Applied research on serum protein fingerprints for prediction of Qi deficiency syndrome and phlegm and blood stasis in patients with non-small cell lung cancer

Zhizhen Liu, Zongyang Yu, Xuenong OuYang, Jian Du, Xiaopeng Lan, Meng Zhao

Abstract

OBJECTIVE: This study screened serum tumor biomarkers by surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) to establish a subset which could be used for the prediction of Qi deficiency syndrome and phlegm and blood stasis in patients with non-small cell lung cancer; and as diagnostic model of Chinese medicine.

METHODS: Serum samples from 63 lung cancer patients with Qi deficiency syndrome and phlegm and blood stasis, and 28 lung cancer patients with non-Qi deficiency syndrome and phlegm and blood stasis were analyzed using SELDI-TOF-MS with a PBS II-C protein chip reader. Protein profiles were generated using immobilized metal affinity capture (IMAC3) protein chips. Differentially-expressed proteins were screened. Protein peak clustering and classification analyses were performed using Biomarker Wizard and Biomarker Pattern software packages, respectively.

RESULTS: A total of 268 effective protein peaks were detected in the 1,000-10,000 Da molecular range for the 15 serum proteins screened ($P<0.05$). The decision tree model was M 2284.97, with a sensitivity of 96.2% and a specificity of 66.7%.

CONCLUSION: SELDI-TOF-MS techniques, combined with a decision tree model, can help identify serum proteomic biomarkers related to Qi deficiency syndrome and phlegm and blood stasis in lung cancer patients; and the predictive model can be used to discriminate between Chinese medicine diagnostic models of disease.

INTRODUCTION

Lung cancer is the most common malignant tumor, with more than 1.5 million new cases and more than 1 million deaths per year worldwide. About two thirds of lung cancer cases are recognized only when in the ad-
advanced stages. We previously found that advanced lung cancer was caused by a weakness of vital Qi and phlegm and blood stasis obstructing lung collaterals, which is a syndrome of deficiency in origin and excessive in superficiality.

Surface enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS) is a newly developed proteomics technique which consists of a protein chip and mass spectrometry. The technique has attracted much attention because of its use in the early detection of cancer, and as an anti-cancer therapy including drug exploitation. SELDI-TOF-MS can be used to examine the expression patterns of corelative proteins in a certain physiologic or pathologic course, and shows significant value and great potential in biological research, especially cancer research.

In the present study, serum samples from lung cancer patients with Qi deficiency syndrome and phlegm and blood stasis were analyzed using an immobilized metal affinity capture (IMAC) protein chip with SELDI-TOF-MS technology to screen serum for differentially-expressed markers of Qi deficiency syndrome and phlegm and blood stasis to develop a diagnostic model. Such a model will aid in the theory of Traditional Chinese Medicine (TCM).

**MATERIALS AND METHODS**

**Participant selection**
Criteria for enrollment: 1) Patients with non-small cell lung cancer (NSCLC), confirmed histopathologically; 2) Patients without any treatment; 3) TCM syndrome differentiation, using the diagnostic criteria for Qi deficiency syndrome and phlegm and blood stasis stipulated in the TCM Clinical Diagnosis and Treatment Terminology as being: cough, spatum, lassitude, shortness of breath and a disinclination to talk, lusterless complexion, spontaneous sweating, fear of fresh wind, a stinging or smarting sensation, a dark and gloomy face, tongue displaying a white-sticky or petechia coating, and a weak-thready or rough-thready pulse.

Criteria for exclusion: 1) Patients with small cell lung cancer; 2) Patients with any acute disease; 3) Patients with liver disease or illness affecting the blood serum protein content; 4) Patients who have previously undergone a surgical operation, chemotherapy, radiotherapy or biological therapy; and 5) Pregnant or breast-feeding women.

General data: Ninety-one cases of NSCLC met the criteria and were enrolled in the study. Sixty-three cases were in the Qi deficiency syndrome and phlegm and blood stasis group, which included 44 males and 19 females. The age of participants ranged between 38-80 years, with the mean being 55 years. Of the participants, 7 were at stage I, 14 were at stage II, 14 were at stage III, and 37 where at stage IV. For histopathological classification, 26 cases had squamous carcinoma; 25 had adenocarcinoma; 7 had adenocarcinoma and alveolar carcinoma; and 2 had large cell carcinoma.

The control group (non-Qi deficiency and phlegm and blood stasis) consisted of 28 cases, including 20 males and 8 females. The age of participants ranged between 35-82 years, with the mean being 55 years. Of the participants, 4 were at stage I, 5 where at stage II, 8 where at stage III, and 11 where at stage IV. For histopathological classification, 11 cases had squamous carcinoma; 12 had adenoid carcinoma; 2 had adenocarcinoma and alveolar carcinoma; and 1 had large cell carcinoma.

There were no significant differences in the age, sex or the clinical stages between the two groups (P>0.05).

**Samples**
Peripheral blood was collected from all pre-surgical patients in the morning after overnight fasting. Samples were kept at room temperature for 30 min before being centrifuged at 3,500 rpm for 10 min at 4°C to isolate the serum. Serum was stored at -80°C until use.

**Preparation of protein chips**
Results from preliminary serum protein profiling experiments using three different chip chemistries (a weak cation-exchanger CM10 chip and an immobilized metal affinity capture-copper IMAC3-Cu chip, all from Ciphergen, CA, USA) led us to choose the IMAC3-Cu chip for this study because it yielded the best protein resolution.

Protein chips were prepared according to the manufacturer’s instructions. Briefly, each spot on a chip was activated using 100µL of copper sulfate (100 mM). Excessive copper was removed by washing 3 times with deionized water, incubating with 50 mM sodium acetate (pH 3.0), followed by rinsing 3 more times with deionized water. Buffering solution (0.01MPBS with 500 mM NaCl) was applied to each spot and shaken 3 times for 5 min. The surface around the spots was wiped dry and excessive buffer was removed. Serum samples were thawed on ice and diluted 1:40 in PBS. Diluted sample solution was added to a spot on the IMAC3-Cu chip and shaken for 90 min. Each spot was then washed twice with PBS and once with deionized water. When the chip was dry, 0.5µL of energy-absorbing molecules (EAM) solution (50% sinapinic acid in 50% acetonitrile and 0.1% trifluoroacetic acid) was applied to the spots twice as a matrix. Chips were analyzed with a Protein Biological System II (PBS II) ProteinChip Reader (Ciphergen Biosystems Inc., CA, USA).

All protein chips used in the study went through a single-sample quality control. The coefficients of variance (CV) for the peak intensity were between 0.02 and 0.25, suggesting that any experimental errors were within the acceptable range.
Mass spectrometer analysis

Calibration of the ProteinChip Reader was achieved with an NP20 chip using All-In-One Peptide Molecular Mass Standards (Ciphergen Biosystems Inc., CA, USA). The maximum value for molecular weight collected was set at 100,000 Da, and the optimal range was set between 1,000-10,000 Da.

Data collection and processing

Mass spectra data obtained from the ProteinChip Reader were processed using Ciphergen ProteinChip software (Version 3.1.1) for baseline subtraction and common peak detection. Baseline subtraction was performed on each individual spectrum to eliminate any baseline signal (caused mainly by chemical noise). Peaks in all spectra were normalized using peaks at 5915.6 m/z. Noise filtering was achieved using Biomarker Wizard software (Version 3.0). Peaks with a signal/noise ratio of >5 were considered real peaks. After peak normalization and noise filtering, peaks of similar molecular weight from the same group were grouped together into a peak cluster, so that each cluster then represented a particular protein. Average peak intensities for a particular cluster were compared between groups using the student’s t test, and the resulting P values were used to determine whether the protein was differentially-expressed between two groups. A P value of <0.05 was considered statistically significant.

Decision tree development

A diagnostic decision tree was generated using Biomarker Pattern software (Version 3.1.1). The serum protein expression spectra for all samples were uploaded onto the Biomarker Pattern software to generate decision trees. The sensitivity, specificity, and positive and negative prediction values for each decision tree were blindly tested using the serum protein spectra and clinical diagnosis for the samples in the testing groups.

RESULTS

Fifteen proteins differentially-expressed between Qi deficiency syndrome and phlegm and blood stasis and non-Qi deficiency syndrome and phlegm and blood stasis

Comparisons of serum protein expression profiles between Qi deficiency syndrome and phlegm and blood stasis and non-Qi deficiency syndrome and phlegm and blood stasis in patients revealed 15 proteins were differentially-expressed at a statistically significant level (P<0.05). Among them, 6 proteins, with molecular weights of 1490.14, 1518.06, 2209.18, 4745.96, 5205.92 and 6840.99 Da were downregulated Qi deficiency syndrome and phlegm and blood stasis (QDPBS) compared with non-Qi deficiency syndrome and phlegm and blood stasis (non-QDPBS) (Table 1). Nine proteins with molecular weights of 1260.60, 1528.54, 2075.74, 2083.15, 2284.98, 2510.17, 2857.69, 2894.21 and 6970.05 Da were upregulated (Table 2).

Table 1 Proteins upregulated in patients with Qi deficiency syndrome and phlegm and blood stasis

<table>
<thead>
<tr>
<th>Molecular weight</th>
<th>Protein peak</th>
<th>QDPBS patients</th>
<th>Non-QDPBS patients</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1260.60</td>
<td>1.2±1.4</td>
<td>0.3±0.7</td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>1528.54</td>
<td>1.4±1.2</td>
<td>0.6±0.6</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>2075.74</td>
<td>1.7±2.4</td>
<td>0.2±0.5</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>2083.15</td>
<td>2.5±2.3</td>
<td>1.1±0.6</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>2284.98</td>
<td>2.5±2.4</td>
<td>0.8±0.6</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>2510.17</td>
<td>2.2±1.8</td>
<td>1.1±0.6</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>2857.69</td>
<td>5.0±5.8</td>
<td>1.3±1.3</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>2894.21</td>
<td>1.7±1.5</td>
<td>0.8±0.3</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>6970.05</td>
<td>4.1±5.3</td>
<td>1.3±1.3</td>
<td>0.012</td>
<td></td>
</tr>
</tbody>
</table>

Notes: All protein peak values are shown as mean±standard deviation. Sixty-three QDPBS, and 28 non-QDPBS patients.

Table 2 Proteins downregulated in patients with Qi deficiency syndrome and phlegm and blood stasis

<table>
<thead>
<tr>
<th>Molecular weight</th>
<th>Protein peak</th>
<th>QDPBS patients</th>
<th>Non-QDPBS patients</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1490.14</td>
<td>0.9±1.3</td>
<td>2.2±1.3</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>1518.06</td>
<td>2.0±2.9</td>
<td>5.1±5.0</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>2209.18</td>
<td>0.5±0.9</td>
<td>1.1±0.6</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>4745.96</td>
<td>0.5±0.7</td>
<td>1.0±0.9</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>5205.92</td>
<td>0.8±0.9</td>
<td>1.8±1.8</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>6840.99</td>
<td>2.0±2.0</td>
<td>5.1±4.3</td>
<td>0.022</td>
<td></td>
</tr>
</tbody>
</table>

Notes: All protein peak values are shown as mean±standard deviation. Sixty-three QDPBS, and 28 non-QDPBS patients.

Decision tree development to predict Qi deficiency syndrome and phlegm and blood stasis with NSCLC

To evaluate the clinical significance of differentially-expressed proteins identified in this study, we developed a decision tree using Biomarker Pattern software based on the protein expression profiles from 63 samples from patients with Qi deficiency syndrome and phlegm and blood stasis and samples from 28 patients with non-Qi deficiency syndrome and phlegm and blood stasis (Figure 3).

The decision tree had one judgment node and two terminal nodes. When blindly applied to the test group containing the remaining 63 samples from patients with Qi deficiency syndrome and phlegm and blood stasis and the 28 samples from patients with non-Qi de-
Liu ZZ et al. SELDI-TOF-MS on Qi deficiency syndrome prediction in non-small cell lung cancer

In the study, the forecasting accuracy was 83.52% (76/91), the sensitivity was 96.2%, and the specificity was 66.7%. The detection ratios of QDPBS and non-QPBS patients were 79.4% (50/63) and 92.86% (26/28), respectively. The area under the ROC curve of the decision tree was 0.870.

DISCUSSION

Syndrome differentiation in TCM is a method used to diagnose and treat diseases based on the observation of symptoms and signs. In the study, the decision tree for detecting Qi deficiency syndrome and phlegm and blood stasis controls showed good diagnostic accuracy. The results suggest that SELDI-TOF-MS can be a useful tool in the early detection and treatment of non-small cell lung cancer.
recognize and analyze syndromes of disease. Its unique concept of "dialectic wholism theory" provides for the possibility of using proteomic techniques. Applying proteomics technology, analyzing protein compositions of different syndromes of a same disease and different diseases of a same syndrome, screening expression profiles of protein, and using bioinformatics and statistical analyses, can build a specific protein data mine of a disease syndrome type. To explore the relationship between syndromes and the expression profile of proteins, we set up a forecasting model,7 which revealed the scientific connotation of traditional Chinese medical syndrome and provides a reliable basis and method for diagnosis.

SELDI-TOF-MS integrates a protein chip with mass spectrometry technology, which can directly analyze a sample, such as blood, urine, saliva and other clinical samples, without purification and with convenient operation and sample consumption, and also obtain more single and repeatable maps. In recent years, advanced developments have been made in SELDI-TOF-MS technology, making the detection of cancer faster and more effective. In 2008, Zhao Jian et al8 used the weak cation exchanger protein chip (WCX2) to analyze serum from 12 lung cancer patients with Qi and yin deficiency and 12 lung cancer patients with Qi stagnation and blood stasis. They identified 6 different markers and created a decision tree (diagnostic model) with a sensitivity of 66.67% and a specificity of 100%. These results suggest that SELDI-TOF-MS technology may establish new TCM syndrome differentiation models for lung cancer.

The present study analyzed serum from 63 lung cancer patients with Qi deficiency syndrome and phlegm and blood stasis and compared it with 28 lung cancer patients with non-Qi deficiency syndrome and phlegm and blood stasis using an IMAC3 chip, and successfully discriminated the changing processes of Qi deficiency syndrome and phlegm and blood stasis by using a panel of 15 biomarkers. Nine of these markers were upregulated and 6 were downregulated in patients with Qi deficiency syndrome and phlegm and blood stasis. Two candidate protein peaks, with the m/z value of 2284.97, were selected to establish a predictive model using biomarker pattern software (BPS) with a sensitivity of 96.2% (25/26) and a specificity of 66.7% (51/62). These 15 differentially-expressed protein peaks were different from previous reports because of the differences in the ages of participants, histopathological classification and TNM stage, and the specific ionization and chip surface conditions used (WCX2 and IMAC3 ProteinChip). It will be necessary to further extend the sample size, or use a new mathematical model (such as an artificial neural network), to improve on this classification diagnosis model by giving it a higher accuracy.

SELDI-TOF-MS is an automatic proteomic technique with high throughput and sensitivity and has recently been applied to research on traditional Chinese medicine syndrome types. However, some aspects of the SELDI-TOF-MS technology are still under development, such as isolation, identification and building a protein database. It is believed that with the developments in life sciences, proteomic techniques and bioinformatics, understanding the mechanisms of the internal relationship between disease and syndrome will improve to such a point at which they can provide a new basis for TCM therapy.

REFERENCES