Integrative characterization and validation of tumor-associated immune signature combined with immune-infiltration analysis in clear cell renal cell carcinoma

Hongxi Chen
Hunan Normal University
Jinliang Xie
Central South University Xiangya School of Medicine
Peng Jin (jinpeng2077@sina.com)
Central South University Xiangya School of Medicine https://orcid.org/0000-0003-4266-0034

Primary research

Keywords: Immune signature, Tumor mutation burden, Immune infiltrates, Prognosis, ccRCC

Posted Date: April 3rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-20246/v1

Abstract

Background We aimed to explore and validate a prognostic immune signature for predicting the prognosis of ccRCC patients, combined with immune-infiltration analysis.

Methods We obtained the multi-omics data from public datasets. Differential analysis was performed by edgeR package. 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. Functional analysis was conducted to uncover the related crosstalk. Importantly, we implemented the CIBERSORT algorithm to estimate the immune cell fractions in ccRCC samples and analyzed the differential abundance of tumor-infiltrating immune cells in two TAIG groups using Wilcoxon rank-sum test. The prognostic role of differential immune cells was further assessed by Kaplan-Meier analysis. In addition, we investigated the associations of single immune signature with specific immune cells.

Results A total of 628 ccRCC patients were included in our integrative analysis, including 537 ccRCC patients in discovery group and 91 patients in validation group. Then, we identified the 14 key immune signatures. The AUC was 0.802 and patients with higher TAIG suffered from worse prognosis. Correlation analysis indicated that TAIG correlated tightly with clinical variables and TMB. Moreover, functional analysis also implicated the immune-related GO items or crosstalk. Therefore, we discovered the relationships of TAIG with tumor-infiltrating immune cells. The differential abundance of immune cells showed significant prognostic difference consisted of memory activated CD4 + T cell, T follicular helper cells, T regulatory cell, and so on. Moreover, we also characterized the associations between identified signature with specific immune cells. Finally, the 5-year AUC in ICGC cohort was 0.72, suggesting the robustness of TAIG that we constructed.

Conclusions Totally, our team characterized the tumor-associated immune signature in ccRCC and further uncovered the prognostic tumor-infiltrating immune cells related with TAIG, providing a comprehensive foundation for investigating mechanisms or individualized immunotherapy.

1. Background

Kidney cancer has become an increased solid malignancy in urology system with high mortality and poor survival outcomes (1), of which clear cell renal cell carcinoma (ccRCC) is the most common subtype accounting for nearly 70% (2, 3). Despite the developed strategies in detection or cancer management, the incidence of ccRCC still increased and the newly estimated cases in the United states is up to 73,820 in 2019 (4). So far, the surgical intervention consisted of laparoscopic partial nephrectomy or radical nephrectomy is still the mainstream treatment for ccRCC patients. Nevertheless, for approximately 30% of cases inevitably progressed into advanced pathological stages or tumor recurrence, the ccRCC patients still suffered from poor overall survival prognosis. The current determinants of prognosis in ccRCC mainly included tumor size (T stage), pathological grades, as well as histological subtypes, yet the heterogeneity still existed in cases with similar clinical characteristics implicating the insufficiency of

traditional prognostic factors in predicting clinical outcomes. Hence, present emphasis is warranted to place on the identification of robust and stable biomarkers with credible accuracy and sensitivity to reflect the comprehensive biological features in ccRCC.

Aberrant immune regulation has been well recognized as a vital component in tumor microenvironment (5, 6), participating in tumorigenesis (7), progression, and even metastasis (8). With the intensive investigations in immunity, immunotherapies emerged as a promising role in tumor treatment theses years, such as 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 the combination of Avelumab (PD-L1 antibody) or Pembrolizumab (PD-1 antibody) plus with Axitinib showed superiority versus the single sunitinib (13, 14). However, the efficacy of PD-1/PD-L1 inhibitors revealed incompletely definite in patients with various objective response rates, indicating the importance of characterizing tumor immune environment. Previous studies already reported that overall survival (OS) (15) and prognosis in patients could be influenced by fractions of immune environment, including tumor-associated macrophage (16, 17), mast cells, or stromal cells. Nevertheless, immune gene biomarkers discovered based on large samples were less reported and characterization of immune signature in tumor microenvironment was of significance for our understanding on ccRCC. Moreover, tumor mutation burden (TMB) was also reported as an effective biomarker for discriminating the responsiveness of immunotherapy in patients (18, 19), but whether TMB was associated with improved prognosis or promotion of immunotherapies remained in consistence across multiple malignancies. Hence, it shed light on the discussion on the potential associations of immune signature in tumor microenvironment with TMB, even immune-infiltrating cells.

In our study, we mainly obtained the transcriptome data and mutation profiles from The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC) databases to screen the most significantly prognostic immune signature in ccRCC. We revealed the associations between the identified tumor-associated immune Genes (TAIG) and clinical features. Moreover, we further explored the relationships between TAIG and genomic alterations, TMB. In addition, we continued to characterize the differential immune infiltrates associated with TAIG and prognostic value of significant infiltration cells. Besides, we utilized another data set to validate the predictive accuracy of our identified signature. Our team intended to explore the tumor-associated immune signature biomarkers in microenvironment for predicting prognosis or immunotherapy, meanwhile providing a comprehensive characterization of immune components in ccRCC including tumor mutation burden, immune infiltrates and the potential relationships among them.

2. Methods And Materials

2.1 Data collection and preprocessing

The transcriptome expression data of ccRCC samples were obtained from TCGA database (https://portal.gdc.cancer.gov/) and ICGC database (https://icgc.org/). Besides, the somatic mutation data were downloaded from the“Masked Somatic Mutation” category in TCGA which were analyzed by

VarScan software. The differential analysis and normalization of transcriptome profiles were conducted by edgeR package. The files with Mutation Annotation Format (MAF) was prepared and dealt with maftools (20) package, which was commonly used in analysis of cancer genomics with functions of customizable visualizations. A list of 4678 immune signature was acquired from the InnateDB database (https://www.innatedb.ca/), which was a publicly available resource for immunity research (Table S1). Furthermore, clinical features of age, gender, TNM stages, tumor grades, follow-up with vital status were searched from the database via TCGA biolinks package and the patients with insufficient clinical data were excluded.

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

The differentially expressed genes (DEGs) in tumor versus normal samples were obtained from the edgeR package. Then, univariate Cox analysis was conducted using survival package with $P < 0.01$ to identify key prognostic immune genes. Then, we conducted stepwise regression analysis to find the independent prognostic factors in multivariate Cox method with $P < 0.05$. The process of selection was illustrated in Venn graph by VennDiagram package. The risk TAIG model based on hub immune signature was thus constructed as the following: $TAIG = \sum(\beta_i * EXP_i)$, where β_i, the coefficients, meant the weight of each signature and EXP_i represented the expression data. We could accordingly classify the patients into two groups using the median TAIG score as the cutoff. We assessed the differential cluster of hub signature in heatmap plot in two groups by pheatmap package. In addition, the 3-year and 5-year of ROC curve was shown by timeROC package to evaluate the predictive value of TAIG in OS prediction. Kaplan-Meier analysis was conducted to compare the difference of OS in two TAIG levels. Similarly, we also analyzed the clinical value of TAIG in tumor recurrence or progression.

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

We extracted the expression data of each identified signature and merged with OS time and PFS time for 537 ccRCC patients in TCGA cohort. Kaplan-Meier analysis with log-rank test was utilized for assessing the prognostic difference of each hub signature in OS or progression free survival (PFS), respectively. Meanwhile, we explored the potential associations of TAIG with clinical variables, where Wilcoxon rank-sum test was utilized for comparing differential levels of TAIG between two groups, yet Kruskal-Wallis was appropriate when it comes to three or more groups (Table S2). What is more, we conducted the 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. As the N stage included lots of missing cases, we disregarded it in subsequent analysis. To make sure whether the TAIG level maintained an independent risk factor compared with other clinical variables, we selected the significant ones and performed multivariate Cox regression to assess the clinical significance of TAIG with $P < 0.05$.

2.4 Tumor mutation burden and correlation analysis

Since we obtained the mutation profiles in ccRCC, we written the Perl scripts based on JAVA platform to extract the specific genomic alterations for each patient. The detected mutants included deletions, insertions, or substitutions across bases and we defined the TMB as following: 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, TMB score was calculated for each patient and the association with TAIG was determined with Pearson correlation analysis with estimated P value. We further investigated the differential distributions of TAIG in high- and low-TMB groups with Wilcoxon rank-sum test. Additionally, the prognostic value of TMB in OS or PFS was evaluated by Kaplan-Meier analysis as a supplementary analysis for TAIG.

2.5 Functional analysis, GSEA

Based on the differential and prognostic analysis in Fig. 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. Then, we used the cluster Profiler, enrichplot and ggplot2 packages to search the significantly enriched GO items related with hub prognostic immune genes. Given we have classified the cohort into two groups with high- and low-TAIG levels, we further conducted the GSEA between two groups using the TAIG as the phenotype. The GSEA software was running based on 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). Enriched pathways with FDR < 0.05 was thought to be statistically significant.

2.6 Estimation of tumor immune infiltrates through CIBERSORT or gene markers

CIBERSORT was a developed computational method used to quantify cell fractions from bulk tissue gene expression profiles, which was well validated by flow cytometry with large tumor biopsy samples. We used the normalized expression data with FPKM format in TCGA cohort and obtained the COBERSORT R package from the public website (https://cibersort.stanford.edu/). The inferred immune cell fractions were shown by boxplot and the differential abundance of immune infiltrates in high- or low-TAIG levels was assessed by Wilcoxon rank-sum test. What is more, we merged the immune factions with survival information and performed the Kaplan-Meier analysis to uncover whether the differentially distributed immune infiltrating cells in two TAIG levels possessed the prognostic values. Besides, to reduce the statistical bias in analysis of cell compositions, we exploited another commonly recognized method to quantify the tumor immune infiltrating cells based on already identified marker genes listed in Table S6 using the GSVA package. Given the matrix of immune cells and the expression data of hub immune signature, we further discussed the specific associations of single gene with tumor immune infiltrating cells. The Spearman correlation analysis with estimated P value was shown in dotplot.

2.7 Validation of TAIG in another independent cohort

Since we analyzed the clinical significance of TAIG and association with TMB or immune infiltrates, we intended to demonstrate the robustness of model in an independent data set. From the ICGC database, we acquired the 91 ccRCC patients with complete survival information and transcriptome sequencing data (Table S3). We extracted the expression profiles of identified signature in TCGA and conducted the ROC curve to assess the predictive power of TAIG model in ICGC cohort. Besides, the differential survival outcomes between two TAIG levels in ICGC cohort was further compared by Kaplan-Meier analysis with log-rank test.

2.8 Statistical analysis

Cox regression models or Kaplan-Meier analysis were conducted by survival package. The Student's t test was used for continuous variables, while categorical variables were dealt with Chi-square (χ^2) test. Wilcoxon rank-sum test was utilized to compare ranked data and Kruskal-Wallis test was used for comparisons among three or more groups. All statistical analysis was performed in R studio (Version 3.5.2), and we thought the statistical significance only with $P < 0.05$.

3. Results

3.1 Construction and assessment of TAIG in ccRCC patients

We obtained a total of 537 ccRCC patients from TCGA database with transcriptome profiles and the mutation data of 336 patients. The clinical baseline was summarized in Table 1. Differential analysis based on edgeR package revealed a list of 826 DEGs and we acquired the 253 intersect immune signature from the InnateDB database. Then, univariate Cox analysis was used to find the hub 53 prognostic immune signature (Fig. 1A-B). To further screen the independent factors, stepwise regression method was utilized and we finally got the 14-hub tumor-associated immune signature, which were shown in forest plot with hazard ratio and corresponding 95% confidence interval (CI) (Fig. 1C). The TAIG was thus established from the multivariate Cox analysis as following: 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 could observe the differentially expressed levels of signature in high- and low-TAIG groups in heatmap (Fig. 1D). The distribution plot intuitively revealed the patients with higher TAIG suffered more dead cases (Fig. 2A-B). Besides, The AUC in 3-year OS prediction was 0.778 and that in 5-year OS prediction was 0.802, respectively, suggesting the superior predictive power (Fig. 2C). Correspondingly, the log-rank test showed patients with higher TAIG had more risk in OS with $P < 0.001$ (Fig. 2D). Furthermore, the TAIG also showed better predictive value in tumor progression and the AUC of 5-year PFS was 0.794 (Fig. 2E). Meanwhile, higher TAIG indicated a highly significant predictor of tumor progression or recurrence with $P < 0.001$ (Fig. 2F). Given the obvious clinical significance of TAIG, we specifically continued to explore the prognostic value of single hub gene in OS or PFS, and the log-rank test showed that nearly all hub genes were tightly associated with OS or PFS in ccRCC patients (Fig. 3, S1).

Table 1
Clinical features of patients included in this study.

Variables	TCGA	ICGC
	(N = 537)	(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 (%).

3.2 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. First, we merged the TAIG with other variables and conducted the Cox analysis. 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. However, the TAIG (P < 0.001), tumor grade (P = 0.003) or pathological stage (P = 0.016) still retained the robust significance in multivariate Cox regression analysis, whereas no statistical difference was observed in T stage and M stage (Table 2). Apart from the Cox analysis for demonstrating the prognostic role of TAIG, we also investigated the underlying relationships of TAIG with other variables. The correlation analysis suggested that higher TAIG correlated highly with advanced T stage (P = 4.032e − 17), N stage (P = 2.351e − 04), metastasis (P = 8.708e − 12), pathological stage (P = 4.08e − 19), as well as higher tumor grades (P = 7.03e − 15) (Fig. 4A-E).

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 ***

3.3 TAIG correlated with TMB positively indicating poor prognostic outcomes

Taking into account the fact that tumor mutation burden was reported to be associated with tightly with immunotherapeutic response and tumor prognosis, we thereby intended to figure out the correlation of TMB with identified immune signature. We illustrated the mutation profiling of ccRCC in waterfall plot, in which the different colors annotated at the bottom showed the various mutation types and the TMB for each sample was calculated above the legend (Fig. 5A). As we only extracted the mutation data of only 336 patients, we integrated the TMB score with matched TAIG to perform the correlation analysis (Table

S4). Wilcoxon rank-sum test showed the higher TMB levels in high-TAIG group (Fig. 5B), and the Pearson correlation analysis provided the supplementary proof with r = 0.188 and P = 0.001 (Fig. 5C). Furthermore, we additionally discovered the prognostic value of TMB and found that higher TMB was associated with the poor OS outcomes (P = 0.035) and suffered more hazards in tumor progression or recurrence (P = 0.01) (Fig. 5D-E).

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

Since we have already obtained a list of 53 prognostic differential immune signature in Fig. 1B, we further conducted the GO enrichment analysis (Fig. 6A). In biological process group, immune DEGs were mainly enriched in regulation of cell-cell adhesion or T cell activation. In the cellular component category, these genes were associated with extracellular matrix. While in molecular function group, these genes mainly participated in receptor ligand activity, glycosaminoglycan biding, as well as cytokine activity. Moreover, Gene set enrichment analysis for comparing immune phenotype between high- and low-TAIG groups indicated that the higher TAIG was associated with regulation of chemokine signaling pathway, VEGF signaling pathway or T cell receptor signaling pathway, while correlated negatively with PPAR signaling pathway, lysine degradation, as well as TGF-β signaling pathway (Fig. 6B).

3.5 TAIG correlated with several prognostic immune infiltrating cells

Given the potential relationships of TAIG with TMB or several immune-related crosstalk, we considered whether TAIG influenced the density levels of tumor infiltration cells functioning an important role in microenvironment. Based on signature expression data from ccRCC patients and CIBERSORT algorithm, we estimated the specific fractions of 22 immune cells in each sample illustrated in Figure S2, where the sum of various immune types in boxplot equaled 100% (Table S5). Then, Wilcoxon rank-sum test indicated the differential distributions of several immune cells in two TAIG groups. Significant difference was found that higher abundance of memory activated $CD4^+$ T cell (P = 0.002), T follicular helper cells (P = 0.002), T regulatory cell (P < 0.001), M0 macrophage (P < 0.001) were in high-TAIG group, yet M2 macrophage (P = 0.035), resting dendritic cells (P < 0.001), as well as resting mast cells (P < 0.001) showed lower infiltrating levels in high-TAIG group (Fig. 7A). In accordance with the previous results, the subsequent Kaplan-Meier analysis interestingly indicated that most of these differentially distributed immune infiltration cells in two TAIG groups possessed prognostic significance, in which higher levels of memory activated $CD4^+$ T cell (P = 0.022), T follicular helper cells (P = 0.003), T regulatory cell (P = 0.004) or M0 macrophage (P = 0.029) correlated with poor OS outcomes, whereas resting mast cells (P < 0.001) with resting dendritic cells (P = 0.002) may serve as tumor suppressors in prognosis (Fig. 7B-H).

Additionally, we utilized another algorithm to quantify tumor-infiltrating immune cells from transcriptomics data based on the marker genes in Table S6. Then, we integrated the output immune cell

matrix with extracted specific signature expression. Spearson correlation analysis with estimated P value showed the statistical associations of single signature with specific immune infiltrating cells, providing another strong evidence between our identified signature with tumor immune cells (Fig. 8).

3.6 Validation of the robust TAIG in ICGC

In ICGC cohort consisting of 91 ccRCC samples, we further demonstrated the robust signature using Cox regression method and the AUC of ROC curve was 0.72 implicating the stable predictive power in an independent data set (Fig. 9A). Meanwhile, a marginally statistical difference was found in Kaplan-Meier analysis with P value of log-rank test = 0.083 indicating the patients with high-TAIG suffered poor survival outcomes (Fig. 9B).

4. Discussion

Previous studies attempted to investigate the meaningful biomarkers in prediction of prognosis in ccRCC including lncRNA (21), microRNA (22, 23), circRNA or high-frequency mutants (24). However, the tumor associated immune signature has been less reported. In our study, we identified a total of 14 hub immune signature associated with survival and constructed an integrative TAIG model from the multivariate Cox regression results. We systematically assessed the prognostic value of TAIG that was demonstrated to be an independent prognostic factor versus other risk clinical features via Cox regression models. Not only the patients with high TAIG showed poor survival outcomes, but the TAIG also correlated positively with AJCC-TNM stages, pathological stages or tumor grades. Besides, we further calculated the TMB for each patient and illustrated the mutation profiles in ccRCC. Wilcoxon rank-sum test revealed the superior relativity between TAIG and TMB, and high TMB levels predicted poor survival outcomes and progression. Given the high correlation of TAIG with TMB and several immune-related crosstalk enriched significantly in differential immune signature, we attempted to figure out whether theses signature correlated with immune infiltrates in ccRCC tumor microenvironment. Firstly, we estimated the abundance of immune cells in each ccRCC sample based on CIBERSORT algorithm, and we conducted the differential analysis using Wilcoxon rank-sum test. Then, we additionally assessed the prognostic value of differentially distributed immune cells in two TAIG groups. Interestingly, we observed that most of these differential immune infiltrating cells possessed significantly 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 included a total of 14 genes possessing prognostic ability. We found most of genes belonged to cytokines or their corresponding receptors, and the GO enriched items revealed the associated pathways consisted of regulation of cell-cell adhesion, activation of T cell or proliferation. Previous researches uncovered the vital roles of cytokines or chemokines, such as IL-4, IL-18 or CXCL family (25–28) in promotion of tumor inflammatory response associated with prognosis. Among these genes, HLA-G was reported as an immune checkpoint molecular and functioning an inhibitor

especially for cytotoxic activity of infiltrating NK cells through ILT2 (29), in accordance with our results. We further established a quantitative model named TAIG as an immune risk score for assessing the hazard levels of each patient. The distributions of all identified 14 immune signatures in two TAIG groups were in agreement with the subsequent Kaplan-Meier analysis, where the hazard immune signature showed higher expression profiles in high-TAIG group. But protective immune genes tend to reveal low expression levels. Though we figured out the clinical significance of TAIG and the tight correlations with TNM stages or tumor grades, whether the combination of TAIG with other risk clinical features could further optimize the predictive model warranted large samples to validate and clinical feasibility should be evaluated again. It is worth mentioning that we discussed the role of TMB in prognosis of ccRCC and analyzed the associations of TMB with TAIG. It has been well recognized that high TMB may yield many neoantigens to stimulate immune response thus correlating with better effect of immunotherapy. Given that previous studies across 33 cancer types have already implicated higher-TMB patients could gain a more favorable prognosis if treated with immunotherapy, otherwise would reveal a worse prognosis compared to lower-TMB patients (30), we hypothesized ccRCC patients with high TAIG and TMB levels might be considered the preferable options of immunotherapies.

Apart from the characterization of immune signature in ccRCC, we also investigated the tumor-infiltrating immune cells that accounted for the indispensable components in immune microenvironment. CIBERSORT, a newly computational approach developed by scholars in Stanford University (31), implemented a deconvolution algorithm to characterize diverse cell types based on gene signature matrix. Avoiding the defects of large material resources or time needed in flow cytometry and immunohistochemistry, we could quantify the immune cell fractions in each patient, especially 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 also distinguished the risk T cell subsets, including $CD4^+$ T cell, T follicular helper cells, as well as T regulatory (Treg) cell. Since the tumor associated macrophage (TAM) was recognized as a promoter in tumor progression and reported currently as powerful predictors for outcomes with Tyrosine kinase inhibitors (TKI) therapy in ccRCC, we also found that the M0 macrophage subset correlated positively with OS prognosis which was less reported (33). However, resting mast cells and dendritic cells showed the protective factors in ccRCC, and the dendritic cells was reported as an immune enhancer utilized as baseline in immunotherapy for solid tumors (34, 35). What is more, we also used another method to infer the fractions of immune cells based on the characteristic immune cells marker gene (36). From another aspect, we also demonstrated the underlying relationships between infiltrating immune cells with our identified signature and we illustrated the specific associations between one gene with single cell subset. These prognostic tumor-infiltrating immune cells all correlated with our immune signature, and we proposed the hypothesis that these immune signatures impact the differential infiltrating density of immune cells, thus influencing the prognosis in ccRCC.

Accordingly, we validated our risk signature in another data set from the ICGC, which is publicly available database providing the international community with comprehensive genomic data for various cancer

types. The predictive value of TAIG still maintained superior with AUC = 0.72. Though the P value of log-rank test in Kaplan-Meier analysis was 0.083, we considered the marginal difference resulted from the smaller sample size with only 91 patients and the median cutoff was defined improperly that need to be further optimized. Taken together, it is our first attempt to uncover the risk immune signature in ccRCC based on large samples with high-throughput data. Additionally, we discussed the TMB and TAIG-related infiltrating immune cells. Characterization of immune landscape from tumor-associated immune gene signature to relative prognostic immune cell profiles in microenvironment favor our comprehensive understanding of prognosis, even immunotherapy strategies in ccRCC.

However, there existed several problems in our work as the following. Firstly, the correlation between TAIG with TMB or immune infiltrates was calculated based on statistics, the actual regulatory mechanisms among them are warranted for further demonstration. Secondly, the fractions or prognostic value of TAIG-related immune cells might be validated by flow cytometry. Lastly, the clinical significance of TAIG needs to be determined by our own cohort which we are preparing to conduct in next work.

Conclusion

Tumor-associated immune signatures were screened and characterized in our study in ccRCC, and its potential prognostic association with TMB was analyzed. Moreover, we further uncovered the prognostic tumor-infiltrating immune cell related with TAIG, providing a comprehensive foundation for investigating mechanisms or individualized immunotherapy.

Abbreviations

ccRCC: clear cell renal cell carcinoma; TCGA: The Cancer Genome Atlas; ICGC: International Cancer Genome Consortium; TAIG: tumor-associated immune Genes; OS: overall survival; PFS: progression free survival; TMB: tumor mutation burden; GO: Gene ontology; GSEA: Gene Set Enrichment Analysis; AUC: area under curve; PD-1: programmed death-1 (PD-1) ; PD-L1: programmed death ligand 1; MAF: Mutation Annotation Format; DEGs: differentially expressed genes

Declarations

Acknowledgements

Not applicable.

Authors’ contributions

Hongxi Chen1, Jinliang Xie2, Peng Jin conceived the concept of the paper. Hongxi Chen and Jinliang Xie conducted the analysis and draft writing.

Funding

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

Availability of data and materials

Additional data not presented in the manuscript can be obtained by contacting the authors.

Ethics approval and consent to participant

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Frew IJ, Moch H. A clearer view of the molecular complexity of clear cell renal cell carcinoma. Annu Rev Pathol. 2015;10:263-89.
3. Delahunt B, Eble JN, McCredie MR, Bethwaite PB, Stewart JH, Bilous AM. Morphologic typing of papillary renal cell carcinoma: comparison of growth kinetics and patient survival in 66 cases. Hum Pathol. 2001;32(6):590-5.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7-34.
5. Vitale M, Cantoni C, Pietra G, Mingari MC, Moretta L. Effect of tumor cells and tumor microenvironment on NK-cell function. Eur J Immunol. 2014;44(6):1582-92.
6. Wang YT, Chen J, Chang CW, Jen J, Huang TY, Chen CM, et al. Ubiquitination of tumor suppressor PML regulates prometastatic and immunosuppressive tumor microenvironment. J Clin Invest. 2017;127(8):2982-97.
7. Sunkara KP, Gupta G, Hansbro PM, Dua K, 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(2):181-90.
9. Shrimali RK, Janik JE, Abu-Eid R, Mkrtichyan M, Khleif SN. Programmed death-1 & its ligands: promising targets for cancer immunotherapy. Immunotherapy. 2015;7(7):777-92.
10. Lazar-Molnar E, Scandiuzzi L, Basu I, Quinn T, Sylvestre E, Palmieri E, et al. 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, et al. 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.
12. Marginean EC, 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(1):26-34.
13. Motzer RJ, Penkov K, Haanen J, Rini B, Albiges L, Campbell MT, et al. Avelumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N Engl J Med. 2019;380(12):1103-15.
14. Rini BI, Plimack ER, Stus V, Gafanov R, Hawkins R, Nosov D, et al. Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N Engl J Med. 2019;380(12):1116-27.
15. Song M, Chan AT. The Potential Role of Exercise and Nutrition in Harnessing the Immune System to Improve Colorectal Cancer Survival. Gastroenterology. 2018;155(3):596-600.
16. Sawa-Wejksza K, Kandefer-Szerszen M. Tumor-Associated Macrophages as Target for Antitumor Therapy. Arch Immunol Ther Exp (Warsz). 2018;66(2):97-111.
17. Jeannin P, Paolini L, Adam C, Delneste Y. The roles of CSFs on the functional polarization of tumor-associated macrophages. FEBS J. 2018;285(4):680-99.
18. Brown SD, Warren RL, Gibb EA, Martin SD, Spinelli JJ, Nelson BH, et al. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. Genome Res. 2014;24(5):743-50.
19. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature. 2013;502(7471):333-9.
20. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747-56.
21. Ellinger J, Alam J, Rothenburg J, Deng M, Schmidt D, Syring I, et al. The long non-coding RNA lnc-ZNF180-2 is a prognostic biomarker in patients with clear cell renal cell carcinoma. Am J Cancer Res. 2015;5(9):2799-807.
22. Nofech-Mozes R, Khella HW, Scorilas A, Youssef L, Krylov SN, Lianidou E, et al. MicroRNA-194 is a Marker for Good Prognosis in Clear Cell Renal Cell Carcinoma. Cancer Med. 2016;5(4):656-64.
23. Heinzelmann J, Unrein A, Wickmann U, Baumgart S, Stapf M, Szendroi A, et al. 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(3):1046-54.
24. Ho TH, Park IY, Zhao H, Tong P, Champion MD, Yan H, et al. High-resolution profiling of histone h3 lysine 36 trimethylation in metastatic renal cell carcinoma. Oncogene. 2016;35(12):1565-74.
25. Quandt D, Jasinski-Bergner S, Muller U, Schulze B, 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, et al. CXCL13 expression is prognostic and predictive for postoperative adjuvant chemotherapy benefit in patients with gastric cancer. Cancer Immunol Immunother. 2018;67(2):261-9.
27. Chang WS, Shen TC, Yeh WL, Yu CC, Lin HY, Wu HC, et al. Contribution of Inflammatory Cytokine Interleukin-18 Genotypes to Renal Cell Carcinoma. Int J Mol Sci. 2019;20(7).
28. Lee NH, Nikfarjam M, He H. Functions of the CXC ligand family in the pancreatic tumor microenvironment. Pancreatology. 2018;18(7):705-16.
29. Naji A, Menier C, Morandi F, Agaugue S, Maki G, Ferretti E, et al. Binding of HLA-G to ITIM-bearing Ig-like transcript 2 receptor suppresses B cell responses. J Immunol. 2014;192(4):1536-46.
30. Wang X, Li M. Correlate tumor mutation burden with immune signatures in human cancers. BMC Immunol. 2019;20(1):4.
31. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018;1711:243-59.
32. Giraldo NA, Becht E, Vano Y, Petitprez F, Lacroix L, Validire P, et al. Tumor-Infiltrating and Peripheral Blood T-cell Immunophenotypes Predict Early Relapse in Localized Clear Cell Renal Cell Carcinoma. Clin Cancer Res. 2017;23(15):4416-28.
33. Hakimi AA, Voss MH, Kuo F, Sanchez A, Liu M, Nixon BG, et al. 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(4):510-25.
34. Jung NC, Lee JH, Chung KH, Kwak YS, Lim DS. Dendritic Cell-Based Immunotherapy for Solid Tumors. Transl Oncol. 2018;11(3):686-90.
35. Laurell A, Lönnemark M, Brekkan E, Magnusson A, Tolf A, Wallgren AC, et al. 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, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39(4):782-95.

Supplemental Figures

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.

Figures

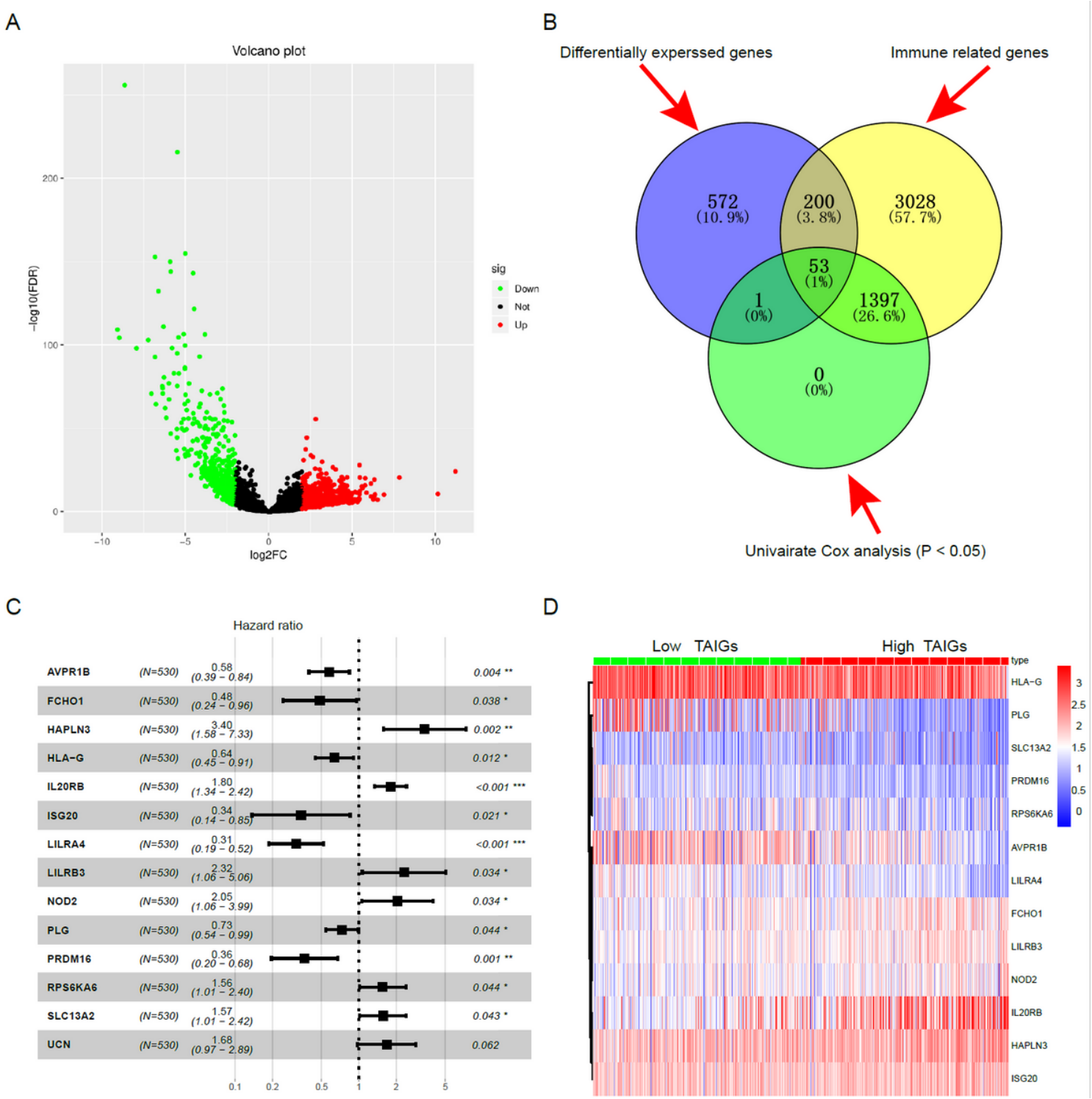

Figure 1

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.

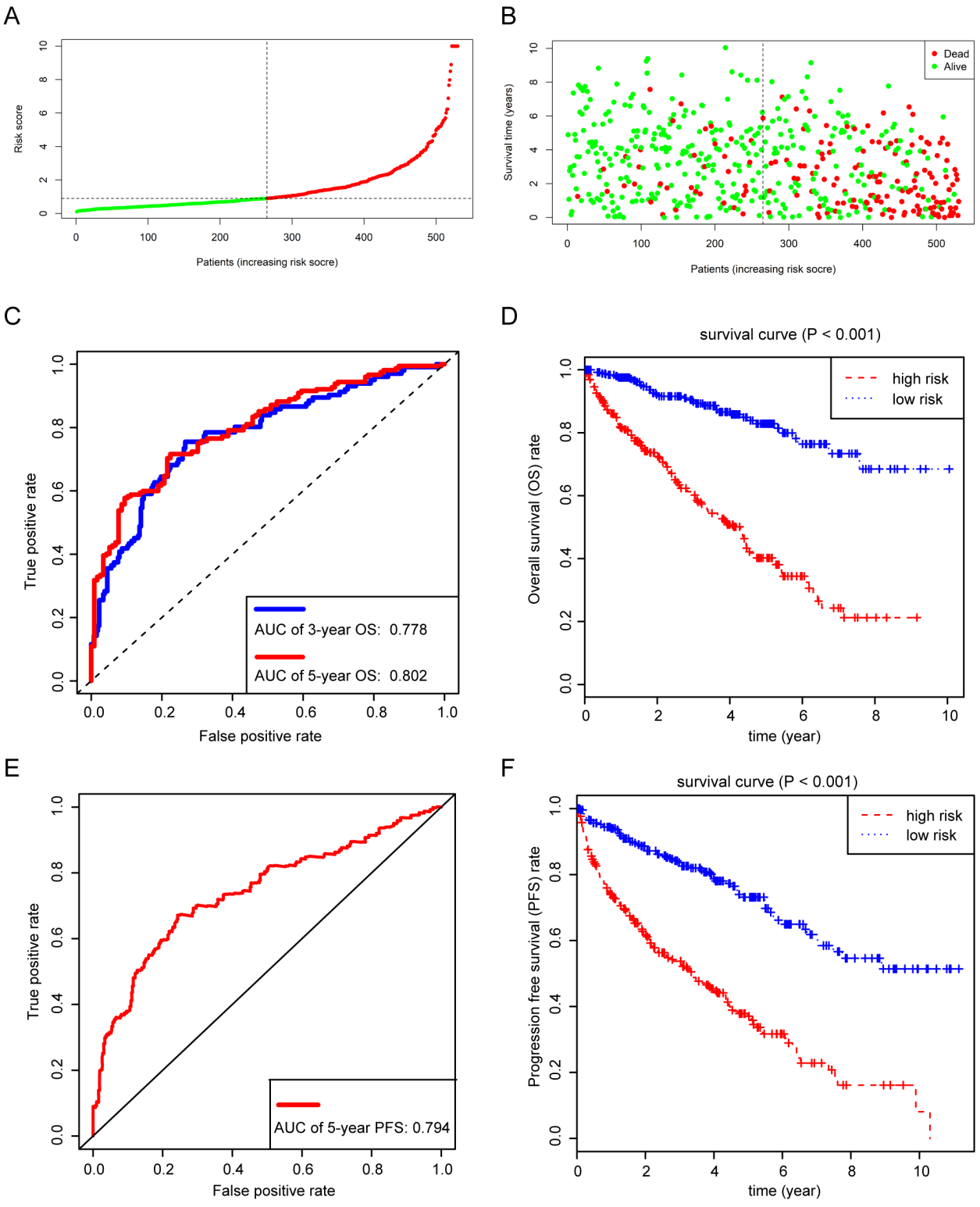

Figure 2

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.

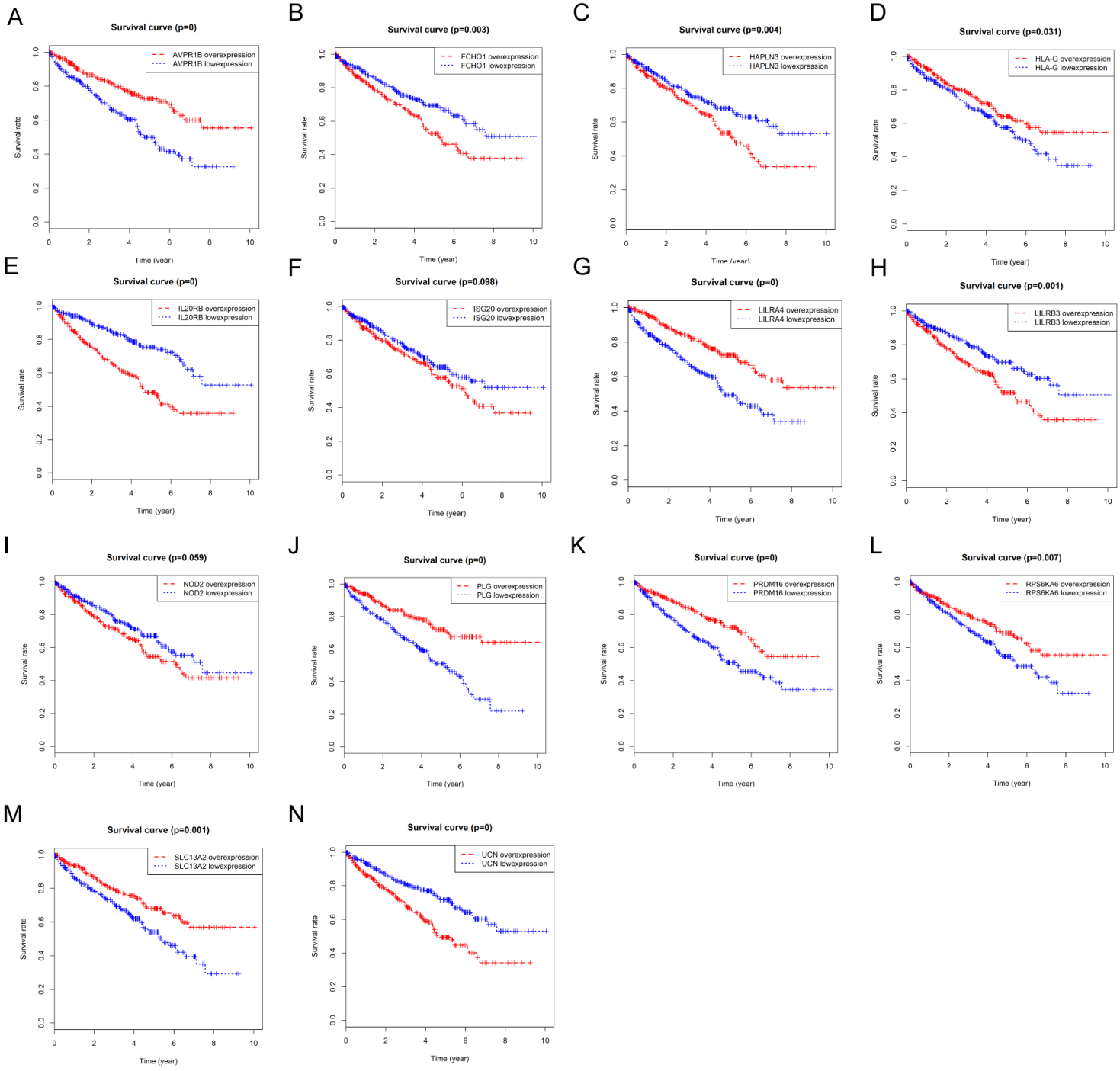

Figure 3

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

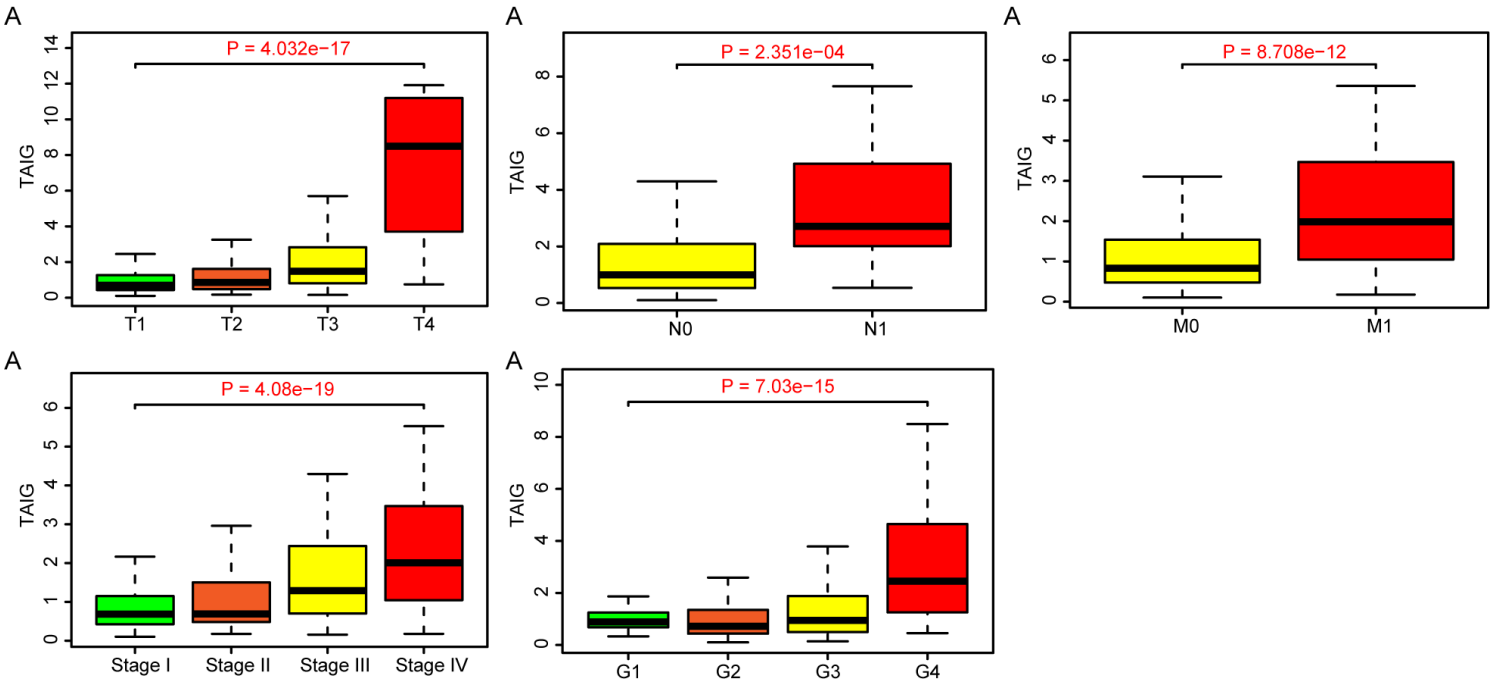

Figure 4

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), higher N stage (P = 2.351e-04), metastasis (P = 8.708e-12), pathological stages (P = 4.08e-19), as well as advanced tumor grades (P = 7.03e-15).

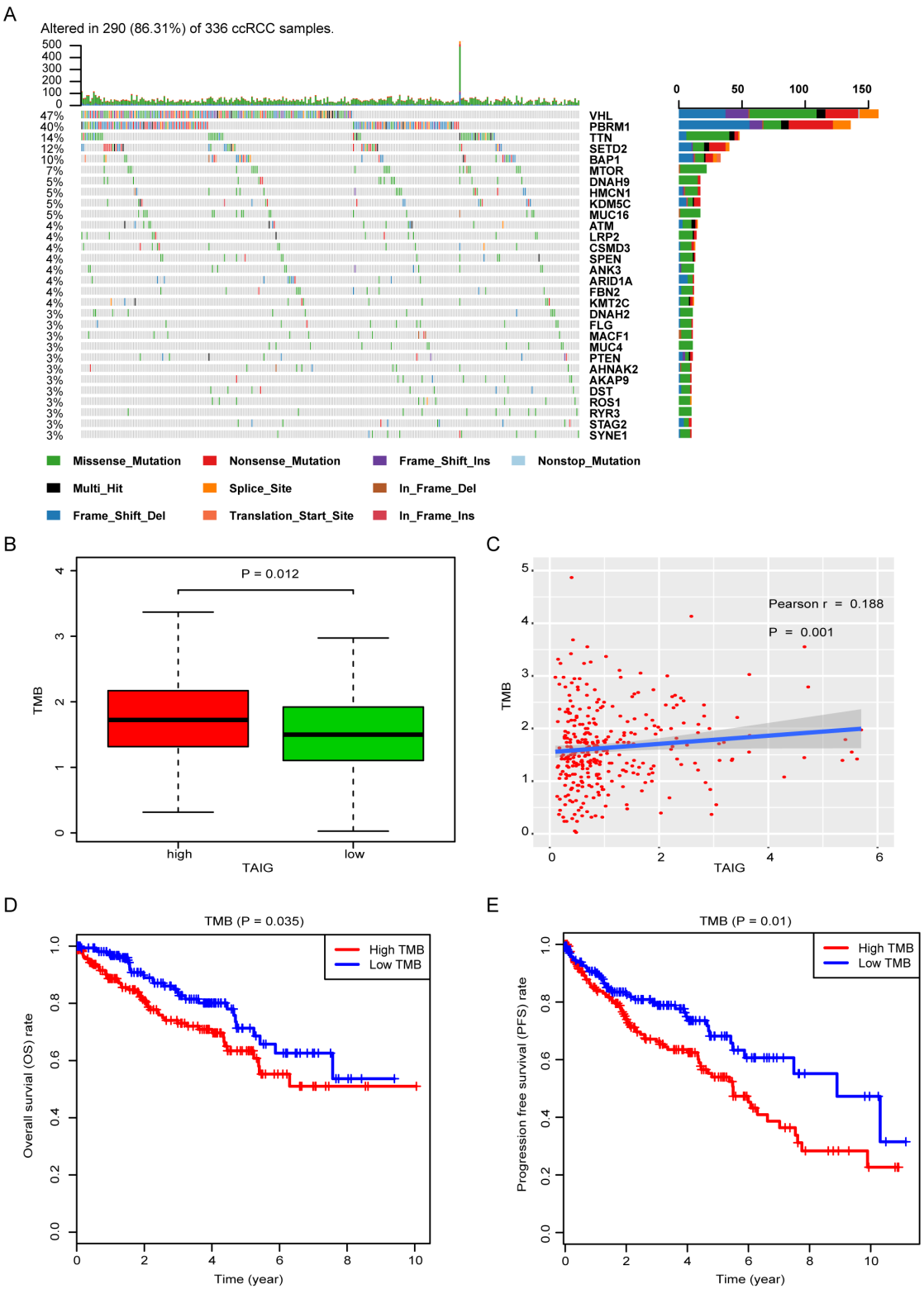

Figure 5

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).

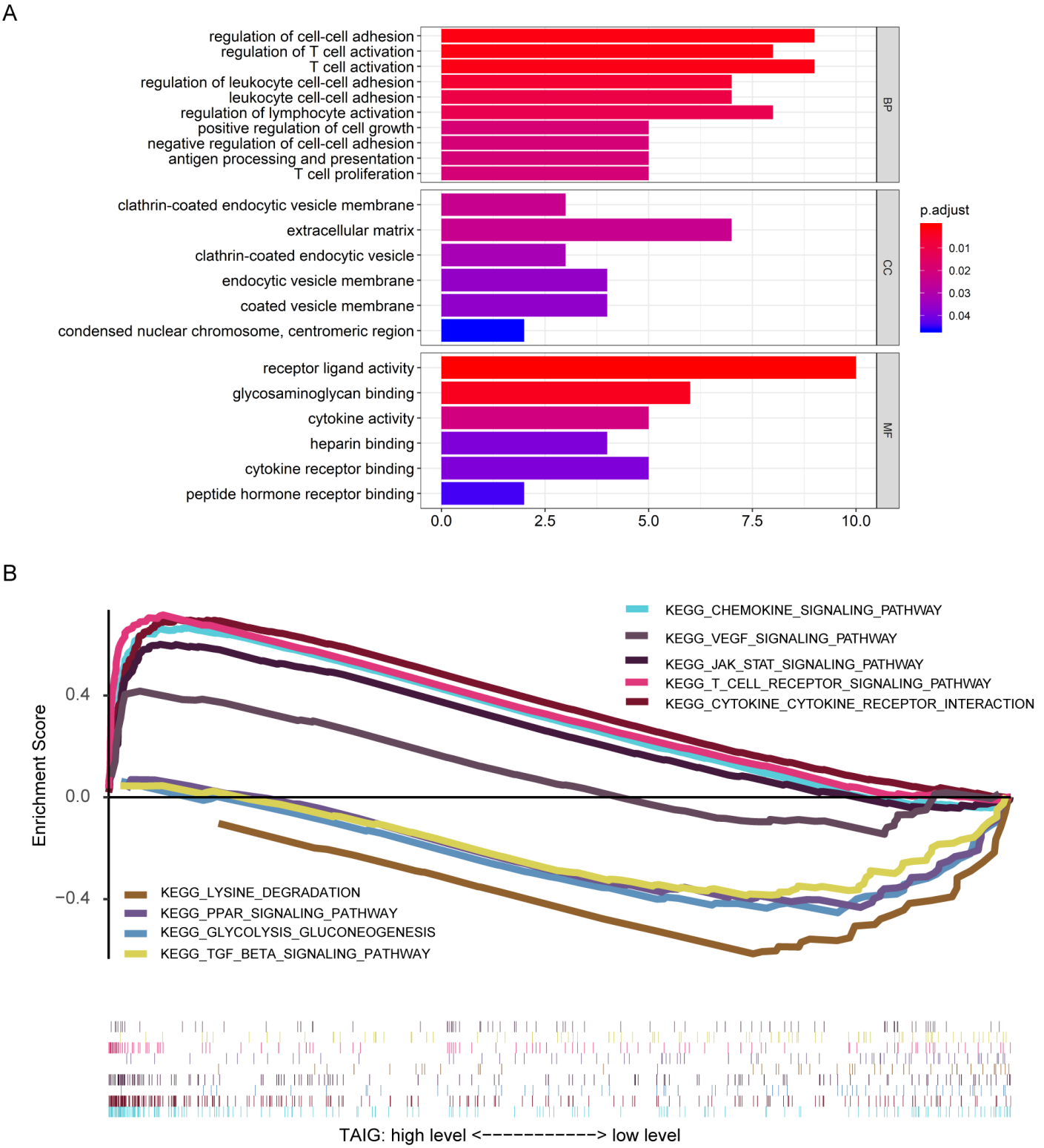

Figure 6

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.

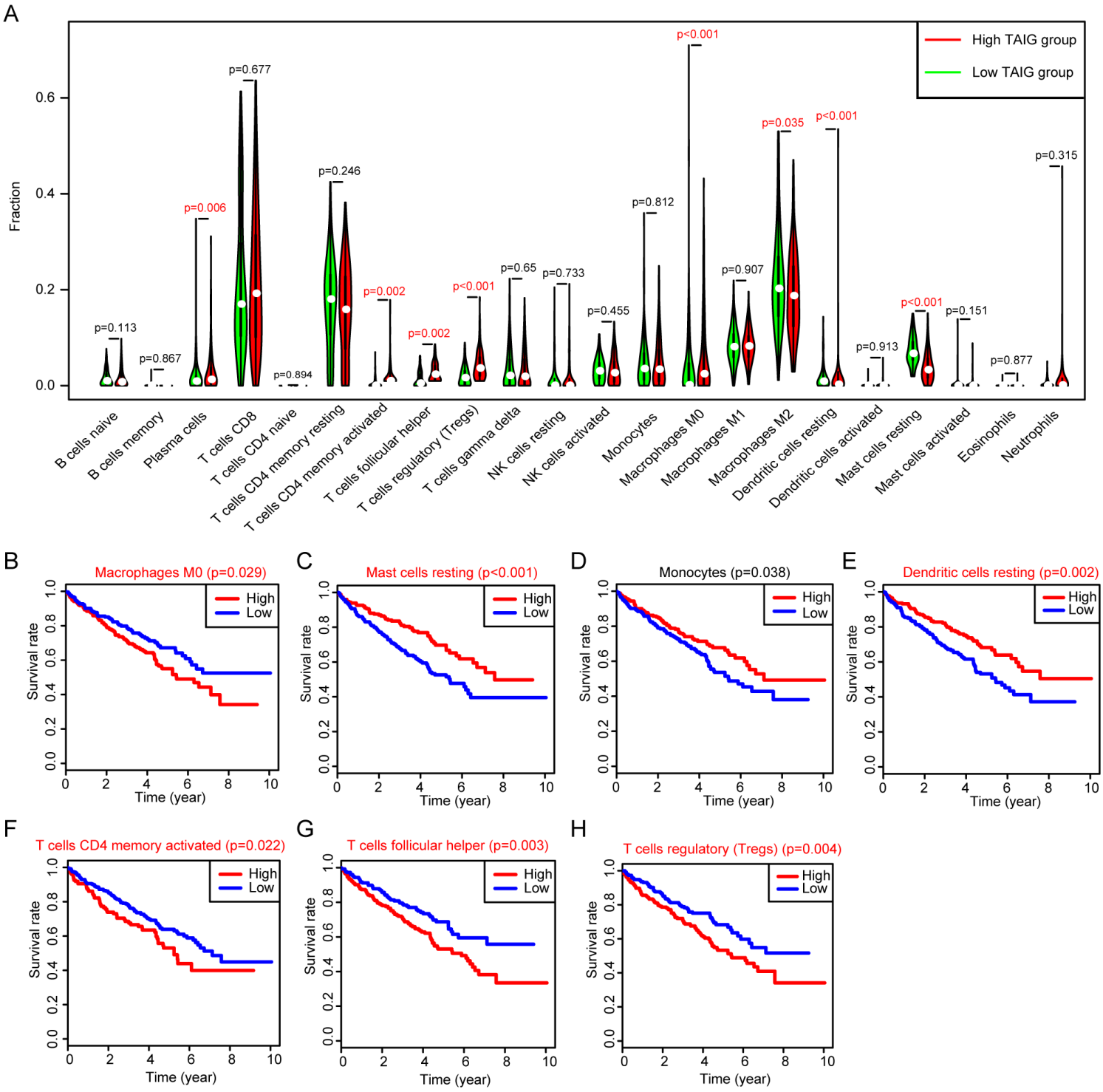

Figure 7

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.

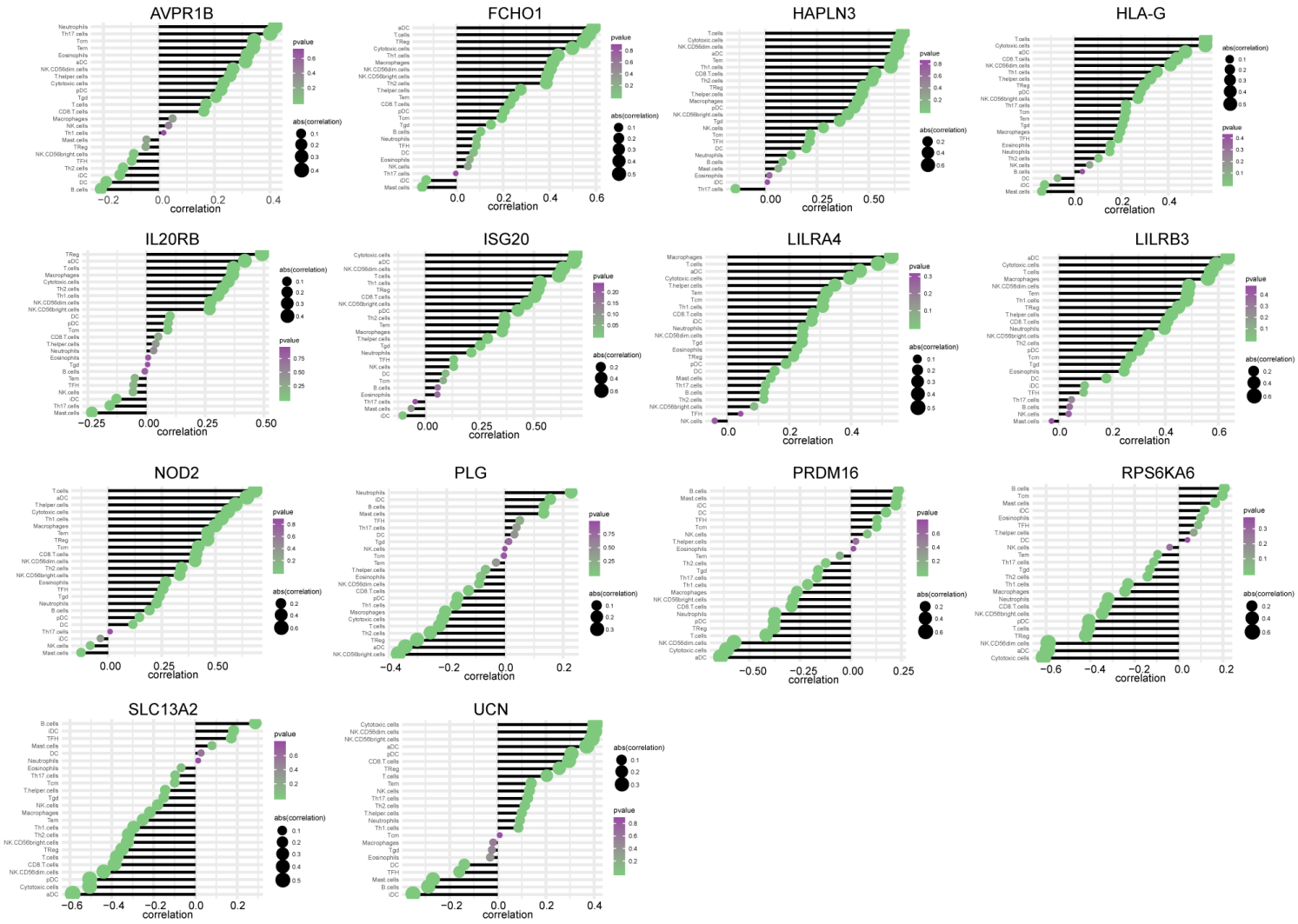

Figure 8

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

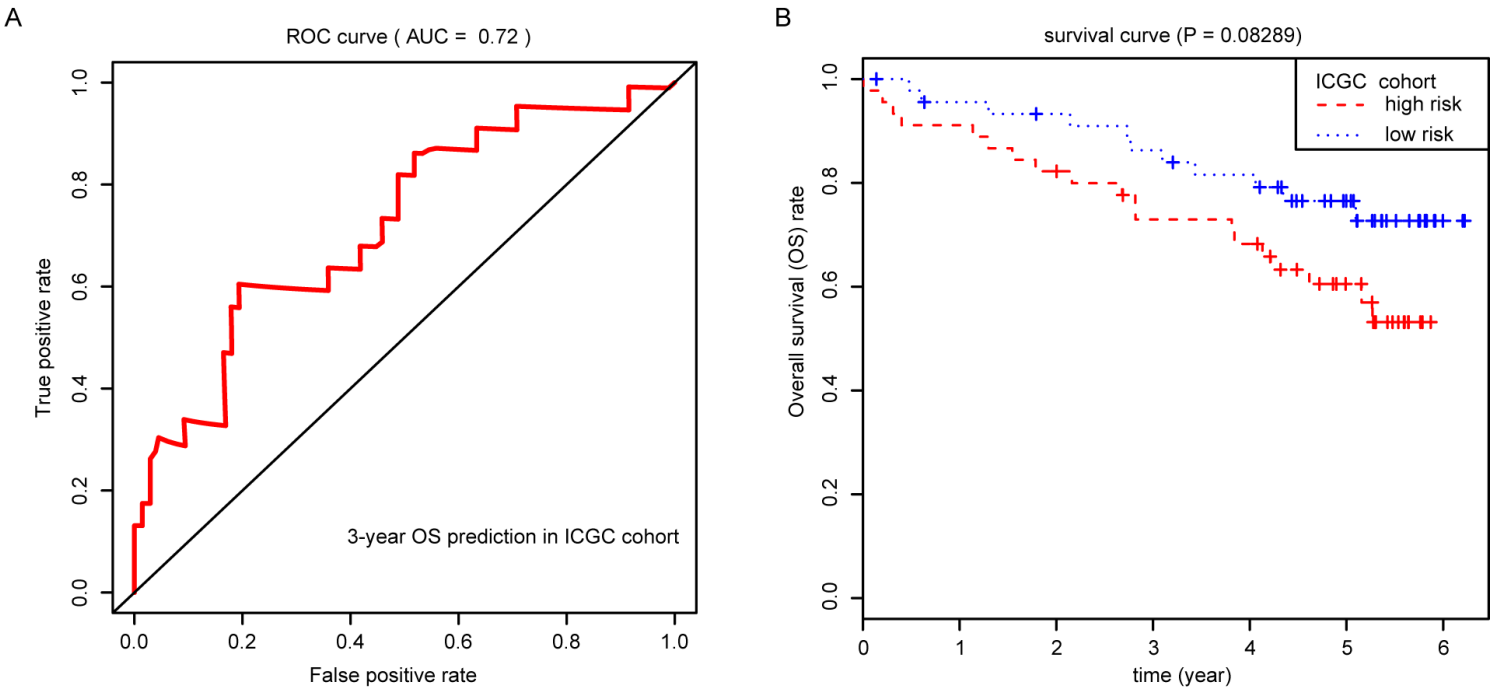

Figure 9

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.

Supplementary Files

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

- TableS2.xls
- FigureS2.tif
- TableS1.xls
- TableS6.xls
- FigureS1.tif
- TableS5.xls
- TableS4.xls
- TableS3.xls