

## Original Article

# Development of a novel five-gene immune-related risk model for the prognosis evaluation of prostate adenocarcinoma patients

Hongjun Fei, Xiongming Chen

*Department of Reproductive Genetics, International Peace Maternity and Child Health Hospital, Shanghai Key Laboratory of Embryo Original Diseases, Shanghai Municipal Key Clinical Specialty, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China*

Received April 8, 2022; Accepted April 30, 2022; Epub May 15, 2022; Published May 30, 2022

**Abstract:** The interaction between the immune cells and the host immune system with the tumor cells is significantly associated with the initiation and progression of prostate adenocarcinoma (PRAD), whereas the application of immune-related genes (IRGs) for the prognosis evaluation of PRAD patients is still lacking. In this study, we aimed to identify IRGs with prognostic values and to develop a clinically effective risk model. Wilcoxon rank-sum test and univariate Cox analysis were applied to identify the differentially expressed immune-related genes (DEIRGs) related to the survival of PRAD patients. The Least absolute shrinkage and selection operator (LASSO) analysis was performed to identify the independent prognostic DEIRGs and to establish an immune risk score prognostic model. The reliability and veracity of the prognostic model were validated in PRAD patients from the internal cohort (The Cancer Genome Atlas, TCGA dataset) and the external cohort (International Cancer Genome Consortium, ICGC dataset), respectively. Six of the 193 identified DEIRGs were survival-associated in PRAD patients. Five prognostic DEIRGs (*SLPI*, *NOX1*, *DES*, *BIRC5* and *AMH*) were selected to construct the immune-related prognostic model with optimal robustness. In the 2 independent cohorts we chose, PRAD patients could be effectively stratified according to our risk model. Patients with high risk scores had worse survival. Clinical correlation analysis proved that the risk score was associated with advanced clinicopathologic features. Multivariate analysis indicated that the risk model was an independent prognostic indicator. We also established a nomogram based on the risk score model for clinical application. Additionally, the risk score model was correlated with immune cell infiltration and reflected the status of the immune microenvironment. The prognostic value of the five immune-related genes used in the prognostic model was also validated. Our immune-related prognostic model was an effective tool that could not only serve as a predictor for prognosis, but also provide potential prognostic and therapeutic molecular biomarkers for optimizing personalized therapies in clinical practice.

**Keywords:** Prostate adenocarcinoma, immune-related prognostic signature, immune-related genes, prognosis evaluation, risk score

## Introduction

Prostate adenocarcinoma (PRAD) is the most common type of prostate cancer and the second leading cause of cancer related death in men, accounting for 26% of diagnosis and 11% of death in men [1]. According to the estimate by American Cancer Society, 248,530 new cases and 34,130 associated deaths of PRAD were projected in the United States in 2021 [2]. There is an urgency to explore the biomarkers for effective early detection and accu-

rate prognosis evaluation [3]. Most PRAD is slow-growing, while other PRAD progresses quickly; they were considered as indolent or potentially lethal PRAD, respectively [4]. Accurately distinguishing between indolent PRAD and potentially lethal PRAD will assist clinician in selecting the most appropriate treatment, not only to avoid excessive treatment to patients predicted to have favorable outcomes, but also to provide more effective targeted treatment and surveillance to patients with a poor prognosis [5-7].

## Immune-related signature for PRAD patients

Traditionally, clinicopathologic indicators such as histological grade and TNM stage are used to stratify PRAD patients, to evaluate the risk of progression, and to provide the guidance for treatment [8]. However, prostate-specific antigen (PSA) screening is more commonly used currently to provide prognostic information, which even causes a decline in the prognostic assessment power of the traditional clinicopathologic indicators [9, 10]. Although PSA has the advantage of sensibility, it is compromised by its low specificity [11]. To identify the molecular prognostic determinants of PRAD for better prognosis prediction and treatment selection, extensive research has been conducted [12, 13]. However, identifying the reliable molecular biomarkers to predict the prognostic outcome remains challenging, which requires the cumulative efforts from diverse research fields.

Recent studies have indicated that the interaction between the immune cells and the host immune system with the tumor cells is significantly associated with the initiation and progression of PRAD [14]. Indeed, tumor-associated immune cells have already been explored in the immunotherapy of PRAD [15, 16], and the manipulation of the immune system to boost its ability is undoubtedly a promising therapy for PRAD treatment [17, 18]. Therefore, exploring the association between the immune molecular features and the prognosis of PRAD patients, and then constructing the diagnostic and prognostic immune characteristics are highly significant for the early detection and individual management of PRAD patients.

In this study, we attempted to incorporate the molecular features of the immune system into the prognostics of PRAD patients, to improve the accuracy of prognostic prediction, to increase the management efficiency of PRAD patients, and to facilitate targeted therapy. Hence, the immune-related genes (IRGs) with prognostic values were identified and used to establish an immune-related prognostic signature for PRAD patients. The clinical availability, reliability and potential value of the immune-related prognostic signature were detailed and assessed. We found that this immune-related prognostic signature could be used for prognosis prediction and to stratify patients for personalized management and targeted therapy.

These promising findings offered valuable insights into the role of IRGs in PRAD and verified the importance of immune-related prognostic signatures in the prognosis evaluation of PRAD patients.

### Methods

#### *Data sources*

The transcriptome sequencing data and the corresponding clinical information of 499 PRAD and 52 adjacent normal samples were obtained from the TCGA data portal (<https://portal.gdc.cancer.gov>) and used as the training cohort. Another RNA-Seq dataset containing 144 PRAD patients and the corresponding survival information was downloaded from the ICGC portal (<https://dcc.icgc.org/>) and used as an external validation cohort for the risk model. The samples of PRAD patients without clinical follow-up information in the TCGA training cohort and without survival information in the ICGC validation cohort were removed, and finally, 495 PRAD samples from TCGA and 137 PRAD samples from ICGC were included. Altogether 2498 immune-related genes (IRGs) that contained 17 immune categories were obtained from the ImmPort database (<https://www.ImmPort.org/home>).

#### *Screening DEIRGs and performing functional enrichment analyses*

The differentially expressed IRGs (DEIRGs) between PRAD and the adjacent tissues in the TCGA training cohort were detected using the R package of “limma” with a significant cut-off value of  $|\log_2\text{FoldChange}|$  ( $|\log_2\text{FC}|$ )  $>1$  and adjusted  $P < 0.05$ . To assess the possible biological functions and the associated molecular pathways of DEIRGs, Gene ontology (GO) and pathway enrichment analysis were performed by R package of “cluster Profiler”, “enrich plot” and “GO plot”. Next, a protein-protein interaction (PPI) network of DEIRGs was constructed based on the Search Tool for the Retrieval of Interacting Genes (STRING) website (<https://string-db.org/>).

#### *Establishment of the prognostic risk model based on DEIRGs with a prognostic role*

The PRAD patients from the TCGA cohort were utilized as a training set to establish the im-

## Immune-related signature for PRAD patients

immune-related prognostic model. The prognostic value of DEIRGs in PRAD patients was investigated using a univariate Cox model. Only those DEIRGs with a  $P$ -value  $<0.05$  were found to be statistically significant and were considered prognostic DEIRGs. Then, the relationship between the expression of the prognostic DEIRGs and the overall survival (OS) was further evaluated by the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm for variable selection and optimization of the model. After that, the possibility of overfitting was minimized, and highly related genes were deleted. Finally, the benefit model for prognosis evaluation was established based on the identified prognostic DEIRGs, and the risk score of the immune related signatures for each patient was calculated using the formula:

$$\text{The risk score} = \sum_{i=1,2,\dots,n} \text{regression coefficient } t(\text{gene}_i) \times \text{expression value of } (\text{gene}_i)$$

The regression coefficient was obtained from the LASSO regression analysis. The PRAD patients were divided into two risk groups (high/low-risk group) by the median value of the risk score. We also established a nomogram using the identified prognostic DEIRGs in the risk model to visualize the risk score and the survival rate of PRAD patients. The graph was drawn by “rms” and “survival” packages in R.

### *Internal and external validation of the prognostic risk model*

The training cohort (495 PRAD patients) from the TCGA database and the validation cohort (137 PRAD patients) from the ICGC database were applied to validate the prognostic performance of the immune-related prognostic model separately. The risk scores of the PRAD patients from these two cohorts were calculated according to the immune-related risk model and were used to divide patients into high or low-risk groups. Firstly, principal component analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE) were performed to explore the distribution of the PRAD patients in different risk groups by “stats” and “Rtsne” R packages, respectively. Then, we plotted the Kaplan-Meier (K-M) survival curve and conducted the log-rank test using R packages of “survival” to compare OS differences between these two risk groups. Finally, the scatter plots of risk scores and the survival status of PRAD patients, and a heatmap of risk genes used in

the risk model were plotted to visualize the corresponding relationship between the risk score and the survival status to evaluate the performance of the immune-related prognostic model.

### *Evaluation of the independent prognostic value of the immune-related risk model*

The clinicopathological characteristics of the PRAD patients in the TCGA training cohort were extracted. The receiver operating characteristic (ROC) curves of the clinicopathological characteristics and the risk model were drawn by R packages of “survivalROC” to evaluate their performance on prognostic prediction. The higher the area under the curve (AUC) value of the ROC curve, the better the prognostic performance. The independent predictive power of our risk model compared with other clinicopathological indicators was further assessed by univariate and multivariate proportional hazards regression analyses and presented through forest plots. Meanwhile, the correlation between our risk model and the clinicopathological indicators was also evaluated via the R package of “beeswarm”. The clinicopathological indicators of age, T classification and N classification were used in the above analyses.

### *Association analysis between the immune-related model and the immune cell infiltration*

The calculation of the abundance of 6 types of tumor infiltrating immune cells (macrophages, neutrophils, CD4+ T cells, CD8+ T cells, B cells, and dendritic cells) in PRAD patients was performed using an online portal called Tumor IMMune Estimation Resource database (TIMER, <https://cistrome.shinyapps.io/timer/>), and the association of our risk model with the abundance of the infiltrating immune cells was determined by a Pearson's correlation test in R.

### *Validation of the prognostic value of every DEIRG used in the immune-related risk model*

At the end, five optimal DEIRGs were used in our prognostic model. Regression coefficient  $>0$  was considered high risk gene, and regression coefficient  $<0$  was low risk gene. To determine whether the high-risk genes were upregulated in PRAD samples than in normal tissues, and the low-risk genes had opposite expression

trend, the expression of the 5 DEIRGs was analyzed, and the correlation between the expression of these 5 DEIRGs and OS of PRAD patients was also evaluated. The protein level of the 5 prognostic related DEIRGs on prostate and PRAD tissue was adopted from The Human Protein Atlas (<http://www.proteinatlas.org>).

### Results

#### *DEIRGs identification and the interaction network as well as the function enrichment analyses*

In total, 2969 genes (1759 upregulated and 1210 downregulated) were differentially expressed between 499 PRAD samples and 52 normal samples (**Figure 1A, 1B**). Among these 2969 differentially expressed genes, 193 immune-related genes (IRGs) were identified as differentially expressed IRGs (DEIRGs), including 77 upregulated and 116 downregulated genes (**Figure 1C, 1D**).

Protein-protein interaction (PPI) network was constructed for all DEIRGs to visualize their interactions (**Figure 2A**). GO analysis revealed that “leukocyte migration” was one of the primary biological processes (BP) that DEIRGs were involved in. In addition, the most enriched molecular function (MF) of all DEIRGs was “receptor ligand activity”, and the major enriched cellular component (CC) was “external side of the plasma membrane” (**Figure 2B**). KEGG pathway analysis indicated that the pathway DEIRGs were primarily enriched in was “cytokine-cytokine receptor interaction” (**Figure 2C**).

#### *Screening of DEIRGs for prognosis value and construction of the prognostic risk model*

A total of 495 PRAD patients from the TCGA cohort were adopted as the training set. To determine the association between DEIRGs with the prognostic characteristics, a univariate Cox regression analysis was performed, and 6 DEIRGs were identified to be associated with PRAD patients' survival. Among the 6 prognostic DEIRGs, 3 were low risk genes with the function of prognostic protection ( $HR < 1$ ), and 3 were high risk genes with prognostic risk ( $HR > 1$ ) (**Figure 3A**). To avoid overfitting, LASSO penalized Cox regression analysis of these 6 prognostic DEIRGs was further con-

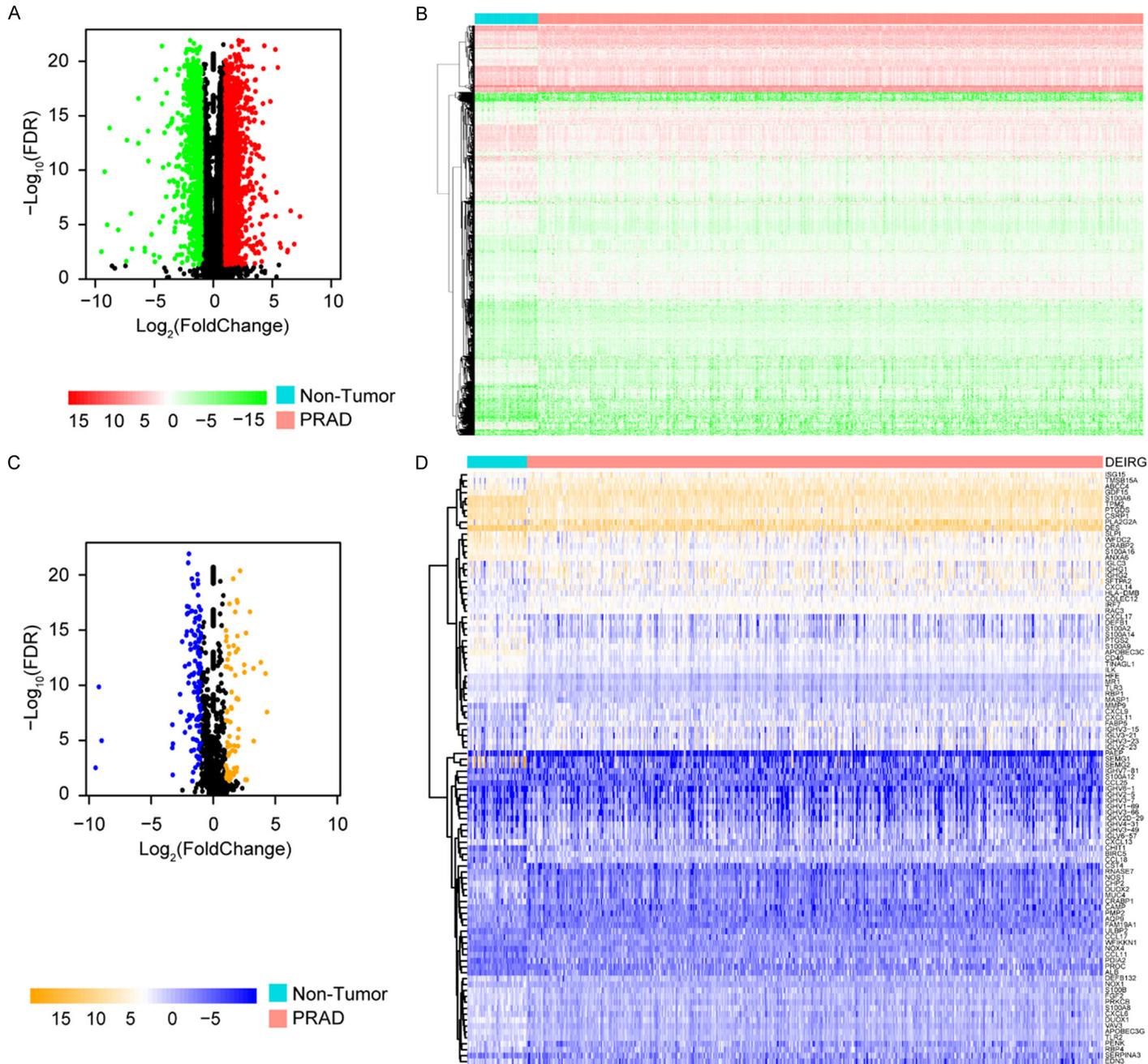
ducted, and 5 candidate genes (*SLPI*, *NOX1*, *DES*, *BIRC5* and *AMH*) were selected to construct the immune-related prognostic model with optimal robustness (**Figure 3B, 3C**). The risk score of each PRAD patient was calculated by the expression level of the 5 candidate genes multiplied by its corresponding regression coefficients. The regression coefficients of the 5 candidate genes used in the prognostic model were listed in **Table 1**. The prognostic risk model was then used for risk scoring and the prognostic evaluation of the PRAD patients in 2 separate cohorts: the entire TCGA cohort and the independent ICGC cohort. The median risk score in the training set was used to stratify patients into high and low-risk groups. For quantitative prognosis evaluation in clinical usage, we built a prognostic nomogram model according to the immune-related risk model. The total points in the nomogram corresponded to the survival probability of PRAD patients (**Figure 3D**).

#### *Validation of the immune related risk model in two independent cohorts*

Based on the median risk score, 495 PRAD samples in the entire TCGA cohort were assigned to the low ( $n=248$ ) and high- ( $n=247$ ) risk groups. PCA plot and the t-SNE plot exhibited that patients in different risk groups were distributed in discrete directions (**Figure 4A**). Kaplan-Meier analysis proved that compared with the high-risk group, PRAD patients in the low-risk group showed remarkably better OS ( $P=1.237e-03$ ) (**Figure 4B**). The risk scores, survival overview of the 495 PRAD patients, and the heatmap of the 5 risk genes used in the risk model were presented in **Figure 4C-E**. The results showed that all patients who died were in the high-risk group, and in contrast to low-risk genes, high-risk genes were upregulated in the high-risk group.

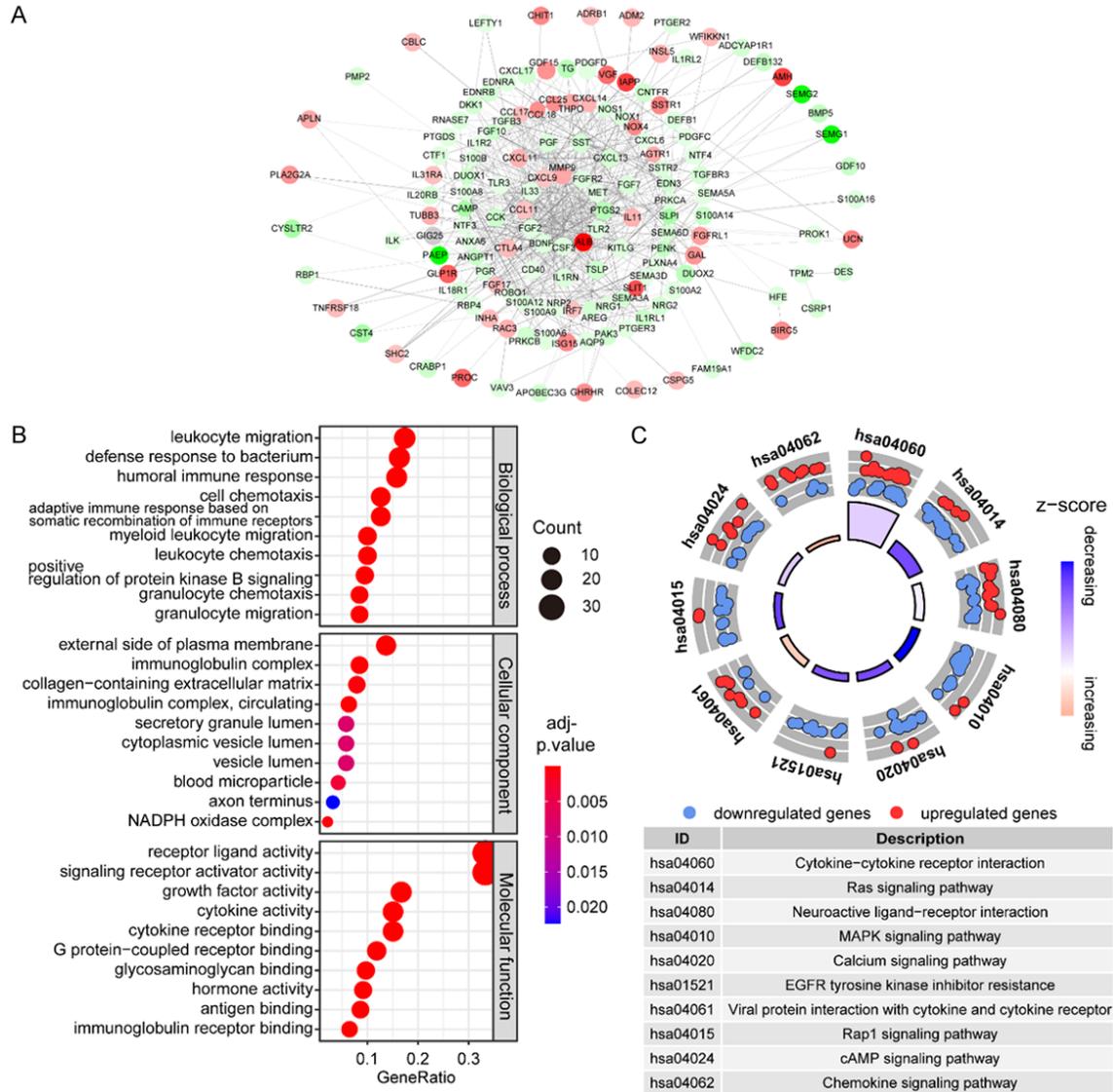
Furthermore, we used an external independent cohort from ICGC to assess the reliability and reproducibility of the immune-related risk model. The ICGC database including 137 samples was used to validate the performance of the risk model. These 137 samples were divided into the low-risk group ( $n=27$ ) and the high-risk group ( $n=110$ ) according to the same cut-off value from the training cohort. As expected, PCA and the t-SNE analyses found that PRAD

# Immune-related signature for PRAD patients



## Immune-related signature for PRAD patients

**Figure 1.** Identification of the differentially expressed immune-related genes (DEIRGs) in PRAD samples. Volcano plot (A) and heatmap (B) of significantly differentially expressed genes (DEGs). Red and green represented higher expression and lower expression in PRAD tissues, respectively. Volcano plot (C) and heatmap (D) of significantly differentially expressed immune-related genes (DEIRGs). Orange and blue represented higher expression and lower expression, respectively.

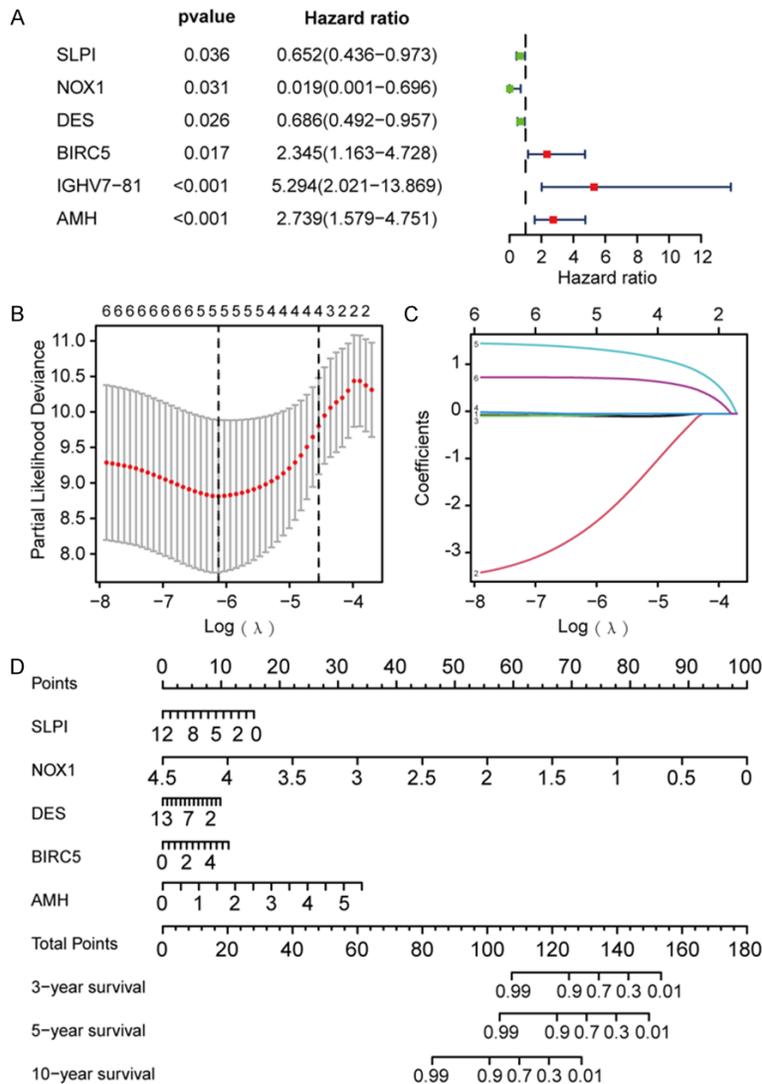


**Figure 2.** Interactions and the functional enrichment analysis of 193 DEIRGs. A. Protein-protein interaction (PPI) network encoded by 193 DEIRGs. Proteins in the most inner circle represented those with more than 20 interactive proteins. Red nodes denoted up-regulated genes, while green nodes denoted down-regulated genes. The size and color depth of nodes were associated with  $P$ -value and  $\log_2$ FoldChange. The larger the combined score of protein interaction, the heavier the lines. B. Gene ontology analysis of 193 DEIRGs. C. KEGG pathway analysis revealed the most significantly enriched pathways that 193 DEIRGs were associated with.

patients in different risk groups could be well separated into two clusters (Figure 4F). Likewise, the K-M curve demonstrated that the PRAD patients in the high-risk group exhibited a markedly reduced survival time than those in

the low-risk group ( $P=7.012e-01$ ) (Figure 4G). Figure 4H, 4I showed that the patients in the low-risk group were mostly alive compared to the high-risk group. The expression profile of these 5 risk genes between these two risk

## Immune-related signature for PRAD patients



**Figure 3.** Identification of prognostic DEIRGs and construction of an immune-related risk score model. **A.** Univariate analysis to determine the prognosis related DEIRGs in the TCGA training set. Genes with a hazard ratio (HR) >1 were regarded as high-risk genes, while HR<1 denoted low-risk genes. **B.** Cross-validation for optimal parameter selection in the LASSO regression. **C.** LASSO regression analysis of 6 prognostic DEIRGs. We identified the 5 best DEIRGs with prognostic values for risk score model establishment. **D.** Nomogram for survival prediction of PRAD patients based on the risk score model.

**Table 1.** Coefficients of the 5 independent key prognostic immune-related genes (IRGs) that formed the risk model

IRGs, immune-related genes	coefficients
SLPI	-0.1433
NOX1	-1.4010
DES	-0.0845
BIRC5	0.1243
AMH	0.7946

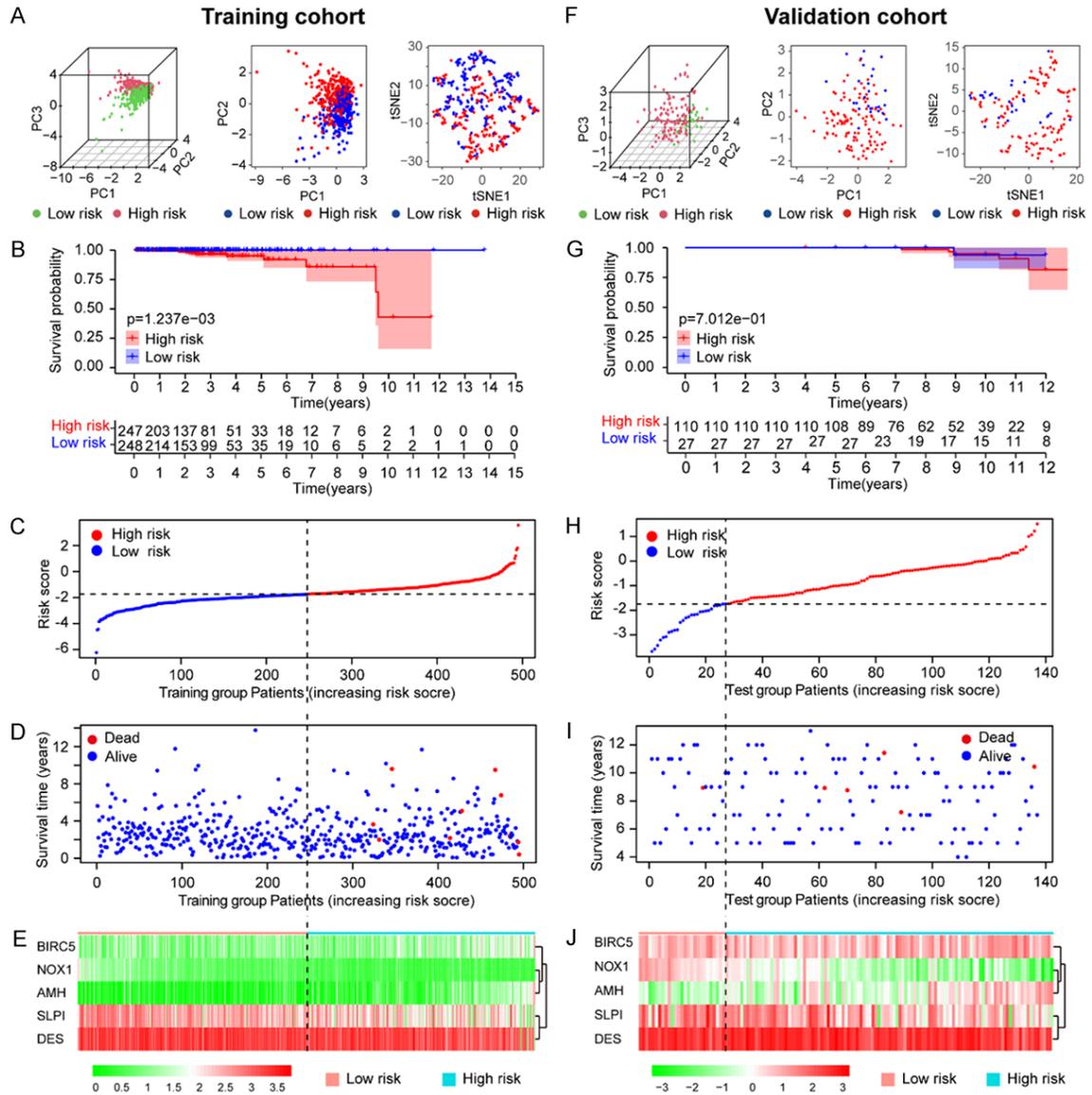
groups was drawn, and the results were the same as in the TCGA cohort (**Figure 4J**).

Together, our internal and external validation results showed that patients with higher risk scores from these 2 independent cohorts had a worse prognosis, which proved that our prognostic model was generally applicable and had high reproducibility for the prognosis evaluation of PRAD patients.

### *Clinical relevance and the independent predictive power of the immune-related risk model*

To further clarify the relevance between the risk model and the prognosis, 495 PRAD patients with clinical information from the TCGA cohort were included for further analysis. We assessed the performance of the risk model for prognosis prediction by operating a ROC curve, and the AUCs of risk score, age, T classification and N classification were 0.799, 0.519, 0.571 and 0.560 respectively, demonstrating the favorable prognostic efficiency of the risk score model (**Figure 5A**). We also analyzed the correlation between the risk score and the clinicopathological features and found patients in the high-risk group generally had advanced tumor phenotype ( $P<0.05$ ) (**Figure 5B**). Subsequently, to further assess the independent predictive power of the risk model, we conducted univariate (**Figure 5C**, left panel) and multivariate Cox regression (**Figure 5C**, right panel) analyses, and the results indicated that only the risk score model was an independent prognostic factor ( $P<0.05$ ). All these results demonstrated the outstanding predictive values of our risk score model for prognosis evaluation.

# Immune-related signature for PRAD patients



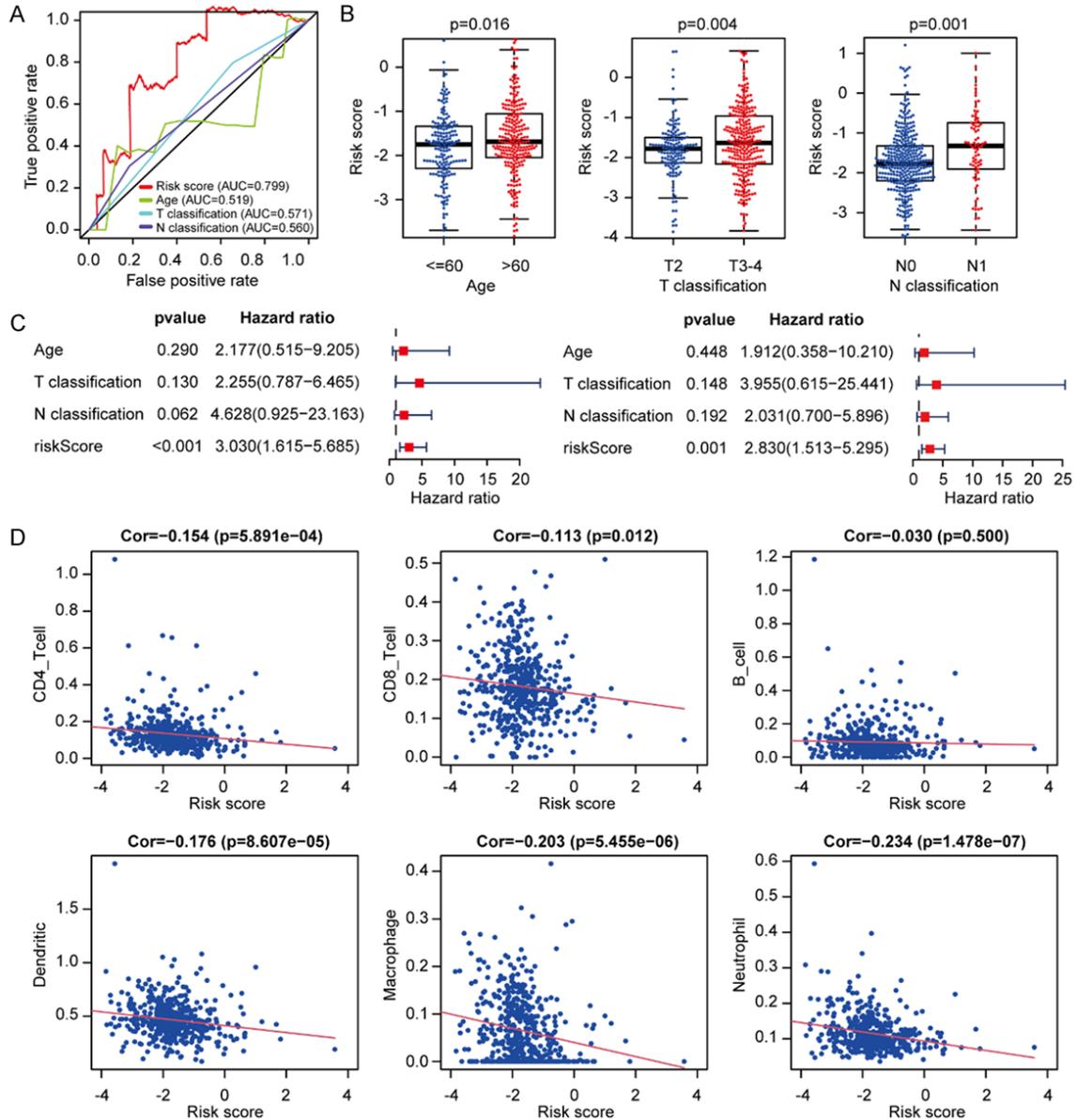
**Figure 4.** Validation of the immune-related risk score model in the TCGA training set and the ICGC validation set. Principal component analysis (PCA) and the t-SNE analysis of PRAD patients in the training cohort (A) and validation cohort (F). In these two independent cohorts, patients in different risk groups were distributed in two discrete directions. Survival analysis of PRAD patients in different risk groups in the training cohort (B) and validation cohort (G). In two independent cohorts, compared with patients in the low-risk group, high-risk group patients showed significantly shorter survival time ( $P < 0.05$ ). The distribution of risk scores and the corresponding survival time of PRAD patients in the training cohort (C, D) and validation cohort (H, I). In two independent cohorts, higher mortality rate was found in high-risk group. Heatmap of the five risk genes that were used in the risk model in the training cohort (E) and validation cohort (J). In two independent cohorts, in contrast to low-risk genes, the expressions of high-risk genes were up-regulated in high-risk samples.

The association between the prognostic risk model and tumor immune microenvironment was also investigated, and the results revealed the potential roles of immune infiltrating cells in prognostic evaluation. As shown in **Figure 5D**, except for B cells, all immune infiltrating cells were negatively correlated with risk score.

## Prognostic value of the 5 risk genes used in our risk model

The expression of the five risk genes used in our risk model was analyzed and compared between adjacent prostate gland samples and PRAD tissues. The results indicated that the

## Immune-related signature for PRAD patients

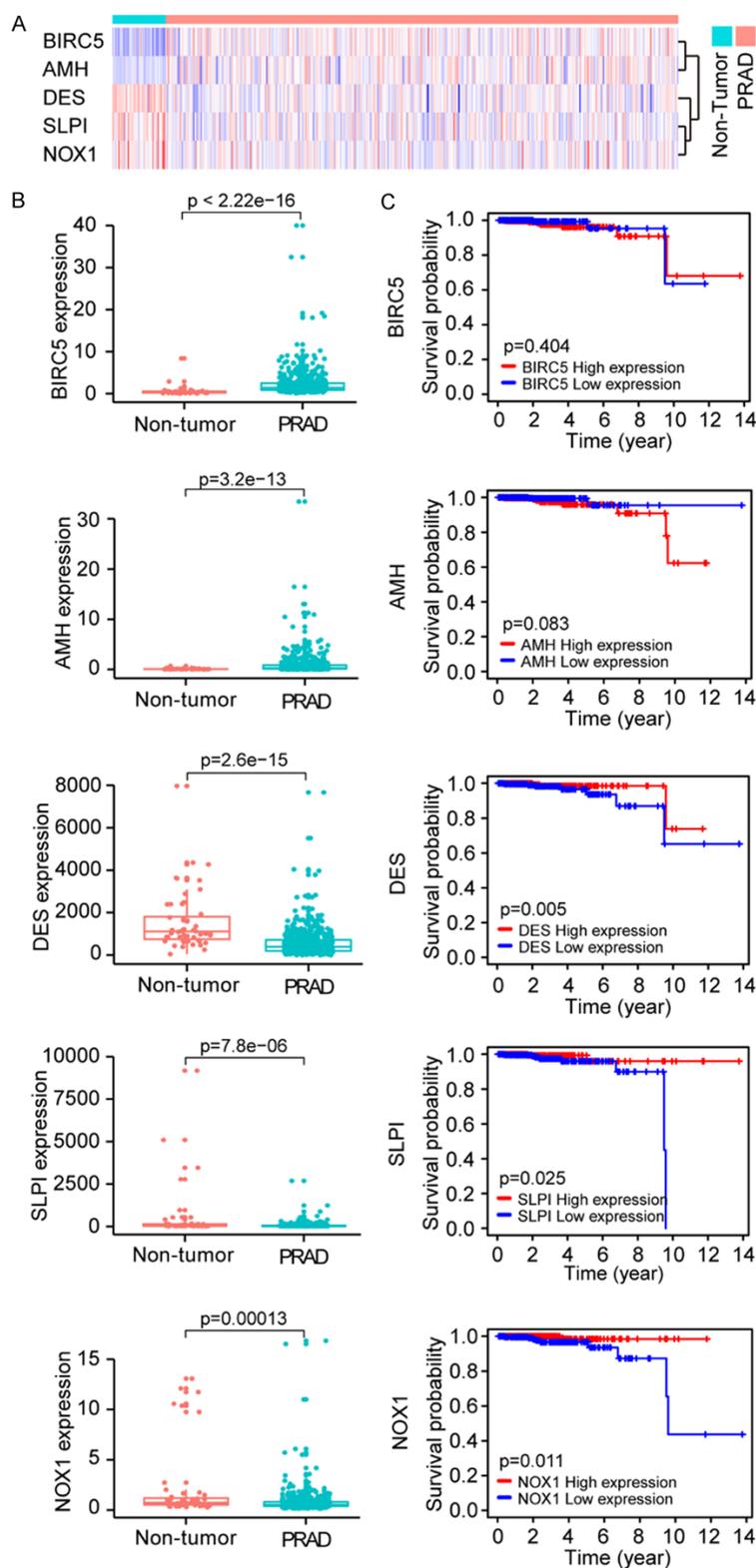


**Figure 5.** Clinical correlation, independent predictive power, and immune activity correlation of the immune-related risk score model. **A.** Five-year Receiver operating characteristic curve (ROC) of the risk score model and the clinicopathological indicators for assessing its prognostic performance. **B.** Association analyses between risk scores and the clinicopathological indicators revealed that high risk scores had a positive correlation with advanced tumor stage. **C.** Evaluation on the independent predictive power of the risk score model. Univariate (left panel) and multivariate Cox regression analyses (right panel) of the risk score model and the clinicopathological indicators. **D.** The correlation between the risk score model and the immune cell infiltration in PRAD.

expression of the high risk genes (*BIRC5* and *AMH*) was significantly increased in PRAD tissues, and the low risk genes (*SLPI*, *NOX1* and *DES*) were dramatically downregulated in PRAD tissues (**Figure 6A, 6B**). **Figure 6C** showed that in PRAD patients, high risk genes were related to worse survival, and low risk genes were

associated with higher survival probability. Immunohistochemistry (IHC) data from The Human Protein Atlas indicated that the expression of *BIRC5* protein was higher in PRAD tissues than in normal prostate tissues, and the expression of *DES* protein was higher in normal prostate tissues than in PRAD tissues (**Figure**

## Immune-related signature for PRAD patients



**Figure 6.** Evaluation of the prognostic value of the five risk genes used in the risk score model. A, B. The expression of the five risk genes in PRAD and non-tumor samples. C. The correlation between the expression of the five risk genes and the survival of PRAD patients.

S1). Immunohistochemistry (IHC) data for AMH, SLPI and NOX1 were missing in The Human Protein Atlas.

### Discussion

PRAD has the highest morbidity and the second highest mortality among men [19]. Since the widespread screening of PRAD with prostate-specific antigen (PSA), the diagnosis of early-stage PRAD has been substantially improved, and total mortality has been declined [20]. However, according to statistics, a large proportion of the screened PRAD has indolent tumors that are clinically innocuous and will not lead to a lethal outcome [21, 22]. Nevertheless, the other proportion of PRAD has aggressive tumors that can metastasize rapidly and have huge potential to be lethal. To this end, there is an urgency to understand their etiological differences and then distinguish indolent tumors from aggressive PRAD [23].

Clinicians need to accurately identify men in need of treatment through reliable and reproducible assays, thereby reducing unnecessary death due to misdiagnosis and unnecessary treatment to optimize treatment benefits [24, 25]. However, due to the lack of efficient and accurate biomarkers, clinicians can only stratify the risk of PRAD patients according to the clinicopathologic indicators to provide guidance for treatment [26]. The histopathological evaluation Gleason grading is still the main tool for risk stratifying and determining the treatment course for PRAD patients [27]; however, the

## Immune-related signature for PRAD patients

recommendation of treatment for PRAD patients with Gleason score of lower than 8 is still ambiguous. Thus, identifying efficient biomarkers to distinguish indolent PRAD from potentially lethal PRAD is critically needed. Molecular drivers integrated with clinicopathologic indicators could improve the risk stratification.

Through extensive high-throughput sequencings, numerous prognostic biomarkers for PRAD patients have been identified and reported; however, a single prognostic biomarker has limitation with low reproducibility and poor predictive power. In contrast, prognostic risk models are dependent on a mathematically precise approach using several prognostic biomarkers with different weighting coefficients for risk stratification; hence, the reproducibility and accuracy is improved [28]. Efforts have been devoted to establishing universally reliable risk models with excellent performance and clinical availability for early diagnosis, prognosis predicting, clinical treatment instruction, and so on. But it is worth noting that most researchers constructed and validated their risk models only based on a training cohort [29], whereas our risk model was validated in two independent cohorts and had been proved to perform better in risk stratification of PRAD patients than other clinicopathologic indicators. In addition, although some researchers constructed risk models based on several genes that they identified, establishing the risk model using immune-related independent prognostic DEIRGs is novel and has not been previously reported.

To explore biomarkers that can distinguish indolence PRAD from aggressive PRAD, using genomic sequencing technology combined with aberrant mRNA expression-based gene signature has shown a tremendous potential in prognosis prediction of PRAD patients. As a pivotal part of the tumor microenvironment (TME), immune cells play an important role in the survival and propagation of PRAD [30]. Multiple studies have shown that the secretory product of macrophages, macrophage inhibitory cytokine-1, is part of the innate immunity and is correlated with inflammation, which contributes to the development of PRAD [31, 32]. Many studies report that cyclooxygenase 2 can affect prostate carcinogenesis by promoting immune suppression [33]. Understanding the difference in the immune status during the

development of PRAD is beneficial for the prognosis evaluation of PRAD.

In our study, the immune-related prognostic factors were analyzed and incorporated into the prognostic evaluation system of PRAD to construct a novel risk prediction model. We validated our novel immune-related prognostic risk model and showed its general applicability and high reliability for prognosis prediction of PRAD patients. The five DEIRGs used in our immune-related prognostic model were also confirmed as independent prognostic genes in PRAD. The five-gene immune-related risk model enhanced the accuracy of prognosis prediction and demonstrated its clinical implication.

### Conclusions

In summary, we identified several DERIGs that were significantly associated with the prognosis of PRAD patients and, for the first time, established a novel five-gene immune-related risk model as an independent prognostic predictor for PRAD patients. The risk model was confirmed to be valid and reliable by comprehensive analyses, which would provide clinical decision support in the prognosis evaluation and targeted therapy. The genes applied in our model indicate that the immune system plays a significant role in PRAD progression.

### Acknowledgements

The study was supported by the clinical research special projects from Shanghai Municipal Health Commission (No. 20204Y0230) and Shanghai Sailing Program from Shanghai Science and Technology Committee (No. 19-YF1452200).

### Disclosure of conflict of interest

None.

### Abbreviations

AUC, Area under the curve; DEIRGs, Differentially expressed Immune related genes; GO, Gene Ontology; HR, Hazard Ratio; ICGC, International Cancer Genome Consortium; IRGs, Immune related genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, Least absolute shrinkage and selection operator; Log<sub>2</sub>FC, Log<sub>2</sub>Fold Change/Logarithm of Fold Change; OS, Overall survival; PPI, Protein-

## Immune-related signature for PRAD patients

protein interaction; PRAD, Prostate Adenocarcinoma; TCGA, The Cancer Genome Atlas.

**Address correspondence to:** Dr. Hongjun Fei, Department of Reproductive Genetics, International Peace Maternity and Child Health Hospital, Shanghai Jiao Tong University School of Medicine, No. 910 Hengshan Road, Shanghai 200030, China. E-mail: feihongjun@alumni.sjtu.edu.cn

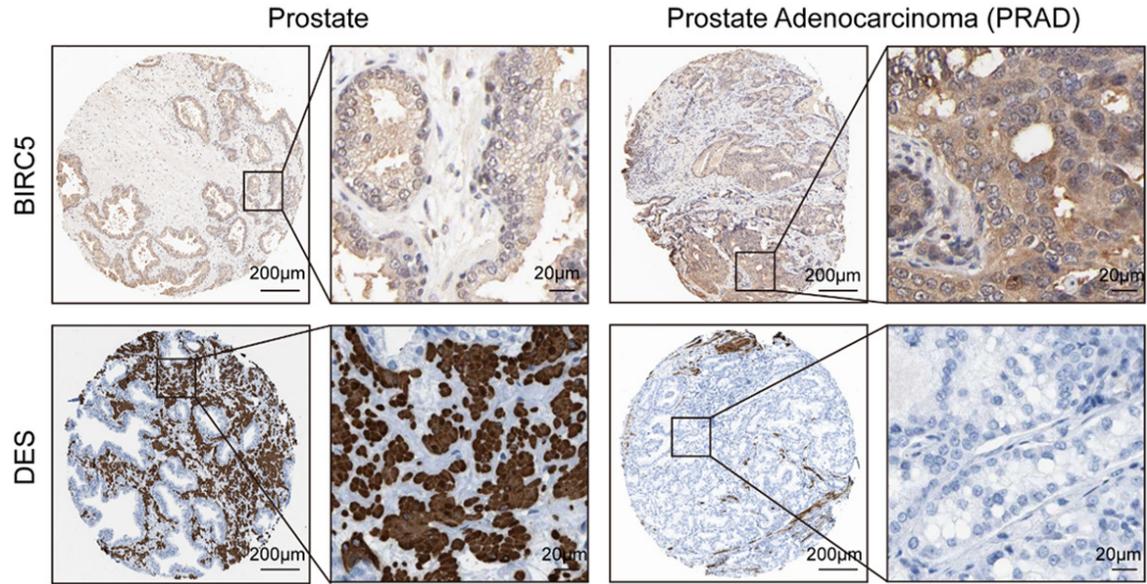
### References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71: 209-249.
- [2] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* 2021; 71: 7-33.
- [3] Pini G, Collins J, Ghadjar P and Wiklund P. Guidance on patient consultation. Current evidence for prostate-specific antigen screening in healthy men and treatment options for men with proven localised prostate cancer. *Curr Urol Rep* 2015; 16: 28.
- [4] Ma X and Huang J. Predicting clinical outcome of therapy-resistant prostate cancer. *Proc Natl Acad Sci U S A* 2019; 116: 11090-11092.
- [5] Muralidhar V, Mahal BA and Nguyen PL. Conditional cancer-specific mortality in T4, N1, or M1 prostate cancer: implications for long-term prognosis. *Radiat Oncol* 2015; 10: 155.
- [6] van Poppel H. Locally advanced and high risk prostate cancer: the best indication for initial radical prostatectomy? *Asian J Urol* 2014; 1: 40-45.
- [7] Klusa D, Lohaus F, Furesi G, Rauner M, Benesova M, Krause M, Kurth I and Peitzsch C. Metastatic spread in prostate cancer patients influencing radiotherapy response. *Front Oncol* 2021; 10: 627379.
- [8] Heo JE, Park JS, Lee JS, Kim J, Jang WS, Rha KH, Choi YD, Hong SJ and Ham WS. Outcomes of pathologically localized high-grade prostate cancer treated with radical prostatectomy. *Medicine (Baltimore)* 2019; 98: e17627.
- [9] Filella X, Albaladejo MD, Allue JA, Castano MA, Morell-Garcia D, Ruiz MA, Santamaria M, Torrejon MJ and Gimenez N. Prostate cancer screening: guidelines review and laboratory issues. *Clin Chem Lab Med* 2019; 57: 1474-1487.
- [10] Martin NE, Mucci LA, Loda M and Depinho RA. Prognostic determinants in prostate cancer. *Cancer J* 2011; 17: 429-437.
- [11] Adhyam M and Gupta AK. A review on the clinical utility of PSA in cancer prostate. *Indian J Surg Oncol* 2012; 3: 120-129.
- [12] Fraser M, Livingstone J, Wrana JL, Finelli A, He HH, van der Kwast T, Zlotta AR, Bristow RG and Boutros PC. Somatic driver mutation prevalence in 1844 prostate cancers identifies ZNRF3 loss as a predictor of metastatic relapse. *Nat Commun* 2021; 12: 6248.
- [13] Dawson NA, Zibelman M, Lindsay T, Feldman RA, Saul M, Gatalica Z, Korn WM and Heath EI. An emerging landscape for canonical and actionable molecular alterations in primary and metastatic prostate cancer. *Mol Cancer Ther* 2020; 19: 1373-1382.
- [14] Wang C, Zhang Y and Gao WQ. The evolving role of immune cells in prostate cancer. *Cancer Lett* 2022; 525: 9-21.
- [15] Wang C, Chen J, Zhang Q, Li W, Zhang S, Xu Y, Wang F, Zhang B, Zhang Y and Gao WQ. Elimination of CD4low HLA-G+ T cells overcomes castration-resistance in prostate cancer therapy. *Cell Res* 2018; 28: 1103-1117.
- [16] Wang C, Peng G, Huang H, Liu F, Kong DP, Dong KQ, Dai LH, Zhou Z, Wang KJ, Yang J, Cheng YQ, Gao X, Qu M, Wang HR, Zhu F, Tian QQ, Liu D, Cao L, Cui XG, Xu CL, Xu DF and Sun YH. Blocking the feedback loop between neuroendocrine differentiation and macrophages improves the therapeutic effects of enzalutamide (MDV3100) on prostate cancer. *Clin Cancer Res* 2018; 24: 708-723.
- [17] Ammirante M, Luo JL, Grivnickov S, Nedospasov S and Karin M. B-cell-derived lymphotoxin promotes castration-resistant prostate cancer. *Nature* 2010; 464: 302-305.
- [18] Cha HR, Lee JH and Ponnazhagan S. Revisiting immunotherapy: a focus on prostate cancer. *Cancer Res* 2020; 80: 1615-1623.
- [19] Elmehraath AO, Afifi AM, Al-Husseini MJ, Saad AM, Wilson N, Shohdy KS, Pilie P, Sonbol MB and Alhalabi O. Causes of death among patients with metastatic prostate cancer in the US from 2000 to 2016. *JAMA Netw Open* 2021; 4: e2119568.
- [20] Torre LA, Siegel RL, Ward EM and Jemal A. Global cancer incidence and mortality rates and trends-an update. *Cancer Epidemiol Biomarkers Prev* 2016; 25: 16-27.
- [21] Yatani R, Chigusa I, Akazaki K, Stemmermann GN, Welsh RA and Correa P. Geographic pathology of latent prostatic carcinoma. *Int J Cancer* 1982; 29: 611-616.
- [22] Albertsen PC. Observational studies and the natural history of screen-detected prostate cancer. *Curr Opin Urol* 2015; 25: 232-237.
- [23] Platz EA, De Marzo AM and Giovannucci E. Prostate cancer association studies: pitfalls

## Immune-related signature for PRAD patients

- and solutions to cancer misclassification in the PSA era. *J Cell Biochem* 2004; 91: 553-571.
- [24] Welch HG and Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986-2005. *J Natl Cancer Inst* 2009; 101: 1325-1329.
- [25] Daskivich TJ, Chamie K, Kwan L, Labo J, Palvolgyi R, Dash A, Greenfield S and Litwin MS. Overtreatment of men with low-risk prostate cancer and significant comorbidity. *Cancer* 2011; 117: 2058-2066.
- [26] Prensner JR, Rubin MA, Wei JT and Chinnaiyan AM. Beyond PSA: the next generation of prostate cancer biomarkers. *Sci Transl Med* 2012; 4: 127rv3.
- [27] Gleason DF. Histologic grading of prostate cancer: a perspective. *Hum Pathol* 1992; 23: 273-279.
- [28] Vrieze SI. Model selection and psychological theory: a discussion of the differences between the akaike information criterion (AIC) and the bayesian information criterion (BIC). *Psychol Methods* 2012; 17: 228-243.
- [29] Li XY, You JX, Zhang LY, Su LX and Yang XT. A novel model based on necroptosis-related genes for predicting prognosis of patients with prostate adenocarcinoma. *Front Bioeng Biotechnol* 2022; 9: 814813.
- [30] Jiao S, Subudhi SK, Aparicio A, Ge Z, Guan B, Miura Y and Sharma P. Differences in tumor microenvironment dictate T helper lineage polarization and response to immune checkpoint therapy. *Cell* 2019; 179: 1177-1190, e13.
- [31] Condeelis J and Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell* 2006; 124: 263-266.
- [32] De Marzo AM, Nakai Y and Nelson WG. Inflammation, atrophy, and prostate carcinogenesis. *Urol Oncol* 2007; 25: 398-400.
- [33] Wang W, Bergh A and Damber JE. Cyclooxygenase-2 expression correlates with local chronic inflammation and tumor neovascularization in human prostate cancer. *Clin Cancer Res* 2005; 11: 3250-3256.

## Immune-related signature for PRAD patients



**Figure S1.** Immunohistochemistry analysis of the 5 genes (*BIRC5*, *AMH*, *DES*, *SLPI* and *NOX1*) that were used to develop the risk prognostic model. *BIRC5* protein was detected by antibody HPA002830. *DES* protein was detected by antibody HPA018803. Immunohistochemistry of *AMH*, *SLPI* and *NOX1* was missing in The Human Protein Atlas.