Plasma Messenger RNAs Identified Through Bioinformatics Analysis are Novel, Non-Invasive Prostate Cancer Biomarkers

Aim: To identify new biomarkers of prostate cancer (PCa) for the diagnosis and prediction of clinical outcomes.

Materials and Methods: Existing microarray data of PCa tissues in the Oncomine database were analyzed and candidate differentially expressed genes (DEGs) that may be novel and noninvasive biomarkers were obtained. On this basis, plasma mRNA was extracted from PCa patients and healthy donors. Furthermore, plasma mRNA expression of DEGs was evaluated by qRT-PCR. Finally, the diagnostic power of the biomarkers was evaluated in comparison to the clinical characteristics of the patients.

Results: In this study, the top five significantly overexpressed mRNA (AMACR, PPIHR14b, PCA3, DLX1, and RPL22L1) and the top five significantly underexpressed mRNA (DUOX1, EFS, GSTP1, S100A16, and NCRNA00087) were selected for further validation in PCa patients and healthy donors by qRT-PCR. The results showed that AMACR, DLX1, PCA3, DUOX1, and GSTP1 mRNA expression were significantly different between PCa circulating free mRNA samples and healthy donors. These mRNAs may be useful biomarkers for PCa diagnosis.

Conclusion: Analysis of the expression of genes in the Oncomine database showed that DLX1, PCA3, and DUOX1 expressions have a cancer specific pattern in PCa. Collectively, DLX1, PCA3, and DUOX1 may be useful candidate biomarkers for PCa diagnosis.

Keywords: prostate cancer, PCA3, DLX1, GSTP1, DUOX1

Introduction
Prostate cancer (PCa) is a major public health problem as it is the second most common cancer and the sixth leading cause of cancer-related deaths in males worldwide. In China, the occurrence rate of PCa is relatively low compared with other types of cancers. However, due to advances in diagnostic tools and the fact that people are living longer, the number of PCa diagnoses is continually increasing. Moreover, PCa is an indolent cancer, and localized forms can be well managed and successfully treated by traditional surgical treatment. However, the five-year survival rate of PCa patients with metastasis is approximately 30%. Early diagnosis contributes to the increased survival rate of PCa patients.

The serum marker prostate-specific antigen (PSA) has been widely used to diagnose PCa and identify PCa relapse, and it is a standard for use in treatment selection. However, some limitations exist in the PSA assay. Although it is organ specific, it is not cancer specific. Some prostate diseases including benign prostate hyperplasia (BPH),
prostatitis, and prostate manipulations (such as DRE and bicycling) lead to increased PSA levels. In addition, PSA is a conventional prognostic marker of PCa, and accumulating studies have demonstrated that patients diagnosed with PCa and equivalent PSA levels may have a variable natural history such as age, race, geographic location, familial history, and genetic background. Therefore, it is hard to predict the initiation, progression, and prognosis of PCa. Currently, there is a need for new biomarkers that can be used to diagnose PCa and predict the clinical outcome.

Since it was first demonstrated in the 1990s by Lo et al, circulating cell-free messenger RNA has been used to detect diagnostic and prognostic biomarkers in various cancers, including lung, nasopharyngeal, and colorectal cancers. In addition, March-Villalba et al showed that plasma telomerase reverse transcriptase (hTERT) mRNA can be a useful biomarker for the diagnosis and prediction of prognosis in PCa patients. The above studies suggested that cell-free RNA in plasma may play a vital role in cancer diagnosis and prognosis. However, due to the limited number of studies, more investigation is required to identify useful circulating cell-free messenger RNA.

Microarray data of cancer transcriptome analyses have been widely applied to explore useful candidate biomarkers from various samples, including tumor tissues and patient bodily fluids. Oncomine is a cancer microarray database and web-based data-mining platform aimed at facilitating useful oncogene and antioncogene discovery from genome-wide expression analyses. Oncomine has been successfully used to initially explore novel, noninvasive biomarkers through bioinformatics analysis in lung cancer.

In this study, existing microarray data from PCa tissues in the Oncomine database were compared with data from normal tissues to obtain useful candidate differently expressed genes (DEGs) as potential novel, noninvasive biomarkers. Next, plasma mRNA was extracted from PCa patients and healthy donors and plasma mRNA expression of DEGs was evaluated by qRT-PCR. Finally, the diagnostic power of these markers was validated in comparison to the clinical and pathological characteristics of these patients. The results from this study showed that some plasma messenger RNAs may be useful biomarkers for PCa diagnosis.

**Materials and Methods**

**Ethics Statement**

This study was approved by the institutional Ethical Committee of Shanghai Ninth Peoples’ Hospital. Signed informed consent was obtained from all human participants.

**Gene Expression Analysis via the Oncomine Database**

PCa microarray data in the Oncomine database were analyzed according to the schematic diagram shown in Figure 1. Concrete description of the Oncomine database was described previously. For this study, mRNA expression in PCa tissues as compared with adjacent normal tissues or was analyzed, and the cut-offs were determined with a P value of $\leq 10^{-4}$ and a fold change of $\geq 2$.

**Sample Collection**

Blood samples from 50 PCa patients and 30 healthy donors (inclusion criteria: No prostate disease and Normal PSA value) were collected by technical nurses in the Changhai hospital between October 2015 and October 2016. Patient parameters are shown in Table 1. Then, samples were centrifuged at 1400 g for 10 min at 4°C and 10,000 g for 15 min at 4°C to obtain plasma. Plasma was stored at −80°C until use.

**RNA Extraction**

Total RNA was extracted from 1 mL plasma using Trizol LS regent (Thermo Fisher Scientific Inc., USA) according to the manufacturers’ instructions. The Nanodrop 1000 (Nanodrop, USA) was used to evaluate the quality of RNA extracted from plasma; concentrations ranged from 3.34 to 15.7 ng/μL.

**Quantitative Real-Time PCR**

The extracted total RNA was used for cDNA synthesis using the PrimeScript RT reagent Kit (TAKARA, JAPAN) according to the manufacturer’s instructions. The cDNA was stored for the next PCR amplification. The reverse transcriptase polymerase chain reactions (RT-PCR) with rational forward and reverse primers were used to detect the transcripts of interest. The primer sequences were:

- **PCA3 F** (5’TTCGAGACCTTGGGCTGTCG 3’)
- **PCA3 R** (5’ATCTTGAATGGCCCTGCGTGTG 3’)
- **DLX1 F** (CAATGGCAAGGGAAAAAAGG)
- **DLX1 R** (GAACCAACATTCCTTGGATCTG)
- **GSTP1 F** (GCCTCCTGCCUUGGGGCA)
- **GSTP1 R** (CGAGAGGAGATCTGTCTCCACAAC)
- **DUOX1 F** (CGAGAGGACCATGCTGTGGT)
- **DUOX1 R** (TGCCGGAAACTTCAGTGGG)

The PCR conditions were as follows: 95°C for 5 min, then 40 cycles of 94°C for 30 sec, 58°C for 1 min. PCR was done with the ABI ViiA 7 Real-Time PCR System (Applied Biosystems). Standard curves were plotted for each optimized assay, each of which produced...
a linear plot of the threshold cycle (Ct) against the log (dilution). The quantity of each gene was normalized against GAPDH.

Statistical Analysis
Statistical analyses were performed using GraphPad 7.0 software. Data are expressed as means ± SD. Differences in expression levels of plasma cell-free mRNAs between PCa patients and healthy donors were assessed by the Mann–Whitney U-test. Analysis of receiver operating characteristic (ROC) curves were used to evaluate the candidate biomarker diagnostic value as previously described. Differences were considered significant at P < 0.05.

Results
Genes Screened from the Oncomine Database
Gene mRNA expression was identified in PCa samples and compared with normal samples by analyzing the Oncomine database, as shown in Figure 1. Sixteen studies that included PCa samples versus normal samples were analyzed, the results are shown in Figure 2. The top five significantly overexpressed mRNA were AMACR, PPP1R14b, PCA3, DLX1, and RPL22L1. The top five significantly underexpressed mRNA were: DUOX1, EFS, GSTP1, S100A16, and NCRNA00087.
Validation of Candidate mRNA Expression in the Plasma of PCa Patients and Healthy Donors

Based on the data from the Oncomine database analysis, 10 mRNAs (the top 5 overexpressed mRNAs and the top 5 underexpressed mRNAs) from the independently recruited PCa patients and healthy donors were selected for further validation. Among the selected mRNAs, PPP1R14b and S100A16 were barely detectable due to too low expression level. Detection of NCRNA significantly varied across technical triplicates, possibly due to its high Ct value which was beyond 35. Levels of RPL22L1 and EFS were undetectable in most of the samples, however AMACR, DLX1, PCA3, DUOX1, and GSTP1 were stably amplified. Of these mRNAs, AMACR expression in plasma was not significantly different in PCa patients compared with healthy donors (data not shown). As shown in Figure 3 (top), PCA3 and DLX1 mRNA expression in PCa patients were higher than in healthy donors, and DUOX1 and GSTP1 mRNA expression in PCa patients were lower than in healthy donors. In addition, the diagnostic power of these mRNAs in PCa diagnosis was determined by receiver-operating characteristic (ROC) curves. Area under the ROC curve (AUC) of PCA3 was 0.762, DLX1 was 0.821, DUOX1 was 0.7072 and GSTP1 was 0.643 (Figure 3, bottom). These results suggested that the deviations in plasma PCA3, DLX1, DUOX1, and GSTP1 levels may be useful as biomarkers for PCa diagnosis.

Candidate Biomarker Expression in Various Cancers as Compared with Normal Donors Was Analyzed in the Oncomine Database

To select the most appropriate biomarkers for PCa diagnosis, mRNAs expression was analyzed in various cancers and compared with normal controls via Oncomine database screening. The results showed that the expression of PCA3 was high only in PCa and was not expressed in other types of cancer (Figure 4). Consistent with the above

Table 1 Clinical Characteristics of Patients and Healthy Donors

<table>
<thead>
<tr>
<th></th>
<th>PCa Patients</th>
<th>Healthy Donors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>Age, mean±SD</td>
<td>65.7±6.9</td>
<td>64.3±7.5</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I–II</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>III–IV</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>Positive</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Gleason scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤6</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>&gt;6</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 Heat map of differentially expressed genes in PCa compared with normal controls. (A) Top 20 over-expressed genes (red). (B) Top 20 under-expressed genes (blue).
studies, PCA3 was shown to be a prostate specific molecule. DLX1 and GSTP1 exhibited the same expression tendency in various cancers, they were intensely overexpressed or underexpressed in PCa, with a contrast of expression types in most of cancers (Figure 4). DUOX1 is underexpressed in 10 of 20 cancer types, including gastric cancer, head and neck cancer, and prostate cancer. However, these results showed that DUOX1 is not a cancer-specific biomarker of PCa. These results showed that PCA3, DLX1, and GSTP1 are better candidate biomarkers for PCa diagnosis.

Discussion

In this study, several candidate biomarkers were identified by integrating diverse public microarray data in the Oncomine database. The top five significantly overexpressed mRNA (AMACR, PPP1R14b, PCA3, DLX1, and RPL22L1) and the top five significantly underexpressed mRNA (DUOX1, EFS, GSTP1, S100A16, and NCRNA00087) were selected for further validation in PCa patients and healthy donors by qRT-PCR. Results showed that AMACR, DLX1, PCA3, DUOX1, and GSTP1 mRNA were stably amplified in plasma. Additionally, DLX1, PCA3, DUOX1, and GSTP1 mRNA expression was significantly different between PCa circulating free mRNA samples and healthy donors and may be useful biomarkers for PCa diagnosis. Finally, analysis of the expression of these genes in various cancers using the Oncomine database showed that the expression of DLX1, PCA3, and DUOX1 is specific to PCa. Collectively, DLX1, PCA3, and DUOX1 may be useful candidate biomarkers for PCa diagnosis.

Noncoding RNAs have been broadly described as key regulators of progression in various cancers. They act by modulating cancer-related genes and their corresponding signaling pathways. A number of studies have shown that Prostate cancer antigen-3 (PCA3), a prostate-specific long noncoding RNA (lncRNA), is highly expressed in prostate cancer (PCa) tissues compared to normal tissues. Previous publications have demonstrated that molecular mechanisms that promote PCa progression modulate the expression of AR and androgen-responsive genes (ARGs) that are specifically expressed in PCa. In addition, it is overexpressed in PCa patients’ urine, compared with controls. Furthermore, the analyzing urine PCA3/PSA ratio is a feasible way to determine whether a repeat biopsy is needed after a previously negative initial biopsy. This assay was recommended by the EAU guidelines and approved by the Food and Drug Administration in 2012 as the first molecular test for PCa screening, detection, and diagnosis. However, its diagnostic value in aggressive prostate cancer is limited, and the value of this assay when using the patient serum is unclear. The current study showed that PCA3 mRNA is overexpressed in PCa patient plasma when compared with controls and may be a useful biomarker for PCa diagnosis.

Distal-less Homeobox 1 (DLX1), located at 2q32, is a group of transcription factors that share sequence homology with the Drosophila distal-less genes (DII). Although the oncogenic functions of DLX1 remain unclear, DLX genes are essential in the development of various organs, including the brain, bones, and blood. In addition, DLX1 is involved in the mediation of epithelial–
neuroendocrine differentiation, a characteristic associated with aggressive cancer.\textsuperscript{17} Although the molecular mechanism of DLX1 in cancer progression is completely unknown, several studies have reported that it may involve the enhancement of TGF-\textbeta/SMAD4 signaling. This may be particularly relevant when considering DLX1 as a biomarker for PCa diagnosis and prognosis. Leyten et al showed that urine DLX1 expression was higher in PCa patients compared with normal controls and may be a potential biomarker for the early diagnosis of PCa.\textsuperscript{19,20} This also was demonstrated in other studies using PCa tissues.\textsuperscript{21,22} In addition, further studies showed that

<table>
<thead>
<tr>
<th>Analysis Type by Cancer</th>
<th>Cancer vs. Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA3</td>
<td>DLX1</td>
</tr>
<tr>
<td>Bladder Cancer</td>
<td>3</td>
</tr>
<tr>
<td>Brain and CNS Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Breast Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Cervical Cancer</td>
<td>2</td>
</tr>
<tr>
<td>Colorectal Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Esophageal Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Gastric Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Head and Neck Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Kidney Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Leukemia</td>
<td>7</td>
</tr>
<tr>
<td>Liver Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Lung Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>1</td>
</tr>
<tr>
<td>Melanoma</td>
<td>1</td>
</tr>
<tr>
<td>Myeloma</td>
<td>4</td>
</tr>
<tr>
<td>Other Cancer</td>
<td>2</td>
</tr>
<tr>
<td>Ovarian Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Pancreatic Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Prostate Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>2</td>
</tr>
</tbody>
</table>

Significant Unique Analyses
Total Unique Analyses

Figure 4 PCA3, DLX1, DUOX1, and GSTP1 expression in various cancers were analyzed. Red represents overexpression in various cancers compared with normal controls; blue represents underexpression in various cancers compared with normal controls.
DLX1 silencing inhibited PCa growth. In the present study, we showed that DLX1 mRNA is overexpressed in PCa patient plasma when compared with controls and may be a useful biomarker for PCa diagnosis.

Glutathione S-transferase π (GTSP1) is one of the glutathione S-transferases (GSTs), a superfamily of phase II enzymes that are mainly involved in detoxifying carcinogens, including tobacco, alcohol, antineoplastic drugs, and molecules, related to oxidative stress, which is an important antineoplastic mechanism induced by radiotherapy.23 Although GSTP1 has been shown to be overexpressed in various cancers, GSTP1 expression is downregulated in PCa compared with controls.24 GSTP1 expression is inhibited by GSTP1 DNA promoter methylation, which is associated with different PCa stages and PCa recurrence, and has been detected in more than 90% of PCas and approximately 70% of PIN lesions.25 GSTP1 silencing was reversed via demethylation using histone deacetylase inhibitors, resulting in a significant gain of function in the inhibition of PCa progression.26 In contrast, the sensitivity of chemotherapeutic drugs in the treatment of PCa can be increased by GSTP1 gene silencing.27 The above studies suggested that GTSP1 might not only be a prognostic biomarker of PCa but also could be used in guiding the selection of chemotherapeutic drugs. In this study, plasma GSTP1 mRNA was overexpressed in PCa patient plasma compared to controls and may be a useful biomarker for PCa diagnosis. In addition, plasma DUOX1 mRNA was overexpressed in PCa when compared with controls.28 DUOX1 is a subfamily of NADPH oxidase (NOX), which transfers electrons across membranes to molecular oxygen, producing ROS.29–31 However, DUOX1 overexpression has also been shown in various cancers, including lung cancer, hepatocellular cancer, and oral squamous cancer.29–31 Therefore, DUOX1 may be not a PCa-specific biomarker.

In conclusion, analyzing the plasma expression of mRNAs identified from the Oncomine database is a feasible way to determine biomarkers for PCa diagnosis. The results showed that plasma PCA3, DLX1, and GSTP1 mRNA expression, respectively, were useful biomarkers for PCa diagnosis, although further validation in larger-scale studies is needed.

However, there were still some shortages and imminent limitations in this study. Firstly, validation samples size is too small, only including 50 PCa and 30 healthy control. This conclusion needs more samples to revalidate. Secondly, gene expression was analyzed in PCa and adjacent normal tissues in oncomine database, however, to explore novel and non-invasive prostate cancer biomarkers, DEGs were validated in blood samples. Maybe gene expression has a difference between PCa tumor tissues and blood samples of PCa patients. In addition, due to RNA is unstable and easy to degrade, successful RNA measurement need higher preservation conditions and operation requirement.

Acknowledgments
The authors greatly thank Dr. Yi Hao (from Shanghai Ninth People’s Hospital, College of Stomatology, Shanghai Jiao Tong University School of Medicine) for contributions to this paper, especially bioinformatics analysis. This work was supported by a grant from the National Natural Science Foundation of China (No. 81603634).

Disclosure
The authors report no conflicts of interest in this work.

References


