Development of a Multiplex RNA Urine Test for the Detection and Stratification of Transitional Cell Carcinoma of the Bladder

Andrew Holyoake,[1] Paul O'Sullivan,[1] Rob Pollock,[1] Terry Best,[1] Jun Watanabe,[3] Yoichiro Kajita,[3] Yoshiyuki Matsui,[3] Masaaki Ito,[3] Hiroyuki Nishiyama,[3] Natalie Kerr,[1] Fernanda da Silva Tatley,[1] Lisa Cambridge,[1] Tumi Toro,[1] Osamu Ogawa,[3] and Parry Guilford[1,2]

Abstract

Purpose: New markers that enable the percentage of transitional cell carcinomas (TCC) of the bladder that are diagnosed before invasion of the bladder muscle layers to be increased would reduce the morbidity and mortality associated with this disease. The purpose of this study was to develop a simple, accurate urine test based on mRNA markers and simple gene signatures that (*a*) could detect TCC before muscle invasion while maintaining high specificity in patients with hematuria or urinary tract infections and (*b*) identify patients most likely to have grade 3 or stage ≥T1 disease.

Experimental Design: RNA markers with high overexpression in stage Ta tumors and/or T1 to T4 tumors but low expression in blood or inflammatory cells were characterized by quantitative reverse transcription-PCR using 2 mL of voided urine from 75 TCC patients and 77 control patients with other urological diseases.

Results: A combination of the RNAs *CDC2, MDK, IGFBP5*, and *HOXA13* detected 48%, 90%, and 100% of stage Ta, T1, and ›T1 TCCs, respectively, at a specificity of 85%. Detection of Ta tumors increased to 60% for primary (non-recurrent) Ta tumors and 76% for Ta tumors ≥1 cm in diameter. Test specificity was 80% for the 20 control patients with urinary tract infections. The combination of *CDC2* and *HOXA13* distinguished between grade 1 to 2 TCCs and grade 3 or stage ≥T1 TCCs with ~80% specificity and sensitivity.

Conclusions: Simple gene expression signatures can be used as urine markers for the accurate detection and characterization of bladder cancer.

There are ~360,000 new cases of bladder cancer worldwide annually and ~145,000 deaths from the disease (1). About 30% of tumors have invaded into or beyond the muscularis propria at the time of diagnosis. Patients with these muscle-invasive tumors (stages T2-T4) have 10-year recurrence-free survival rates ranging from >80% for organ-confined tumors to ~30% for tumors with regional lymph node metastases (2). Muscle-invasive tumors are generally treated by radical cystectomy with bilateral pelvic lymph node dissection followed by urinary diversion. In contrast, tumors that have not invaded the muscularis propria (stage Ta, T1, and *in situ* carcinomas) are usually resected by organ-preserving transurethral resection and present 10-year survival rates of 70% to 85% (2). Screening high-risk groups, such as the elderly, cigarette smokers, and certain occupational groups including truck drivers, painters, and workers in the textile dye and rubber tire industries (3, 4), would be predicted to lead to a higher proportion of tumors being identified at the pre–muscle-invasive stage and an overall decrease in the bladder cancer burden.

Recent studies using high-throughput microarray analysis have shown that mRNA expression profiles are a powerful method for discriminating between bladder tumors and normal urothelium and defining subgroups of bladder tumors (5–8). The informativeness of these expression profiles and the high frequency of tumor cell exfoliation into urine have raised the prospect of using simple gene expression signatures obtained from urine total RNA to better diagnose bladder cancer and provide more information on disease characteristics at the time of diagnosis.

We hypothesized that mRNAs with high overexpression in transitional cell carcinomas (TCC) compared with normal urothelium and low expression in blood and inflammatory cells would constitute sensitive urine markers that also showed high specificity in patients presenting with hematuria or urinary tract infections. Further, we predicted that using a combination of markers would provide greater test accuracy by accounting for tumor heterogeneity and the marked differences in gene expression that occur between low-grade, well-differentiated TCCs and poorly differentiated, high-grade cancers (8, 9).

Here, we describe the selection and validation of mRNA markers for the detection of TCC that are significantly overexpressed

Authors' Affiliations: [1]Pacific Edge Biotechnology Ltd., Centre for Innovation; [2]Cancer Genetics Laboratory, Department of Biochemistry, University of Otago, Dunedin, New Zealand; and [3]Department of Urology, Kyoto University, Kyoto, Japan
Received 7/6/07; revised 9/4/07; accepted 10/31/07.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Requests for reprints: Parry Guilford, Pacific Edge Biotechnology Ltd., Centre for Innovation, University of Otago, P. O. Box 56, Dunedin, New Zealand. Phone: 64-3-4795803; Fax: 64-3-4795801; E-mail: parry.guilford@peblnz.com.

doi:10.1158/1078-0432.CCR-07-1672

in both early- and late-stage disease, have low expression in blood and inflammatory cells, and are stable in urine. Combination of these markers into a multiplex test provides for greater sensitivity than can be achieved with single markers and also provides a means to differentiate between low-grade Ta tumors and tumors that are either high grade or have an invasive phenotype.

Materials and Methods

Patients/samples. All samples were obtained from patients treated at the Department of Urology, Kyoto University Graduate School of Medicine between 2000 and 2006. Cancer stages were assigned according to the International Union Against Cancer tumor-node-metastasis staging system (10). Grades were assigned using the WHO criteria (11) and reviewed by two pathologists. Fifty-eight TCC samples were obtained by cold-cup biopsy at the time of transurethral resection for use in the mRNA quantification study. A further 18 normal urothelium samples were obtained by stripping the epithelium of ureters extracted at the time of radical nephrectomy for renal cell carcinoma (12). Voided urine samples were obtained from an independent series of TCC patients; 36 samples were from patients with no prior history of bladder cancer, and 39 samples were from patients under surveillance for bladder cancer recurrence. Seventy-seven voided urine samples were also obtained from patients with nonmalignant diseases in whom bladder tumors were excluded by flexible cystoscopy. These samples were from patients with urinary tract infections (n = 20), other nonmalignant urological diseases including benign prostate hyperplasia, urolithiasis, and neurogenic bladder (n = 24), and nonbladder urological cancers (n = 33). Five of 20 of the urinary tract infection patients had hematuria and 20 of 20 had cystitis, as evidenced by positive urine WBC counts. Urine cytology results were available for patients with confirmed TCC; samples with cytology grades 4 to 5 were classified as positive for TCC. This study was approved by the Institutional Review Board of Kyoto University Graduate School. Written informed consent to participate in the study was obtained from each patient before surgery according to the local ethical guidelines.

RNA extraction. Tumor and nonmalignant tissues were collected into RNAlater (Qiagen), homogenized in RLT lysis buffer (Qiagen), and stored at -80°C. The RNA was purified using Trizol extraction (Invitrogen) and RNeasy mini kits (Qiagen). The methodology for the extraction of RNA from urine was based on the method of Menke and Warnecke (13). Briefly, voided urine samples were first mixed with an equal volume of RNA extraction buffer [5.64 mol/L guanidine thiocyanate, 0.5% sarkosyl, 50 mmol/L sodium acetate (pH 6.5), 1 mmol/L β-mercaptoethanol; pH adjusted to 7.0 with 1.5 mol/L HEPES (pH 8.0)]. Total RNA was purified from the urine/extraction buffer mix by Trizol extraction and the RNeasy procedure. Briefly, 4 mL urine/extraction buffer was mixed with 3 mL Trizol and 5 μg poly(A) carrier RNA (Roche) in 15 mL heavy Phaselock tubes (Eppendorf). Chloroform (0.6 mL) was added and the mixture was centrifuged at 3,200 × g for 10 min at 4°C. Absolute ethanol (0.53 volume) was added to the aqueous layer and loaded onto RNeasy mini columns on a vacuum manifold. RNA was eluted in 35 μL water by centrifugation. The extraction of RNA from blood was done using Trizol/RNeasy extractions on cells enriched from whole blood by sedimentation in 3.6% dextran (14).

Tissue/blood quantitative PCR. Single-stranded cDNA synthesis was done on 2 μg RNA using oligo(dT) priming and SuperScript II (Invitrogen). Quantitative PCR (qPCR) was done on 50 ng cDNA in 16 μL reactions using either 300 nmol/L primers, 275 nmol/L FAM/TAMRA dual-labeled Taqman probes (Supplementary Table S1), or Applied Biosystems Assays-on-Demand and 2× ROX qPCR master mix (Abgene). Reactions were heated to 95°C for 15 min before being cycled 40 times between 95°C for 15 s and 60°C for 1 min. Reactions were done in duplicate on an ABI Prism 7000 qPCR machine and analyzed using the SDS version 1.2.3 software.

Urine quantitative reverse transcription-PCR. All primers and probes (Supplementary Table S1) were obtained from Integrated DNA Technologies. The fluorophore/quencher pair for all markers was JOE/BHQ-1, with the exception of *HOXA13*, which used FAM/BHQ-1. Both singleplex and multiplex probe quantitative reverse transcription-PCRs (qRT-PCR) were run in 16 μL single-well reactions incorporating RNase OUT (Invitrogen), RNase H- SuperScript III (Invitrogen), and FastStart Taq Polymerase (Roche) using MicroAmp optical 96-well reaction plates (Applied Biosystems). Each reaction contained the following: 3 μL RNA, 0.275 μmol/L each primer, 0.225 μmol/L each probe, 10 mmol/L DTT, 4.5 mmol/L $MgSO_4$, 2.64 μL FastStart GC Rich, 0.375 mmol/L deoxynucleotide triphosphate mix (dATP, dGTP, dCTP, and dUTP), 0.625 μmol/L ROX, 1× FastStart PCR buffer, 1.5 units FastStart Taq Polymerase, 1.25 units SuperScript III, and 1 unit RNase OUT. RNA templates were plated using a Biomek NX liquid handling multiaxis robot with subsequent addition of qRT-PCR master mix. Plates were sealed using MicroAmp optical adhesive film (Applied Biosystems) and centrifuged at 1,000 rpm for 1 min. Thermocycling was conducted on an ABI Prism 7000 series qPCR machine using ROX as a passive reference unless otherwise stated. Cycling conditions were as follows: 50°C for 15 min, 95°C for 8 min, 10 cycles of 95°C for 15 s and 60°C for 2 min followed by 30 cycles of 95°C for 15 s and 60°C for 1 min. All reactions were done in duplicate.

qPCR data analysis. To avoid variability in the analysis of individual plates using the autoanalysis function of the ABI Prism SDS version 1.2.3 software, a threshold for each detector was set manually, whereby the slope of the standard curve approached -3.32 and the R^2 value approached 1. The data set for each probe was then analyzed using these manual thresholds. This method minimizes detector bias on a plate with large numbers of high or low C_T samples. Dilutions (1:16) of a reference RNA (derived from pooled cell line RNAs) were included on each plate to generate a standard curve ranging from 0.3 pg/μL to 20 ng/μL. To provide standardization across PCR plates, C_Ts were expressed as ΔC_Ts, where ΔC_T = C_T (sample) - C_T [reference RNA (20 ng/μL)]. ΔC_Ts were averaged across the duplicates.

Statistical procedures. qRT-PCR ΔC_T values were used to generate classification scores by Fisher linear discriminant analysis (LDA). This analysis generates a formula (the linear discriminant, LD1) that is a linear combination of ΔC_T values and can be used to discriminate between the two classes of samples (i.e., non-TCC versus TCC). LD1 was determined for each marker or combination of markers using 10 rounds of 10-fold cross-validation to test performance. When only a single marker is analyzed, LDA reduces to a scaling of the ΔC_T values. Each round of 10-fold cross-validation was carried out by first randomly dividing the samples into 10 stratified groups (across TCC and non-TCC samples), and then LDA was carried out to determine the value of LD1 for 1 of the 10 groups using the remaining samples to train the classifier. This was repeated for each of the 10 groups so that, during a particular round of cross-validation, each sample is "left out" of the training exactly once. Our approach was to do the 10 rounds of 10-fold cross-validation training using samples from patients with primary TCC and those with nonmalignant urological conditions or other urological cancers. Samples from TCC patients with recurrent disease were included in the 10-fold cross-validation testing phase but were not used in the training of the classifiers.

Values for sensitivity and specificity, as well as 95% confidence ranges for these values, were obtained using the LD1 values (for each marker or combination of markers) and the distribution of these values across training runs. The 95% confidence range for sensitivity at a given specificity was obtained by first calculating the 95% confidence limit for the cutoff value (corresponding to that specificity) and then calculating the 95% confidence interval for the sensitivity at the extreme values for the cutoff. The values of LD1 for each sample were assumed to be normally distributed across different training runs.

The ability of *CDC2-HOXA13* to discriminate between tumors of different stage and grade was determined in a similar manner. In this case, the choice of training group was kept to those primary tumors that

were positive in the uRNA-D diagnostic (TCC versus non-TCC) assay. Again, 10 rounds of 10-fold cross-validation were conducted using LDA to generate discriminants.

All analyses were done using the R package for statistical computing.[4] The LDA and logistic regression (generalized linear models) analysis used the MASS package, and ROC curves (15) and sensitivity and specificity analyses were carried out using the ROCR package (16).

Results

Selection of candidate diagnostic markers. Candidate urine markers were selected from a bladder cancer gene expression array data set obtained using custom-printed arrays containing 30,000 oligonucleotides (MWG Biotech) corresponding to ~26,600 different genes (11). The array data for genes with high overexpression in either stage Ta or T1 to T4 tumors (compared with nonmalignant urothelium) were validated by qRT-PCR (Table 1) using RNA extracted from 18 normal urothelium samples and up to 28 Ta tumors and 30 stage T1 to T4 tumors. Of 14 candidate markers, the DNA replication genes *TOP2A* and *CDC2* showed the highest median overexpression in both stage T1 to T4 tumors and stage Ta tumors but also a pronounced differential between these tumor stages. *TOP2A* and *CDC2* also showed the least overlap between normal and tumor tissues, with 93% of T1 to T4 tumors and 88% of Ta tumors having expression greater than the 95th percentile of the expression in normal urothelium. *BIRC5* showed high median overexpression in T1 to T4 tumors but poor overall separation between tumors and normals, with only 11% of the T1 to T4 tumors having expression higher than the 95th percentile of the normal samples. Five candidate markers (*ENG*, *IGFBP5*, *NRP1*, *SEM2*, and *HOXA13*) showed greater overexpression in stage Ta tumors than stage T1 to T4 tumors, with the greatest differential being shown by *HOXA13* and *IGFBP5*. *MDK* showed consistent, but relatively low, overexpression in all tumor types and no differential between Ta and T1 to T4 cancers.

To test the accumulation of these markers in urine, qRT-PCR was carried out with each candidate using total urine RNA extracted from a preliminary series of patients with confirmed TCC, other urological cancers, and nonmalignant urological indications, including infection and hematuria. A decision to include each marker in further analysis was made based on the reproducibility of each qRT-PCR assay, the differential between marker accumulation in TCC patients and controls, and the level of expression in blood or inflammatory cells. Blood or inflammatory cell expression was estimated by qRT-PCR assays of RNA extracted from whole blood (results not shown) and reference to the EST counts in blood and lymph found in the UniGene EST Profile Viewer.[5] Following this triage, *CDC2*, *MDK*, *IGFBP5*, and *HOXA13* were selected for the final development of assays for both TCC detection and stratification of patients with low- and high-risk disease.

Detection of TCC. To characterize the selected markers and develop an algorithm for their combined use, a cohort of 75 urine samples from TCC patients and 77 samples from patients presenting with nonmalignant urological disease or urological cancer other than TCC was analyzed. The mean age of TCC cases and controls in the cohort was 72 and 69 years, respectively. Eighty-two percent and 79% of the TCC patients and controls were male, respectively.

qRT-PCR was done on all samples using each of the markers *CDC2*, *MDK*, *IGFBP5*, and *HOXA13* and the ΔC_Ts were determined. In addition to assessing the performance of individual markers, we used LDA to develop a multivariate test that consisted of all four markers (designated uRNA-D). Sensitivities and specificities for each individual marker and uRNA-D were determined and the characteristics of each test were further evaluated using ROC curves (16). Because of the potential effect on test characteristics of samples from patients with urinary tract infections, test sensitivities and specificities were calculated both in the absence and presence of 20 infection samples. Figure 1 shows ROC curves for the detection of TCCs of stage Ta, stage ≥T1, grade 1 to 2, and grade 3 in the absence of the infection samples (75 TCCs and 57 non-TCC controls). Separate curves are shown for each of *MDK*, *IGFBP5*, *HOXA13*, *CDC2*, and the test that combined all four markers (uRNA-D). At high specificity (>90%), *IGFBP5* and *HOXA13* showed greater sensitivity of detection of low-stage/low-grade tumors than *MDK* and *CDC2*. However, *MDK* and *CDC2* were markedly more sensitive in the detection of high-grade and high-stage TCCs across the entire ROC curve. The ROC curve for uRNA-D was comparable with both *CDC2* and *MDK* for the high-grade/high-stage TCCs and to *IGFBP5* and *HOXA13* for the low-stage/low-grade tumors at high specificity.

The sensitivity of each individual marker and the combination assay uRNA-D in this cohort of 75 TCCs and 57 infection-negative controls is shown in Table 2 at a specificity of 85%. At this specificity, *MDK* and *CDC2* were the single markers with the highest sensitivity for high-stage/high-grade TCCs. *MDK*, for example, detected 87% of grade 3 tumors, 90% of stage T1 tumors, and 100% of stage >T1 tumors. uRNA-D detected 91% of grade 3 tumors, 90% of stage T1 tumors, and 100% of stage >T1 tumors. *IGFBP5* and *MDK* showed the greatest sensitivity for low-stage/low-grade TCC, although both only detecting 41% of all stage Ta cancers in this study (comprising TCCs detected in both the primary and surveillance settings). This detection rate increased to >50% for the detection of primary Ta tumors alone and to ~60% for the detection of larger Ta tumors (>1 cm). uRNA-D showed better detection of low-stage/low-grade disease, with 48% of all Ta tumors in this cohort being detected. This increased to 60% for the detection of primary Ta tumors and 76% for the detection of Ta tumors ≥1 cm. Urine cytology compared poorly to the RNA markers, detecting only 13%, 40%, and 80% of the Ta, T1, and stage >T1 TCCs in this series, respectively (Table 2).

At the same overall test specificity used in Table 2 (85%), the inclusion of the infection samples into the non-TCC control group had minor effects on the sensitivity of *IGFBP5*, *HOXA13*, and *MDK: IGFBP5* and *HOXA13* showed small improvements in sensitivity and MDK showed no more than a 3% loss in performance for any tumor stage or grade (Table 3). There was, however, a significant change in *CDC2*, with an average 10% decrease in sensitivity across the different tumor stage and grade classifications. The effect on the combination assay, uRNA-D, was intermediate to the single marker tests, showing, for example, 1% and 3% decreases in the sensitivity of detection of

[4] http://www.r-project.org/

[5] http://www.ncbi.nlm.nih.gov/UniGene

Table 1. qPCR analysis of candidate bladder cancer markers in tumors and normal urothelium

Candidate	Symbol	Stage T1-T4		Stage Ta	
		Median T:N fold change	%T >95th percentile of normal	Median T:N fold change	%T >95th percentile of normal
Sperm-associated antigen 5	*SPAG5*	40	27/29 (93%)	8	20/24 (83%)
Topoisomerase (DNA) IIα 170 kDa	*TOP2A*	128	27/29 (93%)	39	21/24 (88%)
Cell division cycle 2, G_1 to S and G_2 to M	*CDC2*	121	27/29 (93%)	27	21/24 (88%)
Endoglin (Osler-Rendu-Weber syndrome 1)	*ENG*	4	16/28 (57%)	8	23/27 (85%)
Insulin-like growth factor binding protein 5	*IGFBP5*	n/c*	n/c*	7	15/25 (60%)
Neuropilin 1	*NRP1*	7	25/28 (90%)	10	20/25 (80%)
Epidermal growth factor–like domain, multiple 6	*EGFL6*	4	15/27 (55%)	4	16/24 (67%)
Matrix Gla protein	*MGP*	3	17/29 (59%)	n/c*	n/c*
Semaphorin sem2	*SEM2*	6	21/29 (72%)	9	18/24 (75%)
Ubiquitin-conjugating enzyme E2C	*UBE2C*	44	14/30 (47%)	11	4/27 (15%)
Homeobox A13	*HOXA13*	4	13/28 (46%)	11	18/27 (67%)
Midkine (neurite growth-promoting factor 2)	*MDK*	2	16/27 (59%)	2	15/22 (68%)
Baculoviral IAP repeat-containing 5 (survivin)	*BIRC5*	28	3/28 (11%)	9	4/26 (15%)
SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	*SMC4L1*	6	17/29 (59%)	3	12/28 (43%)

Abbreviations: T, tumor; N, normal urothelium; n/c, not confirmed by qPCR.
*Gene expression array data showed no differential between *IGFBP5* expression in stage T1 to T4 tumors and nonmalignant urothelium nor between MGP expression levels in stage Ta tumors and nonmalignant urothelium.

grade 1 to 2 and grade 3 TCCs, respectively. Analysis of ROC curves generated using the 75 TCC samples and either the 57 noninfection controls or the 20 infection samples alone confirmed that infection samples have no negative effect on overall test sensitivity for *IGFBP5* and *HOXA13* in the 80% to 100% specificity range and only a minor effect on *MDK* and uRNA-D (Supplementary Fig. S1).

The specificity of each marker and uRNA-D for the different non-TCC control groups is shown in Table 4. With the specificity for all 77 control samples set at 85%, the specificity for the infection samples alone ranged from 75% for *CDC2* to 100% for *HOXA13*; uRNA-D had an intermediate specificity for the infection samples of 80%. There were no major differences in the marker specificities for the non-TCC cancers nor for the other types of benign urinary tract disease.

Overall, these results indicate that combining the individual markers *CDC2*, *MDK*, *IGFBP5*, and *HOXA13* into a multiplex test (uRNA-D) leads to better sensitivity than when the constituent markers are used alone. Further, the infection study supports our approach of triaging markers based on their expression in blood and inflammatory cells to provide a urine test, which is more specific when these cell types are present in urine.

Severity prediction. The reciprocal expression patterns of *CDC2* and *HOXA13* in tumor tissue from stage Ta TCCs and stage T1 to T4 TCCs (Table 1) and the different characteristics of these markers in urine samples from patients with low-stage/low-grade compared with high-stage/high-grade disease (Fig. 1) led us to hypothesize that a combination of these markers would enable patients with TCC to be stratified into groups with high and low probability of having TCC that is advanced in grade (grade 3) or stage (stage ≥T1). This stratification would allow the identification of patients in primary care who are likely to require more urgent clinical workup.

To test this hypothesis, qRT-PCR data for *HOXA13* and *CDC2* from TCC patients who were positive for uRNA-D at 85% specificity were reanalyzed. ROC curves showing the discrimination between stage, grade, and combined stage/grade classes were drawn separately for *MDK*, *HOXA13*, *CDC2*, and *HOXA13-CDC2* combined (Fig. 2A-C). In each case, the greatest accuracy in discriminating between classes was provided by the combined *HOXA13-CDC2* test. At 80% specificity, this combination (designated uRNA-S) showed 42% and 73% accuracy when differentiating between low-grade/high-grade TCC and stage Ta/stage T1 to T3 TCC, respectively. The sensitivity for discriminating between grade 1 to 2 TCCs and high-risk TCCs, which had progressed to either grade 3 or stage ≥T1, was 82%.

Discussion

The detection of bladder cancer using urine samples raises two distinct problems. First, the target cancers range from small highly differentiated polyps of low malignant potential to high-grade *in situ* carcinomas and highly invasive, poorly differentiated tumors. These different tumor types will be marked by major differences in gene and protein expression patterns. Second, the cellular content of a urine sample is affected by the size and histologic characteristics of the tumor, the presence of blood and inflammatory cells, and variation in the number of nonmalignant urothelial cells exfoliated into the urine. The challenge is to develop a test that not only accounts for the tumor heterogeneity but also exhibits high specificity in a clinical setting where frank and occult urinary tract infections and hematuria are common. We reasoned that a combination of overexpressed markers that show low expression in blood and inflammatory cells would provide the basis for a urine test with these attributes.

We elected to use a genome-wide gene expression database as our starting point for marker selection to avoid any selection biases, with the most overexpressed epithelial genes being selected for qPCR validation. The qPCR studies confirmed the

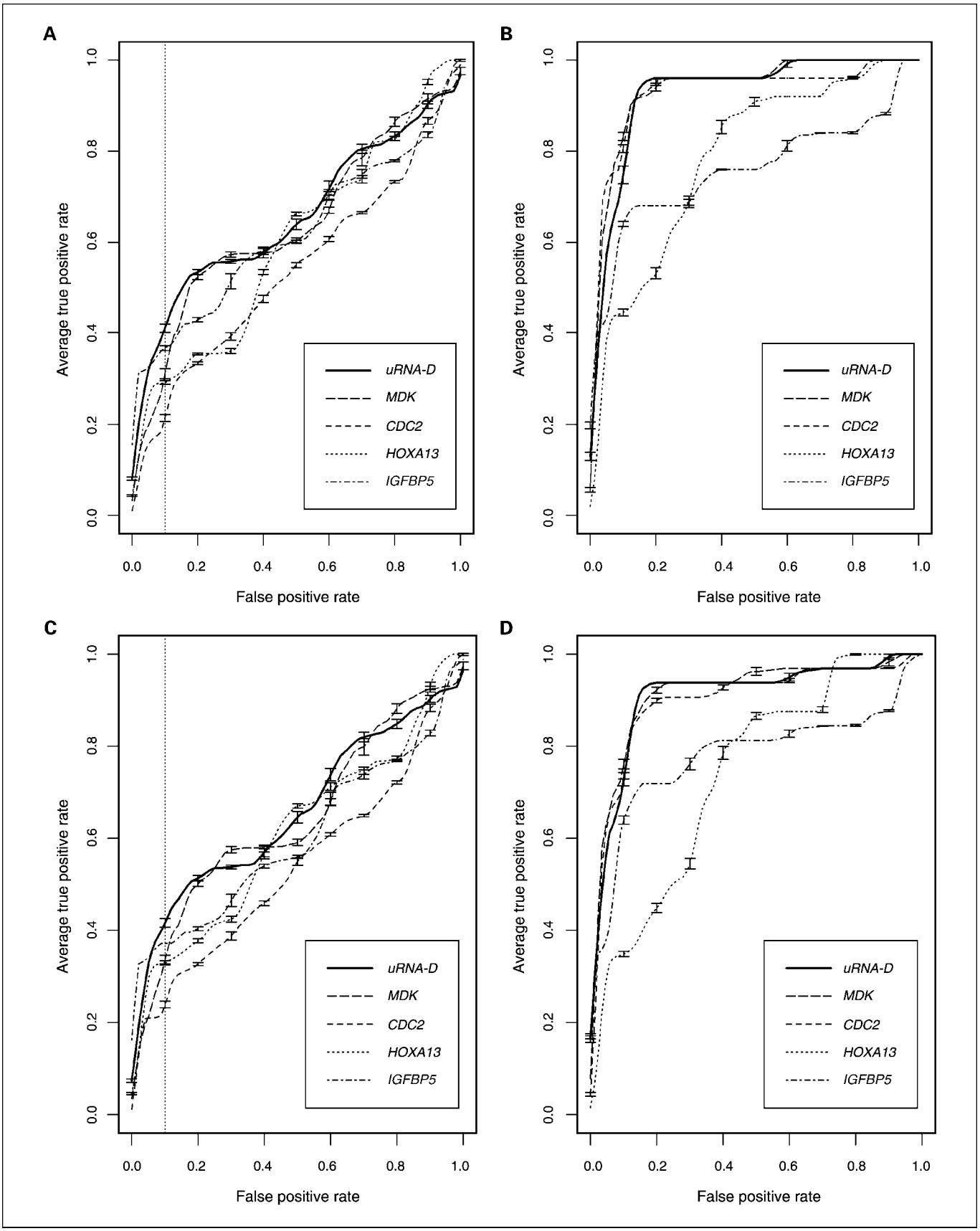

Fig. 1. ROC curves for the detection of stage Ta tumors (*A*), stage ≥T1 tumors (*B*), grade 1 or 2 tumors (*C*), and grade 3 tumors (*D*). Seventy-five TCC and 57 non-TCC samples were used in each analysis. The dotted vertical lines in *A* and *C* mark 90% specificity.

Table 2. Sensitivity of detection of TCC at 85% specificity by stage and grade using individual markers and the combined assay uRNA-D

TCC presentation	*n*	Mean percentage (95% confidence interval)					Cytology positives (%)
		MDK	CDC2	*IGFBP5*	*HOXA13*	uRNA-D	
Stage Ta	45	41 (39-42)	30 (28-32)	41 (40-42)	31 (29-31)	48 (47-49)	6/45 (13%)
Stage Ta (primary)	18	53 (51-56)	36 (33-41)	52 (50-56)	33 (33-33)	60 (57-62)	5/18 (28%)
Stage Ta (≥1 cm)	9	61 (57-67)	39 (35-49)	56 (56-56)	33 (33-33)	76 (67-79)	4/9 (44%)
Stage T1	20	90 (88-90)	90 (90-90)	65 (65-65)	36 (32-37)	90 (90-94)	8/20 (40%)
Stage T1 (primary)	13	85 (81-85)	92 (89-92)	69 (69-69)	38 (34-39)	92 (90-92)	4/13 (31%)
Stage >T1	5	100 (100-100)	100 (100-100)	80 (80-80)	100 (10-100)	100 (100-100)	4/5 (80%)
Tumor *in situ*	5	80 (80-80)	80 (80-80)	80 (80-80)	20 (20-20)	80 (80-80)	3/5 (60%)
Grade 1-2	43	41 (39-42)	31 (28-32)	38 (37-40)	34 (33-35)	47 (47-47)	7/43 (16%)
Grade 1-2 (primary)	18	56 (56-56)	41 (36-42)	56 (56-56)	44 (44-44)	61 (61-61)	5/18 (28%)
Grade 3	32	87 (85-88)	85 (84-88)	70 (69-72)	38 (36-39)	91 (89-93)	14/32 (44%)
Grade 3 (primary)	18	87 (83-89)	90 (89-94)	74 (72-78)	44 (41-45)	93 (89-95)	9/18 (50%)

NOTE: The mean sensitivities and 95% confidence intervals determined by LDA and cross-validation are shown. Fifty-seven non-TCC samples were included in this analysis. The cytology results illustrate the numbers of TCC patients with positive (grade 4-5) urine cytology tests. Cytology specificity figures were unavailable.

identity of several genes with high overexpression, particularly in the T1 to T4 tumors. Genes such as *TOP2A* and *CDC2*, which are involved in DNA synthesis and cell cycle control, showed very high overexpression across the majority of tumors examined, whereas others such as *BIRC5* (survivin) were more variable and only strongly overexpressed in a subset of the tumors. Of the four genes selected for the final combination test (uRNA-D), *HOXA13* and *IGFBP5* showed higher overexpression in stage Ta tumors than stage T1 to T4 tumors. HOXA13 is a transcription factor that is involved in morphogenesis and differentiation of the genitourinary tracts (17); these roles may explain the higher overexpression observed in the more differentiated Ta tumors. IGFBP5 is a multifunctional protein that acts both directly and through the insulin-like growth factor axis in a tissue- and time-dependent manner to affect diverse processes, such as development, differentiation, and survival (18). MDK is a heparin-binding growth factor that is highly expressed during embryogenesis and displays a range of activities, including antiapoptotic, angiogenic, and mitogenic (19).

Analysis of the ROC curves derived from the urine data showed that, in the 90% to 100% specificity range, *IGFBP5* and *HOXA13* showed better performance for the detection of Ta/low-grade tumors than either *MDK* or *CDC2*. In contrast, *MDK* and *CDC2* performed far more strongly with the more invasive, high-grade tumors. To enable us to develop a test with strong performance for both early- and late-stage cancers, we used LDA to develop algorithms that incorporated all four markers. The combined test, uRNA-D, showed better overall performance than any gene alone. In addition to LDA, the analysis was also done using logistic regression and support vector machines, with comparable results (results not shown).

The sensitivity of uRNA-D was markedly better than urine cytology across all stages and grades of TCC. For example, uRNA-D detected 90% of stage T1 cancers (at 85% specificity) compared with 40% by cytology (Table 2). This performance difference would be slightly attenuated by the higher specificity of cytology, which, although not available for this series, is typically ~95% (20). Comparisons with other urine markers

Table 3. Sensitivity of detection of TCC at 85% specificity by stage and grade in the presence of samples from patients with urinary tract infections

TCC presentation	*n*	Mean percentage (95% confidence interval)				
		MDK	*CDC2*	*IGFBP5*	*HOXA13*	uRNA-D
Stage Ta	45	39 (37-41)	22 (20-23)	42 (41-42)	36 (34-36)	47 (45-48)
Stage Ta (primary)	18	52 (49-55)	27 (24-28)	56 (56-56)	39 (36-39)	56 (56-58)
Stage Ta (≥1 cm)	9	59 (55-65)	33 (33-33)	56 (56-56)	44 (38-44)	67 (67-72)
Stage T1	20	88 (86-90)	74 (71-78)	65 (65-65)	40 (40-40)	89 (85-92)
Stage T1 (primary)	13	82 (78-85)	67 (63-74)	69 (69-69)	38 (38-38)	90 (85-93)
Stage >T1	5	100 (100-100)	100 (100-100)	80 (80-80)	100 (100-100)	100 (100-100)
Tumor *in situ*	5	80 (80-80)	80 (80-80)	80 (80-80)	20 (20-20)	80 (80-80)
Grade 1-2	43	40 (38-41)	25 (24-26)	39 (38-40)	37 (37-37)	46 (45-47)
Grade 1-2 (primary)	18	56 (54-56)	33 (30-33)	56 (56-56)	44 (44-44)	61 (61-61)
Grade 3	32	85 (82-87)	71 (69-74)	72 (72-72)	44 (43-45)	88 (84-90)
Grade 3 (primary)	18	84 (80-87)	71 (68-75)	78 (78-78)	50 (47-50)	88 (84-91)

NOTE: The mean sensitivities and 95% confidence intervals determined by LDA and cross-validation are shown. All 77 non-TCC samples, including the 20 infection samples, were included in the analysis.

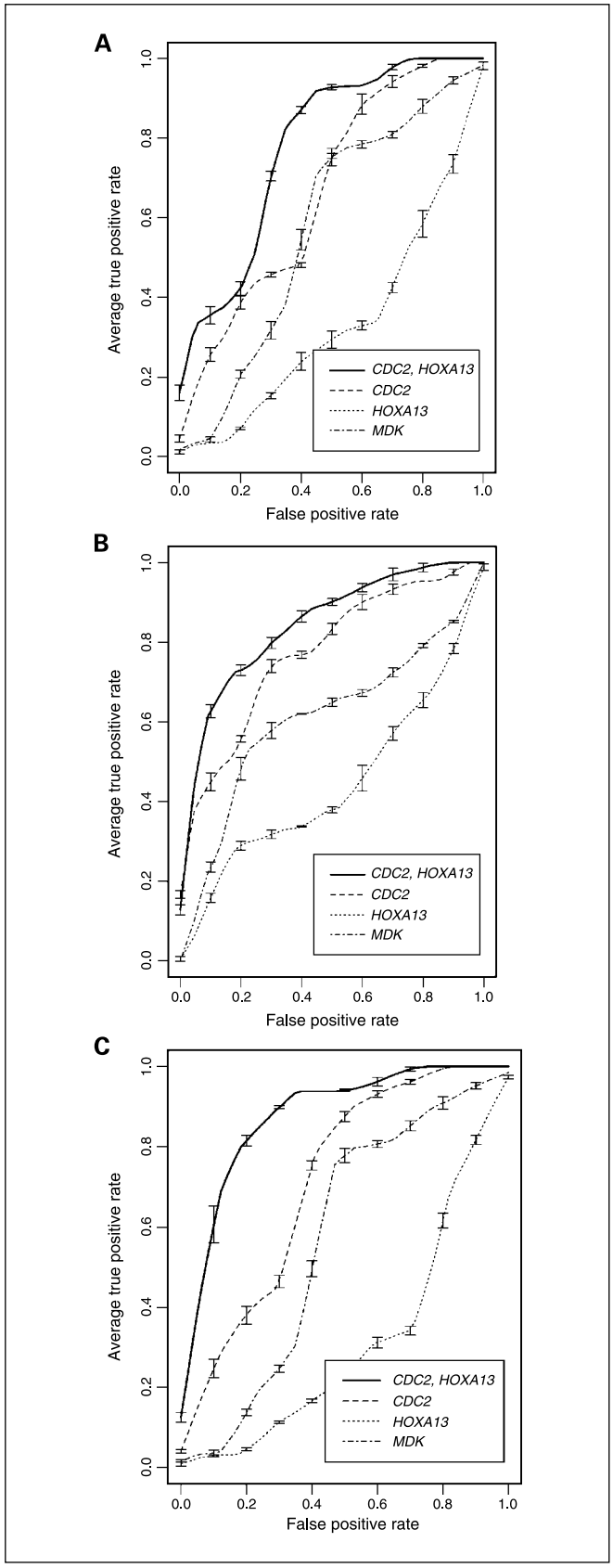

Fig. 2. ROC curves showing the discrimination between grade 1/2 tumors and grade 3 tumors (*A*), stage Ta tumors and stage ≥1 tumors (*B*), and grade 1/2 tumors and either grade 3 or stage ≥1 tumors (*C*).

reported in the literature can only be used as a guideline due to marked differences in the clinical characteristics of both the TCC patients and non-TCC controls in different study populations. With this caution in mind, comparison of the data presented here with published results suggests that uRNA-D may have better performance than the NMP22 BladderChek test (Matritech, Inc.) for the detection of primary TCC at equivalent specificities: uRNA-D detected 60% and 92% of primary Ta and primary T1 tumors, respectively, compared with the 47% and 49% detected by NMP22 (21). Comparison of uRNA-D to the UroVysion fluorescence *in situ* hybridization assay suggests comparable performance across grades (22). The sensitivity of uRNA-D for the detection of Ta tumors increased from 48% for a combination of primary and surveillance tumors to 60% for primary tumors alone and then increased again to ~76% for the detection of tumors ≥1 cm in diameter (Table 2). This variation may indicate the increasing numbers of cells exfoliated from larger TCCs. Alternatively, the expression of some or all of the markers may be correlated with the increasing size of Ta tumors.

The presence of contaminating blood or inflammatory cells is limiting for the clinical application of many bladder cancer screening tests, including urine cytology, BTA-Stat, and NMP22 (23–26). To enable the development of a simple test with better performance in the presence of these contaminating cells, we first selected markers with low expression in blood and inflammatory cells. Of the final four markers, *HOXA13* and *IGFBP5* were the most robust to the presence of infection samples in the study cohort, showing no decrease in test specificity. The low expression of these two markers in blood and inflammatory cells is likely to contribute strongly to the overall specificity and clinical utility of the combination test (uRNA-D).

Using gene expression signatures, considerably more information can be obtained from a tumor biopsy than simple diagnosis, such as likely drug or radiation response, predicted survival (6, 27, 28), and risk of recurrence (5). To explore the potential of using urine samples as a source of tumor cells for gene expression signature determination, we have used the urine levels of the marker pair *CDC2-HOXA13* (uRNA-S) to predict with ~80% accuracy if a TCC is advanced in stage or grade. Given the known inaccuracy of discriminating between stage Ta and stage T1 TCC with surgical staging (29), the actual accuracy of uRNA-S for stage prediction may exceed the figure observed here. The ability to predict disease severity at the time of an initial urine screening test may enable the prioritization of patients for urgent cystoscopy. Treatment delays are often lengthy for bladder cancer, with median delays of 15 to 20 weeks from onset of symptoms to treatment not uncommon in many centers (30, 31). Delays of even 12 weeks can result in muscle invasion of T1 tumors and an increased risk of nodal involvement with more advanced tumors (32, 33), resulting in major reductions in patient survival. The greater accuracy of uRNA-S compared with the individual markers (Fig. 2) argues that the differential between the expression of *CDC2* and *HOXA13* is contributing to the prediction accuracy rather than the effect being explained simply by increasing numbers of exfoliated cells in the high-grade/high-stage TCCs.

In conclusion, this study describes the rational design of a RNA-based test for the diagnosis and initial characterization of

Table 4. The specificity of individual markers and uRNA-D

	Test negatives/total true negatives* (%)				
	MDK	*CDC2*	*IGFBP5*	*HOXA13*	uRNA-D
Urinary tract infections	16/20 (80%)	15/20 (75%)	17/20 (85%)	20/20 (100%)	16/20 (80%)
Other benign urinary tract disease	21/24 (88%)	22/24 (92%)	21/24 (88%)	19/24 (79%)	20/24 (83%)
Other urological cancers	29/33 (88%)	29/33 (88%)	28/33 (85%)	27/33 (82%)	30/33 (91%)
Overall	85%	85%	85%	85%	85%

*True negatives were defined as samples from patients who were negative for TCC by flexible cystoscopy.

bladder cancer using urine samples. Because of the flexibility and scalability of RNA detection platforms, the potential exists to expand the test to include additional transcripts that can enhance test performance and informativeness. The results presented here will need to be further validated in a prospective setting to more accurately determine test characteristics, particularly in patients presenting with hematuria and other urological conditions.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
2. Parekh DJ, Bochner BH, Dalbagni G. Superficial and muscle-invasive bladder cancer: principles of management for outcomes assessments. J Clin Oncol 2006;24:5519–27.
3. Dryson E, Walls C, McLean D, Pearce N. Occupational bladder cancer in New Zealand: a 1-year review of cases notified to the New Zealand Cancer Registry. Intern Med J 2005;35:343–7.
4. Kirkali Z, Chan T, Manoharan M, et al. Bladder cancer: epidemiology, staging and grading, and diagnosis. Urology 2005;66:4–34.
5. Dyrskjot L, Thykjaer T, Kruhoffer M, et al. Identifying distinct classes of bladder carcinoma using microarrays. Nat Genet 2003;33:90–6.
6. Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. J Clin Oncol 2006; 24:778–89.
7. Dyrskjot L, Zieger K, Kruhoffer M, et al. A molecular signature in superficial bladder carcinoma predicts clinical outcome. Clin Cancer Res 2005;11:4029–36.
8. Thykjaer T, Workman C, Kruhoffer M, et al. Identification of gene expression patterns in superficial and invasive human bladder cancer. Cancer Res 2001;61:2492–9.
9. Sanchez-Carbayo M, Socci ND, Charytonowicz E, et al. Molecular profiling of bladder cancer using cDNA microarrays: defining histogenesis and biological phenotypes. Cancer Res 2002;62:6973–80.
10. Fleming ID, Cooper JS, Henson DE, et al. AJCC cancer staging manual. 5th ed. Philadelphia: Lippincott-Raven; 1997.
11. Ito M, Nishiyama H, Kawanishi H, et al. P21-activated kinase 1: a new molecular marker for intravesical recurrence after transurethral resection of bladder cancer. J Urol 2007;178:1073–9.
12. Booth C, Harnden P, Trejdosiewicz LK, Scriven S, Selby PJ, Southgate J. Stromal and vascular invasion in an human *in vitro* bladder cancer model. Lab Invest 1997;76:843–57.
13. Menke TB, Warnecke JM. Improved conditions for isolation and quantification of RNA in urine specimens. Ann N Y Acad Sci 2004;1022:185–9.
14. Natvig J, Perlmann P, Wigzell H, editors. Lymphocytes. Isolation, fractionation and characterization. Oslo: Universitetsforlaget, 1976. p. 184.
15. Fawcett T. ROC graphs: notes and practical considerations for researchers. Palo Alto (CA): Hewlett-Packard Laboratories; 2004.
16. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005;21:3940–1.
17. Scott V, Morgan EA, Stadler HS. Genitourinary functions of Hoxa13 and Hoxd13. J Biochem (Tokyo) 2005;137:671–6.
18. Beattie J, Allan GJ, Lochrie JD, Flint DJ. Insulin-like growth factor-binding protein-5 (IGFBP-5): a critical member of the IGF axis. Biochem J 2006;395:1–19.
19. Kadomatsu K, Muramatsu T. Midkine and pleiotrophin in neural development and cancer. Cancer Lett 2004;204:127–43.
20. Glas AS, Roos D, Deutekom M, Zwinderman AH, Bossuyt PM, Kurth KH. Tumor markers in the diagnosis of primary bladder cancer. A systematic review. J Urol 2003;169:1975–82.
21. Grossman HB, Messing E, Soloway M, et al. Detection of bladder cancer using a point-of-care proteomic assay. JAMA 2005;293:810–6.
22. Sarosdy MF, Kahn PR, Ziffer MD, et al. Use of a multitarget fluorescence *in situ* hybridization assay to diagnose bladder cancer in patients with hematuria. J Urol 2006;176:44–7.
23. Sarosdy MF, Schellhammer P, Bokinsky G, et al. Clinical evaluation of a multi-target fluorescent *in situ* hybridization assay for detection of bladder cancer. J Urol 2002;168:1950–4.
24. Sanchez-Carbayo M, Urrutia M, Silva JM, Romani R, De Buitrago JM, Navajo JA. Comparative predictive values of urinary cytology, urinary bladder cancer antigen, CYFRA 21-1 and NMP22 for evaluating symptomatic patients at risk for bladder cancer. J Urol 2001;165:1462–7.
25. Smellie WS, Forth JO, McNulty CA, et al. Best practice in primary care pathology: review 2. J Clin Pathol 2006;59:113–20.
26. van Rhijn BW, van der Poel HG, van der Kwast TH. Urine markers for bladder cancer surveillance: a systematic review. Eur Urol 2005;47:736–48.
27. Wang Y. Gene expression-driven diagnostics and pharmacogenomics in cancer. Curr Opin Mol Ther 2005;7:246–50.
28. Dyrskjot L, Zieger K, Real FX, et al. Gene expression signatures predict outcome in non-muscle-invasive bladder carcinoma: a multicenter validation study. Clin Cancer Res 2007;13:3545–51.
29. Maruniak NA, Takezawa K, Murphy WM. Accurate pathological staging of urothelial neoplasms requires better cystoscopic sampling. J Urol 2002; 167:2404–7.
30. Liedberg F, Anderson H, Mansson A, Mansson W. Diagnostic delay and prognosis in invasive bladder cancer. Scand J Urol Nephrol 2003;37:396–400.
31. Wallace DM, Bryan RT, Dunn JA, Begum G, Bathers S. Delay and survival in bladder cancer. BJU Int 2002; 89:868–78.
32. Sanchez-Ortiz RF, Huang WC, Mick R, Van Arsdalen KN, Wein AJ, Malkowicz SB. An interval longer than 12 weeks between the diagnosis of muscle invasion and cystectomy is associated with worse outcome in bladder carcinoma. J Urol 2003;169:110–5; discussion 5.
33. Chang SS, Hassan JM, Cookson MS, Wells N, Smith JA, Jr. Delaying radical cystectomy for muscle invasive bladder cancer results in worse pathological stage. J Urol 2003;170:1085–7.