PORTLAND PRESS

## Research Article

# Development of a novel prognostic signature for predicting the overall survival of bladder cancer patients

**Huamei Tang**[*], **Lijuan Kan**[*], **Tong Ou**[*], **Dayang Chen, Xiaowen Dou, Wei Wu, Xiang Ji, Mengmeng Wang, Zengyan Zong, Hongmei Mo, Xiuming Zhang** and (iD) **Dan Xiong**

Medical Laboratory of the Third Affiliated Hospital of Shenzhen University, Shenzhen 518001, China

**Correspondence:** Dan Xiong (xiongdan930@email.szu.edu.cn) or Xiuming Zhang (zxm0760@163.com)

OPEN ACCESS

Background: Bladder cancer is one of the most common malignancies. So far, no effective biomarker for bladder cancer prognosis has been identified. Aberrant DNA methylation is frequently observed in the bladder cancer and holds considerable promise as a biomarker for predicting the overall survival (OS) of patients. Materials and methods: We downloaded the DNA methylation and transcriptome data for bladder cancer from The Cancer Genome Atlas (TCGA), a public database, screened hypo-methylated and up-regulated genes, similarly, hyper-methylation with low expression genes, then retrieved the relevant methylation sites. Cox regression analysis was used to identify a nine-methylation site signature of a training group. Predictive ability was validated in a test group by receiver operating characteristic (ROC) analysis. Results: We identified nine bladder cancer-specific methylation sites as potential prognostic biomarkers and established a risk score system based on the methylation site signature to evaluate the OS. The performance of the signature was accurate, with area under curve was 0.73 in the training group and 0.71 in the test group. Taking clinical features into consideration, we constructed a nomogram consisting of the nine-methylation site signature and patients' clinical variables, and found that the signature was an independent risk factor. Conclusions: Overall, the significant nine methylation sites could be novel prediction biomarkers, which could aid in treatment and also predict the overall survival likelihoods of bladder cancer patients.

## Introduction

Bladder cancer is a genomically heterogeneous malignancy with nearly 430,000 cases diagnoses worldwide and over 165,000 deaths from this disease in 2012 [1]. Urothelial carcinoma accounts for >90% of all bladder cancers and is classified into two clinical phenotypes: superficial nonmuscle-invasive bladder cancer and muscle-invasive bladder cancer [2,3].

A variety of factors, such as environmental or occupational exposure to tobacco and biological infection by Schistosoma haematobium, were proved to be associated with the onset of bladder cancer [4,5]. Some hereditary genetic alterations that are involved in carcinogen bioactivation and detoxification may possibly make an individual more susceptible to extrinsic carcinogens [6,7]. Moreover, the accumulation of genetic changes and epigenetic alterations may contribute to the initiation and progression of human cancers, including bladder cancer [8–11].

DNA methylation is conventionally negatively associated with gene expression [12]. The dynamic regulation of DNA methylation is a key process involved in a wide variety of functions that include the maintenance of genomic integrity and stability, gene regulation and cellular differentiation [13,14]. In cancer and other pathological conditions, the methylation levels of DNA may change to hypo-methylation (the state

**1**

of decreased methylation) or hyper-methylation (the state of increased methylation) level [15]. Therefore, DNA methylation-based biomarkers hold considerable promise for predicting the biological behavior of cancer and the survival of patients [16].

The Cancer Genome Atlas (TCGA) is one of the most comprehensive sequencing databases of cancer, which contains methylation data for 28 cancers [17]. In the present study, transcriptome and methylation data for bladder cancer were downloaded from the TCGA. Using a series of statistical methods, we screened a nine-gene site signature that was significantly correlated with bladder cancer patients' overall survival (OS) in a training group, established a risk score system based on the signature and illustrated the risk score in a nomogram. We also validated the signature with information from a test group. Furthermore, take the clinical characteristic into consideration, a more comprehensive nomogram was constructed that integrated the clinical characteristics.

# Materials and methods
## Data retrieval and screen differently methylated genes (DMGs)
Level 3 transcriptome data and DNA methylation profiles were downloaded from TCGA (http://cancergenome.nih.gov/) data portal using the R project (version 3.5.1, http://www.r-project.org/) [17]. DNA methylation data contained preprocessed data in the form of β values, which indicate the degree of methylation. DNA methylation data evaluated on the Illumina Infinium Human Methylation 450 platform (San Diego, CA) were included in our study. We identified differentially expressed genes (DEGs) by using the edgeR package and DMGs by using the Limma package in the R software with fold-changes $\geq 2$ [18]. $P$ values and false discovery rates (FDRs) less than 0.05 were considered significant [19]. We screened genes coexisting in the two groups with overlapping hypo-methylated and up-regulated genes, similarly, hyper-methylation with low expression genes, then retrieved the corresponding methylation sites of the genes.

## Functional and pathway enrichment analyses of the DMGs
Biological annotation of DMGs was performed using the Database for Annotation Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/), which is a widely used web-based tool for analyzing enriched Gene Ontology (GO) terms and pathways. In the present study, it was used to perform GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of the DMGs. Moreover, the criterion for significant enrichment of GO terms and KEGG pathways was $P$ value $< 0.05$.

## Formulation of DMGs signatures
Bladder cancer samples from patients with clinical characteristics and survival data were randomly divided into two groups: a training group (265 patients) and a test group (133 patients). We performed a univariate Cox proportional hazards regression analysis to assess the correlation between patient OS and methylation sites in the training group with $P<0.5$. To access a more powerful predictive diagnostic and prognostic signature, we performed a multivariate Cox regression. The model with minimum AIC was screened by R. A risk scoring formula was construct based on the Cox regression analysis. The risk scoring formula is shown in the following equation: Risk Score (RS) $= \sum_{i=1}^{N}(Exp * Coef)$, which was used to calculate the signature RS for a patient. $N$ is the representative number of prognostic DNA methylation sites; Exp is the degree of DNA methylation; and Coef is a single factor means Cox regression coefficient. Hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated based on the Cox regression analysis. Moreover, a nomogram model was established based on the results of the multivariate Cox regression analyses to predict 1-, 2- and 3-year OS. The performance of the nomogram was evaluated by Harrel's concordance index (C-index), which comparing the predicted and actual probabilities for OS. The value of the C-index ranges from 0.5 to 1.0, the closer to 1.0 the better ability to correctly discriminate the outcome with the model [20]. To provide more comprehensive guidance for clinical diagnosis and treatment, multivariate Cox regression analysis was used to establish a nomogram that integrated the risk scoring model and the clinical information.

## Statistical analysis
Receiver operating characteristic (ROC) and Kaplan–Meier survival analyses were used to validate prognostic ability of the DNA methylated site signature in the training and test groups. We used ROC curve to analyze the specificity and sensitivity across a range of cutoff points for the predictive signature and calculated the area under curve (AUC). We chose the median risk score as a cutoff point in the training and test cohort and divided the bladder cancer patients into low- and high-risk groups. Kaplan–Meier survival analysis was used to estimate the OS of the two groups, and log-rank test was used to evaluate the statistical significance.

**Table 1 Summary of patient demographics and characteristics**

| Characteristic | Training (*N*=265, three patients without stage information) | Test (*N*=133, one patient without stage information) |
| --- | --- | --- |
| Gender | | |
| Female | 65 (24.5%) | 42 (31.6%) |
| Male | 200 (75.5%) | 91 (68.4%) |
| Age | | |
| <60 years | 64 (24.2%) | 25 (18.8%) |
| ≥60 years | 201 (75.8%) | 108 (81.2%) |
| Stage | | |
| I | 2 (0.8%) | 2 (1.5%) |
| II | 77 (29.1%) | 46 (34.6%) |
| III | 89 (33.6%) | 45 (33.8%) |
| IV | 94 (35.5%) | 39 (29.3%) |
| Vital status | | |
| Living | 165 (62.3%) | 75 (56.4%) |
| Dead | 100 (37.7%) | 58 (43.6%) |

# Results

## Data processing and patient characteristics

The transcriptome profile included RNA sequencing data from 421 bladder cancer samples and 19 normal tissue samples; the methylation dataset contained 416 bladder cancer samples and 21 adjacent tissues. Among them, 409 samples had clinical information, such as sex, age and stage. A total of 398 patients clinically diagnosed with bladder cancer were randomly divided into a training group (*n*=265) and a test group (*n*=133). According to the International Union Against Cancer tumor-node-metastasis classification system for bladder tumors, clinical stages of the tumors were classified into stages I to IV (Table 1). The flowchart for the present study is shown in Figure 1.

## Identification and enrichment analysis of DMGs

We identified 4776 DEGs from the expression profile and 368 DMGs from the methylation profile, of which 88 genes were overlapped and 19 genes had inverse relationships between methylation and expression, which indicated the hypo-methylated with up-regulated genes as well as hyper-methylation with low expression genes (Figure 2). To access more information about the DMGs, enrichment analysis with the DAVID was used to elucidate their biological functions [21]. The top significant terms emerging from the GO analysis are shown in Figure 3A. DMGs were enriched in the 'transcription from the RNA polymerase II promoter', 'positive regulation of transcription' and 'transcriptional activator activity' terms, which suggested that DMGs may play important roles in regulating transcription in bladder cancer. KEGG analysis indicated that DMGs were significantly enriched in pathways related to cancer, such as 'viral carcinogenesis' and 'transcriptional misregulation in cancer' (Figure 3B).

## Identifying DNA methylation gene sites in the training group

About 318 methylation sites were located within the 19 selected DMGs. We screened 398 bladder cancer patients with clinical information, such as the sex, age and clinical stage of patients, and randomly divided them into a training group (*n*=265) and a test group (*n*=133). The training group was used to explore the relationship between OS and methylated gene sites. A Cox univariate analysis of the methylation sites with OS was performed, and 32 methylation sites were showed a significant correlation with OS ($P<0.05$, Figure 4A). We next used a Cox multivariate regression analysis to access an optimal prognostic model, composed of cg02948884, cg16331920, cg06397837, cg00049033, cg04101060, cg27096779, cg00379720, cg14068328, cg27580050 from the 32 methylation sites with the minimum Akaike information criterion (AIC) and $P<0.05$ (log-rank test; Figure 4B). Harrel's concordance index (C-index) for OS prediction was 0.66. In general, a predictive model with a lower AIC shows a better model fit, while a higher C-index indicates a better discriminating ability. The nine methylation sites were closely related to patient survival. Six DNA methylation sites were presumed to be risk factors (HR > 1), while three were protective factors (HR < 1) (Figure 4B). A risk score formula was constructed to assess the total risk as follows:

Risk score = (10.28 * methylation level value of cg02948884) + (6.56 * methylation level value of cg16331920) + (2.52 * methylation level value of cg06397837) + (1.49 * methylation level value of cg00049033) + (−2.99 * methylation
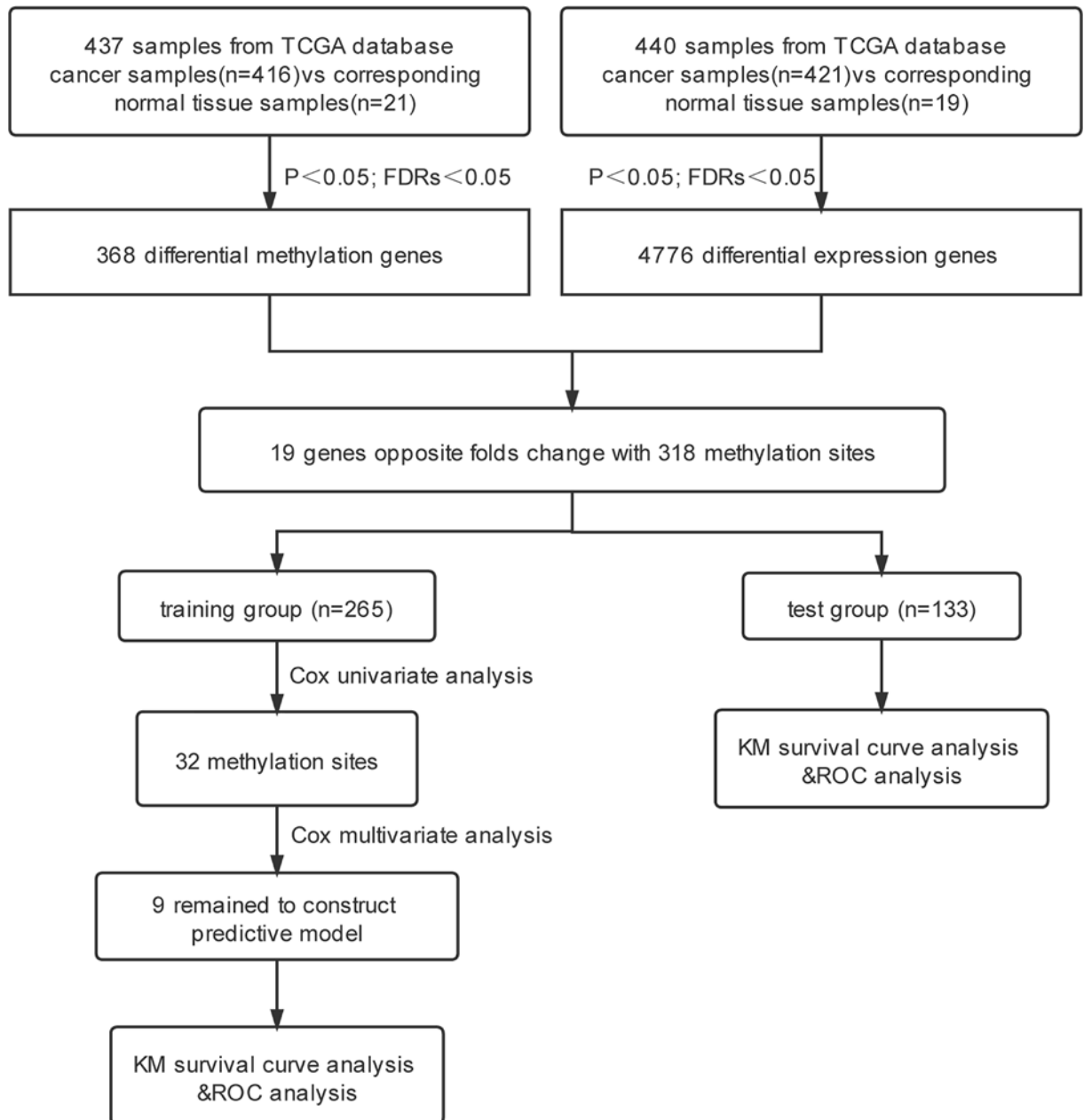
**Figure 1. Flowchart of the research**
The flowchart showing the scheme of developing the prognostic gene signature and validating its efficiency to predict survival.
FDRs, false discovery rates; KM, Kaplan-Meier; ROC, receiver operating characteristic; TCGA, The Cancer Genome Atlas.

level value of cg04101060) + (−4.67 * methylation level value of cg27096779) + (−1.58 * methylation level value of cg00379720) + (1.28 * methylation level value of cg14068328) + (29.13 * methylation level value of cg27580050).

## Prognostic value of the nine-methylation site signature in the training and test groups

To validate the predictive power of the methylation site signature, we conducted a separate ROC analysis in the training and test groups, considering that the larger AUC of ROC calculated a better model for prediction. Figure 5A,B shows the ROC analysis of the training group and the test group, respectively, with AUC values of 0.73 and 0.71,
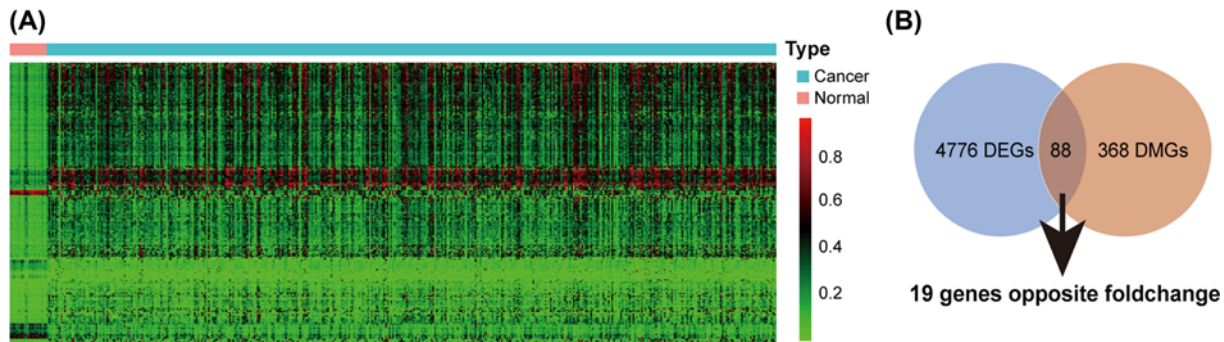
**Figure 2. Identification of DMGs in bladder cancer**

(**A**) Heat map depicting methylation levels of genes in cancer and normal samples. The columns represent the cases, pink represent the cancer samples and green means the control. The lines represent the genes. Red color indicates hyper-methylation levels; green color indicates hypo-methylation level. (**B**) Venn diagram showing there were 88 genes overlapped between DEGs and DMGs, among them 19 genes with opposite folds changed; DEGs, differentially expressed genes; DMGs, differently methylated genes.
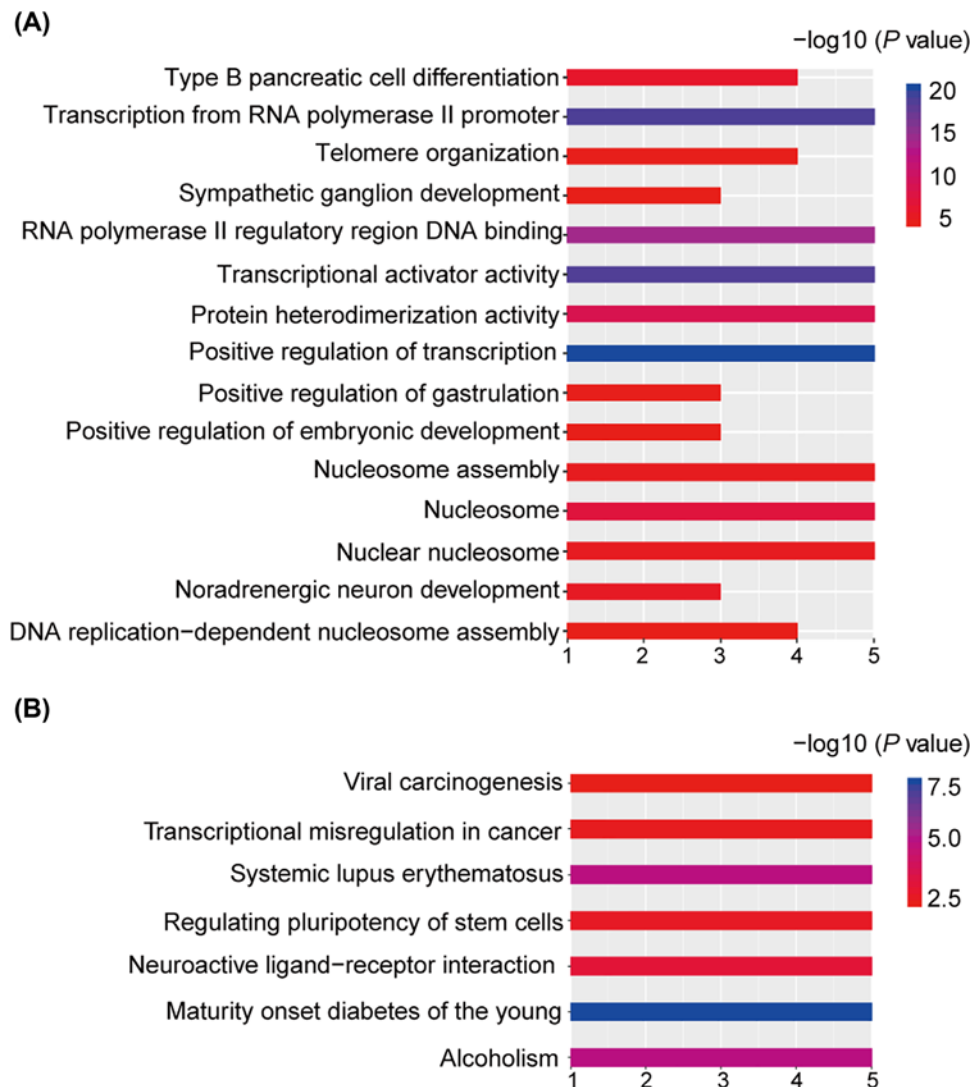


**Figure 3. Gene Ontology (A) and KEGG Pathway (B) enrichment analysis of DMGs**

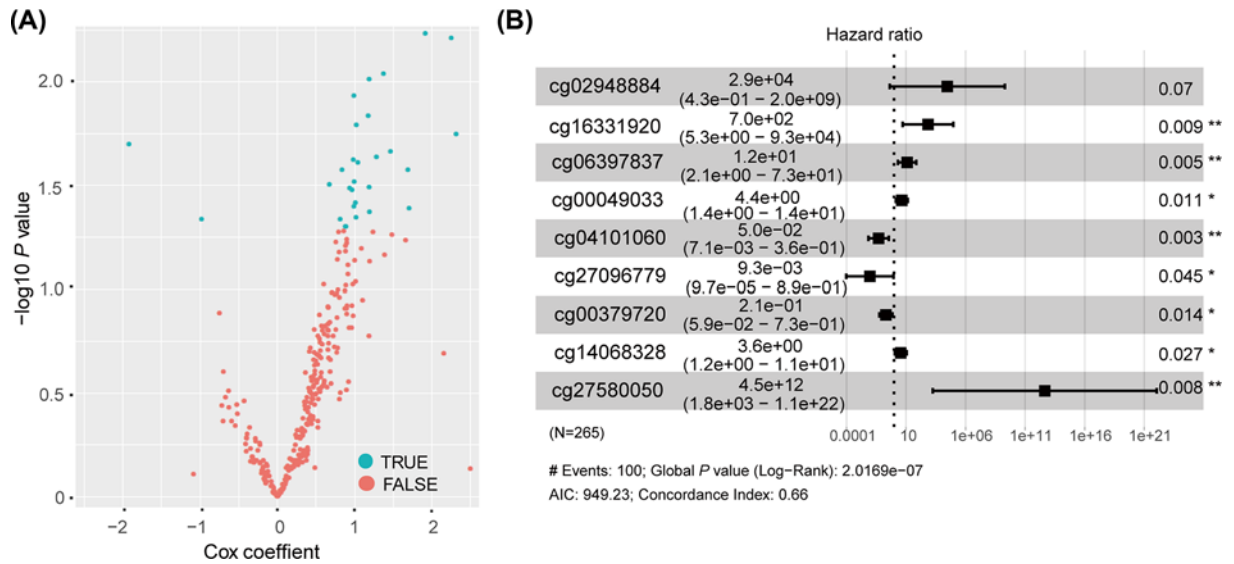Horizontal axis represent the amount of genes enriched on the pathway.

**Figure 4. Identification of the methylation sites in the training group**

(**A**) Cox univariate analysis of the genes' methylation profiling data. Genes with $P<0.05$ were screened. (**B**) Identification of nine methylation sites by Cox multivariate regression analysis. The hazard ratio and 95% confidence interval was showed next to the locus on the picture. AIC shows the fitness of model, concordance index indicates the ability of discriminating; AIC, Akaike information criterion.

which demonstrated that the predictive ability of the nine-methylation site signature was high and further indicated that the signature in our study was a novel and highly accurate prognostic model.

## Establishment of a nomogram for OS prediction in bladder cancer

We developed a bladder cancer nomogram, accepted as a reliable tool to construct a simple intuitive graph of a statistical predictive model that quantified the risk of every patient, to predict a specific individual's probability of 1-, 2- and 3-year survival based on the nine-methylation site gene signature. For every site, the value was evaluated on the scale equal to Exp * Coef. Exp is the degree of DNA methylation, and Coef is the corresponding Cox regression coefficient. The variable got the point corresponding to the point axis upward. We estimated median 1-, 2-, and 3-year survival rates by the total points downward (Figure 5C).

## Confirm the predictive power of the signature and construct a clinically associated nomogram

According to the risk-score formula, the risk scores of all patients in the training group were calculated. We ranked the patients based on the risk scores and classified the training group patients into either a high-risk group or low-risk group using the median risk score as a cut-off point (Figure 6A, below). Through Kaplan–Meier survival analysis, we found that the high-risk group of patients had a lower median overall survival than those in the low-risk group (log-rank test, $P<0.001$, Figure 6A, upper).

To confirm the predictive ability, we applied the risk score formula to the testing group and used the median risk score as the cut-off point (Figure 6B, below). The results in the testing set were similar to those in the training set. The Kaplan–Meier survival analysis showed that the high-risk score patients had a lower median overall survival than those with a low-risk score (log-rank test, $P<0.001$, Figure 6B, upper).

To provide a clinically associated nomogram, a comprehensive nomogram was established in which the risk score integrated with three clinical features, age, sex and stage (Figure 7A). C-index of the model was 0.74, which reveal that the performance of nomogram integrated with clinical features has more powerful predictive ability than the model only based on the risk score (C-index was 0.66). We constructed the nomogram with point scales of these four variables (Figure 7B). The nomogram illustrated that the risk score sharing the largest contribution to prognosis, followed by stage. Each patient for whom with complete clinical information would obtain total points reflecting the probability of 1-, 2- and 3-year survival.
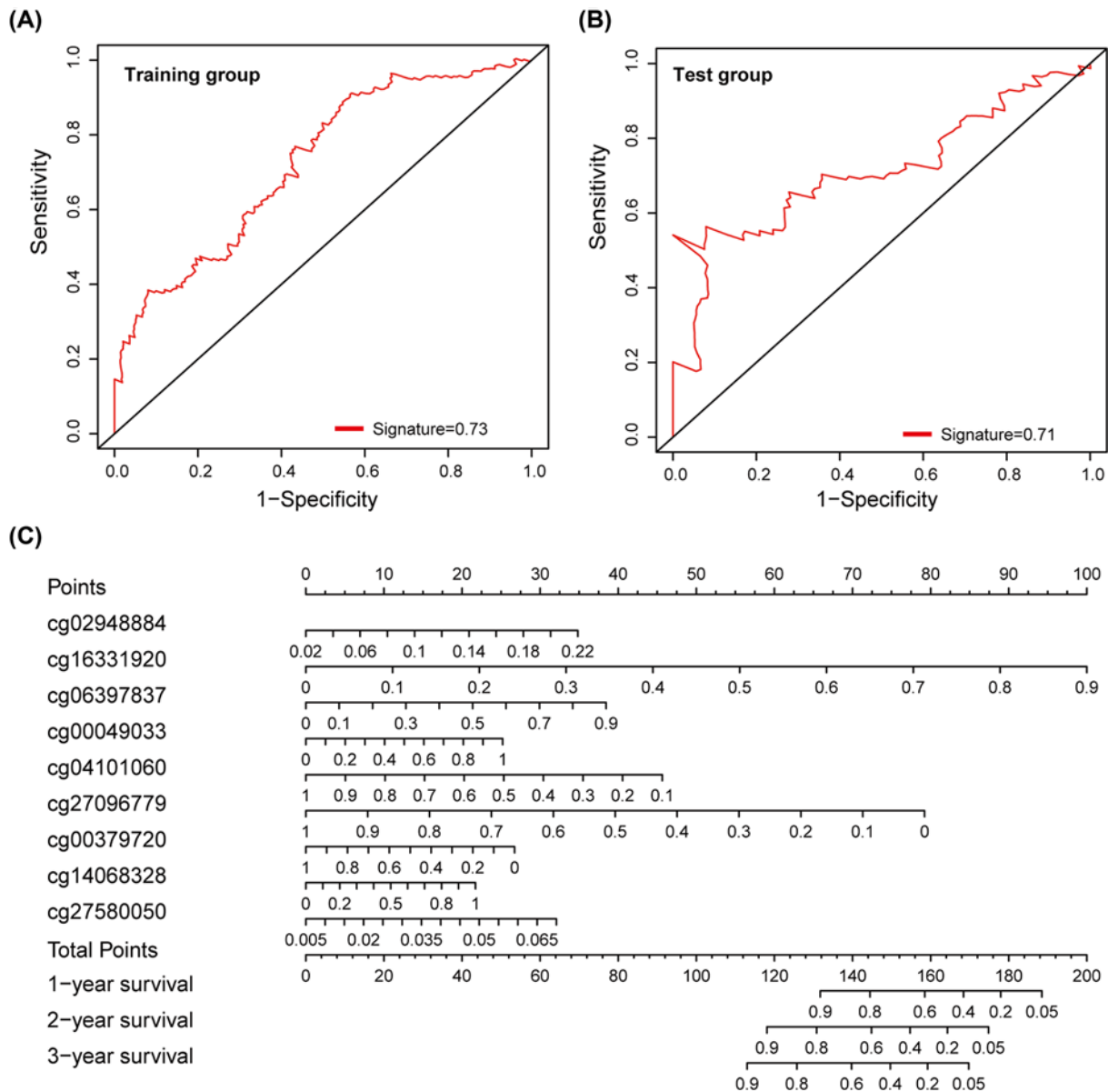
**Figure 5. Predictive ability of the DNA methylation site signature and establishment of a nomogram for bladder cancer**
(**A**) ROC analysis of training group, the AUC was 0.73. (**B**) ROC analysis of training group, the AUC was 0.71. (**C**) Nomogram including nine methylation sites for predicting OS of bladder. Point of every methylated site was got according to the line drawn upward. The total points of the nine components of an individual patient calculated on the 'Total Points' downward, which corresponds to the probability of 1-, 2-, 3-year survival plotted below; AUC, area under curve; OS, overall survival; ROC, receiver operating characteristic.

# Discussion

Numerous studies have found that epigenetic regulation, mainly aberrant DNA methylation, plays a key role in the initiation, development and progression of all human neoplasias, including bladder cancer [22]. Aberrant DNA methylation usually occurs at a GpC (a dinucleotide site that cytosine followed by guanine) island. DNA hypermethylation promotes the development or progression of bladder cancer by silencing genes involved in tumor suppression, DNA repair, cell control, cell invasion and other regulatory pathways, which occur in approximately 50–90% of bladder cancers [22,23].

In the present study, data downloaded from TCGA were used to explore the methylation of bladder cancer and 368 DMGs were identified. Enrichment analysis of the DMGs revealed they were involved in vital biological processes,
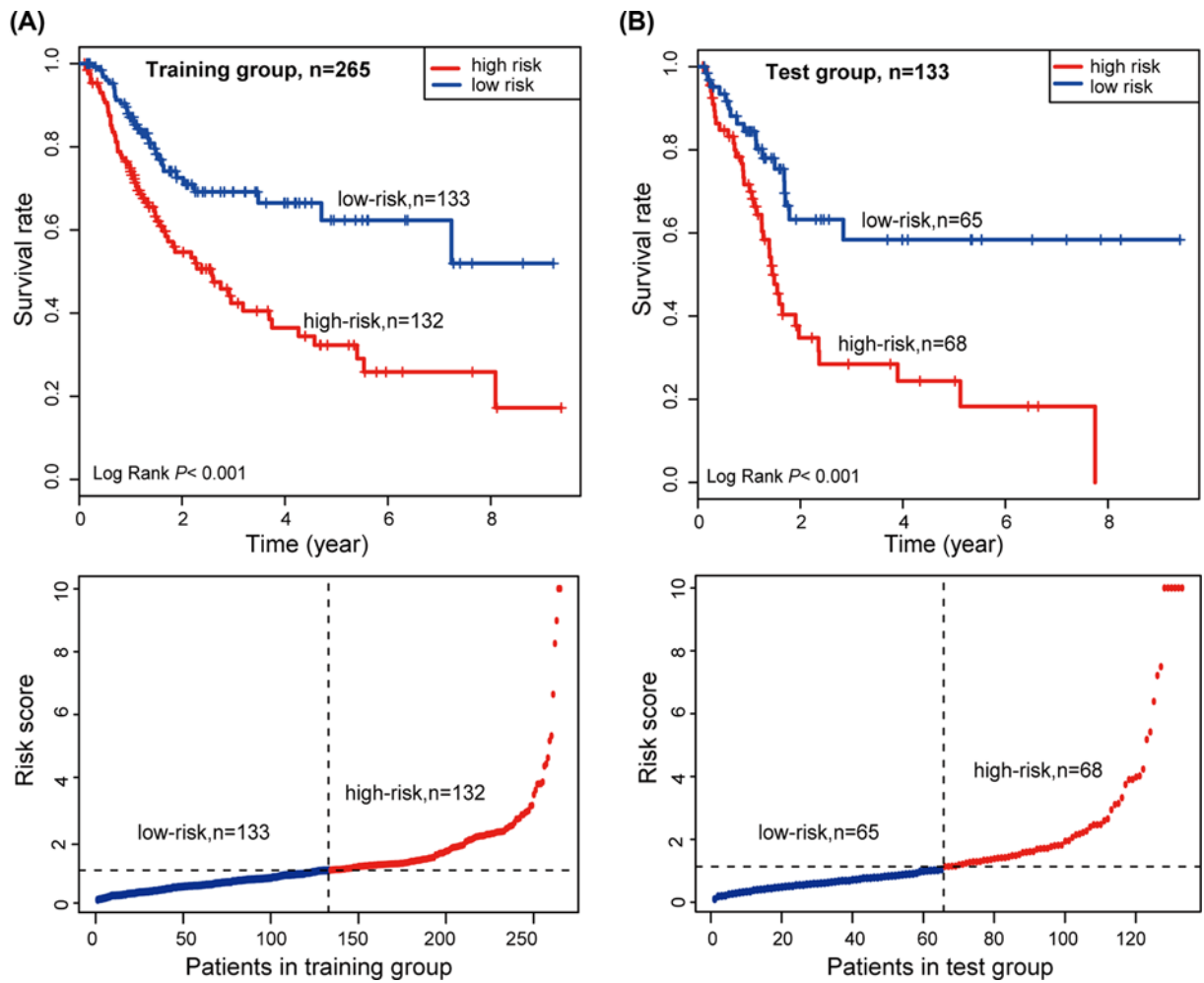
**Figure 6. Kaplan–Meier survival curve analysis and risk score of the patients**

(**A**) According to the median risk score, the training group was classified into high- and low-risk groups. Kaplan–Meier survival curve analysis shows the survival rates of the high- and low-risk groups. *P* values was calculated by log-rank test. (**B**) According to the median risk score, the test group was classified into high- and low-risk groups. Kaplan–Meier survival curve analysis shows the survival rates of the high- and low-risk groups. *P* values was calculated by log-rank test.

including transcription from RNA polymerase II promoter, positive regulation of transcription, and transcriptional activator activity, which suggests that active transcriptions meet the needs for the rapid division of tumor cells. In addition, KEGG enrichments showed that significantly enriched pathways include viral carcinogenesis, transcriptional misregulation in cancer. Abundant expressions of BK polyomavirus genes and virus DNA integration into the host genome were detected in some bladder tumor cases, which suggested that viral infection was a critical factor in the pathogenesis, progression and metastasis of bladder cancer [24,25].

We identified nine methylation sites, which were mapped to seven genes *PITX1* (cg02948884, cg04101060), *MEOX2* (cg16331920, cg27096779), *HOTTIP* (cg06397837), *CBLN1* (cg00049033), *ZNF728* (cg00379720), *TBX18* (cg14068328) and *HSPB6* (cg27580050) by using a series of different statistical approaches. Most of the methylation sites are localized to dinucleotide CpG islands. Among the sites, both cg02948884 and cg04101060 are in the GpC island of *PITX1* (paired-like homeodomain), a tumor-suppressor gene. *PITX1* suppresses tumorigenicity by inhibiting the RAS pathway via transcriptional regulation of RASAL1, a RAS-GTPase-activating protein [26]. Liu *et al.* identified *RASAL1* as a major tumor suppressor gene in thyroid cancer with frequently hypermethylated, which was coupled with its silencing in thyroid cancer cells [27]. However, the relationship between *PITX1* and *RASAL1* in bladder cancer has remained elusive. Several studies previously found that DNA hypermethylation of *PITX1* is associated
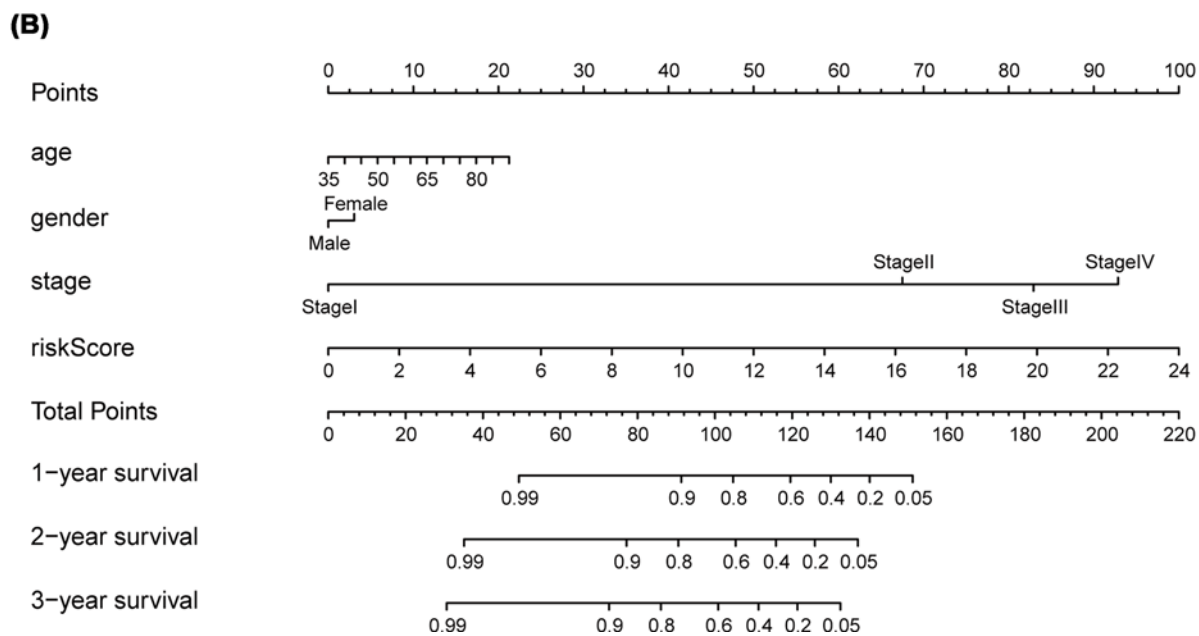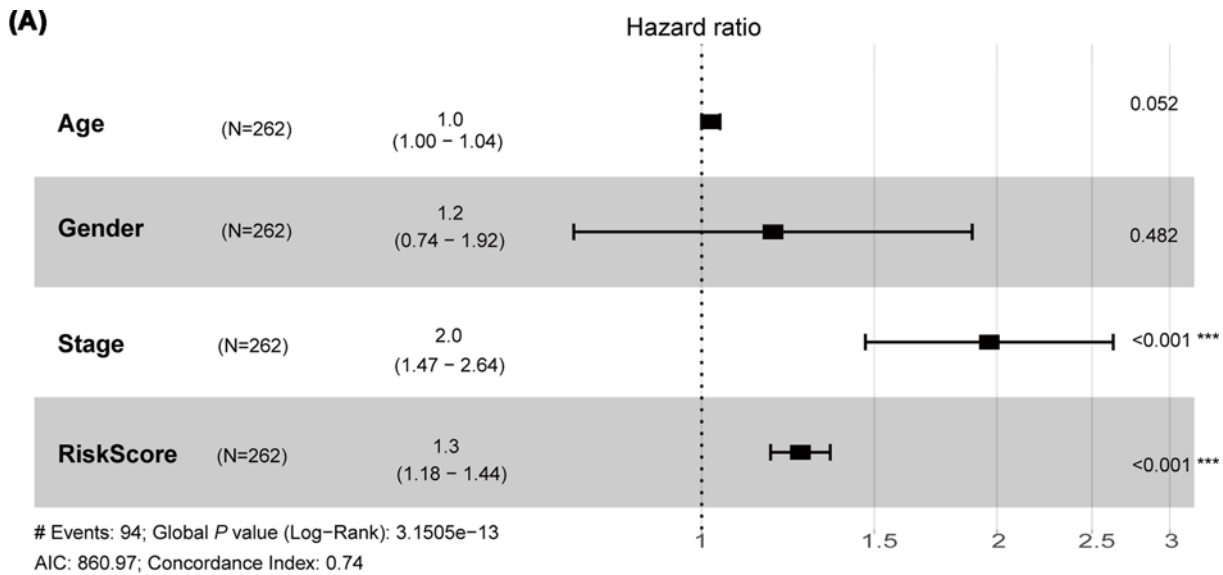
**Figure 7. Construct a clinically associated nomogram combining clinical data and the signature**
(**A**) Cox multivariate regression analysis of the risk score and clinical features with OS of bladder cancer patients. The risk score integrated three clinical features, age, sex and stage were screened and construct the predictive model. (**B**) A nomogram was established based on the four factors; OS, overall survival.

with the growth and progression of various cancers and may predict a poor prognosis [28–30]. ChIP-seq and transcriptomic analyses reveal *PITX1* as a regulator of squamous carcinogenesis in mice and humans, and *PITX1* expression controls bi-stable transcriptional circuits to govern self-renewal and differentiation in squamous cell carcinoma [31]. The transcription factor *MEOX2* (mesenchyme homeobox 2) may play a role in the regulation of vertebrate limb myogenesis. Previous studies revealed that its high expression is relevant to malignant progression and clinical prognosis in lung cancer, hepatocarcinoma and glioma, but little is known about its role in urologic tumor [32]. The Meox2/Tcf15 heterodimers modulate FA transfer to the heart and remedy cardiac dysfunction resulting from altered energy substrate usage [33]. *HOTTIP* is mainly expressed in the colon, prostate, pancreas and urinary bladder and produces a long RNA that regulates the expression of *HOXA* genes in trans. *HOTTIP* aberration promotes hematopoietic stem cell self-renewal leading to acute myeloid leukemia-like disease in mice with reprogramming

leukemic-associated chromatin and gene transcription [34]. *HOTTIP* was overexpressed in small cell lung cancer (SCLC) tissues, and its expression was associated with the disease progression and the shorter survival time of SCLC patients [35]. Compared with patients with lower expression, pancreatic cancer patients with higher *HOTTIP* expression had shorter disease-free survival and overall survival [36]. *CBLN1* encodes a cerebellum-specific protein related to the tumor necrosis factor families of proteins [37,38]. *TBX18* encodes a member of the vertebrate-specific Tbx1 subfamily, which plays a crucial role in embryonic development. Combined with Tbx18, Six1 synergistically regulates smooth muscle cell development and ureter function. It is possible that Tbx18 modulates a signaling pathway that negatively regulates ureteric branching morphogenesis [39]. However, little is known about the function of *CBLN1* and *TBX18* in tumorigenesis and the progression of tumor. *HSPB6*, a small heat shock protein, plays a crucial role in muscle function and has emerged as a novel cardioprotector against stress-induced injury [40,41]. Genome-wide screening of promoter methylation has identified that the frequency of *HSPB6* promoter methylation increases moderately in early stage melanomas and significantly in advanced-stage melanomas [42]. However, not much is known about the function of *ZNF728* (zinc finger protein 728) now. Taken together, there is relatively little research on the seven signature genes in bladder cancer, and our study provides insight for further investigation into the function of these seven genes.

However, there are some limitations to our study that should be noted. Most importantly, there is a lack of information on the mechanism behind the prognostic values of these gene methylation sites in bladder cancer, and experimental studies will further enhance our understanding of their functional roles. Furthermore, although we believe that the nine-site gene signature is promising for predicting patient survival, its potential value still needs to be verified in larger groups of bladder cancer patients.

## Conclusions

Overall, our study systematically screened a DNA methylation signature for predicting of the OS of bladder cancer patients. Integrated with clinical information, we constructed a more comprehensive nomogram.

### Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

### Author Contribution

Conceptualization: Dan Xiong, Tong Ou; Methodology: Huamei Tang, Lijuan Kan; Software: Dayang Chen, Xiaowen Dou; Validation: Xiuming Zhang; Formal analysis: Wang Mengmeng and Zengyan Zong; Data curation: Mo Hongmei; Writing—original draft preparation: Tanghuamei; Writing-review and editing: Dan Xiong, Tong Ou; Project administration: Wei Wu and Xiang Ji; Funding acquisition: Dan Xiong, Tong Ou, Xiuming Zhang.

### Abbreviations

AIC, Akaike information criterion; AUC, area under curve; CI, confidence interval; DAVID, Database for Annotation Visualization and Integrated Discovery; DEG, differentially expressed gene; DMG, differently methylated gene; FDR, false discovery rate; GO, Gene Ontology; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; OS, overall survival; ROC, receiver operating characteristic; TCGA, The Cancer Genome Atlas.

### References

1  Ferlay, J. et al. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, E359–E386, https://doi.org/10.1002/ijc.29210

2  Nagata, M., Muto, S. and Horie, S. (2016) Molecular Biomarkers in Bladder Cancer: Novel Potential Indicators of Prognosis and Treatment Outcomes. *Dis. Markers* **2016**, 8205836, https://doi.org/10.1155/2016/8205836

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

4   Antoni, S. et al. (2017) Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends. *Eur. Urol.* **71**, 96–108,
    https://doi.org/10.1016/j.eururo.2016.06.010

5   Cumberbatch, M.G.K. et al. (2018) Epidemiology of Bladder Cancer: A Systematic Review and Contemporary Update of Risk Factors in 2018. *Eur. Urol.*
    **74**, 784–795, https://doi.org/10.1016/j.eururo.2018.09.001

6   Perez-Losada, J., Castellanos-Martin, A. and Mao, J.H. (2011) Cancer evolution and individual susceptibility. *Integr. Biol. (Camb.)* **3**, 316–328,
    https://doi.org/10.1039/c0ib00094a

7   Carbone, M. et al. (2018) Consensus report of the 8 and 9th Weinman Symposia on Gene x Environment Interaction in carcinogenesis: novel
    opportunities for precision medicine. *Cell Death Differ.* **25**, 1885–1904, https://doi.org/10.1038/s41418-018-0213-5

8   Baylin, S.B. et al. (1998) Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv. Cancer Res.* **72**, 141–196,
    https://doi.org/10.1016/S0065-230X(08)60702-2

9   Jones, P.A. and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3**, 415–428, https://doi.org/10.1038/nrg816

10  Kunze, E. and Schlott, T. (2007) High frequency of promoter methylation of the 14-3-3 sigma and CAGE-1 genes, but lack of hypermethylation of the
    caveolin-1 gene, in primary adenocarcinomas and signet ring cell carcinomas of the urinary bladder. *Int. J. Mol. Med.* **20**, 557–563

11  Weisenberger, D.J. (2014) Characterizing DNA methylation alterations from The Cancer Genome Atlas. *J. Clin. Invest.* **124**, 17–23,
    https://doi.org/10.1172/JCI69740

12  Huang, W.Y. et al. (2015) MethHC: a database of DNA methylation and gene expression in human cancer. *Nucleic Acids Res.* **43**, D856–D861,
    https://doi.org/10.1093/nar/gku1151

13  Michalak, E.M. et al. (2019) The roles of DNA, RNA and histone methylation in ageing and cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 573–589,
    https://doi.org/10.1038/s41580-019-0143-1

14  Greenberg, M.V.C. and Bourc'his, D. (2019) The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* **20**,
    590–607, https://doi.org/10.1038/s41580-019-0159-6

15  Mahmood, N. and Rabbani, S.A. (2019) Targeting DNA Hypomethylation in Malignancy by Epigenetic Therapies. *Adv. Exp. Med. Biol.* **1164**, 179–196,
    https://doi.org/10.1007/978-3-030-22254-3˙14

16  Olkhov-Mitsel, E. et al. (2017) Epigenome-Wide DNA Methylation Profiling Identifies Differential Methylation Biomarkers in High-Grade Bladder Cancer.
    *Transl. Oncol.* **10**, 168–177, https://doi.org/10.1016/j.tranon.2017.01.001

17  Wang, Z., Jensen, M.A. and Zenklusen, J.C. (2016) A Practical Guide to The Cancer Genome Atlas (TCGA). *Methods Mol. Biol.* **1418**, 111–141,
    https://doi.org/10.1007/978-1-4939-3578-9˙6

18  Yu, G. et al. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287,
    https://doi.org/10.1089/omi.2011.0118

19  Reiner, A., Yekutieli, D. and Benjamini, Y. (2003) Identifying differentially expressed genes using false discovery rate controlling procedures.
    *Bioinformatics* **19**, 368–375, https://doi.org/10.1093/bioinformatics/btf877

20  Liang, W. et al. (2015) Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J. Clin.*
    *Oncol.* **33**, 861–869, https://doi.org/10.1200/JCO.2014.56.6661

21  Dennis, Jr., G. et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, P3,
    https://doi.org/10.1186/gb-2003-4-5-p3

22  Porten, S.P. (2018) Epigenetic Alterations in Bladder Cancer. *Curr. Urol. Rep.* **19**, 102, https://doi.org/10.1007/s11934-018-0861-5

23  Sánchez-Carbayo, M. (2012) Hypermethylation in bladder cancer: biological pathways and translational applications. *Tumour Biol.* **33**, 347–361,
    https://doi.org/10.1007/s13277-011-0310-2

24  Tang, K.W. et al. (2013) The landscape of viral expression and host gene fusion and adaptation in human cancer. *Nat. Commun.* **4**, 2513,
    https://doi.org/10.1038/ncomms3513

25  Starrett, G.J. and Buck, C.B. (2019) The case for BK polyomavirus as a cause of bladder cancer. *Curr. Opin. Virol.* **39**, 8–15,
    https://doi.org/10.1016/j.coviro.2019.06.009

26  Kolfschoten, I.G. et al. (2005) A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* **121**, 849–858,
    https://doi.org/10.1016/j.cell.2005.04.017

27  Liu, D. et al. (2013) Identification of RASAL1 as a major tumor suppressor gene in thyroid cancer. *J. Natl. Cancer Inst.* **105**, 1617–1627,
    https://doi.org/10.1093/jnci/djt249

28  Otsubo, T. et al. (2017) DNA hypermethyation and silencing of correlated with advanced stage and poor postoperative prognosis of esophageal
    squamous cell carcinoma. *Oncotarget* **8**, 84434–84448, https://doi.org/10.18632/oncotarget.21375

29  Tai, W.T. et al. (2016) Protein tyrosine phosphatase 1B dephosphorylates PITX1 and regulates p120RasGAP in hepatocellular carcinoma. *Hepatology* **63**,
    1528–1543, https://doi.org/10.1002/hep.28478

30  Chen, Y. et al. (2007) Decreased PITX1 homeobox gene expression in human lung cancer. *Lung Cancer* **55**, 287–294,
    https://doi.org/10.1016/j.lungcan.2006.11.001

31  Sastre-Perona, A. et al. (2019) De Novo PITX1 Expression Controls Bi-Stable Transcriptional Circuits to Govern Self-Renewal and Differentiation in
    Squamous Cell Carcinoma. *Cell Stem Cell* **24**, 390e8–404e8, https://doi.org/10.1016/j.stem.2019.01.003

32  Tachon, G. et al. (2019) Prognostic significance of MEOX2 in gliomas. *Mod. Pathol.* **32**, 774–786, https://doi.org/10.1038/s41379-018-0192-6

33  Coppiello, G. et al. (2015) Meox2/Tcf15 heterodimers program the heart capillary endothelium for cardiac fatty acid uptake. *Circulation* **131**, 815–826,
    https://doi.org/10.1161/CIRCULATIONAHA.114.013721

34  Luo, H. et al. (2019) HOTTIP lncRNA Promotes Hematopoietic Stem Cell Self-Renewal Leading to AML-like Disease in Mice. *Cancer Cell* **36**,
    645e8–659e8, https://doi.org/10.1016/j.ccell.2019.10.011

35  Sun, Y. et al. (2017) A long non-coding RNA HOTTIP expression is associated with disease progression and predicts outcome in small cell lung cancer patients. *Mol. Cancer* **16**, 162, https://doi.org/10.1186/s12943-017-0729-1

36  Fu, Z. et al. (2017) LncRNA HOTTIP modulates cancer stem cell properties in human pancreatic cancer by regulating HOXA9. *Cancer Lett.* **410**, 68–81, https://doi.org/10.1016/j.canlet.2017.09.019

37  Hirai, H. et al. (2005) Cbln1 is essential for synaptic integrity and plasticity in the cerebellum. *Nat. Neurosci.* **8**, 1534–1541, https://doi.org/10.1038/nn1576

38  Jones, R.M. et al. (2013) Genome-wide association study of autistic-like traits in a general population study of young adults. *Front. Hum. Neurosci.* **7**, 658, https://doi.org/10.3389/fnhum.2013.00658

39  Nie, X. et al. (2010) SIX1 acts synergistically with TBX18 in mediating ureteral smooth muscle formation. *Development* **137**, 755–765, https://doi.org/10.1242/dev.045757

40  Dreiza, C.M. et al. (2010) The small heat shock protein, HSPB6, in muscle function and disease. *Cell Stress Chaperones* **15**, 1–11, https://doi.org/10.1007/s12192-009-0127-8

41  Liu, G.S. et al. (2018) Regulation of BECN1-mediated autophagy by HSPB6: Insights from a human HSPB6 mutant. *Autophagy* **14**, 80–97, https://doi.org/10.1080/15548627.2017.1392420

42  Kalari, S. and Pfeifer, G.P. (2010) Identification of driver and passenger DNA methylation in cancer by epigenomic analysis. *Adv. Genet.* **70**, 277–308, https://doi.org/10.1016/B978-0-12-380866-0.60010-1