



# Tumor Expression Profile Analysis Developed and Validated a Prognostic Model Based on Immune-Related Genes in Bladder Cancer

Bingqi Dong<sup>1†</sup>, Jiaming Liang<sup>2,3†</sup>, Ding Li<sup>4†</sup>, Wenping Song<sup>4</sup>, Shiming Zhao<sup>1</sup>, Yongkang Ma<sup>1</sup>, Jinbo Song<sup>1</sup>, Mingkai Zhu<sup>1</sup> and Tiejun Yang<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Xiangqian Guo,  
Henan University, China

### Reviewed by:

Shuangyu Lv,  
Henan University, China  
Jiateng Zhong,  
Xinxiang Medical University, China

### \*Correspondence:

Tiejun Yang  
tiejunyang@126.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Cancer Genetics and Oncogenomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 18 April 2021

**Accepted:** 28 July 2021

**Published:** 27 August 2021

### Citation:

Dong B, Liang J, Li D, Song W,  
Zhao S, Ma Y, Song J, Zhu M and  
Yang T (2021) Tumor Expression  
Profile Analysis Developed and  
Validated a Prognostic Model  
Based on Immune-Related Genes in  
Bladder Cancer.  
Front. Genet. 12:696912.  
doi: 10.3389/fgene.2021.696912

<sup>1</sup>Department of Urology, Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou, China, <sup>2</sup>The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, <sup>3</sup>State Key Laboratory of Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University, National Clinical Research Center for Respiratory Disease, Guangzhou, China, <sup>4</sup>Department of Pharmacy, Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou, China

**Background:** Bladder cancer (BLCA) ranks 10th in incidence among malignant tumors and 6th in incidence among malignant tumors in males. With the application of immune therapy, the overall survival (OS) rate of BLCA patients has greatly improved, but the 5-year survival rate of BLCA patients is still low. Furthermore, not every BLCA patient benefits from immunotherapy, and there are a limited number of biomarkers for predicting the immunotherapy response. Therefore, novel biomarkers for predicting the immunotherapy response and prognosis of BLCA are urgently needed.

**Methods:** The RNA sequencing (RNA-seq) data, clinical data and gene annotation files for The Cancer Genome Atlas (TCGA) BLCA cohort were extracted from the University of California, Santa Cruz (UCSC) Xena Browser. The BLCA datasets GSE31684 and GSE32894 from the Gene Expression Omnibus (GEO) database were extracted for external validation. Immune-related genes were extracted from InnateDB. Significant differentially expressed genes (DEGs) were identified using the R package “limma,” and Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for the DEGs were performed using R package “clusterProfiler.” Least absolute shrinkage and selection operator (LASSO) regression analysis were used to construct the signature model. The infiltration level of each immune cell type was estimated using the single-sample gene set enrichment analysis (ssGSEA) algorithm. The performance of the model was evaluated with receiver operating characteristic (ROC) curves and calibration curves.

**Results:** In total, 1,040 immune-related DEGs were identified, and eight signature genes were selected to construct a model using LASSO regression analysis. The risk score of BLCA patients based on the signature model was negatively correlated with OS and the immunotherapy response. The ROC curve for OS revealed that the model had good accuracy. The calibration curve showed good agreement between the predictions and actual observations.

**Conclusions:** Herein, we constructed an immune-related eight-gene signature that could be a potential biomarker to predict the immunotherapy response and prognosis of BLCA patients.

**Keywords:** bladder cancer, immune-related signature, The Cancer Genome Atlas, Gene Expression Omnibus, immunotherapy

## INTRODUCTION

Bladder cancer (BLCA) ranks 10th in incidence among malignant tumors and 6th in incidence among malignant tumors in males (Bray et al., 2018). The disease may present as muscularly invasive bladder cancer (MIBC), non-muscularly invasive bladder cancer (NMIBC) or as a metastatic form of other diseases (Tran et al., 2021). With the development of technologies such as electrosurgery, chemotherapy, and radical surgery, the overall survival (OS) rate of BLCA patients has greatly improved. However, primary BLCA are prone to recurrence after systematic treatment, the prognosis is not satisfactory (Rouprêt et al., 2021), and there are no clinically meaningful diagnostic markers. Based on the fact that the incidence of BLCA is gradually increasing, valuable biomarkers are particularly urgently needed.

Current studies have revealed that the tumor microenvironment is closely correlated with tumorigenesis, progression and prognosis, and immune cells of the tumor microenvironment exhibit complex interactions with tumor cells (Hanahan and Weinberg, 2011). As the landmark developments of immune checkpoint inhibitors (represented by anti-PD-1/PD-L1 antibodies), Immunotherapy is playing an increasingly important role in the treatment of BLCA (Powles et al., 2014; Pettenati and Ingersoll, 2018; Rouanne et al., 2018). Thus, immune molecules associated with the tumor microenvironment have a tremendous role in serving as prognostic markers for BLCA. Previous studies have proposed immune-related biomarkers in thyroid and ovarian cancers for risk stratification and clinical outcome prediction (Kim et al., 2018; Shen et al., 2019). A few studies have been conducted to assess the potential of immune-related genes to predict clinical outcomes and the immunotherapy response in BLCA (Qiu et al., 2020; Zhu et al., 2020; Lv et al., 2021), but the depth and results of these studies are not satisfactory.

Herein, the purpose of this study is to find suitable biomarkers of BLCA with high sensitivity and strong specificity and molecular targets that affect the clinicopathological process of BLCA, then provide an important reference for the diagnosis of BLCA. We identified immune-related DEGs for BLCA, and constructed an immune-related eight-gene signature model. The signature model showed good prognostic value for predicting

OS and could be used to predict the immunotherapy response in BLCA patients.

## MATERIALS AND METHODS

### Data Collection and Preprocessing

The RNA sequencing (RNA-seq) data, probe annotation files and clinical data of the BLCA patients were extracted from The Cancer Genome Atlas (TCGA) and used to acquire the expression profiles of the BLCA patients. After screening, samples with no clinical data were excluded. A total of 406 tumor samples and 18 normal samples were included in the analysis. BLCA cohorts GSE31684 and GSE32894 were obtained from the Gene Expression Omnibus (GEO) database using the R package “GEOquery” (Davis and Meltzer, 2007). Immune-related genes were obtained from InnateDB.<sup>1</sup>

### Identification of Differentially Expressed Genes and Functional Annotation

Significant DEGs between normal and BLCA samples were identified with screening criteria of adjusted value of  $p < 0.05$  and  $|\log \text{fold change (FC)}| > 1$  by the R package “limma” (Ritchie et al., 2015). Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for the DEGs were performed using R package “clusterProfiler” (Yu et al., 2012).

### Estimation of the Infiltration Degree of Each Immune Cell in BLCA

A group of immune cell gene markers, consisting of 782 genes, which represent 28 immune cell types related to innate and adaptive immunity, were obtained from previous studies to estimate the infiltration level of different immune cell types in the tumor microenvironment (Charoentong et al., 2017). The assessed immune cell types included, B cells, natural killer (NK) cells, dendritic cells, myeloid-derived suppressor cells (MDSCs), neutrophils, and T cells. Subsequently, the expression

<sup>1</sup><https://www.innatedb.ca/>

profiles of each sample were used to estimate the infiltration level of each immune cell type in BLCA using the single-sample gene set enrichment analysis (ssGSEA) algorithm with the R package “GSVA” (Hänzelmann et al., 2013).

## Survival Analysis

Univariate Cox proportional hazards regression analysis was carried out to evaluate the association between the expression level of the immune-related DEGs and the OS of BLCA patients. Immune-related DEGs with a value of  $p < 0.001$  based on the log-rank test were selected as candidate genes for construction of the prognostic model. The risk score of each patient was calculated based on the signature model and was used to evaluate the association between the gene signature model and the prognosis of BLCA patients. The samples were assigned to the high-risk or low-risk group based on the median risk score. Kaplan–Meier curves and log-rank tests were performed to compare the differences in OS and progression-free survival between the high-risk and low-risk groups. A statistically significant difference was defined as  $p < 0.05$ . Survival analysis and log-rank tests were performed using the R package “survival,” and the R package “survminer” was used to plot the Kaplan–Meier curve.

## Establishment and Evaluation of the Immune-Related Signature Model

The TCGA BLCA cohort was randomly divided into a training set ( $n = 285$ ) and a testing set ( $n = 121$ ) at a ratio of 7:3. The R package “glmnet” was used to perform LASSO regularization to reduce the coefficients from the training set. An immune related eight-gene model was constructed. The following formula was used:

$$\begin{aligned} \text{Risk score} = & -0.15744 \times DCCHS1 + 0.05310 \\ & \times PTGIS - 0.32469 \times PTPN6 \\ & - 0.07170 \times AIFM3 + 0.11454 \\ & \times FLRT2 + 0.04565 \times PCSK5 \\ & + 0.15854 \times CLSTN2 - 0.11663 \times HSH2D \end{aligned}$$

The risk score was calculated for each sample using the signature model. The BLCA cohort was assigned to high- and low-risk groups based on the risk score. Receiver operating characteristic (ROC) curves for 1-, 3-, and 5-year OS were generated for the two groups using the R package “survivalROC.” Calibration curves were derived from the R package “rms” to evaluate the precision of the 1-, 3-, and 5-year OS prediction. Decision curve analysis was performed using the R package “ggDCA” to quantify the net benefits at different threshold probabilities and evaluate the clinical usefulness of the immune-related signature model.

## Prediction of the Immunotherapy Response

The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm<sup>2</sup> was used to estimate the response of each sample to anti-PD-1/

PD-L1 and anti-CTLA4 immunotherapy based on the gene expression profiles of the BLCA cohort (Jiang et al., 2018).

## Statistical Analysis

Univariate survival analysis was performed using the log-rank test. The Pearson correlation formula was used to calculate the correlations between the risk score and immune markers, the risk score and characteristic gene expression, characteristic gene expression and the immune cell infiltration score, and the risk score and immune cell infiltration score. Two-tailed Student’s *t* tests were used for two-group comparisons. A statistically significant difference was defined as  $p < 0.05$ . All statistical analyses were performed in R version 4.0.2.

## RESULTS

### Identification of Immune-Related DEGs in BLCA

The overall design is shown in **Figure 1**. First, to identify DEGs in BLCA, a total of 406 BLCA samples and 18 normal samples were downloaded from the University of California, Santa Cruz (UCSC). A total of 3,677 up-regulated and 3,182 down-regulated genes were obtained (**Figure 2A**). Among these DEGs, there were 1,040 immune-related genes, of which 385 were up-regulated and 655 were down-regulated, respectively (**Figures 2B,C**).

### GO and KEGG Annotation of the DEGs

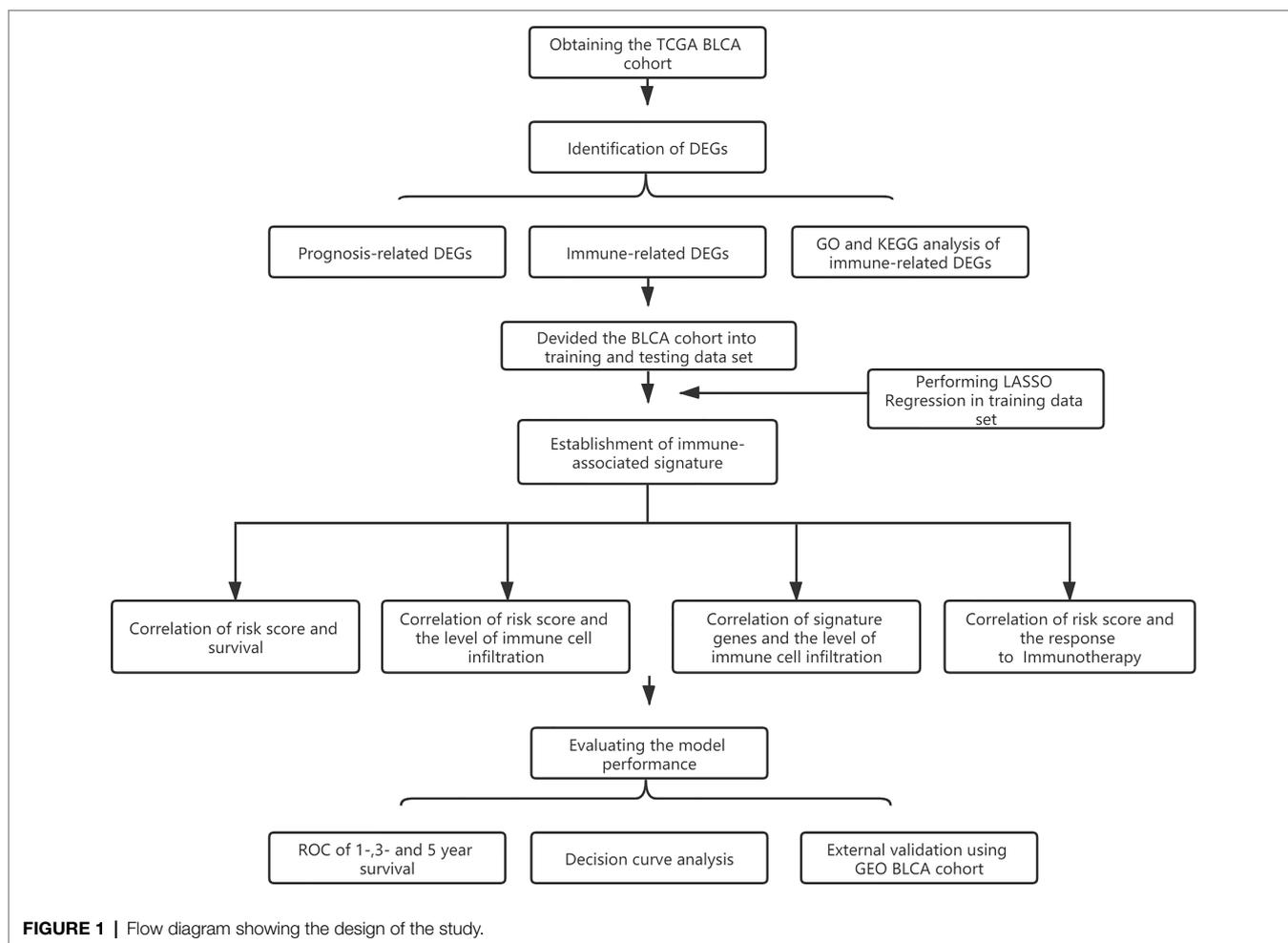
The results of GO functional annotation analysis of the immune-related DEGs showed that the most significantly enriched biological processes (BPs) included homophilic cell adhesion *via* plasma membrane adhesion molecules, extracellular structure organization, and extracellular matrix organization. The most significantly enriched cellular components (CCs) included collagen trimers, the endoplasmic reticulum lumen, and cell–cell junctions, and the molecular functions (MFs) included integrin binding, cytokine receptor binding, and cytokine activity (**Figure 2D**). KEGG pathway enrichment analysis for the immune-related DEGs showed that the significant pathways included the Rap1 signaling pathway, the JAK–STAT signaling pathway, transcriptional misregulation in cancer and proteoglycans in cancer (**Figure 2E**). Most of the above results were related to immunity.

### Construction of the Immune-Related Signature

Univariate Cox regression based on the survival and gene expression data of BLCA patients was used to evaluate the prognostic value of the immune-related DEGs. In total, 13 immune-related DEGs were selected as candidate genes to construct the prognostic model with the criterion of value of  $p < 0.001$ . A forest plot showing the value of  $p$  and hazard ratio of the candidate genes is shown in **Figure 3A**.

The BLCA cohort was randomly divided into training and testing sets at a ratio of 7:3. LASSO regression analysis was

<sup>2</sup><http://tide.dfci.harvard.edu>



performed using the training data set to identify genes that were most significantly related to survival outcomes (Figures 3B,C). A total of eight signature genes were identified: DCHS1, PTGIS, PTPN6, AIFM3, FLRT2, PCSK5, CLSTN2, and HSH2D (Figure 3D); the coefficients of the signature genes are shown in Supplementary Table 1. Next, we investigated the expression profile of these eight genes in the BLCA cohort. The results indicated that AIFM3, HSH2D, and PTPN6 were significantly up-regulated, while DCHS1, PTGIS, FLRT2, PCSK5, and CLSTN2 were significantly down-regulated in tumor samples compared with normal samples (Figure 3E).

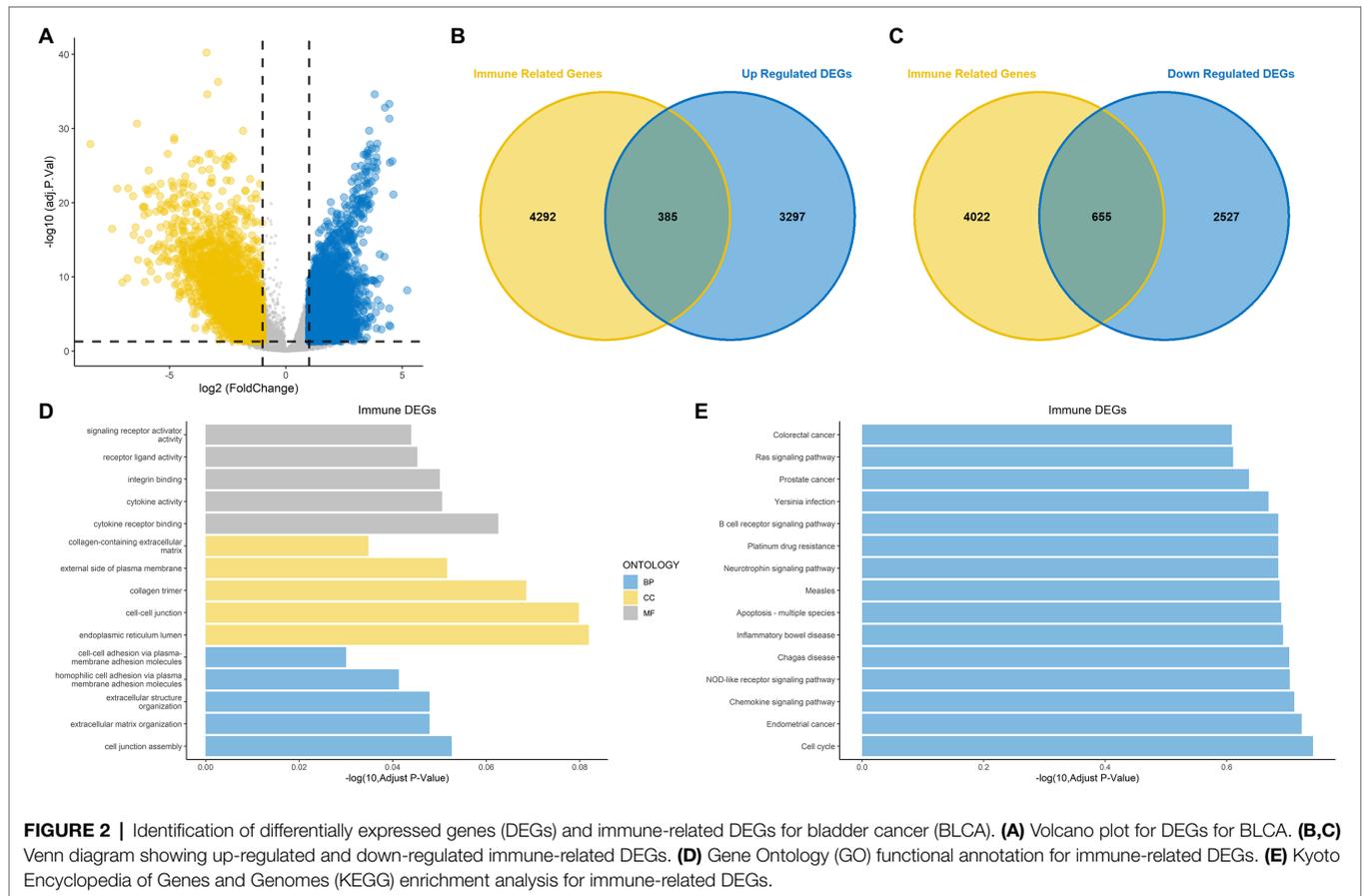
### A High-Risk Score Is Correlated With Poor Clinical Outcome in BLCA

The risk score of each sample was calculated and ranked on the basis of the signature model in the training set (Figure 4A). The scatter plot represented the OS status of BLCA patients according to the risk score, and it suggested that the high-risk group had higher mortality than the low-risk group (Figure 4B). The expression profiles of the signature genes showed that tumors with higher risk scores tended to exhibit elevated PCSK5, DCHS1, CLSTN2, PTGIS, and FLRT2 levels, while those with lower risk scores tended to exhibit elevated AIFM3,

PTPN6, and HSH2D levels (Figure 4C). The same analysis was performed using the testing set, and the results were consistent with those derived from the training data set (Figures 4D–F). Compared with the low-risk group, the high-risk group presented a significantly poorer clinical outcome in both the training ( $p < 0.001$ , Figure 5A) and testing sets ( $p < 0.001$ , Figure 5B). Then, the association between the risk score and OS was evaluated, and Kaplan–Meier analysis showed that the low-risk groups had a longer survival time than the high-risk group in both the training (Figure 5C,  $p < 0.0001$ ) and testing sets (Figure 5D,  $p = 0.0018$ ). These results demonstrated that the risk score was associated with OS and that a low-risk score predicts better survival outcomes for BLCA patients.

### Estimating the Degree of Each Immune Cell Infiltration and Predicting the Response to Immunotherapy

Since the efficacy of immunotherapy depends on the level of immune cell infiltration, we analyzed the correlation between the level of immune cell infiltration and risk score. The results demonstrated that there were differences in the infiltration of most immune cells between the high-groups and low-risk

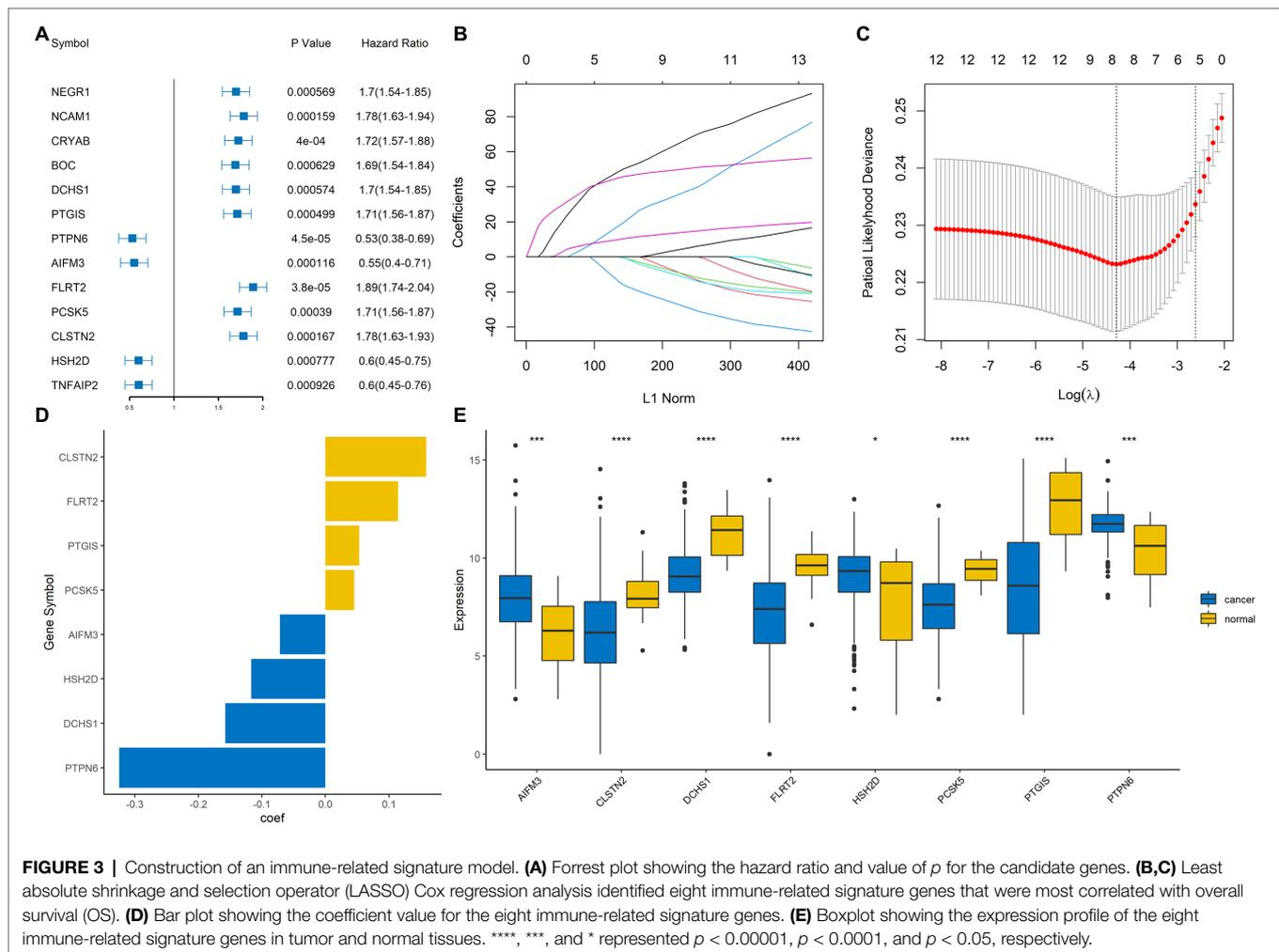


groups, except for CD56dim NK cells and Eosinophil cells, which demonstrated that signature was significantly correlated with immune infiltration (**Figure 6A**). In addition, we explored the correlation between each of these eight genes and immune cell infiltration. DCHS1, PTGIS, FLRT2, PCSK5, and CLSTN2 were positively related to the infiltration almost all immune cells, and AIFM3 was negatively correlated with the infiltration of most immune cells (**Figure 6B**). Since the eight signature genes were significantly correlated with the level of immune cell infiltration, the correlation of the risk score and the immunotherapy response was explored. Firstly, the relationship between the risk score derived from the risk signature and the expression of immune checkpoint molecules was analyzed. The risk score was positively correlated with the expression of the checkpoint markers, PD-1 (**Figure 6C**), PD-L1 (**Figure 6D**) and CTLA-4 (**Figure 6E**), implicating the potential roles of the signature model in the response to immunotherapy in BLCA patients. The immunotherapy response of the BLCA cohort was estimated by the TIDE algorithm, and the ability of the model to predict the immunotherapy response was evaluated. Furthermore, we then calculated the immunotherapy response rate of samples in the high-risk and low-risk groups. Overall, 49 and 66% patients in the high-risk and low-risk groups, respectively, were predicted to respond to immunotherapy (**Figure 6F**). Patients were divided into responsive and nonresponsive groups, then their risk scores were calculated.

The results indicated that the responsive group had significantly lower scores than the nonresponsive group (**Figure 6G**,  $p=0.001$ ). An ROC curve was generated to determine the efficacy of the risk score in predicting the response to immunotherapy. The area under the ROC curve (AUC) was 0.595, suggesting that the immune-related signature model predicted the response to immunotherapy for BLCA with modest accuracy (**Figure 6H**). In summary, the eight-gene signature model was associated with immune cell infiltration and immunotherapy response in BLCA.

## Evaluation of the Model Performance

Time-dependent ROC curves and AUCs were plotted to determine the prognostic values of the eight-gene risk score in the training and testing sets. The AUCs of the risk score for predicting 1-, 3-, and 5-year OS were 0.621, 0.7, and 0.737 in the training set (**Figure 7A**) and 0.655, 0.68, and 0.695 in testing set (**Figure 7B**). To compare the consistency of the model prediction with actual clinical outcomes, calibration curves for 1-, 3-, and 5-year OS in the training (**Supplementary Figures 1A–C**) and testing sets (**Supplementary Figures 1D–F**) were constructed. The results suggested that the calibration curves showed satisfactory agreement between the predicted and observed values for 1-, 3-, and 5-year OS. To determine the clinical usefulness of the risk signature, decision curve analysis in the training and testing sets was performed. Decision curve analysis



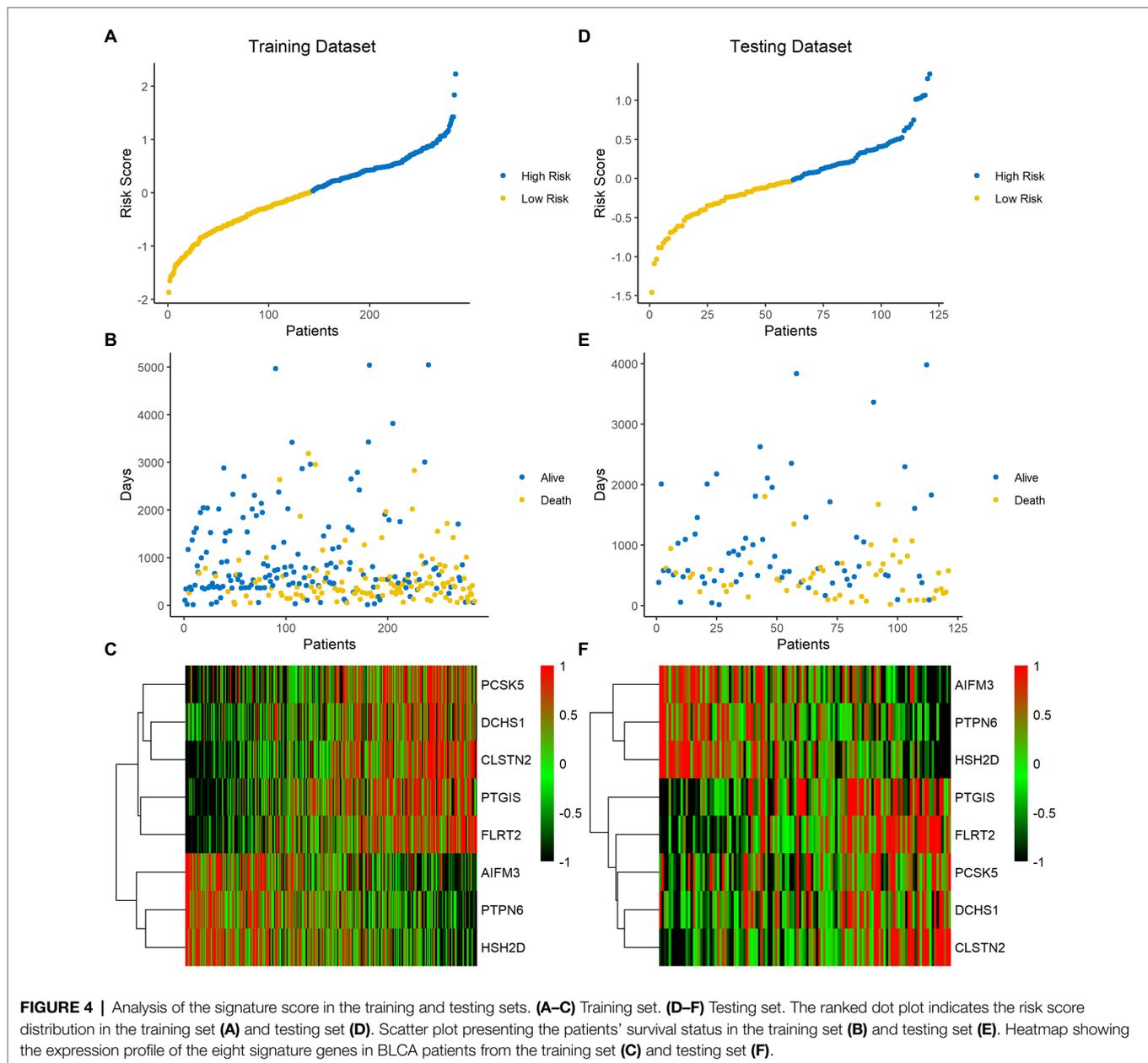
showed that the risk scores offered a net benefit over the “treat-all” or “treat-none” strategy, which indicated that the model was clinically useful (Figures 7C,D). External validation was performed using the BLCA GEO database cohort (GSE31684 and GSE32894). All the samples were divided into high-risk and low-risk groups based on the optimal cutoff point of the risk score, then Kaplan–Meier analysis was performed. The results suggested that the prognosis of the low-risk group was better than that of the high-risk group in the BLCA cohorts GSE31684 (Figure 7E,  $p=0.047$ ) and GSE32894 (Figure 7F,  $p=0.012$ ). The results above indicate that this model had good predictive power in both the TCGA and other external cohorts.

## DISCUSSION

Tumors arise as a result of the accumulation of genetic mutations (Williams and Stein, 2004), and a large number of point mutations and structural changes occur during the development of tumors (Stratton et al., 2009), which inevitably stimulates the production of corresponding tumor antigens and thus induces an immune response (Zhang and Zhang, 2020). In recent

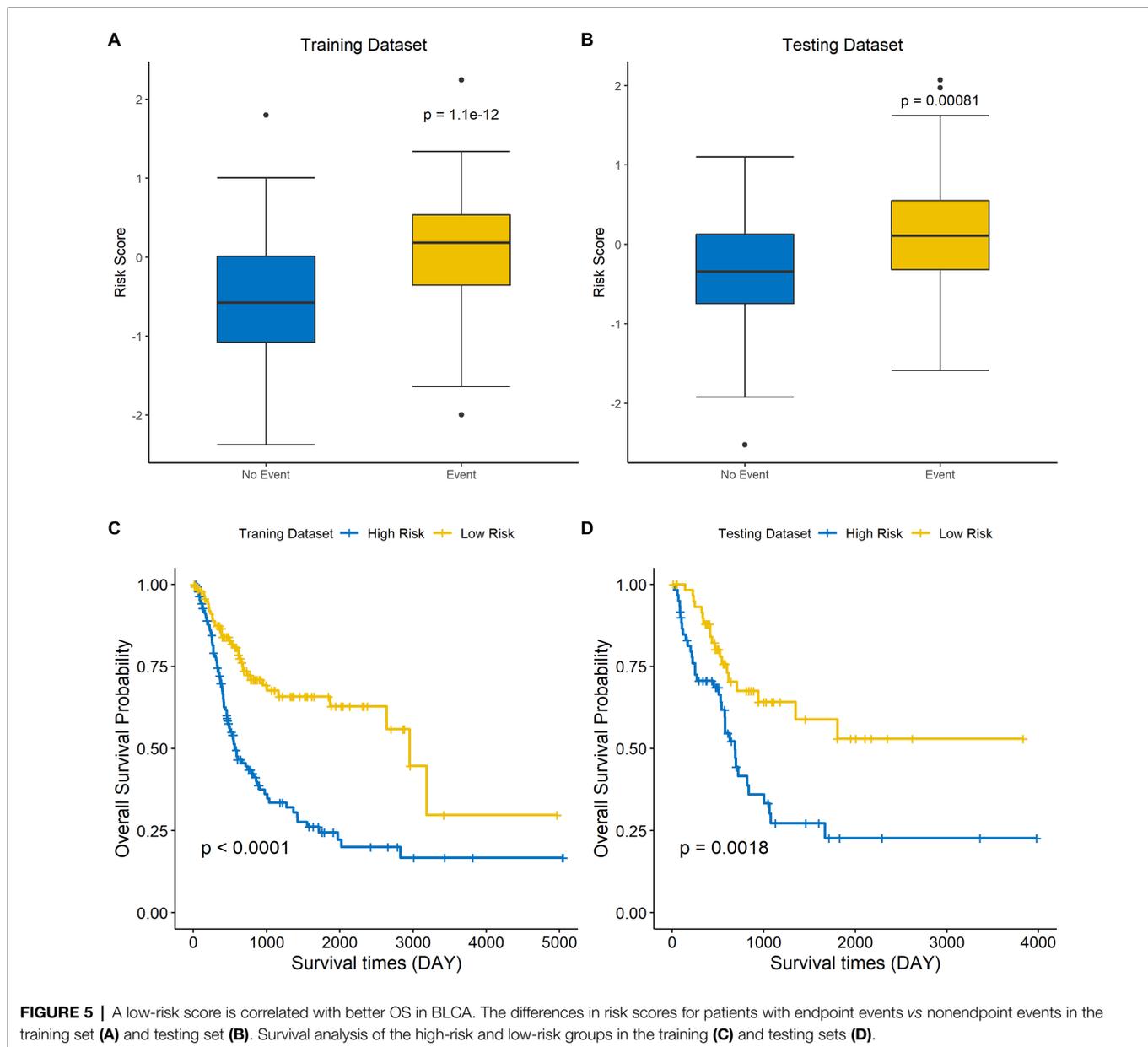
years, with in-depth research on tumor immunology and molecular biology, immunotherapy has provided a new direction for tumor treatment (Ma et al., 2021). A variety of immune checkpoint (PD-1, PD-L1, and CTLA-4) inhibitors (ICIs) are being used in the treatment of advanced BLCA, and clinical studies have shown that they are reliable in terms of safety and efficacy. However, not every BLCA patient benefits from immunotherapy (Fumet et al., 2020). Consequently, it is imperative to identify biomarkers that can predict patient response to immunotherapy. In the present study, we aimed to construct an immune-related DEGs model, to investigate the relationship between this model and patient prognosis as well as the immunotherapy response, and to assess the potential clinical applications of the model.

In the present study, we obtained 1,040 immune-related BLCA DEGs by analyzing the TCGA database; 385 were up-regulated and 655 were down-regulated. Thirteen immune-related DEGs with significant prognostic value were screened by a one-way Cox regression model ( $p < 0.001$ ). Subsequently, we used LASSO regression analysis to screen eight signature genes from the 13 candidate genes in the univariate Cox regression model. The eight genes were CLSTN2, FLRT2, PTGIS,



PCSK5, AIFM3, HSH2D, DCHS1, and PTPN6. FLRT2 is involved in the development of ovarian and uterine carcinogenesis, and FLRT2 has been reported to be a tumor suppressor in breast and prostate cancer (Donninger et al., 2004; Santin et al., 2005; Wu et al., 2016; Bae et al., 2017). The MethHc database show that PTGIS has a high level of DNA methylation in BLCA. PTGIS is a HIF-1 $\alpha$  target gene that plays a primary regulatory role in hypoxic tumor progression by activating the transcription of various oncogenes (Lu et al., 2018). PCSK5 is one member of the chymosin-like proprotein convertase family, which can regulate the cleavage and activation of the prestructural domain of TGF $\beta$ /bmp family members. PCSK5 plays a major role in the mouse skeleton and organogenesis (Szumska et al., 2017). AIFM3 express in a variety of tissues and aberrantly expressed

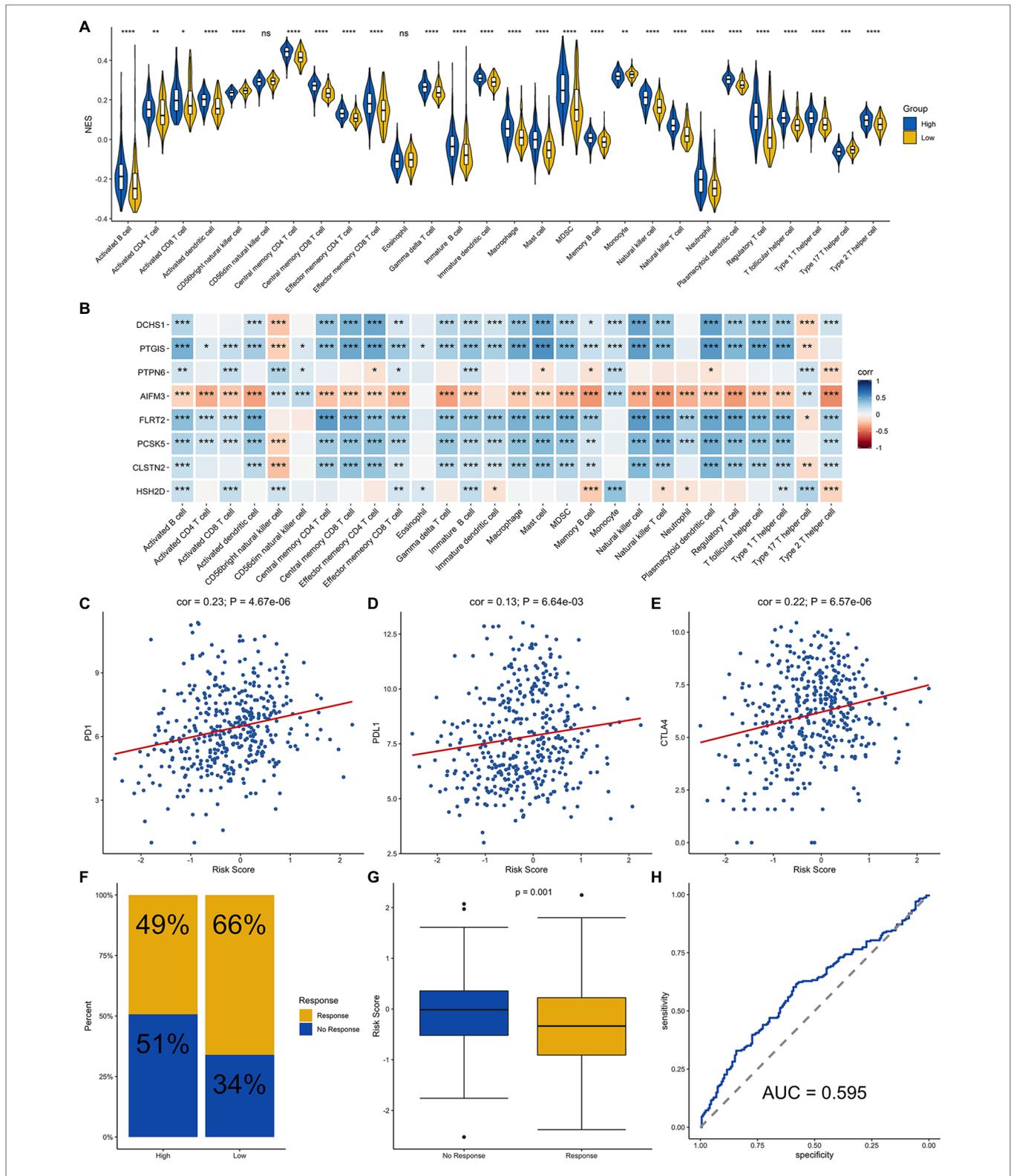
in several cancers, widely. AIFM3 is a direct target of miR-210 and is associated with the proliferation of human liver cancer cells (Yang et al., 2012), and overexpression of AIFM3 predicts stronger proliferative and invasive behavior in breast cancer (Zheng et al., 2019). HSH2D is an important signaling molecule that can affect T cell activation (Oda et al., 2001; Lapinski et al., 2009). HSH2D inhibits the transcriptional activation of the IL-2 promoter, specifically at the RE/AP element of IL-2, which is regulated by CD28, and HSH2D expression contributes to methotrexate resistance in human T-cell acute lymphoblastic leukemia (Shapiro et al., 2004; Pegram et al., 2015). DCHS1 plays a significant role in innate immunity in the human kidney and bladder according to immunostaining studies, and DCHS1 also can participate in cell adhesion, growth, planar cell polarity



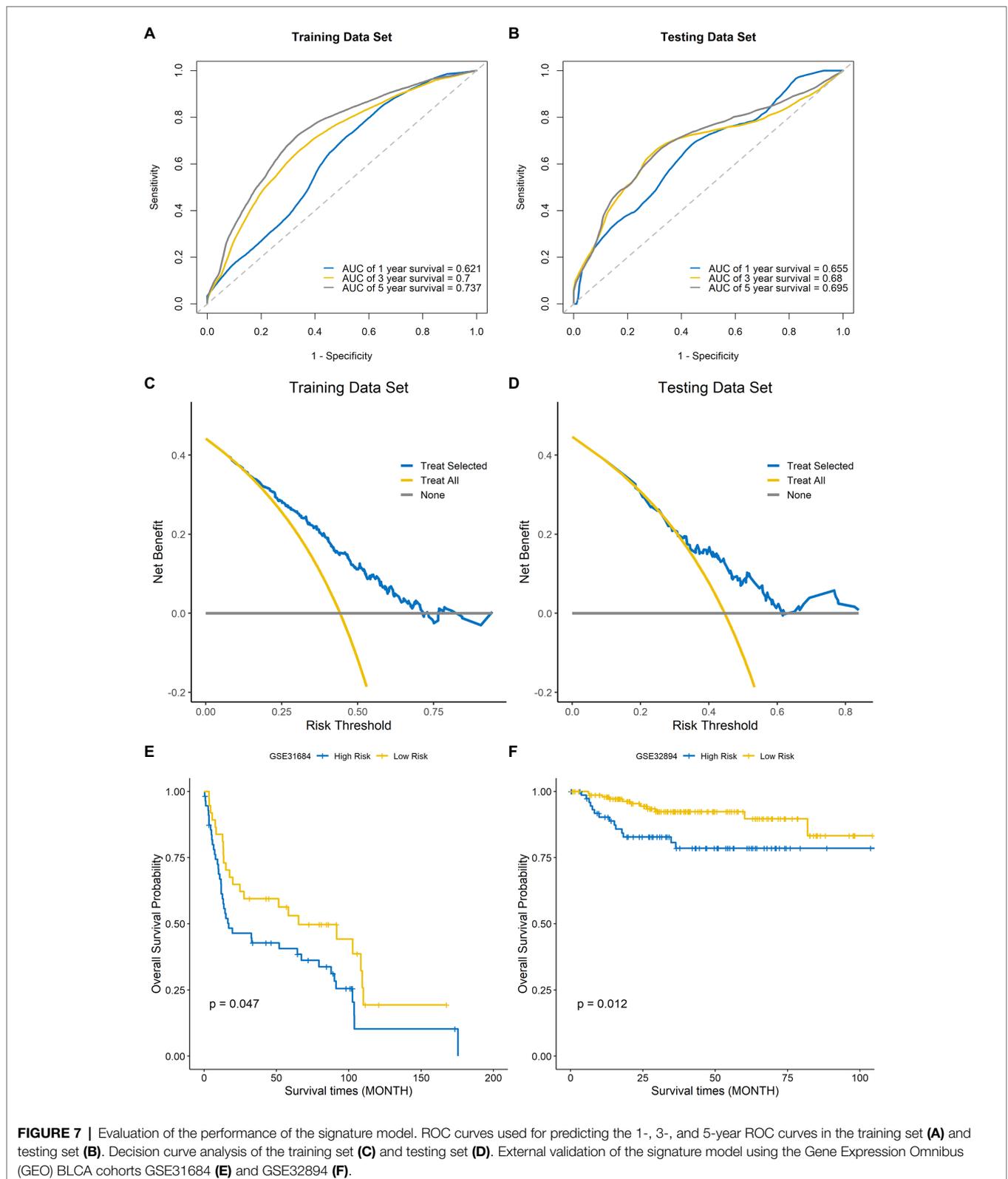
and tissue pattern (Liang et al., 2019; Qureshi et al., 2020). PTPN6 is thought to be a signaling molecule that can regulate many of cellular processes, including cell growth, differentiation, oncogenic transformation and mitotic cycle, as well as can act as a tumor suppressor. PTPN6 may improve the chemotherapy efficacy and can be used in combination with blocking antibodies in immunotherapy; moreover, some studies suggest that PTPN6 may be an immune-related prognostic biomarker for BLCA (Shen et al., 2020).

There is growing evidence that the immune microenvironment, in which immune cells and molecules are important components, acts an important role in tumor development and the degree of immune cell infiltration is highly correlated with patient prognosis (Grivennikov et al., 2010; Seager et al., 2017). With the recent development of technologies such as RNA-seq, it

is possible to systematically analyze the tumor microenvironment and the functional diversity of tumor-infiltrating immune cells, the sensitivity of patients to immunotherapy and the prognosis (Zhang and Zhang, 2019). In recent study, we constructed an eight-gene signature model. The risk scores were significantly associated with the infiltration level of immune cells. Furthermore, six of the eight signature genes showed a significant positive or negative correlation with the infiltration level of immune cells. We also assessed the correlation of the eight signature genes with the response to immunotherapy, and the results suggested that the risk score was significantly and positively correlated with the expression of the checkpoint markers, PD-1, PD-L1, and CTLA-4. The immune-related signature model predicted the response to immunotherapy for BLCA with good accuracy. We applied this signature model



**FIGURE 6 |** Analysis of immune cell infiltration and prediction of the response to immunotherapy in BLCA patients. **(A)** Infiltration level of immune cells in the high-risk and low-risk groups. **(B)** Correlation analysis for the eight signature genes and the infiltration level of the immune cells. Correlation analysis of the risk score and PD-1 **(C)**, PD-L1 **(D)**, and CTLA4 **(E)**. **(F)** Proportion of patients who responded to immunotherapy in the high-risk group and the low-risk group estimated by the TIDE algorithm. **(G)** Comparison of the risk score of the response group with that of the nonresponse group. **(H)** Receiver operating characteristic (ROC) curve for predicting the response to immunotherapy based on the risk score for BLCA. \*\*\*\*, \*\*\*, \*\*, \* and ns represented  $p < 0.00001$ ,  $p < 0.0001$ ,  $p < 0.001$ ,  $p < 0.05$ , and not significant, respectively.



to evaluate the clinical data, and the results showed that patients in the low-risk group had better clinical outcomes than those in the high-risk group. We performed Kaplan–Meier analysis

on the training set, and the results showed that the low-risk group had a longer survival time than the high-risk group ( $p < 0.001$ ). The above data suggest that this model has good

clinical feasibility. We validated the accuracy and clinical usefulness of the signature model using several methods, including ROC curves for 1-, 3-, and 5-year OS and decision curve analysis. Finally, external validation of the model using the GEO BLCA cohorts further verified the prognostic ability of the model.

However, there were some limitations of this study. This study was retrospective, and further prospective studies are required to validate our findings, and in this study, only innate immunity genes were analyzed and adaptive immunity was not involved. In addition, some clinical characteristics, such as age and clinical stages, were not included in our model.

## CONCLUSION

In this study, we established an immune-related eight-gene signature model for BLCA, that could be used to predict the immune response and prognosis of BLCA patients.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

TY made substantial contributions to the conception, design, interpretation, and preparation of the final manuscript.

## REFERENCES

- Bae, H., Kim, B., Lee, H., Lee, S., Kang, H. S., and Kim, S. J. (2017). Epigenetically regulated Fibronectin leucine rich transmembrane protein 2 (FLRT2) shows tumor suppressor activity in breast cancer cells. *Sci. Rep.* 7:272. doi: 10.1038/s41598-017-00424-0
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Charoentong, P., Finotello, E., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18, 248–262. doi: 10.1016/j.celrep.2016.12.019
- Davis, S., and Meltzer, P. S. (2007). GEOquery: a bridge between the gene expression omnibus (GEO) and bioconductor. *Bioinformatics* 23, 1846–1847. doi: 10.1093/bioinformatics/btm254
- Donninger, H., Bonome, T., Radonovich, M., Pise-Masison, C. A., Brady, J., Shih, J. H., et al. (2004). Whole genome expression profiling of advanced stage papillary serous ovarian cancer reveals activated pathways. *Oncogene* 23, 8065–8077. doi: 10.1038/sj.onc.1207959
- Fumet, J. D., Truntzer, C., Yarchoan, M., and Ghiringhelli, F. (2020). Tumour mutational burden as a biomarker for immunotherapy: current data and emerging concepts. *Eur. J. Cancer* 131, 40–50. doi: 10.1016/j.ejca.2020.02.038
- Grivennikov, S. I., Greten, F. R., and Karin, M. (2010). Immunity, inflammation, and cancer. *Cell* 140, 883–899. doi: 10.1016/j.cell.2010.01.025
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi: 10.1016/j.cell.2011.02.013

BD, JL, WS, SZ, YM, JS, and MZ participated in the coordination of data acquisition and data analysis and reviewed the manuscript. DL reviewed and revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by a project cosponsored by the Henan Province and Ministry of Health (201601025) and medical science and technology program of Henan Provincial Science and Technology Department (162102310040).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.696912/full#supplementary-material>

**Supplementary Figure 1** | Calibration curves for evaluating the accuracy of the model in predicting 1-, 3-, and 5-year OS in the training (A–C) and testing sets (D–F).

**Supplementary Figure 2** | Analysis of immune cell infiltration in GEO BLCA cohorts GSE31684 (A) and GSE32894 (B).

**Supplementary Table 1** | The coefficients of the signature genes.

**Supplementary Table 2** | The clinical information of TCGA BLCA data.

**Supplementary Table 3** | The clinical information of GSE32894.

**Supplementary Table 4** | The clinical information of GSE31684.

- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* 14:7. doi: 10.1186/1471-2105-14-7
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi: 10.1038/s41591-018-0136-1
- Kim, K., Jeon, S., Kim, T. M., and Jung, C. K. (2018). Immune gene signature delineates a subclass of papillary thyroid cancer with unfavorable clinical outcomes. *Cancer* 10:494. doi: 10.3390/cancers10120494
- Lapinski, P. E., Oliver, J. A., Bodie, J. N., Marti, F., and King, P. D. (2009). The T-cell-specific adapter protein family: TSAd, ALX, and SH2D4A/SH2D4B. *Immunol. Rev.* 232, 240–254. doi: 10.1111/j.1600-065X.2009.00829.x
- Liang, D., McHugh, K. M., Brophy, P. D., Shaikh, N., Manak, J. R., Andrews, P., et al. (2019). DNA copy number variations in children with vesicoureteral reflux and urinary tract infections. *PLoS One* 14:e0220617. doi: 10.1371/journal.pone.0220617
- Lu, M., Ge, Q., Wang, G., Luo, Y., Wang, X., Jiang, W., et al. (2018). CIRBP is a novel oncogene in human bladder cancer inducing expression of HIF-1 $\alpha$ . *Cell Death Dis.* 9:1046. doi: 10.1038/s41419-018-1109-5
- Lv, Z., Pang, C., Wang, J., Xia, H., Liu, J., Yan, Q., et al. (2021). Identification of a prognostic signature based on immune-related genes in bladder cancer. *Genomics* 113, 1203–1218. doi: 10.1016/j.ygeno.2021.03.012
- Ma, W., Liang, J., Liu, J., Tian, D., and Chen, Z. (2021). Establishment and validation of an eight-gene metabolic-related prognostic signature model for lung adenocarcinoma. *Aging* 13, 8688–8705. doi: 10.18632/aging.202681
- Oda, T., Muramatsu, M. A., Isogai, T., Masuho, Y., Asano, S., and Yamashita, T. (2001). HSH2: a novel SH2 domain-containing adapter protein involved in tyrosine kinase signaling in hematopoietic cells. *Biochem. Biophys. Res. Commun.* 288, 1078–1086. doi: 10.1006/bbrc.2001.5890

- Pegram, H. J., Purdon, T. J., van Leeuwen, D. G., Curran, K. J., Giralt, S. A., Barker, J. N., et al. (2015). IL-12-secreting CD19-targeted cord blood-derived T cells for the immunotherapy of B-cell acute lymphoblastic leukemia. *Leukemia* 29, 415–422. doi: 10.1038/leu.2014.215
- Pettenati, C., and Ingersoll, M. A. (2018). Mechanisms of BCG immunotherapy and its outlook for bladder cancer. *Nat. Rev. Urol.* 15, 615–625. doi: 10.1038/s41585-018-0055-4
- Powles, T., Eder, J. P., Fine, G. D., Braitheh, F. S., Loriot, Y., Cruz, C., et al. (2014). MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* 515, 558–562. doi: 10.1038/nature13904
- Qiu, H., Hu, X., He, C., Yu, B., Li, Y., and Li, J. (2020). Identification and validation of an individualized prognostic signature of bladder cancer based on seven immune related genes. *Front. Genet.* 11:12. doi: 10.3389/fgene.2020.00012
- Qureshi, A. H., Liang, D., Canas, J., Hooks, J., Arrrregui, S. W., Saxena, V., et al. (2020). DCHS1 DNA copy number loss associated with pediatric urinary tract infection risk. *Inmate Immun.* 26, 473–481. doi: 10.1177/175342529017193
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Rouanne, M., Roumiguié, M., Houédé, N., Masson-Lecomte, A., Colin, P., Pignot, G., et al. (2018). Development of immunotherapy in bladder cancer: present and future on targeting PD(L)1 and CTLA-4 pathways. *World J. Urol.* 36, 1727–1740. doi: 10.1007/s00345-018-2332-5
- Rouprêt, M., Babjuk, M., Burger, M., Capoun, O., Cohen, D., Compérat, E. M., et al. (2021). European association of urology guidelines on upper urinary tract urothelial carcinoma: 2020 update. *Eur. Urol.* 79, 62–79. doi: 10.1016/j.eururo.2020.05.042
- Santin, A. D., Zhan, F., Cane, S., Bellone, S., Palmieri, M., Thomas, M., et al. (2005). Gene expression fingerprint of uterine serous papillary carcinoma: identification of novel molecular markers for uterine serous cancer diagnosis and therapy. *Br. J. Cancer* 92, 1561–1573. doi: 10.1038/sj.bjc.6602480
- Seager, R. J., Hajal, C., Spill, F., Kamm, R. D., and Zaman, M. H. (2017). Dynamic interplay between tumour, stroma and immune system can drive or prevent tumour progression. *Converg. Sci. Phys. Oncol.* 3:034002. doi: 10.1088/2057-1739/aa7e86
- Shapiro, M. J., Powell, P., Ndubuizu, A., Nzerem, C., and Shapiro, V. S. (2004). The ALX Src homology 2 domain is both necessary and sufficient to inhibit T cell receptor/CD28-mediated up-regulation of RE/AP. *J. Biol. Chem.* 279, 40647–40652. doi: 10.1074/jbc.M404198200
- Shen, C., Liu, J., Wang, J., Yang, X., Niu, H., and Wang, Y. (2020). The analysis of PTPN6 for bladder cancer: an exploratory study based on TCGA. *Dis. Markers* 2020:4312629. doi: 10.1155/2020/4312629
- Shen, S., Wang, G., Zhang, R., Zhao, Y., Yu, H., Wei, Y., et al. (2019). Development and validation of an immune gene-set based prognostic signature in ovarian cancer. *EBioMedicine* 40, 318–326. doi: 10.1016/j.ebiom.2018.12.054
- Stratton, M. R., Campbell, P. J., and Futreal, P. A. (2009). The cancer genome. *Nature* 458, 719–724. doi: 10.1038/nature07943
- Szumaska, D., Cioroch, M., Keeling, A., Prat, A., Seidah, N. G., and Bhattacharya, S. (2017). Pcsk5 is required in the early cranio-cardiac mesoderm for heart development. *BMC Dev. Biol.* 17:6. doi: 10.1186/s12861-017-0148-y
- Tran, L., Xiao, J. F., Agarwal, N., Duex, J. E., and Theodorescu, D. (2021). Advances in bladder cancer biology and therapy. *Nat. Rev. Cancer* 21, 104–121. doi: 10.1038/s41568-020-00313-1
- Williams, S. G., and Stein, J. P. (2004). Molecular pathways in bladder cancer. *Urol. Res.* 32, 373–385. doi: 10.1007/s00240-003-0345-y
- Wu, Y., Davison, J., Qu, X., Morrissey, C., Storer, B., Brown, L., et al. (2016). Methylation profiling identified novel differentially methylated markers including OPCML and FLRT2 in prostate cancer. *Epigenetics* 11, 247–258. doi: 10.1080/15592294.2016.1148867
- Yang, W., Sun, T., Cao, J., Liu, F., Tian, Y., and Zhu, W. (2012). Downregulation of miR-210 expression inhibits proliferation, induces apoptosis and enhances radiosensitivity in hypoxic human hepatoma cells in vitro. *Exp. Cell Res.* 318, 944–954. doi: 10.1016/j.yexcr.2012.02.010
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, L., and Zhang, Z. (2019). Recharacterizing tumor-infiltrating lymphocytes by single-cell RNA sequencing. *Cancer Immunol. Res.* 7, 1040–1046. doi: 10.1158/2326-6066.CIR-18-0658
- Zhang, Y., and Zhang, Z. (2020). The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell. Mol. Immunol.* 17, 807–821. doi: 10.1038/s41423-020-0488-6
- Zheng, A., Zhang, L., Song, X., Wang, Y., Wei, M., and Jin, F. (2019). Clinical implications of a novel prognostic factor AIFM3 in breast cancer patients. *BMC Cancer* 19:451. doi: 10.1186/s12885-019-5659-4
- Zhu, J., Wang, H., Ma, T., He, Y., Shen, M., Song, W., et al. (2020). Identification of immune-related genes as prognostic factors in bladder cancer. *Sci. Rep.* 10:19695. doi: 10.1038/s41598-020-76688-w

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Dong, Liang, Li, Song, Zhao, Ma, Song, Zhu and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.