

A novel prognostic predictor of immune microenvironment and therapeutic response in clear cell renal cell carcinoma based on angiogenesis-immune-related gene signature

Guixin Ding¹, Tianqi Wang¹, Gonglin Tang¹, Qingsong Zou, Gang Wu^{**}, Jitao Wu^{*}

Department of Urology, Yantai Yuhuangding Hospital, Qingdao University, Yantai, Shandong, China

ARTICLE INFO

Keywords:

Prognostic model
Tumor immune microenvironment
Immune
Immunotherapy
Angiogenic factor
Clear cell renal cell carcinoma

ABSTRACT

Background: Clear cell renal cell carcinoma (ccRCC), the most common type of RCC, typically produces no symptoms initially. Patients with ccRCC are at increased risk of developing advanced metastatic disease due to the absence of dependable and effective prognostic biomarkers. Therefore, it is particularly urgent to find optimal stratification of patients with ccRCC to distinguish the clinical benefits of different malignant degrees. Angiogenesis has a profound impact on the malignant behavior of renal cancer cells, and anti-angiogenic drugs have been applied to metastatic renal cancer patients. Moreover, immune function dysregulation is also a significant factor in tumorigenesis. We aim to construct a predictive model that combines angiogenesis and immune-related genes (AIRGs) to aid clinicians in predicting ccRCC prognosis. **Methods:** We gathered transcriptome and clinicopathology data from two datasets, the E-MTAB-1980 dataset and the Cancer Genome Atlas (TCGA). We utilized consensus clustering to find new molecular subgroups. A predictive model for the prognosis of angiogenesis-immune-associated genes (AIRGs) was conducted by the lasso and multivariate Cox regression analysis. The signature's predictive ability was then tested in different datasets. Meticulous scrutiny and comprehensive assessment were undertaken, both internally and externally, to establish the prognostic model. Analyses of immunogenomics were carried out to examine the relationship between risk scores and clinical/immune features, including immune cell infiltration, genomic alterations, and response to targeted and immunotherapy therapy.

Results: Our prognostic signature, comprising 4 AIRGs, stood as an independent prognostic factor for ccRCC, while risk scores emerged as a novel indicator for forecasting overall survival. Risk scores exhibited significant associations with various immunophenotypic factors, such as oncogenic pathways, antitumor response, different immune cell infiltration, antitumor immunity, and response to targeted and immunotherapy therapy.

Conclusions: AIRGs-based prognostic prediction model could effectively predict immunotherapy responses and survival outcomes of ccRCC.

* Corresponding author. Department of Urology, Yantai Yuhuangding Hospital, Qingdao University, NO. 20 East Yuhuangding Road, Yantai 264000, Shandong, China.

** Corresponding author. Department of Urology, Yantai Yuhuangding Hospital, Qingdao University, NO. 20 East Yuhuangding Road, Yantai 264000, Shandong, China.

E-mail addresses: wugd1209@163.com (G. Wu), wjturology@163.com (J. Wu).

¹ Guixin Ding, Tianqi Wang and Gonglin Tang contributed equally to this work as co-first authors.

1. Introduction

Renal cell carcinoma (RCC), originating from the renal tubular epithelium, ranks among the most prevalent urological malignancies. Renal cell carcinoma stands as the second most prevalent cancer affecting the urinary system, with a global incidence estimated at approximately 2 % [1,2]. Among its histological subtypes, clear cell renal cell carcinoma (ccRCC) prevails as the most frequent, constituting around 75 % of all RCCs [1,3,4]. Poor prognosis is typically associated with renal cancer patients who have advanced or metastatic stages [5]. The proportion of patients with a diagnosis of ccRCC who develop distant metastases at diagnosis is approximately 30 % , while the survival rate beyond 2 years post-diagnosis is only 40 % [6]. Surgical resection has become the first choice for treatment of ccRCC [7]. While surgical resection of early-stage ccRCC effectively resolves the condition, there are still 20%–30 % of patients developing tumor recurrences or metastasis following tumor resection surgically, and patients diagnosed with stage IV ccRCC have a five-year survival rate below 20 % [8–10]. For this reason, it is particularly urgent to search for an optimal stratification approach for ccRCC patients, which can accurately distinguish the clinical benefits associated with degrees of malignancy.

Solid malignant tumors require an adequate blood supply to support occurrence, development, and metastasis, called tumor angiogenesis, which encompasses the sprouting and proliferation of existing blood vessels in close proximity to the tumor [11,12]. Previous research indicated that ccRCC was correlated with a hyperangiogenic condition caused by vascular endothelial growth factor (VEGF) overproduction caused by inactivation of the von Hippel-Lindau tumor-suppressor gene [13]. Angiogenesis factors (AFs) consist of anti-angiogenic and pro-angiogenic factors to maintain a delicate balance in angiogenesis. However, tumors are characterized by a disruption of this balance and induce activate angiogenesis of equilibrium, which acts as a condition for growth and metastasis [14,15]. Treatments for metastatic RCC have been revolutionized by antiangiogenic targeted therapy [16,17]. Immunotherapy with checkpoint inhibitors is recognized as a promising strategy for improving survival outcomes in ccRCC [18,19]. However, on consideration of intratumoral heterogeneity, the sensitivity of patients to immunotherapy fluctuates widely, and immunotherapy can hardly achieve ideal therapeutic outcome [20,21]. At present, plenty clinical studies have been carried out on the combination of immunotherapy and anti-angiogenesis (AA) drugs for the treatment of ccRCC and preliminary curative effects have been obtained [22]. Further study is still required to obtain an improved understanding of the resistance of AA drugs or immunotherapeutic agents.

ccRCC has been closely linked to the infiltration of immune cells and angiogenesis, making it possible to predict patients' prognosis based on the expression of angiogenesis-immune-related genes (AIRGs) in tumor tissues. Differential expression analysis of AIRGs was performed on ccRCC patients whose expression profiles were incorporated from the TCGA database. Thus, we constructed this prognostic model to clarify the predictive significance of AIRGs signatures in ccRCC and determined its relationship with tumor immunity and therapeutic response. A four-AIRGs (IL4, BIRC5, RNASE2 and CCR10) signature could be indicative of prognosis, immune microenvironment, targeted medicine therapeutic response and tumor mutational burden (TMB) in ccRCC, followed by the verification of the E-MTAB-1980 dataset.

2. Materials and methods

2.1. Collecting and preprocessing the data

525 ccRCC samples and 72 tumor adjacent normal samples were downloaded from the TCGA-KIRC dataset as a training cohort through UCSC Xena (<http://xena.ucsc.edu/>). The validation cohort consisted of gene expression profile and precise clinical survival information of 101 ccRCC samples from the E-MTAB-1980 dataset, which were downloaded from the ArrayExpress database (<https://www.ebi.ac.uk/arrayexpress/>). Cases without corresponding pathological or clinical information were excluded. The characteristics of validation and training cohorts, including clinical information and demographic, were summarized in Table 1. From the ImmPort

Table 1

Clinical features of all eligible 626 ccRCC patients from TCGA and E-MTAB-1980 cohorts and 348 anti-PD-1 therapy patients with metastatic urothelial cancer IMvigor210 cohorts.

Variables	TCGA cohort	E-MTAB-1980 cohort	IMvigor210 cohort
Gender			
Male	343	77	272
Female	182	24	76
Age (years)			
≤65	347	57	NA
>65	178	44	NA
Grade			
1	17	13	NA
2	229	60	NA
3	205	23	NA
4	74	5	NA
Stage			
I	261	66	118
II	58	10	95
III	124	13	69
IV	82	12	66

database (<https://www.immport.org/home>), 1793 human immune-related genes (IRGs) were identified. We retrieved 2158 gene annotations linked to “angiogenesis” from the NCBI database (<http://www.ncbi.nlm.nih.gov/>). Then, we collected 3174 genes from the MSigDB (<http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>) database associated with “angiogenesis”. We excluded 532 genes that were duplicates, resulting in a final set of 4800 angiogenesis factors by combining genes from both resources. 348 metastatic urothelial cancer patients with anti-PD-1 therapy in the IMvigor210 database. In addition, the gene expression profile of a metastatic urothelial cancer cohort that administered atezolizumab was included in our analysis using the “IMvigor210CoreBiologies” package (<http://research-pub.gene.com/IMvigor210CoreBiologies>) to examine the predictive efficacy of AIRGs score on immunotherapy response [23]. R software was used to apply background correction and perform logarithmic conversion on all data.

DESeq2 package [24] was utilized to determine the false discovery rate (FDR) and log-transformed fold change (FC) for genes, enabling the identification of differentially expressed genes (DEGs) between normal and tumor sample. DEGs were identified by $|\log_2FC| \geq 1$, with the cutoff conditions established as the adjusted $P < 0.05$. Ultimately, the intersection of IRGs, DEGs and AFs were selected as potential genes to construct our prognosis model. Volcano and Venn diagrams were visualized by ggplot2.

2.2. Molecular subgroup identification

61 genes involved in ccRCC prognosis according to the univariate Cox regression analysis. We classified ccRCC patients using the consensus clustering analysis conducted using the “ConsensusClusterPlus” package in accordance with the 61 genes’ expression matrices. Furthermore, the Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression algorithm (ESTIMATE) [25] was employed to calculate the ESTIMATE score, tumor purity, stromal score, and immune score.

2.3. Immune analyses

The abundance of six distinct types of immune infiltrating cells (CD8⁺ T cells, CD4⁺ T cells, Neutrophil cells, Dendritic cells, B cells, and Macrophage cells) was determined through the implementation of the TIMER immune infiltrating analysis [26]. CIBERSORT [27], a deconvolution technique, was used to ascertain 22 different immune cell categories ratios. We obtained datasets from the molecular signature database [28] containing 28 kinds of immune cells and their associated 782 genes. Subsequently, we utilized single sample gene set enrichment analysis (ssGSEA) to investigate the abundance of 28 immune infiltrating cells in ccRCC samples. Besides, estimate the proportion of immune and cancer cells (EPIC) [29] and microenvironment cell populations-counter (MCP-COUNTER) [30] were additionally employed to evaluate the tumor immune microenvironment.

2.4. Functional enrichment and survival analysis of DEGs by clustering

The “limma” R package was used to identify DEGs between two clusters. The “ClusterProfiler” package [31] was used to perform gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) analysis for functional annotation. We conducted a difference and prognostic analysis between two clustering groups by “survival” package. Gene set enrichment analysis (GSEA) was performed using Broad Institute’s GSEA program. The GSEA analysis was conducted on the hallmark gene set “c2.cp.kegg.v6.2.symbols.gmt”.

2.5. Construction and validation of a prognostic model for AIRGs

The prognostic genes, which were initially filtered by the “glmnet” package, underwent downsizing through the LASSO algorithm (implementation of the least absolute shrinkage and selection operator) [32]. We proceeded obtained genes to perform multivariate Cox analysis and applied $P < 0.05$ to identify differential genes with independent prognostic capability. Subsequently, we proceeded to construct an elaborate multivariate Cox proportional hazards regression model based on these independent prognostic angiogenesis-immune-related differentially expressed genes (AIR-DEGs), aiming to forecast ccRCC patients prognosis. Using the following method, the risk score was calculated in the training and validation cohorts.

$$\text{Risk Score} = (4.4337 * \text{Exp IL4}) + (0.5463 * \text{Exp BIRC5}) + (0.4397 * \text{Exp RNASE2}) + (0.6688 * \text{Exp CCR10}).$$

Patients diagnosed with ccRCC were classified into two groups, namely high- and low-risk, based on the median risk score. The Kaplan-Meier (KM) curves were applied to determine whether there is a difference in overall survival between two groups. The model’s predictive efficiency was measured by generating ROC curves generated by the “survival ROC” package.

In this formula, β denotes the regression coefficient that corresponds to each gene, while Exp signifies the expression value associated with said gene. To evaluate the model stability, we take the TCGA cohort consisting of 539 samples as the training cohort, and the external validation cohort consisted of 101 samples from the E-MTAB-1980 database.

2.6. Clinicopathological correlation analysis and construction of nomogram

The correlation of 4 AIRGs with clinicopathological features was analyzed by the “survival” R package. Survival state, survival time, and four characteristics was integrated by the “rms” package. Nomograms and calibration curves based on the regression were established. Combining risk scores and clinical pathological characteristics, ccRCC patients’ prognosis was predicted by the nomogram.

2.7. Single nucleotide variants (SNV) and copy number variants (CNV) mutation analysis

The Mutation Annotation Format (MAF) file containing information on mutations in ccRCC patients was obtained by downloading it from the GDC client (<https://gdc-portal.nci.nih.gov/>). Fisher’s exact test was employed between high- and low-risk cohorts to detect genes with differential mutations. The “maftool” package [33] was used to calculate and visualize the somatic variants data of patients, aiming to ascertain the mutation burden of ccRCC in both high- and low-risk score cohorts.

2.8. Chemotherapeutic and immunotherapy response between different subtypes

The correlation between immunotherapeutic efficacy and risk score was analyzed via analyzing immune checkpoint genes (CTLA4, PD-L1 and PD-1), predictive markers of immunotherapy, in high-risk and low-risk score cohorts. The IMvigor210 dataset, which includes patients who treated with anti-PD-L1 therapy for urothelial cancer, was utilized to evaluate therapeutic efficacy between two subtypes.

Machine learning-based immunophenoscore (IPS) is a scoring system based on the weighted averaged Z scores that indicate immune-related gene expression in different immune cell types [34]. Patients can be predicted to the respond of immunotherapy with immune checkpoint inhibitors (ICIs) by IPS. High IPS scores reflect increased immunogenicity.

To forecast the IC50 (half-maximal inhibitory concentrations) of targeted therapy for drug sensitivity in low-risk and high-risk cohorts of ccRCC, the “pRRophetic” package was employed [35,36]. The drug sensitivity testing was conducted on common target drugs including Gefitinib, Sunitinib, Vorinostat and Sorafenib.

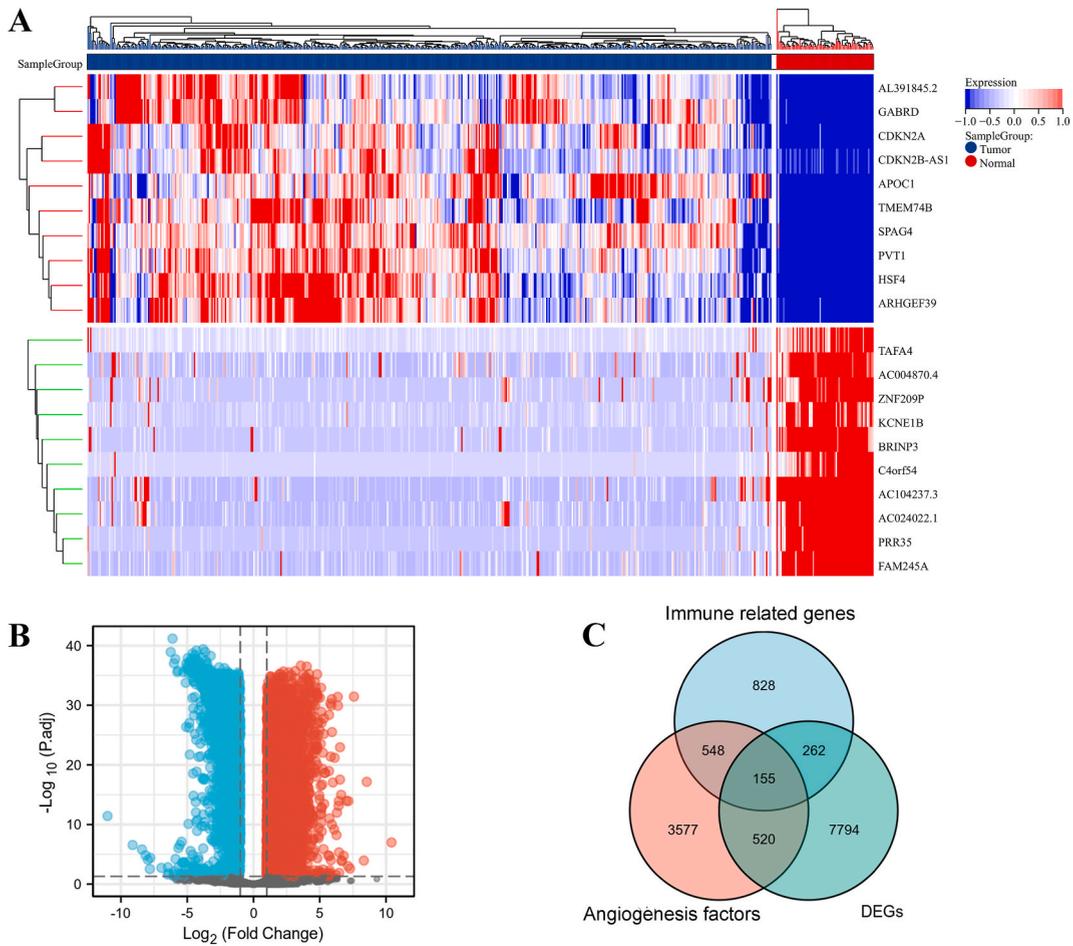


Fig. 1. Screening of differentially expressed genes. (A) Heatmap showing top 10 up-regulated and down-regulated genes expression between normal and tumor tissues. (B) Volcano map of DEGs between normal renal tissue and renal clear cell carcinoma in TCGA database. Red represents up-regulated genes, blue represents down-regulated genes. (C) Venn diagram showing 155 differential genes related to AIRGs based on three datasets. DEG, differentially expressed genes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.9. Statistical analysis

The survival analysis was performed by using the Kaplan-Meier plotter, and the predictive performance of the risk model was evaluated using the “survival ROC” package, which employs time-dependent receiver operating characteristics (ROC) analysis. Correlation between variables was determined by Pearson’s algorithm. Comparisons involving multiple groups were conducted using the Kruskal-Wallis test, while pairwise comparisons between two groups were performed using the Wilcoxon test. Subgroup analysis was conducted by regrouping patients based on characteristics including T, M, N stage, age, and sex, etc. A statistically significant difference was defined as $P < 0.05$. IBM SPSS 26.0 and R (version 3.6.3) were used for the statistical analysis.

3. Results

3.1. Differential expression of angiogenesis-immune-related genes

We have performed the DEG analysis between normal and tumor samples from ccRCC patients. The analysis identified a total of 8731 genes that exhibited a significant difference in expression, with $\text{adjust } P < 0.05$ and $|\log_2(\text{FC})| > 1$. Fig. 1A and B shows that tumor samples contain 2116 down-regulated DEGs and 6615 up-regulated DEGs compared with normal samples. Following this, Venn diagrams were used to analyze 8731 DEGs, 1793 immune-related genes (IRGs), and 4800 angiogenesis factors to determine which 155 angiogenesis-immune related genes were co-expressed (Fig. 1C).

3.2. Identification of two molecular subtypes of 155 co-expressed differential AIRGs

The 155 co-expressed angiogenesis-immune related genes (AIRGs) were evaluated by univariate cox analysis and it was determined that 61 of these genes had a significant correlation with OS ($P < 0.01$) (Supplementary Table 1). The TCGA database separated ccRCC patients into two groups by consensus clustering of the aforementioned 61 prognostic genes. When $K = 2$, the highest level of agreement within each group was observed, with cluster 1 consisted of 240 patients and cluster 2 comprised 267 patients (Fig. 2A–D, Supplementary Table 2). The heat map was employed to depict variations in AIRGs between two subtypes. Significant disparities were discovered between two clusters (Fig. 2E). Patients in cluster 1 had a significantly longer overall survival rate than those in cluster 2 ($P = 2.8\text{e-}05$; Fig. 2F). According to the results, 61 AIRGs categorized ccRCC individuals into two distinct molecular subcategories, each with varying survival rates.

3.2.1. Molecular subtypes differed in immune status and TIME

We then conducted an immune analysis to determine if there were any differences in immunity between two subtypes. Cluster 2 patients exhibited significantly elevated scores in terms of ESTIMATE score ($P < 0.001$), immune response ($P < 0.001$) and stromal

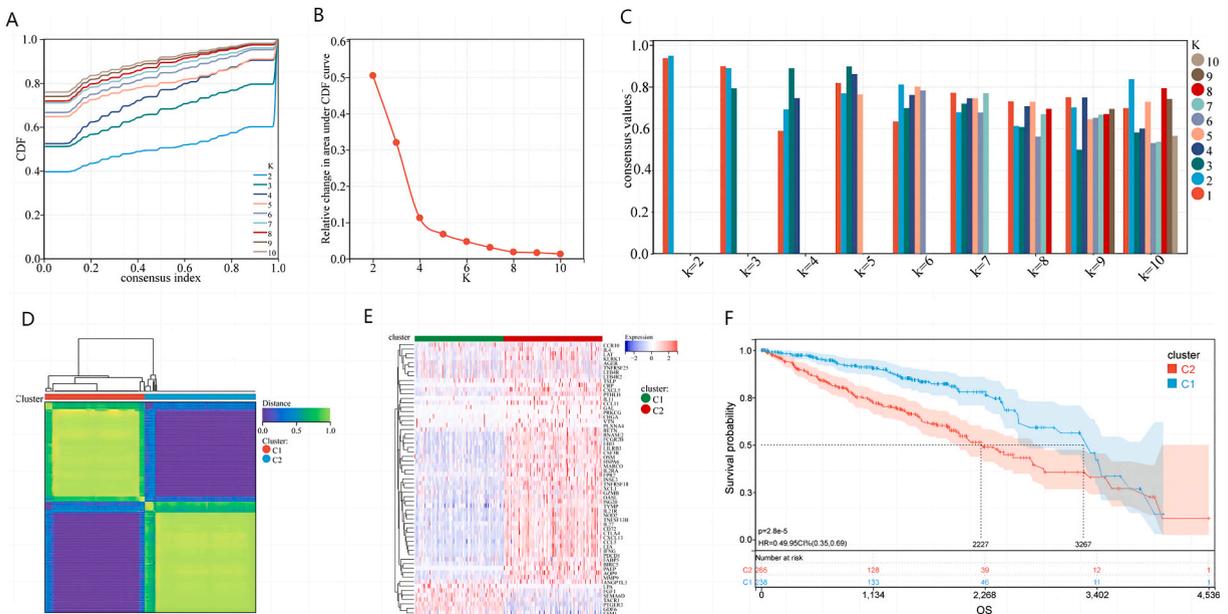


Fig. 2. Consensus cluster. The optimal consensus clustering value was identified as $K = 2$ by (A) cumulative distribution curve, (B) area under the distribution curve, (C) Sample clustering consistency and (D) clustering heat map. The optimal number of clusters was determined using the empirical cumulative distribution function plot. (E) Heatmap visualizing the expression of 61 AIRGs in the two subgroups. (F) Comparison of survival curves between the two subgroups.

score ($P < 0.001$) in comparison to cluster 1 using ESTIMATE algorithm (Fig. 3A). The TIMER algorithm revealed that the infiltration of B cells ($P = 6.0e-6$), CD8⁺ T cells ($P = 2.4e-16$), CD4⁺ T cells ($P = 1.8e-7$), macrophages ($P = 3.0e-4$), neutrophils ($P = 1.7e-28$) and dendritic cells ($P = 4.7e-39$) in cluster 2 was considerably higher compared to cluster 1 (Fig. 3B). Additionally, the Wilcoxon rank-sum test demonstrated that cluster 2 exhibited elevated levels of infiltration for T cells, B cells, CD8⁺ T cells, lymphocytes, monocytes, fibroblasts, and regulatory T cells compared to cluster 1 (Fig. 3C). Moreover, the EPIC algorithm demonstrated that the concentration of macrophages ($P < 0.001$), B cells ($P < 0.001$), cancer-associated fibroblasts (CAFs) ($P < 0.001$), CD4⁺ T cells ($P < 0.05$), and CD8⁺ T cells ($P < 0.001$) was significantly greater in cluster 2 compared to cluster 1. On the other hand, cluster 1 showed a significantly greater prevalence of endothelial cells compared to cluster 2 ($P < 0.001$) (Fig. 3D). Regarding 22 immune cells infiltration in Fig. 3E–F, the violin plot reveals that patients in cluster 2 have higher levels of T cells CD8, neutrophils, plasma cells, and macrophages M0 and lower levels of mast cells activated, NK cells resting, and monocytes (Fig. 3E–G). Furthermore, cluster 1 was found to have a lower immune profile, as indicated in the heat map (Fig. 3H), compared to cluster 2, as determined by the MCP counter algorithm. In this study, both molecular subtypes display distinctly different immune statuses.

3.3. Differentially expressed genes (DEGs) and functional analyses

Limma differential analysis was utilized to identify DEGs with $|\log_2(FC)| \geq 1$ and $FDR < 0.05$ between two clusters. This analysis led to the detection of 245 DEGs, with 63 down-regulated genes and 182 up-regulated genes in cluster 2 (Fig. 4A). The heat map illustrates the variations in gene expression observed between two clusters (Fig. 4B). Functional enrichment analysis was conducted to determine the mechanisms underlying signaling. Further analysis of DEGs indicated their enrichment in biological processes associated with the renal system and immune system, including cell activation, immune response pathways, regulation of cell adhesion, cell chemotaxis, and immunoglobulin production (Fig. 4C–F). Additionally, KEGG enrichment analysis revealed numerous immune-related signaling pathways, such as T cell receptor signaling pathways, NF kappa B signaling, PD-1 checkpoint, cytokine and cytokine receptor interactions in cancer (Fig. 4G). The results revealed that the development of ccRCC is associated with enriched biological processes and immune related signaling pathways. To further understand this relationship, a PPI analysis was conducted, resulting in the identification of six sub-models linked to tumor development and immunity (Fig. 4H). The GSEA and GSVA analyses indicated that cluster 2

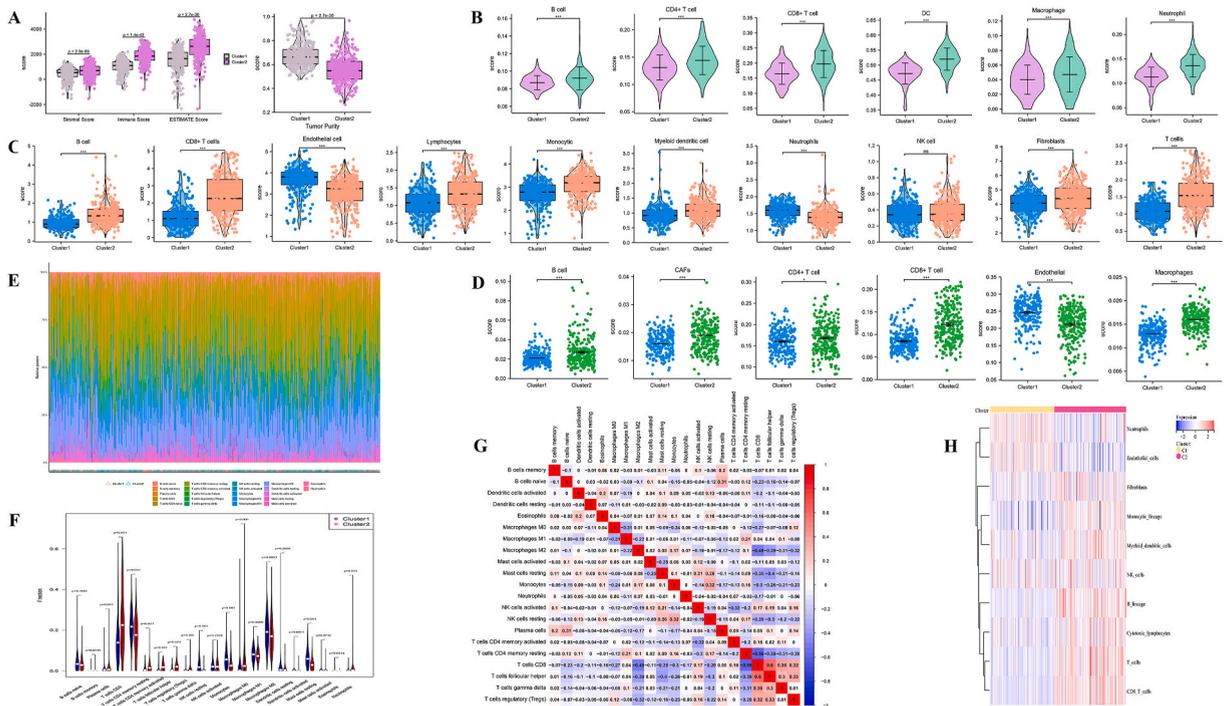


Fig. 3. Two clustered subgroups were analyzed for immune responses. (A) Stromal score, immune score, ESTIMATE score and tumor purity was calculated by ESTIMATE algorithm. (B) The abundance of immune infiltrating cells was evaluated by TIMER. (C) The abundance of 10 immune infiltrating cells was evaluated by the MCP algorithm. (D) The abundance of immune infiltrating cells was evaluated by the EPIC algorithm. (E) The specific 22 immune fractions represented by various colors in each sample through CIBERSORT algorithm were shown in stacked plots. Below Fig. 3E panel, the red area represents cluster 1, and the blue area represents cluster 2. This graph illustrates the relative proportions of immune cells in the two groups at the individual patient level, while right panel illustrates the disparities in immune cell composition at the risk group level. (F) The differences between cluster 1 and cluster 2 of the 22 immune cells by CIBERSORT algorithm are shown in the violin plot. (G) Correlation between 21 types of immune cells. (H) The heatmap depicts the enriching level of 10 immune related cells evaluated by the MCP algorithm. * $p < 0.05$; ** $p < 0.001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

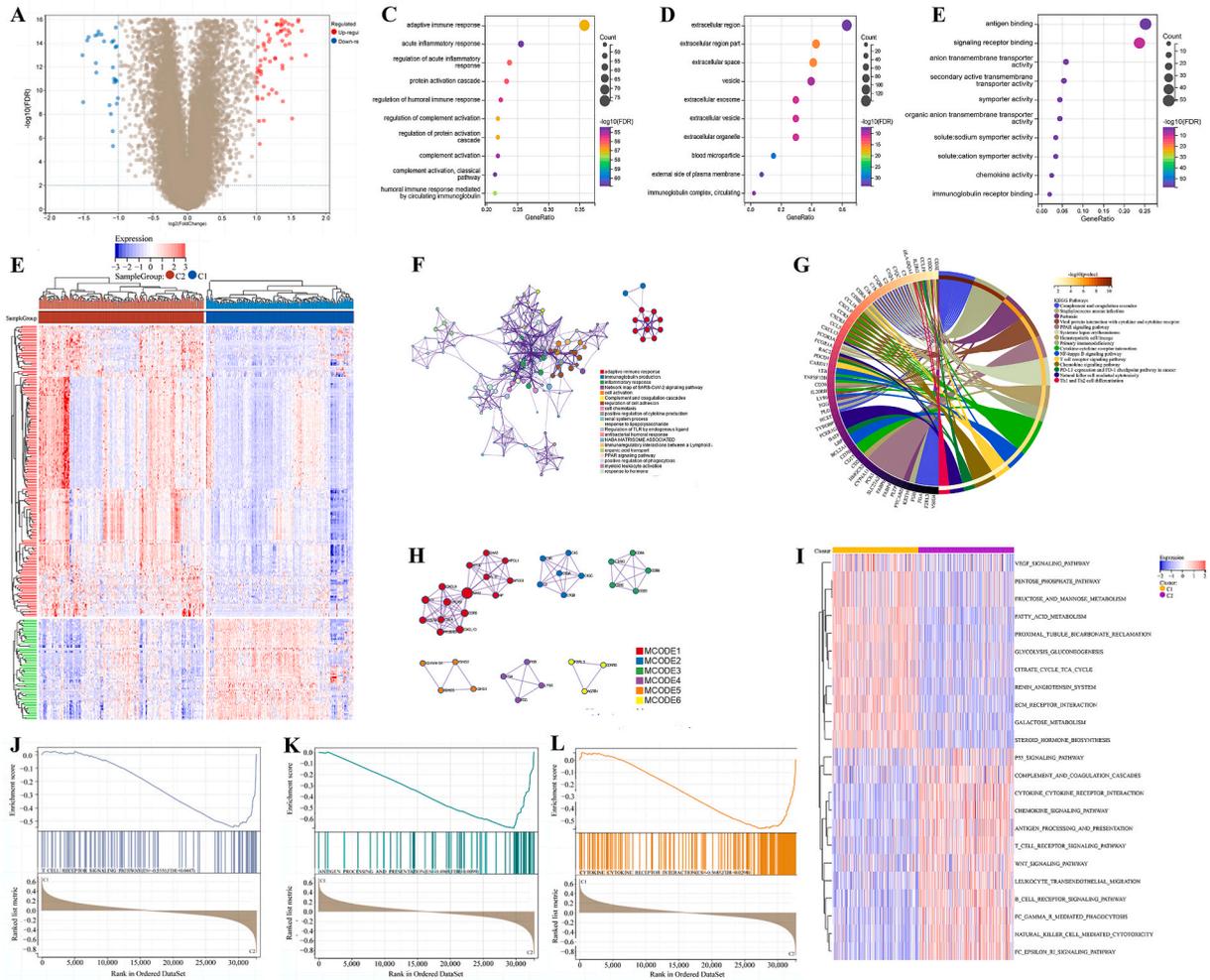


Fig. 4. DEGs analysis and functional analyzes. (A) Volcano plot illustrating the differences between the two subgroups. (B) The heat map shows the expression of 245 DEGs between the two subgroups. (C, D, E) Bubble diagram showing the biological processes (BP), cell component (CC) and molecular function (MF) enriched by gene ontology (GO) analysis. (F) GO analysis enriched biological processes through network visualization. (G) Signaling pathways enriched by the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis are shown in the circle plot. (H) PPI analysis of DEGs. (I) Results of GSEA analysis illustrated as a heatmap. (J–L) A visual representation of the result of the GSEA analysis is represented by GSEA plots.

exhibited significantly elevated expression of immune-related pathways including chemokine signaling pathways, antigen presentation and processing, cytokine-cytokine receptor interaction, and T cell receptor signaling pathways (Fig. 4I–L). Our findings indicate that DEGs are involved in immune regulation and may potentially contribute to the unfavorable prognosis observed in ccRCC.

3.4. Construction and validation of AIRGs prognostic risk model

We constructed a risk characteristic model to evaluate AIRGs’ prognostic value. The lasso analysis screened potential genes and identified 7 genes (Fig. 5A and B). Multivariate cox analysis was constructed to identify 4 genes, IL4, BIRC5, RNASE2, and CCR10, which were used to construct this model (Fig. 5C). Patients classified as high-risk exhibited elevated expression levels of the 4 candidate genes in comparison to those classified as low-risk (Fig. 5D). Meanwhile, individuals in the low-risk category demonstrated a superior overall survival rate when compared to those in the high-risk category (Fig. 5E). The hazard ratio of 4 genes exceeded 1, indicating their status as risk genes. The ROC curve for the risk model diagnosis produced satisfactory results (Fig. 5F). Time-dependent ROC analysis determined that the risk model had precise predictive capability for a five-year period, with AUC of 0.79, 0.73, and 0.76 for 1, 3, and 5 years respectively (Fig. 5F). Univariate Cox regression analysis was employed to calculate the correlation between risk scores and clinical characteristics factors (Fig. 5G). Incorporating risk scores, grade, gender, and age of the model, a multivariate Cox regression model was executed. Risk scores were demonstrated to be the standalone prognosticator of overall survival, with a hazard ratio of 2.10, 95 % CI:1.568–2.812, and $P < 0.001$ (Fig. 5H, Table 2). The multi-indicator ROC curve also indicated

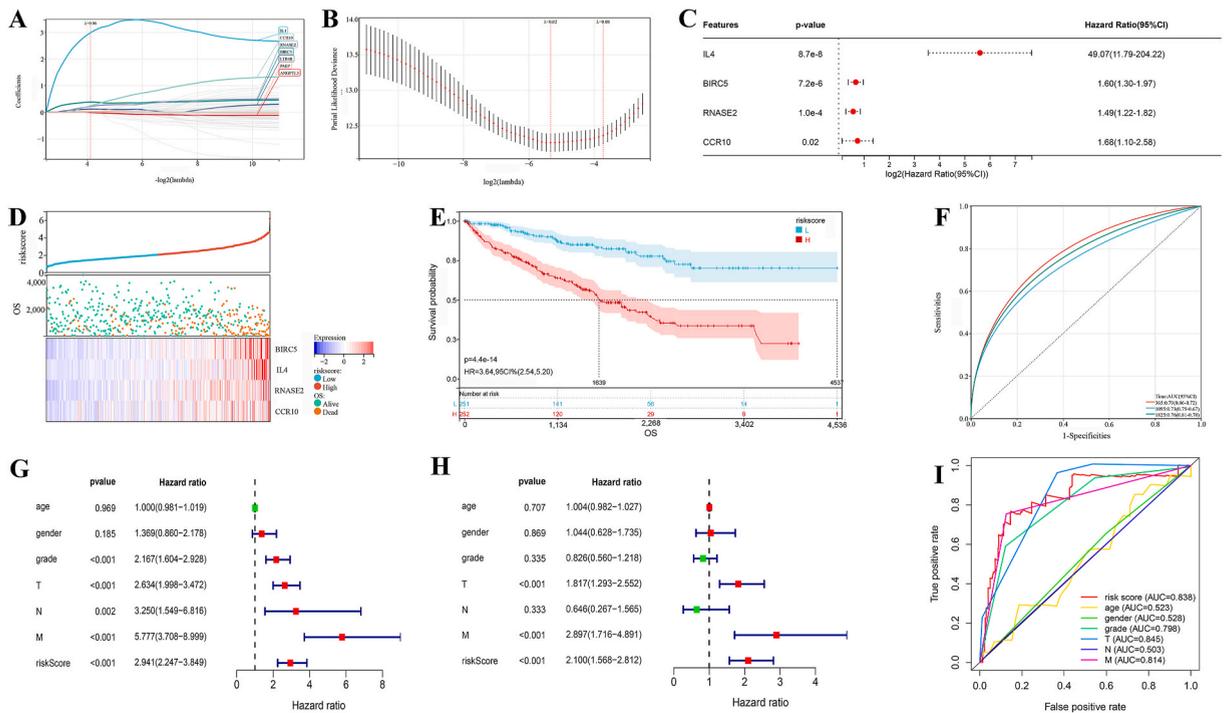


Fig. 5. Risk model construction for the training cohort. (A, B) LASSO analysis. (C) Multivariate Cox analysis. (D) Risk factor association diagram. There were nodes in the upper and middle graph that represented patients with high-risk score, and nodes in the lower and middle graph that represented patients with low-risk score. Each patient is represented by a cell and the color of the cells indicates whether a gene is up-regulated or down-regulated. (E) The survival curve of patients with ccRCC in the two groups. (F) Risk model ROC curve with time dependence. (G) Univariate and (H) multivariate Cox regression analysis of risk score and clinical factors. (I) Multi-indicator ROC curve of the risk score prognostic model. T, tumor; N, node; M, metastasis. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

that the prognostic model had excellent stability, with the AUC of 0.838 (Fig. 5I). In conclusion, we concluded that the AIRGs risk model provides an independent, reliable prediction of survival in ccRCC.

3.5. Survival analysis and differential expression of four AIRGs in ccRCC

The expression levels of four AIRGs (IL4, RNASE2, BIRC5 and CCR10) were compared between normal kidney tissues using the TCGA database and it was found that they were upregulated in the tumor tissue (Supplementary Fig. 1A). The results were further verified by the analysis of paired ccRCC samples that showed high expression of these genes in tumor samples (Supplementary Fig. 1B). The prognostic significance of these four AIRGs was subsequently assessed by employing the Kaplan-Meier plotter to examine the relationship between overall survival and gene expression. These findings revealed that four AIRGs were significantly linked to patient prognosis, with high expression being associated (Supplementary Fig. 1C).

3.6. Different immune status was observed in two groups

To compare the immune status of two groups, the ESTIMATE algorithm was employed. The analysis revealed that high-risk group

Table 2

Univariate and multivariate analysis showed that risk score was a significant independent prognostic factor.

Factors	Univariate cox regression			Multivariate cox regression		
	HR	95%CI	p-value	HR	95%CI	p-value
age	1	(0.981–1.019)	0.969	1.004	(0.982–1.027)	0.707
gender	1.369	(0.860–2.178)	0.185	1.044	(0.628–1.735)	0.869
grade	2.167	(1.604–2.928)	< 0.001	0.826	(0.560–1.218)	0.335
T	2.634	(1.998–3.472)	< 0.001	1.817	(1.293–2.552)	< 0.001
N	3.25	(1.549–8.816)	0.002	0.646	(0.267–1.565)	0.333
M	5.777	(3.708–8.999)	< 0.001	2.897	(1.716–4.891)	< 0.001
riskScore	2.941	(2.247–3.849)	< 0.001	2.1	(1.568–2.812)	< 0.001

exhibited significantly elevated stromal scores ($P = 7.8e-09$), immune scores ($P = 1.3e-20$) and estimate scores ($P = 3.6e-19$) (Fig. 6A). Furthermore, tumor microenvironment analysis indicated higher tumor purity in the low-risk group, indicating a lower density of ccRCC cells in this group (Fig. 6A). The heat map of the seven immune algorithms further indicated the high-risk cohort had a greater abundance of immune cells (Fig. 6B). TIMER algorithm was used to calculate the immune status of both groups and found that neutrophil, dendritic cell, CD4⁺ T cell and CD8⁺ T cell were significantly higher in the high-risk cohort, while there was no notable disparity in B cells and macrophages between two groups (Supplementary Fig. 2B). The results from EPIC and MCP Counter algorithms indicated that high-risk cohort exhibited a notable rise in immune cells in comparison to low-risk cohort (Supplementary Figs. 2B and C). The EPIC and MCP Counter algorithms also showed an increase in cancer-associated fibroblasts (CAFs) in high-risk cohort and an increase in endothelial and neutrophil cells in low-risk cohort (Supplementary Figs. 2B and C). To present the proportion of 22 different

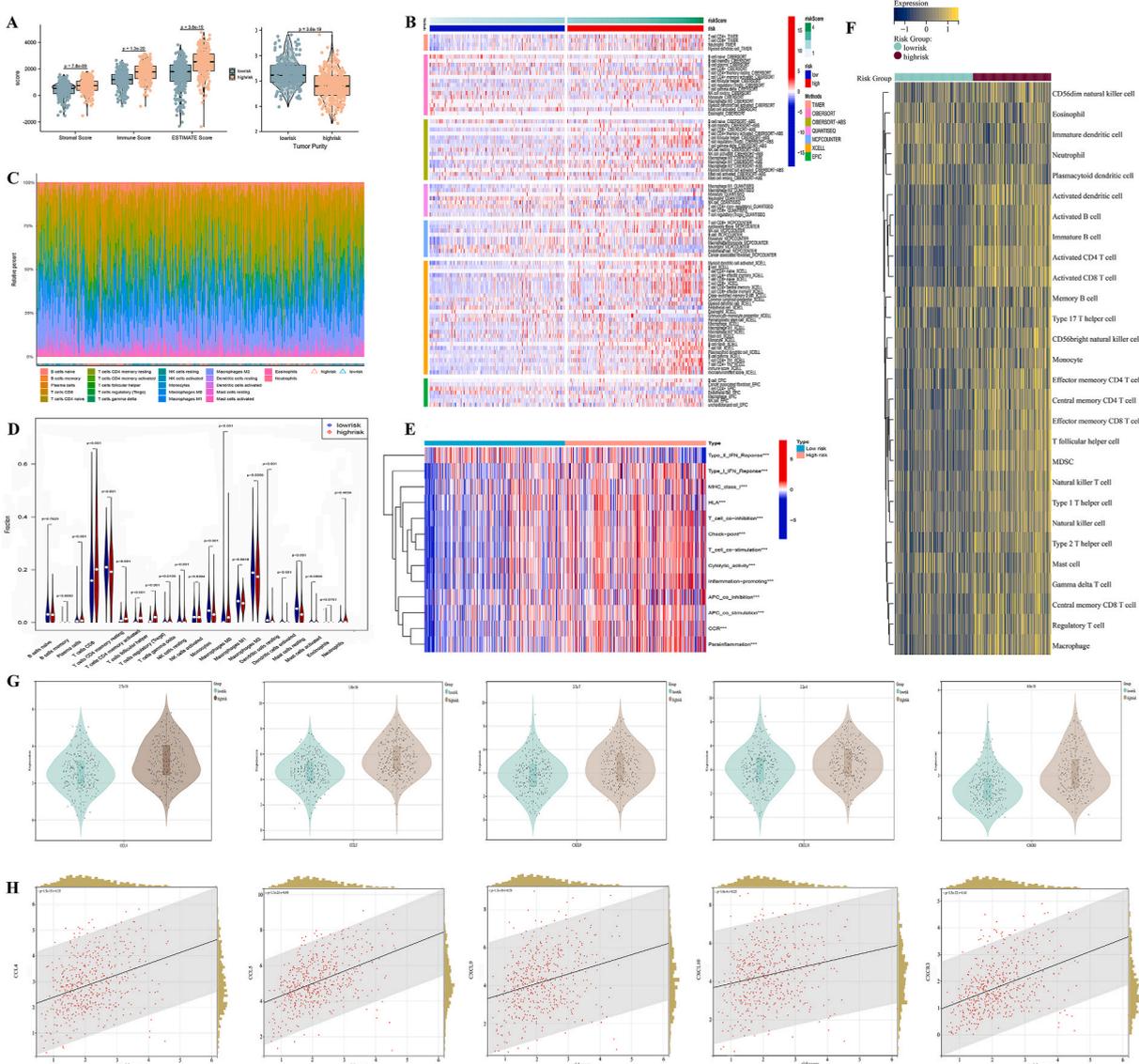


Fig. 6. (A) Stromal score, immune score, ESTIMATE score and tumor purity calculated by ESTIMATE algorithm. (B) Heat map of seven immune algorithms. (C) The specific 22 immune fractions represented by various colors in each sample through CIBERSORT algorithm were shown in stacked plots. Below the Figure, the red area represents high-risk group, and the blue area represents low-risk group. This graph illustrates the relative proportions of immune cells in the two groups at the individual patient level, while right panel illustrates the disparities in immune cell composition at the risk group level. (D) The differences between the low- and high-risk groups of the 22 immune cells by CIBERSORT algorithm are shown in the violin plot. (E) An illustration of the difference between low- and high-risk groups in terms of immune function. (F) The infiltration of immune cells in low- and high-risk group was analyzed by ssGSEA algorithm. (G) Violin plots of expression level of CCL4, CCL5, CXCR3, CXCL9 and CXCL10 between two risk score subtypes. (H) Correlation of the risk score with T cell and NK cell-recruiting chemokines. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

immune cells in ccRCC (Fig. 6C and D), the CIBERSORT algorithm employed a Box and violin plot, while the heat map was utilized to demonstrate the correlation among 21 varieties of immune cells (Supplementary Fig. 2D). The heat map depicted that the immune function of high-risk cohort displayed a generally elevated state ($P < 0.001$) (Fig. 6E). Our ssGSEA algorithm indicated that high-risk cohort displayed high infiltration of the majority of immune cells (Fig. 6F). Researches have indicated that cDC1 cells have an important function in activating and driving CD8⁺ T cell by activating CXCL10 [37,38]. Chemokines CCL4 and CCL5, which are key modulators of CCR5 expression, influence cDC1 recruitment to tumors by activating CCR5 [39,40]. The chemokines CXCR3, CXCL9, and CXCL10 have been documented to affect T cell infiltration and NK cell recruitment [41]. The expression levels of the chemokines CXCL10, CCL5, CCL4, CXCR9 and CXCL3 were compared between the two subtypes and the correlation between genes and risk scores were investigated. Results indicated that high-risk cohort exhibited elevated levels of expression compared to the low-risk cohort ($P < 0.05$) (Fig. 6G), and significant positive correlations were observed between the chemokines (CXCL10, CXCL9, CCL4, CCL5, and CXCR3) and risk scores ($P < 0.05$) (Fig. 6H). Based on these results, the constructed prognostic model offers potential immune prediction in patients with ccRCC.

3.7. Stratified prognostic model of OS based on clinical parameters

To evaluate clinical characteristics of the four AIR-DEG signatures in ccRCC, the stratified analysis was performed on ccRCC patients in the TCGA dataset, with adjustments made for lymph node involvement, gender, age, tumor stage, and distant metastases. The results, as shown by Kaplan-Meier plotter, revealed that the overall survival for patients with ccRCC in high-risk cohort was shorter compared to the low-risk cohort. (Supplementary Fig. 3). The pathological and clinical features are susceptible by the study population size. However, it should be noted that Kaplan-Meier survival curves of Stage II (Supplementary Fig. 3F), T stage 4 (Supplementary

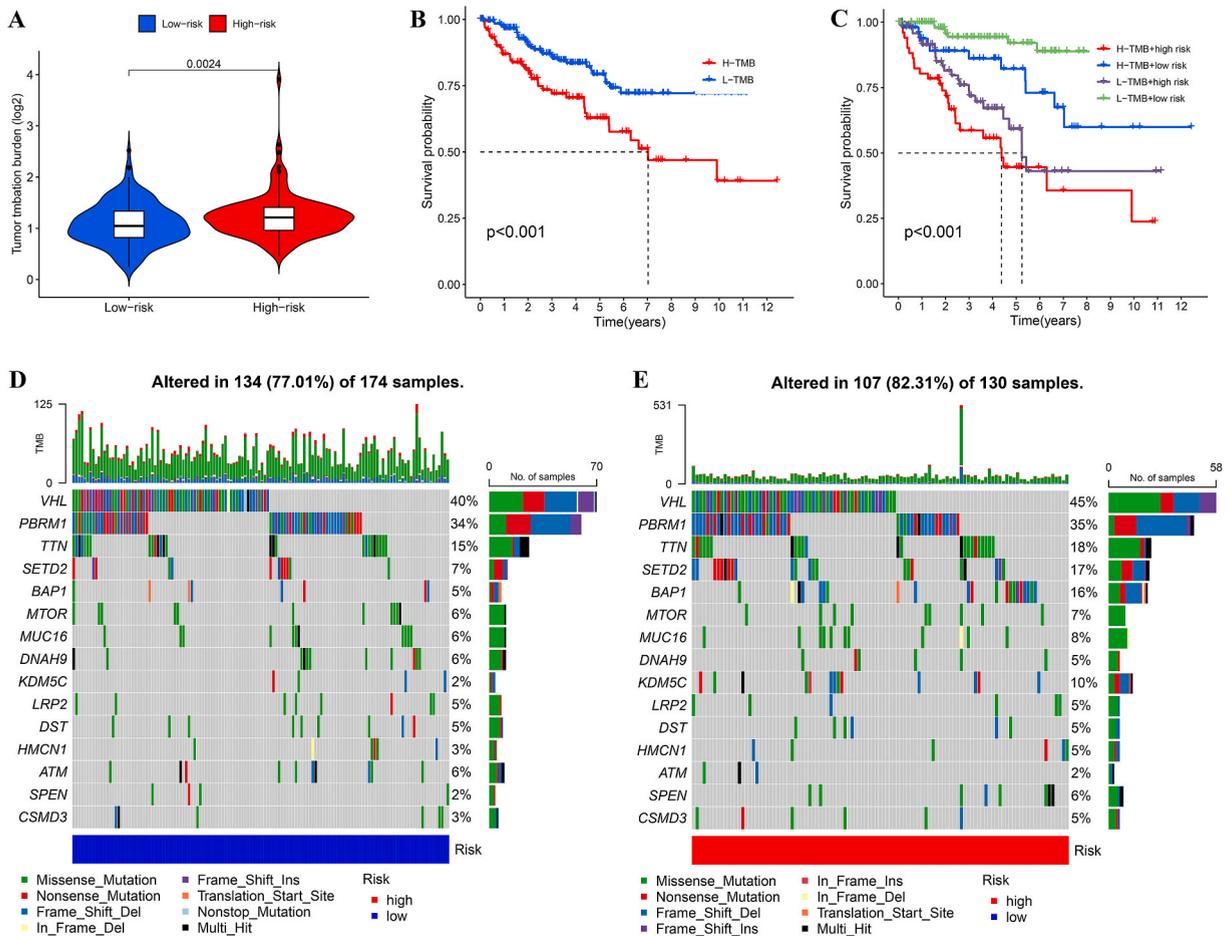


Fig. 7. The correlation between the AIRGs prognostic signature and TMB. (A) The box plot for TMB levels among patients in the low- and high-risk groups. (B) Kaplan-Meier curves for the low- and high-TMB of ccRCC patients. (C) Kaplan-Meier curves for ccRCC patients by TMB status in the low-risk and high-risk groups. (D) The oncoPrint was constructed based on CNV profile in the low-risk scores of ccRCC patients. (E) The oncoPrint was constructed based on CNV profile in the high-risk scores of ccRCC patients. Individual patients are represented in each column. TMB, tumor mutational burden; CNV, copy number variation.

Fig. 3L) and N stage 1 (Supplementary Fig. 3P) samples were not significant, which may be attributed to a relatively small number of these three groups. These results indicate that ccRCC patients can be predicted based on their signatures without taking into account their clinical characteristics.

3.8. Relationship between clinical parameters and prognostic model for OS

Our study aimed to investigate the potential impact of the prognostic model utilizing these four AIRGs on the progression of ccRCC by looking at the association between clinical parameters and the prognostic model. The study indicated the prognostic models was not significantly correlated with age (Supplementary Fig. 4A). Nevertheless, the risk score exhibited a notably greater magnitude in males compared to females (Supplementary Fig. 4B). Additionally, as T, M, and N stages increase, the risk score increases as well ($P < 0.05$) (Supplementary Figs. 4C–E). The study also identified an upward trend in the risk score corresponding to higher grades, stages, and unfavorable prognosis of ccRCC (Supplementary Figs. 4F–H). The results indicate ccRCC progression and the prognostic model for overall survival appear to be correlated significantly.

4. Prognostic signature of AIRGs in relation to tumor mutational load

The objective of our study was to investigate the relationship between the AIRG prognostic signatures and the level of tumor mutational burden (TMB) in ccRCC. The results revealed the high-risk category exhibited higher TMB levels in contrast to the low-risk category ($P < 0.01$, Fig. 7A). Additionally, the Kaplan-Meier plotter revealed a worse survival probability in high-TMB patients as compared to low-TMB patients ($P < 0.001$, Fig. 7B). Predictive profiles of AIRGs were evaluated in relation to TMB status. In this study, survival analyses were conducted on four groups: low-TMB/high-risk group, high-TMB/high-risk group, low-TMB/low-risk group, and high-TMB/low-risk group. The four groups showed significant differences ($P < 0.001$) (Fig. 7C). Results indicate the relationship between the somatic mutation rate and risk scores. To further explore the correlation between risk scores and genomic characteristics, we conducted a copy number variation (CNV) analysis. Results indicated that the high-risk category exhibited an elevated mutation rate of BAP1 ($P = 0.048$) and KDM5C ($P = 0.023$) and markedly low frequency of PBRM1 mutations ($P = 0.007$) (Fig. 7D and E and Table 3).

4.1. Verification cohort prognosis was correlated with ccRCC risk model

To validate the constructed model, further validation was performed of ccRCC patients with a verification cohort. Using the previously mentioned formula, the verification group was classified into low- and high-risk categories (Supplementary Fig. 5A). Heatmaps were used to demonstrate the expression of the four candidate genes (Supplementary Fig. 5A). Survival analysis indicated that patients in the high-risk category had a worse prognosis compared to the low-risk category ($P = 0.021$; Supplementary Fig. 5B). Additionally, the AUC of the 4 AIRGs signature was 0.68, 0.70, and 0.65 at 1, 3, and 5 years respectively (Supplementary Fig. 5C). These results demonstrate a correlation between the established risk model and the prognosis of ccRCC in the verification cohort.

Table 3
Difference of gene and pathway mutations with significance between high-risk and low-risk groups of TCGA.

Gene	low-risk	high-risk	P value
VHL (%)			0.064
WT	162 (64.5)	182 (72.2)	
Mutation	89 (35.5)	70 (27.8)	
TTN (%)			0.084
WT	204 (81.3)	219 (86.9)	
Mutation	47 (18.7)	33 (13.1)	
PBRM1 (%)			0.007
WT	174 (69.3)	201 (79.8)	
Mutation	77 (30.7)	51 (20.2)	
SETD2 (%)			0.202
WT	234 (93.2)	227 (90.1)	
Mutation	17 (6.8)	25 (9.9)	
BAP1 (%)			0.048
WT	240 (95.6)	230 (91.3)	
Mutation	11 (4.4)	22 (8.7)	
MUC16 (%)			0.427
WT	240 (95.6)	237 (94.0)	
Mutation	11 (4.4)	15 (6.0)	
KDM5C (%)			0.023
WT	246 (98.0)	237 (94.0)	
Mutation	5 (2.0)	15 (6.0)	
ATM (%)			0.083
WT	239 (95.2)	247 (98.0)	
Mutation	12 (4.8)	5 (2.0)	

4.2. Construction and calibration of an integrated nomogram

To enhance the precision of prognostication for ccRCC survival, a nomogram incorporating clinical features and the established risk model was developed. The nomogram, shown in Fig. 8A, assigned specific scores to pathological characteristics and the risk score based on their impact on patient prognosis. Both the training and verification cohorts validated the nomogram to ensure its accuracy and reliability. Our nomogram demonstrated precise accuracy in diagnosing the model, as evidenced by the calibration and C-index curve (Fig. 8B–I). Our nomogram achieved the C-index of 0.7933 (95%CI:0.7596–0.8271) in the training cohort (Fig. 8B–D), and the AUC values were 0.887, 0.814, and 0.802, respectively, for the training cohort at 1, 3, and 5 years (Fig. 8E). The nomogram achieved a C-index of 0.7231 (95%CI: 0.6167–0.8295) (Fig. 8F–H), and the AUC values for ROC curves were 0.843, 0.748, and 0.742, respectively, for the validation cohort at 1, 3, and 5 years (Fig. 8I). Indeed, our results provide strong evidence that the integrated nomogram is capable of accurately predicting ccRCC prognosis.

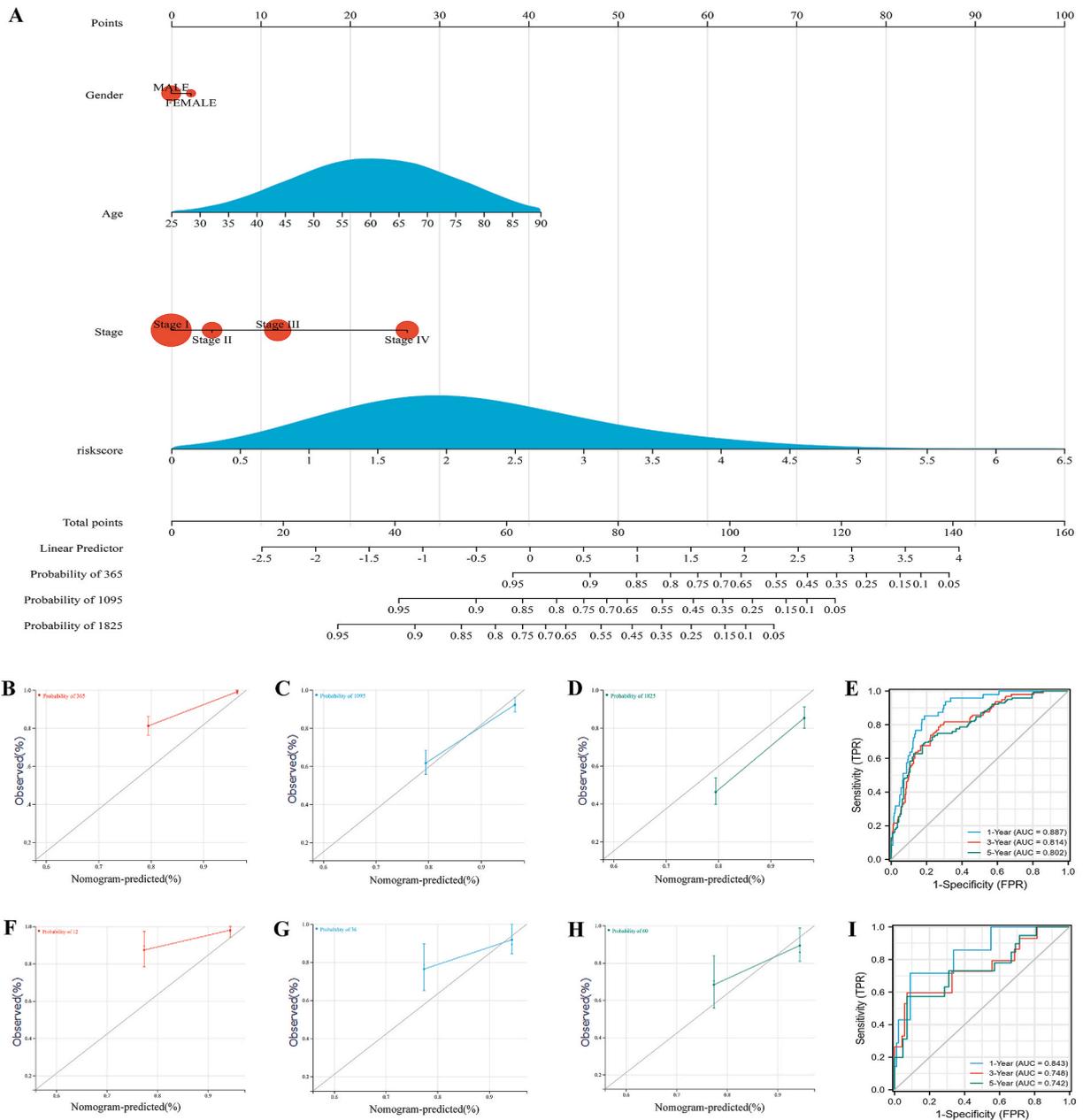


Fig. 8. Construction and calibration of nomogram. (A) Risk score and clinical features integrated into a nomogram. (B–D) Calibration of the nomogram at 1, 3 and 5 years in the training cohort. (E) ROC curves in the training cohort. (F–H) Calibration of the nomogram at 1, 3 and 5 years in the verification cohort. (I) ROC curves in the validation cohort.

4.3. Drug sensitivity and immune checkpoint

Furthermore, the study determined an immunophenoscore (IPS) to forecast the reaction of ccRCC patients to immune checkpoint inhibitor (ICI) treatment. The focus of immunotherapy has mainly been on anti-CTLA-4, anti-PD-L1, and anti-PD-1 treatment for solid tumors. Results showed that CTLA-4, PD-L1 and PD-1 mRNA expression levels were elevated in the high-risk cohort ($P < 0.001$) (Fig. 9A). To understand the correlation between ICI genes and immune infiltration, the study investigated the impact of immune infiltration on ICI gene clinical prognosis. Clinical outcomes were better in patients with high PD-1 levels as compared to low PD-1 levels ($P = 0.017$) (Fig. 9B). The results were similar when considering risk score and CTLA4 ($P = 0.023$) (Fig. 9B). Patients with high-risk score and elevated levels of PD-1 experienced more unfavorable outcomes compared to low-risk patients with high PD-1 levels ($P = 0.005$). This finding supports the notion that elevated expression of immune checkpoint genes was associated with a favorable prognosis. Furthermore, the prognostic significance of the risk score was explored in relation to chemotherapy and targeted therapy, employing the IC50 values of eight distinct drugs. High-risk patients exhibited significantly lower estimated IC50 values for Cisplatin, Lenalidomide, Elesclomol, and Paclitaxel when compared to low-risk patients, suggesting a weaker drug resistance in high-risk patients (Fig. 9C). Additionally, patients with low-risk were found to be greater sensitive to Sunitinib, Vinblastine, and Gefitinib

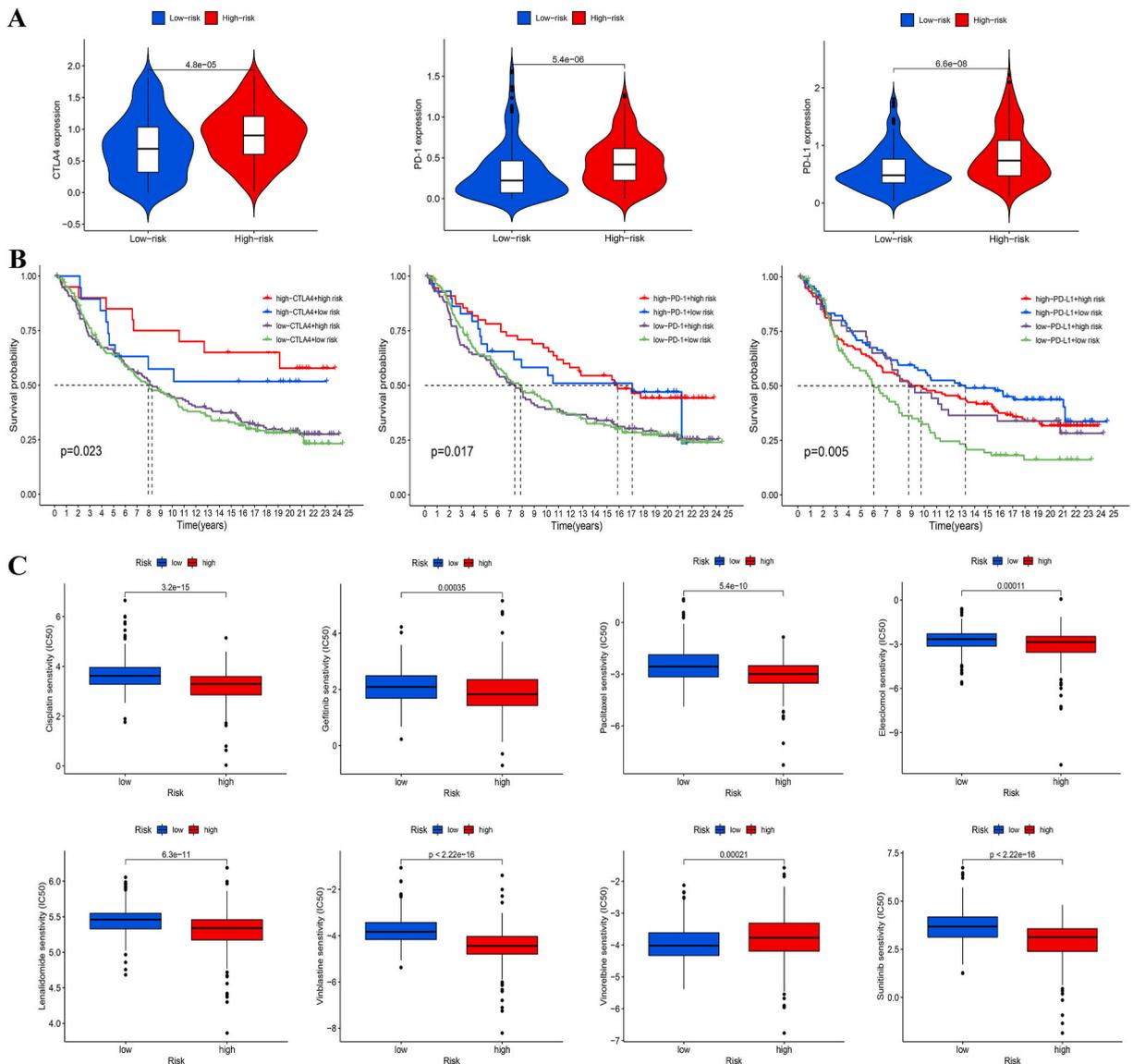


Fig. 9. Responses to immune checkpoint inhibitors. (A) Violin plots for expression levels of PD-1, CTLA-4, and PD-L1 between high- and low-risk patients. (B) Kaplan-Meier curves of PD-1, PD-L1 and CTLA-4. (C) Kaplan-Meier curves for OS outcomes among four groups, according to risk score and PD-1, CTLA-4, and PD-L1. (D) Boxplots illustrate the immunotherapeutic and chemotherapeutic responses of Cisplatin, Gefitinib, Elesclomol, Lenalidomide, Paclitaxel, Vinblastine, Vinorelbine, and Sunitinib in the low- and high-risk patients.

compared to patients with high-risk ($P < 0.05$) (Fig. 9C). The sensitivities of other medications between two groups can be seen in Supplementary Fig. 6. The Vinorelbine response was also observed to be more rapidly in high-risk groups ($P < 0.05$) (Fig. 9C).

5. Discussion

There are various pathological types of RCC with ccRCC being the most prevalent [42]. Prognostic biomarkers, serving as valuable tools in clinical decision-making, have demonstrated their efficacy in predicting the survival outcomes of RCC patients, rendering them a highly desirable option. Despite this, the commonly used TNM staging system for prognosticating RCC patients does not account for gene-level influences and can result in inaccurate predictions [43]. Factors such as angiogenesis and immune cell infiltration play a role in tumor development [44,45]. Advances in immunotherapy and anti-angiogenic therapy have shown promising results in treating advanced or aggressive cancers [46–48]. The objective of our study was to construct a prognosis model for AIRGs that can accurately forecast ccRCC survival and verified the risk model of AIRGs in the E-MTAB-1980 dataset successfully.

The tumor microenvironment (TME) is diverse and composed of a variety of extracellular matrix and cells, which influences the behavior and growth of tumor cells [49,50]. Tumor angiogenesis is regulated not only by the tumor cells themselves, but also by the immune cells and CAFs present within the TME [51]. Hence, the progression and development of ccRCC relies on the crucial involvement of angiogenesis and the immune environment.

Targeted therapy and immunotherapy stand as the foremost treatment modalities for advanced clear cell renal cell carcinoma [52]. Therefore, our model is constructed based on the current treatment landscape of advanced ccRCC and is validated using an independent validation dataset. We conducted a database search and found no reported prognostic models related to renal cancer angiogenesis. This research examined the impact of angiogenesis and immune response on predicting the prognosis of individuals with ccRCC. We applied the Venn diagram analysis of differential genes, angiogenesis factors, and immune genes to identify 155 AIRGs. Subsequently, lasso and Cox regression analyses were employed to determine relationships between prognosis abilities and these AIRGs. Four AIRGs (BIRC5, RNASE2, CCR10, and IL4) were found to be particularly significant. The risk score model built using these four AIRGs showed moderate to good performance in predicting one-year, three-year, and five-year overall survival (AUC = 0.79, 0.73, and 0.76, respectively). The pathological and clinical features are susceptible to be affected by the study population size. This prognostic model is constructed by the TCGA database. The p-value does not show significance because of the limited patient population in stage II. Nevertheless, it still reflects a trend of worse prognosis in high-risk cohort among stage II patients. It is speculated that with an increase in sample size, the p-value will become more significant. Additionally, most of the prognostic models related to immune genes lacked an independent validation cohort to assess diagnostic performance [53,54]. Li et al. constructed an immune-related prognostic model using kidney renal papillary cell carcinoma (KIRP) as an independent validation set. In this validation cohort, the AUC was 0.601, 0.571, and 0.518 at 1, 3, and 5 years, respectively [55], demonstrating the superior diagnostic performance of our risk score model.

Several studies have elucidated their pivotal involvement in inflammatory response, immune microenvironment and angiogenesis. There is evidence that BIRC5 promotes tumor cell proliferation and is associated with immune infiltration, as well as angiogenesis, chemotherapy resistance, and metastasis [56]. Studies have also demonstrated the important role of BIRC5 in the tumorigenesis and progression of various types of tumors [57–61]. CCR10 is expressed by certain types of T cells involved in tumor immune cell infiltration, angiogenesis, and lymphangiogenesis [62,63]. IL4 regulates macrophage polarization in the tumor immune process [64,65]. RNASE2 is an immune-related gene found in human eosinophilic leukocytes and has been shown to promote tumor development [66, 67]. Additionally, it has the potential to serve as an indicator in immune risk models for predicting ccRCC prognosis [53]. Due to the fact that both IL4 and RNASE2 are secreted proteins that function in the bloodstream, there are discrepancies between their qPCR experimental results and the TCGA prediction results. What's more, they have been used as prognostic models for renal clear cell carcinoma in multiple studies [53,68]. qRT-PCR revealed that mRNA level of CCR10 and BIRC5 were consistent with the TCGA results and this has been validated in several studies [69]. This indicates that our prognostic model genes are crucial to tumorigenesis and progression of ccRCC.

The examination of the tumor microenvironment revealed significant variations in the immune ratings among two risk categories, indicating distinctiveness in the number and activity of immune cells within two groups. The TME landscape revealed that two subgroups had distinct immune cells, providing insights into ways to enhance the immune response to improve therapeutic outcomes. High-risk patients had enriched Tregs, T cells, activated memory CD8⁺ T cells, CAFs, Tfh, and macrophage cell, while low-risk patients had more enriched neutrophils and endothelial cells. Among the immune cells infiltrating the TME, T cells, specifically CD8⁺ T cells, play a crucial role as significant immunological components against tumors [70]. The immunosuppressive functions of Tregs are widely regarded as a significant obstacle to effective antitumor immune response in ccRCC [71]. CAFs, the predominant stromal cells within the tumor microenvironment, wield a significant impact on tumor progression [72]. The immune landscape we have described could serve as a foundation for future studies on potential treatments for ccRCC. Risk stratification based on AIRGs can help guide targeted therapy and immunotherapy that targeting the TME.

Through the combination of immune infiltration analyses and tumor mutational burden, we discovered that high-risk individuals exhibit elevated TMB scores, which implies that their immune system exhibited a greater capacity to identify cancer cells [73]. Our tumor mutational burden analysis aligned with prior studies. Mutations in BAP1, SETD2, TTN, PBRM1 and VHL are the most frequently in ccRCC [73,74]. The high prevalence of BAP1 and KDM5C gene mutations observed in ccRCC was particularly pronounced in the high-risk cohort. In contrast to the low-risk cohort, the high-risk cohort exhibited greater immunological infiltrations, distinct drug sensitivity profiles, and increased tumor mutational burden, which is essential for providing therapeutic guidance of ccRCC.

The nomogram is widely used in the fields of oncology and medicine as a prognostic tool [75]. The nomogram developed in this study, which integrated risk score, gender, stage, and age, demonstrated a positive predictive effect. While evaluating the impact of tumor heterogeneity on personal prognosis may pose challenges, the application of a practical risk score amplifies the dependability of the nomogram and provides valuable guidance for clinical decision-making. Our research indicates new perspectives on epigenetic mechanisms which drive the progression of ccRCC and offer directions for personalized cancer treatments in the future.

There are still some constraints to the research. Firstly, the analysis relied on TCGA gene expression data, with validation restricted to the transcriptome level, which may not reflect the actual protein levels or the functional activities of these genes. Furthermore, our study relied on retrospective data analysis to construct and validate the prognostic model and therefore, a large prospective cohort study across multiple centers is needed to validate our findings. Finally, while our bioinformatics analysis identified four key AIR-DEGs in ccRCC as potential diagnostic and prognostic predictors, their specific functions and mechanisms remain unclear and warrant further investigation. We expected that accumulated datasets in the future will allow for additional verification and exploration of their mechanisms.

6. Conclusion

We have established a precise predictive diagnostic model for prognostic prediction in ccRCC patients based on four AIRGs. This signature exhibited a robust correlation with tumor microenvironment, survival time, drug sensitivity, and risk scores. The link between specific immune cell types and the risk score of this signature was supported by immunological research. Our research provides novel insights into the mechanisms underlying progression of ccRCC and presents direction personalized treatments in the future. Our constructed immune and angiogenesis prognostic, predictive model may be used comprehensively to predict survival outcomes and provide personalized treatment recommendations for ccRCC, encompassing targeted therapy and immunotherapy.

Data availability statement

All data are obtained in the article.

Ethics statement

This study was approved by the ethics committee of the affiliated Yantai Yuhuangding hospital of Qingdao university.

Funding

This research was supported by the Taishan Scholars Program of Shandong Province (No. tsqn201909199), Natural Science Foundation of Shandong Province (No. ZR2023MH241) and the National Natural Science Foundation of China (no. 82370690, 82000649).

CRediT authorship contribution statement

Guixin Ding: Writing – original draft, Visualization, Project administration, Methodology, Formal analysis, Data curation. **Tianqi Wang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources. **Gonglin Tang:** Investigation, Formal analysis, Data curation, Conceptualization. **Qingsong Zou:** Project administration, Methodology, Funding acquisition. **Gang Wu:** Validation, Resources, Project administration, Methodology, Investigation. **Jitao Wu:** Writing – review & editing, Resources, Project administration, Methodology, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

All authors have no acknowledgements to disclose.

Abbreviations

RCC	renal cell carcinoma
AIRGs	angiogenesis and immune-related genes
ccRCC	clear cell renal cell carcinoma
FDR	false discovery rate
TCGA	the Cancer Genome Atlas
MCP-COUNTER	microenvironment cell populations-counter

DEGs	differentially expressed genes
TMB	tumor mutational burden
ESTIMATE	Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression
TME	tumor microenvironment
KM	Kaplan-Meier
AIR-DEGs	angiogenesis-immune-related differentially expressed genes
GO	Gene Ontology analysis
IRGs	immune-related genes
CNV	copy number variants
KEGG	Kyoto Encyclopedia of Genes and Genomes analysis
MAF	Mutation comment file
ssGSEA	single-sample gene set enrichment
SNV	Single nucleotide variants
OS	overall survival
IC50	half-maximal inhibitory concentrations
AUC	area under the curve
ICIs	immune checkpoint inhibitors
IPS	immunophenoscore
ROC	receiver operating characteristics
GSEA	gene set enrichment analysis
LASSO	least absolute shrinkage and selection operator analysis
CAFs	cancer-associated fibroblasts
Tregs	regulatory cells
Tfhs	T follicular helper cells
IL4	Interleukin 4
BIRC5	Baculoviral IAP repeat containing 5
RNASE2	Ribonuclease A family member 2
MUC16	Mucin 16
PBRM1	Polybromo 1
CCR10	C-C motif chemokine receptor 10
VHL	Von Hippel-Lindau Tumor Suppressor
BAP1	BRCA1 associated protein 1
TTN	Titin
ATM	ATM serine/threonine kinase
KDM5C	Lysine demethylase 5C
SETD2	SET Domain Containing 2

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e23503>.

References

- [1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, *CA A Cancer J. Clin.* 70 (1) (2020) 7–30.
- [2] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA A Cancer J. Clin.* 71 (3) (2021) 209–249.
- [3] C. Tito, E. De Falco, P. Rosa, A. Iaiza, F. Fazi, V. Petrozza, et al., Circulating microRNAs from the molecular mechanisms to clinical biomarkers: a focus on the clear cell renal cell carcinoma, *Genes* 12 (8) (2021) 1154.
- [4] A.A. Lalani, B.A. McGregor, L. Albiges, T.K. Choueiri, R. Motzer, T. Powles, et al., Systemic treatment of metastatic clear cell renal cell carcinoma in 2018: current paradigms, use of immunotherapy, and future directions, *Eur. Urol.* 75 (1) (2019) 100–110.
- [5] J.J. Hsieh, M.P. Purdue, S. Signoretti, C. Swanton, L. Albiges, M. Schmidinger, et al., Renal cell carcinoma, *Nat. Rev. Dis. Prim.* 3 (2017), 17009.
- [6] R.L. Siegel, K.D. Miller, H.E. Fuchs, A. Jemal, Cancer statistics, 2021, *CA A Cancer J. Clin.* 71 (1) (2021) 7–33.
- [7] Q.K. Li, C.P. Pavlovich, H. Zhang, C.R. Kinsinger, D.W. Chan, Challenges and opportunities in the proteomic characterization of clear cell renal cell carcinoma (ccRCC): a critical step towards the personalized care of renal cancers, *Semin. Cancer Biol.* 55 (2019) 8–15.
- [8] E. Jonasch, J. Gao, W.K. Rathmell, Renal cell carcinoma, *BMJ (Clinical research ed)* 349 (2014) g4797.
- [9] H. Moch, A.L. Cubilla, P.A. Humphrey, V.E. Reuter, T.M. Ulbright, The 2016 WHO classification of tumours of the urinary system and male genital organs-Part A: renal, penile, and testicular tumours, *Eur. Urol.* 70 (1) (2016) 93–105.
- [10] J.R. Srigley, B. Delahunt, J.N. Eble, L. Egevad, J.I. Epstein, D. Grignon, et al., The international society of urological pathology (ISUP) vancouver classification of renal neoplasia, *Am. J. Surg. Pathol.* 37 (10) (2013) 1469–1489.
- [11] J. Folkman, What is the evidence that tumors are angiogenesis dependent? *J. Natl. Cancer Inst.* 82 (1) (1990) 4–6.
- [12] W. Zheng, S. Zhang, H. Guo, X. Chen, Z. Huang, S. Jiang, et al., Multi-omics analysis of tumor angiogenesis characteristics and potential epigenetic regulation mechanisms in renal clear cell carcinoma, *Cell Commun. Signal. : CCS* 19 (1) (2021) 39.
- [13] T.K. Choueiri, R.J. Motzer, Systemic therapy for metastatic renal-cell carcinoma, *N. Engl. J. Med.* 376 (4) (2017) 354–366.

- [14] V. Baeriswyl, G. Christofori, The angiogenic switch in carcinogenesis, *Semin. Cancer Biol.* 19 (5) (2009) 329–337.
- [15] G. Bergers, L.E. Benjamin, Tumorigenesis and the angiogenic switch, *Nat. Rev. Cancer* 3 (6) (2003) 401–410.
- [16] P.C. Barata, B.I. Rini, Treatment of renal cell carcinoma: current status and future directions, *CA A Cancer J. Clin.* 67 (6) (2017) 507–524.
- [17] J.C. Chappell, L.B. Payne, W.K. Rathmell, Hypoxia, angiogenesis, and metabolism in the hereditary kidney cancers, *J. Clin. Investig.* 129 (2) (2019) 442–451.
- [18] F. Peinemann, S. Unverzagt, A.V. Hadjinicolaou, I. Moldenhauer, Immunotherapy for metastatic renal cell carcinoma: a systematic review, *J. Evidence-based Med.* 12 (4) (2019) 253–262.
- [19] S.L. Topalian, F.S. Hodi, J.R. Brahmer, S.N. Gettinger, D.C. Smith, D.F. McDermott, et al., Five-year survival and correlates among patients with advanced melanoma, renal cell carcinoma, or non-small cell lung cancer treated with nivolumab, *JAMA Oncol.* 5 (10) (2019) 1411–1420.
- [20] R.J. Motzer, B. Escudier, D.F. McDermott, S. George, H.J. Hammers, S. Srinivas, et al., Nivolumab versus everolimus in advanced renal-cell carcinoma, *N. Engl. J. Med.* 373 (19) (2015) 1803–1813.
- [21] P. Goldstraw, D. Ball, J.R. Jett, T. Le Chevalier, E. Lim, A.G. Nicholson, et al., Non-small-cell lung cancer, *Lancet (London, England)* 378 (9804) (2011) 1727–1740.
- [22] A. Pircher, D. Wolf, A. Heidenreich, W. Hilbe, R. Pichler, I. Heidegger, Synergies of targeting tumor angiogenesis and immune checkpoints in non-small cell lung cancer and renal cell cancer: from basic concepts to clinical reality, *Int. J. Mol. Sci.* 18 (11) (2017) 2291.
- [23] S. Mariathasan, S.J. Turley, D. Nickles, A. Castiglioni, K. Yuen, Y. Wang, et al., TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells, *Nature* 554 (7693) (2018) 544–548.
- [24] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (12) (2014) 550.
- [25] K. Yoshihara, M. Shahmoradgoli, E. Martínez, R. Vegesna, H. Kim, W. Torres-García, et al., Inferring tumour purity and stromal and immune cell admixture from expression data, *Nat. Commun.* 4 (2013) 2612.
- [26] M. Binnewies, E.W. Roberts, K. Kersten, V. Chan, D.F. Fearon, M. Merad, et al., Understanding the tumor immune microenvironment (TIME) for effective therapy, *Nat. Med.* 24 (5) (2018) 541–550.
- [27] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, *Nat. Methods* 12 (5) (2015) 453–457.
- [28] Y. Cao, Adipocyte and lipid metabolism in cancer drug resistance, *J. Clin. Investig.* 129 (8) (2019) 3006–3017.
- [29] J. Racle, K. de Jonge, P. Baumgaertner, D.E. Speiser, D. Gfeller, Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data, *Elife* 6 (2017), e26476.
- [30] E. Becht, N.A. Giraldo, L. Lacroix, B. Buttard, N. Elarouci, F. Petitprez, et al., Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression, *Genome Biol.* 17 (1) (2016) 218.
- [31] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, *OMICS A J. Integr. Biol.* 16 (5) (2012) 284–287.
- [32] I.S. Ramsay, S. Ma, M. Fisher, R.L. Loewy, J.D. Ragland, T. Niendam, et al., Model selection and prediction of outcomes in recent onset schizophrenia patients who undergo cognitive training, *Schizophrenia research Cognition* 11 (2018) 1–5.
- [33] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, *Genome Res.* 28 (11) (2018) 1747–1756.
- [34] P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efreanova, D. Rieder, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, *Cell Rep.* 18 (1) (2017) 248–262.
- [35] P. Geeleher, N.J. Cox, R.S. Huang, Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines, *Genome Biol.* 15 (3) (2014) R47.
- [36] P. Geeleher, N. Cox, R.S. Huang, pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels, *PLoS One* 9 (9) (2014), e107468.
- [37] M.B. Fuentes, A.K. Kacha, J. Kline, S.R. Woo, D.M. Kranz, K.M. Murphy, et al., Host type I IFN signals are required for antitumor CD8+ T cell responses through CD8 α + dendritic cells, *J. Exp. Med.* 208 (10) (2011) 2005–2016.
- [38] K. Hildner, B.T. Edelson, W.E. Purtha, M. Diamond, H. Matsushita, M. Kohyama, et al., Batf3 deficiency reveals a critical role for CD8 α + dendritic cells in cytotoxic T cell immunity, *Science (New York, NY)* 322 (5904) (2008) 1097–1100.
- [39] J.P. Böttcher, E. Bonavita, P. Chakravarty, H. Blees, M. Cabeza-Cabrero, S. Sammiceli, et al., NK cells stimulate recruitment of cDC1 into the tumor microenvironment promoting cancer immune control, *Cell* 172 (5) (2018), 1022–37.e14.
- [40] S. Spranger, D. Dai, B. Horton, T.F. Gajewski, Tumor-residing Batf3 dendritic cells are required for effector T cell trafficking and adoptive T cell therapy, *Cancer Cell* 31 (5) (2017), 711–23.e4.
- [41] H. Harlin, Y. Meng, A.C. Peterson, Y. Zha, M. Tretiakova, C. Slingluff, et al., Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment, *Cancer Res.* 69 (7) (2009) 3077–3085.
- [42] S. Li, Y. Cheng, G. Cheng, T. Xu, Y. Ye, Q. Miu, et al., High SAA1 expression predicts advanced tumors in renal cancer, *Front. Oncol.* 11 (2021), 649761.
- [43] R.J. Motzer, E. Jonasch, M.D. Michaelson, L. Nandagopal, J.L. Gore, S. George, et al., NCCN guidelines insights: kidney cancer, version 2.2020, *J. Natl. Compr. Cancer Netw. : J. Natl. Compr. Cancer Netw.* 17 (11) (2019) 1278–1285.
- [44] X. Jiang, J. Wang, X. Deng, F. Xiong, S. Zhang, Z. Gong, et al., The role of microenvironment in tumor angiogenesis, *J. Exp. Clin. Cancer Res. : CR* 39 (1) (2020) 204.
- [45] Y.E. Du, G. Tu, G. Yang, G. Li, D. Yang, L. Lang, et al., MiR-205/YAP1 in activated fibroblasts of breast tumor promotes VEGF-independent angiogenesis through STAT3 signaling, *Theranostics* 7 (16) (2017) 3972–3988.
- [46] S. Menon, S. Shin, G. Dy, Advances in cancer immunotherapy in solid tumors, *Cancers* 8 (12) (2016) 106.
- [47] C. Voena, R. Chiarle, Advances in cancer immunology and cancer immunotherapy, *Discov. Med.* 21 (114) (2016) 125–133.
- [48] C.H. Lee, R.J. Motzer, Kidney cancer in 2016: the evolution of anti-angiogenic therapy for kidney cancer, *Nat. Rev. Nephrol.* 13 (2) (2017) 69–70.
- [49] S. Maman, I.P. Witz, A history of exploring cancer in context, *Nat. Rev. Cancer* 18 (6) (2018) 359–376.
- [50] G. Biffi, D.A. Tuveson, Diversity and biology of cancer-associated fibroblasts, *Physiol. Rev.* 101 (1) (2021) 147–176.
- [51] R.S. Watnick, The role of the tumor microenvironment in regulating angiogenesis, *Cold Spring Harbor perspectives in medicine* 2 (12) (2012) a006676.
- [52] R.J. Motzer, E. Jonasch, S. Boyle, M.I. Carlo, B. Manley, N. Agarwal, et al., NCCN guidelines insights: kidney cancer, version 1.2021, *J. Natl. Compr. Cancer Netw. : J. Natl. Compr. Cancer Netw.* 18 (9) (2020) 1160–1170.
- [53] B. Wan, B. Liu, Y. Huang, G. Yu, C. Lv, Prognostic value of immune-related genes in clear cell renal cell carcinoma, *Aging* 11 (23) (2019) 11474–11489.
- [54] J. He, Y. Zhong, Y. Sun, C. Xie, T. Yu, Construction of an immune-related prognostic model by exploring the tumor microenvironment of clear cell renal cell carcinoma, *Anal. Biochem.* 643 (2022), 114567.
- [55] Y. Liu, D. Wu, H. Chen, L. Yan, Q. Xiang, Q. Li, et al., Construction and verification of a novel prognostic risk model for kidney renal clear cell carcinoma based on immunity-related genes, *Front. Genet.* 14 (2023), 1107294.
- [56] S.P. Wheatley, D.C. Altieri, Survivin at a glance, *J. Cell Sci.* 132 (7) (2019), jcs223826.
- [57] L. Cao, C. Li, S. Shen, Y. Yan, W. Ji, J. Wang, et al., OCT4 increases BIRC5 and CCND1 expression and promotes cancer progression in hepatocellular carcinoma, *BMC Cancer* 13 (2013) 82.
- [58] L. Li, X. Chen, L. Hao, Q. Chen, H. Liu, Q. Zhou, Exploration of immune-related genes in high and low tumor mutation burden groups of chromophobe renal cell carcinoma, *Biosci. Rep.* 40 (7) (2020), BSR20201491.
- [59] P. Giménez-Bonafé, A. Tortosa, R. Pérez-Tomás, Overcoming drug resistance by enhancing apoptosis of tumor cells, *Curr. Cancer Drug Targets* 9 (3) (2009) 320–340.
- [60] D.S. O'Connor, J.S. Schechner, C. Adida, M. Mesri, A.L. Rothermel, F. Li, et al., Control of apoptosis during angiogenesis by survivin expression in endothelial cells, *Am. J. Pathol.* 156 (2) (2000) 393–398.

- [61] A.C. Mita, M.M. Mita, S.T. Nawrocki, F.J. Giles, Survivin: key regulator of mitosis and apoptosis and novel target for cancer therapeutics, *Clin. Cancer Res. : an official journal of the American Association for Cancer Research* 14 (16) (2008) 5000–5005.
- [62] N. Xiong, Y. Fu, S. Hu, M. Xia, J. Yang, CCR10 and its ligands in regulation of epithelial immunity and diseases, *Protein & cell* 3 (8) (2012) 571–580.
- [63] J. Korbecki, S. Grochans, I. Gutowska, K. Barczak, I. Baranowska-Bosiacka, CC chemokines in a tumor: a review of pro-cancer and anti-cancer properties of receptors CCR5, CCR6, CCR7, CCR8, CCR9, and CCR10 ligands, *Int. J. Mol. Sci.* 21 (20) (2020) 7619.
- [64] M. Genin, F. Clement, A. Fattaccioli, M. Raes, C. Michiels, M1 and M2 macrophages derived from THP-1 cells differentially modulate the response of cancer cells to etoposide, *BMC Cancer* 15 (2015) 577.
- [65] X. Song, B. Traub, J. Shi, M. Kornmann, Possible roles of interleukin-4 and -13 and their receptors in gastric and colon cancer, *Int. J. Mol. Sci.* 22 (2) (2021) 727.
- [66] H.L. Tiffany, J.S. Handen, H.F. Rosenberg, Enhanced expression of the eosinophil-derived neurotoxin ribonuclease (RNS2) gene requires interaction between the promoter and intron, *J. Biol. Chem.* 271 (21) (1996) 12387–12393.
- [67] T. Wu, Y. Chen, L. Yang, X. Wang, K. Chen, D. Xu, Ribonuclease A family member 2 promotes the malignant progression of glioma through the PI3K/akt signaling pathway, *Front. Oncol.* 12 (2022), 921083.
- [68] P. Terrematte, D.S. Andrade, J. Justino, B. Stransky, D.S.A. de Araújo, A.D. Dória Neto, A novel machine learning 13-gene signature: improving risk analysis and survival prediction for clear cell renal cell carcinoma patients, *Cancers* 14 (9) (2022) 2111.
- [69] J. Wang, M. Chen, C. Dang, H. Zhang, X. Wang, J. Yin, et al., The early diagnostic and prognostic value of BIRC5 in clear-cell renal cell carcinoma based on the cancer genome Atlas data, *Urol. Int.* 106 (4) (2022) 344–351.
- [70] M. Philip, A. Schietinger, CD8(+) T cell differentiation and dysfunction in cancer, *Nat. Rev. Immunol.* 22 (4) (2022) 209–223.
- [71] E.N. Scott, A.M. Gocher, C.J. Workman, D.A.A. Vignali, Regulatory T cells: barriers of immune infiltration into the tumor microenvironment, *Front. Immunol.* 12 (2021), 702726.
- [72] P.M. Galbo Jr., X. Zang, D. Zheng, Molecular features of cancer-associated fibroblast subtypes and their implication on cancer pathogenesis, prognosis, and immunotherapy resistance, *Clin. Cancer Res. : an official journal of the American Association for Cancer Research* 27 (9) (2021) 2636–2647.
- [73] C. Zhang, Z. Li, F. Qi, X. Hu, J. Luo, Exploration of the relationships between tumor mutation burden with immune infiltrates in clear cell renal cell carcinoma, *Ann. Transl. Med.* 7 (22) (2019) 648.
- [74] J. Brugarolas, Molecular genetics of clear-cell renal cell carcinoma, *J. Clin. Oncol. : official journal of the American Society of Clinical Oncology* 32 (18) (2014) 1968–1976.
- [75] V.P. Balachandran, M. Gonen, J.J. Smith, R.P. DeMatteo, Nomograms in oncology: more than meets the eye, *Lancet Oncol.* 16 (4) (2015) e173–e180.