Biology of Human Tumors

Clinical Cancer Research

Identification of a Candidate Gene Panel for the Early Diagnosis of Prostate Cancer

Gisele H.J.M. Leyten[1], Daphne Hessels[2], Frank P. Smit[2], Sander A. Jannink[2], Hans de Jong[2], Willem J.G. Melchers[1,2], Erik B. Cornel[3], Theo M. de Reijke[4], Henk Vergunst[5], Paul Kil[6], Ben C. Knipscheer[7], Christina A. Hulsbergen-van de Kaa[8], Peter F.A. Mulders[1], Inge M. van Oort[1], and Jack A. Schalken[1,2]

Abstract

Purpose: Serum PSA (sPSA) testing has led to the identification of patients with indolent prostate cancer and inevitably overtreatment has become a concern. Progensa PCA3 urine testing was shown to improve the diagnosis of prostate cancer, but its diagnostic value for aggressive prostate cancer is limited. Therefore, urinary biomarkers that can be used for prediction of Gleason score ≥7 prostate cancer in biopsies are urgently needed.

Experimental Design: Using gene expression profiling data, 39 prostate cancer biomarkers were identified. After quantitative PCR analysis on tissue specimens and urinary sediments, eight promising biomarkers for the urinary detection of prostate cancer were selected (*ONECUT2*, *HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *NKAIN1*, *MS4A8B*, *PPFIA2*). The hypothesis that biomarker combinations improve the diagnostic value for aggressive prostate cancer was tested on 358 urinary sediments of an intention-to-treat cohort.

Results: A urinary three-gene panel (*HOXC6*, *TDRD1*, and *DLX1*) had higher accuracy [area under the curve (AUC), 0.77; 95% confidence interval (CI), 0.71–0.83] to predict Gleason score ≥7 prostate cancer in biopsies compared with Progensa PCA3 (AUC, 0.68; 95% CI, 0.62–0.75) or sPSA (AUC, 0.72; 95% CI, 0.65–0.78). Combining the three-gene panel with sPSA further improved the predictive accuracy (AUC, 0.81; 95% CI, 0.75–0.86). The accuracy of the three-gene predictive model was maintained in subgroups with low sPSA concentrations.

Conclusion: The urinary three-gene panel (*HOXC6*, *TDRD1*, and *DLX1*) represents a promising tool to identify patients with aggressive prostate cancer, also in those with low sPSA values. The combination of the urinary three-gene panel with sPSA bears great potential for the early diagnosis of patients with clinically significant prostate cancer. *Clin Cancer Res; 1–10. ©2015 AACR.*

Introduction

With the introduction of serum prostate-specific antigen (sPSA) testing in the late 1980s, the incidence of prostate cancer has increased considerably. Worldwide, 1,111,689 men are diagnosed with prostate cancer every year of whom 307,471 men die from this disease (1). However, in patients with sPSA values between 3 and 10 ng/mL, the sPSA test has a low specificity for prostate cancer, resulting in a high negative biopsy rate of 60% to 75% (2). The specificity is low because, in addition to prostate cancer, a number of benign conditions [such as benign prostatic hyperplasia (BPH) and prostatitis] can cause elevated sPSA levels. Although sPSA-based screening reduces prostate cancer mortality by 20%, it is associated with a high risk of diagnosing clinically insignificant prostate cancer that would not have been diagnosed in the patient's lifetime in the absence of screening (3–5). Currently, it is difficult to predict which tumor will become potentially life-threatening and which one will not. Therefore, overtreatment of localized prostate cancer is a serious clinical issue with attendant, burdensome morbidities, and substantial health care costs (6). Prostate cancer–specific biomarkers that can distinguish between the aggressive prostate cancer tumor type and the indolent prostate cancer form are urgently needed to avoid the problem of overtreatment.

For the diagnosis of prostate cancer, biomarkers should ideally be detectable in body fluids that can be obtained noninvasively and therefore urine has emerged as the substrate for the noninvasive detection of prostate cancer. However, in urine, the biomarkers can be so diluted that they can only be detected when they are sufficiently present. The hypothesis is that the most promising biomarkers for the detection of a disease are those that are markedly upregulated in the disease compared with noncancerous conditions and in case of prostate cancer are therefore most likely to be detected in urine as well.

In search of such prostate cancer–specific biomarkers, 2 promising candidates have already been identified: *P*rostate *CA*ncer gene 3 (*PCA3*) and the fusion of the androgen-regulated gene *TMPRSS2* with the ETS gene family member *ERG*. The *PCA3* gene is highly overexpressed in prostate cancer (7).

[1]Radboud University Nijmegen Medical Centre, Department of Urology, Nijmegen, the Netherlands. [2]Noviogendix, Department of Research and Development, Nijmegen, the Netherlands. [3]ZGT Hospital, Department of Urology, Hengelo, the Netherlands. [4]AMC University Medical Centre, Department of Urology, Amsterdam, the Netherlands. [5]CWZ Hospital, Department of Urology, Nijmegen, the Netherlands. [6]St Elisabeth Hospital, Department of Urology, Tilburg, the Netherlands. [7]Scheper Hospital, Department of Urology, Emmen, the Netherlands. [8]Radboud University Nijmegen Medical Centre, Department of Pathology, Nijmegen, the Netherlands.

G.H.J.M. Leyten and D. Hessels contributed equally to this article.

Corresponding Author: Jack A. Schalken, Department of Experimental Urology, Radboud University Nijmegen Medical Centre, Internal Postal Code 267, P.O. Box 9101, Nijmegen 6500 HB, the Netherlands. Phone: 31-24-3614146; Fax: 31-24-3541222; E-mail: jack.schalken@radboudumc.nl

doi: 10.1158/1078-0432.CCR-14-3334

Translational Relevance

There is an urgent need for biomarkers that can be used to identify patients with significant prostate cancer. We present a stepwise selection of a three-gene panel (*HOXC6*, *TDRD1*, and *DLX1*) for the detection of prostate cancer and, in particular, biopsy Gleason score ≥7 prostate cancer in urinary sediments. The three genes are upregulated in prostate cancer tissue and have been associated with prostate cancer development, explaining the high specificity of the urine test for the disease. After testing this three-gene panel in urinary sediments of a clinical intention-to-treat cohort, it was shown that it represents a promising tool to identify patients with aggressive prostate cancer, also in those with low serum PSA values. Unfortunately, limited data are available on the value of urinary prostate cancer biomarkers in patients with serum PSA values < 3 ng/mL. The combination of the urinary three-gene panel with serum PSA bears great potential for the early diagnosis of patients with clinically significant prostate cancer.

The Progensa PCA3 test is an FDA-approved molecular diagnostic test for the detection of prostate cancer in urine (8). Although *PCA3* has diagnostic value to predict biopsy outcome, its value for distinguishing indolent from aggressive prostate cancer is limited (9–12). Gene fusions between androgen-regulated *TMPRSS2* and members of the ETS transcription factor family are prostate cancer–specific events. *TMPRSS2–ERG* gene fusions are present in about 50% of patients with prostate cancer (13, 14). Similar to *PCA3*, *TMPRSS2–ERG* gene fusions can be detected in urine (15, 16). Owing to the heterogeneity of the disease, a panel of biomarkers will improve the diagnosis of prostate cancer. Improved detection of prostate cancer in urine could be gained when *TMPRSS2–ERG* gene fusions were combined with *PCA3* (16–18). However, the value of this gene fusion for distinguishing indolent from aggressive prostate cancer is controversial (19–21).

Therefore, the need for novel prostate cancer–specific biomarkers, which can be used as an adjunct to sPSA, persists to enable the more accurate detection of prostate cancer and improve the prediction of tumor aggressiveness. In the current report, gene expression profiling was used for the identification of these biomarkers followed by stepwise biomarker selection and testing of a 3-gene panel for the diagnosis of biopsy Gleason score ≥7 prostate cancer in urinary sediments.

Materials and Methods

Patient sample collection and preparation

Retrospective tissue collection (biomarker discovery). Human prostate tissue specimens were collected from patients who underwent radical prostatectomy or transurethral resection of the prostate (TURP) according to an approved IRB protocol at the Radboud University Nijmegen Medical Centre and Canisius Wilhelmina Hospital Nijmegen (Nijmegen, The Netherlands). Normal prostate tissue was obtained from cancer-free regions in radical prostatectomy specimens. BPH tissue was obtained by either TURP or an open adenectomy. Prostate cancer tissues of patients with castration-resistant prostate cancer (CRPC) were obtained by TURP from patients who had progressive disease under endocrine therapy. Prostate cancer metastases were obtained from positive lymph nodes after lymph node dissection (LND).

Gleason scores and tumor–node–metastasis (TNM) classification of the tumors were determined at the Department of Pathology of both hospitals. The specimens were snap-frozen in liquid nitrogen, processed by step sectioning, and at regular intervals. A hematoxylin and eosin staining was performed to determine the percentage of normal, BPH, and tumor cells in the tissue sections. Tumor- and tumor-free areas were microdissected and total RNA was extracted by using TRIzol reagent (Invitrogen) according to manufacturer's instructions. Total RNA was DNase treated and purified using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The integrity of the RNA was determined using the Agilent 2100 Bioanalyzer. Samples with RNA integrity number (RIN) ≥6 were included for microarray analysis.

In total, tissue specimens of 133 patients were collected: normal prostate (NP; $n = 12$), BPH ($n = 16$), low-grade prostate cancer (LG-PCa; $n = 33$), high-grade prostate cancer (HG-PCa; $n = 32$), CRPC ($n = 32$), and metastatic prostate cancer (PCa-M+; $n = 8$). LG-PCa was defined at Gleason score ≤ 6 and HG-PCa was defined as Gleason score ≥ 7.

Urine samples (clinical biomarker testing). As was described by Leijten and colleagues, first-catch urine samples were collected after digital rectal examination (DRE) from men who were scheduled for (initial or repeat) prostate biopsies, on the basis of elevated sPSA levels, a family history of prostate cancer or an abnormal DRE according to an approved IRB protocol at 6 urology clinics in the Netherlands (Radboud University Nijmegen Medical Centre, Nijmegen; Academic Medical Centre, Amsterdam; ZGT Hospital, Hengelo; Canisius Wilhelmina Hospital, Nijmegen; Scheper Hospital, Emmen; and St. Elisabeth Hospital, Tilburg; ref. 22). All the subjects involved in this study signed the IRB-approved consent form. Exclusion criteria were history of prostate cancer, medical therapy known to affect sPSA levels, prostate biopsies within 3 months before enrolment, or invasive treatment for BPH within 6 months before enrolment. The urine samples were processed according to procedures for whole urine as described by Groskopf and colleagues and urinary sediments as was described by Hessels and colleagues (8, 23). From the 443 urine samples previously described, 358 were selected for analysis in this study based on sufficient HPRT mRNA (>4,000 copies) content (22). Prostate biopsies (9–12 core needle biopsies) were performed and evaluated per hospital's standard procedure. In addition, one experienced genitourinary pathologist reviewed all biopsy Gleason scores independently, being blinded for the biomarker scores.

For each patient, clinicopathologic data were collected, including age, sPSA, DRE, and transrectal ultrasound (TRUS) results, prostate volume, biopsy results (current and history), radiologic results, clinical TNM stage (if diagnosed with prostate cancer), and radical prostatectomy results (if applicable). These data and the assay results were entered in a secured preset web-based database with audit trail (in compliance with the International Conference on Harmonization-Good Clinical Practice guidelines). Assay results were not provided to the clinical sites for patient care and the technicians who performed the assays were blinded for patient characteristics.

Gene expression profiling

Retrospectively collected tissue samples ($n = 99$) were used for gene expression profiling on the GeneChip Human Exon 1.0 Sense Target (ST) arrays (Affymetrix) according to the manufacturer's protocol. One microgram of RNeasy purified total RNA was used to generate amplified and biotinylated sense-strand DNA targets from the entire expressed genome. According to the protocol, the majority of ribosomal RNA was removed using the RiboMinus Human/Mouse Transcriptome Isolation Kit (Invitrogen). The generated amplified sense-strand cDNA targets were fragmented by incubation with a mixture of UDG (uracil DNA glycosylase) and APE1 (apurinic/apyrimidinic endonuclease 1) restriction endonucleases and end-labeled via terminal transferase reaction incorporating a biotinylated dideoxynucleotide. Of the fragmented, biotinylated cDNA, 5.5 μg was added to a hybridization mixture, loaded on a Genechip Human Exon 1.0 ST (Affymetrix) and hybridized for 16 hours at 45°C and 60 rpm. Following hybridization, the array was washed in a GeneChip Fluidics station FS450 (Affymetrix) and stained according to the Affymetrix protocol. The array was scanned at 532 nm using a GeneChip Scanner 3000 7G (Affymetrix), generating CEL files for each array.

Gene-level and exon-level expression values were derived from the CEL file using the model-based Robust Multiarray Average (RMA) algorithm as implemented in Partek software (Partek Genomics Suite 6.6). RMA is a normalization approach that includes background correction, normalization, and data summarization and was performed on the core meta-probesets and the extended meta-probesets.

ANOVA was performed for the identification of upregulated genes in prostate cancer (LG-PCa + HG-PCa + CRPC + PCa-M+) compared with nonmalignant prostate tissue (NP + BPH); HG-PCa compared with LG-PCa; PCa-M+ compared with prostate cancer (LG-PCa + HG-PCa); and CRPC compared with prostate cancer (LG-PCa + HG-PCa). The fold changes (FC) of gene expression and *P* values in these different classes of samples were calculated. For each comparison, a list of the 100 most upregulated genes was created. For all the genes in these 4 lists, scatterplots were created. In these plots, the individual prostate samples were ordered in the categories normal, BPH, LG-PCa, HG-PCa, CRPC, and PCa-M+. It has been shown that the fold change–based selection of genes leads to more reproducible results (24). Therefore, the selection of the 39 biomarkers for further testing with qPCR on TaqMan Low Density Arrays (TLDA) was primarily based on fold changes followed by nonstringent *P* values and related scatterplot patterns.

TaqMan low-density arrays

Further selection of the biomarkers was done using Applied Biosystems TLDAs on a case mix of 73 samples used in the microarray experiments and 34 new prostate tissue specimens. Furthermore, 16 urinary sediments obtained from 9 men with prostate cancer–positive biopsies and 7 men without cancer in their biopsies were used in TLDA analysis. Two micrograms of RNA was used in cDNA synthesis using SuperScript II (Invitrogen) according the manufacturer's instructions. One twentieth of the cDNA was mixed with TaqMan Universal PCR Mastermix (Applied Biosystems) and was loaded on the TLDA card. The card was run on an Applied Biosystems 7900 HT thermal cycler with 384-well TaqMan Low-Density Array default thermal-cycling conditions. $\Delta\Delta C_t$ analysis was performed using SDS RQ study software (Applied Biosystems). In the TLDA analysis, *PCA3*, *CRISP*, *FOLH1*, *ERG*, and *AMACR* were used as controls and *GAPDH* and *HPRT* were used as reference genes. The genes *ONECUT2*, *HOXC4*, and *HOXC6* were not tested on TLDA, but for these genes, qPCR assays were developed. *ONECUT2* was chosen from the extended meta-probe set analysis and no TLDA assay was available. The array data for *HOXC6* was obtained through the summarization of the result of 15 probe sets of which only 7 were specific for *HOXC6*, 4 were specific for *HOXC4*, 3 were specific for *HOXC5*, and 1 probe set was specific for the 5′ noncoding exon shared by the 3 genes. Because *HOXC4* and *HOXC6* were both upregulated in prostate cancer and no suitable TLDA assay was available at the time, qPCRs were developed for both genes. On the basis of the obtained results, the most promising biomarkers were selected for qPCR analysis on 358 urinary sediments.

qPCR

Fluorescence-based real-time PCR assays were designed and manufactured by TIB molbiol Berlin for *HPRT*, *ONECUT2*, *HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *NKAIN1*, *MS4A8B*, *PPFIA2*. PCR products were cloned in either the pCR-Blunt cloning vector (Invitrogen) or the pCR2.1-TOPO cloning vector (Invitrogen). Calibration curves with a wide linear dynamic range (10–1,000,000 copies) were generated using serial dilutions of the plasmids. The calibration curve was used to determine the amplification efficiency of each primer pair. For each primer combination, the efficiency ranged from 1.85 to 2.10. The cDNA of several prostate cancer tissue specimens was pooled and used as a reference.

For each cDNA sample, 20 μL qPCR master mix was prepared by combining 2 μL of cDNA, optimized amounts of template-specific forward and reverse primer, 2 pmol of hydrolysis probe, and 1× LightCycler 480 Probe Master mix (Roche). The following amplification conditions were used: 95°C for 10 minutes, then 50 cycles at 95°C for 10 seconds, 60°C for 30 seconds, and a final cooling step at 40°C for 55 seconds (LightCycler LC 480, Roche). The crossing point (Cp) values were determined using the Lightcycler 480 SW 1.5 software (Roche). The Cp values of the samples were converted to copy numbers by interpolation in the generated calibration curve. The assay performance of the real-time PCR experiments was evaluated during in-study validation. The reference control samples had an inter- and intra-assay variation <30%.

Clinical urinary sediments study

Total RNA was extracted from the urinary sediments using a modified TriPure isolation reagent protocol (Roche). After the chloroform-induced phase separation, GlycoBlue (Ambion) was added to the aqueous phase to precipitate the RNA using isopropanol (Merck). The RNA was DNase-treated before the amplification protocol using DNase I enzyme (Invitrogen). Ethanol (Merck)/sodium acetate precipitation (Ambion) was used to purify the RNA. Using the Whole Transcriptome (WT) Expression Kit (Ambion), amplified sense-strand cDNA was generated.

The gene expression of *HPRT*, *ONECUT2*, *HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *NKAIN1*, *MS4A8B*, and *PPFIA2* was measured in the cDNA of urinary sediments using the developed qPCR assays and protocol described above.

The Progensa PCA3 test was performed on whole urine samples collected in urine specimen transport tubes (Progensa PCA3, Hologic; ref. 8). The PCA3 score was calculated as [*PCA3* mRNA]/[*KLK3* mRNA] × 1,000.

Statistical analyses

Statistical analyses were performed with SPSS version 20.0. Two-sided $P \leq 0.05$ was considered to indicate statistical significance. The nonparametric Mann–Whitney tests (for continuous variables) were used to test whether biomarker levels were significantly correlated with prostate cancer and Gleason score. sPSA, Progensa PCA3 score, and the novel biomarkers were assessed as continuous biomarkers. Backward logistic regression analysis was used to test whether the novel biomarkers had independently additional predictive value to sPSA and *PCA3* for diagnosis prostate cancer and Gleason score $\geq$7 prostate cancer. The area under ROC curve (AUC) and corresponding 95% confidence intervals (CI) of the final model were determined. Bootstrapping analysis was used for internal validation of the model. Bootstrap samples were drawn with replacement and with the same size as the original sample. Regression models were created in each bootstrap sample and were tested on the original sample. This procedure was tested 100 times to obtain stable estimates of the optimism of the model.

Results

Biomarker discovery

In Fig. 1, the stepwise approach of the biomarker discovery is schematically illustrated. To identify tumor-specific candidate biomarkers, gene expression was examined in a total of 99 normal prostate and prostate cancer tissue samples using the Human Exon 1.0 ST Array. ANOVA was performed for the identification of upregulated genes in 4 separate comparisons: prostate cancer (LG-PCa + HG-PCa + CRPC + PCa-M+) compared with nonmalignant prostate tissue (NP + BPH); HG-PCa compared with LG-PCa; PCa-M+ compared with prostate cancer (LG-PCa + HG-PCa); and CRPC compared with prostate cancer (LG-PCa + HG-PCa). The FC of gene expression and *P* values in these different classes of samples were calculated. For each comparison, a list of the 100 most upregulated genes was created. For all the genes in these 4 lists, scatterplots were created, which formed the base for the selection process (data not shown).

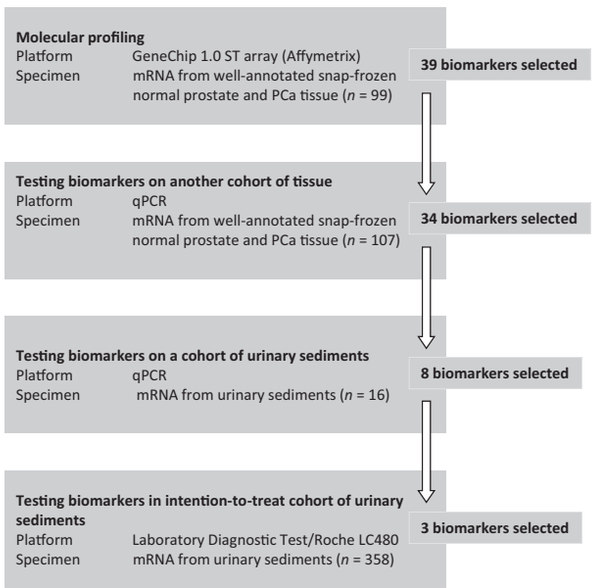

Figure 1.
Study design for the stepwise selection of a panel of biomarkers for the detection of prostate cancer (PCa) and in particular biopsy Gleason score $\geq$ 7 prostate cancer in urinary sediments. To address this unmet need, gene expression profiling was performed on snap-frozen microdissected tissue specimen. The 39 most promising biomarkers were tested on another set of tissue specimen using qPCR (TLDA). The 34 biomarkers that could discriminate prostate cancer from normal prostate were tested on a small set of 16 urinary sediments. The 8 biomarkers that could best detect prostate cancer in urinary sediments were selected and tested in 358 urinary sediments of an intention-to-treat cohort. This resulted in a urinary 3-gene panel to predict Gleason score $\geq$ 7 prostate cancer in biopsies.

Bioinformatics combining FC followed by nonstringent *P* values and related scatterplot patterns led to the identification of 39 biomarkers (Table 1). These 39 candidate biomarkers were tested using qPCR (TLDA) on 107 tissue RNA samples and 16 urinary sediments RNA samples. The resulting expression patterns of these biomarkers in the new set of tissue specimens confirmed the gene expression data, supporting the robustness of the biomarker discovery method used.

In the next step, biomarkers were selected on the basis of their expression profiles in urinary sediments. The first selection was based on the most differentially expressed genes in urinary sediments between cancer and noncancer patients. The second selection was based on the prostate-specific expression of the biomarkers in the urinary sediments. For this purpose, the 16 urinary sediments were preselected, based on either high expression levels of *KLK3* mRNA (prostate cells) and low expression of *HPRT* (background) or high expression levels of *HPRT* and low expression levels of *KLK3*. If the biomarker was high in $KLK3^{-}/HPRT^{+}$ and low in $KLK3^{+}/HPRT^{-}$ samples, the biomarker was more likely associated with the background expression in the urinary sediments and not correlated to the prostate cells. Sixteen biomarkers (Table 1, marked yes in the columns "Difference Prostate Cancer/NP" as well as "not HPRT correlated") met this requirement being *HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *ONECUT2*, *NKAIN1*, *MS4A8B*, *PPFIA2*, *PTPRT*, *GLYATL1*, *C19orf48*, *ALDH3B2*, *UGT2B15*, *COMP*, *CGREF1*, and *ACSM1*. Because *PPFIA2* and *PTPRT* showed similar results, *PPFIA2* was selected on the basis of the highest difference in Cp values in urinary sediments between prostate cancer and no cancer and the highest signal. The mRNA levels of *UGT2B15* and *COMP* were very low (Cp $\geq$ 40) in urinary sediments [Table 1, marked in column "Cp(NP)"] and were not selected for further analysis. The other 13 biomarkers were candidates to be tested in a larger number of urinary sediments obtained from an intend-to-treat patient cohort.

Clinical testing of biomarkers in urinary sediments

Of the 358 men with evaluable urinary sediments, 157 (44%) were diagnosed with prostate cancer and 93 (26%) were diagnosed with Gleason score $\geq$ 7 prostate cancer. Patient characteristics are shown in Table 2. Because only 93 patients with biopsy Gleason score $\geq$ 7 prostate cancer were included in this study, only 8 novel biomarkers could be added to sPSA in the logistic regression analysis to meet the minimal requirement of 10 events per variable. For this purpose, the 8 of the 13 biomarkers were selected that could distinguish best between prostate cancer and no prostate cancer in urinary sediments based on average Cp values (Table 1). The expression of these 8 potential biomarkers (*HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *ONECUT2*, *NKAIN1*, *MS4A8B*, and *PPFIA2*) was determined in urinary sediments using a real-time qPCR approach on the larger cohort of urinary sediments.

Q5 **Table 1.** Biomarker selection from microarray and qPCR (TLDA) data

Q6

		Tissue specimens								Urinary sediments				
		FC LG + HG + CRPC + META vs. NP + BPH		FC CRPC vs. LG + HG		FC META vs. LG + HG		FC HG vs. LG						
	Gene symbol	Array	qPCR	Array	qPCR	Array	qPCR	Array	qPCR	Cp (NP)	Cp (PCa)	Cp (NP) − Cp (PCa)	Difference PCa/NP	Not HPRT1 correlated
c	*AMACR*	7.4	16.3		−2.3		−2.9		1.2	28.0	28.1	0.0	No	No
c	*CRISP3*	14.3	45.5		−1.2		−1.5		−4.0	27.1	26.8	0.3	No	Yes
c	*ERG*	4.6	10.1		−3.3		−1.6		−1.9	39.8	33.0	6.8	Yes	Yes
c	*FOLH1*	5.2	3.4		−1.2		1.0		1.8	36.9	32.2	4.8	Yes	Yes
c	*PCA3*	12.6	48.3		−20.3		−6.2		1.1	30.2	26.1	4.1	Yes	Yes
1	*TDRD1*	10.2	30.2		−1.5		−2.7		−2.1	33.6	29.5	4.1	Yes	Yes
2	*RRM2*	9.8	18.6		2.9		6.9		1.7	24.8	27.0	−2.2	No	No
3	*ONECUT2*	8.0	7.5		2.2		2.9		−1.1	32.1	29.2	2.9	Yes	Yes
4	*ACSM1*	6.2	3.7		−4.2		−3.7		2.6	35.6	34.8	0.8	Yes	Yes
5	*TMEM45B*	5.8	5.4		1.0		1.1		1.1	28.2	27.6	0.6	Yes	No
6	*HOXC6*	5.7	10.4		−1.2		1.3		−1.2	35.1	32.1	3.0	Yes	Yes
7	*GLYATL1*	4.3	3.5		−2.6		−6.2		−1.8	32.1	30.5	1.5	Yes	Yes
8	*FASN*	4.1	2.1		−1.4		−1.4		−1.8	24.8	24.9	−0.1	No	Yes
9	*C19orf48*	4.1	2.5		1.1		1.1		1.0	27.7	27.5	0.2	Yes	Yes
10	*MS4A8B*	4.1	10.8		−2.5		−1.9		3.3	36.1	33.2	2.9	Yes	Yes
11	*NETO2*	4.0	2.8		1.2		−1.4		−1.4	27.4	29.3	−1.9	No	No
12	*TLCD1*	4.0	1.7		1.0		1.0		−1.2	29.8	29.6	0.2	No	Yes
13	*TOP2A*	4.0	3.3		2.0		2.4		1.3	35.4	37.4	−2.0	No	No
14	*TPX2*	3.9	4.2		1.7		2.7		−1.2	27.4	28.7	−1.3	No	No
15	*CGREF1*	3.9	3.8		−1.7		1.0		1.4	34.1	32.0	2.2	Yes	Yes
16	*PTPRT*	3.8	9.9		1.3		1.3		1.8	38.2	33.4	4.8	Yes	Yes
17	*PPFIA2*	3.6	5.0		2.6		1.5		2.9	37.3	31.9	5.4	Yes	Yes
18	*MKI67*	3.4	3.5		1.7		3.8		−1.1	27.1	28.6	−1.5	No	Yes
19	*FAM111B*	3.4	1.9		1.3		1.4		1.2	29.6	30.8	−1.3	No	Yes
21	*CDC20*	3.2	3.6		2.6		3.9		1.0	30.3	31.7	−1.4	No	No
21	*NKAIN1*	3.2	8.9		−2.9		−2.4		−2.4	38.5	33.6	4.9	Yes	Yes
22	*DLX1*	3.1	13.3		−2.0		−1.4		−2.2	39.4	35.0	4.4	Yes	Yes
23	*ALDH3B2*	2.8	5.7		−2.6		−3.3		1.8	33.0	31.4	1.6	Yes	Yes
24	*CKS2*	2.8	1.2		1.4		1.2		1.0	27.8	28.7	−0.9	No	No
25	*CDK1*	2.8	3.0		2.1		1.4		1.4	31.5	32.9	−1.4	No	No
26	*HOXC4*	2.3	6.9		0.8		1.9		0.6	29.8	27.0	2.8	Yes	Yes
27	*UGT2B15*		30.3	4.1	24.9		−2.9		3.2	40.0	38.2	1.8	Yes	Yes
28	*KIF4A*		9.3	3.1	2.9		3.4		1.6	33.6	35.8	−2.2	No	No
29	*PTTG1*		2.3	2.8	1.7		2.8		−1.2	31.1	32.6	−1.5	No	No
30	*ANLN*		2.8	2.8	2.1		8.9		1.2	28.4	29.4	−1.0	No	No
31	*KIF20A*		4.4	2.8	1.8		2.8		1.0	28.3	29.2	−0.9	No	No
32	*BUB1*		2.1	2.6	2.1		3.4		1.3	28.5	29.7	−1.2	No	No
33	*CYP4F8*		−1.1		−1.1	4.1	4.9		−5.1					
34	*PKP1*		−1.6		−1.2	3.5	2.4		−1.6					
35	*FAM110B*		1.7		4.5	2.4	2.4		−1.3	32.8	32.8	0.0	No	Yes
36	*SFRP2*		−1.6		1.0		−7.4	1.4	1.2					
37	*COMP*		4.9		2.0		−3.1	1.7	1.9	42.0	39.0	3.0	Yes	Yes
38	*ABI3BP*		−1.9		1.9		−3.0	1.7	1.0	31.5	31.3	0.1	No	Yes
39	*CDH2*		−1.6		−1.4		−1.3	1.5	6.3					

Abbreviations: c, control genes, META, prostate cancer metastases.

First, univariate analyses were performed to understand the potential associations between the biomarkers selected and prostate cancer and in particular Gleason score ≥ 7 in urinary sediments. All biomarkers could discriminate prostate cancer from no prostate cancer in urinary sediments with a $P < 0.05$ in univariate analysis (Table 3). Furthermore, all biomarkers could discriminate Gleason score ≥ 7 prostate cancer from Gleason score ≤ 6 prostate cancer and Gleason score ≥ 7 prostate cancer from REST (no prostate cancer and Gleason score ≤ 6 prostate cancer) in urinary sediments with a $P < 0.05$ in univariate analysis (Table 4). Therefore, all biomarkers were included in the multivariable logistic regression model.

Backwards logistic regression analysis was performed to test whether the novel biomarkers (*HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *ONECUT2*, *NKAIN1*, *MS4A8B*, and *PPFIA2*) had additional predictive value to sPSA for the diagnosis of biopsy Gleason score ≥ 7 prostate cancer in urinary sediments. Before the analysis, the raw values were visually inspected by histograms to assess normality. Biomarker values were log-transformed to improve normality. In urinary sediments, the markers that had independent additional predictive value to sPSA (OR, 2.84; 95% CI, 1.78–4.52; $P < 0.001$) for the detection of Gleason score ≥ 7 prostate cancer were *HOXC6* (OR, 1.40; 95% CI, 1.14–1.72; $P = 0.001$), *TDRD1* (OR, 1.13; 95% CI, 1.01–1.28; $P = 0.038$), and *DLX1* (OR, 1.15; 95% CI, 1.01–1.30; $P = 0.030$). Using logistic regression analysis, a predictor model was obtained for the diagnosis of a biopsy Gleason score ≥ 7 prostate cancer diagnosis in urinary sediments using sPSA, *HOXC6*, *TDRD1*, and *DLX1*. Hosmer–Lemeshow's

Table 2. Patient characteristics of the clinical intention-to-treat cohort

Descriptives	Cohort ($n = 358$) Median (range)/n (%)
Age, y	65 (44–86)
Prostate cancer in family	66 (18%)
No previous biopsies	280 (78%)
Abnormal DRE	103 (61%)
TRUS prostate volume, cc	48 (15–200)
Prostate cancer upon biopsy	157 (43.9%)
Gleason score ≥ 7	93 (26%)

goodness-of-fit test was used to assess calibration of the model. The significance of this test was 0.490 indicating that the model fits the data well.

Defining a predictor model for the diagnosis of biopsy Gleason score ≥ 7 prostate cancer in urinary sediments

A bootstrap resampling tool was used to test the robustness of these biomarkers for the diagnosis of biopsy Gleason score ≥ 7 prostate cancer in urinary sediments as an adjunct to sPSA. *HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *ONECUT2*, *NKAIN1*, *MS4A8B*, and *PPFIA2* were included in the analysis. After 100 bootstrap replications, *HOXC6*, *TDRD1*, and *DLX1* were predominantly present in the model (appearance of ≥ 76 times) compared with the other markers (appearance of <32 times). This yielded the following predictor model for the diagnosis of a biopsy Gleason score ≥ 7 prostate cancer diagnosis in urinary sediments in which the units of measurements for *HOXC6*, *DLX1*, and *TDRD1* are mRNA copy numbers and for sPSA is ng/mL: Probability $= 1/[1 + \mathrm{EXP}^{-(-5.007 + 0.069 \times \mathrm{sPSA} + 0.345 \times \mathrm{LnHOXC6} + 0.136 \times \mathrm{lnDLX1} + 0.137 \times \mathrm{lnTDRD1})}]$. The average correlation of the bootstrap models with the original model obtained by logistic regression analysis was 0.960.

Evaluation of the predictor model for the 3-gene panel in adjunct to sPSA

Consequently, sPSA, *HOXC6*, *TDRD1*, and *DLX1* were selected for evaluating the model discrimination using the receiver operating characteristic (ROC) AUC. Using ROC analysis, the predictive accuracy for the diagnosis of Gleason score ≥ 7 prostate cancer was higher for the combination of *HOXC6*, *DLX1*, and *TDRD1* (AUC, 0.77; 95% CI, 0.71–0.83) compared with Progensa PCA3 (AUC, 0.68; 95% CI, 0.62–0.75; Fig. 2A). The predictive accuracy of sPSA (AUC, 0.72; 95% CI, 0.65–0.78) could be improved when *HOXC6*, *TDRD1*, and *DLX1* were combined with sPSA (AUC, 0.81; 95% CI, 0.75–0.86 (Fig. 2B). Bootstrap analysis was used for internal validation of the AUC of the model. The mean AUC of 100 bootstrap samples was 0.81 and the mean AUC of 100 tests on the original sample was 0.80. This indicated an optimism of 0.01. Therefore, the internally validated AUC was therefore estimated as 0.80.

On the basis of the predictor model for *HOXC6*, *TDRD1*, *DLX1*, and sPSA, a probability of 0.34 (top left point in the ROC curve) was the cutoff point that maximized the sensitivity (68.5%) and specificity (82.3%) for the detection of biopsy Gleason score ≥ 7 in urinary sediments. For sPSA alone, the cutoff corresponded to a concentration of 9.5 ng/mL maximizing the sensitivity (58.1%) and specificity (74%) for the detection of biopsy Gleason score ≥ 7.

Predictive accuracy of the 3-genel panel for the diagnosis of Gleason score ≥ 7 prostate cancer in sPSA cohorts

In Fig. 3, the AUCs are indicated which measures the predictive accuracy for the diagnosis of Gleason score ≥ 7 prostate cancer of the combination of *HOXC6*, *TDRD1*, and *DLX1* and sPSA alone. At different sPSA cutoff values, the AUC of the combination *HOXC6*, *TDRD1*, and *DLX1* was higher than sPSA and ranged from 0.72 to 0.76. The AUC of sPSA for the detection of Gleason score ≥ 7 prostate cancer was more variable ranging from 0.57 to 0.72. The AUC of sPSA was highest when men with sPSA values > 10 ng/mL were included. On the basis of these results, the predictive accuracy of the model is hardly affected by sPSA for the detection of biopsy Gleason score ≥ 7 in urinary sediments. At lower sPSA levels, the gene panel outperforms sPSA for the detection of biopsy Gleason score ≥ 7 prostate cancer. These results indicate that the gene panel might be useful in identifying aggressive cancers at lower PSA ranges. However, the numbers at the lower PSA ranges are small and the data need to be confirmed.

Table 3. Biomarker characteristics in urinary sediments of the clinical intent-to-treat population

	Prostate cancer		
	No ($n = 201$) Median (Q1–Q3)	Yes ($n = 157$) Median (Q1–Q3)	P
Serum PSA, ng/mL	6.8 (5.1–9.4)	9.2 (6.1–13.7)	<0.001[a]
PCA3 score	24 (12–57)	60 (31–107)	<0.001[a]
HOXC4	5,260 (1,560–9,930)	12,600 (4,140–24,100)	<0.001[a]
HOXC6	321 (84–838)	962 (390–2,760)	<0.001[a]
DLX1	1 (1–1)	1 (1–231)	<0.001[a]
TDRD1	124 (1–383)	367 (60–1,560)	<0.001[a]
ONECUT2	776 (259–2,020)	1,280 (570–2,860)	<0.001[a]
NKAIN1	162 (37–440)	291 (94–891)	<0.001[a]
MS4A8B	168 (1–592)	612 (126–2,100)	<0.001[a]
PPFIA2	167 (1–684)	534 (111–1,240)	<0.001[a]

[a]Mann–Whitney test.

Discussion

In this study, a pragmatic approach was used to identify candidate biomarkers for an important clinical unmet need, that is, biomarkers that can be used to predict the presence of Gleason score ≥ 7 prostate cancer in the biopsy, using a noninvasive substrate, that is, the urine. It is well known that quality of the clinical samples is important to obtain reliable and reproducible data by molecular profiling of clinical specimens. Our many years of experience with molecular profiling has resulted in optimized and reproducible protocols for fresh-frozen tissue specimens collection, microdissection of prostate cancer, RNA isolation, and quality assessment and profiling of RNA from fresh-frozen tissue specimens. With the focus on high-quality RNA, the chance of introducing technical bias in the molecular profiling was minimized.

Bioinformatics combining FC followed by nonstringent P values and related-scatterplot patterns led to the identification of 39 prostate cancer–associated biomarkers that were upregulated in 4 predefined groups of samples. We did not restrict ourselves to the averaged data of the several groups (ranging from normal prostate, BPH, LG-PCa, HG-PCa, CRPC, and PCa-M+) but also examined the individual sample data (scatterplots). By doing so, subgroups, trends and patterns in expression levels of the biomarkers could be identified. For the selection of biomarkers, heterogeneity of biomarker expression levels in the groups

Table 4. Biomarker characteristics in urinary sediments of the clinical intent-to-treat population

	Gleason score			Gleason score			
	≤6 (*n* = 64) Median (Q1–Q3)	≥7 (*n* = 93) Median (Q1–Q3)	*P*	REST (*n* = 265) Median (Q1–Q3)	≥7 (*n* = 93) Median (Q1–Q3)	*P*	AUC (95% CI)
Serum PSA, ng/mL	8 (5.3–10.1)	10.8 (7–20.1)	<0.001[a]	6.9 (5.2–9.5)	10.8 (7–20.1)	<0.001[a]	0.72 (0.65–0.78)
PCA3 score	55.5 (29–93)	61 (32–111)	0.278[a]	31 (15–65)	61 (32–111)	<0.001[a]	0.68 (0.62–0.75)
HOXC4	8,120 (3,600–21,525)	14,700 (4,820–30,700)	0.034[a]	5,940 (1,880–12,300)	14,700 (4,820–30,700)	<0.001[a]	0.69 (0.62–0.75)
HOXC6	633 (309–1,410)	1,550 (520–3,970)	<0.001[a]	392 (110–985)	1,550 (520–3,970)	<0.001[a]	0.76 (0.70–0.82)
DLX1	1 (1–22)	35 (1–758)	<0.001[a]	1 (1–1)	35 (1–758)	<0.001[a]	0.70 (0.63–0.77)
TDRD1	159 (1–481)	843 (146–8,065)	<0.001[a]	130 (1–416)	843 (146–8,065)	<0.001[a]	0.73 (0.67–0.80)
ONECUT2	1,020 (355–1,802)	1,790 (710–5,270)	<0.001[a]	804 (276–1,950)	1,790 (710–5,270)	<0.001[a]	0.69 (0.62–0.75)
NKAIN1	192 (77–438)	392 (128–1,900)	0.006[a]	163 (41–440)	392 (128–1,900)	<0.001[a]	0.66 (0.59–0.73)
MS4A8B	472 (69–1,070)	1,010 (196–3,250)	0.001[a]	204 (1–775)	1,010 (196–3,250)	<0.001[a]	0.70 (0.63–0.76)
PPFIA2	353 (56–722)	713 (147–1,790)	0.004	210 (1–704)	713 (147–1,790)	<0.001[a]	0.67 (0.61–0.74)

NOTE: REST, no prostate cancer + Gleason score ≤ 6.
[a]Mann–Whitney test.

normal prostate and BPH was not allowed, whereas some heterogeneity in the different prostate cancer groups was allowed and expected.

Well-known prostate cancer–associated genes such as *PCA3*, *AMACR*, *CRISP3*, *FOLH1*, and *ERG* were also in the top 15 list of extended or core genes based on fold changes of the microarray data, supporting the robustness of the biomarker discovery method used. The reliability of the data was also confirmed by qPCR (TLDA) on another set of tissue samples showing that the selected 39 genes indeed can discriminate between the 4 predefined groups. The 34 biomarkers that could discriminate prostate cancer from no prostate cancer were tested on a small set of urinary sediments after which 8 biomarkers (*HOXC4*, *HOXC6*, *DLX1*, *TDRD1*, *ONECUT2*, *NKAIN1*, *MS4A8B*, and *PPFIA2*) were selected on the basis of their overexpression in urinary sediments from patients with prostate cancer.

The hypothesis that a combination of these 8 biomarkers can improve the diagnosis of biopsy Gleason score ≥ 7 prostate cancer compared with sPSA or Progensa PCA3 was tested in an intention-to-treat cohort of 358 urinary sediments obtained from men who were scheduled for (initial or repeat) prostate biopsies. Of the 8 biomarkers, *HOXC6*, *TDRD1*, and *DLX1* had independent additional predictive value to sPSA for the detection of biopsy Gleason score ≥ 7 in urinary sediments. Interestingly, these genes have been associated with prostate cancer development.

Homeobox C6 (*HOXC6*), located at 12q13.3 in humans, regulates genes with both oncogenic and tumor suppressor activities as well as several genes important for prostate branching morphogenesis and metastasis to the bone microenvironment (25). In the prostate, there is convincing evidence for an oncogenic function of *HOXC6* (26). Its frequent overexpression

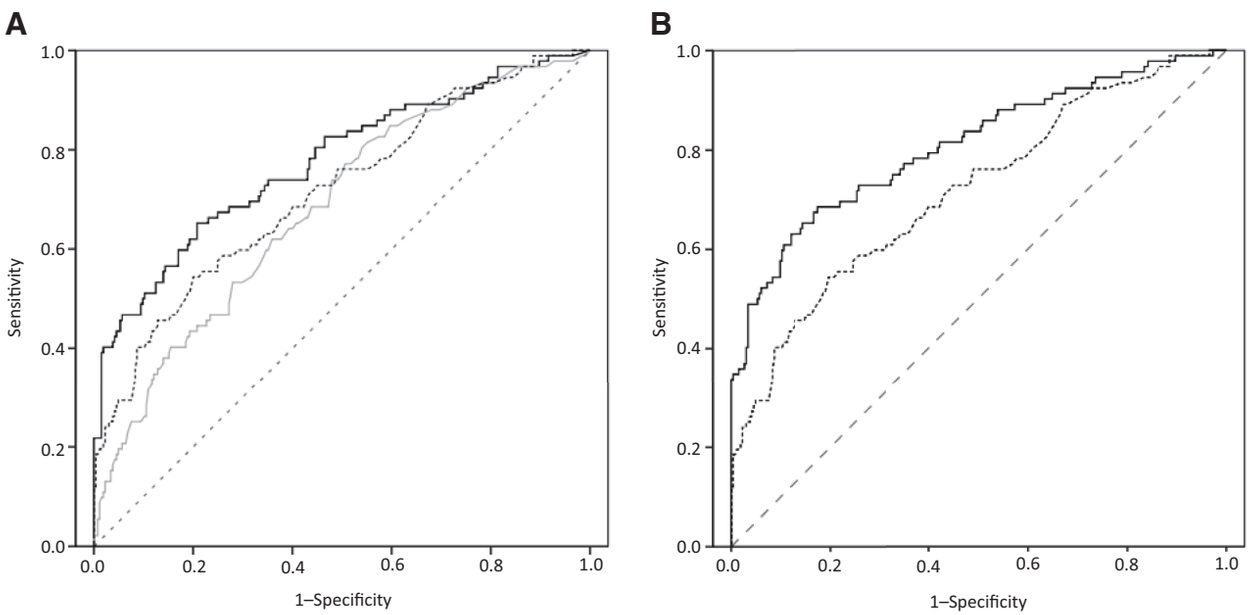

Figure 2.
ROC curves for the combination of urinary *HOXC6*, *TDRD1*, *DLX1* (black line; AUC, 0.77; 95% CI, 0.71–0.83), Progensa PCA3 (gray line; AUC, 0.68; 95% CI, 0.62–0.75), and sPSA (dotted line; AUC, 0.72; 95% CI, 0.65–0.78) to predict Gleason score ≥ 7 prostate cancer in biopsies (A). B, the ROC curve for the predictive model *HOXC6*, *TDRD1*, *DLX1*, sPSA (black line: AUC, 0.81; 95% CI, 0.75–0.86) is shown compared with sPSA alone (dotted line) for predicting Gleason score ≥ 7 prostate cancer in biopsies.

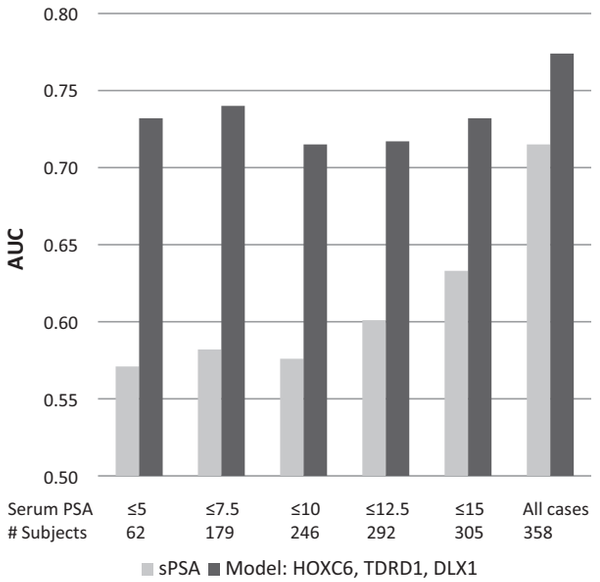

Figure 3.
The effect of sPSA on the predictive accuracy (AUC) of the urinary 3-gene panel *HOXC6*, *TDRD1*, and *DLX1* is shown.

in prostate cancer may predispose tumor cells to androgen independence by necessitating adaptation to diminished androgen signaling (27, 28). The degree of *HOXC6* overexpression is correlated with several clinical parameters of tumor progression, including Gleason scores (27–29). Tudor domain containing 1 (*TDRD1*), a male germline-specific gene located at 10q25.3, belongs to a large family of tudor domain containing proteins. Recently, *TDRD1* was identified as a direct *ERG* target gene that is strongly associated with *ERG* overexpression in primary prostate cancer. *ERG* activates *TDRD1* transcription by inducing loss of DNA methylation at the *TDRD1* promoter–associated CpG island (30). Like *ERG*, *TDRD1* is hardly expressed in normal adult prostatic tissue (31). Low methylation of *TDRD1* appeared to be significantly associated with a higher risk for biochemical recurrence in patients with high-risk prostate cancer (32). Distal-less Homeobox 1 (*DLX1*), located at 2q32, is involved in the acquisition of epithelial–neuroendocrine differentiation, a characteristic associated with aggressive cancer (33). It was shown that *DLX1* is upregulated in $CD26^{+}$ cancer cells isolated from Gleason 3 + 3 (G3) and Gleason 4 + 4 (G4) tumors compared with prostate luminal cells (34).

The principal goal of this study was the identification of genes in prostate cancer that could serve as early detection markers of prostate cancer and in particular biopsy Gleason score $\geq$ 7 prostate cancer. To avoid overdiagnosis and overtreatment of patients with prostate cancer due to the low specificity and unclear benefit of sPSA testing, a prostate cancer -specific biomarker test is required that uses noninvasive substrates such as urine. The first fully translated RNA-based molecular diagnostic test for the detection of prostate cancer in urine is the CE-marked Progensa PCA3 test. Several studies in the urine demonstrated that Progensa PCA3 was superior to sPSA in predicting prostate cancer on repeat prostate biopsy. Given the heterogeneous nature of prostate cancer, the use of a panel of biomarkers can further improve the diagnosis of this disease. The combined use of the Progensa PCA3 test and *TMPRSS2–ERG* could significantly improve the sensitivity for prostate cancer diagnosis as adjunct to sPSA testing (22). However, the value of this combination for predicting biopsy Gleason score $\geq$ 7 in urine is controversial (17, 18, 22, 35).

The fact that *HOXC6*, *TDRD1*, and *DLX1* are functionally implicated in prostate carcinogenesis strengthens the suggestion that the combination of these genes may comprise a promising method of detecting prostate cancer and in particular biopsy Gleason score $\geq$ 7 prostate cancer. It was shown that 3-gene panel was superior to Progensa PCA3 for the diagnosis of a biopsy Gleason score $\geq$ 7 prostate cancer. The predictor model that was obtained by logistic regression analysis for the combination of sPSA with *HOXC6*, *TDRD1*, and *DLX1* in the diagnosis of a biopsy Gleason score $\geq$ 7 prostate cancer diagnosis in urinary sediments correlated well with the predictor model obtained by bootstrap analysis. Furthermore, the predictive accuracy of sPSA alone (AUC, 0.72; 95% CI, 0.65–0.78) for the diagnosis of biopsy Gleason score $\geq$ 7 prostate cancer in urinary sediments could be improved when *HOXC6*, *TDRD1*, and *DLX1* were combined with sPSA (AUC, 0.81; 95% CI, 0.75–0.86).

It was shown that in a sPSA prescreened population, the urine-based Progensa PCA3 test could improve the identification of serious disease especially in the low PSA ranges (36). In this study, it is shown that the combination of urinary *HOXC6*, *TDRD1*, and *DLX1* is superior to Progensa PCA3 in the diagnosis of Gleason score $\geq$ 7 prostate cancer. Furthermore, using ROC analysis at different sPSA cut-offs for the predictive model for *HOXC6*, *TDRD1*, and *DLX1*, it was shown that the AUC for the detection of biopsy Gleason score $\geq$ 7 prostate cancer remained constant and was unaffected by sPSA concentrations. These preliminary data imply that urinary *HOXC6*, *TDRD1*, and *DLX1* can improve the detection of serious prostate cancer compared with sPSA and Progensa PCA3 and may be an important tool to prevent overtreatment. The value of this gene panel needs to be explored further. This study was done on urinary sediments as a substrate. Currently, we are developing a whole urine-based assay for HOXC6, *TDRD1*, and *DLX1*. The next step will be testing this gene panel in whole urine and validating the predictive model obtained in whole urine in an independent cohort of men.

In conclusion, in this study, a stepwise selection of a 3-gene panel for the detection of prostate cancer and in particular biopsy Gleason score $\geq$ 7 prostate cancer in urinary sediments is described. *HOXC6*, *TDRD1*, and *DLX1* were shown to have independent additional predictive value to sPSA for predicting biopsy Gleason score $\geq$ 7 prostate cancer. Our data suggest that *HOXC6*, *TDRD1*, and *DLX1* are useful for the sensitive and noninvasive detection of individuals at risk for Gleason score $\geq$ 7 prostate cancer also in those with low sPSA values. The 3-gene panel may comprise a more promising method of detecting prostate cancer and in particular biopsy Gleason score $\geq$ 7 prostate cancer in urinary sediments than Progensa PCA3 as an adjunct to sPSA testing.

Disclosure of Potential Conflicts of Interest

W.J.G Melchers has ownership interest (including patents) in NovioGendix Holding BV. J.A. Schalken reports receiving speakers bureau honoraria from Astellas and Sanofi, has ownership interest (including patents) in

NovioGendix Holding BV, and is a consultant/advisory board member for Hologic. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: G.H.J.M. Leyten, D. Hessels, F.P. Smit, S.A. Jannink, H. de Jong, W.J.G. Melchers, P.F.A. Mulders, I.M. van Oort, J.A. Schalken

Development of methodology: G.H.J.M. Leyten, F.P. Smit, S.A. Jannink, H. de Jong, W.J.G. Melchers, P.F.A. Mulders, I.M. van Oort, J.A. Schalken

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): G.H.J.M. Leyten, F.P. Smit, S.A. Jannink, H. de Jong, E.B. Cornel, T.M. de Reijke, H. Vergunst, P. Kil, B.C. Knipscheer, C.A.H. van de Kaa, P.F.A. Mulders, I.M. van Oort

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): G.H.J.M. Leyten, D. Hessels, F.P. Smit, S.A. Jannink, H. de Jong, W.J.G. Melchers, P. Kil, P.F.A. Mulders, I.M. van Oort, J.A. Schalken

Writing, review, and/or revision of the manuscript: G.H.J.M. Leyten, D. Hessels, F.P. Smit, S.A. Jannink, H. de Jong, W.J.G. Melchers, E.B. Cornel, T.M. de Reijke, H. Vergunst, P. Kil, C.A.H. van de Kaa, P.F.A. Mulders, I.M. van Oort, J.A. Schalken

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D. Hessels, F.P. Smit, S.A. Jannink, H. de Jong, E.B. Cornel

Study supervision: D. Hessels, I.M. van Oort, J.A. Schalken

Acknowledgments

The authors thank Annelies Hoebers (NovioGendix) for performing Progensa PCA3 testing and Pim Peelen (NovioGendix) for the clinical sample collection and processing. They also thank Ton de Haan (Radboud University Nijmegen Medical Centre) for statistical assistance. They also thank PCMM (03-O-203) for their contribution to the project.

Grant Support

This study was supported in part by grants from EFRO, Ultrasense MR, and NovioGendix.

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.

Received December 23, 2014; revised February 23, 2015; accepted March 8, 2015; published OnlineFirst xx xx, xxxx.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC CancerBase no.11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013 [cited 2013]. Available from: http://globocaniarcfr.
2. Roddam AW, Duffy MJ, Hamdy FC, Ward AM, Patnick J, Price CP, et al. Use of prostate-specific antigen (PSA) isoforms for the detection of prostate cancer in men with a PSA level of 2-10 ng/ml: systematic review and meta-analysis. Eur Urol 2005;48:386–99; discussion 398–9.
3. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 2009;360:1320–8.
4. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Prostate-cancer mortality at 11 years of follow-up. N Engl J Med 2012;366:981–90.
5. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. Lancet 2014;384:2027–35.
6. Heijnsdijk EA, der Kinderen A, Wever EM, Draisma G, Roobol MJ, de Koning HJ. Overdetection, overtreatment and costs in prostate-specific antigen screening for prostate cancer. Br J Cancer 2009;101:1833–8.
7. Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF, Schalken JA, et al. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Res 1999;59:5975–9.
8. Groskopf J, Aubin SM, Deras IL, Blase A, Bodrug S, Clark C, et al. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. Clin Chem 2006;52:1089–95.
9. Haese A, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PF, et al. Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. Eur Urol 2008;54:1081–8.
10. Marks LS, Fradet Y, Deras IL, Blase A, Mathis J, Aubin SM, et al. PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. Urology 2007;69:532–5.
11. Auprich M, Chun FK, Ward JF, Pummer K, Babaian R, Augustin H, et al. Critical assessment of preoperative urinary prostate cancer antigen 3 on the accuracy of prostate cancer staging. Eur Urol 2011;59:96–105.
12. Hessels D, van Gils MP, van Hooij O, Jannink SA, Witjes JA, Verhaegh GW, et al. Predictive value of PCA3 in urinary sediments in determining clinico-pathological characteristics of prostate cancer. Prostate 2010;70:10–6.
13. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.
14. Esgueva R, Perner S, C JL, Scheble V, Stephan C, Lein M, et al. Prevalence of TMPRSS2-ERG and SLC45A3-ERG gene fusions in a large prostatectomy cohort. Mod Pathol 2010;23:539–46.
15. Laxman B, Tomlins SA, Mehra R, Morris DS, Wang L, Helgeson BE, et al. Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. Neoplasia 2006;8:885–8.
16. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. Clin Cancer Res 2007;13:5103–8.
17. Dimitriadis E, Kalogeropoulos T, Velaeti S, Sotiriou S, Vassiliou E, Fasoulis L, et al. Study of genetic and epigenetic alterations in urine samples as diagnostic markers for prostate cancer. Anticancer Res 2013;33:191–7.
18. Robert G, Jannink S, Smit F, Aalders T, Hessels D, Cremers R, et al. Rational basis for the combination of PCA3 and TMPRSS2:ERG gene fusion for prostate cancer diagnosis. Prostate 2013;73:113–20.
19. Fleischmann A, Saramaki OR, Zlobec I, Rotzer D, Genitsch V, Seiler R, et al. Prevalence and prognostic significance of TMPRSS2-ERG gene fusion in lymph node positive prostate cancers. Prostate 2014;74:1647–54.
20. Mounir Z, Lin F, Lin VG, Korn JM, Yu Y, Valdez R, et al. TMPRSS2:ERG blocks neuroendocrine and luminal cell differentiation to maintain prostate cancer proliferation. Oncogene. 2014 Sep 29. [Epub ahead of print].
21. Hagen RM, Adamo P, Karamat S, Oxley J, Aning JJ, Gillatt D, et al. Quantitative analysis of ERG expression and its splice isoforms in formalin-fixed, paraffin-embedded prostate cancer samples: association with seminal vesicle invasion and biochemical recurrence. Am J Clin Pathol 2014;142:533–40.
22. Leyten GH, Hessels D, Jannink SA, Smit FP, de Jong H, Cornel EB, et al. Prospective multicentre evaluation of PCA3 and TMPRSS2-ERG gene fusions as diagnostic and prognostic urinary biomarkers for prostate cancer. Eur Urol 2014;65:534–42.
23. Hessels D, Klein Gunnewiek JM, van Oort I, Karthaus HF, van Leenders GJ, van Balken B, et al. DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. Eur Urol 2003;44:8–15; discussion 16.
24. Guo L, Lobenhofer EK, Wang C, Shippy R, Harris SC, Zhang L, et al. Rat toxicogenomic study reveals analytical consistency across microarray platforms. Nat Biotechnol 2006;24:1162–9.
25. Moreno CS. The Sex-determining region Y-box 4 and homeobox C6 transcriptional networks in prostate cancer progression: crosstalk with the Wnt, Notch, and PI3K pathways. Am J Pathol 2010;176:518–27.
26. Vinarskaja A, Yamanaka M, Ingenwerth M, Schulz WA. DNA methylation and the HOXC6 paradox in prostate cancer. Cancer 2011;3:3714–25.
27. McCabe CD, Spyropoulos DD, Martin D, Moreno CS. Genome-wide analysis of the homeobox C6 transcriptional network in prostate cancer. Cancer Res 2008;68:1988–96.

28. Miller GJ, Miller HL, van Bokhoven A, Lambert JR, Werahera PN, Schirripa O, et al. Aberrant HOXC expression accompanies the malignant phenotype in human prostate. Cancer Res 2003;63:5879–88.
29. Ramachandran S, Liu P, Young AN, Yin-Goen Q, Lim SD, Laycock N, et al. Loss of HOXC6 expression induces apoptosis in prostate cancer cells. Oncogene 2005;24:188–98.
30. Kacprzyk LA, Laible M, Andrasiuk T, Brase JC, Borno ST, Falth M, et al. ERG induces epigenetic activation of Tudor domain-containing protein 1 (TDRD1) in ERG rearrangement-positive prostate cancer. PLoS One 2013;8:e59976.
31. Boormans JL, Korsten H, Ziel-van der Made AJ, van Leenders GJ, de Vos CV, Jenster G, et al. Identification of TDRD1 as a direct target gene of ERG in primary prostate cancer. Int J Cancer 2013;133:335–45.
32. Litovkin K, Joniau S, Lerut E, Laenen A, Gevaert O, Spahn M, et al. Methylation of PITX2, HOXD3, RASSF1 and TDRD1 predicts biochemical recurrence in high-risk prostate cancer. J Cancer Res Clin Oncol 2014;140:1849–61.
33. Merlo GR, Zerega B, Paleari L, Trombino S, Mantero S, Levi G. Multiple functions of Dlx genes. Int J Dev Biol 2000;44:619–26.
34. Pascal LE, Vencio RZ, Page LS, Liebeskind ES, Shadle CP, Troisch P, et al. Gene expression relationship between prostate cancer cells of Gleason 3, 4 and normal epithelial cells as revealed by cell type-specific transcriptomes. BMC Cancer 2009;9:452.
35. Tomlins SA, Aubin SM, Siddiqui J, Lonigro RJ, Sefton-Miller L, Miick S, et al. Urine TMPRSS2:ERG fusion transcript stratifies prostate cancer risk in men with elevated serum PSA. Sci Transl Med 2011;3:94ra72.
36. Roobol MJ, Schroder FH, van Leeuwen P, Wolters T, van den Bergh RC, van Leenders GJ, et al. Performance of the prostate cancer antigen 3 (PCA3) gene and prostate-specific antigen in prescreened men: exploring the value of PCA3 for a first-line diagnostic test. Eur Urol 2010;58:475–81.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

Q1: Page: 1: AU: Per journal style, genes, alleles, loci, and oncogenes are italicized; proteins are roman. Please check throughout to see that the words are styled correctly. AACR journals have developed explicit instructions about reporting results from experiments involving the use of animal models as well as the use of approved gene and protein nomenclature at their first mention in the manuscript. Please review the instructions at http://www.aacrjournals.org/site/InstrAuthors/ifora.xhtml#genenomen to ensure that your article is in compliance. If your article is not in compliance, please make the appropriate changes in your proof.

Q2: Page: 1: Author: Please verify the drug names and their dosages used in the article.

Q3: Page: 1: Author: Note that AACR journals do not use such abbreviations as "BC" for "breast cancer," "CRC" for colorectal cancer, or "Pca" for "prostate cancer," or other ad hoc two-word abbreviations for diseases that incorporate a part of the body. However, these abbreviations are acceptable in tables or figures where space is an issue as long as they are defined in the legend or table notes.

Q4: Page: 4: Author: Please confirm quality/labeling of all images included within this article. Thank you.

Q5: Page: 5: Author: Please verify the layout of Tables 1–4 for correctness.

Q6: Page: 5: Author: Please specify to what the color and gray-shaded values refer to in Table 1.

Q7: Page: 6: Author: Tables 3A and 3B have been renumbered as Tables 3 and 4. Please verify.

Q8: Page: 9: AU:/PE: The conflict-of-interest disclosure statement that appears in the proof incorporates the information from forms completed and signed off on by each individual author. No factual changes can be made to disclosure information at the proof stage. However, typographical errors or misspelling of author names should be noted on the proof and will be corrected before publication. Please note if any such errors need to be corrected. Is the disclosure statement correct?

Q9: Page: 9: Author: The contribution(s) of each author are listed in the proof under the heading "Authors' Contributions." These contributions are derived from forms completed and signed off on by each individual author. As the corresponding author, you are permitted to make changes to your own contributions. However, because all authors submit their contributions individually, you are not permitted to make changes in the contributions listed for any other authors. If you feel strongly that an error is being made, then you may ask the author or authors in question to contact us about making the changes. Please note, however, that the manuscript would be held from further processing until this issue is resolved.

Q10: Page: 9: AU: Please provide the complete cited date for ref. 1.

Q11: Page: 9: AU: Note that refs. 20 and 22 have been updated according to PubMed. Please verify.

AU: Below is a summary of the name segmentation for the authors according to our records. The First Name and the Surname data will be provided to PubMed when the article is indexed for searching. Please check each name carefully and verify that the First Name and Surname are correct. If a name is not segmented correctly, please write the correct First Name and Surname on this page and return it with your proofs. If no changes are made to this list, we will assume that the names are segmented correctly, and the names will be indexed as is by PubMed and other indexing services.

First Name	**Surname**
Gisele H.J.M.	Leyten
Daphne	Hessels
Frank P.	Smit
Sander A.	Jannink
Hans	de Jong
Willem J.G.	Melchers
Erik B.	Cornel
Theo M.	de Reijke
Henk	Vergunst
Paul	Kil
Ben C.	Knipscheer
Christina A	Hulsbergen-van de Kaa
Peter F.A.	Mulders
Inge M.	van Oort
Jack A.	Schalken