Research article

The prognostic significance and clinical relevance of stem cell characteristic in bladder cancer

Xia Chen [a,1], Yuanyuan Yin [a,1], Yuming He [b], Fanqi Meng [b], Jing Zhao [b], Fang Liu [b], Yu Xu [b], Guoqiang Wang [b], Xin Zhu [b], Sijia Ma [b], Huafei Lu [b], Shangli Cai [b], Yanping Song [c], Junyong Dai [c,*]

[a] *Clinical Research Center, Chongqing University Cancer Hospital, Chongqing, China*
[b] *Burning Rock Biotech, Guangzhou, China*
[c] *Department of Urological Oncology, Chongqing University Cancer Hospital, Chongqing, China*

ARTICLE INFO

ABSTRACT

Background: Bladder cancer (BLCA) is a common malignant tumor of urinary system and prognostic biomarkers are needed for better clinical decision-making and patient management. Cancer stem cells (CSCs) are involved in carcinogenesis, development, metastasis and recurrence of BLCA. This study explored the prognostic and predictive value of CSCs-related genes and laid the groundwork for precision treatment development in BLCA.

Methods: The mRNA data and corresponding clinical information obtained from TCGA-BLCA cohort was used to discover biomarkers and develop CSCs-related prognostic model, which was further validated in GSE32548 and GSE32894 datasets. In addition, the association between CSCs-related risk score and therapeutic efficacy was analyzed to explore the potential predictive value of the prognostic model.

Results: We identified four CSCs-related subtypes and 900 differentially expressed genes (DEGs) among subtypes. Then the CSCs-related prognostic model was built based on 16 CSCs-related DEGs with the most significant prognostic value. Patients in the low-risk group had better overall survival than those in high-risk group (P < 0.001; HR, 0.42; 95 %CI, 0.31–0.57). Multivariable Cox analysis in training and test sets confirmed the independence of CSCs-related risk score as a prognostic factor (P < 0.05). The difference of survival between two risk groups were probably due to the significantly varied immune microenvironment based on the analysis of infiltrated immune cells. Additionally, the risk score was significantly associated with chemotherapy sensitivity and the response to anti-PD-L1 therapy (P < 0.05) which suggested a potential predictive value of CSCs-related risk model.

Conclusion: We established a risk classifier based on 16 CSCs-related genes for predicting survival in patients with BLCA. The CSCs-related risk model has both prognostic value and potential predictive value for therapeutic efficacy, which brings us closer to understanding the important role of CSCs in BLCA and may provide guidance for clinical treatment decision-making and patient management. The clinical utility of the CSCs-related risk classifier warrants further studies.

---

* Corresponding author.
   E-mail address: urology0218@163.com (J. Dai).
[1] Xia Chen and Yuanyuan Yin contributed equally to this work.

1. Introduction

Bladder cancer (BLCA) is a common malignant tumor of urinary system with more than 573,000 new cases and 213,000 deaths worldwide in 2020 [1]. BLCA accounts for the fourth highest number of estimated new cases in men in the United States in 2023 [2]. About 70 % of BLCA patients are non-muscle-invasive bladder cancer (NMIBC) and the others are muscle-invasive bladder cancer (MIBC) [3]. Endoscopic resection and intravesical Bacillus Calmette-Guerin (BCG) are the mainstay therapies of NMIBC, and the therapeutic options of MIBC range from radical cystectomy, chemotherapy, radiation, to immune checkpoint inhibitors (ICIs) [3]. The 5-year recurrence free survival rates of patients with low-risk, intermediate-risk and high-risk NMIBC are 43 %, 33 % and 21 %, respectively, and most of high-risk NMIBC patients progress to MIBC with worse prognosis [3]. Thus, a prognostic model of BLCA for clinical management is needed.

Cancer stem cells (CSCs) have been indicated playing a crucial role in chemotherapy resistance and tumor relapse [4]. CSCs are a sort of malignant tumor cells which can self-renew, differentiate and tumor-initiate, leading to the abilities of resisting anticancer therapy, promoting relapse and metastasis, and increasing tumor heterogeneity [4–6]. CSCs were firstly described in acute myeloid leukemia, and soon identified across multiple cancers, such as breast cancer [7], colon cancer [8], ovarian cancer [9], and lung cancer [10]. A group of specific cell surface markers can be used to distinguish CSCs from normal cells. Various markers including CD44, ALDH1A1, SOX2, and SOX4 have been identified and isolated from BLCA patient specimens [11]. These involved genes promote bladder cancer progression via regulating epithelial to mesenchymal transition (EMT), hedgehog signaling pathway, MAPKs, and JAK-STAT pathway [4,11].

It has been demonstrated that high-stemness signatures are negatively associated with immune reaction across 21 types of solid tumors [12]. Emerging evidences suggest that CSCs, infiltrating immune cells, and other cell lineages, interact with each other complexly and dynamically in tumor microenvironment (TME) [12–18]. For example, hypoxic TME induces the development of CSCs through activating Wnt and Notch signaling pathways in breast cancer [19]. CSCs can impair the activity of cytotoxic T cell and promote immune evasion of tumor cells by elevating inhibitory checkpoint receptors levels [20]. Chan et al. reported that the expression level of CD47 which bound to signal-regulatory protein alpha (SIRPα) and inhibited macrophage phagocytosis was higher
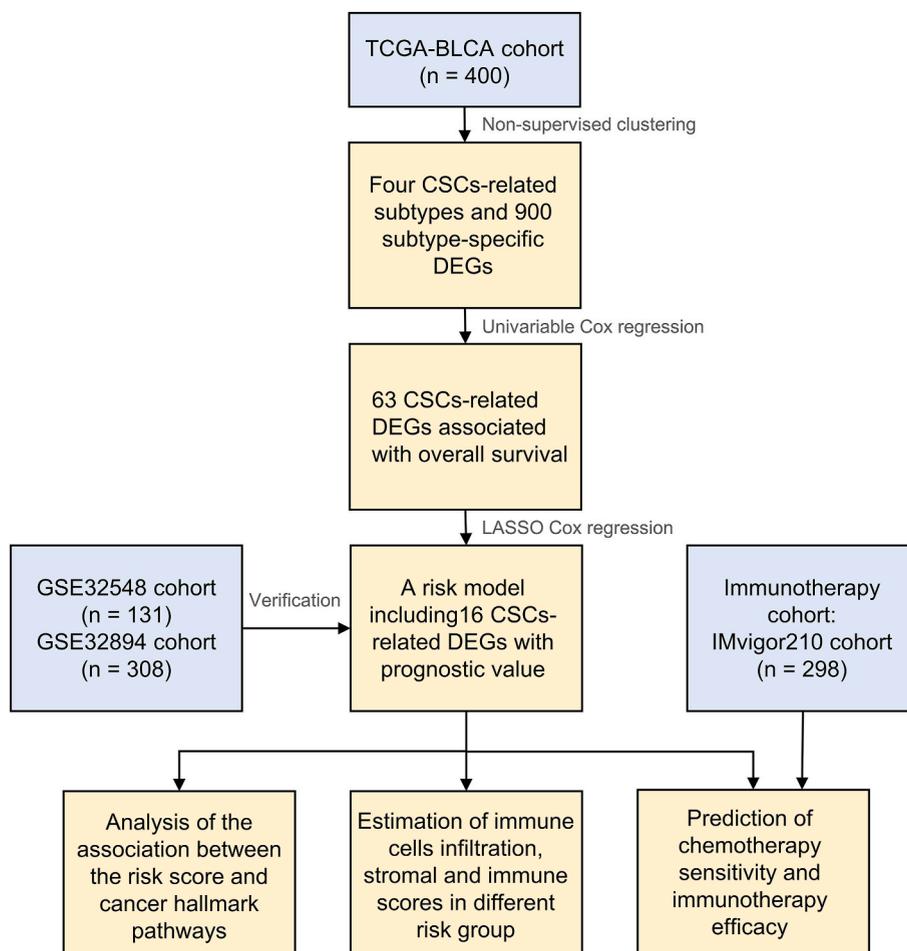
Fig. 1. **Workflow chart.** Abbreviations: BLCA = Bladder cancer, CSCs = Cancer stem cells, DEGs = differentially expressed genes.

in bladder CSCs than that of normal cells [21]. Also, Chan demonstrated that a BLCA-CSCs signature based on differentially expressed genes (DEGs) from $CD47^+$ and $CD47^-$ BLCA cells could predict the invasive progression of NMIBC to MIBC [21]. Altogether, these results suggest that CSCs are important in BLCA, however, whether CSCs-related genes can be used as a prognostic tool in BLCA has not been well reported.

In our study, the prognostic value of CSCs-related genes in BLCA patients was explored, which might provide a better understanding of CSCs features in BLCA for clinical practice. We divided BLCA patients into four subtypes based on 187 CSCs-related genes in TCGA-BLCA cohort using unsupervised clustering method. A classifier including 16 core CSCs-related genes from DEGs among CSCs-subtypes was built and validated. Additionally, we systematically analyzed the distribution of clinical traits, prognostic significance, and the association between the risk classifier and immune cells infiltration, sensitivity to ICIs therapy and chemotherapy, and changes in biological pathways. These efforts can potentially guide clinical treatment decision based on distinct risk levels and bring new insights of CSCs in BLCA.

2. Method and material

2.1. Samples and patients

Datasets used in this study were acquired from TCGA, gene expression omnibus (GEO), and IMvigor210 datasets. The TCGA-BLCA cohort comprising of 400 cancer samples (with somatic mutation, clinical and mRNA expression matrix data) were obtained from TCGA (https://tcga-data.nci.nih.gov/tcga/) and UCSC XENA database (https://xenabrowser.net/datapages/) [22]. We also down-loaded GSE32548 (n = 131) [23] and GSE32894 (n = 308) [24] cohorts from the GEO database harnessed as the external validation sets. Above three cohorts of BLCA were included in the present study, and the corresponding clinical information was summarized in Supplementary Table S1. Moreover, the clinical and mRNA gene expression data of 298 metastatic urothelial carcinoma (mUC) patients treated with atezolizumab from IMvigor210 cohort was retrieved via R package "IMvigor210CoreBiologies" and used to explore the difference in response to anti-PD-L1 therapy [25]. The design for this study is shown in Fig. 1.

2.2. Identification of BLCA subtypes based on CSCs-related genes

To identify the different stemness status of BLCA samples, "ConsensusClusterPlus" R package was used for consensus clustering based on a total of 187 CSCs-related genes which was obtained from a previous research (Supplementary Table S2) [26]. We extracted the expression data of CSCs-related genes from the TCGA-BLCA cohort. Then partition around medoids cluster was performed using 80 % of samples each try with the standard of Euclidean distance for 100 iterations. Clustering score was used to choose the best cluster number for the cumulative distribution function curve with area under the curve monitored.

2.3. Identification of DEGs among BLCA subtypes

To identify genes with differential expression among subtypes, we sequentially screened the DEGs between samples of one subtype and the rest subtypes using "limma" R package [27]. The up-regulated and down-regulated genes which were generated with |log ration of fold change| > 1 and adjusted P value < 0.05 as cutoff values, were considered as subtype-specific DEGs. To explore the biological significances of these DEGs, we performed function enrichment analysis with R package "clusterProfile" [28]. The significant enriched genes were accepted with adjusted P value < 0.05 as threshold.

2.4. Construction and validation of the CSCs-related prognostic model

To explore the prognostic significance of CSCs-related genes, we constructed a CSCs-related prognostic model in TCGA-BLCA cohort. Firstly, all DEGs from different cluster subtypes were used to perform univariable Cox analysis for overall survival (OS) to screen candidate genes (P value < 0.01 as threshold). Secondly, the least absolute shrinkage and selection operator (LASSO) Cox regression analysis was used to screen core prognostic signature genes with standard parameter. A risk model originated from 16 CSCs-related DEGs was established based on the formula: *The risk score* $= \sum_i^{16}$ coefficients$_i * exp_i$ (coefficients$_i$ represents the Cox coefficients for gene i; exp$_i$ represents the expression value for gene i). Based on the median risk score, patients were split into low-risk and high-risk groups. The OS between patients in different risk group was compared by Kaplan-Meier survival analysis and log-rank test. To confirm the prognostic power of the CSCs-related risk stratification model, we also utilized the receiver operating characteristic (ROC) analysis and area under curve (AUC) for evaluation with R package "timeROC". Notably, GSE32548 and GSE32894 datasets were analyzed as independent validation sets to confirm the performance of the CSCs-related prognostic model. Specially, the risk score and other clinicopathological features (including age, sex, grade and TNM stage) were included in successive univariable and multivariable Cox regression analysis to test the independence of the risk score as a prognostic factor of BLCA.

The RNA-seq data used in this study was transformed into fragments per kilobase million (FPKM) data and then processed by log (FPKM+1) transformation before analysis both in the model training cohort and model validation cohort. CSCs-related risk score was calculated according to the same formula, and patients were all split into low-risk and high-risk groups based on the median risk score. Statistical analysis methods were consistent across all cohorts to ensure comparability and reproducibility.
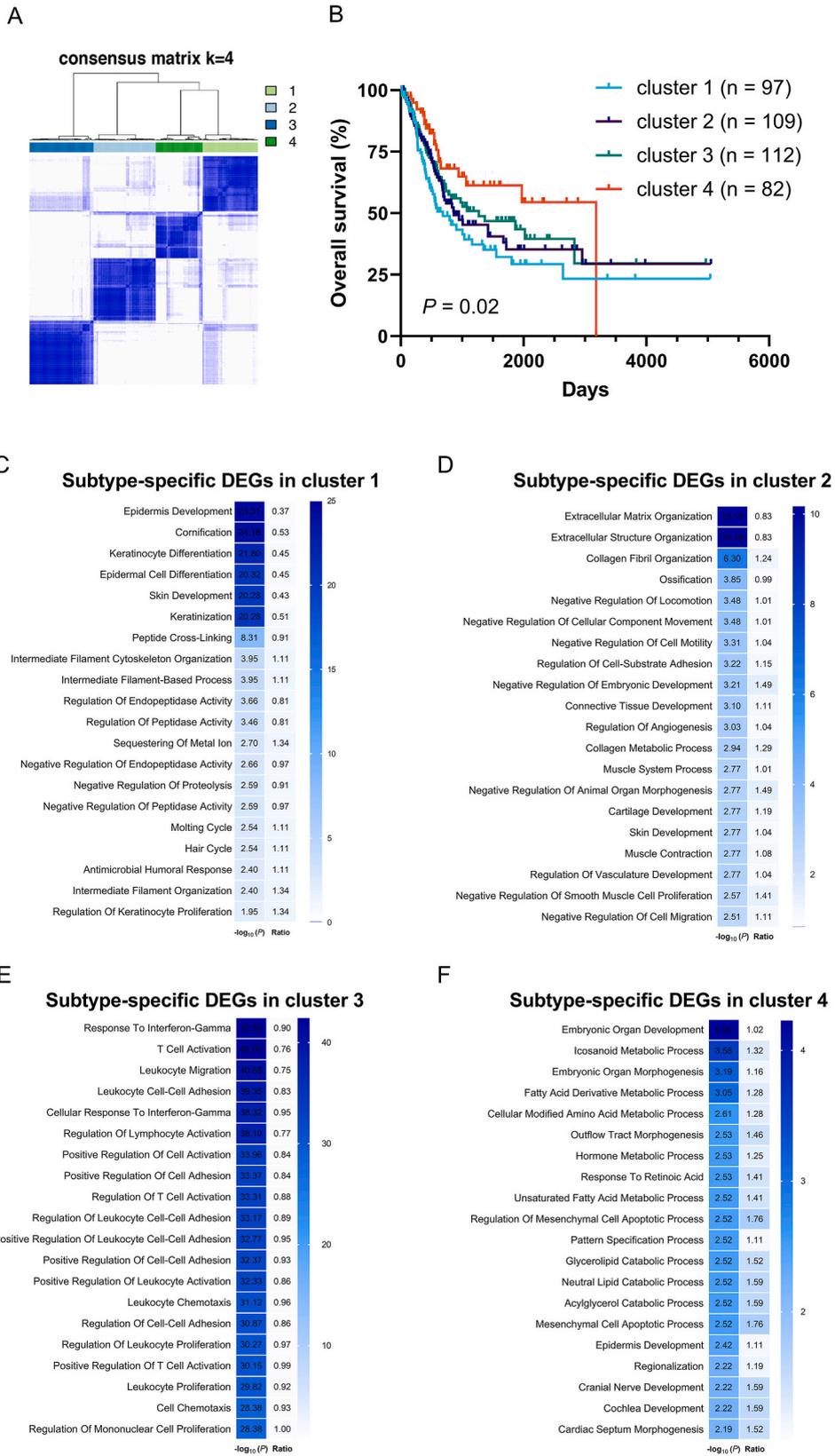
Fig. 2. Non-supervised clustering of CSCs-related genes and gene enrichment results of DEGs.
(A) The non-supervised clustering of CSCs-related genes in TCGA-BLCA cohort. (B) Kaplan-Meier curves for the four CSCs-subtypes of patients. (C–F) Enrichment analysis of subtype-specific DEGs in cluster 1 (C), cluster 2 (D), cluster 3 (E) and cluster 4 (F).

2.5. Correlation analysis between the risk score and cancer hallmark pathways

The Molecular Signature Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/; v7.4) is one of the largest and most popular repositories of gene sets [29]. Single sample gene set enrichment analysis (ssGSEA) algorithm was applied to calculate the enrichment scores of target pathways based on corresponding mRNA data from TCGA-BLCA cohort and gene sets (hallmark and KEGG pathway gene sets) retrieved from MSigDB. The ssGSEA which derived from the GSEA method was implemented by "GSVA" R package [30]. Spearman correlation analysis was applied to evaluate the correlation between risk scores and cancer hallmark pathways.

2.6. Estimation of immune cells infiltration and immune score

Tumor infiltrating leukocytes (TILs) are integral components of TME and associate with prognosis and response to therapy. The deconvolution approach Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) (http://cibersort. stanford.edu/) was used to measure the abundances of 22 distinct leukocyte subsets via mRNA profile [31,32]. CIBERSORT gene signature matrix, termed LM22, encompasses 547 genes and differentiates 22 human immune cells including seven T cell types, naive and memory B cells, plasma cells, natural killer (NK) cells, and myeloid subsets. The proportion of immune cells in TCGA-BLCA cohort was calculated to explore TILs pattern in different risk groups.

Besides, immune and stromal cells are the two main non-cancer components in TME, and they have important implications for tumor diagnosis and prognosis [15,33–35]. Thus we estimated the ratio of immune and stromal cells base on Immune Score, Stromal Score and Estimation of STromal and Immune cells in MAlignant Tumors using Expression data (ESTIMATE) Score, which were calculated from specific gene expression profiles by ESTIMATE algorithm [34].

2.7. Prediction of drug sensitivity

We collected the data of molecular markers associated with drug sensitivity and the response data to drugs in cancer cells from the public database referred as Genomics of Drug Sensitivity in Cancer (GDSC) (https://www.cancerrxgene.org/) [36], including 138 anti-cancer drugs and almost 700 cancer cell lines. We evaluated half-maximal inhibitory concentration (IC50) value to predict drug sensitivity in each patient with "pRRophetic" R package [37]. To assess the association between CSCs-related genes and the sensitivity to anti-cancer drugs, we compared the difference of IC50 of patient samples between high- and low-risk groups.

2.8. Statistical analysis

All statistical tests were performed using R software, version 4.0.1 (R Foundation for Statistical Computing Vienna, Austria). Differences in OS between risk groups were compared using Kaplan-Meier curves, with *P*-values calculated via the log-rank test using the "survival" R package. Hazard's ratio and its 95 % confidence interval (CI) were determined by univariable Cox regression. The heat maps were plotted by "pheatmap" R package. The differences of drug sensitivity, immune cells infiltration, Stromal Score, Immune Score, and ESTIMATE Score between risk groups were compared by Mann Whitney test. The differences of immunotherapy response and BRCA1/BRCA2 mutation between risk groups were compared by Chi-square test. All $P < 0.05$ was considered statistically significant, if otherwise stated.

3. Results

3.1. Identification of CSCs-related subtypes and subtype-specific DEGs in BLCA

To explore the molecular function of CSCs-related genes, we performed non-supervised clustering analysis based on the expression of CSCs-related genes in TCGA-BLCA cohort and divided samples into four CSCs-subtypes named as cluster 1–4 (Fig. 2A). To analyze the relevance of CSCs-subtypes and OS, we performed Kaplan-Meier analysis and found that the OS of four clusters were significantly different (P = 0.02, Fig. 2B).

Then, we identified a total of 900 subtype-specific DEGs among the four subtypes, including 65 DEGs in cluster 1 vs cluster 2/3/4, 162 DEGs in cluster 2 vs cluster 1/3/4, 494 DEGs in cluster 3 vs cluster 1/2/4, and 259 DEGs in cluster 4 vs cluster 1/2/3. To reveal the underlying biological characteristics of each CSCs-subtype, we conducted functional enrichment analysis in each group of subtype-specific DEGs. The results showed that the DEGs in cluster 1 were mainly involved in development and cell differentiation (Fig. 2C). DEGs in cluster 2 were mainly involved in extracellular matrix related pathways (Fig. 2D). DEGs in cluster 3 mainly regulate immunity (Fig. 2E). And the main involved regulation of DEGs in cluster 4 was metabolism (Fig. 2F). Altogether, we identified four clusters in BLCA patients with different survival and enriched pathways based on the CSCs-related genes.

3.2. Development and validation of a CSCs-related risk stratification model in BLCA

As prognostic monitoring is important for BLCA clinical management, we focused our efforts on the discovery of CSCs-related biomarkers, which could serve as potential effective prognostic indicators. Using univariable Cox regression analysis, 63 CSCs-related DEGs that were significantly associated with OS in BLCA patients were identified (P < 0.01, Supplementary Table S3). The candidate genes were utilized to establish the CSCs-related prognostic model to predict the OS of patients in TCGA-BLCA cohort. We penalized the unimportant features using LASSO Cox regression analysis and finally selected 16 CSCs-related biomarkers (including CD109, SERPINB2, PRELP, HNF1B, ZNF683, ETV7, FN1, SMAD6, SPOCD1, SLC45A3, HOXB3, OVGP1, GOLGA8A, CTSE, PTPRR, and MST1R) to build the CSCs-related prognostic model (Fig. 3A–B).
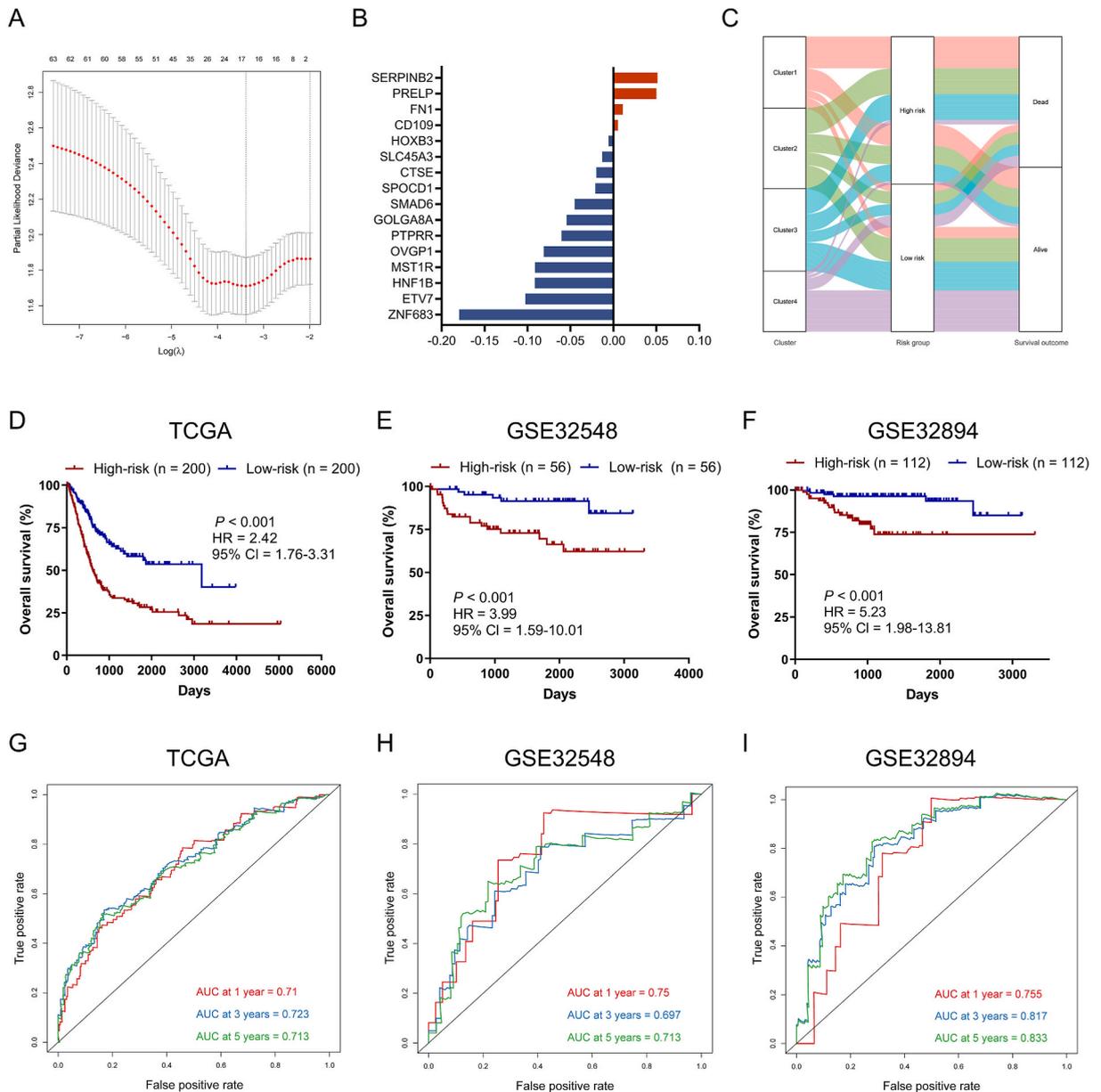
Fig. 3. Construction and validation of the CSCs-related prognostic model
(A) The results of the least absolute shrinkage and selection operator Cox regression analysis for core prognostic gene selection. (B) Regression coefficients of the 16 genes included in the prognostic model. (C) The relationship of four CSCs-subtypes, risk groups, and survival outcome of patients. (D–F) Kaplan-Meier curves of overall survival (OS) between patients with high and low CRCs-related prognostic score in TCGA-BLCA cohort (D), GSE32584 cohort (E) and GSE32894 cohort (F). (G–I) Time-dependent ROC curve analysis for risk score in TCGA cohort (G), GSE32584 cohort (H) and GSE32894 cohort (I).

By weighting the expression levels of the 16 genes with corresponding regression coefficients, a risk stratification model was developed to predict patients' survival. Patients were divided into high-risk and low-risk groups based on the median value of risk scores. The association between four CSCs-subtypes and the two risk groups is depicted in Fig. 3C, with most of the patients in cluster 1 stratified into high-risk while most of the patients in cluster 4 stratified into low-risk. The Kaplan-Meier survival analysis revealed that the low-risk patients had a better OS than those with high-risk (median OS, 3183 vs 623 days; P < 0.001; HR, 0.42; 95 % CI, 0.31–0.57, Fig. 3D). Meanwhile, we validated the prognostic value of CSCs-related prognostic model using GSE32894 and GSE32548 validation sets, and similar results were observed (GSE32894: P < 0.001; HR, 5.23; 95 % CI, 1.98–13.81, Fig. 3E; GSE32548: P < 0.001; HR, 3.99; 95 % CI, 1.59–10.01, Fig. 3F). To further evaluate the performance of CSCs-related prognostic model, we applied ROC analysis for OS and found that the AUC was 0.710, 0.723 and 0.713 at 1-, 3- and 5-years, respectively in TCGA-BLCA cohort (Fig. 3G). Meanwhile, the AUC was 0.75, 0.697, 0.713 for 1-, 3- and 5-years in GSE32548 dataset, and 0.755, 0.817, 0.833 for 1-, 3- and 5-years in GSE32894 dataset (Fig. 3H–I).

To further explore the independence of the risk score as a prognostic factor, multivariable Cox regression was performed adjusting several clinicopathological variables including sex, age, stage, and grade. The results indicated that the risk score was independently associated with OS in TCGA-BLCA, GSE32548, and GSE32894 cohorts (Table 1).

3.3. Association between the risk score from CSCs-related risk model and the sensitivity to anti-cancer drugs

To explore whether the risk score can potentially predict the response to chemotherapy in BLCA patients, we compared the IC50 values of multiple anti-cancer drugs between the two groups with different risk. The IC50 value of cisplatin, gemcitabine, vinblastine, and doxorubicin in the patients of high-risk group was lower than that of low-risk group (P < 0.05, Fig. 4A–D), suggesting a higher responsiveness when treated with these chemotherapeutic drugs in patients with high risk. On the contrary, the IC50 value of methotrexate of the high-risk group was higher than that of the low-risk group (P < 0.05, Fig. 4E).

To investigate whether the risk score could be used to potentially guide ICIs therapy, we conducted survival analysis in mUC patients treated with anti-PD-L1 regimen in IMvigor210 cohort. The low-risk group was significantly associated with favorable OS (median OS, not reached vs 10.9 months; P = 0.01; HR, 0.51; 95 % CI, 0.29–0.87; Fig. 4F). Additionally, the objective response rate (ORR) of low-risk patients was also significantly higher than that of high-risk patients (45 % vs 24 %, Chi-square test, P = 0.04), suggesting that low-risk patients could gain more benefit from ICIs (Fig. 4G). Particularly, patients with complete response (CR) had the lowest risk scores and the risk scores among patients with CR, PR, SD, and PD were significantly different (Kruskal-Wallis, P = 0.03, Fig. 4H). Overall, the predictive value of risk score from CSCs-related risk model for chemotherapeutics and ICIs was observed.

3.4. The CSCs-related risk groups reflected different molecular characteristics

To explore possible contribution of the differences between risk groups, we further analyzed the molecular characteristics and somatic mutations by R package "Maftools" in TCGA-BLCA cohort. The top 5 genes with highest mutation frequency were ERCC2, TET1, CHEK2, ASCC3, and NDC80 in high-risk group (Fig. 5A), and ERCC2, CENPE, CHEK2, PTPRC, and TET1 in low-risk group (Fig. 5B). In high-risk group, the mutation frequencies of ASCC3 and NDC80 were significantly higher, and that of CENPE and PTPRC were lower than low-risk group.

We also conducted ssGSEA analysis in TCGA-BLCA cohort to study the correlation between risk score and the activity of cancer hallmarks pathway. The top 30 significant cancer hallmarks pathways correlated with risk score are exhibited in Fig. 5C. The hallmark

Table 1
Multivariable Cox regression analysis of risk score and clinicopathological characteristics for OS in TCGA, GSE32584, and GSE32894.

Data	Clinical Factor	Comparison	Hazard Ratio (95% Cl)	P value
	Grade	High vs Low	1.10 (0.27 - 4.58)	0.89
TCGA	Sex	Male vs Female	1.01 (0.73 - 1.40)	0.95
	Age	>60 vs <=60	1.70 (1.09 - 2.64)	0.02
	Risk score	High vs Low	4.24 (2.90 - 6.22)	<0.001
	Grade	High vs Low	1.54 (0.50 - 4.78)	0.45
GSE32548	Sex	Male vs Female	1.35 (0.49 - 3.73)	0.51
	Age	>60 vs <=60	2.11 (0.63 - 7.05)	0.22
	Risk score	High vs Low	5.8 (0.99 - 34.06)	0.04
	Grade	High vs Low	5.91 (1.60 - 21.84)	<0.01
GSE32894	Sex	Male vs Female	1.31 (0.47 - 3.64)	0.60
	Age	>60 vs <=60	0.55 (0.2 - 1.48)	0.24
	Risk score	High vs Low	7.25 (1.59 - 33.19)	0.01
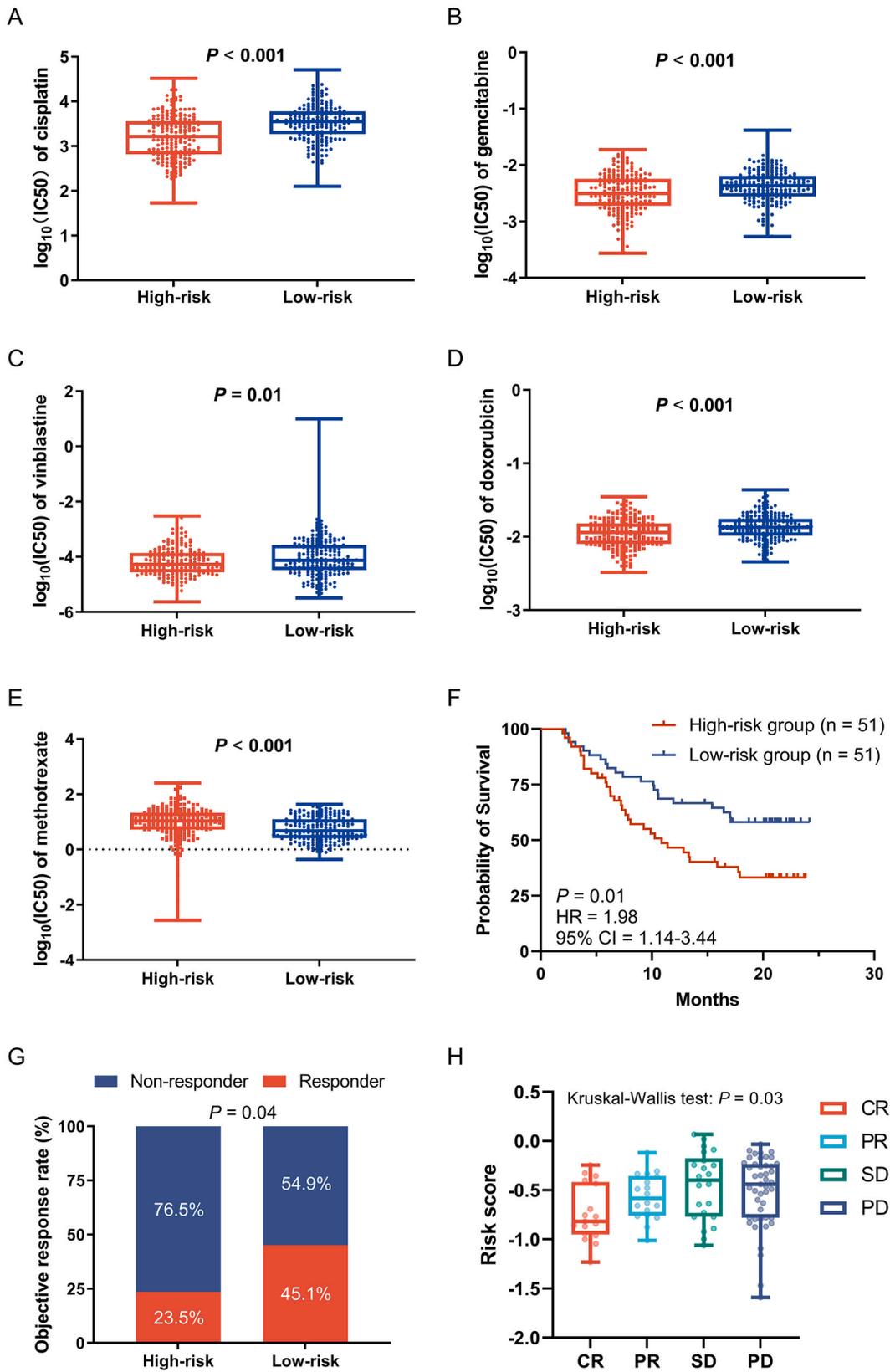
(caption on next page)

Fig. 4. Prediction of chemotherapeutic sensitivity and the response to ICI therapy.
(A–E) Comparison of the IC50 value of cisplatin (A), gemcitabine (B), vinblastine (C), doxorubicin (D) and methotrexate (E) between high- and low-risk groups. (F) Kaplan-Meier curves of OS between patients treated with immune checkpoint inhibitors (ICIs) in high- and low-risk groups in IMvigor210 cohort. (G) Comparison of objective response rate between high- and low-risk groups. (H) Comparison of the risk scores between patients with different responses to immunotherapy. Abbreviations: CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease.
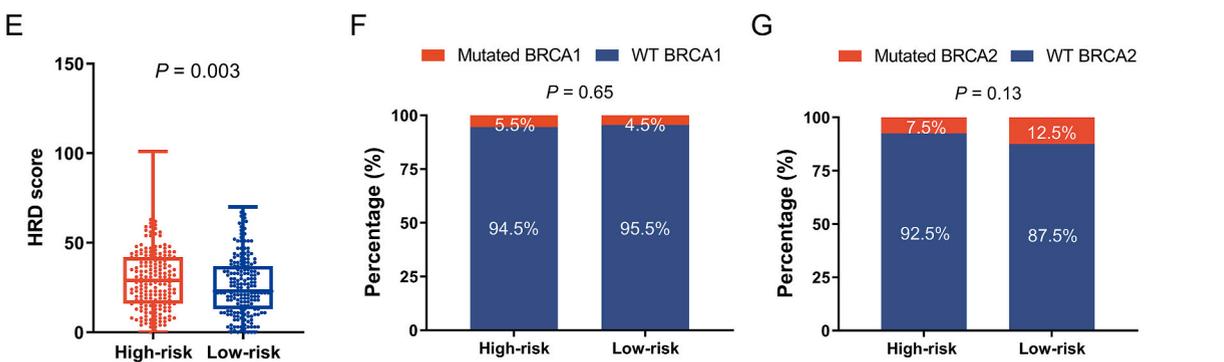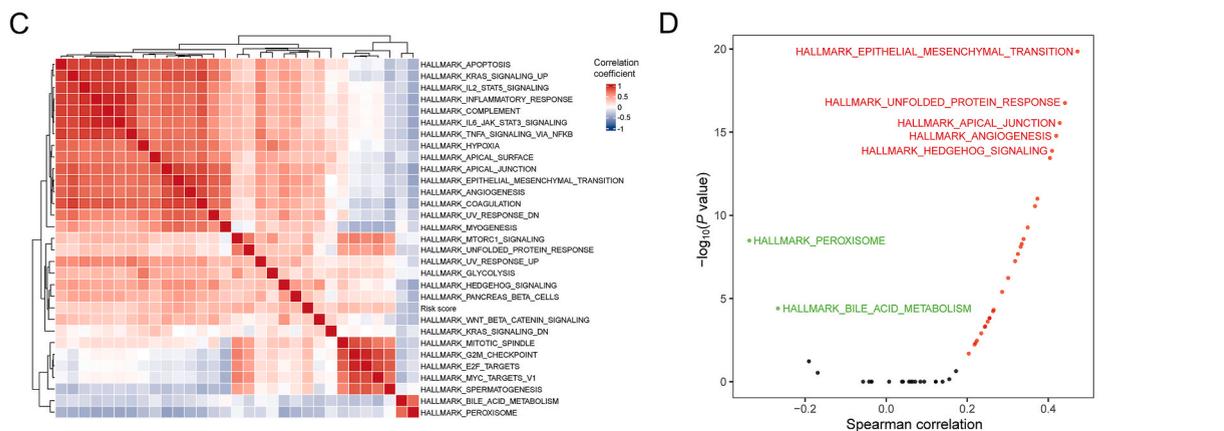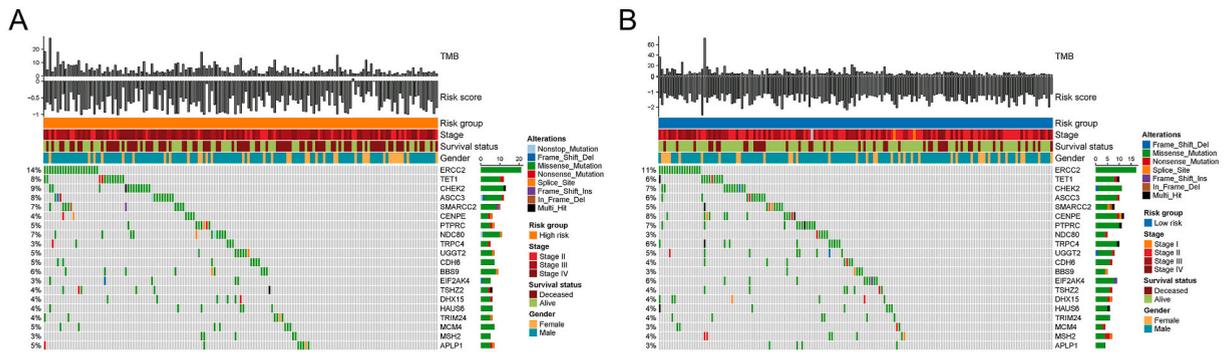
Fig. 5. Comparison of molecular characteristics between CSCs-related risk groups.
(A–B) Distribution of somatic mutations in high- (A) and low- (B) risk groups in TCGA-BLCA cohort. (C) The heat map of Spearman correlation between risk score and the activity of cancer hallmark pathway. (D) Cancer hallmark pathways with statistically significant correlation with risk score were shown. (E) Comparison of the homologous recombination deficiency (HRD) score between high- and low-risk groups. (F–G) Comparison of the percentages of mutated and wild-type BRCA1 (F) or BRCA2 (G) between high-risk and low-risk group pf patients.

epithelial to mesenchymal transition (EMT), hallmark unfolded protein response (UPR), hallmark apical junction, angiogenesis and hedgehog signaling pathway were most significantly positively correlated to the risk score, while the hallmark peroxisome and hallmark bile acid metabolism were most significantly negatively correlated to the risk score (Fig. 5D). Taken together, the results

suggested that CSCs-related genes included in the risk model were involved in complex and multifaceted biological processes in BLCA.

Since homologous recombination deficiency (HRD) has been demonstrated to be associated with neo-adjuvant chemotherapy in BLCA [38], we further compared the HRD score between high-risk and low-risk groups. It was found that HRD score in high-risk group was higher than that in low-risk group (P = 0.003, Fig. 5E), suggesting that the genome in high-risk patients were more unstable. Considering that the key molecules involved in HRD process are BRCA1 and BRCA2, we further analyzed the association between risk score and BRCA gene mutation, however, there was no significant difference in the risk score between BRCA mutated and wild groups (Fig. 5F–G).

3.5. The CSCs-related risk groups reflected different immune features

To further explore the association between the CSCs-related risk stratification model and immune system, the proportions of 22 immune cell infiltrates between two risk groups were compared by CIBERSORT analysis. It was found that the proportions of activated dendritic cells, resting dendritic cells, activated NK cells, plasma cells, regulatory T cells (Tregs), T follicular helper cells, and CD8$^+$ T cells were significantly higher in low-risk group, while the proportions of macrophages M0, M1 and M2 and neutrophils were significantly higher in high-risk group (P < 0.05, Fig. 6A). Spearman correlation analysis also demonstrated that the CSCs-related risk score was significantly positively correlated with the infiltration level of macrophages M0 and M2, whereas negatively correlated with the infiltration level of Tregs, T follicular helper cells, CD8$^+$ T cells, and activated dendritic cells (P < 0.05, Fig. 6B).

Based on the three scores generated by the ESTIMATE algorithm, the difference of the overall ratio of stromal and immune cells were respectively analyzed. The association between the three scores (stromal, estimate and immune) and CSCs-related risk score is presented in Fig. 6C–E. Both the stromal score and estimate score were observed significantly higher in high-risk group (P < 0.001, Fig. 6C–D), and the immune score was borderline significantly higher in high-risk group (P = 0.05, Fig. 6E), suggesting a lower tumor purity in the patients of high risk group.

4. Discussion

CSCs have been hypothesized as the main reason of high recurrent rate in BLCA. In our study, a risk model based on 16 CSCs-related signature genes was built and validated. It was found that the prognosis of two risk groups was significantly different in the training and validation sets. To further explore the biological and clinical association of the risk score, the clinical, immune, and mutational
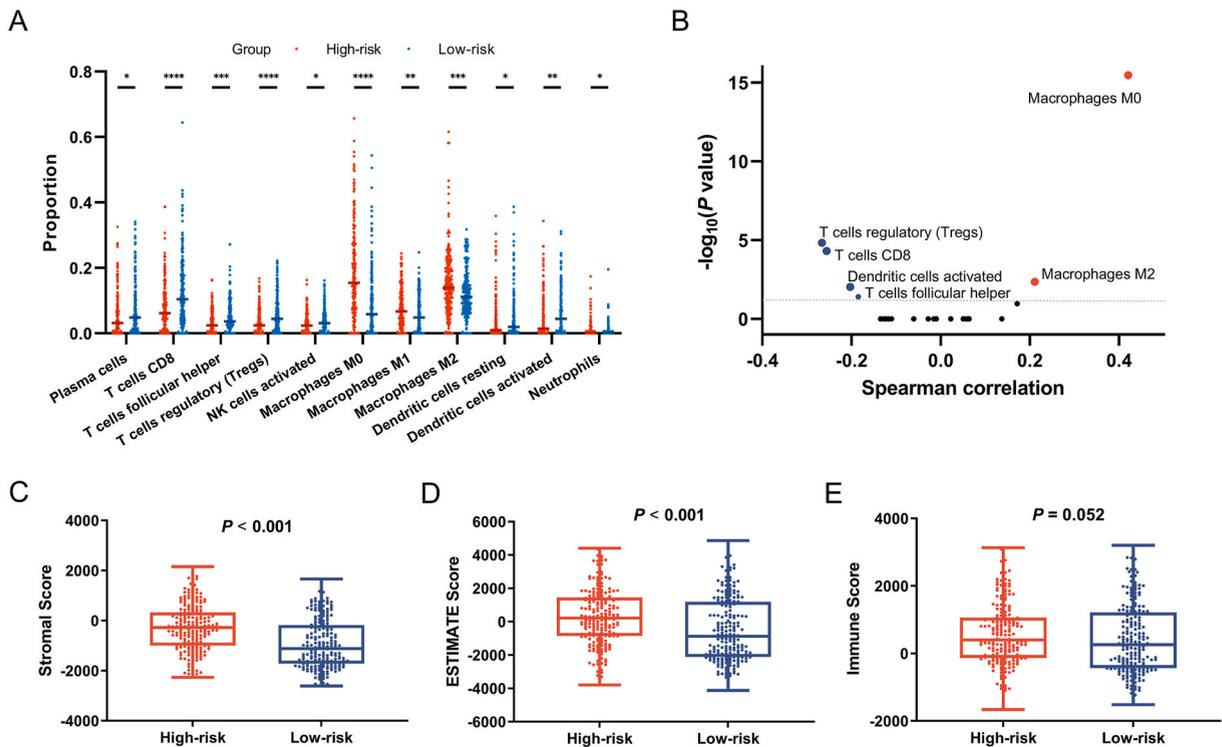
Fig. 6. Immune infiltration analysis between CSCs-related risk groups.
(A) Comparison of the proportions of the 22 immune cell infiltrates between low- and high-risk groups in TCGA-BLCA cohort. (B) Spearman correlation analysis between the risk score and immune cell infiltrates. Immune cells with statistically significant correlation were shown. (C–E) Comparison of the Stromal score (C), ESTIMATE score (D), and Immune score (E) between high- and low-risk groups.

characteristics between the high-risk and low-risk groups were compared. The most frequently mutated genes in each group were related to multifaceted biological processes such as cell growth, mitosis, and oncogenic transformation. Anti-tumor immune cells including activated dendritic cells, resting dendritic cells, activated NK cells, T follicular helper cells, and CD8$^+$ T cells were significantly higher in low-risk group, in contrast macrophages M0, M1 and M2 were higher in high-risk group. The chemotherapy sensitivity and response to ICIs therapy of the two groups was also significantly different.

In our study, a CSCs-related prognostic model with good performance was developed. There are several other studies which also focused on the prognostic role of tumor stemness (TS)-related biomarkers. Zhang et al. reported a risk model including 61 TS-related genes which were associated with EMT and CSCs in BLCA [39]. The TS-related risk model could distinguish prognosis and also was associated with immunotherapy response. Our results are consistent with what they found in their study that patients with high TS scores demonstrated a lower ICIs therapy response rate. In their study, the immunotherapy response of patients was predicted based on tumor immune dysfunction and exclusion (TIDE), while in our study, data from a phase II IMvigor210 study was used to analyze immunotherapy response, which made our conclusion more reliable and robust. The comparison of our study and other studies exploring the TS-related prognostic value was summarized in Supplementary Table S4 [40–43].

It has been reported that CSCs are one of the main reasons of the discrepancies in the treatment outcomes of patients with BLCA receiving chemotherapy [44]. Since the standard first-line treatment strategy for BLCA is chemotherapy including MVAC regimen (methotrexate, vinblastine, doxorubicin and cisplatin), GC regimen (gemcitabine and cisplatin), and CMV regimen (cisplatin, methotrexate and vinblastine) [45], the predicted IC50 value of these drugs was compared to assess the association between chemotherapy sensitivity and the CSCs-related risk score. High-risk patients have a higher chemotherapy sensitivity to cisplatin, gemcitabine, vinblastine and doxorubicin, which all inhibit cancer cells' growth by affecting DNA replication. We suspected that this was due to a greater genetic instability in high-risk group. To test the hypothesis, we compared HRD scores and the frequencies of gene mutation between two risk groups. HRD scores were significantly higher in high-risk group, while there was no significant difference in the frequency of BRCA mutation between the two groups. Besides, ERCC2 and CHEK2 with the highest mutation frequency in both risk groups, are related to the pathways of response and repair damage to DNA [46–48]. These results indicated that the difference of HRD score between risk groups was probably due to multiple genes related to DNA repair response, rather than BRCA1/2. The above results suggested that CSCs-related risk score might imply genome instabilities, thus could indicate the chemotherapy benefit for patients with different risk stratification.

Moreover, considering the survival advantage of immunotherapy, we analyzed the prognostic value of the risk score for ICIs therapy by assigning patients in the IMvigor210 cohort. As expected, results showed that low-risk patients had longer OS and increased response rate when treated with ICIs. Taken together, the risk stratification model based on CSCs-related genes had potential predictive value for chemotherapy and immunotherapy. The relationship between the risk score and immune cells infiltrated in TME was further studied. Researchers have proved that patients who had high numbers of infiltrating macrophages M0, M1, M2 and neutrophils obtained a significantly poor prognosis [17,49], whereas activated dendritic cells, resting dendritic cells, activated natural killer cells, T follicular helper cells, and CD8$^+$ T cells are often recognized as anti-tumor immune cells indicating favorable prognosis [50,51]. These conclusions supported our findings that activated dendritic cells, resting dendritic cells, activated natural killer cells, plasma cells, T follicular helper cells and CD8$^+$ T cells were higher in low-risk group. Surprisingly, Tregs were higher in low-risk group which was in contradiction to that Treg cells are generally referred as pro-tumorigenic immune cell [52]. However, recently, a study demonstrated that Tregs displayed extreme heterogeneity in bladder TME and had varied contribution to anti-tumor responses [53], implying that the proportion of TILs should be carefully regarded as prognostic and predictive indicators because of the plasticity of immune cells. To sum up, the infiltration of more anti-tumor immune cells formed a relative immune-activated TME in the BLCA patients with low-risk, which probably contributed to the better prognosis observed in low-risk group.

Our study has several limitations that should not be ignored. First, The CSCs-related genes in BLCA have not been fully studied, which may limit the performance of our risk stratification model due to incomplete CSCs gene list included in this study. Besides, the underlying mechanism of bladder CSCs interacting with immune infiltrated cells, stromal cells, and other non-CSC cells needs to be further studied to better interpret the results of our study.

In conclusion, we established a risk classifier based on 16 CSCs-related genes for predicting survival in patients with BLCA. The low-risk patients had better prognosis compared to high-risk patients, and the survival benefit might be derived from the relative immune-activated TME. This observation was further confirmed in patients with ICI therapy that low-risk patients were more responsive to anti-PD-L1 regimen. In contrast to immunotherapy, high-risk patients were more likely benefit from chemotherapy. The CSCs-related risk model has both prognostic value and potential predictive value for therapeutic efficacy, which brings us closer to understanding the important role of CSCs in BLCA and may provide guidance for clinical treatment decision-making and patient management. The clinical utility of the CSCs-related risk classifier warrants further studies.

Author Contributions

Conceptualization: Xia Chen, Yuanyuan Yin, Yanping song, and Junyong Dai; Methodology: Xia Chen, and Yuanyuan Yin; Formal analysis: Xia Chen, Yuanyuan Yin, Yuming He, Fanqi Meng, Jing Zhao, and Fang Liu; Visualization: Yuming He, and Fanqi Meng; Writing - Original Draft: Xia Chen, Yuanyuan Yin, Yuming He, Fanqi Meng, Jing Zhao, Fang Liu, and Yanping song; Writing - Review & Editing: Yu Xu, Guoqiang Wang, Xin Zhu, Sijia Ma, Huafei Lu, Shangli Cai, and Junyong Dai; Supervision: Junyong Dai; Funding acquisition: Junyong Dai.

Ethics approval and consent to participate

Not applicable. This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not applicable.

Data availability

All data used are from public databases including TCGA database (https://tcga-data.nci.nih.gov/tcga/), UCSC XENA database (https://xenabrowser.net/datapages/), GEO database (https://www.ncbi.nlm.nih.gov/geo/), IMvigor210 cohort, MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/), and GDSC database (https://www.cancerrxgene.org/), and the sources have all been noted in this paper.

Funding information

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e24858.

References

[1] H. Sung, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J Clin 71 (3) (May 2021) 209–249, https://doi.org/10.3322/caac.21660.
[2] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, CA Cancer J Clin 73 (1) (Jan 2023) 17–48, https://doi.org/10.3322/caac.21763.
[3] A.T. Lenis, P.M. Lec, K. Chamie, M.D. Mshs, Bladder cancer: a review, JAMA 324 (19) (Nov 17 2020) 1980–1991, https://doi.org/10.1001/jama.2020.17598.
[4] N. Aghaalikhani, N. Rashtchizadeh, P. Shadpour, A. Allameh, M. Mahmoodi, Cancer stem cells as a therapeutic target in bladder cancer, J. Cell. Physiol. 234 (4) (Apr 2019) 3197–3206, https://doi.org/10.1002/jcp.26916.
[5] E. Batlle, H. Clevers, Cancer stem cells revisited, Nat. Med. 23 (10) (Oct 6 2017) 1124–1134, https://doi.org/10.1038/nm.4409.
[6] L.T.H. Phi, et al., Cancer stem cells (CSCs) in drug resistance and their therapeutic implications in cancer treatment, Stem Cell. Int. 2018 (2018) 5416923, https://doi.org/10.1155/2018/5416923.
[7] M. Al-Hajj, M.S. Wicha, A. Benito-Hernandez, S.J. Morrison, M.F. Clarke, Prospective identification of tumorigenic breast cancer cells, Proc. Natl. Acad. Sci. U.S. A. 100 (7) (Apr 1 2003) 3983–3988, https://doi.org/10.1073/pnas.0530291100.
[8] L. Ricci-Vitiani, et al., Identification and expansion of human colon-cancer-initiating cells, Nature 445 (7123) (Jan 4 2007) 111–115, https://doi.org/10.1038/nature05384.
[9] S. Zhang, et al., Identification and characterization of ovarian cancer-initiating cells from primary human tumors, Cancer Res. 68 (11) (Jun 1 2008) 4311–4320, https://doi.org/10.1158/0008-5472.CAN-08-0364.
[10] A. Eramo, et al., Identification and expansion of the tumorigenic lung cancer stem cell population, Cell Death Differ. 15 (3) (Mar 2008) 504–514, https://doi.org/10.1038/sj.cdd.4402283.
[11] A. Abugomaa, M. Elbadawy, H. Yamawaki, T. Usui, K. Sasaki, Emerging roles of cancer stem cells in bladder cancer progression, tumorigenesis, and resistance to chemotherapy: a potential therapeutic target for bladder cancer, Cells 9 (no. 1) (Jan 17 2020), https://doi.org/10.3390/cells9010235.
[12] A. Miranda, et al., Cancer stemness, intratumoral heterogeneity, and immune response across cancers, Proc. Natl. Acad. Sci. U.S.A. 116 (18) (Apr 30 2019) 9020–9029, https://doi.org/10.1073/pnas.1818210116.
[13] M. Jinushi, et al., Tumor-associated macrophages regulate tumorigenicity and anticancer drug responses of cancer stem/initiating cells, Proc. Natl. Acad. Sci. U. S.A. 108 (30) (Jul 26 2011) 12425–12430, https://doi.org/10.1073/pnas.1106645108.
[14] B. Sainz Jr., E. Carron, M. Vallespinos, H.L. Machado, Cancer stem cells and macrophages: implications in tumor biology and therapeutic strategies, Mediat. Inflamm. 2016 (2016) 9012369, https://doi.org/10.1155/2016/9012369.
[15] M. Ligorio, et al., Stromal microenvironment shapes the intratumoral architecture of pancreatic cancer, Cell 178 (1) (Jun 27 2019) 160–175 e27, https://doi.org/10.1016/j.cell.2019.05.012.
[16] J.A. Clara, C. Monge, Y. Yang, N. Takebe, Targeting signalling pathways and the immune microenvironment of cancer stem cells - a clinical update, Nat. Rev. Clin. Oncol. 17 (4) (Apr 2020) 204–232, https://doi.org/10.1038/s41571-019-0293-2.
[17] P. Allavena, E. Digifico, C. Belgiovine, Macrophages and cancer stem cells: a malevolent alliance, Mol. Med. 27 (1) (Sep 28 2021) 121, https://doi.org/10.1186/s10020-021-00383-3.

[18] D. Bayik, J.D. Lathia, Cancer stem cell-immune cell crosstalk in tumour progression, Nat. Rev. Cancer 21 (8) (Aug 2021) 526–536, https://doi.org/10.1038/s41568-021-00366-w.

[19] Y. Yan, et al., HIF-2alpha promotes conversion to a stem cell phenotype and induces chemoresistance in breast cancer cells by activating Wnt and Notch pathways, J. Exp. Clin. Cancer Res. 37 (1) (Oct 19 2018) 256, https://doi.org/10.1186/s13046-018-0925-x.

[20] J.M. Hsu, et al., STT3-dependent PD-L1 accumulation on cancer stem cells promotes immune evasion, Nat. Commun. 9 (1) (May 15 2018) 1908, https://doi.org/10.1038/s41467-018-04313-6.

[21] K.S. Chan, et al., Identification, molecular characterization, clinical prognosis, and therapeutic targeting of human bladder tumor-initiating cells, Proc. Natl. Acad. Sci. U.S.A. 106 (33) (Aug 18 2009) 14016–14021, https://doi.org/10.1073/pnas.0906549106.

[22] M.J. Goldman, et al., Visualizing and interpreting cancer genomics data via the Xena platform, Nat. Biotechnol. 38 (6) (Jun 2020) 675–678, https://doi.org/10.1038/s41587-020-0546-8.

[23] D. Lindgren, et al., Integrated genomic and gene expression profiling identifies two major genomic circuits in urothelial carcinoma, PLoS One 7 (6) (2012) e38863, https://doi.org/10.1371/journal.pone.0038863.

[24] G. Sjodahl, et al., A molecular taxonomy for urothelial carcinoma, Clin. Cancer Res. : an official journal of the American Association for Cancer Research 18 (12) (Jun 15 2012) 3377–3386, https://doi.org/10.1158/1078-0432.CCR-12-0077-T.

[25] S. Mariathasan, et al., TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells, Nature 554 (7693) (Feb 22 2018) 544–548, https://doi.org/10.1038/nature25501.

[26] N.P. Palmer, P.R. Schmid, B. Berger, I.S. Kohane, A gene expression profile of stem cell pluripotentiality and differentiation is conserved across diverse solid and hematopoietic cancers, Genome Biol. 13 (8) (Aug 21 2012) R71, https://doi.org/10.1186/gb-2012-13-8-r71.

[27] M.E. Ritchie, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (Apr 20 2015) e47, https://doi.org/10.1093/nar/gkv007.

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

[29] A. Liberzon, C. Birger, H. Thorvaldsdottir, M. Ghandi, J.P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst 1 (6) (Dec 23 2015) 417–425, https://doi.org/10.1016/j.cels.2015.12.004.

[30] A. Subramanian, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U.S.A. 102 (43) (Oct 25 2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.

[31] A.M. Newman, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (May 2015) 453–457, https://doi.org/10.1038/nmeth.3337.

[32] B. Chen, M.S. Khodadoust, C.L. Liu, A.M. Newman, A.A. Alizadeh, Profiling tumor infiltrating immune cells with CIBERSORT, Methods Mol. Biol. 1711 (2018) 243–259, https://doi.org/10.1007/978-1-4939-7493-1_12.

[33] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (5) (Mar 4 2011) 646–674, https://doi.org/10.1016/j.cell.2011.02.013.

[34] K. Yoshihara, et al., Inferring tumour purity and stromal and immune cell admixture from expression data, Nat. Commun. 4 (2013) 2612, https://doi.org/10.1038/ncomms3612.

[35] Z. Liu, et al., Tumor stroma-infiltrating mast cells predict prognosis and adjuvant chemotherapeutic benefits in patients with muscle invasive bladder cancer, OncoImmunology 7 (9) (2018) e1474317, https://doi.org/10.1080/2162402X.2018.1474317.

[36] T. Cokelaer, et al., GDSCTools for mining pharmacogenomic interactions in cancer, Bioinformatics 34 (7) (Apr 1 2018) 1226–1228, https://doi.org/10.1093/bioinformatics/btx744.

[37] 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) e107468, https://doi.org/10.1371/journal.pone.0107468.

[38] J. Borcsok, et al., Detection of molecular signatures of homologous recombination deficiency in bladder cancer, Clin. Cancer Res. : an official journal of the American Association for Cancer Research 27 (13) (Jul 1 2021) 3734–3743, https://doi.org/10.1158/1078-0432.CCR-20-5037.

[39] Y. Zhang, et al., Tumor stemness score to estimate epithelial-to-mesenchymal transition (EMT) and cancer stem cells (CSCs) characterization and to predict the prognosis and immunotherapy response in bladder urothelial carcinoma, Stem Cell Res. Ther. 14 (1) (Feb 1 2023) 15, https://doi.org/10.1186/s13287-023-03239-1.

[40] D. Conconi, et al., Analysis of copy number alterations in bladder cancer stem cells revealed a prognostic role of LRP1B, World J. Urol. 40 (9) (Sep 2022) 2267–2273, https://doi.org/10.1007/s00345-022-04093-1.

[41] Y. Kang, et al., Identification and validation of the prognostic stemness biomarkers in bladder cancer bone metastasis, Front. Oncol. 11 (2021) 641184, https://doi.org/10.3389/fonc.2021.641184.

[42] M. Zhang, et al., Development and validation of cancer-associated fibroblasts-related gene landscape in prognosis and immune microenvironment of bladder cancer, Front. Oncol. 13 (2023) 1174252, https://doi.org/10.3389/fonc.2023.1174252.

[43] Z.H. Feng, et al., m6A-immune-related lncRNA prognostic signature for predicting immune landscape and prognosis of bladder cancer, J. Transl. Med. 20 (1) (Oct 29 2022) 492, https://doi.org/10.1186/s12967-022-03711-1.

[44] Y. Hu, L. Fu, Targeting cancer stem cells: a new therapy to cure cancer patients, Am. J. Cancer Res. 2 (3) (2012) 340–356 [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22679565.

[45] M.A. Knowles, C.D. Hurst, Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity, Nat. Rev. Cancer 15 (1) (Jan 2015) 25–41, https://doi.org/10.1038/nrc3817.

[46] D. Bernard-Gallon, et al., DNA repair gene ERCC2 polymorphisms and associations with breast and ovarian cancer risk, Mol. Cancer 7 (May 2 2008) 36, https://doi.org/10.1186/1476-4598-7-36.

[47] J. Boyle, et al., Persistence of repair proteins at unrepaired DNA damage distinguishes diseases with ERCC2 (XPD) mutations: cancer-prone xeroderma pigmentosum vs. non-cancer-prone trichothiodystrophy, Hum. Mutat. 29 (10) (Oct 2008) 1194–1208, https://doi.org/10.1002/humu.20768.

[48] W. Roeb, J. Higgins, M.C. King, Response to DNA damage of CHEK2 missense mutations in familial breast cancer, Hum. Mol. Genet. 21 (12) (Jun 15 2012) 2738–2744, https://doi.org/10.1093/hmg/dds101.

[49] P.L. Crispen, S. Kusmartsev, Mechanisms of immune evasion in bladder cancer, Cancer Immunol. Immunother. 69 (1) (Jan 2020) 3–14, https://doi.org/10.1007/s00262-019-02443-4.

[50] P. Charoentong, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (1) (Jan 3 2017) 248–262, https://doi.org/10.1016/j.celrep.2016.12.019.

[51] A.K. Schneider, M.F. Chevalier, L. Derre, The multifaceted immune regulation of bladder cancer, Nat. Rev. Urol. 16 (10) (Oct 2019) 613–630, https://doi.org/10.1038/s41585-019-0226-y.

[52] D.E. Speiser, P.C. Ho, G. Verdeil, Regulatory circuits of T cell function in cancer, Nat. Rev. Immunol. 16 (10) (Oct 2016) 599–611, https://doi.org/10.1038/nri.2016.80.

[53] D.Y. Oh, et al., Intratumoral CD4(+) T cells mediate anti-tumor cytotoxicity in human bladder cancer, Cell 181 (7) (Jun 25 2020) 1612–1625, https://doi.org/10.1016/j.cell.2020.05.017, e13.