Am J Clin Exp Immunol 2024;13(3):88-104
www.ajcei.us /ISSN:2164-7712/AJCEI0156764

Original Article

Identification of macrophage-related genes in bladder cancer patients using single-cell sequencing and construction of a prognostic model

Weizhuo Wang[1*], Junheng Shen[2*], Dalong Song[3*], Kai Fu[4], Xu Fu[5]

[1]Department of Urology, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou 215002, Jiangsu, PR China; [2]Department of Reproduction, Suzhou TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Suzhou 215009, Jiangsu, PR China; [3]Department of Urology, Guizhou Provincial People's Hospital, Guiyang 550002, Guizhou, PR China; [4]Center of Reproduction, Second Affiliated Hospital of Soochow University, Suzhou 215002, Jiangsu, PR China; [5]Center of Reproduction, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou 215002, Jiangsu, PR China. [*]Equal contributors.

Received March 24, 2024; Accepted June 19, 2024; Epub June 25, 2024; Published June 30, 2024

Abstract: Single-cell sequencing is an emerging technology that can effectively identify cell types in tumors. In the tumor microenvironment of bladder cancer, macrophages play a crucial role in invasion and immune escape. This study aimed to assess the expression of macrophage-related genes (MRGs) in the tumor microenvironment of bladder cancer patients and construct a prognostic model based on MRGs using bioinformatics methods. Methods: Single-cell sequencing data from bladder cancer patients was downloaded from the GEO. After quality control and cell type identification, macrophages in the samples were extracted for re-clustering. Feature genes were then identified, and MRGs were assessed. Genetic data from TCGA database bladder cancer patients was also downloaded and organized. The intersection of MRGs and the TCGA gene set was determined. Clinical information was connected with this intersection, and the data was divided into training and validation sets. The training set was used for model construction and the validation set for model verification. A prognostic model based on MRGs was built using the LASSO algorithm and Cox regression. Patients were divided into high-risk and low-risk groups based on their prognostic features, and survival information in the training and validation sets was observed. The predictive ability of the model was assessed using a ROC curve, followed by a calibration plot to predict 1-, 3-, and 5-year survival rates. Results: Four cell types were identified, and after extracting macrophages, three cell subgroups were clustered, resulting in 1,078 feature genes. The top 100 feature genes from each macrophage subgroup were extracted and intersected with TCGA expressed genes to construct the model. A risk prediction model composed of CD74, METRN, PTPRR, and CDC42EP5 was obtained. The survival and ROC curves showed that this model had good predictive ability. A calibration curve also demonstrated good prognostic ability for patients. Conclusion: This study, based on single-cell data, TCGA data, and clinical information, constructed an MRG-based prognostic model for bladder cancer using multi-omics methods. This model has good accuracy and reliability in predicting the survival and prognosis of patients with bladder cancer, providing a reference for understanding the interaction between MRGs and bladder cancer.

Keywords: Single-cell sequencing, macrophages, bladder cancer, prognostic model

Introduction

Single-cell sequencing is a powerful technique to measure gene expression, DNA methylation, chromatin accessibility, and various other molecular features at the level of individual cells. Through single-cell RNA sequencing (scRNA-seq), various cell subtypes present in bladder cancer can be identified, which includes both cancer cells and non-tumor cells within the tumor microenvironment. This technique could potentially aid in understanding the heterogeneity of bladder cancer, as well as the roles of different cell subtypes in disease progression. Concurrently, with single-cell DNA sequencing (scDNA-seq), the evolutionary history of bladder

https://doi.org/10.62347/VLDZ7581

cancer can be depicted, including the accumulation of mutations, variations in gene copy numbers, and the expansion of clones. Through scRNA-seq, researchers can gain deep insights into the cellular components and interactions within the tumor microenvironment, such as immune cells, fibroblasts, and vascular cells, etc. [1]. Furthermore, single-cell sequencing can help researchers understand the mechanisms of resistance to chemotherapy or immunotherapy in bladder cancer, and how such resistance evolves within cell populations [2, 3].

Bladder cancer is one of the top ten most common malignant tumors worldwide. The incidence is significantly higher in men, and disease risk increases with age, with the majority of newly diagnosed cases being >65 years old. Smoking, occupational exposure (especially to aromatic amines and other chemicals), chronic cystitis, certain hereditary diseases, and the use of certain anticancer drugs or radiation therapy are all considered major risk factors for bladder cancer [4]. Regarding the treatment options for bladder cancer, these mainly depend on the stage and grade of the tumor, as well as the overall health status of the patient. Common treatments include surgery, chemotherapy, radiotherapy, and immunotherapy. Surgery is the primary treatment, including transurethral resection of the bladder tumor (TURBT) and radical cystectomy. Chemotherapy can be used before or after surgery, or as the main treatment for metastatic bladder cancer. Radiotherapy is usually used in patients who cannot undergo surgery, or as a supplementary treatment to surgery [5]. More recently, immunotherapy, particularly immune checkpoint inhibitors, has become an important part of the treatment for bladder cancer; it has shown significant value, especially in the treatment of metastatic bladder cancer [6]. Immunotherapy works by activating the patient's immune system, enabling it to more effectively recognize and attack cancer cells. Notably, although existing treatments can improve survival and quality of life in some patients, several challenges remain in the treatment of bladder cancer, including disease recurrence, treatment resistance, and side effects. Therefore, basic research and clinical trials are currently being conducted to identify more effective and safer treatments. This includes the development of new drugs and treatment strategies, as well as the use of molecular diagnostics and precision medicine technologies to treat bladder cancer in a personalized manner [6].

Macrophages are an important part of the immune system that are capable of engulfing and eliminating invading pathogens and damaged cells. Macrophages also participate in processes such as the inflammatory response and wound healing. In the context of cancer, such as the occurrence and development of bladder cancer, the role of macrophages becomes much more complex [7]. On the one hand, they can eliminate normal cells that have transformed into cancer cells, thereby preventing the development of cancer. However, some macrophages can be "hijacked" by cancer cells and reprogrammed into "tumor-associated macrophages" (TAMs), which promote tumor growth and metastasis. Studies have shown that a high density of TAMs is often associated with disease worsening and a poorer prognosis [8]. This is because TAMs can promote tumor growth and metastasis in multiple ways, including promoting angiogenesis, inhibiting T cell responses, and altering the tumor microenvironment through the secretion of various growth factors and cytokines [9].

Given that single-cell sequencing can be used to investigate diseases at the level of individual cells, and that macrophages play a significant role in the onset and progression of bladder cancer, our study aimed to analyze the cellular composition of single-cell tumor samples from patients with bladder cancer. We will identify cell types and extract the specific genes of different macrophage subgroups. By combining this with clinical data from The Cancer Genome Atlas (TCGA), we aimed to construct a prognostic model for bladder cancer patients, which could provide new targets and strategies for the treatment of bladder cancer.

Materials and methods

Data collection

Single-cell data (GSE135337) was downloaded from the Gene Expression Omnibus (GEO) database (GSE135337) [10]. Bulk sequencing data and corresponding clinical data for patients with bladder cancer were downloaded from TCGA. Single-cell analysis was conducted using R language. Prior to analysis, the single-cell data were quality control-filtered and selected, followed by cell type identification.

Selection of highly variable genes

The Seurat package in R language was used to identify highly variable genes in the single-cell data, and the ElbowPlot function was used to determine the number of principal components for subsequent dimension reduction in the PCA.

Cell type identification and extraction of characteristic genes

After dimension reduction of the single-cell data from patients with bladder cancer, cell clusters were divided based on PCA numbers. The SINGER package in R language was used to identify cell types in the cell clusters. After filtering out macrophages, new cell clusters were further divided. The FindAllMarkers function was then used to screen characteristic macrophage-related genes (MRGs) in the macrophage cluster. Genes with a *P*-value <0.05 considered as subgroup-specific genes.

Enrichment analysis of MRGs

For the obtained MRGs, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) REST-style Application Programming Interface (API) (https://www.kegg.jp/kegg/rest/keggapi.html) to obtain the latest gene annotations for KEGG Pathways as the background set and mapped the genes to the background set. The R package cluster Profiler was used for enrichment analysis to obtain the results of gene set enrichment. The minimum gene set was set to 5, the maximum gene set to 5,000, and P<0.05 and a false discovery rate (FDR) <0.25 were considered significant. Gene Ontology (GO) was used to further explore potential molecular functions and cellular components involved associated with the MRGs. The screening conditions for GO analysis were P<0.01 and q<0.01.

Construction of the risk prediction model

Initially, a univariate Cox regression analysis was employed to assess the prognostic value of the MRGs, thereby acquiring as set of prognosis-related macrophage genes. Subsequently, the Least Absolute Shrinkage and Selection Operator (LASSO) regression was utilized to select predictive variables and to avoid overfitting. A multivariate Cox regression analysis was then performed to ascertain the final candidates involved in the risk model. Based on the final set of MRGs, a risk signature was constructed to predict the prognosis of patients with bladder cancer. The method for the risk score calculation is as follows:

riskscore = Σcoef gene i × expression level of gene i.

The risk value was obtained by weighting the regression coefficients and expression levels of the MRGs. According to the median value of the risk score, patients with bladder cancer in TCGA were divided into high-risk and low-risk groups.

Evaluation of the risk prediction model and nomogram

The correlation between the risk score and clinical characteristics was analyzed by comparing sex, age, and staging by the American Joint Committee on Cancer (AJCC) with the risk score. We also analyzed the relationship between the risk score and patient gene expression profiles across different subsets. Subsequently, we compared whether there was a significant difference in survival time and status between high-risk and low-risk patients with the same clinical pathological characteristics using R software to plot Kaplan-Meier survival curves. When then generated the receiver operating characteristic (ROC) curve to observe the diagnostic value of the model. Simultaneously, to further verify the prognostic value of the model, we divided patients into groups according to the stage (I-II, III-IV), and used the R package to plot Kaplan-Meier survival curves. Furthermore, we downloaded the patients' progression-free survival time to observe the accuracy of the model's prediction in the high and low-risk groups. Based on the evaluation of the ROC curve, we incorporated clinical data with a significant impact on prognosis, constructing a nomogram together with the risk score to predict the 1-, 3-, and 5-year survival rates. A calibration curve was also constructed to evaluate the predictive power of the nomogram.

Results

Data collection

Single-cell sequencing data from seven patients with bladder cancer were analyzed using the Seurat package in R. The criteria for

cell inclusion were as follows: the number of genes obtained from sequencing was >300, the percentage of mitochondrial genes was <10%, and the percentage of erythrocyte genes was <5%. Subsequently, a total of 36,787 patient cells were included (**Figure 1**). In addition, the expression data of patients with bladder cancer in TCGA were collated, totaling 392 individuals. For specific information, please refer to **Table 1**.

Identification of highly variable genes and PCA dimension reduction

We used the NormalizeData and ScaleData functions to standardize and centralize cell data. Subsequently, we integrated the seven single-cell datasets using an anchor-based method, selecting 2,000 genes for integration. Following integration, we used the FindVariableFeatures function to identify highly variable genes in the single-cell data, setting the number of highly variable genes at 3,000. We then applied PCA to perform dimension reduction on these highly variable genes and generated an ElbowPlot to determine the number of principal components to retain in the PCA dimension reduction process. As the number of principal components increased, the proportion of variance explained gradually decreased (**Figure 2A**). When the variance explanation ratio shows an obvious inflection point on the graph, it can be considered an appropriate number of principal components. This point is called the "elbow point". We selected PCA = 20 for subsequent analyses. We next used the FindNeighbors and FindClusters functions to classify the cells, and then displayed them using t-SNE dimension reduction maps grouped by source after integration (**Figure 2B**) and by cell clusters after integration (**Figure 2C**). After identifying the cell clusters, we used the HumanPrimaryCellAtlasData dataset in the SingleR package of R language [11] to identify the cell types. We finally identified four cell types (**Figure 2D**), namely epithelial cells, endothelial cells, macrophages, and tissue stem cells.

Macrophage subgroup and characteristic gene extraction

Because the histological origin of bladder cancer is comprised almost entirely of epithelial cells, we separately extracted macrophages, totaling 561 cells. We displayed the cell clusters extracted on a t-SNE map (**Figure 3**). Subsequently, we further classified the macrophage subgroups and used the FindMarkers function to extract their characteristic genes, resulting in a total of 1,078 characteristic genes. The names and expression conditions of the top 10 genes with the log2FC differential expression are shown in **Table 2**.

GO and KEGG analysis of MRGs

We used R language for GO enrichment analysis and KEGG pathway analysis of the MRGs (**Figure 4**), focusing on biological process (BP), cellular component (CC), and molecular function (MF) subsets. The BP involved chromosome segregation, positive regulation of the cell cycle process, sister chromatid segregation, etc. The CC mainly involved chromosomal region, condensed chromosome, chromosome, centromeric region, etc. The MF mainly involved ubiquitin-like protein ligase binding, MHC class II protein complex binding, structural constituent of cytoskeleton, etc. In the KEGG analysis, we observed that in addition to "pathways in cancer", the following pathways were enriched: "cell cycle", "Th1 and Th2 cell differentiation", "Th17 cell differentiation", "cell adhesion molecules (CAMs)", "p53 signaling pathway", and "HIF-1 signaling pathway". These results suggest that MRGs are mainly associated with cell division, cell cycle genes, and ubiquitination, indicating that the macrophages in bladder cancer samples are active, that protein ubiquitination is at an active level and has not developed to the terminal stage, and that the cells are actively reshaping their structure. Furthermore, the KEGG analysis showed that these cell activities are closely related to the occurrence and development of tumors. This is because the "pathways in cancer", "cell adhesion molecules (CAMs)", "p53 signaling pathway", and "HIF-1 signaling pathway" have all been found to play a role in the occurrence and development of various tumors [12]. In addition, the HIF-1 signaling pathway is often related to a lack of oxygen in the cellular environment, which corresponds to the high metabolism and hypoxia of the tumor microenvironment.

Prognostic model construction based on MRGs and TCGA expression data

The age, gender, pT stage, pN stage, pM stage, AJCC staging, survival status, and survival time

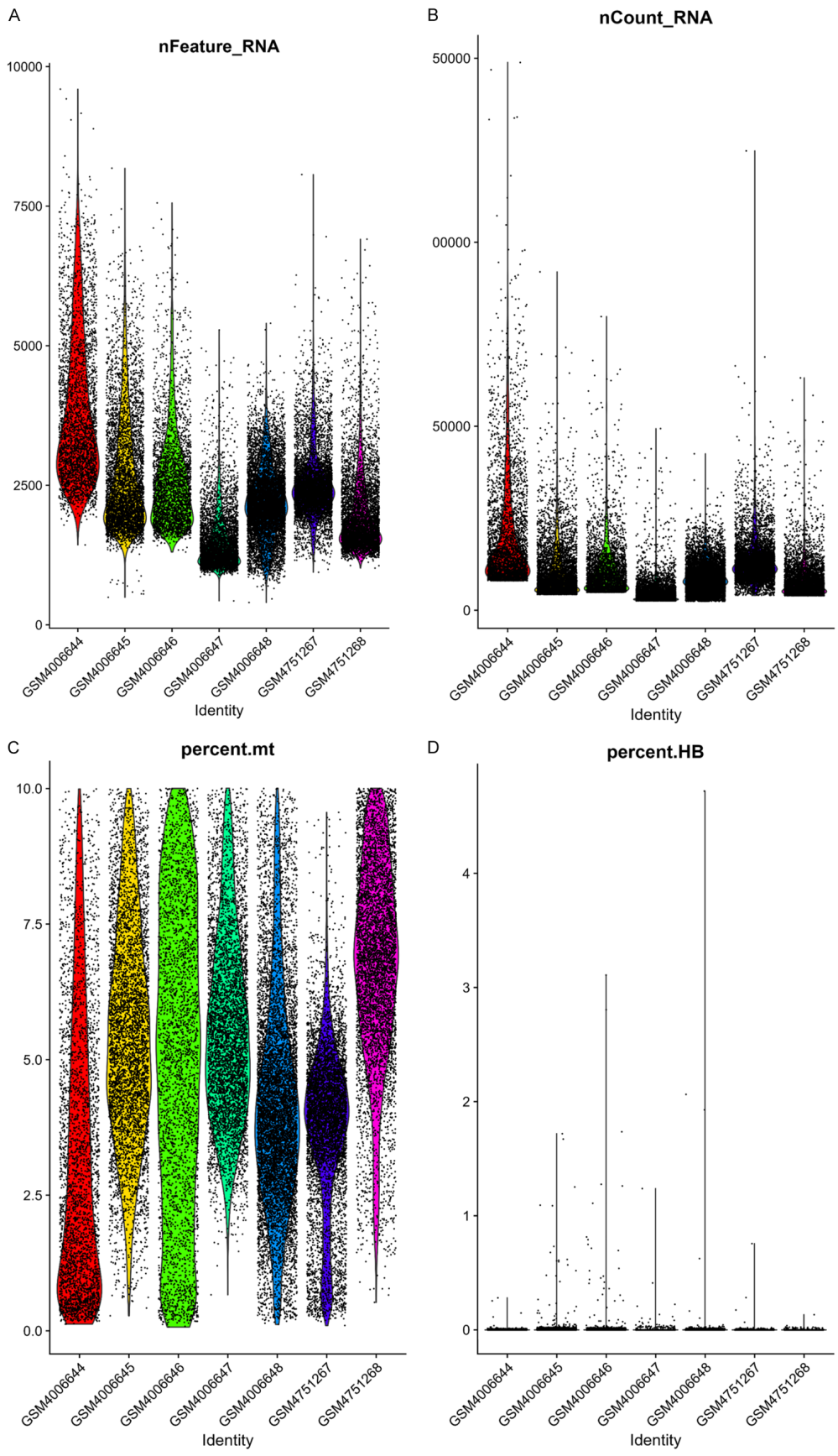
A
nFeature_RNA
10000
7500
5000
2500
0
GSM4006644
GSM4006645
GSM4006646
GSM4006647
GSM4006648
GSM4751267
GSM4751268
Identity
B
nCount_RNA
50000
00000
50000
0
GSM4006644
GSM4006645
GSM4006646
GSM4006647
GSM4006648
GSM4751267
GSM4751268
Identity
C
percent.mt
10.0
7.5
5.0
2.5
0.0
GSM4006644
GSM4006645
GSM4006646
GSM4006647
GSM4006648
GSM4751267
GSM4751268
Identity
D
percent.HB
4
3
2
1
0
GSM4006644
GSM4006645
GSM4006646
GSM4006647
GSM4006648
GSM4751267
GSM4751268
Identity

Figure 1. A. Number of genes detected in the seven samples. B. Number of sequencing counts obtained in the seven samples. C. Mitochondrial gene ratio in the cells of the seven samples. D. Red blood cell gene ratio in the seven samples.

Table 1. Patient clinical information and grouping of clinical information

Covariates	Type	Total	Test	Train	*P*-value
Age	≤65	157 (40.05%)	83 (42.35%)	74 (37.76%)	0.4096
	>65	235 (59.95%)	113 (57.65%)	122 (62.24%)	
Gender	Female	103 (26.28%)	45 (22.96%)	58 (29.59%)	0.1685
	Male	289 (73.72%)	151 (77.04%)	138 (70.41%)	
M	M0	188 (47.96%)	90 (45.92%)	98 (50%)	0.1962
	M1	11 (2.81%)	8 (4.08%)	3 (1.53%)	
	Unknown	193 (49.23%)	98 (50%)	95 (48.47%)	
N	N0	227 (57.91%)	111 (56.63%)	116 (59.18%)	0.8279
	N1	46 (11.73%)	26 (13.27%)	20 (10.2%)	
	N2	74 (18.88%)	37 (18.88%)	37 (18.88%)	
	N3	6 (1.53%)	3 (1.53%)	3 (1.53%)	
	Unknown	39 (9.95%)	19 (9.69%)	20 (10.2%)	
Stage	Stage II	124 (31.63%)	57 (29.08%)	67 (34.18%)	0.5167
	Stage III	136 (34.69%)	69 (35.2%)	67 (34.18%)	
	Stage IV	132 (33.67%)	70 (35.71%)	62 (31.63%)	
T	T2	115 (29.34%)	54 (27.55%)	61 (31.12%)	0.3744
	T3	190 (48.47%)	93 (47.45%)	97 (49.49%)	
	T4	56 (14.29%)	34 (17.35%)	22 (11.22%)	
	Unknown	31 (7.91%)	15 (7.65%)	16 (8.16%)	

of TCGA patients were all included in the clinical data (**Table 1**). Eventually, 392 patients with both clinical survival information and sample sequencing information were recruited and randomly divided into the training set (n = 186) and validation set (n = 186) using the "caret" package in R. Subsequently, we intersected the identified MRGs genes with the expression file genes in TCGA, resulting in 207 intersecting genes. We connected these 207 genes with the clinical data, and then took the log2+1 of their Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values for normalization. We then associated them with the patient's clinical prognosis outcomes, performed univariate and multivariate Cox regression, Lasso regression is used for variable compression (**Figure 5A**, **5B**), and built a patient prognostic model composed of four genes: *CD74, METRN, PTPRR,* and *CDC42EP5*. Their corresponding model coefficients, hazard ratio, and *P* values are shown in **Figure 5C** and **5D**. Furthermore, according to the median risk score, they were divided into high-risk and low-risk groups. A heatmap of the expression differences of MRGs in the high-risk and low-risk groups was drawn (**Figure 6A**, **6B**). The results showed that, whether in the training set or the validation set, the expression of the four genes in the high-risk group was significantly different from that in the low-risk group. Subsequently, scatter plots and risk curves were used to show the survival status and risk scores of each patient with bladder cancer (**Figure 6E**, **6F**); the overall survival rates of the high-risk and low-risk patients in the validation and test set were significantly different (P<0.05), with the mortality rate and hazard ratio of the high-risk group being higher than those of the low-risk group (**Figure 6G**, **6H**).

Validation of the risk model and nomogram construction

To assess whether the model score, age, gender, and pathological grade are independent prognostic factors for patients with bladder cancer, we performed univariate and multivariate Cox regression analyses and generated forest plots. The results showed that AJCC stage

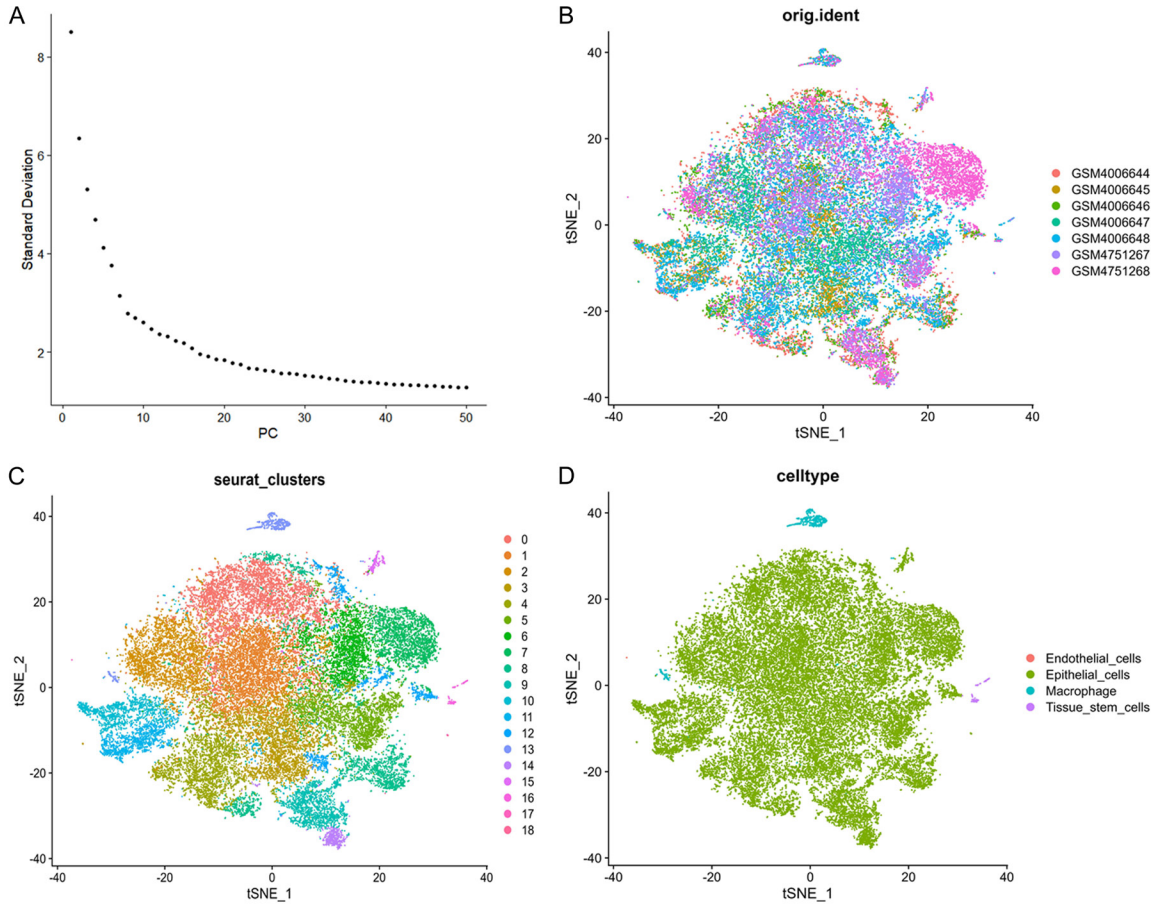

Figure 2. A. Elbow plot of variance explained vectors. B. t-SNE visualization of single cells colored by the origin of the samples. C. t-SNE visualization of single cells colored by cell clusters. D. t-SNE visualization of single cells colored by cell type identification.

and risk score were independent prognostic factors for patients with bladder cancer (P<0.05) (**Figure 7A**, **7B**). This indicates that our prognostic model has diagnostic value for patient prognosis independent of other clinical features. To evaluate the accuracy of the risk score and clinical features in predicting the prognosis of patients with bladder cancer, we plotted ROC curves and time-dependent ROC curves for clinical information in the training set, validation set, and overall dataset (**Figure 6C**, **6D**). The training set exhibited an area under the curve (AUC) of 0.752 for 1-year survival, 0.677 for 3-year survival, and 0.606 for 5-year survival. The test set showed AUC values of 0.729, 0.666, and 0.717 for 1-, 3-, and 5-year survival, respectively. Additionally, we plotted the ROC curve for predicting the 1-year survival probability using the risk score and clinical features in the training and validation sets together. The AUC of the model was the largest, indicating that our model has a greater diagnostic value in assessing prognosis compared to other clinical features (**Figure 8A-D**). In addition, gender, age, and grade were found to be unrelated to the risk score (P<0.05). Furthermore, we conducted progression-free survival analysis (**Figure 8E**, **8F**). The results showed that patients with a higher risk score had a significantly longer progression-free survival than those with a lower risk score, indicating that the genes in our model may impact progression-free survival. To further investigate the prognostic value of the model, we grouped the patients by stage (stage I-II and III-IV) and plotted Kaplan-Meier survival curves (**Figure 9A**, **9B**). The survival curves demonstrated that patients with the same AJCC stage had shorter survival times and a poorer prognosis as the risk score increased (P<0.05). This indicates that our model has some diagnostic value for predicting patient prognosis. Additionally, we

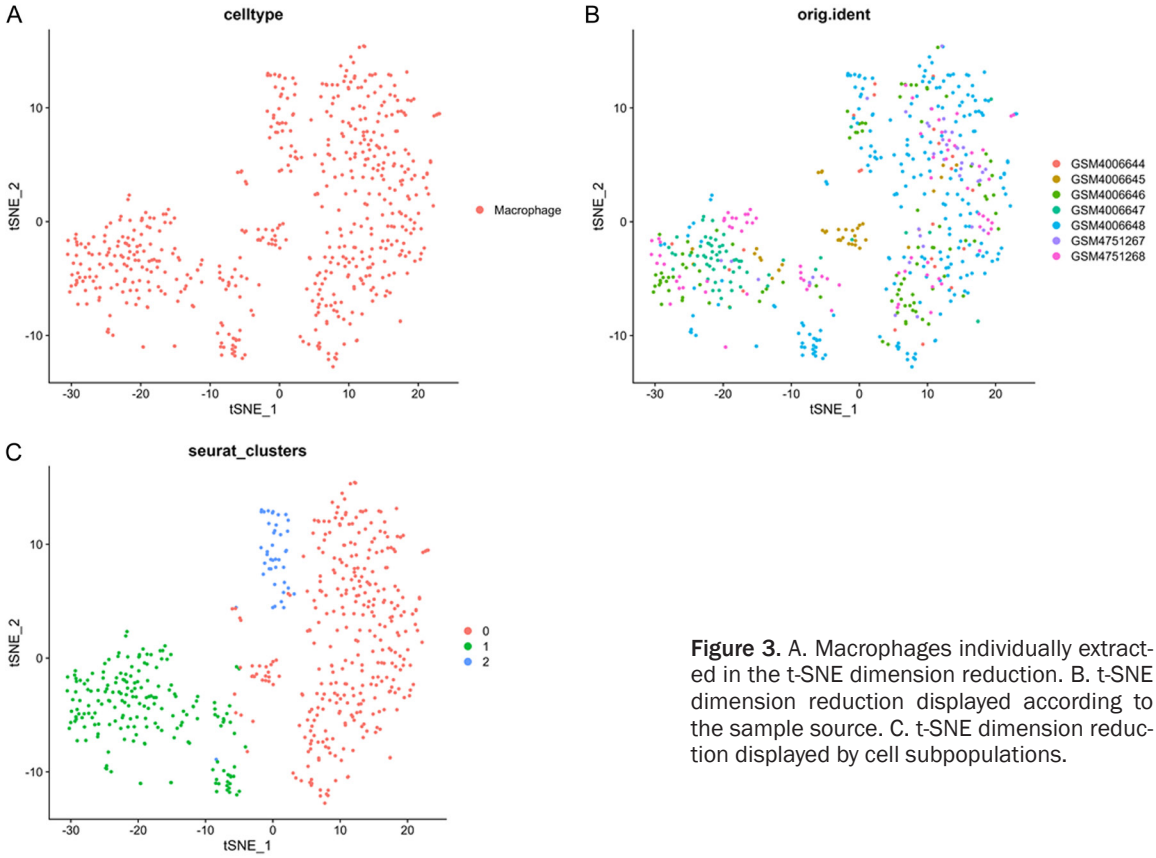

Figure 3. A. Macrophages individually extracted in the t-SNE dimension reduction. B. t-SNE dimension reduction displayed according to the sample source. C. t-SNE dimension reduction displayed by cell subpopulations.

found that age, stage, and risk score all had an impact on patient prognosis. Therefore, we constructed nomograms based on these three indicators to assign scores to patients and predict their 1-, 3-, and 5-year survival probabilities. The calibration curves of the nomograms showed that the predicted values were close to the actual survival probabilities, indicating good diagnostic value for the nomograms and the overall model (**Figure 9C**, **9D**).

Discussion

Bladder cancer is a common malignant tumor, the occurrence and development of which are closely related to the tumor microenvironment [13]. Macrophages, as one of the main immune cells in bladder cancer, play an important role in the tumor microenvironment of bladder cancer. Macrophages exhibit diversity and plasticity, differentiating into subgroups with different functions based on environmental signals. Typically, macrophages are divided into two types: M1 type (pro-inflammatory) and M2 type (anti-inflammatory) [14]. In the bladder cancer tumor microenvironment, macrophages tend to differentiate into the M2 type, promoting tumor growth, metastasis, and angiogenesis by secreting growth factors, cytokines, and chemokines [15]. Tumors require new blood vessels to supply nutrients and oxygen to maintain growth. TAMs promote the formation of new blood vessels by secreting various molecules, such as vascular endothelial growth factor (VEGF) [16]. At the same time, they can suppress the immune system, preventing it from attacking cancer cells by secreting immunosuppressive molecules such as transforming growth factor (TGF)-β and interleukin (IL)-10 [17, 18]. Furthermore, TAMs can help cancer cells penetrate the basement membrane and interstitial cells and invade surrounding tissues and blood vessels, leading to tumor metastasis, through the secretion of various proteinases and cytokines, such as matrix metalloproteinases (MMPs) and epidermal growth factor (EGF) [19, 20].

Single-cell sequencing is a high-throughput sequencing technology that can measure the genome, transcriptome, or epigenome at the single-cell level. Single-cell sequencing can

Table 2. The top 10 specific genes and their related expression conditions in the three macrophage subgroups

Gene	p_val	avg_log2FC	pct.1	pct.2	p_val_adj	Cluster
TYROBP	2.96E-62	1.697220889	1	0.966	5.92E-59	0
CD74	1.52E-58	1.685945728	1	0.983	3.04E-55	0
HLA-DRA	3.14E-58	1.828062408	0.997	0.991	6.28E-55	0
HLA-DPB1	3.30E-57	1.795460403	0.988	0.949	6.60E-54	0
CCL4L2	2.84E-26	1.748558902	0.982	0.739	5.69E-23	0
APOE	6.12E-26	2.225427906	0.969	0.868	1.22E-22	0
APOC1	7.37E-24	1.867185273	0.997	0.936	1.47E-20	0
CCL3	6.74E-21	1.86520665	0.92	0.803	1.35E-17	0
CCL4	2.85E-19	1.743460094	0.887	0.705	5.71E-16	0
CXCL8	1.06E-08	1.750896013	0.841	0.791	2.12E-05	0
KRT19	1.64E-74	2.403691076	1	0.938	3.28E-71	1
CD24	1.75E-71	2.683857451	1	0.995	3.51E-68	1
KRT18	6.91E-69	2.92271572	0.995	0.668	1.38E-65	1
CAMK2N1	1.79E-68	2.130147691	0.99	0.751	3.59E-65	1
SNCG	2.36E-67	2.640455738	1	1	4.72E-64	1
ADIRF	8.99E-66	2.473192258	1	1	1.80E-62	1
DHRS2	4.10E-64	2.480354274	0.984	0.635	8.19E-61	1
HPGD	3.99E-61	2.242237934	0.995	0.916	7.98E-58	1
CLDN4	6.09E-60	2.452498569	0.995	0.776	1.22E-56	1
PSCA	1.24E-57	2.105113368	0.979	0.835	2.48E-54	1
H2AFZ	2.52E-23	1.870332237	1	0.992	5.04E-20	2
STMN1	3.42E-23	2.264808355	1	0.888	6.84E-20	2
TUBB	6.71E-23	2.250717786	1	0.992	1.34E-19	2
HMGB2	2.66E-21	2.334086929	0.953	0.643	5.33E-18	2
CKS1B	1.89E-20	1.869258383	1	0.865	3.77E-17	2
TUBA1B	1.69E-19	1.938305196	1	0.969	3.38E-16	2
KIAA0101	3.02E-19	2.098884223	1	0.875	6.04E-16	2
HMGN2	2.57E-18	1.696403128	1	0.988	5.13E-15	2
UBE2C	1.10E-15	1.941262234	0.977	0.876	2.19E-12	2
PTTG1	1.44E-09	1.645495691	0.907	0.795	2.87E-06	2

resolve the heterogeneity between different cell types within cancer tissues, providing an in-depth understanding of the tumor microenvironment and intercellular interactions. In addition, single-cell sequencing can detect cell types in the tumor microenvironment, such as T cells, B cells, macrophages, NK cells, etc. [21]. This has significant implications for cancer treatment. Through single-cell sequencing, personalized genomic information can be provided for patients, which can help to identify drug targets and achieve precision treatment [22]. Although single-cell sequencing has many advantages in cancer research, it also has certain limitations, such as high cost and complex data processing. Therefore, combining multiple technical means for cancer research can help us to understand the biological characteristics of tumors more comprehensively. This article uses a multi-omics analysis method and uses bladder cancer single-cell sequencing data. After quality control of the data, screening of high-variance genes, cell clustering, and cell type identification, the results show that MRGs are mainly associated with cell division, the cell cycle, and ubiquitination, indicating that macrophages in bladder cancer samples are active and are actively reshaping their structure. Furthermore, KEGG analysis showed that pathways involved in tumor invasion and development include “pathways in cancer”, “p53 signaling pathway”, “Th1 and Th2 cell dif-

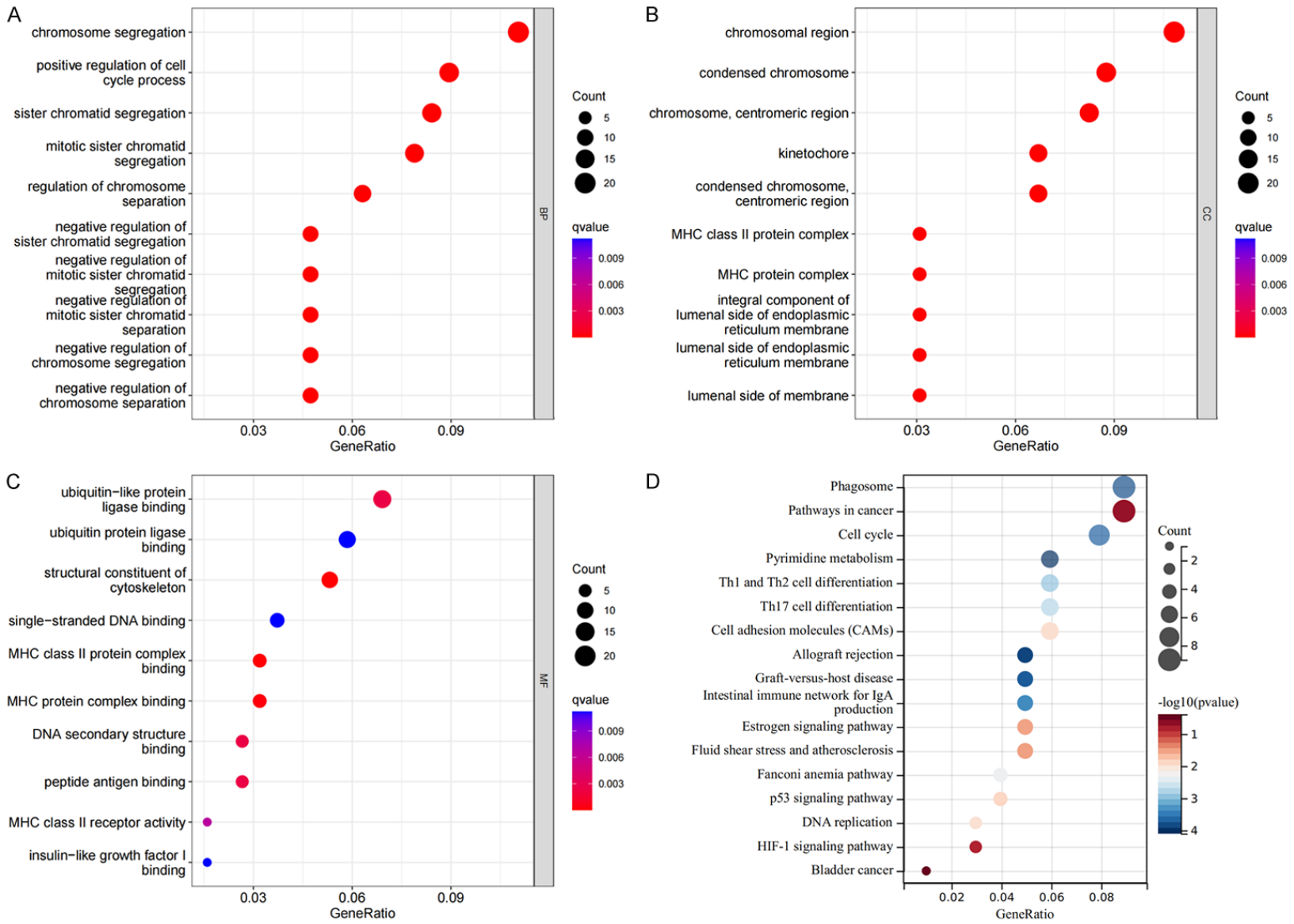

Figure 4. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of macrophage-related genes. (A) GO analysis of biological processes (BP), (B) GO analysis of cellular components (CC), (C) GO analysis of molecular functions (MF), and (D) KEGG pathway analysis.

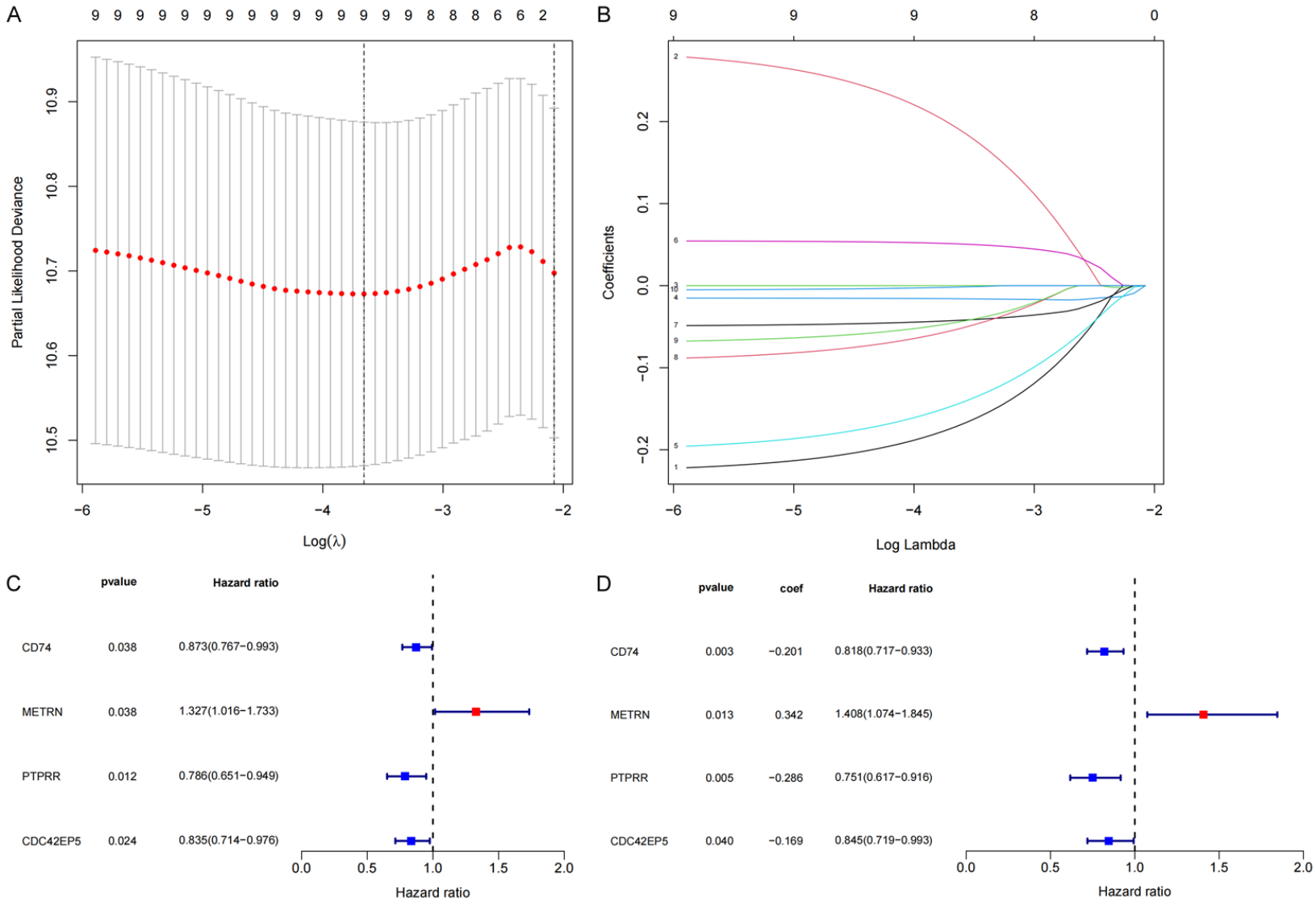

Figure 5. A, B. Number of genes selected by LASSO regression. C. Univariate Cox regression. D. Multivariate Cox regression and model coefficients.

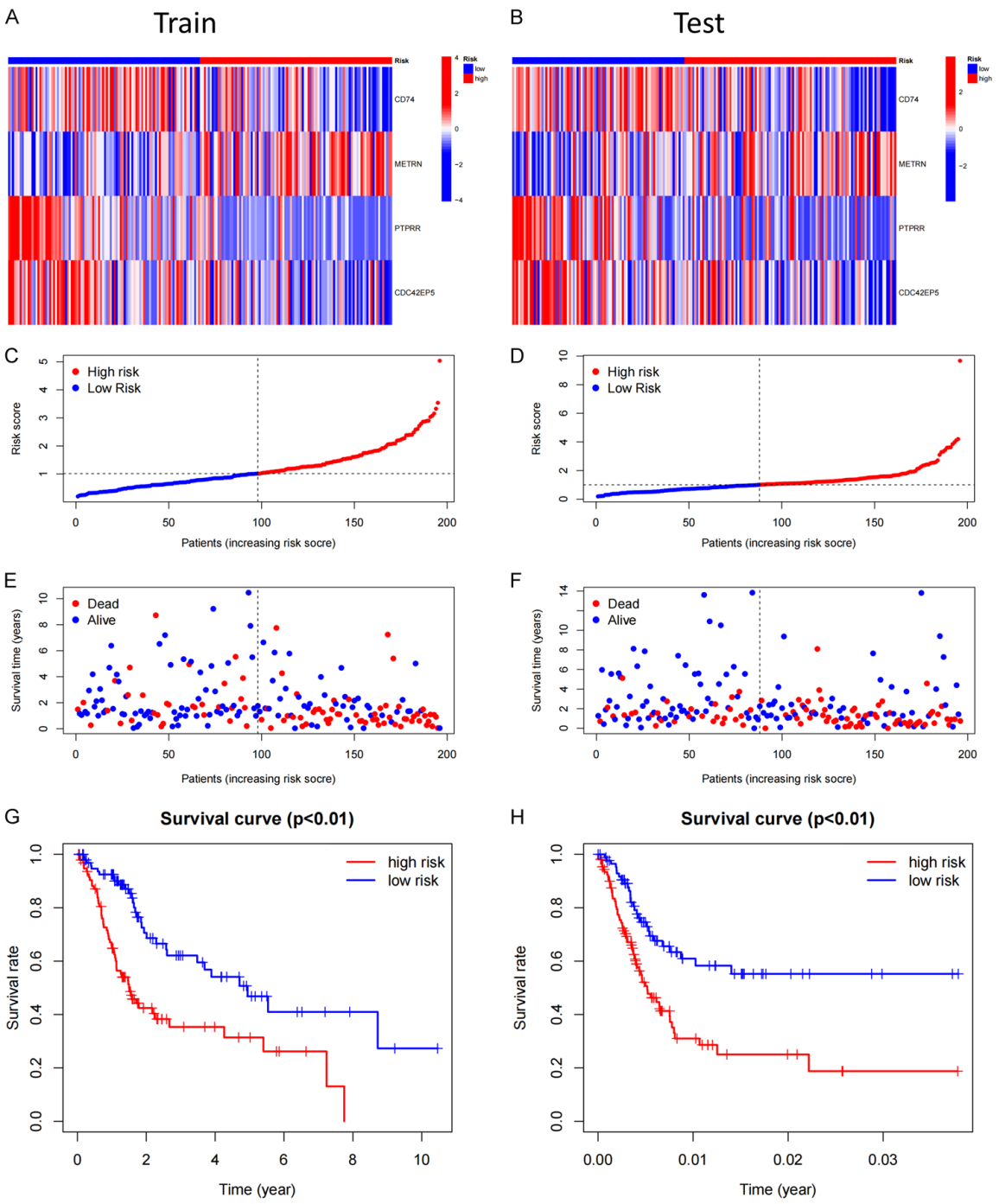

Figure 6. A, C, E, G. Heatmaps of gene expression in the training set with respect to risk scores, changes in patient risk scores, scatter plot of patient risk scores and survival status, and Kaplan-Meier survival curves for high and low-risk groups. B, D, F, H. Heatmaps of gene expression in the validation set with respect to risk scores, changes in patient risk scores, scatter plot of patient risk scores and survival status, and Kaplan-Meier survival curves for high and low-risk groups.

ferentiation", and "HIF-1 signaling pathway". The analysis results indicate that macrophages are active and are involved in cell communication; thus, they actively participate in the occurrence and development of bladder cancer. Subsequently, we intersected these genes with genes contained in the second-generation sequencing in TCGA, and after removing the

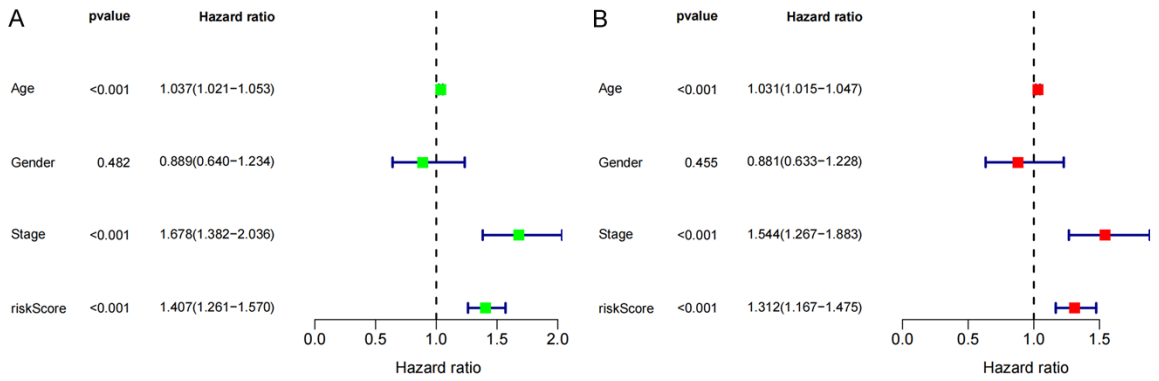

Figure 7. A. Univariate Cox regression analysis for independent prognostic factors. B. Multivariate Cox regression analysis for independent prognostic factors.

low-expression genes, we conducted a prognosis study combined with the clinical information of patients with bladder cancer, and screened out four key prognosis genes: *CD74, METRN, PTPRR,* and *CDC42EP5*. Based on these four genes, we constructed a prognostic model and a nomogram. Both the survival analysis and ROC curves demonstrated that this model is accurate, and the calibration curve indicated good consistency between the predicted and observed survival rates.

Among the genes involved in the model construction, the protein encoded by *CD74* is associated with the class II major histocompatibility complex (MHC) and is an important partner in regulating the antigen presentation of immune responses. It also serves as a cell surface receptor for the cytokine macrophage migration inhibitory factor (MIF). When bound to its coding protein, it initiates survival pathways and cell proliferation. This is consistent with the results of the MF in our previous GO analysis, which was found to be associated with a better prognosis in immune diseases and hepatocellular carcinoma [23]. This is consistent with our study, as the hazard ratio value of *CD74* in the model was <1, suggesting that its high expression may serve as a protective factor for patients.

Previous research has shown that *METRN* regulates the differentiation of neuroglial cells and promotes the formation of axonal networks in neurogenesis [24]. Its study in cancer is relatively limited; however, recent research has shown that its high expression is associated with poor prognosis in colorectal cancer [25], which is consistent with our study results. In the model, its hazard ratio value is >1, which is a risk factor for patient prognosis.

The protein encoded by *PTPRR* is a member of the protein tyrosine phosphatase (PTP) family. PTPs are signaling molecules that regulate a variety of cellular processes, including cell growth, differentiation, mitotic cycle, and oncogenic transformation. This PTP possesses an extracellular region, a single transmembrane region, and a single intracellular catalytic domain, and thus represents a receptor-type PTP. Silencing of *PTPRR* has been associated with colorectal cancer [26]. Furthermore, existing studies have shown that the silencing of *PTPRR* is also associated with the metastasis and deep infiltration of cervical cancer, which can activate MAPK signal transduction and promote metastasis [27]. This is consistent with our model, that is, the hazard ratio value is <1. Furthermore, similar to bladder cancer, the prognosis of cervical cancer is clinically poor after the occurrence of muscle layer infiltration. Therefore, future research can refer to the direction of cervical cancer in the hope of discovering specific mechanisms in bladder cancer.

The protein encoded by *CDC42EP5* is a member of the Borg (binder of Rho GTPases) family of CDC42 effector proteins. There is limited research on this protein; however, some studies have revealed that it may be a potential gene involved in the development of prostate cancer and neurosystem tumors. Other research has proposed that it has a role in the migration of macrophages [28]. In our study, its

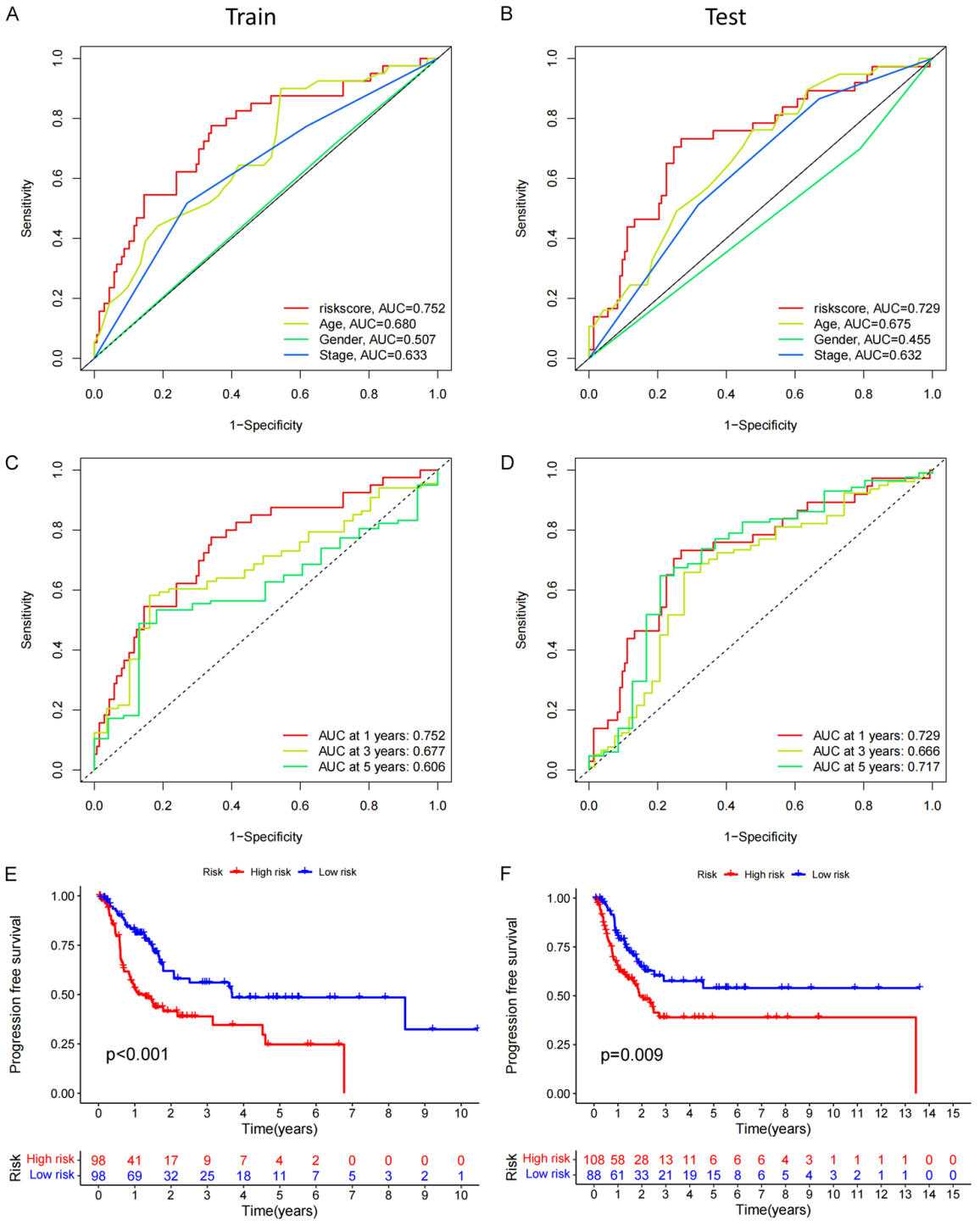

Figure 8. A, C. Receiver operating characteristic (ROC) curves for the clinical characteristics and the 1-, 3-, and 5-year survival in the training set. B, D. ROC curves for the clinical characteristics and the 1-, 3-, and 5-year survival in the validation set. E. Progression-free survival curve in the training set. F. Progression-free survival curve in the validation set.

hazard ratio is <1, indicating that it is a protective factor for patient prognosis; however, the specific mechanism remains to be studied.

In summary, by analyzing the single-cell sequencing data of patients with bladder cancer, we identified MRGs and used TCGA bladder

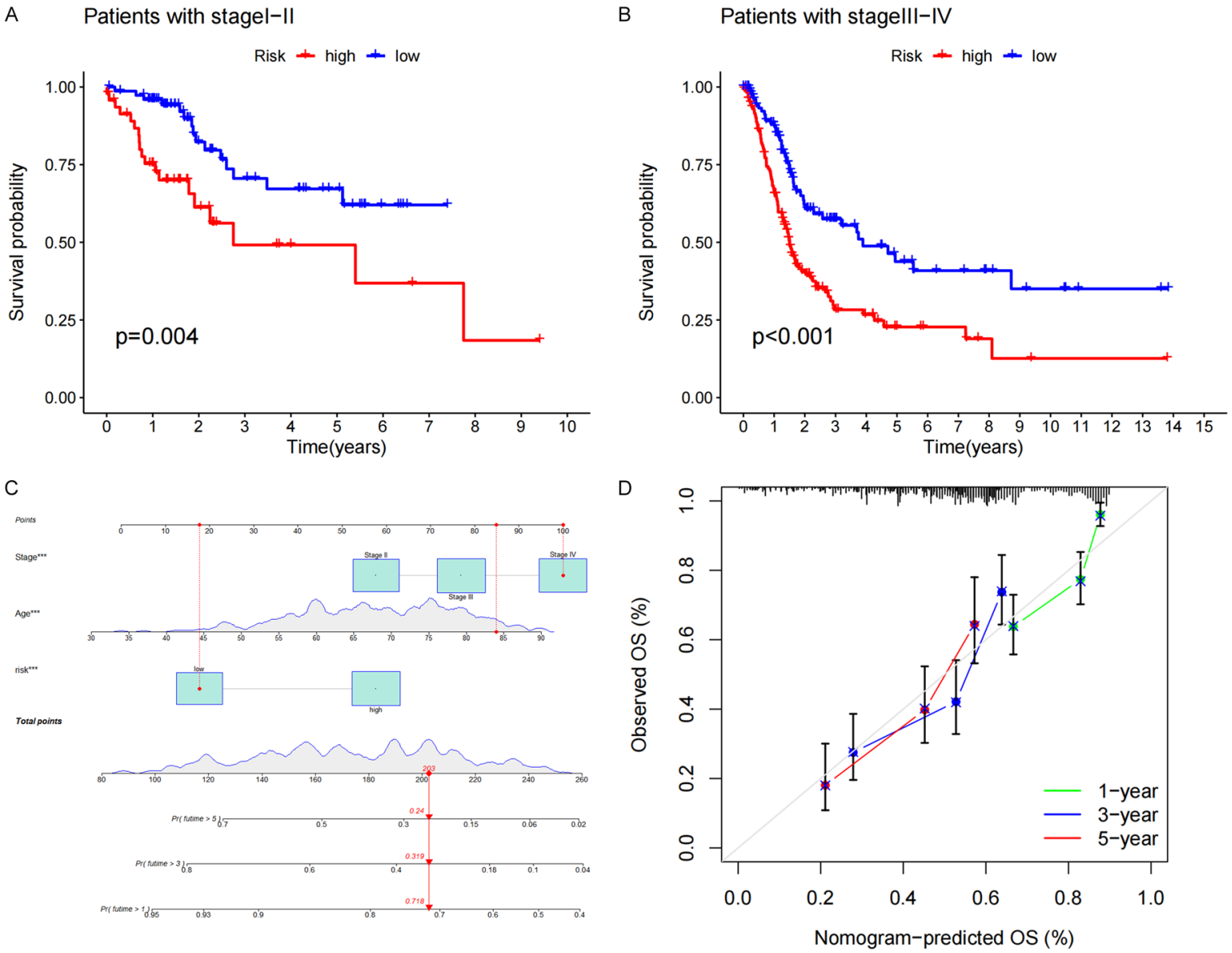

Figure 9. A. Survival curve for patients with stage I-II. B. Survival curve for patients with stage III-IV. C. Prognostic nomogram for patients with bladder cancer. D. Calibration curve of the prognostic nomogram.

cancer second-generation sequencing data and clinical data to construct a prognostic model. This provides a promising approach for individualized survival prediction and clinical outcome prediction in patients with bladder cancer. However, our study has certain limitations as it is based on a public database, without clinical trial research and basic research verification. The four genes, *CD74, METRN, PTPRR*, and *CDC42EP5*, have been found to be involved in tumor invasion, heterogeneity, and immune regulation in multiple cancers; however, focused research on bladder cancer is relatively limited. In the future, our research will focus on these four genes, with the hope to uncover their key biological steps in the development of bladder cancer.

Acknowledgements

We are very grateful for the support of this research by the Guizhou Science and Technology Department for Basic Research Project (No. QKHJC-2020-1Y342).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Xu Fu, Center of Reproduction, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, No. 26 Dao Qian Road, Suzhou 215002, Jiangsu, PR China. E-mail: 1010306126@qq.com

References

[1] Suvà ML and Tirosh I. Single-cell RNA sequencing in cancer: lessons learned and emerging challenges. Mol Cell 2019; 75: 7-12.

[2] Wang H, Mei Y, Luo C, Huang Q, Wang Z, Lu GM, Qin L, Sun Z, Huang CW, Yang ZW, Chen J, Yin W, Qian CN, Zeng J, Chen L, Leng Q, Guo Y and Jia G. Single-cell analyses reveal mechanisms of cancer stem cell maintenance and epithelial-mesenchymal transition in recurrent bladder cancer. Clin Cancer Res 2021; 27: 6265-6278.

[3] Oh DY, Kwek SS, Raju SS, Li T, McCarthy E, Chow E, Aran D, Ilano A, Pai CS, Rancan C, Allaire K, Burra A, Sun Y, Spitzer MH, Mangul S, Porten S, Meng MV, Friedlander TW, Ye CJ and Fong L. Intratumoral CD4+ T cells mediate anti-tumor cytotoxicity in human bladder cancer. Cell 2020; 181: 1612-1625, e13.

[4] Dobruch J and Oszczudłowski M. Bladder cancer: current challenges and future directions. Medicina (Kaunas) 2021; 57: 749.

[5] Lenis AT, Lec PM, Chamie K and Mshs MD. Bladder cancer: a review. JAMA 2020; 324: 1980-1991.

[6] Chen X, Xu R, He D, Zhang Y, Chen H, Zhu Y, Cheng Y, Liu R, Zhu R, Gong L, Xiao M, Wang Z, Deng L and Cao K. CD8+ T effector and immune checkpoint signatures predict prognosis and responsiveness to immunotherapy in bladder cancer. Oncogene 2021; 40: 6223-6234.

[7] Xu Y, Zeng H, Jin K, Liu Z, Zhu Y, Xu L, Wang Z, Chang Y and Xu J. Immunosuppressive tumor-associated macrophages expressing interlukin-10 conferred poor prognosis and therapeutic vulnerability in patients with muscle-invasive bladder cancer. J Immunother Cancer 2022; 10: e003416.

[8] Jiang Z, Zhang Y, Zhang Y, Jia Z, Zhang Z and Yang J. Cancer derived exosomes induce macrophages immunosuppressive polarization to promote bladder cancer progression. Cell Commun Signal 2021; 19: 93.

[9] Kersten K, Hu KH, Combes AJ, Samad B, Harwin T, Ray A, Rao AA, Cai E, Marchuk K, Artichoker J, Courau T, Shi Q, Belk J, Satpathy AT and Krummel MF. Spatiotemporal co-dependency between macrophages and exhausted CD8+ T cells in cancer. Cancer Cell 2022; 40: 624-638, e9.

[10] Lai H, Cheng X, Liu Q, Luo W, Liu M, Zhang M, Miao J, Ji Z, Lin GN, Song W, Zhang L, Bo J, Yang G, Wang J and Gao WQ. Single-cell RNA sequencing reveals the epithelial cell heterogeneity and invasive subpopulation in human bladder cancer. Int J Cancer 2021; 149: 2099-2115.

[11] Natarajan KN, Miao Z, Jiang M, Huang X, Zhou H, Xie J, Wang C, Qin S, Zhao Z, Wu L, Yang N, Li B, Hou Y, Liu S and Teichmann SA. Comparative analysis of sequencing technologies for single-cell transcriptomics. Genome Biol 2019; 20: 70.

[12] Infantino V, Santarsiero A, Convertini P, Todisco S and Iacobazzi V. Cancer cell metabolism in hypoxia: role of HIF-1 as key regulator and therapeutic target. Int J Mol Sci 2021; 22: 5703.

[13] Cao R, Yuan L, Ma B, Wang G and Tian Y. Tumour microenvironment (TME) characterization identified prognosis and immunotherapy response in muscle-invasive bladder cancer (MIBC). Cancer Immunol Immunother 2021; 70: 1-18.

[14] Yunna C, Mengru H, Lei W and Weidong C. Macrophage M1/M2 polarization. Eur J Pharmacol 2020; 877: 173090.

[15] Martínez VG, Rubio C, Martínez-Fernández M, Segovia C, López-Calderón F, Garín MI, Teijeira A, Munera-Maravilla E, Varas A, Sacedón R, Guerrero F, Villacampa F, de la Rosa F, Castellano D, López-Collazo E, Paramio JM, Vicente Á and Dueñas M. BMP4 induces M2 macrophage polarization and favors tumor progression in bladder cancer. Clin Cancer Res 2017; 23: 7388-7399.

[16] Pan Y, Yu Y, Wang X and Zhang T. Tumor-associated macrophages in tumor immunity. Front Immunol 2020; 11: 583084.

[17] Zhu L, Fu X, Chen X, Han X and Dong P. M2 macrophages induce EMT through the TGF-β/Smad2 signaling pathway. Cell Biol Int 2017; 41: 960-968.

[18] Zhang H, Li R, Cao Y, Gu Y, Lin C, Liu X, Lv K, He X, Fang H, Jin K, Fei Y, Chen Y, Wang J, Liu H, Li H, Zhang H, He H and Zhang W. Poor clinical outcomes and immunoevasive contexture in intratumoral IL-10-producing macrophages enriched gastric cancer patients. Ann Surg 2022; 275: e626-e635.

[19] Chi PL, Cheng CC, Hung CC, Wang MT, Liu HY, Ke MW, Shen MC, Lin KC, Kuo SH, Hsieh PP, Wann SR and Huang WC. MMP-10 from M1 macrophages promotes pulmonary vascular remodeling and pulmonary arterial hypertension. Int J Biol Sci 2022; 18: 331-348.

[20] Lian G, Chen S, Ouyang M, Li F, Chen L and Yang J. Colon cancer cell secretes EGF to promote M2 polarization of TAM through EGFR/PI3K/AKT/mTOR pathway. Technol Cancer Res Treat 2019; 18: 1533033819849068.

[21] Lei Y, Tang R, Xu J, Wang W, Zhang B, Liu J, Yu X and Shi S. Applications of single-cell sequencing in cancer research: progress and perspectives. J Hematol Oncol 2021; 14: 91.

[22] Zhang Z, Wang ZX, Chen YX, Wu HX, Yin L, Zhao Q, Luo HY, Zeng ZL, Qiu MZ and Xu RH. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. Genome Med 2022; 14: 45.

[23] Xiao N, Li K, Zhu X, Xu B, Liu X, Lei M and Sun HC. CD74+ macrophages are associated with favorable prognosis and immune contexture in hepatocellular carcinoma. Cancer Immunol Immunother 2022; 71: 57-69.

[24] Nishino J, Yamashita K, Hashiguchi H, Fujii H, Shimazaki T and Hamada H. Meteorin: a secreted protein that regulates glial cell differentiation and promotes axonal extension. EMBO J 2004; 23: 1998-2008.

[25] Xu X, Zhang C, Xia Y and Yu J. Over expression of METRN predicts poor clinical prognosis in colorectal cancer. Mol Genet Genomic Med 2020; 8: e1102.

[26] Menigatti M, Cattaneo E, Sabates-Bellver J, Ilinsky VV, Went P, Buffoli F, Marquez VE, Jiricny J and Marra G. The protein tyrosine phosphatase receptor type R gene is an early and frequent target of silencing in human colorectal tumorigenesis. Mol Cancer 2009; 8: 124.

[27] Su PH, Lin YW, Huang RL, Liao YP, Lee HY, Wang HC, Chao TK, Chen CK, Chan MW, Chu TY, Yu MH and Lai HC. Epigenetic silencing of PTPRR activates MAPK signaling, promotes metastasis and serves as a biomarker of invasive cervical cancer. Oncogene 2013; 32: 15-26.

[28] Farrugia AJ, Rodríguez J, Orgaz JL, Lucas M, Sanz-Moreno V and Calvo F. CDC42EP5/BORG3 modulates Sept9 to promote actomyosin function, migration, and invasion. J Cell Biol 2020; 219: e201912159.