scientific reports

Check for updates

OPEN

A novel efferocytosis-related gene signature for predicting prognosis and therapeutic response in bladder cancer

Weitao Yu[1,2,3,5], Dongnuan Yao[1,2,3,5], Xueming Ma[1,2,3], Juanjuan Hou[4] & Junqiang Tian[1,2,3]✉

Efferocytosis, the process by which phagocytes like macrophages and dendritic cells clear apoptotic cells, is crucial for maintaining tissue homeostasis. However, its function in bladder cancer (BLCA) remains unclear and warrants further exploration. This study seeks to establish a prognostic and treatment response signature based on efferocytosis-related genes (EFRGs) for bladder cancer patients. BLCA-related datasets were sourced from the Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). A comprehensive analysis was performed on 28 prognostic EFRGs. Clustering analysis was carried out using ConsensusClusterPlus. Prognostic differentially expressed genes (DEGs) were identified based on expression variations across the subtypes. A prognostic model was subsequently developed using least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression. Lastly, a thorough analysis was conducted to explore the relationship between risk scores and the tumor immune microenvironment, somatic mutations, as well as responses to immunotherapy and chemotherapy. Consensus clustering revealed two efferocytosis subtypes, Cluster A and Cluster B, and identified 61 prognostic DEGs between them. A risk scoring model, incorporating four key DEGs—SERPINE2, DPYSL3, CTSE, and KRT16—was constructed and validated. This model successfully stratified patients into high-risk and low-risk groups, with high-risk patients showing worse prognosis, increased immune infiltration, and higher immune checkpoint gene expression. The risk scores also provide insights into patient responsiveness to treatment. In conclusion, we identified four key genes—SERPINE2, DPYSL3, CTSE, and KRT16—that can be used to develop a prognostic model for bladder cancer. These findings may provide valuable molecular targets for the clinical diagnosis and therapeutic strategies of bladder cancer.

Keywords Bladder Cancer, Efferocytosis, Prognostic features, Bioinformatics, Treatment

Bladder cancer (BLCA) ranks as the ninth most common cancer globally, with 613,791 new cases and 220,349 deaths reported in 2022[1]. Approximately 75% of newly diagnosed bladder cancer cases are non-muscle invasive bladder cancer (NMIBC), while around 25% are classified as muscle-invasive bladder cancer (MIBC)[2]. Previous studies indicate that NMIBC has a high 5-year recurrence and progression rate of 78% and 45%, respectively, complicating treatment and increasing the risk of recurrence[3]. Patients with NMIBC who progress to muscle-invasive stages tend to have a worse prognosis compared to those initially diagnosed with primary muscle-invasive disease[4,5]. Bacillus Calmette-Guérin (BCG) immunotherapy, the only established and widely adopted conservative treatment, is effective in preventing the progression of high-risk NMIBC[6]. MIBC, characterized by high-grade tumors, locally advanced disease, or metastasis, typically requires neoadjuvant chemotherapy and radical cystectomy, along with additional biomarker-guided immune checkpoint inhibitors and targeted combination therapies[7]. Despite advancements in targeted therapy, radiotherapy, and immunotherapy, challenges remain in treating bladder cancer due to tumor recurrence and metastasis, leading to treatment failure and poor survival outcomes[8,9].

[1]Department of Urology, The Second Hospital of Lanzhou University, No.82 Cuiyingmen, Lanzhou 730030, China. [2]Gansu Province Clinical Research Center for urinary system disease, Lanzhou, China. [3]The Second Hospital and Clinical Medical School, Lanzhou University, Lanzhou, China. [4]Department of Molecular Diagnostic Medicine, The People's Hospital of Qingyang City, Qingyang 745000, China. [5]Weitao Yu and Dongnuan Yao contributed equally to thiswork and share first authorship. ✉email: ery_tianjq@lzu.edu.cn

Efferocytosis is the process by which professional efferocytes, such as macrophages and dendritic cells, along with non-professional efferocytes, engulf apoptotic and necrotic cells from damaged tissues[10]. This tightly regulated process is crucial for homeostasis, and its dysfunction can result in the buildup of dead cells, secondary necrosis, and inflammatory conditions like atherosclerosis, age-related inflammation, infections, lupus, obesity, and cancer[11,12]. Conversely, the anti-inflammatory effects of efferocytosis-related tumor-associated macrophages (TAMs) may inhibit antitumor responses, indicating that shifting efferocytosis towards a more inflammatory state could be advantageous[13]. Mechanistically, Efferocytosis activates the PPARγ and LXR signaling pathways, promoting a metabolic preference for fatty acid oxidation in TAMs and driving their polarization toward the M2 phenotype[14]. M2 TAMs foster an immunosuppressive microenvironment by secreting anti-inflammatory cytokines such as IL-10 and TGF-β, while suppressing pro-inflammatory cytokines like IL-12 and TNF-α[15]. In addition, metabolites released during efferocytosis—such as the cholesterol derivative 27-hydroxycholesterol—activate LXR signaling, which in turn inhibits the proliferation and function of $CD8^+$ T cells[16]. Efferocytosis also promotes the expression of immune checkpoint ligands PD-L1 and PD-L2 in cancer cells via the MerTK/Axl signaling axis, thereby dampening T cell activity[17]. Through these coordinated signaling events, efferocytosis contributes to the establishment of an immunosuppressive and tolerogenic tumor microenvironment (TME), ultimately promoting tumor initiation and progression. Furthermore, recent studies have shown that blockade of CD276 impairs the efferocytic function of TAMs and reverses their immunosuppressive state via the AXL/MerTK–JUN pathway. Notably, combined blockade of CD276 and PD-1 significantly enhances antitumor immune responses in a murine bladder cancer model, offering strong potential for clinical translation[18]. However, the role of efferocytosis-related genes (EFRGs) in shaping clinical outcomes remains underexplored, underscoring the need to define their prognostic and therapeutic relevance in human bladder cancer.

In recent years, advances in computer science and statistics have enabled their broad application across scientific disciplines[19–21]. Bioinformatics, an interdisciplinary field combining biology, computer science, mathematics, and statistics, plays a central role in developing computational algorithms, analytical tools, and databases for the interpretation of large-scale biological data[22]. These technological advances have made it possible to classify molecular subtypes of bladder cancer based on genetic profiles, enabling more accurate prognostic assessments[23,24]. Consensus clustering, a widely used unsupervised learning method, allows stratification of patients into molecular subtypes using gene expression and survival data, facilitating subtype discovery and comparative analysis[25]. Moreover, the identification and analysis of efferocytosis-related molecular subtypes have already been validated and applied in various cancers, including gastrointestinal tumors, gliomas, and leukemia, demonstrating significant predictive potential[26–28]. In this study, the primary objective is to develop a prognostic model based on EFRGs to predict survival outcomes and therapeutic responsiveness in bladder cancer patients. Secondary objectives include characterizing the immune microenvironmental differences between EFRGs clusters, exploring the association of the EFRGs signature with immunotherapy and chemotherapy responses, and identifying potential therapeutic targets for high-risk subpopulations. Our findings offer valuable insights into the role of efferocytosis, helping to identify potential biomarkers and opening new avenues for treatment strategies and prognosis in BLCA.

Materials and methods

Data collection

The mRNA expression data, mutation data, and clinical information for 403 bladder cancer samples, along with 19 normal bladder samples, were retrieved and downloaded from the Cancer Genome Atlas database (TCGA, https://portal.gdc.cancer.gov/). Gene expression profiles and corresponding clinical data from the GSE13507 (165 tumor and 68 normal samples), GSE32548 (130 tumor samples), and GSE32894 (224 tumor samples) datasets were downloaded from Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/). Samples with complete clinical and expression data, including non-missing follow-up information, were retained for downstream analysis. The combined dataset of TCGA-BLCA and GSE13507 served as the internal cohort, while GSE32548 and GSE32894 were independently used as external validation cohorts. The original data were preprocessed by adjusting the background, performing quantile normalization, and applying the "Combat" algorithm to eliminate batch effects, using the "limma" and "sva" R packages[29,30]. For TCGA, Transcripts Per Million (TPM)-normalized gene expression data were log2-transformed and then quantile-normalized to ensure consistent distribution across samples. This approach aligned genes with identical quantile ranks to exhibit comparable expression values. For the GEO microarray datasets, raw expression data were similarly log2-transformed and quantile-normalized to reduce intra-dataset variability. To further address inter-platform technical heterogeneity, batch effect correction was performed using the ComBat algorithm, which effectively adjusts for platform-specific biases while preserving biologically meaningful variation across cohorts.

Prognostic, tumor mutation burden, copy number variation, and co-expression analyses of EFRGs

We retrieved 156 EFRGs from the Kyoto Encyclopedia of Genes and Genomes database (KEGG, https://www.genome.jp/kegg/) and utilized univariate Cox regression analysis to identify prognostic efferocytosis-related genes. The "maftools" R package was used to generate a waterfall plot, examining mutation frequency and types in these prognostic efferocytosis genes[31]. Additionally, copy number matrix files were acquired from the Xena website (https://xena.ucsc.edu/) to construct a copy number variation (CNV) frequency plot, displaying the amplification or deletion frequency of prognostic efferocytosis genes. The "Reshape2" and "ggpubr" R packages were utilized to analyze and identify significant differences between normal and tumor tissues. Gene expression levels were analyzed using the Wilcoxon test, with $P < 0.05$ deemed statistically significant. Finally, co-expression analysis of prognostic EFRGs was performed using the STRING database (https://cn.string-db.org/).

Consensus clustering analysis of EFRGs

Consensus clustering was carried out using the "ConsensusClusterPlus" R package. Using the expression profiles of prognostic EFRGs, the optimal K value was determined to divide the samples into distinct clusters. The selection criteria included low coefficient of variation, high clustering stability, and relatively flat cumulative distribution function (CDF) curves[25].

Principal component analysis and clinical characteristics of efferocytosis subtypes

Principal Component Analysis (PCA) was performed using the "ggplot2" R package to visualize the distribution of samples across efferocytosis subtypes[32]. The "pheatmap" R package was used to create a heatmap, displaying the distribution of efferocytosis genes and clinical characteristics across the clusters. Survival differences between the subtypes were compared using Kaplan-Meier analysis, performed with the "survminer" and "survival" R packages.

Gene set variation analysis and single-sample gene set enrichment analysis (GSVA and ssGSEA)

GSVA, an unsupervised and non-parametric method, was used to analyze gene sets and identify enriched biological pathways, translating them into functional levels[33]. The gene sets "c2.cp.kegg_legacy.v2023.2.Hs.symbols.gmt" and "c5.go.v2023.2.Hs.symbols.gmt" were obtained from the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb) for enrichment analysis[34–38]. Significant enrichment was defined by an adjusted $P < 0.05$. Additionally, the ssGSEA algorithm analyzed immune gene sets, assessing immune cell types, pathways, functions, and checkpoints[39].

DEGs identification and consensus clustering analysis

DEGs between the efferocytosis subtypes were identified using the "limma" R package, applying the thresholds of $|logFC| > 1.5$ and $FDR < 0.05$. Univariate Cox regression analysis identified prognostic DEGs. Using the expression profiles of prognostic DEGs, bladder cancer patients were also classified into distinct gene subtypes with the "ConsensusClusterPlus" R package.

Construction and validation of the efferocytosis-related risk model

We merged bladder samples from the TCGA-BLCA and GSE13507 cohorts. The bladder cancer samples were randomly split into training (284 patients) and test (283 patients) sets in a 1:1 ratio. A risk scoring model was constructed using the training set by applying Least Absolute Shrinkage and Selection Operator (LASSO) and multivariate Cox regression on the prognostic DEGs. To identify independent prognostic genes while addressing potential multicollinearity, we applied LASSO Cox proportional hazards regression using the "glmnet" R package. This method incorporates L1 regularization, which effectively shrinks the coefficients of correlated variables toward zero, thereby reducing redundancy and multicollinearity among genes. The optimal value of the regularization parameter (λ) was determined through 10-fold cross-validation, using the value of λ that minimized the partial likelihood deviance. Genes with non-zero coefficients at this optimal λ were selected for inclusion in the final multivariate Cox regression model. Stepwise multivariate Cox regression based on the Akaike Information Criterion (AIC) was employed to further refine the model and identify the final set of target genes. Additionally, we applied stringent selection criteria, requiring the 3-year AUC to exceed 0.65 with $P < 0.01$ in the internal training set and the 3-year AUC to exceed 0.60 with $P < 0.01$ in the internal test set. The formula is as follows: $Risk = \Sigma (Coefi * Exp)$.

Coefi denotes the risk coefficient, while Exp refers to the expression level. Training set samples were split into high-risk and low-risk groups based on the median score, and Kaplan-Meier survival analysis was used to compare overall survival (OS) between them. ROC curves were generated to assess the model's prognostic accuracy. The model was further validated in the test set and external validation cohorts (GSE32548 and GSE32894) using Kaplan-Meier survival analysis and ROC curves with the "timeROC," "survival," and "survminer" R packages. A nomogram was constructed using the "survival," "rms," and "regplot" R packages to estimate 1-year, 3-year, and 5-year survival rates.

Analysis of cancer stem cell index, tumor mutation burden, and TME between risk groups

The cancer stem cell (CSC) index, reflecting the similarity between tumor cells and stem cells, was analyzed for its correlation with the risk score. The "maftools" R package was used to calculate the tumor mutational burden (TMB) for each bladder cancer sample, followed by a differential analysis between the two risk clusters. We also used the ESTIMATE algorithm to calculate the ImmuneScore, StromalScore, and ESTIMATEScore for each sample, then conducted a differential analysis between the high-risk and low-risk clusters[40].

Immune infiltration analysis, immune checkpoint evaluation, and drug sensitivity correlation analysis between risk groups

The CIBERSORT algorithm was employed to evaluate the relative proportions of 23 immune cell types in each bladder cancer sample[41]. Next, we performed correlation analyses between the risk score and immune cell abundance, as well as between immune cell abundance and the expression of genes involved in the model. We analyzed differential expression of immune checkpoints between high-risk and low-risk groups using the "limma," "ggplot2," "ggpubr," and "ggExtra" R packages. Additionally, we used the "pRRophetic" R package to calculate half maximal inhibitory concentration (IC50) values for commonly used chemotherapeutic drugs, predicting chemotherapy response based on tumor gene expression profiles[42].

Statistical analysis

All statistical analyses were conducted using R software (version 4.3.3) within the RStudio environment (version 2024.09.0). Relevant R packages were downloaded from the Comprehensive R Archive Network (CRAN) and Bioconductor repositories, and applied for data processing, statistical modeling, and visualization. Spearman's correlation analysis was applied to evaluate correlations, while independent sample t-tests or Mann-Whitney U tests were used to compare differences between two groups. Kaplan-Meier analysis was employed for survival analysis. $P < 0.05$ was considered statistically significant.

Results

Tumor mutation burden, copy number variation and co-expression analyses of prognostic EFRGs in bladder cancer

A total of 156 EFRGs were sourced from the KEGG database, and 28 prognostic EFRGs were identified through Kaplan-Meier survival analysis and univariate Cox regression, using data from the integrated TCGA-BLCA and GSE13507 cohorts (Fig. 1A) (Supplementary Table 1). In the TCGA-BLCA dataset, we examined somatic mutations and CNVs in these 28 genes. Among the 404 samples analyzed, 106 (26.24%) exhibited mutations. RXRA had the highest mutation frequency at 5%, while no mutations were detected in CALR and RAB17 in these samples (Fig. 1B). CNVs were prevalent across the 28 EFRGs. Among the genes with CNV frequencies exceeding 10%, RAB17 had the highest frequency of deletions, whereas ARNT, CPT1A, ATP11A, GAS6, and SPHK1 exhibited the highest frequencies of gains (Fig. 1C). We also analyzed the expression differences of these genes between cancerous and normal tissues in the TCGA-BLCA and GSE13507 cohorts (Fig. 1D). Sixteen genes exhibited differential expression: CAMK2B, TYRO3, CALR, RAB17, CPT1A, and CPT1B were upregulated in tumors, while MAPK3, MFGE8, GAS6, PTGER2, MAPK12, DUSP2, ATP11A, PROS1, GULP1, and SIRPA were upregulated in normal tissues. Finally, we constructed an interaction network for the 28 prognostic EFRGs using STRING (Fig. 1E).

Identification of efferocytosis subtypes in bladder cancer

The efferocytosis network demonstrated interactions among EFRGs and highlighted their prognostic relevance in bladder cancer (Fig. 2A). We conducted consensus clustering analysis on the merged cohorts. Based on the expression levels of 28 EFRGs, we found K = 2 to offer the most stable classification, identifying two distinct clusters: EFRGs_cluster A ($n = 251$) and EFRGs_cluster B ($n = 316$) (Fig. 2B) (Supplementary Table 2). The CDF curve is shown in Supplementary Fig. 1A. We also validated the clustering performance based on the 28 EFRGs in the external datasets GSE32548 and GSE32894 (Supplementary Fig. 1C-F). PCA further demonstrated clear differences between the two clusters (Fig. 2C). Kaplan-Meier survival curves revealed significant OS differences between the clusters, with patients in EFRGs_cluster B showing a better prognosis (Fig. 2D). Heatmap displayed the distribution of the 28 EFRGs and clinical characteristics across clusters (Fig. 2E). EFRGs_Cluster B had lower staging, grading, and age, while EFRGs_cluster A had a higher proportion of female patients. Boxplots indicated significant differences in EFRG expression between the subtypes, with 27 genes showing notable differences. Specifically, PANX1, SPHK1, ARNT, MAPK3, MFGE8, GAS6, MAP2K1, PTGER2, CAMK2B, MAPK11, MAPK12, ATP11A, PROS1, CALR, GULP1, VPS18, CPT1A, CPT1C, SIRPA, and SIRPG were highly expressed in EFRGs_cluster A, while PLA2G6, RXRA, ALOX5, DUSP2, XKR6, RAB17, and CPT1B were highly expressed in EFRGs_cluster B (Fig. 2F).

Immune cell infiltration and GSVA enrichment analysis between efferocytosis subtypes

Using the ssGSEA algorithm, we assessed the relative abundance of 23 immune cells in EFRGs_clusters A and B. Significant differences were observed in 22 immune cells, with EFRGs_cluster A showing higher infiltration of various immune cells, including activated B cells, CD4 T cells, CD8 T cells, dendritic cells, eosinophils, γδ T cells, macrophages, and others. In contrast, CD56 dim NK cells, monocytes, and Type 17 T helper cells were more abundant in EFRGs_cluster B (Fig. 3A). Interestingly, despite its poorer prognosis, Cluster A displayed a markedly higher level of immune infiltration.

GO analysis revealed that EFRGs_cluster A was significantly enriched in biological processes and molecular functions, particularly those related to cytoskeletal reorganization, ECM remodeling, and response to amyloid-beta (Fig. 3B). The KEGG analysis results reveal significant differences in pathway enrichment between EFRGs_Cluster A and B (Fig. 3C). EFRGs_Cluster A is mainly enriched in pathways associated with cell structure, tumor progression, cardiomyopathies, and immune responses, including Gap Junction, Glioma, Actin Cytoskeleton Regulation, Cancer Pathways, ECM-Receptor Interaction, and Chemokine Signaling. In contrast, EFRGs_Cluster B is primarily enriched in metabolic pathways, such as Retinol, Linoleic Acid, α-Linolenic Acid Metabolism, Peroxisome, and Glycerophospholipid Metabolism.

Identification of prognostic DEGs between efferocytosis subtypes and gene subtype classification

We used the "limma" R package to identify 141 DEGs associated with EFRGs_subtypes, and univariate Cox regression analysis uncovered 61 prognostic DEGs. Consensus clustering, based on these DEGs, classified bladder cancer samples into two gene_subtypes: A ($n = 267$) and B ($n = 300$) (Fig. 4A) (Supplementary Table 3). The CDF curve is shown in Supplementary Fig. 1B. Kaplan-Meier survival curves indicated that gene_cluster A had significantly shorter OS compared to gene_cluster B ($P = 0.001$; Fig. 4B). Clinically, gene_cluster A was associated with higher staging, grading, and age, with a higher proportion of female patients (Fig. 4C).

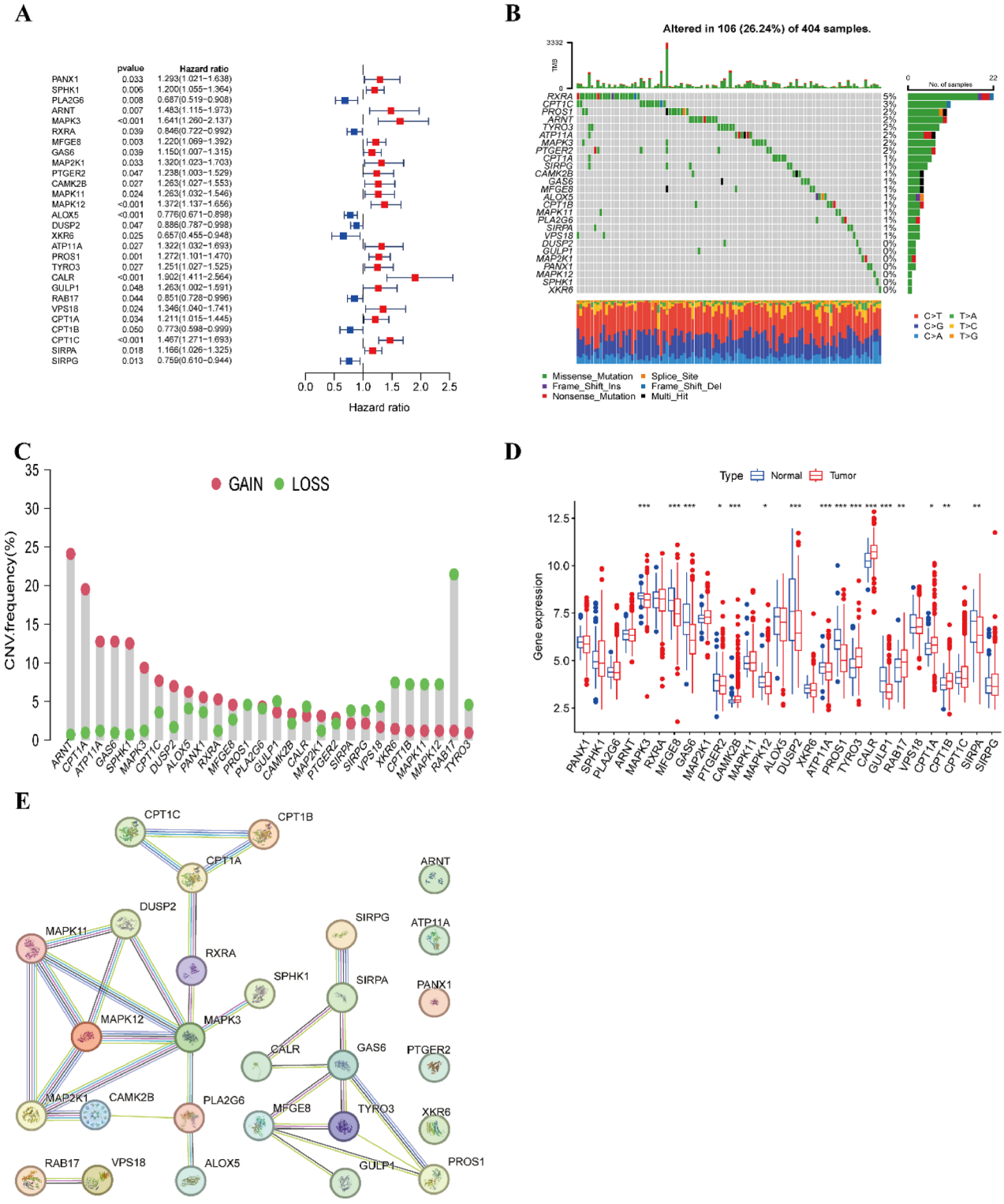

Fig. 1. Characteristics of Efferocytosis in Bladder Cancer. (**A**) 28 EFRGs associated with prognosis. (**B**) Somatic mutation status of prognosis-related EFRGs in 404 bladder cancer samples. (**C**) Frequency of CNVs of prognosis-related EFRGs. (**D**) Expression levels of prognosis-related EFRGs in normal (blue) and tumor (red) groups from the TCGA-BLCA dataset. (**E**) Co-expression analysis of prognosis-related EFRGs using the online tool STRING. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$.

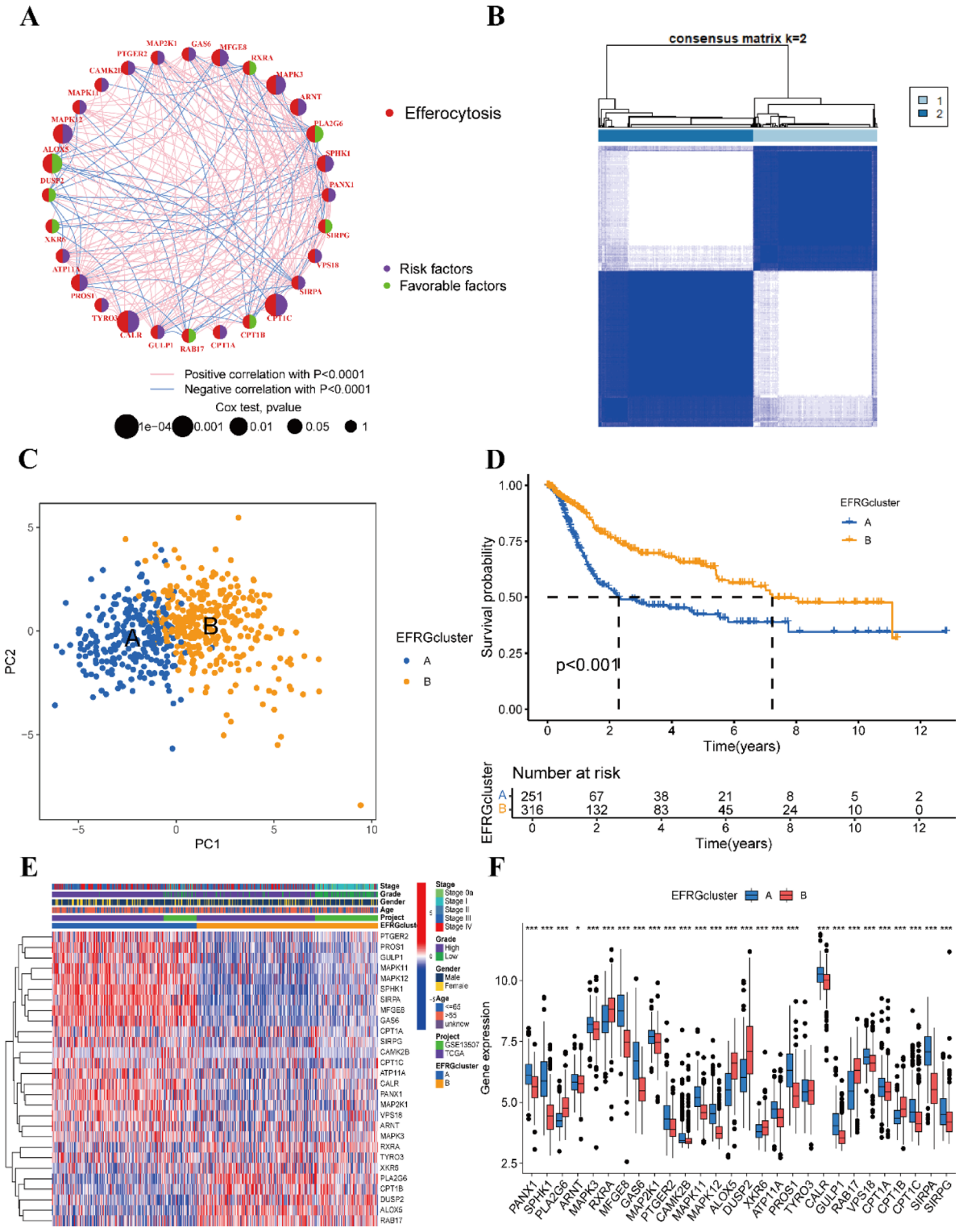

Fig. 2. Identification of Efferocytosis Subtypes and Comparison of Clinical Characteristics and Expression Levels of Prognosis-Related EFRGs between Two Subtypes. (**A**) Prognostic value and gene interaction of prognosis-related EFRGs in bladder cancer patients. (**B**) Unsupervised clustering analysis of prognosis-related EFRGs. The consensus matrix heatmap defines two EFRGs_clusters (k = 2) and their associated areas. (**C**) PCA of the expression distribution of prognosis-related EFRGs. (**D**) Kaplan-Meier curves of overall survival (OS) for the two EFRGs_subtypes ($P < 0.001$). (**E**) Heatmap showing differences in clinical characteristics and expression levels of EFRGs between EFRGs_cluster A and B. (**F**) Box plot showing differences in expression levels of prognosis-related EFRGs between EFRGs_cluster A and B. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

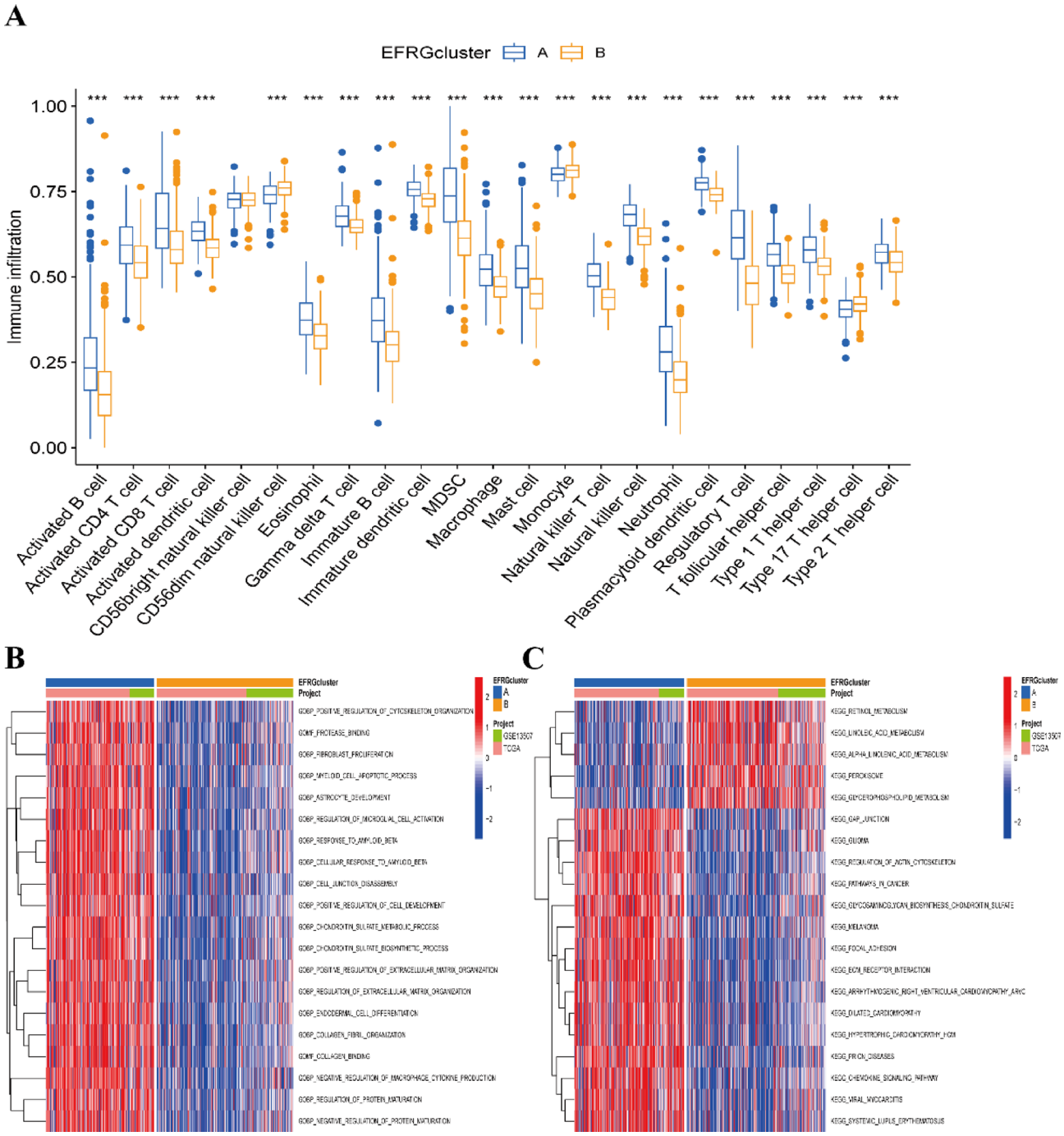

Fig. 3. Immune Cell Infiltration and Functional Analysis between EFRGs_Clusters. (**A**) Differences in infiltration of 23 immune cell types between EFRGs_clusters. (**B**) Differential enrichment of GO terms based on GSVA analysis. (**C**) Differential enrichment of the KEGG pathway based on GSVA analysis. In the heatmap, red indicates activated pathways, and blue indicates inhibited pathways. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$.

Construction and validation of the risk scoring model

A Sankey diagram was used to visualize the distribution of bladder cancer samples across various classifications (Fig. 5A). Samples from the merged cohorts were randomly split into training and internal validation groups at a 1:1 ratio (Supplementary Table 4). LASSO analysis and multivariate Cox regression analysis of 61 prognostic DEGs led to the selection of four model genes: SERPINE2, DPYSL3, CTSE, and KRT16 (Fig. 5B, C). The formula is as follows: Risk Score = (0.1145 × SERPINE2 expression) + (0.1344 × DPYSL3 expression) - (0.1021 × CTSE expression) + (0.0834 × KRT16 expression).

Patients were categorized into high-risk and low-risk groups according to the median risk score in the training set, with high-risk patients showing poorer prognosis across all cohorts (Fig. 5D-K). Kaplan-Meier

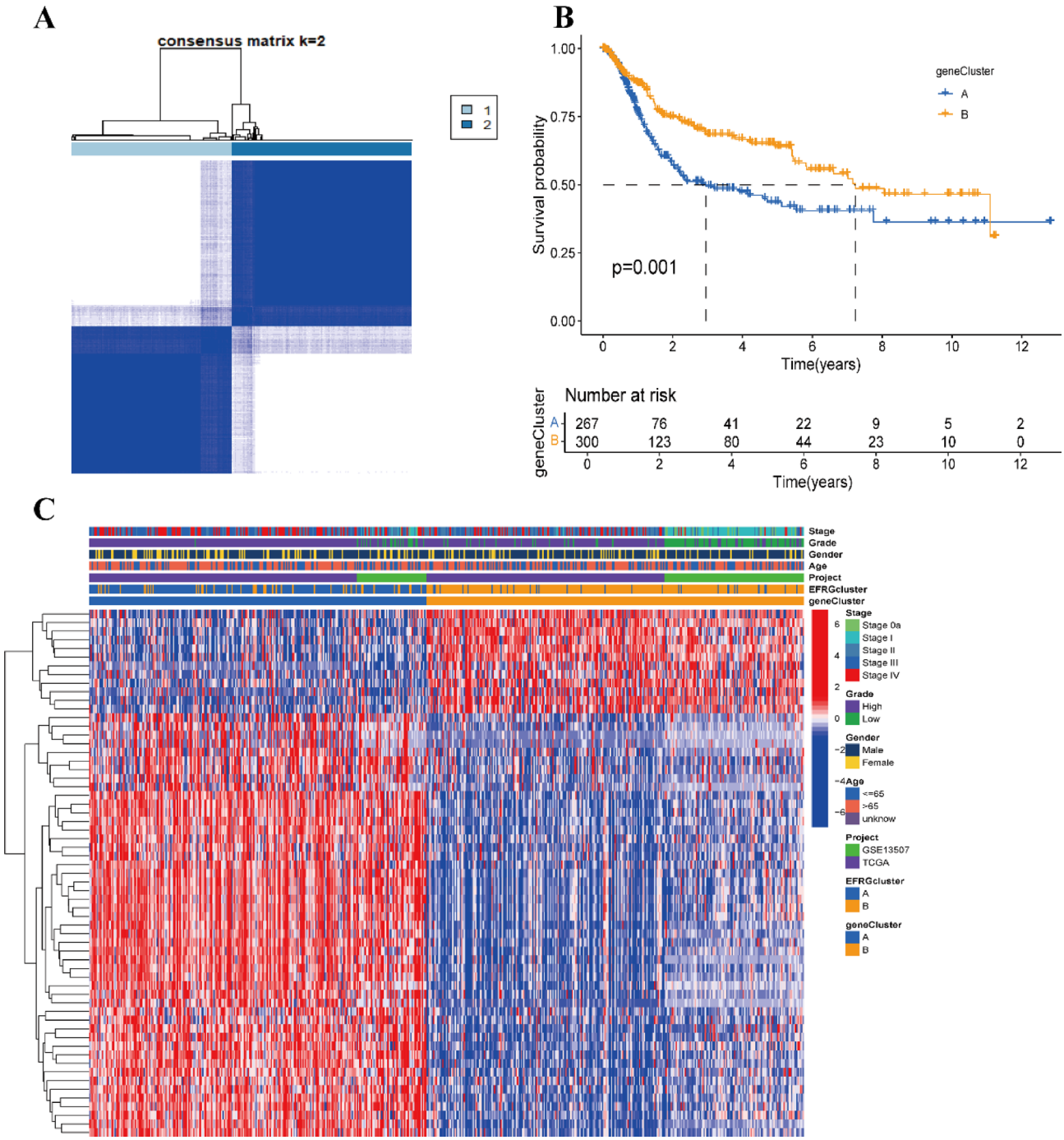

Fig. 4. Identification of Gene Subtypes and Comparison of Clinical Characteristics between Two Gene Subtypes. (**A**) Unsupervised clustering analysis of prognostic differentially expressed genes (DEGs) between two EFRGs_subtypes. The consensus matrix heatmap defines two Gene_clusters (k = 2) and their associated areas. (**B**) Kaplan-Meier curves of overall survival (OS) for the two gene subtypes ($P = 0.001$). (**C**) Heatmap showing differences in clinical characteristics and expression levels of prognostic DEGs between Gene_cluster A and B. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

analysis confirmed shorter OS in high-risk patients ($P < 0.001$, $P = 0.023$, $P < 0.001$, $P < 0.001$, Fig. 5L-O). ROC analysis demonstrated the model's robust prognostic capability, with AUCs exceeding 0.60 in all cohorts and surpassing 0.70 in external validation cohorts (Fig. 5P-S). Specifically, the 1-, 3-, and 5-year AUCs in the internal training set were 0.674, 0.718, and 0.707, respectively; in the internal validation set, they were 0.635, 0.619, and 0.602. For the external validation cohorts, the AUCs were 0.853, 0.830, and 0.772 in GSE32548, and 0.821, 0.889, and 0.799 in GSE32894, respectively. The AUCs exceeded 0.65 in the internal training set and remained above 0.70 in the external cohorts, indicating favorable predictive accuracy and generalizability of the model.

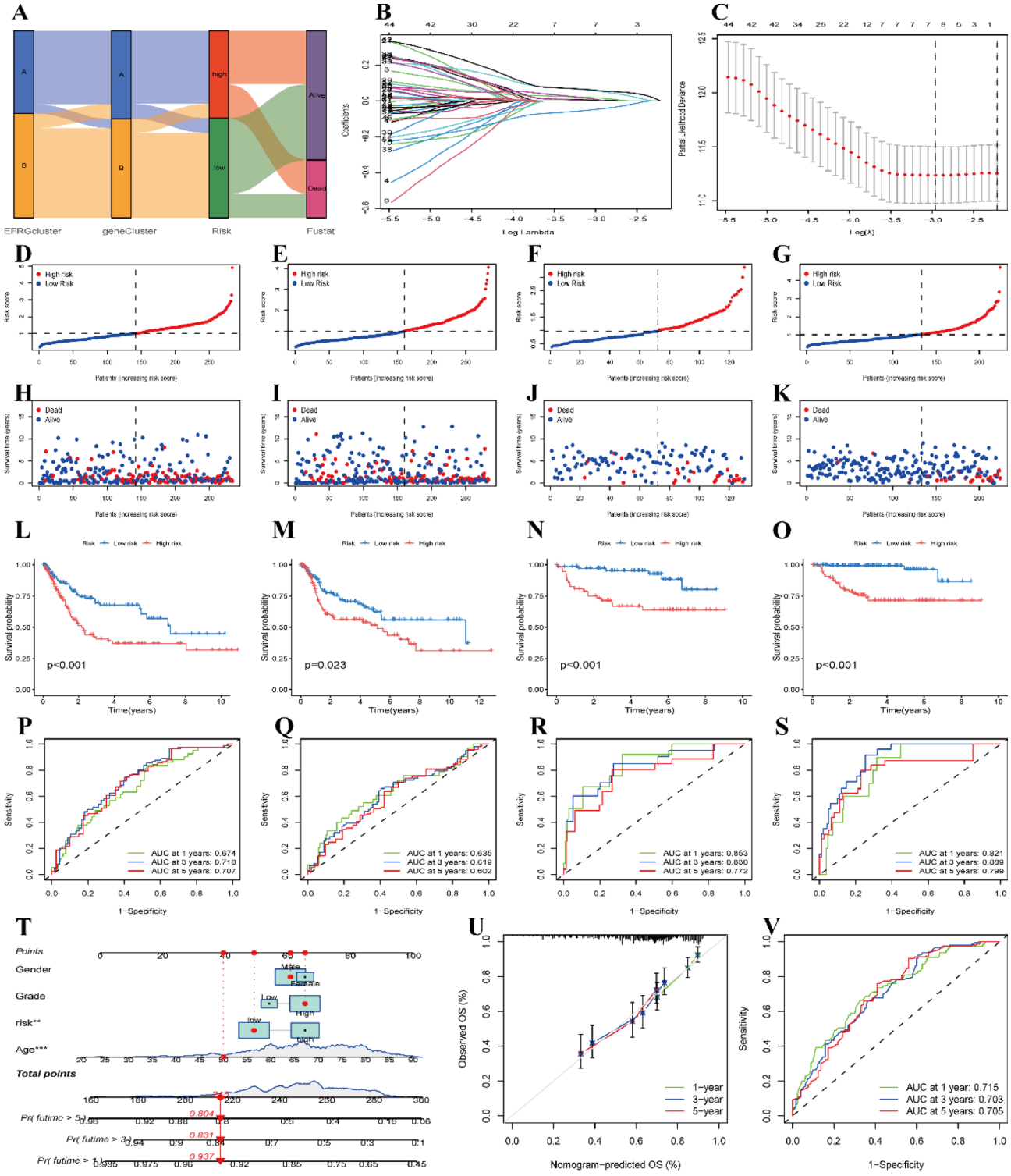

Fig. 5. Construction of Prognostic Features for Bladder Cancer Patients Based on DEGs between EFRGs_ Clusters. (**A**) Sankey diagram showing the distribution of bladder cancer patient subtypes and their prognosis. (**B**) Lasso coefficient plot. (**C**) Selection of the optimal log Lambda value in the training set using ten-fold cross-validation. (**D**–**G**) Classification of bladder cancer patients in the training set, internal validation set, and external validation sets (GSE32548 and GSE32894) into high-risk and low-risk groups based on the median risk score in the training set. (**H**–**K**) Distribution of survival status of bladder cancer patients with different risks in the training set, internal validation set, and external validation sets (GSE32548 and GSE32894). (**L**–**O**) Kaplan-Meier survival curves of high-risk and low-risk groups in the training set, internal validation set, and external validation sets (GSE32548 and GSE32894). (**P**–**S**) Time-dependent ROC curves in the training set, internal validation set, and external validation sets (GSE32548 and GSE32894). (**T**) Nomogram predicting 1-year, 3-year, and 5-year overall survival rates. (**U**,**V**) Calibration curves and ROC curves for the accuracy of the nomogram predictions. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$.

To improve clinical applicability, we developed a nomogram that integrates age, gender, pathological grade, and risk score, achieving accurate predictions for 1-year, 3-year, and 5-year OS with AUCs of 0.715, 0.703, and 0.705, respectively (Fig. 5T-V). The risk scores for samples from each dataset are provided in Supplementary Table 5.

Analysis of cancer stem cell indices, tumor mutation burden, and tumor microenvironment between risk groups

The presence of CSCs is a major factor contributing to tumor treatment difficulty and recurrence. Correlation analysis between the CSC index and risk score revealed a negative correlation (R = -0.27, P = 3.5e-08) (Fig. 6A).

Although tumors with high TMB are generally thought to respond better to immunotherapy and have improved prognoses, differential analysis revealed no significant difference in TMB between the high- and low-risk groups (P = 0.47). However, In the low-risk group, TMB showed a positive correlation with the risk score (R = 0.27, P = 0.00013), while in the high-risk group, it exhibited a negative correlation (R = -0.29, P = 3e-06) (Fig. 6B-E). Waterfall plots for both groups revealed that the top 10 genes with the highest mutation frequencies were identical: TP53, TTN, KMT2D, MUC16, KDM6A, ARID1A, PIK3CA, SYNE1, RB1, and RYR2 (Fig. 6F-G). The ESTIMATE algorithm was used to assess StromalScore, ImmuneScore, and ESTIMATEScore, revealing that

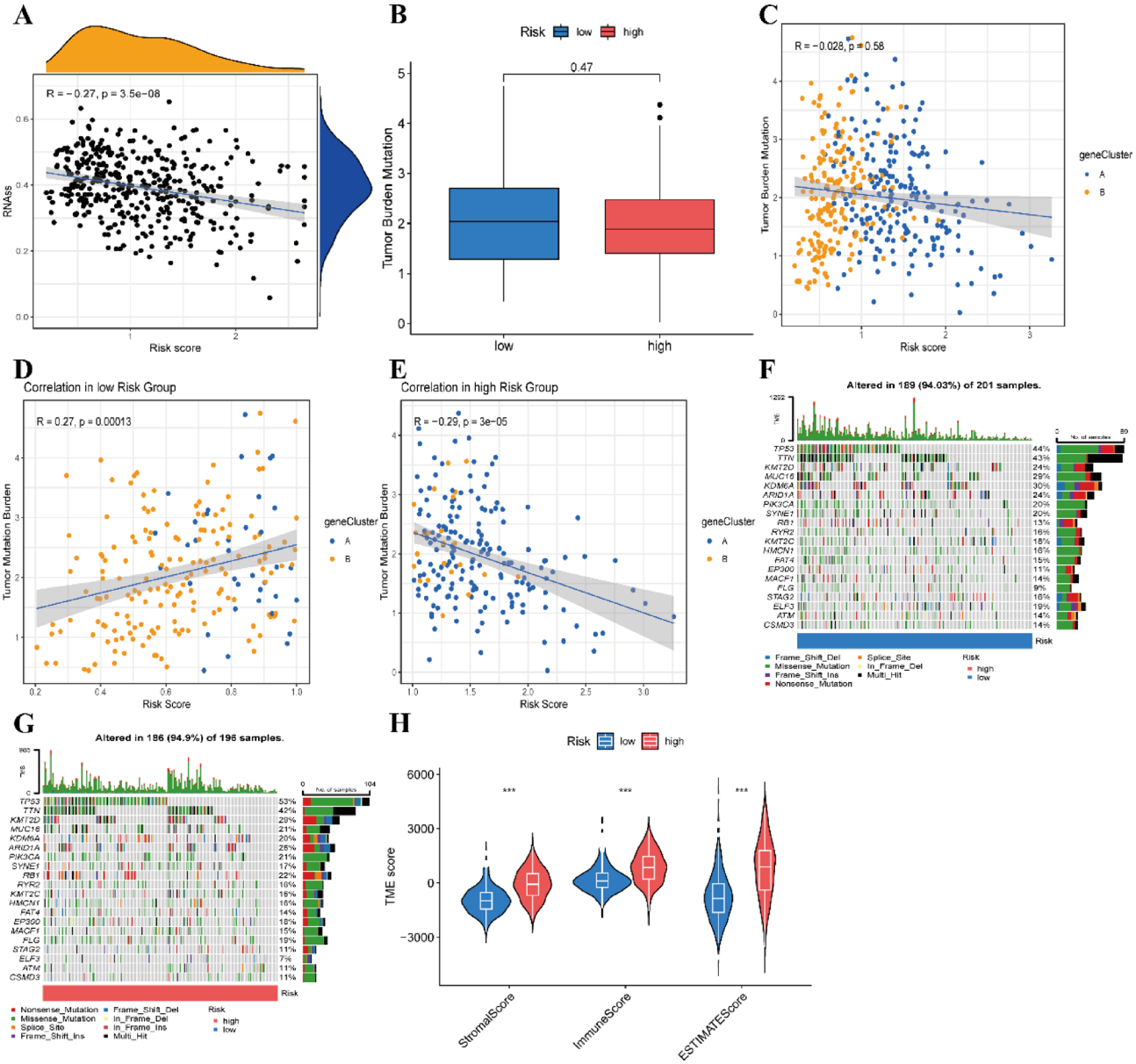

Fig. 6. Association Analysis of Risk Scores with CSC Index, TMB, and TME. (**A**) Correlation analysis between CSC index and risk scores. (**B**) Differential analysis of TMB between the two risk groups. (**C**–**E**) Spearman correlation analysis between TMB and risk scores. (**F**,**G**) Waterfall plots of somatic mutation characteristics in high-risk and low-risk groups. (**H**) Violin plot analysis of three TME scores in high-risk and low-risk groups. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

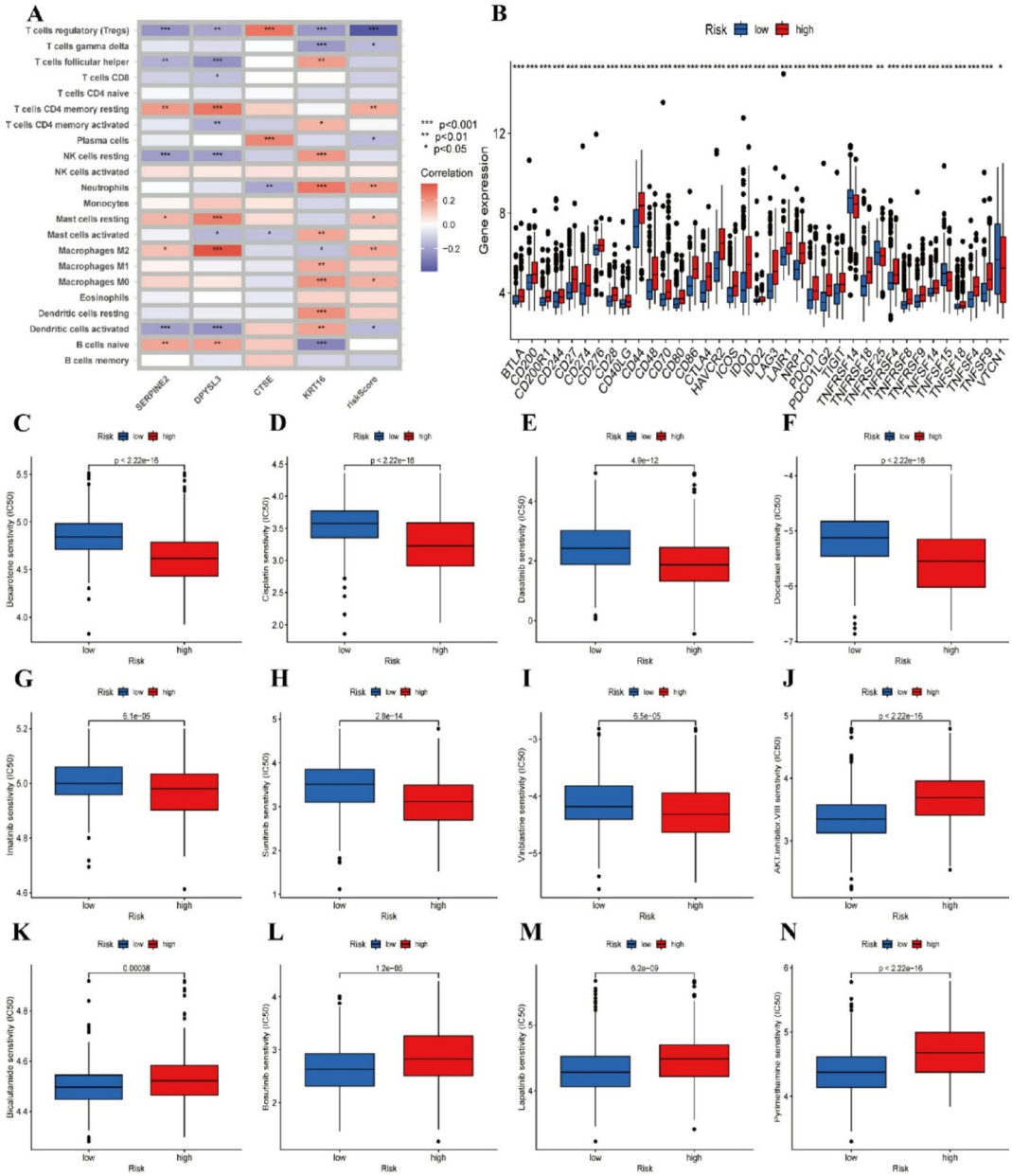

Fig. 7. Association Analysis of Risk Scores with Immune Cell Infiltration, Immune Checkpoints, and Drug Sensitivity. (**A**) Heatmap of Spearman correlation analysis between four model-related genes and the content of immune cells. (**B**) Differential expression analysis of immune checkpoints between high-risk and low-risk groups. (**C**–**N**) Differential analysis of IC50 values for commonly used drugs or compounds. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$.

low-risk bladder cancer samples had significantly lower scores in all three categories compared to the high-risk group (Fig. 6H).

Immune infiltration analysis, immune checkpoint evaluation, and drug sensitivity correlation analysis between risk groups

Correlation analysis between risk score and immune cell infiltration revealed positive associations with neutrophils, resting mast cells, M0 and M2 macrophages, and resting CD4 memory T cells, while showing negative correlations with regulatory T cells, gamma delta T cells, plasma cells, and activated dendritic cells (Fig. 7A). Additionally, we found significant differences in the expression of 37 immune checkpoints between the risk groups. Genes like TNFRSF14, TNFRSF25, TNFSF15, and VTCN1 were highly expressed in the low-risk group, whereas most genes, including PD-1 (PDCD1) and PD-L1 (CD274), were more highly expressed in the high-risk group (Fig. 7B).

Finally, we selected commonly used drugs or compounds to detect the association between drug sensitivity and risk score. Differential analysis of IC50 values between the two risk clusters indicated that the high-risk cluster had lower IC50 values for bexarotene, cisplatin, dasatinib, docetaxel, imatinib, paclitaxel, sunitinib and vinblastine compared to the low-risk cluster (Fig. 7C-I). Conversely, the low-risk cluster showed lower IC50 values for AKT inhibitor VIII, bicalutamide, bosutinib, lapatinib, and pyrimethamine (Fig. 7J-N).

Discussion

Efferocytosis-related genes and pathways, such as phosphatidylserine, TYRO3, MerTK, indoleamine-2,3-dioxygenase 1, TGF-β, and IL-10, are gaining attention as potential targets for cancer therapy[43]. These signaling molecules are frequently upregulated in bladder cancer and are linked to poor prognosis. Despite significant advancements in the diagnosis, prognosis prediction, and treatment strategies for bladder cancer, patients still face unfavorable outcomes due to high recurrence rates, distant metastasis, and drug resistance[44–46]. Considering the pivotal role of efferocytosis in tumor progression, this study developed a novel efferocytosis-related scoring system to assess prognosis and predict personalized treatment strategies.

In this study, we thoroughly investigated the roles and prognostic relevance of efferocytosis-related genes in bladder cancer. The first discovery revealed two distinct clusters of bladder cancer patients, each defined by unique clinical outcomes, immune cell infiltration patterns, and biological functions. The second discovery was the creation of a prognostic model based on four differential genes between clusters, which could help predict the prognosis of bladder cancer patients. Additionally, based on the risk score, we identified differential immune checkpoint expression between high- and low-risk groups, as well as associations between TME and TMB alterations and the risk score. These findings suggest possible mechanisms linking efferocytosis to immune responses and highlight potential targets for bladder cancer treatment. Efferocytosis in bladder cancer involves specific molecules and associated signaling pathways, offering a new perspective for future in-depth molecular mechanistic studies of bladder cancer.

We identified 28 EFRGs and explored their genetic variations, including somatic mutations and CNVs. These genetic alterations highlight the complexity and heterogeneity of bladder cancer, with the RXRA gene showing the highest mutation frequency, while the RAB17 gene exhibited significant CNV alterations. Previous research has shown that RXRA mutations linked to bladder cancer can activate peroxisome proliferator-activated receptors, leading to increased urothelial proliferation[47]. RXRA plays a pivotal role in bladder cancer by functioning as a transcription factor that regulates lipid metabolism. Specifically, it promotes the transcription of the long noncoding RNA UCA1, leading to aberrant lipid accumulation and increased resistance to epirubicin in bladder cancer cells. Moreover, RXRA mutations have been associated with a significantly higher risk of tumor recurrence, suggesting its potential utility as a prognostic biomarker. However, the underlying molecular mechanisms remain to be fully elucidated. In contrast, reduced RAB17 expression has been linked to poor prognosis in clear cell renal cell carcinoma, where its expression is associated with DNA methylation status and immune cell infiltration[48]. Nonetheless, studies investigating the role of RAB17 in bladder cancer are currently lacking.

Our study identified two EFRGs_clusters (A and B), which exhibited significant differences in OS and clinical characteristics. EFRGs_cluster A is characterized by higher immune cell infiltration and enrichment in pathways related to cytoskeletal reorganization and extracellular matrix remodeling, suggesting a more active or aggressive tumor phenotype[49,50]. The poor prognosis observed in Cluster A, despite its high level of immune infiltration, may be attributed to factors such as immune exhaustion, the accumulation of immunosuppressive cells (Regulatory T cells and myeloid-derived suppressor cells), low tumor immunogenicity, or defects in antigen presentation. Additionally, persistent chronic inflammation can generate reactive oxygen species and pro-inflammatory cytokines, which may paradoxically promote tumor proliferation, angiogenesis, and metastasis. These findings are consistent with previous studies[51–53]. In contrast, EFRGs_cluster B is enriched in metabolic pathways, such as lipid metabolism, and exhibits negative regulation of immune processes, suggesting a more suppressive or regulatory phenotype that may result in a less invasive tumor profile. These results indicate that efferocytosis plays a key role in shaping the TME and impacting disease outcomes, with potential implications for therapeutic targeting.

Given the notable differences between EFRGs_clusters, we performed differential gene expression and prognostic analyses between the two subtypes, leading to the identification of 61 prognostic DEGs. Finally, we established a risk scoring model comprising four genes: SERPINE2, DPYSL3, CTSE, and KRT16. In the internal validation set and external validation sets (GSE32548 and GSE32894), the 1-year AUC values were 0.635, 0.853, and 0.821, respectively. Overall, all AUC values exceeded 0.6. Notably, the external datasets showed higher AUC values compared to the internal dataset. This discrepancy may be attributed to the use of cancer-specific survival in the external cohorts, whereas overall survival was used in the internal cohort. Differences in dataset selection may also have contributed to this variation. To further enhance the model's predictive performance, we created a nomogram with AUC values exceeding 0.7 for 1, 3, and 5 years in the internal cohort. These results demonstrate that the newly developed efferocytosis-related prognostic feature provides an effective tool for assessing prognosis in bladder cancer patients.

SERPINE2, which acts against thrombin, urokinase, and plasminogen, has been shown to be inhibited by miR-186-5p and miR-378a-3p activated by BARX2, suppressing invasion and metastasis in oral squamous cell carcinoma, and its overexpression is linked to poor pathological features in bladder cancer[54–56]. DPYSL3 is a cytoplasmic phosphoprotein highly expressed in the nervous system that promotes cancer cell invasion and metastasis by mediating cytoskeletal modifications and altering the biological behavior of UC cells, potentially serving as a novel therapeutic target[57–59]. CTSE (cathepsin E) is involved in antigen processing and innate immunity, and its overexpression promotes BLCA cell proliferation, migration, and invasion, making it a potential target for enhancing immune responses[60–62]. Wild et al. discovered that elevated CTSE expression is

linked to shorter disease-specific survival[63]. However, other studies have shown that the expression of CTSE in the cytoplasm is associated with a lower risk of disease progression[64]. In our model, CTSE is recognized as a protective factor, aligning with the findings of Wild et al. KRT16, located on chromosome 17q21.2, encodes type I cytoskeletal protein 16, and its overexpression is strongly linked to advanced cancer, poor prognosis, as well as invasion and squamous differentiation in bladder cancer, making it a potential therapeutic target[65,66]. The roles and mechanisms of these four genes in different cancers highlight their significance in tumor invasion, metastasis, and prognosis, suggesting they could serve as potential therapeutic targets in the future.

CSC index analysis revealed a negative correlation between the CSC index and risk score, indicating that low-risk patients exhibit more pronounced stem cell characteristics, potentially having a higher risk of recurrence[67]. Targeting CSCs in these patients could therefore improve treatment outcomes. TMB is closely related to immunotherapy[68]. TMB analysis showed no significant differences between the high-risk and low-risk groups; however, within-group correlation analysis revealed opposite trends. In the low-risk group, TMB was positively correlated with the risk score, while in the high-risk group, it was negatively correlated. These findings suggest a more complex role for TMB in bladder cancer, requiring further investigation.

The TME consists of a diverse population of immune cells, stromal cells, endothelial cells, and cancer-associated fibroblasts, all of which play a crucial role in tumor development[69]. Our study revealed a positive correlation between the risk score and neutrophil abundance in bladder cancer samples, while showing a negative correlation with regulatory T cells (Tregs). Previous research has demonstrated that tumor cells secrete CXCL1 and CXCL8 to recruit neutrophils, which in turn activate ERK and JNK signaling pathways and upregulate VEGFA and MMP9 expression in neutrophils, promoting tumor lymphangiogenesis and facilitating lymphatic metastasis[70]. Thus, targeting neutrophils in tumors presents an effective therapeutic strategy to enhance the prognosis of bladder cancer patients. In most tumors, Tregs are immunosuppressive cells that may impede the ability of cytotoxic T lymphocytes (CTLs) to cancer cells in the TME[71]. However, other studies in bladder cancer have shown a positive correlation between Treg infiltration and survival rates, finding that Treg-mediated MMP2 inhibition in the TME may explain the paradoxical positive prognostic effect of tumor-infiltrating Tregs in BLCA[72,73].

Our study identified 37 immune checkpoint genes with differential expression between risk groups, with most immune checkpoints showing elevated expression in the high-risk group. This indicates that high-risk patients may derive greater benefit from immunotherapy. Malignant tumors establish an immunosuppressive microenvironment by activating immune checkpoints, with PD-1 and PD-L1 being the most extensively studied checkpoints implicated in tumor immune evasion[74]. Recent clinical studies have demonstrated that immune checkpoint inhibitors enhance treatment response rates and improve outcomes in patients with MIBC[75,76].

Despite providing valuable findings, our study has several limitations. It relies heavily on public datasets such as TCGA-BLCA and GEO (GSE13507, GSE32548, and GSE32894), which may introduce selection bias, batch effects, and data imbalance—such as unequal distributions of clinical variables—that could affect model performance and generalizability. Although we validated the model across multiple cohorts, further confirmation in larger, balanced, and prospectively collected clinical datasets is needed. Additionally, while the model shows potential for personalized prediction, its clinical utility requires validation through independent experimental data and mechanistic studies, including in vivo and in vitro experiments, to confirm the biological relevance of efferocytosis-related genes and their connection to immune and drug response.

Conclusion

In summary, our findings demonstrate that efferocytosis is linked to the progression and survival outcomes of BLCA. We constructed a prognostic model incorporating four key genes—SERPINE2, DPYSL3, CTSE, and KRT16—that accurately predicts prognosis and drug sensitivity in bladder cancer patients. Moreover, our study revealed the association between the model's risk score and tumor immune microenvironment in bladder cancer. This research provides new insights for prognostic prediction and precision treatment in BLCA.

Data availability

The data during the current study are available from TCGA (https://portal.gdc.cancer.gov/) and GEO (https://www.ncbi.nlm.nih.gov/geo/) database. Further inquiries can be directed to the corresponding author.

Received: 15 October 2024; Accepted: 23 May 2025
Published online: 06 June 2025

References

1. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 (2024).
2. Babjuk, M. et al. European association of urology guidelines on Non-muscle-invasive bladder Cancer (Ta, T1, and carcinoma in Situ). *Eur. Urol.* **81**, 75–94 (2022).
3. Sylvester, R. J. et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: A combined analysis of 2596 patients from seven EORTC trials. *Eur. Urol.* **49**, 466 – 465; discussion 475–477 (2006).
4. Schrier, B. P., Hollander, M. P., van Rhijn, B. W., Kiemeney, L. A. & Witjes, J. A. Prognosis of muscle-invasive bladder cancer: difference between primary and progressive tumours and implications for therapy. *Eur. Urol.* **45**, 292–296 (2004).
5. Moschini, M. et al. Comparing long-term outcomes of primary and progressive carcinoma invading bladder muscle after radical cystectomy. *BJU Int.* **117**, 604–610 (2016).
6. Dobruch, J. & Oszczudłowski, M. Bladder cancer: current challenges and future directions. *Med. (Kaunas Lithuania).* **57**, 749 (2021).
7. Mohanty, S. K., Lobo, A., Mishra, S. K. & Cheng, L. Precision medicine in bladder cancer: present challenges and future directions. *J. Pers. Med.* **13**, 756 (2023).

8. Wang, L. et al. EMT- and stroma-related gene expression and resistance to PD-1 Blockade in urothelial cancer. *Nat. Commun.* **9**, 3503 (2018).
9. Sundahl, N. et al. Randomized phase 1 trial of pembrolizumab with sequential versus concomitant stereotactic body radiotherapy in metastatic urothelial carcinoma. *Eur. Urol.* **75**, 707–711 (2019).
10. Dutta, S., Mukherjee, A. & Nongthomba, U. Before the cytokine storm: boosting efferocytosis as an effective strategy against SARS-CoV-2 infection and associated complications. *Cytokine Growth Factor. Rev.* **63**, 108–118 (2022).
11. Korns, D., Frasch, S. C., Fernandez-Boyanapalli, R., Henson, P. M. & Bratton, D. L. Modulation of macrophage efferocytosis in inflammation. *Front. Immunol.* **2**, 57 (2011).
12. Morioka, S., Maueröder, C. & Ravichandran, K. S. Living on the edge: efferocytosis at the interface of homeostasis and pathology. *Immunity* **50**, 1149–1162 (2019).
13. Mehrotra, P. & Ravichandran, K. S. Drugging the efferocytosis process: concepts and opportunities. *Nat. Rev. Drug Discov.* **21**, 601–620 (2022).
14. Liu, S. et al. S100A4 enhances protumor macrophage polarization by control of PPAR-γ-dependent induction of fatty acid oxidation. *J. Immunother Cancer*. **9**, e002548 (2021).
15. Shapouri-Moghaddam, A. et al. Macrophage plasticity, polarization, and function in health and disease. *J. Cell. Physiol.* **233**, 6425–6440 (2018).
16. Ma, L. et al. 27-Hydroxycholesterol acts on myeloid immune cells to induce T cell dysfunction, promoting breast cancer progression. *Cancer Lett.* **493**, 266–283 (2020).
17. Nguyen, K. Q. et al. Overexpression of MERTK receptor tyrosine kinase in epithelial cancer cells drives efferocytosis in a gain-of-function capacity. *J. Biol. Chem.* **289**, 25737–25749 (2014).
18. Cheng, M. et al. CD276-dependent efferocytosis by tumor-associated macrophages promotes immune evasion in bladder cancer. *Nat. Commun.* **15**, 2818 (2024).
19. Giannelos, S., Bellizio, F., Strbac, G. & Zhang, T. Machine learning approaches for predictions of CO2 emissions in the Building sector. *Electr. Power Syst. Res.* **235**, 110735 (2024).
20. Pereira, A. G., Barbosa, G. F., Filho, M. G., Shiki, S. B. & da Silva, A. L. Quality control in Extrusion-Based additive manufacturing: A review of machine learning approaches. *IEEE Trans. Cybern* (2025).
21. Cetinkaya, C. et al. Artificial intelligence-empowered functional design of semi-transparent optoelectronic and photonic devices via deep Q-learning. *Sci. Rep.* **15**, 13508 (2025).
22. Bourne, P. E. Is bioinformatics dead. *PLoS Biol.* **19**, e3001165 (2021).
23. Guo, S., Su, Q., Yang, Z. & Mo, W. Development and validation of a novel liquid-liquid phase separation gene signature for bladder cancer. *Sci. Rep.* **14**, 22552 (2024).
24. Huang, J. et al. Integrating oxidative-stress biomarkers into a precision oncology risk-stratification model for bladder cancer prognosis and therapy. *Front. Cell. Dev. Biol.* **12**, 1453448 (2024).
25. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
26. Zhong, F. et al. A novel molecular classification based on efferocytosis-related genes for predicting clinical outcome and treatment response in acute myeloid leukemia. *Inflamm. Res.* **73**, 1889–1902 (2024).
27. Gao, M. et al. Identification of efferocytosis-related subtypes in gliomas and elucidating their characteristics and clinical significance. *Front. Cell. Dev. Biol.* **11**, 1295891 (2023).
28. Wang, L. et al. A signature based on efferocytosis-related genes for the evaluation of prognosis and the tumour microenvironment in gastric cancer. *Sci. Rep.* **15**, 14226 (2025).
29. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
30. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The Sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
31. Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C. & Koeffler, H. P. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **28**, 1747–1756 (2018).
32. Giuliani, A. The application of principal component analysis to drug discovery and biomedical data. *Drug Discov Today*. **22**, 1069–1076 (2017).
33. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14**, 7 (2013).
34. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* **25**, 25–29 (2000).
35. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res.* **53**, D672–D677 (2025).
36. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
37. Gene Ontology Consortium. The gene ontology knowledgebase in 2023. *Genetics* **224**, iyad031 (2023).
38. Thomas, P. D. et al. PANTHER: making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
39. Jin, Y. et al. Identification of novel subtypes based on SsGSEA in immune-related prognostic signature for tongue squamous cell carcinoma. *Cancer Med.* **10**, 8693–8707 (2021).
40. Yoshihara, K. et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
41. Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biology (Clifton N J)*. **1711**, 243–259 (2018).
42. Geeleher, P., Cox, N. & Huang, R. S. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS ONE*. **9**, e107468 (2014).
43. Tajbakhsh, A. et al. The complex roles of efferocytosis in cancer development, metastasis, and treatment. *Biomed. Pharmacotherapy = Biomedecine Pharmacotherapie*. **140**, 111776 (2021).
44. Biswas, P. K. et al. TTYH3 modulates bladder Cancer proliferation and metastasis via FGFR1/H-Ras/A-Raf/MEK/ERK pathway. *Int. J. Mol. Sci.* **23**, 10496 (2022).
45. Chow, P. M. et al. The UCHL5 inhibitor b-AP15 overcomes cisplatin resistance via suppression of cancer stemness in urothelial carcinoma. *Mol. Ther. Oncolytics*. **26**, 387–398 (2022).
46. Pan, S. et al. Immune status for monitoring and treatment of bladder cancer. *Front. Immunol.* **13**, 963877 (2022).
47. Halstead, A. M. et al. Bladder-cancer-associated mutations in RXRA activate peroxisome proliferator-activated receptors to drive urothelial proliferation. *Elife* **6**, e30862 (2017).
48. Zeng, Z. et al. Downregulation of RAB17 have a poor prognosis in kidney renal clear cell carcinoma and its expression correlates with DNA methylation and immune infiltration. *Cell. Signal.* **109**, 110743 (2023).
49. Yoshida, G. J. Regulation of heterogeneous cancer-associated fibroblasts: the molecular pathology of activated signaling pathways. *J. Exp. Clin. Cancer Res.* **39**, 112 (2020).
50. Peng, J. M. et al. Evasion of NK cell immune surveillance via the vimentin-mediated cytoskeleton remodeling. *Front. Immunol.* **13**, 883178 (2022).

51. Yang, L. et al. Immune profiling reveals molecular classification and characteristic in urothelial bladder Cancer. *Front. Cell. Dev. Biol.* **9**, 596484 (2021).
52. Pan, S., Zhan, Y., Chen, X., Wu, B. & Liu, B. Bladder Cancer exhibiting high immune infiltration shows the lowest response rate to immune checkpoint inhibitors. *Front. Oncol.* **9**, 1101 (2019).
53. Oshi, M. et al. Abundance of reactive oxygen species (ROS) is associated with tumor aggressiveness, immune response, and worse survival in breast cancer. *Breast Cancer Res. Treat.* **194**, 231–241 (2022).
54. Bouton, M. C. et al. Emerging role of serpinE2/protease nexin-1 in hemostasis and vascular biology. *Blood* **119**, 2452–2457 (2012).
55. Chuang, H. W. et al. SERPINE2 overexpression is associated with poor prognosis of urothelial carcinoma. *Diagnostics (Basel Switzerland)*. **11**, 1928 (2021).
56. Zhang, S. et al. SERPINE2 promotes liver cancer metastasis by inhibiting c-Cbl-mediated EGFR ubiquitination and degradation. *Cancer Commun. (London England)*. **44**, 384–407 (2024).
57. Kawahara, T. et al. Quantitative proteomic profiling identifies DPYSL3 as pancreatic ductal adenocarcinoma-associated molecule that regulates cell adhesion and migration by stabilization of focal adhesion complex. *PLoS ONE.* **8**, e79654 (2013).
58. Rubtsova, S. N., Zhitnyak, I. Y. & Gloushankova, N. A. Phenotypic plasticity of Cancer cells based on remodeling of the actin cytoskeleton and adhesive structures. *Int. J. Mol. Sci.* **22**, 1821 (2021).
59. Liang, P. I. et al. Upregulation of dihydropyrimidinase-like 3 (DPYSL3) protein predicts poor prognosis in urothelial carcinoma. *BMC Cancer*. **23**, 599 (2023).
60. Goto, S. et al. Role of cathepsin E in decidual macrophage of patients with recurrent miscarriage. *Mol. Hum. Reprod.* **20**, 454–462 (2014).
61. Gonçalves, N. P. et al. Differential expression of cathepsin E in transthyretin amyloidosis: from neuropathology to the immune system. *J. Neuroinflammation.* **14**, 115 (2017).
62. Mengwasser, J. et al. Cathepsin E deficiency ameliorates Graft-versus-Host disease and modifies dendritic cell motility. *Front. Immunol.* **8**, 203 (2017).
63. Wild, P. J. et al. Gene expression profiling of progressive papillary noninvasive carcinomas of the urinary bladder. *Clin. Cancer Res.* **11**, 4415–4429 (2005).
64. Fristrup, N. et al. Cathepsin E, Maspin, Plk1, and survivin are promising prognostic protein markers for progression in non-muscle invasive bladder cancer. *Am. J. Pathol.* **180**, 1824–1834 (2012).
65. Huang, W. C. et al. A novel miR-365-3p/EHF/keratin 16 axis promotes oral squamous cell carcinoma metastasis, cancer stemness and drug resistance via enhancing β5-integrin/c-met signaling pathway. *J. Exp. Clin. Cancer Res.* **38**, 89 (2019).
66. Mehus, A. A. et al. Activation of PPARγ and Inhibition of cell proliferation reduces key proteins associated with the basal subtype of bladder cancer in As3+-transformed UROtsa cells. *PLoS ONE.* **15**, e0237976 (2020).
67. Zhuang, J. et al. Cancer-Associated Fibroblast-Derived miR-146a-5p generates a niche that promotes bladder Cancer stemness and chemoresistance. *Cancer Res.* **83**, 1611–1627 (2023).
68. Chan, T. A. et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann. Oncol.* **30**, 44–56 (2019).
69. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
70. Zhang, Q. et al. ETV4 mediated Tumor-Associated neutrophil infiltration facilitates lymphangiogenesis and lymphatic metastasis of bladder cancer. *Adv. Sci.* **10**, e2205613 (2023).
71. Lei, X. et al. Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett.* **470**, 126–133 (2020).
72. Winerdal, M. E. et al. FOXP3 and survival in urinary bladder cancer. *BJU Int.* **108**, 1672–1678 (2011).
73. Winerdal, M. E. et al. Urinary bladder Cancer Tregs suppress MMP2 and potentially regulate invasiveness. *Cancer Immunol. Res.* **6**, 528–538 (2018).
74. Alsaab, H. O. et al. PD-1 and PD-L1 checkpoint signaling Inhibition for Cancer immunotherapy: mechanism, combinations, and clinical outcome. *Front. Pharmacol.* **8**, 561 (2017).
75. Szabados, B. et al. Toxicity and surgical complication rates of neoadjuvant Atezolizumab in patients with Muscle-invasive bladder Cancer undergoing radical cystectomy: updated safety results from the ABACUS trial. *Eur. Urol. Oncol.* **4**, 456–463 (2021).
76. Li, R. et al. Comparative effectiveness of neoadjuvant pembrolizumab versus Cisplatin-based chemotherapy or upfront radical cystectomy in patients with Muscle-invasive urothelial bladder Cancer. *Eur. Urol. Oncol.* **7**, 614–624 (2024).

Acknowledgements

We acknowledge TCGA, GEO, KEGG, MSigDB, and STRING databases and analysis tools for providing their platforms and contributors for uploading their meaningful datasets.

Author contributions

The study design was conceptualized by W. T. Yu, D. N. Yao, J. J. Hou, and J. Q. Tian. Data acquisition, analysis, and interpretation were conducted by W. T. Yu, D. N. Yao, and X. M. Ma. The initial draft of the manuscript was written by W. T. Yu, and all authors provided feedback on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by (1) Science and Technology Project of Chengguan District, Lanzhou City, Gansu Province Science and Technology Bureau (Project Number 2017KJGG0052); (2) Cuiying Graduate Supervisor Applicant Training Program of Lanzhou University Second Hospital (Project Number 201704); (3) Lanzhou City Talent Innovation and Entrepreneurship Project (Project Number 2019-RC-37) ; (4) The Major Scientific Research Project and Key Medical Technology Problem-Solving Project of 2022 China Medical Education Association(Project Number 2022KTZ007) ; (5) Natural Science Foundations Project of Gansu Province (23JRRM0747).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-04037-w.

Correspondence and requests for materials should be addressed to J.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.