Am J Transl Res 2020;12(11):7096-7113
www.ajtr.org /ISSN:1943-8141/AJTR0121115

Original Article

Assessment of hazard immune-related genes and tumor immune infiltrations in renal cell carcinoma

Hongxi Chen[1], Jinliang Xie[2], Peng Jin[2]

[1]*Department of Gastrointestinal Surgery, Hunan Provincial People's Hospital (The First Affiliated Hospital of Hunan Normal University), Changsha 410005, Hunan Province, China;* [2]*Organ Transplant Center, Xiangya Hospital, Central South University, Changsha 410008, Hunan Province, China*

Received August 26, 2020; Accepted September 23, 2020; Epub November 15, 2020; Published November 30, 2020

Abstract: Background: The present study aimed to explore and validate a prognostic immune signature, to formulate a prognosis for ccRCC patients combined with immune-infiltration analysis. Methods: Public datasets were used as our source of multi-omics data. Differential analysis was performed via the edgeR package. A prognostic immune signature was identified by univariate Cox analysis, and we constructed an integrative tumor-associated immune genes (TAIG) model from the multivariate Cox results. In order to interrogate and identify the related crosstalk, functional analysis was deployed. Significantly, we implemented the CIBERSORT algorithm to estimate the immune cell fractions in the ccRCC samples, and analyzed the differential abundance of tumor-infiltrating immune cells in two TAIG groups, using a Wilcoxon rank-sum test. The prognostic role of differential immune cells was further assessed via a Kaplan-Meier analysis. In addition, we investigated the associations of a single immune signature with specific immune cells. Results: A total of 628 ccRCC patients were comprised in our integrative analysis, including 537 ccRCC patients in the discovery group and 91 patients in the validation group. Fourteen key immune signatures were subsequently identified. A figure of 0.802 was registered for AUC, and worse prognosis was evinced for those patients with a higher TAIG. Correlation analysis indicated that TAIG correlated closely with both clinical variables and TMB. Moreover, functional analysis implicated the immune-related GO items or crosstalk. Hence, we were able to identify the relationships obtaining between tumor-infiltrating immune cells and TAIG. The differential abundance of immune cells showed a significant prognostic difference and consisted of memory-activated $CD4^+$ T cells, T follicular helper cells, T regulatory cells, and so on. Moreover, we also characterized the associations between identified signatures and specific immune cells. Finally, the five-year AUC in the ICGC cohort was 0.72, suggesting the robustness of the TAIG that we constructed. Conclusions: Overall, our team characterized the tumor-associated immune signature in ccRCC, and further identified the prognostic tumor-infiltrating immune cells related to TAIG. This in turn provided a solid foundation for investigating individualized immunotherapy, as well as other relevant mechanisms.

Keywords: Immune signature, tumor-mutation burden, immune infiltrates, prognosis, ccRCC

Introduction

Cancer of the kidney may be defined as an expanding solid malignancy in the urology system. Survival outcomes are poor; mortality levels are high [1]. Generally speaking, clear-cell renal-cell carcinoma (ccRCC) is the most prevalent subtype of the condition, and represents almost 70% of cases [2, 3]. Despite relatively advanced strategies for detection or cancer management, the incidence of ccRCC has continued to increase, and the level of estimated cases in the United States rose to 73,820 in 2019 [4]. The typical surgical treatment for ccRCC patients, thus far, remains radical nephrectomy or laparoscopic partial nephrectomy. Nevertheless, since approximately 30% of cases inevitably progress into advanced pathological stages or tumor recurrence, ccRCC patients still suffer from poor overall survival prognosis. The current determinants of prognosis in ccRCC primarily comprise tumor size (T stage), pathological grades, and histological subtypes. Nevertheless, the fact that cases with clinically similar characteristics evince heterogeneous outcomes reflects the inadequacy of traditional prognostic methods. Hence, an increased emphasis is now warranted on the

identification of robust and stable biomarkers, with credible accuracy and sensitivity, that more comprehensively reflect the biological features of ccRCC.

Aberrant immune regulation has been clearly recognized as a vital component in tumor microenvironments [5, 6]. It is, for instance, pertinent to tumorigenesis [7], as well as progression, and even metastasis [8]. Recently, immunotherapies have been highlighted as promising potential avenues for tumor treatment, partly as a result of intensive research into immunity *per se.* Areas of interest include programmed death-1 (PD-1) [9-11] or programmed death ligand 1 (PD-L1) [9, 12] blockade. In particular, a large cohort has recently demonstrated that the combination of Avelumab (PD-L1 antibody) or Pembrolizumab (PD-1 antibody) with Axitinib evinced superior results, compared with Sunitinib alone [13, 14]. Nonetheless, the effectiveness of PD-1/PD-L1 inhibitors was shown to be less than total in the case of patients with a varying range of objective response rates. This reflected, *inter alia,* the fact that the tumor immune environment should be carefully characterized.

Previous studies have already reported that overall survival (OS) [15] and prognosis in patients could be influenced by fractions of the immune environment, including tumor-associated macrophage [16, 17], mast cells, or stromal cells. Nonetheless, less attention has been paid to immune gene biomarkers discovered via large samples. Meanwhile, the characterization of immune signatures in the tumor microenvironment was significant for our understanding of ccRCC. Moreover, the tumor-mutation burden (TMB) was also reported as an effective biomarker for discriminating the responsiveness of immunotherapy in patients [18, 19]. Conversely, across a range of malignancies, there were inconsistent findings regarding the association of TMB with immunotherapy promotion or better prognoses. Taken together, these factors shed light on the discussion regarding the potential associations of immune signatures in the tumor microenvironment with TMB, even in the case of immune-infiltrating cells.

In our study, we obtained transcriptome data and mutation profiles primarily from the Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) databases. This information was duly deployed in order to screen the most significant immune signatures in ccRCC, in terms of prognosis formation. We revealed the links between the identified tumor-associated immune genes (TAIG) and certain clinical features. Moreover, we further explored the relationships between TAIG, genomic alterations and TMB. In addition, we continued to characterize the differential immune infiltrates associated with TAIG, and the prognostic value of significant infiltration cells. The predictive reliability of the signature we identified was validated by recourse to another data set. Our team intended to explore the tumor-associated immune signature biomarkers in the microenvironment, in order to make predictions regarding prognosis or immunotherapy. Meanwhile, we also sought to provide a thorough characterization of the immune components within ccRCC. The latter included immune infiltrates, the tumor-mutation burden and, indeed, the potential connections between these elements.

Methods and materials

Data collection and pre-processing

The transcriptome expression data of the ccRCC samples were obtained from the TCGA database (https://portal.gdc.cancer.gov/) and the ICGC database (https://icgc.org/). Meanwhile, the somatic mutation data were downloaded from the “Masked Somatic Mutation” category in TCGA, and were analyzed via VarScan software. The edgeR package was used to conduct the normalization and differential analysis of the transcriptome profiles. The files in Mutation Annotation Format (MAF) were prepared with the maftools package [20]. The latter is frequently deployed in the analysis of cancer genomics and comprises customizable visualization functionality.

A list of 4,678 immune signatures was acquired from the InnateDB database (https://www.innatedb.ca/), which is a publicly available resource for immunity research. Furthermore, clinical features of age, gender, TNM stages, tumor grades, and follow-ups, with vital status, were obtained from the database via the TCGA biolinks package. Patients were subsequently excluded if they evinced insufficient clinical data.

Screening of the hub immune signature and construction of a Tumor Associated Immune Signature (TAIG) model

The edgeR package was utilized to obtain the differentially expressed genes (DEGs) in normal samples, versus tumors. Subsequently, a univariate Cox analysis was conducted using a survival package, with $P < 0.01$ to identify key prognostic immune genes. Then, we conducted a stepwise regression analysis to find the independent prognostic factors in the multivariate Cox method, with $P < 0.05$. The process of selection was illustrated in a Venn graph via the VennDiagram package.

The risk TAIG model, based on the hub immune signature, was thus constructed as follows: $TAIG = \sum(\beta_i * EXP_i)$, where β_i, the coefficients, signified the weight of each signature, and EXP_i represented the expression data. Hence, by deploying the median TAIG score as a cutoff factor, we were able to classify each patient in one of two groups. We assessed the differential cluster of the hub signature in a heatmap plot in the two groups via the pheatmap package. In addition, the three-year and five-year aspects of the ROC curve were illustrated via the timeROC package. This was done in order to assess the predictive value of TAIG in the context of an OS prediction. A Kaplan-Meier analysis was conducted to compare the differences of OS in two TAIG levels. Similarly, we analyzed the clinical value of TAIG in tumor recurrence or progression.

Prognostic analysis of hub immune signature and correlation with other clinical variables

We extracted the expression data of each identified signature and merged these with OS time and PFS time for 537 ccRCC patients within the TCGA cohort. To evaluate the prognostic discrepancy for each hub signature in either progression-free survival (PFS) or OS, a Kaplan-Meier analysis with log-rank test was utilized. Meanwhile, we explored the potential associations of TAIG with clinical variables, whereby a Wilcoxon rank-sum test was utilized to compare differential levels of TAIG between two groups. In the context of three or more groups, conversely, a Kruskal-Wallis analysis was deemed appropriate. Moreover, we conducted a univariate Cox regression analysis to determine the prognostic value of TAIG with other clinical characteristics, such as age, gender, AJCC-TNM stages, pathological stages or tumor grades. The N stage was discarded in subsequent analyses, since it comprised a substantial number of missing cases. To ensure the TAIG level maintained an independent risk factor compared with other clinical variables, we selected the more significant ones, and performed a multivariate Cox regression to assess the clinical significance of TAIG with $P < 0.05$.

Tumor-mutation burden and correlation analysis

Since we had obtained the mutation profiles within ccRCC, we wrote the Perl scripts based on the JAVA platform to extract the specific genomic alterations for each patient. Insertions, deletions, and cross-base substitutions were among the mutants detected. Moreover, we defined the TMB as follows: TMB = (total count of variants)/(the whole length of exons). We implemented the maftools package to exhibit the mutation profiles in ccRCC by waterfall plot. Then, a TMB score was calculated for each patient, and the association with TAIG was determined via a Pearson correlation analysis with an estimated *P* value. A Wilcoxon rank-sum test was also deployed to interrogate the differential TAIG distributions in the low- and high-TMB groups. Additionally, the prognostic value of TMB in OS or PFS was evaluated by a Kaplan-Meier analysis, as a supplementary analysis for TAIG.

Functional analysis, GSEA

Based on the differential and prognostic analysis in **Figure 1B**, we selected the 53 intersect genes to conduct the Gene Ontology (GO) analysis. The org.Hs.eg.db package was used to transfer the gene symbol with entrezIDs. Subsequently, to identify the significantly enriched GO items related to hub prognostic immune genes, we deployed the enrichplot, Profiler and ggplot2 packages. Given that we had classified the cohort into two groups, with high- and low-TAIG levels, we also conducted a GSEA between these two groups, using the TAIG as the phenotype. The GSEA software ran on the JAVA platform, and we obtained the "c2.cp.kegg.v6.2.symbols.gmt gene sets" from the MSigDB database (http://software.broadinstitute.org/gsea/msigdb). We deemed enriched pathways significant if they comprised FDR < 0.05.

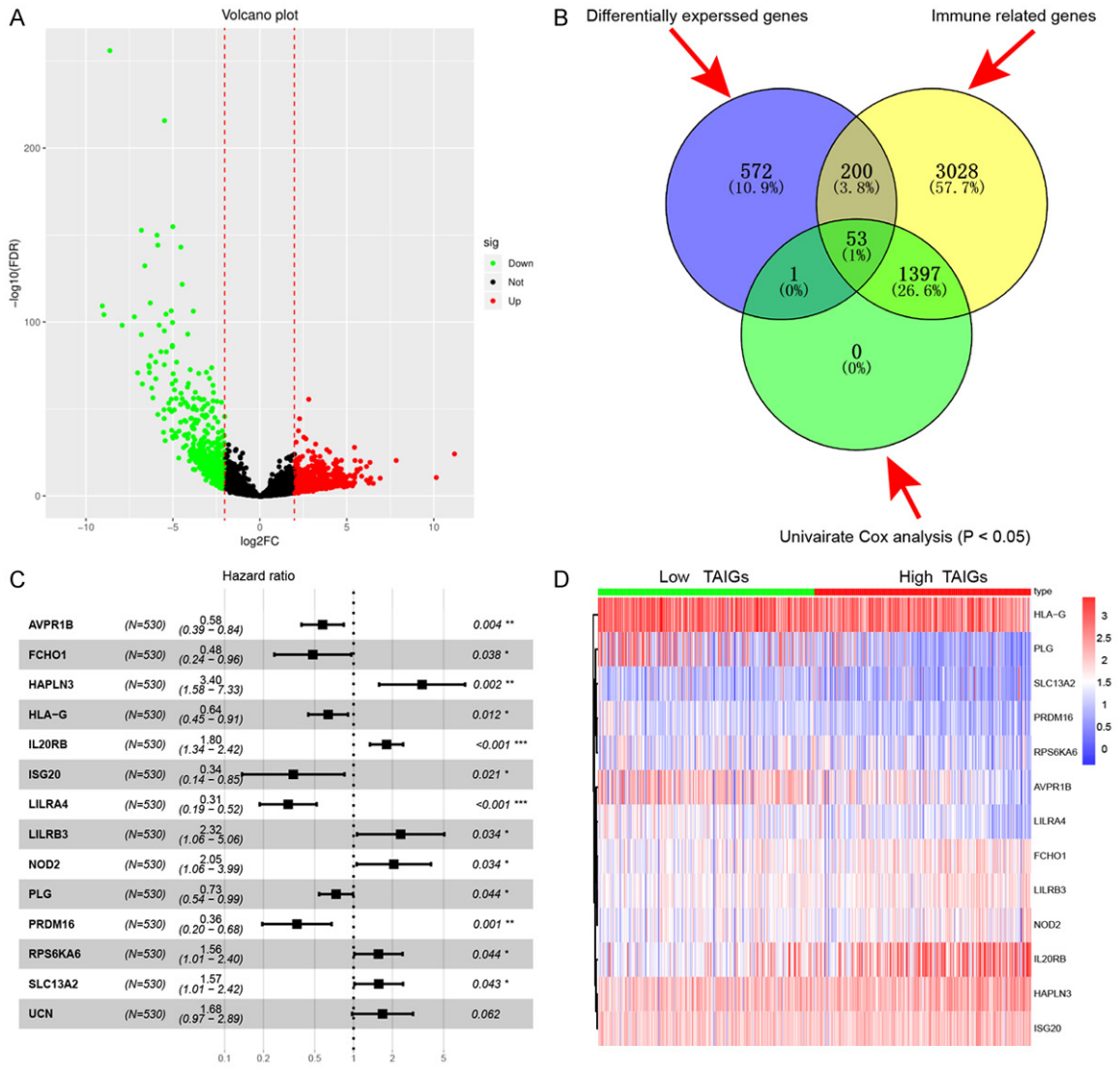


Figure 1. Identification of hub tumor-associated immune signature and construction of TAIG in ccRCC. A. Volcano plot representing the differentially expressed genes. B. The screening procedure by Venn diagram for identifying hub prognostic immune signature. C. Hazard ratios with 95% CI of each hub signature from the stepwise regression model illustrated in forest plot. D. Cluster analysis revealing the differential distributions of immune signature in two TAIG groups.

Estimation of tumor immune infiltrates through CIBERSORT or gene markers

CIBERSORT is a computational method developed to quantify cell fractions from bulk-tissue gene expression profiles. Flow cytometry with large tumor-biopsy samples provided good levels of validation. We used the normalized expression data with the FPKM format in the TCGA cohort and obtained the COBERSORT R package from a public website (https://cibersort.stanford.edu/). The inferred immune cell fractions were illustrated via boxplot. Moreover, a Wilcoxon rank-sum test was used to evaluate the differential abundance of immune infiltrates in low- or high-TAIG levels.

We additionally merged the immune fractions with survival information and performed a Kaplan-Meier analysis to discover whether the differentially distributed immune infiltrating cells in the two TAIG levels possessed prognostic values. Meanwhile, we sought to minimize statistical bias in our cell-composition analysis. We did so by exploiting another commonly recognized method to quantify the tumor-immune infiltrating cells, based on already identified marker genes listed via the GSVA

Table 1. Clinical features of patients included in this study

Variables	TCGA (N = 537)	ICGC (N = 91)
Age (Mean ± SD)	60.59 ± 12.14	60.47 ± 9.97
Follow-up (y)	3.12 ± 2.23	4.14 ± 1.73
Status		
Alive	367 (68.34)	61 (67.03)
Dead	170 (31.66)	30 (32.97)
Gender		
Male	346 (64.43)	52 (57.14)
Female	191 (35.57)	39 (42.86)
AJCC-T		
T1	275 (51.21)	54 (59.34)
T2	69 (12.85)	13 (14.28)
T3	182 (33.89)	22 (24.18)
T4	11 (2.05)	2 (2.20)
AJCC-N		
N0	240 (44.69)	79 (86.81)
N1	17 (3.17)	2 (2.20)
Unknow	280 (52.14)	10 (10.99)
AJCC-M		
M0	426 (79.33)	81 (89.01)
M1	79 (14.71)	9 (9.89)
Unknow	32 (5.96)	1 (1.10)
Pathological stage		
I	269 (50.09)	-
II	57 (10.61)	-
III	125 (23.28)	-
IV	83 (15.46)	-
Unknow	3 (0.56)	-
Grade		
G1	14 (2.61)	-
G2	230 (42.83)	-
G3	207 (38.54)	-
G4	78 (14.53)	-
Unknow	8 (1.49)	-
TAIG score		
Low	265 (49.35)	46 (50.55)
High	265 (49.35)	45 (49.45)
Unknow	7 (1.30)	0

Data are shown as n (%). Abbreviations: TCGA, The Cancer Genome Atlas; ICGC, International Cancer Genome Consortium; AJCC, American Joint Committee on Cancer.

package. In light of the matrix of immune cells and the expression data of the hub immune signature, we further discussed the specific associations of a single gene with tumor-immune infiltrating cells. We deployed dotplot to illustrate the Spearman correlation analysis, together with the estimated *P* value.

Validation of TAIG in another independent cohort

We had analyzed the clinical significance of TAIG and the association with TMB or immune infiltrates. We therefore set out to use an independent data set to illustrate the robustness of the model. From the ICGC database, we acquired 91 ccRCC patients with complete survival information and transcriptome sequencing data. We extracted the expression profiles of the identified signature in TCGA and implemented the ROC curve to assess the predictive power of the TAIG model in the ICGC cohort. A Kaplan-Meier analysis with log-rank test was also used to compare the varying survival outcomes between the two TAIG levels within the ICGC cohort.

Statistical analysis

The Cox regression models and Kaplan-Meier analyses were conducted via the survival package. The student t-test was used for continuous variables, while categorical variables were dealt with via a Chi-square (χ^2) test. A Wilcoxon rank-sum test was utilized to compare ranked data. Meanwhile, comparisons between three groups or more were undertaken via a Kruskal-Wallis test. R Studio (Version 3.5.2) was used for all statistical analysis; only $P < 0.05$ was seen as providing statistical significance.

Results

Construction and assessment of TAIG in ccRCC patients

We obtained a total of 537 ccRCC patients from the TCGA database with transcriptome profiles, together with the mutation data of 336 patients. The clinical baseline was summarized in **Table 1**. Differential analysis based on the edgeR package revealed a list of 826 DEGs, and we acquired 253 intersect immune signatures from the InnateDB database. In order to locate the hub 53 prognostic immune signature (**Figure 1A**, **1B**), a univariate Cox analysis was deployed. Further to screen the independent factors, a stepwise regression method was utilized, and we finally obtained the 14-hub tumor-associated immune signature. The latter was illustrated via forest plot, complete with corresponding 95% confidence interval (CI) and hazard ratio (**Figure 1C**). The TAIG was thus

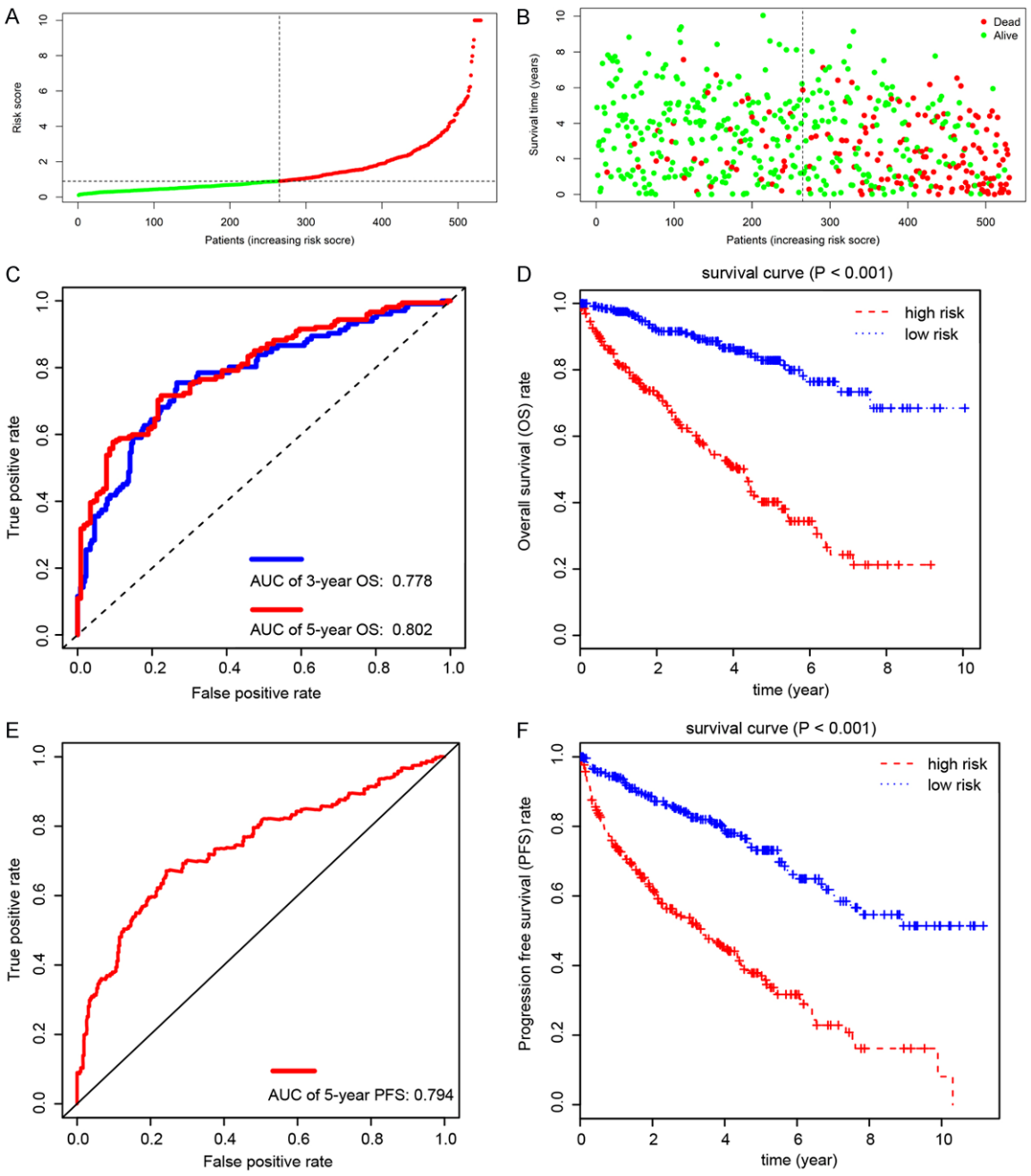


Figure 2. Prognostic assessment of TAIG in ccRCC. A, B. The median cutoff of TAIG and the distributions of vital status according to TAIG scores. C. ROC curve conducted to show the power of 3-year or 5-year OS prediction. D. Kaplan-Meier analysis of ccRCC patients in two TAIG groups. E. ROC curve performed to show the power of 5-year PFS prediction. F. The ccRCC patients with high TAIG correlated with more hazards in tumor progression or recurrence.

established from the multivariate Cox analysis, as follows: TAIG = (-0.5526*AVPR1B - 0.7276*FCHO1 + 1.2240*HAPLN3 - 0.4534*HLA-G + 0.5895*IL20RB - 1.0766*ISG20 - 1.1680*LILRA4 + 0.8415*LILRB3 + 0.7190*NOD2 - 0.3117*PLG - 1.0108*PRDM16 + 0.4443*RPS6KA6 + 0.4492*SLC13A2 + 0.5179*UCN).

We were able to observe the differentially expressed levels of signature in high- and low-TAIG groups via heatmap (**Figure 1D**). The fact that higher-TAIG patients evinced elevated mortality levels was intuitively illustrated by the distribution plot (**Figure 2A**, **2B**). Moreover, the AUC in three-year OS predictions was 0.778,

and that in five-year OS predictions was 0.802. This was indicative of superior predictive capacity (**Figure 2C**). Correspondingly, the log-rank test showed patients with higher TAIG were at greater risk in the context of OS, with $P < 0.001$ (**Figure 2D**). Furthermore, the TAIG also showed better predictive value in tumor progression, and the AUC of five-year PFS was 0.794 (**Figure 2E**).

Simultaneously, with $P < 0.001$ (**Figure 2F**), higher TAIG constituted a notably significant indicating factor for tumor recurrence or progression. Given the obvious clinical significance of TAIG, we specifically continued to explore the prognostic value of the single hub gene in OS or PFS, and the log-rank test showed that nearly all hub genes were closely associated with OS or PFS in ccRCC patients (**Figures 3**, S1).

Association of TAIG with other clinical variables

Considering the potential clinical significance of TAIG in ccRCC, we wanted to clarify the correlation of TAIG with other traditional clinical features, including tumor grades, pathological stages or TNM stages. Initially, TAIG was merged with other variables; a Cox analysis was subsequently undertaken. Then, the univariate Cox analysis indicated that age ($P < 0.001$), tumor grade ($P < 0.001$), T stage ($P < 0.001$), M stage ($P < 0.001$) and TAIG score ($P < 0.001$) were all risk factors. Nonetheless, the TAIG ($P < 0.001$), tumor grade ($P = 0.003$) or pathological stage ($P = 0.016$) still retained a robust significance in the multivariate Cox regression analysis. By contrast, the T and M stages evinced no statistical differences (**Table 2**). In addition to using the Cox analysis to demonstrate the prognostic role of TAIG, we also investigated the underlying relationships of TAIG with other variables. The correlation analysis suggested that higher TAIG correlated strongly with advanced T stage ($P = 4.032e\text{-}17$), N stage ($P = 2.351e\text{-}04$), metastasis ($P = 8.708e\text{-}12$), pathological stage ($P = 4.08e\text{-}19$), and higher tumor grades ($P = 7.03e\text{-}15$) (**Figure 4A-E**).

TAIG correlated positively with TMB, indicating poor prognostic outcomes

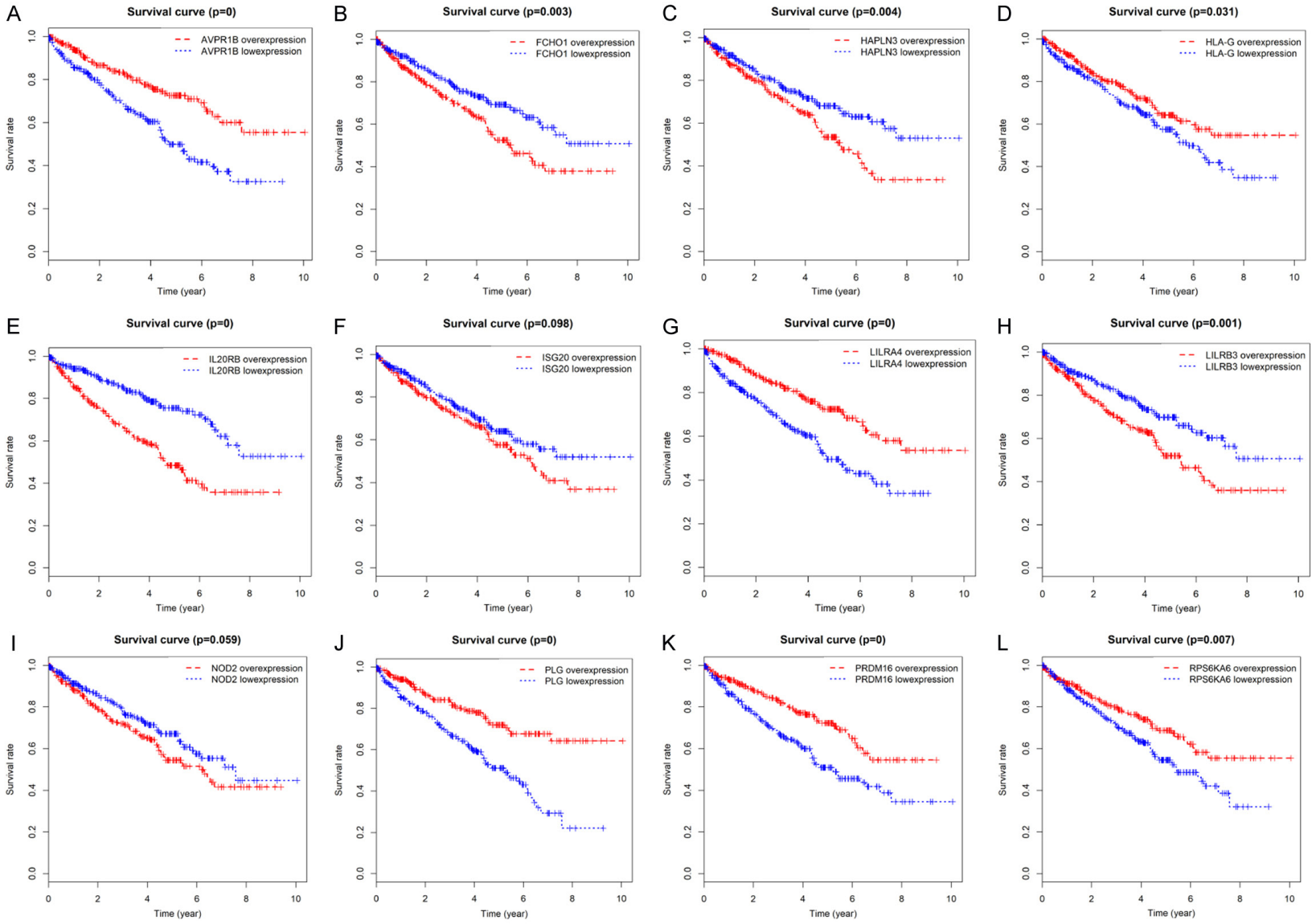
We noted that the tumor-mutation burden was reported to be closely associated with immunotherapeutic response and tumor prognosis. Thus, we planned to interrogate the correlation between the identified immune signature and TMB. We illustrated the mutation profiling of ccRCC via waterfall plot, in which the different colors annotated at the bottom showed the various mutation types. Meanwhile, above the legend, the calculated TMB for each sample was presented (**Figure 5A**). In order to perform the correlation analysis, we integrated the matched TAIG with the TMB score, since we had only extracted mutation data for 336 patients. A Wilcoxon rank-sum test illustrated the higher TMB levels in the high-TAIG group (**Figure 5B**), and the Pearson correlation analysis provided supplementary proof with $r = 0.188$ and $P = 0.001$ (**Figure 5C**). Furthermore, we discovered the prognostic value of TMB, and found that higher TMB was associated with poor OS outcomes ($P = 0.035$). It also correlated with greater risk in terms of tumor recurrence and progression ($P = 0.01$) (**Figure 5D**, **5E**).

Immune-related GO items or crosstalk associated with immune signature and TAIG phenotype

We proceeded to implement a GO enrichment analysis (**Figure 6A**), since we had already generated a list comprising 53 prognostic differential immune signatures (**Figure 1B**). In the biological process group, immune DEGs were mainly enriched in the regulation of cell-cell adhesion or T-cell activation. In the cellular component category, these genes were associated with an extracellular matrix. Conversely, in the molecular-function group, these genes principally contributed to receptor-ligand activity, cytokine activity and glycosaminoglycan binding. Moreover, gene-set enrichment analysis, undertaken to compare immune phenotypes between the high- and low-TAIG groups, indicated that higher TAIG was associated with regulation of the chemokine signaling pathway, the VEGF signaling pathway or the T-cell-receptor signaling pathway. Meanwhile, it correlated *negatively* with lysine degradation, the PPAR signaling pathway and the TGF-β signaling pathway (**Figure 6B**).

TAIG correlated with several prognostic immune infiltrating cells

Given the potential relationships of TAIG with TMB or several instances of immune-related


A
Survival curve (p=0)
AVPR1B overexpression
AVPR1B lowexpression
B
Survival curve (p=0.003)
FCHO1 overexpression
FCHO1 lowexpression
C
Survival curve (p=0.004)
HAPLN3 overexpression
HAPLN3 lowexpression
D
Survival curve (p=0.031)
HLA-G overexpression
HLA-G lowexpression
E
Survival curve (p=0)
IL20RB overexpression
IL20RB lowexpression
F
Survival curve (p=0.098)
ISG20 overexpression
ISG20 lowexpression
G
Survival curve (p=0)
LILRA4 overexpression
LILRA4 lowexpression
H
Survival curve (p=0.001)
LILRB3 overexpression
LILRB3 lowexpression
I
Survival curve (p=0.059)
NOD2 overexpression
NOD2 lowexpression
J
Survival curve (p=0)
PLG overexpression
PLG lowexpression
K
Survival curve (p=0)
PRDM16 overexpression
PRDM16 lowexpression
L
Survival curve (p=0.007)
RPS6KA6 overexpression
RPS6KA6 lowexpression
Survival rate
Time (year)

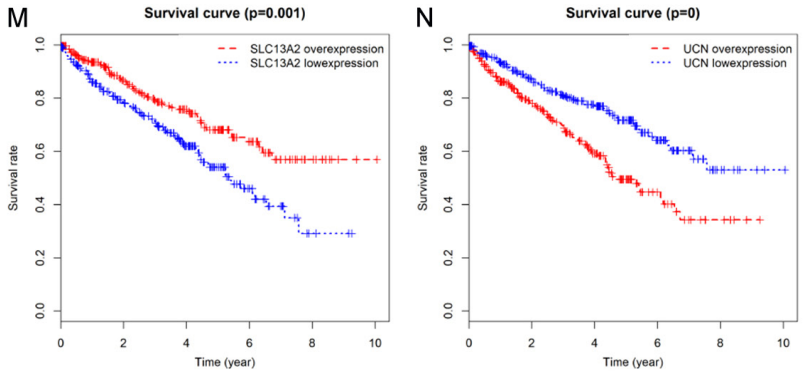


Figure 3. Kaplan-Meier analysis of 14 hub immune signature in ccRCC.

Table 2. Univariate and multivariate Cox analysis for TAIG scores and other clinical characteristics in TCGA cohort

Variables	Univariate Cox regression				Multivariate Cox regression			
	Hazard ratio	95% confidence interval		*P* value	Hazard ratio	95% confidence interval		*P* value
Age	1.033	1.019	1.047	< 0.001	1.035	1.019	1.050	< 0.001
Gender	0.931	0.675	1.284	0.663	-	-	-	-
Tumor grade	2.293	1.854	2.836	< 0.001	1.436	1.126	1.829	0.003
Pathological stage	1.889	1.649	2.164	< 0.001	1.741	1.110	2.730	0.016
T stage	1.941	1.639	2.299	< 0.001	0.852	0.563	1.290	0.449
M stage	4.284	3.106	5.908	< 0.001	1.121	0.570	2.204	0.740
TAIG score	1.107	1.085	1.130	< 0.001	1.060	1.034	1.086	< 0.001

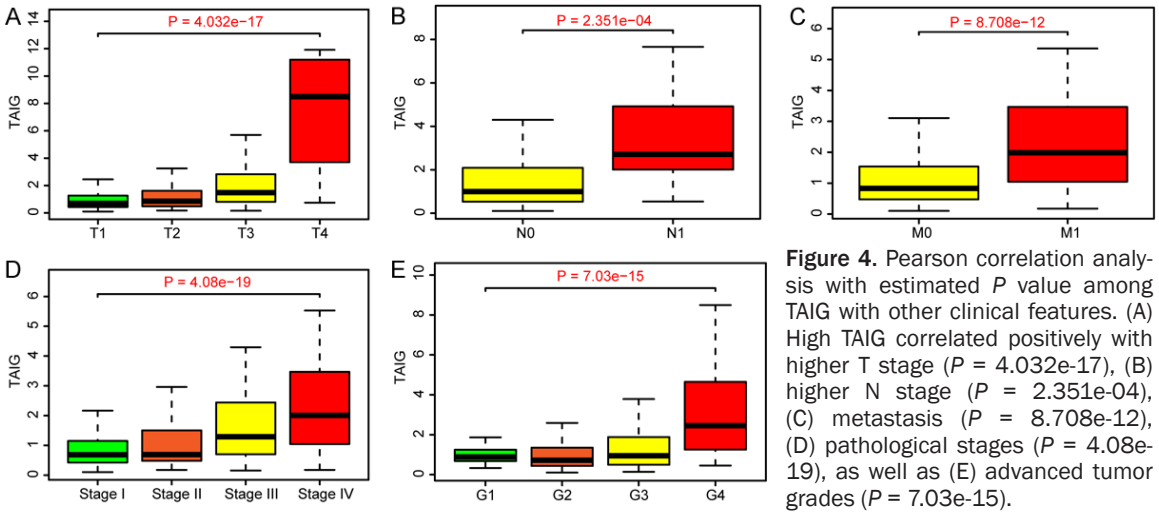


Figure 4. Pearson correlation analysis with estimated *P* value among TAIG with other clinical features. (A) High TAIG correlated positively with higher T stage (*P* = 4.032e-17), (B) higher N stage (*P* = 2.351e-04), (C) metastasis (*P* = 8.708e-12), (D) pathological stages (*P* = 4.08e-19), as well as (E) advanced tumor grades (*P* = 7.03e-15).

crosstalk, we considered whether TAIG influenced the density levels of tumor-infiltration cells, and whether this constituted an important role in the microenvironment. Based on signature expression data from ccRCC patients and the CIBERSORT algorithm, we estimated the specific fractions of 22 immune cells in each sample, as illustrated in Figure S2, where the sum of various immune types in boxplot equaled 100%. A Wilcoxon rank-sum test was subsequently deployed to illustrate the differential distributions of several immune cells within the two TAIG groups. A significant difference was found, in terms of a higher abundance of memory-activated $CD4^+$ T cells (*P* = 0.002), T follicular helper cells (*P* = 0.002), T regulatory cells (*P* < 0.001), and M0 macrophages (*P* < 0.001) in the high-TAIG group. Conversely, M2 macrophages (*P* = 0.035), resting dendritic cells (*P* < 0.001), and resting mast cells (*P* < 0.001) showed lower infiltrating levels in the high-TAIG group (**Figure 7A**). In line with the previous findings, it was interesting to observe that the Kaplan-Meier analysis illustrated prognostic significance for the majority of these differentially distributed immune-infiltration cells, in the two TAIG groups. Within this context, higher levels of memory-activated $CD4^+$ T cells (*P* = 0.022), T follicular helper cells (*P* = 0.003), T regulatory cells (*P* = 0.004) or M0 macrophages (*P* = 0.029) correlated with poor OS outcomes. Conversely, the role of tumor suppressor in prognosis terms may be taken by resting mast cells (*P* < 0.001) with resting dendritic cells (*P* = 0.002) (**Figure 7B-H**).

An alternate algorithm was utilized to enumerate the tumor-infiltrating immune cells from transcriptomics data, constructed on the marker genes. Subsequently, the output immune cell matrix was assimilated with extracted specific signature expression. Pearson correla-

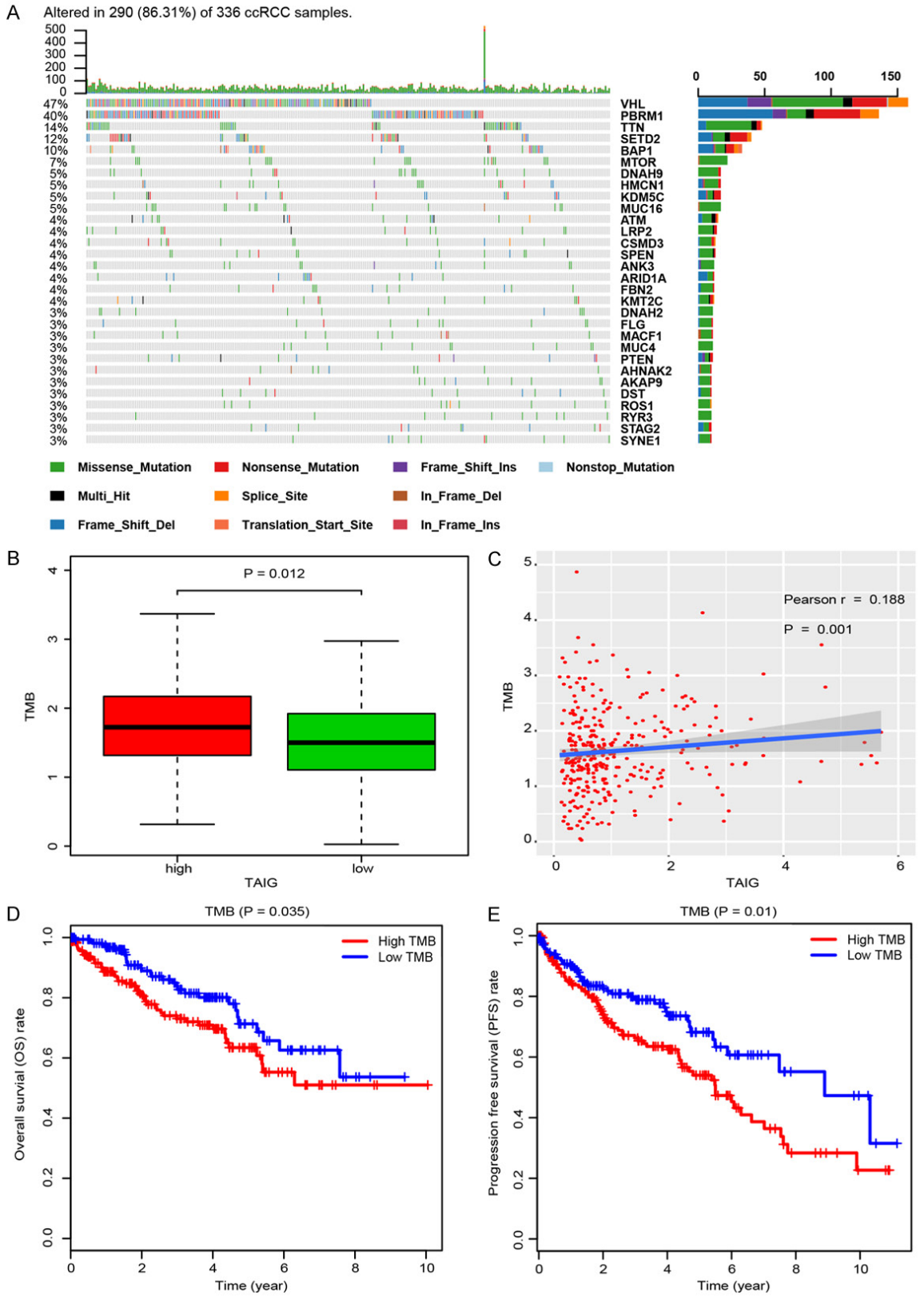


Figure 5. Landscape of mutation profiles in ccRCC and correlation with TAIG. A. Mutation profiling illustrated in the waterfall plot, where various colors with corresponding annotations represented the different mutation types. The barplot above the legend exhibited the mutation burden. B. TMB levels were significantly high in high-TAIG group by

Wilcoxon rank-sum test. C. Correlation analysis showed the associations between TMB and TAIG with $P = 0.001$. D, E. Prognostic analysis showed the high TMB correlated with poor OS outcomes ($P = 0.035$) and high risk in tumor recurrence ($P = 0.01$).

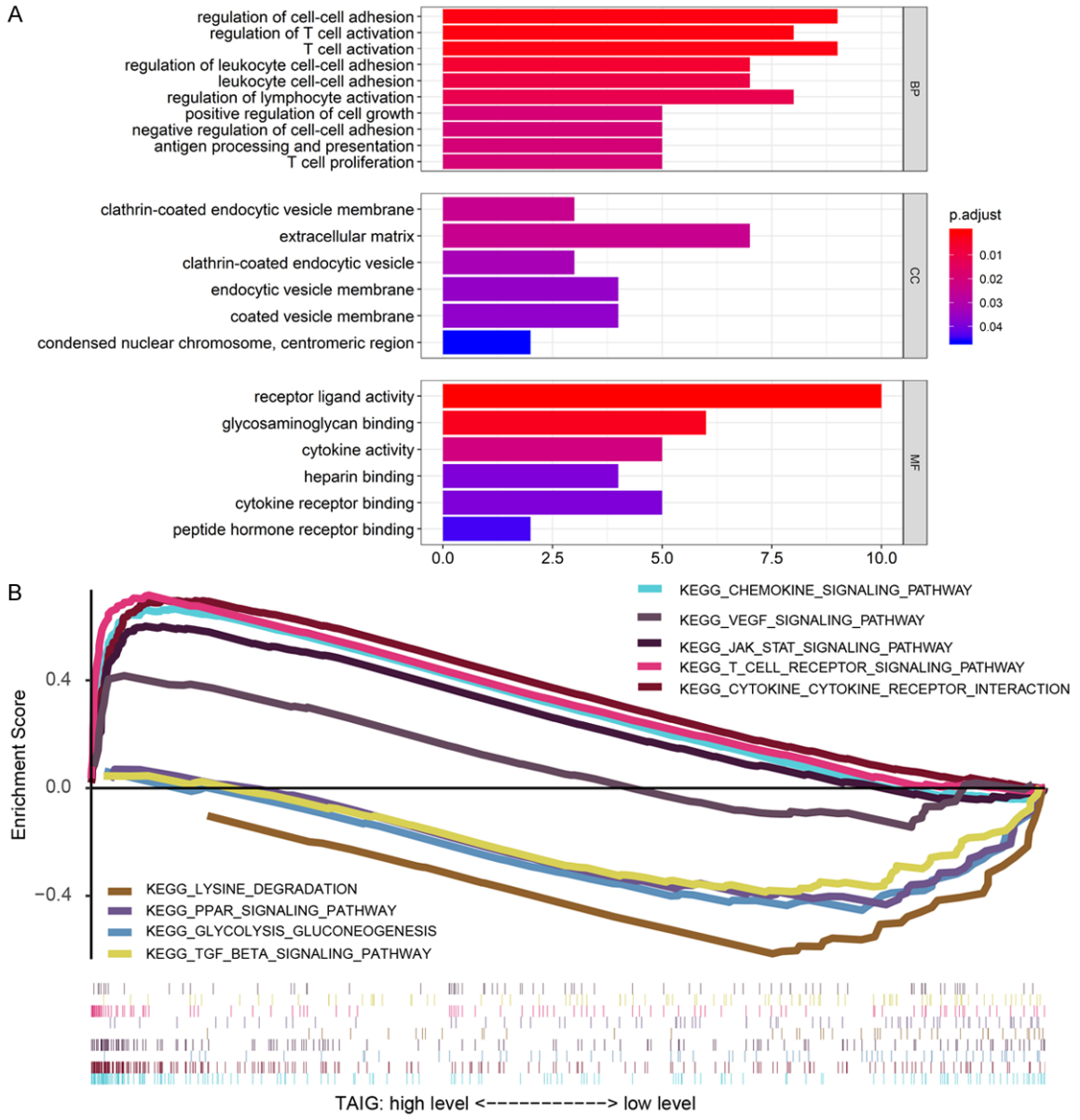


Figure 6. Functional enrichment analysis. A. Enriched GO items for 53 prognostic immune genes associated with survival in three groups consisted of biological process (BP), cellular components (CC), and molecular function (MF) categories. B. GSEA conducted using the TAIG as the phenotype suggesting the upregulated or downregulated crosstalk associated with TAIG.

tion analysis, with estimated *P* value depicted the statistical affiliations of single signature, with specific immune infiltrating cells, bestowing an additional solid substantiation between our identified signature with the tumor immune cells (**Figure 8**).

Validation of the robust TAIG in ICGC

In ICGC cohort consisting of 91 ccRCC samples, we additionally validated the robust signature employing the Cox regression method, with the AUC of ROC curve being 0.72, inferring the sta-

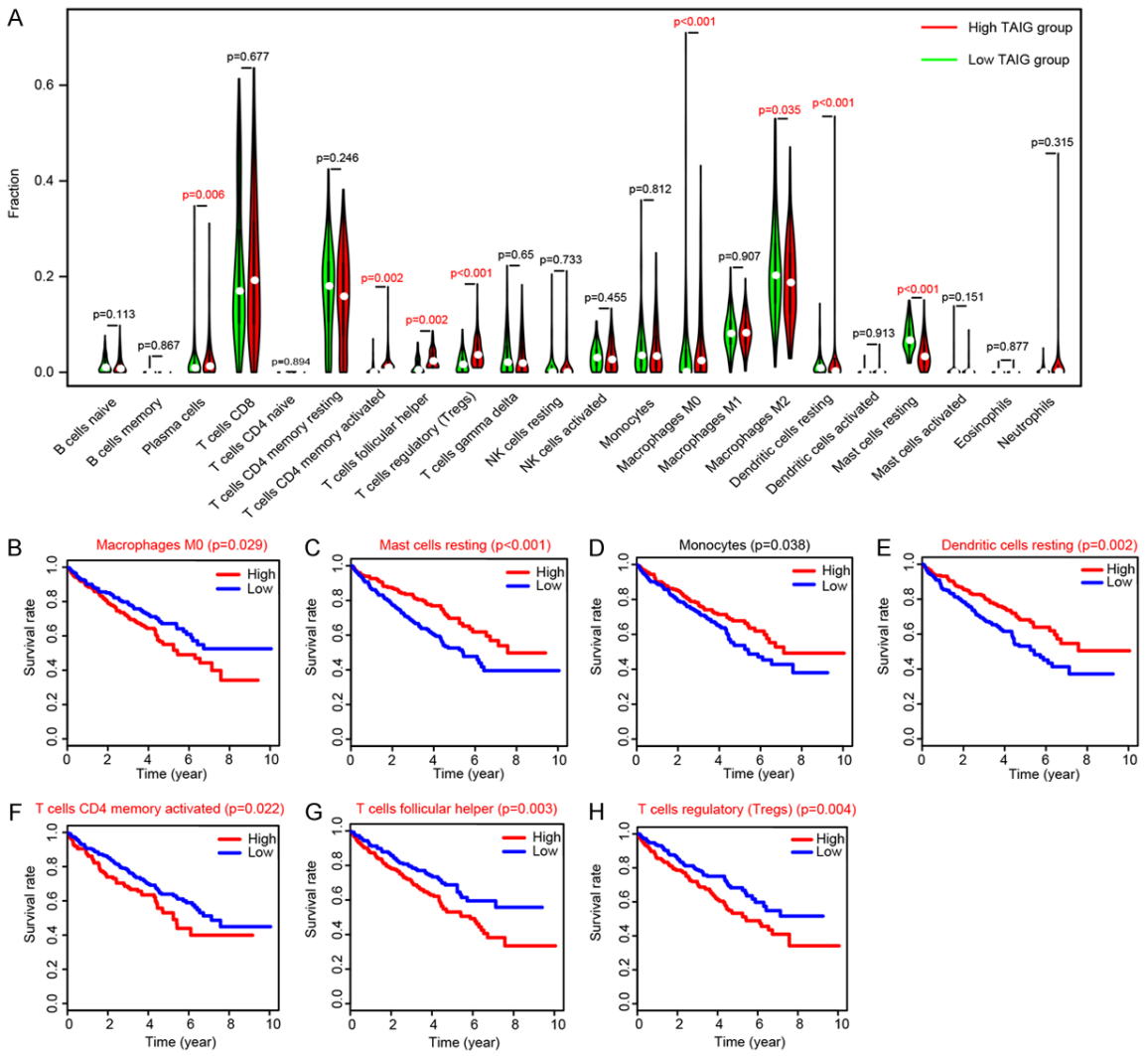


Figure 7. Estimation of Tumor-infiltrating immune cells in ccRCC from the CIBERSORT algorithm. A. Wilcoxon rank-sum test revealed the differential infiltration levels of immune cells in two TAIG groups. B-H. Survival analysis for all immune cells in ccRCC and selecting the significant ones, where the cells annotated in red represented the differential distributions in two TAIG groups.

ble predictive power in an independent data set (**Figure 9A**). In the interim, a marginally statistical variance was observed in the Kaplan-Meier analysis with *P* value of log-rank test = 0.083, indicating the patients with high-TAIG suffered poor survival outcomes (**Figure 9B**).

Discussion

Historic investigations in this field have endeavored to examine the significant biomarkers in the prediction of the prognosis in ccRCC including lncRNA [21], microRNA [22, 23], circRNA or high-frequency mutants [24]. Notwithstanding, however, the tumor associated immune signature has been significantly less testified. In our study, a total of 14 hub immune signatures, were identified, correlated with survival, and an integrative TAIG model was formulated from the multivariate Cox regression results. Methodically, the prognostic value of TAIG was evaluated, which was determined to be an autonomous prognostic factor, versus other risk clinical features via the Cox regression models. It was not simply those patients with a high TAIG that illustrated poor survival outcomes, but the TAIG also correlated positively with the AJCC-TNM stages, pathological stages or tumor grades. Moreover, we further computed the TMB for each patient and depicted the

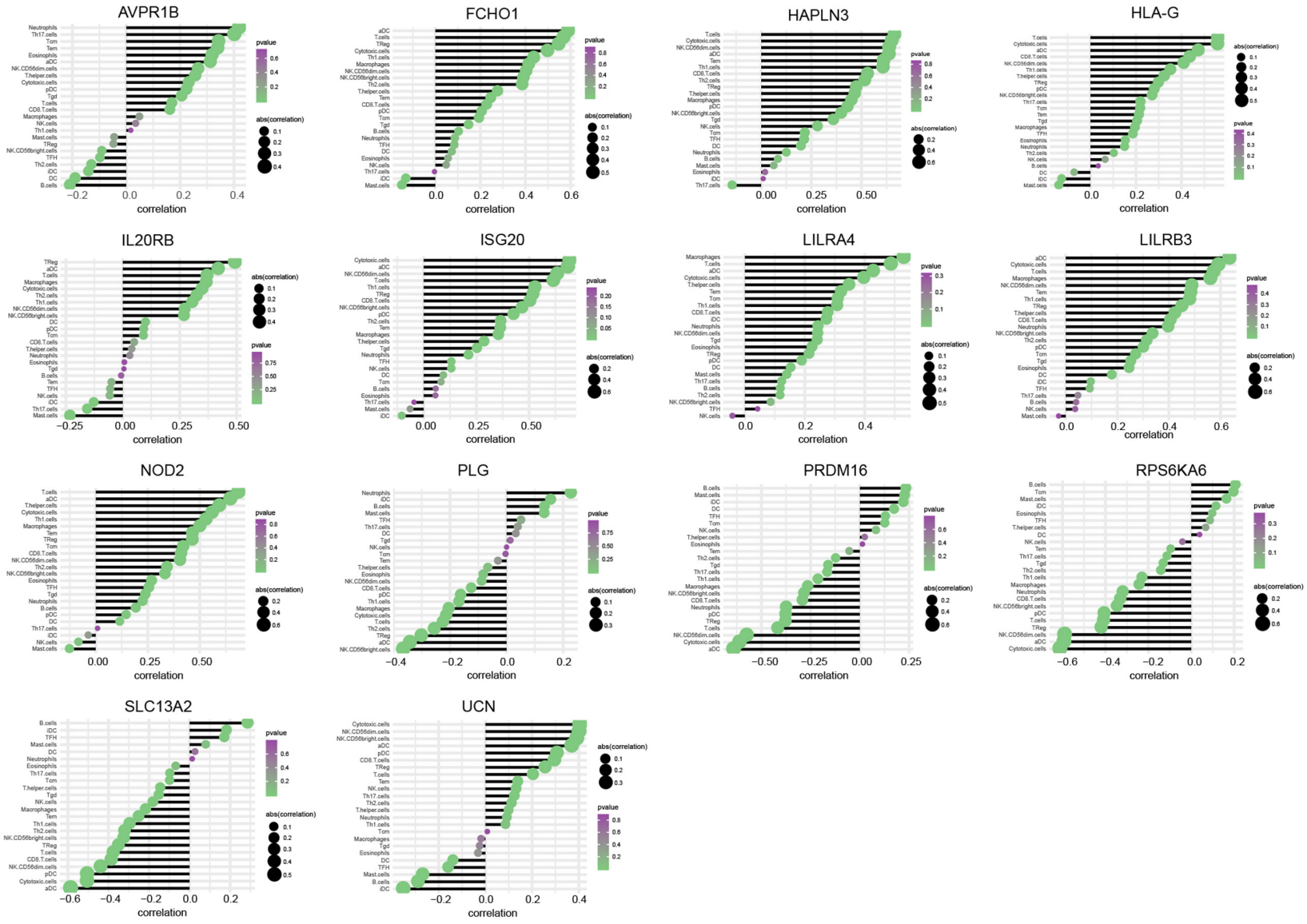


Figure 8. Inferred immune cell fractions based on another approach using marker genes. The Pearson correlation analysis and estimated *P* value revealed the specific associations between identified tumor associated immune signature with immune cells.

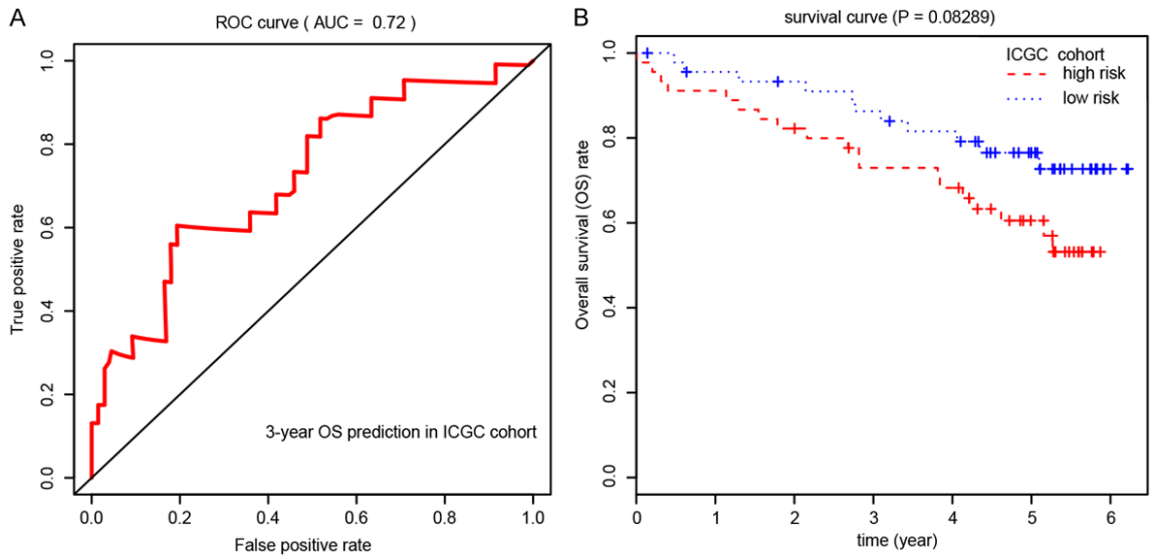


Figure 9. Validation of TAIG in another ICGC cohort. A. The AUC of ROC curve in 3-year OS prediction was 0.72. B. Kaplan-Meier analysis with log-rank test showed the marginal survival difference in two TAIG groups using the median TAIG as the cutoff.

mutation profiles in ccRCC. The Wilcoxon rank-sum test demonstrated the exceptional dependence between TAIG and TMB, and additionally, the high TMB levels projected unfavorable survival outcomes and progression. Given the high correlation of TAIG with TMB, and several immune-related crosstalk, enriched significantly in differential immune signatures, we attempted to determine whether these signatures were concurrent with immune infiltrates in the ccRCC tumor microenvironment. To begin with, we anticipated the abundance of immune cells in each ccRCC sample, based on the CIBERSORT algorithm. Subsequently, we executed the differential analysis, adopting the Wilcoxon rank-sum test. Succeeding this, the prognostic value of differentially distributed immune cells in two TAIG groups was assessed, and remarkably, we discerned that the preponderance of these differential immune infiltrating cells, possessed a significant prognostic value in ccRCC, among which higher infiltrating density of memory activated $CD4^+$ T cell, T follicular helper cells, T regulatory cell or M0 macrophage were hazard factors in high-risk TAIG group, yet higher levels of resting mast cells with resting dendritic cells in low-TAIG group may contribute to tumor suppressors in ccRCC.

The tumor associated immune signature encompassed an aggregate of 14 genes with prognostic ability. It was determined that the majority of the genes were derived from cytokines or their corresponding receptors, whilst the GO enriched items divulged that the affiliated pathways comprised of the regulation of cell-cell adhesion, activation of T cell or proliferation. Previous studies have unveiled the imperative functions of cytokines or chemokines, including IL-4, IL-18 or CXCL family [25-28] in the promotion of tumor inflammatory response, associated with prognosis. Among these genes, HLA-G was reported as an immune checkpoint molecular and functioning an inhibitor particularly for cytotoxic activity of infiltrating NK cells through ILT2 [29], corroborating our results. In addition, we established a quantitative model named TAIG as an immune risk score to evaluate each patient's hazard levels. The distributions of all identified 14 immune signatures, in two TAIG groups, agreed with the subsequent Kaplan-Meier analysis. This illustrated that the hazard immune signature demonstrated higher expression profiles in the high-TAIG group. Notwithstanding, protective immune genes tend to reveal low expression levels and so although we computed the clinical significance of TAIG and the tight correlations with TNM stages or tumor grades, the question of whether the combination of TAIG, with other risk clinical features could further optimize the predictive model, necessitated vast samples to authenticate and

confirm clinical feasibility, therefore, repeat evaluations are required. It must be acknowledged that the role of TMB in the prognosis of ccRCC was debated, along with an analysis of the associations of TMB with TAIG. It is renowned that a high TMB can yield many neoantigens to stimulate immune response, hereby correlating with an enhanced effect of immunotherapy. Given that previous studies across 33 cancer types have already alluded to higher-TMB patients achieving a more favorable prognosis if treated with immunotherapy, otherwise would reveal a more graver prognosis, compared with lower-TMB patients [30]. Consequently, we hypothesized that ccRCC patients with high TAIG and TMB levels might be considered the preferable options of immunotherapies.

Aside from the process of categorizing the immune signature in the ccRCC, we also investigated the tumor-infiltrating immune cells that accounted for the indispensable components in the immune microenvironment. CIBERSORT, a newly computational approach, established by scholars in Stanford University [31], implemented a deconvolution algorithm to characterize diverse cell types, based on their gene signature matrix. Negating the limitations of vast material resources or time needed in flow cytometry and immunohistochemistry, we were able to determine the immune cell fractions in each patient, which was particularly appropriate for dealing with large samples. In agreement with research by Giraldo NA et al., revealing that tumor-infiltrating and peripheral blood T-cell immunophenotypes predict early relapse in localized ccRCC [32], our study was also capable of extricating the risk T cell subsets, including $CD4^+$ T cell, T follicular helper cells, as well as T regulatory (Treg) cells. Since the tumor associated macrophage (TAM) was recognized as a catalyst in tumor progression and has been widely reported of late as being a prevailing predictors for outcomes with Tyrosine kinase inhibitors (TKI) therapy in ccRCC. The study also illustrated that the M0 macrophage subset correlated positively with OS prognosis, which was less reported [33]. Resting mast cells and dendritic cells, however, depicted the protective factors in ccRCC, and the dendritic cells were reported as an immune enhancer utilized as baseline in immunotherapy for solid tumors [34, 35]. More specifically, we also employed another technique to infer the fractions of immune cells, based on the characteristic immune cells marker gene [36]. From an alternate perspective, we elucidated the underlying affiliations between infiltrating immune cells with our identified signature, and we illustrated the specific associations between one gene with single cell subsets. These prognostic tumor-infiltrating immune cells all correlated with our immune signature, and we suggested the hypothesis that these immune signatures impact the differential infiltrating density of immune cells, thus influencing the prognosis in ccRCC.

Accordingly, we confirmed our risk signature in another data set from the ICGC, which is, in essence, a publicly available database that provides the international community with comprehensive genomic data for various cancer types. The predictive value of TAIG still remained superior with AUC = 0.72. Though the *P* value of the log-rank test in the Kaplan-Meier analysis was 0.083, we decided that the marginal variance was caused from a smaller sample size, consisting of only 91 patients. Additionally, the median cutoff was defined improperly, and therefore needs to be optimized, moving forward. Regarded in collaboration, it was the initial endeavor to uncover the risk immune signature in ccRCC, based on large samples with high-throughput data. We also discussed the TMB and TAIG-related infiltrating immune cells. Characterization of the immune landscape from tumor-associated immune gene signatures, to relative prognostic immune cell profiles, in microenvironments, support our comprehensive understanding of prognosis, even regarding immunotherapy strategies in ccRCC.

There were, however, some considerable limitations to our study. First, the correlation between TAIG with TMB or immune infiltrates was computed based on statistics, whereas the actual regulatory mechanisms among them are warranted for further demonstration. Second, the fractions or prognostic value of TAIG-related immune cells might be validated by flow cytometry and finally, the clinical significance of TAIG must be established by our own cohort, this is an area we are formulating in preparation for our next study.

In conclusion, tumor-associated immune signatures were screened and characterized in our study in ccRCC, along with an analysis of its potential prognostic association with TMB. Furthermore, we explicated the prognostic tumor-infiltrating immune cell related with TAIG,

providing a comprehensive foundation for investigating mechanisms or individualized immunotherapy.

Acknowledgements

This research was funded by the Natural Science Foundation of Hunan Province (No. 2018JJ3813).

Disclosure of conflict of interest

None.

Address correspondence to: Peng Jin, Organ Transplant Center, Xiangya Hospital, Central South University, No. 87 Xiangya Road, Kaifu District, Changsha 410008, Hunan Province, China. E-mail: PENGJIN@csu.edu.cn

References

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and 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: 394-424.

[2] Frew IJ and Moch H. A clearer view of the molecular complexity of clear cell renal cell carcinoma. Annu Rev Pathol 2015; 10: 263-289.

[3] Delahunt B, Eble JN, McCredie MR, Bethwaite PB, Stewart JH and Bilous AM. Morphologic typing of papillary renal cell carcinoma: comparison of growth kinetics and patient survival in 66 cases. Hum Pathol 2001; 32: 590-595.

[4] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019; 69: 7-34.

[5] Vitale M, Cantoni C, Pietra G, Mingari MC and Moretta L. Effect of tumor cells and tumor microenvironment on NK-cell function. Eur J Immunol 2014; 44: 1582-1592.

[6] Wang YT, Chen J, Chang CW, Jen J, Huang TY, Chen CM, Shen R, Liang SY, Cheng IC, Yang SC, Lai WW, Cheng KH, Hsieh TS, Lai MZ, Cheng HC, Wang YC and Chen RH. Ubiquitination of tumor suppressor PML regulates prometastatic and immunosuppressive tumor microenvironment. J Clin Invest 2017; 127: 2982-2997.

[7] Sunkara KP, Gupta G, Hansbro PM, Dua K and Bebawy M. Functional relevance of SATB1 in immune regulation and tumorigenesis. Biomed Pharmacother 2018; 104: 87-93.

[8] Weber GF. Molecular mechanisms of metastasis. Cancer Lett 2008; 270: 181-190.

[9] Shrimali RK, Janik JE, Abu-Eid R, Mkrtichyan M and Khleif SN. Programmed death-1 & its ligands: promising targets for cancer immunotherapy. Immunotherapy 2015; 7: 777-792.

[10] Lazar-Molnar E, Scandiuzzi L, Basu I, Quinn T, Sylvestre E, Palmieri E, Ramagopal UA, Nathenson SG, Guha C and Almo SC. Structure-guided development of a high-affinity human programmed cell death-1: implications for tumor immunotherapy. EBioMedicine 2017; 17: 30-44.

[11] Cortellini A, Buti S, Santini D, Perrone F, Giusti R, Tiseo M, Bersanelli M, Michiara M, Grassadonia A, Brocco D, Tinari N, De Tursi M, Zoratto F, Veltri E, Marconcini R, Malorgio F, Garufi C, Russano M, Anesi C, Zeppola T, Filetti M, Marchetti P, Botticelli A, Antonini Cappellini GC, De Galitiis F, Vitale MG, Sabbatini R, Bracarda S, Berardi R, Rinaldi S, Tudini M, Silva RR, Pireddu A, Atzori F, Chiari R, Ricciuti B, Iacono D, Migliorino MR, Rossi A, Porzio G, Cannita K, Ciciarelli V, Fargnoli MC, Ascierto PA and Ficorella C. Clinical outcomes of patients with advanced cancer and pre-existing autoimmune diseases treated with anti-programmed death-1 immunotherapy: a real-world transverse study. Oncologist 2019; 24: e327-e337.

[12] Marginean EC and Melosky B. Is there a role for programmed death ligand-1 testing and immunotherapy in colorectal cancer with microsatellite instability? Part II-the challenge of programmed death ligand-1 testing and its role in microsatellite instability-high colorectal cancer. Arch Pathol Lab Med 2018; 142: 26-34.

[13] Motzer RJ, Penkov K, Haanen J, Rini B, Albiges L, Campbell MT, Venugopal B, Kollmannsberger C, Negrier S, Uemura M, Lee JL, Vasiliev A, Miller WH Jr, Gurney H, Schmidinger M, Larkin J, Atkins MB, Bedke J, Alekseev B, Wang J, Mariani M, Robbins PB, Chudnovsky A, Fowst C, Hariharan S, Huang B, di Pietro A and Choueiri TK. Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. N Engl J Med 2019; 380: 1103-1115.

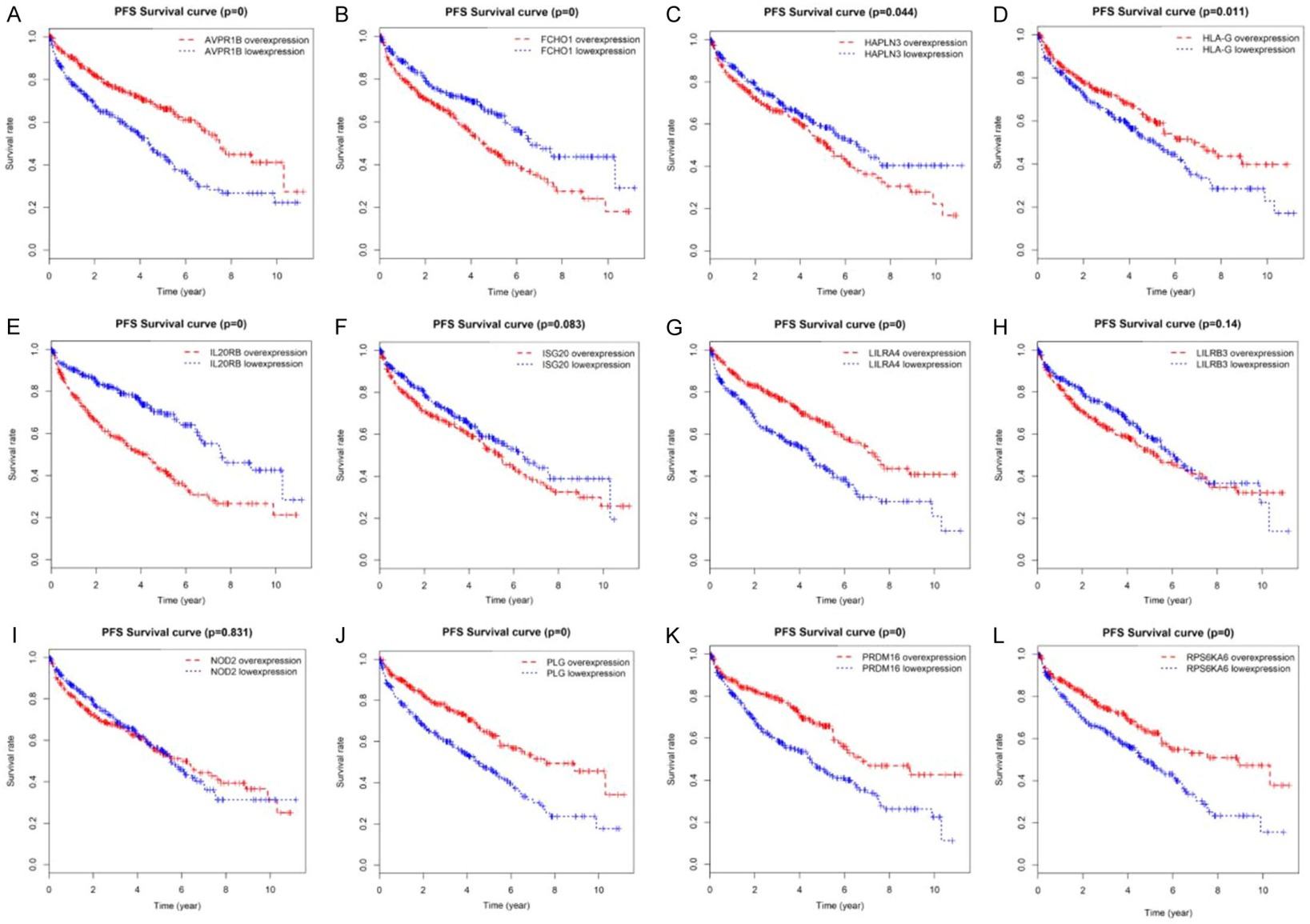
[14] Rini BI, Plimack ER, Stus V, Gafanov R, Hawkins R, Nosov D, Pouliot F, Alekseev B, Soulières D, Melichar B, Vynnychenko I, Kryzhanivska A, Bondarenko I, Azevedo SJ, Borchiellini D, Szczylik C, Markus M, McDermott RS, Bedke J, Tartas S, Chang YH, Tamada S, Shou Q, Perini RF, Chen M, Atkins MB and Powles T; KEYNOTE-426 Investigators. Pembrolizumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. N Engl J Med 2019; 380: 1116-1127.

[15] Song M and Chan AT. The potential role of exercise and nutrition in harnessing the immune system to improve colorectal cancer survival. Gastroenterology 2018; 155: 596-600.

[16] Sawa-Wejksza K and Kandefer-Szerszen M. Tumor-associated macrophages as target for antitumor therapy. Arch Immunol Ther Exp (Warsz) 2018; 66: 97-111.

[17] Jeannin P, Paolini L, Adam C and Delneste Y. The roles of CSFs on the functional polarization of tumor-associated macrophages. FEBS J 2018; 285: 680-699.

[18] Brown SD, Warren RL, Gibb EA, Martin SD, Spinelli JJ, Nelson BH and Holt RA. Neoantigens predicted by tumor genome meta-analysis correlate with increased patient survival. Genome Res 2014; 24: 743-750.
[19] Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MDM, Miller CA, Welch JS, Walter MJ, Wendl MC, Ley TJ, Wilson RK, Raphael BJ and Ding L. Mutational landscape and significance across 12 major cancer types. Nature 2013; 502: 333-339.
[20] Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018; 28: 1747-1756.
[21] Ellinger J, Alam J, Rothenburg J, Deng M, Schmidt D, Syring I, Miersch H, Perner S and Müller SC. The long non-coding RNA lncZNF180-2 is a prognostic biomarker in patients with clear cell renal cell carcinoma. Am J Cancer Res 2015; 5: 2799-2807.
[22] Nofech-Mozes R, Khella HW, Scorilas A, Youssef L, Krylov SN, Lianidou E, Sidiropoulos KG, Gabril M, Evans A and Yousef GM. MicroRNA-194 is a marker for good prognosis in clear cell renal cell carcinoma. Cancer Med 2016; 5: 656-664.
[23] Heinzelmann J, Unrein A, Wickmann U, Baumgart S, Stapf M, Szendroi A, Grimm MO, Gajda MR, Wunderlich H and Junker K. MicroRNAs with prognostic potential for metastasis in clear cell renal cell carcinoma: a comparison of primary tumors and distant metastases. Ann Surg Oncol 2014; 21: 1046-1054.
[24] Ho TH, Park IY, Zhao H, Tong P, Champion MD, Yan H, Monzon FA, Hoang A, Tamboli P, Parker AS, Joseph RW, Qiao W, Dykema K, Tannir NM, Castle EP, Nunez-Nateras R, Teh BT, Wang J, Walker CL, Hung MC and Jonasch E. High-resolution profiling of histone h3 lysine 36 trimethylation in metastatic renal cell carcinoma. Oncogene 2016; 35: 1565-1574.
[25] Quandt D, Jasinski-Bergner S, Muller U, Schulze B and Seliger B. Synergistic effects of IL-4 and TNFalpha on the induction of B7-H1 in renal cell carcinoma cells inhibiting allogeneic T cell proliferation. J Transl Med 2014; 12: 151.
[26] Wei Y, Lin C, Li H, Xu Z, Wang J, Li R, Liu H, Zhang H, He H and Xu J. CXCL13 expression is prognostic and predictive for postoperative adjuvant chemotherapy benefit in patients with gastric cancer. Cancer Immunol Immunother 2018; 67: 261-269.
[27] Chang WS, Shen TC, Yeh WL, Yu CC, Lin HY, Wu HC, Tsai CW and Bau DT. Contribution of inflammatory cytokine interleukin-18 genotypes to renal cell carcinoma. Int J Mol Sci 2019; 20: 1563.
[28] Lee NH, Nikfarjam M and He H. Functions of the CXC ligand family in the pancreatic tumor microenvironment. Pancreatology 2018; 18: 705-716.
[29] Naji A, Menier C, Morandi F, Agaugue S, Maki G, Ferretti E, Bruel S, Pistoia V, Carosella ED and Rouas-Freiss N. Binding of HLA-G to ITIM-bearing Ig-like transcript 2 receptor suppresses B cell responses. J Immunol 2014; 192: 1536-1546.
[30] Zhao Y, Du T, Du L, Li P, Li J, Duan W, Wang Y and Wang C. Long noncoding RNA LINC02418 regulates MELK expression by acting as a ceRNA and may serve as a diagnostic marker for colorectal cancer. Cell Death Dis 2019; 10: 568.
[31] Chen B, Khodadoust MS, Liu CL, Newman AM and Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. Methods Mol Biol 2018; 1711: 243-259.
[32] Giraldo NA, Becht E, Vano Y, Petitprez F, Lacroix L, Validire P, Sanchez-Salas R, Ingels A, Oudard S, Moatti A, Buttard B, Bourass S, Germain C, Cathelineau X, Fridman WH and Sautes-Fridman C. Tumor-infiltrating and peripheral blood t-cell immunophenotypes predict early relapse in localized clear cell renal cell carcinoma. Clin Cancer Res 2017; 23: 4416-4428.
[33] Hakimi AA, Voss MH, Kuo F, Sanchez A, Liu M, Nixon BG, Vuong L, Ostrovnaya I, Chen YB, Reuter V, Riaz N, Cheng Y, Patel P, Marker M, Reising A, Li MO, Chan TA and Motzer RJ. Transcriptomic profiling of the tumor microenvironment reveals distinct subgroups of clear cell renal cell cancer: data from a randomized phase III trial. Cancer Discov 2019; 9: 510-525.
[34] Jung NC, Lee JH, Chung KH, Kwak YS and Lim DS. Dendritic cell-based immunotherapy for solid tumors. Transl Oncol 2018; 11: 686-690.
[35] Laurell A, Lönnemark M, Brekkan E, Magnusson A, Tolf A, Wallgren AC, Andersson B, Adamson L, Kiessling R and Karlsson-Parra A. Intratumorally injected pro-inflammatory allogeneic dendritic cells as immune enhancers: a first-in-human study in unfavourable risk patients with metastatic renal cell carcinoma. J Immunother Cancer 2017; 5: 52.
[36] Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman WH, Becker C, Pagès F, Speicher MR, Trajanoski Z and Galon J. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013; 39: 782-795.


A
PFS Survival curve (p=0)
AVPR1B overexpression
AVPR1B lowexpression
B
PFS Survival curve (p=0)
FCHO1 overexpression
FCHO1 lowexpression
C
PFS Survival curve (p=0.044)
HAPLN3 overexpression
HAPLN3 lowexpression
D
PFS Survival curve (p=0.011)
HLA-G overexpression
HLA-G lowexpression
E
PFS Survival curve (p=0)
IL20RB overexpression
IL20RB lowexpression
F
PFS Survival curve (p=0.083)
ISG20 overexpression
ISG20 lowexpression
G
PFS Survival curve (p=0)
LILRA4 overexpression
LILRA4 lowexpression
H
PFS Survival curve (p=0.14)
LILRB3 overexpression
LILRB3 lowexpression
I
PFS Survival curve (p=0.831)
NOD2 overexpression
NOD2 lowexpression
J
PFS Survival curve (p=0)
PLG overexpression
PLG lowexpression
K
PFS Survival curve (p=0)
PRDM16 overexpression
PRDM16 lowexpression
L
PFS Survival curve (p=0)
RPS6KA6 overexpression
RPS6KA6 lowexpression
Survival rate
Time (year)

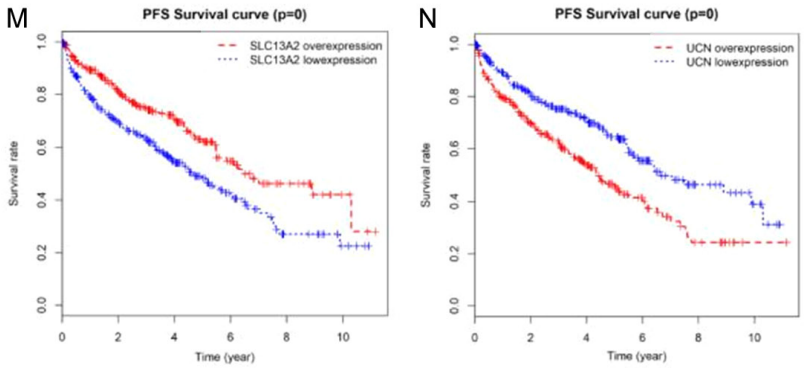


Figure S1. Survival analysis for comparing the PFS difference of identified 14 hub immune signature.

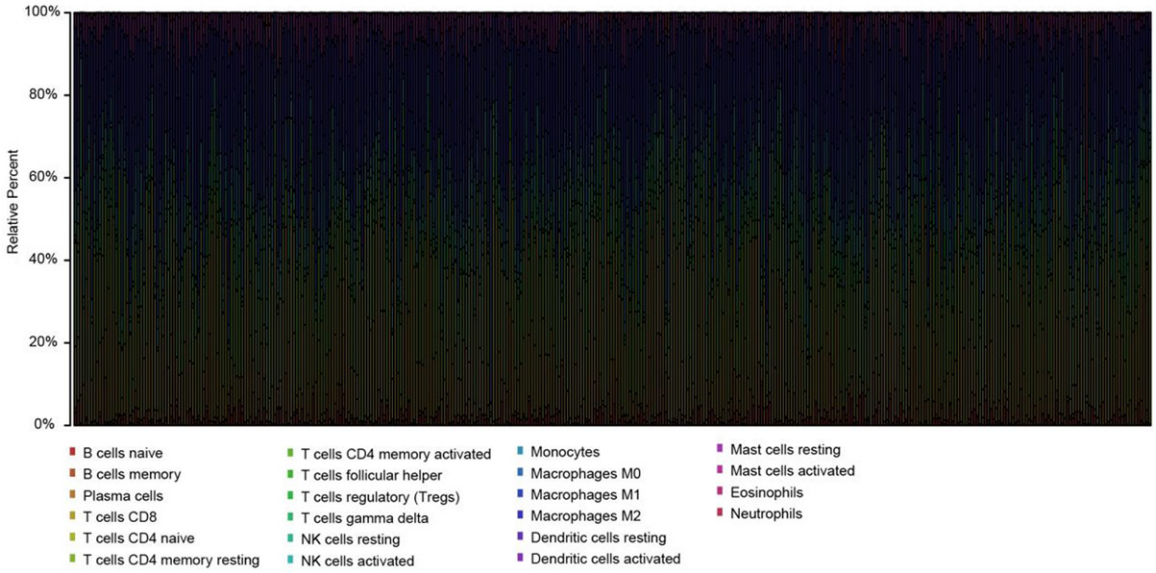


Figure S2. Estimation of 22 immune cell subsets fractions using the CIBERSORT algorithm. Each Bar chart exhibited the cell proportions of each patient and various colors annotated below the legend represented the 22 immune cells.