Int J Clin Exp Med 2015;8(2):2077-2085
www.ijcem.com /ISSN:1940-5901/IJCEM0003180

Original Article

Comparative proteomic analysis of serum diagnosis patterns of sputum smear-positive pulmonary tuberculosis based on magnetic bead separation and mass spectrometry analysis

Jiyan Liu[1], Tingting Jiang[2], Feng Jiang[3], Dandan Xu[2], Liliang Wei[4], Chong Wang[2], Zhongliang Chen[2], Xing Zhang[2], Jicheng Li[2]

[1]College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, P. R. China; [2]Institute of Cell Biology, Zhejiang University, Hangzhou, P. R. China; [3]Dongzhimen Hospital Affiliated to Beijing University of Chinese Medicine, Beijing, P. R. China; [4]Department of Respiratory Medicine, The Sixth Hospital of Shaoxing, Shaoxing, P. R. China

Received October 15, 2014; Accepted January 21, 2015; Epub February 15, 2015; Published February 28, 2015

Abstract: A major challenge in pulmonary tuberculosis (TB) control is early and accurate diagnosis of sputum smear negative pulmonary TB (SSN-PTB). The patients with SSN-PTB have to wait for a longer period of time before receiving proper treatment than sputum smear positive pulmonary TB (SSP-PTB) patients due to delay in diagnosis. The purpose of this study is to discover potential serum protein biomarkers for SSN-PTB. Surface-enhanced laser desorption/ionization-time of flight mass spectrometry (SELDI-TOF MS) combined with weak cation exchange (WCX) magnetic beads was used to screen serum samples from SSN-PTB patients (N = 66), SSP-PTB patients (N = 49), and healthy volunteers (N = 80). The serum protein profiles were analyzed with Biomarker Wizard system. A classification model was established using Biomarker Pattern Software (BPS). Fifty-eight protein peaks were identified to exhibit significant differences between SSN-PTB, SSP-PTB and healthy control groups ($P < 0.05$), among which 6 peaks were found to be down-regulated, while 10 peaks were up-regulated gradually in the healthy control, SSN-PTB, and SSP-PTB groups. Twenty-three discriminating m/z peaks were detected between SSN-PTB patients and healthy controls ($P < 0.01$, Fold ≥ 1.5). The classification tree combined with three protein peaks (2747.0, 4480.0, and 9410.1 Da) could distinguish SSN-PTB patients from healthy controls with a sensitivity of 83.33% and a specificity of 82.50%. Early diagnosis of SSN-PTB disease is critical in order to reduce morbidity and mortality associated with TB. The study will help to clarify the role of differential proteins in the pathogenesis of TB.

Keywords: Proteomic, Biomarker, surface-enhanced laser desorption/ionization, sputum smear negative pulmonary tuberculosis

Introduction

Tuberculosis (TB) is a global public health problem, and is ranked as the second leading cause of death among infectious diseases. In 2012, an estimated 8.6 million people developed TB and 1.3 million died from the disease. The global burden of TB is concentrated in developing countries. China ranks second among the 22 high TB burden countries worldwide. According to the World Health Organization (WHO) report, in 2012, there were an estimated 1.4 million prevalent cases and 1.0 million newly diagnosed cases of TB in China [1].

Early and accurate diagnosis is important for controlling and preventing TB [2]. Pulmonary TB (PTB), the most common form of TB, is diagnosed by detecting *Mycobacterium tuberculosis* complex (MTBC) bacilli in samples of sputum expectorated by the patient. Sputum culture is the gold standard for diagnosis of PTB. However, the slow growth of the bacteria can lead to delay in diagnosis and medical intervention. Culture tests take as long as 2-6 weeks to produce results [3]. The microscopic examination of sputum for acid-fast bacilli (AFB) is a simple and rapid diagnostic test for TB. However, it exhibits low sensitivity and requires

at least 5 × 10^3 bacilli per ml of sputum [4]. Some patients presenting with active pulmonary TB may exhibit negative sputum AFB smears. According to the Global TB report, there were an estimated 2.5 million sputum smear positive pulmonary TB (SSP-PTB) patients, and 1.9 million sputum smear negative pulmonary TB (SSN-PTB) patients in 2013 [1]. Patients with SSN-PTB are also capable of transmitting the infection [5]. The appropriate treatment of patients with SSN-PTB is often delayed. Therefore, there is a need for early diagnosis of SSN-PTB disease [6].

Development of high throughput proteomic technology provides a new pathway to large-scale screening and identification of biomarkers in serum [7, 8]. Proteomic is defined as the systematic, comprehensive and quantitative analysis of the proteins present in a biological sample at a defined time point [9]. Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) can detect proteins with low molecular weights and is considered as a powerful proteomic technology for serum protein profiling [10, 11].

This study sought to screen differentially expressed proteins in serum of SSN-PTB patients. The study also explores the potential biomarkers for the early diagnosis of SSN-PTB.

Materials and methods

Patients and controls

In this study, we tested a total of 195 human serum samples from 115 patients with active PTB and 80 healthy volunteers. The patients with active PTB were recruited randomly from the Sixth Hospital of Shaoxing (Shaoxing, China) and Hangzhou Red Cross Hospital (Hangzhou, China). All PTB patients were diagnosed according to criteria from the WHO [12], including clinical, radiological and histopathological analysis. The patients with hepatic, renal, metabolic and autoimmune disorders, endocrine, blood, nervous system diseases, malignant tumors, and long-term use of immunosuppressive agents were not included in the study. According to the acid-fast staining of bacilli in sputum smear, the PTB patients were divided into two groups: the SSN-PTB group (N = 66) and the SSP-PTB group (N = 49). All the blood samples were collected and preserved upon the first visit and before any treatment. Eighty sex-matched controls were recruited from healthy population who came to the hospital for regular health examination. Both patients and controls were from the same ethnic (Han) population and lived in the same region (Southeast China).

The study was conducted according to the principles expressed in the Declaration of Helsinki and was approved by the Faculty of Medicine (Zhejiang University, China). Informed consent was obtained from all subjects prior to collection of blood. The peripheral blood samples were collected from the PTB patients and the healthy controls in early morning without anticoagulation. Then the blood samples were allowed to clot for 1-2 h prior to 4,000 g centrifugation for 10 min at 4°C to separate the serum out. The serum samples were aliquoted and stored at -80°C for further analysis.

SELDI-TOF MS analysis combined with WCX magnetic beads

The subjects and serum samples were randomly divided into two groups: the training set and the testing set (**Table 1**). The training set was used to detect discriminating peaks and construct the classification tree of SSN-PTB. The discriminatory ability of the classification algorithm was then challenged with the testing set. The serum samples were analyzed according to the standard protocol [13]. Briefly, WCX magnetic beads (Beijing SED Science & Technology, China) were pre-activated with binding buffer in a magnetic separator. Each serum sample was initially denatured with U9 solution at 4°C. Denatured serum samples were further diluted 1:40 in binding buffer. Then, the diluted serum sample was added to the activated magnetic beads, and incubated for 1 h at 4°C, after which the beads were washed twice with binding buffer to remove non-selectively bound proteins. Following binding and washing, the bound proteins were eluted from the magnetic beads using 10 μL of 0.5% trifluoroacetic acid. Then, 5 μL of the eluted sample was diluted 1:2-fold in 5 μL of SPA. Next, 2 μL of the resulting mixture was aspirated and spotted onto an 8-spot pre-structured sample Au-chip.

After air drying, protein crystals on the chip were scanned with the ProteinChip reader (model PBS IIc) (Ciphergen Biosystems, USA) to determine the masses and intensities of all

Table 1. Characteristics of the study population

	Training set	Testing set	Statistics
SSP-PTB			
Total number of patients	32	17	-
Years range, age (median ± SD)	20-65 (44.7 ± 11.8)	21-63 (44.5 ± 12.4)	$t = 0.056$, $P = 0.956$[a]
Sex (female : male)	13:19	8:9	$\chi^2 = 0.188$, $P = 0.665$[b]
Abnormal chest radiograph	32 (100%)	17 (100%)	-
Sputum culture-positive	32 (100%)	17 (100%)	-
BCG vaccination	15 (46.9%)	8 (47.1%)	$\chi^2 = 1.51\times10^{-4}$, $P = 0.990$[b]
HIV-negative	32 (100%)	17 (100%)	-
SSN-PTB			
Total number of patients	42	24	$t = 0.568$, $P = 0.572$[a]
Years range, age (median ± SD)	18-64 (42.5 ± 12.2)	19-62 (44.3 ± 12.7)	$\chi^2 = 0.055$, $P = 0.815$[b]
Sex (female : male)	18:24	11:13	-
Abnormal chest radiograph	42 (100%)	24 (100%)	-
Sputum culture-positive	42 (100%)	24 (100%)	$\chi^2 = 0.020$, $P = 0.889$[b]
BCG vaccination	20 (47.6%)	11 (45.8%)	-
HIV-negative	42 (100%)	24 (100%)	
Healthy volunteers			
Total number of control subjects	40	40	$t = 0.449$, $P = 0.655$[a]
Years range, age (median ± SD)	18-65 (43.6 ± 12.8)	19-62 (44.9 ± 13.1)	$\chi^2 = 0.051$, $P = 0.822$[b]
Sex (female : male)	17:23	18:22	$\chi^2 = 0.202$, $P = 0.653$[b]
BCG vaccination	19 (47.5%)	17 (42.5%)	-
HIV-negative	40 (100%)	40 (100%)	

SSP-PTB: Sputum smear positive pulmonary tuberculosis; SSN-PTB: Smear negative pulmonary tuberculosis; BCG: Bacillus Calmette-Guérin; HIV: Human Immunodeficiency Virus. [a]Difference between the training set and the testing set (*t*-test). [b]Difference between the training set and the testing set (χ^2 test).

peaks. The reader was set up as follows: mass range was set from 1,000 to 50,000 Da, optimized mass range was set from 1,000 to 15,000 Da, laser intensity was set at 265, and laser sensitivity was set at 7.

In the sample pre-treatment and proteomic analysis process, the serum samples from the patients and control groups were randomized, and the investigator was blinded to their clinical manifestations. Strict standard operating procedures, internal and external control were combined for data quality and reproducibility. In internal control method, one point was randomly selected for each Au chip to perform the same experiment with quality control serum. The "All-in-one protein standard II" (Bio-Rad, USA) was used as the external control to obtain protein standard spectra for mass accuracy calibration. All PTB patients and healthy control samples were detected by SELDI-TOF MS using the same batch of magnetic beads, the same Au-chip and on the same equipment. The same procedures were followed within one week to ensure experimental repeatability and reliability.

Statistical analysis

The profiling spectra of serum samples from the training set were normalized using total ion current normalization by Ciphergen ProteinChip Software (version 3.1). Peak labeling was performed by Biomarker Wizard software, version 3.1 (Ciphergen Biosystems). Comparisons between the three groups were tested by one-way ANOVA and least significant difference (LSD) test using SPSS software version 16.0. *P* value < 0.05 was considered as statistically significant. Proteins of the SSN-PTB and healthy control groups with low *P*-values were selected, and the intensities of the selected peaks were transferred to Biomarker Pattern Software (BPS, Ciphergen Biosystems) to construct the classification tree of SSN-PTB.

The tree construction process primarily splits spectral data of the training set into two nodes,

Table 2. Gradient difference peaks from healthy controls to SSN-PTB to SSP-PTB

m/z	HV (Mean ± SE)	SSN-PTB (Mean ± SE)	SSP-PTB (Mean ± SE)
Down-regulated peaks			
4480.0	10.85 ± 0.51[a]	6.34 ± 0.41[b]	4.69 ± 0.41[c]
6437.0	21.01 ± 1.13[a]	12.70 ± 1.15[b]	9.90 ± 0.76[b]
2747.0	23.43 ± 1.79[a]	14.39 ± 2.02[b]	11,73 ± 1.20[b]
3321.2	8.32 ± 0.47[a]	5.38 ± 0.49[b]	4.53 ± 0.43[b]
9192.8	23.36 ± 4.24[a]	12.88 ± 3.78[b]	11.31 ± 2.72[b]
1625.7	5.12 ± 0.59[a]	3.61 ± 0.67[a,b]	2.57 ± 0.56[b]
Up-regulated peaks			
8607.5	1.86 ± 0.13[a]	4.74 ± 0.40[b]	4.93 ± 0.94[b]
5876.4	2.45 ± 0.16[a]	3.36 ± 0.48[a]	5.69 ± 0.38[b]
5847.5	1.67 ± 0.26[a]	2.76 ± 0.33[a]	5.44 ± 1.05[b]
5804.3	2.43 ± 0.27[a]	5.37 ± 0.71[b]	7.57 ± 0.96[b]
5953.1	4.70 ± 0.27[a]	7.36 ± 0.67[b]	10.14 ± 0.76[c]
6117.2	7.54 ± 0.46[a]	12.04 ± 1.22[b]	14.67 ± 1.19[b]
2688.4	4.83 ± 0.24[a]	4.86 ± 0.23[a]	6.38 ± 0.44[b]
2958.2	10.66 ± 0.43[a]	13.61 ± 1.11[b]	16.55 ± 1.18[b]
3979.1	1.91 ± 0.20[a]	2.17 ± 0.34 [a]	3.91 ± 0.69[b]
8939.2	3.12 ± 0.21[a]	3.69 ± 0.33[a]	4.61 ± 0.41[b]

SSP-PTB: Sputum smear positive pulmonary tuberculosis; SSN-PTB: Sputum smear negative pulmonary tuberculosis. Significance level: $P = 0.05$, [a] > [b] > [c].

using one rule at a time in the form of a question. The splitting decision is based on the intensity of a peak. The process of splitting was continued until the terminal nodes were produced. Multiple classification trees were generated using this process, and the best performing tree was chosen for testing. The discriminatory ability of the classification algorithm was then challenged with a blinded testing set. The sensitivity was defined as the probability of predicting SSN-PTB cases, and the specificity was defined as the probability of predicting healthy control samples. Accuracy was defined as the proportion of correct state classifications.

Results

Study population

The clinical characteristics of SSN-PTB, SSP-PTB, and healthy control groups are shown in **Table 1**. There were no statistically significant differences between the three groups with regard to age and gender ($P > 0.05$). No differences were found in the number of individuals with BCG vaccination and HIV-negative between the three groups ($P > 0.05$). All PTB patients showed different changes as inflammation, opacities, fibrosis and cavities on chest X-ray. Sputum culture for patients with PTB (including the SSN-PTB cases and the SSP-PTB cases) was positive in 115 (100%) subjects with bacteriological analysis. The detailed characteristics of the PTB group are shown in **Table 1**.

A total of 195 serum samples (SSN-PTB = 66, SSP-PTB = 49, healthy controls = 80) were selected for both the training and testing sets. The characteristics between the two sets were tested for statistical significance to ensure that nothing outside the main differentiating factor confounded the results.

Proteomic profiling by SELDI-TOF mass spectrometry

Samples collected from PTB patients and healthy controls were subjected to SELDI-TOF mass spectrometry and analyzed as described above. Up to 138 protein peaks per spot were detected between *m/z* 1,500 and *m/z* 15,000 and the peaks showed the effectiveness of the SELDI technology separation of low molecular weight proteins (< 15,000) (Supplementary Figure 1).

Serum protein peaks associated with SSN-PTB, SSP-PTB and healthy controls

SSN-PTB versus SSP-PTB versus healthy controls: Protein profiling of the 116 serum samples from the training set (42 cases of SSN-PTB, 32 cases of SSP-PTB, and 40 healthy volunteers) were analyzed, and a total of 58 discriminating m/z peaks were detected among the three groups ($P < 0.05$) (Supplementary Table 1). Among these peaks, 19 peaks were up-regulated and 23 peaks were down-regulated in patients with PTB (including the SSN-PTB cases and the SSP-PTB cases) compared to the healthy control subjects. Moreover, 6 protein peaks were found to be down-regulated, while 10 protein peaks were found to be up-regulated

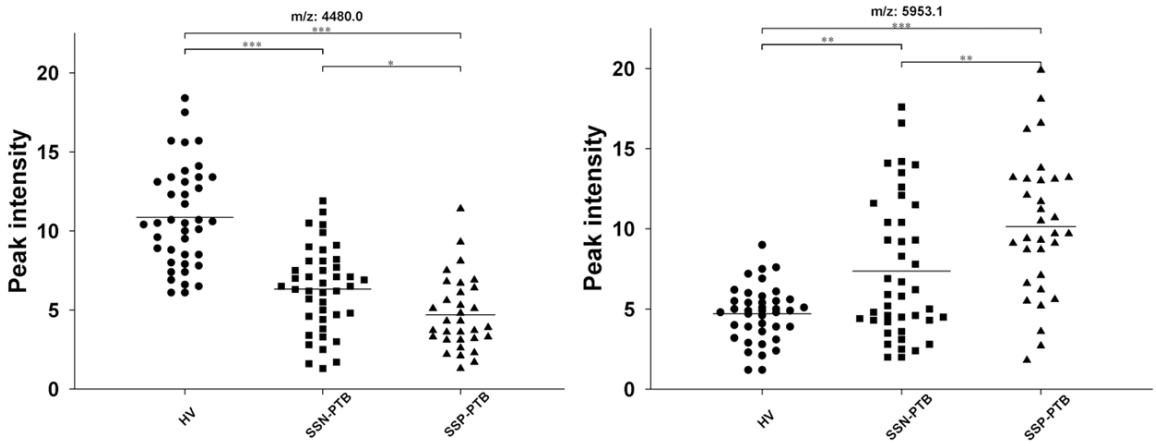


Figure 1. Differential m/z peaks observed by SELDI-TOF MS among SSN-PTB, SSP-PTB and healthy controls. Spectra from SSN-PTB, SSP-PTB and healthy control samples were obtained by SELDI-TOF MS analysis using WCX magnetic beads. Statistical significance ($P < 0.05$) was calculated by one-way ANOVA and least significant difference (LSD) test. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Two SELDI-TOF MS peaks (4480.0, 5953.1) showed significant difference from healthy controls to SSN-PTB to SSP-PTB. Median value (-) of peak intensities are depicted for SSN-PTB, SSP-PTB and healthy control samples. HV: Healthy volunteer; SSP-PTB: Sputum smear positive pulmonary tuberculosis; SSN-PTB: Sputum smear negative pulmonary tuberculosis.

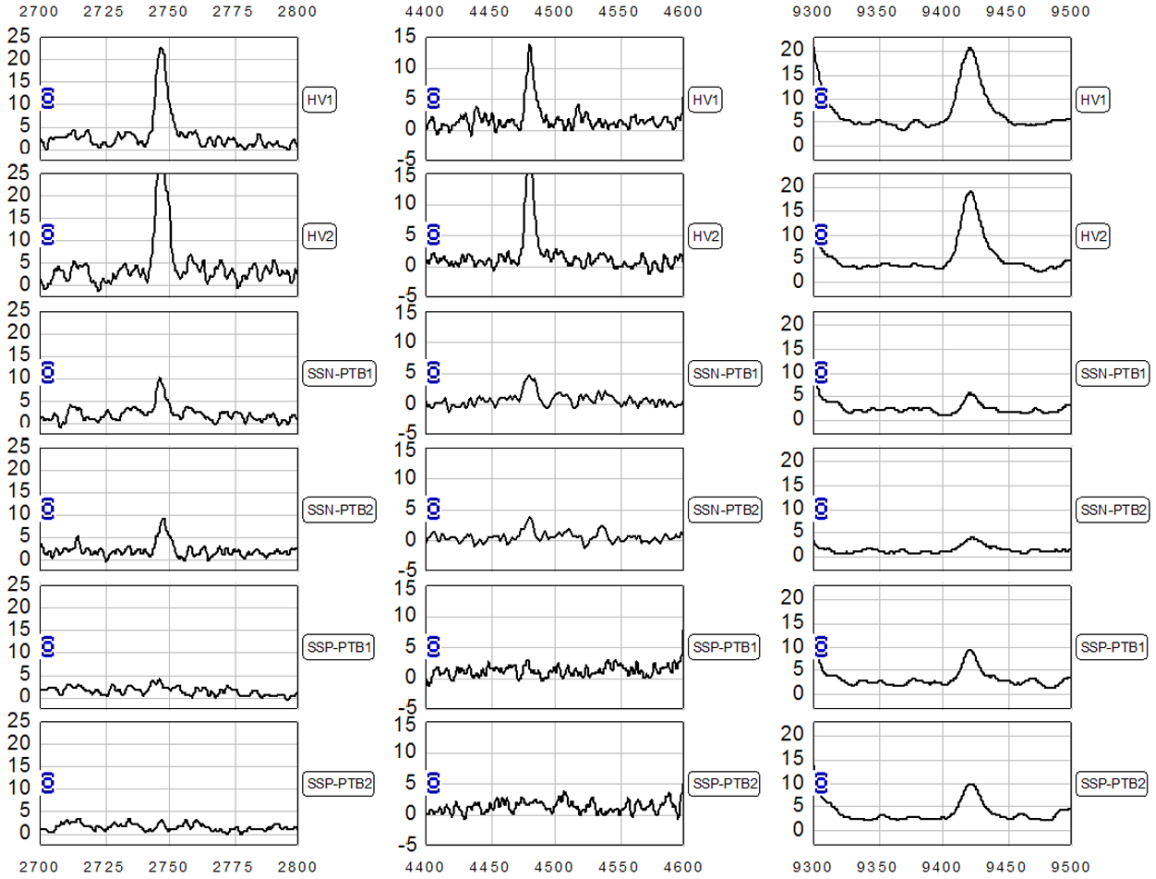


Figure 2. Differential expression of SELDI-TOF MS peaks in serum samples of SSN-PTB patients and healthy volunteers. Peaks with mass/charge of 2747.0, 4480.0, and 9410.1 were detected by SELDI-TOF MS in serum samples from patients with SSN-PTB patients and healthy volunteers. HV: Healthy volunteer; SSN-PTB: Sputum smear negative pulmonary tuberculosis; SELDI-TOF MS: surface-enhanced laser desorption/ionization-time of flight mass spectrometry.

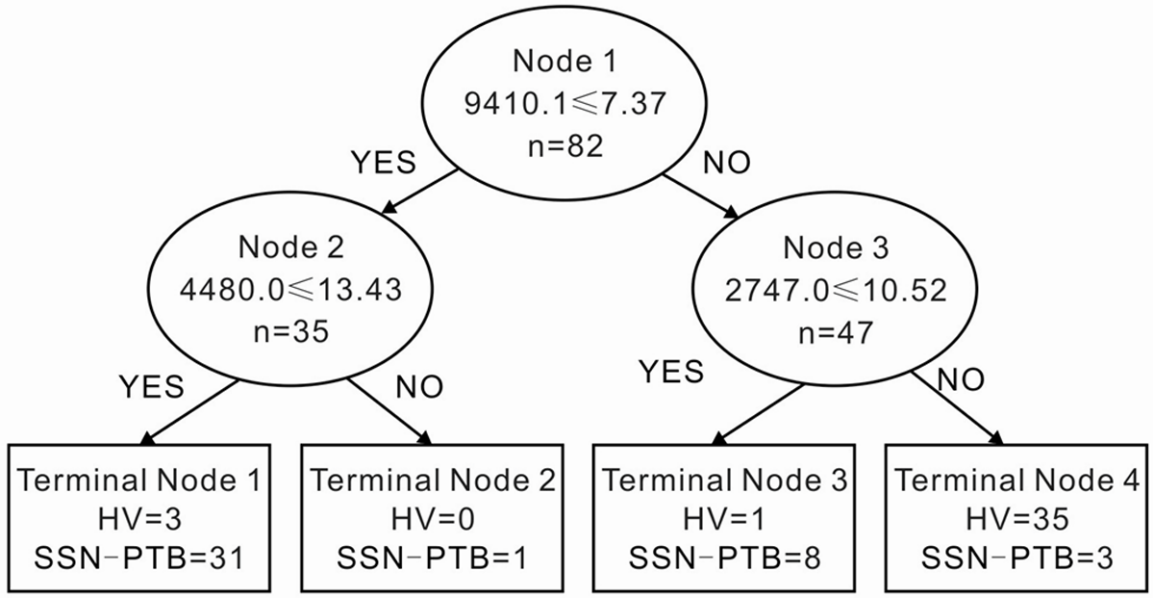


Figure 3. Decision trees in the diagnostic model for SSN-PTB. Three peaks, mass/charge 2747.0, 4480.0, and 9410.1 were chosen to set up the decision tree by the Biomarker Patterns Software. The diagnostic model shows the tree structure and sample distribution of the training set. HV: Healthy volunteer; SSN-PTB: Sputum smear negative pulmonary tuberculosis.

Table 3. Prediction results of the diagnostic model for SSN-PTB

Group	*Samples*	*Cases*	*Correct-classed*	*Accurate %*
Training set	SSN-PTB	42	39	92.86%
	HV	40	36	90.00%
Testing set	SSN-PTB	24	20	83.33%
	HV	40	33	82.50%

HV: Healthy volunteer; SSN-PTB: Sputum smear negative pulmonary tuberculosis.

gradually in healthy controls, SSN-PTB, and SSP-PTB (**Table 2**). The peak of 4480.0 m/z was down-regulated gradually from the healthy controls to SSN-PTB, and then to SSP-PTB ($P < 0.05$). The peak of 5953.1 m/z was up-regulated gradually from the healthy controls to SSN-PTB, and then to SSP-PTB ($P < 0.05$) (**Figure 1**).

SSN-PTB versus healthy controls: By comparing the protein profile data of the serum samples between 42 patients with SSN-PTB and 40 healthy volunteers, 23 peaks were found to be significantly different ($P < 0.01$, Fold ≥ 1.5) (Supplementary Table 2). Among these peaks, 10 were up-regulated and 13 were down-regulated in patients with SSN-PTB compared with healthy control subjects.

Construction and test of the diagnostic model for SSN-PTB

To develop biomarker patterns for the diagnosis of SSN-PTB, a total of three peaks (2747.0, 4480.0, and 9410.1 Da) were selected to construct a classification tree (**Figure 2**). The tree structure and sample distribution is shown in **Figure 3**. The classification tree using the combination of the three peaks identified 42 patients with SSN-PTB and 40 healthy subjects with a calculated sensitivity of 92.86% and a specificity of 90.00% (**Table 3**).

For validation, the SSN-PTB diagnostic model was tested using an independent data set of 64 samples, including 24 patients with SSN-PTB and 40 healthy subjects. The sensitivity and specificity of the combined three-peak model were 83.33% and 82.50%, respectively, and overall accuracy was 82.81% (**Table 3**).

Discussion

Recent advances in mass spectrometry technology enable the detection of hundreds of proteins from a few microliters of serum. As a high-throughput and non-invasive method, SELDI-TOF MS has been successfully applied in the discovery of serum biomarkers for many diseases [14-16]. In order to explore effective

biomarkers for early diagnosis of TB, SELDI-TOF MS has been applied to analyze the serum proteome [3, 13, 17, 18]. These reports revealed that the serum proteome could distinguish TB patients from healthy controls with high specificity and accuracy.

Disease occurrence and development involves change in proteins. Proteins in human serum may have a correlation with the physiologic and pathologic processes [19]. Therefore, serum proteome are expected to yield promising biomarkers for the pathophysiology of the disease [20]. TB infection can lead to the synthesis of TB-associated proteins which appear in the blood circulation through a variety of pathways [13]. The pathological state of TB patients can also be reflected by the serum proteome [21, 22]. In the present study, we examined the protein profile of 114 serum samples from 42 cases of SSN-PTB, 32 cases of SSP-PTB, and 40 healthy volunteers. For convenient and efficient enrichment of proteins in serum samples, we applied the WCX magnetic beads instead of protein chips. Up to 138 peaks were detected between *m/z* 1,500 and *m/z* 15,000 by WCX magnetic beads combined with SELDI-TOF MS. For the reproducibility of peak intensities detected by SELDI-TOF MS, strict standard operating procedures, internal and external control were combined for data quality and reproducibility [13].

As detected by bioinformatics, 58 peaks showed significant differences betwen the three groups ($P < 0.05$), among which 6 protein peaks were found to be down-regulated, while 10 protein peaks were up-regulated gradually in the healthy control, SSN-PTB, and SSP-PTB groups. Moreover, the peak of 4480.0 m/z was down-regulated ($P < 0.05$), whereas, the peak of 5953.1 m/z was up-regulated gradually ($P < 0.05$). The existence of differential peaks among SSN-PTB patients, SSP-PTB patients and healthy controls indicates a broad pathological change of TB in serum proteome.

The SSN-PTB patients have to wait for a longer period of time to receive proper treatment than the SSP-PTB patients due to delay in diagnosis [23, 24]. Therefore, a rapid diagnosis of SSN-PTB is needed. In our experiment, a total of 23 discriminating m/z peaks were detected between SSN-PTB patients and healthy controls ($P < 0.01$, Fold ≥ 1.5). Decision tree is a flowchart-like tree structure that repeatedly splits data sets into subsets in accordance with the given SSN-PTB patients versus healthy controls. The decision classification tree combined with the three candidate protein peaks with m/z values of 2747.0, 4480.0, and 9410.1 were produced by the BPS with a sensitivity of 83.33% and a specificity of 82.50%.

In our diagnostic model, the three peaks may be biomarkers unique for SSN-PTB. Serum protein peaks at m/z values similar to 9410.1 have been previously reported in the literature from other SELDI-TOF MS-based studies. Kim et al. (2012) [25] identified a 9412 Da protein marker in Gestational Diabetes as isoforms of apolipoprotein C-III (apoC-III). The mass of our marker (9410.1 Da) for SSN-PTB is very similar to that of apoC-III. Apolipoproteins (APOs) are lipid carriers and the expression of APOs is associated with lipid metabolism. Recent reports have indicated that the levels of APOs in the blood could be potential biomarkers for different diseases. ApoC-III has been identified as a potential serum biomarker for breast cancer, papillary thyroid carcinoma, and pancreatic cancer [26-28]. Our previous study has indicated that apoC-III is a potential biomarker for Traditional Chinese Medicine (TCM) syndromes of TB [29]. Profiling of serum proteins is the initial step in identifying such biomarkers, and will help in further investigations.

Conclusions

In our experiments, a set of protein peaks has been found to differ significantly in patients with SSN-PTB, SSP-PTB, and healthy controls. It will help to clarify TB pathogenesis by identifying the differential proteins as novel biomarkers. The model we constructed using the protein peaks at 2747.0, 4480.0, and 9410.1 m/z could successfully distinguish the SSN-PTB patients from the healthy controls. Early detection of SSN-PTB patients is critical to reduce the overall morbidity and mortality of TB, and would be a great benefit from a public health standpoint.

Acknowledgements

This study was supported by grants from National Special Sci-Tech Projects (2012ZX10005001-006), the National Basic Research Program of China (2014CB543002), the National Natural Science Foundation of China (81273882, 81403280), and the Natural Science

Foundation of Zhejiang Province (LQ13H31-0005).

Disclosure of conflict of interest

None.

Abbreviations

TB, tuberculosis; SSN-PTB, sputum smear negative pulmonary TB; SSP-PTB, sputum smear positive pulmonary TB; SELDI-TOF MS, surface-enhanced laser desorption/ionization time-of-flight mass spectrometry; WCX, weak cation exchange; BPS, Biomarker Pattern Software; PTB, pulmonary TB; MTBC, *Mycobacterium tuberculosis* complex; AFB, acid-fast bacilli; LSD, least significant difference; APOs, apolipoproteins; apoC-III, apolipoprotein C-III.

Address correspondence to: Dr. Jicheng Li, Institute of Cell Biology, Zhejiang University School of Medicine, 866, Yuhangtang Road, Hangzhou 310058, P. R. China. Tel: +86-571-88208088; Fax: +86-571-88208088; E-mail: lijichen@zju.edu.cn

References

[1] World Health Organization. Global tuberculosis report 2013. Available: http://www.who.int/tb/publications/global_report/en/ Accessed 2 September 2014.

[2] Feyzioglu B, Dogan M, Sanli OO, Ozdemir M and Baykan M. Comparison of the performance of TK system with LJ and MGIT methods in the diagnosis of tuberculosis. Int J Clin Exp Med 2014; 7: 1084-1088.

[3] Agranoff D, Fernandez-Reyes D, Papadopoulos MC, Rojas SA, Herbster M, Loosemore A, Tarelli E, Sheldon J, Schwenk A, Pollok R, Rayner CF and Krishna S. Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. Lancet 2006; 368: 1012-1021.

[4] Kumar P, Pandya D, Singh N, Behera D, Aggarwal P and Singh S. Loop-mediated isothermal amplification assay for rapid and sensitive diagnosis of tuberculosis. J Infect 2014; 69: 607-615.

[5] Caliskan T, Ozkisa T, Aribal S, Kaya H, Incedayi M, Ulcay A and Ciftci F. High resolution computed tomography findings in smear-negative pulmonary tuberculosis patients according to their culture status. J Thorac Dis 2014; 6: 706-712.

[6] Sreeramareddy CT, Rahman M, Harsha Kumar HN, Shah M, Hossain AM, Sayem MA, Moreira JM and Van den Ende J. Intuitive weights of harm for therapeutic decision making in smear-negative pulmonary Tuberculosis: an interview study of physicians in India, Pakistan and Bangladesh. BMC Med Inform Decis Mak 2014; 14: 67.

[7] Liu Y, Hüttenhain R, Collins B and Aebersold R. Mass spectrometric protein maps for biomarker discovery and clinical research. Expert Rev Mol Diagn 2013; 13: 811-825.

[8] Cardoza JD, Parikh JR, Ficarro SB and Marto JA. Mass spectrometry-based proteomics: qualitative identification to activity-based protein profiling. Wiley Interdiscip Rev Syst Biol Med 2012; 4: 141-162.

[9] Cho WC. Proteomics technologies and challenges. Genomics Proteomics Bioinformatics 2007; 5: 77-85.

[10] Liu Z, Yuan Z and Zhao Q. SELDI-TOF-MS proteomic profiling of serum, urine, and amniotic fluid in neural tube defects. PLoS One 2014; 9: e103276.

[11] Cao XL, Li H, Yu XL, Liang P, Dong BW, Fan J, Li M and Liu FY. Predicting early intrahepatic recurrence of hepatocellular carcinoma after microwave ablation using SELDI-TOF proteomic signature. PLoS One 2013; 8: e82448.

[12] World Health Organization; International Union Against Tuberculosis and Lung Disease; Royal Netherlands Tuberculosis Association. Revised international definitions in tuberculosis control. Int J Tuberc Lung Dis 2001; 5: 213-215.

[13] Liu J, Jiang T, Wei L, Yang X, Wang C, Zhang X, Xu D, Chen Z, Yang F and Li JC. The discovery and identification of a candidate proteomic biomarker of active tuberculosis. BMC Infect Dis 2013; 13: 506.

[14] Chung L, Moore K, Phillips L, Boyle FM, Marsh DJ and Baxter RC. Novel serum protein biomarker panel revealed by mass spectrometry and its prognostic value in breast cancer. Breast Cancer Res 2014; 16: R63.

[15] Li X, Liu W, Dong B, Sheng L, Ren H, Han Z, Lu T and Liang P. Identification of serum biomarkers for lung cancer using protein mass spectrometry. Mol Med Rep 2012; 6: 531-534.

[16] Garrisi VM, Tommasi S, Facchiano A, Bongarzone I, De Bortoli M, Cremona M, Cafagna V, Abbate I, Tufaro A, Quaranta M and Paradiso A. Proteomic profile in familial breast cancer patients. Clin Biochem 2013; 46: 259-265.

[17] Zhang J, Wu X, Shi L, Liang Y, Xie Z, Yang Y, Li Z, Liu C and Yang F. Diagnostic serum proteomic analysis in patients with active tuberculosis. Clin Chim Acta 2012; 413: 883-887.

[18] Liu JY, Jin L, Zhao MY, Zhang X, Liu CB, Zhang YX, Li FJ, Zhou JM, Wang HJ and Li JC. New serum biomarkers for detection of tuberculosis using surface- enhanced laser desorption/ionization time-of-flight mass spectrometry. Clin Chem Lab Med 2011; 49: 1727-1733.

[19] Marko-Varga G, Lindberg H, Löfdahl CG, Jönsson P, Hansson L, Dahlbäck M, Lindquist E, Johansson L, Foster M and Fehniger TE. Discovery of biomarker candidates within disease by protein profiling: principles and concepts. J Proteome Res 2005; 4: 1200-1212.

[20] Cho WC, Yip TT, Chung WS, Leung AW, Cheng CH and Yue KK. Potential biomarkers found by protein profiling may provide insight for the macrovascular pathogenesis of diabetes mellitus. Dis Markers 2006; 22: 153-166.

[21] Wallis RS, Wang C, Doherty TM, Onyebujoh P, Vahedi M, Laang H, Olesen O, Parida S and Zumla A. Biomarkers for tuberculosis disease activity, cure, and relapse. Lancet Infect Dis 2010; 10: 68-69.

[22] Phillips PP, Davies GR and Mitchison DA. Biomarkers for tuberculosis disease activity, cure, and relapse. Lancet Infect Dis 2010; 10: 69-70.

[23] Fan L, Zhang Q, Cheng L, Liu Z, Ji X, Cui Z, Ju J and Xiao H. Clinical diagnostic performance of the simultaneous amplification and testing me-thods for detection of the *Mycobacterium tuberculosis* complex for smear –negative or sputum-scarce pulmonary tuberculosis in China. Chin Med J (Engl) 2014; 127: 1863-1867.

[24] Liu Q, Chen X, Hu C, Zhang R, Yue J, Wu G, Li X, Wu Y and Wen F. Serum protein profiling of smear-positive and smear-negative pulmonary tuberculosis using SELDI-TOF mass spectrometry. Lung 2010; 188: 15-23.

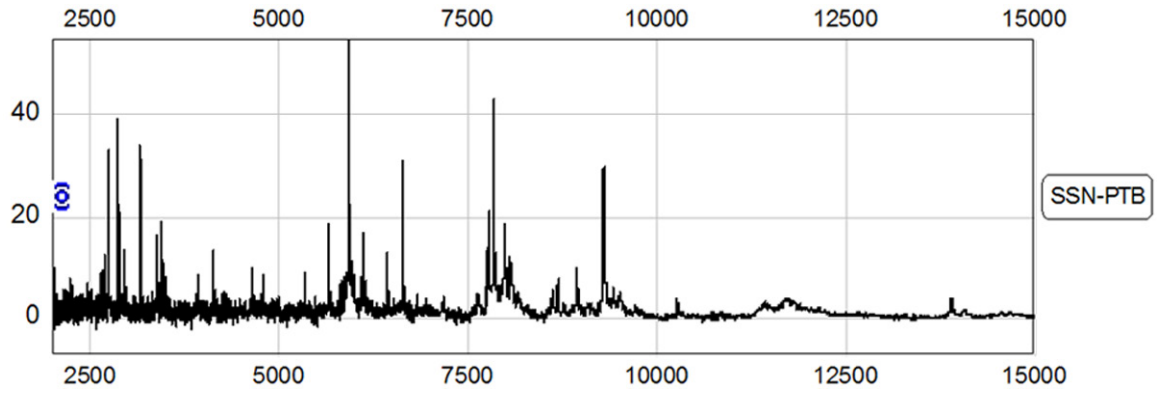
[25] Kim SM, Park JS, Norwitz ER, Lee SM, Kim BJ, Park CW, Jun JK, Kim CW and Syn HC. Identification of proteomic biomarkers in maternal plasma in the early second trimester that predict thesubsequent development of gestational diabetes. Reprod Sci 2012; 19: 202-209.

[26] Chen J, Anderson M, Misek DE, Simeone DM and Lubman DM. Characterization of apolipoprotein and apolipoprotein precursors in pancreatic cancer serum samples via two-dimensional liquid chromatography and mass spectrometry. J Chromatogr A 2007; 1162: 117-125.

[27] Fan Y, Shi L, Liu Q, Dong R, Zhang Q, Yang S, Fan Y, Yang H, Wu P, Yu J, Zheng S, Yang F and Wang J. Discovery and identification of potential biomarkers of papillary thyroid carcinoma. Mol Cancer 2009; 8: 79.

[28] McComb ME, Perlman DH, Huang H and Costello CE. Evaluation of an ontarget sample preparation system for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in conjunction with normal-flow peptide highperformance liquid chromatography for peptide mass fingerprint analyses. Rapid Commun Mass Spectrom 2006; 21: 44-58.

[29] Liu J, Li Y, Wei L, Yang X, Xie Z, Jiang T, Wang C, Zhang X, Xu D, Chen Z, Yang F and Li JC. Screening and identification of potential biomarkers and establishment of the diagnostic serum proteomic model for the Traditional Chinese Medicine Syndromes of tuberculosis. J Ethnopharmacol 2014; 155: 1322-1331.



Supplementary Figure 1. Representative protein spectrum of SSN-PTB sample detected by SELDI-TOF MS with WCX magnetic beads. Protein spectrum of smear-negative pulmonary tuberculosis sample detected by SELDI-TOF MS with WCX magnetic beads showing the protein mass/charge between 1,500 and 15,000. SELDI-TOF MS: surface-enhanced laser desorption/ionization time-of-flight mass spectrometry; SSN-PTB: Sputum smear negative pulmonary tuberculosis.

Supplementary Table 1. The 58 discriminating m/z peaks among SSN-PTB, SSP-PTB and healthy controls ($P < 0.05$)

m/z	*P*	*F*	*m/z*	*P*	*F*	*m/z*	*P*	*F*
4480.0	0	49.51	5847.5	6.9×10^{-5}	10.46	2672.9	0.003	6.30
6437.0	0	28.67	2872.9	7.5×10^{-5}	10.37	3979.1	0.003	6.21
14038.1	0	28.66	8055.1	1.2×10^{-4}	9.8	2399.8	0.003	6.06
4824.3	8.1×10^{-9}	22.15	5030.6	2.3×10^{-4}	9.06	3733.9	0.003	6.02
4537.9	2.1×10^{-8}	20.85	2958.2	2.3×10^{-4}	9.08	2623.1	0.004	5.94
5953.1	3.7×10^{-8}	20.04	3178.5	2.6×10^{-4}	8.90	5647.2	0.004	5.88
5876.4	4.8×10^{-8}	19.70	16094.7	3.1×10^{-4}	8.71	3896.2	0.006	5.40
3321.2	2.1×10^{-7}	17.74	6848.2	3.3×10^{-4}	8.61	8939.2	0.007	5.23
9410.1	8.8×10^{-7}	15.85	8147.5	3.3×10^{-4}	8.61	1864.9	0.008	4.99
4716.8	9.0×10^{-7}	15.83	3960.4	3.7×10^{-4}	8.50	1625.7	0.020	4.08
3402.4	2.7×10^{-6}	14.41	16301.8	3.9×10^{-4}	8.42	2133.1	0.021	3.99
5804.3	3.5×10^{-6}	14.08	2688.4	5.4×10^{-4}	8.06	4679.6	0.027	3.73
8217.8	5.1×10^{-6}	13.62	3623.3	5.5×10^{-4}	8.04	3448.5	0.027	3.72
8692.8	1.1×10^{-5}	12.66	3845.9	8.2×10^{-4}	7.59	16535.8	0.028	3.68
3067.7	1.2×10^{-5}	12.55	3512.5	8.9×10^{-4}	7.50	18504.5	0.030	3.63
6117.2	1.5×10^{-5}	12.30	4846.4	0.001	7.37	9192.8	0.031	3.60
3163.8	2.0×10^{-5}	11.93	14556.9	0.001	6.95	2336.1	0.033	3.51
2747.0	2.5×10^{-5}	11.65	2111.0	0.002	6.80	7976.1	0.037	3.40
8607.5	5.1×10^{-5}	10.82	2844.9	0.002	6.63	N/A	N/A	N/A
9709.0	5.4×10^{-5}	10.75	8767.1	0.002	6.42	N/A	N/A	N/A

m/z means mass-to-charge ratio. $P < 10^{-10}$ regarded as 0. SSP-PTB: Sputum smear positive pulmonary tuberculosis; SSN-PTB: Sputum smear negative pulmonary tuberculosis.

Supplementary Table 2. The 23 discriminating m/z peaks between patients with SSN-PTB and healthy controls

m/z	*P*[a]	Fold	*m/z*	*P*[a]	Fold	*m/z*	*P*[a]	Fold
8607.5	0	+2.5	4716.8	2.9×10^{-6}	-1.8	8153.4	2.2×10^{-4}	+1.5
3402.4	1.1×10^{-8}	-1.9	2747.0[b]	5.2×10^{-6}	-1.6	16301.8	2.9×10^{-4}	+2.0
4480.0[b]	2.0×10^{-8}	-1.7	4824.3	1.9×10^{-5}	+6.3	16097.0	3.1×10^{-4}	+2.3
9410.1[b]	3.1×10^{-8}	-2.0	3321.4	4.0×10^{-5}	-1.6	9348.4	5.0×10^{-4}	-1.5
14038.1	3.9×10^{-7}	-1.8	8767.1	7.5×10^{-5}	-1.7	2111.0	5.6×10^{-4}	-1.6
4537.9	5.0×10^{-7}	+2.5	6847.9	1.0×10^{-4}	-1.6	5804.3	0.001	+2.2
6437.1	1.3×10^{-6}	-1.7	8056.2	1.6×10^{-4}	+2.3	4846.4	0.008	+2.1
8217.8	2.5×10^{-6}	+1.8	9192.5	2.1×10^{-4}	-1.9	N/A	N/A	N/A

m/z means mass-to-charge ratio. [a]Generated by peak comparison between SSN-PTB and healthy controls. [b]Peak selected as biomarkers for SSN-PTB diagnostic model. +fold change (up-regulated), -fold change (down-regulated). $P < 10^{-10}$ regarded as 0. SSP-PTB: Sputum smear positive pulmonary tuberculosis; SSN-PTB: Sputum smear negative pulmonary tuberculosis.