A promising prognostic risk model for advanced renal cell carcinoma (RCC) with immune-related genes

Peng Cao
Beijing Chaoyang Hospital, Capital Medical University
Jian-Dong Zhang
Beijing Chaoyang Hospital, Capital Medical University
Ze-Jia Sun
Beijing Chaoyang Hospital, Capital Medical University
Xiang Zheng
Beijing Chaoyang Hospital, Capital Medical University
Bao-Zhong Yu
Beijing Chaoyang Hospital, Capital Medical University
Hao-Yuan Cao
Beijing Chaoyang Hospital, Capital Medical University
Fei-Long Zhang
Beijing Chaoyang Hospital, Capital Medical University
Zi-Hao Gao
Beijing Chaoyang Hospital, Capital Medical University
Wei Wang (zico73@medmail.com.cn)
Beijing Chaoyang Hospital, Capital Medical University https://orcid.org/0000-0003-2642-3338

Research article

Keywords: renal cell carcinoma, RCC, immunologic signature, tumor immunity, prognosis, prognostic risk model, gene signatures

Posted Date: December 14th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-44822/v2

Abstract

Background

Renal cell carcinoma (RCC) is a common tumor of the urinary system. Nowadays, immunotherapy is a hot topic in treatment of solid tumors, especially those with pre-activated immune state.

Methods

In this study, we have downloaded genomic and clinical data of RCC samples from TCGA database. Four immune-related genetic signatures were used to predict the prognosis of RCC by Cox regression analysis. We have established a prognostic risk model. The model consists of the genes most related to prognosis from the four genes signatures and aims at prognosis of the RCC samples via Kaplan-Meier survival analysis. Independent data from the ICGC database were used to test the predictive stability of the model. Furthermore, we have performed landscape analysis to assess the presence of mutations in the genes of interest in the RCC samples from the TCGA. Finally, we have explored the correlation between the selected genes and the level of tumor immune infiltration via TIMER platform.

Results

We have used four genetic signatures to construct prognostic risk models and found that each of the models divide the RCC samples into high- and low-risk groups, each of the groups correlating with significantly different prognosis, especially in the advanced RCC cases. A comprehensive prognostic risk model was constructed with eight candidate genes from four signatures (HLA-B, HLA-A, HLA-DRA, IDO1, TAGAP, CIITA, PRF1 and CD8B) dividing the advanced RCC samples from the TCGA database into high-risk and low-risk groups with a significant difference in the overall survival. The stability of the model was verified by independent data from the ICGC database. The samples from different subgroups. Landscape analysis showed that the mutation ratios in some genes were different between two risk groups. Besides, the expression levels of the selected genes were interrelated with the infiltration degree of the immune cells in the advanced RCC.

Conclusions

Eight immune-related genes were screened in our study to construct a promising prognostic risk model with a great predictive value for the prognosis of advanced RCC. The selected genes were associated with infiltrating immune cells in tumors which presents a chance for personalized treatment for advanced RCC.

1. Background

Renal cell carcinoma (RCC) is the 14th most common cancer accounting for 2.2% of all cancers worldwide. 403262 new cases have been reported in 2018 with a ratio of males to females being estimated as 1.5:1 [1]. RCC is not a single disease, but rather comprises several different types of tumors

that arise from the renal epithelium [2]. It can be divided roughly into clear cell RCC (ccRCC) and non-clear-cell RCC (nccRCC). ccRCC is the most common subtype of RCC and accounts for > 80%. Papillary RCC (pRCC) and chromophobe cell renal carcinoma (ccRC) are the most represented subtypes in nccRCC accounting for 10-15% and 4-5% of RCC, respectively [3]. Despite the diagnosis and the improved treatment of RCC, its overall survival remains low. The problem may be attributed to the lack of specific treatments for different subtypes of RCC and tumor progression. Most of the available treatments are focused on ccRCC. Additionally, the same methods are used for nccRCC due to lack of effective treatments for the disease. The systemic therapies are the most recommended therapies for localized and advanced RCC. Surgery is the first choice for localized RCC. However, the increase in the tumor stage makes prognosis worse. Data indicate that the age-standardized 5-year relative survival of RCC patients decreases with the increase in the clinical stage [4]. Cytoreductive nephrectomy (CN) can be performed for the advanced RCC and need to be supplemented with adjuvant therapy whether in the cytokine or targeted therapy eras [5, 6]. It should be noted that the advanced RCC patients with poor physical condition could not benefit from CN [7]. Therefore, other treatments, such as embolization, targeted therapy and immunotherapy as a supplement or alternative for surgery provide new visions for successful treatment and better prognosis of RCC [3, 8].

RCC is a malignant tumor which is insensitive to traditional radiotherapy and chemotherapy. It has strong immunogenicity and is considered as a hot tumor in which a large number of B cells, T cells, macrophages and other immune cells infiltrate the tumor tissue. Therefore, immunotherapy is a good choice for its treatment. At present, immunotherapy has leapt to the forefront of cancer research. Endless new immunotherapy drugs have been approved for a variety of solid tumors. In particular, the overall therapeutic effect of patients with advanced and metastatic RCC has improved in recent years [9, 10]. With the development of RCC genomic research and the new progress about the mechanism of the immune response to cancers, the immunotherapy of RCC has shifted from non-specific immunotherapy (cytokine therapy) to new types of immunotherapy (immune checkpoint inhibitor, combined immunotherapy), which opens a new era of immunotherapy for RCC. For example, PD1/PD-L1, CTLA-4 and other immune checkpoints, which are negative costimulatory molecular control signals, inhibit the activation and function of T cells and promote tumor immune escape and self-proliferation [11]. Immune checkpoint inhibitors block the immunomodulatory effect of these inhibitory immune checkpoints and indirectly strengthen the anti-tumor immune response and improve the therapeutic effect. However, the incidence of immune-related adverse events (irAEs) in patients receiving immunosuppressive therapy was very high, up to more than 70% [12]. Therefore, it is highly prerequisite to discovery biomarkers which better evaluate RCC prognosis and are correlated with the immune cells infiltrating in the tumor. This will help providing potential targets for immunotherapy of RCC and thus to a therapeutic effect.

Here, we have downloaded 758 different pathological types of RCC samples from the TCGA database and used four reported immune-related genetic signatures to evaluate RCC prognosis. We have selected eight candidate genes (*HLA-B*, *HLA-A*, *HLA-DRA*, *IDO1*, *TAGAP*, *CIITA*, *PRF1* and *CD8B*) from each signature and combined them to construct a comprehensive prognostic risk model that divides the advanced

kidney cancer into high- and low-risk groups. We have detected that the overall survival (OS) of the high-risk group was significantly lower than that of the low-risk group. After the verification of the model, we found that the mutation ratios of some genes were different among the two types of risk groups and that there were some correlations between the expression of the selected genes and the degree of the tumor-infiltrating immune cells.

2. Methods

2.1 Data acquisition

Patients' clinical information and mRNA expression profiles from The Cancer Genome Atlas (TCGA) database of three main pathologic types of RCC, i.e. kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP) and kidney chromophobe (KICH) were downloaded from the UCSC Xena. The R packet TCGA biolinks were used to obtain the genetic mutation information. 91 RCC samples were used for validating the prognostic risk model and they were taken from the International Cancer Genome Consortium (ICGC) database (Table 1). Genes used for the analysis in the present study were from four immune-related genetic signatures and were as follows: *HLA-A* and *HLA-B* in HLA class I molecules, IFN gamma signature, expanded immune gene signature and cytotoxic T lymphocyte (CTL) level signature. The corresponding genes in the signatures were reported to be closely related to the clinical outcomes and prognosis of solid tumors (Table 2).

Table 1. Data of the three main subtypes of RCC: the kidney renal clear cell carcinoma (KIRC), the kidney renal papillary cell carcinoma (KIRP) and the kidney chromophobe (KICH) from the TCGA and ICGC databases.

Data	Number of RCC sample
KICH	62
KIRC	429
KIRP	267
RCC from ICGC	91

Table 2. The immune-related signatures used in the study, including HLA-A and -B, IFN gamma signature, the expanded immune gene signature, the cytotoxic T lymphocyte (CTL) signature and their corresponding genes.

Signature	Gene
HLA class I molecules	HLA-A, HLA-B
IFN gamma signature	IDO1, CXCL10, CXCL9, HLA-DRA, IFNG
Expanded immune gene signature	CD30(TNFRSF8), IDO1, CIITA, CD3E, CCL5, GZMK, CD2, HLA-DRA, CXCL13, NKG7, HLA-E, CXCR6, LAG3, TAGAP, CXCL10, STAT1, GZMB
Cytotoxic T lymphocyte (CTL) level signature	CD8A, CD8B, GZMA, GZMB, PRF1

2.2 Survival analysis via univariate COX regression analysis

The survival time and survival status of the patients with RCC were extracted from the TCGA database. The samples with incomplete clinical data were removed. Taken together, a total number of 730 samples with complete prognostic outcome were selected. According to the clinical stage of the tumor, recommended by the American Joint Committee on Cancer (AJCC) [13], all samples were divided into two groups that comprised four stages. The first group comprised stages I and II and was designated as an early stage RCC group while the second group was the advanced RCC group and included RCC in stages III and IV. We have used the coxph function in the R package to conduct the univariate COX regression analysis and to explore the association between the corresponding genes in each immune-related genetic signature and the disease-free survival (DFS) and overall survival (OS) rates of the two groups of RCC samples.

2.3 Establishment and verification of the prognostic model

Genes in the signatures for the multiple COX regression analysis constructed four prognostic risk models for the early and advanced RCC, respectively. The two genes most related to the prognosis in the advanced RCC group were selected from each of the signatures. The selected eight genes were used for the multiple COX regression analysis to construct a new comprehensive prognostic risk model. Then the surv_cutpoint function in the survminer R package was applied to determine the best threshold point to distinguish between the low-risk and high-risk RCC. The Kaplan-Meier (KM) survival analysis was used to evaluate the predictive ability of the prognostic model. We then built receiver operating characteristic curves (ROC) to evaluate the specificity and sensitivity of the model *via* the survival ROC in the R package. The prognosis risk model was then tested by independent data obtained from the ICGC database. Last but not the least, we have applied the model to different clinical subtypes such as age, gender, clinical stage and pathological pattern to assess its stability.

2.4 Mutation analysis in high and low-risk groups

The R package maftools have been used to calculate the gene mutations for each patient with RCC genetic data from the TCGA database. We have screened sixteen genes in the low- and high-risk samples, respectively, according to the mutation ratio, and then have built a waterfall map to show the distribution of the mutations in the genes in the two groups of the RCC samples.

2.5 Association among the tumor-infiltrating immune cells and selected genes

The Tumor Immune Estimation Resource, TIMER [14] is a comprehensive database to study systematically the tumor-infiltrating immune cells in various malignancies. The web contains a large number of different cancer samples in the TCGA database. We have investigated the association of six types of immune infiltrating cells (B cells, CD8 + T cells, CD4 + T cells, macrophages, neutrophils, and dendritic cells) with eight selected genes to evaluate the immune status of the tumor in the low and high-risk groups *via* the TIMER platform.

2.6 Statistical analysis

All statistical analysis was carried out in the R studio (version 3.6.2). The R package survival and survminer were used to perform the univariate/multivariate Cox analysis, the Kaplan-Meier survival analysis and to determine the optimal threshold points to classify the low-risk and high-risk renal cancer patients. The R package maftools, which was used for statistical analysis of gene mutations in the renal cancer patients, was used for the analysis in Figure 4. The R package pheatmap was used for drawing the heatmap in the analysis. The R package survival ROC was used to draw the ROC curves. P-values < .05 were considered statistically significant.

3. Results

3.1 Association of the genes in the described immune-related signatures with the disease-free and overall survival in RCC

In our study, we have downloaded the data of 758 RCC samples from the TCGA database and independent data of 91 RCC samples from the ICGC database. The analysis of the correlation between the expression levels of the immune-related genes and the prognosis of RCC allowed us to select genes derived from the IFN-gamma signature, the expanded immune gene signature, the cytotoxic T lymphocyte (CTL) signature and the HLA-A and HLA-B in HLA I molecules. These immune-related signatures were reported to be related to the prognosis of solid tumors, such as melanoma, ovarian cancer, breast cancer[15-19].

The univariate COX regression analysis was used to correlate gene expression levels with DFS and the OS of RCC. First, according to the clinical stage, we have divided the samples into two groups: an early-stage group that comprised the RCC in stages I and II and an advanced stage group containing RCC in stages III and IV. After excluding the invalid samples, 499 early RCC and 231 advanced RCC samples were further analyzed. In the two groups of RCC subsets, we found that few of the immune-related genes were

significantly associated with the DFS and OS of RCC patients. For the early-stage RCC, we found that the high expression levels of *CXCL13* and *STAT1* resulted in poor DFS while the high expression levels of *IDO1*, *CXCL13* and *GZMB* were related to detrimental OS. For the advanced RCC, the high expression levels of *TNFRSF8* and *CXCL13* were shown to be good predictors of adverse DFS and OS, respectively (Supplementary Figure 1).

3.2 Construction of prognostic risk models for RCC based on the genes from each of the immune-related signatures

Genes from the four gene signatures were studied to perform a multiple COX regression analysis in the early and advanced RCC groups, to construct prognostic models for the OS of RCC and to evaluate the performance of each model in the two groups of samples. The model was used to calculate the risk score of each sample. It determined the division threshold according to the surv_cutpoint function, dividing the samples into high-risk and low-risk groups. The KM survival analysis according to the high and low-risk groups of the samples was further conducted. All four models allowed discrimination of the RCC samples into high and low-risk groups. The OS was worse in the high-risk RCC group than in the low-risk one. Contrary to the early-stage RCC, the survival curves for the four models of the immune-related signatures indicated more significant differences in the OS between the two groups from the advanced RCC. The differences were statistically significant: for the HLA-A and HLA-B the p-value was 0.0015, for the IFN-gamma signature the calculated p-value was 9.787e-6, whereas for the expanded immune gene signature the p-value was1.137e-11, as for the cytotoxic T cell lymphocyte signature the p-value was 0.00011 (Figure 1).

3.3Establishment and examination of the prognostic model with the selected genes for the advanced RCC

The four risk models constructed by using the four immune-related signatures in the advanced RCC samples divided the samples into high and low-risk groups with significant statistical differences in the overall survival rate. We have used 8 genes that were most likely to be associated with OS in the advanced RCC samples. These genes were *HLA-B*; *HLA-A*; *HLA-DRA*; *IDO1*; *TAGAP*; *CIITA* and *PRF1*. They were combined to make a multiple COX regression analysis and a comprehensive prognostic risk model reflecting the genes' weight coefficient (see Supplementary Table 1). The advanced RCC samples were divided into high- and low-risk groups according to the risk score of each sample (Supplementary Table 2). The estimated division threshold was with a cutoff of -2.20465 (Figure 2b). The OS in the high-risk group was lower than that in the low-risk group and there were significant differences in the overall survival between the two groups with a p-value equal to 0.032 (Figure 2a). The ROC curves suggested that though the predictive value of risk score for prognosis of advance RCC was not high enough (area under the curve (AUC) = 0.64). however, the stage of the RCC combined with risk score could increase the predictive value for the prognosis (AUC = 0.77) as seen in Figure 2e.

We have further used 91 of RCC samples from the ICGC database to test the comprehensive prognostic risk model. The division threshold determined by the above method was -2.622015 (Figure 4b). The samples were divided into high and low-risk groups (Supplementary Table 3) according to the threshold.

The results showed that the OS in the high-risk group was significantly lower than that in the low-risk group (p-value = 0.013, Figure 4a). The prediction result of the model was consistent with the previous results (Figure 3 and Figure 4), and the stability of the model was effectively verified.

3.4 Landscape analysis of gene mutation in the high- and low-risk advanced RCC groups based on the TCGA database

Among the advanced RCC samples in the TCGA database, the genes with the top ten mutation rates in the high-risk group included *TTN*, *MUC4*, *PBRM1, VHL, CHECK2, ATRX, DNAM2, FAT1, FRG1B, KMT2C* (Figure 5a), while in the low-risk group these genes were *PBRM1*, *VHL, TTN, SETD2, MUC4, BAP1, MUC16, MT-CYB, MUC2, CSMD3* (Figure 5b). The distribution and annotation of mutations of the top sixteen mutant genes in the two groups of samples are shown in Figure5. The frequencies of the mutant genes, such as *VHL*, *CHEK2 and ATRX,* were different between the high- and low-risk groups. Among them, the frequency of *ATRX* in the high-risk group was significantly higher than that in the low-risk group (p-value = 0.0455).

3.5 Stability assessment of the prognostic risk model

The stability of model risk score in the different RCC clinical characteristic subgroups of the TCGA database was evaluated. There were significant differences between the high and low-risk groups according to the age, gender, clinical stage and pathological pattern (Figure 5a-d). Moreover, it was indicated that the high-risk groups in all subgroups led to adverse prognosis. This showed that the comprehensive prognostic model constructed by the eight selected genes had a good stability.

3.6 Association of the genes involved in the model with the tumor immune infiltrates

The Tumor Immune Assessment Resource (TIMER) platform was used to download the immune score (Supplementary Table 4) of the advanced RCC samples. Then we have explored the relationship between the expression of *HLA-B*, *HLA-A*, *HLA-DRA*, *IDO1*, *TAGAP, CIITA*, *PRF1* and *CD8B* at the transcriptional level and the tumor-infiltrating immune cell populations (B cells, CD8+ T cells, CD4+, T cells, neutrophils and dendritic cells). We found that the expression levels of *PRF1*, *CIITA*, *TAGAP* and *HLA-DRA* were positively correlated with infiltrates of six types of immune cells in the tumors. Additionally, higher infiltration levels of CD8+ T cells, neutrophils and myeloid dendritic cells were significantly correlated with higher expression of the eight selected genes, respectively (Figure 6).

4. Discussion

The renal parenchyma malignant tumors originate from the renal tubular epithelial cells. The clear cell RCCs, the papillary and the chromophobe and are the three main subtypes of renal cell carcinomas. In recent years, with the development of genetic and high-throughput sequencing technology, data on RCC pathogenesis, prognosis and treatment have accumulated, thus making a considerable progress [20]. RCC is regarded as a tumor in a pre-activated immune state and is believed

to have a better response to immunotherapy. Our research is of importance to the research field as the research in the tumor immune-related biomarkers holds the potential to discover potential targets for the diagnosis and treatment of RCC which can serve as predictors for the disease progression. The last is prerequisite for the quality of life improvement and long-term survival of patients with RCC.

In the present article, we discuss the role of four previously described immune-related gene signatures [21, 22], namely the IFN gamma, the expanded immune gene, the CTL signature, and the HLA-A and HLA-B molecules, in the prognosis of the early stage RCC (stages I and II) and the advanced stage group (stages III and IV). We have found that each of the four signatures established a prediction model dividing the RCC samples into high- and low-risk groups. Especially in the advanced RCC samples, the high-risk group had significantly worse OS than the low-risk group. Thereafter, we chose eight genes, *HLA-B*, *HLA-A*, *HLA-DRA*, *IDO1*, *TAGAP*, *CIITA*, *PRF1* and *CD8B*, from the four signatures which were most likely to be related to OS in the advanced RCC. These genes were combined to construct a comprehensive prognostic risk model to assess the OS of the advanced RCC. It similarly implied an unfavorable OS prognosis in the high-risk group of the advanced RCC. As a next step, we have used 91 RCC samples from the ICGC database to verify the stability of the model. The eight selected immune-related genes in this combination played pivotal roles in different biological processes of the tumor growth such as proliferation, apoptosis, metastasis and metabolism. There were many subtypes in the human leukocyte antigen (HLA) system participating in the human immune response. According to the structure and function, the HLA genes are divided into two classes: class I and II. Goebel et al. found that the frequency of HLA subtypes impacts RCC development [23]. It was indicated that the high expression of HLA-A and HLA-B which belong to class I in the ccRCC showed better prognosis than those with low expression [24]. HLA-DRA is one of the HLA class II alpha chain paralogues [23]. And Class II transactivator (CIITA) is one of the HLA class II regulatory genes playing a role in inducing the expression of other immune system genes. Butler and Blanck suggested that the expression of the two HLA class II molecules had a high-level correlation with pRCC [25]. Indoleamine 2,3-Dioxygenase 1 (IDO1) is a tryptophan catabolic enzyme that modifies inflammation and promotes cancer. IDO inhibitors can be used as immune-metabolic adjuvants which safely and potently facilitate the efficacy of immunotherapy [26]. T-cell activation Rho GTPase-activating protein (TAGAP) is a GAP-domain containing protein and was found to exert a role in T-cell differentiation [27]. ZHAO et al. reported that the expression level of TAGAP was related to the positive number of lymph nodes in prostate cancer [27]. Perforin 1 (PRF1) encodes a protein with structural similarities to complement component C9 that is important in immunity. The protein can form membrane pores releasing granzymes, thus leading to the cytolysis of the target cells [28]. CD8B and CD8A are heterodimers of CD8 (a glycoprotein) expressed only on those cytotoxic T cells to regulate maturation of T cells. Lee found that CD8B gene expression was closely correlated with tumor-infiltrating lymphocytes (TILs) in breast cancer [29]. The above-discussed data support the novel prognostic risk model composed of eight selected immune-related genes. The model was proved to divide the advanced RCC samples into low and high-risk groups with a significant difference in the OS based on the cutoff value. The question whether the cutoff value could become another evaluation index to improve the efficacy of IMDC risk classification needs further verification.

We have constructed an integrative prediction model with eight of the discussed genes able to clearly and precisely divide the advanced RCC samples into high-risk and low-risk groups. Notably, the predictive stability of the model was verified not only by analysis of external data from ICGC database but also in the subtypes of the samples, including age, gender and clinical stage and pathological pattern. Afterwards, we analyzed the gene mutations in the high- and low-risk groups of the advanced kidney cancer and discovered the mutation of some genes, such as *VHL*, *CHEK2, BAP1, PBRM1,* which were closely related to RCC[30-32]. Both ccRCC and nccRCC are heterogeneous cancers with different histologic, molecular, and genetic alterations. The most common subtype ccRCC is strongly associated with the mutation of von Hippel-Lindau (VHL). Targeted therapies for VHL and its related molecular targets, such as vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF) and mammalian target of rapamycin (mTOR), are currently extensively studied [33]. Though most of the genetic mutations in ccRCC can be found in nccRCC, we can still detect some characteristic differences in the subtypes of nccRCC. It was reported that the abnormal mutations of the MET proto-oncogene and the gene encoding the fumarate hydratase occurred in Type I and Type II pRCC, respectively [34]. The mutations of TP53 and PTEN could be detected in chRCC [35]. These mutant genes may be potential therapeutic targets for nccRCC. Notably, we found that the mutant ratio of ATRX was significantly higher in the high-risk group than that in the low-risk group. The ATRX protein is a chromatin remodeling factor functioning as a transcriptional regulator [36]. The mutations in ATRX was found in various cancers [37]. We hypothesized that the different mutation frequency of the genes may result in different prognosis in high-risk and low-risk groups.

Interestingly, the eight selected genes were not only related to immune response, but are also in connection with other tumors. It suggested that the genes involved in the immune activation might affect the development of RCC. Therefore, we have further investigated the connection between the composing genes and the tumor-infiltrating cells *via* TIMER platform. It is implied that the high expression of the genes favors the immune cells to infiltrate into the tumors. For example, the levels of $CD8^+$ T cells, neutrophils and myeloid dendritic cells positively correlate with the expression levels of all selected genes. $CD8^+$ T cells, which are a subtype of the cytotoxic T lymphocytes, contribute a lot to the antitumor activity through releasing of tumor cytokines such as INF-γ, perforin and granzyme B [38]. In recent years, studies have confirmed that the tumor-related neutrophils can differentiate into neutrophil type 1 (N1) and neutrophil type 2 (N2) under the influence of the tumor microenvironment. For example, N1 induced by IFN-βfunctions as an anti-tumor neutrophil. In contrast, neutrophils are more likely to become tumor-promoting N2 when the TGF-β pathway is activated [39, 40]. Dendritic cells function as antigen-presenting cells and are necessary for the initiation and maintenance of an effective immune response against cancer cells [41]. We suspect that the close relationship between the reported eight genes and various tumor-infiltrating immune cells may be a reason for better prediction of the risk model for the advanced RCC development and progression.

Normally, the study faces some limitations. First, the TCGA database is short of clinical data about the time of patients enrolled or relapsed, the time free interval and the therapeutic regimen for the patients.

This can influence the OS of each patient. It may result in a decrease in the predictive power of the model. Second, the selection of genes in this study was from previously described immune-related genetic signatures, therefore we may have missed genes with a predictive role which are not included in the studied signatures. Moreover, it is an *in silico* analysis without any further experimental verification. Therefore, future independent prospective clinical studies to confirm the capacity of the comprehensive prognostic risk model and research on the mutant genes leading to changes in the molecular mechanism of kidney cancer are further needed.

5. Conclusion

In conclusion, the prognostic risk models composed of genes selected from four immune-related genetic signatures demonstrated the potential to predict the survival prognosis of patients with advanced RCC, and have certain reference values for the prognosis assessment of the disease. The close relationship between the genes and the tumor-infiltrating immune cells helps provide new directions for immunotherapy to suppress tumor immune escape and to develop a personalized therapeutic regimen for the high-risk group of advanced RCC.

Abbreviations

RCC: renal cell carcinoma; TCGA: The Cancer Genome Atlas; ICGC:International Cancer Genome Consortium; KM: Kaplan-Meier; TIMER: Tumor Immune Estimation Resource; OS: overall survival; DFS: disease-free survival; ccRCC; clear cell renal cell carcinoma; pRCC: papillary renal cell carcinoma; ccRC: chromophobe cell renal carcinoma; irAEs: immune-related adverse events; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; KICH: kidney chromophobe; CTL: cytotoxic T lymphocyte; AUC: area under the curve; N1: neutrophil type 1; N2: neutrophil type 2.

Declarations

Acknowledgements

The Genomic data and clinical information of this research are downloaded from the Cancer Genome Atlas database (TCGA), International Cancer Genome Consortium (ICGC), UCSC Xena. We are grateful for their source of data used in our research.

Authors' contributions

WW and PC conceived and designed the study. BZY, HYC and FLZ worked together to search the data. PC performed the data analysis, interpreted the results and drafted the manuscript. XZ and ZHG helped to collect references. JDZ and ZJS proposed construction revisions to the study. All authorship reviewed and agreed on the final version of the manuscript.

Funding

This study was supported by Capital Clinical Specialty and Application Research and Outcome Promotion Program (No. Z171100001017555) the National Natural Science Foundation of China (Grant No. 81771720). Neither of the funders had influence on the design of the study or collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

The data and information downloaded and analyzed during the present study are available in the UCSC Xena, http://xenabrowser.net/datapages/, International Cancer Genome Consortium, http://icgc.org/, and Tumor Immune Estimation Resource, http://timer.cistrome.org/.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare no conflicts of interest in this work.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA,Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018; 68: 394-424.
2. Diaz-Montero CM, Rini BI,Finke JH. The immunology of renal cell carcinoma. Nat Rev Nephrol. 2020; 16: 721-35.
3. Ljungberg B, Albiges L, Abu-Ghanem Y, Bensalah K, Dabestani S, Fernandez-Pello S, Giles RH, Hofmann F, Hora M, Kuczyk MA, Kuusk T, Lam TB, Marconi L, et al. European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update. Eur Urol. 2019; 75: 799-810.
4. Siegel RL, Miller KD,Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020; 70: 7-30.
5. Petrelli F, Coinu A, Vavassori I, Cabiddu M, Borgonovo K, Ghilardi M, Lonati V,Barni S. Cytoreductive Nephrectomy in Metastatic Renal Cell Carcinoma Treated With Targeted Therapies: A Systematic Review With a Meta-Analysis. Clin Genitourin Cancer. 2016; 14: 465-72.
6. Flanigan RC, Salmon SE, Blumenstein BA, Bearman SI, Roy V, McGrath PC, Caton JR, Jr., Munshi N,Crawford ED. Nephrectomy followed by interferon alfa-2b compared with interferon alfa-2b alone for metastatic renal-cell cancer. N Engl J Med. 2001; 345: 1655-9.
7. Culp SH, Tannir NM, Abel EJ, Margulis V, Tamboli P, Matin SF,Wood CG. Can we better select patients with metastatic renal cell carcinoma for cytoreductive nephrectomy? Cancer. 2010; 116: 3378-88.

8. Barata PC,Rini BI. Treatment of renal cell carcinoma: Current status and future directions. CA Cancer J Clin. 2017; 67: 507-24.
9. Motzer RJ, Rini BI, McDermott DF, Aren Frontera O, Hammers HJ, Carducci MA, Salman P, Escudier B, Beuselinck B, Amin A, Porta C, George S, Neiman V, et al. Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. Lancet Oncol. 2019; 20: 1370-85.
10. Rodriguez-Vida A, Hutson TE, Bellmunt J,Strijbos MH. New treatment options for metastatic renal cell carcinoma. ESMO Open. 2017; 2: e000185.
11. Wolchok JD,Chan TA. Cancer: Antitumour immunity gets a boost. Nature. 2014; 515: 496-8.
12. Michot JM, Bigenwald C, Champiat S, Collins M, Carbonnel F, Postel-Vinay S, Berdelou A, Varga A, Bahleda R, Hollebecque A, Massard C, Fuerea A, Ribrag V, et al. Immune-related adverse events with immune checkpoint blockade: a comprehensive review. Eur J Cancer. 2016; 54: 139-48.
13. Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW,Amin MB. Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers. Eur Urol. 2018; 73: 560-9.
14. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B,Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res. 2017; 77: e108-e10.
15. Rooney MS, Shukla SA, Wu CJ, Getz G,Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015; 160: 48-61.
16. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, Albright A, Cheng JD, Kang SP, Shankaran V, Piha-Paul SA, Yearley J, Seiwert TY, et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017; 127: 2930-40.
17. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018; 24: 1550-8.
18. Alcaraz-Sanabria A, Baliu-Pique M, Saiz-Ladera C, Rojas K, Manzano A, Marquina G, Casado A, Cimas FJ, Perez-Segura P, Pandiella A, Gyorffy B,Ocana A. Genomic Signatures of Immune Activation Predict Outcome in Advanced Stages of Ovarian Cancer and Basal-Like Breast Tumors. Front Oncol. 2019; 9: 1486.
19. Noblejas-Lopez MDM, Nieto-Jimenez C, Morcillo Garcia S, Perez-Pena J, Nuncia-Cantarero M, Andres-Pretel F, Galan-Moya EM, Amir E, Pandiella A, Gyorffy B,Ocana A. Expression of MHC class I, HLA-A and HLA-B identifies immune-activated breast tumors with favorable outcome. Oncoimmunology. 2019; 8: e1629780.
20. Linehan WM,Ricketts CJ. The Cancer Genome Atlas of renal cell carcinoma: findings and clinical implications. Nat Rev Urol. 2019; 16: 539-52.
21. Chowell D, Morris LGT, Grigg CM, Weber JK, Samstein RM, Makarov V, Kuo F, Kendall SM, Requena D, Riaz N, Greenbaum B, Carroll J, Garon E, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. Science. 2018; 359: 582-7.

22. Vitale I, Sistigu A, Manic G, Rudqvist NP, Trajanoski Z,Galluzzi L. Mutational and Antigenic Landscape in Tumor Progression and Cancer Immunotherapy. Trends Cell Biol. 2019; 29: 396-416.
23. Goebel S, Kehlen A, Bluemke K, Altermann W, Schlaf G, Fischer K, Fornara P, Wullich B, Wach S,Taubert H. Differences in the frequencies of HLA-class I and II alleles between German patients with renal cell carcinoma and healthy controls. Cancer Immunol Immunother. 2017; 66: 565-71.
24. Yuan J, Liu S, Yu Q, Lin Y, Bi Y, Wang Y,An R. Down-regulation of human leukocyte antigen class I (HLA-I) is associated with poor prognosis in patients with clear cell renal cell carcinoma. Acta Histochem. 2013; 115: 470-4.
25. Butler SN,Blanck G. Immunoscoring by correlating MHC class II and TCR expression: high level immune functions represented by the KIRP dataset of TCGA. Cell Tissue Res. 2016; 363: 491-6.
26. Prendergast GC, Mondal A, Dey S, Laury-Kleintop LD,Muller AJ. Inflammatory Reprogramming with IDO1 Inhibitors: Turning Immunologically Unresponsive 'Cold' Tumors 'Hot'. Trends Cancer. 2018; 4: 38-58.
27. Chen J, He R, Sun W, Gao R, Peng Q, Zhu L, Du Y, Ma X, Guo X, Zhang H, Tan C, Wang J, Zhang W, et al. TAGAP instructs Th17 differentiation by bridging Dectin activation to EPHB2 signaling in innate antifungal response. Nat Commun. 2020; 11: 1913.
28. Voskoboinik I, Whisstock JC,Trapani JA. Perforin and granzymes: function, dysfunction and human pathology. Nat Rev Immunol. 2015; 15: 388-400.
29. Lee HJ, Song IH, Park IA, Heo SH, Kim YA, Ahn JH,Gong G. Differential expression of major histocompatibility complex class I in subtypes of breast cancer is associated with estrogen receptor and interferon signaling. Oncotarget. 2016; 7: 30119-32.
30. Ricketts CJ, De Cubas AA, Fan H, Smith CC, Lang M, Reznik E, Bowlby R, Gibb EA, Akbani R, Beroukhim R, Bottaro DP, Choueiri TK, Gibbs RA, et al. The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. Cell Rep. 2018; 23: 3698.
31. Ye K, Wang J, Jayasinghe R, Lameijer EW, McMichael JF, Ning J, McLellan MD, Xie M, Cao S, Yellapantula V, Huang KL, Scott A, Foltz S, et al. Systematic discovery of complex insertions and deletions in human cancers. Nat Med. 2016; 22: 97-104.
32. Carlo MI, Mukherjee S, Mandelker D, Vijai J, Kemel Y, Zhang L, Knezevic A, Patil S, Ceyhan-Birsoy O, Huang KC, Redzematovic A, Coskey DT, Stewart C, et al. Prevalence of Germline Mutations in Cancer Susceptibility Genes in Patients With Advanced Renal Cell Carcinoma. JAMA Oncol. 2018; 4: 1228-35.
33. Roskoski R, Jr. Vascular endothelial growth factor (VEGF) and VEGF receptor inhibitors in the treatment of renal cell carcinomas. Pharmacol Res. 2017; 120: 116-32.
34. Maher ER. Hereditary renal cell carcinoma syndromes: diagnosis, surveillance and management. World J Urol. 2018; 36: 1891-8.
35. Joshi S, Tolkunov D, Aviv H, Hakimi AA, Yao M, Hsieh JJ, Ganesan S, Chan CS,White E. The Genomic Landscape of Renal Oncocytoma Identifies a Metabolic Barrier to Tumorigenesis. Cell Rep. 2015; 13: 1895-908.

36. Picketts DJ, Higgs DR, Bachoo S, Blake DJ, Quarrell OW,Gibbons RJ. ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. Hum Mol Genet. 1996; 5: 1899-907.
37. Watson LA, Goldberg H,Berube NG. Emerging roles of ATRX in cancer. Epigenomics. 2015; 7: 1365-78.
38. Farhood B, Najafi M,Mortezaee K. CD8(+) cytotoxic T lymphocytes in cancer immunotherapy: A review. J Cell Physiol. 2019; 234: 8509-21.
39. Fridlender ZG, Sun J, Kim S, Kapoor V, Cheng G, Ling L, Worthen GS,Albelda SM. Polarization of tumor-associated neutrophil phenotype by TGF-beta: "N1" versus "N2" TAN. Cancer Cell. 2009; 16: 183-94.
40. Kim J,Bae JS. Tumor-Associated Macrophages and Neutrophils in Tumor Microenvironment. Mediators Inflamm. 2016; 2016: 6058147.
41. Wculek SK, Cueto FJ, Mujal AM, Melero I, Krummel MF,Sancho D. Dendritic cells in cancer immunology and immunotherapy. Nat Rev Immunol. 2020; 20: 7-24.

Figures

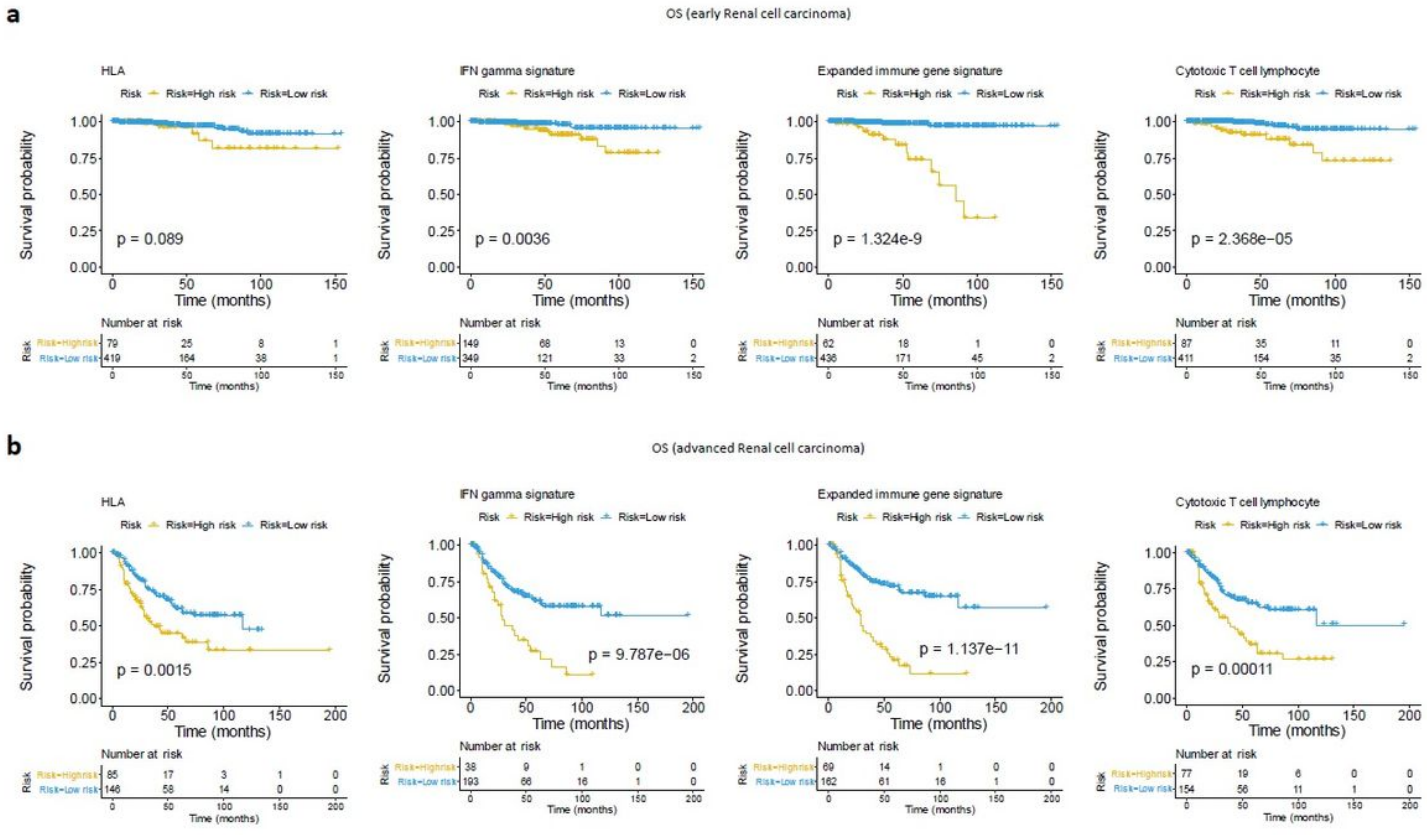

Figure 1

Prognostic risk models constructed by four immune-related signatures for overall survival (OS) in early and advanced RCC. (a) Classified efficiency of prognostic risk models constructed by four immune-related signatures (IFN-gamma signature, extended immune gene signature, cytotoxic T lymphocyte signature and HLA-A and HLA-B) in stage I + II RCC. (b) Classified efficiency of the four prognostic risk models for stage III + IV RCC. The p-value was shown in the survival plots.

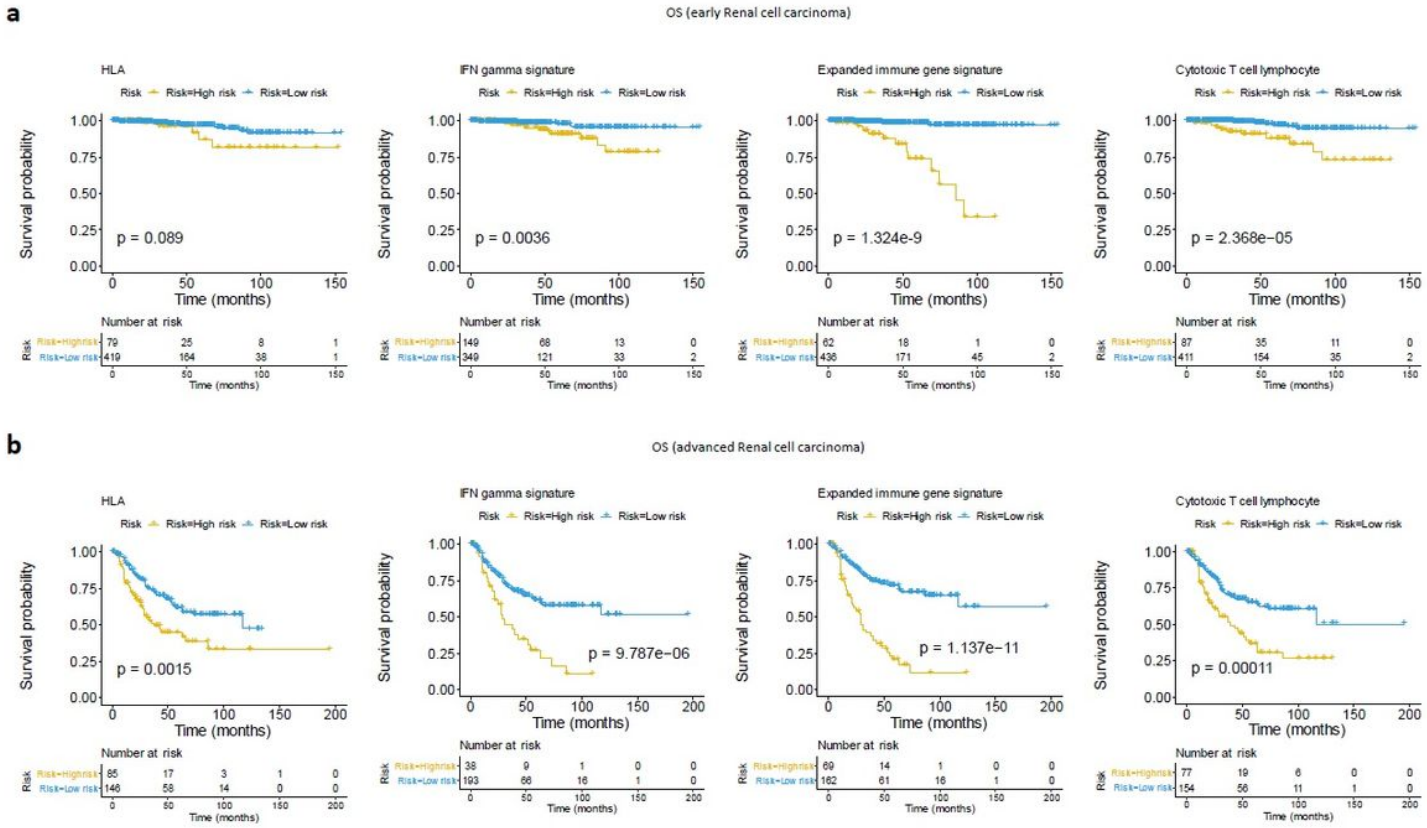

Figure 1

Prognostic risk models constructed by four immune-related signatures for overall survival (OS) in early and advanced RCC. (a) Classified efficiency of prognostic risk models constructed by four immune-related signatures (IFN-gamma signature, extended immune gene signature, cytotoxic T lymphocyte signature and HLA-A and HLA-B) in stage I + II RCC. (b) Classified efficiency of the four prognostic risk models for stage III + IV RCC. The p-value was shown in the survival plots.

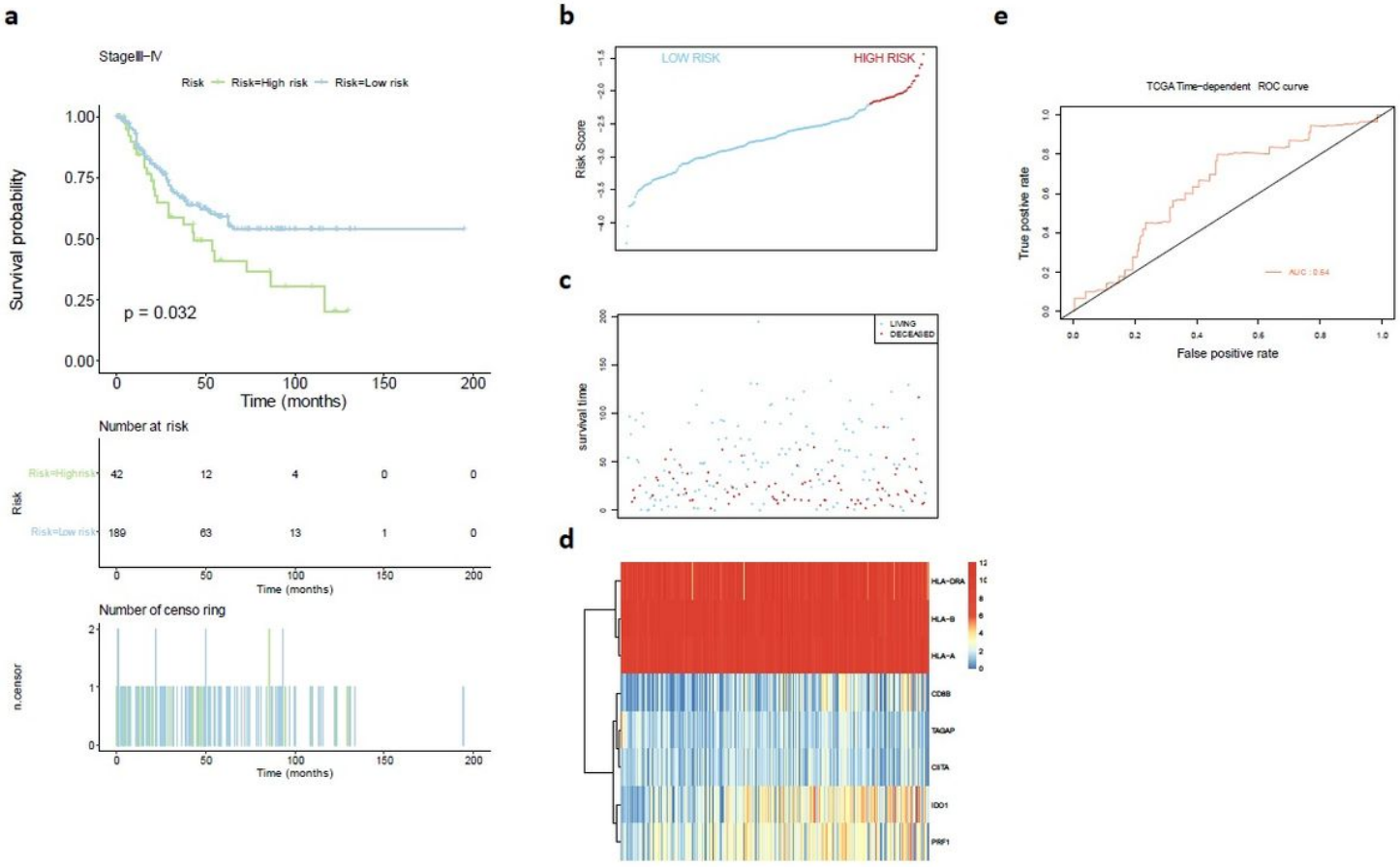

Figure 2

Prognostic risk models constructed by 8 genes combination for OS in advanced RCC. (a) Survival plots showed OS of high-risk group and low-risk group in advanced RCC. The risk score curve (b) and the scatter plot (c) were drawn according to risk score of every advanced RCC samples calculated by the model. (d) The heatmap indicated the expression levels of selected genes in the advanced RCC samples. High and low expressions were highlighted in red and blue respectively. (e) The predicted value of the model was assessed by Time-dependent ROC curve. The p-value was shown in the survival plot.

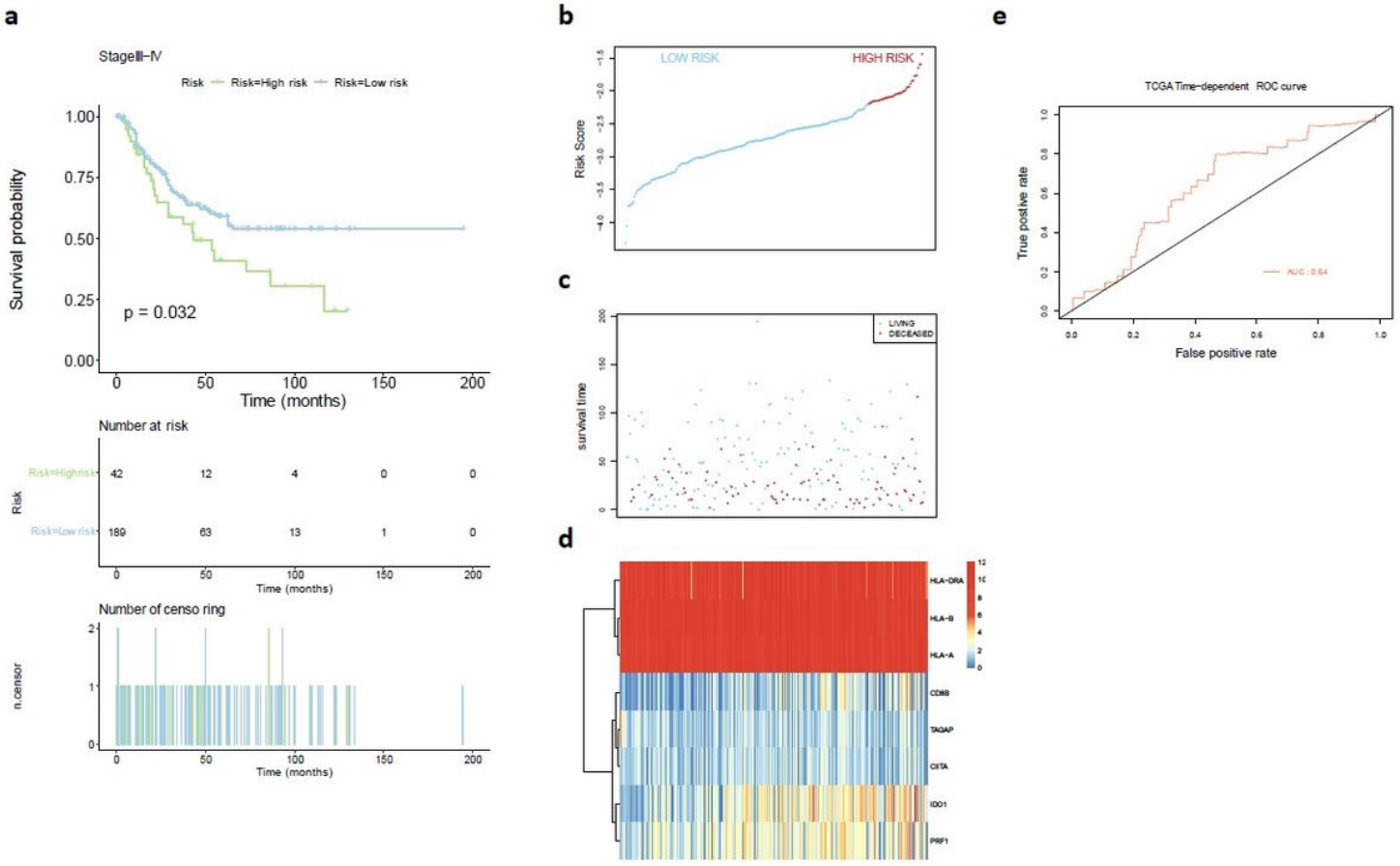

Figure 2

Prognostic risk models constructed by 8 genes combination for OS in advanced RCC. (a) Survival plots showed OS of high-risk group and low-risk group in advanced RCC. The risk score curve (b) and the scatter plot (c) were drawn according to risk score of every advanced RCC samples calculated by the model. (d) The heatmap indicated the expression levels of selected genes in the advanced RCC samples. High and low expressions were highlighted in red and blue respectively. (e) The predicted value of the model was assessed by Time-dependent ROC curve. The p-value was shown in the survival plot.

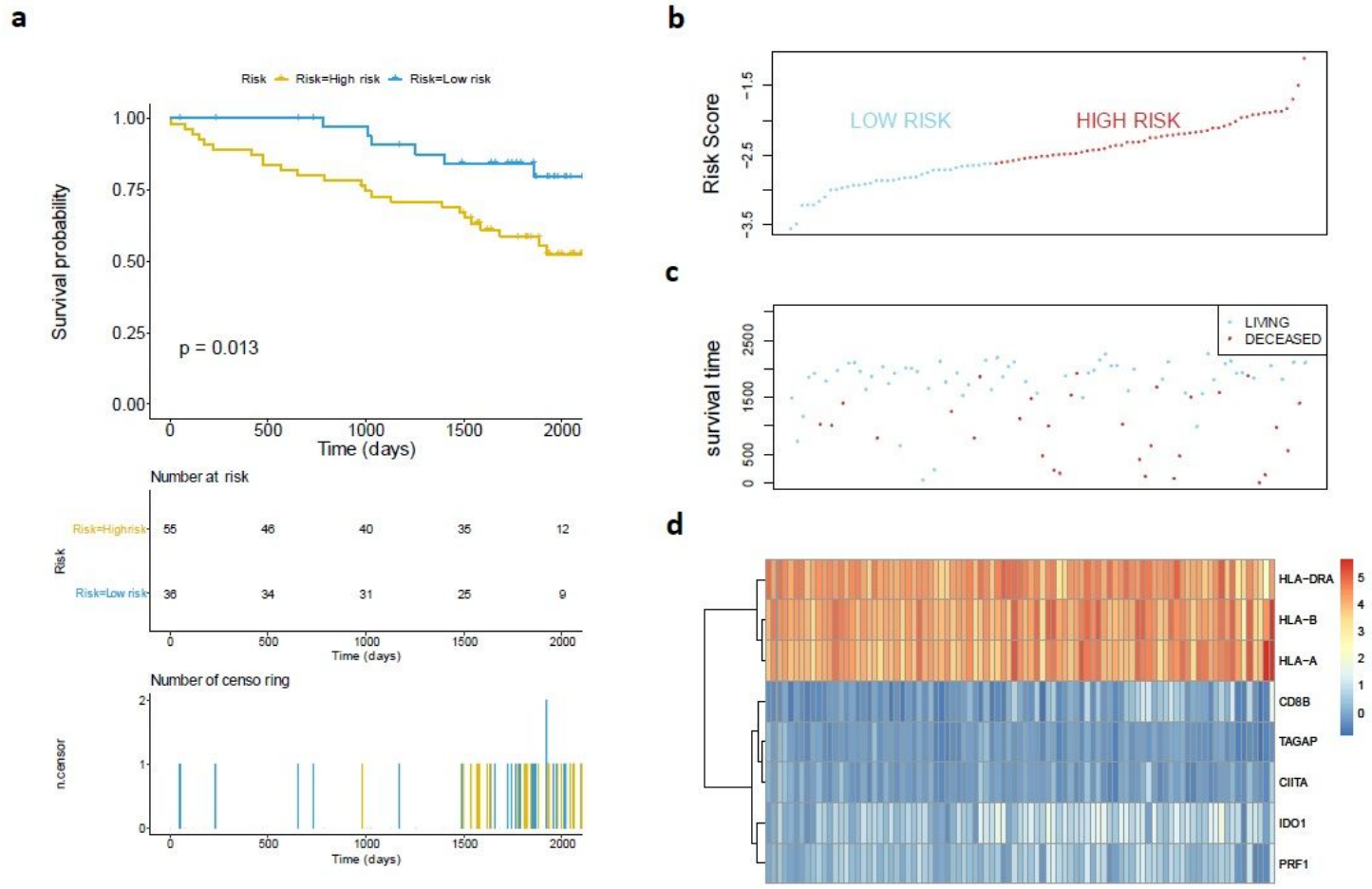

Figure 3

Validating the classified efficiency of the prognostic risk model constructed by 8 selected genes combination via data from ICGC. (a) Survival plots showed OS of high-risk group and low-risk group in advanced RCC from ICGC. The risk score curve (b) and the scatter plot (c) were drawn according to risk score of each RCC samples calculated by the model. (d) The heatmap indicated the expression levels of selected genes in the RCC samples. The p-value was shown in the survival plot.

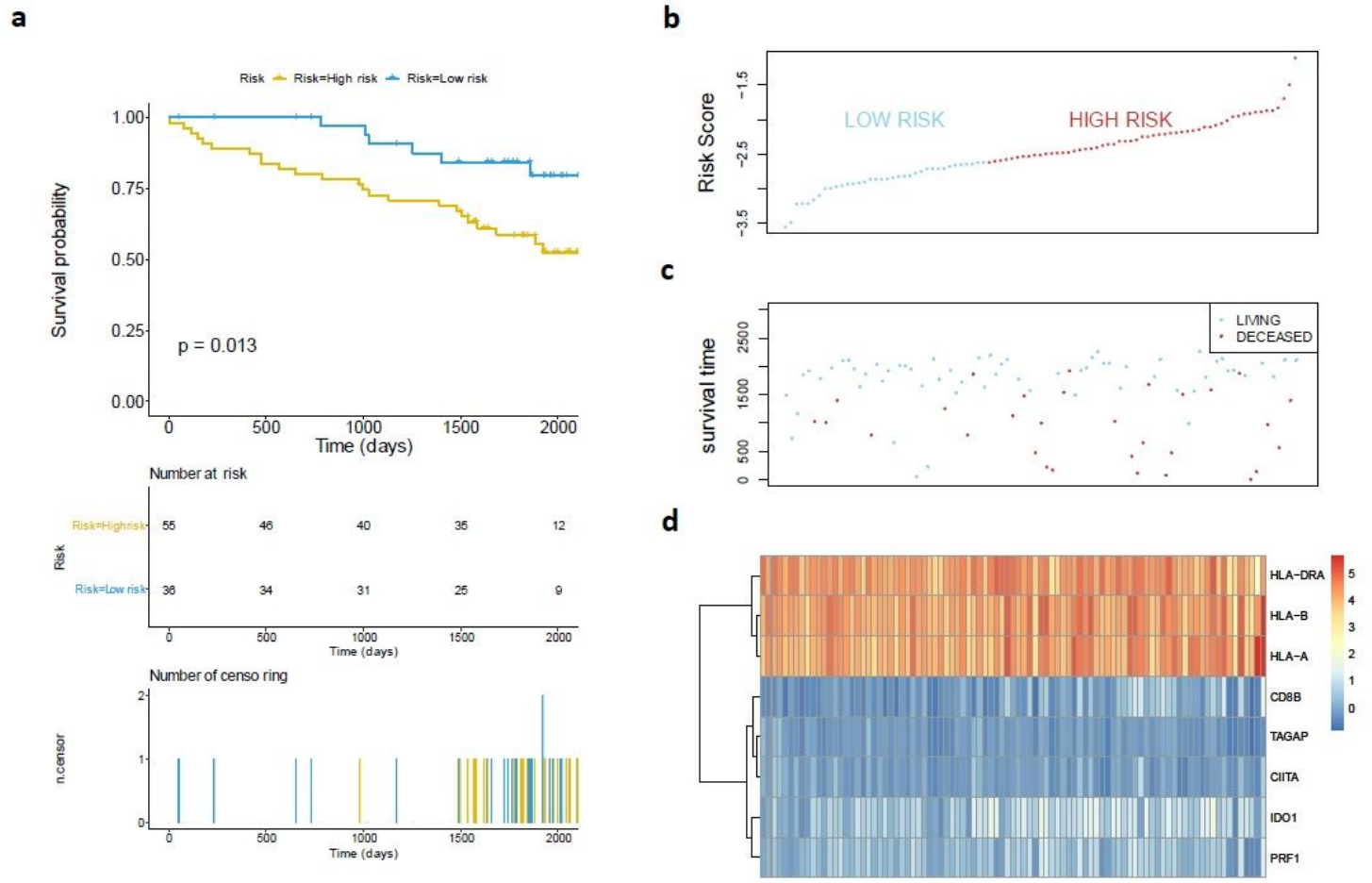

Figure 3

Validating the classified efficiency of the prognostic risk model constructed by 8 selected genes combination via data from ICGC. (a) Survival plots showed OS of high-risk group and low-risk group in advanced RCC from ICGC. The risk score curve (b) and the scatter plot (c) were drawn according to risk score of each RCC samples calculated by the model. (d) The heatmap indicated the expression levels of selected genes in the RCC samples. The p-value was shown in the survival plot.

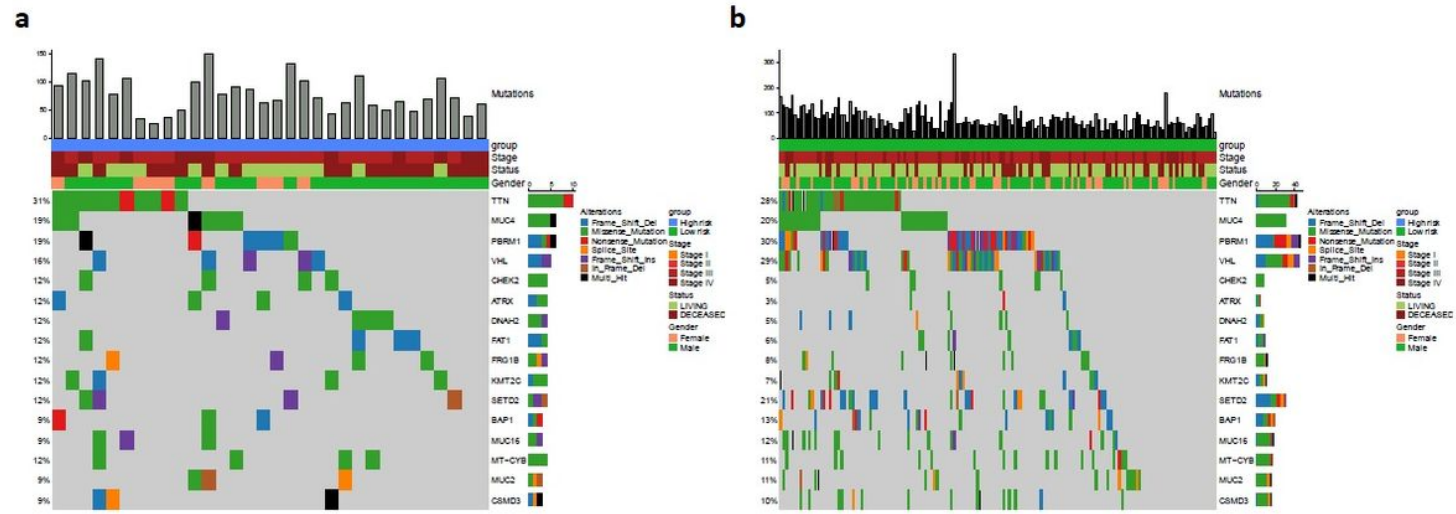

Figure 4

Frequencies and distribution of gene mutations in advanced RCC samples from high and low risk groups. The landscape analysis showed the top 16 genes with mutation frequency in high-risk group (a) and low-risk group (b) of the advanced RCC. The histogram showed the number of mutations in the RCC samples. Annotation information of the samples included risk groups, clinical stages, living status and genders. Different colors represented different mutation types.

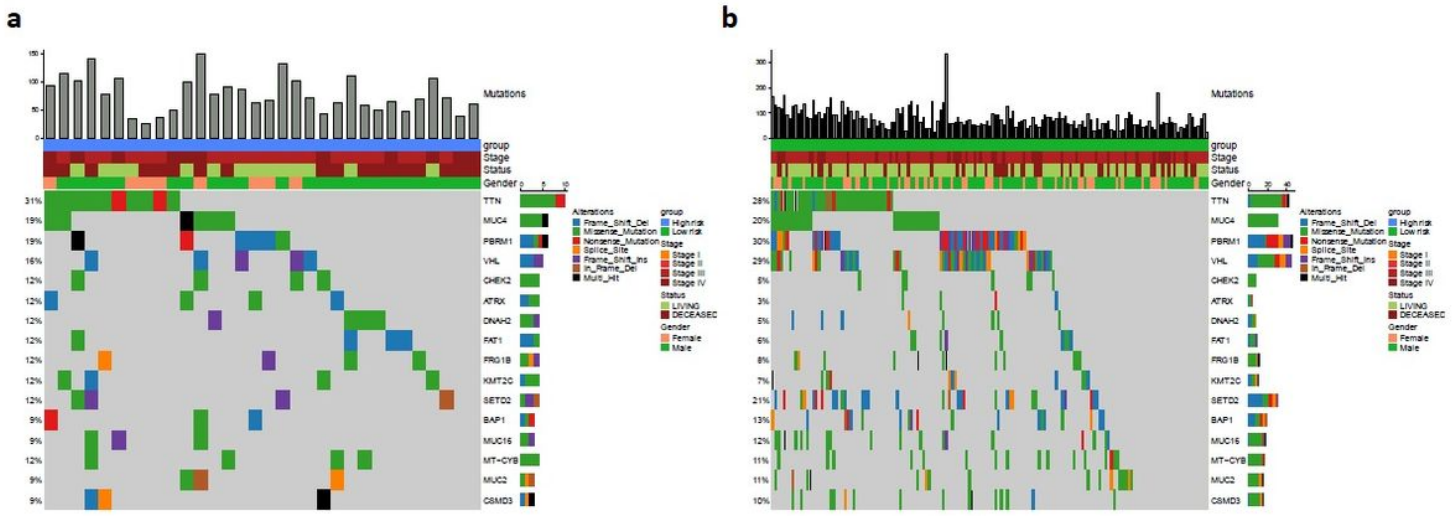

Figure 4

Frequencies and distribution of gene mutations in advanced RCC samples from high and low risk groups. The landscape analysis showed the top 16 genes with mutation frequency in high-risk group (a) and low-risk group (b) of the advanced RCC. The histogram showed the number of mutations in the RCC samples. Annotation information of the samples included risk groups, clinical stages, living status and genders. Different colors represented different mutation types.

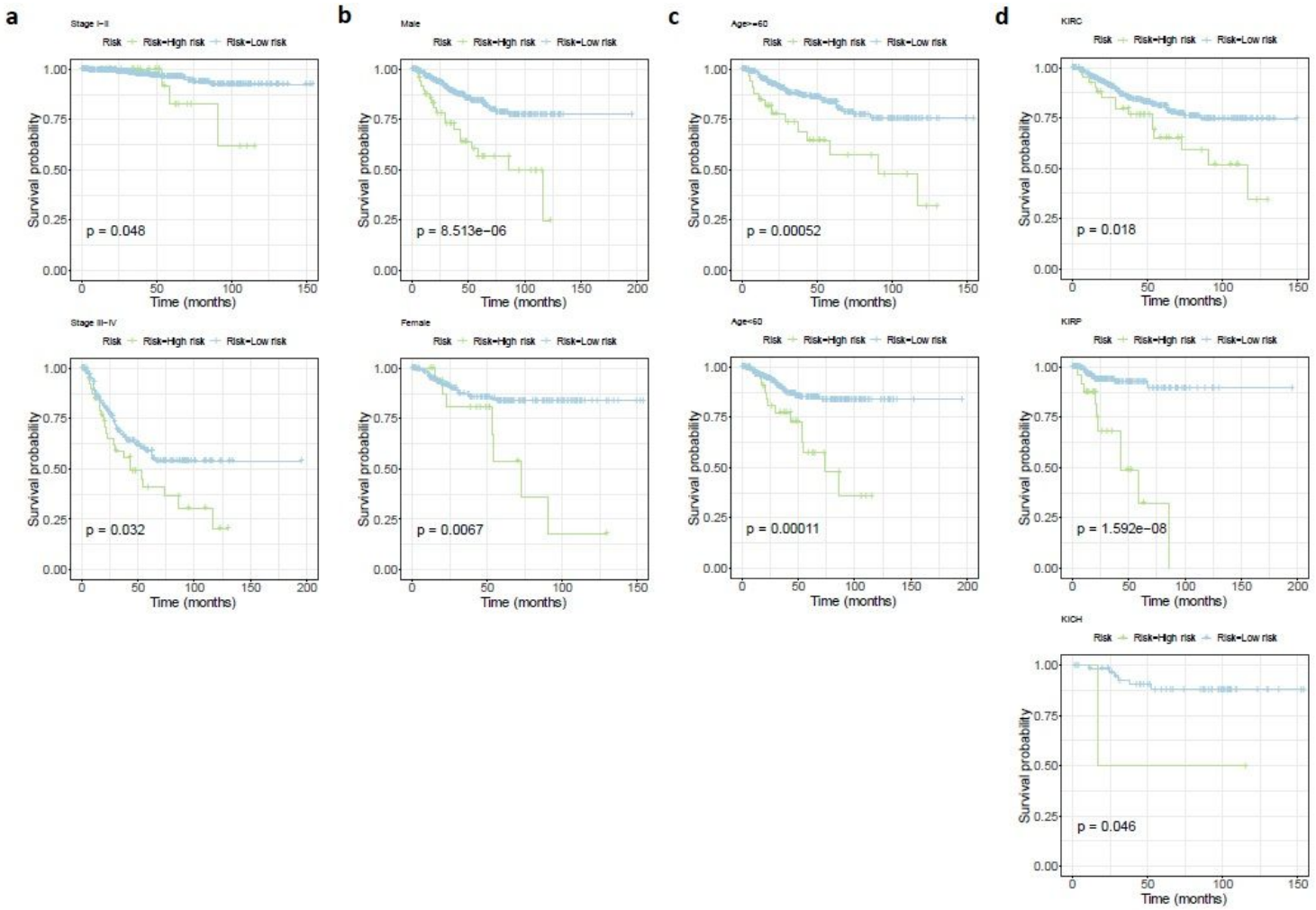

Figure 5

Validating the stability of the prognostic risk model constructed by the 8 selected genes for different subtypes of advanced RCC. Survival plots all showed that high-risk RCC classified by the model resulted in unfavorable OS in different stages (a); genders (b); ages (c) and pathological patterns (d). The p values were shown in the survival plots.

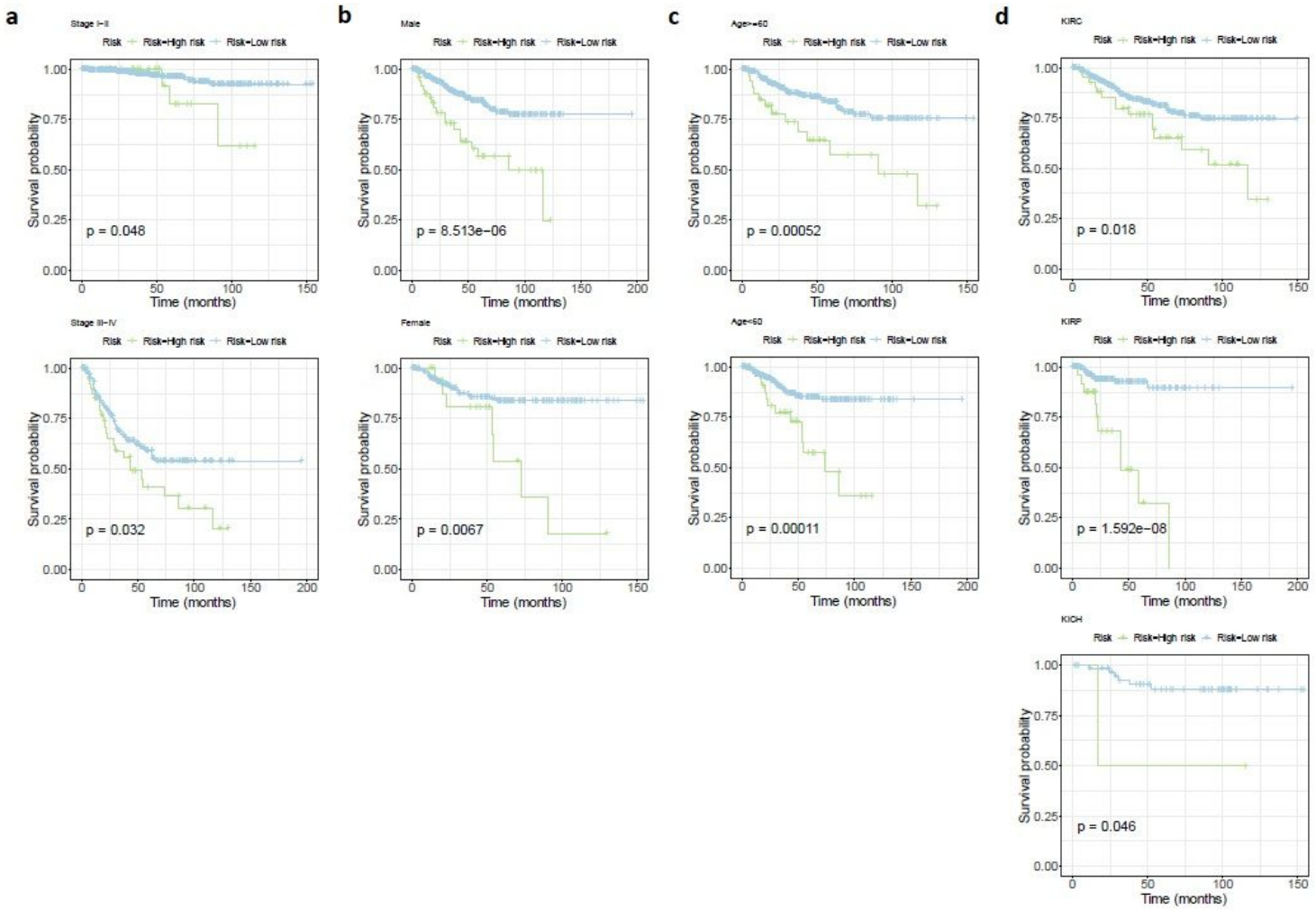

Figure 5

Validating the stability of the prognostic risk model constructed by the 8 selected genes for different subtypes of advanced RCC. Survival plots all showed that high-risk RCC classified by the model resulted in unfavorable OS in different stages (a); genders (b); ages (c) and pathological patterns (d). The p values were shown in the survival plots.

Figure 6

Association of the expression of the 8 selected genes with immune infiltrates in advanced RCC. Correlation analysis between the 8 genes expression (HLA-B, HLA-A, HLA-DRA, IDO1, TAGAP, CIITA, PRF1 and CD8B) and the level of tumor immune infiltrates (B-cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils and dendritic cells) via Tumor Immune Assessment Resource (TIMER) platform. The correlation indexes and p values were shown in the figure.

Figure 6

Association of the expression of the 8 selected genes with immune infiltrates in advanced RCC. Correlation analysis between the 8 genes expression (HLA-B, HLA-A, HLA-DRA, IDO1, TAGAP, CIITA, PRF1 and CD8B) and the level of tumor immune infiltrates (B-cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils and dendritic cells) via Tumor Immune Assessment Resource (TIMER) platform. The correlation indexes and p values were shown in the figure.

Supplementary Files

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

- supplementarytable1.docx
- supplementarytable1.docx
- supplementarytable2.docx
- supplementarytable2.docx
- supplementarytable3.docx
- supplementarytable3.docx
- supplementarytable4.docx
- supplementarytable4.docx
- Sfigure1.pdf
- Sfigure1.pdf