Science in China Series C: Life Sciences

Discovery and identification of Serum Amyloid A protein elevated in lung cancer serum

DAI SongWei[1,2], WANG XiaoMin[1,2], LIU LiYun[1,2], LIU JiFu[3], WU ShanShan[3], HUANG LingYun[1,2], XIAO XueYuan[1,2†] & HE DaCheng[1,2†]

[1] Key Laboratory of Cell Proliferation and Regulation of Ministry of Education, Beijing Normal University, Beijing 100875, China;
[2] Universities' Confederated Institute of Proteomics, Beijing 100875, China;
[3] Department of Thoracic Surgery, General Hospital of Beijing Unit, PLA, Beijing 100500, China

Two hundred and eighteen serum samples from 175 lung cancer patients and 43 healthy individuals were analyzed by using Surface Enhaced Laser Desorption/Ionization Time of Flight Mass Spectrometry (SELDI-TOF-MS). The data analyzed by both Biomarker Wizard™ and Biomarker Patterns™ software showed that a protein peak with the molecular weight of 11.6 kDa significantly increased in lung cancer. Meanwhile，the level of this biomarker was progressively increased with the clinical stages of lung cancer. The candidate biomarker was then obtained from tricine one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis by matching the molecular weight with peaks on WCX2 chips and was identified as Serum Amyloid A protein (SAA) by MALDI/MS-MS and database searching. It was further validated in the same serum samples by immunoprecipitation with commercial SAA antibody. To confirm the SAA differential expression in lung cancer patients, the same set of serum samples was measured by ELISA assay. The result showed that at the cutoff point 0.446（OD value）on the Receiver Operating Characteristic (ROC) curve, SAA could better discriminate lung cancer from healthy individuals with sensitivity of 84.1% and specificity of 80%. These findings demonstrated that SAA could be characterized as a biomarker related to pathological stages of lung cancer.

lung cancer, Serum Amyloid A protein, Surface Enhanced Laser Desorption/Ionizaion Time of Flight Mass Spectrometry

Up to now, lung cancer has become the leading cause of malignancy-related deaths in China, and the five-year patient survival rates remain at 14%[1]. According to pathological changes, lung cancer can be divided into two types: small cell lung carcinoma (sclc) and non-small cell lung carcinoma. And the latter further comprises squamous cell carcinoma (scc), adenocarcinoma (ad), and large cell carcinoma[2]. Lung cancer has complicated pathological and clinical characteristics owing to its different subtypes. Few biomarkers have been accepted for clinical diagnosis because they are less satisfied in sensitivity and specificity. SELDI-TOF-MS is a relatively new and highly sensitive proteomic tool used in the study of proteins. Compared with the conventional technology, SELDI has fast screening ability and high throughput capability. This technology proclaims its advantages especially in raw mixture samples, since it can analyze those samples directly, such as serum, urine, tears and so forth. An increasing number of cancer-related biomarkers for diagnosis, progression and prognosis have been successfully identified by SELDI technique[3−6]. SAA is an acute phase

Received December 20, 2006; accepted January 28, 2007
doi: 10.1007/s11427-007-0053-x
† Corresponding authors (email: dhe@bnu.edu.cn; xyxiao@bnu.edu.cn)
Supported in part by the Major State Basic Research Program of China (Grant No. 2006CB910100), National Natural Science Foundation of China (Grant No. 30370712), Natural Science Foundation of Beijing (Grant No. 7051002), and Science Technology Committee of Beijing (Grant No.Y0204002040111)
Abbreviations: SAA, Serum Amyloid A protein; ELISA, enzyme linked immunosorbent assay; ROC, receiver operating characteristic; SELDI-TOF-MS, Surface-Enhanced Laser Desorption/Ionization-Time of Flight-Mass Spectrometry

reactant, synthesized mainly in liver, whose level in the serum is elevated in head and neck cancer, renal carcinoma, ovarian cancer. SAA was also reported as biomarkers of some diseases and tumors[5,6]. The level of SAA measured by ELISA was reportedly increased in lung cancer patients, while the small number of samples (25) and lack of the detailed clinical stages analysis in those patients were less satisfied in their studies[7]. In order to further analyze how SAA correlates with the clinical stages of lung cancer, we adopted SELDI and other proteomic tools in this study and found that the level of SAA progressively increased with the clinical stage. The same set of serum samples was measured by ELISA assay, and the results showed the consistency with SELDI. The findings demonstrated SAA could be a useful biomarker to monitor the progression of lung cancer.

1 Materials and methods

1.1 Materials

1.1.1 Serum samples. Serum samples of lung cancer were collected from two hospitals, the Department of Thoracic Surgery, General Hospital of Beijing Unit, PLA and the Department of Respiratory Medicine, Second Hospital of Xi'an Jiaotong University. A total of 175 sera included 59 adenocarcinoma, 78 squamous cell carcinoma and 38 small cell lung carcinoma. 43 sera coming from general survey of healthy individuals were used as normal controls and were provided by the State Sport and Physical Culture Administrator. All the sera were collected during October 2003 and July 2004. Samples were immediately aliquoted and stored at −80℃ until the analysis was carried out.

1.1.2 Reagent, protein chip array and instrument. Urea, protein G beads, OGP, CHAPS, acetonitrile and sinapinic acid (SPA) were purchased from Sigma. SAA antibody was ordered from Santa Cruz Biotech. Weak cation exchange chips (WCX2) were obtained from Ciphergen Biosystems Inc. ProteinChip Reader was a laser desorption/ionization time-of-flight mass spectrometer and was made by Ciphergen Biosystems Inc., Fremont, CA, USA.

1.2 Methods

1.2.1 Protein Chip Profiling Analysis. WCX2 chips were pretreated with 10 μL of 10 mmol/L HCl on each spot and stayed at room temperature for 10 min. Then chips were performed in bioprocessors and pretreated with 200 μL of 100 mmol/L sodium acetate according to manufacturer's instruction (Ciphergen Biostystems, CA). Serum/urea mixture samples of 100 μL each diluted with binding buffer were applied to each well, and the bioprocessors were sealed and shaken on a platform shaker at 250 r/min for 60 min. Each well was washed with 200 μL washing buffer (100 mmol/L NaAc, pH 4) three times. The chips were removed from the bioprocessors, washed with 1 mmol/L HEPES（pH 7） quickly, and dried in air. Saturated solution of SPA 0.5 μL was applied to each chip twice and then assayed on a PBS-II protein chip reader. Data were collected by averaging 125 laser shots at an intensity of 210 in the positive mode and a detector sensitivity of 9, and in the optimization range from 3000 to 50000 Da with high mass of 200000 Da. Mass accuracy was calibrated externally using all-in-one protein (Ciphergen Biosystem Inc.).

1.2.2 Tricine-SDS-PAGE and image procurement. Tricine-SDS-PAGE contains Stacking gel and Separating gel. Stacking gel monomer was prepared by mixing 96.0 g acrylamide and 3.0 g N, N′-methylene-bisacrylamide in a total volume of 200 mL ddH_2O. Separating gel monomer was prepared by mixing 96.0 g acrylamide and 3.0 g N, N′-methylene-bisacrylamide in a total volume of 200 mL ddH_2O. Ammonium persulfate and TEMED were added before use. Ten μL lung cancer and normal individual sera were run in tricine sodium dodecyl-sulfate buffer for 1 h at 40 V, and then the voltage was switched to 60 V and continued to run until the bottom of the gel was reached. The gels were stained with CBB R250. Electric images of the gels were obtained using the MagicScan densitometer.

1.2.3 In-gel digestion followed by MALDI-MS/MS identification. The differential expression bands of about 11 kDa were manually excised from the gel and destained with 400 μL of 50% acetonitrile in 25 mmol/L NH_4HCO_3 (pH 8.0) and then dried in a vacuum centrifugation with 50 μL 100% ACN. The dried gel pieces were rehydrated with 10 μL of 25 mmol/L NH_4HCO_3, containing 0.1 g/L trypsin. After incubation for 16−20 h at 37℃, the tryptic peptide mixture was desalted with Millipore ZIPTIP C18 column (Millipore, Bedford, MA, USA) and then analyzed by a Qstar Pulser I Quadrupole time-flight mass spectrometer. The MALDI-TOF MS-MS data were used to identify each protein by

searching the NCBI database with the MASCOT search engine. The mass tolerance was set up as 50 ppm.

1.2.4 Immunoprecipitation. Wash 25 μL of Protein A/G Sepharose beads with PBS for 3 times and then remove supernatant carefully. 20 μL SAA antibody (polyclonal rabbit anti-human) was diluted in immunoprecipitation buffer (1×PBS and PBS with 0.1% Triton X-100, pH 7.4), then blended with 25 μL of pre-cleared Protein A/G Sepharose beads for 2 h at RT with tumbling. After washing beads with PBS for 3 times, 20 μL lung cancer serum was diluted in immunoprecipitation buffer and then added to beads with tumbling overnight at 4℃. The mixture was centrifuged at 1000 r/min for 2 min, and then the supernatant was analyzed on SELDI-TOF-MS with WCX2 chips. After washing beads with PBS for 3 times, the beads with the captured material were eluted with the organic elution buffer (33.3% isopropanol/16.7% acetonitrile/0.1% trifluoracetic acid). The elution was also analyzed on SELDI-TOF-MS with WCX2 chips.

1.2.5 ELISA assay. Serum samples were diluted with coating buffer (0.05 mol/L carbonate buffer, pH 9.6) at a final dilution of 1:10, with 100 μL added to each well of the 96-well plate. Affter overnight at 4℃, the wells were washed with 200 μL of washing buffer for 3 times (1× PBS, 0.05 % Tween 20, pH 7.4) and blocked with 100 μL of 3 % BSA. Then, 100 μL 1:1000 diluted primary antibody (polyclonal rabbit anti-human SAA) was added and incubated for 1 h at 37℃. After washing for 3 times, each well was added with 100 μL secondary antibody (peroxidase-conjugated affinipure goat anti-rabbit IgG-HRP, Golden Bridge Co., China), diluted at 1:3000 and incubated for 1 h at 37℃. After washing, the reaction was developed with o-phenylenediamine dihydrochloride. The reaction was terminated by the addition of 50 μL 3M H_2SO_4. Absorbencies were measured on a microplate reader (Bio-Rad model 550, Bio-Rad, Hercules, CA, USA) at 570 nm.

1.2.6 Bioinformatics and statistics. After normalization, SELDI-TOF-MS data were analyzed by Biomarker Wizard software version 3.2 (Ciphergen Biosystems. Inc.). SPSS software version 13.0 (SPSS INC., Chicago, IL, USA) was used to conduct all statistical comparisons. A nonparametric test (Mann-Whitney U test) was employed to compare the results between different groups. Two-tailed *P* values of 0.05 or less were considered statistically significant. ROC analysis was used to detect the optimal cut-off points (i.e. those with the highest total accuracy) for separating lung cancer patients from the control group.

2 Results

2.1 The peak at 11.61 kDa significantly increased in lung patients

Two hundred and eighteen serum samples were screened by WCX2 chips. Peaks were identified after mass calibration, background subtraction and normalization using the clustering and alignment function of ProteinChip Biomarker Wizard software 3.1. The average peak intensity (mean ± SD) of the 11.6 kDa biomarker in lung cancer (18.48 ± 21.22) was significantly higher than that of normal individuals (2.15 ± 0.73) ($P<0.01$) (Figure 1(a)). The peak information was exported into ProteinChip Biomarker Pattern 4.0 for statistical analysis and a classification tree was constructed (Figure 1(b)). The splitting decision was made by the presence or absence and the intensity levels of one peak. Peaks at 11627, 9431, 10966 and 3294 Da were automatically selected by Biomarker Pattern software as the "node" in the classification tree. The 11627 Da was used to split the root node, and the rule used in the form of a question was whether or not the intensity was equal to or less than 1.027. If the answer was "yes", the samples were discriminated as "normal"; otherwise they were identified as "lung cancer". After this process, 77 samples in 218 were primarily defined as "normal" and went down to "node 2", while 141 samples were primarily discriminated as "lung cancer" and went down to "node 3". The splitting process continued till terminal nodes or leaves were produced or further splitting had no gain. The classification tree with multiple biomarkers can distinguish lung cancer patients from the normal control individuals with sensitivity of 95% and specificity of 81%. We analyzed the level of the 11.6 kDa peak in lung cancer and found that the elevation of 11.6 kDa biomarker correlated with the clinical stage of lung cancer. This phenomenon was especially obvious in squamous carcinomas and adenocarcinomas.

2.2 The biomarkers were identified as SAA by MALDI-MS/MS

Ten μL lung cancer and normal individual sera were separated using Tricine 1-D SDS-PAGE and the differ-

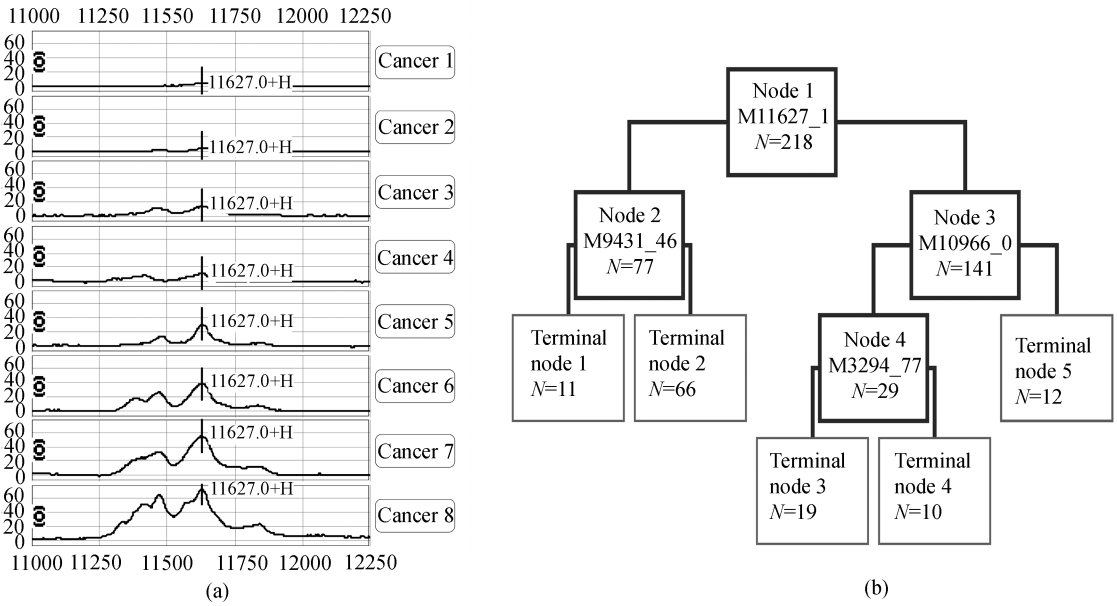

Figure 1 (a) SELDI spectra showing a protein peak of 11.6 kDa on WCX2 chips, with molecular mass from 11000 Da to 13000 Da. Lung cancer sera samples were arranged from top to bottom according to the TNM stage (I, II, III, IV). Two patients in each stage were selected. (b) Classification of the lung cancer and normal control by the decision tree algorithm. The root node is 11627 Da. If the cases of peak intensity are under or equal to 1.027, then they are split into node 2 and diagnosed as normal control, otherwise into node 3 and recognized as lung cancer patient . The split criteria value of 9431 Da is under or equal to 3.483 and the split criteria value of 10966 Da is under or equal to 0.117. Finally the classification tree generates five terminal nodes and 218 samples are split into two groups. M represents the *m*/*z* value of biomarkers, and *N* represents the number of the samples.

ential protein bands were detected on the R250 stained gels according to the image analysis software. One band at about 11 kDa was showed to be significantly up-regulated in lung cancer patients compared with that of normal individuals (Figure 2). The mass value of this band was close to that of the differentially expressed proteins previously determined by SELDI technique. Thus, the differential expression bands of about 11 kDa were cut off and followed by trypsin digestion and MALDI-MS/MS identification. The PMF result showed that the candidate biomarker at about 11 kDa was identified as Serum Amyloid A protein precursor (SAA) (SAA, MW 13553 Da, Mascot accession No. P02735-00-07-00, scored 91) (Figure 3). One peptide of 2178 Da was selected to do tandem MS/MS. Tandem MS/MS spectra were used for database searching with the QSTAR Pulsar I Protean Proteomics software and Internet Mascot software. The searching results also showed that the biomarker was an amyloid A protein precursor, whose *m/z* value was 11682 Da when a signal peptide that contained 18 amino acid was lost.

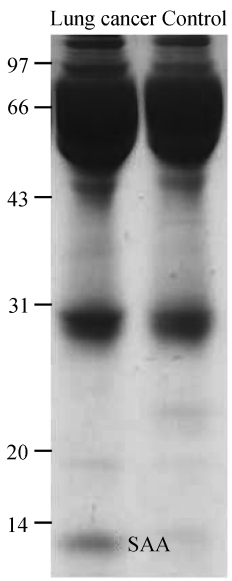

Figure 2 Tricine-SDS-PAGE of lung cancer patient and normal control. The figure shows that the band at about 11 kDa was up-regulated in lung cancer patients.

2.3 Candidate peaks/bands were confirmed by SAA antibody

Considering that the band of about 11 kDa may not be a single protein, we used the immunoprecipitation approach with the commercial SAA antibody to confirm the identification result. The supernatant and elute resulting from the immunoprecipitation with SAA antibody were respectively analyzed by WCX2 chips. It was obviously showed that the peak intensities of 11.6 kDa

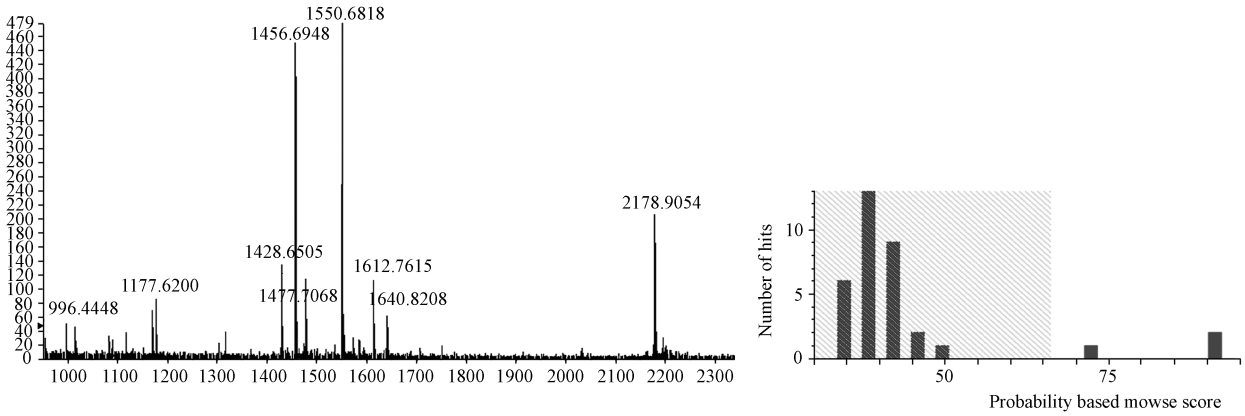

Figure 3 PMF of a protein at about 11 kDa after trypsin digestion and the score by database searching.

were strongly decreased, whereas the peak intensities of other proteins in the same samples remained the same (Figure 4). In contrast, the specific peaks at 11.6 kDa were well demonstrated in the elutes, whereas no peaks at the same *m*/*z* were detected when the SAA antibody was replaced with PBS buffer (figure not showed). We further added the elutes to the control sera and analyzed the mixture by WCX2 chips. Compared with control sera, the mixture demonstrated a unique peak of 11.6 kDa on the spectrum (figure not showed). These results confirmed that the peaks/bands of 11.6 kDa were SAA.

2.4 ELISA-validated SAA increased in lung cancer

We randomly selected 63 lung cancer sera and 25 control cancer sera from the same set of sera with which WCX2 chips were run and performed ELISA assay on the sera. It was found that the levels of SAA (OD value) in lung cancer were significantly higher than those in the normal control group ($P< 0.01$). Considering there was overlap between lung cancer and the control groups, ROC curve analysis was used to detect the optimal cut-off points of SAA for better discrimination of lung cancer from the normal control group. At an optimal cut-off point of 0.446 (OD value) for SAA levels, the sensitivity and the specificity were 84.1% and 80%, respectively. Parallel to ELISA assay, at the selected cut-off points of 2.89 (the intensity of peak) for the SAA on WCX2 chips, the sensitivities of the SAA were 74.6% and the specificities were 81%. Hence, the two approaches were well correlated (Figure 5(a) and (b)). The results revealed that through half-quantitive comparisons of peak intensity, we could carry out the screening of differential proteins.

3 Discussion

Lung cancer is the first cause of death by cancer in the world and its incidence is steadily rising in China. Most of the patients were diagnosed in the late stage of the disease due to lacking of an early specific diagnosis so that the patients lost an opportunity for surgical resection. That is why lung cancer patients have very high mortality rates at all stages. In recent years, screening cancer biomarkers have become a hot field in proteomics.

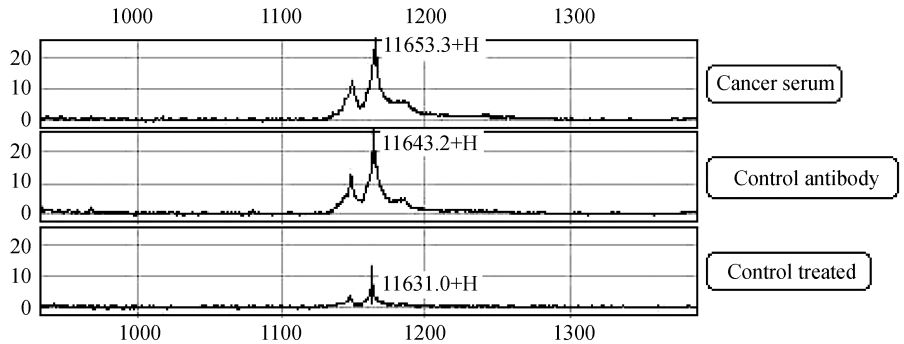

Figure 4 WCX2 protein profilings showing the successful depletion of SAA peaks in the lung cancer serum immunoprecipitated with SAA antibody. (a) The untreated cancer serum; (b) the supernatant of serum without SAA antibody (SAA antibody was replaced with PBS buffer); (c) the supernatant of serum with SAA antibody.

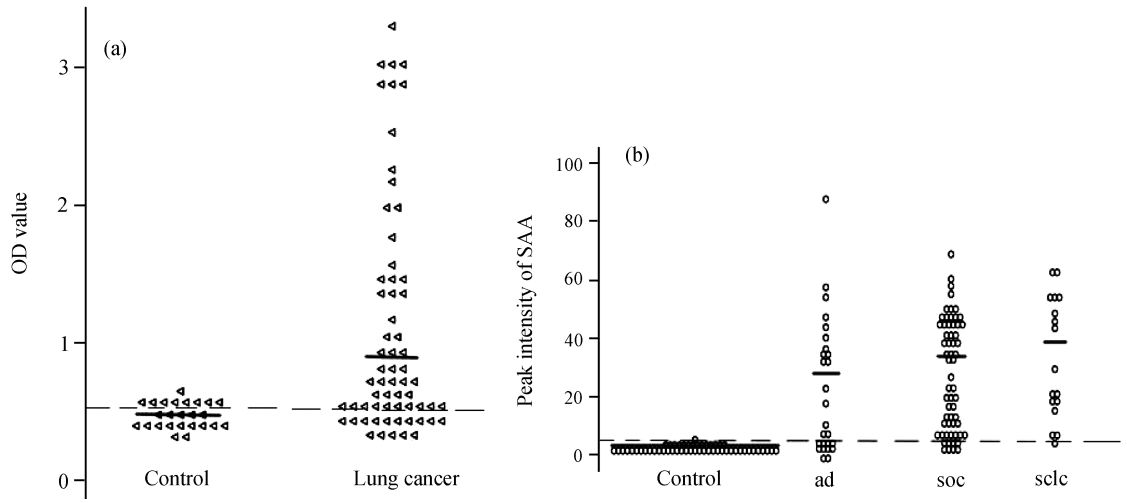

Figure 5 (a) Distribution of the OD value of the 11.6 kDa biomarker in normal control and lung cancer patients was determined by ELISA assay. The real line represents the average OD value of the samples, while dashed line labels the cut-off point. (b) Distribution of the peak intensity of the 11.6 kDa biomarker in the normal control, and three subtypes of lung cancer patients were determined by SELDI. The real line represents the average intensity of the peaks, while dashed line labels the cut-off point. ad, Adenocarcinoma; scc squamous cell carcinoma; sclc small cell lung carcinoma.

SELDI is a promising technique used in study of differential proteomics. It is highly speedy and capable of analyzing the complex biological mixtures directly, but requires only a small amount of samples. The efficacy of the SELDI technology for discovery of biomarkers has been proven in ovarian cancer, prostate cancer, bladder cancer and laryngeal carcinoma[8−11], as well as other diseases[12].

SAA, synthesized mainly in liver, is an acute-phase protein existing as various isoforms. The level of SAA in blood is elevated in response to trauma, infection, inflammation and other acute-phase responses[13−15]. It was also reported previously that the elevation of SAA was detected in many types of tumors[16−18], including lung cancer[7]. The study of SAA level in lung cancer was carried out with ELISA, and only 25 samples were available. Therefore, this study analyzed SAA as a lung cancer biomarker from a new perspective. First of all, 218 serum samples were screened by WCX2 chips, and the results showed that the protein with the *m*/*z* value of 11627 Da was picked up as a candidate biomarker. The candidate biomarker was then obtained from Tricine-SDS gel bands by matching the molecular weight with peaks on WCX chips and was identified as the Serum Amyloid A protein by MALDI/MS-MS and database searching. In our study, very supportive results were yielded from immunoprecipitation, which further confirmed that SAA could be a potential useful biomarker for lung cancer diagnosis. Interestingly we noticed that the SAA level not simply increased in lung cancer patients, but was demonstrated to be well correlated with the clinical stage. This phenomenon was remarkable both in squamous carcinomas and adenocarcinomas. Though the same trend occurred in small cell carcinomas, it had no statistical significance. Therefore, it is meaningful to pay attention to the SAA level when we monitor the clinical stage of lung cancer.

SELDI has the limitation in quantitative analysis of protein, and only yields half-quantitative results, but it selects candidate biomarkers rapidly and minimizes our focus range. Based on the aforementioned results, the differential levels of SAA were re-measured in the same sets of sera by ELISA assay, which was normally used in clinical research. The results were well consistent with that of the SELDI assay. SELDI can screen multiple biomarkers at the same time, and further combine with bioinformatics. It can construct various classification models, such as classification tree, clustering, neuron-network and so forth. These bioinformatics tools improve the sensitivity and specificity in disease diagnosis. In this study the classification tree that was reflected through four masses (M11627, M9431, M10966, M3249) calculated by the Biomarker Pattern program generated five terminal nodes. The classification tree splits the samples into two groups with sensitivity of 95%, much higher than 74.6% when using a single biomarker. Therefore, using the combination of multiple biomarkers in diagnosis yields better sensitivity, moreover, due to

the different mechanism and progression of different diseases, there are pathological changes and unique individual responses, so the combination of multiple biomarkers also improves the specificity in the early diagnosis.

The aforementioned advantages of SELDI show its priority over ELISA especially in differential protein profilings, but the expensive instrument cost constrains its prevalence in clinical use. It is worth mentioning that SELDI has superior visibility so that it provides us not only the accurate *m*/*z* value and relative concentration of proteins, but also some valuable three-dimensional information, such as protein modification and degradation. ELISA assay is easy to operate and does not require the aid of expensive instruments, so it is preferable in general survey of health or clinical use, but it can only measure the total concentration of a protein, which leads to a low discrimination specificity. Other proteins selected by the classification tree are under identification in our lab, and we expect that using other candidate biomarkers which combine with the SAA could improve the sensitivity and specificity in discrimination, and provide a new perspective to facilitate the clinical application.

1 Xiao X Y, Wei X P, He D C. Proteomics approaches to biomarker discovery in lung cancers by SELDI Technology. Sci China Ser C-Life Sci , 2003, 46 (5): 531－537
2 Xiao X Y, Liu D H, He D C, et al. Development of proteomic patterns for detecting lung cancer. Disease Marker, 2004, 19 (1): 33－39
3 Pan Y Z, Xiao X Y, Zhao X J, et al. Application of surface-enhanced laser desorption/ionization time-of-flight-based serum proteomic array technique for the early diagnosis of prostate cancer. Asian J Androl, 2006, 8 (1): 45－51
4 Volker S, Matthias P A E, Thomas B. Advances in clinical cancer proteomics: SELDI-TOF-mass spectrometry and biomarker discovery. Brief Funct Genom Proteom, 2005, 4: 16－26
5 Kokubun M, Imafuku Y, Okada M, et al. Serum amyloid A (SAA) concentration varies among rheumatoid arthritis patients estimated by SAA/CRP ratio. Clin Chim Acta, 2005, 360: 97－102
6 Centre L B, Unite′d O M, et al. Protein chip array profiling analysis of sera from neuroblastoma patients. Cancer Lett, 2005, 228: 91－96
7 Cromer C J, Campa M, Patz E F. Clinical utility of serum amyloid A and macrophage migration inhibitory factor as serum biomarkers for the detection of 16 nonsmall cell lung carcinoma. Cancer, 2004, 101(2): 379－384
8 Petricoin E F, Ardekani A M, Hitt B A, et al. Use of proteomic patterns in serum to identify ovarian cancer. The Lancet, 2002, 359: 572－577
9 William E G, Semmes O J, et al. The early detection research network surface-enhanced laser desorption and ionization prostate cancer detection study. Urologic Oncol, 2004, 22: 337－343
10 Mueller J, Eggeling F V, Driesch D, et al. Protein chip technology reveals distinctive protein expression profiles in the urine of bladder cancer patients. Eur Urol, 2005, 47: 885－894
11 Xiao X Y, Zhao X, Liu J K, et al. Discovery of laryngeal carcinoma by serum proteomic pattern analysis. Sci China Ser C-Life Sci, 2004, 47(3): 219－223
12 Pia D, Magnus S, et al. Proteome studies of CSF in AD patients. Mech Ageing Devel, 2006, 127: 133－137
13 Malle E, De Beer F C. Human serum amyloid A (SAA) protein: A prominent acute-phase reactant for clinical practice. Eur J Clin Invest, 1996, 26(6): 427－435
14 Yip T, Chan J, Cho W, et al. Protein chip array profiling analysis in patients with severe acute respirator syndrome identified serum amyloid A protein as a biomarker potentially useful in monitoring the extent of pneumonia. Clin Chem, 2005, 51: 47－55
15 Kokubun M, Imafuku Y, Okada M, et al. Serum amyloid A (SAA) concentration varies among rheumatoid arthritis patients estimated by SAA/CRP ratio. Clin Chim Acta, 2005, 360 (1-2): 97－102
16 Okoi K, Shih L C, Kobayashi R, et al. Serum amyloid A as a tumor marker in sera of nude mice with orthotopic human pancreatic cancer and in plasma of patients with pancreatic cancer. Int J Oncol, 2005, 27(5): 1361－1369
17 Le L L, Chi K, Tyldesley S, et al. Identification of serum amyloid A as a biomarker to distinguish prostate cancer patients with bone lesions. Clin Chem, 2005, 51: 695－707
18 Cho W, Yip T, Yip C, et al. Identification of serum amyloid A protein as a potential useful biomarkers to monitor relapse of nasopharyngeal cancer by serum proteomic profiling. Clin Cancer Res, 2004, 10: 43－52