Improved Urine DNA Methylation Panel for Early Bladder Cancer Detection

Qixun Fang
Yaneng Bioscience, Co., Ltd, Fosun Pharma, Shenzhen
Xu Zhang
Department of Urology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou;
Qing Nie
Yaneng Bioscience, Co., Ltd, Fosun Pharma, Shenzhen
Jianqiang Hu
South China University of Technology, Guangzhou
Shujun Zhou (zhoushujun@yanengbio.com)
Yaneng Bioscience, Co., Ltd, Fosun Pharma, Shenzhen
Chaojun Wang
Department of Urology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou;

Research Article

Keywords: Urine, Bladder cancer, Biomarkers, Methylation, Diagnosis, RRBS, qMSP

Posted Date: July 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-653839/v1

Version of Record: A version of this preprint was published at BMC Cancer on March 3rd, 2022. See the published version at https://doi.org/10.1186/s12885-022-09268-y.

Abstract

Purpose

Bladder cancer is one of the most common malignancies but the corresponding diagnostic methods are either invasive or limited in specificity and/or sensitivity. Detecting DNA methylation using urine sample has been reported as a potential method for bladder cancer detection. This study aimed to develop a bladder cancer detecting panel by improving published panels and validated the panel performance with clinical samples.

Materials and Methods

Related researches were reviewed and the biomarkers of most potential were selected into a new panel. A total of 92 samples were included in this study, 45 of which had their methylation status measured by RRBS and others measured by qMSP. RRBS data was used to compare performance of the new panel and the panels it derided from while qMSP data was used to validate panel performance in occasions of clinical use. To further investigate panel detection limit and result reproducibility, synthesized samples were repeatedly measured in different time and concentration.

Results

Three biomarkers (PCDH17, POU4F2 and PENK) were selected to construct a new panel P3. P3 panel achieved an accuracy of 0.836 and an AUC of 0.964 with RRBS while the accuracy and AUC with qMSP were 0.915 and 0.948, showing more balanced specificity and sensitivity and higher accuracy compared to individual marker. Detection limit study showed high sensitivity in plasmid sample and the results can be reproduced with acceptable deviations.

Conclusion

Overall, the P3 panel achieved relatively high sensitivity and accuracy in bladder cancer detection.

1. Background

Bladder cancer (BC) is known as one of the most common malignancies in the world and it ranks just after prostate cancer in genitourinary malignancies [1]. It is reported that 75% of the primary tumors present as a non-muscle-invasive stage Ta or T1 tumor while other tumors show bladder muscle invasion in stages T2-4. Clinically, stage Ta BC is characterized by frequent recurrence after resection, up to 60% of patients [2]. Typically, within 8 to 10 years, one or more tumors will appear each year with no progression, but, as many as 25% of tumors will eventually develop an aggressive invasive phenotype [3].

In current, cystoscopy/biopsy is the gold standard for the BC detection of suspicious lesions. Unfortunately, 10–40% of malignancies were failed to be detected by this procedure. In addition, cystoscopy/biopsy it is not only invasive, painful and costly but also misses up to 15% of the papillary and up to 30% of the flat recurrences. On the other hand, although urine cytology possesses a high specificity, it lacks of sensitivity, particularly in low-risk tumors [4].

Recently, many researchers are committed to discovering better markers for disease diagnosis and prognosis by employing non-invasive methods to collect samples, such as urine sediments. Although the sensitivity of cytology has been improved by the addition of nuclear matrix protein 22 (NMP-22), bladder tumor antigen, or UroVysion FISH, the proposed markers have been rarely adopted in clinical practice due to their limited specificity or sensitivity[1, 5, 6].

DNA methylation plays an important role in transcription regulation [7, 8]. It has been found that the changes in DNA methylation are chemically stable and can be accurate quantified, making them competitive candidates as tumor markers [9, 10]. Inactivation of tumor suppressor genes by gain of DNA methylation (hypermethylation) or global loss of DNA methylation (hypomethylation), activating genes that were normally not expressed, had been both observed in bladder

tumors [11–13]. Further studies also demonstrated methylation changes found in urine sediments reflected those found in tumor tissues [14–16]. Reduced representation bisulfite sequencing (RRBS) has become increasingly for analysis of genome-wide methylation profiles with single nucleotide precision. One of the main goals for RRBS study is to discover differentially methylated regions (DMRs) between different biological conditions.

In last decades, more and more DNA methylation markers had been found for BC detection and these proposed markers were evaluated with different methods such as MS-MLPA, quantitative methylation-specific PCR (qMSP) and pyrosequencing. However, in most cases, the performance of marker panels were not satisfying with sensitivity in the range of 44%-90% and specificity varying from 31–94% (Table 1). Thus, it is still urgently to seek a reliable DNA methylation marker set operating on a low-cost platform with high sensitivity and specificity for BC detection.

Table 1
Reviewed panels for bladder cancer detection

ID	Gene Panel	Method	AUC	AC	SP	SN	ref
1	TIMP3,APC,CDKN2A,MLH1,ATM,RARB, CDKN2B,HIC1,CHFR,BRCA1,CASP8,CDKN1B, PTEN,BRCA2,CD44,RASSF1,DAPK1,FHIT,VHL, ESR1,TP73,IGSF4,GSTP1,CDH13	MS-MLPA	-	-	-	-	[17]
2	HIC1,RASSF1,GSTP1	MS-MLPA	0.696	0.72	0.66	0.78	[17]
3	HOXA9,ISL1	qMSP	-	-	0.91	0.44	[18]
4	PCDH17,POU4F2	qMSP	-	-	0.94	0.90	[19]
5	E2F3,CCND1,UTP6,CDADC1,SLC35E3,METRNL,TPCN2, NACC2,VGLL4,PTEN	metadata	-	-	-	-	[20]
6	CDH13,CFTR,NID2,SALL3,TMEFF2,TWIST1,VIM2	pyrosequencing	-	-	-	-	[21]
7	CFTR,SALL3,TWIST1	pyrosequencing	0.741	-	0.31	0.90	[21]
*AUC: area under ROC curve, AC: accuracy, SP: specificity, SN: sensitivity							

In this study, we aimed to develop a reliable DNA methylation panel for early BC detection in clinical practise. We composed an improved urine DNA methylation panel based on other published methylation panels. This newly proposed panel was compared with published panels using RRBS and the panel performance was further applied to clinical diagnosis with qMSP and systematically evaluated.

2. Methods

Study design

In this study, panel development went through two stages: panel design and validation. In panel design stage, literatures related to methylation panels for bladder cancer detection were reviewed and a panel collection consisting of both published panels and panels derided from prior was established. The performance of these panels were assessed using RRBS with 28 BC patient samples with 17 non-BC controls. In validation phase, 29 patients were enrolled with 18 non-BC controls (Fig. 1). All patients with BC were confirmed with cystoscopy. All samples were collected from volunteers of both sex above 18 years old. Non-BC controls included healthy samples and other non-BC patient samples. BC patients were not balanced on any aspect although male patients and early stage patients were clearly oversampled. Besides, BC patients with prior chemotherapeutic treatment could be included.

2.1 Patients and sample collection

The fresh urine samples were obtained from patients at the First Affiliated Hospital, Zhejiang University School of Medicine. Cell pellet were centrifuged and kept frozen until used for DNA extraction. Clinical, demographic, and pathological data were collected for all specimens (Table 2). Informed consent was obtained from all participants. This research was approved by Zhejiang University Institutional Review Board including that all methods were performed in accordance with the relevant guidelines and regulations.

Table 2
Clinicopathological and demographical information of involved population

	Verification phase (RRBS)	Validation phase (qMSP)
Sample number	45	47
BC patient	28	29
Non-BC control	17	18
Among patients		
Age range	48–92	44–84
Mean age	67.9	64.3
Stage		
Ta	11	7
T1	6	13
T2	5	3
T3	4	4
T4	2	2
MIBC	13	7
NMIBC	10	13
High grade	18	13
Low grade	10	9
Primary	18	12
Recurrence	10	10

* MIBC: Muscle invasive bladder cancer, NMIBC: Non-muscle invasive bladder cancer. Pathological details were unclear or lost for some samples, these samples were excluded in relative analyses.

2.2 DNA extraction

DNAs from the urinary cell pellets were extracted using QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. DNA quality and quantity were assessed using a NanoDrop spectrophotometer and 1% agarose gel electrophoresis.

2.3. RRBS library construction

100 ng genomic DNA from each sample was used for the construction of each RRBS library. Genomic DNA was digested overnight with MspI restriction enzyme (recognition site C^CGG, NEB, Ipswich, MA, USA) at 37°C. After purification, digested DNA was subsequently treated with a mix of T4 DNA polymerase, Klenow Fragment and T4 polynucleotide kinase to repair, blunt and phosphorylate ends. The mixture DNA fragments were subsequently 3' adenylated using Klenow Fragment (3'-5' exo-) which followed by ligation to adaptors with 5'-methylcytosine substituting cytosine using T4 DNA Ligase. Bisulfite conversion was performed with a ZYMO EZ DNA Methylation-Gold Kit (ZYMO, Irvine, CA, USA) and lambda DNA was added as a conversion marker. The final libraries were generated by PCR amplification of 13 cycles. RRBS libraries were analysed by an Agilent 2100 Bioanalyzer (Agilent Technologies). Sequencing were performed on an Illumina Hiseq platform with a pair-end 300 cycles setting, yielding ~ 10G sequencing data.

2.4 Data analysis

Raw sequencing data were processed by the Illumina base-calling pipeline. Low-quality reads that contained more than 30% 'N's or over 10% low-quality calls (quality value < 20) were removed. Adapter contamination was removed by cutadapt (version 1.9) [22]. The clean reads were aligned to the reference genome hg19 using BSMAP (version 2.73) [23]. Conversion ratio was calculated on lambda DNA and samples with a conversion ratio lower than 99% were considered unqualified for further analysis and corresponding libraries were rebuilt for sequencing. Only uniquely aligned reads containing the *Msp* I enzyme digestion sites were used for further analysis. Only CpG sites with sequencing depths ≥ 5 were selected as candidate sites. After bisulfite treatment, cytosines were read as "T" if unmethylated and otherwise as "C" if methylated. Methylation level of the sample was defined as the ratio of "C" counts to total "C" and "T" counts on the site. Differential methylation analysis on these sites were performed with ANOVA (analysis of variance) and the acquired p-values were adjusted to Q-values with Benjamini-Hochberg method [24]. Differential methylation sites (DMSs) were defined as sites with a Q-value lower than 0.05.

2.5 Performance verification with RRBS results

To verify panel performance, distinct DMS related to marker genes (within 5,000 bp to gene region) were extracted from differential analysis results from RRBS. These DMS were grouped on individual panel and used as sample features for model training and sample classification. Support vector machine (SVM) with radial basis function (RBF) kernel was applied as the main classification model and the penal performance was evaluated in 4-fold cross validation manner where the cohort were divided into 4 groups with equal healthy/cancer ratio and trained model with 3 groups and test model performance with the last group in rotation. The final panel performance was obtained by averaging performance estimation of all 4 tests. Confusion matrixes and Receiver Operating Characteristic (ROC) curves were preserved for further performance analysis.

2.6 Methylation measuring with qMSP

DNA samples were treated by EZ DNA Methylation-Gold™ Kit (Zymo Research). The bisulfite modification reaction was executed by 96-Well GeneAmp PCR System 9700 (Applied iosystems) with the mixture of 150 µl that contains 130 µl of CT conversion reagent (Zymo Research) and 200 ng of DNA template. The condition of the reaction was configured to 98°C for 10 minutes followed by 64°C for 2.5 hours and then on hold at 4°C. 20 µl of M-Elution Buffer (Zymo Research) was used to purify each DNA sample. The purified DNA was stored at -20°C until use.

Primers of candidate biomarkers (PCDH17, POU4F2, and PENK) and β-actin (reference gene) were used in the qMSP assay of the bisulfite modified DNA samples. Each reaction mixture (25µl in total) was set to be consisted with 5 µl of 5X probe qPCR Buffer (Tiagen), 2.5 µl of 10X HA Buffer, 2 µl of dNTP (2.5mM), 9.5uL primer and probe mix, 0.75uL HA Taq, 0.25uL Taq-Antibody, and 5µl of DNA template. The reactions were performed on Real-Time PCR System (Applied Biosystems) after a pre-incubation of 5 minutes at 95°C, and were executed for 48 cycles at 95°C for 15 seconds each, situated at 60°C for 40 seconds. The fluorescence signal was measured at the end of each extension step at 60°C. The reaction was

repeated three times for each sample. The mean Ct value from the triplicate were used as the final Ct value. According to [25], the site methylation rate (SMR) referencing ACTB can be calculated as:

$$SMR = 2^{[Ct(ACTB)-Ct]} \times 100$$

where Ct(ACTB) and Ct are the Ct values of the ACTB gene and target gene respectively. Thus, to simplify, $Ct - Ct(ACTB)$ as ΔCt calculated by final Ct values was used as a proxy of the SMR. However, in this case, ΔCt increases with the decrease of SMR and the correlation between is logarithmic.

2.7 Cancer detection with qMSP results

Two third of the validation samples were used to train a Random Forest (RF) model with 500 decision trees whose maximum depth was set to 3, and the rest samples were used to test the model performance. QMSP sample cohort was first split into train set and test set in a random manner maintaining identical healthy/cancer ratio within both set. While Random Forest model was applied for classification with the three-marker panel, SVM with RBF kernel was applied to individual marker for further investigation on single marker detection power. After training and testing, confusion matrixes and ROC curves on entire validation cohort were preserve as results.

2.8 Reproducibility and limit of detection (LOD) study

Investigation on reproducibility and LOD requires stable samples with certain methylation ratio. Therefore, plasmids with methylated target sequences were synthesized for each marker, including ATCB, and standard samples were obtained by mixing plasmids of different concentration (details shown on Table 3). For reproducibility study, testing with standard samples were repeated 10 times on two date with same protocol. For LOD study, standard samples with single plasmids in 4 concentration were applied.

Table 3
Component concentration in standard samples

Sample	Concentration(ng/μL)			
	ATCB	PCDH17	POU4F2	PENK
RP-1	5.00E-06	6.00E-08	4.00E-07	2.00E-07
LDR-1	5.00E-09	0.00	0.00	0.00
LDR-2	1.00E-08	0.00	0.00	0.00
LDR-3	5.00E-08	0.00	0.00	0.00
LDR-4	1.00E-07	0.00	0.00	0.00
LDP1-1	0.00	6.00E-09	0.00	0.00
LDP1-2	0.00	3.00E-08	0.00	0.00
LDP1-3	0.00	6.00E-08	0.00	0.00
LDP1-4	0.00	3.00E-07	0.00	0.00
LDP2-1	0.00	0.00	7.12E-09	0.00
LDP2-2	0.00	0.00	3.56E-08	0.00
LDP2-3	0.00	0.00	7.12E-08	0.00
LDP2-4	0.00	0.00	3.56E-07	0.00
LDP3-1	0.00	0.00	0.00	2.00E-09
LDP3-2	0.00	0.00	0.00	4.00E-09
LDP3-3	0.00	0.00	0.00	2.00E-08
LDP3-4	0.00	0.00	0.00	4.00E-08

3. Results

A graphic outline of this study is given in Fig. 1.

3.1 Novel panel design with relevant literature

According to the reported performance of panels reviewed, the panel consisting PCDH17 and POU4F2 with high scores on both sensitivity and specificity was selected as a starting point for optimisation. This panel, in spite of consisting only two markers, was reported best performance within panels reviewed. To further improve the panel performance with least addition to the panel, PENK was introduced to the panel as complement forming a new panel, namely P3. PENK was a biomarker used by Genomictree as the sole marker for bladder cancer which alone reaches 0.89 sensitivity and 0.875 specificity. We hypothesized the extra marker would improve the performance of the origin panel.

3.2 Panel performance analysis with RRBS on cell lines and clinical samples

To verify our hypothesis, RRBS was performed on the verification sample set to address the correlation between marker methylation and cancer condition. Other panel reviewed were also verified alone with the proposed panel to evaluate panel performance on specificity, sensitivity, accuracy and area under ROC curve (AUC) (Table 4). Overall, significant difference

between verification results and published performance was observed. Compared to the published figures, higher specificity and/or sensitivity were achieved by most panels except the two marker panel from Wang et al. [19]. Such improvement could be attributed to higher region coverage and resolution of RRBC and the use of SVM optimization in criteria design. In simple words, SVM classifies samples by difference on feature values and balances the specificity and sensitivity. However, another fact worth noticing is that most of these panels have relatively high specificity while compromised sensitivity. It could be caused by model preference in optimizing panel accuracy.

Best overall performance was achieved by a panel with 24 markers from Casadio et al. [17] with highest accuracy, specificity and sensitivity. It was followed by the P3 panel containing significantly lower marker number with slightly lower accuracy and sensitivity but higher AUC. Compared to the panel it derived from, which is the panel from Wang et al. [19], P3 panel achieved overall better performance in terms of accuracy, sensitivity and AUC with an equivalent specificity of 1.0. However, P3 also failed in certain samples which were also missed by most, if not all, of other panels (Fig. 2).

Table 4

Verified performance of reviewed panels and P3 panel

Panel ID	AUC	Accuracy	Specificity	Sensitivity	reference
1	0.941	0.893	1.000	0.829	[17]
2	0.816	0.737	0.906	0.671	[17]
3	0.904	0.817	0.988	0.714	[18]
4	0.933	0.822	1.000	0.729	[19]
5	0.851	0.742	0.976	0.629	[20]
6	0.921	0.760	1.000	0.629	[21]
7	0.882	0.786	1.000	0.650	[21]
P3	0.964	0.836	1.000	0.736	N/A

3.3 Performance validation with qMSP on clinical samples

The proposed P3 panel achieved a promising result only second to another highly complicated panel in RRBC verification, but different error would be introduced by different methylation measuring methods and qMSP is a more widely used measuring platform in clinical practise. Therefore, further validation was performed with qMSP which is low on cost and instrument requirement. RF models were applied instead of SVM for positive/negative classification and the performance was also measured by accuracy, sensitivity, specificity and AUC. The RF model is a collection of decision trees which can individually provide classification decisions on given samples after training. An example of decision tree can be found on Fig. 3. Although 500 decision trees were used in this RF model, the difference among these trees were not significant and mostly follow the principle that BC correlates to the hypermethylation (low ΔCt) of these markers. Difference of decision trees mainly occurs on marker importance to decision making and the grey area between normal and low ΔCt.

Compared to RRBS results, P3 panel with qMSP platform showed increased accuracy (0.915 vs 0.836 on RRBS) and sensitivity (0.931 vs 0.736 on RRBS) and similar AUC (0.948 vs 0.964 on RRBS) but decayed specificity (0.889 vs 1.000 on RRBS) (see Table 5 for details). It is expected that performance parameters change when applied to qMSP but the overall performance was still acceptable (with accuracy > 90%).

Table 5
Performance of individual marker and p3 panel in BC detection

Gene	SP	SN	AC	AUC
PCDH17	0.778	0.724	0.745	0.852
POU4F2	0.778	0.966	0.894	0.952
PENK	0.944	0.828	0.872	0.969
Combine	0.889	0.931	0.915	0.948
SP: specificity, SN: sensitivity, AC: accuracy, AUC: area under ROC curve				

For better understanding of the P3 panel property, detection power of individual marker was also investigated with ROC curve (Fig. 4). Three markers displayed different detection features but none performed unanimously better than others. Compared other individual markers, PENK reached highest specificity and AUC while POU4F2 achieved highest accuracy and sensitivity. When combining three markers, although subtle decrease was observed on specificity, sensitivity and AUC compared to the individual best, a higher accuracy was achieved. It is clear that the P3 panel improves overall performance by balancing specificity and sensitivity.

3.4 Reproducibility and LOD study result

For understanding of the stability in panel performance, the reproducibility and LOD were investigated. To assess the reproducibility of the measuring results, tests on synthesized standard sample were conducted in two different date by different operator for 10 times each. Although no significant difference (on level of 0.05) was observed between results from different date (Fig. 5), significant difference in in-group deviant was observed, indicating a fact that the measurement is mostly stable but the possibility of inaccurate measurement should not be neglected. Therefore, more works should be carried on to reduce inaccurate measurement.

In LOD study, plasmid samples of different concentration were measured and result deviation and correlation between Ct value and concentration were considered. Measurement of PCDH17 and PENK did not significantly scatter until plasmid concentration below 1e-8 ng/μL while POU4F2 started scattering from concentration below 1e-7 ng/μL (Fig. 6). Ct values yielded by ACTB measurement did not significantly deviate throughout all concentration but lost alignment with the concentration below 1e-8 ng/μL. Thus, the LOD was set to 1e-8 ng/μL for marker PCDH17 and PENK and 1e-7 ng/μL for ACTB and POU4F2. In term of Ct value, no major deviate was observed within the same concentration group for ACTB and POU4F2 while deviate notably increased with Ct value above 33 for PCDH17 and PENK and major deviates appeared in groups whose mean Ct value close to 36. Most of the critical results from clinical samples clustered below a Ct value of 36 and therefore should not significantly deviate in repeat tests. However, it should be note that the use of plasmid samples could lead to over-optimistic LOD estimation since PCR efficiency could be different between plasmid and genome sequence. Further validation with cell line or clinical samples will required. Overall, the reproducibility and LOD results showed that within the Ct value range observed from clinical samples, the panel would produce stable results.

3.5 Penal diagnostic potential in BC subtype detection

To further understand the detection power of P3 panel, we investigated its performance on distinguishing muscle invasive/non-muscle invasive (MIBC/NMIBC), high/low grade, early/late stage and primary/recurrent BCs (Table 6). Same metrics was used to evaluate the performance (Table 5). Compared to detecting BC from non-BC, panel performance on these subtype classification decays. According to AUC values, the penal might be used in separating MIBC and NMIBC or high and low grade BCs with AUCs close to 0.8 but showed no significant improvement (AUC < 0.7) than a random classifier on early/late stage and primary/recurrent BCs. In simple words, given all AUC values higher than 0.5, higher AUC

indicates higher possibility in getting a well-perform threshold for separating different groups. Penal perform similarly on MIBC/NMIBC and high/low grade detection is likely attribute to the high correlation of low grade with NMIBC. The penal achieved similar specificity (close to 0.9) in these two occasions as that in BC/non-BC detection but the sensitivity drop to 0.222 and 0.769 respectively leading to decrease on accuracy.

Table 6

Subtype detection performance of individual marker.

	PCDH17		POU4F2		PENK		P3
	Training	Testing	Training	Testing	Training	Testing	
BC/non-BC							
Specificity	0.80	0.77	0.80	0.77	1.00	0.92	0.889
Sensitivity	0.80	0.68	1.00	0.95	0.90	0.79	0.931
Accuracy	0.80	0.72	0.93	0.88	0.93	0.84	0.915
AUC	0.92	0.82	0.92	0.96	1.00	0.96	0.948
MI/NMI							
Specificity	0.50	0.43	0.50	0.57	0.50	0.29	1.000
Sensitivity	0.50	0.86	0.50	0.57	0.50	0.14	0.222
Accuracy	0.50	0.64	0.50	0.57	0.50	0.21	0.65
AUC	0.50	0.71	0.38	0.69	0.50	0.16	0.778
Grade							
Specificity	0.67	0.83	0.67	0.83	1.00	0.67	0.889
Sensitivity	0.75	0.78	0.75	0.78	0.75	0.67	0.769
Accuracy	0.71	0.80	0.71	0.80	0.86	0.67	0.818
AUC	0.66	0.87	0.92	0.78	0.75	0.78	0.803
Stage E/L							
Specificity	0.60	0.36	0.40	0.64	0.80	0.55	0.688
Sensitivity	0.50	0.00	0.50	1.00	0.50	0.25	0.500
Accuracy	0.57	0.27	0.43	0.73	0.71	0.47	0.636
AUC	0.60	0.32	0.30	0.77	0.60	0.66	0.677
P/R							
Specificity	0.67	0.67	0.67	0.33	0.67	0.22	0.583
Sensitivity	0.50	0.67	0.75	0.33	0.50	0.17	0.600
Accuracy	0.57	0.67	0.71	0.33	0.57	0.20	0.591
AUC	0.66	0.72	0.75	0.19	0.50	0.15	0.675

We then further decomposed the panel and exam metrics in training and testing cohorts to investigate marker contribution for subtype detection (Table 6). In term of BC/non-BC detection, three markers contributed similarly which was consistent with results described on Sect. 3.3. Detecting power decreases in all three markers for cancer grade but the accuracies are still mostly between 0.7 and 0.8, resulting an accuracy above 0.8 in combined panel. Interestingly, all three markers showed no detecting power for distinguishing MIBC and NMIBC in training set individually but demonstrated a perfect specificity when combined. Meanwhile, PENK expressed clear bias in testing set, yielding an accuracy of 0.21 and an AUC of 0.16 which significantly different from expected results (close to 0.5 by random prediction). This indicates these three markers, especially PENK, could possess certain detection power for MIBC and NMIBC which is disrupted by noise and increasing sample size might manifest such power. However, in case of developing a panel for MIBC/NMIBC detection, it might be more valuable to replace PCDH17 and POU4F2 with other markers. Apart from this, no marker displayed individual detecting power on stage and primary/recurrence detection.

4. Discussion

In this study, P3 panel designed from reviewing published works was proposed for detecting BC using urine sample based on methylation. The proposed panel was then verified with RRBS in comparison of other panels reviewed and validated with qMSP. P3 panel showed promising performance in both verification and validation.

In verification phase, P3 panel ranked on top amidst panels reviewed with relatively low number of marker and displayed performance superior to panels it derived from, although the improvement was not significant in every aspect. Also, overall performance in these panels suggested there is space for further improvement. The best performing panel only reached an accuracy (of 0.893) below 0.90 as a result of low sensitivity (best being 0.829). It appeared that the loss of sensitivity was caused by serval difficult BC samples which possess genotypes more similar to healthy controls than other BCs (Fig. 2). Due to low number of this kind of sample and absent of clear common intragroup feature, a reasonable explanation has yet to be concluded. However, these sample could be a breaking point to further panel improvement.

In validation phase, qMSP was applied not just for measuring methylation in another way but also as a trial for clinical application of P3 panel. Overall, the panel achieve an accuracy of 0.915 with 0.899 specificity and 0.931 sensitivity which were considered highly promising. However, the panel only demonstrated moderate (on muscle invasive and grade detection) or unsatisfying (on stage and primary/recurrence detection) performance on cancer status detection. In practice, distinguishing primary and recurrence is unnecessary and more of research interest while detecting other features could be diagnostically useful, so improving on primary/recurrence detection was not considered. Detail analysis on individual marker presented similar characteristics in subtype detection: a higher performance on BC grade but deceased performance on muscle invasion and stage. Note that we oversampling on early stage BCs (Ta, T1 and T2, 23/29, 79.3%) for more critical validation for BC detection, since imageological methods such as cystoscopy are generally less sensitive for early stage cancer and we attempted to tackle this issue with P3 panel. Given such design, the panel managed to achieve a high sensitivity on BC detection with no significant detecting power regarding to stage. Therefore, it is premature to claim that the panel cannot be used on stage detection and further validation with more balanced samples is worthy for clinical applications. Meanwhile, sampling on other BC subtypes was properly balanced so improvement should be made on panel design and RRBS could provide viable options.

Both RRBS and qMSP were included and compared in this study. It is worth noticing that the proposed panel performed differently with RRBS and qMSP. We attribute such difference to 1) different experiment errors generated by measuring methods and 2) different coverage and selection of related regions. RRBS allows measurement of large range of regions related to marker genes in single nucleotide resolution while qMSP is only able to detect average methylation changes on specific regions. However, in practice, qMSP has relatively lower requirement in cost and equipment compared to RRBS so that it is more preferred in cancer screening and detection. On the other hand, RRBS is highly preferred in researches such as novel biomarker discovery, but this method must be accompanied by sophisticate analysis tools to highlight distinct

methylation activities and remove undesired noise which compromises specificity and/or sensitivity. Therefore, improvement of BC detection panel could be achieved by adding novel markers discovered with RRBS but adjustment would be required in switching to qMSP for desired results. In this process, large number of samples would be necessary to establish related parameters and validate performance. Proper statistic models and machine learning models can improve efficiency of this process.

Compared to cystoscopy, nucleotide methylation test on urine sample with qMSP provides a low-cost, non-invasive, comfortable measure to detecting or/and monitor BCs. The discovery of P3 panel facilitates this application with high accuracy and high sensitivity. Urine qMSP results from P3 panel could be used to accompany cystoscopy as compliment of sensitivity for early stage tumours or as primary screen for BCs. However, reproducibility study indicated that there are unidentified factors causing uncertainty in detection result and the panel LOD requires validation with clinical samples. More work need to be done on test protocol to meet the IVD (in vitro diagnostics) product standards. On the other hand, although improvement will be required for BC subtype detection, a panel with effective subtype detection would be crucial for BC progress monitoring and therefore would be a part of the next step of panel development. Also, to further validate the performance of this panel for clinical application, more clinical samples from multicentre will be required.

5. Conclusion

In conclusion, we developed a three-gene methylation panel P3 for BC detection. The panel demonstrated promising results in both RRBS and qMSP tests and showed tremendous potential in clinical application. Although further performance improvement will be required for commercialising the panel on qMSP platform, it has shown great potential value in clinical application.

Abbreviations

BC: Bladder cancer

NMP-22: nuclear matrix protein 22

FISH: Fluorescence in situ hybridization

DNA: deoxyribonucleic acid

RRBS: Reduced representation bisulfite sequencing

DMR: differentially methylated region

PCR: Polymerase Chain Reaction

qMSP: quantitative methylation-specific PCR

MS-MLPA: Methylation-Specific Multiplex Ligation-Dependent Probe Amplification

MIBC: Muscle invasive bladder cancer

NMIBC: Non-muscle invasive bladder cancer

ANOVA: analysis of variance

DMS: Differential methylation site

SVM: Support vector machine

ROC: Receiver Operating Characteristic

AUC: area under the curve

IVD: In vitro diagnostics

Declarations

Ethics approval and consent to participate

The fresh urine samples were obtained from patients at the First Affiliated Hospital, Zhejiang University School of Medicine, which was approved by Zhejiang University Institutional Review Board.

Consent for publication

Not applicable

Availability of data and materials

Sequencing data was deposited to NCBI SRA database and can be accessed via https://www.ncbi.nlm.nih.gov/sra/PRJNA715028.

Competing interests

Shujun Zhou and Qixun Fang are postdoctors in South China University of Technology, and Jianqiang Hu is the supervisor and guide. Shujun Zhou and Qixun Fang are also the technical directors of the department of R&D in Yaneng Bioscience, Co., Ltd, and receive research funding from this company. Qing Nie is an employee of Yaneng Bioscience, Co., Ltd. Xu Zhang is a junior doctor in the First Affiliated Hospital, Zhejiang University School of Medicine, and Chaojun Wang is the head of the Department of Urology. Also, as an urological surgeon, he is part of the Yaneng Bioscience Medical Advisory Board.

Funding

None

Authors' contributions

Shujun Zhou and Chaojun Wang made substantial contributions to the concept and design of the work. Qixun Fang analyzed and interpreted the data and was a major contributor in writing the manuscript, and Jianqiang Hu substantively revised it. Xu Zhang and Qing Nie performed the sample collection and major technical work. All authors read and approved the final manuscript.

Acknowledgements

We thank Bingfeng Leng and Qi Hu for helpful discussion on this study.

References

1. Parker J, Spiess PE. Current and Emerging Bladder Cancer Urinary Biomarkers. Thescientificworldjo. 2011;11:1103-12.
2. Millan-Rodriguez F, Chechile-Toniolo G, Salvador-Bayarri J, Palou J, Algaba F, Vicente-Rodriguez J. Primary superficial bladder cancer risk groups according to progression, mortality and recurrence. J Urology. 2000;164:680-84.

3. Babjuk M, Oosterlinck W, Sylvester R, Kaasinen E, Bohle A, Palou-Redorta J, et al. EAU Guidelines on Non-Muscle-Invasive Urothelial Carcinoma of the Bladder, the 2011 Update. Eur Urol. 2011;59:997-1008.
4. Tritschler S, Sommer ML, Straub J, Hocaoglu Y, Tilki D, Strittmatter F, et al. Urinary Cytology in Era of Fluorescence Endoscopy: Redefining the Role of an Established Method With a New Reference Standard. Urology. 2010;76:677-80.
5. Pisitkun T, Johnstone R, Knepper MA. Discovery of urinary biomarkers. Mol Cell Proteomics. 2006;5:1760-71.
6. Schroeder GL, Lorenzo-Gomez MF, Hautmann SH, Friedrich MG, Ekici S, Huland H, et al. A side by side comparison of cytology and biomarkers for bladder cancer detection. J Urology. 2004;172:1123-26.
7. Esteller M. Molecular origins of cancer: Epigenetics in cancer. New Engl J Med. 2008;358:1148-59.
8. Schubeler D. Function and information content of DNA methylation. Nature. 2015;517:321-26.
9. Wolff EM, Liang GN, Jones PA. Mechanisms of disease: genetic and EPIGENETIC alterations that drive bladder cancer. Nat Clin Pract Urol. 2005;2:502-10.
10. Laird PW. The power and the promise of DNA methylation markers. Nat Rev Cancer. 2003;3:253-66.
11. Kim WJ, Kim YJ. Epigenetic biomarkers in urothelial bladder cancer. Expert Rev Mol Diagn. 2009;9:259-69.
12. Reinert T, Modin C, Castano FM, Lamy P, Wojdacz TK, Hansen LL, et al. 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;17:5582-92.
13. Vallot C, Stransky N, Bernard-Pierrot I, Herault A, Zucman-Rossi J, Chapeaublanc E, et al. A Novel Epigenetic Phenotype Associated With the Most Aggressive Pathway of Bladder Tumor Progression. Jnci-J Natl Cancer I. 2011;103:47-60.
14. Friedrich MG, Weisenberger DJ, Cheng JC, Chandrasoma S, Siegmund KD, Gonzalgo ML, et al. Detection of methylated apoptosis-associated genes in urine sediments of bladder cancer patients. Clin Cancer Res. 2004;10:7457-65.
15. Seifert HH, Schmiemann V, Mueller M, Kazimirek M, Onofre F, Neuhausen A, et al. In situ detection of global DNA hypomethylation in exfoliative urine cytology of patients with suspected bladder cancer. Exp Mol Pathol. 2007;82:292-97.
16. Kim YK, Kim WJ. Epigenetic markers as promising prognosticators for bladder cancer. Int J Urol. 2009;16:17-22.
17. Casadio V, Molinari C, Calistri D, Tebaldi M, Gunelli R, Serra L, et al. DNA Methylation profiles as predictors of recurrence in non muscle invasive bladder cancer: an MS-MLPA approach. J Exp Clin Canc Res. 2013;32:1-9.
18. Kitchen MO, Bryan RT, Haworth KE, Emes RD, Luscombe C, Gommersall L, et al. Methylation of HOXA9 and ISL1 Predicts Patient Outcome in High-Grade Non-Invasive Bladder Cancer. Plos One. 2015;10:1-12.
19. Wang YQ, Yu Y, Ye R, Zhang D, Li QL, An D, et al. An epigenetic biomarker combination of PCDH17 and POU4F2 detects bladder cancer accurately by methylation analyses of urine sediment DNA in Han Chinese. Oncotarget. 2016;7:2754-64.
20. Shivakumar M, Lee Y, Bang LS, Garg T, Sohn KA, Kim D. Identification of epigenetic interactions between miRNA and DNA methylation associated with gene expression as potential prognostic markers in bladder cancer. Bmc Med Genomics. 2017;10:65-75.
21. van der Heijden AG, Mengual L, Ingelmo-Torres M, Lozano JJ, van Rijt-van de Westerlo CCM, Baixauli M, et al. Urine cell-based DNA methylation classifier for monitoring bladder cancer. Clin Epigenetics. 2018;10:1-10.
22. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal. 2011;17:10-12.
23. Li YX, Li W. BSMAP: whole genome bisulfite sequence MAPping program. Bmc Bioinformatics. 2009;10:1-9.
24. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met. 1995;57:289-300.
25. De Strooper LMA, Meijer CJLM, Berkhof J, Hesselink AT, Snijders PJF, Steenbergen RDM, et al. Methylation Analysis of the FAM19A4 Gene in Cervical Scrapes Is Highly Efficient in Detecting Cervical Carcinomas and Advanced CIN2/3

Lesions. Cancer Prev Res. 2014;7:1251-57.

Figures

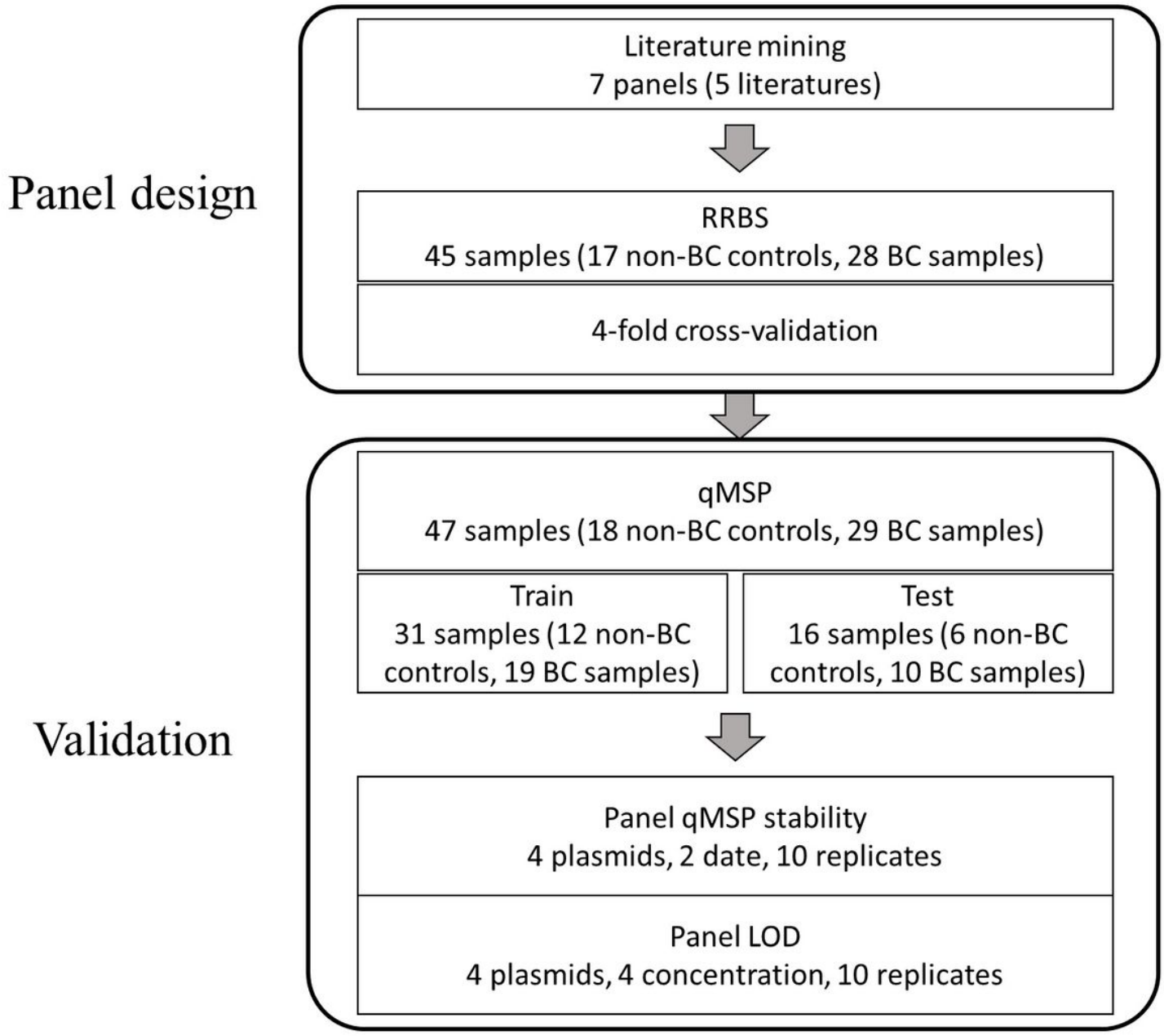

Figure 1

Outline of study design. BC: bladder cancer, qMSP: quantitative methylation-specific PCR, LOD: limit of detection. Sample splitting in cross-validation maintained equivalent BC/non-BC ratio for every group. Support vector machine was used for RRBS data and random forest was used for qMSP data as predicting model.

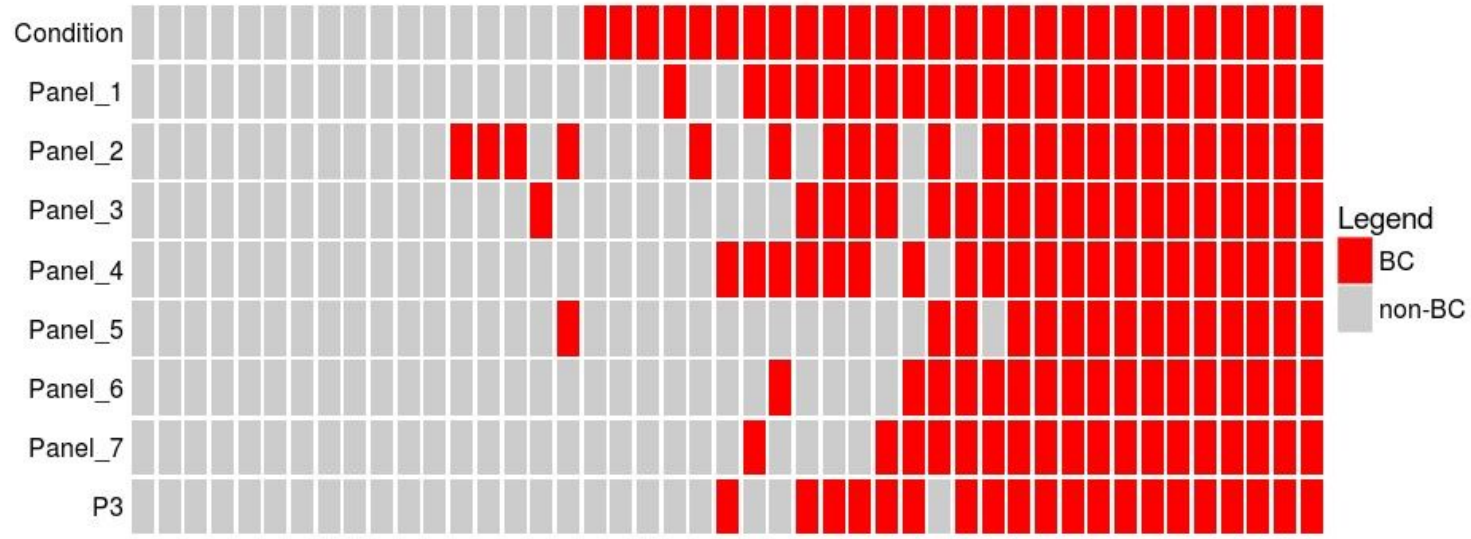

Figure 2

Bladder cancer predictions of reviewed panel and P3 panel. The condition row presents the true status of samples. Other rows show predictions made by panels. Predictions were made by one of the models generated in cross-validation process for presentation. Note that predictions could be varied by different models on the same panel.

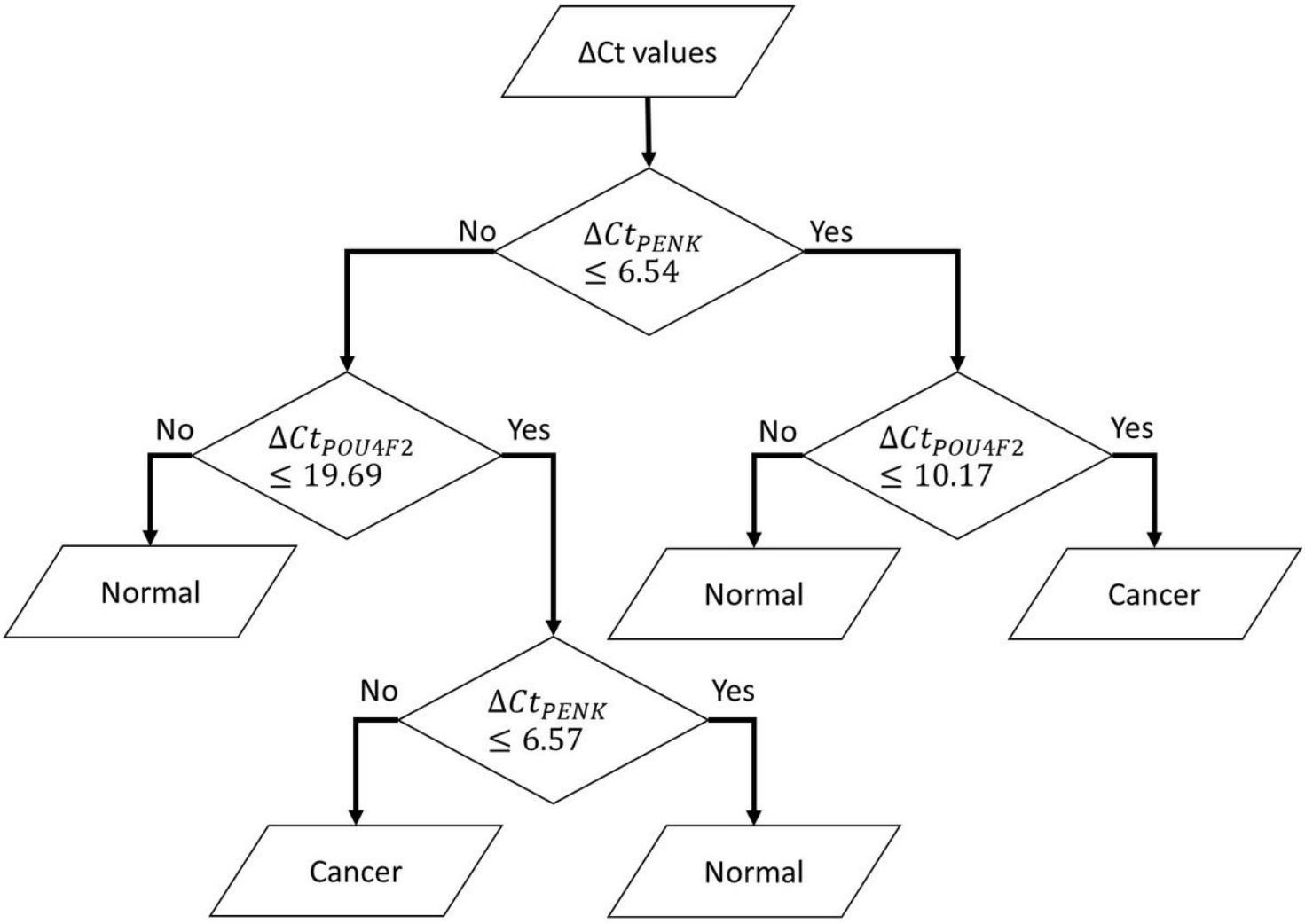

Figure 3

Sample decision tree of cancer detection criteria.

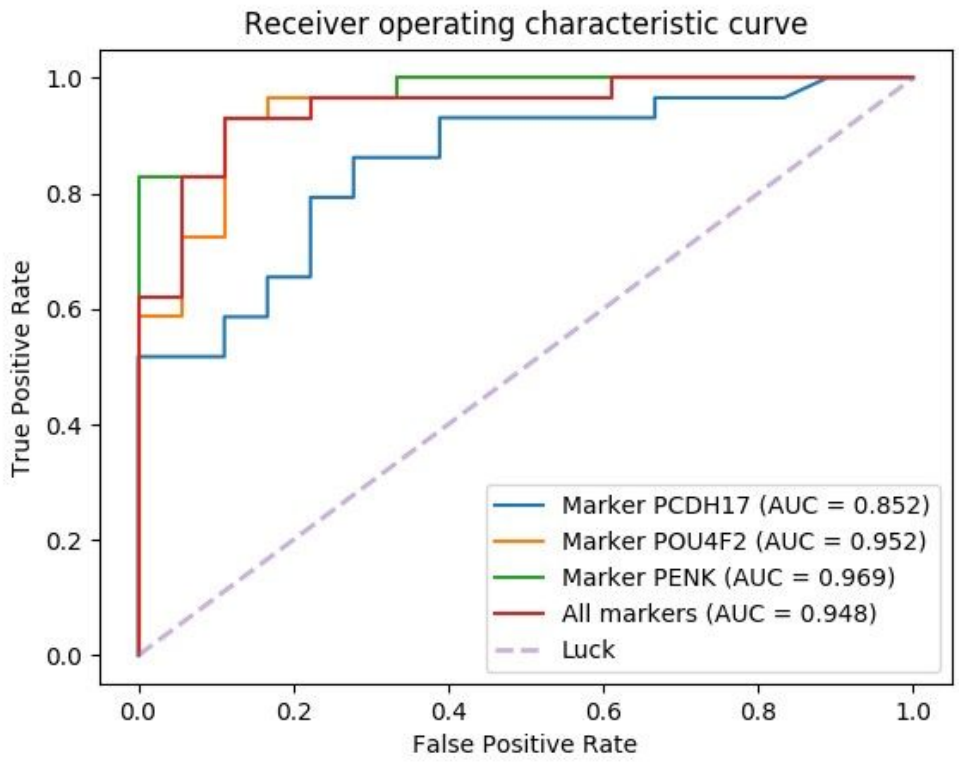

Figure 4

Receiver Operating Characteristic (ROC) curve of individual marker and combined P3 panel.

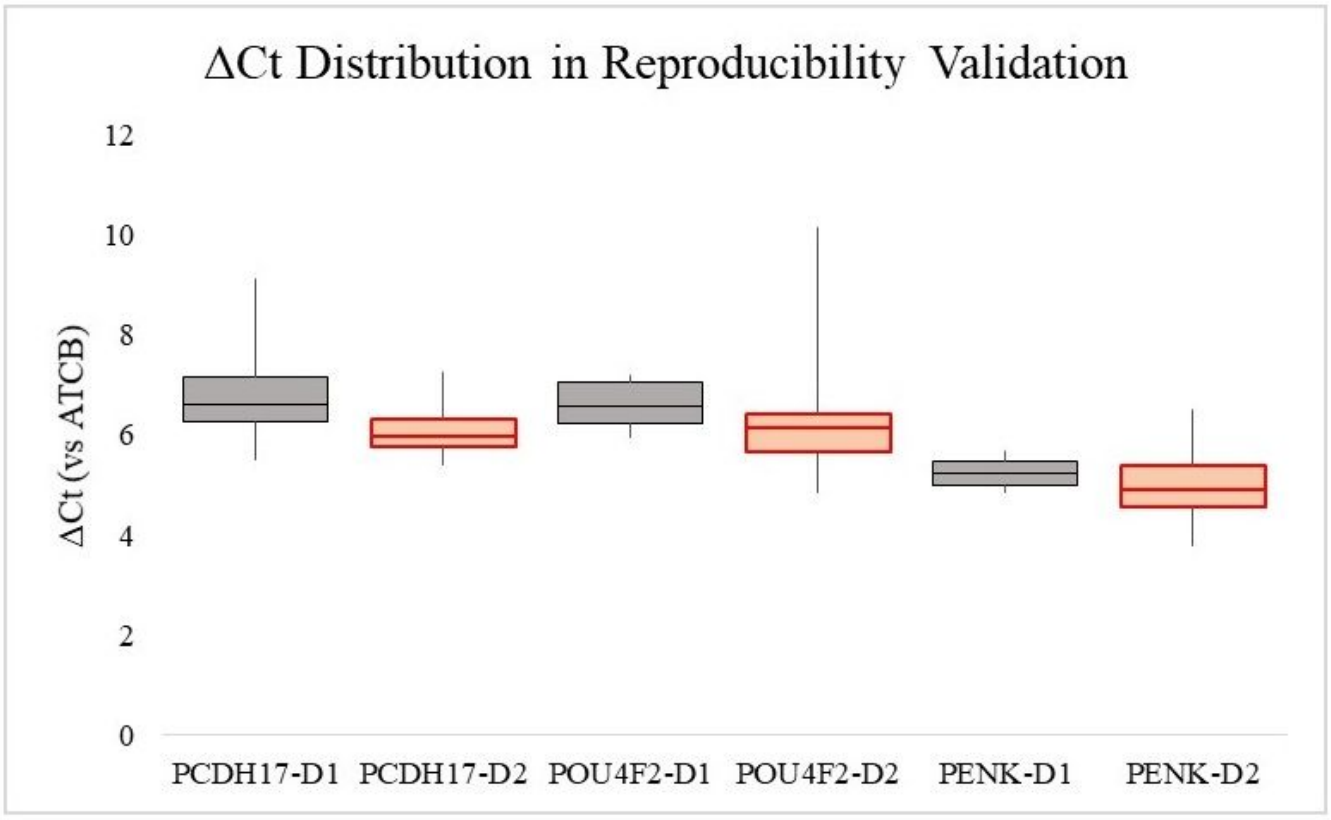

Figure 5

ΔCt distribution in reproducibility study. Tests on each sample were repeated 10 times in 2 separate dates. Statistics difference between measurements of the same marker were estimated by two-tail t-test. In a significance level of 0.05, no significant difference appeared.

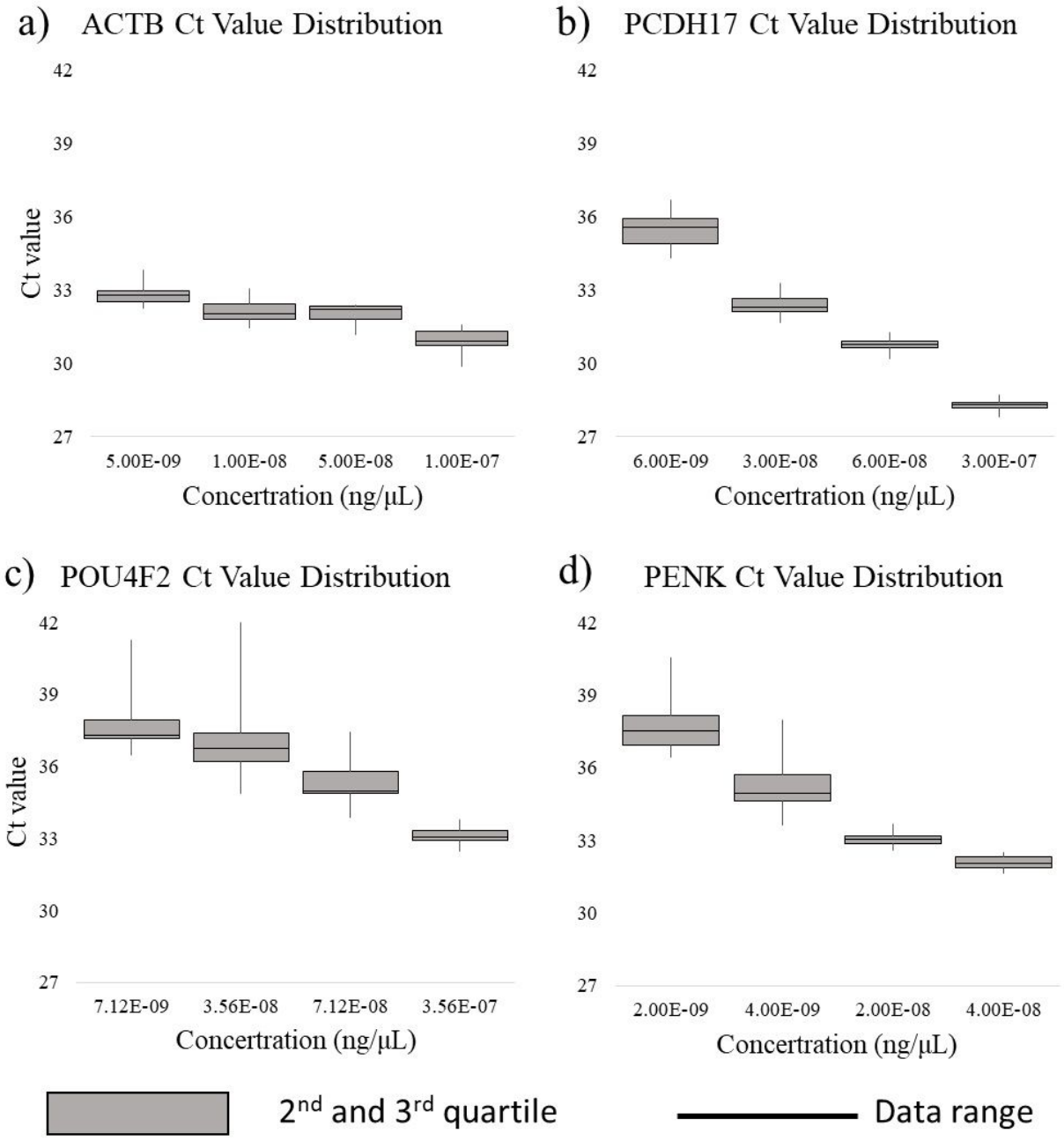

Figure 6

LOD study for individual P3 markers and reference marker. Tests of each sample of each concentration were repeated 20 times. Mean, second and third quartiles and the full range were illustrated.