

Identification of serum biomarkers for lung cancer using protein mass spectrometry

XIN LI^{1,2}, WEIXING LIU², BAOWEI DONG¹, LIN SHENG¹, HE REN¹, ZHIYU HAN¹,
TONG LU¹ and PING LIANG¹

¹Department of Interventional Ultrasound, Chinese PLA General Hospital, Beijing 100853; ²Department of Ultrasound, Chinese General Hospital of Armed Police Forces, Beijing 100039, P.R. China

Received February 8, 2012; Accepted June 8, 2012

DOI: 10.3892/mmr.2012.955

Abstract. Prior to pathological changes becoming apparent in any disease, the component and amount of intracellular proteins may undergo alteration. Thus, monitoring of proteins may be used to screen indicators in order to identify prognostic markers. The aim of this study was to investigate the feasibility of identification of serum biomarkers for lung cancer using protein mass spectrometry. Surface-enhanced laser desorption/ionization (SELDI) and weak cation exchange 2 (WCX2) protein chip array were employed for protein profiling of the sera of 17 healthy rabbits and 23 cancer-bearing rabbits, of which 15 developed cancer in the lung (cancer group) and 8 developed lung cancer in the follow-up period (pre-cancer group). Data were obtained using a PBSII-C protein chip reader and analyzed using Biomarker Wizard and Proteinchip 3.1 software. Compared with the healthy rabbits, a total of 5 biomarkers were identified to be differentially expressed among 32 proteins screened from the sera in the cancer group and the pre-cancer group ($P < 0.05$): 3 of the 5 biomarkers were upregulated and 2 were downregulated. Protein mass spectrometry can be used to identify specific molecules closely correlated with the progression of lung cancer and, thus, this method may become an effective tool for the early diagnosis or prediction of lung cancer.

Introduction

Before pathological changes become apparent in any disease, the component and amount of intracellular proteins may undergo alteration. Theoretically, the dynamic monitoring of proteins may be used to screen indicators for the early identification of diseases. A number of intracellular proteins and their lytic or modified forms are known to be reflectors of cellular physiology and pathology. These proteins and protein fragments may enter

the blood and lymph nodes and undergo circulation. Thus, in the serum, with the exception of resident proteins with large molecular weights and high abundance (including albumin and other carrier proteins), proteins or protein fragments with small molecular weights and low abundance are also present. These proteins carry a large amount of information for the diagnosis of disease (1). These proteins or protein fragments may undergo slight alterations at the molecular level (2). The proteins in the serum may be repeatedly collected, presenting an effective way of identifying biomarkers in the serum proteins (1,3,4). However, since screening of these markers is high-throughput, and proteins with small molecular weights and in low abundance are difficult to identify using traditional methods (5,6), the early diagnosis of disease requires that specific markers be identified. Proteomics allows the large-scale screening of potential biomarkers and provides a new technique for the molecular diagnosis of clinical diseases, particularly cancers. In the present study, surface-enhanced laser desorption and ionization time of flight mass spectrometry (SELDI-TOF-MS) was employed to detect the biomarkers for lung cancer in healthy and lung cancer-bearing rabbits, with the aim of establishing a favorable method for the early, rapid and simple diagnosis of lung cancer using minimal samples.

Materials and methods

Experimental animals. The experimental protocol was approved by the animal research committee of the Chinese PLA General Hospital prior to study. A total of 17 healthy New Zealand rabbits were included in the control group and 23 rabbits inoculated with lung cancer cells (VX2 cells) were included in the experimental group. Cancer-bearing rabbits underwent CT scans and those with evident lung cancer were defined as the cancer group ($n=15$), while those without cancer were defined as the pre-cancer group ($n=8$). In the pre-cancer group, lung cancer was observed during the follow-up period.

Reagents. Acetonitrile, trifluoroacetic acid, urea, Hepes buffer, water (HPLC), Tris base, CHAPS, DTT and sinapinic acid (SPA) were purchased from Sigma (St. Louis, MO, USA).

Protein chip array. The weak cation exchange 2 (WCX2; Ciphergen Biosystems Inc., Fremont, CA, USA) protein chip

Correspondence to: Dr Ping Liang, Department of Interventional Ultrasound, Chinese PLA General Hospital, 28 Fuxing Road, Beijing 100853, P.R. China
E-mail: liaorzi999@sina.com

Key words: protein mass spectrometry, lung cancer, biomarker

array was selected, having been pre-coated with weak anionic carboxyl. This array was used to interact with the positively charged groups on the surface of analytes (including lysine, arginine and histidine) in order to capture the proteins.

Protein chip reader. The PBSII-C protein chip reader (Ciphergen Biosystems Inc.), which is designed according to principles in SELDI-TOF-MS, was employed to obtain the data and was used to detect molecules with molecular weights from 1000 to 500 kDa.

Serum processing. In brief, 1 ml of blood was collected from the ear vein and stored at 4°C for 2 h followed by centrifugation at 3000 rpm for 5 min. Aliquots of serum (20 µl/tube) were stored at -70°C until use.

Loading and elution. The sera had been stored in a freezer at a temperature of -70°C. These sera were then removed, thawed and centrifuged at 1000 rpm for 4 min at 4°C. The supernatant was collected for protein chip array profiling analysis. Spots A-H on the WCX2 protein chip array were pre-treated with 5 µl of 10 mmol/l HCl and then washed with MQ water three times. The protein chip array was loaded onto the bioprocessor and 150 µl of binding buffer (100 mmol/l sodium acetate, pH=4.0) was added to spots A-H followed by incubation at room temperature for 5 min under continuous agitation. The liquid was removed, and 150 µl of elution buffer (100 mmol/l, pH=4.0) was used to wash the protein chip array twice (5 min for each). The array was removed from the bioprocessor and washed twice with MQ water. When the protein chip array was dried at room temperature, 0.5 µl of SPA was added to each spot followed by drying.

Obtaining data. The protein chip reader was used to obtain the data. Each day the reader was corrected with standard polypeptides and proteins <200 kDa. The systemic mass deviation was approximately 0.1%. The detection used conditions were as follows: laser intensity, 195; detection sensitivity, 8; optimized range of molecular weight, 3-30 kDa; and maximum molecular weight, 50 kDa. Considering the presence of the matrix peak, peaks <1000 were filtered to avoid any influence by the matrix peak. Proteinchip 3.1 software (Ciphergen Biosystems) was used to obtain the data and Biomarker Wizard software was applied to analyze the protein fingerprinting.

Statistical analysis. Statistical analysis was performed with Proteinchip 3.1 software (Ciphergen Biosystems Inc.). Comparisons of the protein peaks among the three groups were performed using the one-way analysis of variance test. $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Detection of differential protein peaks. The raw data were adjusted using Proteinchip software (homogenization of total ion intensity and molecular weight). For peaks with an M/Z ratio of 1000-50000, Biomarker Wizard software was employed to filter the noise. The filtered value of the initial

Table I. Differentially expressed proteins in the lung cancer bearing rabbits.

M/Z	↑↓	Control	Pre-cancer	Cancer
3139	↑	0.69±1.52	2.57±1.80	5.68±3.07
3971	↑	2.78±3.10	3.78±2.41	15.96±7.01
4154	↓	14.18±6.05	9.86±4.56	5.00±3.12
5563	↑	4.51±5.87	9.53±6.43	26.31±11.01
7960	↓	6.29±3.09	4.27±1.15	3.25±1.88

↑, upregulated; ↓, downregulated.

noise was 5 and that of the second noise was 2, and 10% was used as the minimal threshold for clustering. After the primary data were obtained, the means of 32 single-charged peaks were analyzed using the one-way analysis of variance test. Five protein peaks with differential expression were identified ($P < 0.05$ among groups). Of these proteins, 3 were upregulated (control group < pre-cancer group < cancer group) and 2 were downregulated (control group > pre-cancer group > cancer group) (Table I).

Protein peaks and appearance of the proteins on the analog gel view. The 5 proteins with differential expression and the appearance of the proteins on the analog gel were significantly different among the three groups. The two proteins with an M/Z ratio of 5563 and 4154 were selected for comparison of the protein peak and the gel view (Figs. 1 and 2).

Discussion

Protein is the final product of gene expression, directly reflects the function of specific genes and is the main stakeholder of life activity and biological structure. Proteins not only direct cell metabolism and regulation, but are the targets of pathogenic genes and the majority of drugs. Prior to the pathological changes becoming apparent in any disease, the component and amount of intracellular proteins may experience corresponding alteration. Thus, theoretically, the dynamic monitoring of proteins may serve as a method to screen for certain indicators for the early diagnosis of diseases and prediction of metastasis and recurrence. Numerous intracellular proteins and their lytic or modified forms are known to be reflectors of cellular physiology and pathology. These proteins and protein fragments may enter the blood and lymph nodes undergoing circulation. Thus, in the serum, with the exception of resident proteins with large molecular weights and high abundance (including albumin and other carrier proteins), proteins or protein fragments with small molecular weights and low abundance are also present and carry a large amount of information for disease diagnosis (1). These proteins or protein fragments might undergo slight alterations at the molecular level (2). The proteins in the serum can be repeatedly collected, thus, it presenting an effective way to identify biomarkers in the serum proteins (1,3,4). However, proteins with small molecular weights and in low abundance are difficult to identify using traditional methods (5,6).

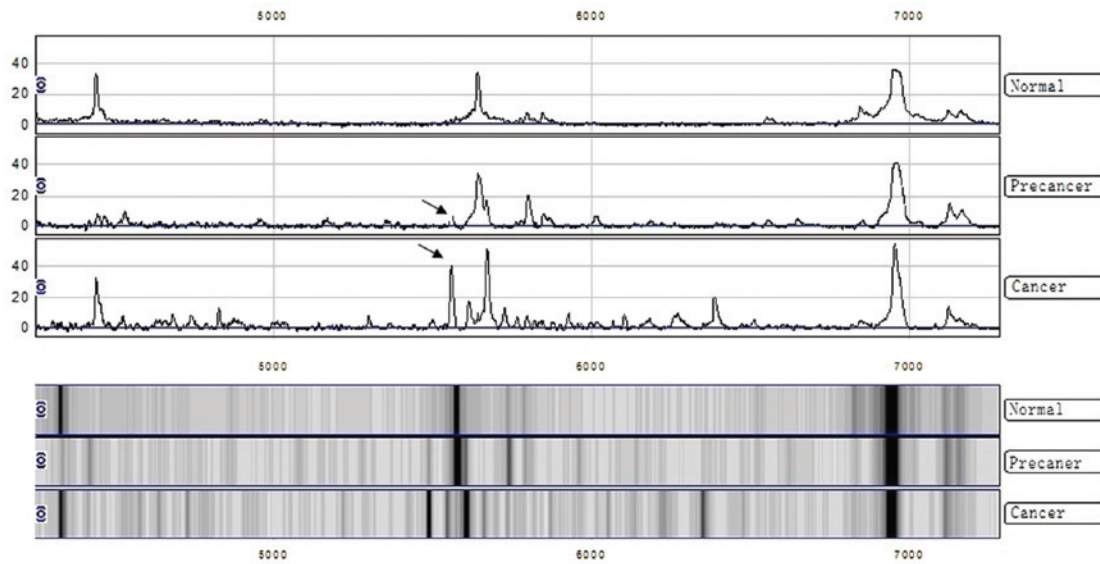


Figure 1. Mass spectrum of protein with a molecular weight of 5563. Arrow, protein peak; normal, control group; pre-cancer, pre-cancer group; cancer, cancer group.

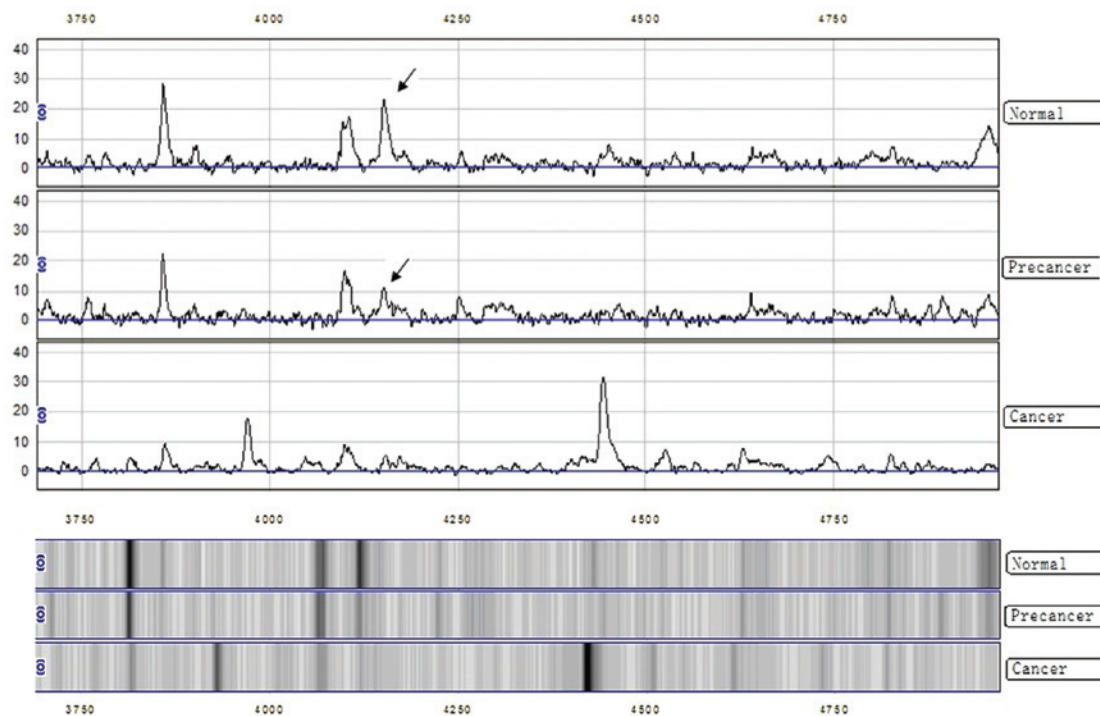


Figure 2. Mass spectrum of protein with a molecular weight of 4154. Arrow, protein peak; normal, control group; pre-cancer, pre-cancer group; cancer, cancer group.

Advances in basic science have led to the development of the SELDI-TOF-MS protein chip array, which integrates recent advances in various disciplines, including Life Science, Microelectronics and Hydrology. The protein chip array provides a reliable platform for the investigation of the structure, property and function of proteins. In this technique, specific probes are used to capture the proteins on the surface. The proteins bind to the SELDI protein chip array. In the desorption/ionization time-of-flight mass spectrometry,

proteins with different M/Z ratios are separated and form different peaks. Each protein has a mass spectrum and software can be employed for bioinformatic analysis of these data to identify the mass spectrums for differentiation (7). The SELDI protein chip array with different surface modifications is capable of selectively capturing proteins, which reduces the complexity of the protein component. Thus, this method may be used to rapidly, simply, sensitively and simultaneously detect multiple proteins with extremely low abundance. In

addition, this detection does not require the purification of samples, including serum, urine or cell lysates. Thus, SELDI protein chip array profiling is particularly suitable for the detection of samples with low abundance (cerebrospinal fluid) and those with complex components (serum). Although the SELDI protein chip array has only been developed in recent years, it is widely applied in the early diagnosis of malignancies. Adam *et al* (8) applied SELDI protein chip array profiling and bioinformatic methods in the screening of prostate cancer specific markers. The results revealed 9 biomarkers for prostate cancer with a sensitivity of 83% and a specificity of 97%. The specificity of these biomarkers was significantly higher than that of prostate-specific antigen (25%) for the detection of prostate cancer. Petricoin *et al* (9) applied SELDI protein chip array profiling and neural network analysis in the screening of biomarkers from the serum of ovarian cancer patients. The results revealed 5 biomarkers with a sensitivity of 100%, a specificity of 95% and a positive detection rate of 94%.

However, previous studies have emphasized the identification of differentially expressed proteins between healthy subjects and patients, and little attention has been paid to the dynamic monitoring of disease. In the present study, healthy rabbits, rabbits without evident lung cancer after inoculation (confirmation of lung cancer in follow-up) and those with evident lung cancer after inoculation were recruited for the screening of biomarkers for lung cancer from the serum. The aim was to explore the feasibility of SELDI protein chip array profiling in the early diagnosis of lung cancer.

In our study, the rabbits in the pre-cancer group developed lung cancer in the follow-up period. Thus, the stage of lung cancer in these rabbits was earlier than that in lung cancer rabbits (cancer group). Our findings revealed 5 differentially expressed proteins in the serum ($P < 0.05$ among the three groups), of which 3 were upregulated (control group < pre-cancer group < cancer group) and 2 were downregulated (control group > pre-cancer group > cancer group). These findings indicate that this method may be crucial for monitoring the occurrence of lung cancer. We also identified a total of 117 protein peaks between the control group and the cancer group, some of which indicated a marked difference ($P < 0.000$). However, the levels of the majority of the peaks were comparable to those in the pre-cancer group ($P > 0.05$) and the control group ($P > 0.05$). Thus, we speculate that these proteins are not suitable for prediction. Proteins without differences between the control group and the pre-cancer group might serve as markers for cancer diagnosis, while proteins without differences between the cancer group and the pre-cancer group suggest that the tumor is or could be present. Determining the stage of disease also requires further study.

Our findings demonstrate that specific biomarkers correlated with disease progression can be identified from the serum. SELDI protein chip array profiling may therefore become an effective tool for the early diagnosis or prediction of lung cancer in clinical practice.

In conclusion, protein mass spectrometry can be used to identify specific biomarkers in the serum associated with disease progression. SELDI is capable of rapidly and simply detecting proteins with low abundance and may become an effective tool for the early diagnosis or prediction of lung cancer.

References

1. Liotta LA, Ferrari M and Petricoin E: Clinical proteomics: written in blood. *Nature* 425: 905, 2003.
2. Liotta LA and Kohn EC: The microenvironment of the tumour-host interface. *Nature* 411: 375-379, 2001.
3. Poon TCW and Johnson PJ: Proteome analysis and its impact on the discovery of serological tumor markers. *Clin Chim Acta* 313: 231-239, 2001.
4. Kas K: On the technicalities of discovering and applying protein biomarkers for cancer prevention. *Eur J Cancer Prev* 13: 437-446, 2004.
5. Tirumalai RS, Chan KC, Prieto DA, *et al*: Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics* 2: 1096-1103, 2003.
6. Petricoin EF and Liotta LA: SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr Opin Biotechnol* 15: 24-30, 2004.
7. Ball G, Mian S, Holding F, *et al*: An integrated approach utilizing artificial neural networks and SELDI mass spectrometry for the classification of human tumours and rapid identification of potential biomarkers. *Bioinformatics* 18: 395-404, 2002.
8. Adam BL, Qu Y, Davis JW, *et al*: Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign hyperplasia and healthy men. *Cancer Res* 62: 3609-3614, 2002.
9. Petricoin EF, Ardekani AM, Hitt BA, *et al*: Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359: 572-577, 2002.