

Development and validation of an individualized immune-related gene pairs prognostic signature in papillary renal cell carcinoma

Xianghong Zhou

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

Shi Qiu

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

Di Jin

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

Kun Jin

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

Xiaonan Zheng

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

Lu Yang

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

Qiang Wei (✉ weiqiang933@126.com)

1.Department of Urology, Institute of Urology, National Clinical Research Center for Geriatrics and Center of Biomedical big data, West China Hospital of Sichuan University, Chengdu, Sichuan Province, China

<https://orcid.org/0000-0002-3380-4932>

Research

Keywords: papillary renal carcinoma, immune-related gene pairs, prognosis signature

Posted Date: May 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-29603/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Papillary renal carcinoma (PRCC) is one of the important subtypes of kidney cancer, with a high degree of heterogeneity. At present, there is still a lack of robust and accurate biomarkers for the diagnosis, prognosis and treatment selection of PRCC. Considering the important role of tumor immunity in PRCC, we aim to construct a signature based on immune-related gene pairs (IRGPs) to estimate the prognostic of patients with PRCC.

Methods: We obtained gene expression profiling and clinical information of patients with PRCC from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), which were divided into discovery and validation cohorts, respectively. The immune-related genes in the samples were used to construct gene pairs, and the immune-related genes pairs (IRGPs) with robust impact for overall survival (OS) were screened out to construct the signature by univariate analysis, multivariate Cox analysis, and least absolute shrinkage and selection operator (Lasso) analysis. Then we verified the prognostic role of the signature, and assessed the relationship between this signature with tumor immune infiltration and functional pathways.

Results: A total of 315 patients were included in our study, and divided to discovery (n=287) and validation (n=28) cohorts. Finally, we selected 14 IRGPs with a panel of 22 unique genes to construct the prognostic signature. According to the signature, we stratified patients into high-risk group and low-risk group. In both discovery and validation cohorts, the results of Kaplan-Meier analysis showed that there were significant differences in OS between the two groups ($p < 0.001$). Combined with multiple clinical pathological factors, the results of multivariate analyses confirmed that this signature was an independent predictor of OS (HR, 3.548; 95%CI, 2.096–6.006; $p < 0.001$). The results of immune infiltration analysis demonstrated that the abundance of multiple tumor-infiltration lymphocytes such as CD8⁺ T cells, Tregs, and T follicular cell helper were significantly higher in the high-risk group. Functional analysis showed that multiple immune-related signaling pathways were enriched in the high-risk group.

Conclusions: We successfully established an individualized prognostic immune-related gene pairs signature, which can accurately and independently predict the OS of patients with PRCC.

Introduction

Kidney cancer is one of the most common malignant tumors in the urinary system, and it is estimated that there will be 73,750 new cases diagnosed of kidney cancer in the United States in 2020 [1]. Papillary Renal Cell Carcinoma (PRCC) is a relatively rare histological subtype in kidney cancer, second to clear cell renal carcinoma (ccRCC), and accounts for about 10% -20% of kidney cancer [2]. PRCC is a heterogeneous disease, and the outcomes of different patients vary greatly in terms of disease progression, survival and response to therapy [3]. Currently, the common classification of PRCC is based on histology and mainly includes two main sub-types: PRCC type 1 and PRCC type 2 [3-5]. However, several previous studies have demonstrated that the classification has limited discrimination for the

clinical outcomes of PRCC [6-8]. Meanwhile, due to the relatively small number of PRCC cases, many clinical and molecular studies on kidney cancer have not included patients with PRCC. Thus, in order to provide more specific and accurate biomarkers for the diagnosis, treatment and prognosis of PRCC, we need more researches on molecular profiling of PRCC to provide reliable data.

In recent years, immunotherapy based on immune checkpoint inhibitors (ICIs) has been applied in various tumors, greatly improving the survival outcomes of patients with advanced tumors. These ICIs include multiple agents that target programmed cell death protein 1 (PD-1), programmed death ligand 1 (PD-L1) or cytotoxic T lymphocyte associated protein 4 (CTLA4). [9] For ccRCC, nivolumab, a PD-L1 inhibitor, has been approved for patients with metastatic tumors because of the encouraging results of random control trials. [10] In a subsequent retrospective study, the researchers analyzed the application of nivolumab in 16 PRCC patients and found that the response rate was over 30%, which was higher than that of ccRCC [11]. It can be seen that the successful immunotherapy on ccRCC may also change the treatment of PRCC. Meanwhile, some researchers have also reported the correlation of immune infiltration with immunotherapeutic response and clinical outcome in kidney cancer. Their findings demonstrated that the proportion of specific tumor-infiltration lymphocytes (TILs) such as tumor-associated macrophages and T CD8⁺ cells in the tumor immune microenvironment could independently predict the clinical outcome of ccRCC and PRCC patients.[12, 13] These studies have shown that tumor immunology characteristics are closely related to the prognosis and treatment response of patients with PRCC. Prognostic markers based on immunology may help the risk stratification and treatment selection of patients with PRCC. Currently, there is only one research team to identify and validate a 15 immune-related gene based risk signature for prognosis of PRCC [14]. Although the signature has shown a good discrimination, this study lacked external validation and did not analyze the impact of PRCC's traditional histological typing. More stable and reliable immune-related markers are still urgently needed.

To circumvent these defects, we utilized the transcriptome data from The Cancer Genome Atlas (TCGA) as the discovery cohort to construct an individualized immune-related gene pairs (IRGPs) prognostic signature for predicting overall survival (OS) of PRCC patients. The microarray data from Gene Expression Omnibus (GEO) as the external validation cohort set was used to verify the reliability of the signature.

Materials And Methods

Data collection and processing

This was a retrospective study based on two independent datasets. The first dataset was the discovery cohort including 287 samples from TCGA. The second dataset was the validation cohort including 28 samples from GEO. The gene expression quantification data and clinical data of these samples were obtained from TCGA (<https://portal.gdc.cancer.gov/>), GEO (<http://www.ncbi.nlm.nih.gov/geo/>, GSE2748) and related publication [15]. A total of 315 samples were included in our analyses. All the samples contained completed follow-up information.

For the data from TCGA, the transcriptome profiling was obtained by RNA-seq and measured by Fragments per Kilobase Million (FPKM) values and genes with zero of FPKM values in more than half of the samples were removed. As for the data from GEO, the transcriptome profiling was converted from the probe level to the corresponding gene symbol according to the annotation file.

Construction of Immune-related Gene Pairs Prognostic Signature

Firstly, we obtained information of 2,498 immune-related genes (IRGs) from the ImmPortdatabase (<https://import.niaid.nih.gov>). These IRGs were related to natural killer cell cytotoxicity, cytokines, cytokine receptors, antigen processing, T-cell receptor signaling pathway, B-cell antigen receptor signaling pathway and so on. We measured the IRGs involved in this study, and only IRGs with a median absolute deviation (MAD) greater than 0.5, that is, with a large degree of variation, were retained. Then, by comparing the expression levels of all IRGs in each sample in pairs, we obtained a series of immune-related Gene Pairs (IRGPs). When comparing two IRGs in an IRGP, if the expression level of the previous IRG was higher than the latter IRG, then the value of this IRGP was considered to be 1, otherwise it is 0. After removing IRGPs with small variation and imbalanced distribution, the remaining IRGPs were used as candidate signatures to predict the OS of PRCC.

Subsequently, log-rank test was used to preliminarily assess the correlation between IRGPs and the OS of PRCC patients in the discovery cohort. Then, we applied a Cox proportional hazards regression model combined with the least absolute shrinkage and selection operator (Lasso) and 10-fold cross validation to minimize the risk of overfitting. After the above two screenings, the remaining IRGPs were used as the prognostic IRGPs to predict the OS of PRCC patients. Finally, using the values and coefficients of these prognostic IRGPs, we could build a model for calculating the immune-related Gene Pairs index (IRGPI) risk score of each sample. To separate patients into low or high-risk groups, time-dependent receiver operating characteristic (ROC) curve was used to find the optimal cut-off value of IRGPI at 1 year in the training cohort for OS. The point closest to the 100% true positive rate and 0% false-positive rate could be seen as the cut-off point.

Validation of the IRGPs prognostic signature

In order to verify the accuracy of IRGPs signature on the stratification of patients' prognosis, we calculated the IRGPI risk score of each sample of the discovery cohort (TCGA) and the validation cohort (GSE2748) separately, and divided them into different risk groups according to the cut-off value. Kaplan-Meier curve and log-rank test were used to verify whether the OS between the two groups was significantly different. In both discovery cohort and validation cohort, we calculated the area under the curve (AUC) to evaluate the prognostic accuracy of the immune-related risk signature, which ranges from 0 to 1 and 0.5 represents a random prediction[16].

Further, for confirming that the IRGPs signature was an independent prognostic factor, we combined the IRGPs signature with available clinicopathological in univariable and multivariable Cox proportional hazards regression model analyses, including age, gender, stage and PRCC type.

Analysis of correlation between IRGPs signature and immune cells infiltration

Tumor-infiltrating lymphocytes (TILs) in PRCC samples were assessed by applying the “Cell type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT)” deconvolution algorithm [17]. By analyzing the relative expression levels of 547 genes in samples, CIBERSORT could predict the proportion of 22 types of TILs in each PRCC sample. The gene expression signature matrix of 22 tumor-infiltrating immune cells was obtained from the CIBERSORT platform (<https://cibersortx.stanford.edu/>). We set 1000 permutations and $P < 0.05$ as the criteria for the successful deconvolution of a sample. Then, we compared the proportions of the immune cell subsets between the IRGPI risk groups using the Mann–Whitney U test.

Gene ontology and gene set enrichment analysis

GO analysis was performed for enhancing the biological understanding of the prognostic IRGPs signature. GSEA was conducted using the Bioconductor package “fgsea” with 100,000 permutations. We obtained and compared the log₂ fold change between the gene expression profiles of different IRGPI risk groups. All the biological processes involved in our study were obtained from the Molecular Signature Database (MSigDB C5 databases, version 7). Gene sets with FDR-adjusted $P < 0.05$ or nominal (NOM) $P < 0.05$ were selected.

Statistical analysis

The statistical software R (version 3.6.2), Perl (version 5.24.3) were used in the above analyses. Wilcoxon tests were used to compare the differences between two groups. A p-value < 0.05 was considered statistically significant.

Results

IRGPs prognostic signature construction

A total of 315 patients with PRCC were included in our study. We assigned the samples from TCGA ($n = 287$) to the discovery cohort and the samples from GEO ($n = 28$) to the validation cohort. Genes with relatively lower degree of variation ($MAD \leq 0.5$) were removed firstly. Then we obtained 2,498 immune-related genes (IRGs) from the ImmPort database, among them 172 IRGs were measured in filtered discovery cohort. Based on the 172 IRGs, we successfully constructed 3683 IRGPs as candidates. We evaluated the correlation between all IRGPs and OS using univariate Cox analysis, of which there were 48 significant prognostic IRGPs. Furtherly, we used Lasso Cox proportional hazard regression on the discovery cohort, and finally selected 14 IRGPs with more stable prognostic significance to construct the model for calculating the IRGPI (Table 1). The 14 IRGPs included a panel of 22 unique genes, and main of them were associated with antigen processing and presentation, antimicrobials, and cytokines. After calculating the IRGPI risk score of each patient, according to the time-dependent ROC curve analysis, the

cut-off value for distinguishing patients into high- or low-risk groups was determined to be 0.184 (Figure 1 and Table S1).

Validation of the IRGPs signature as an independent prognostic factor

In the discovery cohort, Kaplan-Meier curves showed that the OS of the patients in high-risk group was significantly poorer than that of the patients in low-risk group ($p < 0.001$) (Figure 2a). Subsequently, we jointly evaluated the effects of IRGPI risk score, age, gender, stage, and PRCC type on OS in univariate and multivariate Cox regression analyses. The results of the multivariate analyses demonstrated that IRGPI risk score and stage were the independent prognostic factors. (IRGPI risk score: HR, 3.548; 95%CI, 2.096–6.006; $p < 0.001$; Stage: HR, 1.880; 95%CI, 1.101–3.209; $p = 0.021$). (Figure 3a, 3b) The results of AUC demonstrated the excellent predictive accuracy of the signature for OS of PRCC patients (1-year AUC, 0.957; 3-year AUC, 0.825; 5-year AUC, 0.760).

To validate the consistency of prognostic value of the IRGPs signature, we applied it in an independent validation cohort from GSE2748 ($n = 29$). Similarly, we calculated the IRGPI risk score for each patient in the validation cohort and stratified patients according to the cut-off value obtained in the discovery cohort (Table S1). Same as the previous result, the high-risk group was associated with poorer OS than the low-risk group ($p < 0.001$) (Figure 2b). The results of univariate and multivariate Cox regression analyses demonstrated IRGPI risk score was the independent prognostic factor again (Univariate: HR, 2.721; 95%CI, 1.321–5.604; $p = 0.007$; and Multivariate: HR, 2.667; 95%CI, 1.105–6.435; $p = 0.029$) (Figure 3c, 3d). The accuracy of the application of the IRGPs signature in the validation cohort was still promising (1-year AUC, 0.786; 3-year AUC, 0.791; 5-year AUC, 0.820).

Immune cells infiltration in different risk groups

Based on CIBERSOFT algorithm, we systematically estimated the proportions of 22 kinds of TILs for each PRCC patient in different risk groups. Detailed information of the output of the algorithm was shown in Figure 4. We found that different TILs were significantly enriched in different risk groups. In the high-risk group, T CD8⁺ cells ($p < 0.001$), T regulatory cells (Tregs, $p = 0.001$), T follicular helper cells ($p < 0.001$), B cells naive ($p < 0.001$), Plasma cells ($p = 0.009$), T CD4⁺ memory cells activated ($p < 0.001$), Macrophage M1 ($p < 0.001$) were highly expressed, while Macrophage M0 ($p = 0.027$), Macrophages M2 ($p < 0.001$) were lowly expressed, compared with the low-risk group.

Functional assessment of the IRGPs signature

We performed GO analysis and GSEA for functional annotation of the IRGPs signature (Table S2). Figure 5 showed a total of top 50 GO terms with FDR < 0.05 , sorted by FDR. We found that the IRGPs signature genes in the discovery cohort were mostly involved in “mitotic sister chromatid segregation”. The results of GSEA demonstrated multiple immune-related pathways that differed between high- and low-risk groups significantly, including “adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains”, “lymphocyte mediated immunity”, “B cell

mediated immunity”, “immunoglobulin production”, “regulation of immune effector process”, and “regulation of lymphocyte activation” (Figure 6). Thus, the IRGPs signature demonstrated an intensive immune phenotype.

Discussion

Considering the important impact of tumor immunity in PRCC, our study analyzed a discovery cohort from TCGA to establish a robust prognostic signature based on 14 immune-related gene pairs for predicting OS of PRCC patients. The signature can accurately distinguish the prognosis of patients with PRCC and is a prognostic factor independent of other clinical pathological factors. An external validation cohort from GEO confirmed the reliability of the prognostic signature. Furtherly, we found that the signature was associated with various proportion of specific TILs in the tumor immune microenvironment, and is involved in many immune-related GO terms. The prognostic signature can be used as an important marker for risk stratification in patients with PRCC, and may be a potential target for immunotherapy.

In view of the obvious heterogeneity of PRCC, some research groups have studied the methods of classification or prognostic stratification of PRCC. At present, the most widely used PRCC classification method divides PRCC into type 1 and type 2 according to histological characteristics. Type 1 is characterized by papillary and tubular structures covered by small cells containing a basophilic cytoplasm and a small uniform oval nucleus, while type 2 is characterized by Large cells with eosinophilic cytoplasm and large spherical nuclei.[4, 19] However, the role of histological sub-types of PRCC in distinguishing patients from different clinical outcomes is controversial. A multicenter retrospective study included 486 patients undergoing partial nephrectomy with the two PRCC histological sub-types (76% type 1 and 24% type 2). The results showed that there were no demographic, clinical or tumor differences between the two types of PRCC [7]. Another research group performed a retrospective study of 88 PRCC patients and studied the prognostic factors of PRCC. The results of multivariate analysis demonstrated that the stage and grade were independent prognostic factors, excluding histological sub-types.[6] Our findings confirmed the results of previous studies again.

In both the discovery cohort and the validation cohort, the results of univariate and multivariate analyses suggested that the histological sub-types of PRCC were not independent prognostic factors for OS of patients with PRCC, while our IRGPs signature shows a robust independent prognostic ability. Therefore, although we already knew, different gene mutations had been associated with the 2 papillary histological sub-types, including FH gene with type 2 and c-met with type 1, the molecular factors that determine the clinical manifestations of tumors still required more exploration [20, 21].

Some researchers have developed several signatures on the prognosis of PRCC at the molecular level, including mRNA, lncRNA, alternative splicing, mutation and etc. [22-25]. However, these signatures have not yet reached a robust high accuracy rate, and have not considered the important impact of tumor immunology on the prognosis and treatment of PRCC. A signature based on immune-related genes of

PRCC has recently been reported [26]. Using data from TCGA, the authors constructed a prognostic signature of 15 immune genes to predict the survival outcome of KIRP patients, showing the value of immune-related prognostic signatures in PRCC. However, there were still some deficiencies in this study, such as the lack of external validation cohort, the lack of multivariate analysis including histological subtypes, etc. In our research, we further added important independent external validation sets to make the signature more robust, and at the same time proved the signature's prognostic role independent of age, gender, stage, and PRCC histological subtypes. As far as the accuracy of prediction was concerned, in the discovery cohort and validation cohort, our signature had an advantage in predicting AUC of 1, 3, and 5 years (0.957, 0.825, 0.760 versus 0.934, 0.796, 0.662). In addition, our signature was built on gene pairs, and this method had some additional advantages. The biological heterogeneity of tumors and differences in sequencing platforms often caused technical bias, thus standardizing gene expression profiles was necessary and difficult. We used a novel method based on gene pairs to construct the prediction model. Data preprocessing such as scaling and normalization was not required, instead, we compared the relative ranking and pairing of gene expression values. This method could reduce the impact of the technical bias of different platforms on the results and improve the robustness of the signature.[27] In recent years, this method has been applied to the construction of various tumor prognosis models with excellent results, including non-small cell lung cancer, colorectal cancer, and serous ovarian carcinoma [28-30]

The prognostic signature was constructed of 14 IRGPs with 22 unique genes. Most of this signature genes were related to antigen processing and presentation, and cytokines, and were enriched in multiple immune-related GO terms. Among the 22 unique genes, previous studies demonstrated NRF2A was an important gene regulating tumor cell dormancy. Down-regulated NRF2A was associated with the occurrence and recurrence of various tumors.[31, 32] In ccRCC, high expression of APRIL (TNFSF13) was closely related to poor prognosis, and VEGFA was significantly upregulated compared with normal tissue. [31, 32] High expression of CTSS is a predictor of poor prognosis and tumor metastasis in papillary carcinoma of the thyroid [35]. Meanwhile, overexpression of HSPA2 was related to tumor angiogenesis and poor prognosis of pancreatic cancer, while the survival prognosis of breast cancer patients with high expression of NOX4 was poor, too [36] [37]. By the CIBERSORT algorithm, we found that some TILs' proportions were significantly different between the two risk groups.

Between the two risk groups, we observed significant differences in the proportion of specific TILs. There were some interesting findings. In general, CD8⁺ T cells can recognize tumor specific antigens and played an important role in tumor immunity. Higher CD8⁺ T cell infiltration in multiple cancer types is associated with a better prognosis.[41]. However, in our study, we could find that the high-risk group had obvious higher CD8⁺ T cell infiltration than the low-risk group did. There are some research results that can explain this rare phenomenon to some degrees. First, previous research found that CD8⁺ T cells were not only specific for tumor-derived antigens, but also specific for non-tumor antigens. The enrichment of CD8⁺ T cells may not always play an anti-tumor effect, and has phenotypic heterogeneity in tumors and patients. Therefore, the prognostic effect of CD8⁺ T cells is not necessarily the same in different tumor

types and patients [42]. Meanwhile, At the same time, some researchers also found similar results in this study of RCC, that was, higher CD8⁺ T cell infiltration was associated with poor prognosis. The possible reason is the dysfunction of CD8 + T cells caused by various factors, such as high DGK-alpha, disabled MAPK pathways and JAK3/STAT5/6 pathway alterations [43]. And studies have confirmed that abnormal dendritic cells are involved in the process of CD8⁺ T cell suppression, which may cause CD8⁺ T cell to have higher abundance, but not to exert the corresponding anti-tumor function [44, 45]. In addition, researchers have found that the abundance of M2 macrophages and the abundance of CD8 + T cells in RCC are negatively correlated, which supports our findings [45] In our study, the M2 macrophage abundance of the high-risk group was significantly lower than that of the low-risk group. Similarly, the abundance of Treg cell and T cells follicular helper in the high-risk group was significantly higher than those in the low-risk group. These two TILs are considered to be factors that promote tumor progression and are related to the poor prognosis of patients [46, 47]. Previous publication found that the abundance of CD8⁺ T cells was positively correlated with the abundance of Tregs and T cells follicular helper, and negatively correlated with the abundance of M2 macrophages, which is consistent with our research results [45]. Macrophage M2, T regs and T cell follicular helper may play a role in the balance of the exhaustion or inhibition of T cells, and balance each other [48].

There are still some limitations to our study. First, although we have tried to introduce an external validation to improve the robustness of our results, our research is still retrospective in nature. In the future, we need more prospective research to further apply and verify our findings. Second, our research data is based on RNA-seq and microarray, the high price and complicated analysis process limit the clinical application of our results. We need more researches to explore how to simplify the IRGPs signature and how to combine it with existing clinical pathological factor to improve the ease of use and accuracy of clinical applications.

Conclusion

All in all, we have established an individualized prognostic immune-related gene pairs signature, which can accurately assess and predict the OS of patients with PRCC. The signature we developed is an independent prognostic factor, a practical tool for stratifying the prognosis risk of patients, and may provide a reference when screening PRCC patients to receive immunotherapy.

Abbreviations

AUC: the area under the curve; ccRCC: clear cell renal carcinoma; CI: confidence interval; GEO: Gene Expression Omnibus; HR: hazard ratio; IRG: immune-related Gene; IRGPs: immune-related gene pairs; IRGPI: immune-related Gene Pairs index; Lasso: least absolute shrinkage and selection operator; MAD: median absolute deviation; OS: overall survival; PRCC: Papillary renal carcinoma; ROC: receiver operating characteristic; TCGA: The Cancer Genome Atlas; TILs: tumor-infiltration lymphocytes;

Declarations

Ethics approval and consent to participate

This study used two datasets from the two public databases TCGA and GEO, and therefore did not require institutional review board approval and the patients' data were atomized.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analysed during the current study are available in the TCGA-KIRP, (<https://portal.gdc.cancer.gov/>) and GSE2748, (<http://www.ncbi.nlm.nih.gov/geo/>).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National key research and development program of China (Grant No. 2017YFC0908003), National Natural Science Foundation of China (Grant No.81902578, 81974098), China Postdoctoral Science Foundation (2017M612971), Post-doctoral Science Research Foundation of Sichuan University (2020SCU12041), Post-Doctor Research Project, West China Hospital, Sichuan University (2018HXBH085), National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University (Z2018C01).

Authors' contributions

QW, LY designed the study. XHZ, SQ are responsible for writing, collecting data, analysis, interpretation, revision present article. DJ, KJ are responsible for data collecting and analysis partly. XNZ are responsible for data analysis partly.

XHZ, SQ contributed to the study equally. All authors have read and approved the final manuscript.

Acknowledgements

The authors wish to thank all researchers and staff who uploaded and maintained data in TCGA and GEO databases for their great efforts.

References

1. Siegel, R.L., K.D. Miller and A. Jemal. Cancer statistics, 2020. *CA Cancer J Clin* 2020; 70: 7-30.

2. Courthod, G., M. Tucci, M. Di Maio, and G.V. Scagliotti. Papillary renal cell carcinoma: A review of the current therapeutic landscape. *Crit Rev Oncol Hematol* 2015; 96: 100-12.
3. Linehan, W.M., P.T. Spellman, C.J. Ricketts, C.J. Creighton, S.S. Fei, and C. Davis, et al. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *The New England journal of medicine* 2016; 374: 135-145.
4. Delahunt, B. and J.N. Eble. Papillary renal cell carcinoma: a clinicopathologic and immunohistochemical study of 105 tumors. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 1997; 10: 537-544.
5. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *New England Journal of Medicine* 2016; 374: 135-145.
6. Mejean, A., V. Hopirtean, J.P. Bazin, F. Larousserie, H. Benoit, and Y. Chrétien, et al. Prognostic factors for the survival of patients with papillary renal cell carcinoma: meaning of histological typing and multifocality. *The Journal of urology* 2003; 170: 764-767.
7. Bigot, P., J. Bernhard, I.S. Gill, N.S. Vuong, G. Verhoest, and V. Flamand, et al. The subclassification of papillary renal cell carcinoma does not affect oncological outcomes after nephron sparing surgery. *World journal of urology* 2016; 34: 347-352.
8. Crépel, M., H. Isbarn, U. Capitanio, D. Liberman, C. Jeldres, and M. Sun, et al. Does histologic subtype affect oncologic outcomes after nephron-sparing surgery? *Urology* 2009; 74: 842-845.
9. Del Paggio, J.C. Immunotherapy: Cancer immunotherapy and the value of cure. *Nature reviews. Clinical oncology* 2018; 15: 268-270.
10. Motzer, R.J., B. Escudier, D.F. McDermott, S. George, H.J. Hammers, and S. Srinivas, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *The New England journal of medicine* 2015; 373: 1803-1813.
11. McKay, R.R., D. Bossé, W. Xie, S.A.M. Wankowicz, A. Flaifel, and R. Brandao, et al. The Clinical Activity of PD-1/PD-L1 Inhibitors in Metastatic Non-Clear Cell Renal Cell Carcinoma. *Cancer immunology research* 2018; 6: 758-765.
12. Zhang, S., E. Zhang, J. Long, Z. Hu, J. Peng, and L. Liu, et al. Immune infiltration in renal cell carcinoma. *Cancer science* 2019; 110: 1564-1572.
13. Pan, Q., L. Wang, S. Chai, H. Zhang, and B. Li. The immune infiltration in clear cell Renal Cell Carcinoma and their clinical implications: A study based on TCGA and GEO databases. *Journal of Cancer* 2020; 11: 3207-3215.

14. Wang, Z., Q. Song, Z. Yang, J. Chen, J. Shang, and W. Ju. Construction of immune-related risk signature for renal papillary cell carcinoma. *Cancer medicine* 2019; 8: 289-304.
15. Yang, X.J., M. Tan, H.L. Kim, J.A. Ditlev, M.W. Betten, and C.E. Png, et al. A molecular classification of papillary renal cell carcinoma. *Cancer research* 2005; 65: 5628-5637.
16. Heagerty, P.J., T. Lumley and M.S. Pepe. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000; 56: 337-344.
17. Newman, A.M., C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, and Y. Xu, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature methods* 2015; 12: 453-457.
18. Zbar, B., K. Tory, M. Merino, L. Schmidt, G. Glenn, and P. Choyke, et al. Hereditary papillary renal cell carcinoma. *The Journal of urology* 1994; 151: 561-566.
19. Tomlinson, I.P.M., N.A. Alam, A.J. Rowan, E. Barclay, E.E.M. Jaeger, and D. Kelsell, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nature genetics* 2002; 30: 406-410.
20. Schmidt, L., F.M. Duh, F. Chen, T. Kishida, G. Glenn, and P. Choyke, et al. Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas. *Nature genetics* 1997; 16: 68-73.
21. Zuo, S., L. Wang, Y. Wen, and G. Dai. Identification of a universal 6-lncRNA prognostic signature for three pathologic subtypes of renal cell carcinoma. *J Cell Biochem* 2018.
22. Duan, Y. and D. Zhang. Identification of novel prognostic alternative splicing signature in papillary renal cell carcinoma. *J Cell Biochem* 2020; 121: 672-689.
23. Zhang, C., Y. Zheng, X. Li, X. Hu, F. Qi, and J. Luo. Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. *Ann Transl Med* 2019; 7: 427.
24. Cao, D.L., W.X. Dai, Y.Q. Huang, L.J. Yu, J.L. Wu, and G.H. Shi, et al. Development and validation of a robust multigene signature as an aid to predict early relapse in stage I-III clear cell and papillary renal cell cancer. *J Cancer* 2020; 11: 997-1007.
25. Wang, Z., Q. Song, Z. Yang, J. Chen, J. Shang, and W. Ju. Construction of immune-related risk signature for renal papillary cell carcinoma. *Cancer Med* 2019; 8: 289-304.
26. Eddy, J.A., J. Sung, D. Geman, and N.D. Price. Relative expression analysis for molecular cancer diagnosis and prognosis. *Technology in cancer research & treatment* 2010; 9: 149-159.
27. Li, B., Y. Cui, M. Diehn, and R. Li. Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer. *JAMA oncology* 2017; 3:

1529-1537.

28. Wu, J., Y. Zhao, J. Zhang, Q. Wu, and W. Wang. Development and validation of an immune-related gene pairs signature in colorectal cancer. *Oncoimmunology* 2019; 8: 1596715.
29. Zhang, L., P. Zhu, Y. Tong, Y. Wang, H. Ma, and X. Xia, et al. An immune-related gene pairs signature predicts overall survival in serous ovarian carcinoma. *OncoTargets and therapy* 2019; 12: 7005-7014.
30. Sosa, M.S., F. Parikh, A.G. Maia, Y. Estrada, A. Bosch, and P. Bragado, et al. NR2F1 controls tumour cell dormancy via SOX9- and RAR β -driven quiescence programmes. *Nature communications* 2015; 6: 6170.
31. Borgen, E., M.C. Rypdal, M.S. Sosa, A. Renolen, E. Schlichting, and P.E. Lønning, et al. NR2F1 stratifies dormant disseminated tumor cells in breast cancer patients. *Breast cancer research : BCR* 2018; 20: 120.
32. Lee, C., J. Park, J.H. Suh, and K.C. Moon. High expression of APRIL correlates with poor prognosis in clear cell renal cell carcinoma. *Pathology, research and practice* 2015; 211: 824-828.
33. Wang, X., J. Zhang, Y. Wang, M. Tu, Y. Wang, and G. Shi. Upregulated and act as potential prognostic genes for clear cell renal cell carcinoma. *OncoTargets and therapy* 2018; 11: 1697-1706.
34. Tan, J., X. Qian, B. Song, X. An, T. Cai, and Z. Zuo, et al. Integrated bioinformatics analysis reveals that the expression of cathepsin S is associated with lymph node metastasis and poor prognosis in papillary thyroid cancer. *Oncology reports* 2018; 40: 111-122.
35. Zhai, L., Q. Xie, C. Zhou, D. Huang, Z. Tang, and T. Ju. Overexpressed HSPA2 correlates with tumor angiogenesis and unfavorable prognosis in pancreatic carcinoma. *Pancreatology : official journal of the International Association of Pancreatology (IAP) ... [et al.]* 2017; 17: 457-463.
36. Ham, J., S. Lee, H. Lee, D. Jeong, S. Park, and S.J. Kim. Genome-Wide Methylation Analysis Identifies NOX4 and KDM5A as Key Regulators in Inhibiting Breast Cancer Cell Proliferation by Ginsenoside Rg3. *The American journal of Chinese medicine* 2018; 46: 1333-1355.
37. Chen, D.S. and I. Mellman. Oncology meets immunology: the cancer-immunity cycle. *Immunity* 2013; 39.
38. Simoni, Y., E. Becht, M. Fehlings, C.Y. Loh, S. Koo, and K.W.W. Teng, et al. Bystander CD8 T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* 2018; 557: 575-579.
39. Prinz, P.U., A.N. Mendler, I. Masouris, L. Durner, R. Oberneder, and E. Noessner. High DGK- α and disabled MAPK pathways cause dysfunction of human tumor-infiltrating CD8+ T cells that is reversible by pharmacologic intervention. *Journal of immunology (Baltimore, Md. : 1950)* 2012; 188: 5990-6000.

40. Giraldo, N.A., E. Becht, F. Pagès, G. Skliris, V. Verkarre, and Y. Vano, et al. Orchestration and Prognostic Significance of Immune Checkpoints in the Microenvironment of Primary and Metastatic Renal Cell Cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2015; 21: 3031-3040.
41. Pan, Q., L. Wang, S. Chai, H. Zhang, and B. Li. The immune infiltration in clear cell Renal Cell Carcinoma and their clinical implications: A study based on TCGA and GEO databases. *Journal of Cancer* 2020; 11: 3207-3215.
42. Finotello, F. and Z. Trajanoski. New strategies for cancer immunotherapy: targeting regulatory T cells. *Genome medicine* 2017; 9: 10.
43. Long, J., A. Wang, Y. Bai, J. Lin, X. Yang, and D. Wang, et al. Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma. *EBioMedicine* 2019; 42: 363-374.
44. Speiser, D.E., D.T. Utzschneider, S.G. Oberle, C. Münz, P. Romero, and D. Zehn. T cell differentiation in chronic infection and cancer: functional adaptation or exhaustion? *Nature reviews. Immunology* 2014; 14: 768-774.

Table 1

Table 1 Detail information about 14 immune-related gene pairs used to calculate immune-related gene pairs index

IRG 1	Immune process	IRG 2	Immune process	Coefficient
CTSS	Antigen_Processing_and_Presentation	ADM	Antigen_Processing_and_Presentation	-0.13784168
HLA-DPA1	Antigen_Processing_and_Presentation	IFITM1	BCRSignalingPathway	-0.536023127
HSPA2	Antigen_Processing_and_Presentation	NR2F1	Cytokine_Receptors	-0.813388891
MICB	Antigen_Processing_and_Presentation	CX3CR1	Chemokine_Receptors	+0.547230703
RBP4	Antimicrobials	TNFRSF19	Cytokine_Receptors	-0.479900288
NOX4	Antimicrobials	TNFSF13B	Cytokines	-0.536067102
CHIT1	Antimicrobials	CCL4	Antimicrobials	-0.055801169
VEGFA	Antimicrobials	AR	Cytokine_Receptors	+0.295949021
VEGFA	Antimicrobials	ITGB2	NaturalKiller_Cell_Cytotoxicity	+0.313650746
ITGAV	Antimicrobials	TNFSF13	Cytokines	+0.033017443
WNT5A	Antimicrobials	NR2F1	Cytokine_Receptors	-0.421788958
BTK	Antimicrobials	TNFSF13B	Cytokines	-0.212282282
IFITM1	BCRSignalingPathway	TNFSF13	Cytokines	+0.012878529
TNFSF13B	Cytokines	CSF3R	Cytokine_Receptors	+0.448335019

Abbreviations: IRG: immune-related gene

Figures

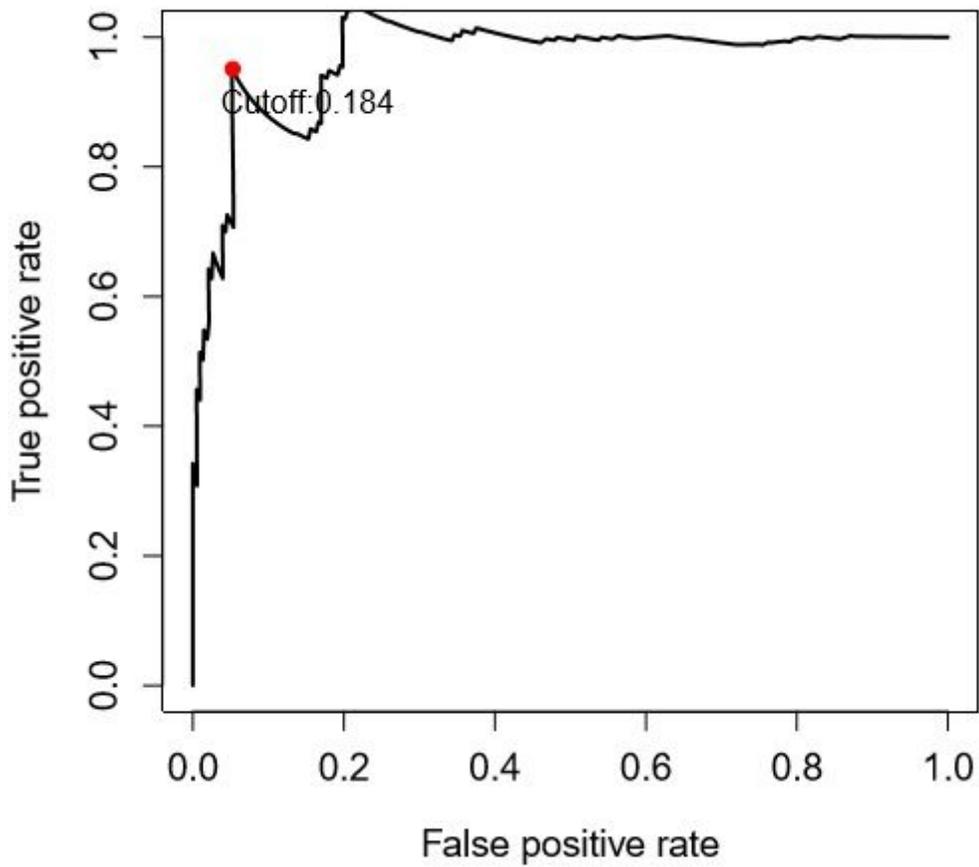


Figure 1

Time-dependent ROC curve for IRGPI risk score in the discovery cohort. 0.184 was used as a cut-off for IRGPI risk score to stratify patients into low- or high-risk groups.

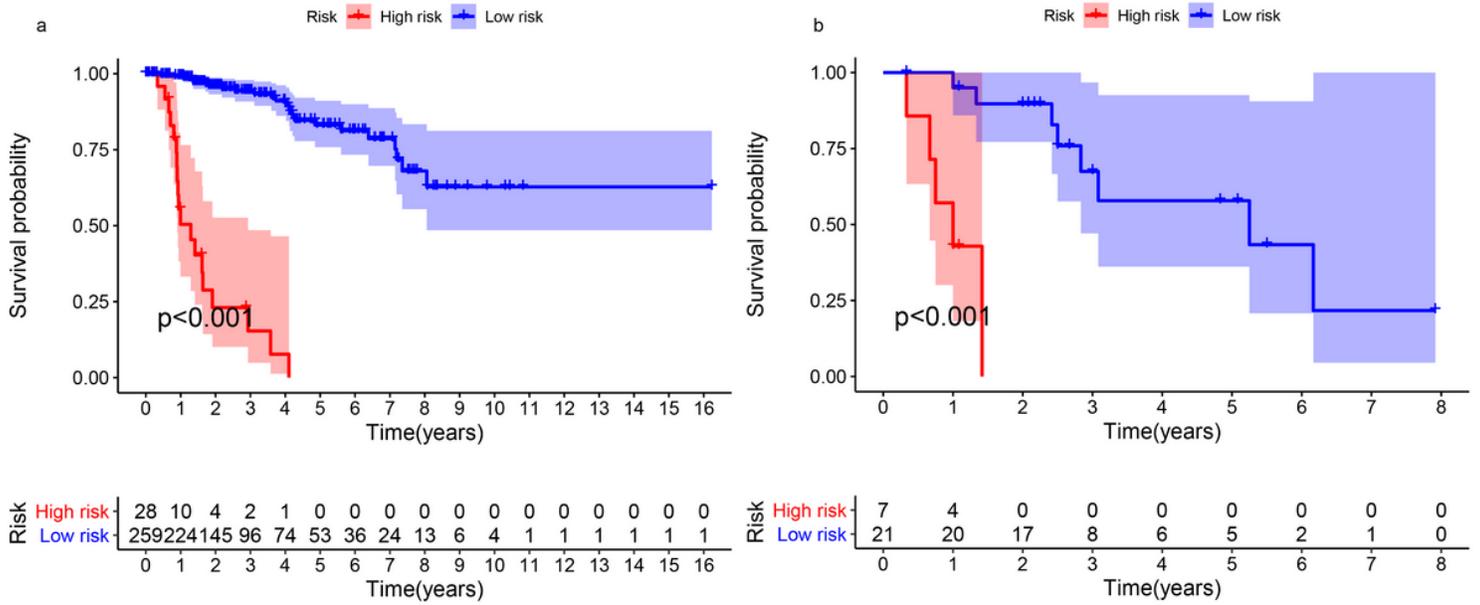


Figure 2

Kaplan-Meier curves of OS between different IRGPI risk groups in discovery cohort (a) and validation cohort (b). OS: overall survival; IRGPI: immune-related gene pair index

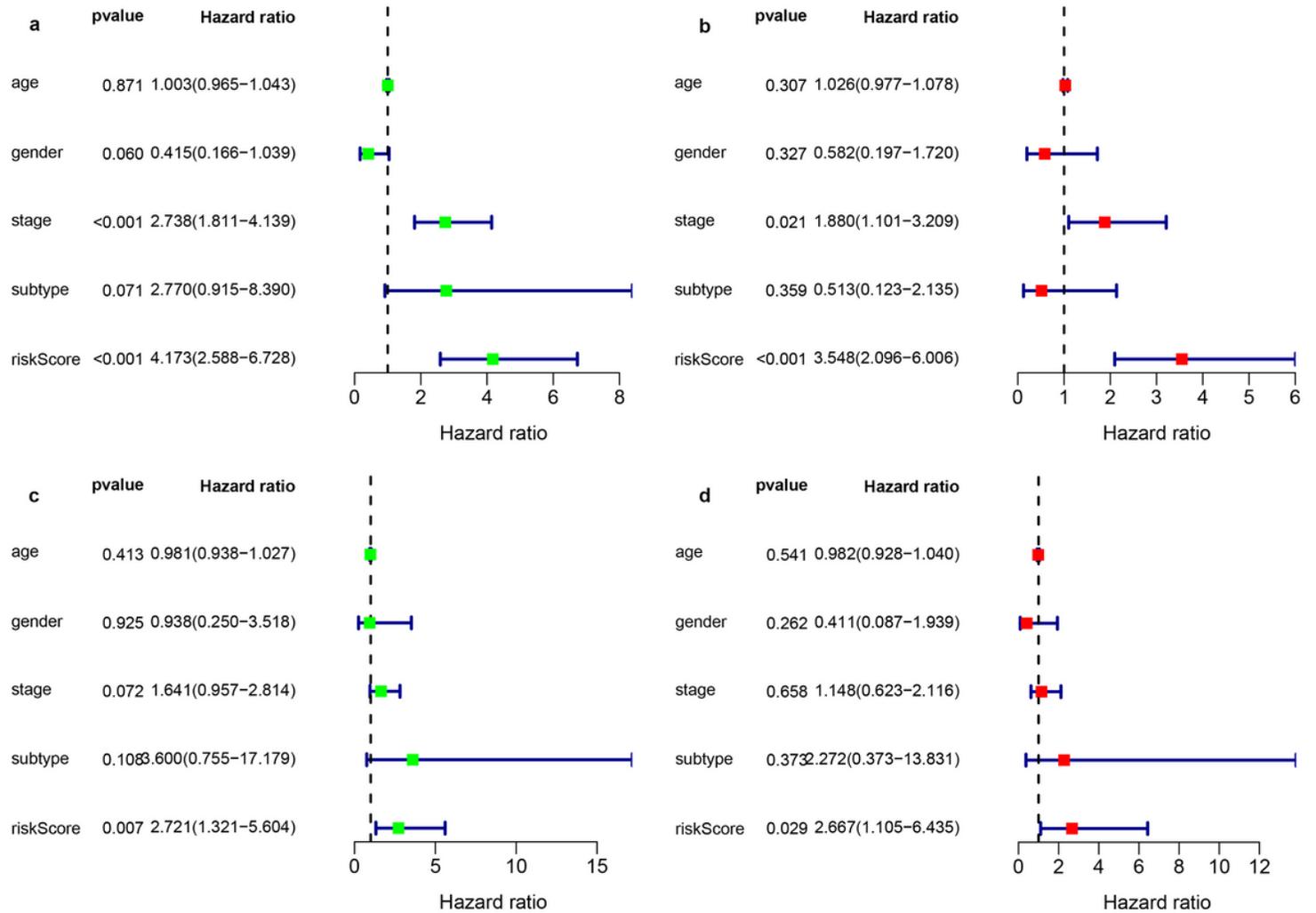


Figure 3

Univariate (a) and multivariate (b) analyses of prognostic factors in discovery cohort and univariate (c) and multivariate (d) analyses of prognostic factors in validation cohort. Note: subtype, histological subtypes of papillary renal carcinoma; riskScore, immune-related gene pair index (IRGPI) risk score

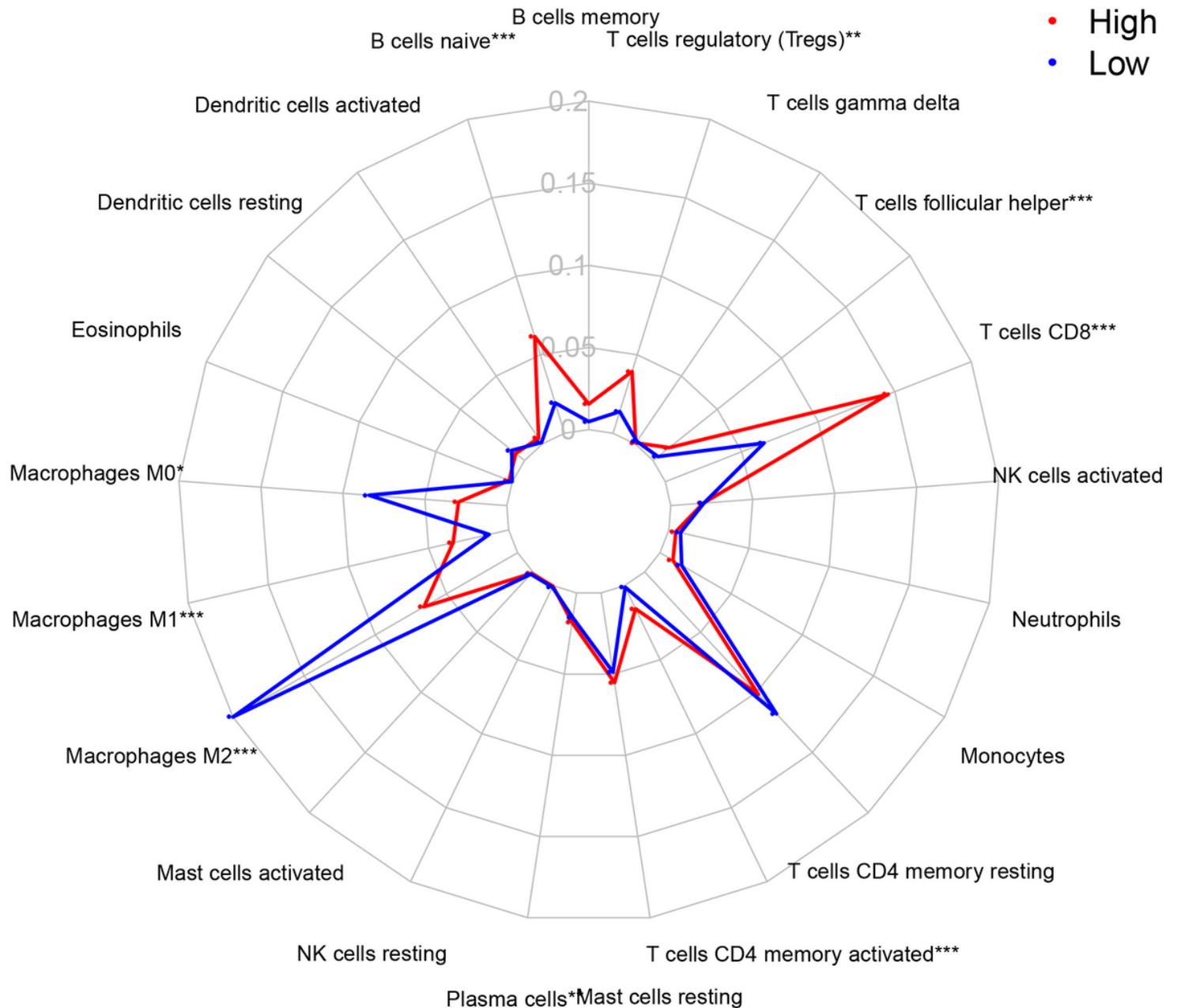


Figure 4

The proportion of 22 TILs in tumor immune microenvironment in two RIGPI risk groups. Note 1: TILs, tumor-infiltration lymphocytes; IRGPI, immune-related index; *, p<0.05; **, p<0.01; ***, p<0.001; Note 2: T cells CD4 + naive is not shown in the picture because of its low abundant.

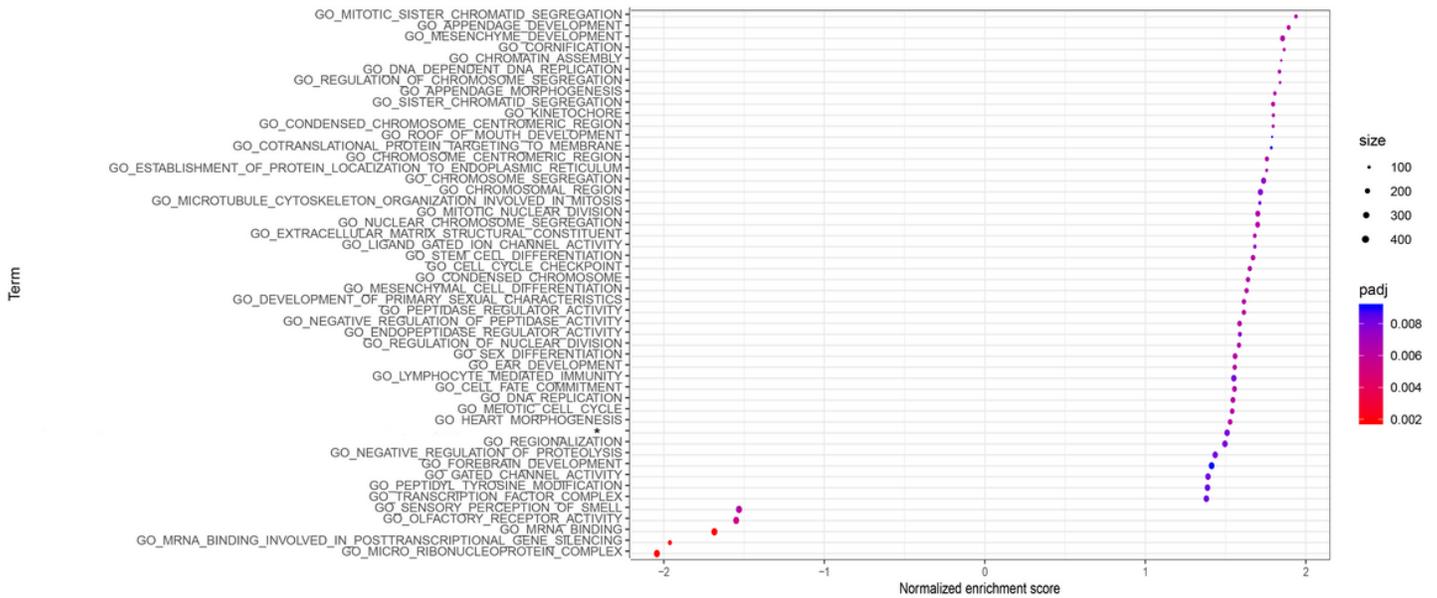


Figure 5

GO analysis of the 22 immune signature genes. The top 50 GO terms with FDR <0.05 are shown in the figure. Note: GO, gene ontology; FDR, false discovery rate

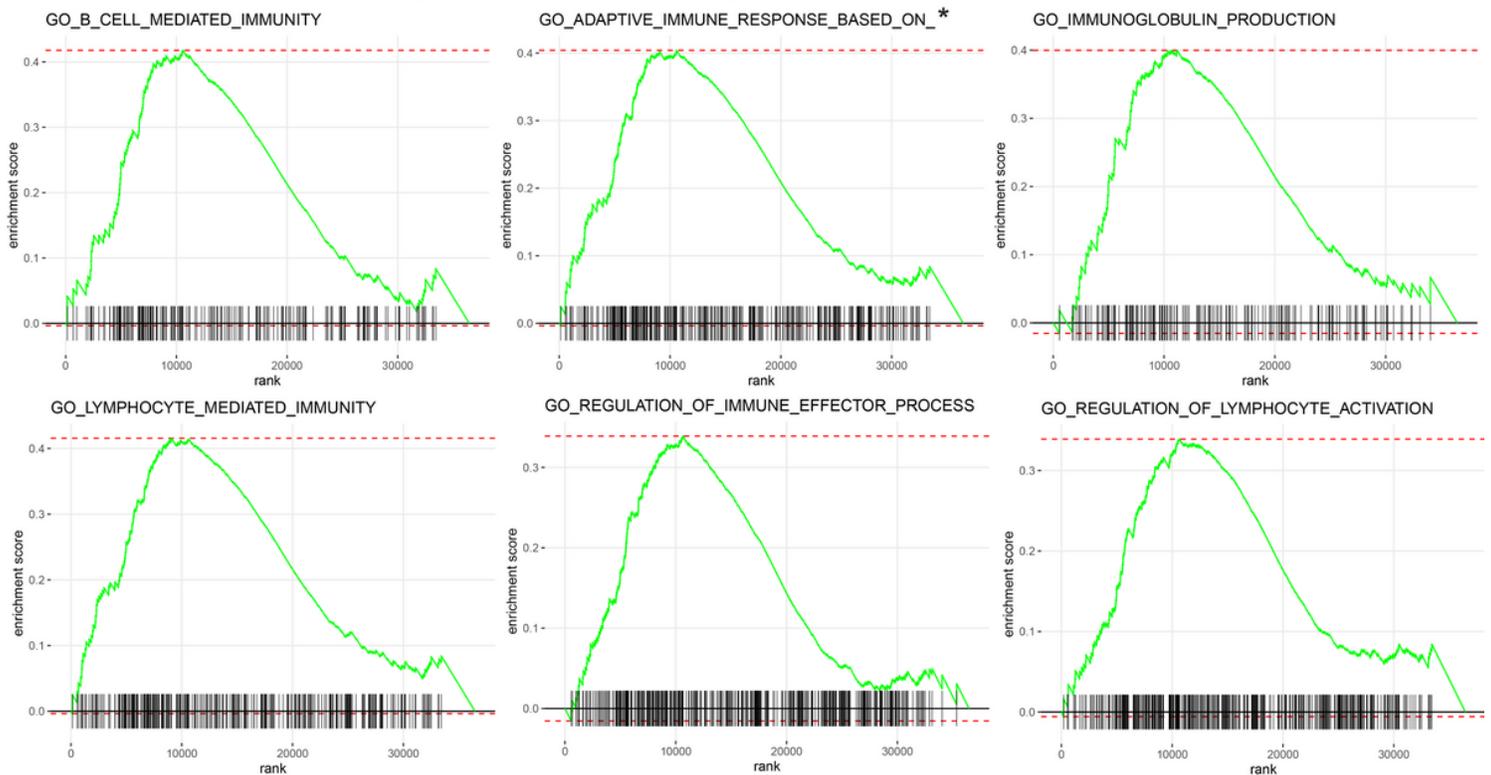


Figure 6

GSEA of the IRGPI risk groups. GSEA found 6 immune-related pathways with significant differences between two groups. Note: GSEA, gene set enrichment analysis; IRGPI, immune-related index

Supplementary Files

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

- [Supplementary.pdf](#)