Original Article
An eight-gene-methylation-based signature for prognosis of bladder cancer

Jianwu Liu, Xuebing Han

Urinary Surgery, Shanxi Provincial Cancer Hospital, Taiyuan 030013, P. R. China

Abstract: Bladder cancer (BC) is one of the most common malignancies of the urological system with heterogeneous prognosis among patients. This study aimed to identify a novel prognostic model for BC using a joint analysis of DNA methylation and gene expression. A total of 412 cases with methylation data and 408 with mRNA expression data of BC patients were retrieved from the TCGA-BLCA Project. The data were preprocessed and the methylation data were randomly divided into a training set and a test set. Differentially methylated CpG sites were identified and a sure independence screening (SIS) model for BC survival prediction was built, and then was verified in the test set. The correlation of methylated CpG loci and associated gene expression was determined in the model. In addition, the relationship between expression levels of characteristic genes and overall survival rate was analyzed. As a result, the age, race and TNM staging were not significantly different between the training group and the test group. A total of 485 hypo-methylated sites and 808 hyper-methylated sites were identified. A SIS model containing 8 CpG sits was established and verified in the test set. Seven genes mapped by 8 CpG sits had a close relationship with overall survival of BC patients. Collectively, 8 CpG loci mapping to 7 genes were found to be a reliable and practical classifier to improve prediction accuracy of prognosis and survival in BC patients.

Keywords: Bladder cancer, prognostic classifier, DNA methylation, gene expression, survival

Introduction

Bladder cancer (BC) occurring in the tissues of the urinary bladder is one of the most common cancers. The number of deaths caused by BC increased from 114,000 in 1990 to 170,000 in 2010 [1]. There were approximately 81,000 cases and 17,000 deaths from BC in the United States in 2018, and the average five-year survival rate of BC patients was 77% [2]. BC with has a recurrence rate and requires accurate prognostication and risk assessment for treatment-related decisions. Risk categories using the American Joint Committee on Cancer/tumor, nodes, and metastases (AJCC/TNM) staging system is widely used to predict survival outcomes [3], but it has been shown to be less accurate. Integrating multiple information platforms is beneficial for prognostic-based patient stratification and helping patients to follow personalized medicine and individualized therapy.

Biological markers which are recognized as molecular factors, can be used in diagnose, monitoring of disease states, and predict responses and outcomes. Kim et al. identified a progression-related gene classifier for non-muscle-invasive BC (NMIBC) to predict disease outcome that may help in selecting NMIBC patients who might benefit from more aggressive therapeutic intervention or surveillance [4, 5]. The long-term validation of this classifier could be used for clinical prediction of the development of muscle invasion in NMIBC patients [6]. Ingelmo-Torres et al. found a 2-microRNA (miRNA)-based prognostic classifier for NMIBC [7]. These biomarkers may be novel potential classifiers for prognosis of NMIBC patients, but the safety and efficacy of them in clinical applications have not been rigorously evaluated.

DNA methylation in BC patients would be an early event in cancer pathogenesis. For instance, aberrant promoter hypermethylation is a major mechanism to inhibit the expression of tumor suppressor genes [8]. Kim et al. identified three gene (*HOXA9*, *ISL1* and *ALDH1A3*)

methylation patterns, which may be used in the assessment of NMIBC recurrence and progression [8]. Some genes, such as *CDH1*, *SYNPO2* [9], *CDKN2A*, *ARF* [10], *GDF15*, *TMEFF2* and *VIM* [11] with abnormal methylation levels were associated with BC progression. The genome-wide DNA methylation patterns and the resulting aberrant gene expression at two different types of epigenetic modifications would be a valuable tool for prognosis of BC and guiding appropriate management strategies.

In this study, bioinformatics methods were used to investigate the prognostic value of methylation biomarkers for the overall survival of BC patients. Data retrieved from the Cancer Genome Atlas (TCGA) project was used to build a prognostic model. In addition, the correlation between the CpG methylation level of the prognostic model and gene expression level were analyzed, and the role of the prognostic model in the development and progression of BC from was further explained.

Materials and methods

DNA methylation data and mRNA expression data of BC

The methylation data of BC was available in the TCGA database (https://portal.gdc.cancer.gov/) [12]. The data obtained from 412 BC patients in the TCGA-BLCA Project was used, of which 21 samples had both the cancer and the normal tissues. The microarray platform was the Illumina Human Methylation 450 (HM450) arrays.

The mRNA expression data was also obtained from the TCGA database. There were 408 mRNA expression samples of BC patients retrieved from the TCGA-BLCA Project. The HTSeq-Counts data was downloaded and used for subsequent analysis.

Data preprocessing

The mRNA expression data and methylation data were normalized by the DESeq2 package [13] and the Illumina Methylation Analyzer (IMA) package [14] of R, respectively. The main step of normalizing the DESeq2 function package was as follows: first, a logarithmic conversion (log2 or log10) was performed on the reads matrix and average values of each gene (after

logarithmic conversion) was taken; then the logarithmic mean was subtracted from the logarithmic matrix to obtain the ratio matrix, and the median of the logarithmic ratio matrix of each sample was calculated. Finally, the median of the logarithm was converted to a true number, that was, the standardized factor of each sample, and the initial expression matrix was divided by this normalization factor. The function package of Illumina Methylation Analyzer (IMA) obtained the β-value representing the degree of methylation according to the original methylation data. The following analysis was based on the β-value. From 412 samples, 206 of them were randomly selected as a training set to construct a prognostic model based on methylation level of CpG sites, and the remaining 206 were used as a test set to verify the reliability of the model. Meanwhile, statistical description was conducted on 412 samples combined with clinical data, including age, gender, race, and TNM stage.

Differentially methylated CpG site detection

The 21 methylation samples from both the cancer tissue and normal tissues of BC patients were screened for differentially methylated CpG sites detection. First, the sex chromosome-derived probes (chrX, chrY) and the probes with missing values were filtered out. Then, the differentially methylated CpG sites were screened using the limma package [15] in R by the t-test method. The *p*-value after t-test were corrected using the Benjamini & Hochberg method [16], and finally the corrected *p*-value (BH_p) were obtained. The threshold of $|\Delta\beta|>0.4$ and $BH\_p < 1 \times 10^{-7}$ was used to identify differentially methylated CpG sites.

Sure independence screening (SIS) and model building

The univariate Cox regression of the "survival" Bioconductor package was used to assess the association between differentially methylated CpG loci and overall survival in the training set. The β-value of the methylation level of CpG sites was used as a continuous variable for cox regression analysis of single factor. SIS of lasso penalized regression in SIS package was used to identify candidate stable and reliable CpG sites and a model based on the methylation values of those CpG sites was built to assess the overall survival of the sample [17].
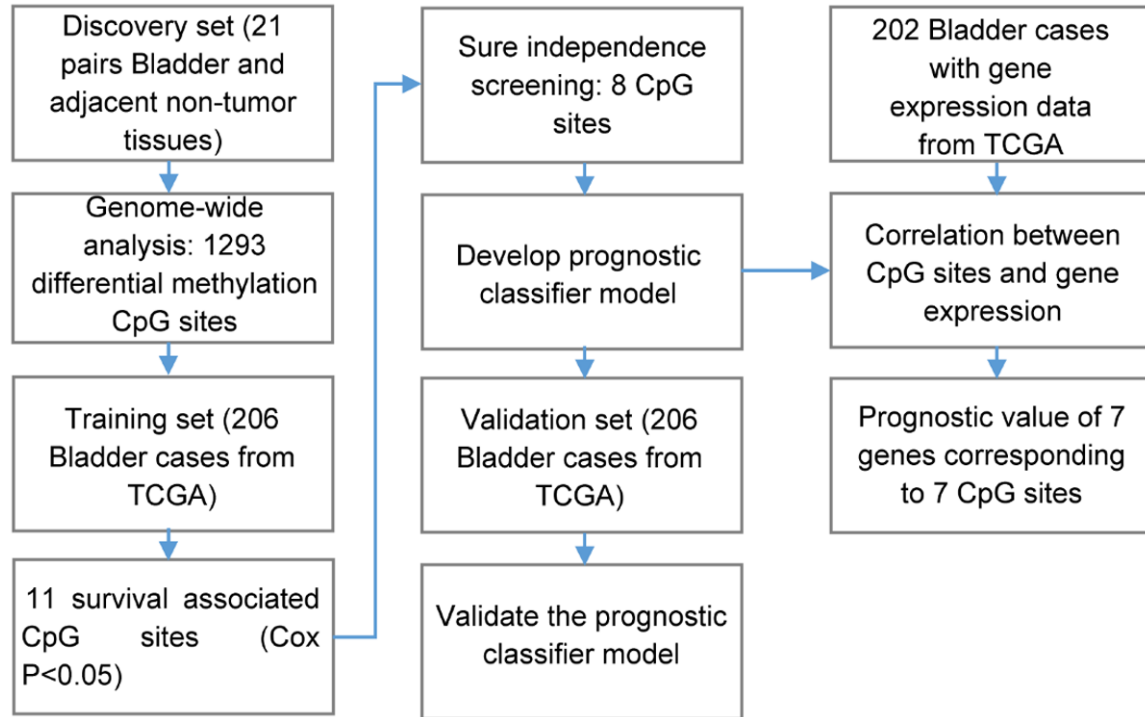
Figure 1. Flow chart describing the study design.

Model verification

The model classifier obtained from the training set was applied to calculate the scores of each sample in the training set and the test set. The training set was divided into two groups based on the scores and the Kaplan-Meier survival curves were drawn. The log-rank tests were used to compare the survival between the two groups. The same analysis was also carried out for BC samples in the testing set to validate the reliability of the model.

Correlation analysis between CpG loci and associated gene expression

The characteristic CpGs loci of the model classifier were mapped to genes. Then the correlation between methylation levels of the CpG loci and expression levels of genes was analyzed. Furthermore, the relationship between expression levels of characteristic genes and overall survival rate was also explored through survival analysis.

Statistical analysis

Kaplan-Meier method was used to estimate the overall survival rate of different groups, and log-rank was used to test the significance of the difference in survival rate between different groups, with P<0.05 as the significance threshold. Statistical analysis was carried out by R 3.5.2 software.

Results

Data characteristics

The pipeline of the study is shown in **Figure 1**. The clinical data of 412 BC patients, including age, gender, race, and TNM staging are shown in **Table 1**. The result of chi-square test revealed that age, race and TNM staging were not significantly different between the training group and the test group (P>0.05) but that of gender was P=0.0187.

Identification of differentially methylated CpG sites

Genome-wide differential methylation analysis was conducted on both tumor and adjacent non-tumor tissues of 21 BC patients. As a result, 1293 differentially methylated CpGs were obtained with the thresholds of $|\Delta\beta|>0.4$ and BH_p$<1 \times 10^{-7}$, including 485 hypo-methyl-

Table 1. Demographic and clinical characteristic of bladder cancer patients in sets

Characteristic	Training set (n=206)	Validation set (n=206)	Chi-square test (P)
Age			0.7985
Range	38-90	35-90	
Median years	71	68	
Gender, n (%)			0.0187
Male	141 (68.5)	163 (79.1)	
Female	65 (31.5)	43 (20.9)	
Race, n (%)			
White	166 (80.6)	161 (78.2)	0.4882
Asian	22 (10.7)	22 (10.7)	
Black or African American	8 (3.8)	15 (7.3)	
No reported	10 (4.9)	8 (3.8)	
TNM stage, n (%)			0.3658
Early (I-II)	67(32.5)	66 (32.0)	
Advance (III-IV)	139 (67.5)	138 (67.0)	
Not reported	0 (0.0)	2 (1.0)	

ated sites and 808 hyper-methylated sites (**Figure 2A** and **2B**).

Construction of SIS model

Using univariate Cox regression, the association between differentially methylated CpG sites and overall survival of BC patients in the training set, found 11 CpG sites were identified (P<0.05, **Table 2**; **Figure 2C**), including cg18-693395, cg10040329, cg12421755, cg232-44488, cg23900203, cg11827910, cg0569-7849, cg16886987, cg16787600, cg2001-4398, cg09671258. Allocation of those CpG sites to genes is provided in **Table 3**. **Figure 2C** illustrates the methylation level of those 11 CpG sites in normal and tumor bladder tissues. In addition, 8 CpG sites were identified as the model characteristic quantities through SIS analysis. The final predictive model consisted of the methylation levels and the multivariate weights of the eight CpG sites: $\text{score}_{\text{methylation}}$ = 0.08595633 × $\text{cg12421755}_{ONECUT1}$ + 0.1160-8698 × $\text{cg18693395}_{PTGDR}$ + 0.11535279 × $\text{cg2-3244488}_{TWIST1}$ + 0.07405796 × $\text{cg239002-03}_{RP11}$ + 0.41573496 × $\text{cg05697849}_{ELAVL4}$ - 0.35073355 × $\text{cg16886987}_{RIPPLY3}$ + 0.38-575295 × g20014398_{PAX6} - 0.31391318 × g09671258_{LHX4}.
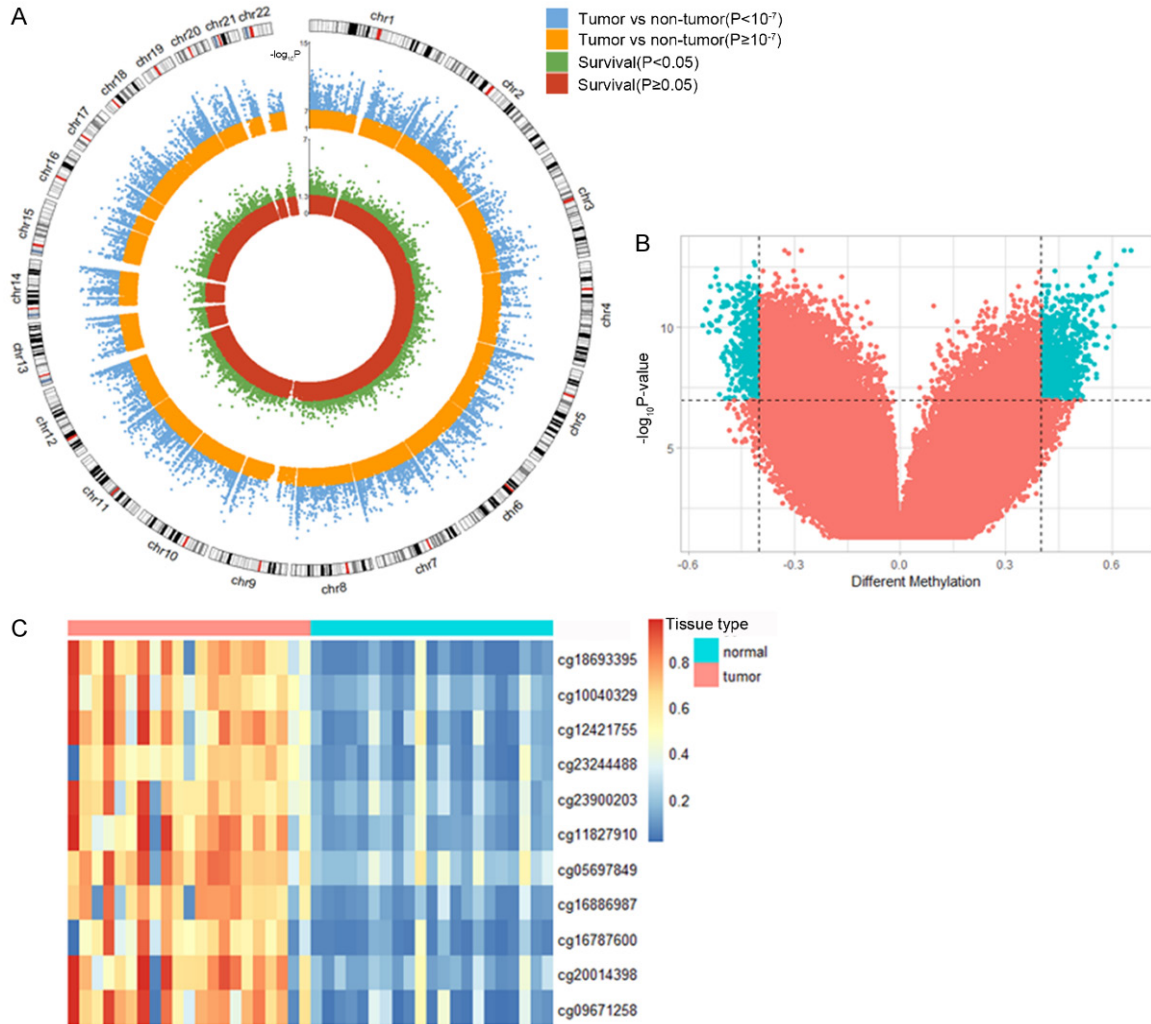
Verification of SIS model

We used 0.35 as the threshold for model scoring and divided the training samples into the low-risk and the high-risk group, which was determined by Kaplan-Meier survival analysis and log-rank test according to the highest χ^2 value. **Figure 3** displayed that a hazard ratio of 4.72 for the high-risk group and the low-risk group was obtained in the training set (P=5.82 × 10^{-8}). Similarly, the test set also divided patients into two groups using the same scoring, and the risk of death in the high-risk group was 1.93 times higher than that in the low-risk group (P=0.015).

The close correlation analysis of CpG methylation, gene expression and prognosis

The expression data were available in 202 cases of training set. The correlation analysis between expression level of seven CpG sits-mapped genes and methylation levels was performed. It was observed that the methylation levels of CpG sites in PTGDR, *TWIST1* and *ELAVL4* were negatively correlated with their expression levels, while the methylation levels of CpG sites in *ONECUT1*, *RIPPLY3*, *PAX6* and *LHX4* were positively correlated with their expression levels (**Figure 4A-G**, left panel). Moreover, the relationship between the expression levels of these seven genes and the overall survival of BC patients was further analyzed. The highly expressed *ELAVL4*, *RIPPLY3*, and *LHX4* were significantly associated with a shorter survival of patients, while highly expressed *PAX6* was associated with a longer survival of patients (**Figure 4A-G**, right panel).

Figure 2. The eight-CpG-based prognosis classifier. A. Circos plot of epigenome-wide DNA methylation CpG sites. Results are presented as *p* values ordered by genomic position, the t-test in the differential methylation analysis (orange and blue symbols) and the univariate Cox regression analysis of the training set (green and red symbols). B. Volcano plot comparing CpG methylation for BC tumor and non-tumor tissues. A total of 1293 CpG sites had an absolute value of differential methylation ($|\Delta\beta|$) of >0.4 and t test *p* value (BH_p) of $<1 \times 10^{-7}$ (blue dots). C. Heatmap showing methylation levels of 11 CpG sites in tumor tissues and normal tissues.

Table 2. The coefficient of association in the Cox-regression analysis

|	coef	exp (coef)
cg18693395	-0.9905	0.3714
cg10040329	-1.4568	0.233
cg12421755	-0.5975	0.5502
cg23244488	-0.8879	0.4115
cg23900203	-1.1659	0.3116
cg11827910	-0.9216	0.3979
cg05697849	-0.9481	0.3875
cg16886987	-0.6672	0.5131
cg16787600	-0.8699	0.419
cg20014398	-0.8807	0.4145
cg09671258	-1.0853	0.3378

Discussion

Cancer is a complicated and multifaceted disease involving multiple regulatory network [18]. Integrating many biomarkers into an aggregated model could increase the prognostic value compared to a single biomarker [19]. Biomarkers have been discovered in bladder cancer [20, 21], However, few people apply the TCGA dataset to investigate biomarkers in different omics. In this study, a BC prognostic classifier model that included 8 CpG loci was built, and the model was validated using a test set. The results showed that certain prognostic features were significantly associated with overall survival in patients with BC, and patients with high-

Table 3. CpGs that significantly associated with BC overall survival in univariate Cox-regression analysis and their corresponding genes

CpGs	Gene symbol	Gene name
cg18693395	PTGDR	prostaglandin D2 receptor
cg10040329	RFX4	regulatory factor X4
cg12421755	ONECUT1	one cut domain, family member 1
cg23244488	NA	NA
cg23900203	NA	NA
cg11827910	PAX6	paired box 6
cg05697849	ELAVL4	ELAV like RNA binding protein
cg16886987	DSCR6	ripply transcriptional repressor 3
cg16787600	SORCS3	sortilin related VPS10 domain containing receptor 3
cg20014398	PAX6	paired box 6
cg09671258	LHX4	LIM homeobox 4
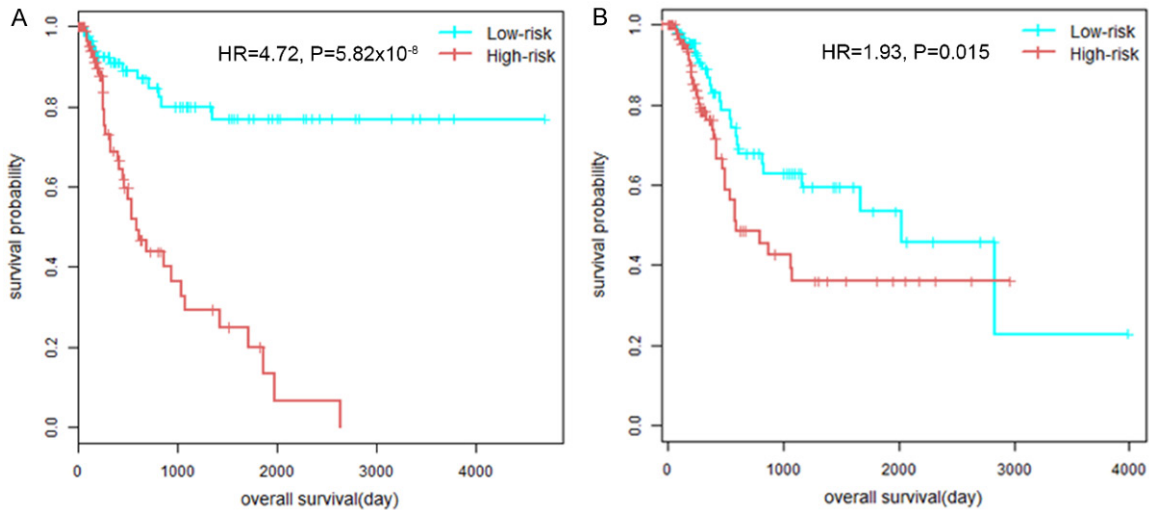
Figure 3. Kaplan-Meier survival analysis of training set (A) and test set (B). Patients were divided into two groups by the risk score obtained through the Sure independence screening (SIS) model. *P* values were calculated using log-rank test, and HR indicates hazard ratio.
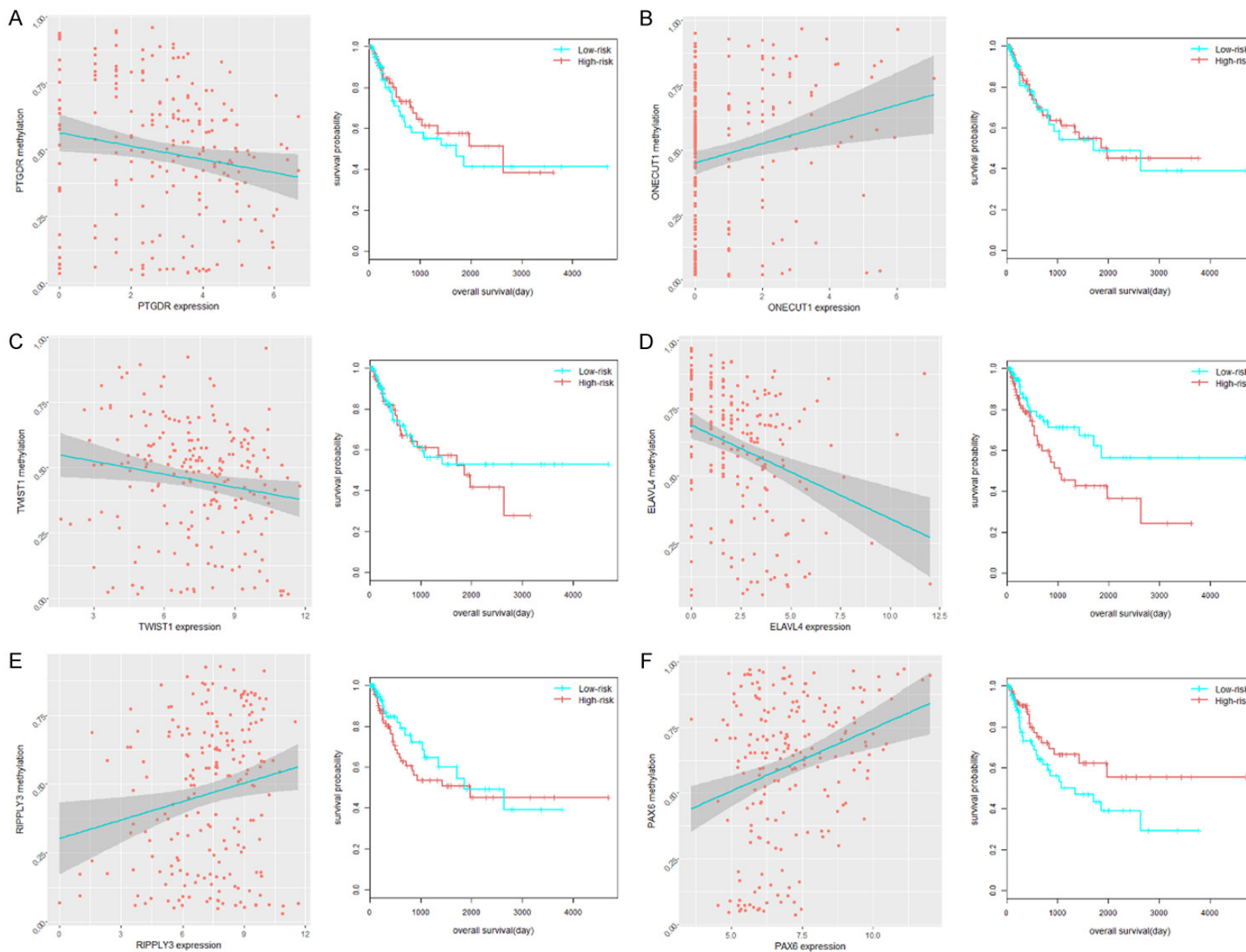
er prognostic scores tended to have poorer clinical outcomes.

Kandimalla et al. found that the methylation of TBX2, TBX3, GATA2, and ZIC4 could be considered to predict pTa-specific progression in bladder cancer [22]. Another study revealed that the methylation of GRIA1 was strongly related to bladder cancer and could be recognized as a potential noninvasive biomarker for basal-like bladder cancer [23]. An additional study validated the methylation of HOXA9, PCDH17, POU4F2, and ONECUT2 as urinary biomarkers for the detection of bladder cancer in Chinese patients with hematuria [24]. These studies paid attention to the discovery of noninvasive

diagnostic biomarkers and establishment of combined models for the detection of bladder cancer. In this study, we investigated differentially methylated genes between bladder cancer and normal tissues and to assess their relevance as prognostic biomarkers. According to the previous studies, seven genes mapped by the CpG loci in prognostic classifier model were associated with BC or other types of cancers.

The methylation levels of CpG sites in *ONECUT1*, *RIPPLY3*, *PAX6* and *LHX4* were positively correlated with their expression levels. There was no relationship between transcription factor *ONECUT1* (*HNF6*) and BC, while down-regulated *ONECUT1* was correlated with pancreatic
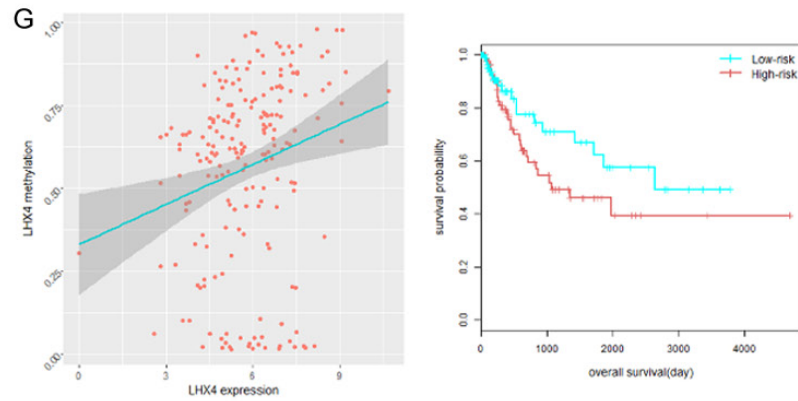
Figure 4. Correlation analysis and survival analysis of characteristic CpG loci and corresponding gene expression. Left panels show correlation of PTGDR (A), ONE-CUT1 (B), TWIST1 (C), and ELAVL4 (D) RIPPLY3 (E), PAX6 (F) and LHX4 and (G) expression (X-axis) with CpG sites methylation (Y-axis). Right panels show Kaplan-Meier survival plots of gene expression from the TCGA-BLCA Project. HR indicates hazard ratio. Patients were categorized into high-risk and low-risk groups based on the gene's expression value by an optimum cutoff point according to the highest χ^2 value.

development and pancreatic cancer progression [25]. *ONECUT1* was positively related to tumor suppressor p53 regulating lung cancer cell plasticity [26]. *RIPPLY3* is a transcriptional corepressor that negatively regulates the transcriptional activity of TBX1 which is involved in parathyroid tumorigenesis [27]. Previous study indicated that the methylation level of *PAX6*-promoters was increased in early BC and may represent a biomarker for the disease [28]. Frequent CpG hypermethylation of *LHX4* was discovered in tumor samples [29]. While, the methylation levels of *PTGDR*, *TWIST1* and *ELAVL4* were negatively correlated with gene expression levels. Thomas et al. validated that methylation of eight genes including *PTGDR* that may be promising markers for early BC detection [30]. Hypermethylated *TWIST1* was identified as a prognostic factor for BC [31, 32]. It was found that EL*AVL4* was associated with small cell lung cancer [33] and neuroblastoma [34]. These results suggest the correlation between methylation levels of seven genes and BC progression.

In addition to CpG methylation, the mRNA levels of *ELAVL4*, *RIPPLY3*, *LHX4* and *PAX6* were found to be significantly related to the prognosis of BC patients. *LHX4* which functions as a potential tumor suppressor was overexpressed in tumors of soft tissue or muscle tissue, such as in human hepatocellular carcinoma [35] and colorectal cancer [36]. The transcription factor *PAX6* is a tumor suppressor that is overexpressed in cancers [37, 38]. Our results indicate the reliability of 8 CpG loci as a prognostic molecular classifiers for BC patients.

However, the relationship between gene expression levels and patient survival was not particularly significant. Three genes of *ONECUT1*, *PTGDR*, and *TWIST1* had no significant characteristics from the survival curve, which may be due to the absence of gender-related differences in the grouping process. Besides, this study analyzed the effects of gene methylation levels and gene expression levels on the prognosis of bladder cancer from the statistical analyses. It was interesting to note that most of the methylation's effect might function beyond just affecting expression, but may also be related to gene function [39]; investigation of which will require further functional experiments.

In conclusion, 8 CpG loci mapping to 7 genes could be a reliable and practical classifier to improve the prediction accuracy of prognosis and survival. This SIS model may be a novel prediction tool for appropriate clinical adjuvant trials. Further external validation, including consideration of age, BC stage and other clinical characteristics to incorporate into the prognostic classifiers for routine clinical practice will be conducted.

Disclosure of conflict of interest

None.

Address correspondence to: Xuebing Han, Urinary Surgery, Shanxi Provincial Cancer Hospital, No. 3 Zhigong Xinjie, Xinghualing District, Taiyuan, Shanxi Province, P. R. China. Tel: +86-0351-4651314; E-mail: hxbdoctor@21cn.com

References

[1] Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R, Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo JP, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KM, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA 3rd, Porrini E, Pourmalek F, Raju M, Ranganathan D,

Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, De León FR, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA, Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh PH, Yip P, Zabetian A, Zheng ZJ, Lopez AD, Murray CJ, AlMazroa MA and Memish ZA. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012; 380: 2095-2128.

[2] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019; 69: 7-34.

[3] Tran B, Roshan D, Abraham E, Wang L, Garibotto N, Wykes J, Campbell P and Ebrahimi A. An analysis of the American joint committee on cancer 8th edition T staging system for papillary thyroid carcinoma. J Clin Endocrinol Metab 2018; 103: 2199-2206.

[4] Kim WJ, Kim EJ, Kim SK, Kim YJ, Ha YS, Jeong P, Kim MJ, Yun SJ, Lee KM, Moon SK, Lee SC, Cha EJ and Bae SC. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. Mol Cancer 2010; 9: 3.

[5] 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.

[6] Kang HW, Seo SP, Jeong P, Ha YS, Kim WT, Kim YJ, Lee SC, Yun SJ and Kim WJ. Long-term validation of a molecular progression-associated gene classifier for prediction of muscle invasion in primary non-muscle-invasive bladder cancer. Oncol Lett 2017; 14: 2468-2474.

[7] Ingelmo-Torres M, Lozano JJ, Izquierdo L, Carrion A, Costa M, Gomez L, Ribal MJ, Alcaraz A and Mengual L. Urinary cell microRNA-based prognostic classifier for non-muscle invasive bladder cancer. Oncotarget 2017; 8: 18238-18247.

[8] Wang X, Chen E, Yang X, Wang Y, Quan Z, Wu X and Luo C. 5-azacytidine inhibits the proliferation of bladder cancer cells via reversal of the aberrant hypermethylation of the hepaCAM gene. Oncol Rep 2016; 35: 1375-1384.

[9] Kandimalla R, van Tilborg AA and Zwarthoff EC. DNA methylation-based biomarkers in bladder cancer. Nat Rev Urol 2013; 10: 327-335.

[10] Hoque MO, Begum S, Topaloglu O, Chatterjee A, Rosenbaum E, Van Criekinge W, Westra WH, Schoenberg M, Zahurak M, Goodman SN and Sidransky D. Quantitation of promoter methyl-ation of multiple genes in urine DNA and bladder cancer detection. J Natl Cancer Inst 2006; 98: 996-1004.

[11] Costa VL, Henrique R, Danielsen SA, Duarte-Pereira S, Eknaes M, Skotheim RI, Rodrigues A, Magalhaes JS, Oliveira J, Lothe RA, Teixeira MR, Jeronimo C and Lind GE. Three epigenetic biomarkers, GDF15, TMEFF2, and VIM, accurately predict bladder cancer from DNA-based analyses of urine samples. Clin Cancer Res 2010; 16: 5842-5851.

[12] Tomczak K, Czerwinska P and Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015; 19: A68-77.

[13] Yang L, Guell M, Niu D, George H, Lesha E, Grishin D, Aach J, Shrock E, Xu W, Poci J, Cortazio R, Wilkinson RA, Fishman JA and Church G. Genome-wide inactivation of porcine endogenous retroviruses (PERVs). Science 2015; 350: 1101-1104.

[14] Rochtus A, Izzi B, Vangeel E, Louwette S, Wittevrongel C, Lambrechts D, Moreau Y, Winand R, Verpoorten C, Jansen K, Van Geet C and Freson K. DNA methylation analysis of Homeobox genes implicates HOXB7 hypomethylation as risk factor for neural tube defects. Epigenetics 2015; 10: 92-101.

[15] Neumeyer S, Popanda O, Edelmann D, Butterbach K, Toth C, Roth W, Blaker H, Jiang R, Herpel E, Jakel C, Schmezer P, Jansen L, Alwers E, Benner A, Burwinkel B, Hoffmeister M, Brenner H and Chang-Claude J. Genome-wide DNA methylation differences according to oestrogen receptor beta status in colorectal cancer. Epigenetics 2019; 14: 477-493.

[16] van der Meer AJ, Scicluna BP, Moerland PD, Lin J, Jacobson EW, Vlasuk GP and van der Poll T. The selective Sirtuin 1 activator SRT2104 reduces endotoxin-induced cytokine release and coagulation activation in humans. Crit Care Med 2015; 43: e199-202.

[17] Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society 2011; 73: 273-282.

[18] Fillon M. Making sense of the mountains of new cancer data. J Natl Cancer Inst 2017; 109: 1-2.

[19] Cohen-Solal A, Laribi S, Ishihara S, Vergaro G, Baudet M, Logeart D, Mebazaa A, Gayat E, Vodovar N, Pascual-Figal DA and Seronde MF. Prognostic markers of acute decompensated heart failure: the emerging roles of cardiac biomarkers and prognostic scores. Arch Cardiovasc Dis 2015; 108: 64-74.

[20] Jin F, Thaiparambil J, Donepudi SR, Vantaku V, Piyarathna DWB, Maity S, Krishnapuram R, Putluri V, Gu F, Purwaha P, Bhowmik SK, Ambati CR, von Rundstedt FC, Roghmann F, Berg

S, Noldus J, Rajapakshe K, Gödde D, Roth S, Störkel S, Degener S, Michailidis G, Kaipparettu BA, Karanam B, Terris MK, Kavuri SM, Lerner SP, Kheradmand F, Coarfa C, Sreekumar A, Lotan Y, El-Zein R and Putluri N. Tobacco-specific carcinogens induce hypermethylation, DNA adducts, and DNA damage in bladder cancer. Cancer Prev Res (Phila) 2017; 10: 588-597.

[21] Cheng X, Liu X, Liu X, Guo Z, Sun H, Zhang M, Ji Z and Sun W. Metabolomics of non-muscle invasive bladder cancer: biomarkers for early detection of bladder cancer. Front Oncol 2018; 8: 494.

[22] Kandimalla R, van Tilborg AA, Kompier LC, Stumpel DJ, Stam RW, Bangma CH and Zwarthoff EC. Genome-wide analysis of CpG island methylation in bladder cancer identified TBX2, TBX3, GATA2, and ZIC4 as pTa-specific prognostic markers. Eur Urol 2012; 61: 1245-1256.

[23] Tilley SK, Kim WY and Fry RC. Analysis of bladder cancer tumor CpG methylation and gene expression within The Cancer Genome Atlas identifies GRIA1 as a prognostic biomarker for basal-like bladder cancer. Am J Cancer Res 2017; 7: 1850-1862.

[24] Wu Y, Jiang G, Zhang N, Liu S, Lin X, Perschon C, Zheng SL, Ding Q, Wang X, Na R, Sun J and Xu J. HOXA9, PCDH17, POU4F2, and ONECUT2 as a urinary biomarker combination for the detection of bladder cancer in Chinese patients with hematuria. Eur Urol Focus 2020; 6: 284-291.

[25] Pekala KR, Ma X, Kropp PA, Petersen CP, Hudgens CW, Chung CH, Shi C, Merchant NB, Maitra A, Means AL and Gannon MA. Loss of HNF6 expression correlates with human pancreatic cancer progression. Lab Invest 2014; 94: 517-527.

[26] Yuan XW, Wang DM, Hu Y, Tang YN, Shi WW, Guo XJ and Song JG. Hepatocyte nuclear factor 6 suppresses the migration and invasive growth of lung cancer cells through p53 and the inhibition of epithelial-mesenchymal transition. J Biol Chem 2013; 288: 31206-31216.

[27] Verdelli C, Avagliano L, Guarnieri V, Cetani F, Ferrero S, Vicentini L, Beretta E, Scillitani A, Creo P, Bulfamante GP, Vaira V and Corbetta S. Expression, function, and regulation of the embryonic transcription factor TBX1 in parathyroid tumors. Lab Invest 2017; 97: 1488-1499.

[28] Hellwinkel OJ, Kedia M, Isbarn H, Budaus L and Friedrich MG. Methylation of the TPEF- and PAX6-promoters is increased in early bladder cancer and in normal mucosa adjacent to pTa tumours. BJU Int 2008; 101: 753-757.

[29] Rauch T, Li H, Wu X and Pfeifer GP. MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. Cancer Res 2006; 66: 7939-7947.

[30] Reinert T, Modin C, Castano FM, Lamy P, Wojdacz TK, Hansen LL, Wiuf C, Borre M, Dyrskjot L and Orntoft TF. Comprehensive genome methylation analysis in bladder cancer: identification and validation of novel methylated genes and application of these as urinary tumor markers. Clin Cancer Res 2011; 17: 5582-5592.

[31] Fantony JJ, Abern MR, Gopalakrishna A, Owusu R, Jack Tay K, Lance RS and Inman BA. Multi-institutional external validation of urinary TWIST1 and NID2 methylation as a diagnostic test for bladder cancer. Urol Oncol 2015; 33: 387, e381-386.

[32] Song YH, Shiota M, Yokomizo A, Uchiumi T, Kiyoshima K, Kuroiwa K, Oda Y and Naito S. Twist1 and Y-box-binding protein-1 are potential prognostic factors in bladder cancer. Urol Oncol 2014; 32: 31, e31-37.

[33] Wang F, Lu J, Li S, Huo X, Liu X, Du X, Li C, Wang J and Chen Z. Application of serum ELAVL4 (HuD) antigen assay for small cell lung cancer diagnosis. Anticancer Res 2017; 37: 4515-4522.

[34] Swerts K, De Moerloose B, Dhooge C, Vandesompele J, Hoyoux C, Beiske K, Benoit Y, Laureys G and Philippe J. Potential application of ELAVL4 real-time quantitative reverse transcription-PCR for detection of disseminated neuroblastoma cells. Clin Chem 2006; 52: 438-445.

[35] Hung TM, Hu RH, Ho CM, Chiu YL, Lee JL, Jeng YM, Shih DT and Lee PH. Downregulation of alpha-fetoprotein expression by LHX4: a critical role in hepatocarcinogenesis. Carcinogenesis 2011; 32: 1815-1823.

[36] Cha N, Liu W, Yang N, Xie S, Gao Y, Chen X, Wang X and Ren J. Oncogenicity of LHX4 in colorectal cancer through Wnt/beta-catenin/TCF4 cascade. Tumour Biol 2014; 35: 10319-10324.

[37] Hegge B, Sjottem E and Mikkola I. Generation of a PAX6 knockout glioblastoma cell line with changes in cell cycle distribution and sensitivity to oxidative stress. BMC Cancer 2018; 18: 496.

[38] Little EC, Kubic JD, Salgia R, Grippo PJ and Lang D. Canonical and alternative transcript expression of PAX6 and CXCR4 in pancreatic cancer. Oncol Lett 2017; 13: 4027-4034.

[39] Schubeler D. Function and information content of DNA methylation. Nature 2015; 517: 321-326.