Blackwell Publishing

Asian J Androl 2006; 8 (1): 45–51
DOI: 10.1111/j.1745-7262.2006.00103.x

AJA

·Original Article·

Application of surface-enhanced laser desorption/ionization time-of-flight-based serum proteomic array technique for the early diagnosis of prostate cancer

Yu-Zhuo Pan[1†], Xue-Yuan Xiao[2†], Dan Zhao[1], Ling Zhang[1], Guo-Yi Ji[1], Yang Li[1], Bao-Xue Yang[1], Da-Cheng He[2‡], Xue-Jian Zhao[1‡]

[1]*Research Center of Prostate Diseases, Department of Reproductive Pathophysiology, School of Basic Medicine, Jilin University, Changchun 130021, China*
[2]*Research Institute of Cytobiology of the Academy of Life Science, Research Institute of Proteomics, Normal University, Beijing, 100875, China*

Abstract

Aim: To identify the serum biomarkers of prostate cancer (PCa) by protein chip and bioinformatics. **Methods:** Serum samples from 83 PCa patients and 95 healthy men were taken from a mass screening in Changchun, China. Protein profiling was carried out using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS). The data of spectra were analyzed using two bioinformatics tools. **Results:** Eighteen serum differential proteins were identified in the PCa group compared with the control group ($P < 0.01$). There were four proteins at the higher serum level and 14 proteins at the lower serum level in the PCa group. A decision tree classification algorithm that used an eight-protein mass pattern was developed to correctly classify the samples. A sensitivity of 92.0 % and a specificity of 96.7 % for the study group were obtained by comparing the PCa and control groups. **Conclusion:** We identified new serum biomarkers of PCa. SELDI-TOF MS coupled with a decision tree classification algorithm will provide a highly accurate and innovative approach for the early diagnosis of PCa. *(Asian J Androl 2006 Jan; 8: 45–51)*

Keywords: prostate cancer; early diagnosis; protein chip; biomarker; serum

Correspondence to: Dr Xue-Jian Zhao, Prostate Diseases Prevention and Treatment Research Center, Jilin University, Changchun 130021, China.
Tel/Fax: +86-431-563-2348
E-mail: pro_2@jlu.edu.cn
Dr Da-Cheng He, Research Institute of Cytobiology of Academy of Life Science, Research Institute of Proteomics, Normal University, Beijing 100875, China
Tel: +86-10-5880-9729, Fax: +86-10-5880-5042
E-mail: dhe@bnu.edu.cn
†These authors contributed equally to this work.
‡Corresponding authors who contributed equally to this work according to the cooperating protocol.
Received 2005-01-19 Accepted 2005-10-14

1 Introduction

Prostate cancer (PCa) is one of the main forms of cancer affecting old men worldwide. In the USA, in 2003, PCa had the highest cancer incidence with 220 900 new cases and the second-highest cancer-related mortality rate, with 28 900 deaths [1]. The retrospective research of Gu *et al*. [2] showed that PCa incidence in Beijing, in the 1990s was five times higher than that in the 1950s. We retrospectively investigated all PCa cases from eight hospitals in Changchun, China from 1986 to 2001, and found the number of cases had increased 4.6-fold in the 1999–

2001 period *vs*. the 1986–1989 period [3]. We also organized a mass screening in Changchun for PCa in men aged 50 years or older with routine serum prostate-specific antigen (PSA) tests. The detection rate of PSA was greater than 1.7 %, and 18.8 % of the cases involved osseous metastasis [4, 5].

Currently, the serum PSA test for mass screening for PCa appears to be controversial. Recently, Stamey *et al.* [6] announced that the serum PSA test as the standard detection test for PCa was out of date in the USA, based on 20 years of experience; in our experience, the biomarker still plays an important role in the mass screening of PCa in China. However, with the rapid emergence of proteomics, it is appealing to many investigators to explore new and more effective specific proteins to accurately detect PCa. The use of surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS), for example, requires less labor, and has a high throughput and excellent reproducibility. Initial researches [7–9] on the identification of biomarkers for cancers using this technique have been very promising, which has encouraged us to further verify the availability of MS. This technique has been successfully applied for exploring new protein markers for early detection of PCa in the USA. However, the proteomic approach has not yet been used to identify the protein markers of PCa in a cohort of Chinese men.

In this study, we analyzed 178 serum samples using SELDI-TOF MS to explore the marker proteins for the detection of PCa for the first time in China.

2 Materials and methods

2.1 Serum samples

One hundred and seventy-eight serum samples were randomly taken from a serum bank in the Research Center of Prostate Diseases at Jilin University (Changchun, China) and were kept at –70 ºC. Of the 178 samples, 95 were taken from healthy men and 83 from PCa patients, respectively, confirmed by PSA test and histology examination through prostate needle biopsy. The healthy control (n = 95, age range 52-81 years, mean 63 years) and PCa (n = 83) groups were age-matched. The PCa group was subsequently divided into the organ-confined PCa group (n = 45; T1/T2; age range 50–89 years, mean 67 years) and the non-organ-confined PCa group (n = 37; T3/T4; age range 48–91 years, mean 70 years). The PSA ranges were 0.01–1.00 ng/mL (mean 0.59 ng/mL) for the control, 0.00–22.76 ng/mL (mean 9.62 ng/mL) for the organ-confined PCa group, and 0.00–108.00 ng/mL (mean 77.21 ng/mL) for the non-organ-confined PCa group. The study was conducted in accordance with the ethics of the Declaration of Helsinki and with the approval of the Jilin University Bethune Medical Services Ethics Committee.

2.2 SELDI-TOF protein profiling

Immobilized metal affinity capture array (IMAC)-Cu metal binding chips (Ciphergen Biosystems, Fremont, CA, USA) were used and put into a bioprocessor (Ciphergen Biosystems, Fremont, CA, USA), a device that holds eight chips and allows for the application of large volume of serum to each chip array. The chips were coated with 50 μL of 100 mmol/L $CuSO_4$ on each array and agitated for 5 min at room temperature. The chips were rinsed three times with deionized H_2O, then 150 μL of 100 mmol/L sodium acetate buffer (pH 4.0) was added to each array and shaken for 5 min to remove the unbound copper. Serum samples for SELDI-TOF analysis were prepared by vortexing 5 μL of serum with 10 μL of 8 mol/L urea and 1 % 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonic acid in phosphate-buffered saline (PBS) in a tube at 4 ºC for 30 min. Ten microliters of the serum/urea mixture was applied to each well with 90 μL PBS, and the bioprocessor was sealed and shaken on a platform shaker for 30 min. The serum/urea mixture was discarded, and the chips were washed twice, 5 min each wash cycle, in 150 μL of 100 mmol/L sodium acetate buffer as described above. The chips were removed from the bioprocessor, washed twice with deionized H_2O, air-dried, and stored in the dark at room temperature until subjected to SELDI-TOF analysis.

Before SELDI-TOF MS analysis, 0.5 μL of saturated solution of sinapinic acid (Fluka, Milwaukee, WI, USA) in 50 % (v/v) acetonitrile, and 0.5 % trifluoroacetic acid was applied onto each spot twice and air-dried between applications. The chips were then placed in the PBS-II SELDI-TOF MS (Ciphergen Biosystems, Fremont, CA, USA) operated in the positive ion mode. Time-of-flight spectra were generated by averaging 90 laser shots collected on each spot with a laser intensity setting of 195, detector sensitivity setting of 9, and a lag time focusing of 900 ns. The spectra were calibrated using the All-in-1 protein molecular mass standard (Ciphergen Biosystems, Fremont, CA, USA). To compensate for slight spot-to-spot variations, if any, the spectra were also normalized

using the total ion current method in the mass to charge (*m*/*z*) range of 1 500–30 000 (IMAC-Cu surface) with subtracted baseline. The reproducibility of the SELDI-TOF system was determined using the pooled normal serum quality control sample.

Peaks were identified after mass calibration, background subtraction and normalization using the clustering and alignment function of ProteinChip Biomarker Wizard software 3.1 (Ciphergen Biosystems, Fremont, CA, USA). The settings used were as follows: signal/noise ratio in the first pass: 3; minimum peak threshold: 5 %; cluster mass window: 0.3 %; and signal/noise in the second pass: 1.5. The peaks detected within ± 0.15 % *m*/*z* units of each other across the spectra were considered one cluster, and a particular cluster was represented by its average *m*/*z* value. The peak information, including *m*/*z* and intensity values, was exported into ProteinChip Biomarker Pattern 4.0 (Ciphergen Biosystems, Fremont, CA, USA) for statistical analysis. The reproducibility of the SELDI-TOF system was determined, through which each chip included a control sample (aliquoted from a single collection from each individual), permitting estimation of assay variability and reproducibility.

2.3 Bioinformatics and biostatistics

Construction of the classification tree was accomplished with the 308 peaks per spectrum by the ProteinChip Biomarker Wizard software (Ciphergen Biosystems, Fremont, CA, USA). Classification trees split up a dataset into two bins or nodes, using one rule once in the form of a question. The splitting decision was defined by the presence or absence and the intensity levels of one peak. The splitting process continued till terminal nodes or leaves were produced or further splitting had no gain. The classification of terminal nodes was determined by the group ("class") of samples (i.e. PCa or control) representing the majority of samples in the corresponding node. The data from the Biomarker Wizard software were analyzed with the Biomarker Pattern software based on the Gini clustering classification tree (sensitivity: 0.05; method: no independent testing–exploratory tree).

Comparisons between groups were performed by analysis of variance with the Biomarker Wizard software. Statistical significance was defined as $P < 0.05$.

Variable importance scores reflect the contribution of each variable makes in classifying or predicting the target variable, with the contribution stemming from both the variable's role as a primary splitter and its role as a surrogate to any of the primary splitters. The variable used to split the root node is ranked as the most important. This variable received a zero score, indicating that it did not play any role in the analysis as either primary splitters or surrogates.

3 Results

3.1 Differences in protein mass spectrum

The SELDI-TOF protein profiles of the PCa and control groups are shown in Figure 1. The SELDI-TOF technology is particularly effective in resolving low molecular weight (< 10 kDa). Figure 1A is a representative protein spectrum showing the protein masses between 2 000 kDa and 20 000 kDa of a single serum specimen. The SELDI-TOF Biomarker Wizard program detected 308 peaks per spectrum for each corresponding sample. Eighteen serum differential proteins were clearly identified in the PCa group compared with the control group ($P < 0.01$). However, the use of each single peak from the 18 proteins could not completely differentiate the PCa group from the control group. The 18 proteins from the PCa

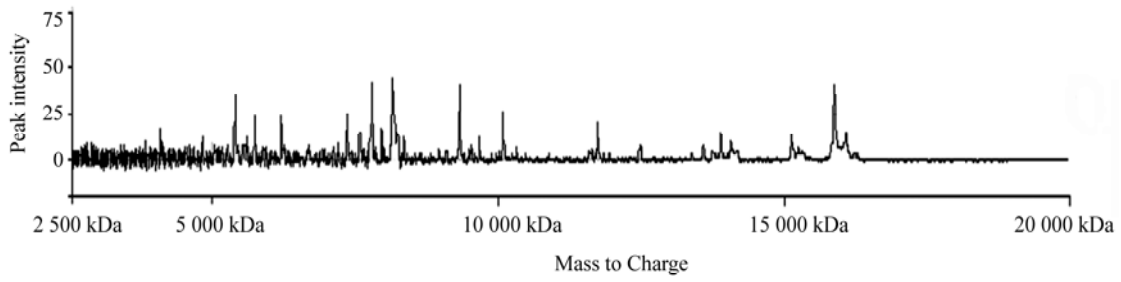

Figure 1. Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SEIDI-TOF MS) serum spectrum obtained using immobilized metal affinity capture (IMAC)-Cu protein array. The spectrum represents the overview of protein profiling on IMAC-Cu chips of serum from one sample (P3). The x-axis represents the molecular mass to charge from 2 500 kDa to 20 000 kDa; the y-axis represents relative peak intensity.

group included 14 proteins in low abundance and 4 proteins in high abundance (Table 1).

3.2 Construction and analysis of taxonomic tree

As Figure 2 shows, the classification tree that was reflected through eight masses (M1 669, M4 300, M5 923, M6 250, M6 652, M7 782, M15 868 and M145 875) calculated by the Biomarker Pattern program generated nine terminal nodes (Figure 2). Of the eight masses, six (M4 300, M5 923, M6 250, M6 652, M7 782 and M15 868) were identified as being consistent with those in the corresponding 18 differential proteins for PCa (Table 1). Figure 3 shows the spectrum and gel maps representing the six masses. As a result of their absence among the 18 differential proteins, the remaining two masses (M1 669 and M145 875) are not shown. The split criteria from Node 1 to Node 8 (7.66 [M5 923], 2.535 [M15 868], 24.46 [M7 782], 9.852 [M6 250], 0.1 [M145 875], 1.66 [M4 300], 4.303 [M6 652] and 0.854 [M1 669]) directed the nine terminal nodes. The terminal nodes 1, 3, 4, 6 and 8 represented PCa (Figure 2). The remaining terminal nodes were

Table 1. Serum biomarker proteins in prostate cancer group. " ↓ "means mass is low abundance; " ↑ "means mass is high abundance.

Mass (Da)	Serum biomarker level
3 700	↓
3 904	↓
3 978	↓
4 300	↓
5 085	↓
5 923	↓
6 250	↓
6 569	↓
6 652	↓
7 121	↓
7 782	↓
7 850	↓
7 991	↓
9 503	↓
15 265	↑
15 878	↑
16 003	↑
16 068	↑

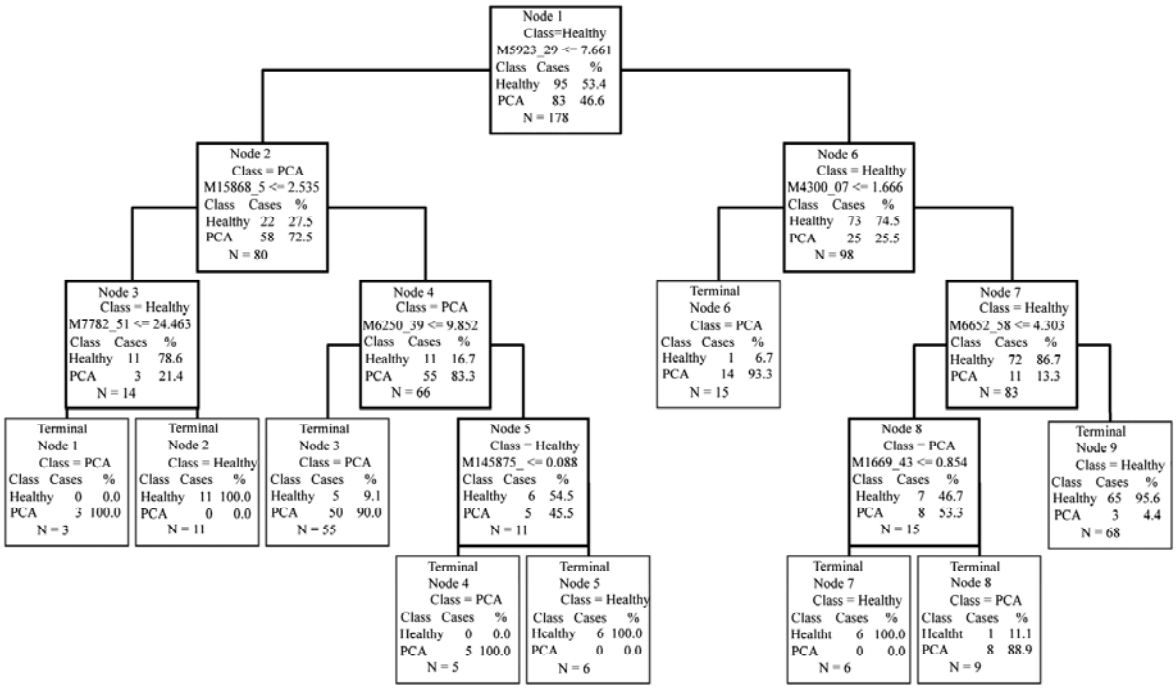

Figure 2. Classification tree structure and sample distribution. A diagram of decision tree analyses shows the root node (top), the descendant nodes and the terminal nodes (Terminal Node 1–Terminal Node 10) as rectangles. The numbers in each node represent the classes (top number, number of control [CONT] samples; bottom number, number of prostate cancer samples). The first number under the root and descendant nodes is the mass value followed by the peak intensity value. For example, the mass value under the root node is 5 923.29 kDa, and the intensity is 7.661.

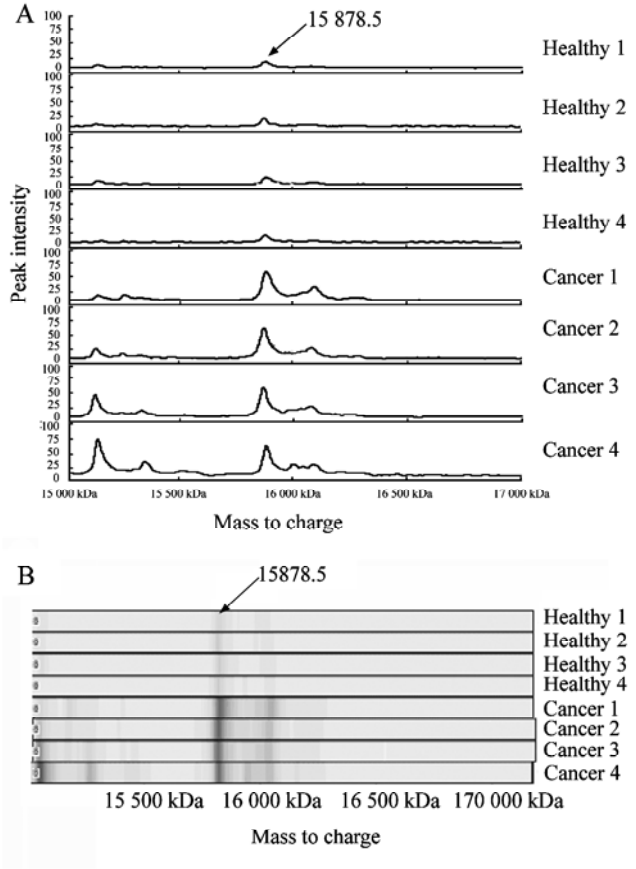

Figure 3. Representative surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) serum spectra of 15 878. Differential expression of SELDI peak 15 878 using immobilized metal affinity capture (IMAC)-Cu in comparison of the control (CONT) group (top 4) and prostate cancer (PCa) group (bottom 4). (A): spectra of the peak with the molecular mass to charge (*m/z*) of 15 878.5 kDa. A protein with mass 15 878 kDa was of high abundance in the PCa group compared with the CONT group by spectrum; (B): gel maps of the peak with *m/z* of 15 878.5 kDa.

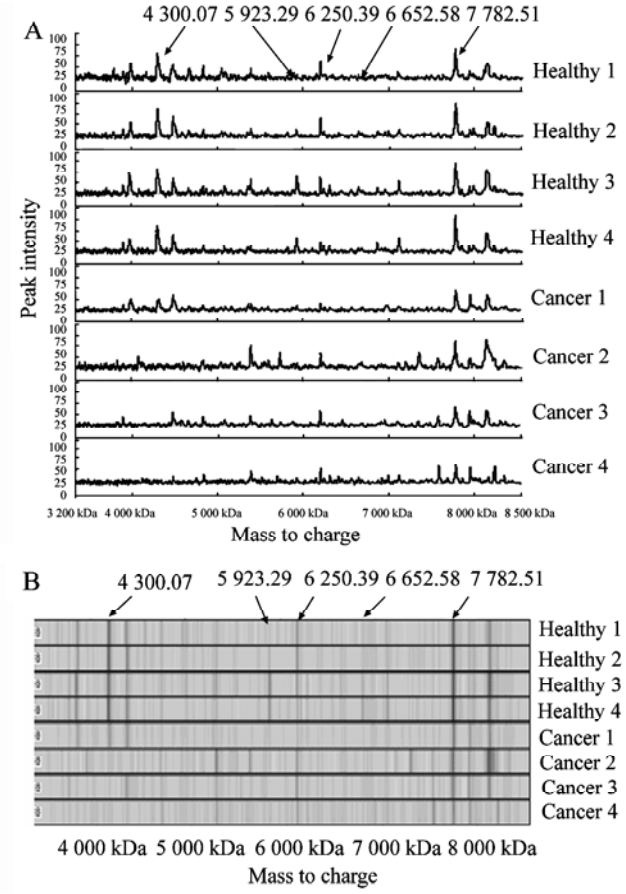

Figure 4. Representative surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) serum spectra of 4 300.07 kDa, 5 923.29 kDa, 6 250.39 kDa, 6 652.58 kDa and 7 782.51 kDa. The average molecular mass of the five proteins identified to be unique or overexpressed in the normal control specimens are: 4 300.07 kDa, 5 923.29 kDa, 6 250.39 kDa, 6 652.58 kDa and 7 782.51 kDa. Differential expression of SELDI peak using immobilized metal affinity capture (IMAC)-Cu to compare the control (CONT) group (top 4) and prostate cancer (PCa) group (bottom 4). (A): spectra of the peak with the molecular mass to charge (*m/z*) of 4 300.07 kDa, 5 923.29 kDa, 6 250.39 kDa, 6 652.58 kDa and 7 782.51 kDa. The proteins with 4 300.07 kDa, 5 923.29 kDa, 6 250.39 kDa, 6 652.58 kDa and 7 782.51 kDa were of high abundance in the PCa group compared with CONT group by spectrum; (B): gel maps of the peak with the *m/z* of 4 300.07 kDa, 5 923.29 kDa, 6 250.39 kDa, 6 652.58 kDa and 7 782.51 kDa.

normal (Figure 2). The diagnosis sensitivity and specifity for PCa were 92.0 % and 96.7 %, respectively. The variable importance scores of the eight masses and the corresponding diagnosis sensibility and speciality of each mass are summarized in Table 2. Node 1 (M5 923) had the highest importance scores.

4 Discussion

PSA is a tissue-specific antigen of the prostate. The serum PSA test has been used as the standard approach for the detection of PCa for many years. Such a simple approach has been successful for mass screening of PCa [9]. A PSA level of 4.0–10.0 ng/mL is a significant indication for PCa in the early period of the disease, when clinical treatments are optimal [10]. However, many cases of non-cancer diseases are covered in this interval, which

Table 2. Node variable importance scores of eight protein biomarkers.

Node	Mass (kDa)	Sensitivity (%)	Specificity (%)	Variable importance scores
1	5 923	56.2	61.3	100.0
2	15 878	43.5	70.1	83.6
3	4 300	42.8	52.4	85.4
4	7 782	35.6	36.2	82.4
5	6 250	28.9	40.3	17.6
6	6 652	34.8	39.5	33.5
7	145 875	15.6	2.4	28.1
8	1 669	12.3	34.7	29.4

influences the differentiation between non-cancer diseases and PCa. In order to overcome such drawbacks, researchers have attempted to make use of new parameters, such as PSA density and free PSA/total PSA, to replace the original PSA test [8]. However, these attempts are just adjuvant for the optimization of the PSA test and cannot resolve the essentially low specificity of PSA. Most studies have shown that the specificity of the PSA test is 25 % [8]. Stamey *et al*. [6] pointed out that it is necessary to explore more specific proteins and more efficient methods with the gradual disappearance of the predominance of the PSA approach.

Proteomics techniques, especially protein array, have been widely applied in the field of searching biomarkers for early cancer detection [7, 11–15]. Adam *et al*. [7] reported promising results with sensitivity of 83 % and specificity of 97 % using the SELDI-TOF pattern for serum biomarkers of PCa detection. Petricoin *et al.* [14] and Qu *et al*. [16] reported a 95 % sensitivity at 78–83 % specificity and a 97–100 % sensitivity at 97–100 % specificity, respectively. These data indicate that the proteomic technique for PCa detection is superior to the conventional PSA technique in terms of sensitivity.

Adam *et al*. [7] used nine masses at *m/z* ratios of 4 475, 5 074, 5 382, 7 024, 7 820, 8 141, 9 149, 9 507 and 9 656, whereas Petricoin *et al*. [14] selected seven masses at *m/z* ratios of 2 092, 2 367, 2 582, 3 080, 4 819, 5 439 and 18 220. Qu *et al*. [16] identified 12 major masses at *m/z* ratios of 9 656, 9 720, 6 542, 6 797, 6 949, 7 024, 8 067, 8 356, 3 963, 4 080, 7 885 and 6 991 in differentiating non-cancer from cancer, and nine masses at *m/z* ratios of 7 820, 4 580, 7 844, 4 071, 7 054, 5 298, 3 486, 6 099 and 8 943 in differentiating healthy individuals from patients with benign prostatic hyperplasia (BPH). Of note is that, in addition to masses at *m/z* ratios of 7 820 identified by Petricoin *et al*. [14] and Qu *et al*. [16], no cross-talk between those identified by Petricoin *et al.* [14], Adam *et al*. [7] or Qu *et al*. [16] was found. Furthermore, although Adam *et al*. [7] and Qu *et al*. [16] used the same chip for serum extraction and the same instrument for peak identification, their distinguishing peaks are clearly different. In this study, we found that only three masses at *m/z* ratios of 5 074, 7 850 and 9 507, as shown in Table 1, are matched with that reported by Adam *et al*. [7]. We also found that the peak mass at an *m/z* ratio of 7 850 is much closer to the peak mass of 7 820 identified by Adam *et al.* [7] and Qu *et al*.[16].

The results from the study of Adam *et al*. [7] demonstrated that proteomic array is an optimal modality for the diagnosis of PCa. In our study, the IMAC-Cu chip was adopted to collect serum proteins from 178 samples. We identified 18 differential expression proteins for the PCa group, including 14 lower-expression proteins and 4 higher-expression proteins. Through further classification using Biomarker Pattern software, we developed a taxonomic tree with eight protein biomarkers and confirmed that a combination of these eight proteins could accurately screen PCa patients, with a sensitivity of 93 % and specificity of 96 %. This method is superior to using each biomarker alone, such as the 5 923 kDa protein with a sensitivity of only 56.2 %. In terms of specificity, our data collected from a cohort of Chinese PCa patients are comparable to results reported by others [7, 12, 14].

In the present study, our results demonstrated that the use of cluster analysis for 308 peaks per spectrum could construct the optimal classification tree model. We found eight biomarkers, six of which were located at superior layers of the taxonomic tree with statistical significance in discriminating between the PCa and normal groups. Deletion of the other two proteins results in decreased sensitivity and specificity of the decision tree to 80 % and 81 %, respectively, indicating the impor-

tance of the combined use of multiple marker proteins. Our data showed that the correct resolution of a single protein with molecular mass 5 923.29 kDa was just 56.23 %. This might explain the low specificity of the PSA test. Our data demonstrated that the SELDI-TOF pattern facilitated the exploration of various proteins simultaneously. The cluster analysis of these proteins could identify special proteins which may be not useful alone, but may be beneficial when incorporated with other proteins for the diagnosis of malignancies. The model of a combination of multiple biomarkers is superior to that of a single biomarker. Other researches [7, 11, 12, 14] have also shown the advantage of multiple biomarkers in different cancers.

In conclusion, this study represents the first demonstration that the SELDI-TOF-based serum proteomic array technique is effective for the diagnosis of PCa in Chinese men. Such an approach is useful for the research of other cancers requiring biomarker analysis.

Acknowledgment

This study was supported by the Special Technical Cooperation Items between the Chinese and Japanese Governments (59th item).

References

1 Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. Cancer statistics, 2003. CA Cancer J Clin 2003; 53: 5–26.
2 Gu FL. Changing constituents of genitourinary cancer in recent 50 years in Beijing. Chin Med J (Engl) 2003; 116: 1391–3. Erratum in: Chin Med J (Engl). 2003; 116:1643.
3 Li XM, Zhang L, Li J, Li Y, Wang HL, Ji GY, *et al*. Measurement of serum zinc improves prostate cancer detection efficiency in patients with PSA levels between 4ng/mL and 10 ng/mL. Asian J Androl. 2005; 7:323-8.
4 Zhang HF, Wang HL, Xu N, Li SW, Ji GY, Li XM, *et al*. Mass screening of 12,027 elderly men for prostate carcinoma by measuring serum prostate specific antigen. Chin Med J (Engl) 2004; 117: 67–70.
5 Kuwahara M, Tochigi T, Kawamura S, Ogata Y, Xu N, Wang H, *et al*. Mass screening for prostate cancer: a comparative study in Natori, Japan and Changchun, China. Urology 2003; 61: 137–41.
6 Stamey TA, Caldwell M, McNeal JE, Nolley R, Hemenez M, Downs J. The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? J Urol 2004; 172: 1297–301.
7 Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, *et al*. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. Cancer Res 2002; 62: 3609–14.
8 Zhang L, Ji G, Li X, Wang W, Gao H, Pan Y, *et al*. Early diagnosis of prostate cancer using free/total prostate-specific antigen ratio with population-based screening data. Zhonghua Nan Ke Xue 2004; 10: 582–5.
9 Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. N Engl J Med 1987; 317: 909–16.
10 Stamey TA, Kabalin JN, McNeal JE, Johnstone IM, Freiha F, Redwine EA, *et al*. Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. II. Radical prostatectomy treated patients. J Urol 1989; 141: 1076–83.
11 Xiao X, Liu D, Tang Y, Guo F, Xia L, Liu J, He D. Development of proteomic patterns for detecting lung cancer. Dis Markers 2003; 19: 33–9.
12 Xiao XY, Zhao X, Liu J, Guo F, Liu D,He DC. Discovery of laryngeal carcinoma by serum proteomic pattern analysis. Sci China C Life Sci, 2004, 47: 219–23.
13 Petricoin EF 3rd, Ornstein DK, Paweletz CP, Ardekani A, Hackett PS, Hitt BA, *et al*. Serum proteomic patterns for detection of prostate cancer. J Natl Cancer Inst 2002, 94: 1576–8.
14 Petricoin EF, Ornstein DK, Liotta LA. Clinical proteomics: applications for prostate cancer biomarker discovery and detection. Urol Oncol 2004; 22: 322–8.
15 Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, *et al*. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002; 359: 572–7.
16 Qu Y, Adam BL, Yasui Y, Ward MD, Cazares LH, Schellhammer PF, *et al*. Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. Clin Chem 2002; 48:1835–43.