High Expression of CTSV in Bladder Cancer as a Predictor of Poor Prognosis: A Study Based on TCGA and GEO Database

Zhiqiang Yao

Lanzhou University Second Hospital

Jiajia Zhang

Lanzhou University Second Hospital

Chaohu Chen

Lanzhou University Second Hospital

Pan Li

Lanzhou University Second Hospital

Jinlong Cao

Lanzhou University Second Hospital

Dali Han

Lanzhou University Second Hospital

Jianpeng Li

Lanzhou University Second Hospital

Lijun Ying

Lanzhou University Second Hospital

Junqiang Tian (✉ ery_tianjq@lzu.edu.cn)

Lanzhou University Second Hospital <https://orcid.org/0000-0002-4529-4650>

Research article

Keywords: CTSV, Bladder cancer, Prognosis, TCGA, GEO, GSEA

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-378298/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Bladder cancer (BLCA) is the most common malignancy of urinary system with a high recurrence rate. We aimed to explore the relationship between cathepsin V (CTSV) expression and prognosis in patients with bladder cancer.

Methods: The RNA-Seq gene expression data and corresponding clinical information with BLCA were downloaded from TCGA database. The gene expression profiles of GSE13507 and GSE133624 were downloaded from GEO database. BLCA patients were divided into high and low expression group according to the cutoff value of CTSV expression. The relationship between clinicopathologic characteristics and CTSV expression was analyzed with the Wilcoxon signed-rank test and logistic regression. Kaplan-Meier analysis and Cox regression were used to analyze the relationship between overall survival and clinicopathologic characteristics. Gene set enrichment analysis (GSEA) was utilized to identify enriched KEGG pathway.

Results: High expression of CTSV was significantly correlated with pathological grade (OR = 1.662 for low vs. high), clinical stage (OR = 1.589 for I-II vs. III-IV), status (OR = 1.435 for normal vs. tumor), T stage (OR = 1.589 for T1-2 vs. T3-4), and M stage (OR = 4.499 for M0 vs M1). The expression of CTSV was significantly increased in BLCA compared with normal tissue ($P < 0.001$). Kaplan-Meier survival analysis showed that BLCA patients with high expression of CTSV had a poorer prognosis than low expression of CTSV patients ($P = 0.0016$). Univariate Cox analysis showed that high expression of CTSV was significantly associated with poorer overall survival (HR:1.662, 95%CI:1.209-2.286, $P = 0.002$). Multivariate Cox regression showed that high expression of CTSV was an independent risk factor for poor prognosis in BLCA patients (HR: 1.495, 95%CI: 1.069-2.089, $P = 0.019$). We also used the GSE13507 and GSE133624 to verify whether CTSV was differently expressed in bladder cancer tissues and normal tissues. The results showed that CTSV expression was significantly increased in BLCA patients ($P < 0.05$). Finally, GSEA was used to show 22 enriched signaling pathways in a high phenotype.

Conclusion: High expression of CTSV in bladder cancer is associated with poor prognosis and may serve as a new biomarker. In addition, the chemokine signaling pathway, MAPK signaling pathway, Wnt signaling pathway, JAK-STAT signaling pathway, tight Junction and cell adhesion molecules may be the key pathway regulated by CTSV in BLCA.

Background

Bladder cancer (BLCA) is the most common malignancy of urinary system, with nearly 550,000 new cases and 200,000 deaths occurred in 2018 worldwide[1]. According to the depth of tumor invasion and histopathological subtypes, BLCA is divided into non-muscle invasive bladder cancer (NMIBC) and muscle invasive bladder cancer (MIBC). Although the patients suffering from NMIBC (pTa-pT1) was performed the transurethral resection of bladder tumor (TURBT), the recurrence rate within five years can be up to 75%[2]. MIBC has a high metastasis rate with a 5-year survival rate of approximately 60%[3].

BLCA is not easily cured due to its high rates of recurrence and metastasis, causing a five-year survival rate of approximately 57%[4, 5]. Therefore, it is particularly important to discover new diagnostic methods and treatment strategies for BLCA.

The most common pathologic type of BLCA is transitional cell carcinoma. Mutations are found in oncogenes and tumor suppressor genes such as FGFR3 and TP53. NMIBC and MIBC are different in genetics. NMIBC is characterized by high frequency mutation in the FGFR3 oncogene, leading to constitutive activation of the RAS/ MAPK pathway. However, MIBC is highly enriched with inactivating mutations in TP53[6-8]. In addition, other mutated genes in BLCA have been identified, including HRAS, RB1, TSC1 and PIK3CA[9].

Cathepsins are a class of lysosomal peptidases and the cathepsin family includes cathepsin A, B, C, D, E, F, G, H, L, K, O, S, V and W, which play a key role in normal tissue homeostasis[10]. The studies have shown that cathepsins are highly expressed in various cancers and closely related to tumor progression and metastasis[11]. Cathepsin V, also known as cathepsin L2 (CTSV/CTSL2), is a cysteine cathepsin that is specifically expressed in the thymus, cornea, and testis[12, 13]. CTSV is over-expressed in tumors such as breast cancer, colorectal cancer, renal cancer and ovarian cancer[14]. It has also been shown that high expression of CTSV is associated with adverse prognosis in hepatocellular carcinoma[15]. However, there are no studies on its significance in the prognosis of BLCA.

Thus, the objective of the current study was to evaluate the prognostic value of CTSV expression in BLCA based on TCGA database. In order to identify the biological pathways related to the CTSV regulatory network in BLCA, we used GSEA. The high expression of CTSV in BLCA is related to poor prognosis, which provided a potential diagnostic marker and therapeutic target for BLCA.

Methods

Data acquisition and preprocessing

RNA-Seq gene expression data and corresponding clinical information (including age, gender, race, histologic grade, clinical stage, T stage, N stage, M stage, overall survival (OS) time and Survival outcome) with BLCA were downloaded from The Cancer Genome Atlas (TCGA) database(<https://portal.gdc.cancer.gov/>). We got 408 tumor samples and 19 normal samples by removing multiple tissue samples corresponding to the same patient. 403 samples with complete survival information were used for survival analysis. These samples with complete clinicopathologic variables were used for logistics regression, univariate and multivariate Cox analysis. The GSE13507 and GSE133624 gene expression profile matrix files were downloaded from Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), which are used to validate the expression level of CTSV mRNA in BLCA patients. The GSE13507 contains 188 bladder cancer tissues and 68 normal bladder tissues (Platform: GPL6102 Illumina human-6 v2.0 expression beadchip). The GSE133624 contains 36 bladder cancer tissues and 29 normal bladder tissues (Platform: GPL20795 HiSeq X Ten (Homo sapiens)). The gene expression data and clinical data were retained and further analyzed.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states[16]. The Molecular Signatures Database (MSigDB) is a collection of annotated gene sets for use with GSEA software. In this study, the samples were divided into the high expression group and the low expression group according to the cutoff value of CTSV expression, and GSEA was utilized to identify enriched KEGG pathway. The low expression group and high expression group were used as a phenotype label. The parameter settings of permutations and minimum setSize are 1000 and 100, respectively. We used the normalized enrichment score (NES) and false discovery rate (FDR) to sort the pathways enriched in each phenotype. NOM p-val <0.05 and FDR q-val < 0.05 were considered statistically significant.

Statistical Analyses

All statistical analyses were conducted by R software (version 4.0.0, <https://www.r-project.org/>). The boxplots were used to visualize the different CTSV expression of all sample for discrete variables. The Wilcoxon signed-rank test and logistic regression were used to analyze the relationship between clinicopathologic characteristics and CTSV. The Kaplan-Meier method was used for survival analysis, and the cutoff value was determined by the survminer package based on CTSV expression value, the survival time, and the survival status. Then univariate Cox analysis was used to analyze the relationship between OS and clinicopathologic characteristics. Multivariate Cox analysis was used to compare the influence of CTSV expression on survival along with other clinical characteristics (age, clinical stage, T stage, N stage, and M stage). $P < 0.05$ was considered statistically significant.

Results

Patient characteristics

As shown in **Table 1**, 408 primary tumors with both clinical data and gene expression data were downloaded from TCGA database in July 2020. The patients ranged in age from 34 to 90 years old, with 301 males and 107 females. The race of the patients was 79.41% white, 10.78% Asian and 5.64% Black or African American. Most tumors (92.65%, n=378) were high grade, while the low grade was only 5.88% (n=24). According to clinical stage information, stage I, II, III and IV were composed of 0.49% (n=2), 31.86% (n=130), 34.32% (n=140), and 32.84% (n=134), respectively. There were 3 (0.74%) patients with T1 stage, 119 (29.17%) patients with T2 stage, 194 (47.55%) patients with T3 stage, and 57 (13.97%) patients with T4 stage. There were 237 patients (58.09%) with N0 stage, 46 patients (11.27%) with N1 stage, 75 patients (18.38%) with N2 stage, and 8 patients (1.97%) with N3 stage. Among the 408 patients, 11 patients (2.70%) had distant metastasis, while 196 patients (48.04%) had no distant metastasis.

Association with CTSV expression and clinicopathologic variables

A total of 427 BLCA samples with CTSV expression data were downloaded from the TCGA database, including 408 tumor samples with complete clinical information and 19 normal samples. As shown in **Fig. 1**, the expression of CTSV in tumor tissues significantly increased compared to normal tissues ($P < 0.001$). High expression of CTSV was associated with gender ($P = 0.046$), pathological grade ($P < 0.001$), race ($P = 0.0039$), and T stage ($P = 0.043$). Univariate logistic regression analysis showed that CTSV expression as the dependent variable (based on the cutoff value of 211) was correlated with the clinicopathological characteristics of poor prognosis (**Table 2**). The high expression of CTSV was significantly correlated with race (OR = 2.519 for white vs. others), pathological grade (OR = 1.662 for low vs. high), clinical stage (OR = 1.589 for I-II vs. III-IV), status (OR = 1.435 for normal vs. tumor), T stage (OR = 1.589 for T1-2 vs. T3-4), M stage (OR = 4.499 for M0 vs M1) (all $P < 0.05$). These results suggested that high expression of CTSV in BLCA is associated with higher pathological grade, clinical stage, T stage, and M stage.

Survival outcomes and univariate and multivariate Cox analysis

BLCA patients were divided into high expression group and low expression group according to the cutoff value of CTSV which determined by survminer package in R software (**Fig. 2B**). Kaplan–Meier survival analysis showed that the prognosis of BLCA patients with high CTSV expression was worse than that of those with low CTSV expression ($P = 0.0016$) (**Fig. 2A**). Univariate Cox analysis showed that age (HR: 1.034, 95%CI: 1.018-1.051, $P < 0.001$), clinical stage (HR:1.709, 95%CI:1.410-2.071, $P < 0.001$), T stage (HR:1.676, 95% CI:1.335-2.105, $P < 0.001$), N stage (HR:1.497, 95% CI:1.222-1.833, $P < 0.001$), M stage (HR:1.188, 95% CI:1.022-1.381, $P = 0.025$), and CTSV (HR:1.662, 95%CI :1.209-2.286, $P = 0.002$) were associated with OS (**Table 3**).

The statistically significant variables by univariate Cox analysis were included in multivariate Cox regression for further analysis. The results showed that age (HR: 1.033, 95%CI:1.016-1.050, $P < 0.001$), T stage (HR: 1.459, 95%CI:1.088-1.958, $P = 0.012$), N stage (HR: 1.433, 95%CI:1.075-1.909, $P = 0.014$), and CTSV (HR: 1.495, 95%CI: 1.069-2.089, $P = 0.019$) were independent prognostic risk factors (**Table 3**).

Validation of CTSV gene expression in BLCA

The GSE13507 and GSE133624 were used to verify the CTSV gene expression in BLCA tissues and normal tissues. As shown in **Fig. 3**, the expression of CTSV in BLCA tissues was significantly higher than normal tissues, which verified the conclusions analyzed above. $P < 0.05$ was considered statistically significant.

GSEA identifies CTSV-related signaling pathways

To explore the underlying molecular mechanisms of CTSV in BLCA, GSEA was conducted to compare between low expression group and high expression group in TCGA data. GSEA reveals significant difference (NOM p-val < 0.05 , FDR < 0.05) in the enrichment of MSigDB Collection (c2.cp.kegg.v7.2.symbols.gmt). In high expression of CTSV phenotype, 22 important tumor-related signaling

pathways were enriched (**Fig. 4 and Table 4**). As shown in **Fig. 4**, cell cycle, purine metabolism, chemokine signaling pathway, regulation of actin cytoskeleton, pathways in cancer, tight junction, focal adhesion, ubiquitin mediated proteolysis, Wnt signaling pathway, endocytosis, MAPK signaling pathway, T cell receptor signaling pathway, axon guidance, Toll-like receptor signaling pathway, spliceosome, cell adhesion molecules (CAMs), natural killer cell mediated cytotoxicity, JAK-STAT signaling pathway, insulin signaling pathway, leukocyte transendothelial migration, lysosome, and cytokine-cytokine receptor interaction were differentially enriched.

Discussion

It has been reported that CTSV is highly expressed in some tumors, including colorectal cancer, breast cancer, oral cancer and endometrial cancer[14, 17, 18], and associated with poor prognosis of liver cancer and breast cancer[19, 15]. High expression of CTSV was significantly correlated with tumor number, pathological grade, vascular invasion, T stage and clinical stage[15]. The purpose of our study was to explore the expression of CTSV in BLCA and its potential effect on prognosis, and whether it can act as a biomarker for prognosis in BLCA patients by bioinformatics.

In this study, RNA sequencing data and relevant clinical information of BLCA were downloaded from the TCGA database for bioinformatics analysis. The results showed that high expression of CTSV was correlated with clinicopathological features (histologic grade, clinical stage, T stage, etc.), OS, and poor prognosis. In order to further identify the function of CTSV in BLCA, we performed gene set enrichment analysis between high and low CTSV expression data sets. GSEA enrichment analysis showed that cell cycle, purine metabolism, chemokine signaling pathway, regulation of actin cytoskeleton, pathways in cancer, tight junction, focal adhesion, ubiquitin mediated proteolysis, Wnt signaling pathway, endocytosis, MAPK signaling pathway, T cell receptor signaling pathway, axon guidance, Toll-like receptor signaling pathway, spliceosome, cell adhesion molecules (CAMs), natural killer cell mediated cytotoxicity, JAK-STAT signaling pathway, insulin signaling pathway, leukocyte transendothelial migration, lysosome and cytokine-cytokine receptor interaction were enriched in a high phenotype. These results suggested that CTSV may be a potential biomarker for prognosis and therapeutic target in BLCA patients.

As a member of the cysteine cathepsin family, CTSV plays an important role in cancer progression, proliferation, invasion and metastasis[20]. It has been found that high expression of CTSV in hepatocellular carcinoma is associated with poor prognosis by immunohistochemistry and quantitative real-time polymerase chain reaction[15]. In our study, it was also confirmed that the high expression of CTSV in BLCA was associated with advanced clinicopathologic features and predicted poor prognosis. Cysteine cathepsin is a key acid hydrolytic enzyme in lysosomes, which can degrade extracellular matrix and promote tumor progression[11]. In our GSEA enrichment results, we also found many biological pathways associated with extracellular matrix, including regulation of actin cytoskeleton, tight junction, focal adhesion, and cell adhesion molecules (CAMs).

In this study, we enriched 22 biological pathways in a high phenotype by using GSEA. It has been reported that chemokine signaling pathway plays an important role in occurrence and development of BLCA and may be a therapeutic target[21]. Sun et al.[22] demonstrated that activation of MAPK signaling pathway promotes proliferation and migration of BLCA cells. Similarly, Wnt signaling pathway is activated in BLCA and regulates tumor growth and proliferation[23, 24]. It has been discovered that cytokine signaling 3 gene (SOCS3) regulates the cancer initiation and progression of BLCA through JAK-STAT signaling pathway[25]. The current evidence suggests that HGF/SF and IL-8 may play an important role in the development of BLCA from superficial to invasive via influencing the structure and function of tight Junction[26]. Cell adhesion molecules (CAMs) play a role in cancer progression and metastasis, and serum level of CAMs is correlated with extent of distant metastasis in BLCA[27]. The above results indicate that CTSV plays a crucial role in the progression of BLCA through these pathways.

This study has some limitations. The sample numbers are not large enough and some clinical data is still unknown, such as nearly half of the patients do not have accurate distant metastasis, which may cause some bias. In addition, the expression and molecular mechanism of CTSV in BLCA patients need to be further verified in vitro. Therefore, more experimental and clinical studies are needed for further study.

Conclusions

In summary, high expression of CTSV in BLCA is associated with poor prognosis and may serve as a new biomarker. In addition, the chemokine signaling pathway, MAPK signaling pathway, Wnt signaling pathway, JAK-STAT signaling pathway, tight Junction and cell adhesion molecules may be the key pathway regulated by CTSV in BLCA.

Abbreviations

BLCA: bladder cancer; CTSV/CTSL2: Cathepsin V/ cathepsin L2; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; GSEA: Gene Set Enrichment Analysis; NMIBC: non-muscle invasive bladder cancer; MIBC: muscle invasive bladder cancer; OS: Overall survival; TURBT: transurethral resection of bladder tumor; MSigDB: The Molecular Signatures Database; NES: the normalized enrichment score; FDR: false discovery rate

Declarations

Acknowledgments

Not applicable.

Author contributions

YZQ, ZJJ and CCH designed the study, drafted and reviewed the manuscript. LP, CJL and HDL downloaded, integrated and analyzed data. LJP and YLJ prepared the figures and tables. TJQ reviewed

and revised the manuscript. All authors read and approved the final manuscript.

Funding This study was supported by the Science and Technology project of Chengguan District, Lanzhou city, Gansu province Science and Technology Bureau (Project number:2017KJGG0052), the Cuiying Graduate Supervisor Applicant Training Program of Lanzhou University Second Hospital (201704), the Gansu Health Industry Research Project (GSWSKY2017-10), the Cuiying Scientific and Technological Innovation Program of Lanzhou University Second Hospital (Project number: CY2017-BJ16, Doctoral supervisor training program), Lanzhou City Talent Innovation and Entrepreneurship Project (Project number:2019-RC-37) and the National Nature Science Foundation of China (NO: 81372732&30800206).

Availability of data and materials

Publicly available datasets were analyzed in this study, which can be found in The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>) and Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>).

Ethics approval and consent to participate

This study did not contain any work with human participants or animals conducted by any of the authors.

Consent for publication

Not applicable.

Competing interests

The authors have no conflicts of interest to declare.

References

1. Bray F, Ferlay J, Soerjomataram I et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424.
2. Pruthi RS, Baldwin N, Bhalani V et al. Conservative management of low risk superficial bladder tumors. *J Urol.* 2008;179(1):87-90.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7-34.
4. Larré S, Catto JW, Cookson MS et al. Screening for bladder cancer: rationale, limitations, whom to target, and perspectives. *Eur Urol.* 2013;63(6):1049-58.
5. Dong F, Xu T, Shen Y et al. Dysregulation of miRNAs in bladder cancer: altered expression with aberrant biogenesis procedure. *Oncotarget.* 2017;8(16):27547-68.
6. McConkey DJ, Choi W. Molecular Subtypes of Bladder Cancer. *Curr Oncol Rep.* 2018;20(10):77.

7. Bakkar AA, Wallerand H, Radvanyi F et al. FGFR3 and TP53 gene mutations define two distinct pathways in urothelial cell carcinoma of the bladder. *Cancer Res.* 2003;63(23):8108-12.
8. Apolo AB, Vogelzang NJ, Theodorescu D. New and promising strategies in the management of bladder cancer. *Am Soc Clin Oncol Educ Book.* 2015:105-12.
9. Martin-Doyle W, Kwiatkowski DJ. Molecular biology of bladder cancer. *Hematol Oncol Clin North Am.* 2015;29(2):191-203.
10. Tan GJ, Peng ZK, Lu JP et al. Cathepsins mediate tumor metastasis. *World J Biol Chem.* 2013;4(4):91-101.
11. Gocheva V, Joyce JA. Cysteine cathepsins and the cutting edge of cancer invasion. *Cell Cycle.* 2007;6(1):60-4.
12. Brömme D, Li Z, Barnes M et al. Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization. *Biochemistry.* 1999;38(8):2377-85.
13. Adachi W, Kawamoto S, Ohno I et al. Isolation and characterization of human cathepsin V: a major proteinase in corneal epithelium. *Invest Ophthalmol Vis Sci.* 1998;39(10):1789-96.
14. Santamaría I, Velasco G, Cazorla M et al. Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas. *Cancer Res.* 1998;58(8):1624-30.
15. Jing J, Wang S, Ma J et al. Elevated CTSL2 expression is associated with an adverse prognosis in hepatocellular carcinoma. *Int J Clin Exp Pathol.* 2018;11(8):4035-43.
16. Subramanian A, Tamayo P, Mootha VK et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545-50.
17. Feng Y, Li Q, Chen J et al. Salivary protease spectrum biomarkers of oral cancer. *Int J Oral Sci.* 2019;11(1):7.
18. Skrzypczak M, Springwald A, Latrich C et al. Expression of cysteine protease cathepsin L is increased in endometrial cancer and correlates with expression of growth regulatory genes. *Cancer Invest.* 2012;30(5):398-403.
19. Toss M, Miligy I, Gorringer K et al. Prognostic significance of cathepsin V (CTSV/CTSL2) in breast ductal carcinoma in situ. *J Clin Pathol.* 2020;73(2):76-82.
20. Olson OC, Joyce JA. Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer.* 2015;15(12):712-29.
21. Nazari A, Khorramdelazad H, Hassanshahi G. Biological/pathological functions of the CXCL12/CXCR4/CXCR7 axes in the pathogenesis of bladder cancer. *Int J Clin Oncol.* 2017;22(6):991-1000.
22. Sun M, Zhao W, Chen Z et al. Circular RNA CEP128 promotes bladder cancer progression by regulating Mir-145-5p/Myd88 via MAPK signaling pathway. *Int J Cancer.* 2019;145(8):2170-81.

23. Pierzynski JA, Hildebrandt MA, Kamat AM et al. Genetic Variants in the Wnt/ β -Catenin Signaling Pathway as Indicators of Bladder Cancer Risk. *J Urol*. 2015;194(6):1771-6.

24. Schmid SC, Sathe A, Guerth F et al. Wntless promotes bladder cancer growth and acts synergistically as a molecular target in combination with cisplatin. *Urol Oncol*. 2017;35(9):544.e1-.e10.

25. Guan HJ, Li XX, Guo YP et al. Methylation of the suppressor of cytokine signaling 3 gene (SOCS3) in bladder cancer. *Int J Clin Exp Pathol*. 2017;10(11):11326-34.

26. Haynes MD, Martin TA, Jenkins SA et al. Tight junctions and bladder cancer (review). *Int J Mol Med*. 2005;16(1):3-9.

27. Perabo F, Sharma S, Gierer R et al. Circulating intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule-1 (VCAM-1) and E-selectin in urological malignancies. *Indian J Cancer*. 2001;38(1):1-7.

Tables

Table 1 Bladder cancer patient characteristics in the TCGA cohort

Characteristics		Total (408)	%
Age (year)		408 (34-90)	100%
Gender	Male	301	73.77%
	Female	107	26.23%
Race	White	324	79.41%
	Asian	44	10.78%
	Black or African American	23	5.64%
	Unknow	17	4.17%
Histologic grade	Low grade	24	5.88%
	High grade	378	92.65%
	Unknow	6	1.47%
Clinical stage	I	2	0.49%
	II	130	31.86%
	III	140	34.32%
	IV	134	32.84%
	Unknow	2	0.49%
T stage	T1	3	0.74%
	T2	119	29.17%
	T3	194	47.55%
	T4	57	13.97%
	TX	35	8.57%
N stage	N0	237	58.09%
	N1	46	11.27%
	N2	75	18.38%
	N3	8	1.97%
	NX	42	10.29%
M stage	M0	196	48.04%
	M1	11	2.70%
	MX	201	49.26%

Table 2 CTSV expression association with clinicopathologic characteristics (logistic regression) in the TCGA cohort

Clinicopathologic variable	Total(N)	Odds Ratio in CTSV expression	P-Value
Age (continuous)	408	1.012(0.993-1.031)	0.213
Gender (male vs. female)	408	1.479(0.937-2.364)	0.093
Race (White vs. others)	391	2.519(1.478-4.349)	6.59e-04
Histologic grade (low vs. high)	402	1.662(1.352-2.051)	8.02e-07
Clinical stage (I -II vs. III-IV)	406	1.589(1.043-2.422)	0.031
Status (normal vs. tumor)	427	1.435(1.181-1.748)	1.45e-04
T1-2 vs. T3-4	373	1.589(1.027-2.463)	0.038
N0 vs. N1-3	366	1.459(0.939-2.286)	0.092
M0 vs. M1	207	4.499(1.124-30.012)	0.032

Table 3 Univariate and multivariate Cox regression analyses of CTSV expression and the clinicopathological characteristics in the TCGA cohort

Clinicopathologic variable	Cox univariate analysis			Cox multivariate analysis		
	HR	95%CI	P-Value	HR	95%CI	P-Value
Age	1.034	1.018-1.051	2.26e-05	1.033	1.016-1.050	7.12e-05
Gender	1.156	0.835-1.601	0.382			
Race	1.024	0.772-1.358	0.870			
Histologic grade	0.347	0.086-1.402	0.137			
Clinical stage	1.709	1.410-2.071	4.489e-08	1.221	0.917-1.625	0.173
T stage	1.676	1.335-2.105	8.854e-06	1.459	1.088-1.958	0.012
N stage	1.497	1.222-1.833	9.636e-05	1.433	1.075-1.909	0.014
M stage	1.188	1.022-1.381	0.025	1.137	0.968-1.336	0.118
CTSV	1.662	1.209-2.286	0.002	1.495	1.069-2.089	0.019

Table 4 GSEA of high expression phenotype in MSigDB Collection (c2.cp.kegg. v7.2.symbols.gmt)

Gene set name	NES	NOM p-val	FDR q-val
KEGG_CELL_CYCLE	2.166	0.000	0.000
KEGG_PURINE_METABOLISM	2.026	0.002	0.004
KEGG_CHEMOKINE_SIGNALING_PATHWAY	2.009	0.000	0.003
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	1.989	0.000	0.003
KEGG_PATHWAYS_IN_CANCER	1.981	0.002	0.003
KEGG_TIGHT_JUNCTION	1.979	0.000	0.003
KEGG_FOCAL_ADHESION	1.942	0.007	0.006
KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS	1.927	0.002	0.006
KEGG_WNT_SIGNALING_PATHWAY	1.899	0.006	0.007
KEGG_ENDOCYTOSIS	1.882	0.002	0.008
KEGG_MAPK_SIGNALING_PATHWAY	1.875	0.002	0.008
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY	1.869	0.004	0.008
KEGG_AXON_GUIDANCE	1.849	0.004	0.009
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	1.847	0.012	0.008
KEGG_SPLICEOSOME	1.844	0.002	0.008
KEGG_CELL_ADHESION_MOLECULES_CAMS	1.804	0.021	0.012
KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY	1.798	0.012	0.011
KEGG_JAK_STAT_SIGNALING_PATHWAY	1.796	0.004	0.011
KEGG_INSULIN_SIGNALING_PATHWAY	1.790	0.004	0.011
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION	1.776	0.013	0.012
KEGG_LYSOSOME	1.767	0.029	0.012
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	1.724	0.025	0.015

Note: NES: normalized enrichment score; NOM: nominal; FDR: false discovery rate. Gene sets with NOM p-val < 0.05 and FDR q-val < 0.05 are considered as statistically significant.

Figures

Figure 1

Association with CTSV expression and clinicopathologic characteristics in the TCGA cohort, including age (A), gender (B), histologic grade (C), race (D), clinical stage (E), T stage (F), M stage (G), all tumor and normal samples (H), paired tumor and normal samples (I).

Figure 2

Kaplan-Meier survival analysis. A Impact of CTSV expression on overall survival (OS) in bladder cancer patients in TCGA cohort. B The cutoff value of CTSV evaluated by survminer package.

Figure 3

CTSV gene expression between bladder cancer tissues and normal tissues in GSE13507 dataset (A) and GSE133624 dataset (B).

Figure 4

Enrichment plots from gene set enrichment analysis (GSEA) of high expression phenotype in BLCA patients. The top 22 KEGG signaling pathways are ranked by FDR q-val. A The 1-8 KEGG signaling pathways; B The 9-15 KEGG signaling pathways; C The 16-22 KEGG signaling pathways.