Title
Six stroma-based RNA markers diagnostic for prostate cancer in European-Americans validated at the RNA and protein levels in patients in China

Permalink
https://escholarship.org/uc/item/21j4x80f

Journal
Oncotarget, 6(18)

Authors
Zhu, J
Pan, C
Jiang, J
et al.

Publication Date
2015

DOI
10.18632/oncotarget.4430

Peer reviewed
Six stroma-based RNA markers diagnostic for prostate cancer in European-Americans validated at the RNA and protein levels in patients in China

Jianguo Zhu1,2,*, Cong Pan1,*, Jun Jiang3,*, Mingsen Deng1,*, Hengjun Gao4, Bozhao Men5, Michael McClelland5, Dan Mercola6, Wei-De Zhong7 and Zhenyu Jia1,8,9

1 Guizhou Provincial Key Laboratory of Computational Nano-Material Science, Guizhou Normal College, Guiyang, China
2 Department of Urology, Guizhou Provincial People’s Hospital, Guizhou, China
3 Department of Chemical Physics, University of Science and Technology of China, Hefei, China
4 National Engineering Center for Biochip at Shanghai, Shanghai, China
5 Department of Microbiology and Molecular Genetics, University of California, Irvine, Irvine, CA, USA
6 Department of Pathology and Laboratory Medicine, University of California, Irvine, Irvine, CA, USA
7 Department of Urology, Guangdong Key Laboratory of Clinical Molecular Medicine and Diagnostics, Guangzhou First People’s Hospital, Guangzhou Medical University, Guangzhou, China
8 Department of Statistics, The University of Akron, Akron, OH, USA
9 Department of Family and Community Medicine, Northeast Ohio Medical University, Rootstown, OH, USA

* These authors are Co-first authors

Correspondence to: Zhenyu Jia, email: zjia@uakron.edu
Weide Zhong, email: zhongwd2009@live.cn

Keywords: prostate cancer, microenvironment, stroma, diagnosis, race

Received: February 04, 2015 Accepted: May 31, 2015 Published: June 19, 2015

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

We previously analyzed human prostate tissue containing stroma near to tumor and from cancer-negative tissues of volunteers. Over 100 candidate gene expression differences were identified and used to develop a classifier that could detect nearby tumor with an accuracy of 97% (sensitivity = 98% and specificity = 88%) based on 364 independent test cases from primarily European American cases. These stroma-based gene signatures have the potential to identify cancer patients among those with negative biopsies. In this study, we used prostate tissues from Chinese cases to validate six of these markers (CAV1, COL4A2, HSPB1, ITGB3, MAP1A and MCAM). In validation by real-time PCR, four genes (COL4A2, HSPB1, ITGB3, and MAP1A) demonstrated significantly lower expression in tumor-adjacent stroma compared to normal stroma (p value ≤ 0.05). Next, we tested whether these expression differences could be extended to the protein level. In IHC assays, all six selected proteins showed lower expression in tumor-adjacent stroma compared to the normal stroma, of which COL4A2, HSPB1 and ITGB3 showed significant differences (p value ≤ 0.05). These results suggest that biomarkers for diagnosing prostate cancer based on tumor microenvironment may be applicable across multiple racial groups.

INTRODUCTION

After years of intensive research, prostate cancer remains one of the major men’s health issues worldwide and the second leading cause of cancer related death in males in United States. In China the disease was comparatively less frequent but the incidence is rapidly increasing to levels comparable to Europe and North America [1, 2].

Accurate diagnosis of the disease, especially the malignant type, is critical for optimal patient care. However, even the best current methods, including
transrectal ultrasound (TRUS) procedures, may miss up to 30% of clinically significant prostate cancers who give false negative results on initial biopsy [3, 4]. These false negative biopsies contain ample stroma and other types of tissues (tumor microenvironment) that may be close to tumors. The tumor influences adjacent tissues through so-called paracrine mechanisms. It has been reported that the interaction and crosstalk between tumor and its microenvironment together contribute to tumor progression [5-12].

We previously developed a gene expression classifier which can diagnose the presence of prostate tumor using tumor adjacent stroma tissue. Almost exclusively European American (EA) cases in North America were used in that study [13]. We compared tumor-adjacent stroma to the stroma from normal subjects using fresh-frozen tissues and detected hundreds of gene expression differences. After correcting for age-related expression changes, we identified 131 genes that are reliably altered in RNA expression between tumor adjacent stroma and normal stroma. We then developed an expression classifier with these 131 genes using PAM [14]. The classifier was then tested on 364 independent cases with an accuracy of 97% (sensitivity = 98% and specificity = 88%) [13].

Racial and ethnic differences in outcome for prostate cancer (PCa) in the United States have been well documented [15-17]. The incidence of prostate cancer is lower in Asia [18]. These differences raise the question of whether the biomarkers for early diagnosis are also influenced by race. To determine if these diagnostic markers were broadly applicable for Chinese patients and, further, if they were applicable at the protein level, we selected six genes (CAV1, COL4A2, HSPB1, ITGB3, MAP1A, and MCAM) from the initial 131 gene expression differences between tumor adjacent stroma and normal stroma in North American cases. Expression of RNA from the six genes was measured using real-time PCR with prostate samples from a collection of Chinese cases (set A). Then, antibodies directed at the proteins encoded by these six genes were employed for immunohistochemistry assays on tissue microarrays which consisted of tissue samples from another collection of Chinese prostate cancer patients and Chinese normal subjects (set B). The results from both assays based on Chinese samples highly correlated with the results of the previous study based on North American cases, and extended them to the protein level, indicating that it may be feasible to develop a diagnostic assay applicable to multiple races. Investigation of the function of these genes in depth (and genes that are closely related to these genes) may help uncover the biology of etiology of the disease.

**RESULTS**

We have previously identified 131 genes that are significantly differentially expressed between tumor adjacent stroma and normal stroma using North American cases [13]. In the current study, we selected six representative genes for validation using Chinese cases. Previous literature indicated that these six genes, i.e., CAV1 [19, 20], COL4A2 [21], HSPB1 [22-24], ITGB3 [25], MAP1A [26] and MCAM [27], were linked to prostate cancer progression. Furthermore, pathway analysis using these six genes showed that they are highly significantly associated with important transcription factors, such as AP-1, AP-2, DPP4, AML1ETO fusion protein, p53, PBX2, PPAR-gamma, and STAT3.

We used a set of Chinese cases from the tissue bank at Guangzhou First People’s Hospital, China (Set A) for an RT-qPCR assay to test whether the RNA expression changes previously observed in primarily European-American patients, could be extended to Chinese patients. Stroma samples were collected from prostate cancer patients and normal prostate glands using LCM. RT-qPCR was used to measure the RNA expression changes of these six genes between tumor adjacent stroma and normal stroma. The results based on t test (Figure 1) indicate that COL4A2, HSPB1, ITGB3, and MAP1A were significantly less expressed in tumor adjacent stroma in comparison with normal stroma (p values < 0.05). The original images for the PCR agarose gel are given in Figure S1.

We next used a Tissue Microarray to investigate

**Figure 1:** The results of RT-qPCR. A. The electrophoresis of PCR product for six genes. B. Histogram for the relative expression levels for six genes. NS: normal stroma; TS: tumor adjacent stroma; *P < 0.05, **P < 0.01
Table 1: Summary statistics based on the t test.

<table>
<thead>
<tr>
<th>Protein</th>
<th>CAV1</th>
<th>COL4A2</th>
<th>HSPB1</th>
<th>ITGB3</th>
<th>MAP1A</th>
<th>MCAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MExp TS</td>
<td>195.3</td>
<td>188.9</td>
<td>177.7</td>
<td>190.3</td>
<td>188.6</td>
<td>194.4</td>
</tr>
<tr>
<td>Sd TS</td>
<td>5.3</td>
<td>10.4</td>
<td>12.1</td>
<td>6.4</td>
<td>5.9</td>
<td>7.6</td>
</tr>
<tr>
<td>MExp NS</td>
<td>197.4</td>
<td>195.7</td>
<td>183.4</td>
<td>196.0</td>
<td>190.0</td>
<td>197.2</td>
</tr>
<tr>
<td>Sd NS</td>
<td>6.6</td>
<td>4.5</td>
<td>8.2</td>
<td>3.3</td>
<td>5.7</td>
<td>5.3</td>
</tr>
<tr>
<td>t statistic</td>
<td>-1.18</td>
<td>-5.06</td>
<td>-2.52</td>
<td>-6.00</td>
<td>-0.89</td>
<td>-1.87</td>
</tr>
<tr>
<td>P value</td>
<td>0.255</td>
<td>3.26E-05</td>
<td>0.021</td>
<td>5.38E-06</td>
<td>0.388</td>
<td>0.077</td>
</tr>
</tbody>
</table>

MExp: mean expression; Sd: standard deviation

**Figure 2:**

a. The stroma component selected for differential expression analysis. The images presented in the figure were obtained from IHC analysis of MCAM. Samples labeled with black/red represent, respectively, tumor-bearing and tumor-adjacent tissues from patients' prostate glands; samples labeled with green are prostate tissue from normal subjects.

b. Representative figures obtained from IHC analysis of CAV1, COL4A2, ITGB3, HSPB1, MAP1A and MCAM. The samples labeled with T represent tumor-bearing/tumor-adjacent stroma area from patients' prostate glands; the samples labeled with N represent tumor-adjacent stroma (tumor free) area from the same patients' prostate glands.
whether these results could be extended, for the first time, to protein (Set B). TMAs were manually viewed by an experienced pathologist. From each tumor-bearing sample, we selected three stroma regions (marked yellow in Figure 2 (a)) representing the stroma that are adjacent to tumor epithelium regions. In our previous studies, we have learned that the field effect related to tumor paracrine mechanism depends on the tumor-stroma distance, for which the closer the stronger signal would be detected. Based on our experience, the optimal tumor-stroma

![Image of immunohistochemistry (IHC) stained prostate tissue microarray (TMA) cores showing heterogeneous compartment loci: lumen, epithelial cell, and stroma.](image)

**Figure 3:** Schematic of an immunohistochemistry (IHC) stained prostate tissue microarray (TMA) cores showing heterogeneous compartment loci: lumen, epithelial cell, and stroma.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patent sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>Set A</td>
</tr>
<tr>
<td>Patients</td>
<td>21</td>
</tr>
<tr>
<td>Normal subjects</td>
<td>8</td>
</tr>
<tr>
<td>Age for patients only</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>55</td>
</tr>
<tr>
<td>Maximum</td>
<td>87</td>
</tr>
<tr>
<td>Median</td>
<td>73</td>
</tr>
<tr>
<td>Gleason Score</td>
<td></td>
</tr>
<tr>
<td>2-6</td>
<td>8 (38%)</td>
</tr>
<tr>
<td>7 (3+4)</td>
<td>3 (14%)</td>
</tr>
<tr>
<td>7 (4+3)</td>
<td>2 (10%)</td>
</tr>
<tr>
<td>8-10</td>
<td>8 (38%)</td>
</tr>
<tr>
<td>Insufficient data</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>5 (24%)</td>
</tr>
<tr>
<td>T2</td>
<td>7 (33%)</td>
</tr>
<tr>
<td>T3</td>
<td>9 (43%)</td>
</tr>
<tr>
<td>Lymph Node Status</td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td></td>
</tr>
<tr>
<td>Insufficient data</td>
<td></td>
</tr>
</tbody>
</table>

The percentage in parenthesis represents the percent of patient cases (not including normal subjects) in each category.
distance for these markers is within 1mm [28-31]. So, we used pen tool in Aperio Imagescope system to select stroma that are close to tumor regions (Figure 2).

Enlarged pictures are shown in Figure 2 (b). For each of five normal tissues, three stroma regions were also selected for comparison (not shown in Figure 2 (a)). The selected areas were analyzed with the application “Positive Pixel Count V9”, by which the image data were translated to numerical data. The average intensity, which is the ratio of the sum of the intensities of positive signal (weak positive, positive and strong positive) and the sum of the number of positive signal (weak positive, positive and strong positive), is calculated and used for further analysis. For each of six proteins, the mean average intensities and standard deviations were calculated and compared between tumor adjacent stroma and normal stroma using the Student t test. The summary statistics based on the t test are presented in Table 1. All six proteins are less expressed in tumor adjacent stroma in comparison with the normal stroma. The differences are statistically significant for COL4A2, HSPB1 and ITGB3 (p values < 0.05), and the difference for MCAM showed the same trend (p value = 0.077, not significant but approaching significant level of 0.05).

We determined if the protein expression levels for these six genes in tumor-adjacent stroma are correlated with clinical variables, such as age, Gleason scores, and stage, in order to explore the potential prognostic power for these biomarkers. There was no association between age and any of the six proteins, indicating the changes of these genes/proteins in tumor adjacent stroma are not due to aging, which is consistent with what has been reported in previous study based on American patient samples[13]. We found that only HSPB1 expression was weakly negatively associated with Gleason score (P value = 0.095) and tumor stage (P value = 0.136). Further validation is needed to prove this correlation. The protein encoded by HSPB1 translocates from the cytoplasm to the nucleus upon stress induction and is involved in stress resistance and actin organization [32].

**DISCUSSION**

Stroma plays an important role in prostate carcinogenesis [33, 34]. It is also a key regulator in prostate function and cancer malignancy [35]. The stroma-epithelium crosstalk in prostate cancer has been intensively investigated in the literature [36]. Through paracrine mechanism, tumor epithelium cells send signals to adjacent stroma which arouses stroma to react by altering expression levels of involved genes/proteins, even if there is little morphological changes in stroma cells. These observations form the basis of continuing studies to identify biomarkers from stroma for disease diagnosis and prognosis [13, 32, 37].

Current diagnosis of presence of prostate tumor based on needle biopsies largely depends on examining the tissue for pathological epithelium. If biopsy samples do not contain recognizable tumor cells, which is very likely based on previous statistics presented [13], false negatives are unfortunately inevitable. However, if changes of genes/proteins in stroma due to the presence of tumor can be identified, instant companion test is possible to fill the void of early detection due to the inefficient biopsy procedure.

Previously, we identified 131 genes that are differentially expressed at the RNA level between tumor adjacent stroma and normal stroma in North American cases (Caucasians and African Americans) [13]. Here, we tested if some of these genes behave in a similar manner in Chinese patients, and extended the study beyond RNA expression to also include protein expression. In this study, we selected six genes, that had additionally been independently reported to be related to prostate cancer progression, CAV1 [19, 20], COL4A2 [21], HSPB1 [22-24], ITGB3 [25], MAP1A [26] and MCAM [27]. We tested these six genes using RT-qPCR based on another collection of 21 Chinese patient cases and 8 Chinese normal cases (Set A). Four genes (COL4A2, HSPB1, ITGB3, and MAP1A) were significantly less expressed in tumor adjacent stroma in comparison with normal stroma (p values < 0.05). These results are consistent with our RNA expression data, based on North American cases[13], indicating that it is possible to use these biomarkers to develop a diagnostic tool that is generic to various races.

By using a tissue microarray (TMA) of 97 tumor cores and 78 associated stroma, and 5 cores from normal donors, we showed that the proteins encoded by all these six genes were down-regulated in tumor-associated stroma of the prostate in Chinese cases (Set B), relative to normal stroma. Three proteins (COL4A2, HSPB1 and ITGB3) showed significant differences (p values < 0.05), and the difference for MCAM was weakly correlated (p value = 0.077).

The incidence of prostate cancer in the United States is significantly higher than in most other countries, particularly Asian countries, even though the incidence of histological lesions (abnormal pathology that is potentially precancerous) has been reported to be similar worldwide [35]. Environmental factors and region-specific diet have therefore been presumed to cause prostate carcinogenesis[36]. Our data indicated that at least one small subset of genes showed a similarly clinical signature between different biological populations. These expression markers have enormous potential for diagnosing the presence of tumor when biopsy samples do not contain recognizable tumors and may be the basis for a companion test that can help avoid unnecessary repeated prostate biopsies and reduce healthcare spending. The current study, aimed to extend this observation to Asian patients and from RNA to protein. These are proof-of-concept studies, and rigorous prospective trials are
needed to confirm the potential application of these genes as diagnostic markers in clinical settings.

MATERIALS AND METHODS

Prostate tissues

The study was approved by the Institutional Review Board at Guizhou Normal College. Two collections of prostate tissue were used in this study, Set A for RT-qPCR validation and Set B for Tissue Microarray assay. The characteristics of all patient cases used in the current study are summarized in Table 2.

Specimens used for RT-qPCR validation (Set A) include frozen samples (-80 °C) of 21 primary prostate cancer tissues (age median = 73, range 55-87) and eight normal prostate tissues (age median = 61.5, range 38-84), selected from the tissue bank at Guangzhou First People’s Hospital, China. Using these samples was approved by the Research Ethics Committee of Guangzhou First People’s Hospital, Guangzhou Medical University, China. Informed consent was obtained from all of the patients. All specimens were handled and made anonymous according to ethical and legal standards. None of the patients or subjects recruited for the study had chemotherapy or radiotherapy before the surgery. The prostate cancer tissues were collected from radical prostatectomy and TURP specimens, and the normal prostate tissues from cystoprostatectomy specimens for bladder cancer. Before RNA processing, all tissues were reconfirmed by HE staining and stroma tissues were collected from these samples using Laser Capture Micro-dissection.

Specimens used for Tissue Microarray assay (Set B) were obtained at the time of initial surgery. Prior consent from patients and approval from the Ethics Committees of Hospital was obtained for using these clinical materials for research purposes. All these specimens had confirmed pathological diagnosis and were classified according to the World Health Organization (WHO) criteria [38].

RNA extraction and RT-qPCR

Stroma tissues were collected from 21 primary prostate cancer tissues and eight normal prostate tissues (Set A) using Laser Capture Micro-dissection. Total RNA of stroma tissues were extracted using RNASimple Total RNA Kit (Cat No. DP419, TianGene, China) and measured by using UV spectrophotometer NanoDrop 2000 (Thermo Scientific, USA). RNA from 21 patient cases and eight normal cases were combined, respectively, and then analyzed using RT-qPCR to examine the mRNA expression levels of the genes of interest. The cDNA templates were synthesized from RNA samples by MMLV. The primer sequence information is given in Supplementary Table S1. Gene expression was determined using SYBR Green Real time PCR Master Mix (Cat No. QPK-201B, TOYOBO, Japan) and 0.2 μg of cDNA template. RT-qPCR was performed on a MyiQ.2 Two-Color RT-qPCR Detection System (Bio-Rad) following below amplification conditions: 5 min, 95°C; followed by 30 cycles of 10 seconds 95°C; 20 seconds 58°C; and 20 seconds 72°C. All assays were carried out by triplicate to control for technical variance. CT-values were determined using the IQ5 software (Bio-Rad). Gene expressions were normalized with GAPDH expression within each sample. Relative quantification of target gene expression was evaluated using the comparative cycle threshold (CT) method.

Tissue microarray assay

Prostate Tissue Microarrays (TMA) were fabricated by Shanghai Outdo Biotech Co., Ltd (Cat. No. HPro-Ade180PG-02). On the TMAs, there are a total of 97 prostate cancer patient cases and five prostate glands from normal subjects (Set B). In addition, 78 out of 97 prostate tumor cases were paired as tumor-bearing tissues and adjacent tumor-free tissues from the same patient. All tissues were re-examined using a microscope by an experienced pathologist after transferred from a local hospital, based on which the pathological indices including Gleason score and stage were given to each patient.

Immunohistochemical staining of formalin-fixed and paraffin-embedded sections was performed using a standard immunohistochemistry (IHC) protocol. Briefly, after deparaffinization and rehydration using a Leica autostainer XL ST5010 system, the TMA slides were pretreated with 10mM sodium citrate buffer (pH 6.0) for 5-10 minutes in a microwave for antigen retrieval. The endogenous peroxidase was quenched by adding the hydrogen peroxide (3% H2O2 in 70% methanol) at room temperature for 15 minutes. After washing, the slides were blocked for 30 minutes. The blocking buffer was removed and the slides were then incubated for 1 hour with primary antibodies (CAV1, COL4A2, HSPB1, ITGB3, MAP1A and MCAM), respectively, with the optimized dilutions instructed (DAB kit, DAKO, Denmark). Slides were counterstained with hematoxylin before microscopic analysis. An H-Score was initially calculated based on scoring of stained cells according to published method [39].
Image analysis

The expressions of each protein in a TMA were measured by analyzing the staining signal intensity using Aperio image scope v11 (Aperio, USA). Briefly, in Aperio Imagescope windows. Epithelial cancer cells and stroma area were compartmentalized by an experienced pathologist using pen tool, based on typical pathological features (Figure 3). The brown staining (positive) in the intensely stained image and the blue staining (negative) in the least intensely stained area were selected for further data processing. The subsequent staining intensity was measured as the densitometry of the digital image (× 400), and the counted positive pixels were transformed to three intensity bins.

The application “Positive Pixel Count V9” of Aperio image scope v11 was used to select areas of interest (stroma component) from each IHC image, and the image data were then translated to numerical data, such as intensities of positive signal, intensities of negative signal, number of positive, number of negative. The average intensity, which is the ratio of the sum of the intensities of positive signal (weak positive, positive and strong positive) and the sum of the number of positive signal (weak positive, positive and strong positive), is calculated and used for further statistical analysis.

Statistical analysis

The Student t test was used to compare the staining intensities between tumor adjacent stroma and normal stroma in IHC assay, as well as the comparison of gene expression levels between tumor adjacent stroma and normal stroma in RT-qPCR assay. P value ≤ 0.05 was used for detecting significant difference in t tests.

Gene ontology software

Using the MetaCore software (GeneGo, Philadelphia, PA), an enrichment analysis was performed to identify significant biological pathways that resulted from our gene list. To limit false discovery and increase biological significance, pathways of interest had to meet the following conventional criteria, i.e., FDR < 5% and p < 0.05.

ACKNOWLEDGMENTS

This work is supported by the Natural Science Foundation of Guizhou Province QKH-J [2013]2238; International Scientific and Technological Cooperation Projects of Guizhou Province QKHW-G [2014]7004; National Natural Science Foundation of China (NO.81170699, 81272813, 81360119); (China Postdoctoral Science Foundation (2014M562157)); Outstanding Youth Science and Technology Talent Cultivating Object of Guizhou Province in 2013 (2013-18); MM and DM were supported by the US National Institute of Health grant U01 CA114810, U01 CA152738.

CONFLICTS OF INTEREST

None of the authors has a conflict of interest to declare except MM and DM, who are the founders of a company. Proveri, Inc, which has an interest in biomarkers for Prostate Cancer, including diagnosis and prognosis.

REFERENCES


