001$ ) by the Kaplan-Meier analysis with the log-rank tests (**Figure 3F**). We performed the univariate and multivariate independent survival tests to evaluate the prediction ability of the riskScore as an independent prognostic risk factor. The univariate independent survival analysis indicated that the riskScore, tumor grade and stage were closely correlated ( $P < 0.001$ ) to the prognosis of ccRCC patients in risk cohorts (**Figure S2A**). The multivariate independent survival analysis demonstrated that the riskScore could be used as the independent prognostic indicator with significant differences ( $P < 0.001$ ) (**Figure 3G**). The ROC (receiver operator characteristic) plot of survival time of the riskScore showed the AUC (area under the curve) values were all greater than 0.7 (1 year=0.790, 2 year=0.745, 3 year=0.761) (**Figure 3H**), declaring the excellence of the sensitivity and specificity of the riskScore-related AS events as prognostic signatures. Meanwhile, the AUC values of the riskScore, the grade and the stage were all greater than 0.7 (risk=0.790, grade=0.715, stage=0.793) in the ROC plot of clinical characteristics (**Figure S2B**). These statistics provided sufficient evidence to prove the important value of the OS-related AS events as prognostic signatures in ccRCC patients.

When further analyzed separately the top 20 of other 7 types of OS-related AS events in the same way in turn, we found that the significant prognostic signatures for each type in these 7 types can be still established (**Figure S3A-G**), illustrating the stability and importance of the riskScore prognostic signatures created with the OS-related AS events in ccRCC patients.

### **3: The predictive value of overall survival of ccRCC patients**

To reveal the correlation between AS-related prognostic signatures and clinical characteristics of ccRCC patients, we measured the correlation between the riskScore and different clinical pathological parameters, including age, gender, grade, and stage-T, M, and N (**Figure S4A-G**), illustrating the significant positive correlation ( $P < 0.05$ ) between the riskScore and these clinical characteristics, excepting the age and gender ( $P > 0.05$ ). In addition, we constructed a nomograph model, which incorporates eight prognostic factors (such as riskScore, age, gender, grade, stage, T, M and N), to predict 1-, 2- and 3-year survival rate of patients with ccRCC (**Figure 4A**). As showed in the calibration curves (**Figure 4B-D**), the red lines were very close to the ideal gray lines, which proved the authenticity and effectiveness of the predictive power of the nomograph model. Collectively, these data demonstrate that the AS events-related prognostic signatures can be applied to predict OS in ccRCC patients.

When further analyzed separately the other 7 types of OS-related AS events in all above way (**Figure S5A-G**), respectively, we testified the stability of these prognostic signatures in different types of OS-related AS events.

### **4: The biological processes and pathways of differential gene sets in ccRCC subgroups with high/low risk.**

In order to further clarify the biological processes and pathways in the high/low risk subgroups, we analyzed the differential gene sets (DGSs) by the GSEA, including GO analysis which composed of BP, MF and CC, and KEGG analysis which was KEGG pathway. We found that in high risk group, the top 10 GO enrichments in the DGSs were all the BP, most of which were closely related to the immunity, such as GOBP\_activation\_of\_immune\_response, GOBP\_B\_cell\_activation, GOBP\_B\_cell\_mediated\_immunity, GOBP\_B\_cell\_receptor\_signaling\_pathway, etc. (**Figure 5A**). However, in low risk group, most of the top 10 GO enrichments in the DGSs were the CC, such as GOCC\_apical\_part\_of\_cell, GOCC\_apical\_plasma\_membrane, GOCC\_cajal\_body, etc. (**Figure 5B**). Meanwhile, in the terms of KEGG pathway, the DGSs in high risk group were also enriched in immune-related pathways, such as KEGG\_chemokine\_signaling\_pathway, KEGG\_cytokine\_cytokine\_receptor\_interactin and KEGG\_intestinal\_immune\_network\_for\_iga\_production (**Figure 5C**). However, the DGSs in low risk group tended to be enriched in metabolism-related pathways, such as KEGG\_citrate\_cycle\_TCA\_cycle, KEGG\_fatty\_acid\_metabolism, KEGG\_glycolysis\_gluconeogenesis, KEGG\_propanoate\_metabolism and KEGG\_retinol\_metabolism (**Figure 5D**). Hence, it was a hypothesis that tumor immune microenvironment (TIME) might play an important role in high risk group. Those subgroups classified according to the 7 risk models built based on the other 7 types of OS-related AS events were further analyzed by the GSEA (**Figure S6A-G**). Interestingly and fortunately, these GO enrichments and KEGG pathways enriched in 7

high risk subgroups were all consistent with those above results, which tended to be closely related to the immunity, so TIME in the high/low risk subgroups was focused on the later research.

## 5 The immune features in ccRCC cohorts with high/low risk.

To investigate thoroughly the effect of TIME in ccRCC subgroups with high/low risk, we calculated the stromal cell score, immune cell score and tumor purity to reflect TME of each ccRCC patients with the riskScore, and found interestingly that ccRCC subgroup with high riskScore tended to manifest the higher immune cell score ( $P < 0.05$ ) (**Figure 6B**), though, no relationship between these clusters and the stromal cell score ( $P > 0.05$ ) (**Figure 6A**), inferring a hypothesis of the higher immune cell infiltration in high riskScore subgroup. To verify the hypothesis, we estimated a comparison of the otherness of the 22 immune cells in high/low riskScore cohorts (**Figure 6C**), discovered amazingly that these immune cells (such as naive B cell, CD8<sup>+</sup> T cell, naive CD4<sup>+</sup> T cell, activated NK cell and activated Mast cell, which were regulating immunological lethality) were indifference relation, and the immune cells (such as resting memory CD4<sup>+</sup> T cell and macrophages M1/M2) were fewer in high riskScore cluster (**Figure S7A-C**), but more the cells regulating immunological suppression such as regulatory T cells (Tregs) (**Figure S7D**). Further to evaluate the immune cell infiltration and their functions in each ccRCC sample by ssGSEA, as shown in **Figures 6D and 6E**, the results were roughly consistent with those above results. Then there was a positive correlation between high-risk patients and classic immune checkpoints such as PD-1 and CTLA4 (**Figure 6F and S7E-F**), and the vast majority of genes related to immune checkpoints were increased in the high riskScore patients group (**Figure 6G**). These results implied that TIME affected by immune cell infiltration and immune checkpoints may be one of the most important reasons that contribute to poor prognosis in the high risk subgroups differentiated based on AS events.

When further analyzed separately the other clusters based on the different riskScore in the 7 types of OS-related AS events in turn (**Figure S8A-G**), we testified the stability of relationships between ccRCC patients and immune features.

## 6: The clinical features and immune features of different parental genes of riskScore-related AS events in ccRCC

To combine all parental genes of AS events used to establish these risk models and analyse the differences of these genes between ccRCC tumor samples and normal samples in TCGA database (**Supplementary Excel 3**), the C4orf19 was confirmed as the only crossed different gene ( $|\log FC| > 0.585$ ,  $P < 0.05$ ) between the model built with top 50 AS events and other independent 7 models built with top 20 AS events in each type. Firstly, compared with normal samples, the expression of C4orf19 was lower in ccRCC tumor samples (**Figure 7A**). And according to the median of the C4orf19 expression, these ccRCC samples were divided into the low-expression group and the high-expression group to research the clinicopathological features of C4orf19, showing that except for age, gender and lymph node metastasis, the expression of C4orf19 was negatively correlated with grade, stage, T and M. (**Figure 7B-C and S9A-E**). Meanwhile, patients with low expression of C4orf19 had a poorer prognosis than patients with high

expression of C4orf19 (**Figure 7D**), but of which the sensitivity and specificity (AUC: 0.661) were ordinary as a prognostic signatures from the ROC curve (**Figure 7E**). In addition to the C4orf19, other different parental genes ( $|\log FC| > 1$ ,  $P < 0.05$ ), including ARHGAP24, DNASE1L3, P4HA1, SLC39A14 and TAF1D (**Supplementary Excel3**), were from their respective models and were analyzed respectively their clinicopathological features in ccRCC data by the preceding similar way. We discovered that the clinicopathological features and trends of ARHGAP24 and DNASE1L3 were similar to those of C4orf19, while those of P4HA1 and TAF1D were opposite to those of the former three genes (**Figure S10A-E**). As a special case, SLC39A14 was obviously high expression in tumor samples compared with normal, but no significant difference in clinicopathological characteristics (**Figure S10D**). What is noteworthy was that the sensitivity and specificity of these 5 genes were weak as a prognostic signatures by itself (**Figure S10A-E**), similar to C4orf19, indicating that these 6 genes can not be the independent prognostic signatures.

To explore the immune features of different parental genes, it was shown that patients with low expression of C4orf19 had more immune cell infiltration (**Figure 7F-G**), less purity of tumor (**Figure S9F-G**), more infiltration of Treg cell (**Figure 7H-I**), and more expression of these immune checkpoint related genes (**Figure 7J**). However, in the perspective of immune cell infiltration, except that the results of ARHGAP24 were coincident with the C4orf19, the results of DNASE1L3, P4HA1, SLC39A14 and TAF1D were no significant (**Figure S10A-E**). Similarly, the expression of immune checkpoint related genes of patients with low expression of ARHGAP24 or DNASE1L3 were consistent with those of patients with low expression of C4orf19, as well as patients with high expression of P4HA1, SLC39A14 or TAF1D (**Figure S10A-E**). In sum, the expression of these genes can cause more immune cell infiltration in ccRCC, especially Treg cells, and further lead to abnormal expression of immune checkpoints by regulating the expression of genes related to immune checkpoints, which may contribute to ineffective immunotherapy or drug resistance.

## 7: The splicing factors regulating riskScore-related AS events.

As we all know, the irregular occurrence of AS events is caused by the SFs which is a protein binding and splicing to the pre-mRNA. In order to explore the SFs which regulates riskScore-related AS events, a total of 212 SFs related to OS were testified by univariate Cox regression analysis. The correlation coefficient between the PSI values of riskScore-related AS events ( $n=46$ ) and the SFs ( $n=212$ ) expression values were calculated by Spearman correlation,  $r > 0.6$  and  $P < 0.05$  as the cutoffs. As shown in the network, which was constituted of the 12 AS events and 52 SFs (**Figure 8**), an AS event could be regulated by multiple SFs, for instance TAF1D|18313|RI and TMEM104|217418|ME, and a SF could regulate multiple AS events, such as SNRNP70, SER31B, etc. There SFs were divided into high expression cluster ( $n=37$ , blue, rhombus) and low expression cluster ( $n=15$ , azure, triangle) in ccRCC samples. And there AS events were divided into high frequency cluster ( $n=10$ , red, round) and low frequency cluster ( $n=2$ , lime, round) based on their PSI values in ccRCC samples. Red line represents the positive correlation, for instance, most of highly expressed SFs (blue) and all high-frequency AS events (red) were connected by the red line. Azure line represents the negative correlation, for example, most of lowly expressed SFs (azure) and the high-frequency AS events (red) such as GIT2|24375|AA were connected by the azure line, meanwhile

the low-frequency AS events (lime) such as SERP2|25779|ME and most of highly expressed SFs (blue) were connected by the azure line. These results expound that the unbalance of the AS events related to high-risk model may be caused by the abnormal expression of SFs, contributing to poor prognosis in ccRCC patients.

## Discussion

Both of initiation and progression of tumor are a complicated process, which is an intricate regulatory network involving many genes or proteins. Recently, the pathogenesis of ccRCC has been elaborated from the perspective of tumor evolution by some scholars, who believed through the TRACERx Renal studies that in ccRCC, driver gene mutations including those in VHL, SETD2, PBRM1 and BAP1, and chromosomal copy number variation including chromosome 3p loss, may engender genomic instability and promote defects in DNA repair pathways, which drive tumor initiation events [33]. Then with emergence of these driver events, the enormous irregular changes of downstream genes regulated by these driver genes at the transcription and translation levels, which promote tumor progression, and even confer lethality [34]. Thus it can be seen that AS, as an indispensable part of the process of post-transcriptional modifications, of which it is particularly important to clarify function and molecular mechanism of ccRCC.

With the tremendous advance of high-throughput sequencing technology, the potential significance of AS profiling obtained from RNA-seq datas of tumor has been great breakthroughs in cancer [8, 9], including in ccRCC. In previous research, Ge Y, et al. kept on this view that DCLK1-ASVs, which is an alternate splice variant of DCLK1, assistantly maintains the stemness of ccRCC cell, and potentially heightens the curative effect of immunotherapy by regulating PD-L1 or CTLA4 [35]. In the early time, PKM2, rather than PKM1, was discovered to play a tumorigenic role in ccRCC [36], and recently, four transcript isoforms of the PKM type were combined to evaluate the survival outcome of ccRCC patients [37]. Moreover, the splicing factor SF3B3 promoted cell growth, migration and proliferation in ccRCC cell lines and even tumorigenicity in a xenograft model through inhibiting expression of EZH2  $\Delta$ 14 and promoting expression of EZH2 [38]. Similarly, the alternative splicing events of CCDC50 regulated by HnRNP A1 have been demonstrated to contribute to cancer progression through the ZNF395 pathway in ccRCC [39]. In addition, AS events, as prognostic signatures, are concerned about and researched. For instance, Song J, et al. established a robust prognostic prediction model based on the seven hub genes, including KRT222, LENG8, APOB, SLC3A1, SCD5, AQP1, and ADRA1A, who believed that this model can be as the potential prognostic biomarkers for ccRCC [40]. Synchronously, Meng T, et al. revealed that RHOT2-32938-RI and TCIRG1-17288-RI regulated by DDX39B, two AS events in RI type, may be intimately associated with the metastasis and poor prognosis of ccRCC [41]. Furthermore, Chen T, et al. systematically analysed the survival-associated AS signatures from the ccRCC samples in TCGA database, and then created prognostic predictors based on AS events and uncovered key SFs in splicing networks [42]. But unsatisfactorily, previous research results remain completely unascertained about the prognostic biomarkers of ccRCC.

Therefore, likewise, in our study, a total of 525 ccRCC samples and 72 matched normal samples from the TCGA database were enrolled in our study cohort, and then 9339 survival-associated AS events in 4480 parental genes were detected from among 31407 AS events and 9734 parental genes by a univariate Cox regression analysis ( $P < 0.05$ ) (**Figure 2**). Secondly, we obtained the most significant 5 events, including C4orf19|69001|AT, UACA|31438|AP, FAM120C|89237|AT, TRIM16L|39629|AP and SEC31A|100881|ES, from the top 50 OS-related AS events through the LASSO regression and the multivariate Cox regression, which were used to establish the prognostic risk model (**Figure 3A-E**). These results, which were evaluated by the Kaplan-Meier analysis, the univariate and multivariate independent survival tests, the ROC curves (**Figure 3F-H and S2A-B**) and a nomograph model (**Figure 4A-D**), provided sufficient evidence to prove the availability of the OS-related AS events obtained by our research as prognostic signatures in ccRCC patients. Differ from the results of other previous studies, we further analyzed separately the top 20 of other 7 types of OS-related AS events in the same way in turn, and found that the significant prognostic signatures for each type in these 7 types can be still established (**Figure S3A-G and S5A-G**). Hence, our research results further illustrated the stability and importance of the riskScore prognostic signatures created with the OS-related AS events in ccRCC patients. And so far, the practicability and effectiveness of AS events as a prognostic biomarker for predicting the prognosis of ccRCC patients requires more clinical trials to verify.

In order to further clarify the biological processes and pathways in the high/low risk subgroups, GO analysis and KEGG analysis were carried out, and we found that in high risk group, most of the top 10 GO enrichments and the KEGG pathway were closely related to the immunity (**Figure 5A-D**), so it was a hypothesis that TIME might play an important role in high risk group. This hypothesis had been successfully confirmed by those results of those GO enrichments and KEGG pathways enriched in 7 high risk subgroups (**Figure S6A-G**). Therefore, we further reviewed the literature to understand the role of AS events in the TME, especially in the TIME. Firstly, Dery KJ, et al. constructed transcriptome sequencing in cell lines of breast carcinoma, to find that IRF-1 can take part in regulation of growth and differentiation of tumor cells through inducing the global AS of genes, as well as augmentation of the immune response by global AS changes of genes of the cytokine family [43]. Next, a study of Sotillo E, et al. showed that alternative splicing of CD19 can reinforce the resistance to CART-19 immunotherapy by compromising expression of the CART-19 epitope [44]. Again, the research results of Georgilis A, et al. manifested that AS events mediated by PTBP-1, a splicing factor, may regulate the inflammatory secretome and accelerate tumorigenic effects of senescent cells [45]. These studies demonstrate that AS events can play a very important role in the TIME. What's more, the relationship between AS events and TIME has been confirmed in a variety of tumor diseases, such as hepatocellular carcinoma [46], esophageal squamous cell carcinoma [47], gastric carcinoma [48], endometrial carcinoma [49], triple negative breast carcinoma [50] and glioma [51]. However, this relationship has not been studied in ccRCC.

In this study, we amazingly found that the higher immune cell infiltration such as regulatory T cells (Tregs), rather than immune cell regulating immunological lethality such as CD8<sup>+</sup> T cell and activated NK cell, etc., and higher expression of classic immune checkpoints such as PD-1 and CTLA4 involved in high

riskScore subgroup (**Figure 6A-G and S7A-F**), which implied that TIME affected by immune cell infiltration and immune checkpoints may be an important reason that contributes to poor prognosis in the high risk subgroups differentiated based on AS events. This phenomenon has been confirmed again by those results between immune features and ccRCC patients who were divided into different risk clusters based on the 7 types of OS-related AS events (**Figure S8A-G**). In addition, we further explored the clinical features and immune features of different parental genes of riskScore-related AS events in ccRCC, and discovered that 6 different parental genes were obtained, of which 3 genes presented low expression in tumor samples, including C4orf19, ARHGAP24 and DNASE1L3, as well as other 3 genes of high expression, such as P4HA1, SLC39A14 and TAF1D. These 6 genes can not be the independent prognostic signatures (**Figure 7A-E and S10A-E**). Similarly, the expression of these genes was closely related to immune cells infiltration and the expression of immune checkpoints in TIME of ccRCC (**Figure 7F-J and S10A-E**), which was consistent with the former results.

Not only that, we also in surprise found that immune cells, such as CD8<sup>+</sup> T cell, resting memory CD4<sup>+</sup> T cell and macrophages M1/M2, themselves infiltration scores were very high in all ccRCC samples (**Figure 6C-E and S7A-G**), which makes us mistakenly believe the significant augment of these immune cells with lethality in ccRCC. However, there is a recent viewpoint from the research of Braun DA, et al. scholars, who performed single-cell RNA and T cell receptor sequencing on 164,722 individual cells from tumor and adjacent non-tumor tissue in patients with ccRCC, then proposed that exhausted CD8<sup>+</sup>T cells and macrophages M2 co-occurred in advanced tumor, and expressed ligands and receptors facilitating T cell dysfunction [52]. The results indicate that immune cells significantly infiltrating in ccRCC do not always have effective immunity lethality, and the possible factors of resistance to immunotherapy is explained. Furthermore, Bi K, et al. believed that immune checkpoint blockade (ICB) might remodel the TIME in ccRCC and modify the interplay between tumor cells and immune cell populations, which explained the reasons of response and resistance to ICB [53]. Meanwhile, Krishna C, et al. performed paired single-cell RNA and T cell receptor sequencing of 167283 cells, and discovered that the enrichment of tissue-resident CD8<sup>+</sup> T cells in a patient responsive to ICB but tumor-associated macrophages in a resistant patient, which emphasizes the existence of heterogeneity of the immune microenvironment within and between ccRCC patients [54]. Accordingly, those study will provide a crucial resource for the discovery of new immunotherapies and the advance of combination regimens in ccRCC.

In this study, we illustrated the stability and importance of the riskScore prognostic signatures from these prognostic risk models which were repetitiously constructed based on survival-related AS events in ccRCC samples from the TCGA database. After dividing into high/low risk groups based on these riskScores, the results from the GO and KEGG analyses showed that TIME might play an important role in high risk group. Hence, we further found that in high-risk patient groups, more abundance of immune cell infiltrations such as Treg cells, etc., but no significant difference in infiltrative CD8<sup>+</sup> T cells, and the higher expression of immune checkpoints such as PD-1, CTLA-4, etc. These comprehensive results indicate the heterogeneity of TIME and the limited effect of ICIs in ccRCC.

# Conclusion

In summary, our study suggested that OS-related AS events mediated by aberrant SFs can be constructed a prognostic risk model to predict prognosis of patients and utilized to index the situation of immune cell infiltration and immune checkpoint expression that impact TIME in ccRCC.

# Abbreviations

RCC: renal cell carcinoma; ccRCC: clear cell renal cell carcinoma; AS: alternative splicing; SFs: splicing factors; mRNA: messenger RNA; OS: overall survival; AA: alternate acceptor site; AD: alternate donor site; AP: alternate promoter; AT: alternate terminator; ES: exon skip; ME: mutually exclusive exons; RI: retained intron; TME: tumor microenvironment; TIME: tumor immune microenvironment; ICIs: immune checkpoint inhibitors; KIRC: kidney renal clear cell carcinoma; TCGA: The Cancer Genome Atlas; PSI: percent-spliced-in; SD: standard deviation; LASSO: least absolute shrinkage and selection operator; GSEA: gene set enrichment analysis; GO: gene ontology; BP: biological process; MF: molecular function; CC: cellular component; KEGG: Kyoto Encyclopedia of Genes and Genomes; ROC: receiver operator characteristic; AUC: area under the curve; DGSs: differential gene sets.

# Declarations

## Acknowledgements

The authors would like to give our sincere appreciation to the research groups for the TCGA and SpliceSeq databases, which provided data for this collection.

## Author contributions:

The original draft was prepared and written by WJ. GW, YY and LC were in charge of the data curation. YY, WX and TW helped to review and modify all figure, YM, ZX and WY helped to review and modify the article. CZ carried out the experiment design and administered the project. All authors contributed to the article and approved the submitted version.

## Funding:

This study was funded partly by the National Natural Science Foundation of China (Grant No. 81872084) and the Postdoctoral Science Foundation of China (Grant NO. 2019M662080).

## Availability of data and materials:

The datasets exhibited in this study can be discovered in TCGA (<https://portal.gdc.cancer.gov/>) and SpliceSeq (<https://bioinformatics.mdanderson.org/TCGASpliceSeq/PSIdownload>) databases. Further inquiries can be directed to the corresponding authors.

## Ethics approval and consent to participate:

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## References

- [1] Capitanio U, Bensalah K, Bex A, et al. Epidemiology of Renal Cell Carcinoma. *Eur Urol.* 2019; 75 (1):74-84. doi:10.1016/j.eururo.2018.08.036
- [2] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020; 70 (1):7-30. doi:10.3322/caac.21590
- [3] Barata PC, Rini BI. Treatment of renal cell carcinoma: Current status and future directions. *CA Cancer J Clin.* 2017; 67 (6):507-524. doi:10.3322/caac.21411
- [4] Braun DA, Bakouny Z, Hirsch L, et al. Beyond conventional immune-checkpoint inhibition - novel immunotherapies for renal cell carcinoma. *Nat Rev Clin Oncol.* 2021; 18 (4):199-214. doi:10.1038/s41571-020-00455-z
- [5] van Vlodrop IJH, Joosten SC, De Meyer T, et al. A Four-Gene Promoter Methylation Marker Panel Consisting of GREM1, NEURL, LAD1, and NEFH Predicts Survival of Clear Cell Renal Cell Cancer Patients. *Clin Cancer Res.* 2017; 23 (8):2006-2018. doi:10.1158/1078-0432.CCR-16-1236
- [6] Khattak MA, Bakr F, Krzystanek M, et al. Prognostic and predictive markers in metastatic renal cell carcinoma. *J Clin Oncol.* 2013; 31 (7):971-972. doi:10.1200/JCO.2012.46.9353
- [7] Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol.* 2017; 18 (7):437-451. doi:10.1038/nrm.2017.27
- [8] Bonnal SC, López-Oreja I, Valcárcel J. Roles and mechanisms of alternative splicing in cancer - implications for care. *Nat Rev Clin Oncol.* 2020; 17 (8):457-474. doi: 10.1038/s41571-020-0350-x
- [9] Rahman MA, Krainer AR, Abdel-Wahab O. SnapShot: Splicing Alterations in Cancer. *Cell.* 2020; 180 (1):208-208.e1. doi:10.1016/j.cell.2019.12.011
- [10] O'Connor CM, Narla G. Splice of Life for Cancer: Missplicing of PPP2R5A by Mutant SF3B1 Leads to MYC Stabilization and Tumorigenesis. *Cancer Discov.* 2020; 10 (6):765-767. doi:10.1158/2159-8290.CD-

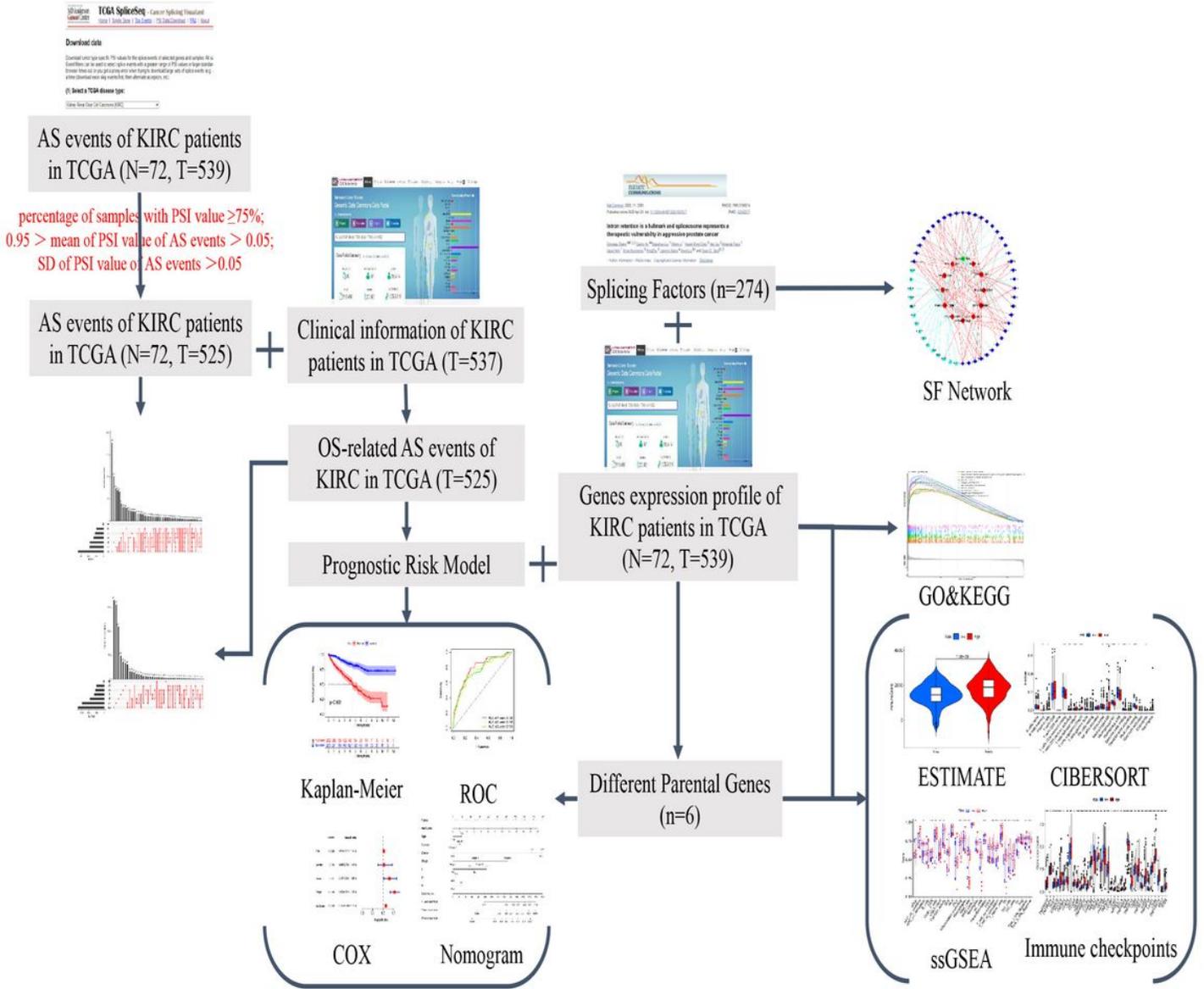
- [11] Seiler M, Peng S, Agrawal AA, et al. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep.* 2018;23(1):282-296.e4. doi:10.1016/j.celrep.2018.01.088
- [12] Yin S, Gambe RG, Sun J, et al. A Murine Model of Chronic Lymphocytic Leukemia Based on B Cell-Restricted Expression of Sf3b1 Mutation and Atm Deletion. *Cancer Cell.* 2019;35(2):283-296.e5. doi:10.1016/j.ccell.2018.12.013
- [13] Bigot J, Lalanne AI, Lucibello F, et al. Splicing Patterns in SF3B1-Mutated Uveal Melanoma Generate Shared Immunogenic Tumor-Specific Neopeptides. *Cancer Discov.* 2021;11(8):1938-1951. doi:10.1158/2159-8290.CD-20-0555
- [14] Kahles A, Lehmann KV, Toussaint NC, et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell.* 2018; 34 (2):211-224.e6. doi:10.1016/j.ccell.2018.07.001
- [15] Zhang J, Manley JL. Misregulation of pre-mRNA alternative splicing in cancer. *Cancer Discov.* 2013; 3 (11):1228-1237. doi:10.1158/2159-8290.CD-13-0253
- [16] Liu Z, Yoshimi A, Wang J, et al. Mutations in the RNA Splicing Factor SF3B1 Promote Tumorigenesis through MYC Stabilization. *Cancer Discov.* 2020; 10 (6):806-821. doi:10.1158/2159-8290.CD-19-1330
- [17] Liu Z, Rabadan R. Computing the Role of Alternative Splicing in Cancer. *Trends Cancer.* 2021; 7(4):347-358. doi:10.1016/j.trecan.2020.12.015
- [18] Yin H, Wang J, Li H, et al. Extracellular matrix protein-1 secretory isoform promotes ovarian cancer through increasing alternative mRNA splicing and stemness. *Nat Commun.* 2021; 12 (1):4230. doi:10.1038/s41467-021-24315-1
- [19] Paschalis A, Sharp A, Welti JC, et al. Alternative splicing in prostate cancer. *Nat Rev Clin Oncol.* 2018; 15 (11):663-675. doi:10.1038/s41571-018-0085-0
- [20] Li X, Kim W, Arif M, et al. Discovery of Functional Alternatively Spliced PKM Transcripts in Human Cancers. *Cancers (Basel).* 2021;13 (2):348. doi:10.3390/cancers13020348
- [21] Luo X, Deng C, Liu F, et al. HnRNPL promotes Wilms tumor progression by regulating the p53 and Bcl2 pathways. *Onco Targets Ther.* 2019; 12: 4269-4279. doi:10.2147/OTT.S203046
- [22] Sciarrillo R, Wojtuszkiewicz A, Assaraf YG, et al. The role of alternative splicing in cancer: From oncogenesis to drug resistance. *Drug Resist Updat.* 2020; 53: 100728. doi:10.1016/j.drug. 2020.100728
- [23] Jia K, Wu Y, Huang J, Wu H. Survival-Associated Alternative Splicing Events in Pan-Renal Cell Carcinoma. *Front Oncol.* 2019; 9: 1317. doi:10.3389/fonc.2019.01317

- [24] Duan Y, Zhang D. Identification of novel prognostic alternative splicing signature in papillary renal cell carcinoma. *J Cell Biochem.* 2020; 121 (1):672-689. doi:10.1002/jcb.29314
- [25] Zhang Y, Wu X, Li J, et al. Comprehensive characterization of alternative splicing in renal cell carcinoma. *Brief Bioinform.* 2021;22 (5):bbab084. doi:10.1093/bib/bbab084
- [26] Frankiw L, Baltimore D, Li G. Alternative mRNA splicing in cancer immunotherapy. *Nat Rev Immunol.* 2019; 19 (11):675-687. doi: 10.1038/s41577-019-0195-7
- [27] Li ZX, Zheng ZQ, Wei ZH, et al. Comprehensive characterization of the alternative splicing landscape in head and neck squamous cell carcinoma reveals novel events associated with tumorigenesis and the immune microenvironment. *Theranostics.* 2019;9(25):7648-7665. Published 2019 Oct 14. doi:10.7150/thno.36585
- [28] Zhang J, Goel A, Zhu L. Identification of Novel Alternative Splicing Events Associated With Tumorigenesis, Protein Modification, and Immune Microenvironment in Early-Onset Gastric Cancer. *Front Oncol.* 2021; 11: 640272. doi:10.3389/fonc.2021.640272
- [29] Shi JY, Bi YY, Yu BF, Wang QF, Teng D, Wu DN. Alternative Splicing Events in Tumor Immune Infiltration in Colorectal Cancer. *Front Oncol.* 2021; 11: 583547. doi:10.3389/fonc.2021. 583547
- [30] Colaprico A, Silva TC, Olsen C, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 2016; 44 (8):e71. doi:10.1093/nar/gkv1507
- [31] Ryan Michael, Wong Wing Chung, Brown Robert et al. TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. [J]. *Nucleic Acids Res*, 2016, 44: D1018-22.
- [32] Xu Q, Xu H, Deng R, et al. Immunological significance of prognostic alternative splicing signature in hepatocellular carcinoma [published correction appears in *Cancer Cell Int.* 2021 Aug 24;21(1):448]. *Cancer Cell Int.* 2021;21(1):190. doi:10.1186/s12935-021-01894-z
- [33] Turajlic S, Xu H, Litchfield K, et al. Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell.* 2018;173(3):595-610.e11. doi:10.1016/j.cell.2018.03.043
- [34] Jonasch E, Walker CL, Rathmell WK. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat Rev Nephrol.* 2021; 17(4):245-261. doi:10.1038/s41581-020-00359-2
- [35] Ge Y, Weygant N, Qu D, et al. Alternative splice variants of DCLK1 mark cancer stem cells, promote self-renewal and drug-resistance, and can be targeted to inhibit tumorigenesis in kidney cancer. *Int J Cancer.* 2018; 143(5):1162-1175. doi:10.1002/ijc.31400
- [36] Jiang J, Chen X, Liu H, et al. Polypyrimidine Tract-Binding Protein 1 promotes proliferation, migration and invasion in clear-cell renal cell carcinoma by regulating alternative splicing of PKM. *Am J Cancer Res.* 2017;7(2):245-259. Published 2017 Feb 1.

- [37] Li X, Turanli B, Juszcak K, et al. Classification of clear cell renal cell carcinoma based on PKM alternative splicing. *Heliyon*. 2020; 6(2):e03440. doi:10.1016/j.heliyon.2020.e03440
- [38] Chen K, Xiao H, Zeng J, et al. Alternative Splicing of EZH2 pre-mRNA by SF3B3 Contributes to the Tumorigenic Potential of Renal Cancer. *Clin Cancer Res*. 2017;23(13):3428-3441. doi:10.1158/1078-0432.CCR-16-2020
- [39] Sun G, Zhou H, Chen K, et al. HnRNP A1-mediated alternative splicing of CCDC50 contributes to cancer progression of clear cell renal cell carcinoma via ZNF395. *J Exp Clin Cancer Res*. 2020;39(1):116. doi:10.1186/s13046-020-01606-x
- [40] Song J, Liu YD, Su J, Yuan D, Sun F, Zhu J. Systematic analysis of alternative splicing signature unveils prognostic predictor for kidney renal clear cell carcinoma. *J Cell Physiol*. 2019; 234(12):22753-22764. doi:10.1002/jcp.28840
- [41] Meng T, Huang R, Zeng Z, et al. Identification of Prognostic and Metastatic Alternative Splicing Signatures in Kidney Renal Clear Cell Carcinoma. *Front Bioeng Biotechnol*. 2019; 7:270. doi:10.3389/fbioe.2019.00270
- [42] Chen T, Zheng W, Chen J, et al. Systematic analysis of survival-associated alternative splicing signatures in clear cell renal cell carcinoma. *J Cell Biochem*. 2020; 121(10):4074-4084. doi:10.1002/jcb.29590
- [43] Dery KJ, Kujawski M, Grunert D, et al. IRF-1 regulates alternative mRNA splicing of carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) in breast epithelial cells generating an immunoreceptor tyrosine-based inhibition motif (ITIM) containing isoform. *Mol Cancer*. 2014;13:64. Published 2014 Mar 21. doi:10.1186/1476-4598-13-64
- [44] Sotillo E, Barrett DM, Black KL, et al. Convergence of Acquired Mutations and Alternative Splicing of CD19 Enables Resistance to CART-19 Immunotherapy. *Cancer Discov*. 2015;5(12):1282-1295. doi:10.1158/2159-8290.CD-15-1020
- [45] Georgilis A, Klotz S, Hanley CJ, et al. PTBP1-Mediated Alternative Splicing Regulates the Inflammatory Secretome and the Pro-tumorigenic Effects of Senescent Cells. *Cancer Cell*. 2018; 34(1):85-102.e9. doi:10.1016/j.ccell.2018.06.007
- [46] Xu Q, Xu H, Deng R, et al. Immunological significance of prognostic alternative splicing signature in hepatocellular carcinoma [published correction appears in *Cancer Cell Int*. 2021 Aug 24;21(1):448]. *Cancer Cell Int*. 2021; 21(1):190. doi:10.1186/s12935-021-01894-z
- [47] Wu Q, Zhang Y, An H, et al. The landscape and biological relevance of aberrant alternative splicing events in esophageal squamous cell carcinoma. *Oncogene*. 2021;40(24):4184-4197. doi:10.1038/s41388-021-01849-8

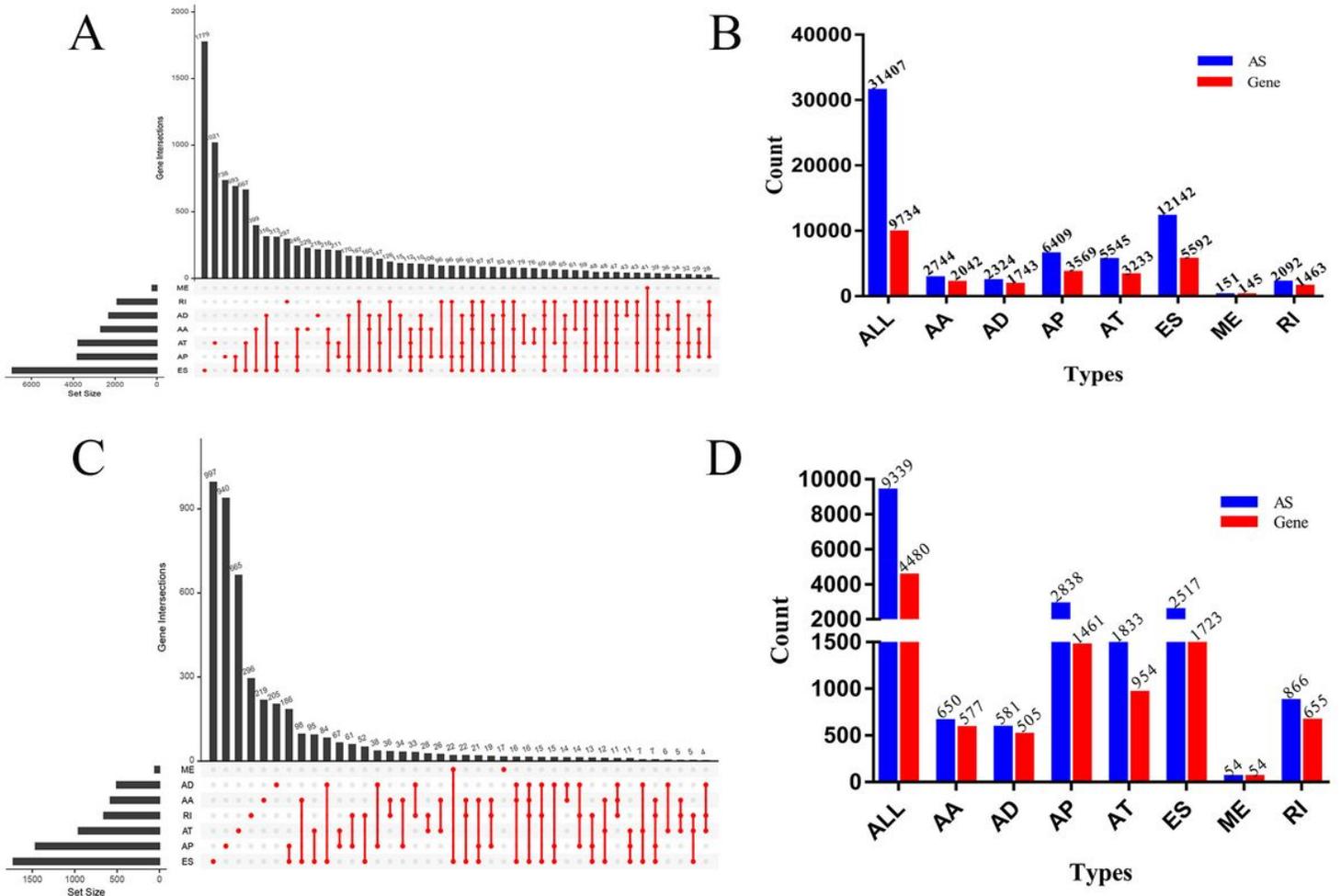
- [48] Lou S, Zhang J, Zhai Z, et al. The landscape of alternative splicing reveals novel events associated with tumorigenesis and the immune microenvironment in gastric cancer. *Aging* (Albany NY). 2021;13(3):4317-4334. doi:10.18632/aging.202393
- [49] Liu X, Liu C, Liu J, et al. Identification of Tumor Microenvironment-Related Alternative Splicing Events to Predict the Prognosis of Endometrial Cancer. *Front Oncol*. 2021; 11:645912. doi:10.3389/fonc.2021.645912
- [50] Yu S, Hu C, Liu L, et al. Comprehensive analysis and establishment of a prediction model of alternative splicing events reveal the prognostic predictor and immune microenvironment signatures in triple negative breast cancer. *J Transl Med*. 2020; 18(1):286. doi: 10.1186/s12967-020-02454-1
- [51] Wang M, Zhou Z, Zheng J, et al. Identification and Validation of a Prognostic Immune-Related Alternative Splicing Events Signature for Glioma. *Front Oncol*. 2021; 11:650153. doi: 10.3389/fonc.2021.650153
- [52] Braun DA, Street K, Burke KP, et al. Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell*. 2021;39(5):632-648.e8. doi:10.1016/j.ccell.2021.02.013
- [53] Bi K, He MX, Bakouny Z, et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell*. 2021; 39(5):649-661.e5. doi:10.1016/j.ccell.2021.02.015
- [54] Krishna C, DiNatale RG, Kuo F, et al. Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy. *Cancer Cell*. 2021;39(5):662-677.e6. doi:10.1016/j.ccell.2021.03.007

## Figures



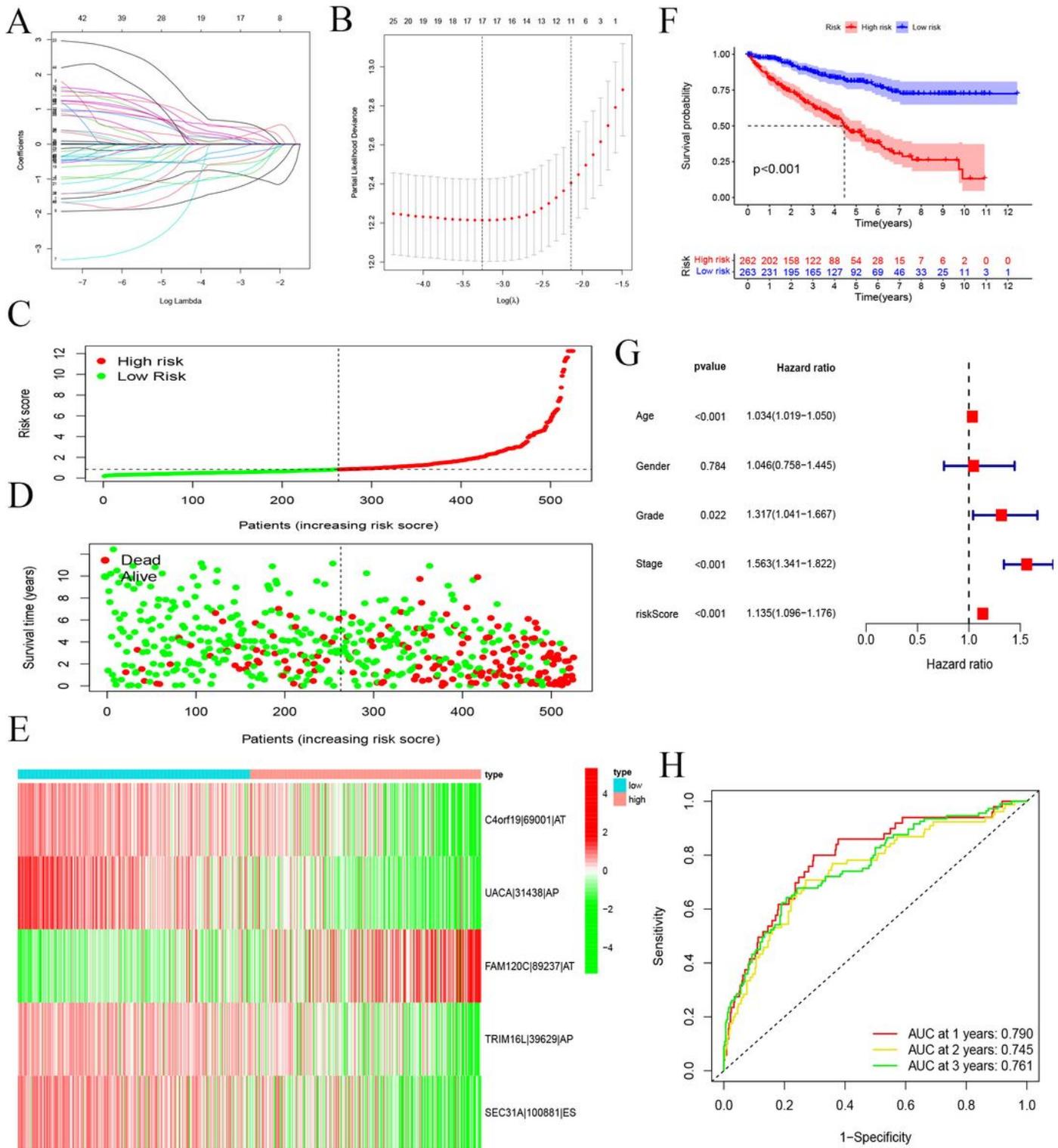
**Figure 1**

**Flow diagram of this study.**



**Figure 2**

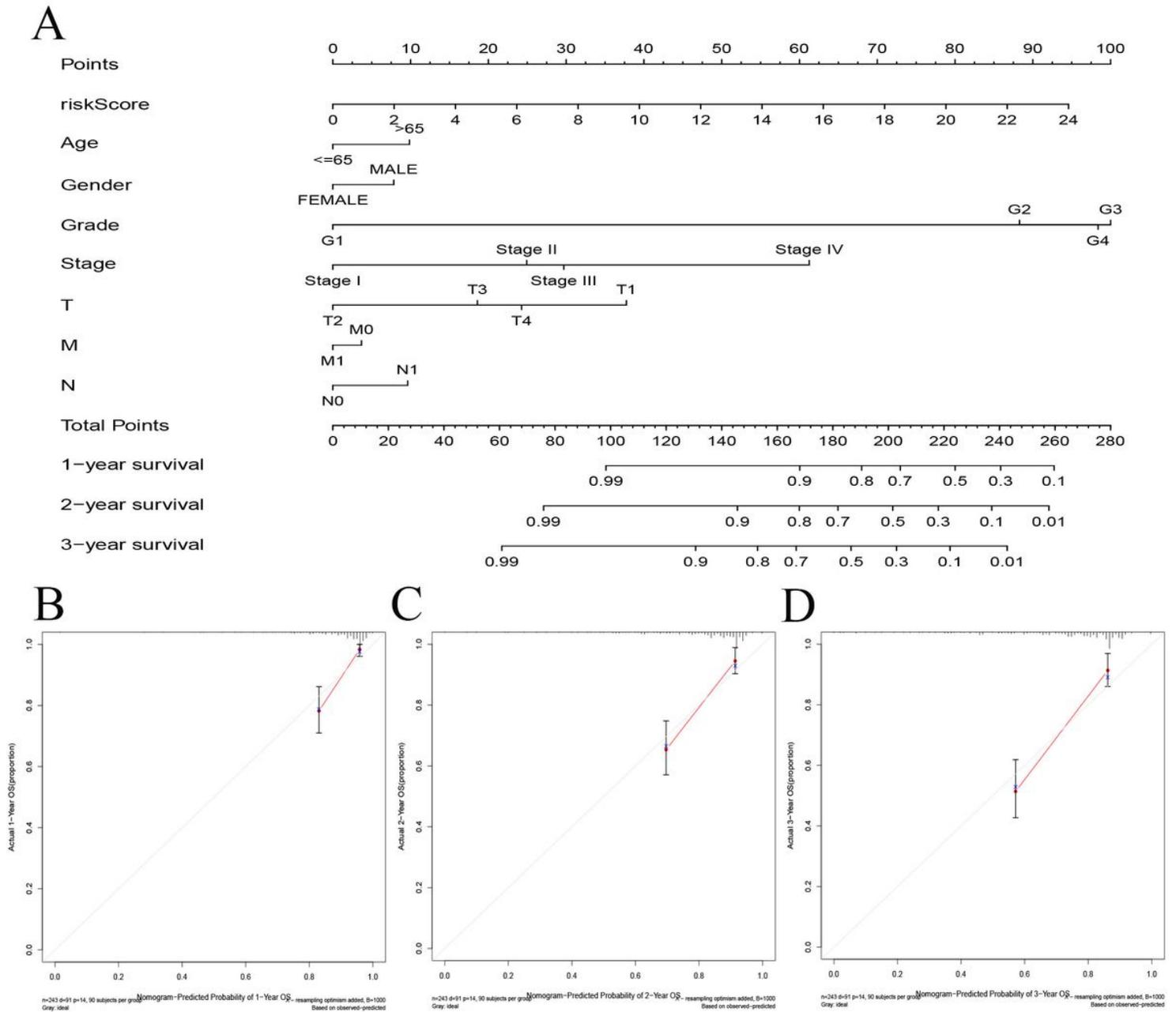
**Exhibition of the type and number of all AS events and OS-related AS events in ccRCC.** (A and C) The UpSet plot showed the top 50 interaction sets of all AS events and OS-related AS events, respectively. (B and D) The histograms displayed the number of events and parental genes of each AS type in all AS events set and OS-related AS events set, respectively, blue: AS events, red: parental gene, ALL: all 7 types, AA: alternate acceptor site, AD: alternate donor site, AP: alternate promoter, AT: alternate terminator, ES: exon skip, ME: mutually exclusive exons, and RI: retained intron.



**Figure 3**

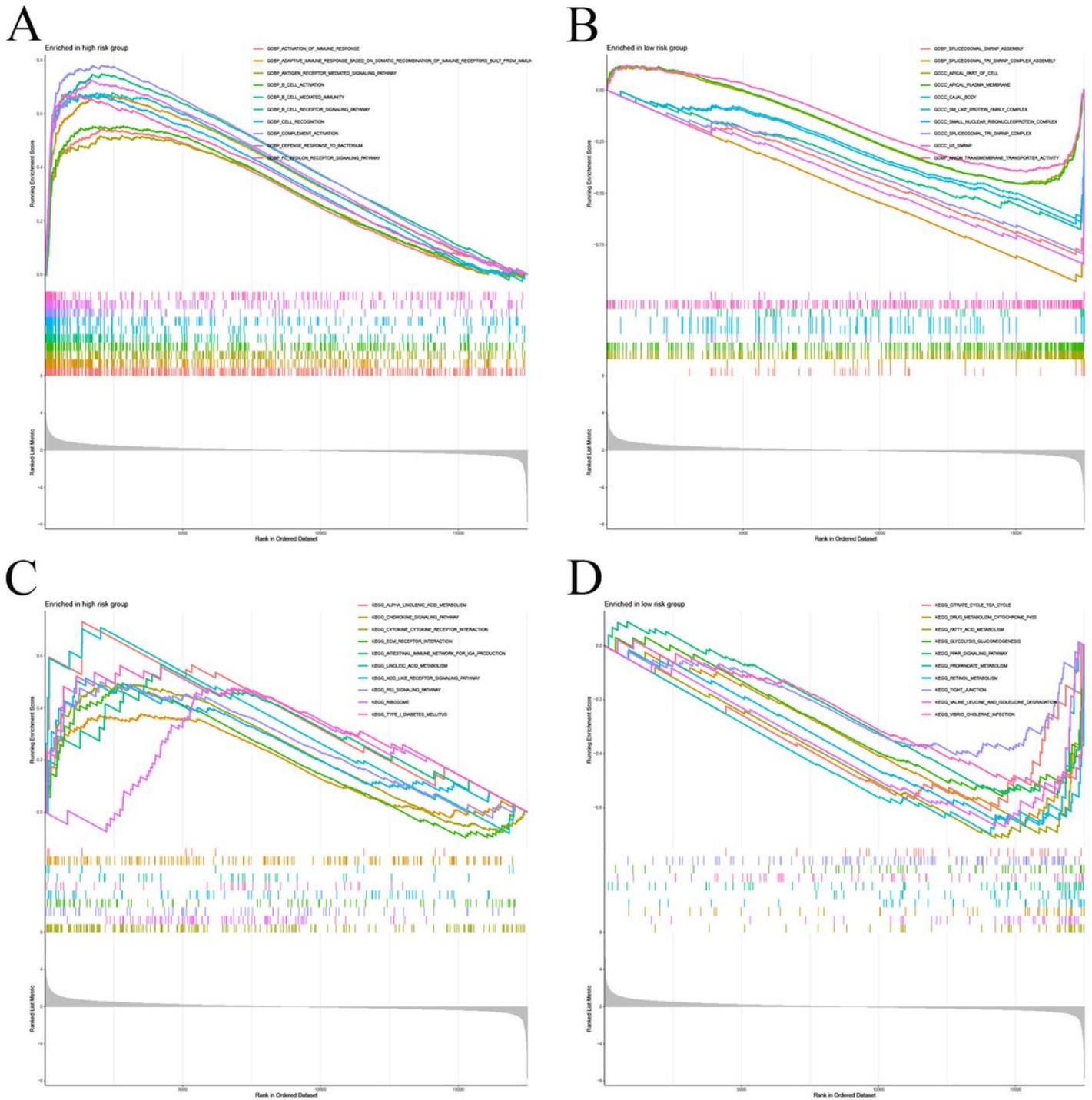
**Establishment and evaluation of the riskScore prognostic signatures about OS-related AS events in ccRCC.** (A and B) The lambda graph and the cross validation graph from the LASSO regression, respectively. (C) The riskScore of each ccRCC samples, red: high risk, green: low risk. (D) The survival state of each ccRCC patients with riskScore, red: dead, green: alive. (E) The heatmap of the PSI value of final OS-related AS events in high/low risk cohorts. (F) The Kaplan-Meier survival curve for ccRCC

samples with high/low riskScore, red: high risk, blue: low risk, ( $P < 0.001$ ). (G) The results of the multivariate independent survival analysis. (H) The ROC curves of prognostic models at 1, 2 and 3 years, red: 1 years, yellow: 2 years, and green: 3 years, ROC: receiver operator characteristic, AUC: area under the curve.



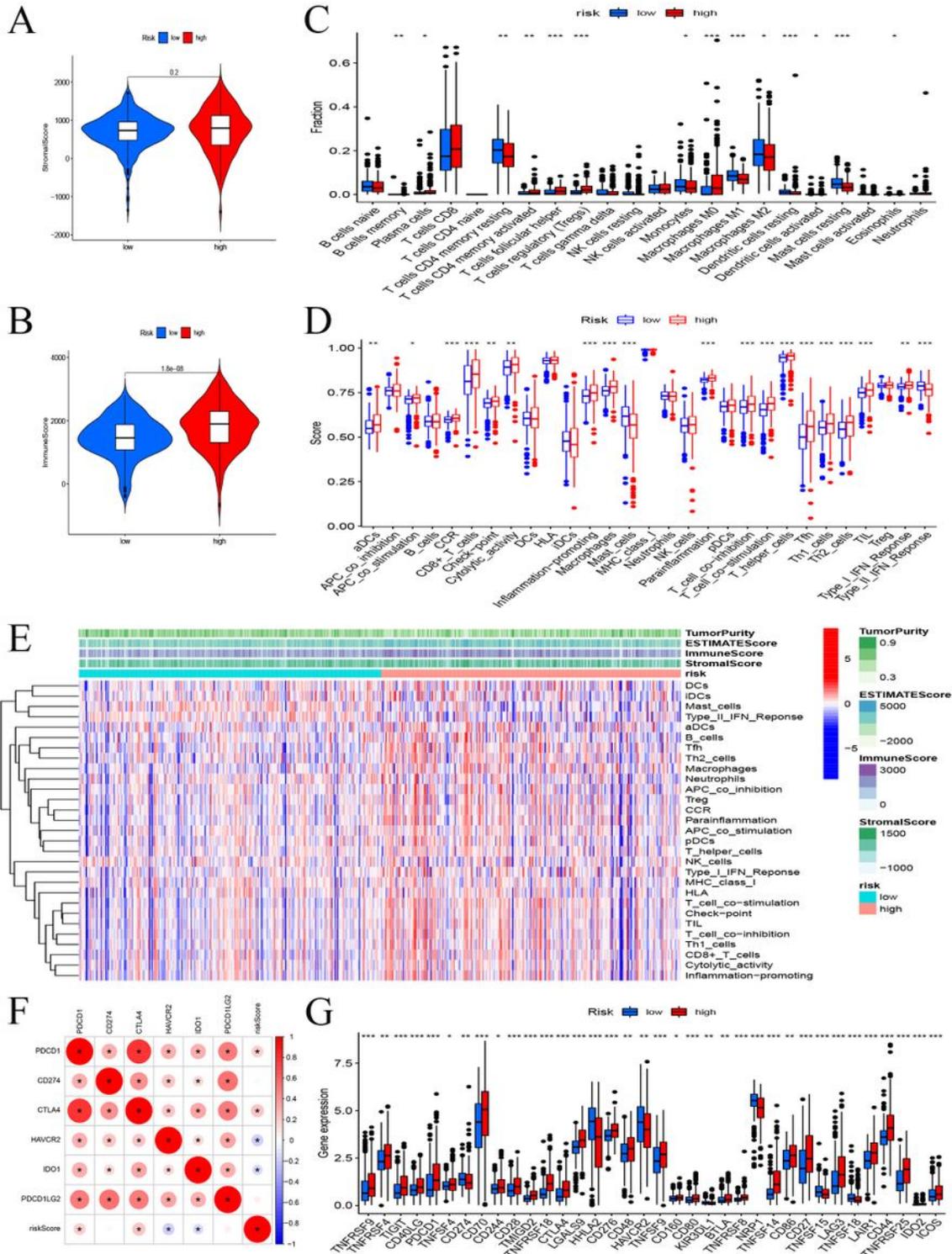
**Figure 4**

**The predictive model of overall survival of ccRCC patients.** (A) A nomogram for predicting OS of ccRCC patients. (B-D) The 1-, 2-, and 3-year OS calibration curves of the nomogram, respectively, red: real, gray: ideal.



**Figure 5**

The GO and KEGG analysis in ccRCC subgroups with high/low risk. (A, B) The top 10 GO enrichments in high/low risk subgroups, respectively. (C, D) The top 10 KEGG pathways enriched in high/low risk subgroups, respectively.



**Figure 6**

**The immune features in ccRCC cohorts with high/low risk.** (A) The relationship between high/low clusters and the stromal cell score, ( $P > 0.05$ ). (B) The relationship between high/low clusters and the immune cell score, ( $P < 0.05$ ). (C) The relationship between high/low clusters and the fractions of 22 types of immune cell infiltration. (D) The relationship between high/low clusters and the scores of immune cells and their functions by ssGSEA. (E) The heat map of the scores of immune cells and their function in each risk

sample. (F) The correlation between the riskScore and significant immune checkpoint. (G) The expression of immune checkpoint related genes in the risk clusters. \*: P<0.05, \*\*: P<0.01, \*\*\*: P<0.001.

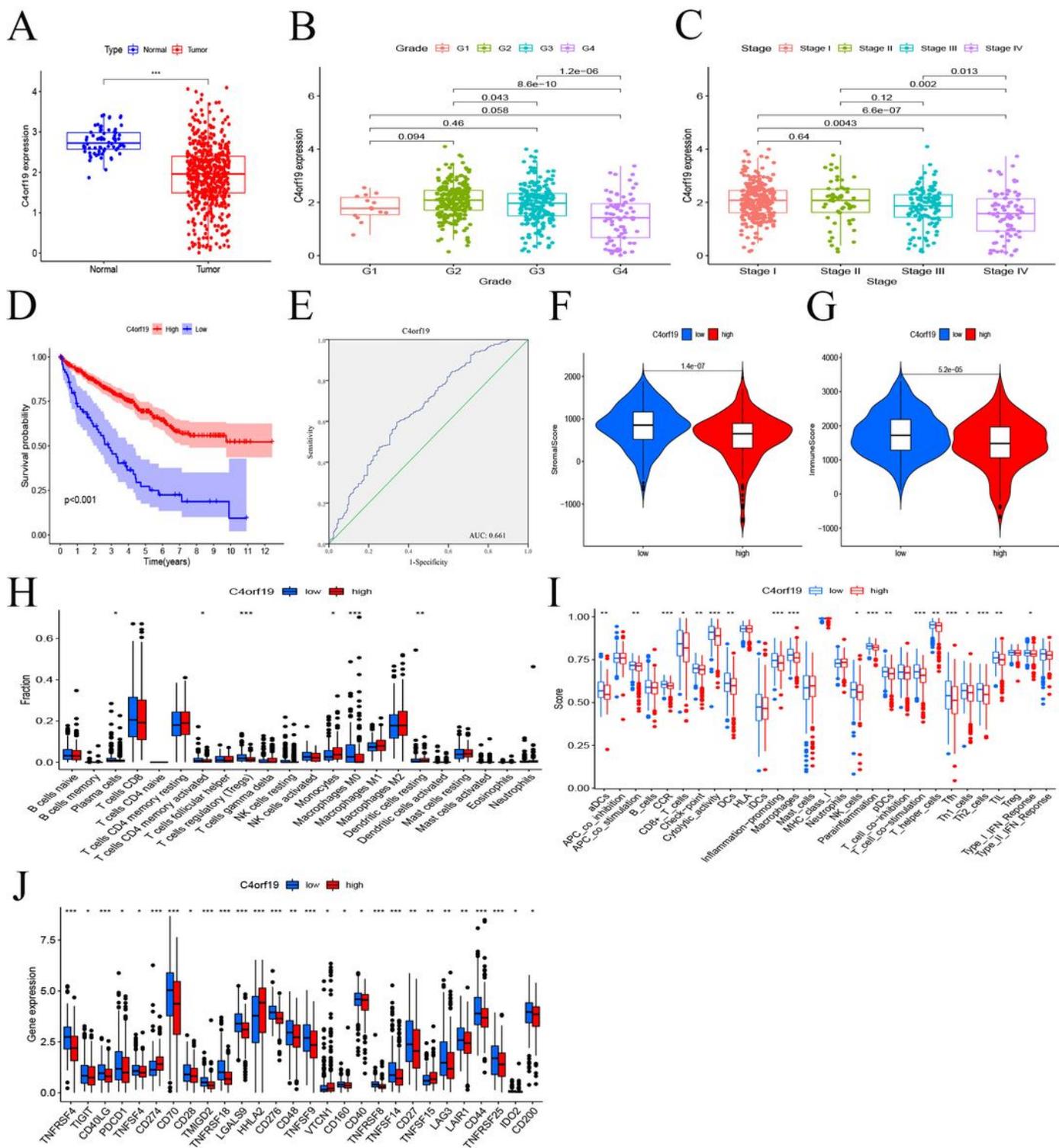


Figure 7

The clinical features and immune features of C4orf19 in ccRCC. (A) The expression of C4orf19 in ccRCC tumor samples and normal samples, red: tumor, blue: normal, (P<0.001). (B-C) The relationship between

the C4orf19 expression and grade and stage, respectively, ( $P < 0.05$ ). (D) The Kaplan-Meier survival curve for ccRCC samples with high/low expression of C4orf19, ( $P < 0.001$ ). (E) The ROC curves of C4orf19. (F-G) The stromal cell score and immune cell score in high/low clusters of C4orf19 expression, respectively, ( $P < 0.05$ ). (H) The fractions of 22 types of immune cell infiltration in high/low clusters of C4orf19 expression. (I) The scores of immune cells and their functions in high/low clusters of C4orf19 expression by ssGSEA. (J) The expression of immune checkpoint related genes in high/low clusters of C4orf19 expression. \*:  $P < 0.05$ , \*\*:  $P < 0.01$ , \*\*\*:  $P < 0.001$ .

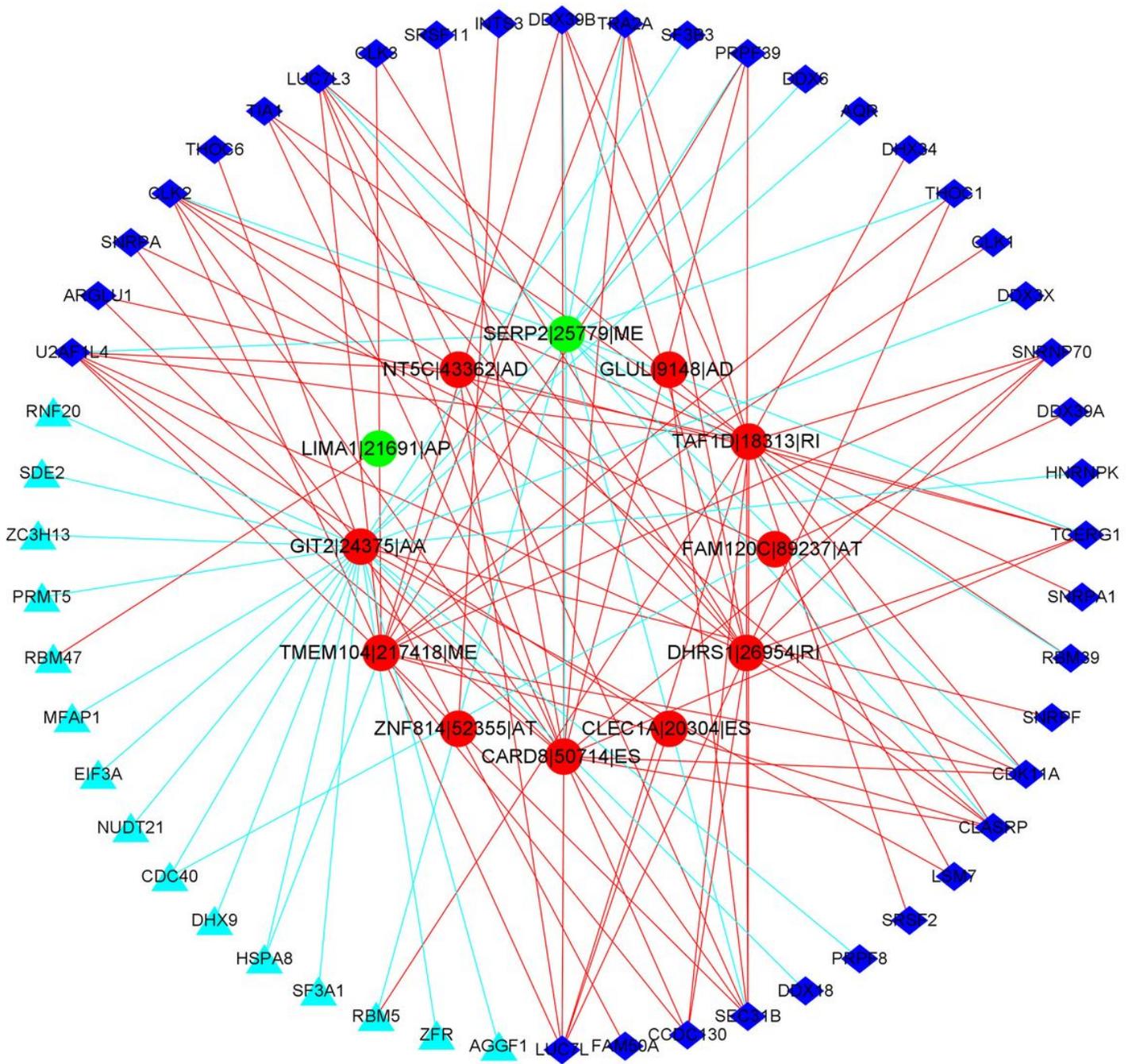


Figure 8

**The network of AS events and SFs.** Blue and rhombus: highly expressed SFs, azure and triangle: lowly expressed SFs, red and round: high-frequency AS events, lime and round: low-frequency AS events, red line: positive, azure line: negative.

## Supplementary Files

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

- [SupplementaryTable1.docx](#)
- [SupplementaryExcel1.xlsx](#)
- [SupplementaryExcel2.xlsx](#)
- [SupplementaryExcel3.xlsx](#)
- [SupplementaryFigure1.tif](#)
- [SupplementaryFigure2.tif](#)
- [SupplementaryFigure3.tif](#)
- [SupplementaryFigure4.tif](#)
- [SupplementaryFigure5.tif](#)
- [SupplementaryFigure6.tif](#)
- [SupplementaryFigure7.tif](#)
- [SupplementaryFigure8.tif](#)
- [SupplementaryFigure9.tif](#)
- [SupplementaryFigure10.tif](#)
- [SupplementaryFigureLegends.docx](#)