



Contents lists available at ScienceDirect

Journal of the National Cancer Center

journal homepage: [www.elsevier.com/locate/jncc](http://www.elsevier.com/locate/jncc)

Full Length Article

# Tumor microenvironment-based signatures distinguish intratumoral heterogeneity, prognosis, and immunogenomic features of clear cell renal cell carcinoma



Aihetaimujiang Anwaier<sup>1,2,†</sup>, Wenhao Xu<sup>1,2,†</sup>, Wangrui Liu<sup>3,†</sup>, Shiyin Wei<sup>4,†</sup>, Xi Tian<sup>1,2</sup>,  
Yuan Yuan Qu<sup>1,2</sup>, Jianfeng Yang<sup>5,\*</sup>, Hailiang Zhang<sup>1,2,\*</sup>, Dingwei Ye<sup>1,2,\*</sup>

<sup>1</sup> Department of Urology, Fudan University Shanghai Cancer Center; Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

<sup>2</sup> Shanghai Genitourinary Cancer Institute, Shanghai, China

<sup>3</sup> Department of Interventional Oncology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

<sup>4</sup> Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China

<sup>5</sup> Department of Surgery, ShangNan Branch of Longhua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

## ARTICLE INFO

### Keywords:

Clear cell renal cell carcinoma  
Tumor microenvironment  
Risk group  
Immunogenomic feature

## ABSTRACT

**Background:** The tumor microenvironment (TME) performs a crucial function in the tumorigenesis and response to immunotherapies of clear cell renal cell carcinoma (ccRCC). However, a lack of recognized pre-clinical TME-based risk models poses a great challenge to investigating the risk factors correlated with prognosis and treatment responses for patients with ccRCC.

**Methods:** Stromal and immune contexture were assessed to calculate the TMErisk score of a large sample of patients with ccRCC from public and real-world cohorts using machine-learning algorithms. Next, analyses for prognostic efficacy, correlations with clinicopathological features, functional enrichment, immune cell distributions, DNA variations, immune response, and heterogeneity were performed and validated.

**Results:** Clinical hub genes, including *INAFM2*, *SRPX*, *DPYSL3*, *VSIG4*, *APLN*, *FHL5*, *A2M*, *SLFN11*, *ADAMTS4*, *IFITM1*, *NOD2*, *CCR4*, *HLA-DQB2*, and *PLAUR*, were identified and incorporated to develop the TMErisk signature. Patients in the TME<sup>high</sup> risk group (category) exhibited a considerably grim prognosis, and the TMErisk model was shown to independently function as a risk indicator for the overall survival (OS) of ccRCC patients. Expression levels of immune checkpoint genes were substantially increased in TME<sup>high</sup> risk group, while those of the human leukocyte antigen (HLA) family genes were prominently decreased. In addition, tumors in the TME<sup>high</sup> group showed significantly high infiltration levels of tumor-infiltrated lymphocytes, including M2 macrophages, CD8<sup>+</sup> T cells, B cells, and CD4<sup>+</sup> T cells. In heterogeneity analysis, more frequent somatic mutations, including pro-tumorigenic BAP1 and PBRM1, were observed in the TME<sup>high</sup> group. Importantly, 19.3% of patients receiving immunotherapies in the TME<sup>high</sup> group achieved complete or partial response compared with those with immune tolerance in the TME<sup>low</sup> group, suggesting that TMErisk prominently differentiates prognosis and responses to immunotherapy for patients with ccRCC.

**Conclusions:** We first established the TMErisk score of ccRCC using machine-learning algorithms based on a large-scale population. The TMErisk score can be utilized as an innovative independent prognosis predictive marker with high sensitivity and accuracy. Our discovery also predicted the efficacy of immunotherapy in ccRCC patients, indicating the intimate link between tumor immune microenvironment and intratumoral heterogeneity.

## 1. Introduction

Clear cell renal cell carcinoma (ccRCC) is the predominant histological subtype of renal cell carcinoma (RCC), representing 70%–85% of all kidney cancers and approximately 3.5%–4.2% of all newly diagnosed

cancers.<sup>1</sup> It predominantly occurs in adults aged more than 60 years, and males are three times as likely to be affected as females.<sup>1</sup> Although early-stage localized ccRCC can be cured with surgical treatment, over a third of cases present with organ, bone, or lymphatic metastases at initial diagnosis or develop metastases after surgery.<sup>2</sup> Advanced ccRCC is

\* Corresponding authors.

E-mail addresses: [yangjianfeng@shutcm.edu.cn](mailto:yangjianfeng@shutcm.edu.cn) (J. Yang), [zhanghl918@alu.fudan.edu.cn](mailto:zhanghl918@alu.fudan.edu.cn) (H. Zhang), [dingwei\\_ye@fudan.edu.cn](mailto:dingwei_ye@fudan.edu.cn) (D. Ye).

† These authors contributed equally to this work.

<https://doi.org/10.1016/j.jncc.2023.08.003>

Received 22 June 2023; Received in revised form 31 July 2023; Accepted 8 August 2023

2667-0054/© 2023 Chinese National Cancer Center. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

biologically heterogeneous and is a major cause of mortality, with a 10% 5-year rate of survival.<sup>3,4</sup> Hence, it is essential to deeply investigate the tumor microenvironment (TME) to understand tumor heterogeneity better and improve the effectiveness of treatment management strategies.

The TME is defined as the cellular milieu surrounding tumor cells consisting of stromal cells, extracellular matrix, immune cells, and different chemokines and cytokines, and the interaction of these components is crucial to carcinogenesis and the response to immunotherapies in ccRCC patients.<sup>5</sup> Genomic alterations within incipient ccRCC not only contribute to neoplastic progression but also affect the cellular composition in TME, including the high infiltration of CD8<sup>+</sup> T cells, which is linked to worse outcomes and response to immunotherapy.<sup>6–9</sup> Biallelic inactivation of the *VHL* tumor suppressor gene on a background of a loss of chromosome 3p is an important genetic event in most ccRCCs.<sup>10,11</sup> Some epigenetic regulatory genes located at chromosome 3p, such as *PBRM1*, *SETD2*, and *BAP1*, are also identified as frequently mutated genes and thought to interact with *VHL* inactivation to promote the tumorigenesis of ccRCC.<sup>12–14</sup> Importantly, these most commonly mutated genes were associated with variations in the TME and affected the effectiveness of immunotherapeutic regimens.<sup>15,16</sup> Therefore, to clarify the tumor heterogeneity, appropriate treatments, and prognosis of ccRCC, it is of interest to study the molecular characteristics and subclassify ccRCC into prognostic risk groups.

Several studies have well characterized the gene expression signatures and suggested some potential therapeutic targets in ccRCC.<sup>17,18</sup> The Cancer Genomic Atlas (TCGA) divided ccRCC with different somatic mutations into four molecular subtypes and differential survival rates based on extensive transcriptional profiles.<sup>10</sup> The m1 subtype showed a frequent mutation in *PBRM1* with the best outcome. The m3 subtype with the worst outcome was characterized by *PTEN* mutation and *CDK2NA* delation. The m4 subtype showed a mutation in DNA repair-related genes, mTOR signaling pathway-related genes, and *BAP1*, and also correlated with poor prognosis.<sup>10</sup> Our previous study on immune infiltration and prognosis of ccRCC also identified three immunophenotyping clusters, which demonstrated the prognostic significance of immunological contexture in ccRCC TME. Our data revealed that *VHL* and *PBRM1* were the most commonly mutated genes in ccRCC, and mutation rates of these genes were greater in Cluster A than in Cluster B and C. Survival rates are much lower in immune-hot clusters B and C, which is indicative of pro-tumor immune infiltration, compared to immune-cold cluster A.<sup>19</sup> We also noticed the differences in myeloid infiltration score (stromal score) and immune infiltration score (immune score) among subgroups. Therefore, to further clarify how genomic changes alter the TME in ccRCC, we established a TME-related risk model based on transcriptomic and proteomic data to predict ccRCC patients' prognoses in the present study.

The TME is a complex ecosystem that affects tumor occurrence, development, and treatment resistance. However, a lack of recognized pre-clinical TME-related risk models of ccRCC poses great challenges to investigating the risk factors correlated with patients' prognosis and treatment responses. To address these challenges, many studies have investigated the interactions between immune cells or identified some prognostic hub genes that alter the TME in ccRCC, but there have been few studies focused on investigating TME-related risk models based on immune and stromal features of ccRCC. We, therefore, sought to identify TME-related differential expression genes to establish a novel risk score that can potentially enhance the prognostic reliability and lay the conceptual groundwork for the development of targeted treatment options.

## 2. Materials and methods

### 2.1. Downloading and preprocessing of data

The fragments per kilobase of exon model per million mapped fragments (FPKM) gene expression level data, clinicopathological and sur-

vival data, and gene mutation data of ccRCC patients were derived from the The Cancer Genome Atlas (TCGA) dataset, with gene IDs transferred from Ensembl ID format into the gene symbol matrix format (<https://portal.gdc.cancer.gov>). Genes with missing values in more than 30% samples were removed to minimize the effect of missing values in gene expression data. RNA-seq data and related clinical data of 520 ccRCC patients were constructed after filtering the samples with survival times less than 30 days. Immune subtypes of cancer were obtained from datasets from the previous study.<sup>20</sup> Treatment response and corresponding RNA-seq data were obtained from the CheckMate cohorts published previously.<sup>21</sup> Proteomic data of paired primary ccRCC tissues and adjacent normal tissues of 232 cases were enrolled from our previous study cohort (FUSCC cohort). To evaluate the immune infiltration and landscape of TME, we used the “Immune” model of the TIMER2.0 dataset (<http://timer.cistrome.org/>), which utilizes six state-of-the-art algorithms to produce a more accurate estimation of the infiltration levels of immune cells.

### 2.2. Construction of stromal and immune risk signature

To quantify the enrichment of immune and stromal cells in TME, tumor purity, stromal, and immune scores of 520 ccRCC cases were computed as per the gene expression levels using the ESTIMATE algorithm<sup>22</sup> by R package “ESTIMATE”. Patients were categorized as having low or high stromal/immune scores depending on the optimal cutoff point, which was computed with the “survminer” R tool (<https://cran.r-project.org/web/packages/survminer/index.html>). The Kaplan-Meier (KM) analysis was adopted to assess the prognosis of each group with overall survival (OS) as the prognostic endpoint, and the significance of the difference (variation) was evaluated by means of the log-rank test (Supplementary Fig. 2A and B).

### 2.3. Identification of differentially expressed genes

Subsequently, differentially expressed genes (DEGs) between the stromal and immune groups were identified with the “limma” R program,<sup>23</sup> with the screening criteria of false discovery rate (FDR) q-value <0.05 and  $|\log_2\text{FC}| > \log_2(1.5)$ . Compared to the low immune/stromal group, the overexpressed genes in the high stromal/immune group were defined as “upregulated DEGs”, and those with low expression in the high immune/stromal group were defined as “downregulated DEGs”.

### 2.4. Establishment of WGCNA

The R package “WGCNA” was utilized to investigate the correlation patterns among genes, and identify stromal and immune scores correlated modules.<sup>24</sup> To ensure a signed scale-free co-expression gene network, we used soft-threshold values of power  $\beta = 2$  and a scale-free  $R^2 = 0.83$ . Based on the optimal soft-threshold 7, the minimum module size of 30, and 31 coexpressed gene modules were discovered in total. Modules with a stromal or immune score correlation coefficient > 0.5 and  $P < 0.05$  were defined as modules associated with the stromal or immune score.

### 2.5. Construction of TMErisk model

To find a set of prognostic genes, we conducted univariate Cox regression analysis on the intersecting genes between stromal and immune score-related DEGs and module genes. Subsequently, lasso regression analysis was performed to reduce the number of prognostic genes, followed by the multivariate Cox stepwise regression analysis to select the TMErisk genes by calculating the differences in the Akaike information criterion (AIC) with the minimum AIC value of 1 779.72. Finally, each case's TMErisk score was determined by applying the following risk equation:  $\text{Riskscore} = \sum_{i=1}^n (\text{Exp } i * \beta_i)$  where “Exp *i*” signifies the expression of each prognostic gene and “ $\beta_i$ ” signifies each prognostic

gene's coefficient (Supplementary Table 1). Patients were classified into low- and high-risk categories depending on their TMErisk scores, with the optimum threshold of 0.7263283 computed using the “survminer” R program.

## 2.6. Evaluation and verification of the TMErisk model

The survival analysis between the low- and high-risk categories of TME was carried out using the KM technique. The TMErisk model's ability to anticipate specificity and sensitivity was assessed utilizing the receiver operating characteristic (ROC) curve. The tumor stage and the C-index of the TMErisk model were applied in conjunction with the “ggplot” R program to assess the stability of the model.<sup>25</sup> Moreover, the TMErisk model's significance as an independent prognosis predictive indicator was examined via univariate analysis, and the findings were shown in a forest plot with the help of the “forestplot” tool in R (<https://cran.r-project.org/web/packages/forestplot/index.html>).

Using an anti-CCR4 antibody (1:500, ab216560, Abcam) and anti-NOD2 antibody (1:500, ab188646, Abcam), immunohistochemical (IHC) staining of CCR4 and NOD2 was implemented to validate the expression levels of CCR4 and NOD2 in 20 paired samples that were collected from the FUSCC tissue bank. Two professional pathologists rated the IHC staining for CCR4 and NOD2 separately, and any disagreements were discussed and resolved to establish a consensus. The proportion of tumor cells that were covered by the stain was represented by a score range of 0–4: 0%, 1–25%, 26–50%, 51–70%, and 76–100% respectively. Levels of staining intensity spanned from 0 (no staining) to 3 (very strong staining), with 3 signifying strong staining, 2 signifying moderate staining, 1 signifying weak staining, and 0 signifying no staining. The cumulative IHC score (between 0 and 12) was determined by multiplying the staining degree value by the staining intensity.

## 2.7. Gene set enrichment analysis and functional enrichment analysis

TME low- and high-risk subgroups were subjected to Gene Set Enrichment Analysis (GSEA) to identify significant enrichment pathways, which were annotated by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) datasets. We defined highly enriched results in the GSEA as those with an FDR q-value < 0.05 and a normalized enrichment score NES > 1.

## 2.8. Correlations between TMErisk model and immune characteristics

Pearson correlation was applied to probe the link between the TMErisk model and the tumor purity, the immune and stromal scores, and the consensus measurement of purity estimations (CPE) score. A Wilcoxon rank-sum test was used to compare the expression patterns of 39 immune checkpoint genes (ICGs) and 19 HLA family genes between the TME low- and high-risk categories. Heat maps of ICGs and the HLA family associated with TMErisk and immune cell infiltration in distinct TMErisk groups were visualized with the aid of the R “ComplexHeatmap” tool.<sup>26</sup>

## 2.9. Mutation analysis

Multiple types of mutations were identified and categorized. Translation\_Start\_Site, Splice\_Site, Nonstop\_Mutation, Nonsense\_Mutation, Missense\_Mutation, In\_Frame\_Ins, In\_Frame\_Del, Frame\_Shift\_Ins, and Frame\_Shift\_Del, were further classified as non-synonymous mutations. Furthermore, synonymous mutations also included silent mutations as well as other mutation types such as 3'Flank, 3'UTR, 5'Flank, 5'UTR, and Intron. In 353 samples from the TCGA mutation cohort, genes with more than 20 mutations were defined as high-frequency mutant genes.

## 2.10. Treatment response

By employing the gene expression profiles for 520 TCGA cohort patients with ccRCC, the 50% inhibitory concentrations (IC<sub>50</sub>) value of 138 drugs was extrapolated using the “pRRophetic” tool in R and the value was normality transformed.<sup>27</sup> Tumor immunodeficiency and exclusion (TIDE) scores were utilized to reflect the patient's potential responsiveness to immunotherapeutic regimens (<http://tide.dfci.harvard.edu/>). Gene expression data and treatment response data from CheckMate cohorts<sup>21</sup> of 172 patients who had received Nivolumab were utilized to assess the link between the TMErisk model and immune response.

## 2.11. Culture of human ccRCC cells and derivation of lentiviral-transfected cell lines

The 786-O human ccRCC cell line was acquired from the Cell Bank of Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences (Shanghai, China). The 786-O cells were grown in a humid chamber with 5% CO<sub>2</sub> at 37 °C in RPMI 1640 solution (Gibco, CA, USA) with 10% fetal bovine serum (FBS; Gibco, Grand Island, NY) and 1% penicillin-streptomycin mixture (Gibco, Grand Island, NY). A 70% confluent cell monolayer was reached after overnight incubation of cells in 6-well plates before transfection. Using the instructed concentration of Lipofectamine 3000 Transfection Reagent (Invitrogen, USA), 786-O cells were transfected with NOD2-siRNA (General Biol) and an overexpression plasmid (GeneChem).

## 2.12. Western blotting assay

In 6-well plates, 786-O cells ( $1 \times 10^5$ ) were seeded to start the experiment. Following an incubation duration of 24 h, the cells were subjected to treatment with DMSO (serving as the vehicle) or FHP01 or XAV939 (serving as Merck) for the durations specified. CcRCC cells were transfected and then rinsed using PBS, and total protein extracts were prepared by introducing 80  $\mu$ L of RIPA lysis buffer into the mixture. As previously described,<sup>28</sup> for the Western blotting, 10  $\mu$ g of proteins generated from whole lysates were deposited onto 8% polyacrylamide gels with 1% Laemmli solution and separated by SDS-PAGE. The proteins were then transferred onto an Immobilon-P PVDF membrane (Millipore, IPVH00010) and examined with anti- $\beta$ -actin antibody (15G5A11/E2, Invitrogen, ThermoFisher) and NOD2 polyclonal antibody (PA5-19991, Invitrogen, ThermoFisher, USA). ECL-plus<sup>TM</sup> western blotting chemiluminescence detection kits (BD Biosciences, New Jersey, USA) were then utilized to visualize the bands.

## 2.13. Cell viability and transwell assays

Transfected 786-O cells (either with siRNA or negative control) were plated at a density of  $3 \times 10^3$  cells/well in 100 $\mu$ L medium in 96-well plates for the Cell Counting Kit-8 (CCK-8, KeyGEN BioTECH, Nanjing, China) test. Next, the cells were incubated in a serum-free medium (SFM) with 10% CCK8 for 2 h, as per the recommendations of the manufacturer, after adhesion to the wall. Then, absorbance values at 450 nm OD were recorded on days 1, 2, 3, 4, and 5 utilizing an automated microplate reader (TEAN, Swiss). The samples were analyzed three times independently. Specifically, 24-well plates equipped with Transwell chambers (Corning Costar, Corning, NY, USA) were employed to conduct the transwell tests. In brief,  $1 \times 10^5$  cells in 100 $\mu$ L serum-free DMEM medium were seeded in the top chamber devoid of matrigel, and the bottom chamber was then introduced with 800 $\mu$ L of DMEM medium with 20% FBS, and the cells were grown for 24 h in an incubator. The cells in the bottom chamber were fixed, washed, and labeled with 0.5% crystal violet before being counted in six separate randomized fields utilizing a light microscope (400 $\times$ ).

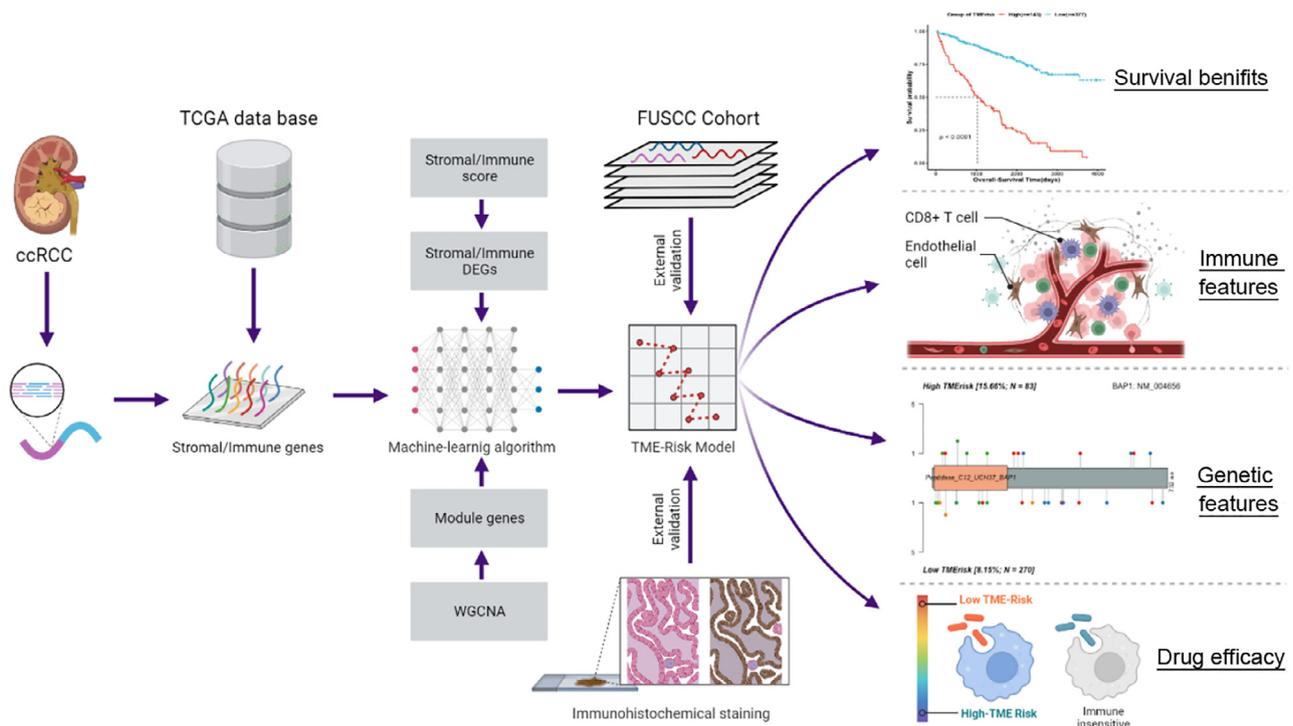


Fig. 1. Computational and experimental workflow for tumor microenvironment risk model. ccRCC, clear cell renal cell carcinoma.

#### 2.14. Statistical analysis

In the statistical analyses, the variations across the two groups of samples were evaluated using the Wilcox test, whereas the variations across multiple groups were evaluated by applying the Kruskal rank test. Moreover, the Pearson correlation test was adopted to examine the links between the groups. All analyses were completed in the R 3.5.2 version. The R tools “ggplot2” and “ggpubr” were employed to generate all of the statistical graphs, whereas the “survival” and “survminer” tools were employed to generate all of the survival graphs. All hypothetical tests were two-sided, and a  $P$ -value of less than 0.05 indicated a significant level in all tests.

### 3. Results

TME is linked to a variety of cancerous biological mechanisms, such as oncogenesis, cellular metabolic irregularities, and aberrant immunological modulation. This investigation into the prognosis-associated TMErisk model and its links to immune characteristics, gene mutations, and drug effectiveness was carried out in three stages (Fig. 1). Firstly, 520 ccRCC cases with available RNA-seq data and respective clinical data from the TCGA cohort were enrolled to construct the TMErisk model using a machine-learning algorithm. Secondly, correlation analyses were undertaken between the TMErisk model and clinical parameters, HLA family genes, immune checkpoint-related genes, immune cell infiltrations, and gene mutations. Lastly, the TMErisk model was used to predict responses to 138 drug efficacy and validate the immunotherapy response for ccRCC patients in 172 cases who received Nivolumab.

#### 3.1. Stromal and immune score landscape in ccRCC

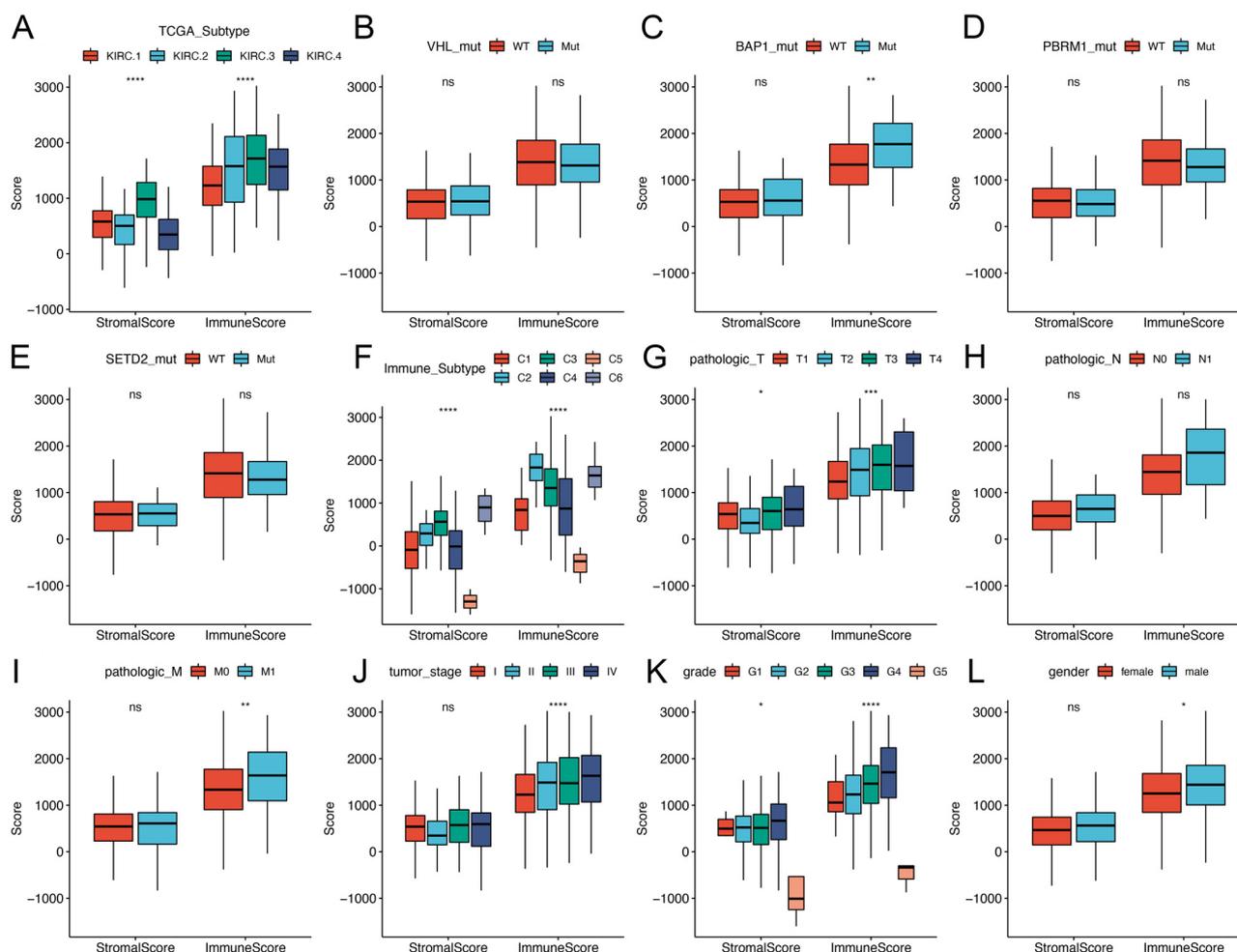
To begin with, we matched transcriptome data with clinical information of ccRCC patients from the TCGA dataset and further obtained 520 cases for the next analyses. Also, the immune and stromal scores were computed by the “ESTIMATE” method to indicate the level of immune cells and stromal cells in the TME, respectively. In the subgroup analysis, the stromal and immune scores revealed significant variations between

the TCGA subtypes ( $P < 0.0001$ ) and immune subtypes ( $P < 0.0001$ ) (Fig. 2A, F). Besides, *BAP1* gene mutation status also suggested a remarkable variation in the immune score ( $P < 0.01$ ), but not significant in the stromal score. *VHL*, *PBRM1*, and *SETD2* mutation status was not significant in both stromal and immune scores (Fig. 2B-E). In the clinicopathological characteristics, higher pT stage ( $P < 0.001$ ), pM stage ( $P < 0.01$ ), tumor stage ( $P < 0.0001$ ), tumor grade ( $P < 0.0001$ ), and gender ( $P < 0.05$ ) demonstrated substantial variations in the immune scores however, only elevated pT stage ( $P < 0.05$ ) and tumor grade ( $P < 0.05$ ) showed distribution significance in the stromal score (Fig. 2G-L).

We conducted survival analysis between high- and low-stromal/immune categories, and the findings illustrated that patients in the high stromal ( $P = 0.0130$ ) and high immune ( $P = 0.0018$ ) categories had dismal prognoses in contrast with those in the low stromal and immune categories (Fig. 3A and B). Furthermore, the “ESTIMATE” technique revealed that the stromal score ( $R = -0.85$ ,  $P < 0.0001$ ) and immune score ( $R = -0.92$ ,  $P < 0.001$ ) were inversely linked to tumor purity (Fig. 3C and D). CPE methods were utilized to further validate the significant contribution of tumor purity to the stromal and immune scores, and the findings indicated that tumor purity was inversely linked to the stromal score ( $R = -0.63$ ,  $P < 0.0001$ ) and immune score ( $R = -0.7$ ,  $P < 0.0001$ ) as well (Fig. 3E and F). As shown in the volcano plot, we screened out 571 stromal score-related DEGs (491 up-regulated, 80 down-regulated) and 573 immune score-related DEGs (530 up-regulated, 43 down-regulated) (Supplementary Fig. 1).

#### 3.2. Development of the ccRCC coexpression modules

Cluster analyses of 520 ccRCC samples from the TCGA dataset carried out using WGCNA revealed the stromal and immune signature-related modules. In particular, we established the power value as  $\beta = 2$  (scale free = 0.83) to construct a scale-free network, which impacted the scale independence and mean connectivity of the coexpression module (Supplementary Fig. 2C and D). As per the outcomes of the WGCNA analysis, genes exhibiting comparable expression profiles were clustered into a module, and 31 coexpression gene modules were found (the genes clustered in gray modules exhibited no coexpression) (Fig. 3G). The coex-



**Fig. 2.** Correlations between stromal and immune scores and clinicopathological features in patients with clear cell renal cell carcinoma. (A) Distribution of stromal and immune scores between different TCGA subtypes. (B-E) Correlations between VHL, BAP1, PBRM1, SETD2 mutation status and stromal/immune score. (F) Distribution of stromal and immune score between different immune subtypes. (G-I) Correlations between pathological TNM stages and stromal/immune score. (J-L) Correlations between tumor stage, pathological grade, gender and stromal/immune score. Mut, mutant; ns, not significant; WT, wild type.

pression modules' relationship to the stromal/immune score is depicted in Fig. 3H. Modules with  $R > 0.5$  and  $P$  values  $< 0.05$  were selected for further study, of which the modules colored in blue, royal blue, salmon, light cyan, dark red, and midnight blue had strong correlations with the stromal score, while the modules colored in the salmon, dark orange, tan, light cyan, and saddle brown showed strong correlations with the immune score.

### 3.3. Construction of TMErisk model

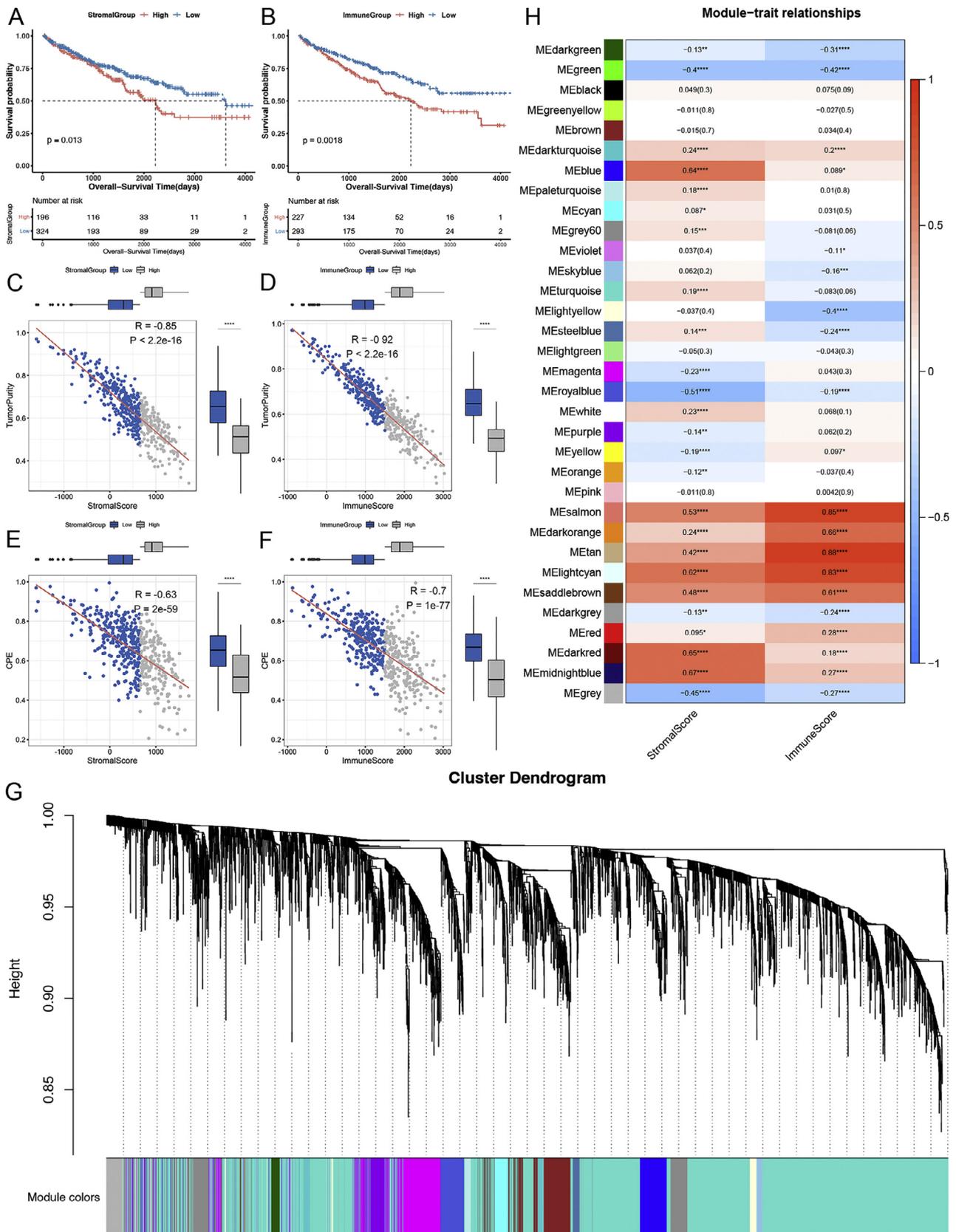
The Venn diagram shows the intersecting DEGs between stromal and immune-associated module genes, of which 442 stromal-related intersecting genes and 455 immune-related intersecting genes were identified (Fig. 4A and B). Prognostic values of these intersecting genes were further assessed. Overall, 162 prognostic genes associated with the stromal score and 165 associated with the immune score were found. Fig. 4C and D displays the top 20 prognostic genes and hazard ratio (HR) with a 95% confidence interval (CI). Moreover, lasso regression analysis was used in reducing the number of prognostic genes, reducing overfitting between genes, and selecting 37 candidate stromal and immune-related prognostic genes (Fig. 4E and F). Also, we verified the expression patterns and prognostic significance of these 37 candidate genes in our proteomic dataset (FUSCC cohort), and the results indicated that a total of 8 corresponding proteins revealed significant differ-

ences in both expression and prognosis, including COL5A3, COL11A1, IFITM1, CPA3, UNC13D, RNASE2, TCIRG1 and ALDH6A1 (Supplementary Fig. 3). Subsequently, the 37 candidate genes were incorporated into the multivariate analysis, and the prognostic genes consisting of *INAFM2*, *SRPX*, *DPYSL3*, *VSIG4*, *APLNR*, *FHL5*, *A2M*, *SLFN11*, *ADAMTS4*, *IFITM1*, *NOD2*, *CCR4*, *HLA-DQB2*, and *PLAUR* were chosen to develop the TMErisk model ( $\text{TMErisk score} = \sum \text{Ni} = 1 (\text{Exp } i * \beta_i)$ ).

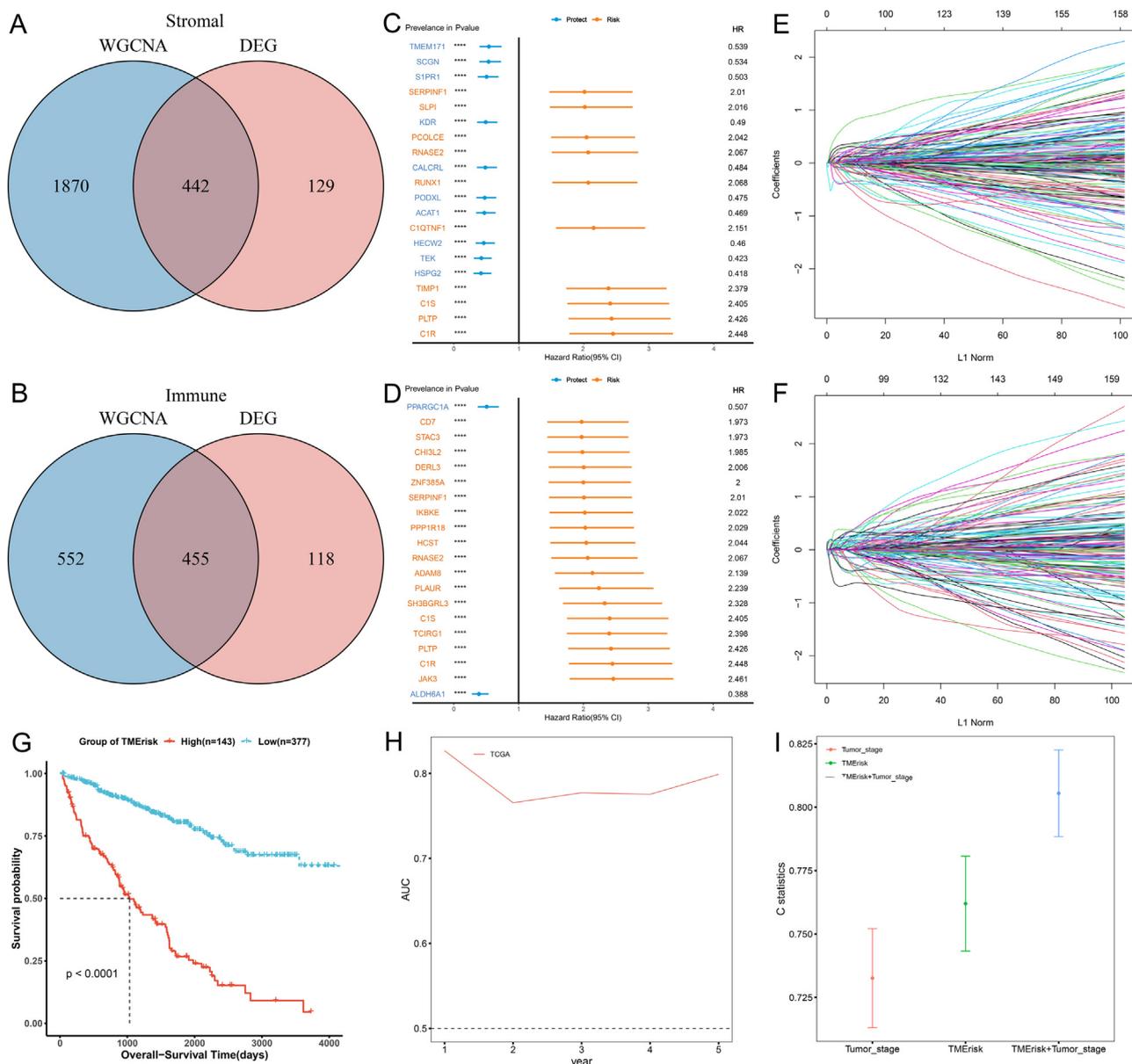
### 3.4. Prognostic value of TMErisk model

After computing the TMErisk score for each case, patients were classified into low- ( $n = 143$ ) and high-risk ( $n = 377$ ) categories as per the optimal threshold of the risk score. Survival studies showed a strong link between patients in the high-risk category and a grim prognosis ( $P < 0.0001$ ) (Fig. 4G), and the ROC curve also indicated the high prediction specificity and sensitivity of the TMErisk model in the OS over 5 years (area under the curve [AUC]  $> 0.75$ ) (Fig. 4H). C-index demonstrated that the TMErisk model had a higher predictive discrimination ability than the traditional prognostic model tumor stage and the predictive ability was further improved when the TMErisk model was integrated with the tumor stage (Fig. 4I).

Further, the expression pattern and prognostic significance of CCR4 and NOD2 in the FUSCC cohort were validated using IHC staining, as the coefficients of CCR4 (coef = 0.737) and NOD2 (coef = 0.717) are the



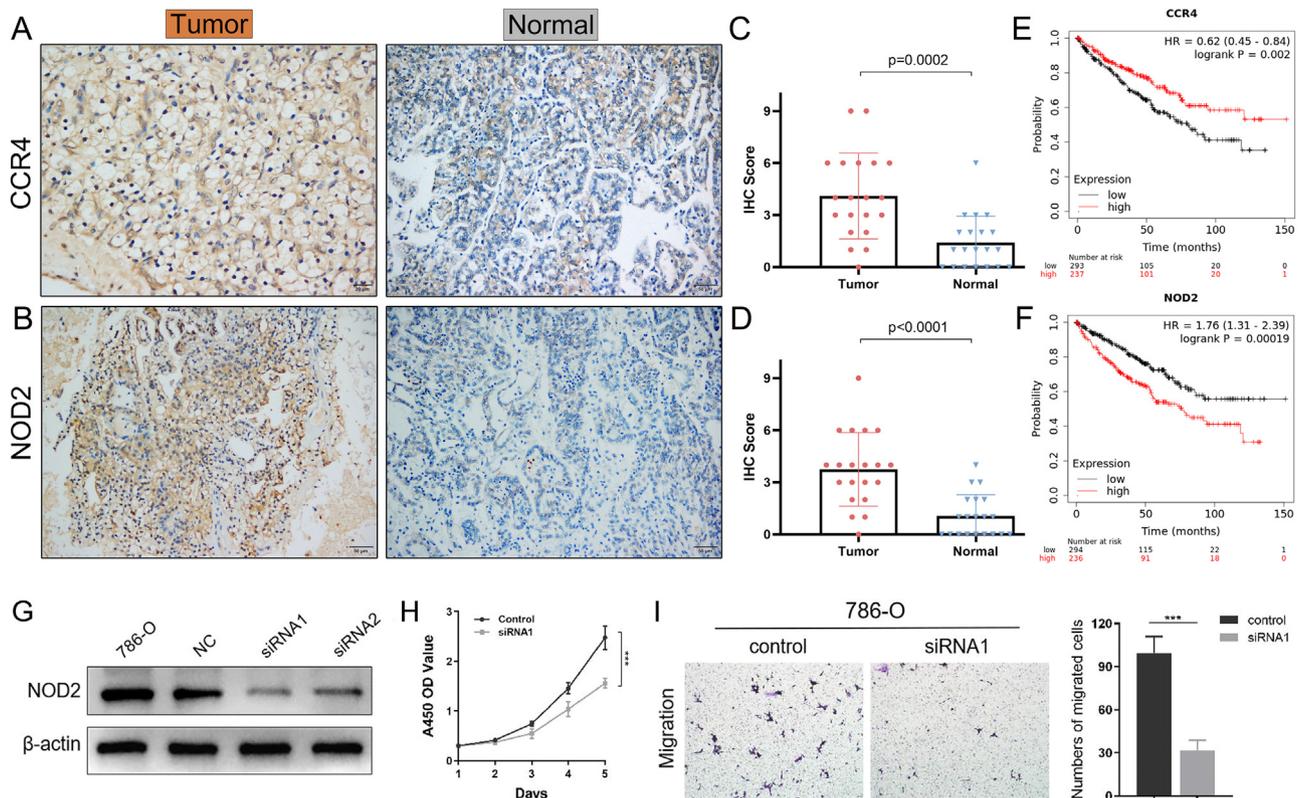
**Fig. 3.** Survival analysis of stromal/immune groups and identification of WGCNA co-expression modules. (A, B) Survival curves of low- and high-stromal (A) and -immune (B) groups. (C, D) Correlation between tumor purity and stromal/immune score. (E, F) Correlation between CPE score and stromal/immune score. (G) Clustering dendrograms of the gene modules based on the results of WGCNA analysis. Highly similar modules are identified by clustering and then merged dynamically. (H) A heatmap of the correlations between module eigengenes and stromal/immune score. The row represents distinct eigengene modules, and the column represents distinct signature. The corresponding correlation and *P* value are shown in each cell. CPE, consensus measurement of purity estimations.



**Fig. 4.** Construction of TMErisk model. (A) Intersection of stromal-related DEGs with WGCNA stromal-related modules. (B) Intersection of immune-related DEGs with WGCNA immune-related modules. (C, D) Forrest plot showed the results of univariate Cox regression analysis of stromal intersection genes (C) and immune intersection genes (D). (E, F) Lasso regression analysis of stromal-related prognostic genes (E) and immune-related prognostic genes (F). (G) Survival analysis between low- and high-TMErisk groups. (H) AUC curve for 1 to 5-year survival of TMErisk model. (I) C-index showed the prognostic predictive ability of tumor stage, TMErisk model and TMErisk model + tumor stage. AUC, area under the curve; DEGs, differentially expressed genes; TME, tumor microenvironment.

highest in the TMErisk model. Fig. 5A-D demonstrates that CCR4 and NOD2 expression levels were elevated in tumor tissues but lowered in normal tissues. Survival curves also indicated that the attenuated expression of CCR4 ( $P = 0.002$ ) and enhanced expression of NOD2 ( $P = 0.0002$ ) were considerably linked to unfavorable outcomes of ccRCC patients (Fig. 5E and F). To reveal malignant biological behaviors of hub oncogene NOD2, we first validated the downregulation of NOD2 in the siRNA-transfected group compared with the control group in 786-O cells (Fig. 5G). Cell viability test results demonstrated that ccRCC cell proliferative capacity was considerably suppressed when NOD2 was downregulated ( $P < 0.001$ ; Fig. 5H). Besides, a Transwell migration experiment showed that ccRCC cell migration was remarkably attenuated due to NOD2 downregulation ( $P < 0.001$ ; Fig. 5I). Taken together, the decreased NOD2 expression significantly restrained the potential of ccRCC cells to proliferate and migrate in vitro.

As shown in Supplementary Fig. 4, the Forest plot indicated that, in age, tumor stage, pTNM stage, tumor grade, and gender stratification analyses, the TMErisk model served as a risk indicator to the OS in an independent way. The TMErisk model suggested significant distribution differences among the TCGA subtypes ( $P < 0.0001$ ), of which subtype 3 was the group with the highest TMErisk score and the worst prognosis among the four subtypes, illustrating that a higher TMErisk score is correlated with poor outcomes (Supplementary Fig. 5A). Besides, our analysis indicated that the TMErisk model also revealed significant distribution differences between immune subtypes ( $P < 0.0001$ ), with subtype C6 having the highest TMErisk score and also being recognized as a high lymphocytic infiltration subtype with the worst outcome, which is consistent with our findings (Supplementary Fig. 5F). We also analyzed the association of the TMErisk model with gene mutation status, and the results showed that patients with the *BAP1* mutation and the *SETD2*



**Fig. 5.** External validation of the hub oncogenes CCR4 and NOD2 in FUSCC cohort. (A, B) Expression of CCR4 (A) and NOD2 (B) in tumor and normal tissues. Immunohistochemical score indicated the significantly different expression of CCR4 (C) and NOD2 (D) between tumor and normal tissues. (E, F) Survival analysis between CCR4 low- and high-expression groups (E) and between NOD2 low- and high-expression groups (F). (G) Expression levels of NOD2 in the control group and the siRNA-transfected group. (H) Down-regulated NOD2 expression significantly inhibited the proliferative ability of ccRCC cells than the control group. (I) Down-regulated expression of NOD2 significantly suppressed the migration capacity of ccRCC cells.

mutation had higher TMErisk scores (Supplementary Fig. 5C and E). In addition, we noticed that elevated pT stage ( $P < 0.0001$ ), pN stage ( $P < 0.01$ ), pM stage ( $P < 0.0001$ ), tumor stage ( $P < 0.0001$ ), and tumor grade ( $P < 0.0001$ ) were correlated with higher TME scores (Supplementary Fig. 5G-L).

### 3.5. Pathway enrichment analysis

Comparative assessment of the functional annotations of those with low and high TMErisk was done by means of GSEA. The results showed that 55 KEGG terms in total (with the high- and low-TMErisk categories containing 10 and 45 terms, correspondingly), 103 molecular functions (MF) terms (with the high- and low-TMErisk categories containing 25 and 78 terms, correspondingly), 424 biological processes (BP) terms (with the high- and low-TMErisk categories containing 192 and 232 terms, respectively), 64 cellular components (CC) terms (with the high- and low-TMErisk categories containing 19 and 45 terms, correspondingly) were significantly enriched (Supplementary Fig. 6). Details and top ten terms of the clustering results were summarized in Supplementary Table 2.

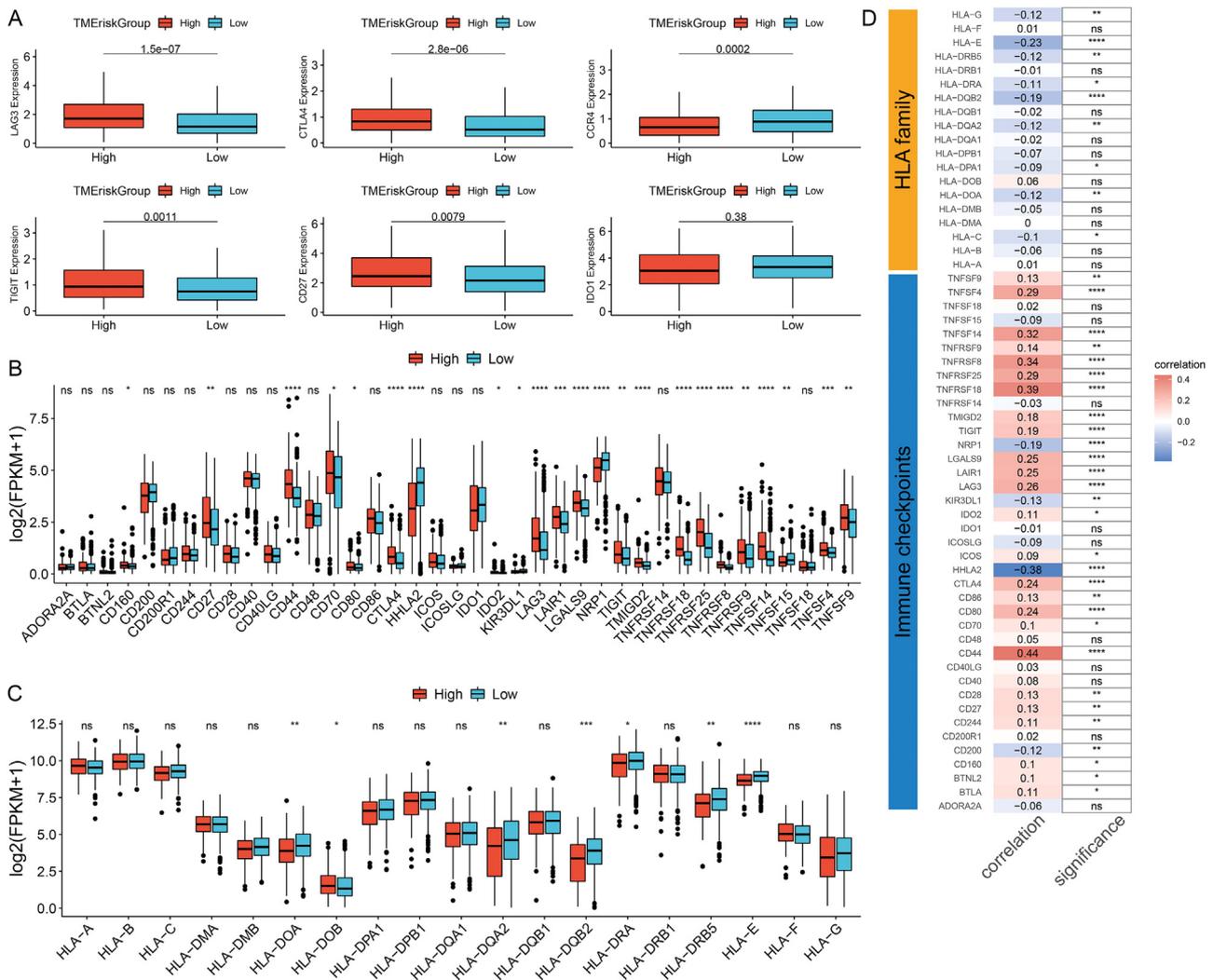
### 3.6. Association of TMErisk model with immune features

We examined links between the TMErisk groups and tumor cell components. High-TMErisk score was strongly linked to greater immune scores and reduced tumor purity (Supplementary Fig. 5), suggesting that the infiltration of immune cells was higher in the advanced tumor, while the tumor cells were less abundant.

The TMErisk model and the six major ICGs were then subjected to comparison. As depicted in Fig. 6A, *LAG3* ( $P < 0.0001$ ), *CTLA4* ( $P <$

$0.0001$ ), *CCR4* ( $P = 0.0002$ ), *TIGIT* ( $P = 0.0011$ ), and *CD27* ( $P = 0.0079$ ) were differently expressed between the low- and high-TMErisk categories. We then further analyzed the expression of 19 HLA family genes and 39 ICGs in the TMErisk model. Per the results of the Wilcoxon test, the expression levels of 7 HLA family genes (Fig. 6B) and 23 ICGs (Fig. 6C) had significant variations in the high- and low-TMErisk categories. Besides, as reflected in Fig. 6D, the heat map revealed the correlations between TMErisk score and HLA family genes together with ICGs, of which *CD44* ( $R = 0.44$ ,  $P < 0.0001$ ), *TNFRSF18* ( $R = 0.39$ ,  $P < 0.0001$ ), *TNFRSF8* ( $R = 0.34$ ,  $P < 0.0001$ ) were the top three positively correlated genes, while *HHLA2* ( $R = -0.38$ ,  $P < 0.0001$ ), *HLA-E* ( $R = -0.23$ ,  $P < 0.0001$ ), *NPR1* ( $R = -0.19$ ,  $P < 0.0001$ ) were the top three negatively correlated genes.

Three algorithms (TIMER, CIBERSORT, Xcell) were adopted to demonstrate the infiltration status of immune and stromal cells in the TME. Furthermore, the results were visualized in the heat map after normalizing the infiltration fractions (Fig. 7). Infiltration levels of B cells ( $P < 0.01$ ) and macrophages ( $P < 0.05$ ) were shown to be substantially varied across the low- and high-TMErisk categories, as depicted by the TIMER method. Through the CIBERSORT algorithm, we noticed that T cell follicular helper ( $P < 0.0001$ ), Tregs ( $P < 0.0001$ ), macrophage M2 ( $P < 0.001$ ), mast cell activated ( $P < 0.001$ ), T cell CD4<sup>+</sup> memory activated ( $P < 0.001$ ), Macrophage M0 ( $P < 0.001$ ), T cell CD4<sup>+</sup> memory resting ( $P < 0.05$ ), NK cell activated ( $P < 0.05$ ) and B cell memory ( $P < 0.05$ ) had substantially varied infiltration levels between low- and high-TMErisk categories. Xcell algorithm indicated considerable variations in the infiltration levels of several stromal and immune cells between the two TMErisk groups, particularly in hematopoietic stem cells ( $P < 0.0001$ ), endothelial cells ( $P < 0.0001$ ), T cell CD4<sup>+</sup> Th1 ( $P < 0.0001$ ), T cell NK ( $P < 0.0001$ ), B cells ( $P < 0.0001$ ), B cell naive ( $P <$



**Fig. 6.** Correlation between TME risk groups and immune checkpoint genes and the HLA family. (A) Differential expression of immune checkpoints, LAG3, CTLA-4, CCR4, TIGIT, CD27, IDO1 in TME low- and high-risk group. (B, C) Expression of immune checkpoint genes and HLA families in TME low- and high-risk groups. (D) Heat map of correlations between TME risk and immune checkpoint genes and HLA family. HLA, human leukocyte antigen; TME, tumor microenvironment.

0.0001), monocytes ( $P < 0.001$ ), B cell plasma ( $P < 0.001$ ), plasmacytoid dendritic cells ( $P < 0.001$ ), common lymphoid progenitors ( $P < 0.01$ ), macrophage M1 ( $P < 0.01$ ), mast cells ( $P < 0.01$ ), myeloid dendritic cells activated ( $P < 0.05$ ), T cell CD4<sup>+</sup> Th2 ( $P < 0.05$ ), macrophage M2 ( $P < 0.05$ ), T cell CD8<sup>+</sup> effector memory ( $P < 0.05$ ), B cell memory ( $P < 0.05$ ), and T cell CD4<sup>+</sup> ( $P < 0.05$ ). Additionally, we also investigated the distribution of tumor mutation burden (TMB), somatic mutations, and clinical characteristics of ccRCC and displayed them in the heat map (Fig. 7).

Prognostic analyses were also performed on cells that had significant differential infiltrations in the TME risk groups, and the results were visualized in Fig. 8. High infiltration of T cell follicular helper ( $P < 0.0001$ ), Tregs ( $P = 0.0058$ ), NK cell activated ( $P = 0.0210$ ), B cell ( $P = 0.0004$ ), monocyte ( $P = 0.0230$ ), B cell naïve ( $P = 0.0033$ ), T cell CD4<sup>+</sup> Th1 ( $P = 0.0002$ ), T cell CD4<sup>+</sup> Th2 ( $P = 0.0280$ ) were remarkably linked to the favorable OS, but high infiltration of mast cell activated ( $P = 0.0002$ ), endothelial cell ( $P = 0.0260$ ), hematopoietic stem cell ( $P < 0.0001$ ) was substantially linked to with poor OS (Fig. 8A-N).

### 3.7. Analyses of somatic mutations of ccRCC

To start with, we compared the mutation frequency between the high- and low-TME risk categories. More somatic mutations, includ-

ing synonymous and non-synonymous mutations, were observed in the high-TME risk category ( $P < 0.0001$ ) (Fig. 9A-C). Next, we selected genes with more than 20 mutations as high-frequency mutation genes in ccRCC, including *VHL* (46% mutation), *PBRM1* (40% mutation), *TTN* (19% mutation), *SETD2* (13% mutation), *BAP1* (10% mutation), *MUC16* (8% mutation), *MTOR* (8% mutation). The high-frequency mutation gene *BAP1* has prognostic efficacy and is a risk factor for mutation status prognosis ( $P = 0.024$ ) (Fig. 9D). Notably, we observed that *BAP1* mutations and *PBRM1* mutations ( $P < 0.01$ ) were significantly mutually exclusive, and *PBRM1* mutations significantly co-occurred with *VHL* mutations and *SETD2* mutations ( $P < 0.01$ ) (Fig. 9E). In addition, we discovered that the high-TME risk category showed higher co-mutation rates (50.6% vs. 41.9%) of the seven high-frequency mutation genes. Moreover, the low-TME risk patients showed greater single mutations (35.6% vs. 31.3%) and wild type (22.6% vs. 18.1%) than those in the high-TME risk patients ( $P < 0.0001$ ) (Fig. 9F). We further compared the different mutation statuses of *VHL*, *SETD2*, and *PBRM1*, and the findings suggested that the high-TME risk subgroup also showed elevated proportions of *VHL*-*PBRM1* (25.3% vs. 21.5%) and *SETD2*-*PBRM1* (12% vs. 7%) co-mutations ( $P < 0.0001$ ) (Fig. 9F). The findings above elaborated the high-TME risk patients exhibited a greater probability of having co-mutations, implying that co-mutations may be a high-risk factor in ccRCC. We also analyzed the mutation sites and mutation types of *BAP1*

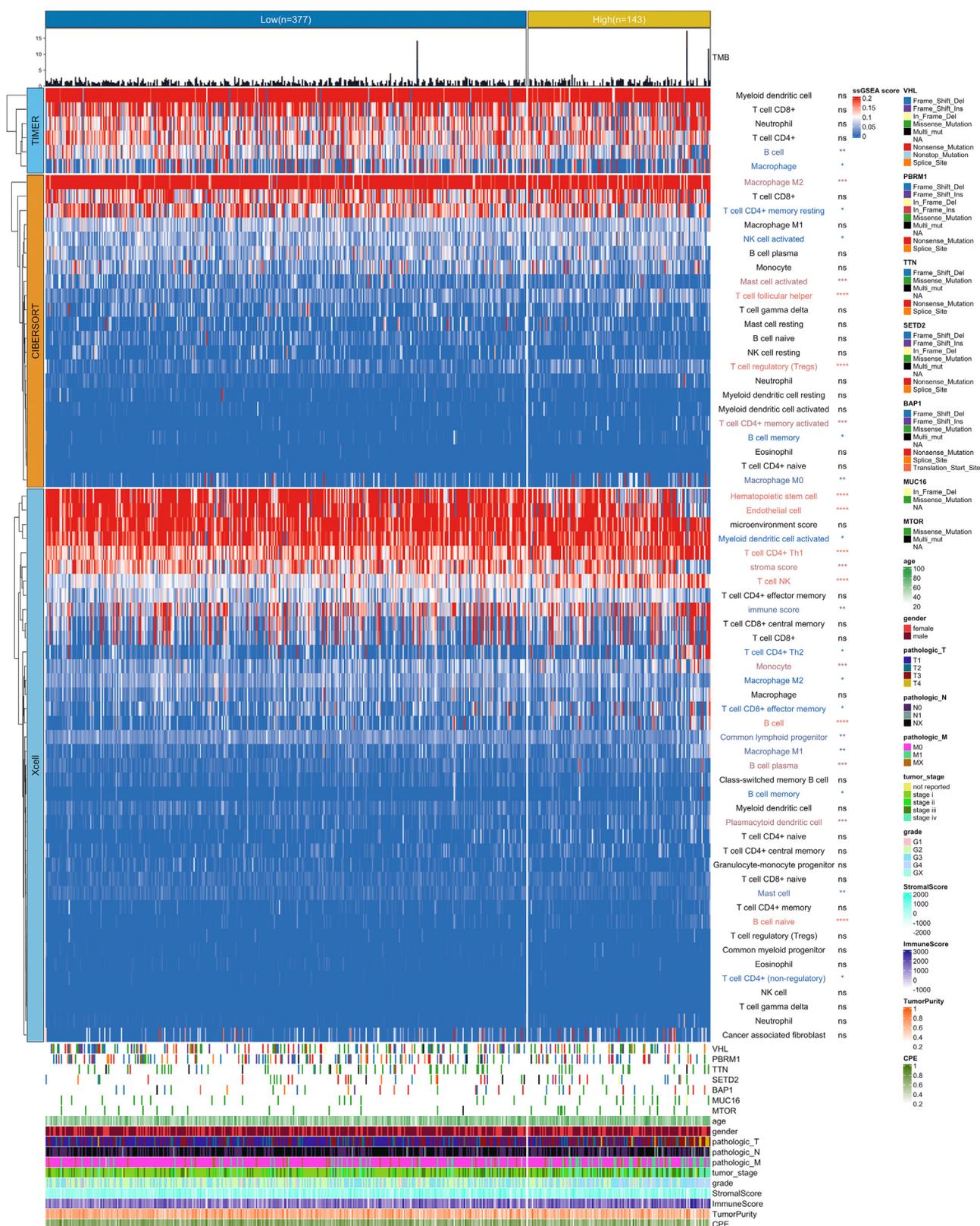
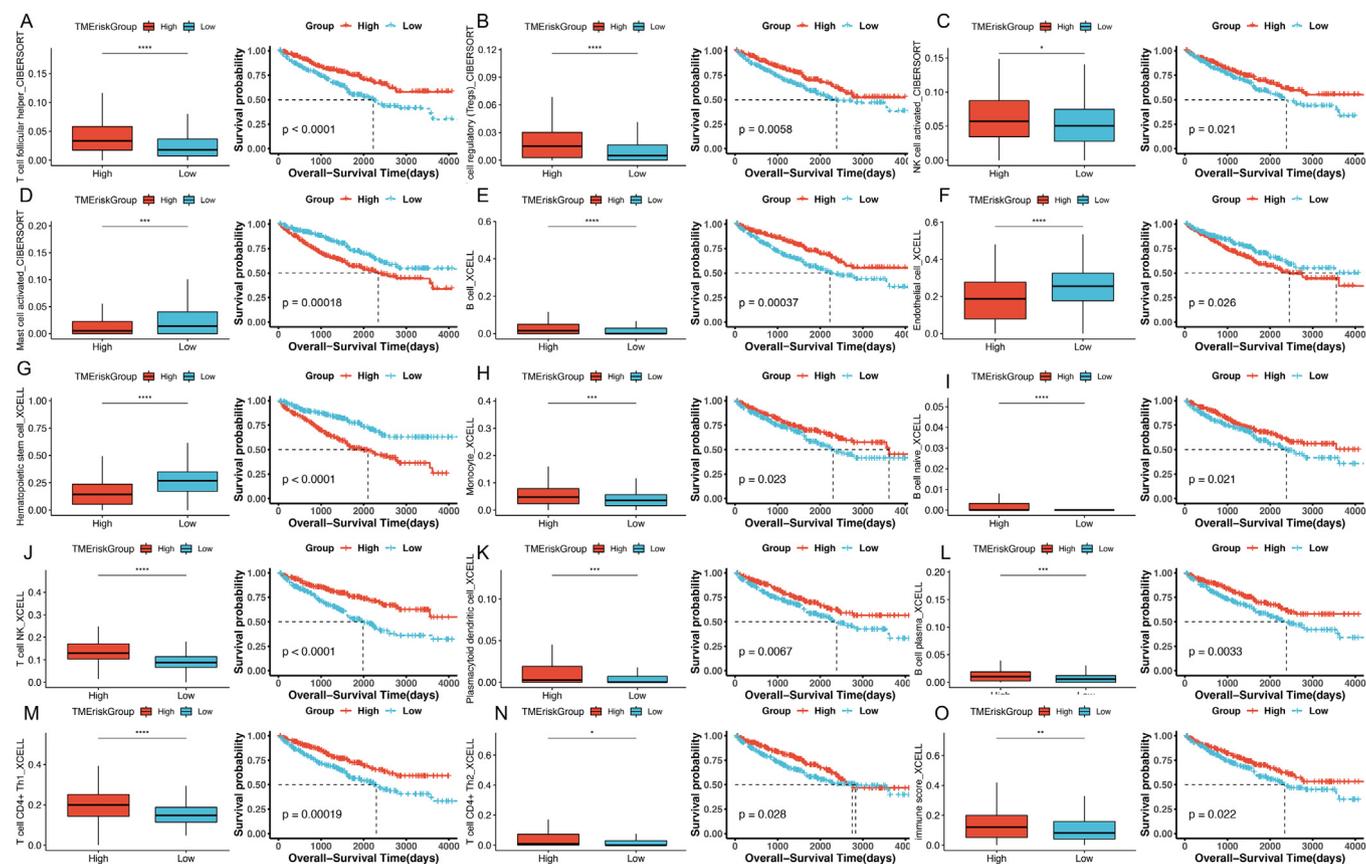


Fig. 7. Heat map of correlation between TMErisk group and immune cell infiltration. The heat map shows standardized fractions for immune and stromal cell infiltration. Blue represents cells with low levels of infiltration and red represents cells with higher levels of infiltration.



**Fig. 8.** Cell infiltration and survival differences in different TMErisk groups. Cells with significantly differential infiltration in low- and high-TMErisk groups and survival differences between low and high infiltrated cellular groups. TME, tumor microenvironment.

between the high- and low-TMErisk categories, and the outcomes were displayed in Fig. 9G.

### 3.8. Drug sensitivity and efficacy analysis

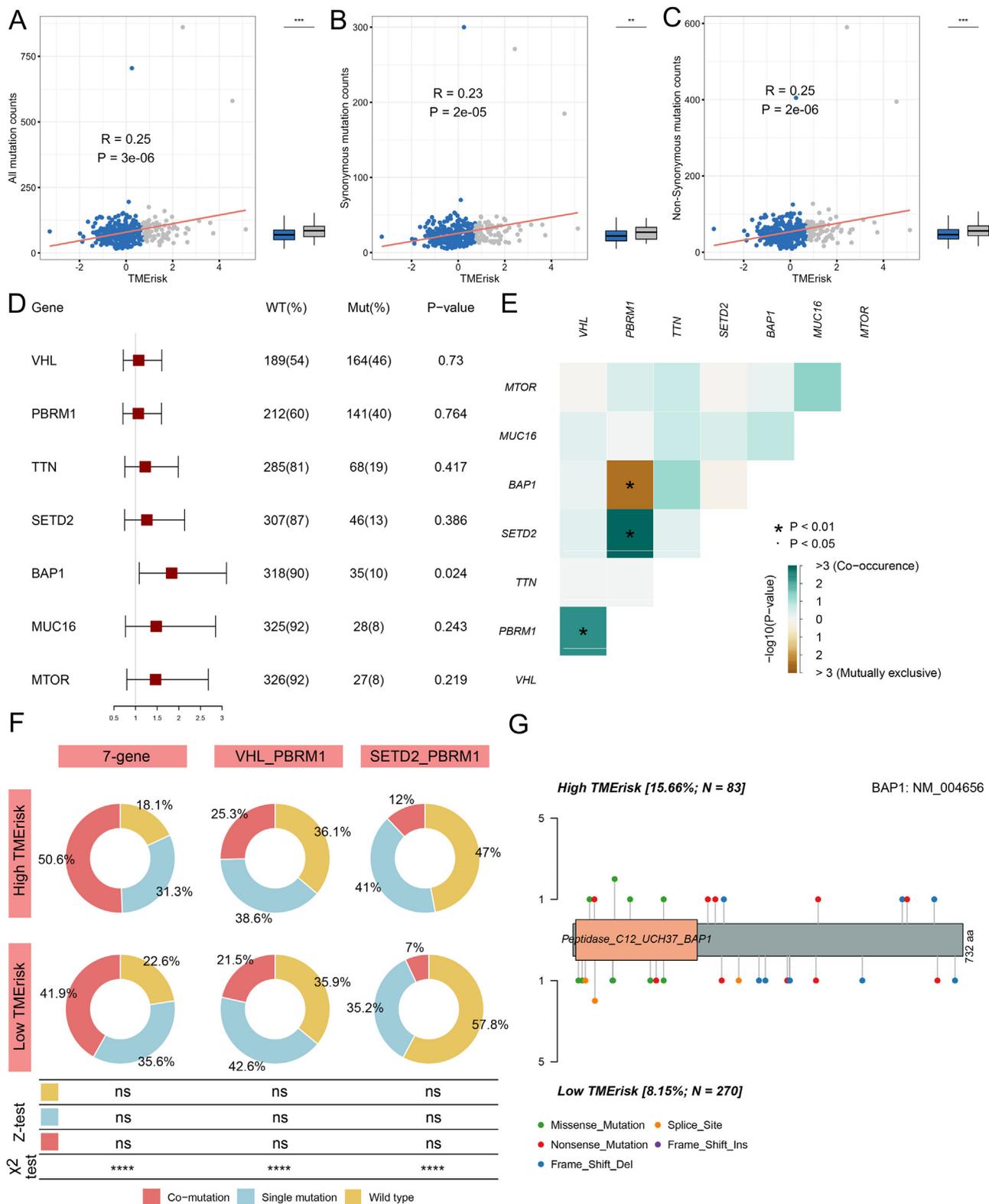
We analyzed the IC<sub>50</sub> of 138 medications as per the gene expression data of ccRCC patients from TCGA, and the results were visualized in Fig. 10A. Tumor immunodeficiency and exclusion (TIDE) scores were utilized to reflect the patient’s potential response to immunotherapy, but we did not observe the differential distribution of TIDE scores between TMErisk groups (Fig. 10B). Additionally, we analyzed the effectiveness of immunotherapeutic regimens in 172 patients who were treated with Nivolumab. Accordingly, the results suggested that 19.3% of the high-TME-risk patients achieved complete (CR) or partial response (PR) as opposed to the low-TME-risk patients (13.3%) (Fig. 10C), showing that high-TME-risk individuals benefited significantly from immunotherapy. Nevertheless, no significant variation was identified in TMErisk scores between patients with PR or CR and those with progressive disease (PD) or stable disease (SD) ( $P = 0.28$ ) (Fig. 10D).

## 4. Discussion

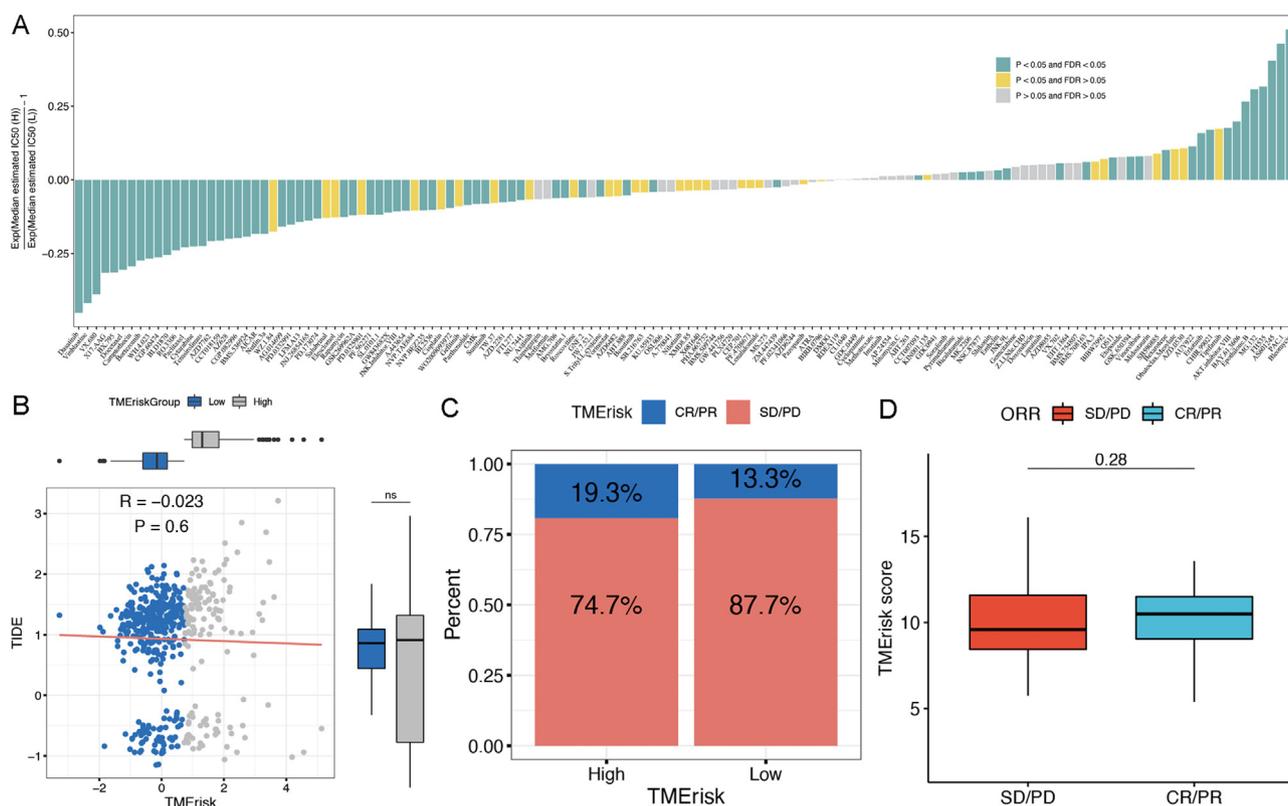
The component of the ccRCC microenvironment can affect disease development, cellular metabolism, and immune regulation. There were differences in prognosis and therapeutic efficacy between individuals, which may be due to the differences in the cell-type-specific TMEs. To reveal the interactions between cells in the TME and identify hub genes that alter the TME in ccRCC, we constructed the TMErisk model and investigated its significance in prognosis and correlations with the immune features, gene mutations, and drug efficacy. The TMErisk model

indicated that the high-TME-risk patients exhibited a grim prognosis, and our model also served as a risk indicator to the OS in an independent way. An increased risk score for TME was highly related to a worse outcome in our study. At the same time, we found that the TMErisk model showed significant distribution differences among the TCGA subtypes of ccRCC, of which subtype 3 exhibited the worst prognosis and also had the highest TMErisk score, proving again that a higher TMErisk score is correlated to poor outcomes. In addition, we noticed that elevated tumor stage, pTNM stage, and tumor grade were remarkably correlated to higher TME scores. Besides, our analysis indicated that the TMErisk model also revealed significant distribution differences among immune subtypes. In particular, subtype C6 had the highest TMErisk score and also was recognized as a high lymphocytic infiltration subtype with the worst outcome, which is consistent with our findings. Angiogenesis may be induced by a range of vascular active chemicals secreted by renal carcinoma tissues, which is also one of the targets of advanced renal cancer treatment.<sup>29</sup> The immune subtype C1 had elevated expression of angiogenic genes<sup>20</sup> and also correlated with a higher TMErisk score in our analysis, indicating that our risk model conforms to the biological characteristics of kidney cancer.

With the increasing evidence of the effectiveness of immunotherapy in the treatment of renal cancer, it is of interest to study the potential molecular biomarkers that potentially affect the efficacy of immunotherapy. HLA family molecules and ICGs are implicated in the recognition, antigen presentation, and T cell-mediated apoptosis of malignant cells, and their impairment may promote tumor immune evasion.<sup>30</sup> In the present study, we identified that some HLA molecules, including HLA-DQB2, HLA-DQA2, HLA-E, and HLA-DOA, showed substantially low expression in the high-TMErisk category, and some immune checkpoint



**Fig. 9.** Somatic mutations and correlation with TMErisk model. (A-C) The association between all mutation counts, synonymous mutation counts, non-synonymous mutation counts, and TMErisk score. Blue and gray represent low- and high-TMErisk group, respectively. (D) Forest plot indicate the prognostic ability of genes with more than 20 mutations. (E) Interaction of high-frequency mutated genes. (F) Distribution of mutation types of high frequency mutated genes, VHL-PBRM1 and SETD2-PBRM1 between low- and high-TMErisk group. (G) Lollipop plot of somatic mutation site in BAP1. TME, tumor microenvironment.



**Fig. 10.** Correlation between TMErisk model and immunotherapy. (A)  $IC_{50}$  value for 138 drugs. (B) Correlation between TIDE score and TMErisk score. (C) Proportion of responses to immunotherapy in the low- and high-TMErisk group. (D) Differences in TMErisk score in the response and non-response groups. CR, complete response;  $IC_{50}$ , 50% inhibition concentration; ORR, objective response rate; PD, progressive disease; PR, partial response; SD, stable disease; TIDE, tumor immunodeficiency and exclusion; TME, tumor microenvironment.

inhibitors, including LAG3, CTLA4, CD44, and CD27, were shown to be remarkably highly expressed in the high-TMErisk category. Existing literature also indicated that downregulated HLA molecules and some upregulated ICGs were linked to lower objective response rates and poor prognosis.<sup>31,32</sup> In mutation analyses, we found that more somatic mutations were observed in the high-TMErisk category in contrast with the low-TMErisk category, signifying that high mutation frequency may affect the TME, which may result in carcinogenesis and tumor development. Notably, we also discovered that *PBRM1* mutations frequently co-occurred with *VHL* mutations and *SETD2* mutations, whereas *BAP1* mutations were mutually exclusive with *PBRM1* mutations. Immune response analysis indicated that immunotherapy was more effective in the high-TMErisk patients, but no significant variation was identified in the TMErisk scores between patients with PR or CR and those with PD or SD, which may be attributed to the fewer patients in the high-TMErisk category.

This study constructed and validated the TMErisk model of ccRCC. To additionally emphasize the link between the RCC microenvironment and patient prognosis, this model was shown to have excellent accuracy and sensitivity in predicting prognoses and may be used as a powerful prognostic indicator in an independent way. We also scrutinized how the TMErisk model correlated with mutation status, immune checkpoints, immune cell infiltration, and the sensitivity to immunotherapeutic regimens, suggesting that a high-TMErisk score was linked to a greater incidence of deleterious mutations and immune cell infiltration that aided in tumorigenesis. Besides, there was solid evidence in this study depicting that the high-TMErisk score of ccRCC strongly predicted the responsiveness of patients to immunotherapy.

Nonetheless, there were several drawbacks to this research. The mechanism of the genes implicated in the TMErisk model has not been elucidated in this work. Our future research will focus on elu-

cidating the real biological functions of these genes and determining how they work with TME to promote oncogenesis. In addition, although the TMErisk model was developed and verified by several publicly accessible and real-world datasets, additional multicenter investigations and prospective studies are needed before they can be used clinically for patients with ccRCC, due to the limitation of retrospective analysis.

#### Declaration of competing interest

The authors declare that they have no conflict of interests.

#### Ethics approval and consent to participate

All of the study designs and test procedures were performed in accordance with the Helsinki Declaration II. The ethics approval and participation consent of this study was approved by the ethics committee of Fudan University Shanghai Cancer Center (Ethical IRB number: 050432-4-1911D). All patients participating in this study signed informed consent forms.

#### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

#### Acknowledgement

We thank researchers for providing public datasets online, and for patients enrolled from the FUSCC cohort. Thanks for the professional language editing by Miss. Li and Bullet Edits. This work was

supported by grants from the National Natural Science Foundation of China (grant numbers: 81802525 and 82172817), the Natural Science Foundation of Shanghai (grant number: 20ZR1413100), Beijing Xisike Clinical Oncology Research Foundation (grant number: Y-HR2020MS-0948), the National Key Research and Development Project (grant number: 2019YFC1316005), the Shanghai “Science and Technology Innovation Action Plan” Medical Innovation Research Project (grant number: 22Y11905100) and the Shanghai Anti-Cancer Association Eyas Project (grant numbers: SACA-CY21A06 and SACA-CY21B01).

#### Author contributions

D.Y., H.Z., and J.Y. conducted the conception and design and administrated the support. A.A., W.X., W.L., S.W., Y.Q and X.T. performed provision of study materials or patients, collection and assembly of data, data analysis and interpretation. A.A. and W.X. conducted graphs plotting. All of the authors drafted and revised this manuscript.

#### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jncc.2023.08.003.

#### References

- Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2022. *CA Cancer J Clin*. 2022;72(1):7–33. doi:10.3322/caac.21708.
- Jonasch E, Gao J, Rathmell WK. Renal cell carcinoma. *BMJ*. 2014;349:g4797. doi:10.1136/bmj.g4797.
- Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012;366(10):883–892. doi:10.1056/NEJMoa1113205.
- Miller KD, Nogueira L, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin*. 2019;69(5):363–385. doi:10.3322/caac.21565.
- Bi K, He MX, Bakouny Z, et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell*. 2021;39(5):649–661 e5. doi:10.1016/j.ccell.2021.02.015.
- Tumeh PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. 2014;515(7528):568–571. doi:10.1038/nature13954.
- Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol*. 2013;14(10):1014–1022. doi:10.1038/ni.2703.
- Fridman WH, Zitvogel L, Sautès-Fridman C, et al. The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol*. 2017;14(12):717–734. doi:10.1038/nr-clinonc.2017.101.
- Braun DA, Street K, Burke KP, et al. Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. *Cancer Cell*. 2021;39(5):632–648 e8. doi:10.1016/j.ccell.2021.02.013.
- Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013;499(7456):43–49. doi:10.1038/nature12222.
- Gerlinger M, Horswell S, Larkin J, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet*. 2014;46(3):225–233. doi:10.1038/ng.2891.
- Varela I, Tarpey P, Raine K, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature*. 2011;469(7331):539–542. doi:10.1038/nature09639.
- Peña-Llopis S, Vega-Rubin-de-Celis S, Liao A, et al. BAP1 loss defines a new class of renal cell carcinoma. *Nat Genet*. 2012;44(7):751–759. doi:10.1038/ng.2323.
- Dalglish GL, Furge K, Greenman C, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature*. 2010;463(7279):360–363. doi:10.1038/nature08672.
- Zhu S, Ding W, Chen Y, et al. High VHL Expression Reverses Warburg Phenotype and Enhances Immunogenicity in Kidney Tumor Cells. *Genomics Proteomics Bioinformatics*. 2022;20(4):657–669. doi:10.1016/j.gpb.2019.12.002.
- Chakiryan NH, Hajiran A, Kim Y, et al. Correlating Immune Cell Infiltration Patterns with Recurrent Somatic Mutations in Advanced Clear Cell Renal Cell Carcinoma. *Eur Urol Focus*. 2022;8(3):784–793. doi:10.1016/j.euf.2021.04.014.
- Brannon AR, Haake SM, Hacker KE, et al. Meta-analysis of clear cell renal cell carcinoma gene expression defines a variant subgroup and identifies gender influences on tumor biology. *Eur Urol*. 2012;61(2):258–268. doi:10.1016/j.eururo.2011.10.007.
- Clark DJ, Dhanasekaran SM, Petralia F, et al. Integrated Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma. *Cell*. 2019;179(4):964–983 e31. doi:10.1016/j.cell.2019.10.007.
- Xu W, Anwaier A, Ma C, et al. Prognostic Immunophenotyping Clusters of Clear Cell Renal Cell Carcinoma Defined by the Unique Tumor Immune Microenvironment. *Front Cell Dev Biol*. 2021;9:785410. doi:10.3389/fcell.2021.785410.
- Thorsson V, Gibbs DL, Brown SD, et al. The Immune Landscape of Cancer. *Immunity*. 2018;48(4):812–830 e14. doi:10.1016/j.immuni.2018.03.023.
- Braun DA, Hou Y, Bakouny Z, et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med*. 2020;26(6):909–918. doi:10.1038/s41591-020-0839-y.
- Yoshihara K, Shahmoradgol M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612. doi:10.1038/ncomms3612.
- Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559. doi:10.1186/1471-2105-9-559.
- Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*. 2017;33(18):2938–2940. doi:10.1093/bioinformatics/btx364.
- Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847–2849. doi:10.1093/bioinformatics/btw313.
- Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome Biol*. 2014;15(3):R47. doi:10.1186/gb-2014-15-3-r47.
- Anwaier A, Zhu SX, Tian X, et al. Large-Scale Proteomics Data Reveal Integrated Prognosis-Related Protein Signatures and Role of SMAD4 and RAD50 in Prognosis and Immune Infiltrations of Prostate Cancer Microenvironment. *Phenomics*. 2022;2(6):404–418. doi:10.1007/s43657-022-00070-1.
- Barata PC, Rini BI. Treatment of renal cell carcinoma: current status and future directions. *CA Cancer J Clin*. 2017;67(6):507–524. doi:10.3322/caac.21411.
- Khong HT, Restifo NP. Natural selection of tumor variants in the generation of “tumor escape” phenotypes. *Nat Immunol*. 2002;3(11):999–1005. doi:10.1038/ni1102-999.
- Wang J, Liu L, Qu Y, et al. HLA class I expression predicts prognosis and therapeutic benefits from tyrosine kinase inhibitors in metastatic renal-cell carcinoma patients. *Cancer Immunol Immunother*. 2018;67(1):79–87. doi:10.1007/s00262-017-2064-1.
- Wang Q, Zhang J, Tu H, et al. Soluble immune checkpoint-related proteins as predictors of tumor recurrence, survival, and T cell phenotypes in clear cell renal cell carcinoma patients. *J Immunother Cancer*. 2019;7(1):334. doi:10.1186/s40425-019-0810-y.