

# A zinc finger protein gene signature enables bladder cancer treatment stratification

Jiandong Zhang<sup>1,2,\*</sup>, Chen Zhang<sup>3,\*</sup>, Peng Cao<sup>1</sup>, Xiang Zheng<sup>1</sup>, Baozhong Yu<sup>1</sup>, Haoyuan Cao<sup>1</sup>, Zihao Gao<sup>1</sup>, Feilong Zhang<sup>1</sup>, Jiyuan Wu<sup>1</sup>, Huawei Cao<sup>1</sup>, Changzhen Hao<sup>1</sup>, Zejia Sun<sup>1</sup>, Wei Wang<sup>1</sup>

<sup>1</sup>Beijing Chaoyang Hospital Affiliated Capital Medical University, Beijing 100020, China

<sup>2</sup>Shanxi Bethune Hospital Affiliated Shanxi Academy of Medical Sciences, Taiyuan 030032, China

<sup>3</sup>State Key Laboratory of Membrane Biology, Institute of Zoology, Chinese Academy of Sciences and University of Chinese Academy of Sciences, Beijing 100101, China

\*Equal contribution

**Correspondence to:** Wei Wang; email: [weiwang0920@163.com](mailto:weiwang0920@163.com), <https://orcid.org/0000-0003-2642-3338>

**Keywords:** bladder cancer, zinc finger proteins, prognosis, treatment

**Received:** December 23, 2020

**Accepted:** March 31, 2021

**Published:** May 7, 2021

**Copyright:** © 2021 Zhang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

Bladder cancer (BC) is a commonly occurring malignant tumor affecting the urinary tract. Zinc finger proteins (ZNFs) constitute the largest transcription factor family in the human genome and are therefore attractive biomarker candidates for BC prognosis. In this study, we profiled the expression of ZNFs in The Cancer Genome Atlas (TCGA) BC cohort and developed a novel prognostic signature based on 7 ZNF-coding genes. After external validation of the model in the GSE48276 dataset, we integrated the 7-ZNF-gene signature with patient clinicopathological data to construct a nomogram that forecasted 1-, 2-, and 3-year OS with good predictive accuracy. We then accessed The Genomics of Drug Sensitivity in Cancer database to predict the therapeutic drug responses of signature-defined high- and low-risk BC patients in the TCGA cohort. Greater sensitivity to chemotherapy was revealed in the low-risk group. Finally, we conducted gene set enrichment analysis of the signature genes and established, by applying the ESTIMATE algorithm, distinct correlations between the two risk groups and the presence of stromal and immune cell types in the tumor microenvironment. By allowing effective risk stratification of BC patients, our novel ZNF gene signature may enable tailoring more intensive treatment for high-risk patients.

## INTRODUCTION

Bladder cancer (BC) is one of the most common forms of malignant tumors affecting the urinary tract. Although there are many methods to diagnose and remove solid neoplasms, the high recurrence rate and mortality of BC are still not well controlled [1]. Therefore, there is an urgent need for effective diagnostic and prognostic biomarkers for BC. Although there are several treatment options, such as chemotherapy, for controlling the disease, there is still no effective guidance on how to choose chemotherapeutic drugs [2]. For this reason, strategies to individualize treatment are being explored actively. However, studies aimed at finding a reliable

measure to stratify treatment for patients with BC have not yet yielded the desired results [3].

Zinc finger proteins (ZNFs) represent the largest transcription factor family in the human genome. These proteins are involved in diverse biological processes, including differentiation, development, metabolism, apoptosis, autophagy, and stemness maintenance. Over the last few decades, increasing evidence has supported key roles for ZNFs in cancer onset, progression, and metastasis [4]. For example, Chen et al. reported that ZNF830 promotes homologous recombination repair and survival of cancer cells in response to DNA damage, and that high expression of ZNF830 is associated with poor

survival in lung and gastric cancer patients by mediating resistance to chemoradiotherapy [5]. Further examples include ZNF281 [6], which acts as an oncogene, ZNF185 [7] and ZNF750 [8], which serve as tumor suppressor genes, and ZEB1 [9] and ZBP89 [4], which appear to act as either oncogenes or tumor suppressor genes in different contexts. Interestingly, a recent study indicated that a single gene, encoding the zinc finger protein SPOP, can predict the prognosis of several tumors and guide stratification therapy [10].

Identification of biomarker signatures represents a valuable approach to mine the wealth of information contained within biological samples [11, 12]. Since the significance of ZNFs in BC diagnosis, treatment, and prognosis remains unclear developing a biomarker signature based on ZNF protein genes might be helpful to guide decision-making to select appropriate treatments and to predict prognosis for BC patients. Moreover, the prognostic performance of the signature can be enhanced by constructing nomograms that integrate, along with the corresponding expression profiles, the patient's clinical variables, e.g. tumor status [13]. Therefore, the goal of this study was to construct a ZNF gene-based signature to stratify patients, predict individual prognosis, and guide BC treatment. We also

developed and tested a nomogram based on the ZNF-gene signature and clinical variables, assessed the signature's association with stromal and immune cells in the tumor microenvironment and predicted, based on the expression of signature genes in low- and high-risk patients, their response to common chemotherapy agents. Our findings shed light on the potential contribution of ZNFs to the pathogenesis of BC and may inform clinical practice to guide individualized treatment. A flow chart depicting the analyses performed in this study is shown in Figure 1.

## RESULTS

### Identification of differentially expressed ZNF-coding genes in BC

To determine the expression patterns of ZNF genes in BC, expression levels of 1818 human ZNF protein-coding genes retrieved from the UniProt database were evaluated in the transcriptional profiles of 403 muscle-invasive BC patients and 19 normal bladder controls, available in the TCGA (Figure 2A). A total of 319 upregulated and 139 downregulated ZNF-coding, differentially expressed genes (DEGs) were thus identified (Figure 2B and Supplementary Table 1).

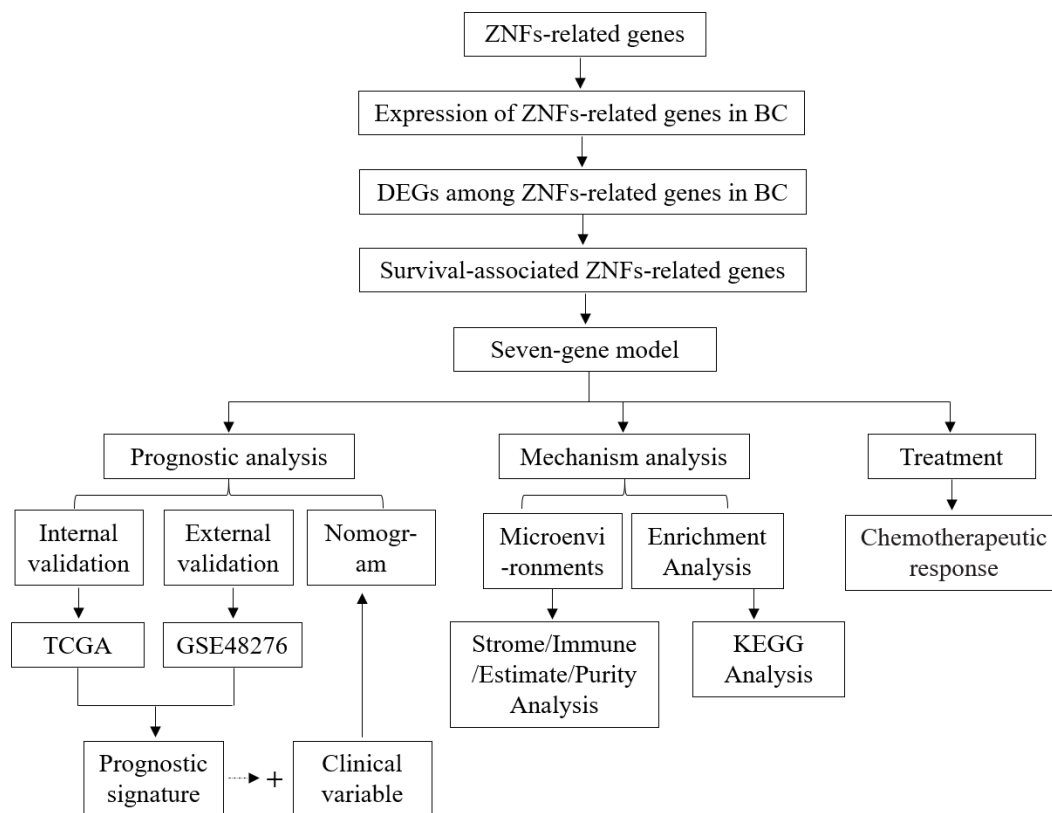
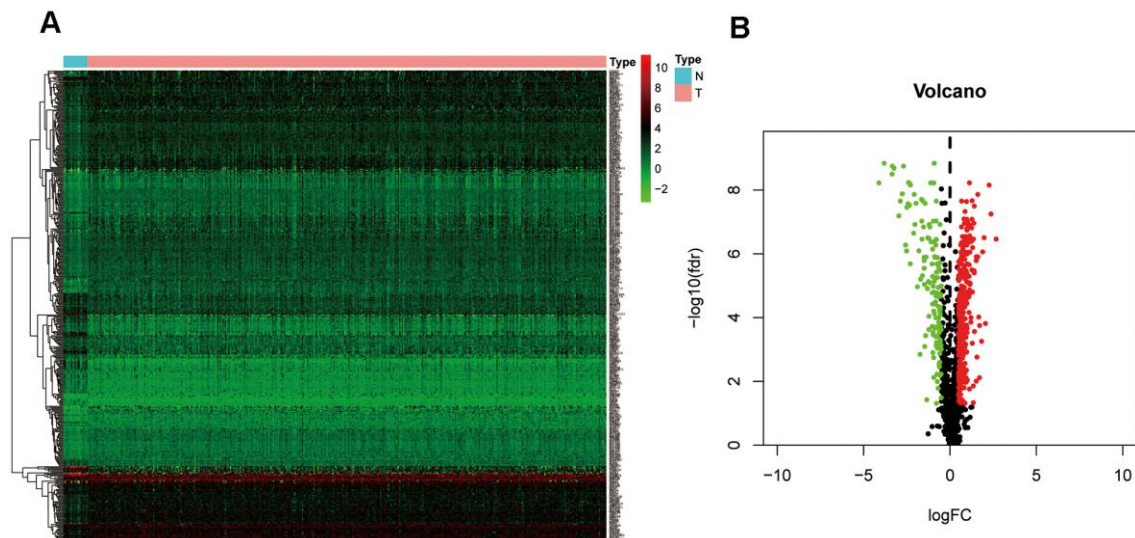


Figure 1. Flow chart of the study.

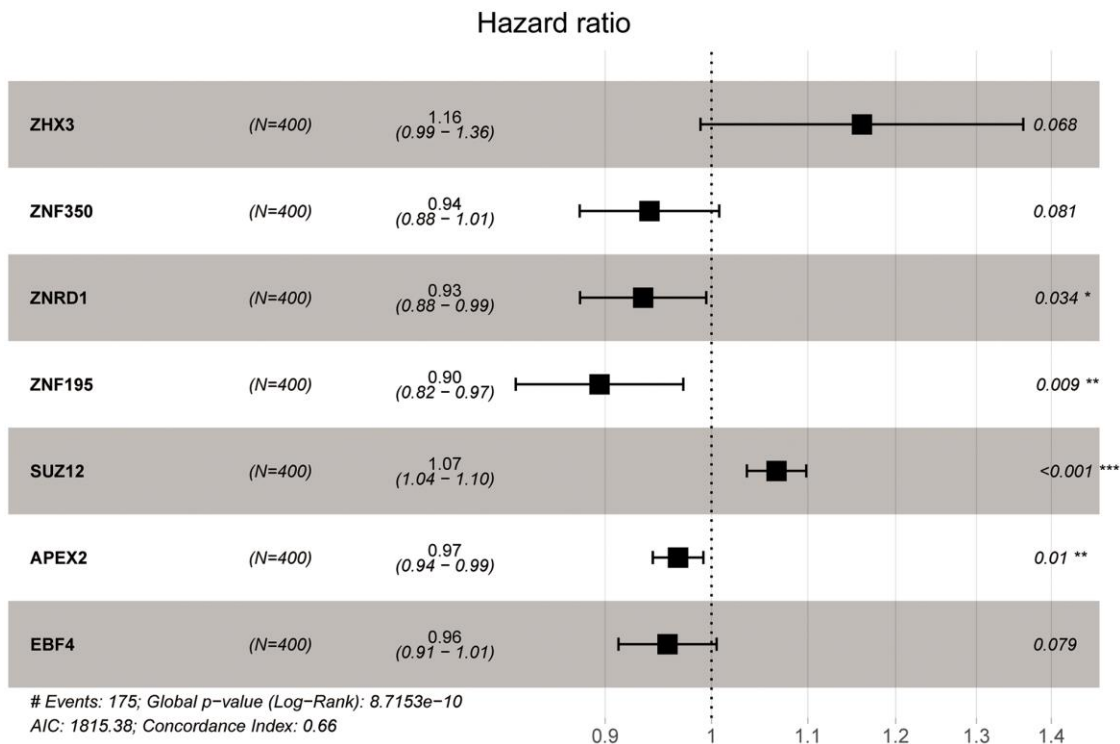
## Development of a ZNF gene-based prognostic signature for BC

A prognostic signature was then established by first identifying survival-associated ZNF-coding DEGs in the

TCGA-BC cohort using univariate Cox regression. After screening out significant DEGs using LASSO regression and multivariate Cox regression, 7 BC-specific, prognostic ZNF genes were selected (Figure 3). After extracting the corresponding coefficient values (Table 1),



**Figure 2. Identification of differentially expressed ZNF genes in the TCGA-BC cohort.** (A) Heatmap depicting the expression levels of ZNF genes in tumor (T) and normal (N) samples. (B) Volcano plot representation of differentially expressed ZNF genes in the TCGA-BC cohort.



**Figure 3. Characteristics of BC-specific ZNF genes.** Forest plot showing hazard ratios (HRs) with 95% confidence interval (95% CI) of prognostic ZNF genes in BC based on multivariate Cox regression results.

**Table 1. The coefficient of selected genes.**

Gene	coefficient	HR	HR.95Low	HR.95High	P value
ZHX3	0.148896	1.160553	0.989119	1.361699	0.067879
ZNF350	-0.0615	0.940356	0.877571	1.007633	0.08111
ZNRD1	-0.06775	0.934497	0.877895	0.994749	0.033577
ZNF195	-0.11088	0.895049	0.823812	0.972446	0.008786
SUZ12	0.064369	1.066486	1.035628	1.098264	1.73E-05
APEX2	-0.03312	0.967418	0.943455	0.99199	0.009642
EBF4	-0.04355	0.957387	0.912023	1.005007	0.078698

individual risk scores were estimated based on coefficient-weighted expression levels of the selected genes.

Following exclusion of three BC patients with no follow-up information in the TCGA cohort, the patients were stratified into a high-risk group (n=200) and a low-risk group (n=200) according to the median cut-off value (Figure 4A). As shown in Figure 4B, patients with high risk had a higher probability of early death than those with low risk. Consistently, the heatmap of expression profiles in the TCGA dataset showed distinct differences between groups (Figure 4C). Moreover, survival analysis indicated that patients in the high-risk group had a significantly worse OS than their low-risk counterparts (Figure 4D,  $P < 0.001$ ). The ZNF gene-based prognostic signature showed good performance, with AUCs of 0.654 and 0.664 at 3- and 5-year follow-up respectively (Figure 4E). Furthermore, the signature was significantly predictive of survival on univariable (Figure 3) and multivariable (Figure 3) analyses that included risk score, age, gender, tumor stage, and TNM ( $P < 0.001$  for all variables, except gender and tumor stage).

### External validation of the prognostic signature

To verify the predictive value of our model, samples from patients in the GSE48276 cohort served as external testing data after categorizing them into high- or low-risk groups using the median value calculated with the same formula and cutoff value applied before (Figure 5A). The expression profiles corresponding to the signature genes are shown in Figure 5C. Similar to the results obtained in the original TCGA cohort, patients in the high-risk group were more likely to die earlier (Figure 5B) and had a reduced survival time compared with those in the low-risk group (Figure 5D;  $P = 0.001$ ). Moreover, the prognostic performance of the ZNF gene-based prognostic signature showed acceptable discrimination, with AUCs of 0.723 and 0.834 at 3- and 5-year follow-up, respectively (Figure 5E).

### Predictive accuracy of a ZNF gene-based nomogram

After asserting the prognostic reliability of the 7-ZNF-gene signature on BC outcomes, we used it along with patient clinicopathological data to construct a nomogram to forecast 1-, 2-, and 3-year OS (Figure 6A). The calibration plot of the nomogram indicated optimal predictive accuracy, with a close overlap between predicted and actual survival rate (Figure 6B).

### The ZNF-gene signature predicts differences in BC microenvironment

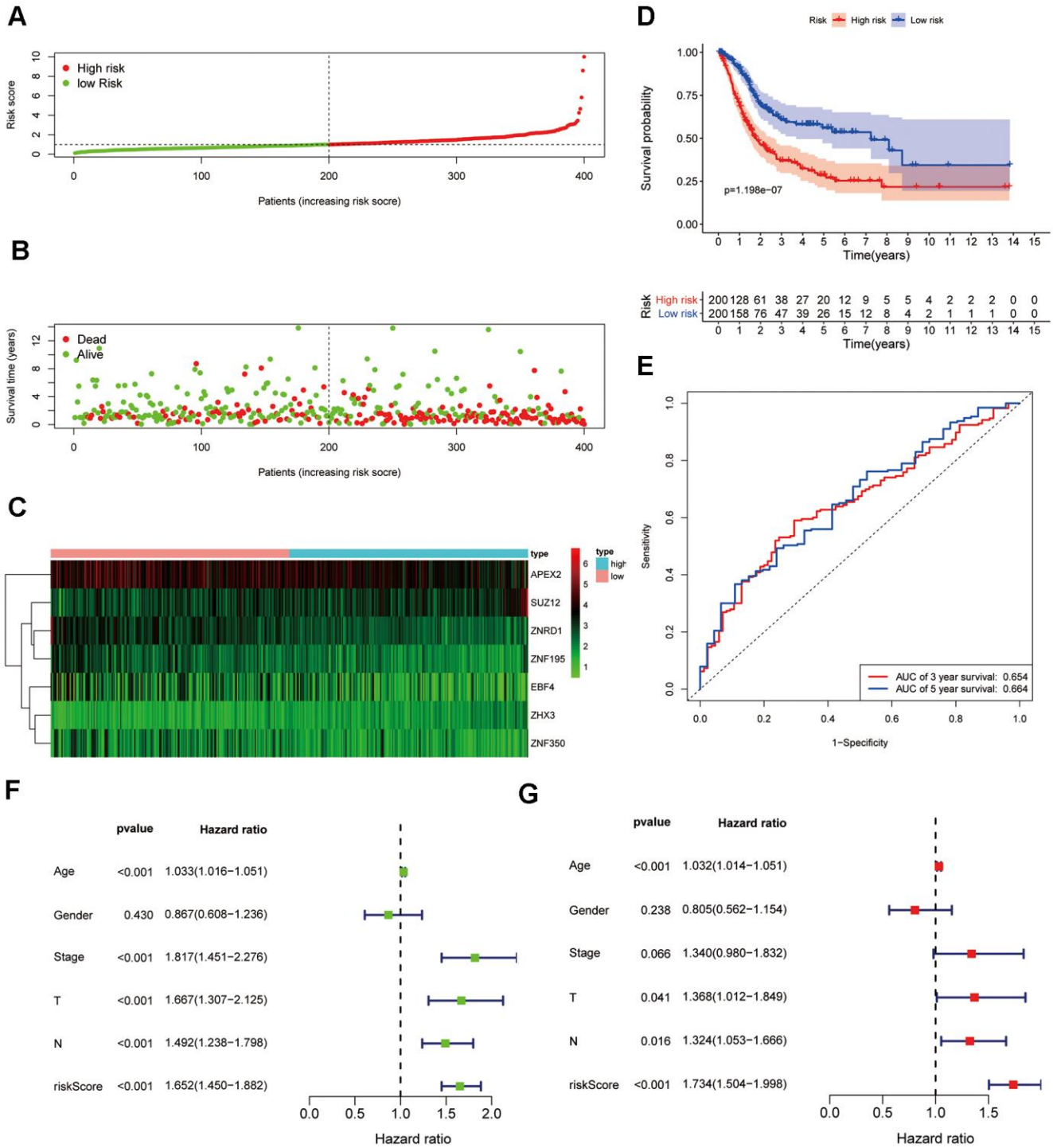
To assess whether the ZNF-gene signature can help distinguish differences in the tumor microenvironment of BC, we employed the ESTIMATE tool to compare gene expression signatures of stromal and immune cells among risk groups. The stromal score ranged from -788.35 to -267.74 (Figure 7A), the immune score ranged from 379.45 to 715.66 (Figure 7B), and the ESTIMATE score ranged from -408.9 to 447.92 (Figure 7C), with statistically significant differences ( $P < 0.001$  for all scores) detected between the high-risk and low-risk groups. Meanwhile, a lower tumor purity, distributed between 0.76 and 0.83 (Figure 7D) was observed in the high-risk group ( $P < 0.001$ ).

### Correlation between tumor-infiltrating immune cells and the ZNF-gene signature

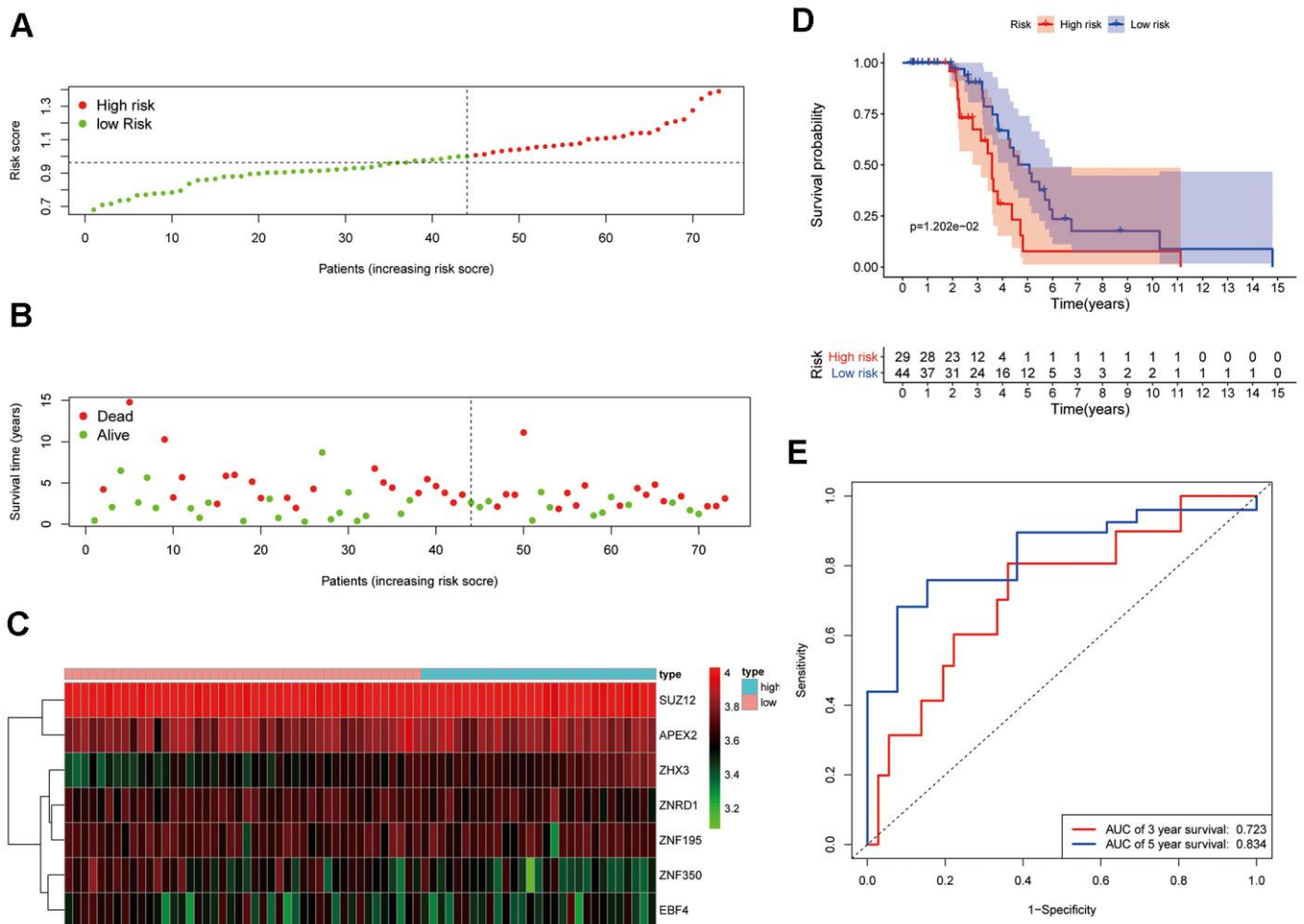
To further investigate the relationship between the ZNF gene signature's risk score and the tumor's immune status, the enrichment scores of diverse immune cell subpopulations, and their related functions or pathways, were quantified in the TCGA-BC cohort using ssGSEA. The results showed that scores for cell types related to the antigen presentation process, including dendritic cells (DCs), activated DCs, plasmacytoid DCs, tumor-infiltrating lymphocytes (TILs), B cells, macrophages, mast cells, neutrophils, CD8 T cells, T helper (Th) cells, Th1 cells, T follicular helper cells (Tfh), and regulatory T cells (TRegs) were significantly different between risk groups (adjusted  $P < 0.05$  for all; Figure 8A). On

KEGG analysis, cytokine-cytokine receptor interaction showed a higher score in the high-risk group (adjusted  $P < 0.05$ ; Figure 8B). Moreover, the high-risk group showed also enrichment in the activity of checkpoint

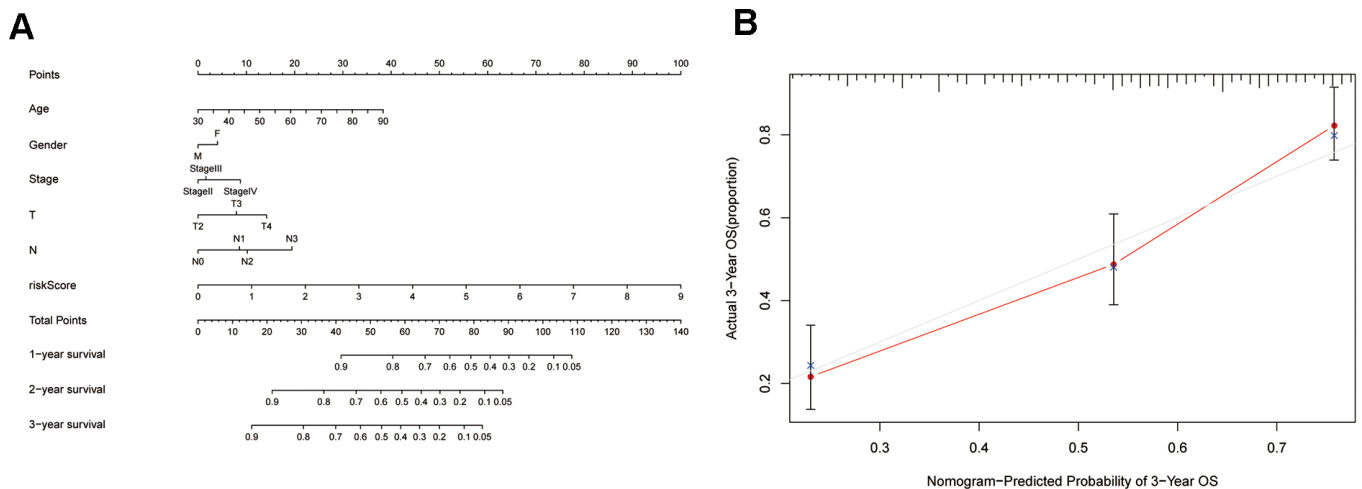
molecules and higher scores for macrophages or Tregs, whereas scores for type II IFN response, type I IFN response, and NK cells were instead lower (adjusted  $P < 0.05$ , Figure 8A, 8B).



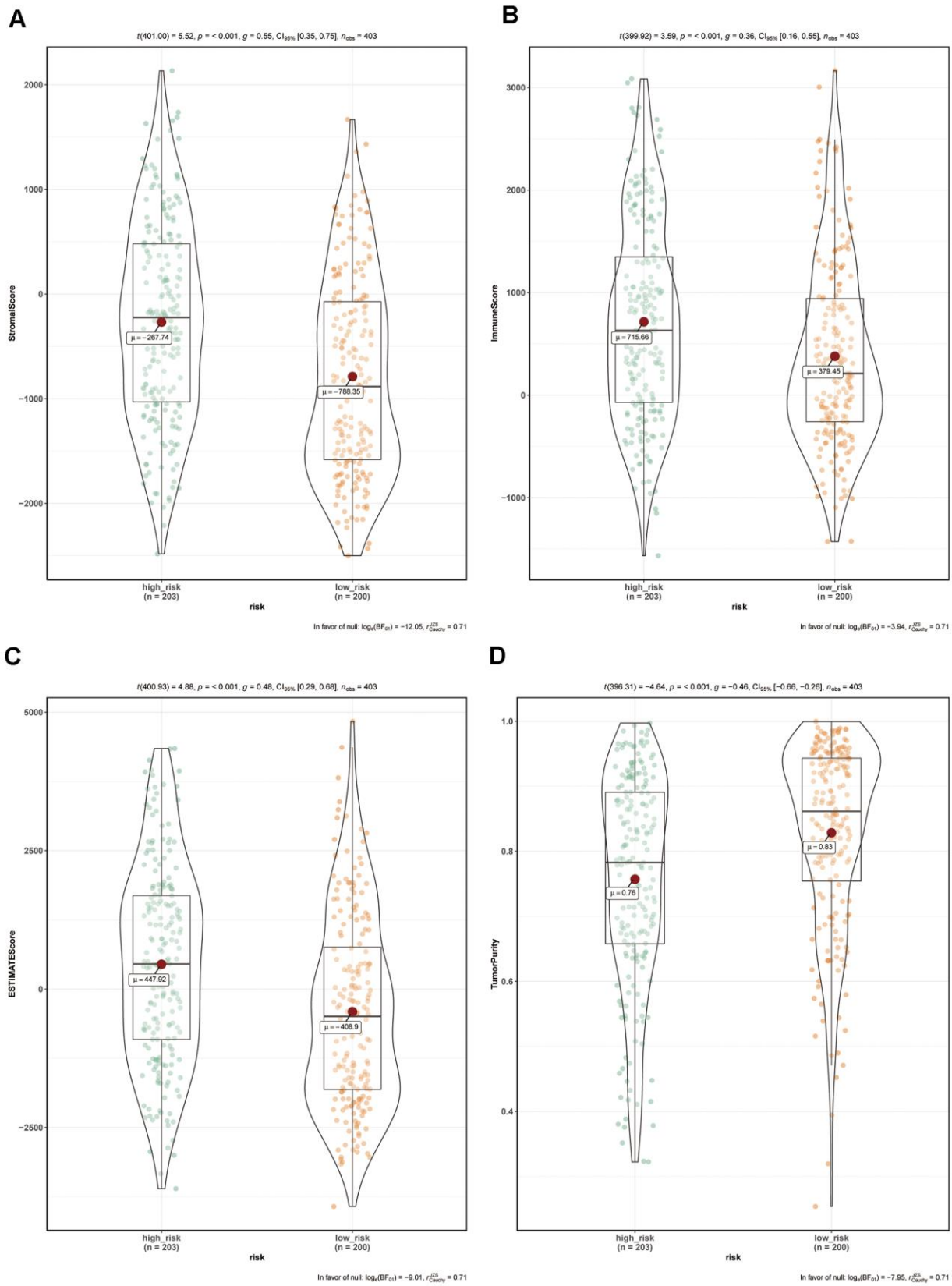
**Figure 4. Development of a prognostic signature for BC based on 7 ZNF genes.** (A) Distribution and median value of the risk scores in the TCGA cohort. (B) Survival status of patients in the different risk groups. (C) Heatmap of the expression profiles of the 7 ZNF genes included in the prognostic signature. (D) Time-dependent ROC curve of the 7-gene signature. (E) Survival analysis of the signature-defined risk groups. (F) Univariable and (G) multivariable analyses adjusting for risk score, age, gender, tumor stage, and TNM.



**Figure 5. Validation of the 7-ZNF-gene prognostic signature in a GEO dataset.** (A) Distribution and median value of the risk scores in the GSE48276 cohort. (B) Survival status of low-risk and high-risk patients. (C) Heatmap of the expression profiles of the 7-ZNF-gene signature. (D) Time-dependent ROC curve of the prognostic signature. (E) Survival analysis of signature-defined risk groups.



**Figure 6. Construction of a nomogram based on the 7-ZNF-gene signature.** (A) Nomogram based on the ZNF-gene signature and clinical information. (B) Decision curve analysis evaluating the clinical utility of the nomogram at 3-year survival.



**Figure 7. Comparison of tumor microenvironment composition between risk groups in the TCGA-BC cohort. (A)** Comparison of stromal scores between risk groups ( $P < 0.001$ ). **(B)** Comparison of immune scores between risk groups ( $P < 0.001$ ). **(C)** Comparison of ESTIMATE scores between risk groups ( $P < 0.001$ ). **(D)** Comparison of tumor purity between risk groups ( $P < 0.001$ ).

## The ZNF-gene signature predicts chemotherapy response in BC

Considering that chemotherapy is still the most effective adjuvant measure to treat BC, we accessed the Genomics of Drug Sensitivity in Cancer (GDSC) database to estimate the response of low-risk and high-risk BC patients to commonly used drugs. The correlation between risk groups and IC50 values for 138 chemotherapeutic agents was visualized using scatterplots. We found significant discrimination between groups in the estimated IC50 values of 28 drugs (Figure 9,  $P < 0.05$  for all). Hence, we concluded that the low-risk group may be more sensitive to common chemotherapies during clinical treatment.

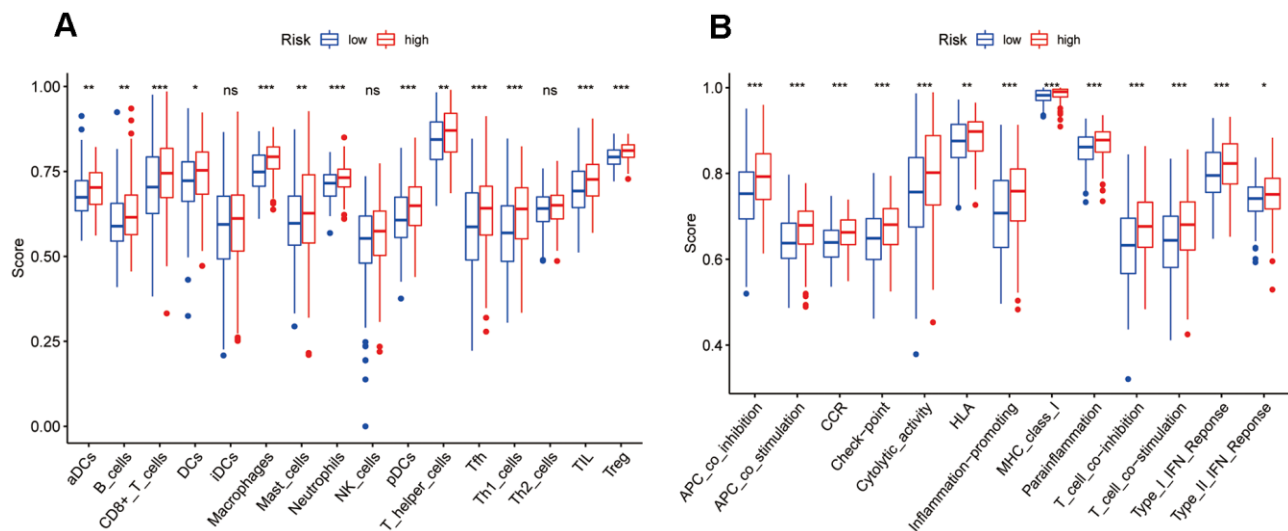
## Association of the ZNF signature genes with BC progression

To gain insight into the functions of the 7 ZNF protein-coding genes included in our signature, we performed KEGG enrichment analysis based on GSEA enrichment scores. The results indicated that the expression patterns that conformed to the high-risk group were enriched in KEGG terms related to tumor progression, such as extracellular matrix (ECM)-receptor interaction, adherens junction, chemokine signaling pathway, and gap junction (Figure 10,  $P < 0.05$  for all). Interestingly, our ZNF-gene signature was closely correlated with other malignancies, for instance melanoma, pancreatic cancer, and glioma. These results suggest that the ZNF protein genes comprising our BC signature may also drive the onset or progression of other types of cancers.

## DISCUSSION

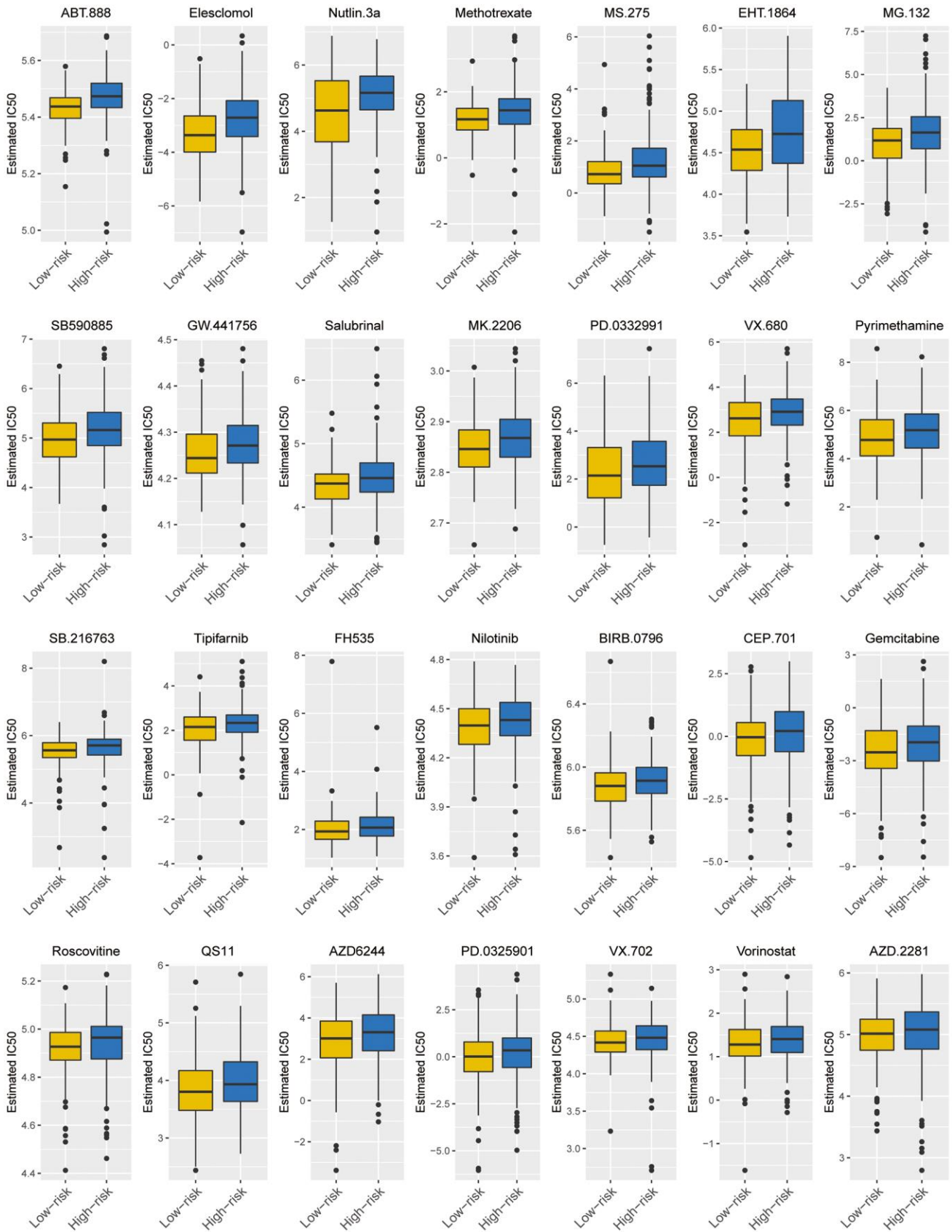
BC, a common urinary malignancy, is associated with poor prognosis in advanced stages. Thus, identification of novel biomarkers will help assess prognosis, screen out patients in need of systemic therapy, and guide individual treatment. ZNFs represent highly promising biomarkers for BC. First, they are one of the most abundant groups of proteins and have a wide range of molecular functions. Second, ZNFs are involved in tumorigenesis, cancer progression, and metastasis formation. Third, ZNFs may act as oncogenes or tumor suppressors and thus serve as valuable prognostic factors. Fourth, ZNF protein genes are closely involved in the oncogenesis of BC. For instance, activation of MDM2, a ZNF domain-containing E3 ubiquitin ligase, leads to ubiquitylation and proteasomal degradation of p53, a major tumor suppressor protein closely involved in the occurrence, progression, and metastasis of BC [14]. Furthermore, ZNFs' genes are involved in telomere maintenance and genome integrity in cancer and aging [15]. Therefore, in this study we conducted bioinformatics analysis to explore whether differential expression of ZNF genes could be exploited to predict BC outcomes and aid risk stratification.

Based on transcriptional data from the TCGA-BC cohort, we identified 7 prognostic ZNF-coding genes that may serve as valuable biomarkers in the clinical setting. Similar to other analyses [16, 17], the prognostic signature based on the 7 ZNF genes categorized BC patients into two subgroups with different survival outcomes. Importantly, the ability of the gene signature



**Figure 8. Comparison of ssGSEA scores between risk groups in the TCGA-BC cohort. (A)** Scores of 16 immune cell types. **(B)** Functions enriched in the 7-ZNF-gene signature. CCR, cytokine-cytokine receptor; ns, not significant; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  (adjusted P values).





**Figure 9. Predicted responses to chemotherapy for risk groups in the TCGA-BC cohort.** Boxplots exhibiting the estimated IC50 values of 28/138 screened drugs for tumors cells from the two risk groups ( $P < 0.05$  for all).

to distinguish high-and low-risk patients, and to estimate OS, was independently validated in the GSE48276 dataset.

The increasing use of nomograms, based on integrated analysis of tumor signatures (gene expression) and patient-specific clinicopathological data, has the potential to drastically improve disease prognostication [13]. Based on the 7-ZNF-gene prognostic signature, we constructed a nomogram that showed high predictive accuracy for OS. Moreover, besides predicting survival outcomes, the ZNF-based prognostic signature also predicted differences in the composition of the tumor microenvironment, determined by differential representation of stromal and immune cells. In addition, by inputting the ZNF-gene signature profiles of the BC risk groups into the GDSC database, we found that patients in the low-risk group may be more sensitive to common chemotherapies during clinical treatment.

Our ZNF protein gene-based signature included 7 genes, i.e., ZHX3, ZNF350, ZNRD1, ZNF195, SUZ12, APEX2, and EBF4. Among these, 4 genes (ZHX3, ZNF350, ZNF195, and SUZ12) have been implicated, as discussed below, in the onset and progression of BC. Two signature genes (ZHX3 and SUZ12) were positively correlated with the occurrence of BC, while a negative correlation was instead determined for the remaining 5 genes (ZNF350, ZNRD1, ZNF195, APEX2, and EBF4). In line with a recent study [18], our results showed that ZHX3

plays an oncogenic role in BC pathogenesis. Interestingly, ZHX3 has also shown to act as a tumor suppressor in renal cell carcinoma [19], breast cancer [20], and liver cancer [21]. A role for SUZ12 overexpression in BC is also supported by previous studies. Fan et al. showed that SUZ12 overexpression promoted BC progression by stimulating colony formation, migration, and invasiveness of BC cells [22]. In turn, Lee et al. reported a gene signature that includes E2F1-EZH2-SUZ12 and shows predictive value for prognosis in BC [23]. All these data strongly indicate that ZHX3 and SUZ12 act as oncogenes in BC, and suggest that a signature based on these two genes may be of significance to guide patients' treatment and improve prognosis.

The 5 protective genes included in our signature were further retrieved and analyzed. Our results showed that ZNF350 expression was associated with reduced BC risk, which is consistent with previous results [24]. Similarly, another report associated high expression levels of ZNF195 with favorable survival in BC [25]. In contrast, the roles of ZNRD1, APEX2, and EBF4 in BC onset and development had not, to our knowledge, been as yet explored. Interestingly, upregulated expression of zinc ribbon domain containing 1 antisense RNA 1 (ZNRD1-AS1), a negative regulator of ZNRD1, was detected in BC [26]. This finding indirectly supports the reduction in ZNRD1 expression detected by our analysis. The protein encoded by APEX2 plays an important role in both

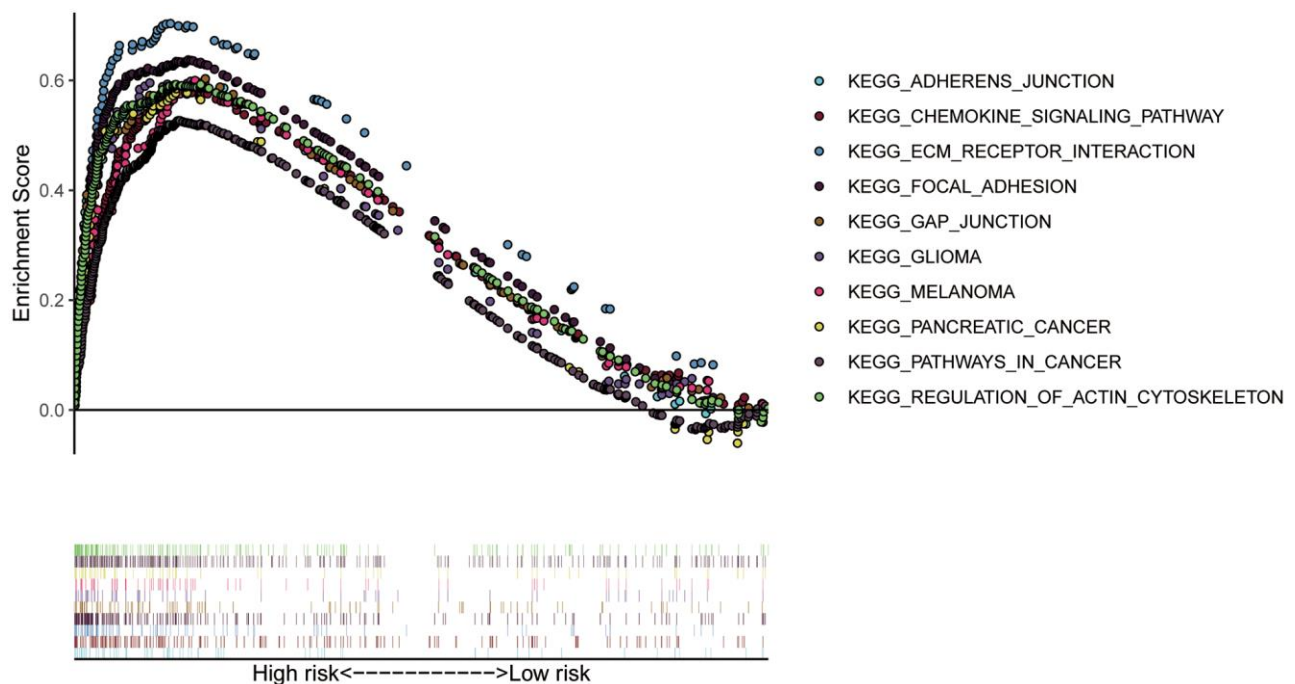


Figure 10. KEGG pathway enrichment analysis of signature genes. ECM, extracellular matrix.

nuclear and mitochondrial base excision repair [27]. Our findings on APEX2 are supported by evidence that this ZNF protein serves as a synthetic lethal target in BRCA1- and BRCA2-deficient colonic and ovarian cancer cell lines [28]. In contrast, APEX2 overexpression has been reported in liver cancer [29] and myeloma [30]. Jensen et al. recently explored the expression of APEX2 in multiple cancers and indicated that this gene possesses tissue-specific characteristics [31]. Little research has been done on the role of EBF4 in cancer [32, 33], and a few evidences suggest that it plays important roles in neural development and B-cell maturation [34]. Based on current knowledge, and pending further investigation, our findings suggest that the 7 genes included in our signature may exert important roles in the tumorigenesis and progression of BC.

The ZNF-gene signature identified by us showed a close association with the tumor microenvironment, as it predicted differential representation of stromal and immune infiltrating cells among risk groups. These cells form the major fraction of non-tumor cells in tumor tissues and establish key interactions that influence growth, survival, and metastasis of tumor cells [35]. Based on expression patterns of our ZNF-gene signature among risk groups, higher stromal, immune, and ESTIMATE scores, as well as lower tumor purity, were calculated for the high-risk group. The presence of distinct subsets of immune cells within the tumor microenvironment does not only influence tumor progression, but impact treatment responses as well. Our results showed that tumor-infiltrating immune cell populations were more abundant in the high-risk than in the low-risk group. In particular, the representation of immune cell types involved in antigen presentation was significantly greater in the high-risk group. A mechanistic explanation of this phenomenon may involve a distinct effect of ZNFs on the expression of chemokines, leading to enhanced recruitment of tumor-infiltrating cells.

Our ZNF-gene signature for BC was able to predict chemotherapy sensitivity and may thus help guide treatment selection. Although there are many chemotherapeutic options for BC treatment, there are so far no consensual guidelines in this regard. GDSC is the largest public database containing information on drug sensitivity of cancer cell lines and molecular markers of drug response based on large genomic data. Here, we built statistical models based on gene expression and drug sensitivity data derived from BC cell lines. Our findings showed that the IC50 of 28 chemotherapeutic agents, including gemcitabine and methotrexate [36], predicted using the GDSC dataset, were lower for the low-risk group. While these data suggest that low-risk BC patients are more sensitive to chemotherapy, additional analyses

implementing new tools like CancerTracer, which allows further assessment of intratumor heterogeneity, will further help guide chemotherapy drug selection for personalized treatment [37].

In light of evidence that links deregulation of ZNFs' expression with either pro-oncogenic or tumor-suppressing activities, the significance of ZNFs in cancer tumorigenesis, progression, and metastasis is a topic of intense research. To our knowledge, our study is the first to document a full ZNF gene-based signature with prognostic ability in BC. Nevertheless, there are multiple limitations to the present study. First, since matched, normal bladder samples were far fewer than the BC specimens analyzed, the results need to be verified by expanding the number of controls. Second, the functional relationship between the ZNF gene signature members and non-tumor cells in the tumor microenvironment, especially infiltrating immune cells, could not be elucidated and requires future *in vitro* and *in vivo* studies.

In summary, we established a novel ZNF gene-based prognostic signature that divides BC patients into two subgroups with different survival outcomes and constructed a nomogram to help clinical decision-makers provide optimal treatment. The prognostic signature is associated with differences in stromal and immune cell components of the tumor microenvironment, and predicts sensitivity to chemotherapeutic agents in low risk and high risk BC patients. Our study may stimulate further research on the role of ZNFs on BC and help guide stratified therapy to provide individualized treatment.

## MATERIALS AND METHODS

### Sample information and data collection

The transcriptional data and corresponding clinical information of 403 chemotherapy-naïve BC patients and 19 normal bladder control samples were downloaded from the TCGA website (<https://www.cancer.gov/tcga>) [38]. Gene expression profiles were normalized by the “limma” R package. The GSE48276 [39] dataset, containing mRNA expression profiles from 73 BC tissues, was downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) and used as external validation data. By accessing the UniProt [40] website (<http://www.uniprot.org>), we retrieved the latest list of ZNF-coding genes, which includes 1818 genes.

### Construction and validation of a prognostic model

Differentially expressed genes (DEGs) between tumor and matched normal tissues were identified in the TCGA cohort by the “limma” R package using a false discovery

rate (FDR) < 0.05. ZNF-coding genes with prognostic value were screened out by univariate Cox analysis of overall survival (OS) and p values were adjusted by Wilcoxon tests. To minimize the risk of overfitting, a prognostic model was constructed using LASSO penalized Cox regression analysis [41]. Variable selection and shrinkage of the prognostic model were achieved by running the LASSO algorithm in the “glmnet” R package. The independent variables of the model were the DEGs with prognostic values, and the response variables were OS and status of patients in the TCGA cohort. To improve the reliability and objectivity of the results, 1000 cross-validation runs were performed to determine the optimal value of the penalty parameter ( $\lambda$ ). The normalized expression level of each gene and its corresponding regression coefficient were used to calculate the risk score of patients. The formula was established as follows: score =  $\sum$  (each gene's expression  $\times$  corresponding coefficient). The patients were stratified into high-risk and low-risk groups according to the median value of the risk score. To evaluate the predictive power of the gene signature, a time-dependent ROC curve was built with the “survivalROC” R package. Clinical characteristics, including age, gender, stage, and tumor-node-metastasis (TNM) status were collected from TCGA database. Univariable and multivariate Cox regression analysis were run using clinical data and risk scores to determine whether the predictive value of the risk scores was independent of the clinical characteristics.  $P < 0.05$  was considered statistically significant.

The prognostic signature, with an identical risk score formula and threshold, was then verified against the BC dataset GSE48276. Performance of the prognostic model on the validating dataset was represented via risk score-based plots depicting prognostic gene expression, risk score distribution, and survival status among patients.

### Construction of a ZNF-based nomogram

A nomogram combining the risk score model derived from the prognostic ZNF signature and clinicopathological factors was constructed using the “rms” R package. Discrimination of the nomogram was verified using ROC analysis at 1-, 2-, and 3-year follow-up, and predictive accuracy was assessed through a calibration plot contrasting predicted vs actual survival.

### Estimation of stromal and immune scores

Stromal and immune cells play a fundamental role in shaping the tumor microenvironment [42]. To further confirm the predictive power of our prognostic signature on tumor progression, the ESTIMATE algorithm in R was used to assign stromal and immune scores to the

high-risk and low-risk groups defined by the model. The ESTIMATE score, reflecting tumor purity, was thereby derived [35].

### Analysis of tumor-infiltrating immune cells

The “clusterProfiler” R package was employed to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses based on the DEGs ( $|\log_2FC| \geq 1$ , FDR < 0.05) between the high-risk and low-risk groups. P values were adjusted with the Wilcoxon test. Tumor infiltration scores for 16 immune cell types and activation status for 13 immune-related pathways were assessed with the single-sample gene set enrichment analysis (ssGSEA) function [43] in the “gsva” R package.

### Prediction of chemotherapeutic response

The chemotherapeutic response of each BC patient in the TCGA cohort was predicted according to the public pharmacogenomic database Genomics of Drug Sensitivity in Cancer (GDSC, <https://www.cancerrxgene.org/>). The GDSC database contains data from a large collection of human cancer cell lines, anticancer compounds, and experimental data on drug sensitivity [44]. The prediction of drug sensitivity (IC50) values was conducted using the R package “pRRophetic” [45], which uses a ridge regression model based on cancer cell lines' expression profiles in the GDSC.

### Abbreviations

BC: bladder cancer; ZNFs: zinc finger proteins; ZNF830: zinc finger protein 830; DEGs: differentially expressed genes; FDR: false discovery rate; OS: overall survival; TNM: tumor-node-metastasis; AUC: area under the curve; ROC: receiver operating characteristic; KEGG: Kyoto Encyclopedia of Genes and Genomes; GEO: Gene Expression Omnibus; GSEA: gene set enrichment analysis; ssGSEA: single-sample gene set enrichment analysis; GDSC: Genomics of Drug Sensitivity in Cancer; ZHX3: zinc fingers and homeoboxes 3; ZNF350: zinc finger protein 350; ZNRD1: zinc ribbon domain-containing 1; ZNF195: zinc finger protein 195; SUZ12: SUZ12 polycomb repressive complex 2 subunit; APEX2: apurinic/apyrimidinic endodeoxyribonuclease 2; EBF4: EBF family member 4; ZNRD1-AS1: zinc ribbon domain containing 1 antisense RNA 1.

### AUTHOR CONTRIBUTIONS

Conceptualization: Wang W. Methodology: Zhang C., Cao P., and Yu B.Z. Bioinformatics analyses: Cao H.Y., Gao Z.H., Zhang F.L., and Wu J.Y. Writing, original draft: Zhang J.D. Writing, review and editing: Zhang J.D., Cao H.W., and Hao C.Z. Supervision: Wang W.

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

## FUNDING

This study was supported by grants from the Natural Science Foundation of Beijing (7212040) and the National Natural Science of China Programs (81771720 and 82070764).

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020; 70:7–30. <https://doi.org/10.3322/caac.21590> PMID:31912902
2. Dudley JC, Schroers-Martin J, Lazzareschi DV, Shi WY, Chen SB, Esfahani MS, Trivedi D, Chabon JJ, Chaudhuri AA, Stehr H, Liu CL, Lim H, Costa HA, et al. Detection and Surveillance of Bladder Cancer Using Urine Tumor DNA. *Cancer Discov.* 2019; 9:500–09. <https://doi.org/10.1158/2159-8290.CD-18-0825> PMID:30578357
3. Pardo JC, Ruiz de Porras V, Plaja A, Carrato C, Etxaniz O, Buisan O, Font A. Moving towards Personalized Medicine in Muscle-Invasive Bladder Cancer: Where Are We Now and Where Are We Going? *Int J Mol Sci.* 2020; 21:6271. <https://doi.org/10.3390/ijms21176271> PMID:32872531
4. Cassandri M, Smirnov A, Novelli F, Pitolli C, Agostini M, Malewicz M, Melino G, Raschella G. Zinc-finger proteins in health and disease. *Cell Death Discov.* 2017; 3:17071. <https://doi.org/10.1038/cddiscovery.2017.71> PMID:29152378
5. Chen G, Chen J, Qiao Y, Shi Y, Liu W, Zeng Q, Xie H, Shi X, Sun Y, Liu X, Li T, Zhou L, Wan J, et al. ZNF830 mediates cancer chemoresistance through promoting homologous-recombination repair. *Nucleic Acids Res.* 2018; 46:1266–79. <https://doi.org/10.1093/nar/gkx1258> PMID:29244158
6. Zhang L, Zhou Y, Cheng C, Cui H, Cheng L, Kong P, Wang J, Li Y, Chen W, Song B, Wang F, Jia Z, Li L, et al. Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma. *Am J Hum Genet.* 2015; 96:597–611. <https://doi.org/10.1016/j.ajhg.2015.02.017> PMID:25839328
7. Smirnov A, Lena AM, Cappello A, Panatta E, Anemona L, Bischetti S, Annicchiarico-Petruzzelli M, Mauriello A, Melino G, Candi E. ZNF185 is a p63 target gene critical for epidermal differentiation and squamous cell carcinoma development. *Oncogene.* 2019; 38:1625–38. <https://doi.org/10.1038/s41388-018-0509-4> PMID:30337687
8. Lin DC, Hao JJ, Nagata Y, Xu L, Shang L, Meng X, Sato Y, Okuno Y, Varela AM, Ding LW, Garg M, Liu LZ, Yang H, et al. Genomic and molecular characterization of esophageal squamous cell carcinoma. *Nat Genet.* 2014; 46:467–73. <https://doi.org/10.1038/ng.2935> PMID:24686850
9. Caramel J, Ligier M, Puisieux A. Pleiotropic Roles for ZEB1 in Cancer. *Cancer Res.* 2018; 78:30–35. <https://doi.org/10.1158/0008-5472.CAN-17-2476> PMID:29254997
10. Song Y, Xu Y, Pan C, Yan L, Wang ZW, Zhu X. The emerging role of SPOP protein in tumorigenesis and cancer therapy. *Mol Cancer.* 2020; 19:2. <https://doi.org/10.1186/s12943-019-1124-x> PMID:31901237
11. Borrebaeck CA. Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nat Rev Cancer.* 2017; 17:199–204. <https://doi.org/10.1038/nrc.2016.153> PMID:28154374
12. Yuan P, Ling L, Fan Q, Gao X, Sun T, Miao J, Yuan X, Liu J, Liu B. A four-gene signature associated with clinical features can better predict prognosis in prostate cancer. *Cancer Med.* 2020; 9:8202–15. <https://doi.org/10.1002/cam4.3453> PMID:32924329
13. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* 2015; 16:e173–80. [https://doi.org/10.1016/S1470-2045\(14\)71116-7](https://doi.org/10.1016/S1470-2045(14)71116-7) PMID:25846097
14. Levine AJ. p53: 800 million years of evolution and 40 years of discovery. *Nat Rev Cancer.* 2020; 20:471–80. <https://doi.org/10.1038/s41568-020-0262-1> PMID:32404993
15. Zhu L, Ding R, Yan H, Zhang J, Lin Z. ZHX2 drives cell growth and migration via activating MEK/ERK signal and induces Sunitinib resistance by regulating the autophagy in clear cell Renal Cell Carcinoma. *Cell Death Dis.* 2020; 11:337. <https://doi.org/10.1038/s41419-020-2541-x> PMID:32382017
16. Chen L, Xiang Z, Chen X, Zhu X, Peng X. A seven-gene signature model predicts overall survival in kidney renal clear cell carcinoma. *Hereditas.* 2020; 157:38. <https://doi.org/10.1186/s41065-020-00152-y> PMID:32883362

17. Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y, Li YH. A Novel Ferroptosis-related Gene Signature for Overall Survival Prediction in Patients with Hepatocellular Carcinoma. *Int J Biol Sci.* 2020; 16:2430–41.  
<https://doi.org/10.7150/ijbs.45050>  
PMID:[32760210](https://pubmed.ncbi.nlm.nih.gov/32760210/)
18. Deng M, Wei W, Duan J, Chen R, Wang N, He L, Peng Y, Ma X, Wu Z, Liu J, Li Z, Zhang Z, Jiang L, et al. ZHX3 promotes the progression of urothelial carcinoma of the bladder via repressing of RGS2 and is a novel substrate of TRIM21. *Cancer Sci.* 2021. [Epub ahead of print].  
<https://doi.org/10.1111/cas.14810> PMID:[33440047](https://pubmed.ncbi.nlm.nih.gov/33440047/)
19. Kwon RJ, Kim YH, Jeong DC, Han ME, Kim JY, Liu L, Jung JS, Oh SO. Expression and prognostic significance of zinc fingers and homeoboxes family members in renal cell carcinoma. *PLoS One.* 2017; 12:e0171036.  
<https://doi.org/10.1371/journal.pone.0171036>  
PMID:[28152006](https://pubmed.ncbi.nlm.nih.gov/28152006/)
20. You Y, Ma Y, Wang Q, Ye Z, Deng Y, Bai F. Attenuated ZHX3 expression serves as a potential biomarker that predicts poor clinical outcomes in breast cancer patients. *Cancer Manag Res.* 2019; 11:1199–210.  
<https://doi.org/10.2147/CMAR.S184340>  
PMID:[30787639](https://pubmed.ncbi.nlm.nih.gov/30787639/)
21. Yamada K, Ogata-Kawata H, Matsuura K, Kagawa N, Takagi K, Asano K, Haneishi A, Miyamoto K. ZHX2 and ZHX3 repress cancer markers in normal hepatocytes. *Front Biosci (Landmark Ed).* 2009; 14:3724–32.  
<https://doi.org/10.2741/3483> PMID:[19273305](https://pubmed.ncbi.nlm.nih.gov/19273305/)
22. Fan Y, Shen B, Tan M, Mu X, Qin Y, Zhang F, Liu Y. TGF- $\beta$ -induced upregulation of malat1 promotes bladder cancer metastasis by associating with suz12. *Clin Cancer Res.* 2014; 20:1531–41.  
<https://doi.org/10.1158/1078-0432.CCR-13-1455>  
PMID:[24449823](https://pubmed.ncbi.nlm.nih.gov/24449823/)
23. Lee SR, Roh YG, Kim SK, Lee JS, Seol SY, Lee HH, Kim WT, Kim WJ, Heo J, Cha HJ, Kang TH, Chung JW, Chu IS, Leem SH. Activation of EZH2 and SUZ12 Regulated by E2F1 Predicts the Disease Progression and Aggressive Characteristics of Bladder Cancer. *Clin Cancer Res.* 2015; 21:5391–403.  
<https://doi.org/10.1158/1078-0432.CCR-14-2680>  
PMID:[26268246](https://pubmed.ncbi.nlm.nih.gov/26268246/)
24. Figueroa JD, Malats N, Rothman N, Real FX, Silverman D, Kogevinas M, Chanock S, Yeager M, Welch R, Dosemeci M, Tardón A, Serra C, Carrato A, et al. Evaluation of genetic variation in the double-strand break repair pathway and bladder cancer risk. *Carcinogenesis.* 2007; 28:1788–93.  
<https://doi.org/10.1093/carcin/bgm132>  
PMID:[17557904](https://pubmed.ncbi.nlm.nih.gov/17557904/)
25. Chen Y, Xu T, Xie F, Wang L, Liang Z, Li D, Liang Y, Zhao K, Qi X, Yang X, Jiao W. Evaluating the biological functions of the prognostic genes identified by the Pathology Atlas in bladder cancer. *Oncol Rep.* 2021; 45:191–201.  
<https://doi.org/10.3892/or.2020.7853> PMID:[33200223](https://pubmed.ncbi.nlm.nih.gov/33200223/)
26. Gao Z, Li S, Zhou X, Li H, He S. Knockdown of lncRNA ZNRD1-AS1 inhibits progression of bladder cancer by regulating miR-194 and ZEB1. *Cancer Med.* 2020; 9:7695–705.  
<https://doi.org/10.1002/cam4.3373> PMID:[32862492](https://pubmed.ncbi.nlm.nih.gov/32862492/)
27. Al-Safi RI, Odde S, Shabaik Y, Neamati N. Small-molecule inhibitors of APE1 DNA repair function: an overview. *Curr Mol Pharmacol.* 2012; 5:14–35.  
<https://doi.org/10.2174/1874467211205010014>  
PMID:[22122462](https://pubmed.ncbi.nlm.nih.gov/22122462/)
28. Mengwasser KE, Adeyemi RO, Leng Y, Choi MY, Clairmont C, D'Andrea AD, Elledge SJ. Genetic Screens Reveal FEN1 and APEX2 as BRCA2 Synthetic Lethal Targets. *Mol Cell.* 2019; 73:885–99.e6.  
<https://doi.org/10.1016/j.molcel.2018.12.008>  
PMID:[30686591](https://pubmed.ncbi.nlm.nih.gov/30686591/)
29. Zheng R, Zhu HL, Hu BR, Ruan XJ, Cai HJ. Identification of APEX2 as an oncogene in liver cancer. *World J Clin Cases.* 2020; 8:2917–29.  
<https://doi.org/10.12998/wjcc.v8.i14.2917>  
PMID:[32775374](https://pubmed.ncbi.nlm.nih.gov/32775374/)
30. Kumar S, Talluri S, Pal J, Yuan X, Lu R, Nanjappa P, Samur MK, Munshi NC, Shammas MA. Role of apurinic/apyrimidinic nucleases in the regulation of homologous recombination in myeloma: mechanisms and translational significance. *Blood Cancer J.* 2018; 8:92.  
<https://doi.org/10.1038/s41408-018-0129-9>  
PMID:[30301882](https://pubmed.ncbi.nlm.nih.gov/30301882/)
31. Jensen KA, Shi X, Yan S. Genomic alterations and abnormal expression of APE2 in multiple cancers. *Sci Rep.* 2020; 10:3758.  
<https://doi.org/10.1038/s41598-020-60656-5>  
PMID:[32111912](https://pubmed.ncbi.nlm.nih.gov/32111912/)
32. Fatima A, Tariq F, Malik MF, Qasim M, Haq F. Copy Number Profiling of MammaPrint™ Genes Reveals Association with the Prognosis of Breast Cancer Patients. *J Breast Cancer.* 2017; 20:246–53.  
<https://doi.org/10.4048/jbc.2017.20.3.246>  
PMID:[28970850](https://pubmed.ncbi.nlm.nih.gov/28970850/)
33. Li W, Liu J, Zhang B, Bie Q, Qian H, Xu W. Transcriptome Analysis Reveals Key Genes and Pathways Associated with Metastasis in Breast Cancer. *Onco Targets Ther.* 2020; 13:323–35.  
<https://doi.org/10.2147/OTT.S226770>  
PMID:[32021278](https://pubmed.ncbi.nlm.nih.gov/32021278/)

34. Wang SS, Betz AG, Reed RR. Cloning of a novel Olf-1/EBF-like gene, O/E-4, by degenerate oligo-based direct selection. *Mol Cell Neurosci*. 2002; 20:404–14. <https://doi.org/10.1006/mcne.2002.1138> PMID:[12139918](https://pubmed.ncbi.nlm.nih.gov/12139918/)
35. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013; 4:2612. <https://doi.org/10.1038/ncomms3612> PMID:[24113773](https://pubmed.ncbi.nlm.nih.gov/24113773/)
36. International Collaboration of Trialists on behalf of the Medical Research Council Advanced Bladder Cancer Working Party (now the National Cancer Research Institute Bladder Cancer Clinical Studies Group), and the European Organisation for Research and Treatment of Cancer Genito-Urinary Tract Cancer Group, and the Australian Bladder Cancer Study Group, and the National Cancer Institute of Canada Clinical Trials Group, and Finnbladder, Norwegian Bladder Cancer Study Group, and Club Urologico Espanol de Tratamiento Oncologico Group. International phase III trial assessing neoadjuvant cisplatin, methotrexate, and vinblastine chemotherapy for muscle-invasive bladder cancer: long-term results of the BA06 30894 trial. *J Clin Oncol*. 2011; 29:2171–77. <https://doi.org/10.1200/JCO.2010.32.3139> PMID:[21502557](https://pubmed.ncbi.nlm.nih.gov/21502557/)
37. Wang C, Yang J, Luo H, Wang K, Wang Y, Xiao ZX, Tao X, Jiang H, Cai H. CancerTracer: a curated database for inpatient tumor heterogeneity. *Nucleic Acids Res*. 2020; 48:D797–806. <https://doi.org/10.1093/nar/gkz1061> PMID:[31701131](https://pubmed.ncbi.nlm.nih.gov/31701131/)
38. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, Hinoue T, Laird PW, Hoadley KA, Akbani R, Castro MA, Gibb EA, Kanchi RS, et al, and TCGA Research Network. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell*. 2018; 174:1033. <https://doi.org/10.1016/j.cell.2018.07.036> PMID:[30096301](https://pubmed.ncbi.nlm.nih.gov/30096301/)
39. Choi W, Porten S, Kim S, Willis D, Plimack ER, Hoffman-Censits J, Roth B, Cheng T, Tran M, Lee IL, Melquist J, Bondaruk J, Majewski T, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell*. 2014; 25:152–65. <https://doi.org/10.1016/j.ccr.2014.01.009> PMID:[24525232](https://pubmed.ncbi.nlm.nih.gov/24525232/)
40. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. 2017; 45:D158–69. <https://doi.org/10.1093/nar/gkw1099> PMID:[27899622](https://pubmed.ncbi.nlm.nih.gov/27899622/)
41. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent. *J Stat Softw*. 2011; 39:1–13. <https://doi.org/10.18637/jss.v039.i05> PMID:[27065756](https://pubmed.ncbi.nlm.nih.gov/27065756/)
42. Gillette MA, Satpathy S, Cao S, Dhanasekaran SM, Vasaikar SV, Petralia F, Li Y, Liang WW, Reva B, Krek A, Ji J, Song X, et al, and Clinical Proteomic Tumor Analysis Consortium. Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. *Cell*. 2020; 182:200–25.e35. <https://doi.org/10.1016/j.cell.2020.06.013> PMID:[32649874](https://pubmed.ncbi.nlm.nih.gov/32649874/)
43. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015; 160:48–61. <https://doi.org/10.1016/j.cell.2014.12.033> PMID:[25594174](https://pubmed.ncbi.nlm.nih.gov/25594174/)
44. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013; 41:D955–61. <https://doi.org/10.1093/nar/gks1111> PMID:[23180760](https://pubmed.ncbi.nlm.nih.gov/23180760/)
45. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One*. 2014; 9:e107468. <https://doi.org/10.1371/journal.pone.0107468> PMID:[25229481](https://pubmed.ncbi.nlm.nih.gov/25229481/)

## SUPPLEMENTARY MATERIALS

### Supplementary Table

Please browse Full Text version to see the data of Supplementary Table 1.

**Supplementary Table 1. The DEGs related-ZNFs in bladder cancer.**