Int J Clin Exp Med 2017;10(1):608-617
www.ijcem.com /ISSN:1940-5901/IJCEM0037114

Original Article

A novel 13-gene signature of TGF-beta pathway correlates with tumor stage and grade and predicts poor survival for bladder cancer patients

Tian-Bao Huang[1,2*], Chuan-Peng Dong[3*], Sheng-Ming Lu[1,2], Yang Luan[1,2], Xiao Gu[1,2], Guang-Chen Zhou[1,2], Xue-Fei Ding[1,2]

[1]Department of Urology, Northern Jiangsu People's Hospital, Yangzhou University, Yangzhou, Jiangsu Province, China; [2]Department of Urology, College of Clinical Medicine, Yangzhou University, Yangzhou, Jiangsu Province, China; [3]Institute of Biomedical Science, Shanghai Medical College, Fudan University, Shanghai, China. [*]Equal contributors.

Received August 2, 2016; Accepted October 6, 2016; Epub January 15, 2017; Published January 30, 2017

Abstract: Purposes: TGF-beta pathway functions as both tumor-inhibitor and tumor-promoter in different phases of various cancers including bladder cancer. Our study was to determine if a TGF-beta pathway associated signature could distinguish more aggressive phenotype with worse survival outcomes. Methods: Gene expression profiling of 791 bladder cancer patients from both TCGA and GEO databases were selected and included in our study. 13-gene signature was generated by using Biometric Research Branch-Array Tools. With the specific risk score formula, patients in each dataset were classified into high risk or low risk group. The following analyses were performed in TCGA dataset and validated in other three independent testing sets (GSE13507, GSE31684 and GSE32548). Results: Patients in high risk group had significant shorter overall survival and disease specific survival, compared with who in low risk group. Multivariable Cox regression analysis revealed that the prognostic value of the 13-gene signature was independent of age, gender and smoking status. The 13-gene signature gave a significant performance in distinguishing patients at high risk of worse survival from those at low risk, as measured by the area under the receiver operating characteristic curve. Further analyses demonstrated that the risk score of this signature positively associated with stage and histologic grade (all $P<0.05$). Conclusions: In present study, a novel TGF-beta pathway associated gene signature that is useful in survival prediction in bladder cancer patients was developed. The identification of high risk subpopulation could assist in selecting patients who need more aggressive therapeutic intervention. Meanwhile, the prognosis value of this signature and its potential as a biomarker deserve further investigation in future studies.

Keywords: TGF-beta, signature, bladder cancer, survival

Introduction

In 2016, approximate 76,960 new cancer cases and 16,390 cancer deaths related to bladder cancer (BC) are estimated to occur in United States [1]. In the past two decades, treatment of BC has not changed significantly and the 5-year relative survival rate for patients stood at approximate 79% [1, 2]. At the time of diagnosis, about 20-30% were diagnosed with muscle-invasive (MI) or metastatic BC. Besides, up to one third of the patients with initially non-muscle-invasive (non-MI) BC later progressed to MIBC or metastatic disease [3]. And approximately one half the patients with MIBC later developed into metastatic disease, which is almost invariably lethal (5-year relative survival rate is nearly 5%) [1, 4]. Clinicopathologic parameters such as TNM stage and grade are strongly related to survival outcomes and play an important role in choosing optimum treatment, but there still exist significant variability in the prognosis of patients with similar characteristics. Therefore, additional predictive and prognostic markers are required to distinguish high risk individual from low risk for proper clinical management of the BC patients.

TGF-beta functions as both tumor-inhibitor and tumor-promoter in different phases of various

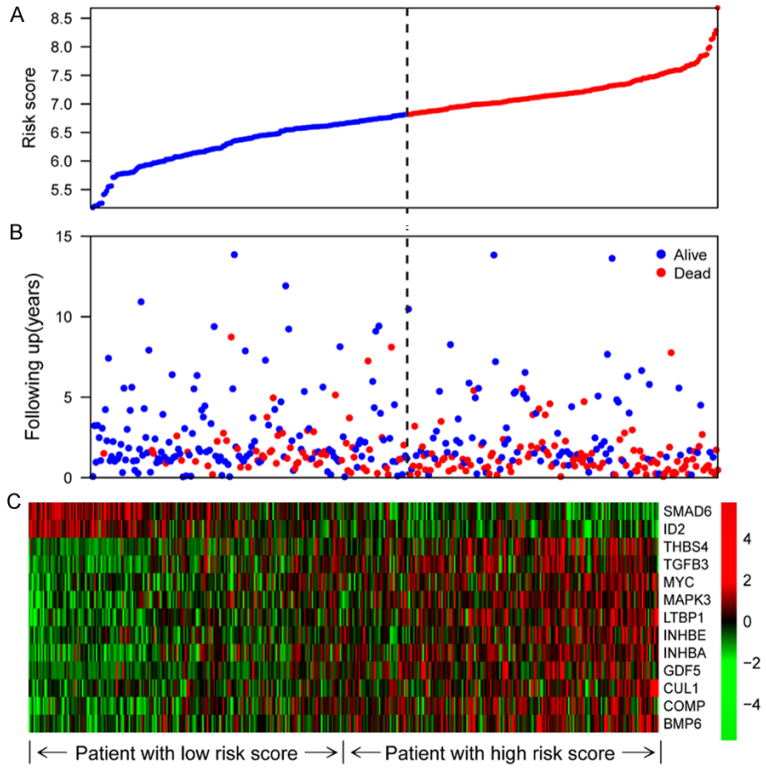


Figure 1. Gene risk score analysis of the TCGA patients. The distribution of 13-gene risk score, patients' survival status and gene expression were analyzed in the entire TCGA database (n = 403). (A) 13-gene risk score distribution; (B) Patients' survival status. The dotted line in the middle of (A) and (B) divided the patients into low risk and high risk group. In high risk group, the patients had a higher mortality rate (113/200 versus 64/203, $P<0.0001$) and a shorter survival time (log rank $P<0.0001$). (C) Heatmap of the 13-gene expression profiles. As the risk score rising, the expression value of SMAD6, ID2 got lower, and the other 11 gene ascended. Rows represent genes in TGF-beta pathway, and columns represent patients.

cancers including BC, which plays a crucial role in malignant evolution, epithelial-to-mesenchymal transition (EMT) and metastasis [5, 6]. Currently, no relevant study of TGF-beta gene on survival outcomes of BC were published, which might be mainly due to its dual character of the TGF-beta and complicacy of the TGF pathway. It seems that the single TGF-beta for risk classification and survival prediction is inappropriate and imprecise. This underlines the importance of gene combinations. Gene signature, based on microarray gene expression profiling, have been recently developed and widely used in prediction of a series of tumor characteristics and outcomes, such as stage, recurrence, progression of non-MIBC and survival [7, 8]. Hence, we asked if a TGF-beta pathway associated signature could discriminate more aggressive phenotype with worse survival outcomes. By using public available datasets, we attempted to generate a novel signature with superior prediction of survival outcomes among BC patients.

Methods

Datasets

Gene expression datasets of BC and corresponding clinical data were downloaded from the publicly available The Cancer Genome Atlas (TCGA) and GEO databases. Gene expression and clinical data of TCGA database are available from the website of Cancer Genomics Browser of University of California Santa Cruz (UCSC) (https://genome-cancer.ucsc.edu) [9]. 403 BC samples with detailed gene expression data and survival data were chosen from the updated TCGA database according to parameters defined in a previous study [10]. Microarray studies from the GEO database are available via the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo). Three datasets, GSE13507 (N = 165), GSE31684 (N = 93) and GSE32548 (N = 130) were selected as testing datasets, which contained either overall survival (OS) or disease specific survival (DSS). 84 TGF-beta pathway-associated genes were obtained from the KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.ad.jp/kegg), including TGF-beta pathway members, target genes and other genes involved in TGF-beta pathway.

Statistical analysis

The association between the gene expression and patient's survival outcomes was assessed by univariate Cox regression analysis along with a permutation test using Biometric Research Branch-Array (BRB-Array) Tools edition 4.5.0 [11]. Genes were considered statistically significant if their permutation p values were less

Table 1. Genes in TGF-beta pathway which correlated with overall survival in The Cancer Genome Atlas bladder cancer dataset

Gene symbols	Gene names	HR	Coefficient	Permutation *p* value
THBS4	Thrombospondin 4	1.082	0.038220873	0.002005
TGFB3	Transforming growth factor beta 3	1.134	-0.164209899	0.006572
SMAD6	SMAD family member 6	0.846	-0.113239657	0.000357
MYC	V-myc avian myelocytomatosis viral oncogene homolog	1.147	0.114266956	0.007135
MAPK3	Mitogen-activated protein kinase 3	1.443	0.317344107	0.003953
LTBP1	Latent transforming growth factor beta binding protein 1	1.249	0.098855692	0.000427
INHBE	Inhibin beta E subunit	1.137	0.155141587	0.005562
INHBA	Inhibin beta A subunit	1.094	-0.013237303	0.006370
ID2	Inhibitor of DNA binding 2, HLH protein	0.866	-0.120674189	0.008203
GDF5	Growth differentiation factor 5	1.135	0.020001652	0.005220
CUL1	Cullin 1	1.577	0.268970151	0.008714
COMP	Cartilage oligomeric matrix protein	1.061	0.039333047	0.007977
BMP6	Bone morphogenetic protein 6	1.151	0.142129333	0.008596

HR: hazard ratio.

than or equal to 0.01. To construct a predictive model, the selected genes were fitted in a multivariable Cox regression model in the training set as described [12]. A risk score formula was then established by including each of these selected genes, weighted by their estimated regression coefficients in the multivariable Cox regression analysis [13]. With this risk score formula, patients in each dataset were classified into high risk or low risk group by using the median risk score as the cutoff point, respectively. The difference of clinicopathological characteristics between the high risk and the low risk group was determined by χ^2 test. Kaplan-Meier survival analyses were used to estimate the survival distributions between the low risk and the high risk group in each set [14, 15]. The log-rank test was used to assess the statistical significance between stratified groups. A two-sided *p* value <0.05 was regarded as significant. The receiver operating characteristic (ROC) curve was constructed using R package pROC. Area under the curve (AUC) values were calculated from the ROC curves. Multivariable Cox regression model was built to further investigate the independent predictive value of the 13-gene signature in each set. The significance was defined as *p* value less than 0.05. All the data were analyzed by R program (www.rproject.org) and statistical software package SPSS for Windows, version 19 (Chicago, SPSS inc, USA).

Results

Identification of prognostic genes in TGF-beta pathway from the TCGA dataset

A total of 84 TGF-beta pathway-associated genes were obtained from the KEGG website. By using the BRB-Array tools, 13 of them were identified as relating to OS (permutation *P*< 0.01). Among these 13 genes, two genes, SMAD6 and ID2, were crudely regarded as protective factors according to hazard ratio (HR, <1), the other 11 genes were, to the contrary, considered as risk factors (**Figure 1C**). The detailed gene information was displayed in **Table 1**.

Predictive value of the 13-gene signature on survival outcomes of bladder cancer patients

The risk score of each patient was estimated according to the 13-gene expression and their corresponding coefficients in TCGA dataset. With this risk score formula, patients in TCGA dataset were divided into high risk (n = 203) or low risk group (n = 200) by using the median risk score as the cutoff point. Compared with the low risk group, the patients in the high risk group had a higher mortality rate (113/200 versus 64/203, *P*<0.0001) and a shorter overall survival time (log rank *P*<0.0001) throughout the follow-up (**Figures 1B** and **2A**). This survival discrepancy were further validated in other two independent BC datasets, GSE13507 and

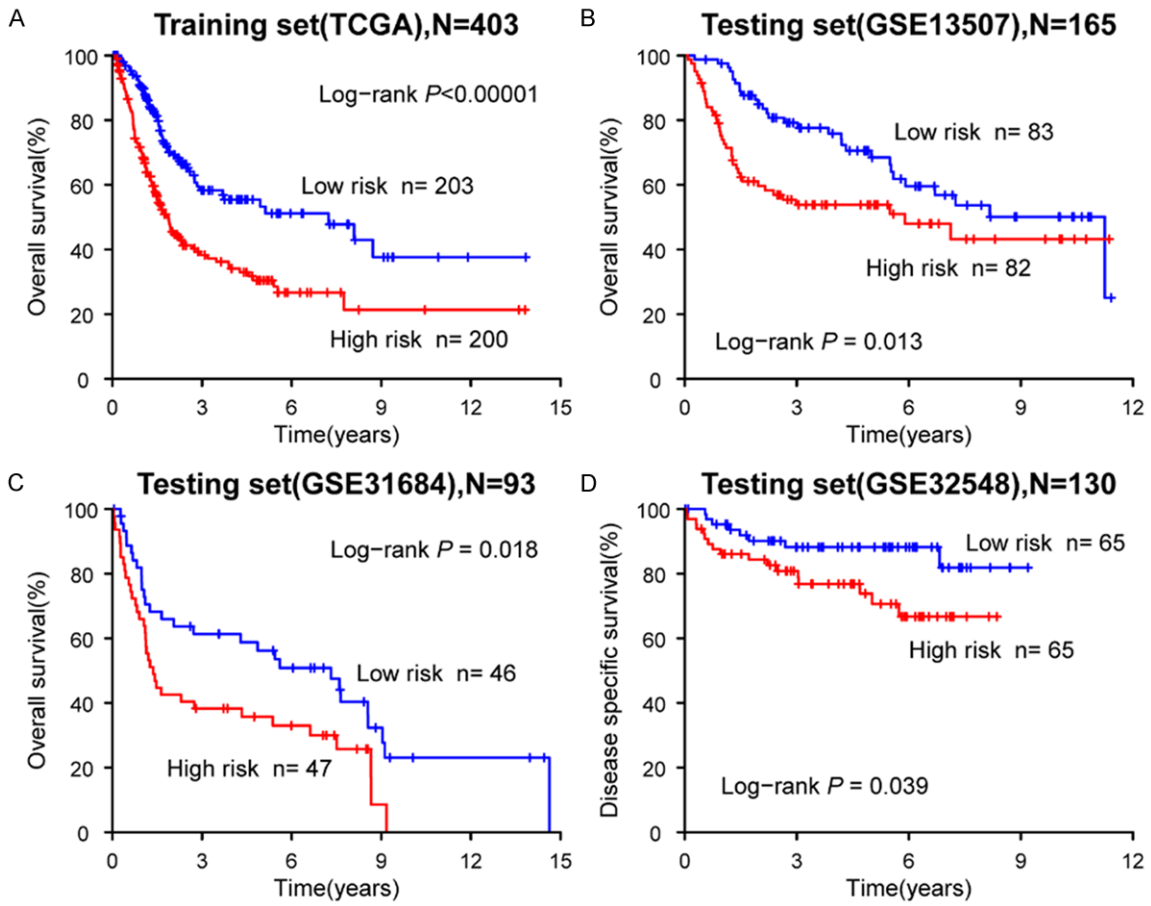


Figure 2. Kaplan-Meier estimates of survival outcomes of patients with bladder cancer in four independent dataset using the 13-gene signature of TGF-beta pathway. Based on the median risk score, patients were divided into two groups: low risk and high risk group. A: Kaplan-Meier curves for TCGA patients. B: Kaplan-Meier curves for GSE13507 patients. C: Kaplan-Meier curves for GSE31684 patients. D: Kaplan-Meier curves for GSE32548 patients. The differences between the two curves were determined by the two-side log-rank test.

GSE31684 (log rank P = 0.013 and log rank P = 0.018, respectively) (**Figure 2B**, **2C**). In addition, multivariate Cox regression analysis of TCGA dataset on OS revealed that the prognostic value of the 13-gene signature was independent of age, gender and smoking status (HR, 2.166, 95% confidence interval (CI), 1.587-2.955, P<0.001), which was validated in another two datasets (both P<0.05) (**Table 2**).

Two independent datasets, including GSE13507 and GSE32548, contained patients' gene expression profile and corresponding DSS data. As shown in Supplementary Figure 1, in GSE13507 dataset, the high risk individuals demonstrated a significantly shorter DSS time than the low risk individuals (log rank P<0.001), which was validated in GSE32548 dataset (log rank P = 0.039). Besides, the multivariate Cox regression analysis of GSE13507 dataset showed that the 13-gene signature was independent of age and gender for DSS prediction (HR, 3.699, 95% CI, 1.657-8.255, P = 0.001) (Supplementary Table 1).

13-gene signature correlates with pathologic stage and histologic grade of bladder cancer

In TCGA dataset, the distribution of pathologic stage, T stage, N stage and histologic grade between the high risk and low risk group were statistically significantly different (P<0.001, P<0.001, P = 0.004 and P<0.001, respectively). As to pathologic stage, we found that patients in stage III/IV had a significant higher risk score than who in stage I/II (P<0.001) (**Figure 3A**). Besides, we also observed that mean risk score ascended as the T stage and N

Table 2. Univariate and multivariable Cox regression analyses on overall survival in the training and testing set

	Univariate model		Multivariable model	
	HR (95% CI)	*p* value	HR (95% CI)	*p* value
Training set (TCGA)				
Risk score (High vs. Low)	2.144 (1.577-2.916)	<0.001	2.166 (1.587-2.955)	<0.001
Age (≥65 vs. <65)	1.723 (1.246-2.381)	0.001	1.766 (1.277-2.443)	0.001
Gender (F vs. M)	1.108 (0.798-1.537)	0.540	1.015 (0.730-1.413)	0.928
Smoking status (Y vs. N)	1.252 (0.915-1.713)	0.161	1.134 (0.826-1.556)	0.438
Testing set (GSE13507)				
Risk score (High vs. Low)	1.814 (1.125-2.925)	0.015	1.777 (1.098-2.875)	0.019
Age (≥65 vs. <65)	3.935 (2.234-6.933)	<0.001	3.981 (2.251-7.040)	<0.001
Gender (F vs. M)	1.560 (0.878-2.772)	0.129	1.559 (0.878-2.770)	0.130
Testing set (GSE31684)				
Risk score (High vs. Low)	1.824 (1.101-3.022)	0.020	1.694 (1.018-2.818)	0.043
Age (≥65 vs. <65)	1.047 (0.605-1.811)	0.871	0.967 (0.551-1.695)	0.906
Gender (F vs. M)	1.012 (0.574-1.785)	0.966	1.045 (0.581-1.881)	0.883
Smoking status		0.063		0.117
Never	Reference		Reference	
Former	2.110 (1.054-4.221)	0.035	1.993 (0.992-4.007)	0.053
Current	1.296 (0.548-3.067)	0.555	1.339 (0.563-3.187)	0.509

HR: hazard ratio, CI: confidence interval.

stage rising in both TCGA and GSE31684 datasets (**Figure 3B**, **3C**). Apart from that, three of the four datasets contained histologic differentiation level of BC. Further comparative analysis revealed that 13-gene expression signature exhibited a significant association with grade of BC (all $P<0.05$, **Figure 3D**).

Finally, we performed ROC analysis to compare the sensitivity and specificity of survival prediction of gene expression signature with age, gender, smoking status, histologic grade and stage in TCGA dataset. The AUROC of the 13 gene signature risk score was 0.688, which was significantly larger than that of age (AUROC = 0.611, $P = 0.030$), gender (AUROC = 0.480, $P<0.001$), smoking status (AUROC = 0.514, $P<0.001$), histologic grade (AUROC = 0.535, $P<0.001$), but insignificantly larger than that of AJCC stage (AUROC = 0.667, $P = 0.530$) (**Figure 4**). These results suggested that 13-gene signature might have a better survival predictive ability.

Discussion

TGF-beta pathway regulates manifold cellular processes, including morphogenesis, embryonic development, adult stem cell differentiation, immune regulation, wound healing and inflammation [16, 17]. Meanwhile, in cancer biology, it also play important roles. TGF-beta functions as both tumor-inhibitor and paradoxical promoter in different phases of various cancers [5, 6], which have vital impact on malignant evolution, invasive, EMT and metastasis [6, 18]. Though the TGF-beta signaling pathway in BC was not fully understood, the existed data revealed a similar dual effect. On one hand, TGF-beta 1 could decrease cell viability, cellular growth and induce apoptosis in BC cell lines [19, 20]. On the other hand, TGF-beta 1 could also induce EMT and invasive via diverse signaling, which might result in the recurrence and progression of BC [21-24]. Recently, Liang et al conditionally knocked out the TGF-beta 2 receptor (TGFBR2) in BC mouse model and found that ablation of TGF-beta signaling could inhibit the cancer cell proliferation, cancer stem cell population and EMT, hence suppressed the invasive cancer progression [25]. Similar to TGFBR1 and TGFBR2, TGFBR3 also played a dichotomous role in human BC, acting as both a tumor suppressor and as a tumor promoter [4]. As a con-

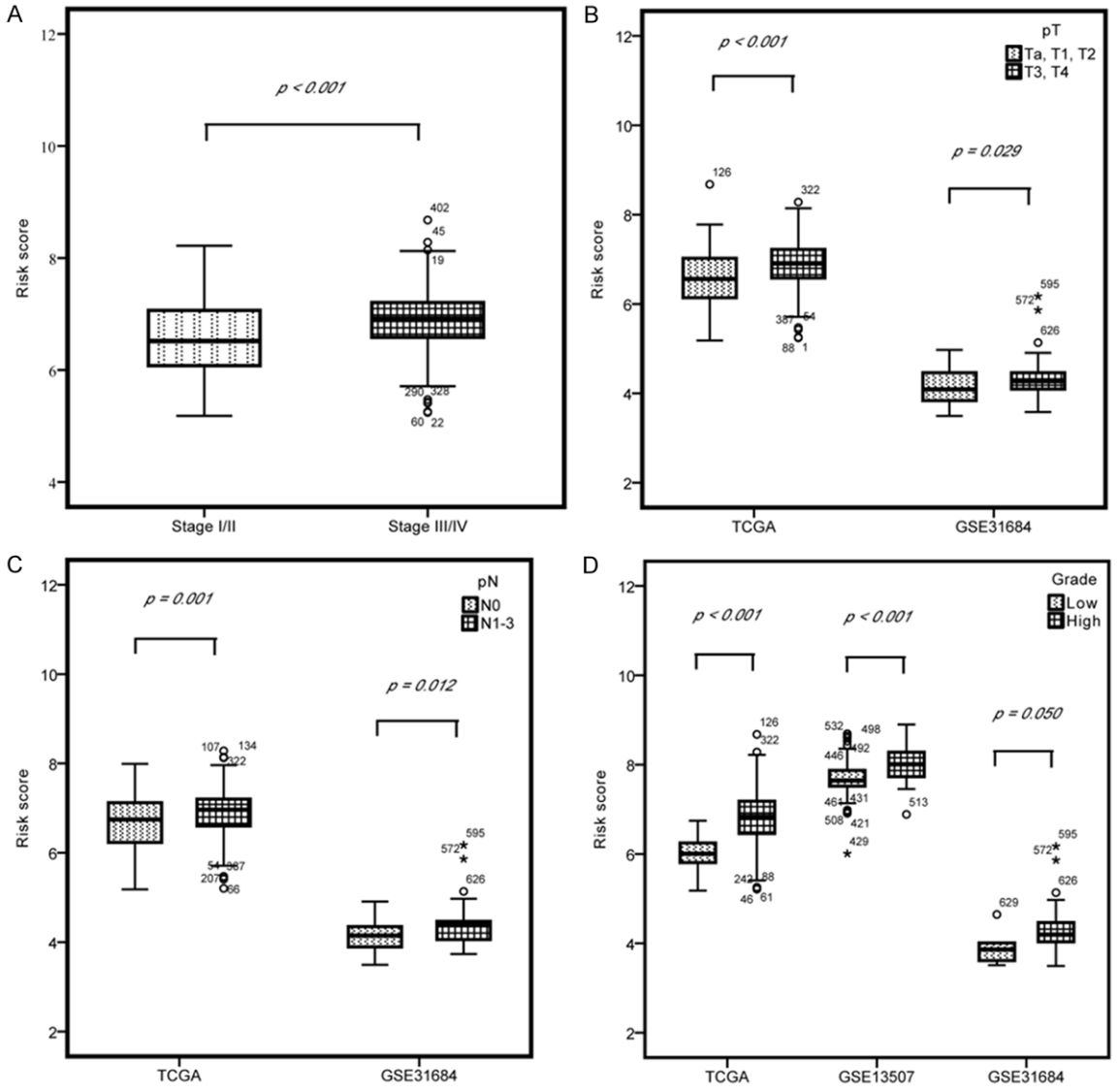


Figure 3. Risk score was correlated with tumor pathological parameters. A: In TCGA dataset, boxplot of risk score in patients with different tumor stage. Mean risk score rose as the tumor stage ascending (P<0.001). B, C: Boxplot of risk score in patients with different T stage and N stage, respectively. The correlation of risk score and T/N stage were analyzed in TCGA and GSE31684. All p values were less than 0.05, which meant that the distribution of risk score differed from tumor T stage and N stage significantly. D: Boxplot of risk score in patients with low grade and high grade bladder cancer. In all three dataset, risk score was significant lower in patients with low grade bladder cancer.

sequence, the output of TGF-beta response in a certain process seemed difficult to predict. In some cases, the TGF-beta pathway served as a tumor inhibitor or tumor promoter due to various alterations of the pathway member or other genes involved in TGF-beta pathway [5]. Therefore, overview of the polygene's expression might somewhat help in further understanding the potential association between TGF-beta pathway and bladder cancer.

Gene signatures based on microarray gene expression profiling have been recently developed to identify subgroups with more aggressive phenotype or poor survival outcomes in BC [2, 7, 26]. For instance, Kim et al [27] identified a four-gene signature with statistically significant correlation with disease progression among patients with MIBC. Jeong et al [28] generated a three-gene signature and validated its performance of disease progression prediction in non-MIBC individuals. Recently, van der

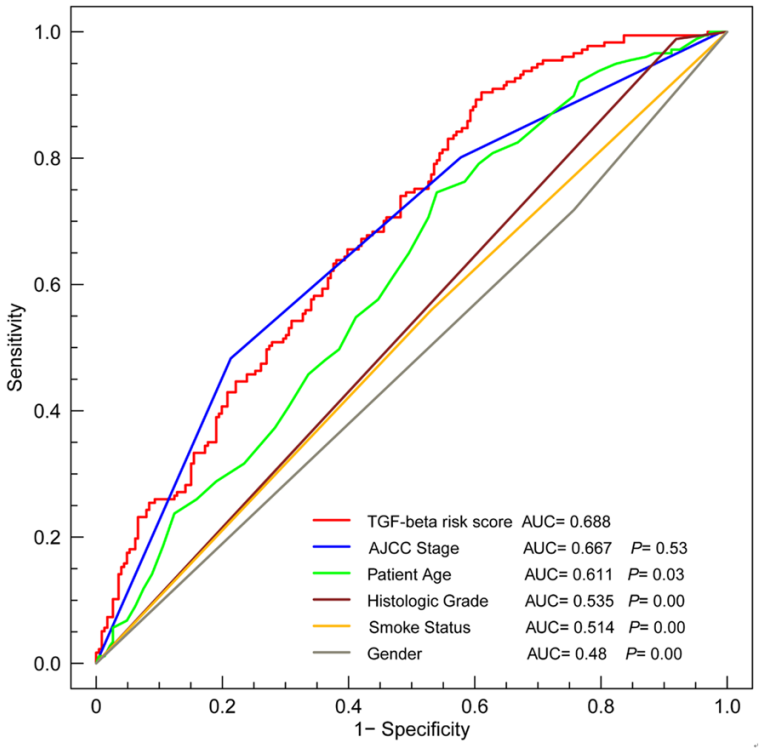


Figure 4. ROC analysis of the sensitivity and specificity of the 13-gene risk score, age, gender, smoking status, grade and stage on overall survival prediction in TCGA data set. As can be observed, the AUROC of the 13 gene expression signature risk score was 0.688, which was significantly larger than that of age (AUROC = 0.611, $P = 0.030$), gender (AUROC = 0.480, $P<0.001$), smoking status (AUROC = 0.514, $P<0.001$) and grade (AUROC = 0.535, $P<0.001$), but insignificantly larger than that of AJCC stage (AUROC = 0.667, $P = 0.530$).

Heijden et al [29] developed a five-gene expression signature to identify T1G3 BC patients with high risk of progression. Currently, there is no relevant study involving in TGF-beta pathway in prediction of BC patients' survival. Hence, we attempted to generate a gene signature of TGF-beta signaling to assist in improving prediction of survival in patients with BC. From KEGG website, we obtained 84 TGF-beta pathway associated genes. All the candidate genes were firstly tested for their potential predictive value among 403 BC patients in TCGA dataset. 13 genes were identified and significantly associated with overall survival of patients with BC. With this risk score formula, 403 BC patients were, subsequently, classified into high risk and low risk group. Then, the clinicopathologic parameters and survival outcomes of this two subgroups were compared. Surprisingly, the 13 TGF-beta pathway genes' signature correlated with pathological stage, T stage, N stage and histologic grade of BC and exhibited a promising independent prognostic value on BC patients' survival. These findings were validated in another three independent dataset, GSE13507, GSE31684 and GSE32548.

13 genes that involved in our novel TGF-beta pathway included THBS4, TGFB3, SMAD6, MYC, MAPK3, LTBP1, INHBE, INHBA, ID2, GDF5, CUL1, COMP and BMP6. Of them, the function and clinical significance had been investigated in CUL1, INHBA, MYC and SMAD6. Mao et al [30] found that CUL1 overexpressed in high-grade urothelial carcinoma and correlated with poor prognosis of the patients with urothelial carcinoma. Similarly, increased expression of INHBA was also significantly associated with advanced clinicopathological features in urothelial carcinoma and significantly implied inferior DSS and metastasis free survival (both $P<0.001$) [31]. Besides, high expression of c-Myc was proved relating to shorter disease-free survival (HR, 3.05, $P = 0.011$) [32]. Riester et al [2] observed that SMAD6 was highly expressed in non-MI BC, as compared to muscle invasive BC. To date, the remaining nine genes were not directly validated their clinical significance in BC. In our study, the 13-gene signature had clinical significance in sorting high risk patients with worse survival. It is possible that these genes' interaction play an important role in progression, recurrence, metastasis of BC, by which the signature significantly influence the survival outcomes. The biological significance of these 13 TGF-beta signaling genes deserves further investigation.

Several limitations should be taken into account for our study. Firstly, considering the inherent discrepancy of the samples and difference of the measurement procedures among four datasets, we either did not combine them together

as a whole to perform these analyses, nor did not use a same risk score cutoff in the analyses of the four independent datasets. Secondly, TCGA and another two datasets (GSE13507, GSE31684) do not collect detailed therapeutic measures (surgery, intravesical chemotherapy, systemic chemotherapy), race [33], insurance status [34], income status, dietary pattern [35], body mass index [36, 37], comorbidities such as diabetes [38], hypertension [39] and coronary heart disease, which might more or less influence the survival of BC patients, without adding these factors in the multivariate analyses might partially bias the independence of the predictive value. Thirdly, in our study, only 84 TGF-beta associated genes were found and tested in the training dataset. The prognostic genes identified here did not represent all the gene candidates that were potentially correlated with BC patients' survival. Finally, this gene signature was inferred by bioinformatics analysis, and the biological roles of several genes in this signature were not clear, which should be investigated in further fundamental researches.

In conclusion, a novel TGF-beta pathway gene signature that is useful in survival prediction in BC patients was developed. Meanwhile, this signature is positively correlated to malignant behavior, such as stage and histological grade. The identification of high risk subpopulation could assist in selecting patients who need more aggressive therapeutic intervention. Meanwhile, the prognosis value of this signature and its potential as a biomarker deserve further investigation in future studies.

Disclosure of conflict of interest

None.

Address correspondence to: Xue-Fei Ding and Guang-Chen Zhou, Department of Urology, Northern Jiangsu People's Hospital, Yangzhou University, No. 98 West Nantong Road, Yangzhou 225001, Jiangsu Province, China. Tel: +8618051061234; Fax: +8618051061234; E-mail: dingxf_sbyy@163.com (XFD); Tel: +8618051062835; Fax: +8618051062835; E-mail: 1032922911@qq.com (GCZ)

References

[1] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016; 66: 7-30.

[2] Riester M, Taylor JM, Feifer A, Koppie T, Rosenberg JE, Downey RJ, Bochner BH and Michor F. Combination of a novel gene expression signature with a clinical nomogram improves the prediction of survival in high-risk bladder cancer. Clin Cancer Res 2012; 18: 1323-1333.

[3] Kaufman DS, Shipley WU and Feldman AS. Bladder cancer. Lancet 2009; 374: 239-249.

[4] Liu XL, Xiao K, Xue B, Yang D, Lei Z, Shan Y and Zhang HT. Dual role of TGFBR3 in bladder cancer. Oncol Rep 2013; 30: 1301-1308.

[5] Zhu H, Luo H, Shen Z, Hu X, Sun L and Zhu X. Transforming growth factor-beta1 in carcinogenesis, progression, and therapy in cervical cancer. Tumour Biol 2016; 37: 7075-7083.

[6] Massague J. TGF beta in cancer. Cell 2008; 134: 215-230.

[7] Sanchez-Carbayo M, Socci ND, Lozano J, Saint F and Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. J Clin Oncol 2006; 24: 778-789.

[8] Birkhahn M, Mitra AP, Williams AJ, Lam G, Ye W, Datar RH, Balic M, Groshen S, Steven KE and Cote RJ. Predicting recurrence and progression of noninvasive papillary bladder cancer at initial presentation based on quantitative gene expression profiles. Eur Urol 2010; 57: 12-20.

[9] Cline MS, Craft B, Swatloski T, Goldman M, Ma S, Haussler D and Zhu J. Exploring TCGA pan-cancer data at the UCSC cancer genomics browser. Sci Rep 2013; 3: 2652.

[10] Lu X, Wan F, Zhang H, Shi G and Ye D. ITGA2B and ITGA8 are predictive of prognosis in clear cell renal cell carcinoma patients. Tumour Biol 2016; 37: 253-262.

[11] Simon R, Lam A, Li MC, Ngan M, Menenzes S and Zhao Y. Analysis of gene expression data using BRB-ArrayTools. Cancer Inform 2007; 3: 11-17.

[12] Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D and Levy R. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med 2004; 350: 1828-1837.

[13] Kawaguchi A, Iwadate Y, Komohara Y, Sano M, Kajiwara K, Yajima N, Tsuchiya N, Homma J, Aoki H, Kobayashi T, Sakai Y, Hondoh H, Fujii Y, Kakuma T and Yamanaka R. Gene expression signature-based prognostic risk score in patients with primary central nervous system lymphoma. Clin Cancer Res 2012; 18: 5672-5681.

[14] Meng J, Li P, Zhang Q, Yang Z and Fu S. A four-long non-coding RNA signature in predicting breast cancer survival. J Exp Clin Cancer Res 2014; 33: 84.

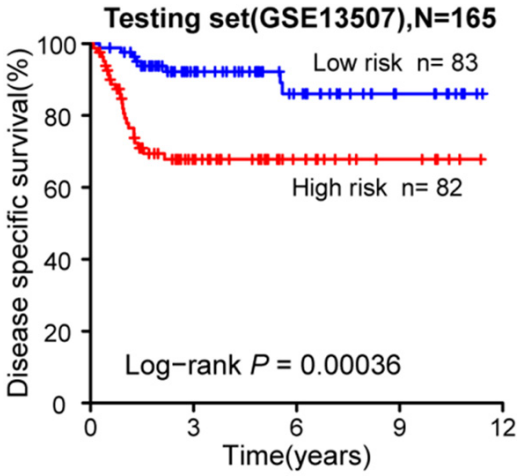
[15] Zhang XQ, Sun S, Lam KF, Kiang KM, Pu JK, Ho AS, Lui WM, Fung CF, Wong TS and Leung GK. A long non-coding RNA signature in glioblastoma multiforme predicts survival. Neurobiol Dis 2013; 58: 123-131.
[16] Nana AW, Yang PM and Lin HY. Overview of transforming growth factor beta superfamily involvement in glioblastoma initiation and progression. Asian Pac J Cancer Prev 2015; 16: 6813-6823.
[17] Morikawa M, Derynck R and Miyazono K. TGF-beta and the TGF-beta family: context-dependent roles in cell and tissue physiology. Cold Spring Harb Perspect Biol 2016; 8.
[18] Liu S, de Boeck M, van Dam H and Ten Dijke P. Regulation of the TGF-beta pathway by deubiquitinases in cancer. Int J Biochem Cell Biol 2016; 76: 135-145.
[19] Al-Azayzih A, Gao F, Goc A and Somanath PR. TGFbeta1 induces apoptosis in invasive prostate cancer and bladder cancer cells via Akt-independent, p38 MAPK and JNK/SAPK-mediated activation of caspases. Biochem Biophys Res Commun 2012; 427: 165-170.
[20] Lee C, Lee SH, Kim DS, Jeon YS, Lee NK and Lee SE. Growth inhibition after exposure to transforming growth factor-beta1 in human bladder cancer cell lines. Korean J Urol 2014; 55: 487-492.
[21] Liao Y, Qiu M, Liu J and Huang J. [Effects of transforming growth factor-beta1 gene silencing on the expression of vascular endothelial growth factor in human bladder cancer cell lines]. Zhonghua Yi Xue Za Zhi 2014; 94: 1274-1276.
[22] Chen MF, Zeng F, Qi L, Zu XB, Wang J, Liu LF and Li Y. Transforming growth factorbeta1 induces epithelial mesenchymal transition and increased expression of matrix metalloproteinase16 via miR200b downregulation in bladder cancer cells. Mol Med Rep 2014; 10: 1549-1554.
[23] Li W, Kidiyoor A, Hu Y, Guo C, Liu M, Yao X, Zhang Y, Peng B and Zheng J. Evaluation of transforming growth factor-beta1 suppress Pokemon/epithelial-mesenchymal transition expression in human bladder cancer cells. Tumour Biol 2015; 36: 1155-1162.
[24] Gupta S, Hau AM, Al-Ahmadie HA, Harwalkar J, Shoskes AC, Elson P, Beach JR, Hussey GS, Schiemann WP, Egelhoff TT, Howe PH and Hansel DE. Transforming growth factor-beta Is an upstream regulator of mammalian target of rapamycin complex 2-dependent bladder cancer cell migration and invasion. Am J Pathol 2016; 186: 1351-1360.
[25] Liang Y, Zhu F, Zhang H, Chen D, Zhang X, Gao Q and Li Y. Conditional ablation of TGF-beta signaling inhibits tumor progression and invasion in an induced mouse bladder cancer model. Sci Rep 2016; 6: 29479.
[26] Blaveri E, Simko JP, Korkola JE, Brewer JL, Baehner F, Mehta K, Devries S, Koppie T, Pejavar S, Carroll P and Waldman FM. Bladder cancer outcome and subtype classification by gene expression. Clin Cancer Res 2005; 11: 4044-4055.
[27] Kim WJ, Kim SK, Jeong P, Yun SJ, Cho IC, Kim IY, Moon SK, Um HD and Choi YH. A four-gene signature predicts disease progression in muscle invasive bladder cancer. Mol Med 2011; 17: 478-485.
[28] Jeong P, Ha YS, Cho IC, Yun SJ, Yoo ES, Kim IY, Choi YH, Moon SK and Kim WJ. Three-gene signature predicts disease progression of non-muscle invasive bladder cancer. Oncol Lett 2011; 2: 679-684.
[29] van der Heijden AG, Mengual L, Lozano JJ, Ingelmo-Torres M, Ribal MJ, Fernandez PL, Oosterwijk E, Schalken JA, Alcaraz A and Witjes JA. A five-gene expression signature to predict progression in T1G3 bladder cancer. Eur J Cancer 2016; 64: 127-136.
[30] Mao SY, Xiong DB, Huang TB, Zheng JH and Yao XD. Expression of CUL1 correlates with tumour-grade and recurrence in urothelial carcinoma. ANZ J Surg 2016; [Epub ahead of print].
[31] Lee HY, Li CC, Huang CN, Li WM, Yeh HC, Ke HL, Yang KF, Liang PI, Li CF and Wu WJ. INHBA overexpression indicates poor prognosis in urothelial carcinoma of urinary bladder and upper tract. J Surg Oncol 2015; 111: 414-422.
[32] Massari F, Bria E, Ciccarese C, Munari E, Modena A, Zambonin V, Sperduti I, Artibani W, Cheng L, Martignoni G, Tortora G and Brunelli M. Prognostic value of beta-tubulin-3 and c-Myc in muscle invasive urothelial carcinoma of the bladder. PLoS One 2015; 10: e0127908.
[33] DeDeugd C, Miyake M, Palacios DA and Rosser CJ. The influence of race on overall survival in patients with newly diagnosed bladder cancer. J Racial Ethn Health Disparities 2015; 2: 124-131.
[34] Niu X, Roche LM, Pawlish KS and Henry KA. Cancer survival disparities by health insurance status. Cancer Med 2013; 2: 403-411.
[35] Parsons JK, Pierce JP, Natarajan L, Newman VA, Barbier L, Mohler J, Rock CL, Heath DD, Guru K, Jameson MB, Li H, Mirheydar H, Holmes MA and Marshall J. A randomized pilot trial of dietary modification for the chemoprevention of noninvasive bladder cancer: the dietary intervention in bladder cancer study. Cancer Prev Res (Phila) 2013; 6: 971-978.
[36] Dabi Y, Rouscoff Y, Anract J, Delongchamps NB, Sibony M, Saighi D, Zerbib M, Peyraumore M and Xylinas E. Impact of body mass index on the oncological outcomes of patients treated

with radical cystectomy for muscle-invasive bladder cancer. World J Urol 2016; [Epub ahead of print].

[37] Kluth LA, Xylinas E, Crivelli JJ, Passoni N, Comploj E, Pycha A, Chrystal J, Sun M, Karakiewicz PI, Gontero P, Lotan Y, Chun FK, Fisch M, Scherr DS and Shariat SF. Obesity is associated with worse outcomes in patients with T1 high grade urothelial carcinoma of the bladder. J Urol 2013; 190: 480-486.

[38] Oh JJ, Kang MY, Jo JK, Lee HM, Byun SS, Lee SE, Lee S and Hong SK. Association between diabetes mellitus and oncological outcomes in bladder cancer patients undergoing radical cystectomy. Int J Urol 2015; 22: 1112-1117.

[39] Dal Moro F, Bovo A, Crestani A, Vettor R, Gardiman MP and Zattoni F. Effect of hypertension on outcomes of high-risk patients after BCG-treated bladder cancer: a single-institution long follow-up cohort study. Medicine (Baltimore) 2015; 94: e589.



Supplementary Figure 1. Kaplan-Meier estimates of the disease specific survival of patients with bladder cancer in GSE13507 dataset using the 13-gene signature of TGF-beta pathway. Based on the median risk score, patients were divided into two groups: low risk and high risk group. The differences between the two curves were determined by the two-side log-rank test.

Supplementary Table 1. Univariate and multivariable Cox regression analyses on disease specific survival in the GSE13507 dataset

	Univariate model		Multivariable model	
	HR (95% CI)	*p* value	HR (95% CI)	*p* value
GSE13507				
Risk score (High vs. Low)	3.873 (1.737-8.632)	0.001	3.699 (1.657-8.255)	0.001
Age (≥65 vs. <65)	3.003 (1.343-6.716)	0.007	2.849 (1.270-6.391)	0.011
Gender (F vs. M)	2.097 (0.969-4.538)	0.060	2.087 (0.965-4.515)	0.062

HR: hazard ratio, CI: confidence interval.