

Identification and validation of immunotypes for clear cell RCC prognosis

Zhaolin Yang (✉ 18369611522@163.com)

Shanghai Jiao Tong University

Jiale Zhou

Shanghai Jiao Tong University

Yizheng Xue

Shanghai Jiao Tong University

Yu Zhang

Shanghai International Studies University

Kaijun Zhou

Shanghai International Studies University

Ying Zhang

Shanghai International Studies University

Jin Zhang

Shanghai Jiao Tong University

Jiwei Huang

Shanghai Jiao Tong University

Yiran Huang

Shanghai Jiao Tong University

Wei Xue

Shanghai Jiao Tong University

Yan Li

Shanghai International Studies University

Wei Zhai

Shanghai Jiao Tong University

Yonghui Chen

Shanghai Jiao Tong University

Research article

Keywords: immune gene, metastasis, Risk-metastasis model, immunotype model, immune cell, prognosis

Posted Date: August 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-847352/v1>

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

[Read Full License](#)

Identification and validation of immunotypes for clear cell RCC prognosis

Zhaolin Yang^{1,#}, Jiale Zhou^{1,#}, Yizheng Xue^{1,#}, Yu Zhang², Kaijun Zhou², Ying Zhang², Jin Zhang¹, Jiwei Huang¹, Yiran Huang¹, Wei Xue¹, Yan Li^{2,*}, Wei Zhai^{1,*} and Yonghui Chen^{1,*}

1 Department of Urology, Renji Hospital, School of Medicine in Shanghai Jiao Tong University, 160 Pujian Road, Pudong District, Shanghai, 200127, China

2 School of Business and Management, Shanghai International Studies University, 550 Dalian Rd. (W), Hongkou District, Shanghai, 200083, China

Contributed equally to this work

*Correspondence to:

Yan Li (liyan0417@shisu.edu.cn)

Wei Zhai (jacky_zw2002@hotmail.com)

Yonghui Chen (cyh1488@163.com)

Abstract

Purpose To develop an immunotype-based prognostic model for predicting the overall survival (OS) of patients with clear cell renal carcinoma (ccRCC). We explored novel immunotypes of patients with ccRCC, particularly those associated with overall survival. A risk-metastasis model was constructed by integrating the immunotypes with immune genes and used to test the accuracy of the immunotype model.

Patients and Methods Patient cohort data were obtained from The Cancer Genome Atlas (TCGA) database, Gene Expression Omnibus (GEO) database, Renji database, and Surveillance, Epidemiology, and End Results (SEER) database. We employed the R software to select 3 immune cells and construct an immunotype-based prediction model. Immune genes selected using random Forest Algorithm were validated by immunohistochemistry (IHC). The H&L risk-metastasis model was constructed to assess the accuracy of the immunotype model through Multivariate COX regression analysis.

Result Patients with ccRCC were categorized into immunotype H subgroup and immunotype L subgroup based on the overall survival rates. The immunotypes were found to be the independent prognostic index for ccRCC prognosis. As such, we constructed a new immunotypes-based SSIGN model. Three immune genes associated with difference between immunotype H and L were identified. An H&L risk-metastasis model was constructed to evaluate the accuracy of the immunotype model. Compared to the W-Risk-metastasis model which did not incorporate immunotypes, the H&L risk-metastasis model was more precise in predicting the survival of ccRCC patients.

Conclusion The established immunotype model can effectively predict the survival of ccRCC patients. Except for mast cells, T cells and macrophages are positively associated with the overall survival of patients. The three immune genes identified, herein, can predict the survival rate of ccRCC patients, and expression of these

immune genes is strongly linked to poor survival. The new SSIGN model provides an accurate tool for predicting the survival of ccRCC patients. H&L risk-metastasis model can effectively predict the risk of tumor metastasis.

Keywords immune gene; metastasis; Risk-metastasis model; immunotype model; immune cell; prognosis

Introduction

Clear renal cell carcinoma (ccRCC), the most common type of kidney cancer, accounting for 2% of all cancer cases and deaths, has seen an unprecedented rise worldwide ^{1, 2}. Due to its high potential for metastasis, ccRCC has a poor prognosis ³, and approximately 30% of RCC patients develop metastasis after surgery ⁴. Emerging evidence shows that common ccRCC metastases sites are lungs, regional lymph nodes, bones, liver, and brain ⁵. A few studies have evaluated the time of metastasis and OS in metastatic ccRCC.

Immune cells infiltration of the tumor microenvironment (TME) has an important role in tumorigenesis. T-cells, macrophages, and mast cells are crucial components of the TME ⁶. More T-cells have been reported in the TME of solid tumors⁷. CD4+ helper and cytotoxic CD8+ T-cells have potential inhibitory effects on tumor growth by targeting antigenic tumor cells ^{8, 9}. Tumor-associated macrophages (TAMs), innate immune effector cells from the myeloid lineage that influence tissue development, homeostasis, and immune surveillance, may also suppress or promote tumor growth ^{13, 14}. Migration of mast cells(derived from bone-marrow hematopoietic progenitors) from vascular to peripheral tissues may promote tumor growth ^{15, 16}. There is evidence that TME not only plays a significant role in cancer initiation, progression, and metastasis but also influences the effectiveness of targeted therapy ¹⁷. Thus, modulation of TME is of promise in improving the effectiveness of existing therapies and the discovery of novel avenues for targeted therapy ^{18, 19}.

Advances in gene therapy have achieved great success in managing various cancers ²⁰ and are now under consideration for severe diseases ²¹. Currently, tyrosine kinase inhibitors (TKIs) are first-line recommended

treatments for mRCC ^{1, 2}. Targeted therapies such as TKI generally improve patient survival. The cytokines interleukin-2 (IL-2) and interferon-alpha (IFN- α), also have significant clinical benefits. However, most mRCC patients eventually develop resistance to TKIs after 8–9 months, on average. Recent evidence shows that cancer immunotherapy has been used in managing carcinoma ²²⁻²⁴; however, in tumor micro-environment, the immune cells associated clinical outcomes are poorly understood ²⁴. Recent developments in the mechanisms of the immune system have brought anti-ccRCC immunotherapy into focus. Particularly, targeted therapy and immunotherapy are being used to manage mRCC ²⁵. Thus, necessitating an urgent development of new therapies, especially for patients not benefiting from TKI, IL-2, and IFN.

Currently, tools for stratifying ccRCC patients are limited to aspects such as tumor size. Evaluation of ccRCC prognosis relies on a few clinical and pathological factors for stratifying patients based on the risk of recurrence and metastasis. Immunotypes can be more precise in predicting the OS of patients. However, no detailed analysis of the correlation between immune types and immune genes in ccRCC has been reported.

Here, our risk-metastasis model aims to elucidate the relationship between immunotype, immune genes, and metastasis. We hypothesized that there are 2 different immunotypes, immunotype H and immunotype L. To test this possibility, 21 immune cells were explored to reveal the most relevant for prognosis. Next, we constructed an immune index, validated the immunotype model, and perfected the previous SSIGN model. Three immune genes associated with immunotype model were identified and could predict survival and improve ccRCC therapy. Finally, we integrated risk parameters, including age, gene, immunotypes, and immune-score to construct a risk-metastasis model and tested the immunotype validity.

Methods and procedures

Collection of patient samples

We conducted a series of experiments to determine whether immunotypes could predict the prognosis of ccRCC patients. In the first experiment, 2 independent patient data sets were obtained from 2 databases, the TCGA database (Renal clear cell carcinoma Dataset, <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>) and Renji Hospital database were analyzed. In the second experiment, an immunotype model comprising 3 independent patient data sets obtained from 3 databases was constructed. The training cohort included 539 patients obtained from the TCGA database and the testing cohort included 130 patients downloaded from the GEO database (GSE46699 dataset) (<https://www.ncbi.nlm.nih.gov/geo/>), whereas the validation cohort comprised 267 ccRCC patients from Renji Hospital database. In the third experiment, we tested the performance of the new SSIGN model, based on 3 independent data sets from patients with ccRCC obtained from 3 databases. The training cohort included 2877 patients obtained from the SEER database (Renal clear cell carcinoma Dataset), whereas the testing cohort was composed of patients from the TCGA database. The validation cohort comprised patient data retrieved from the Renji database.

Human specimens were collected with informed consent from patients, The use of human samples in this study was approved by the Clinical Research Ethics Committee of Renji Hospital. The inclusion criteria included: (a) patients who provided informed consent; (b) patients without comorbidities (e.g., those living with other malignant tumors.); (c) patients with complete follow-up data; (d) patients who underwent nephrectomy, verified by post-operative histopathology diagnosis; (e) Tumor stage classification was conducted or updated in line with the 2016 TNM classification. According to the 2016 Fuhrman RCC morphologic stage system, the Renji database was divided into four groups; F1, F2, F3, and F4. Patients with metastatic tumors were assigned to the M group.

The primary outcome of survival analysis was overall survival (defined as the time from the start of systemic

therapy to death). A follow-up protocol on patient history, physical examination, and laboratory measurement was implemented every 3 months in the first year, and then every 6 months in the next year.

Immunohistochemistry

Immunohistochemistry for the target immune proteins was performed on paraffinized sections using the following primary antibodies: anti-CD68, anti-CD8, anti-AA1, anti-SLN, anti-CPA4, and anti-AQP9. Paraffinized samples were sectioned into 6 pieces (each with 5 μ m thickness) then subjected to immunohistochemical staining for immune cells (T cells, macrophages cells, mast cells) and immune-related genes (SLN, AQP9, CPA4). At least two pathologists were blinded to the clinicopathological features of the tumors assessed the degree of reactivity. The stained ccRCC tissues were histologically scored in randomly selected fields (magnification, 10 \times 2, 10 \times 10, 10 \times 40).

The tissues were examined using the Pannoramic MIDI a tissue chip scanner (3D HISTECH). For data analysis, we employed the Quant center software for Pannoramic viewer. The densito quant software was applied for automated identification and setting of all dark brown spots (strong positives), brown yellows (moderate positives), light yellows (weak positives), and blue cell nuclei (negatives) on the tissue sections. The percentage of positivity was used to calculate the immunoscore.

Briefly, the number of positive cells and the intensity of staining in each section were converted into corresponding values for semi-quantitative straining purposes. $H\text{-SCORE} = \sum (P_i \times i) = (\text{percentage of weak intensity} \times 1) + (\text{percentage of moderate-intensity} \times 2) + (\text{percentage of strong intensity} \times 3)$. P_i represents the number of positive cells as a percentage of the total cells and i represents the intensity of staining.

Statistical analysis

All analyses including, survival analysis, Cox regression analysis, and single-factor Cox analyses were

performed in R software. Three immune cells were selected. R software was also employed for LASSO regression model analysis to calculate the immune coefficient formula, where $II = 2^{(0.40524 \times T\text{-cells} + 0.20646685 \times \text{macrophages} - 0.26974584 \times \text{mast cells})}$. PCA analysis allowed for the calculation of immune-score for 3 different immune cells. The distinction between immunotype H with immunotype L subgroups of the Renji Hospital database was plotted via correlation analysis. The relationship between cells and clinical stages was evaluated on Boxplot based on the Renji database. We performed ROC curve analysis in the training cohort and the logistic model calculated. Using a nomogram, we improved the previous SSIGN into a new SSIGN model. A confusion matrix of the logistic model in TCGA and Renji database was generated.

According to TCGA datasets, gene expression differences between the H&L immunotype groups were identified via single-factor COX regression analysis. Using R, forest maps of the genes were generated, from which 659 genes were further identified. Subsequently, 14 immune-related genes were selected based on the standard of importance >2 . GEPIA (<http://gepia.cancer-pku.cn/>) was employed for KM analysis between gene expression and progression-free survival (PFS), and to plot a gene expression rectangular map [$|\log_2fc|$ cutoff: 1; p-value cutoff: 0.01; database: TCGA normal and GTEX data], and map the relationship between gene expression and clinical stage [use major stage: Yes. Log scale: Yes].

In the experimental group, according to the TCGA database, patients were divided into groups H and L based on the immunotype H & L model. Using the multivariate Cox model, we identified different groups to reconstruct the model. Multivariate Cox regression analysis of immunotype H&L was performed. R software was used to map metastasis risk for patients in groups H and L. In the control group, which comprised the same TCGA dataset without distinguishing immunotype H and L, the tumor metastasis risk model incorporated the whole model factors, T-cells, mast cells, macrophages, CPA4, SLN, AQP9, immune-score, and age. Analysis of the accuracy of the risk model was calculated using the random forest algorithm.

Results

Selecting immune cells associated with ccRCC from TCGA

Kaplan-Meier analysis of the correlation between OS and levels of various immune cells revealed 4 immune cells as significantly correlated. T follicular helper cells and Tregs were negatively correlated with survival while resting dendritic cells and resting mast cells were positively correlated with survival rate (Figure 1A). No significant differences were reported in other immune cells (Supplementary Figure A). Univariate Cox regression analysis revealed that the P-value of Tregs, resting mast cells, M2 macrophages, activated memory CD4 T-cells and follicular helper T-cells were markedly valuable (Table 1). However, there was no useful P-value found in the other 16 immune cells. Analysis of correlation between immune cell infiltration and tumor grade revealed that resting mast cells, activated memory CD4 T-cells, resting NK cells, Tregs, plasma cells, resting memory CD4 T-cells, macrophages M2, CD8 T-cells, monocytes, follicular helper T-cells, activated dendritic cells, M1 macrophages, and M0 macrophages were significantly different based on the relationship between grade and infiltration ($p < 0.05$, Figure 1B, Supplementary Figure B). The other immune cells were not significantly different (Supplementary figure C). M2 macrophages, resting mast cells, Tregs, follicular helper T-cells, resting memory CD4 T-cells, CD8 T-cells, resting NK cells, and activated memory CD4 T-cells were found to be significantly associated with tumor stage ($p < 0.05$, Figure 1C, Supplementary Figure D) but correlation with the other immune cells was not significant (Supplementary figure E). Taken together, methods outlined above demonstrated that T-cells, macrophage cells, and mast cells can be used in immunotype model construction. To verify the accuracy of these immune cells, we evaluated their levels in immunohistochemistry (IHC) data from Renji Hospital (Figure 1D-F). In the Renji database, patients were divided into F1, F2, F3, F4, and metastasis subgroups. IHC results of these 3 immune cells revealed that the levels of T-cells and

macrophages increased with Fuhrman RCC morphologic stage, while mast cells levels decreased with Fuhrman RCC morphologic stage. Thus, T-cells and macrophages correlated negatively with overall survival, while mast cells exhibited a positive correlation. Therefore, the 3 immune cells were selected for further analysis.

Constructing immune index and immunotypes

We used 3 cohorts (the training, testing, and validation cohorts, based on TCGA, GEO, and Renji hospital datasets, respectively) to explore the correlation of these 3 immune cells with ccRCC. Using R, we identified lambda to be 0.00171. Based on LASSO Cox regression analysis of the training cohort, the coefficients for these 3 immune cells were calculated using the formula: $II=2^{(0.40524 \times T \text{ cell} + 0.20646685 \times \text{Macrophages} - 0.26974584 \times \text{Mast cells})}$ (Figure 2A). Values of $II \geq 0.2145$ defined immunotype H, while values of $II < 0.2145$ defined immunotype L. Based on TCGA database, survival rate varied significantly between immunotype L (low immunotype) and immunotype H (high immunotype) (Figure 2B), suggesting that the newly defined immunotype could effectively predict ccRCC prognosis. To validate the accuracy of this immunotype, we classified tumor and normal tissue groups into immunotype H and L subgroups based on the testing cohort (Figure 2C-D). This analysis revealed that these immunotypes not only effectively distinguished tumors from normal controls but also precisely identified high-score and low-score subgroups. A profile built to distinguish between immunotype H and immunotype L subgroups (Figure 2F) showed that immunotypes were suitable for identifying immunotype H and immunotype L based on the 3 selected immune cells according to Renji database (Figure 2E). We proceeded to evaluate the predictive value of the immunotype for overall survival of patients by comparing tumor stage to grade among immunotype H and immunotype L in the validation cohort (Figure 2F). The results revealed that in different immunotype subgroups, immune-score positively correlated with clinical stage and grade, demonstrating the accuracy of

the immunotype in the validation cohort. Finally, analysis of the accuracy of the immunotype model according to the RECA-EU database uncovered immunotype A group with a high immune score and immunotype B group with a low immune score (Figure 2G). The immunotype A group exhibited a lower mean survival rate, which demonstrated the accuracy of the immunotype model.

Refining SSIGN model by adding the immunotypes

Three cohorts were used to improve the SSIGN model (a widely used model for RCC prognosis by Mayo Clinic). The training, testing, and validation cohorts were based on SEER, TCGA, and Renji databases, respectively. To evaluate the prognostic value of immunotype H&L, we constructed a ROC curve and confusion matrix of logistic model for survival rate based on the training cohort (Figure 3A). It was revealed that the prediction accuracy based on the refined immunotype H&L model was higher than the original model. We used confusion Matrix of logistic model to analysis accuracy of immunotype model in predicting survival in the testing and validation cohorts, which demonstrated that the overall survival prediction accuracy by immunotype H and L was more precise than that of the entire sub-group (Table 2-3). Taken together, the findings demonstrate that the immunotype H&L model has higher prediction precision. Based on immunotype H and L divergence, we re-constructed the SSIGN model (Table 4) and created a novel nomogram to improve the SSIGN model (Figure 3B).

Identifying immune genes associated with the immunotypes

To reveal potential immune genes associated with immunotype, 659 genes were selected from the TCGA based on differential gene expression in immunotype H vs immunotype L (Supplementary Table 1). Using the Forest algorithm, we selected 14 significant immune genes (SLN, GJB6, KCNJ11, SLC22A8, OBP2A, SLC22A6, SLC22A12, XCR1, MIXL1, RORB, ADGRA1, F7, AQP9, CPA4) based on an importance

of >2 (Figure 4A). Three immune genes (SLN, AQP9, CPA4) that exhibited negative expression correlation with Disease-Free Survival (DFS) were selected (Figure 4B-D). The remaining immune genes did not significantly correlate with DFS (Supplement Figure F-G). To validate the accuracy of these 3 immune genes, we evaluated the relationship between their expression and the ccRCC clinical stage (Figure 4E-G, Supplementary figure H). The results demonstrated that the expression of these 3 immune genes increased with advancing clinical stage. Additionally, IHC analysis results of the 3 genes in subgroups F1-F4 and metastatic individuals in the Renji database, were consistent with previous findings (Figure 4H-J). Collectively, these data supported the view that the 3 significant immune genes could effectively predict the OS in ccRCC.

Reconstructing the H&L risk-metastasis model

Immunotype H&L correlated with OS. As such, using immunotypes model, we analyzed TCGA data by first dividing it into immunotype H&L database. Based on the immunotype H database, a novel H-risk-metastasis model of tumor metastasis was constructed to predict metastasis risk in immunotype H database. We adopted a risk forest plot of multivariate Cox regression analysis to select key elements in immunotype H database (Figure 5A, Table 5). These 4 factors, including age, immune-score, AQP9 gene, SLN gene, were compared its importance (Figure 5B). With the risk forest plot, we used these 4 factors to construct the H-risk-metastasis model (Figure 5C, Supplementary Figure I). SLN and AQP9 were most significant than other factors, whereas age was least important compared to other factors. Next, we constructed a new L-risk-metastasis model using a risk forest plot of multivariate Cox regression analysis to predict metastasis risk of immunotype L database (Figure 5D, Table 6) and then selected 3 elements to construct an L-risk-metastasis model (Figure 5E-F, Supplementary Figure J)., which further revealed that CPA4 was most significant of the other 2 factors, whereas age was the least important than other factors. We constructed 3 types of ccRCC metastasis risk models: H-risk-metastasis model, comprising SLN, AQP9, immune score, and

age; L-risk-metastasis model, comprising SLN, CPA4, and age; and W-Risk-metastasis model based on the entire TCGA database. The W-Risk-metastasis model included T-cells, macrophages, mast cells, AQP9, SLN, CPA4, and age, without immunotype. We constructed random forest programs to validate the accuracy of the metastasis risk models. The accuracy of the H-Risk-metastasis model and L-Risk-metastasis model was 0.828 and 0.823 respectively, whereas that of the W-Risk-metastasis model was 0.709. Notably, the accuracy of H-Risk-metastasis model and L-Risk-metastasis model was higher than that of the W-Risk-metastasis model (Table 7), demonstrating the necessity and accuracy of our above immunotype model.

Discussion

ccRCC has been associated with significantly dysfunctional immune cell infiltration, which contributes to the failure of the immune system to prevent tumor growth ²⁶. However, its relationship has not been examined based on immune cells with immunotype differentiation ²⁷.

In the last two decades, few studies have examined the role of immunotypes in ccRCC prognosis, necessitating the development of a more detailed immunotype for ccRCC ²⁶. A previous study classified ccRCC patients into immunotypes I, II, and III ²⁷; however, immunotypes were based on consensus clustering to identify 3 subtypes, which may ignore immune cells in exploring the correlation of various subtypes. However, this work lacked a sufficient database to fully prove the accuracy of these 3 sub-types. Relative to previous evidence, our study relied on multi-platform databases and the Renji database ²⁸⁻³⁰. Furthermore, we used immune cells to construct immunotypes by fully taking immune cells into account.

Here, we first designed a program to construct immunotype H&L. Based on overall survival, stage, grade, and univariate Cox regression analysis, we identified mast cells, macrophages, and T-cells as leading indicators for constructing immunotype and verified their accuracy. The results demonstrated that mast cells,

macrophages, and T-cells influence tumor development by releasing several cellular factors. T-cell transfer can trigger tumor regression. Several studies reported the dominance of TAMs in patients with ccRCC, which could explain above difference ³¹. TAMs have been characterized as macrophages with distinctive negativity for CD33, with immunosuppressive properties ³². There is evidence that mast cells induce immunosuppression by releasing TNF- α and IL-10, which promote Treg-mediated immune tolerance ³³. In the Renji database, we found that the levels of T-cell and macrophages increased from F1 to metastasis while mast cells reduced. The findings demonstrated that T-cells and macrophages negatively correlate with overall survival, while mast cells correlate positively with overall survival, hence, supporting our approach of using these 3 immune cells to construct the immunotype model.

Unlike previous reports in which immunotype model H&L was based on 4 different aspects, our immunotype model based on mast cells, macrophages, and T-cells can improve the prediction of ccRCC survival. First, the immunotyping model was validated in GEO datasets, therefore is effective in distinguishing tumors with normal subtypes and can accurately identify the immune H and L subtypes. Secondly, we verified the accuracy of the immunotype model according to the Renji hospital database. The results provide convincing evidence that the 3 indicators (macrophages, mast cells, and T cells) used to construct this model can fully distinguish different immunotypes. Thirdly, using the Renji hospital database, we compared the relationship between tumor and clinical-stage or grade and found that immune score increased with advancing clinical tumor grade and stage. Thus, we strongly ascertain that immune-score is an important index for distinguishing immune H & L groups. Finally, using the TCGA database, we evaluated the OS predictive value of the immunotype. These results fully support the view that that immunotyping H & L may better predict survival rate. Taken together, immunotype is associated with overall survival, T stage, and the Fuhrman grade in RCC patients, therefore, is of great promise as a novel biomarker for RCC diagnosis and treatment.

Considering the immunotypes, our new SSIGN model differs from previously reported RCC models. Previous RCC models for evaluating OS include Karakiewicz nomogram, CCP, R-CCP, and Clearcode 34 ²⁸⁻³⁰. Karakiewicz nomogram was based on clinical factors without incorporating tumor type, while CCP, R-CCP, and Clearcode 34 were based on genes without including clinical and immune factors ²⁸⁻³⁰. Of note, the new SSIGN model includes clinical factors, such as tumor grade, stage, and metastasis, and interpolate immunotypes and immune genes. The model was further validated in 3 independent cohorts (testing cohort, training cohort, and validation cohort) from 3 databases. With this view, the new SSIGN model is valid and effective for survival prediction and treatment.

Moreover, gene expression data were analyzed to explore the potential biological phenotypes associated with the immunotype. In particular, we focused on the correlation between OS and the expression of selected immune-related genes. The results showed that expressions of SLN, AQP9, and CAP4 genes were negatively correlated with survival rate, thereby demonstrating prognostic and therapeutic potential. Upregulation of these 3 immune genes was correlated with poor survival. A previous trial showed that genes were linked with different immunotypes ³⁴⁻³⁹. Several lines of evidence demonstrated that these 3 immune-related genes are closely associated with clinical survival and recurrence ³⁸⁻⁴².

We also constructed 3 metastasis risk models based on multi-factors. Of note, the immune score was revealed to be most influential. The accuracy of the models (divided into H&L risk-metastasis group) was higher than that of the W-Risk-metastasis model. The observations further illustrate the accuracy of our immunotype model. Combined with the H & L-Risk metastasis models, the 3 immune genes play pivotal roles in tumor metastasis. Mounting evidence shows that SLN, CPA4, and AQP9 are related to tumor progression, offering a basis for developing combinational immunotherapies ^{40-46,50}. Thus, targeted therapy based on these genes may be a useful tool for early ccRCC detection and treatment ⁴⁷⁻⁴⁹. We believe that our findings have a

high potential to translate into clinical practice.

In summary, this work, has, for the first time, selected immune cells based on 3 ccRCC patient cohorts. An immune index was constructed and 2 immunotypes - H&L were revealed. This work also demonstrates that immunotypes are of promise as significantly effective prognostic and predictive tools for ccRCC patients. Also, the newly established SSIGN models offer indispensable comprehensive clinical functions for improving personalized ccRCC patient management. Additionally, the constructed H or L-Risk metastasis models show higher accuracy than other models of predicting metastasis and OS. Lastly, the 3 ccRCC-associated immune cells and genes identified, herein, highlights novel, potential, targeted therapeutic strategies.

Author Contributions

YHC, WZ, and YL conceived and designed the study. YL and ZLY performed all the experiments. WX provided support with experimental techniques. ZLY, JLZ, and YZX collected clinical data. YL performed the pathologic analyzes. YZ, KJZ, and YZ analyzed the data and prepared the manuscript. WZ reviewed and revised the manuscript. YRH, WX, JZ, and JWH provided material support and supervised all experiments. All authors read and approved the final manuscript.

Funding

This study was supported by funding from the National Natural Science Foundation of China (No. 81972369) and The Shanghai Youth Talent Support Program (W. Zhai).

References

1. Assi, T. El Rassy, E. Farhat, F., et al., Overview on the role of preoperative therapy in the management of kidney cancer. Clin Transl Oncol, 2019.

2. DiNatale, R.G. Sanchez, A. Hakimi, A. A., et al., Metabolomics informs common patterns of molecular dysfunction across histologies of renal cell carcinoma. *Urol Oncol*, 2019.
3. Wen, J.L. Wen, X. F. Li, R. B., et al., UBE3C promotes growth and metastasis of renal cell carcinoma via activating Wnt/beta-catenin pathway. *PLoS One*, 2015. 10(2): p. e0115622.
4. Mikami, S. Oya, M. Mizuno, R., et al., Invasion and metastasis of renal cell carcinoma. *Med Mol Morphol*, 2014. 47(2): p. 63-7.
5. Beutner, U. Leowardi, C. Bork, U., et al., Survival after renal cell carcinoma metastasis to the thyroid: single center experience and systematic review of the literature. *Thyroid*, 2015. 25(3): p. 314-24.
6. Chevrier, S. Levine, J. H. Zanotelli, V. R. T., et al., An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell*, 2017. 169(4): p. 736-749 e18.
7. Bialojan, A. Levine, J. H. Zanotelli, V. R. T., et al., Transcutaneous immunization with CD40 ligation boosts cytotoxic T lymphocyte mediated antitumor immunity independent of CD4 helper cells in mice. *Eur J Immunol*, 2019.
8. Bohner, P. Chevalier, M. F. Cesson, V., et al., Double Positive CD4(+)CD8(+) T Cells Are Enriched in Urological Cancers and Favor T Helper-2 Polarization. *Front Immunol*, 2019. 10: p. 622.
9. Hid Cadena, R. Reitsema, R. D. Huitema, M. G., et al., Decreased Expression of Negative Immune Checkpoint VISTA by CD4+ T Cells Facilitates T Helper 1, T Helper 17, and T Follicular Helper Lineage Differentiation in GCA. *Front Immunol*, 2019. 10: p. 1638.
10. Bedke, J. Stuhler, V. Stenzl, A., et al., Immunotherapy for kidney cancer: status quo and the future. *Curr Opin Urol*, 2018. 28(1): p. 8-14.
11. Kim, H. Wang, S. Y. Kwak, G., et al., Exosome-Guided Phenotypic Switch of M1 to M2 Macrophages for Cutaneous Wound Healing. *Adv Sci (Weinh)*, 2019. 6(20): p. 1900513.
12. Lagana, A.S. Salmeri, F. M. Ban Frangez, H., et al., Evaluation of M1 and M2 macrophages in ovarian endometriomas from women affected by endometriosis at different stages of the disease. *Gynecol Endocrinol*, 2019: p. 1-4.
13. Li, D. Ji, H. Niu, X., et al., Tumor-associated macrophages secrete CCL2 and induce tamoxifen resistance by activating PI3K/Akt/mTOR in breast cancer. *Cancer Sci*, 2019.
14. Qi, H. Gao, Y. Li, Y., et al., Induction of Inflammatory Macrophages in the Gut and Extra-Gut Tissues by Colitis-Mediated Escherichia coli. *iScience*, 2019. 21: p. 474-489.
15. Strattan, E. Palaniyandi, S. Kumari, R., et al., Mast Cells Are Mediators of Fibrosis and Effector Cell

Recruitment in Dermal Chronic Graft-vs.-Host Disease. *Front Immunol*, 2019. 10: p. 2470.

16. Yamaguchi, T. Ikeda, Y. Tashiro, K., et al., The role of galanin in the differentiation of mucosal mast cells in mice. *Eur J Immunol*, 2019.
17. Allegrezza, M.J. and J.R. Conejo-Garcia, Targeted Therapy and Immunosuppression in the Tumor Microenvironment. *Trends Cancer*, 2017. 3(1): p. 19-27.
18. Wu, T. and Y. Dai, Tumor microenvironment and therapeutic response. *Cancer Lett*, 2017. 387: p. 61-68.
19. Yang, S. and H. Gao, Nanoparticles for modulating tumor microenvironment to improve drug delivery and tumor therapy. *Pharmacol Res*, 2017. 126: p. 97-108.
20. Rogers, G.L. and P.M. Cannon, Gene Therapy Approaches to Human Immunodeficiency Virus and Other Infectious Diseases. *Hematol Oncol Clin North Am*, 2017. 31(5): p. 883-895.
21. Mendenhall, A. Crane, M. M. Tedesco, P. M., et al., *Caenorhabditis elegans* Genes Affecting Interindividual Variation in Life-span Biomarker Gene Expression. *J Gerontol A Biol Sci Med Sci*, 2017. 72(10): p. 1305-1310.
22. Pardoll, D. and C. Drake, Immunotherapy earns its spot in the ranks of cancer therapy. *J Exp Med*, 2012. 209(2): p. 201-9.
23. Zhang, Y., L. Quan, and L. Du, The 100 top-cited studies in cancer immunotherapy. *Artif Cells Nanomed Biotechnol*, 2019. 47(1): p. 2282-2292.
24. Subramanian, P. and N.B. Haas, Recent advances in localized RCC: A focus on VEGF and immuno-oncology therapies. *Urol Oncol*, 2018. 36(1): p. 23-30.
25. Newman, A.M. Liu, C. L. Green, M. R., et al., Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*, 2015. 12(5): p. 453-7.
26. Liang, J. Liu, Z. Zou, Z., et al., The Correlation Between the Immune and Epithelial-Mesenchymal Transition Signatures Suggests Potential Therapeutic Targets and Prognosis Prediction Approaches in Kidney Cancer. *Sci Rep*, 2018. 8(1): p. 6570.
27. Yu, W. Wang, Y. Jiang, Y., et al., Distinct immunophenotypes and prognostic factors in renal cell carcinoma with sarcomatoid differentiation: a systematic study of 19 immunohistochemical markers in 42 cases. *BMC Cancer*, 2017. 17(1): p. 293.
28. Bishoff, J.T. Freedland, S. J. Gerber, L., et al., Prognostic utility of the cell cycle progression score generated from biopsy in men treated with prostatectomy. *J Urol*, 2014. 192(2): p. 409-14.
29. Karakiewicz, P.I. Briganti, A. Chun, F. K., et al., Multi-institutional validation of a new renal cancer-specific

survival nomogram. *J Clin Oncol*, 2007. 25(11): p. 1316-22.

30. Brooks, S.A. Brannon, A. R. Parker, J.S., et al., ClearCode34: A Prognostic Risk Predictor for Localized Clear Cell Renal Cell Carcinoma. *European Urology*, 2014. 66(1): p. 77-84.

31. Erdag, G. Schaefer, J. T. Smolkin, M. E., et al., Immunotype and immunohistologic characteristics of tumor-infiltrating immune cells are associated with clinical outcome in metastatic melanoma. *Cancer Res*, 2012. 72(5): p. 1070-80.

32. van Dalen, F.J. van Stevendaal, Mhme. Fennemann, F. L., et al., Molecular Repolarisation of Tumour-Associated Macrophages. *Molecules*, 2018. 24(1).

33. Ribatti, D., Mast cells as therapeutic target in cancer. *Eur J Pharmacol*, 2016. 778: p. 152-7.

34. Franck, J.P.C., E. Slight-Simcoe, and N.C. Wegner, Endothermy in the smalleye opah (*Lampris incognitus*): A potential role for the uncoupling protein sarcolipin. *Comp Biochem Physiol A Mol Integr Physiol*, 2019. 233: p. 48-52.

35. Yoo, J. Kohlbrenner, E. Kim, O., et al., Enhancing atrial-specific gene expression using a calsequestrin cis-regulatory module 4 with a sarcolipin promoter. *J Gene Med*, 2018. 20(12): p. e3060.

36. Benarroch, L. Aubart, M. Gross, M. S., et al., Marfan Syndrome Variability: Investigation of the Roles of Sarcolipin and Calcium as Potential Transregulator of FBN1 Expression. *Genes (Basel)*, 2018. 9(9).

37. Voit, A. Patel, V. Pachon, R., et al., Reducing sarcolipin expression mitigates Duchenne muscular dystrophy and associated cardiomyopathy in mice. *Nat Commun*, 2017. 8(1): p. 1068.

38. Nie, L. Yuan, X. L. Jiang, K. T., et al., Salsalate Activates Skeletal Muscle Thermogenesis and Protects Mice from High-Fat Diet Induced Metabolic Dysfunction. *EBioMedicine*, 2017. 23: p. 136-145.

39. Huang, X. Yu, X. Li, H., et al., Regulation mechanism of aquaporin 9 gene on inflammatory response and cardiac function in rats with myocardial infarction through extracellular signal-regulated kinase1/2 pathway. *Heart Vessels*, 2019. 34(12): p. 2041-2051.

40. Bestetti, S. Galli, M. Sorrentino, I., et al., Human aquaporin-11 guarantees efficient transport of H₂O₂ across the endoplasmic reticulum membrane. *Redox Biol*, 2019. 28: p. 101326.

41. Liu, X. Gong, J F. Wang, L G., et al., Genome-Wide Profiling of the MicroRNA Transcriptome Regulatory Network to Identify Putative Candidate Genes Associated with Backfat Deposition in Pigs. *Animals*, 2019. 9(6).

42. Cavariani, M.M. de Mello Santos, T. Pereira, D. N., et al., Maternal protein restriction differentially alters the expression of AQP1, AQP9 and VEGFr-2 in the epididymis of rat offspring. *Int J Mol Sci*, 2019. 20(3).

43. Pan, H. Pan, J. Ji, L., et al., Carboxypeptidase A4 promotes cell growth via activating STAT3 and ERK

signaling pathways and predicts a poor prognosis in colorectal cancer. *Int J Biol Macromol*, 2019. 138: p. 125-134.

44. Zhang, H. Hao, C. Wang, H., et al., Carboxypeptidase A4 promotes proliferation and stem cell characteristics of hepatocellular carcinoma. *Int J Exp Pathol*, 2019. 100(2): p. 133-138.

45. Weisheit, C.K., D.R. Engel, and C. Kurts, Dendritic Cells and Macrophages: Sentinels in the Kidney. *Clin J Am Soc Nephrol*, 2015. 10(10): p. 1841-51.

46. Peng, H. Qin, C. Zhang, C., et al., circCPA4 acts as a prognostic factor and regulates the proliferation and metastasis of glioma. *J Cell Mol Med*, 2019. 23(10): p. 6658-6665.

47. Fu, Y. Su, L. Cai, M., et al., Downregulation of CPA4 inhibits non small-cell lung cancer growth by suppressing the AKT/c-MYC pathway. *Mol Carcinog*, 2019. 58(11): p. 2026-2039.

48. Wu, J. Qu, J. Cao, H., et al., Monoclonal antibody AC10364 inhibits cell proliferation in 5-fluorouracil resistant hepatocellular carcinoma via apoptotic pathways. *Onco Targets Ther*, 2019. 12: p. 5053-5067.

49. Liu, K., Breakthroughs in cancer gene therapy. *Seminars in Oncology Nursing*, 2003. 19(3): p. 217-226.

50. Xu WH, Shi SN, Xu Y et al., Prognostic implications of Aquaporin 9 expression in clear cell renal cell carcinoma. *J Transl Med*. 2019 Nov 8;17(1):363.

Key of Definitions for Abbreviations

Renal clear cell carcinoma (ccRCC)

The Cancer Genome Atlas (TCGA)

Gene Expression Omnibus (GEO)

Metastatic renal cell carcinoma (mRCC)

Overall survival (OS)

Immunohistochemistry (IHC)

Tyrosine kinase inhibitors (TKIs)

Microenvironment (TME)

Renal cell carcinoma (RCC)

Tumor-associated macrophages (TAMs)

Kaplan-Meier survival curves (KM survival curves)

Figure legends

Figure 1. Identification of immune cells. (a) K-M survival curve of four immune cells. (b) The expression of immune cells in different grades. (c) The expression of immune cells in different stages. (d) The IHC of three immune cells in the F1-F4 subgroups and individuals with metastases.

Figure 2. Construction of an immunotype model. (a) The LASSO Cox regression models were constructed based on the coefficients of immune cells. Based on II, immunotype H was defined as immune index ≥ 0.2145 and immunotype L was defined as immune index < 0.2145 . (b) The survival rate of immunotype H and immunotype L. (c) The immune scores of 3 immune cells in the tumor and normal tissues. (d) The immune scores of 3 immune cells in immunotype H and immunotype L. (e) Immunescore of different immunotype subgroups in relation to tumor grade and stage. (f) Correlation among the 3 immune cells in the validation cohort. (g) The survival rate of patients in immunotype H and immunotype L based on the RECA-EU database.

Figure 3. Construction of a novel SSIGN model. (a) ROC curve for training cohort and the survival rate. (b) Construction of the SSIGN model.

Figure 4. Identification of immunotype-associated genes. (a) A Forest Algorithm identified 3 immune-related genes. (b-d) The expression and DFS of these 3 genes. (e-g) The relationship between expression levels of the 3 genes and Furman stage. (h-j) Results of IHC assay for the 3 immune genes in the F1-F4 subgroups and patients with metastases.

Figure 5. Construction of the risk-metastasis model. (a) The H-risk-metastasis model. (b) The four factors used to assess the performance of the model. (c) Random forest models constructed from the four factors. (d) A risk forest plot of multivariate Cox regression analysis performed for the L-Risk-metastasis model. (e) Three factors used to assess the performance of the model. (f) Random forest models constructed based on 3 factors.

Table 1 Five immune cells (T cells regulatory, Mast cells, Macrophages M2, T cells CD4 memory, T cells follicular helperactivated) selected by COX regression in 539 patients with ccRCC.

Table 2 The new SSIGN model.

Table 3 The Confusion Matrix logistic model used to assess the survival rate of patients in the Testing cohort.

Table 4 The Confusion Matrix logistic model used to evaluate the survival rate of patients in the Validation cohort.

Table 5 Multivariate Cox regression analysis of immunotype H database

Table 6 Multivariate Cox regression analysis of immunotype L database

Table 7 Comparison of the 3 models

Supplementary Figure A

K-M survival curve of four immune cells

Supplementary Figure B

The expression of immune cells at different grades.

Supplementary figure C

Expression of Immune cells at different grades.

Supplementary Figure D

Expression of immune cells at different stages.

Supplementary figure E

Expression of immune cells at different stages.

Supplementary figure F

Expression of 11 genes; GJB6, KCNJ11, SLC22A8, OBP2A, SLN22A6, SLN22A12, XCR1, MIXL, RORB, ADGRV1, and F7 in normal samples and samples from ccRCC patients.

Supplementary figure G

The DFS of 11 genes: GJB6, KCNJ11, SLC22A8, OBP2A, SLN22A6, SLN22A12, XCR1, MIXL, RORB, ADGRV1, and F7.

Supplementary figure H

Association of expression of 11 genes; GJB6, KCNJ11, SLC22A8, OBP2A, SLN22A6, SLN22A12, XCR1, MIXL, RORB, ADGRV1, and F7 with the Furman stage.

Supplementary figure I

Random forest models for assessment of the H-Risk-metastasis model constructed from 4 factors

Supplementary figure J

Random forest model for evaluation of the L-Risk-metastasis model constructed from 3 factors

Supplementary Table 1

Comparison of genes between immunotype H and immunotype L.

Figures

Figure 1

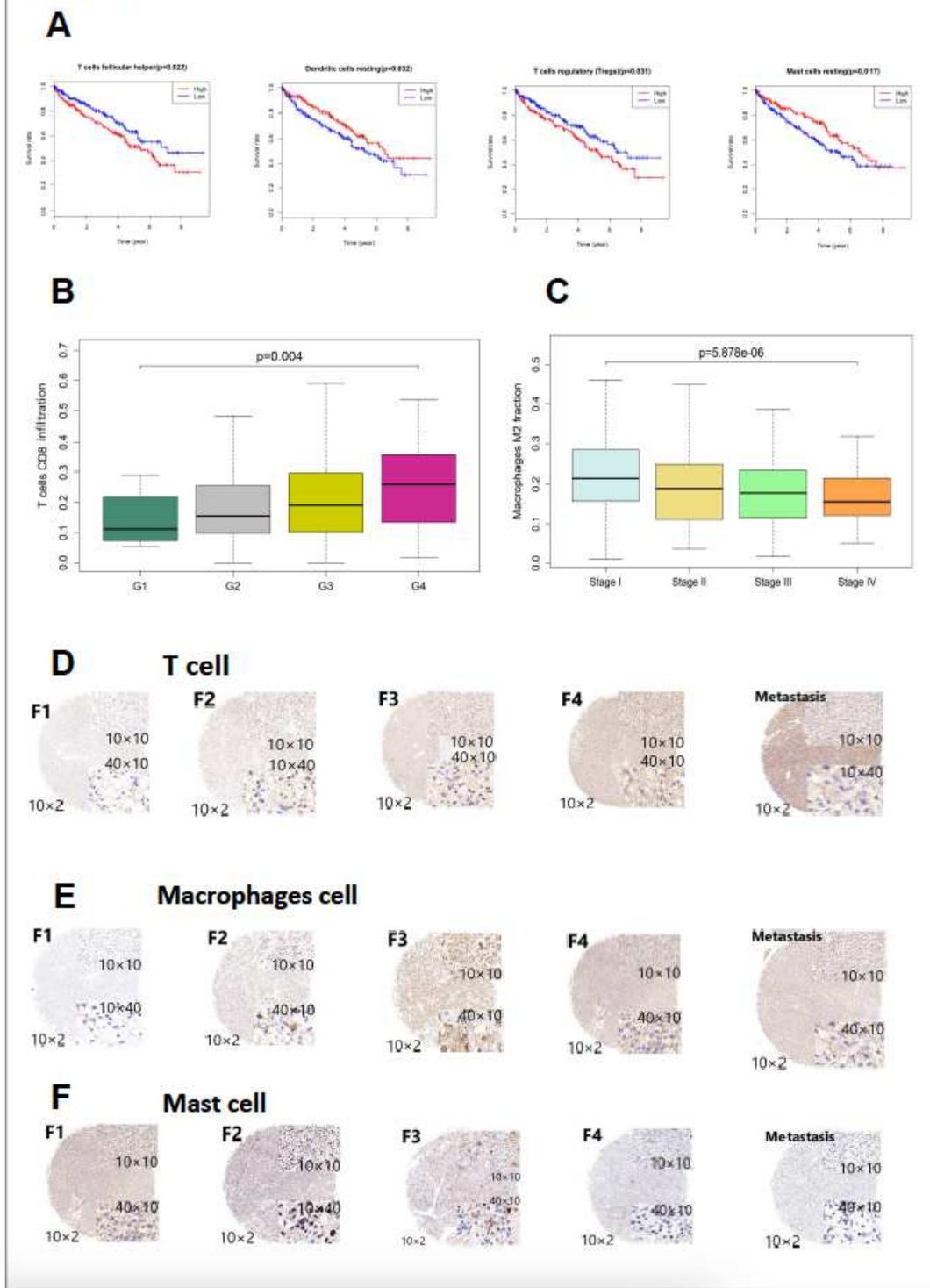


Figure 1

Identification of immune cells. (a) K-M survival curve of four immune cells. (b) The expression of immune cells in different grades. (c) The expression of immune cells in different stages. (d) The IHC of three immune cells in the F1-F4 subgroups and individuals with metastases.

Figure 2

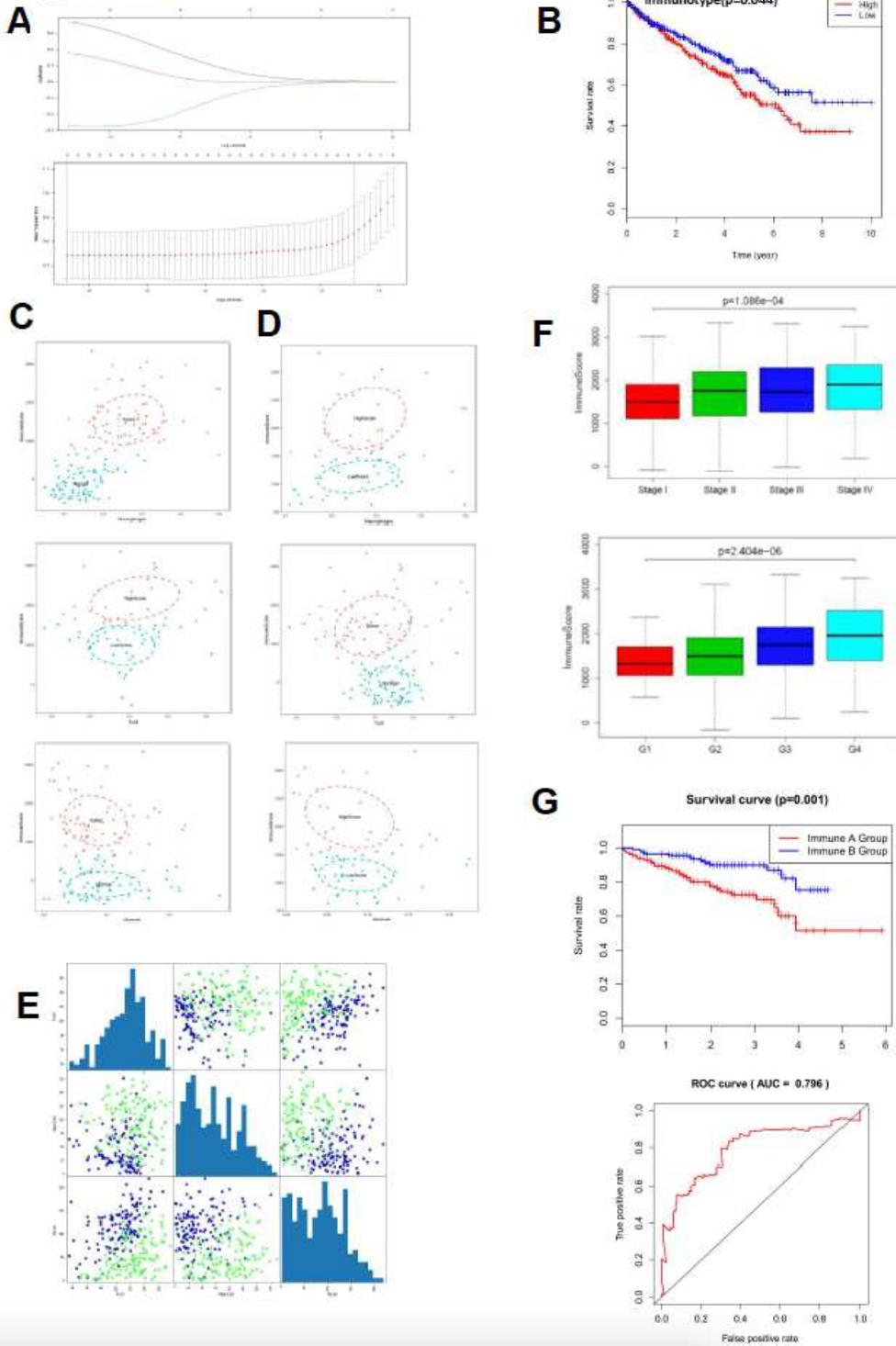


Figure 2

Construction of an immunotype model. (a) The LASSO Cox regression models were constructed based on the coefficients of immune cells. Based on it, immunotype H was defined as immune index ≥ 0.2145 and immunotype L was defined as immune index < 0.2145 . (b) The survival rate of immunotype H and immunotype L. (c) The immune scores of 3 immune cells in the tumor and normal tissues. (d) The immune scores of 3 immune cells in immunotype H and immunotype L. (e) Immunescore of different

immunotype subgroups in relation to tumor grade and stage. (f) Correlation among the 3 immune cells in the validation cohort. (g) The survival rate of patients in immunotype H and immunotype L based on the RECA-EU database.

Figure 3

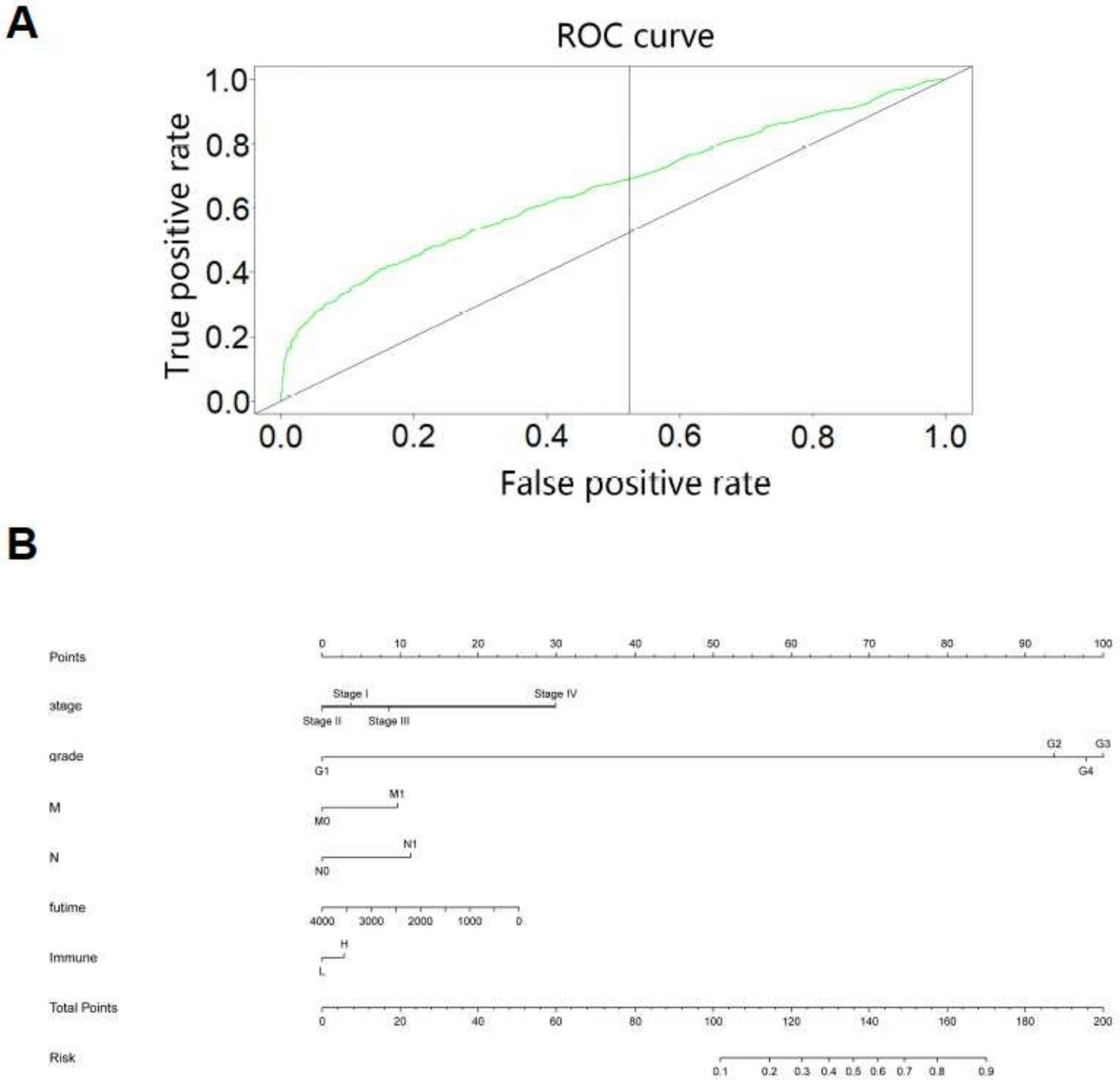


Figure 3

Construction of a novel SSIGN model. (a) ROC curve for training cohort and the survival rate. (b) Construction of the SSIGN model.

Figure 4

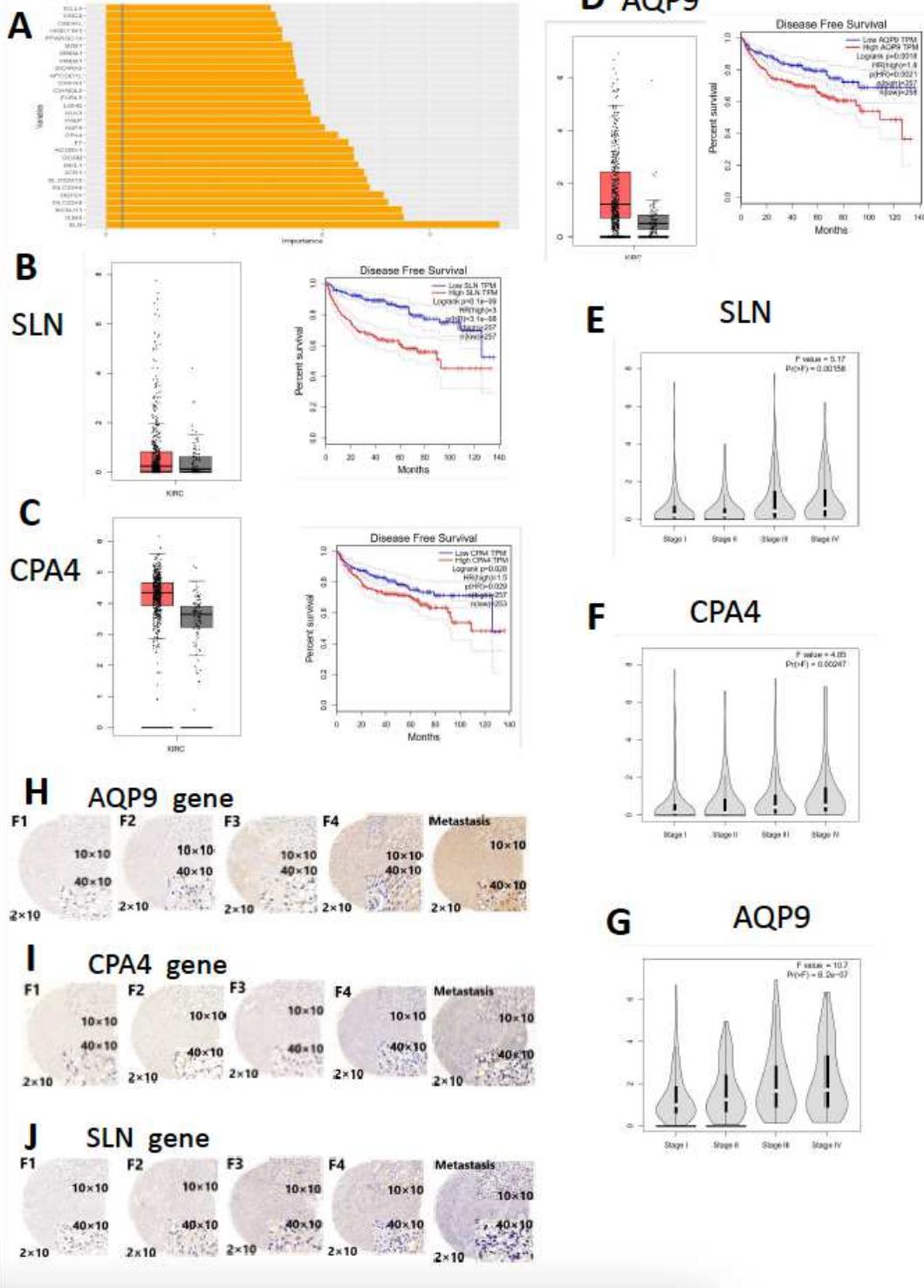
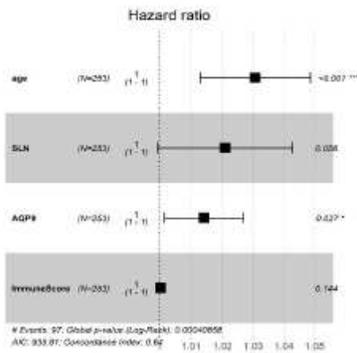


Figure 4

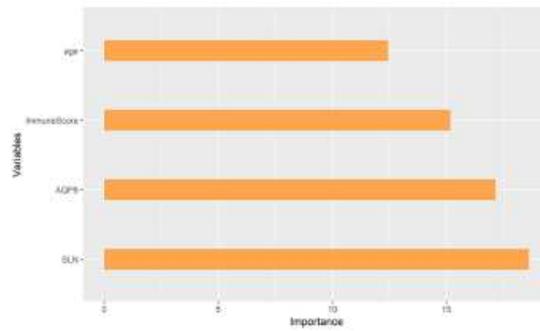
Identification of immunotype-associated genes. (a) A Forest Algorithm identified 3 immune-related genes. (b-d) The expression and DFS of these 3 genes. (e-g) The relationship between expression levels of the 3 genes and Furman stage. (h-j) Results of IHC assay for the 3 immune genes in the F1-F4 subgroups and patients with metastases.

Figure 5

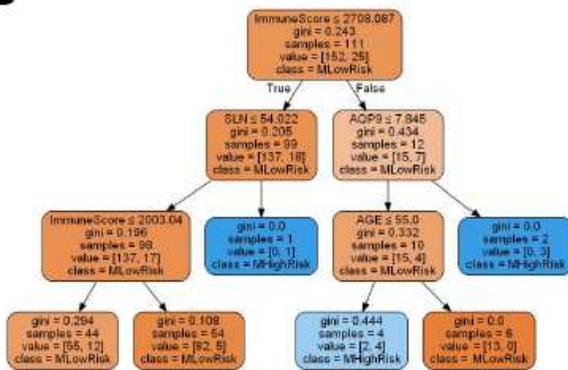
A



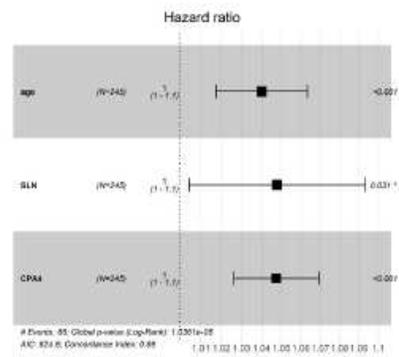
B



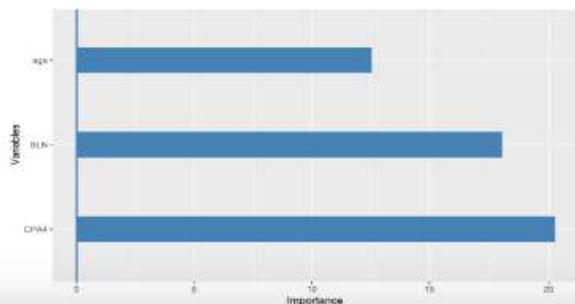
C



D



E



F

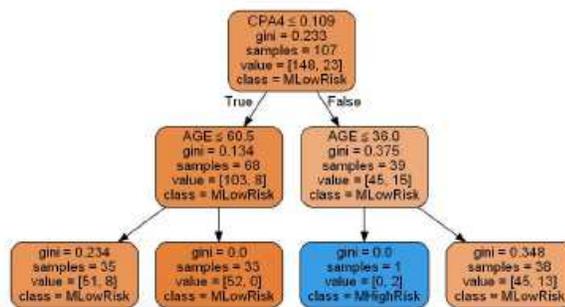


Figure 5

Construction of the risk-metastasis model. (a) The H-risk-metastasis model. (b) The four factors used to assess the performance of the model. (c) Random forest models constructed from the four factors. (d) A risk forest plot of multivariate Cox regression analysis performed for the L-Risk-metastasis model. (e) Three factors used to assess the performance of the model. (f) Random forest models constructed based on 3 factors.

Supplementary Files

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

- [Sullementaryfigure.pdf](#)
- [Supplementarytable.doc](#)