



Published in final edited form as:

*Cancer Res.* 2008 February 1; 68(3): 645–649. doi:10.1158/0008-5472.CAN-07-3224.

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

Bharathi Laxman<sup>1,2</sup>, David S. Morris<sup>1,3</sup>, Jianjun Yu<sup>1,2,4</sup>, Javed Siddiqui<sup>1,2</sup>, Jie Cao<sup>1,2</sup>, Rohit Mehra<sup>1,2,5</sup>, Robert J. Lonigro<sup>1,5</sup>, Alex Tsodikov<sup>1,6</sup>, John T. Wei<sup>1,3,5</sup>, Scott A. Tomlins<sup>1,2</sup>, and Arul M. Chinnaiyan<sup>1,2,3,4,5</sup>

<sup>1</sup> Michigan Center for Translational Pathology, University of Michigan Medical School, Ann Arbor, Michigan

<sup>2</sup> Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan

<sup>3</sup> Department of Urology, University of Michigan Medical School, Ann Arbor, Michigan

<sup>4</sup> Bioinformatics Program, University of Michigan Medical School, Ann Arbor, Michigan

<sup>5</sup> Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan

<sup>6</sup> Department 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 sedimented 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.

### 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

Requests for reprints: Arul M. Chinnaiyan, University of Michigan Medical School, 1400 East Medical Center Drive, 5316 CCGC, Ann Arbor, MI 48109-0602. Phone: 734-615-4062; Fax: 734-615-4498; arul@umich.edu.

B. Laxman, D.S. Morris, J. Yu, and S.A. Tomlins contributed equally to this work.

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

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 *PSA* 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 *PSA* (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 ( $C_t$ ) 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 *PSA*  $C_t$  values of  $>28$  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  $-\Delta C_t$  (to stabilize the variance of testing variables) as opposed to testing markers against control ( $2^{-\Delta C_t}$ ). *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* ( $C_{tPSA} - C_{tPCA3}$ ). All other testing variables were adjusted

against their mean urine *PSA* and *GAPDH* values  $[(C_{tPSA} + C_{tGAPDH})/2 - C_{tVariable}]$ . 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 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 R<sup>7</sup> 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 *GOLPH2* (6, 7),<sup>8</sup> as well as those overexpressed in subsets of prostate cancers, such as *ERG* and *TMPRSS2:ERG*, and *TFF3* and *SPINK1* (8, 23, 24).<sup>9</sup>

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

<sup>7</sup><http://www.r-project.org>

<sup>8</sup>A.M. Chinnaiyan, unpublished observations.

<sup>9</sup>S.A. Tomlins et al., unpublished observations.

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), 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* ( $P = 0.003$ ), *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 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.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**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.

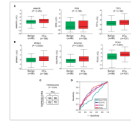
We thank Alex Bond for technical assistance and Mahaveer Bhojani for helpful discussions.

## References

1. Brawer MK, Chetner MP, Beatie J, Buchner DM, Vessella RL, Lange PH. Screening for prostatic carcinoma with prostate specific antigen. *J Urol* 1992;147:841–5. [PubMed: 1371559]
2. Catalona WJ, Smith DS, Ratliff TL, et al. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991;324:1156–61. [PubMed: 1707140]
3. Brawley OW, Knopf K, Merrill R. The epidemiology of prostate cancer part I: descriptive epidemiology. *Semin Urol Oncol* 1998;16:187–92. [PubMed: 9858324]
4. Catalona WJ, Hudson MA, Scardino PT, et al. Selection of optimal prostate specific antigen cutoffs for early detection of prostate cancer: receiver operating characteristic curves. *J Urol* 1994;152:2037–42. [PubMed: 7525995]

5. Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level  $\leq 4.0$  ng per milliliter. *N Engl J Med* 2004;350:2239–46. [PubMed: 15163773]
6. Rubin MA, Zhou M, Dhanasekaran SM, et al.  $\alpha$ -Methylacyl coenzyme A racemase as a tissue biomarker for prostate cancer. *JAMA* 2002;287:1662–70. [PubMed: 11926890]
7. de Kok JB, Verhaegh GW, Roelofs RW, et al. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res* 2002;62:2695–8. [PubMed: 11980670]
8. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644–8. [PubMed: 16254181]
9. van Gils MP, Hessels D, van Hooij O, et al. The time-resolved fluorescence-based PCA3 test on urinary sediments after digital rectal examination; a Dutch multicenter validation of the diagnostic performance. *Clin Cancer Res* 2007;13:939–43. [PubMed: 17289888]
10. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817–26. [PubMed: 15591335]
11. Schmidt U, Fuessel S, Koch R, et al. Quantitative multi-gene expression profiling of primary prostate cancer. *Prostate* 2006;66:1521–34. [PubMed: 16921506]
12. Laxman B, Tomlins SA, Mehra R, et al. Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. *Neoplasia* 2006;8:885–8. [PubMed: 17059688]
13. Tomlins SA, Mehra R, Rhodes DR, et al. Whole transcriptome amplification for gene expression profiling and development of molecular archives. *Neoplasia* 2006;8:153–62. [PubMed: 16611408]
14. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002;3:RESEARCH0034. [PubMed: 12184808]
15. Kumar-Sinha C, Shah RB, Laxman B, et al. Elevated  $\alpha$ -methylacyl-CoA racemase enzymatic activity in prostate cancer. *Am J Pathol* 2004;164:787–93. [PubMed: 14982833]
16. Specht K, Richter T, Muller U, Walch A, Werner M, Hofler H. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol* 2001;158:419–29. [PubMed: 11159180]
17. Perner S, Mosquera JM, Demichelis F, et al. TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. *Am J Surg Pathol* 2007;31:882–8. [PubMed: 17527075]
18. Venables, WN.; Ripley, BD. *Modern applied statistics with S*. 4. New York: Springer; 2002.
19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45. [PubMed: 3203132]
20. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998;280:969–74. [PubMed: 9749478]
21. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst* 1998;90:766–71. [PubMed: 9605647]
22. Kattan MW, Stapleton AM, Wheeler TM, Scardino PT. Evaluation of a nomogram used to predict the pathologic stage of clinically localized prostate carcinoma. *Cancer* 1997;79:528–37. [PubMed: 9028364]
23. Faith DA, Isaacs WB, Morgan JD, et al. Trefoil factor 3 overexpression in prostatic carcinoma: prognostic importance using tissue microarrays. *Prostate* 2004;61:215–27. [PubMed: 15368473]
24. Garraway IP, Seligson D, Said J, Horvath S, Reiter RE. Trefoil factor 3 is overexpressed in human prostate cancer. *Prostate* 2004;61:209–14. [PubMed: 15368472]
25. Hessels D, Klein Gunnewiek JM, van Oort I, et al. DD3(PCA3)-based molecular urine analysis for the diagnosis of prostate cancer. *Eur Urol* 2003;44:8–5. discussion 6. [PubMed: 12814669]
26. Fradet Y, Saad F, Aprikian A, et al. uPM3, a new molecular urine test for the detection of prostate cancer. *Urology* 2004;64:311–5. discussion 315–6. [PubMed: 15302485]

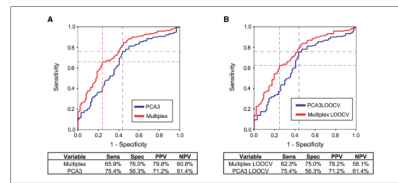
27. Groskopf J, Aubin SM, Deras IL, et al. APTIMA PCA3 molecular urine test: development of a method to aid in the diagnosis of prostate cancer. *Clin Chem* 2006;52:1089–95. [PubMed: 16627561]
28. Marks LS, Fradet Y, Deras IL, et al. PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. *Urology* 2007;69:532–5. [PubMed: 17382159]
29. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999–2009. [PubMed: 12490681]
30. 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. [PubMed: 17785564]



**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.





**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 (0.758 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;  $P = 0.006$ ).

**Table 1**

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

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

NOTE: For the multivariate analysis, AIC-based backward selection was used to drop insignificant terms.