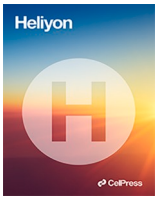
Heliyon 10 (2024) e34632

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Identification of epithelial-related artificial neural network prognostic models for the prediction of bladder cancer prognosis through comprehensive analysis of single-cell and bulk RNA sequencing

Fan Zhao, Kun Zhang, Limin Ma, Yeqing Huang *

Department of Urology, Affiliated Hospital of Nantong University, Nantong, 226001, China

ARTICLE INFO

Keywords:
Bladder cancer
Single-cell analysis
Artificial neural network prognostic models
Machine learning
NFIC

ABSTRACT

Background: Bladder cancer (BLCA) presents as a heterogeneous epithelial malignancy. Progress in the early detection and effective treatment of BLCA relies heavily on the identification of novel biomarkers. Therefore, the primary goal of this study is to pinpoint potential biomarkers for BLCA through the fusion of single-cell RNA sequencing and RNA sequencing assessments. Furthermore, the aim is to establish practical clinical prognostic models that can facilitate accurate categorization and individualized therapy for patients.

Methods: In this research, training sets were acquired from the TCGA database, whereas validation sets (GSE32894) and single-cell datasets (GSE135337) were extracted from the GEO database. Single-cell analysis was utilized to obtain characteristic subpopulations along with their associated marker genes. Subsequently, a novel BLCA subtype was identified within TCGA-BLCA. Furthermore, an artificial neural network prognostic model was constructed within the TCGA-BLCA cohort and subsequently verified utilizing a validation set. Two machine learning algorithms were employed to screen hub genes. QRT-qPCR was performed to detect the gene expression levels utilized in the construction of prognostic models across various cell lines. Additionally, the cMAP database and molecular docking were utilized for searching small molecule drugs.

Results: The results of single-cell analysis revealed the presence of epithelial cells in multiple subpopulations, with 1579 marker genes selected for subsequent investigations. Subsequently, four epithelial cell subtypes were identified within the TCGA-BLCA cohort. Notably, cluster A exhibited a significant survival advantage. Concurrently, an artificial neural network prognostic model comprising 17 feature genes was constructed, accurately stratifying patient risk. Patients categorized in the low-risk group demonstrated a considerable survival advantage. The ROC analysis suggested that the model has strong prognostic ability. Furthermore, the findings of the validation group align consistently with those from the training group. Two types of machine learning algorithms screened NFIC as hub genes. Forskolin, a small molecule drug that binds to NFIC, was identified by employing a cMAP database and molecular docking.

Conclusion: The analysis results supplement the research on the role of epithelial cells in BLCA. An artificial neural network prognostic model containing 17 characteristic genes demonstrates the

* Corresponding author.
E-mail address: huangyeqing@ntu.edu.cn (Y. Huang).

https://doi.org/10.1016/j.heliyon.2024.e34632
Received 20 May 2024; Received in revised form 3 July 2024; Accepted 12 July 2024
Available online 17 July 2024

capability to accurately stratify patient risk, thereby potentially improving clinical decision-making and optimizing personalized therapeutic approaches.

1. Introduction

Bladder cancer (BLCA) is a common urological malignancy that presents a significant global health challenge. Every year, the global incidence of BLCA exceeds 570,000 new cases, resulting in over 200,000 deaths [1,2]. BLCA is characterized by two types, such as muscle-invasive bladder cancer (MIBC) and non-muscle-invasive bladder cancer (NMIBC), on the basis of the depth of invasion into the muscularis propria of the bladder. Approximately 75 % of BLCA cases are attributed to NMIBC, while MIBC accounts for approximately 25 % of BLCA cases [3]. NMIBC is typically treated with transurethral resection, but there is a significant rate of postoperative recurrence and tumor progression [4]. In some cases, patients initially diagnosed with NMIBC advance to MIBC [5]. MIBC is an aggressive disease characterized by a considerable risk of metastasis, often spreading to peripheral organs and lymph nodes [3]. Resultantly, the five-year survival rate of MIBC patients is lower than that of patients with NMIBC. Currently, MIBC has two treatment modalities, including radical cystectomy plus neoadjuvant chemotherapy and maximal transurethral cystectomy of the bladder tumor plus radiotherapy and chemotherapy [6]. Subsequently, most patients with MIBC still experience recurrence or metastasis post-treatment. In many cases, BLCA patients lack specific clinical manifestations, leading to diagnoses at advanced stages. With advancements in medical treatment, immunotherapy offers new options for cancer patients. However, due to the tumor heterogeneity of BLCA, immunotherapy still struggles to enhance the treatment status of BLCA patients. Hence, in this era of personalized medicine, it is crucial to thoroughly comprehend the molecular mechanisms implicated in the progression of BLCA. Moreover, the identification and development of robust and reliable biomarkers and risk models are necessary for improving the current therapeutic status of BLCA.

With the rapid advancements in bioinformatics, transcriptomics has been used to explore biomarkers for various diseases, including tumors and non-tumors. Currently, an increasing number of genes have been identified as effective therapeutic targets for cancer [7]. The report by Fei Xie et al. suggests that circPTPRA in the exon can inhibit cancer progression by endogenously blocking the recognition of m6A-modified RNA via IGF2BP1 [8]. Jianqiang Nie et al. discovered notably upregulated expression of Glycyl-TRNA Synthetase 1 (GARS1) in various cancer types, demonstrating prognostic value in diverse types of cancer. The authors also revealed that GARS1 can enhance the proliferative and migratory capabilities of BLCA cells [9]. In contrast to traditional transcriptomic approaches, single-cell RNA sequencing (scRNA-seq) enables researchers to analyze transcriptomic patterns on a cellular level. This facilitates a thorough examination of cellular diversity within the tumor microenvironment (TME) and provides novel insights into cell interactions [10–12]. Kun et al. identified CXCL14 as a pivotal regulator of lymph node metastasis in breast cancer [13]. The study by Junya Peng et al. revealed multiple malignant and stromal cell types in primary pancreatic ductal adenocarcinoma. Among them, heterogeneous malignant subtypes were composed of several subgroups exhibiting distinct proliferation and migration potentials [14]. Hence, the objective of this investigation is the detection of potential biomarkers of BLCA and the construction of relevant clinical prognostic models by integration of RNA-seq and scRNA-seq analyses. This approach aims to achieve precise stratification and personalized treatment for patients.

In this research, scRNA-seq data and a large amount of RNA-seq data were utilized to screen for BLCA-related characteristic genes. Subsequently, a clinical prognosis model was constructed, and its risk stratification ability was validated using external datasets.

Additionally, an exploration into the correlation between immune system cells and predictive models was conducted in order to enhance comprehension regarding the potential role of immune cells in the advancement of BLCA.

. Overall, this study provides a new insight that may contribute to personalized treatment and surgical accuracy for BLCA patients.

2. Materials and methods

Datasets and patient samples The RNA-seq data, clinical information, and mutation data for BLCA were obtained from the TCGA database. The TCGA-BLCA dataset comprised 19 normal tissue samples and 412 tumor samples. An external validation cohort, the GSE32894 dataset, was acquired from the GEO database. After excluding samples with incomplete clinical and survival information, the GSE32894 dataset consisted of 224 tumor samples. The scRNA-seq dataset GSE135337 for BLCA was obtained from the GEO database. GSE135337 comprised single-cell sequencing data from seven BLCA patients.

2.1. Identification of major cell subpopulations and associated marker genes via scRNA-seq analysis

The single-cell data were analyzed utilizing R packages such as “Seurat” [15,16]. The following filters were implemented to the raw matrix of every cell for retaining high-quality single-cell data, including genes expressed in at least five single cells, eliminating cells expressing fewer than 100 genes, as well as excluding cells with more than 5 % of mitochondrial genes. The “NormalizeData” function of the R package “Seurat” was utilized for normalizing the single-cell data. The normalized data is converted to Seurat objects, which are then employed for the identification of the top 2000 genes exhibiting high variability. Principal Component Analysis (PCA) was performed using the "RunPCA" function in the R package "Seurat" to reduce the dimensionality of single-cell data by analyzing the top 2000 genes. JackStraw analysis was utilized for identifying significant PCs and selecting the appropriate PCs as per the proportion of variance accounted for cell clustering analysis. The “FindClusters” and “FindNeighbors” functions from the “Seurat” package were

employed to carry out the cell clustering analysis. The t-distributed Stochastic Neighborhood Embedding (t-SNE) was utilized for the visualization of nonlinear downscaling [17–19]. The "FindAllMarkers" function was utilized to identify marker genes having logfc of 0.25 (multiplicity of variance) and minpct of 0.25 (expression rate of the least differentiated gene). Ultimately, the marker genes were filtered by employing the corrected *P*-value <0.05.

2.2. Functional enrichment analysis

The R package "clusterProfiler" was employed to conduct the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses. The GO analyses included pathways of molecular functions (MFs), biological processes (BPs), cellular components (CCs), and others [20]. Pathways with FDR <0.25 and $P < 0.05$ exhibited statistical significance.

2.3. Identification of different subgroups in BLCA based on unsupervised clustering

Marker gene expression for key cellular subpopulations identified in single-cell analysis was acquired through the TCGA-BLCA dataset. The Cox analysis (univariate) was conducted to screen candidate prognostic genes, applying a threshold of $P < 0.05$. The "ConsensusClusterPlus" R package was utilized to carry out consensus clustering on the basis of prognosis-related genes. The heat map depicts the relationship between clinical characteristics and different clusters [21].

2.4. Development of protein-protein interaction (PPI) network and artificial neural network prognostic model

The prognostic genes were imported into the STRING database to acquire interaction information. Cytoscape was utilized for visualizing the PPI network and identifying key nodes within the PPI network through CytoHubba, a Cytoscape plugin [22,23]. Subsequently, these key nodes were used to construct the artificial neural network model, which comprised hidden, output, and input layers(Fig. 1). The quantity of nodes in the input layer and output layer aligned with the quantity of input data characteristics and the amount of anticipated labels, respectively. Moreover, the volume of nodes in the concealed layer matched the quantity of covert nodes. The intricacy of the model was fine-tuned by modifying the quantity of covert nodes. It is important to consider that there might be multiple hidden layers within the hidden layer. The hidden layer is comprised of fully connected nodes (neurons), followed by a dropout layer. The output of the network is achieved through simple fully connected integration, where the results of the output nodes are used as risk values. Finally, this model employed the following loss function to train the network:

$$Y(\alpha) : = \frac{1}{N_{E=1}} * \sum i : E_{i=1} \left(h^{\alpha}(x_i) - log \Sigma_{i \in R(T_i)} e^{h^{\alpha}(x_i)} \right) \quad (1)$$

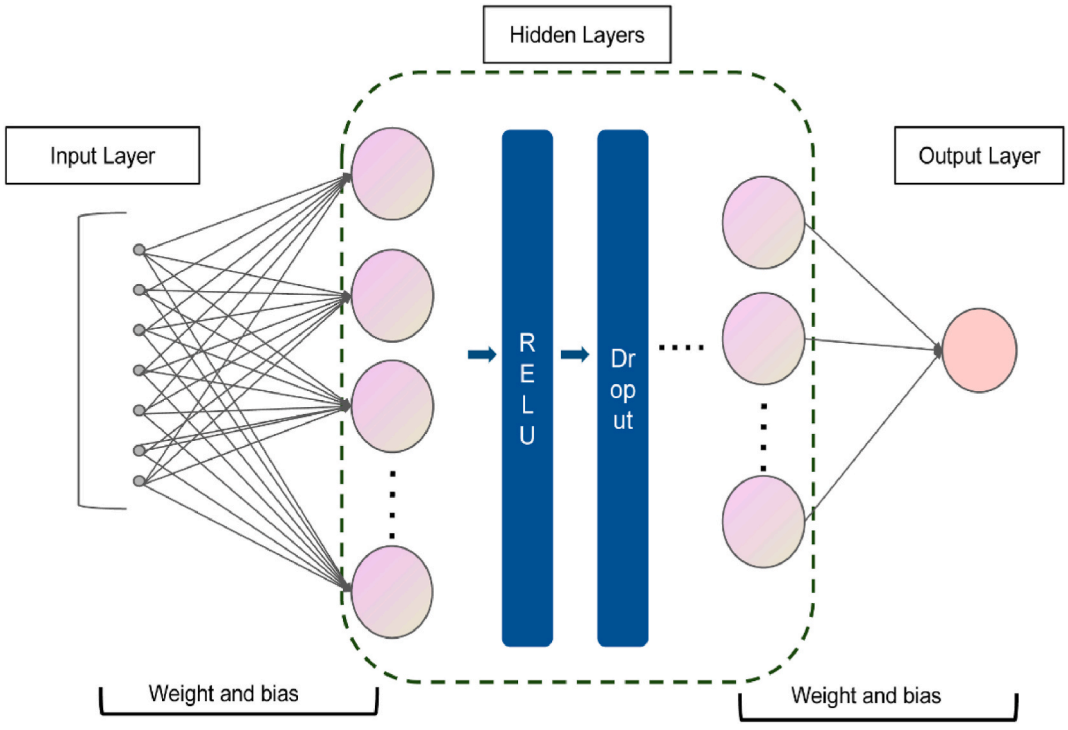


Fig. 1. Visualization of artificial neural network prognostic prediction model.

In the equation, $h^{\alpha}(x_i)$ specifically refers to the output of a neural network, representing the logarithmic estimate of the risk for a given covariate (x_i). $\log \sum_{i \in R(T_i)} e^{h^{\alpha}(x_i)}$ is the logarithmic sum of the risks of all individuals who survived before time (T_i). $N_{E=1}$ denotes the number of individuals who have experienced an event. This process was carried out repeatedly. The application of nonlinear activation and multi-layer functions distinguishes this technique from linear models, which are limited to classifying linear data.

Following this, BLCA patients were categorized into groups with low and high risks based on the median risk score obtained from the neural network prognostic model. An analysis using the Kaplan-Meier method was carried out to assess the association between these groups and the overall survival (OS). Receiver operating characteristic (ROC) curves were employed to evaluate the practical significance of the model in predicting the survival outcomes of BLCA patients. Moreover, in order to validate its clinical utility, GSE32894 was employed as an independent external validation cohort.

2.5. Clinical correlation between two risk groups

Cox regression analyses were carried out to assess the potential of risk scores as an autonomous prognostic indicator for BLCA patients. Furthermore, the association between clinicopathological features and risk score was explored within the TCGA-BLCA cohort. The research also generated column graphs depicting various clinical attributes and risk scores. Furthermore, estimates for the 1-, 3-, and 5-year survival rates of BLCA patients were projected utilizing the "survival" and "rms" software packages.To determine the accuracy and discriminatory power of the column charts, calibration curves were employed.

2.6. Differences in TME infiltration and somatic mutations between low and high-risk groups

The CIBERSORT algorithm, a fundamental tool for the deconvolution of immune cell expression matrices as per linear support vector regression, measures the infiltrating immune cell proportions by labeling gene expression [24]. To assess the relationship between risk groups and immune cell infiltration, this algorithm was employed. The R package "maftools" was employed for the extraction of the mutation annotation format from the TCGA database, enabling the identification of mutations in BLCA patients between different risk groups.

2.7. Drug sensitivity analysis and drug prediction

In order to forecast chemosensitivity disparities between high- and low-risk groups, the IC50 values at half-maximal inhibitory concentration were deduced by referencing the Genomics of Drug Sensitivity in Cancer (GDSC) database. This was achieved by constructing ridge regression models with ten-fold cross-validation utilizing the pRRophetic R package [25,26]. The Connectivity Map database (cMAP,https://portals.broadinstitute.org/cmap/) was employed for predicting potential small molecule drugs.

2.8. Identification of important features and molecular docking

Two machine learning algorithms (XGboost and LightGBM) were utilized to examine the feature importance of model genes in the model, aiding in further identification of the key genes that contribute the most to the model. SHAP values were used to clarify the role of each gene in the model. This explicatory method quantifies the influence of each characteristic on the model's result. The SHAP value average was calculated for each gene across all samples to identify genes that significantly impact the overall model output. In this research, a new blind docking tool, CB-Dock2 (http://clab.labshare.co.uk/cb-dock/php/index.php), was applied, with a focus on

Table 1
The primer pairs targeting genes.

Gene	Forward primer sequence (5′-3′)	Reverse primer sequence (5′-3′)
CKB	GGCAAGCATGAGAAGTTCTCGG	ACCAGCTCCACCTCTGAGAAGC
COL6A2	CGTGGAGACTCAGGACAGCCA	CCTTTCAAGCCAAAGTCGCCTC
COX7A1	CCTCCTCCGTGGTGGACGAAT	CAGCGTGACAGCTGGGGACG
DPYSL2	ACTGCCCAGAAGGCTGTAGGAA	CAGCCACAAACTGGTTCTCATCC
FLNA	CAACAAGTTCACTGTGGAGACCA	TGTAGGTGCCAGCCTCATAAGG
FSTL1	TCGCATCATCCAGTGGCTGGAA	TCACTGGAGTCCAGGCGAGAAT
GADD45B	GCCAGGATCGCCTCACAGTGG	GGATTTGCAGGGCGATGTCATC
GNG11	GTTCTAGGGCGAAAATGCCG	AACGACCTCTTCCACTGAGG
GSN	ATCTGCCATCCTGACTGCTCAG	CTTCCCACCAAACAGGCTCATG
KCNJ15	TGTGCTTGGTGATTCAGGTAGCC	GACAGTGGCTTGGTTGAGGAGA
KRTCAP2	GGTCTCTTCGTGTTCTCGCTCA	CAAAGAGAGCCAACAGGAGGCA
LDLR	GAATCTACTGGTCTGACCTGTCC	GGTCCAGTAGATGTTGCTGTGG
MYADM	CCAGTTCGATGAGAAGTATGGCG	CAGCCACATACGCCAGTAGGTT
NFIC	TGGCGGCGATTACTACACTTCG	GGCTGTTGAATGGTGACTTGTCC
OLFML3	GGAGTACATGGAACGCCGACTA	CCTTCTCTGCCACCTCCAGCA
RGS2	CTCTACTCCTGGGAAGCCCAAA	TTGCTGGCTAGCAGCTCGTCAA
SERINC1	TGATGACCCTTACCAACTGGTAC	CGAGTGTCCAAACATACAGCACG
GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA

enhancing docking accuracy. This software anticipates the binding area of a given protein. It also calculates the dimensions and midpoint using curvature-focused cavity identification methods. Moreover, it interconnects with the state-of-the-art docking software. Furthermore, CB-Dock2 evaluates binding modes using Vina scores and provides an interactive 3D visualization of these binding modes. The absolute values of binding energy are arranged in descending order, from high to low [27–29]. A binding energy < -5 kcal/mol denotes good binding ability, whereas a binding energy < -7 kcal/mol indicates strong activity. The small molecule structures were obtained from the PubChem database and saved in saf format. Then, the 3D structure of the protein was extracted from the AlphaFold Protein Structure Database and saved in PDB format. This study utilized the online CB-Dock2 database for molecular docking and complete visual analysis.

2.9. Real-time fluorescence quantitative polymerase chain reaction (qRT-PCR)

TRIzol reagent (Thermo Fisher, USA) was utilized for extracting total RNA from cell lines, such as SV-HUC-1, EJ-1, and BT-B. The cDNA was synthesized via 500 ng RNA utilizing HiScript II SuperMix (Vazyme, China). The PCR amplification conditions are 94 °C for 10 min, 94 °C for 10 s, and 60 °C for 45 s, respectively. GAPDH serves as an internal reference. Table 1 depicts the primer pairs targeting genes.

2.10. Cell culture and transient transfection

SV-HUC-1, EJ-1 and BT-B cells were purchased from Beijing Bena Biotechnology Co. (Beijing, China). Cells were cultured in DEME F-12 medium. The negative control (NC) and siRNA for KRTCAP2 (Sagon, China) were transfected into the cells utilizing Lipofectamine 2000 (Invitrogen, Thermo Fisher, USA).

Gene	Target sequences (5′-3′)
si-KRTCAP2#1	AGGATTCCAAGCAAAGATCTTCC
si-KRTCAP2#2	GTCTGTACTACATCAACAAGATC
si-KRTCAP2#3	GAGAAACTGACCCTGAATGTTCA

2.11. Cell viability

Cell viability was determined through the Cell Counting Kit-8 test (Beyotime, China), following the guidelines specified by the manufacturer. Various treatments were administered to cells, which were then grown in 96-well dishes with 1×103 cells in each well. The CCK-8 solution was administered at specified intervals. Following a 2-h incubation at 37 °C, the optical density values at 450 nm were measured using a microplate reader (BioTeK, USA).

2.12. Transwell assay

Migration assays using Transwell were carried out with EJ-1 and BT-B cell lines. In summary, cells ($1x10^5$) were seeded onto Matrigel-free chambers (BD Biosciences, USA) for the migration process. The upper chamber was filled with serum-free medium, while the lower chamber contained complete DMEM medium. Following a 24-h incubation period, fixed with 4 % paraformaldehyde and stained with 0.1 % crystal violet, the migrating or invading cells were observed.

2.13. 5-Ethynyl-2′-deoxyuridine (EdU) assay

The EdU assay was carried out using the BeyoClick™ EdU Cell Proliferation Kit with Alexa Fluor 594 by Biotek in Shanghai, China. Following a rinse with PBS, the cells were exposed to the EdU solution for a duration of 2 h, after which they were stained using DAPI solution for the identification of nuclei. Subsequently, the samples underwent washing before being observed using an inverted microscope from Olympus.

2.14. Western-blot

Protein was extracted using RIPA buffer. Following quantification, protein samples were loaded onto a 10 % SDS-PAGE gel, then transferred onto a 0.45-μm PVDF membrane. Blocking was performed using 5 % skim milk for 2 h. Primary antibodies were then added and left to incubate at 4 °C overnight. Primary antibodies were: anti-P53 (60283-2-Ig, Proteintech, 1/500), anti-PI3K (20584-1-AP, Proteintech, 1/500), anti-*P*-PI3K (310164, Zenbio, 1/500), anti-AKT (60203-2-Ig, Proteintech, 1/500), anti-*P*-AKT (80455-1-RR, Proteintech, 1/500) and anti-GAPDH (ab8245, abcam, 1/5000). Afterward, secondary antibodies were applied for 1 h at room temperature. The immunoblots were developed using an ECL fluorescence detection system (Beyotime, Shanghai) and analyzed with a Tanon 4600 imaging system (Tanon Science and Technology Co., Ltd.).

2.15. Statistical analyses

R software (version 4.1.3) was employed to conduct all statistical analyses and visualizations. The Wilcoxon test was employed for comparisons of differences between two groups, whereas the Kruskal-Wallis test and one-way ANOVA were employed for comparisons

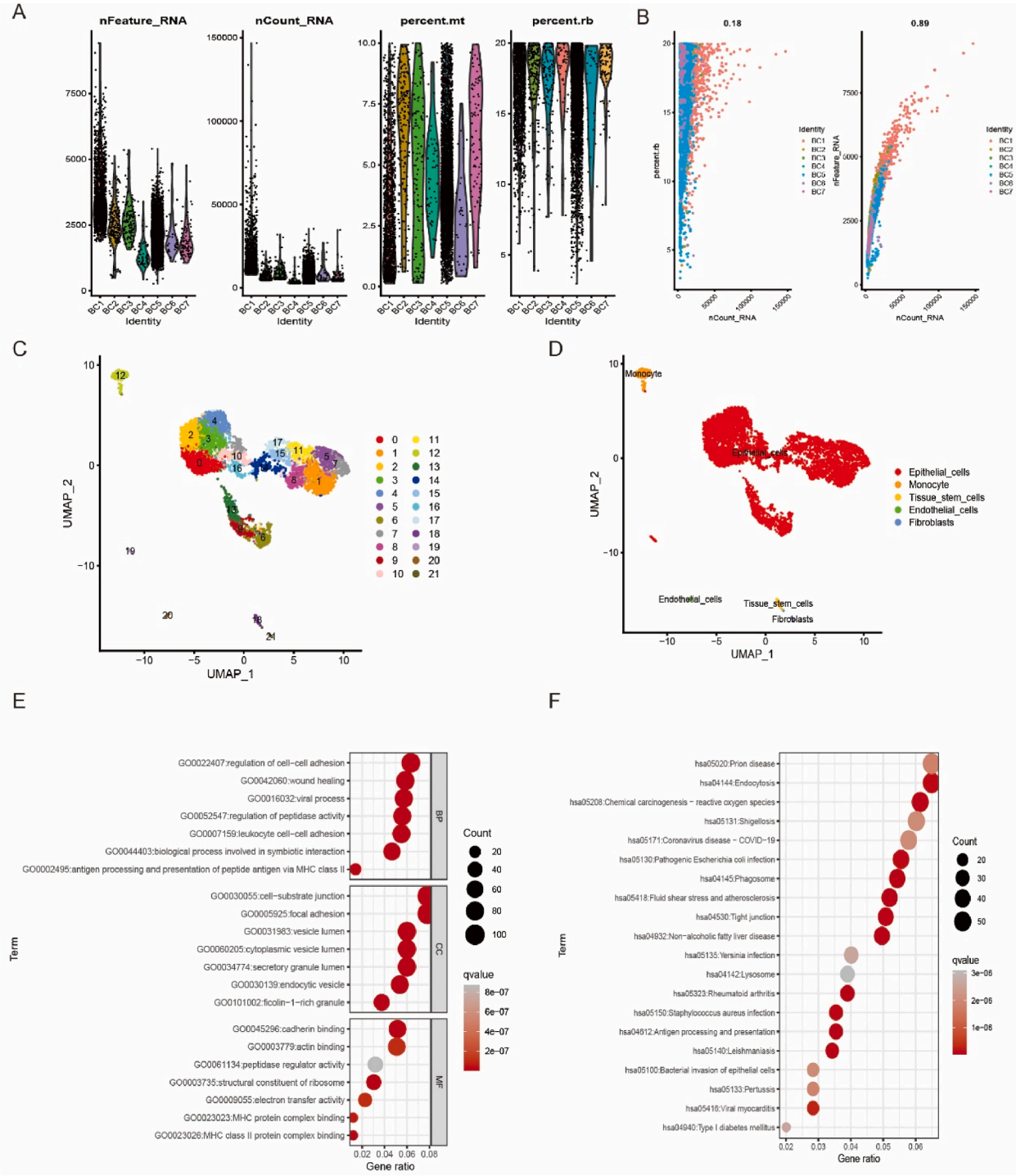


Fig. 2. Identification of marker genes for epithelial cell subpopulations. (A) After quality control of single-cell data, plot it on a violin chart to display nFeature-RNA and nCount-RNA. (B) Correlation analysis between nFeatures and nCounts. (C) The UMAP algorithm displays 22 cell clusters. (D) 22 cell clusters were annotated as 5 immune cell subpopulations. (E) GO analysis. (F) KEGG analysis.

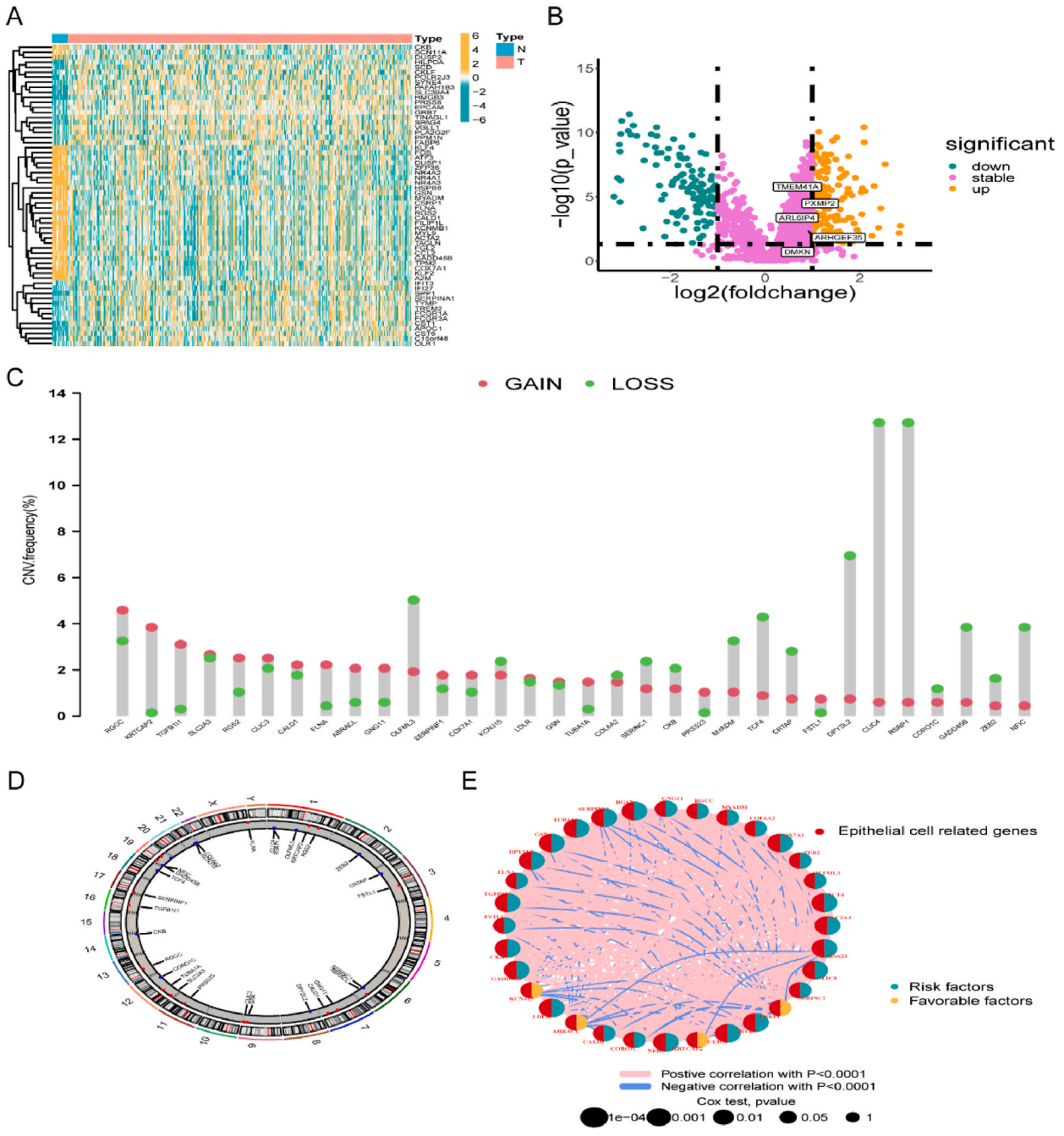


Fig. 3. Genetic characteristics of epithelial cell-associated genes in BLCA. In the TCGA cohort (A, B), 1579 marker genes were examined for differential expression between normal and tumor samples. A: heat map; B: volcano map. (C) The frequency of CNV changes in 32 prognostic genes in TCGA-BLCA. Red dot: Amplification frequency; Green: deletion frequency. (D) The positions of 32 prognostic genes on chromosomes. (E) The interaction network between 32 prognostic genes.

of differences between two or more groups. Survival analyses were conducted utilizing the K-M method, while correlations were assessed utilizing Pearson correlation analyses. *P*-value<0.05 exhibited statistical significance.

3. Results

3.1. Cell composition of bladder cancer

This research comprised a total of 7 single-cell samples from BLCA for analysis. Initially, the unqualified cells were filtered out to ensure quality control (Fig. 2A). Fig. 2B illustrates a strong positive correlation between the number of detected genes (nFeatures) and sequencing depth (UMI total, nCount), exhibiting a coefficient of 0.89. PCA was conducted using the first 2000 variable genes to reduce dimensionality, resulting in the identification of 22 cell clusters (Fig. 2C). After that, the annotation of the cell identity of every cluster was carried out by employing a reference dataset through the human primary cell atlas, with 22 cell clusters annotated as 5 immune cell subpopulations (Fig. 2D). Epithelial cells exist in multiple subpopulations with a higher number of cells. Therefore, 1579 marker genes specific to the epithelial cell subpopulation were included in subsequent studies. Functional enrichment analysis revealed substantial associations of these marker genes with various biological pathways, including antigen processing and presentation, chemical carcinogenesis, reactive oxygen species, symbiotic interactions, and focal adhesion (Fig. 2E and F).

3.2. Genetic characteristics of epithelial cell-associated genes in BLCA

The expression levels of 1579 marker genes were extracted from the TCGA-BLCA cohort for comparing the expression differences between normal and BLCA samples A total of 285 genes showed significant differential expression (DEGs) based on the criteria of $P < 0.05$ and Log2 (Fold Change) > 1 or < -1 (Fig. 3A and B). Through univariate Cox analysis, 32 prognosis-associated genes were acquired from obtained DEGs (Table 2). Subsequently, the copy number variations (CNV) of these 32 genes were analyzed. Notably, the CNV of OLFML3, KCNJ15, COL6A2, SERINC1, CKB, MYADM, TCF4, CRTAP, DPYSL2, CLIC4, RSRP1, CORO1C, GADD45B, ZEB2, and NFIC exhibited significant reduction. Conversely, the copy number of other genes demonstrated significant elevation (Fig. 3C). Fig. 3D illustrates the positions of CNVs of these genes on chromosomes. The gene network diagram depicts the molecular interactions between the 32 genes (Fig. 3E).

Table 2
Univariate Cox analysis of DEG.

id	HR	HR.95L	HR.95H	*P*-value
SLC2A3	1.006	1.004	1.009	<0.001
TCF4	1.050	1.031	1.069	<0.001
CLIC3	1.001	1.000	1.001	<0.001
DPYSL2	1.011	1.007	1.015	<0.001
NFIC	1.011	1.006	1.016	<0.001
SERPINF1	1.002	1.001	1.002	<0.001
GSN	1.001	1.001	1.002	<0.001
CRTAP	1.007	1.004	1.010	<0.001
LDLR	1.007	1.004	1.010	<0.001
RGS2	1.002	1.001	1.002	<0.001
TUBA1A	1.001	1.000	1.001	<0.001
CLIC4	1.002	1.001	1.002	<0.001
GADD45B	1.001	1.000	1.001	<0.001
CKB	1.001	1.000	1.001	<0.001
TGFB1I1	1.006	1.003	1.009	<0.001
PRSS23	1.006	1.003	1.009	<0.001
COX7A1	1.004	1.002	1.005	<0.001
ZEB2	1.037	1.018	1.056	<0.001
CORO1C	1.005	1.002	1.008	<0.001
MYADM	1.003	1.001	1.005	<0.001
KCNJ15	0.985	0.977	0.993	<0.001
RSRP1	0.981	0.971	0.991	<0.001
ABRACL	0.996	0.994	0.998	<0.001
SERINC1	1.004	1.002	1.006	<0.001
GNG11	1.009	1.004	1.014	<0.001
FSTL1	1.002	1.001	1.003	<0.001
CALD1	1.001	1.000	1.001	<0.001
COL6A2	1.000	1.000	1.000	<0.001
OLFML3	1.003	1.001	1.004	<0.001
KRTCAP2	0.866	0.795	0.942	<0.001
FLNA	1.000	1.000	1.000	<0.001
RGCC	1.003	1.001	1.004	<0.001

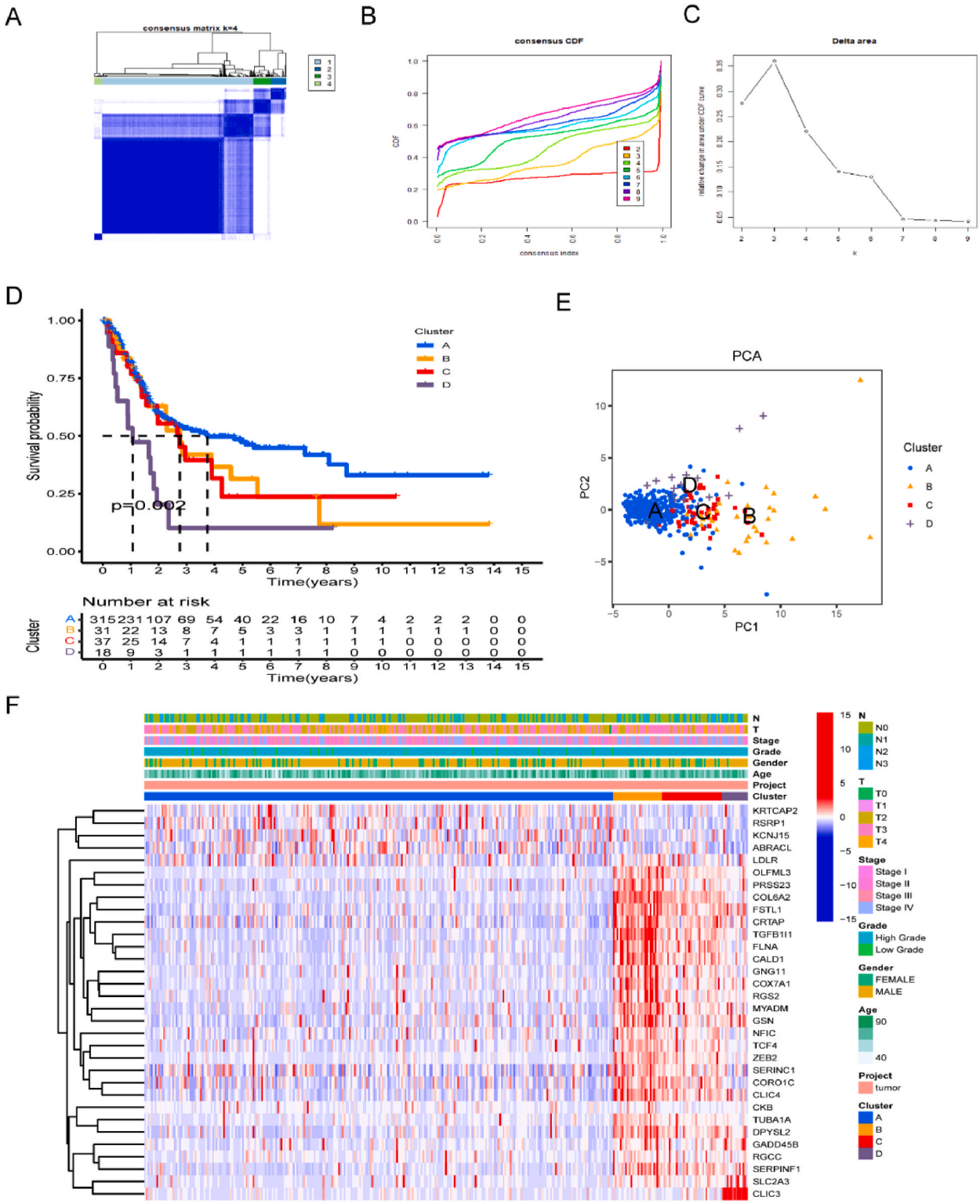


Fig. 4. Identification of four subtypes of BLCA associated with endothelial cells. (A) Consistency clustering cumulative distribution function (CDF) at k = 2 to 9. (B) The relative change in area under the CDF curve when k = 2 to 9. (C) When k = 4, BLCA patients in the TCGA cohort are categorized into four different subgroups. (D) Survival analysis of four different subtypes of BLCA patients. (E) Principal Component Analysis (PCA) between four different subtypes. (F) The heat map shows the relationship between four different subtypes and related clinical features in the TCGA-BLCA cohort.

3.3. Identification of four subtypes of BLCA associated with epithelial cells

Consistency clustering analysis was employed for classifying BLCA patients in TCGA into four different subtypes related to epithelial cells (Fig. 4A–C). K-M analysis assessed the OS of BLCA patients within the TCGA cohort (Fig. 4D). The findings confirmed significant differences in prognosis among the four different subtypes of BLCA patients. Clusters A and D exhibited the best and worst prognoses, respectively. In Fig. 4E, the PCA shows a significantly different distribution among clusters (A-D). Additionally, relevant clinical information was integrated with distinct subgroups. They were depicted through heat maps to visually represent the distribution of clinical attributes within different subgroups (Fig. 4F).

3.4. Functional annotation analysis between four different subtypes

GSEA was performed to analyze the KEGG pathway across the four different subgroups. The results revealed significant downregulation of chemokine signaling pathways, cell adhesion molecules, and cytokine-cytokine receptor interactions in cluster A (Fig. 5A). In clusters B and C, there was a notable increase in the upregulation of extracellular matrix (ECM) receptor interaction, cytokine receptor interaction, and chemokine signaling pathway (Fig. 5B and C). The chemokine signaling pathway and leukocyte cross-endothelial migration were significantly upregulated in cluster D. Moreover, the drug metabolism of cytochrome P450 and cytochrome P450 significantly downregulated the metabolism of exogenous substances in cluster D (Fig. 5D).

3.5. Construction and validation of an artificial neural network prognostic model

After importing the 32 genes related to prognosis into the STRING database, they were visualized using the Cytohubba plugin in Cytoscape. Finally, a PPI network containing 17 genes was obtained, with darker colors indicating more adjacent nodes (Fig. 6A). Then, an artificial neural network prognostic model based on 17 genes was developed and evaluated. In the TCGA-BLCA dataset,

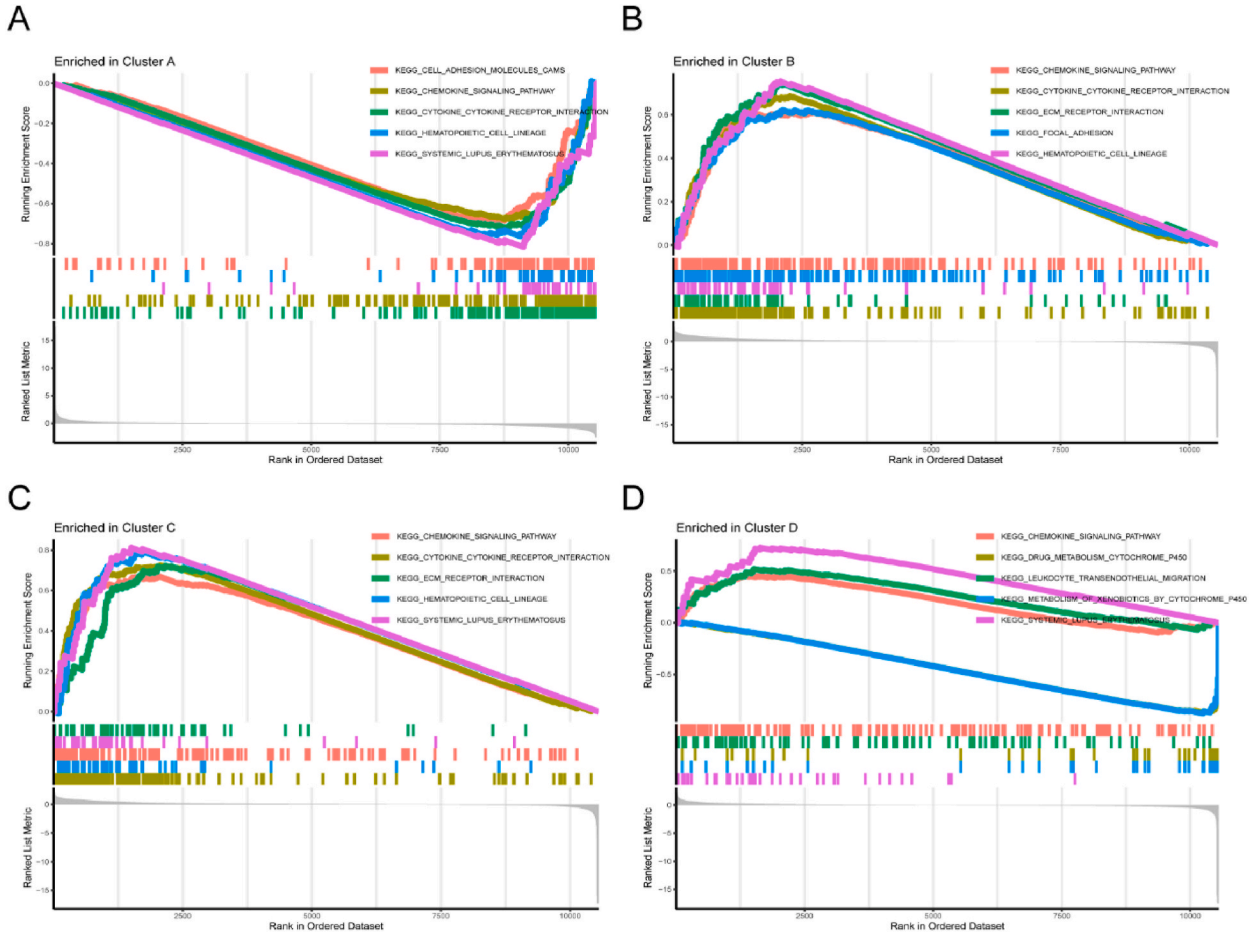


Fig. 5. Functional annotation analysis between four different subtypes. KEGG analysis between four different subtypes (A–D). A: Cluster A; B: Cluster B; C: Cluster C; D: Cluster D.

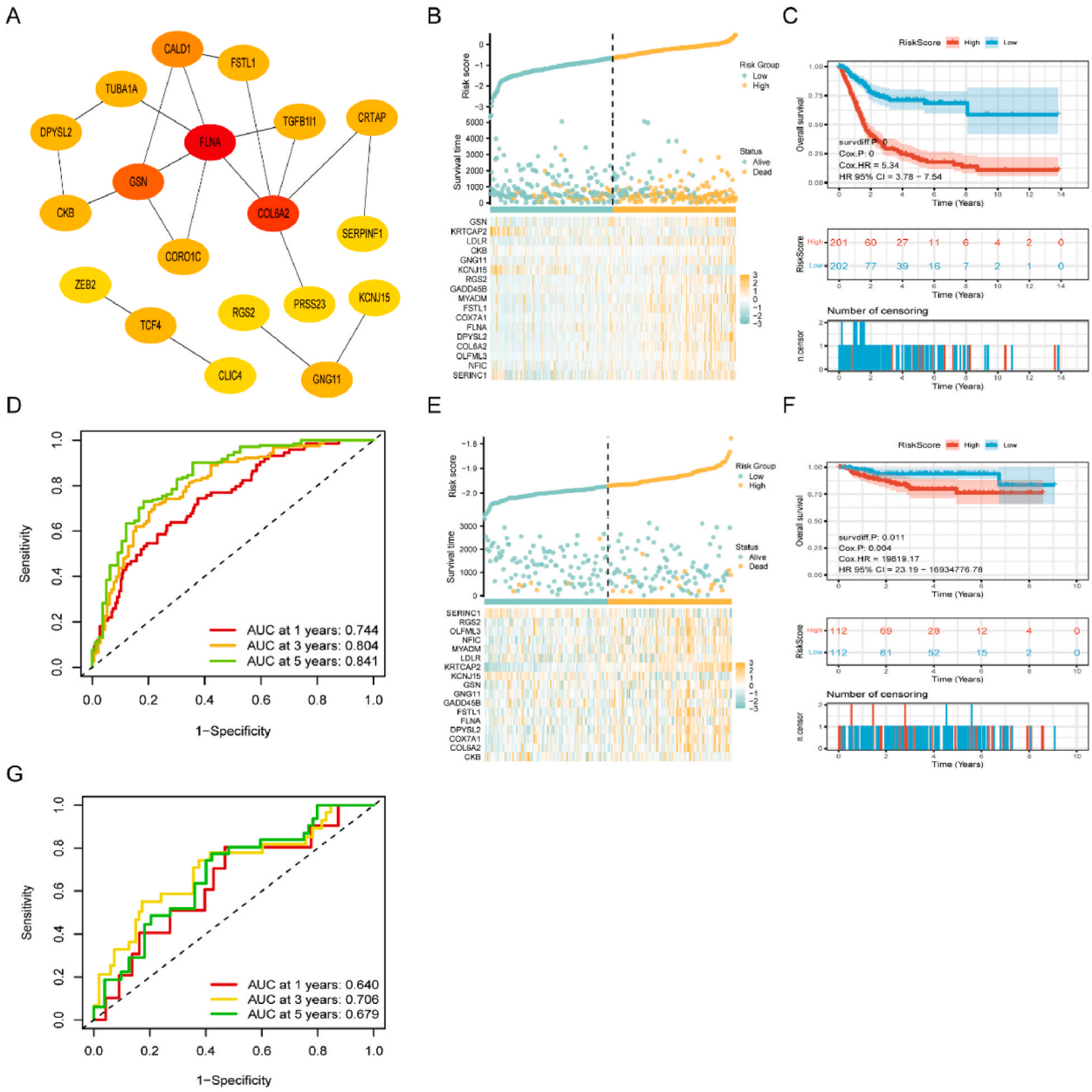


Fig. 6. Construction and verification of an artificial neural network prognostic model. (A) A PPI network containing 17 genes. In the training set, (B) the risk score and survival status distribution of BLCA patients. The heat map demonstrates the difference in expression levels of 17 risk genes between the two groups. (C) The survival time of the high-risk group is substantially lower than low-risk group. (D) The ROC curves and AUC of 1, 3, and 5 years evaluated the clinical predictive power of the model. In the validation set, the risk score and survival status distribution of (E) BLCA patients. The heat map depicts the difference in expression levels of 17 risk genes between the two groups. (F) The survival time of the high-risk group is notably shorter than low-risk group. (G) The ROC curves and AUC of 1, 3, and 5 years validated the clinical predictive power of the model.

individuals were separated into two risk categories (high and low) based on the median risk score (Fig. 6B). Higher risk scores correlated with lower survival rates and increased mortality. The K-M analysis results demonstrated that patients classified as low-risk had a better OS compared to those in the high-risk category (Fig. 6C). A plot of the ROC curve was generated to validate the predictive ability of the model. Following the assessment, the AUC under the ROC curve was determined to be 0.744, 0.804, and 0.841 at 1, 3, and 5-year time intervals, respectively (Fig. 6D). Furthermore, the predictive power of the model was also verified through the validation set. In line with findings from the training group, the validation cohort exhibited a correlation wherein higher risk scores were linked to lower patient survival rates and increased mortality (Fig. 6E). The results from the K-M analysis showed that patients in the low-risk group had a longer survival time compared to those in the high-risk group (Fig. 6F).At time points of 1 year, 3 years, and 5 years, the

AUC was 0.640, 0.706, and 0.679, respectively (Fig. 6G). The findings suggest that the model exhibits reliable prognostic ability.

3.6. The association between risk score and clinical attributes of artificial neural network models

In Fig. 7A, the clinical and pathological features of individuals in the high-risk and low-risk categories are showcased. The risk score demonstrated a significant relationship with clinical features such as tumor size ($P < 0.001$), Grade staging ($P < 0.001$), Stage staging ($P < 0.001$), and lymph node metastasis ($P < 0.001$) (Fig. 7B–E).

3.7. Development of new BLCA-related nomogram

A new nomogram was developed in this research that includes age, gender, Grade staging, Stage staging, tumor size, lymph node metastasis, and risk score (Fig. 8A). Calibration chart analysis revealed the strong predictive ability of the column chart for actual survival outcomes (Fig. 8B). The results of the risk accumulation chart showed that over time, the cumulative risk for high-risk patients considerably exceeded that of low-risk patients (Fig. 8C).

3.8. The correlation between risk score and tumor microenvironment and somatic mutations

Utilizing the CIBERSORT algorithm, the correlation between risk score and immune cell infiltration was analyzed. Substantial differences were observed in numerous immune cells between both risk groups. In the high-risk group, a considerable upregulation was observed in Mast cells resetting ($P = 0.016$) and T cells CD4 memory resetting ($P = 0.008$). In the low-risk group, naïve B cells ($P = 0.041$), Regulatory T cells (Tregs) ($P = 0.045$), and CD8 T cells ($P = 0.037$) exhibited considerable upregulation (Fig. 9A). Additionally,

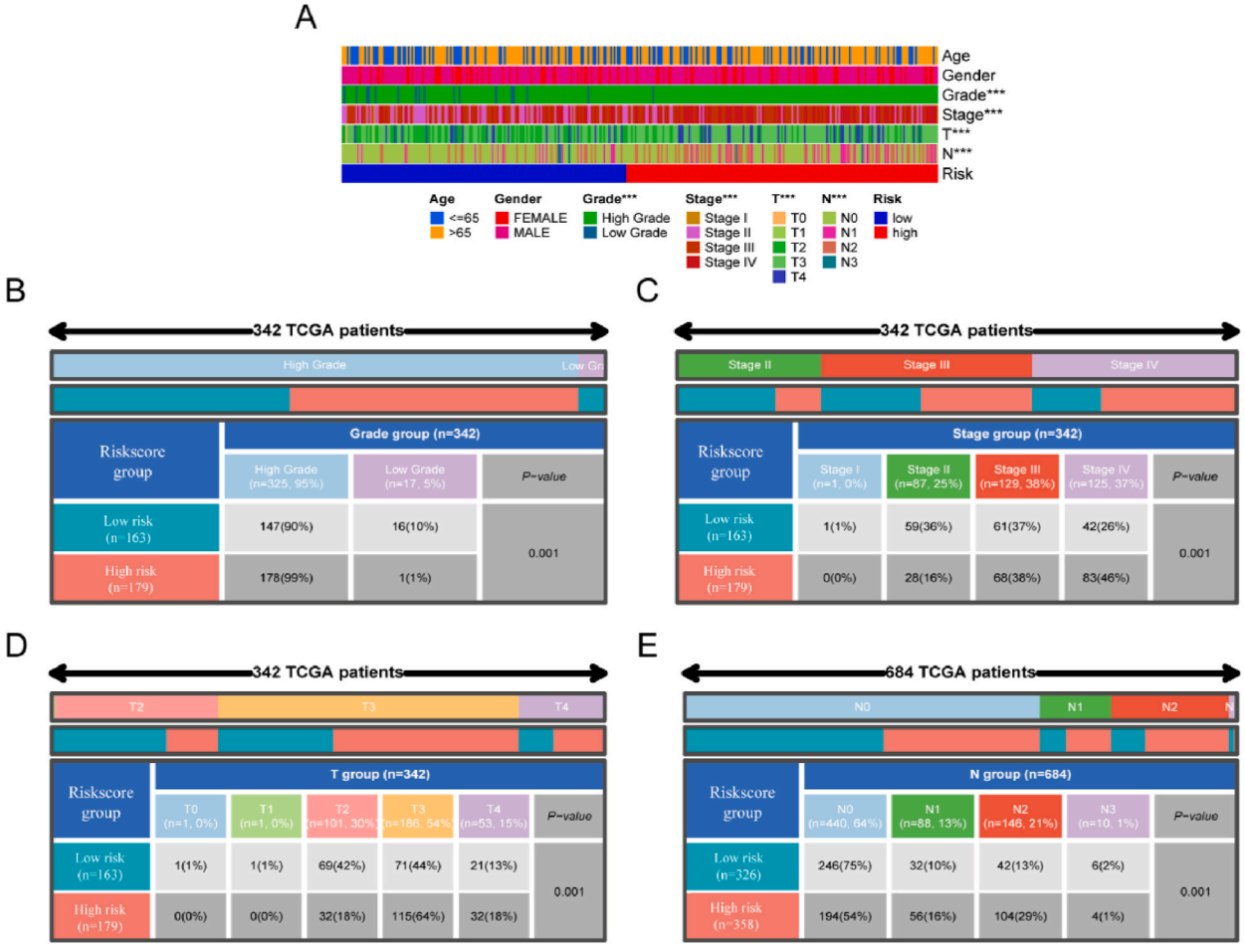


Fig. 7. The relationship between risk scores of artificial neural network models and clinical features. (A) Differences in clinical and pathological characteristics between high-risk and low-risk groups. (B) Variations in the proportion and quantity of Grade staging between high-risk and low-risk groups. (C) Variances in the proportion and quantity of Stage staging between two risk groups. (D) The proportion and quantity differences of tumor size between the above-mentioned groups. (E) Differences in the proportion and quantity of lymph node metastasis between both risk groups.

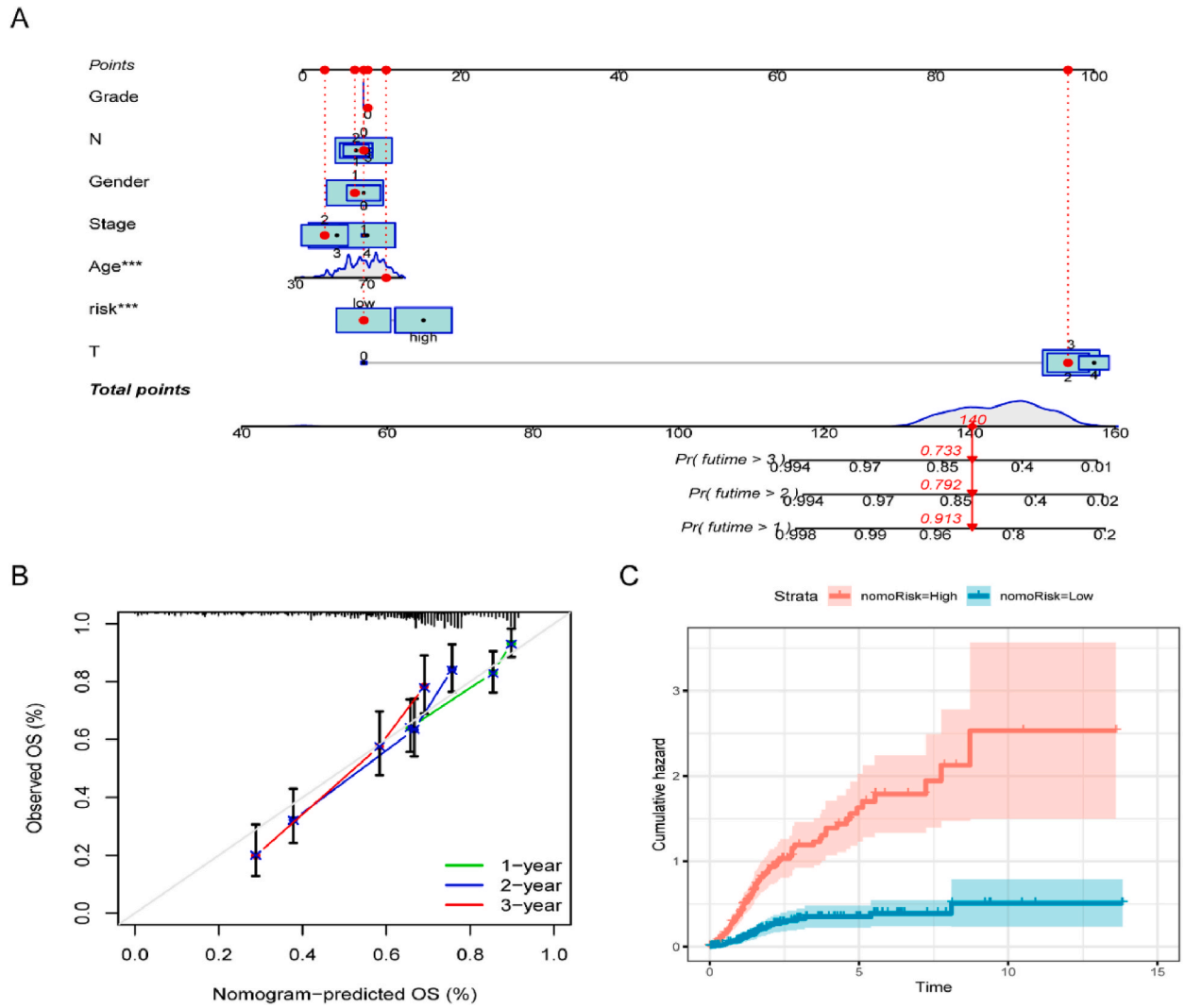


Fig. 8. Development of new nomogram related to BLCA. (A) Development of a nomogram that combines clinical features and risk scores. (B) Nomogram calibration curves for 1 year, 2 years, and 3 years. (C) A risk accumulation chart divided by risk stratification.

an analysis was performed to compare somatic mutations between the two risk groups, aiming to investigate the correlation between BLCA mutations and risk scores. The high-risk group exhibited a greater mutation evaluation rate (94 %) (Fig. 9B and C). Among them, the high-risk group (49 %) exhibited a substantially heightened frequency of TP53 mutations in comparison to the low-risk group (47 %).

3.9. Drug sensitivity analysis and drug prediction analysis

The outcomes of drug sensitivity analysis demonstrated that high-risk patients exhibited higher sensitivity to Axitinib and Cisplatin (Fig. 10A and Fig. B), whereas low-risk patients showed higher sensitivity to Nilotinib and Sorafenib (Fig. 10C and D). Using the TCGA-BLCA cohort, a comparative analysis was carried out to identify potential small molecule targeted medications for high-risk patients by examining participants from two distinct risk categories. A total of 59 genes exhibiting increased expression and 57 genes showing decreased expression were submitted to the cMAP database for further analysis. By analyzing connectivity scores, four major targeted drugs were identified (carbamazepine, forskolin, thioridine, and trichostatin). In summary, the artificial neural network model can provide meaningful guidance for drug selection in BLCA patients, thereby advancing the progress of personalized and precise drug therapy for BLCA patients.

3.10. Identification and molecular docking of important genes in the model

Two machine learning algorithms (XGboost and LightGBM) were utilized to analyze the importance of 17 genes in the model,

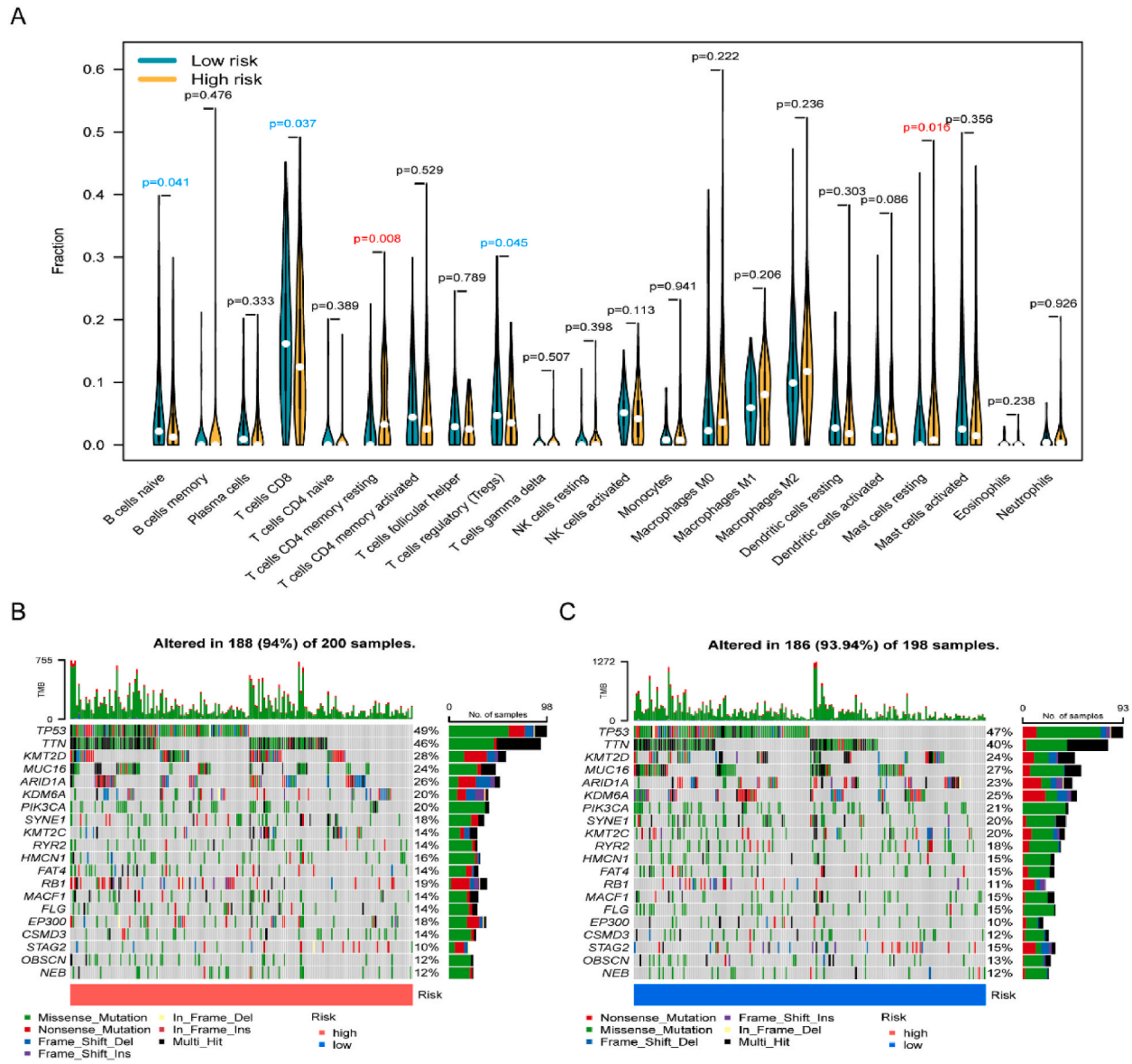


Fig. 9. The correlation between risk score and tumor microenvironment and somatic mutations. (A) Variations in immune cell infiltration between high-risk and low-risk groups. Differences in somatic mutations between aforementioned groups (B, C). B: High-risk group; C: Low-risk group.

aiming to identify feature genes that make significant contributions to the model. After that, SHAP values were employed to explain the involvement of every gene in the model. As shown in Fig. 11 A and Fig. 11B, the SHAP value of NFIC ranked first consistently in the results of both algorithms. This indicates that NFIC contribution remains stable and consistent across various algorithms. Hence, NFIC was designated as the key gene that contributed the most to the model prediction. For subsequent analysis of the most efficient small molecule drugs targeting NFIC proteins, molecular docking analysis was conducted between NFIC proteins and the top five drugs screened. A binding energy < -5 kcal/mol denotes good binding ability, whereas a binding energy < -7 kcal/mol indicates strong activity [30]. Our molecular docking analysis revealed that all small-molecule drug compounds demonstrated binding energies lower than -6 kcal/mol, suggesting a favorable interaction with NFIC, the target protein. The binding energies of carbamazepine, forskolin, thioridine, and trichostatin were -6.8 kcal/mol, -7.0 kcal/mol, -6.2 kcal/mol, and -6.8 kcal/mol, respectively (Fig. 11C–F).

3.11. Validation of model genes in qRT-PCR

Three cell lines, SV-HUC-1, EJ-1, and BT-B, were selected and the gene expression of 17 each was validated. The findings exhibited substantially elevated gene expression of COL6A2, DPYSL2, FLNA, GNG11, GSN, LDLR, MYADM, NFIC, OLFML3, and RGS2 in normal cells than in cancer cells. The KRTCAP2 expression in cancer cells was significantly elevated in comparison to normal cells (Fig. 12A-

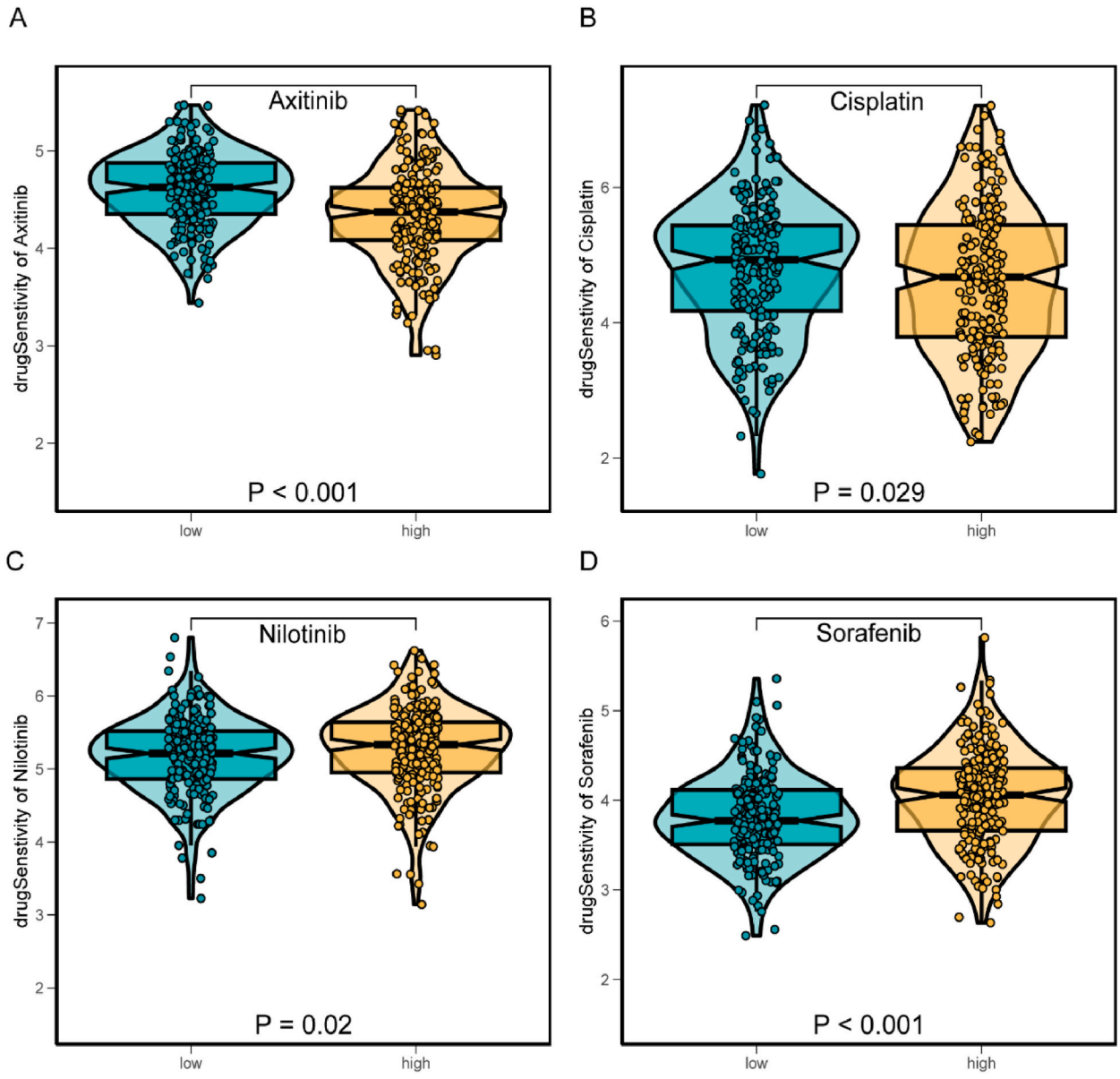


Fig. 10. Drug sensitivity analysis. (A,B) Axitinib and Cisplatin are more sensitive to high-risk patients. (C, D) Nilotinib and Sorafenib are more sensitive to low-risk patients.

Q). In summary, these genes were significantly linked to the development of tumor cells.

3.12. KRTCAP2 promotes the progression of bladder cancer

We evaluated the efficacy of detecting small interfering RNAs in bladder cancer cell lines BT-B and EJ-1. Among them, si-KRTCAP2#1 showed the highest inhibition efficiency (Fig. 13A and B). Suppression of KRTCAP2 gene expression in BT-B and EJ-1 cell lines led to a notable reduction in the viability of malignant bladder cells (Fig. 13C and D). The EdU staining results indicated that the suppression of KRTCAP2 notably reduced the proliferation capacity of bladder cancer cell lines in vitro (Fig. 13E). Transwell results indicated that blocking KRTCAP2 notably reduced the migratory capacity of in vitro bladder cancer cell lines (Fig. 13F). Considering the important role of the PI3K-AKT signaling pathway in tumorigenesis, we examined the alterations in the, PI3K-AKT signaling pathway before and after KRTCAP2 inhibition in the EJ-1 cell line. Upon inhibition of KRTCAP2 in the EJ-1 cell line, phosphorylation of PI3K and AKT was significantly inhibited. Meanwhile the density of P53 protein was restored (Fig. 13G–J).

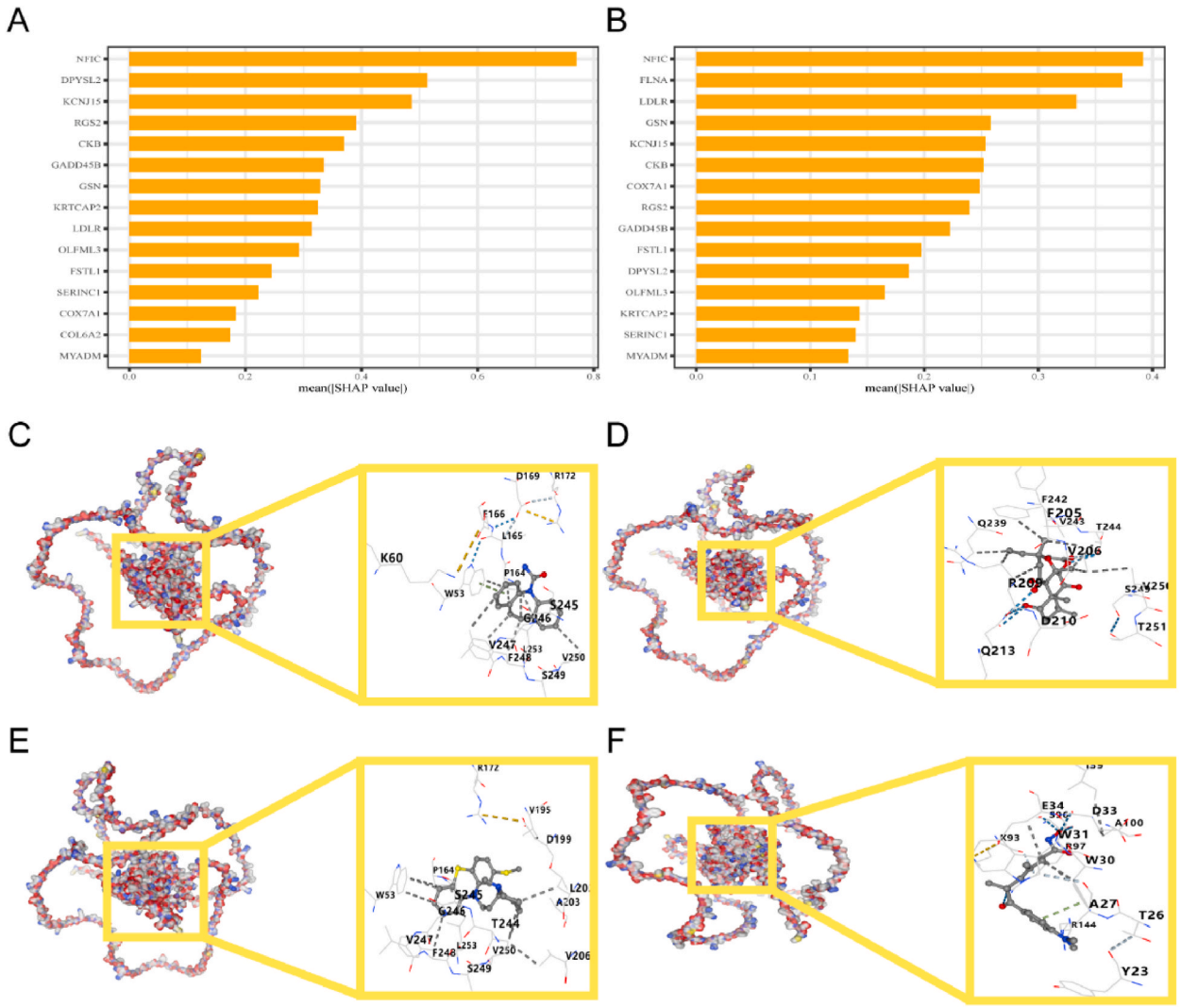
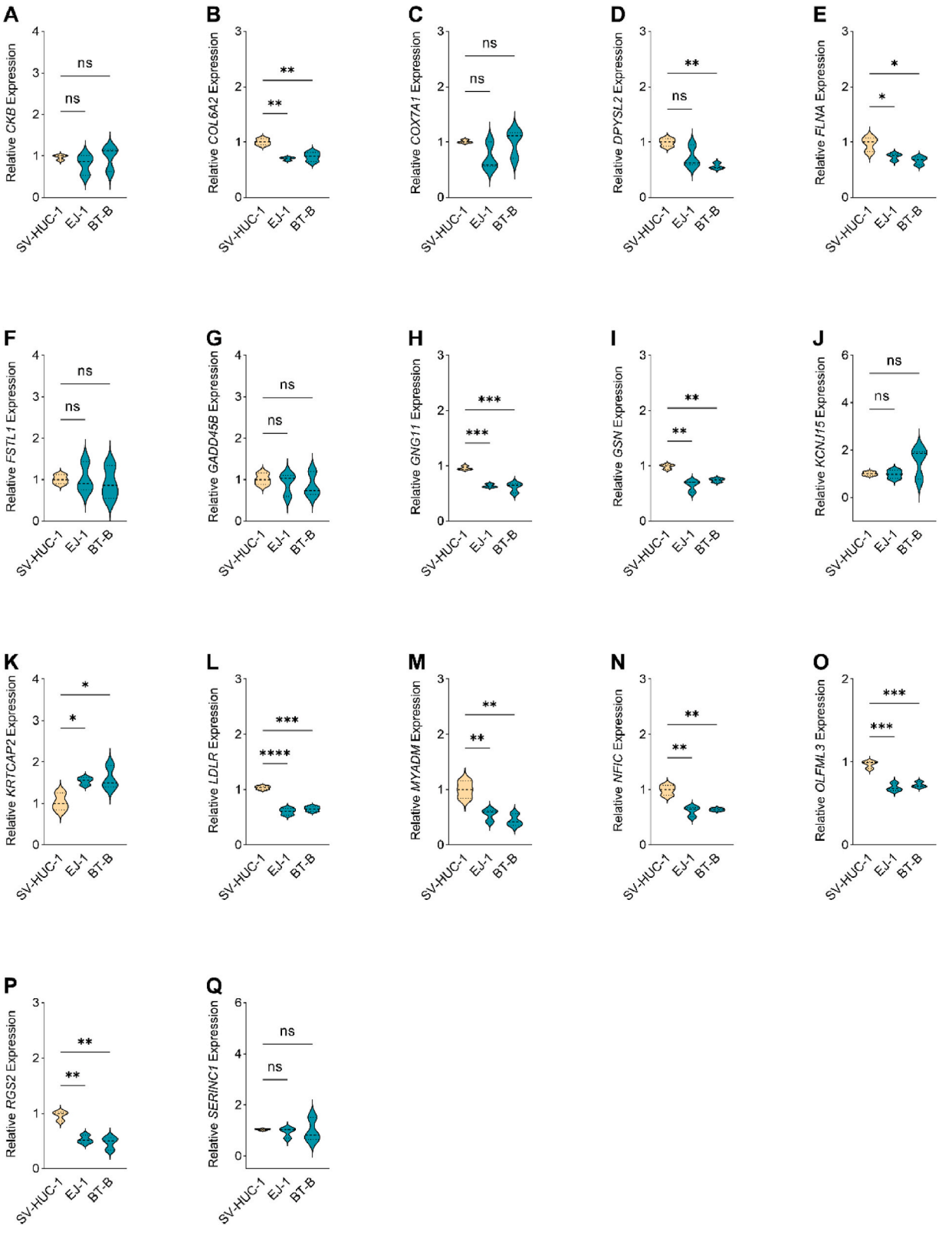


Fig. 11. (A, B) Calculation of the contribution value of every gene in the constructed model using two machine learning algorithms (XGboost and LightGBM), and their arrangement from maximum to minimum. The SHAP value depicts the absolute average impact of each gene on the model. A: XGboost; B: LightGBM. (C–F) depicts the binding positions of ligands in proteins and demonstrates the interactions between ligands and residues. The docking method is a structure that is based on blind docking. C: Carbamazepine; D: Forskolin; E: Thioridazine; F: Trichostatin-a.

4. Discussion

BLCA, a common urological cancer, is known for its diverse range of malignant tumors originating from epithelial cells. The disease poses a high risk of recurrence and advancement, leading to considerable pain and distress for patients [31] The urinary tract epithelium is composed of several cell layers lining the bladder wall, forming an impermeable barrier for urine [32]. It can prevent infection, fluid exchange, and toxin damage [32,33]. Under physiological conditions, the renewal of the urinary tract epithelium is relatively slow compared to other tissues [34]. The study by Minal Garg et al. suggests that due to the accumulation of mutations, the transition from epithelium to stroma may promote the development of urothelial carcinoma stem cells, thereby facilitating cell colonization at a distance and ultimately promoting the onset and progression of tumors [35]. However, the mechanism of action of epithelial cells in BLCA still requires further research to explore. In this research, a comprehensive analysis was carried out on a single-cell dataset containing 7 BLCA samples. The results showed that the epithelial cell subpopulation was the main cell subpopulation, and the marker genes contained in the epithelial cell subpopulation were included in subsequent analyses. Subsequently, based on the marker genes of epithelial cell subpopulations, an artificial neural network prognostic model containing 17 feature genes was constructed using the TCGA-BLCA dataset. This prognostic model can accurately stratify BLCA patients. The model was further validated on independent datasets. The findings of this research may enhance our understanding of the mechanism of epithelial cells in BLCA and pave the way for personalized treatment approaches for BLCA.

A functional annotation analysis was conducted on the marker genes of epithelial cell subpopulations. The outcomes suggest that



(caption on next page)

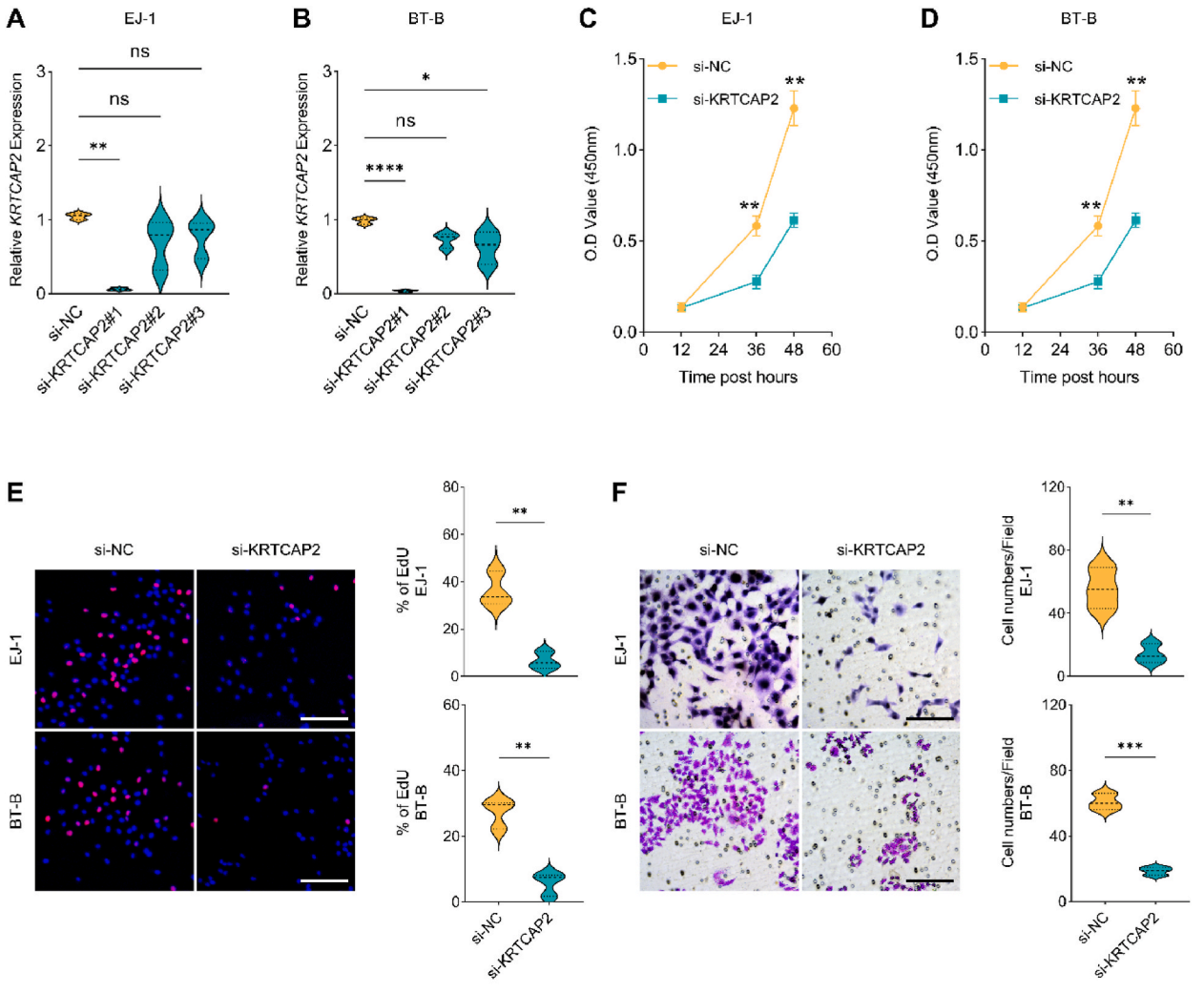
Fig. 12. qRT qPCR was utilized to validate 17 gene expression in SV-HUC-1, EJ-1, and BT-B cell lines (A). CKB gene expression was observed in cells. (B) COL6A2 gene expression in cells. (C) COX7A1 gene expression in cells. (D) DPYSL2 gene expression in cells. (E) FLNA gene expression in cells. (F) FSTL1 gene expression in cells. (G) GADD45B gene expression in cells. (H) GNG11 gene expression in cells. (I) GSN gene expression in cells. (J) KCNJ15 gene expression in cells. (K) KRTCAP2 gene expression in cells. (L) LDLR gene expression in cells. (M) MYADM gene expression in cells. (N) NFIC gene expression in cells. (O) OLFML3 gene expression in cells. (P) RGS2 gene expression in cells. (Q) SERINC1 gene expression in cells. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

these genes are substantially correlated with biological pathways, including chemical carcinogenesis of reactive oxygen species, antigen processing and presentation, symbiotic interactions, and focal adhesion. Adhesive plaques are a group of macromolecular proteins that can connect the end of specialized actin fibers to the ECM and facilitate cell migration. In pathology, most studies have suggested that adhesive plaques are crucial for the process of tumor metastasis [36–39]. Furthermore, adhesive plaques exert a critical impact on epithelial-mesenchymal transition [40–43].

In this study, four different subtypes associated with epithelial cells were identified. Cluster A has a significant survival advantage. The changes in the intercellular adhesion system exert a critical impact on the onset and progression of cancer [44]. Cadherin plays a key role in cell-to-cell binding within epithelial tissue and plays a critical part in numerous significant pathways linked to cancer progression, such as the formation of cancer stem cell traits and the transition from epithelial to mesenchymal states [45–48] E-cadherin serves as the main mediator of intercellular adhesion in epithelial tissue and is expressed by most normal epithelial cells [45,49,50]. The upregulation of immune expression of E-cadherin is related to good prognosis in different cancers [51–53]. The expression of E-cadherin is usually absent with the progression of grading and staging, and this is often accompanied by an increase in the expression of N- or *P*-cadherin. This phenomenon is called calcium-binding protein conversion [49,54–56]. Previous studies reported that CS played a late role in a variety of malignant tumors, including BLCA, which led to stronger tumor invasion and malignant phenotype and worse prognosis of patients [45,54,57–60]. This is consistent with our analysis results, where the cell adhesion molecule signaling pathway is significantly downregulated in cluster A, potentially explaining the better survival time of cluster A patients.

In this study, an artificial neural network prognostic model containing 17 feature genes was constructed. Utilizing the model's risk score, we can accurately categorize BLCA patients into high-risk and low-risk groups. Those classified as low-risk demonstrate a considerable survival benefit. The findings of the ROC and K-M curves in the validation and training groups indicate that the model has good accuracy and higher extrapolation ability. Furthermore, two types of machine learning, XGboost and LightGBM, were utilized to analyze and identify the most significant contributing feature genes among the 17 feature genes, and SHAP was utilized for model interpretation. Among the two machine learning algorithms, the SHAP value of NFIC always ranks first. NFIC, as a member of the transcription factor nuclear factor I family, regulates cell differentiation and proliferation through transcriptional control of target genes [61–63]. CCND1 and FoxF1 are effective inducers of epithelial-mesenchymal transition, invasiveness, and tumorigenicity. NFIC can directly inhibit the expression of CCND1 and FoxF1, thus serving as a tumor inhibitor [64,65]. The research of Isidoro Cobo et al. identified NFIC as a regulator of pancreatic acinar cell function, which can regulate ribosomal biology and endoplasmic reticulum stress in pancreatic acinar cells, thereby inhibiting the progress of pancreatic cancer [66]. The study by Hong Zhang et al. identified NFIC as a new therapeutic target for lung squamous cell carcinoma. NFIC can exert anti-cancer effects in lung squamous cell carcinoma by regulating CASC2 [67]. Furthermore, NFICs still play a pro-cancer role in some tumors. The study by Gang Xu et al. revealed that NFIC can upregulate LBX2-AS1 expression, thereby targeting miR-491-5p/ZNF703 to promote the development of gastric cancer [68]. However, there remains inadequate research on the mechanism of action of NFICs in BLCA. Therefore, more research is needed to explore their mechanism in BLCA. Furthermore, molecular docking was utilized for identifying the small molecule drug, Forskolin, with the highest affinity for the hub gene (NFIC). Although the adenylate cyclase agonist forskolin is known as a cAMP-presenting agent, it also demonstrates potential as a cancer therapeutic agent. Forskolin can also be utilized in combination with other anti-cancer agents. Forskolin can enhance the cytotoxicity induced by paclitaxel and H3K27me2/3 demethylase inhibitor GSKJ-4 in non-small cell lung cancer and acute myeloid leukemia cells [69,70]. The co-delivery of paclitaxel and hairy throat hormone through liquid crystal nanoparticles can hinder the dryness of breast cancer stem cells and reverse the chemoresistance, thus stimulating strong anti-tumor activity [71]. These results highlight the possible importance of Forskolin for cancer therapy. However, further studies are necessary to confirm Forskolin's clinical role in BLCA.

Additionally, this research examined the variations in immune cell infiltration between patients at high and low risk using the CIBERSORT algorithm. Tregs are key mediators in shaping the immune microenvironment of multiple diseases, including cancer [72]. In tumors, Tregs accumulate and inhibit anti-tumor immunity by expressing anti-inflammatory cytokines and co-inhibitory molecules that regulate tumor cells and other tumor-related immune subsets [73–76]. Therefore, Tregs have long been considered as an unresponsive suppressor of anti-tumor immunity. They have been found to accumulate in colorectal cancer, glioma, and hepatocellular carcinoma, thereby inhibiting anti-tumor responses and correlating with poor prognosis in patients [77–79]. Prior investigations have also found that a large number of Tregs exist in tumor specimens and peripheral blood monocytes of patients with BLCA. The high infiltration content of Tregs may lead to an increase in cancer recurrence rate [80,81]. Recent studies have suggested that although high Treg infiltration adversely affects the prognosis of many patients with solid tumors, this prognostic effect varies depending on the type of tumor [82]. Malin E. Winerdal et al. found that Tregs can inhibit the expression of MMP2 and regulate the invasion of BLCA [83]. This seems consistent with our analysis, where Tregs are significantly upregulated in low-risk patients. It is believed that this may be the reason why low-risk patients have a survival advantage. However, additional analysis is required to confirm the involvement of Tregs in BLCA.



(caption on next page)

Fig. 13. KRTCAP2 promotes the progression of bladder cancer. (A–B) PCR was performed to detect the inhibition efficiency of three small interfering RNAs against different targets in EJ-1 versus BT-B cell lines. (C–D) Alterations in cell viability before and after KRTCAP2 inhibition in EJ-1 and BT-B cell lines. (E) Alterations in cell proliferative capacity before and after KRTCAP2 inhibition in EJ-1 and BT-B cell lines. (F) Alterations in cell migration capacity before and after KRTCAP2 inhibition in EJ-1 and BT-B cell lines. (G–J) Alterations in the PI3K-AKT signaling pathway before and after KRTCAP2 inhibition in the EJ-1 cell line. N = 3/group. *≤0.05, **≤0.01, ***≤0.001, ****≤0.0001. The results are presented as mean ± SD. Note bene: The uncut version of the original Fig. 13G can be found in Supplementary Material 1.

The goal of this study is to build an artificial neural network prognostic model and accurately stratify BLCA patients to provide personalized treatment for each patient. Therefore, drug sensitivity was conducted to select drugs that exhibit different activities in distinct risk groups. Patients at a high risk showed increased sensitivity to Cisplatin and Axitinib, whereas those at a low risk exhibited heightened sensitivity to Sorafenib and Nilotinib. These findings can serve as valuable insights for clinical doctors to develop personalized treatment strategies.

There are still some inevitable shortcomings in current research. On the one hand, the analysis work of this research was based on data from public databases, and only qRT PCR validation was conducted. Therefore, the outcomes of this research still need to be further verified in multicenter trials and larger patient cohorts. On the other hand, further experimentation in vivo and in vitro is essential to fully examine and authenticate the operational mechanism of the hub gene NFIC in BLCA.

5. Conclusion

In summary, the artificial neural network prognostic model constructed in this study can accurately stratify the risk of BLCA patients, predict patient prognosis accurately, and provide additional support for the importance and applicability of medication decision-making for BLCA patients. This prognostic model is expected to enhance clinical decision-making and optimize treatment strategies.

Data Availability Statement

All data comes from TCGA database (https://portal.gdc.cancer.gov) and GEO database (https://www.ncbi.nlm.nih.gov/geo/). Readers can download and use it for free.

Funding

This project was supported by National Natural Science Foundation of China Youth Program (81802580).

Ethics statement

Not applicable.

Consent for publication

Not applicable.

CRediT authorship contribution statement

Fan Zhao: Formal analysis, Conceptualization. **Kun Zhang:** Software. **Limin Ma:** Validation. **Yeqing Huang:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e34632.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] E. Compérat, M.B. Amin, R. Cathomas, et al., Current best practice for bladder cancer: a narrative review of diagnostics and treatments, Lancet (London, England) 400 (10364) (2022) 1712–1721, https://doi.org/10.1016/s0140-6736(22)01188-6. Nov 12.

[2] H. Sung, J. Ferlay, R.L. Siegel, et al., Global cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality Worldwide for 36 cancers in 185 Countries, CA: a cancer journal for clinicians 71 (3) (May 2021) 209–249, https://doi.org/10.3322/caac.21660.

[3] V.G. Patel, W.K. Oh, M.D. Galsky, Treatment of muscle-invasive and advanced bladder cancer in 2020, CA: a cancer journal for clinicians 70 (5) (Sep 2020) 404–423, https://doi.org/10.3322/caac.21631.

[4] Y. Lotan, A.M. Kamat, M.P. Porter, et al., Key concerns about the current state of bladder cancer: a position paper from the bladder cancer Think Tank, the bladder cancer Advocacy network, and the Society of urologic Oncology, Cancer 115 (18) (2009) 4096–4103, https://doi.org/10.1002/cncr.24463. Sep. 15.

[5] M. Sun, Q.D. Trinh, Diagnosis and staging of bladder cancer, Hematol. Oncol. Clin. N. Am. 29 (2) (Apr 2015) 205–218, https://doi.org/10.1016/j.hoc.2014.10.013.

[6] S.S. Chang, B.H. Bochner, R. Chou, et al., Treatment of non-Metastatic muscle-invasive bladder cancer: AUA/ASCO/ASTRO/SUO guideline, J. Urol. 198 (3) (Sep 2017) 552–559, https://doi.org/10.1016/j.juro.2017.04.086.

[7] C. You, X.M. Piao, K. Kang, Y.J. Kim, K. Kang, Integrative transcriptome Profiling reveals SKA3 as a novel prognostic marker in non-muscle invasive bladder cancer, Cancers (2021) 13, https://doi.org/10.3390/cancers13184673. Sep. 17.

[8] F. Xie, C. Huang, F. Liu, et al., CircPTPRA blocks the recognition of RNA N(6)-methyladenosine through interacting with IGF2BP1 to suppress bladder cancer progression, Mol. Cancer 20 (1) (2021) 68, https://doi.org/10.1186/s12943-021-01359-x. Apr 14.

[9] J. Nie, T. Liu, T. Mao, et al., Transcriptome sequencing and single-cell sequencing analysis identify GARS1 as a potential prognostic and immunotherapeutic biomarker for multiple cancers, including bladder cancer, Front. Immunol. 14 (2023) 1169588, https://doi.org/10.3389/fimmu.2023.1169588.

[10] R. Ahmed, T. Zaman, F. Chowdhury, et al., Single-cell RNA sequencing with Spatial transcriptomics of cancer tissues, Int. J. Mol. Sci. (2022) 23, https://doi.org/10.3390/ijms23063042. Mar 11.

[11] Z. Chen, L. Zhou, L. Liu, et al., Single-cell RNA sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma, Nat. Commun. 11 (1) (2020) 5077, https://doi.org/10.1038/s41467-020-18916-5. Oct 8.

[12] H. Lai, X. Cheng, Q. Liu, et al., Single-cell RNA sequencing reveals the epithelial cell heterogeneity and invasive subpopulation in human bladder cancer. International journal of cancer, Dec 15 (12) (2021) 149, https://doi.org/10.1002/ijc.33794.

[13] K. Xu, W. Zhang, C. Wang, et al., Integrative analyses of scRNA-seq and scATAC-seq reveal CXCL14 as a key regulator of lymph node metastasis in breast cancer, Hum. Mol. Genet. 30 (5) (2021) 370–380, https://doi.org/10.1093/hmg/ddab042. Apr 27.

[14] J. Peng, B.F. Sun, C.Y. Chen, et al., Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma, Cell Res. 29 (9) (Sep 2019) 725–738, https://doi.org/10.1038/s41422-019-0195-y.

[15] T. Stuart, A. Butler, P. Hoffman, et al., Comprehensive integration of single-cell data, Cell. Jun 13 (7) (2019) 177, https://doi.org/10.1016/j.cell.2019.05.031, 1888-1902.e21.

[16] Y. Chen, X. Zhou, L. Ji, et al., Construction and analysis of a joint diagnostic model of machine learning for cryptorchidism based on single-cell sequencing, Birth Defects Research 116 (3) (2024) e2316, https://doi.org/10.1002/bdr2.2316, 03/01 2024.

[17] B. Zhou, W. Jin, Visualization of single cell RNA-seq data using t-SNE in R, Methods Mol. Biol. 2117 (2020) 159–167, https://doi.org/10.1007/978-1-0716-0301-7_8.

[18] T. Stuart, R. Satija, Integrative single-cell analysis, Nat. Rev. Genet. 20 (5) (May 2019) 257–272, https://doi.org/10.1038/s41576-019-0093-7.

[19] Z. Lu, Z. Wang, Z. Song, et al., Single-cell sequencing of brain tissues reveal the central nervous system's susceptibility to SARS-CoV-2 and the drug. Original Research, Front. Pharmacol. (2022), https://doi.org/10.3389/fphar.2022.971017. September-13 2022.

[20] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (5) (May 2012) 284–287, https://doi.org/10.1089/omi.2011.0118.

[21] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (12) (2010) 1572–1573, https://doi.org/10.1093/bioinformatics/btq170. Jun 15.

[22] P. Shannon, A. Markiel, O. Ozier, et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (Nov 2003) 2498–2504, https://doi.org/10.1101/gr.1239303.

[23] X. Huang, Z. Wang, M. Song, et al., CircIQGAP1 regulates RCAN1 and RCAN2 through the mechanism of ceRNA and promotes the growth of malignant glioma, Pharmacol. Res. 197 (2023), https://doi.org/10.1016/j.phrs.2023.106979.

[24] Z. Xu, J. Song, L. Cao, et al., Improving ovarian cancer treatment decision using a novel risk predictive tool, Aging 14 (8) (2022) 3464–3483, https://doi.org/10.18632/aging.204023. Apr 19.

[25] P. Geeleher, N.J. Cox, R.S. Huang, Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines, Genome Biol. 15 (3) (2014) R47, https://doi.org/10.1186/gb-2014-15-3-r47. Mar 3.

[26] P. Geeleher, N. Cox, R.S. Huang, pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels, PLoS One 9 (9) (2014), https://doi.org/10.1371/journal.pone.0107468, 107468.

[27] Y. Liu, M. Grimm, W.T. Dai, M.C. Hou, Z.X. Xiao, Y. Cao, CB-Dock: a web server for cavity detection-guided protein-ligand blind docking, Acta Pharmacol. Sin. 41 (1) (Jan 2020) 138–144, https://doi.org/10.1038/s41401-019-0228-6.

[28] Y. Cao, L. Li, Improved protein-ligand binding affinity prediction by using a curvature-dependent surface-area model, Bioinformatics 30 (12) (2014) 1674–1680, https://doi.org/10.1093/bioinformatics/btu104. Jun 15.

[29] Z. Wang, Z. Lu, Y. Chen, et al., Targeting the AKT-P53/CREB pathway with epicatechin for improved prognosis of traumatic brain injury, CNS Neurosci. Ther. 30 (2) (2024), https://doi.org/10.1111/cns.14364, 02/01 2024.

[30] L. Pinzi, G. Rastelli, Molecular docking: Shifting Paradigms in drug discovery, Int. J. Mol. Sci. (18) (2019) 20, https://doi.org/10.3390/ijms20184331. Sep. 4.

[31] J. Kim, R. Akbani, C.J. Creighton, et al., Invasive bladder cancer: Genomic insights and therapeutic Promise. Clinical cancer research, an official journal of the American Association for Cancer Research 21 (20) (2015) 4514–4524, https://doi.org/10.1158/1078-0432.Ccr-14-1215. Oct 15.

[32] C. Liu, T. Tate, E. Batourina, et al., Pparg promotes differentiation and regulates mitochondrial gene expression in bladder epithelial cells, Nat. Commun. 10 (1) (2019) 4589, https://doi.org/10.1038/s41467-019-12332-0. Oct 9.

[33] P. Khandelwal, S.N. Abraham, G. Apodaca, Cell biology and physiology of the uroepithelium, Am. J. Physiol. Ren. Physiol. 297 (6) (Dec 2009) F1477–F1501, https://doi.org/10.1152/ajprenal.00327.2009.

[34] S.P. Jost, Cell cycle of normal bladder urothelium in developing and adult mice, Virchows Arch. B Cell Pathol. Incl. Mol. Pathol. 57 (1) (1989) 27–36, https://doi.org/10.1007/bf02899062.

[35] M. Garg, Urothelial cancer stem cells and epithelial plasticity: current concepts and therapeutic implications in bladder cancer, Cancer Metastasis Rev. 34 (4) (Dec 2015) 691–701, https://doi.org/10.1007/s10555-015-9589-6.

[36] J. Hu, L. Wang, L. Li, Y. Wang, J. Bi, A novel focal adhesion-related risk model predicts prognosis of bladder cancer – a bioinformatic study based on TCGA and GEO database, BMC Cancer 22 (1) (2022) 1158, https://doi.org/10.1186/s12885-022-10264-5. Nov 10.

[37] D. Bianconi, M. Unseld, G.W. Prager, Integrins in the Spotlight of cancer, Int. J. Mol. Sci. (12) (2016) 17, https://doi.org/10.3390/ijms17122037. Dec 6.

[38] R. Ata, C.N. Antonescu, Integrins, Cell Metabolism, An intimate relationship impacting cancer, Int. J. Mol. Sci. 18 (1) (2017) 18, https://doi.org/10.3390/ijms18010189. Jan.

[39] J.S. Desgrosellier, D.A. Cheresh, Integrins in cancer: biological implications and therapeutic opportunities, Nat. Rev. Cancer 10 (1) (Jan 2010) 9–22, https://doi.org/10.1038/nrc2748.

[40] F. Xu, J. Zhang, G. Hu, L. Liu, W. Liang, Hypoxia and TGF-β1 induced PLOD2 expression improve the migration and invasion of cervical cancer cells by promoting epithelial-to-mesenchymal transition (EMT) and focal adhesion formation, Cancer Cell Int. 17 (2017) 54, https://doi.org/10.1186/s12935-017-0420-z.

[41] M.A. Eckert, M. Santiago-Medina, T.M. Lwin, J. Kim, S.A. Courtneidge, J. Yang, ADAM12 induction by Twist1 promotes tumor invasion and metastasis via regulation of invadopodia and focal adhesions, J. Cell Sci. 15 (2017) 130, https://doi.org/10.1242/jcs.198200. Jun2036-2048.

[42] P.J. Wipff, B. Hinz, Integrins and the activation of latent transforming growth factor beta1-an intimate relationship, Eur. J. Cell Biol. 87 (8–9) (Sep 2008) 601–615, https://doi.org/10.1016/j.ejcb.2008.01.012.

[43] A. Bianchi-Smiraglia, D. Kunnev, M. Limoge, A. Lee, M.C. Beckerle, A.V. Bakin, Integrin-β5 and zyxin mediate formation of ventral stress fibers in response to transforming growth factor β, Cell Cycle 12 (21) (2013) 3377–3389, https://doi.org/10.4161/cc.26388. Nov 1.

[44] A. Săndulescu, A.E. Stepan, C. Mărgăritescu, V. Enăchescu, G. Mitroi, C.E. Simionescu, The role of cell adhesion molecules in the progression of bladder urothelial carcinomas, Romanian journal of morphology and embryology = Revue roumaine de morphologie et embryologie 63 (1) (Jan-Mar 2022) 145–151, https://doi.org/10.47162/rjme.63.1.15.

[45] U. Cavallaro, B. Schaffhauser, G. Christofori, Cadherins and the tumour progression, is it all in a switch? Cancer letters 176 (2) (2002) 123–128, https://doi.org/10.1016/s0304-3835(01)00759-5. Feb 25.

[46] M. Takeichi, Morphogenetic roles of classic cadherins, Curr. Opin. Cell Biol. 7 (5) (Oct 1995) 619–627, https://doi.org/10.1016/0955-0674(95)80102-2.

[47] R.T. Bryan, C. Tselepis, Cadherin switching and bladder cancer, J. Urol. 184 (2) (Aug 2010) 423–431, https://doi.org/10.1016/j.juro.2010.04.016.

[48] R.T. Bryan, Bladder cancer and cancer stem cells: basic science and implications for therapy, TheScientificWorldJOURNAL 11 (2011) 1187–1194, https://doi.org/10.1100/tsw.2011.117. Jun 9.

[49] M.J. Wheelock, Y. Shintani, M. Maeda, Y. Fukumoto, K.R. Johnson, Cadherin switching, J. Cell Sci. 121 (Pt 6) (2008) 727–735, https://doi.org/10.1242/jcs.000455. Mar 15.

[50] F. Nollet, P. Kools, F. van Roy, Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members, J. Mol. Biol. 299 (3) (2000) 551–572, https://doi.org/10.1006/jmbi.2000.3777. Jun 9.

[51] Z. Li, S. Yin, L. Zhang, W. Liu, B. Chen, Prognostic value of reduced E-cadherin expression in breast cancer, a meta-analysis. Oncotarget 8 (10) (2017) 16445–16455, https://doi.org/10.18632/oncotarget.14860. Mar 7.

[52] X. Ren, J. Wang, X. Lin, X. Wang, E-cadherin expression and prognosis of head and neck squamous cell carcinoma: evidence from 19 published investigations, OncoTargets Ther. 9 (2016) 2447–2453, https://doi.org/10.2147/ott.S98577.

[53] X. Xing, Y.B. Tang, G. Yuan, et al., The prognostic value of E-cadherin in gastric cancer: a meta-analysis, Int. J. Cancer 132 (11) (2013) 2589–2596, https://doi.org/10.1002/ijc.27947. Jun 1.

[54] Y. Shimoyama, S. Hirohashi, S. Hirano, et al., Cadherin cell-adhesion molecules in human epithelial tissues and carcinomas, Cancer Res. 15 (1989) 49. Apr2128-33.

[55] H. Shiozaki, H. Tahara, H. Oka, et al., Expression of immunoreactive E-cadherin adhesion molecules in human cancers, Am. J. Pathol. 139 (1) (Jul 1991) 17–23.

[56] M. Takeichi, The cadherins: cell-cell adhesion molecules controlling animal morphogenesis, Development 102 (4) (Apr 1988) 639–655, https://doi.org/10.1242/dev.102.4.639.

[57] R.T. Bryan, P.A. Atherfold, Y. Yeo, et al., Cadherin switching dictates the biology of transitional cell carcinoma of the bladder: ex vivo and in vitro studies, J. Pathol. 215 (2) (Jun 2008) 184–194, https://doi.org/10.1002/path.2346.

[58] J. Paredes, J. Figueiredo, A. Albergaria, et al., Epithelial E- and P-cadherins: role and clinical significance in cancer, Biochim. Biophys. Acta 1826 (2) (Dec 2012) 297–311, https://doi.org/10.1016/j.bbcan.2012.05.002.

[59] F. van Roy, Beyond E-cadherin: roles of other cadherin superfamily members in cancer, Nat. Rev. Cancer 14 (2) (Feb 2014) 121–134, https://doi.org/10.1038/nrc3647.

[60] A. Usui, S.Y. Ko, N. Barengo, H. Naora, P-cadherin promotes ovarian cancer dissemination through tumor cell aggregation and tumor-peritoneum interactions, Mol. Cancer Res. : MCR 12 (4) (Apr 2014) 504–513, https://doi.org/10.1158/1541-7786.Mcr-13-0489.

[61] K.S. Chen, J.W.C. Lim, L.J. Richards, J. Bunt, The convergent roles of the nuclear factor I transcription factors in development and cancer, Cancer Lett. 410 (2017) 124–138, https://doi.org/10.1016/j.canlet.2017.09.015. Dec 1.

[62] A. Jolma, J. Yan, T. Whitington, et al., DNA-binding specificities of human transcription factors, Cell. Jan 17 152 (1–2) (2013) 327–339, https://doi.org/10.1016/j.cell.2012.12.009.

[63] F. Zhang, M. Liang, C. Zhao, Y. Fu, S. Yu, NFIC promotes the vitality and osteogenic differentiation of rat dental follicle cells, J. Mol. Histol. 50 (5) (Oct 2019) 471–482, https://doi.org/10.1007/s10735-019-09841-z.

[64] J. Eeckhoute, J.S. Carroll, T.R. Geistlinger, M.I. Torres-Arzayus, M. Brown, A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer, Gene Dev. 20 (18) (2006) 2513–2526, https://doi.org/10.1101/gad.1446006. Sep. 15.

[65] J. Nilsson, K. Helou, A. Kovács, et al., Nuclear Janus-activated kinase 2/nuclear factor 1-C2 suppresses tumorigenesis and epithelial-to-mesenchymal transition by repressing Forkhead box F1, Cancer Res. 1 (2010) (2020) 70, https://doi.org/10.1158/0008-5472.Can-09-1677. Mar.

[66] I. Cobo, S. Paliwal, C. Bodas, et al., NFIC regulates ribosomal biology and ER stress in pancreatic acinar cells and restrains PDAC initiation, Nat. Commun. 14 (1) (2023) 3761, https://doi.org/10.1038/s41467-023-39291-x. Jun 23.

[67] H. Zhang, Z. Luo, J. Tang, et al., Transcription factor NFIC functions as a tumor suppressor in lung squamous cell carcinoma progression by modulating lncRNA CASC2, Cell Cycle 21 (1) (Jan 2022) 63–73, https://doi.org/10.1080/15384101.2021.1995130.

[68] G. Xu, Y. Zhang, N. Li, et al., LBX2-AS1 up-regulated by NFIC boosts cell proliferation, migration and invasion in gastric cancer through targeting miR-491-5p/ZNF703, Cancer Cell Int. 20 (2020) 136, https://doi.org/10.1186/s12935-020-01207-w.

[69] M. Illiano, M. Conte, L. Sapio, et al., Forskolin Sensitizes human acute myeloid leukemia cells to H3K27me2/3 demethylases GSKJ4 inhibitor via protein kinase A, Front. Pharmacol. 9 (2018) 792, https://doi.org/10.3389/fphar.2018.00792.

[70] A. Salzillo, A. Ragone, A. Spina, S. Naviglio, L. Sapio, Forskolin affects proliferation, migration and Paclitaxel-mediated cytotoxicity in non-small-cell lung cancer cell lines via adenylyl cyclase/cAMP axis, Eur. J. Cell Biol. 102 (2) (Jun 2023) 151292, https://doi.org/10.1016/j.ejcb.2023.151292.

[71] D. Singh, P. Singh, A. Pradhan, R. Srivastava, S.K. Sahoo, Reprogramming cancer stem-like cells with Nanoforskolin enhances the efficacy of paclitaxel in targeting breast cancer, ACS Appl. Bio Mater. 4 (4) (2021) 3670–3685, https://doi.org/10.1021/acsabm.1c00141. Apr 19.

[72] A. Chaudhry, A.Y. Rudensky, Control of inflammation by integration of environmental cues by regulatory T cells, J. Clin. Invest. 123 (3) (Mar 2013) 939–944, https://doi.org/10.1172/jci57175.

[73] G. Darrasse-Jèze, A.S. Bergot, A. Durgeau, et al., Tumor emergence is sensed by self-specific CD44hi memory Tregs that create a dominant tolerogenic environment for tumors in mice, J. Clin. Invest. 119 (9) (Sep 2009) 2648–2662, https://doi.org/10.1172/jci36628.

[74] S.Z. Josefowicz, L.F. Lu, A.Y. Rudensky, Regulatory T cells: mechanisms of differentiation and function, Annu. Rev. Immunol. 30 (2012) 531–564, https://doi.org/10.1146/annurev.immunol.25.022106.141623.

[75] C. Ménétrier-Caux, T. Curiel, J. Faget, M. Manuel, C. Caux, W. Zou, Targeting regulatory T cells, Targeted Oncol. 7 (1) (Mar 2012) 15–28, https://doi.org/10.1007/s11523-012-0208-y.

[76] T. Yamaguchi, S. Sakaguchi, Regulatory T cells in immune surveillance and treatment of cancer, Semin. Cancer Biol. 16 (2) (Apr 2006) 115–123, https://doi.org/10.1016/j.semcancer.2005.11.005.

[77] Y.C. Lin, J. Mahalingam, J.M. Chiang, et al., Activated but not resting regulatory T cells accumulated in tumor microenvironment and correlated with tumor progression in patients with colorectal cancer, Int. J. Cancer 132 (6) (2013) 1341–1350, https://doi.org/10.1002/ijc.27784. Mar 15.

[78] O.M. Grauer, S. Nierkens, E. Bennink, et al., CD4+FoxP3+ regulatory T cells gradually accumulate in gliomas during tumor growth and efficiently suppress antiglioma immune responses in vivo, Int. J. Cancer 121 (1) (2007) 95–105, https://doi.org/10.1002/ijc.22607. Jul 1.

[79] J. Zhou, T. Ding, W. Pan, L.Y. Zhu, L. Li, L. Zheng, Increased intratumoral regulatory T cells are related to intratumoral macrophages and poor prognosis in hepatocellular carcinoma patients, Int. J. Cancer 125 (7) (2009) 1640–1648, https://doi.org/10.1002/ijc.24556. Oct 1.

[80] A. Loskog, C. Ninalga, G. Paul-Wetterberg, M. de la Torre, P.U. Malmström, T.H. Tötterman, Human bladder carcinoma is dominated by T-regulatory cells and Th1 inhibitory cytokines, J. Urol. 177 (1) (Jan 2007) 353–358, https://doi.org/10.1016/j.juro.2006.08.078.
[81] L.J. Chi, H.T. Lu, G.L. Li, et al., Involvement of T helper type 17 and regulatory T cell activity in tumour immunology of bladder carcinoma, Clin. Exp. Immunol. 161 (3) (Sep 2010) 480–489, https://doi.org/10.1111/j.1365-2249.2010.04215.x.
[82] B. Shang, Y. Liu, S.J. Jiang, Y. Liu, Prognostic value of tumor-infiltrating FoxP3+ regulatory T cells in cancers: a systematic review and meta-analysis, Sci. Rep. 5 (2015) 15179, https://doi.org/10.1038/srep15179. Oct 14.
[83] M.E. Winerdal, D. Krantz, C.A. Hartana, et al., Urinary bladder cancer Tregs suppress MMP2 and potentially regulate invasiveness, Cancer Immunol. Res. 6 (5) (May 2018) 528–538, https://doi.org/10.1158/2326-6066.Cir-17-0466.