Ability of a urine gene expression classifier to reduce the number of follow-up cystoscopies in bladder cancer patients

RUTH MONTALBO, JUAN J. LOZANO, LAURA IZQUIERDO, MERCEDES INGELMO-TORRES, CARMEN BAÑOS, JOAN PALOU, ANTOINE G. VAN DER HEIJDEN, RAFAEL MEDINA, JOERG SCHMIDBAUER, ALEIX PRAT, MARIA J. RIBAL, ANTONIO ALCARAZ, and LOURDES MENGUAL

BARCELONA, AND SEVILLA, SPAIN; NIJMEGEN, THE NETHERLANDS; AND VIENNA, AUSTRIA

This study aimed to improve our previous urine gene expression classifiers focusing on the detection of non−high-risk non−muscle-invasive bladder cancer (NMIBC), and develop a new classifier able to decrease the frequency of cystoscopies during bladder cancer (BC) patients' surveillance. A total of 597 urines from BC patients, controls and patients in follow-up for BC (PFBC) were included. The study has 3 phases. In the urinary biomarker discovery phase, 84 urines from BC and control patients were retrospectively included and analyzed by Ribonucleic Acid (RNA) sequencing. In the classifier development phase, a total of 132 selected genes from previous phase were evaluated by nCounter in 214 prospectively collected urines from PFBC (98 with tumor). A diagnostic classifier was generated by logistic regression. Finally, in the classifier validation phase, a multicentric and international cohort of 248 urines (134 BC and 114 nonrecurrent PFBC) was used to validate classifier performance. A total of 521 genes were found differentially expressed between non−high-risk NMIBC samples and all other groups ($P < 0.05$). An 8-gene diagnostic classifier with an area under curve (AUC) of 0.893 was developed. Validation of this classifier in a cohort of PFBC achieved an overall sensitivity (SN) and a negative predictive value (NPV) of 96% and 97%, respectively (AUC = 0.823). Notably, this accuracy was maintained in non−high-risk NMIBC group (SN = 94%; NPV = 98%). In conclusion, this 8-gene expression classifier has high SN and NPV in a real clinical scenario. The use of this classifier can reduce the number of follow-up cystoscopies in PFBC, although assessing its final place in clinical setting is necessary. (Translational Research 2019; 000:1−12)

Abbreviations: BC = bladder cancer; NMIBC = non-muscle invasive bladder cancer; SN = sensitivity; SP = specificity; LG = low grade; NR-PFBC = non-recurrent patients in follow-up for bladder cancer; C = control; MIBC = muscle invasive bladder cancer; PCA = principal component analysis;

FDR = false discovery rate; FC = fold change; AUC = area under curve; NPV = negative predictive value; PPV = positive predictive value; HG = high grade; RT-qPCR = reverse transcription quantitative PCR; GO = gene ontology

AT A GLANCE COMMENTARY

Ruth M, et al.

Background

Cystoscopy, combined in some cases with urine cytology, is considered the standard of care for bladder cancer (BC) surveillance, although it has some limitations, especially its invasiveness. Cytology, the non-invasive methodology, has inter-observer variability and low sensitivity in low-grade tumors and, precisely, the majority of recurrent patients under surveillance for BC present low grade disease.

Translational Significance

Here we have developed and validated a urine gene expression classifier with an overall high sensitivity and negative predictive value useful to detect BC during patients' surveillance. Notably, this high accuracy is maintained in the detection of low grade disease. The use of this classifier in the clinical setting would allow us to replace cystoscopy during follow-up or lower cystoscopy frequency in a routine fashion. Furthermore, the classifier is based on an effortless and uncomplicated methodology which allows for easy translation into clinical practice.

INTRODUCTION

Seventy to 80 percent of bladder cancer (BC) cases are non−muscle-invasive BC (NMIBC) at diagnosis, of which 20 to 80%, depending of the risk group, will have 1 or several recurrences. Consequently, NMIBC patients require frequent and long-term surveillance. The follow-up (FU) schedule consists of cystoscopy and in case of high-risk NMIBC, cystoscopy combined with urine cytology. Depending on the patient's risk profile, the European Association of Urology guidelines recommend upto 15 cystoscopies during the first 5 years of FU.[1]

Cystoscopy is an invasive technique with patient discomfort and possible complications. Moreover, it has been estimated that it can overlook 10%−20% of the papillary lesions and 50% of flat bladder lesions, and could also be inconclusive some cases.[2,3] For this reason, cystoscopy is associated in some cases with cytology[1] in BC diagnostic and FU schedules as its

specificity (SP) reaches 98%.[4] Cytology is a noninvasive methodology but it has low sensitivity (SN), inter-observer variability, subjective evaluation, and low-accuracy in low-grade (LG) tumors.[5]

Over the last decades, many studies have attempted to identify urinary biomarkers that could replace cystoscopy. In fact, several urinary biomarkers have been reported in the literature and some of them are even Food and Drugs Administration FDA-approved, but none of these tests have been incorporated into routine clinical practice due to their insufficient diagnostic performance.[6]

We have previously reported and validated several gene expression classifiers in exfoliated urinary cells with a diagnostic accuracy equal or superior to the current gold standard (cystoscopy combined with cytology).[7−9] However, these classifiers lack SN in LG recurrences and were only evaluated in case-control cohorts. Here we investigate the urine gene expression profile specifically associated with non−high-risk NMIBC tumors in order to identify a specific set of urine biomarkers in this subgroup of patients. The combination of the newly identified biomarkers with the previously reported by our group could lead to the development of a more accurate classifier for detecting LG tumors. This would allow us to replace cystoscopy during FU or lower cystoscopy frequency in a routine fashion.

MATERIAL AND METHODS

Patients and samples. Multicenter and international study performed according to Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines[10] and reported according to REporting recommendations for tumour MARKer prognostic studies (REMARK) guidelines.[11] The protocol was approved by an institutional review committee at Hospital Clinic, Barcelona (Spain), and by local ethical committees of the 4 external participating institutions; Fundació Puigvert (Spain), Radboud University Nijmegen (The Netherlands), Hospital Vírgen del Rocío (Spain), and University of Vienna (Austria).

The study was divided into 3 phases: identification of non−high-risk NMIBC specific urinary biomarkers (discovery phase), development of a diagnostic classifier (training set phase), and validation of the diagnostic classifier (testing set phase) (Fig 1). A total of 597 urine samples from BC patients, non-recurrent patients in FU for BC (NR-PFBC) and controls without neoplastic urological disease (C) were enrolled between April 2008 and June 2016 in the different centers. External centers
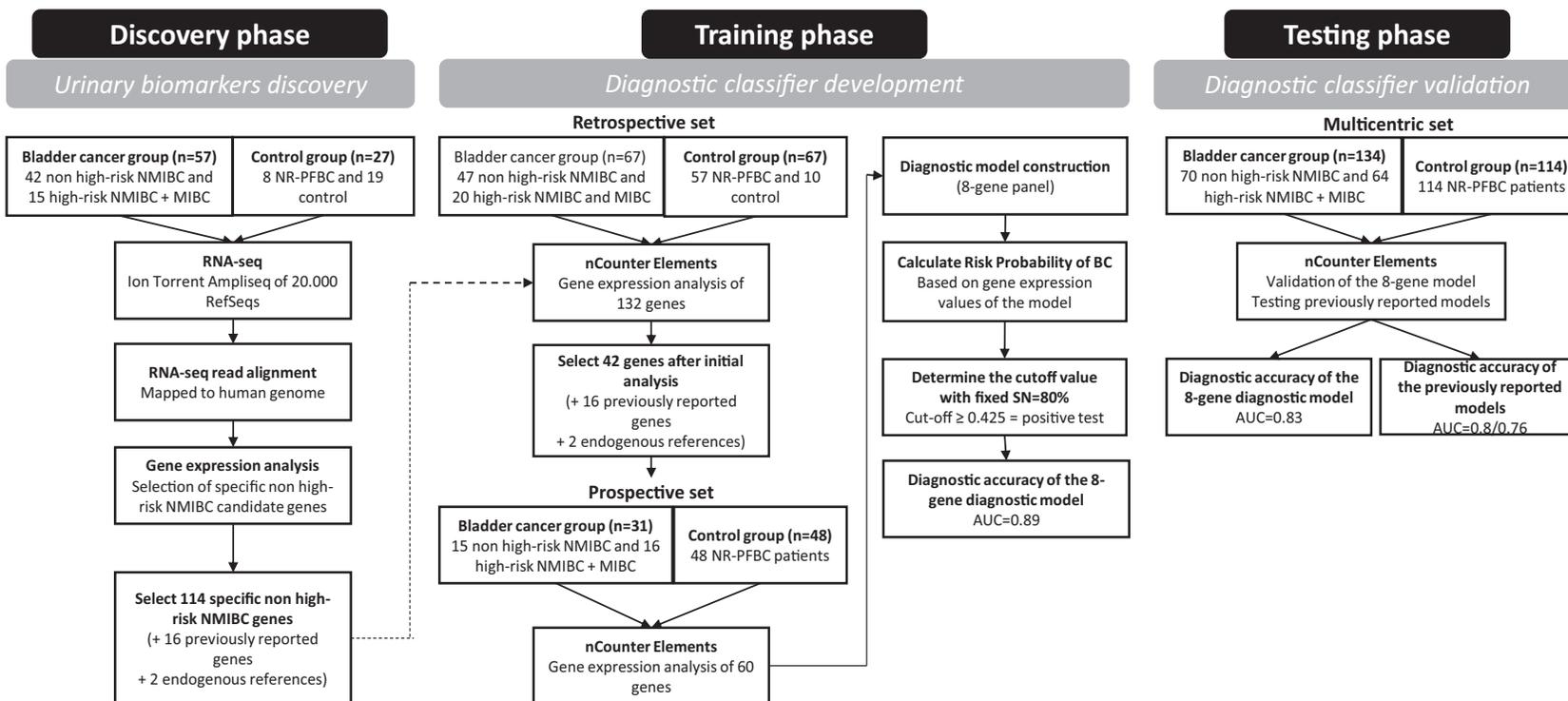
Fig 1. Study design. Urine samples from 57 BC patients and 27 control subjects were collected for RNA-seq analysis. High-risk NMIBC and MIBC patients were included in the analysis as controls to avoid selection of genes related to these groups. A subset of genes differentially expressed specifically in non−high-risk NMIBC was selected to be tested in a total of 215 urine samples from BC patients and controls. Then, an 8-gene classifier for predicting the probability of BC was constructed using multivariate logistic regression. Diagnostic performance of the classifier was validated in an independent multicentric cohort of 249 urine samples from 134 BC patients and 114 NR-PFBC.

Abbreviations: AUC, area under the curve; BC, bladder cancer; NMIBC, non-muscle invasive bladder cancer; NR-PFBC, non-recurrent control patients in follow-up for BC; SN, sensitivity.

were asked to collect and prepare the urine samples for final processing at the Hospital Clinic.

After excluding nonevaluable samples (see nCounter Elements gene expression analysis section), 546 samples were included in the study (Table I). The grade and stage of the tumors were determined according to WHO criteria[12] and Tumor Node Metastasis (TNM) classification,[13] respectively. Tumors were classified into 3 categories according to their risk: non−high-risk NMIBC, high-risk NMIBC and muscle invasive BC. Bacillus Calmette Guérin (BCG) treatments were applied following European Association of Urology guidelines.[1]

Urine cytology. Urine cytologies were performed according to Papanicolau staining and were evaluated by expert pathologists in each participating center blinded to the patient's clinical history. The results were either considered as positive, negative, or suspicious. Suspicious cytology was defined as those samples that contained cells with morphologies that could not be clearly classified as tumor cells or normal cells.

Urine processing and RNA isolation. Around $50-100$ mL of voided urine was collected from all patients of the series. Urine samples were processed as previously described.[14] Briefly, urine samples were mixed with 1 of 25 volumes of 0.5 M EDTA, pH 8.0, stored at 4°C and processed within the next 24 hours. Urines were centrifuged at $1000 \times g$ for 10 minutes, the cell pellets were resuspended in 1 mL of TRIzol reagent (Invitrogen, Carlsbad, California) and frozen at -80°C until Ribonucleic Acid (RNA) extraction. RNAs from the urinary cell pellets were extracted following manufacturer's instructions and quantified with a NanoDrop1000 (NanoDrop Technologies, Wilmington, Delaware). In the discovery phase, RNA integrity was assessed with Agilent Bionalyzer by using Eukaryote Total RNA Nano kit (mean Integrity Number (RIN) value was 2.5; range: $0-8.9$).

Library preparation and sequencing method. RNA sequencing of 20,000 RefSeqs was performed in the discovery phase. Ion AmpliSeq Transcriptome Human Gene Expression Kit (Thermo Fisher Scientific, P/N A26325) was used for library preparation. Briefly, cDNA was synthesised from total RNA by using the SuperScript VILO cDNA Synthesis kit (Thermo Fisher Scientific) from 10 ng of RNA. Then, cDNA was amplified using Ion AmpliSeq technology. Finally, after a partial digestion of the primer sequence with FUPA reagent, ligation of the barcoded adapters, and purification by Agencourt AMPure XP Reagent of the amplified cDNA, the library was quantified with Ion Library TaqMan Quantitation Kit (Thermo Fisher Scientific).

An input concentration of 8 pooled libraries copy/ Ion Sphere Particles (ISPs) was added to the emulsion Polymerase Chanin Reaction (PCR) master mix and the emulsion was generated using the Ion Chef Instrument (Thermo Fisher Scientific) using the Ion PI Hi-Q Chef Kit (Thermo Fisher Scientific). Template-positive Ion Sphere Particles were enriched, and sequencing was undertaken using an Ion PI v3 (Thermo Fisher Scientific) on the Ion Proton sequencer (Thermo Fisher Scientific) using the Ion PI Hi-Q Sequencing 200 Kit (Thermo Fisher Scientific).

Read alignment and differential gene expression analysis. Partek Flow 6.0 (http://www.partek.com/) was used to analyze AmpliSeq transcriptome data. Briefly, primary read alignment for AmpliSeq sequencing data of all samples was performed using the Torrent Mapping Alignment Program. After quantification, features with a minimum ≤1.0 were excluded. Between-sample normalization at gene level was performed using the trimmed mean method followed by quantile normalization. Gene-specific analysis was used to identify a statistical model that is the best for a specific transcript, and then the best model was used to test for differential expression.

nCounter elements gene expression analysis. Gene expression levels in training and testing sets were quantified via the nCounter Elements approach, according to manufacturer's instructions (NanoString Technologies, Seattle, West Coast). Briefly, probes were designed to target sequences of interest and the corresponding Elements TagSet (consisting of a fluorescently labeled specific Reporter Tag and a biotinylated universal Capture Tag). The probes were hybridized with the Elements TagSet and 250 ng of RNA for a minimum of 16 hours at 67°C in a thermocycler. Samples were stored at 4°C for a maximum of 24 hours until purified in the nCounter Prep Station to remove unligated probes. Expression data was collected by direct digital counting of the target molecules in each sample using the nCounter Digital analyzer. Those samples with less than 10 counts were excluded from the analysis. Relative expression values for the genes analyzed were used to calculate the risk of presenting BC. If the predicted probability of the model was higher than the established cut-off point value, the samples were classified for the gene expression signature as tumor sample. All the researchers from the Hospital Clinic involved in this analysis were blinded to the patients' clinical data, ensuring the reliability of the results.

Data analysis. NanoString raw data was processed in the R statistical environment (v3.3.2) and normalization was performed using the NanoStringNorm package.[15] The normalization setting was performed using the geometric mean of the 2 housekeeping genes (*BGUS*, and *PPIA*). Logistic regression was used to generate diagnostic models. Performance was

Table I. Clinicopathological and demographic characteristics of the study population classified by **(A)** the study phase and **(B)** the participating center (testing set)

(A) Bladder cancer patients	Discovery phase *N (%)*	Training set *N (%)*	Testing set *N (%)*	Total *N (%)*
Gender				
Male	45 (79)	73 (74)	100 (75)	**218 (75)**
Female	12 (21)	25 (26)	34 (25)	**71 (25)**
Age				
Mean	69	72	73	**71**
Range	53−95	46−90	30−94	**30−95**
Pathological stage				
pTa	42 (74)	64 (65)	73 (54)	**179 (62)**
pT1	6 (11)	14 (14)	44 (33)	**64 (22)**
pT2	7 (12)	15 (15)	7 (5)	**29 (10)**
pTis	2 (3)	5 (6)	10 (7)	**17 (6)**
Histological grade				
Low	44 (77)	61 (62)	49 (37)	**154 (53)**
High	13 (23)	37 (38)	85 (63)	**135 (47)**
Subtotals	57	98	134	**289**
Non-recurrent patients in follow-up for BC	*N (%)*	*N (%)*	*N (%)*	***N (%)***
Gender				
Male	5 (63)	75 (71)	97 (85)	**177 (77)**
Female	3 (37)	31 (29)	17 (15)	**51 (23)**
Age				
Mean	67	71	73	**70**
Range	53−84	48−90	40−89	**40−90**
Pathological stage before TURB				
pTa	6 (75)	68 (64)	76 (67)	**150 (66)**
pT1	2 (25)	21 (20)	31 (27)	**54 (24)**
pT2	-	12 (11)	5 (4)	**17 (7)**
pTis	-	5 (5)	2 (2)	**7 (3)**
Histological grade before TURB				
Low	6 (75)	67 (63)	78 (68)	**151 (66)**
High	2 (25)	39 (37)	36 (32)	**77 (34)**
Subtotals	8	106	114	**228**
Control patients without neoplasic diseases	N (%)	N (%)	N (%)	**N (%)**
Gender				
Male	9 (47)	6 (60)	-	**15 (52)**
Female	10 (53)	4 (40)	-	**14 (48)**
Age				
Mean	69	63	-	**66**
Range	35−84	39−83	-	**35−84**
Urinary condition				
Urolithiasis	12 (63)	7 (70)	-	**19 (66)**
BPH	7 (37)	3 (30)	-	**10 (24)**
Subtotals	19	10	-	**29**
Total	84	214	248	**546**

(B)	Testing set Hospital Clínic	Fundació Puigvert	Radboud University Nijmegen	Virgen del Rocio	University of Vienna
Bladder cancer patients	*N (%)*	*N (%)*	*N (%)*	*N (%)*	*N (%)*
Gender					
Male	33 (77)	37 (77)	16 (84)	9 (56)	6 (75)
Female	10 (23)	11 (23)	3 (16)	7 (44)	2 (25)
Age					
Mean	74	71	74	69	76
Range	55−91	39−89	30−86	45−82	67−94
Pathological stage					
pTa	18 (42)	31 (65)	15 (79)	10 (63)	4 (50)

(continued)

Table I. (*Continued*)

(B)

	Testing set Hospital Clínic	Fundació Puigvert	Radboud University Nijmegen	Virgen del Rocio	University of Vienna
pT1	14 (33)	14 (29)	4 (21)	5 (31)	2 (25)
pT2	1 (2)	3 (6)	-	1 (6)	2 (25)
pTis	10 (23)	-	-	-	-
Histological grade					
Low	31 (72)	31 (65)	17 (89)	11 (69)	4 (50)
High	12 (28)	17 (35)	2 (11)	5 (31)	4 (50)
Subtotals	43	48	19	16	8
Non-recurrent patients in follow-up for BC	N (%)	N (%)	N (%)	N (%)	N (%)
Gender					
Male	75 (87)	5 (83)	14 (82)	-	3 (60)
Female	11 (13)	1 (17)	3 (18)	-	2 (40)
Age					
Mean	74	70	74	-	77
Range	40−89	60−78	47−89	-	69−80
Pathological stage before TURB					
pTa	65 (76)	-	6 (35)	-	-
pT1	15 (17)	-	7 (41)	-	-
pT2	4 (4)	-	2 (12)	-	-
pTis	2 (3)	-	2 (12)	-	-
Histological grade before TURB					
Low	67 (78)	-	6 (35)	-	-
High	19 (22)	-	11 (65)	-	-
Subtotals	86	6	17	-	5
Total	129	54	36	16	13

Abbreviations: BPH, benign prostate history; MIBC, muscle invasive bladder cancer; NMIBC, non−muscle-invasive bladder cancer; TURB, trans-urethral resection of the bladder.

evaluated by Receiver Operating Characteristics (ROC) curves. Comparisons with a *P* value <0.05 were considered statistically significant.

ToppGene (https://toppgene.cchmc.org/)[16] was used to detect functional enrichment of the 8 genes from the classifier. Gene-gene interaction networks for the genes of the model were built by the GeneMANIA Cytoscape 3.0.0 plugin.[17] Physical, coexpression, and pathway gene-gene interactions were evaluated.

RESULTS

Urinary biomarkers discovery. Initial exploratory assessment of the RNA-seq dataset was performed using principal component analysis. As shown, there is a noticeable overlap of urine samples from non−high-risk NMIBC and all other groups. Nevertheless, 521 genes were found to be specifically differentially expressed (*P* < 0.05) between non−high-risk NMIBC and all the other groups (Fig 2 and Table S1).

Classifier development. A total of 114 key differentially expressed genes selected from previous phase (false discovery rate ≤0.12 and absolute fold change ≥1.5) plus 16 selected genes from our previous classifiers[7−9] (Table S2) were analyzed in 136 urine samples. Thereafter, 60 genes were selected (false discovery rate ≤0.12 and absolute fold change ≥1.5) to be tested in 79 independent additional samples (Table S3). Logistic regression analysis was used to generate an 8-gene expression classifier (*ANXA10, IGF2, KIFC3, KRT20, LCN2, MAGEA3, RPS21,* and *SLC1A6*) that achieves an area under curve (AUC) = 0.893 for detecting BC. At a fixed SN of 80%, the classifier reaches 89% SP.

Gene ontology analyses of the 8 genes included in the classifier revealed highly significant enrichment of diseases associated with bladder malignances (Fig 3, A). The generated network by GeneMANIA shows that there are no direct interactions between the 8 genes of the model although 4 of them show coexpression (Fig 3, B).

Classifier validation. The 8-gene classifier was validated in an independent multicentric and international cohort of 249 voided urine samples from PFBC. The 8-gene classifier performed well in this real clinical scenario (AUC = 0.823), achieving an overall SN of 96%. More importantly, the SN of the classifier in non−high-risk NMIBC group was 94%. Notably, the SN of the classifier in the subset of recurrent tumors (SN = 92%; *n* = 72) was comparable to that of primary tumors (SN = 100%; *n* = 62).

On the other hand, the SP of the classifier decreased to 18% in this validation set, as evidenced by AUC value (Table II). Interestingly, we found that the performance of the 8-gene classifier was not affected by number of tumors (*P* = 0.43) nor by tumor size (*P* = 0.34). Fig 4 depicts the risk probabilities derived from the 8-gene classifier in NR-PFBC, non−high-risk NMIBC and high-risk NMIBC+MIBC, both in training and testing sets.

In this phase, we oversampled patients with BC. We therefore cannot directly calculate predictive values from the test results. In daily practice, it is known that recurrence is detected in approximately 10% of all FU cystoscopies performed (90% of patients previously diagnosed with NMIBC are without recurrence at time of FU cystoscopy). In order to calculate predictive values that reflect values in real clinical practice, we assumed the distribution of recurrent vs nonrecurrent to be 10% vs 90%. For this, we multiplied the NR-PFBC samples by 10. Thus, the positive and negative predictive value (NPV) in the validation phase become 66% and 97%, respectively. If NMIBC patients with a negative classifier do not undergo
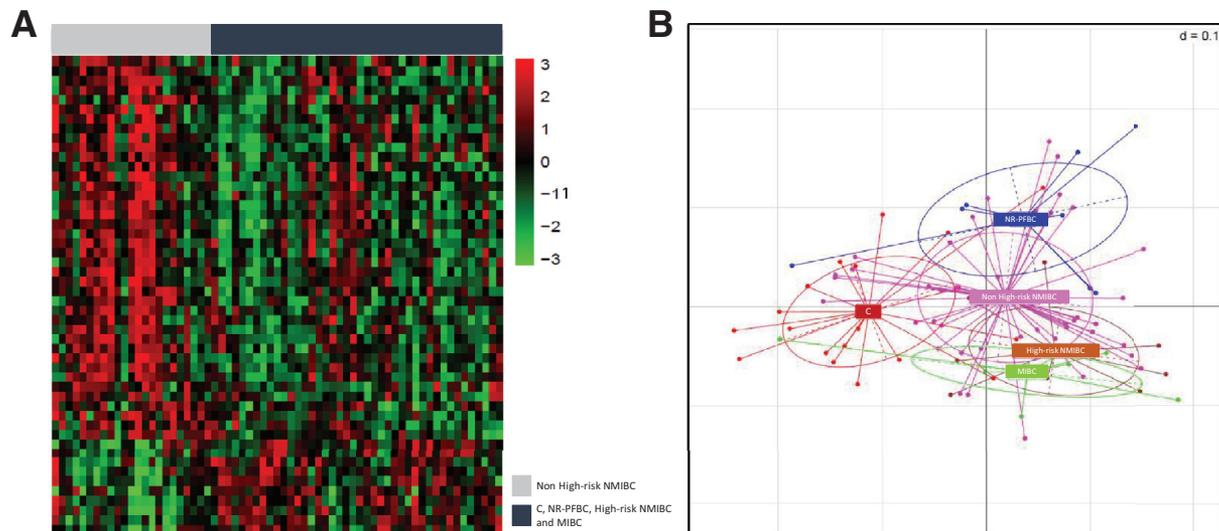
Fig 2. **(A)** Heat map representation of the top 50 differentially expressed genes between low-grade NMIBC and the other groups from the discovery set. Red pixels correspond to an increased abundance of mRNA in the urine samples, whereas green pixels indicate decreased mRNA levels. **(B)** PCA representation of differences between all samples.

Abbreviations: C, controls without neoplasic urological disease; NMIBC, Non-muscle invasive bladder cancer; NR-PFBC, non-recurrent control patients in follow-up for BC. (Color version of figure is available online.)

cystoscopy, this means that ∼17% of all cystoscopies can be prevented at the cost of 2% of recurrences remaining undiagnosed.

Performance comparison of 8-gene classifier and urine cytology. Overall cytology results were available for 198 patients in training and 209 in testing set. Performance of cytology in detecting BC is shown in Table II. In both training and testing set, SN of the 8-gene classifier (80% and 96%, respectively) was higher than that of urine cytology (35% and 43%, respectively) in our cohort of samples. NPV is also higher in both training and testing sets for the 8-gene classifier than for urine cytology (Table II). This means, in the case of the testing set, that 53% of the tumors (51 out of 96) that were missed by urine cytology were detected by the 8-gene classifier. Notably, the classifier only misses four tumors (all were non−high-risk NMIBC) which also were not detected by cytology. On the contrary, positive predictive value and SP are higher for urine cytology than for the 8-gene classifier (Table II).

The combination of the 8-gene model and urinary cytology results does not significantly improve the diagnostic performance of the 8-gene classifier alone, neither in the training set (AUC = 0.903) nor in the testing set (AUC = 0.825) (Fig S1).

DISCUSSION

Patients with NMIBC are submitted to frequent invasive cystoscopies during their surveillance. Cytology is useful, particularly as adjunct to cystoscopy, if high-grade (HG) disease is present, but it presents low SN in LG tumors. Thus, the identification of a reliable non-invasive tool for overall detection of BC would be of great usefulness in the clinical setting, especially for the non−high-risk NMIBC group. Considering this lack of accurate tests for non−high-risk NMIBC, in

this study, we applied an RNA-seq methodology to identify non−high-risk NMIBC specific urinary RNA biomarkers. Moreover, the combination of these new identified biomarkers with our previously described BC urine biomarkers has allowed us to develop and validate an improved 8-gene expression classifier useful for BC monitoring.

To the best of our knowledge, this is the first time that a list of genes differentially expressed specifically in non−high-risk NMIBC tumors is reported. To identify these genes, 5 subgroups of urine samples were evaluated and only those genes differentially expressed specifically in the non−high-risk NMIBC group of patients were selected for further analysis. As we aimed at developing a test for the global diagnosis of BC, the 16 genes reported in our previous classifiers with a high diagnostic performance in HG tumors[7−9] were also included in the subsequent analysis. After 2 validation rounds of all these selected key genes, we have been able to develop a global 8-gene expression classifier that performs equally in all BC risk groups. Overall SN of the classifier is 96%, with no significant difference in LG (96%) and HG (95%) tumors. Of note, the classifier was validated in an international and multicentric cohort of samples. Furthermore, with the aim of translating our classifier to routine clinical setting, we have developed the classifier using a non-observer dependent and non−labor-intensive methodology. The 8 genes included in the classifier
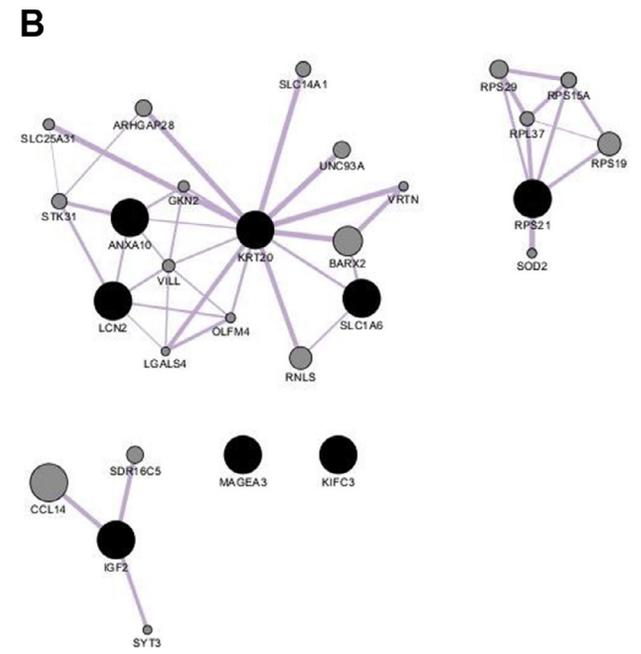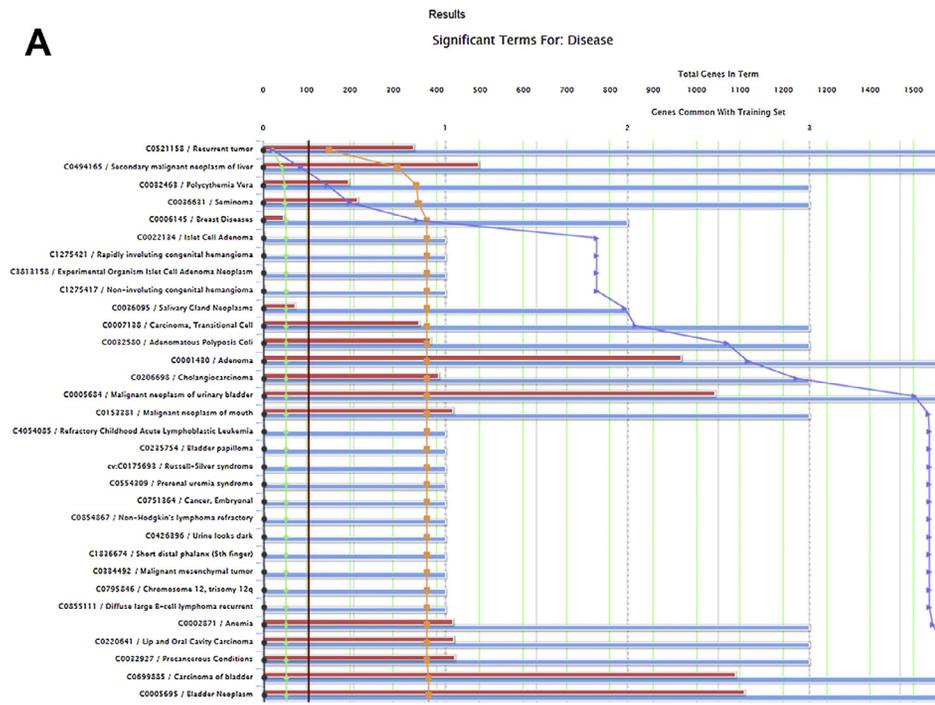
Fig 3. Functional gene enrichment analysis (**A**) and gene-gene interaction network (**B**) for the 8 genes of the classifier. (Color version of figure is available online.)

Table II. Diagnostic performance of the 8-gene test, cytology and combination of both in testing set (at fixed sensitivity of 80% in the training set)

		Training set 8-gene classifier	Cytology*	Testing set 8-gene classifier	Cytology*	Combination
N samples		213	198	249	209	209
AUC		0.893	-	0.823	-	0.825
Sensitivity (%)	Overall	80	35	96	43	96
	Non high-risk NMIBC	86	15	94	40	97
	High-risk NMIBC	96	56	97	54	85
	MIBC	87	75	100	60	100
	LG	87	14	96	33	95
	HG	92	68	95	51	96
NPV (%)	Overall	85	68	97	67	98
	Non high-risk NMIBC	92	73	98	70	99
	High-risk NMIBC	99	91	99	94	99
	MIBC	98	97	100	98	100
	LG	91	72	99	80	99
	HG	97	91	98	80	99
Specificity (%)	NR-PFBC	89	92	18	96	22
	C	80	100	-	-	-

Abbreviations: C, non-neoplasic controls; HG, high grade; LG, low grade; MIBC, muscle invasive bladder cancer; NMIBC, non−muscle-invasive bladder cancer; NR-PFBC, non-recurrent patients in follow-up for bladder cancer.
*Cytology was only available for a subset of samples.

can be easily analyzed in a multiplex format using the nCounter technology.[18]

Recently, several urinary tests for BC FU have been described. Sin et al[19] used RNA-seq to develop a 3-biomarker test to predict BC with an overall SN of 83% and SP of 89%. However, they designed a case-control study and, although it has been tested in 6 PFBC, validation of the results in an independent cohort is necessary to determine the usefulness of the classifier in a real clinical scenario. On the other hand, there are currently 2 commercially available tests based on gene expression quantification by RT-qPCR of 5 markers in urine samples; Cxbladder (Pacific Diagnostics) achieved an AUC of 0.87 in a cohort of first diagnosed BC patients while AUC values drop to 0.66 in a surveillance cohort. This performance in surveillance was improved with the development of CxBladder Monitor test which presents an AUC value of 0.73.[20] Xpert BC (Cepheid) achieves an AUC of 0.87 in a validation cohort of PFBC. Of note, SN in primary and recurrent tumors was 73% and 77%, respectively. However, SN drops to 65% in LG disease.[21] Interestingly, 2 out of the 5 genes included in their final classifier were selected from our previously reported classifiers.[7−9] A test with high SN and NPV is crucial in the BC surveillance scenario to safely lower the number of FU cystoscopies. Our 8-gene expression classifier outperforms the current gold standard and previously developed gene expression tests in terms of SN and NPV.[7−9] Using the test, ∼17% of PFBC from our validation cohort could safely skip cystoscopy. The remaining PFBC should undergo cystoscopy. This is not a major problem, since in normal daily practice PFBC would have undergone a cystoscopy anyway. Using our 8-gene classifier, 6 recurrent PFBC were wrongly diagnosed as not having a recurrence; 4 of them had TaLG, 1 had a TIS and 1 had a T1HG tumor. Of note, cytology also missed these 6 tumors. Furthermore, it is known that cystoscopy, our gold standard, misses upto 15% of the papillary and upto 30% of the flat lesions.[22]

It should be considered that a high proportion of patients who presented non−high-risk tumors are being long-term monitored. In our series, 60% of the BC samples analyzed are non−high-risk NMIBC and a SN of 94% and NPV of 98% was found in this subgroup of patients. This population may be hugely benefited from the avoidance of continuing surveillance procedures which are invasive. Of importance, the 8-gene classifier maintains a high SN and NPV in high-risk NMIBC disease (97% and 99%, respectively) guarantying the detection of potential life-threatening tumors. Lastly, the accuracy of our 8-gene classifier was not affected by number of tumors neither did it correlate with tumor size, further supporting its usefulness in the FU of BC patients.

We found that the 8 genes composing the classifier regulate the carcinogenesis process via different mechanisms. This is shown by the fact that the genes of the model have no direct interactions and derive from different pathways. This 8-gene classifier incorporates 3 genes not previously included in our reported classifiers: *LCN2, RPS21*, and *KIFC3. LCN2* facilitates tumorigenesis by promoting
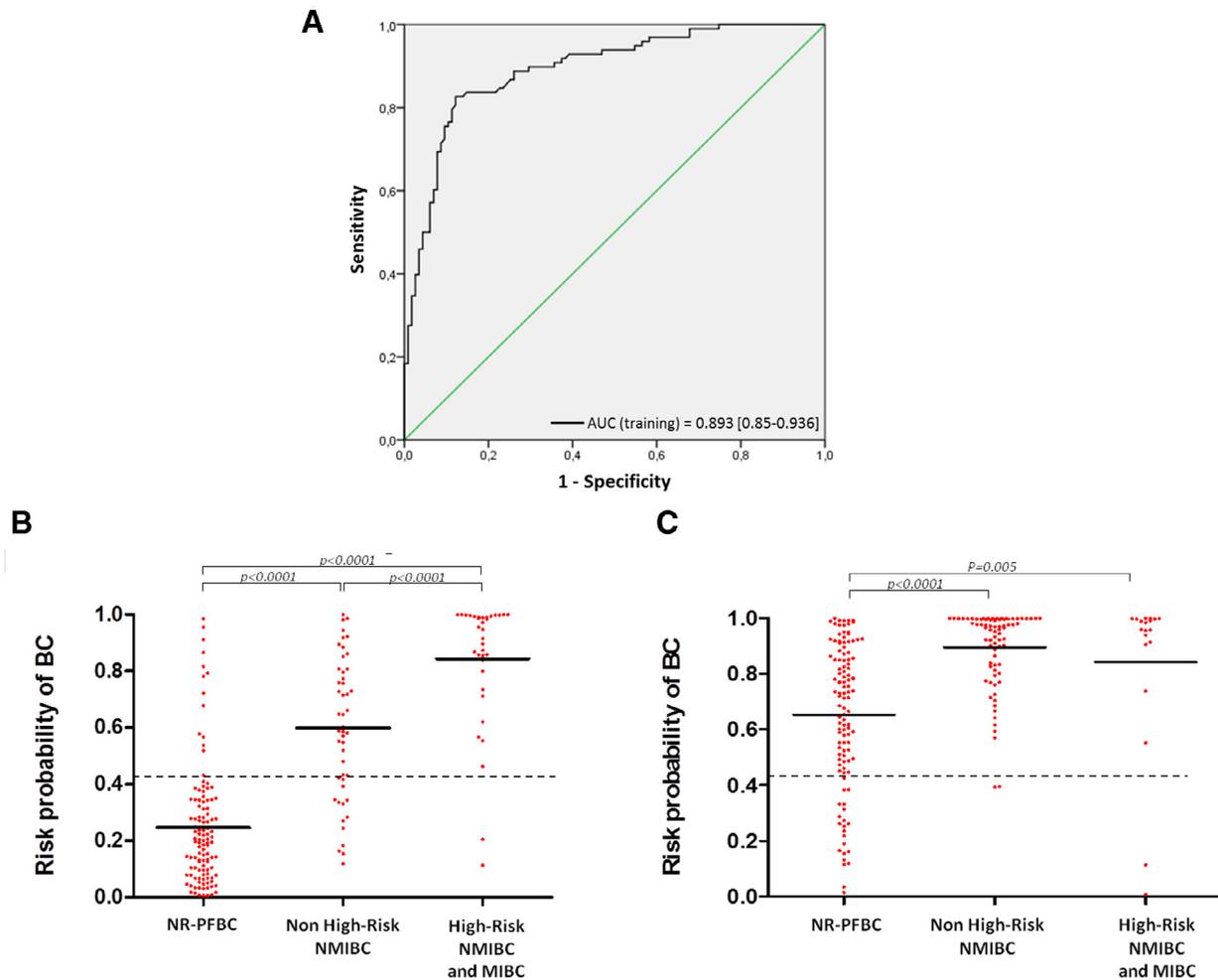
Fig 4. Performance of the 8-gene classifier. **(A)** ROC curve of 8-gene model in training set. **(B)** Risk of tumor probability representation according to the 8-gene expression classifier in training and **(C)** testing sets (cut-off = 0.425).

Abbreviations: AUC, area under the curve; BC, bladder cancer; NMIBC, Non−muscle-invasive bladder cancer; NR-PFBC, non-recurrent control patients in follow-up for BC. (Color version of figure is available online.)

survival, growth, and metastasis[23] and it has been reported several cancers.[24] Preliminary evidence also suggests that *LCN2* may be used as a prognostic urinary biomarker of breast cancer.[25] According to our results, Arthurs et al[26] found that *RPS21* presented significantly raised levels of expression in malignant prostate cancer tissue. Finally, *KIFC3* plays a role in the formation, maintenance, and remodeling of the bipolar mitotic spindle. It has been not related previously with carcinogenetic processes. Gene ontology analysis confirmed that the genes included in the classifier are highly related to bladder malignancies, corroborating the suitability of these genes to be useful for the detection of BC.

The strengths of this study lie in the fact that we have chosen a 3-stage approach using PFBC, the

patients in whom the test should be applied. We have also used a blind approach to identify genes particularly related with non−high-risk BC patients. Furthermore, the use of voided urine samples to analyze gene expression allows for the development of a non-invasive BC diagnostic tool with an easy translation into clinical practice. Lastly, the test is based on a multiplexing format directly quantifiying by a hybridization gene expression. However, some limitations should be mentioned. To avoid inefficiency, patients with recurrence were oversampled by also recruiting patients who were scheduled for a bladder transurethral resection of a proven bladder tumor. Consequently, some of the patients had primary tumors. Moreover, the number of NR-PFBC was misrepresented in the validation series and we

had to make an estimation to calculate the number of cystoscopies that could be skipped. Second, 8.5% of samples had to be excluded due to technical failures. Eventually, the low SP of the test should be improved to prevent unnecessary cystoscopies.

In conclusion, our 8-gene expression classifier has high SN and NPV in a real clinical scenario and this performance is maintained in non−high-risk NMIBC group. If only patients with a positive classifier result would be cystoscopied, 17% of all cystoscopies can be prevented. For a definitive conclusion, these results must be validated through prospective large-scale randomized clinical trials for assessing its final place in clinical setting.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.trsl.2019.02.003.

REFERENCES

1. Babjuk M, Bohle A, Burger M, et al. EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder: update 2016. Eur Urol 2017;71:447–61.
2. Goh AC, Lerner SP. Application of new technology in bladder cancer diagnosis and treatment. World J Urol 2009;27:301–7.
3. Smith ZL, Guzzo TJ. Urinary markers for bladder cancer. F1000Prime Rep 2013;5:21.
4. Mowatt G, Zhu S, Kilonzo M, et al. Systematic review of the clinical effectiveness and cost-effectiveness of photodynamic diagnosis and urine biomarkers (FISH, ImmunoCyt, NMP22) and cytology for the detection and follow-up of bladder cancer. Health Technol Assess 2010;14:1–331, iii−iv.
5. Gakis G, Kruck S, Stenzl A. Can the burden of follow-up in low-grade noninvasive bladder cancer be reduced by photodynamic diagnosis, perioperative instillations, imaging, and urine markers? Curr Opin Urol 2010;20:388–92.
6. Soria F, Droller MJ, Lotan Y, et al. An up-to-date catalog of available urinary biomarkers for the surveillance of non-muscle invasive bladder cancer. World J Urol 2018;36:1981–95.
7. Mengual L, Burset M, Ribal MJ, et al. Gene expression signature in urine for diagnosing and assessing aggressiveness of bladder urothelial carcinoma. Clin Cancer Res 2010;16:2624–33.
8. Mengual L, Ribal MJ, Lozano JJ, et al. Validation study of a noninvasive urine test for diagnosis and prognosis assessment of bladder cancer: evidence for improved models. J Urol 2014;191:261–9.
9. Ribal MJ, Mengual L, Lozano JJ, et al. Gene expression test for the non-invasive diagnosis of bladder cancer: a prospective, blinded, international and multicenter validation study. Eur J Cancer 2016;54:131–8.
10. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. The standards for reporting of Diagnostic Accuracy Group. Croat Med J 2003;44:639–50.
11. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumour MARKer prognostic studies (REMARK). Eur J Cancer 2005;41:1690–6.
12. Lopez-Beltran A, Sauter G, Gasser T, et al. Tumours of the urinary system. In: Eble JN, Sauter G, Epstein JI, Sesterhenn IA, eds. Pathology and genetics of tumours of the urinary system and male genital organs. World Health Organization classification of tumours, Lyon: IARC Press, 2004:89–157.
13. Sobin LH, Gospodariwicz M, Wittekind CH. TNM classification of malignant tumours. UICC International Union Against Cancer, 7th ed. New York: Wiley-Blackwell; 2009:243–8.
14. Mengual L, Burset M, Marin-Aguilera M, Ribal MJ, Alcaraz A. Multiplex preamplification of specific cDNA targets prior to gene expression analysis by TaqMan arrays. BMC Res Notes 2008;1:21.
15. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. Bioinformatics 2012;28:1546–8.
16. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res 2009;37:W305–11.
17. Montojo J, Zuberi K, Rodriguez H, Bader GD, Morris Q. GeneMANIA: fast gene network construction and function prediction for Cytoscape. F1000Res 2014;3:153.
18. Bonneterre J, Prat A, Galvan P, Morel P, Giard S. Value of a gene signature assay in patients with early breast cancer and intermediate risk: a single institution retrospective study. Curr Med Res Opin 2016;32:835–9.
19. Sin MLY, Mach KE, Sinha R, et al. Deep sequencing of urinary RNAs for bladder cancer molecular diagnostics. Clin Cancer Res 2017;23:3700–10.
20. Kavalieris L, O'Sullivan P, Frampton C, et al. Performance characteristics of a Multigene Urine Biomarker Test for monitoring for recurrent urothelial carcinoma in a multicenter study. J Urol 2017;197:1419–26.

21. Pichler R, Fritz J, Tulchiner G, et al. Increased accuracy of a novel mRNA-based urine test for bladder cancer surveillance. BJU Int 2018;121:29–37.

22. Grossman HB, Gomella L, Fradet Y, et al. A phase III, multicenter comparison of hexaminolevulinate fluorescence cystoscopy and white light cystoscopy for the detection of superficial papillary lesions in patients with bladder cancer. J Urol 2007;178:62–7.

23. Rodvold JJ, Mahadevan NR, Zanetti M. Lipocalin 2 in cancer: when good immunity goes bad. Cancer Lett 2012;316:132–8.

24. Iannetti A, Pacifico F, Acquaviva R, et al. The Neutrophil Gelatinase-Associated Lipocalin (NGAL), a NF-kappaB-regulated gene, is a survival factor for thyroid neoplastic cells. Proc Natl Acad Sci USA 2008;105:14058–63.

25. Yang J, Moses MA. Lipocalin 2: a multifaceted modulator of human cancer. Cell Cycle 2009;8:2347–52.

26. Arthurs C, Murtaza BN, Thomson C, et al. Expression of ribosomal proteins in normal and cancerous human prostate tissue. PLoS ONE 2017;12:e0186047.