A First-Generation Multiplex Biomarker Analysis of Urine for the Early Detection of Prostate Cancer

Bharathi Laxman,1,2 David S. Morris,3,1 Jianjun Yu,1,2,4 Javed Siddiqui,1,2 Jie Cao,1,2 Rohit Mehra,1,2,5 Robert J. Lonigro,1,6 Alex Tsodikov,1,7 John T. Wei,1,2,5 Scott A. Tomlins,1,2,8 and Arul M. Chinnaiyan1,2,9

1Michigan Center for Translational Pathology, 2Department of Pathology, 3Department of Urology, 4Bioinformatics Program, and 5Comprehensive Cancer Center, University of Michigan Medical School; and 6Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, Michigan

Abstract

Although prostate-specific-antigen (PSA) serum level is currently the standard of care for prostate cancer screening in the United States, it lacks ideal specificity and additional biomarkers are needed to supplement or potentially replace serum PSA testing. Emerging evidence suggests that monitoring the noncoding RNA transcript PCA3 in urine may be useful in detecting prostate cancer in patients with elevated PSA levels. Here, we show that a multiplex panel of urine transcripts outperforms PCA3 transcript alone for the detection of prostate cancer. We measured the expression of seven putative prostate cancer biomarkers, including PCA3, in sediments of urine using quantitative PCR on a cohort of 234 patients presenting for biopsy or radical prostatectomy. By univariate analysis, we found that increased GOLPH2, SPINK1, and PCA3 transcript expression and TMPRSS2:ERG fusion status were significant predictors of prostate cancer. Multivariate regression analysis showed that a multiplexed model, including these biomarkers, outperformed serum PSA or PCA3 alone in detecting prostate cancer. The area under the receiver-operating characteristic curve was 0.758 for the multiplexed model versus 0.662 for PCA3 alone (P = 0.003). The sensitivity and specificity for the multiplexed model were 65.9% and 76.0%, respectively, and the positive and negative predictive values were 79.8% and 60.8%, respectively. Taken together, these results provide the framework for the development of highly optimized, multiplex urine biomarker tests for more accurate detection of prostate cancer. [Cancer Res 2008;68(3):645–9]

Introduction

Serum prostate-specific antigen (PSA) has been used extensively to screen for prostate cancer in the United States based on early studies showing that PSA levels >4 ng/mL have predictive value for detecting prostate cancer (1, 2). Although PSA testing has led to a dramatic increase in prostate cancer detection (3), PSA has substantial drawbacks. For example, PSA is often elevated in benign conditions, such as benign prostatic hyperplasia and prostatitis, likely accounting for the poor specificity of the PSA test, which has been reported to be only 20% at a sensitivity of 80% (4). Further, the Prostate Cancer Prevention Trial showed that even in patients with PSA levels <4 ng/mL, >15% had biopsy-detectable prostate cancer (5). Together, this supports the identification and characterization of prostate cancer biomarkers that could supplement PSA.

Numerous promising prostate cancer biomarkers have been identified, including genes specific for prostate cancer, such as AMACR (6) and PCA3 (7), and recurrent gene fusions involving TMPRSS2 and ETS family members (such as TMPRSS2:ERG; ref. 8). As prostate cells can be detected in the urine of men with prostate cancer, urine-based diagnostic tests have the advantage of being noninvasive. Although urine-based testing for PCA3 expression has already been documented in large screening programs (9), the feasibility of testing based on other markers has not been rigorously evaluated. Importantly, single marker tests, such as those based on PCA3, ignore the heterogeneity of cancer development and may only capture a proportion of cancer cases. To overcome this limitation, multiplexing, or combining, biomarkers for cancer detection can improve testing characteristics (10, 11). In this study, we sought to explore a multiplexed urine-based diagnostic test for prostate cancer.

Materials and Methods

Urine collection, RNA isolation, amplification, and quantitative PCR. Samples were obtained from 276 patients with informed consent following a digital rectal exam before either needle biopsy (n = 216) or radical prostatectomy (n = 60) at the University of Michigan Health System with Institutional Review Board approval (Supplementary Table S1). The digital rectal examination was done by systematically applying mild digital pressure over the entire palpated surface. Initial voided urine was then collected in urine collection cups containing DNA/RNA preservative (Sierra Diagnostics LLC). Isolation of RNA from urine and TransPlex whole transcriptome amplification (WTA) were as described (12). Quantitative PCR (qPCR) was used to detect seven prostate cancer biomarkers (AMACR, ERG, GOLPH2, PCA3, SPINK1, TFF3, and TMPRSS2:ERG) and the control transcripts PSAD and GAPDH from WTA-amplified cDNA essentially as described (12, 13). The primer sequences for ERG (exon5_6; ref. 8), GAPDH (14), AMACR (15), and PSAD (16) were previously described and for other biomarkers are listed in Supplementary Table S2. Threshold levels were set during the exponential phase of the qPCR using Sequence Detection Software version 1.2.2 (Applied Biosystems), with the same baseline and threshold set for each plate, to generate threshold cycle (Ct) values for all genes for each sample.

Analysis. qPCR was performed on WTA-cDNA from urine collected from 111 biopsy-negative patients and 165 patients with prostate cancer (105 biopsy-positive patients and 60 prostatectomy patients). Samples that had PSAD Ct values >28 were excluded to ensure sufficient prostate cell collection, leading to 105 biopsy negative and 152 samples from patients with prostate cancer in the analysis. We used raw ΔCt (to stabilize the variance of testing variables) as opposed to testing markers against control (2−ΔCt). TMPRSS2:ERG was dichotomized as a binary variable to reflect the
fusion-positive or fusion-negative status observed in tissue samples (8, 17), with positive samples defined as those with $C_t$ values of <37. As PCA3 has been reported to be a prostate tissue–specific marker (7), it was normalized against urine PSA ($C_{t_{PSA}} - C_{t_{PCA3}}$). All other testing variables were adjusted against their mean urine PSA and GAPDH values $[(C_{t_{PSA}} + C_{t_{GAPDH}}) / 2 - C_{t_{urinary}}]$. Normalized PSA alone was not able to discriminate urine from prostate cancer patients and patients with negative needle biopsy samples [area under the curve (AUC) = 0.44; $P$ = 0.11; Supplementary Fig. S1]. We additionally excluded 23 samples showing outlier values, as at least one testing variable (AMACR, PCA3, SPINK1, TFF3, and GOLPH2) in those samples showed an adjusted value below 3 SDs from its sample mean across the entire sample set. Examination of qPCR data confirmed that qPCR failed to detect target gene expression in those samples. This resulted in a final data set of samples from 138 patients with prostate cancer (86 positive needle biopsy and 52 radical prostatectomy) and 96 biopsy-negative patients.

**Statistical analysis.** Univariate and multivariate logistic regressions were used to examine associations between prostate cancer diagnostic status and testing variables. For multivariate logistic regression, the Akaike information criterion (AIC)-based backward selection was used to drop insignificant terms (18). The initial regression model included all testing markers and was further refined by the AIC-based backward selection. After

---

**Figure 1.** Characterization of candidate urine-based biomarkers of prostate cancer. A to C, qPCR was performed on WTA cDNA from urine obtained from patients presenting for needle biopsy or prostatectomy. Biomarker expression in patients with negative needle biopsies (green) or patients with prostate cancer (PCa; positive needle biopsy or prostatectomy; red) is shown. Normalization was performed using $\Delta C_t$, with PCA3 normalized to urine PSA expression as performed previously (25). AMACR, ERG, GOLPH2, SPINK1, and TFF3 were normalized to the average of urine sediment PSA and GAPDH expression. TMPRSS2:ERG gene fusion expression was dichotomized as positive or negative. The $\Delta C_t$ values of genes that were not significant predictors of prostate cancer by univariate analysis (see Table 1) are shown in A, and the expression of those that were significant predictors is shown in B and C. $P$ values from the univariate analysis for the detection of prostate cancer are indicated. D, ROC curves for individual variables for the diagnosis of prostate cancer. AUCs for GOLPH2, PCA3, SPINK1, and serum PSA are 0.664, 0.661, 0.642, and 0.508, respectively.
the final model was determined, the predicted probability for each sample was used as input to generate the receiver-operating characteristic (ROC) curve and the AUC was calculated. As all samples were used for regression model generation, the estimated AUC may be overoptimized. To correct this bias, we further performed a leave-one-out cross-validation (LOOCV). Briefly, one sample was omitted, whereas the regression model was trained on the remaining samples to select optimal markers and estimate their coefficients. The prediction probability is then calculated based on the model prediction for the left-out sample. This was repeated until every sample was left out once and the generated prediction probability values were then used for ROC analysis. Similarly, PCA3 was fitted in a logistic regression model to generate an AUC. The difference of AUCs was examined as described previously (19). All analyses were performed in R² and ROC curves were plotted in Statistical Package for the Social Sciences 11.5 (SPSS, Inc.).

Risk stratification. Clinical information was identified from medical records to determine association with clinical factors and risk categories based on biopsy results and pathologic data (20). Nomograms were used to calculate risk of progression-free survival and pathologic staging (21, 22). All variables were tested for univariate association with each clinical risk group.

Results and Discussion
To develop a multiplexed qPCR-based test for prostate cancer, we assessed seven putative prostate cancer biomarkers in a final cohort of 138 patients with prostate cancer (86 positive needle biopsy and 52 radical prostatectomy patients) and 96 patients with negative needle biopsies (Supplementary Table S1). Biomarkers included those generally overexpressed in prostate cancer, such as PCA3, AMACR, and PCA3 (6, 7), as well as those overexpressed in subsets of prostate cancers, such as ERG and TMPRSS2:ERG, and TFF3 and SPINK1 (8, 23, 24).

All genes were first tested by univariate analysis, with GOLPH2 (P = 0.0002), SPINK1 (P = 0.0002), PCA3 (P = 0.001), and TMPRSS2:ERG fusion (P = 0.034) showing significant association for discriminating patients with prostate cancer from patients with negative needle biopsies (Fig. 1; Table 1). Both AMACR, which has previously been shown to be a sensitive and specific biomarker for prostate cancer in tissues (6, 23), and TFF3, which shows high expression in a subset of prostate cancers (23, 24), were not statistically significant predictors of prostate cancer using urine samples (P = 0.450 and 0.189, respectively). The lack of specificity of these genes in urine may be due to expression of these transcripts in urothelial- or kidney-derived cellular material that shed in the urine. Whereas TMPRSS2:ERG fusion was significantly associated with the presence of prostate cancer (Fig. 1; Table 1), ERG overexpression was not associated with cancer presence on univariate analysis (P = 0.166), suggesting that cells from other tissues may be contributing ERG transcripts in urine. Additionally, serum PSA levels before biopsy or prostatectomy were also not associated with cancer presence in this cohort (P = 0.376). When tested as individual variables for the ability to detect prostate cancer based on the ROC curves, GOLPH2 (AUC = 0.664; P = 2.01E−5), PCA3 (AUC = 0.661; P = 2.84E−5), and SPINK1 (AUC = 0.642; P = 0.0002) outperformed serum PSA (AUC = 0.508; P = 0.837; Fig. 1). Thus, in this study, we have identified multiple biomarkers for urine-based noninvasive detection of prostate cancer. Of the seven markers tested in this study, only PCA3 was previously reported as a urinary diagnostic biomarker (9).

To determine if a multiplex model could improve performance over single biomarkers, tested biomarkers were next analyzed in a multivariate regression analysis using AIC-based backward selection (18) to drop insignificant terms from the model. This analysis resulted in a final model that included SPINK1 (P = 7.41E−5), PCA3 (P = 0.0003), GOLPH2 (P = 0.004), and TMPRSS2:ERG (P = 0.006; Table 1). To evaluate the performance of this model for diagnosing prostate cancer, we then performed ROC analysis based on the predicted probabilities derived from the final model. For our cohort, we compared the ROC curves from the multiplexed model and PCA3 alone, as urine-based detection of PCA3 has previously been evaluated in similar cohorts as a single biomarker using alternative detection technologies (9, 25–28). For example, van Gils et al. (9) showed that, in a cohort of 534 men presenting for prostate biopsy with serum PSA between 3 and 15 ng/mL, urinary PCA3 detection expression had an AUC of 0.66 compared with 0.57 for serum PSA. As shown in Fig. 2A, in our cohort, the AUC for the multiplexed model (0.758; P = 1.91E−11) was significantly improved [P = 0.003 (19)] compared with the AUC for PCA3 alone (0.662; P = 2.58E−5). At the point on the multiplex model ROC with the maximum sum of sensitivity and specificity (65.9% and 76.0%, respectively), the positive and negative predictive values were 79.8% and 60.8%, respectively (Fig. 2A). As we and previous studies used different methodologies to detect PCA3 transcripts in patient urine, directly comparing AUCs is inappropriate; however, we show that PCA3 shows improved AUC compared with serum PSA, consistent with previous reports (9, 25–28). Importantly, we further show that a multiplex model significantly improves predictive ability compared with PCA3 alone. The rationale for the multiplex approach is consistent with tests offered to breast cancer patients to identify patients at high risk for disease recurrence (10, 29).

As all samples were used to select the best subset of variables for regression analysis, this has the potential to overoptimize the reported AUC. Thus, we used LOOCV strategy to generate an unbiased AUC. As shown in Fig. 2B, the AUC for the LOOCV multiplex model (0.736) is again significantly better (P = 0.006) than

Table 1. Univariate and multivariate logistic regression analyses were used to identify urine biomarkers for the detection of prostate cancer

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Univariate logistic regression analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOLPH2</td>
<td>0.4444</td>
<td>0.0002</td>
</tr>
<tr>
<td>SPINK1</td>
<td>0.25</td>
<td>0.0002</td>
</tr>
<tr>
<td>PCA3</td>
<td>0.187</td>
<td>0.001</td>
</tr>
<tr>
<td>TMPRSS2:ERG</td>
<td>0.609</td>
<td>0.034</td>
</tr>
<tr>
<td>ERG</td>
<td>0.043</td>
<td>0.166</td>
</tr>
<tr>
<td>TFF3</td>
<td>0.11</td>
<td>0.189</td>
</tr>
<tr>
<td>PSA (serum)</td>
<td>0.0151</td>
<td>0.376</td>
</tr>
<tr>
<td>AMACR</td>
<td>0.049</td>
<td>0.45</td>
</tr>
<tr>
<td><strong>Multivariate logistic regression analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPINK1</td>
<td>0.308</td>
<td>7.41E-05</td>
</tr>
<tr>
<td>PCA3</td>
<td>0.191</td>
<td>0.003</td>
</tr>
<tr>
<td>GOLPH2</td>
<td>0.372</td>
<td>0.004</td>
</tr>
<tr>
<td>TMPRSS2:ERG</td>
<td>0.924</td>
<td>0.006</td>
</tr>
</tbody>
</table>

NOTE: For the multivariate analysis, AIC-based backward selection was used to drop insignificant terms.
that for LOOCV PCA3 alone (0.645). At the point on the LOOCV multiplex model ROC with the maximum sum of sensitivity and specificity (62.3% and 75.0%, respectively), the positive and negative predictive values were 78.2% and 58.1%, respectively (Fig. 2B).

Lastly, we tested the ability of these genetic markers to predict clinical risk groups based on patient variables. Clinical risk groups were determined by clinical patient data that direct the decision to pursue biopsy, to determine treatment, or to stratify patients for surveillance regimens. We observed only limited association between these prostate cancer biomarkers and clinical risk groups, with GOLPH2, SPINK1, and TMPRSS2:ERG status showing association with risk groups (Supplementary Table S3). As the biomarkers in this study were chosen based on their ability to differentiate benign prostate tissue and prostate cancer, it is not surprising that they did not show strong association with risk stratification measures. Future efforts will be directed toward adding markers that would enable risk stratification based on prebiopsy urine samples. Similar to the previously described PCR-based test for breast cancer recurrence risk, a prostate cancer risk test could drive high-risk patients to therapies more suited for their disease course (10).

In summary, we show that a multiplexed qPCR assay on sedimented urine from patients presenting for prostate biopsy or prostatectomy outperforms serum PSA or PCA3 alone. Notably, the multiplex urine test presented here achieves a specificity and positive predictive value of >75%, establishing a basic framework for the development of a urine multiplex test for the noninvasive detection of prostate cancer. These results support examination of larger cohorts across multiple institutions for further validation. Future studies will be directed at improving the performance of this first-generation urine multiplex test by evaluating additional markers and improving risk stratification and patient counseling before treatment decision making.

Addendum

In a recent report, Hessels et al. (30) showed that a combined test for PCA3 and TMPRSS2:ERG expression in urine using alternative diagnostic assays outperformed serum PSA and PCA3 alone for the detection of prostate cancer.

Acknowledgments

Received 8/20/2007; revised 11/19/2007; accepted 12/6/2007.
Grant support: Department of Defense grants PC040517 and W81XWH-06-1-0224, NIH grants U54 DA021519-01A1 and R01 CA102872, NIH Prostate Specialized Program of Research Excellence grant P50CA69568, Early Detection Research Network grants U01 CA111275-01 and U01 CA113913, Prostate Cancer Foundation, and Gen-Probe Incorporated. A.M. Chinnaiyan is supported by a Clinical Translational Research Award from the Burroughs Wellcome Foundation. S.A. Tomlins is supported by a Rackham Predoctoral Fellowship and is a Fellow of the Medical Scientist Training Program. 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.

We thank Alex Bond for technical assistance and Mahaveer Bhojani for helpful discussions.

References


Figure 2. A multiplexed set of urine biomarkers outperforms PCA3 alone in the detection of prostate cancer. A, multivariate regression analysis resulted in a multiplexed model, including SPINK1, PCA3, GOLPH2, and TMPRSS2:ERG as significant predictors of prostate cancer (see Table 1). ROC analysis was then performed based on the predicted probabilities derived from the final model. The multiplexed model (red) showed significantly greater AUC than PCA3 (blue) alone (0.738 versus 0.662; P = 0.003) for the detection of prostate cancer. The point on the ROC curve with the maximum sum of sensitivity (Sens) and specificity (Spec) is indicated by the dashed line, and the positive (PPV) and negative (NPV) predictive values are given. B, as in A, except LOOCV strategy was used to generate unbiased AUCs. The AUC for the LOOCV multiplex model is significantly better than LOOCV of PCA3 alone (0.736 versus 0.645; P = 0.006).
A First-Generation Multiplex Biomarker Analysis of Urine for the Early Detection of Prostate Cancer

Bharathi Laxman, David S. Morris, Jianjun Yu, et al.