Journal of Clinical Laboratory Analysis 22:131–137 (2008)

Detection and Significance of Serum Protein Markers of Small-Cell Lung Cancer

Mingyong Han,[1] Qi Liu,[1]* Jiekai Yu,[2] and Shu Zheng[2]

[1]*Cancer Therapy Center, Shandong Provincial Hospital, Shandong University, Jinan, China*

[2]*Cancer Institute, Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China*

Currently, no satisfactory biomarkers are available to screen for small-cell lung cancer (SCLC). We applied a surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) ProteinChip system to detect 150 serum samples (including 54 SCLC patients, 24 non-small cell lung cancer [NSCLC] patients, 32 pneumonia patients, and 40 healthy individuals). The spectra data were analyzed by support vector machine (SVM) and potential biomarkers were chosen for the system training and used to construct diagnostic model. Pattern 1, constructed of four protein peaks with mass/charge (m/z) of 4,293 Da, 4,612 Da, 6,455 Da, and 7,582 Da, separated SCLC patients from the healthy individuals with a sensitivity of 88.9% and a specificity of 85.7%. This pattern performed significantly better than the current marker, neuron-specific enolase (NSE) ($P < 0.05$). Pattern 2, constructed of protein peaks with mass/charge (m/z) of 2,764 Da and 1,7368 Da, separated SCLC from pneumonia with a sensitivity of 88.9% and a specificity of 91.7%. Pattern 3, constructed of another three protein peaks with m/z of 3,912 Da, 7,562 Da, and 13,777 Da, separated SCLC from NSCLC. The sensitivity and specificity were 83.3% and 75.0%, respectively. These results suggested that SELDI-TOF MS combined with support vector machine yields significantly higher sensitivity and specificity for the detection of serum protein of SCLC. J. Clin. Lab. Anal. 22:131–137, 2008.

Key words: biomarker; diagnosis; proteomics; lung cancer; SELDI-TOF MS

INTRODUCTION

Lung cancer is the leading cause of cancer death in the world. In 2006, the disease caused over 158,000 deaths, more than colorectal, breast, and prostate cancers combined (1). Human lung cancers comprise two major groups, small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma (NSCLC). Approximately 20% of lung cancers are histologically classified as SCLC, which is characterized by rapid growth and a high metastatic potential. Despite initial radio- and chemosensitivity, patients diagnosed with SCLC display a 5-year survival rate of $<5\%$ (2). The poor prognosis is due largely to lack of sufficient screening and early diagnostic tools (3). Most early-stage lung cancers show no symptoms and are detected as an abnormal shadow on a chest roentgenogram or a chest computed tomography (CT) scan, while the sensitivity of CT for lesions <1 cm is low (4). Neuron-specific enolase (NSE) is a widely investigated neuroendocrine molecule. It increases in the serum of patients suffering from diseases with neuroendocrine differentiation characteristic such as SCLC and is therefore a putative serum marker of SCLC (5,6). But the real clinical applicability of such marker remains controversial. Someone observed that an increase in serum NSE level of SCLC patients at any time during follow-up was strongly associated with an

Abbreviations: CV, coefficient of variance; m/z, mass/charge; H, hydrophobic; N-SCLC, non-small-cell lung cancer; NSE, neuron-specific enolase; QC, quality control; ROC, receiver operator characteristic curve; SCLC, small cell lung cancer; SELDI-TOF MS, surface-enhanced laser desorption/ionization time-of-flight mass spectrometry; SVM, support vector machine; UDWT, undecimated discrete wavelet transform.

*Correspondence to: Qi Liu, Cancer Therapy Center, Shandong Provincial Hospital, Shandong University, 324 jingwuweiqi Road, 250021, Jinan, China. E-mail: liuqi_liuqi2007@163.com

Grant sponsor: Shandong Province Science and Technology Project; Grant number: Z2003C01.

Received 28 November 2007; Accepted 12 December 2007

DOI 10.1002/jcla.20230

Published online in Wiley InterScience (www.interscience.wiley.com).

unfavorable prognosis (7). Thus, it is urgent to search for better methods, which provide more valuable information for identification and diagnosis of SCLC, especially in the early stage.

Advances in the proteomics study have introduced novel techniques for the screening of cancer biomarkers and are taking our technology for early diagnosis of cancer diseases to a new horizon (8). The surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) ProteinChip system is a high-throughput technique for analysis of complex biological specimens such as serum in a relatively short period of time (9–11). Small sample volume is required, and the ProteinChip system appears suitable for profiling low molecular weight proteins (12). SELDI-TOF MS has been successfully used to identify highly sensitive and specific potential biomarkers for the diagnosis of ovarian cancer (13), breast cancer (14,15), pancreatic cancer (11,16), colorectal cancer (17), and brain cancer (18,19), and other diseases such as severe acute respiratory syndrome (SARS) (20). Similar studies of SCLC have not yet been reported.

In this study, we used the SELDI-TOF MS technique to detect serum samples from SCLC, NSCLC, pneumonia, and healthy controls. The aim was to search for potential serum biomarkers in SCLC and establish the patterns for diagnosis of SCLC.

MATERIALS AND METHODS

Samples

A total of 150 serum samples were collected from Shandong Provincial Hospital of Shandong University. Informed consent was obtained from every subject prior to the study. These serum samples were collected from 54 patients with SCLC, 20 patients with NSCLC, 32 patients with pneumonia, and 40 healthy individuals. Staging of SCLC was carried out according to the Veteran's Administration lung cancer group staging system (21), limited stage was defined as a disease as a confined to one hemithorax including mediastinal lymph nodes and/or supraclavicular lymph nodes. Extensive disease was defined by opposition to the criteria of limited disease. Patients with SCLC had an average age of 53.6 years (range, 44−73 years; 33 men and 21 women) and consisted of 19 and 35 patients suffering from limited and extensive disease. The 20 NSCLCs included seven squamous cell carcinomas (all stage III), and 13 adenocarcinomas (stage III and IV). Patients in control groups (NSCLC, pneumonia, and healthy subjects) were matched for age and sex with patients in the SCLC group. All the patients were previously untreated. Pathologic diagnoses of all the lung cancers were confirmed independently by the two pathologists. A total of 2 mL of whole blood from patients and healthy individuals were collected during fasting and stored within 1 hr at 4°C. The blood was later centrifuged for 20 min at 4,000 rpm and distributed into 100 μL aliquots; all samples were stored at –80°C until used.

The study was performed after approval by our institute Human Investigations Committee.

ProteinChip Array Analysis

All serum specimens were thawed in wet ice and centrifuged at 3,000 rpm for 5 min at 4°C, and supernatants were retained. A total of 90 μL of 5 g/L CHAPS (Sigma, St. Louis, MO) (pH 7.4) was added into phosphate buffered saline (PBS) to make up 10 μL of each serum sample, and vortex-mixed. The diluted samples were added to 100 μL Cibacron Blue 3GA (Sigma) (previously equilibrated three times with 5 g/L CHAPS) in 96-well cell culture plates and agitated on a platform shaker at 4°C for 60 min. After centrifugation at 1,000 rpm, 50 μL of supernatants were sampled and further diluted by 150 μL of 20 mmol/L 4-2-hydroxyethyl-1-piperazineethanesulfonic acid (HEPES) (pH 7.4) and applied to each well of a bioprocessor (Ciphergen Biosystems, Fremont, CA) containing hydrophobic surface (H4) chips previously activated with 20 mmol/L HEPES. The bioprocessor was then sealed and agitated on a platform shaker for 60 min at 4°C. The excess serum mixtures were discarded, and the chips were washed three times with 20 mmol/L HEPES and twice with deionized water. The chips were then removed from the bioprocessor and air-dried. Before SELDI analysis, 0.5 μL of α-cyano-4-hydroxycinnamic acid (CHCA) (saturation in 50% acetonitrile and 0.5% trifluoracetic acid) were applied twice to each spot, and then air-dried.

Chips were detected on the Protein Biological System II (PBS-II) plus mass spectrometer reader (Ciphergen Biosystems). Data were collected by averaging 80 laser shots with an intensity of 155, a detector sensitivity of 8, a highest mass of 30,000 Da, and an optimized range of 2,000–20,000 Da. Mass accuracy was calibrated to less than 0.1% using the All-in-1 peptide molecular mass standard (Ciphergen Biosystems).

Bioinformatics and Biostatistics

To establish new diagnostic models for SCLC, we attempted to identify three differential patterns of SCLC biomarkers as follows: SCLC vs. healthy controls, SCLC vs. pneumonia, and SCLC vs. NSCLC. Our method analyzing all the data relies on the undecimated discrete wavelet transform (UDWT) as a first step to denoise of spectra. The UDWT method is based on

version 2.4 of the Rice Wavelet Toolbox (RWT). Wavelets have been used previously to denoise signals in a number of contexts, including magnetic resonance imaging and ultrasound blood flow. It has been reported to yield better visual and qualitative denoising. After denoising, the spectra performed baseline correction by fitting a monotone local minimum curve and mass calibration (adjusting the intensity scale according to three labeled peaks that appears in all the selected spectra). The proteomic peaks detected and quantified by an algorithm locates all local maximal height in each denoised, baseline-corrected, calibrated spectrum. Then the peaks are filtered by a signal-to-noise ratio >3 (the signal-to-noise ratio of a peak is estimated as the height above baseline divided by a wavelet-defined noise). To match peaks across spectra, we pooled the list of detected peaks and combined peaks in relative mass by 0.3%, and the percentage of each peak appears in spectra is specified to 10%. The matched peak across spectra is defined as peaks cluster. The spectra that do not have a peak within a given cluster were assigned a maximal height in the cluster for the peak.

In addition, we constructed a nonlinear support vector machine (SVM) classifier with a radial based function (RBF) kernel, and with the parameter Gamma 0.6 being the cost of the constrain violation 19 to discriminate the different groups. SVM is a new machine learning approach originally proposed and developed by Vapnik (22). SVM applications have recently been actively pursued in various areas, from face recognition to genomics (23). SVM is a powerful tool for analyzing complex data derived from SELDI-TOF MS. A 10-fold cross-validation approach was applied to estimate the accuracy of the classifier. This approach randomly selected the nine out of 10 of all the samples to be the blinded training set, and the remaining one of 10 samples to be the test set and repeated the procedure 10 times. SVM classification is based on the shareware program OSU_SVM v.3.00 Toolbox of Junshui Ma and Yi Zhao.

The power of each peak in discriminating different groups was estimated by receiver operator characteristic curve (ROC). The greater area under the curve value of the peak shows the higher relative importance value of the ability to accurately distinguish the different groups. The peaks with lower area under the curve values are excluded. To further select the set of candidate biomarkers, a stepwise approach was used for training many SVM. The top peak that had the highest ability to predict the two groups (having the highest area under curve values) was selected as single input to build the SVM. The discriminating ability of this SVM was estimated by the accuracy of blind test set. Then, the top two peaks were input to the SVM and the accuracy was calculated. The following peaks were added in input stepwise fashion to train the SVM and the accuracy was calculated. In this way, many models with different peaks were built. The peaks inputted to the model with highest accuracy were selected as the set of potential biomarkers, and the SVM with the highest accuracy was selected for detecting lung cancer. Comparison of rates between SELDI pattern group and NSE group was conducted using the chi-square test and $P<0.05$ was regarded as a significant difference.

Detection of Serum NSE

The serum marker, NSE, was measured in 94 sera (54 SCLC and 40 healthy individuals) using an electrochemiluminescent immunoassay (ECLIA, Elecsys 2010 System; Roche Diagnostics, Switzerland) according to manufacturer's instructions with a cutoff value of 16.3 ng/mL.

RESULTS

Reproducibility of the Experiment

The reproducibility of the SELDI spectra, i.e., mass and intensity intraassay and interassay, was determined with the pooled normal serum quality control (QC) sample. A total of four proteins in the range of 2–30 kDa observed on spectra randomly selected over the course of the study were used to calculate the mean coefficient of variance (CV). The intra- and interassay mean CV for mass were 0.3% and 0.5%, respectively, and the intra- and interassay mean CV for the normalized intensity were 11% and 15%, respectively. There was little variation with day-to-day sampling and instrumentation or chip variations.

Serum SELDI Profiles of Lung Cancer vs. Healthy Controls

After noise filtering and peak cluster identification, 142 mass peaks were detected in the training set. These qualified peaks detected from the SCLC and healthy control groups were ranked by ROC. The top 10 peaks with higher area under curve values were selected, randomly combined, and fed into SVM. The accuracy of each combination in distinguishing SCLC from healthy control was analyzed, and the SVM model with the highest accuracy was used as the diagnostic model. This model, designated Pattern 1, comprised four potential biomarkers with mass/charge (m/z) of 4,293, 4,612, 6,455, and 7,582 Da, respectively. The peaks with m/z of 6,455 Da and 7,582 Da were highly expressed in SCLC but weakly expressed in healthy people; the peaks with m/z of 4,293 Da and 4,612 Da appeared to be expressed

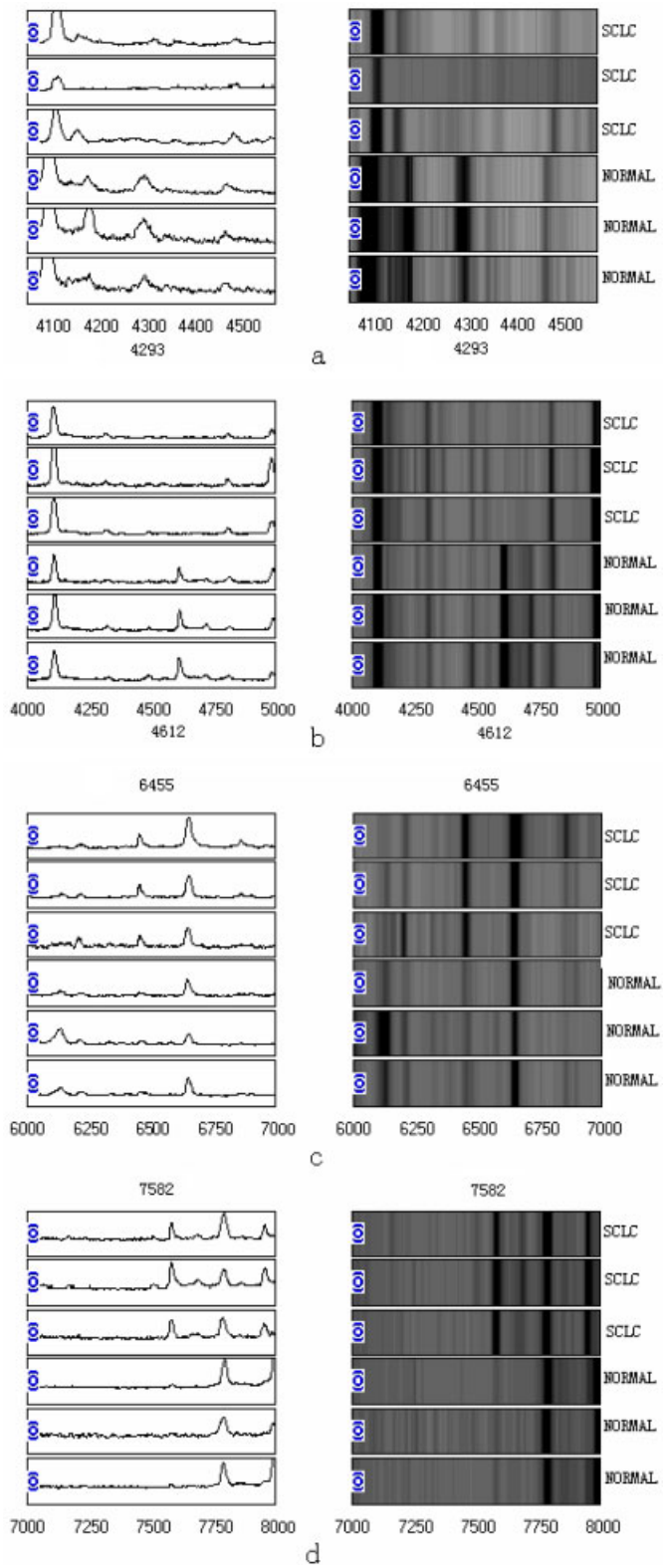

Fig. 1. Representative spectra and gel views of the selected biomarkers of small cell lung cancer patients and healthy individuals. The mass spectrographic profiles reveals upregulation of m/z 6,455 Da (**c**) and m/z 7,582 Da (**d**), and downregulation of m/z 4,293 Da (**a**) and m/z 4,612 Da (**b**) in lung cancer samples.

in a contrasting way, as shown in Fig. 1. The descriptive statistics of these four peaks are shown in Table 1.

The diagnostic model was trained with 62 samples and tested with the remained 32 samples. Through the 10-fold cross-validation SVM, the specificity is 92.3%; the sensitivity is 94.4% in the training set. In the blinded test sets, 16 out of 18 SCLC samples and 12 of 14 healthy individuals were correctly classified by analyzing the blind serum samples. This result yielded a sensitivity of 88.9% and a specificity of 85.7% (Table 2).

NSE levels were available in training and test sets. We found that NSE had a sensitivity of 44.4% (24/54) and the specificity of 77.5% (31/40) for distinguishing SCLC from healthy individuals (Table 2). The four markers identified from our study were significantly better than NSE for distinguishing SCLC patients from healthy controls ($P<0.05$).

The Differential Pattern of SCLC vs. Pneumonia

Pattern 2 was established similarly to distinguish SCLC from pneumonia, and it comprised two potential biomarkers with m/z of 2,764 and 17,368 Da. The two protein peaks were all highly expressed in the SCLC group but all were weakly expressed in the pneumonia group. The specificity and sensitivity of Pattern 2 were 94.4% and 90.0%, respectively, in the training set.

Pattern 2 was tested blindly in another 30 serum samples. The diagnosis was made correctly in 16 out of 18 patients with SCLC and in 11 out of 12 patients with pneumonia. The sensitivity and specificity of the blind test were 88.9% and 91.7%, respectively (Table 2).

TABLE 1. The descriptive statistics of four protein peaks in pattern 1

m/z	AUC	P value ($\times 10^{-5}$)	HP	SCLC	Mean S/N of healthy	Mean S/N of cancer
4293	0.86	0.02	11.66±6.42	2.08±1.17	9.82	1.94
4612	0.85	0.07	8.73±3.96	2.44±0.74	7.67	2.27
6455	0.84	0.11	1.39±4.32	16.39±4.29	1.64	15.27
7582	0.81	0.12	2.47±3.02	7.16±2.24	2.83	7.34

S/N, signal/noise, HP, healthy people.

The Pattern (Pattern 3) to Distinguish SCLC From NSCLC

To distinguish different pathological types, SCLC from NSCLC, Pattern 3, with three peaks, was constructed. The m/z were 3,912, 7,562 and 13,777 Da, respectively. The m/z of 7,562 Da and 13,777 Da were upregulated in the SCLC group, and the m/z of 3,912 Da was upregulated in the NSCLC group. By 10-fold cross-validation, the sensitivity and specificity of Pattern 3 were 88.9% and 87.5%, respectively. Pattern 3 was tested blindly in another 26 serum samples. The diagnosis was made correctly in 15 out of 18 patients with SCLC and in six out of eight patients with NSCLC. In the blind test, a sensitivity of 83.3% and a specificity of 75.0% were obtained (Table 2).

DISCUSSION

Lung cancer is the most common cause of cancer death in industrialized countries and its incidence is steadily increasing in many countries (24). The current common approaches for diagnosis of lung cancer in clinic are mainly based on X-ray and CT, but these kinds of methods have been insufficient in the detection of very small lesion. In many cases, the definite diagnosis for lung cancer was mainstay of pathological diagnosis on biopsy, but the invasive method was not fit for screening and treatment monitoring of lung cancer. The discovery of specific serum biomarkers capable of distinguishing and characterizing lung cancer, subtypes, and different stages was urgent.

The commonly used biomarkers for clinical diagnosis and prognosis in patients with lung cancer today are carcinoembryonic antigen (CEA), cytokeratin-19 fragments (CYFRA-211), and NSE (6). However, all these biomarkers have a poor positive predictive value especially during the early-stage of lung cancer, and some biomarkers are not specific to lung cancer. Serum CEA levels often increase in patients with gastrointestinal tract tumor. In light of the multifactorial nature of

TABLE 2. The predicted results of patterns distinguishing SCLC from controls

		Training set		Test set		NSE	
	Number of protein peaks	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Pattern 1	4	94.4	92.3	88.9	85.7	44.4	78.5
Pattern 2	2	90.0	94.4	88.9	91.7		
Pattern 3	3	88.9	87.5	83.3	75.0		

cancer, it is very likely that a combination of several markers will be necessary to improve the detection and diagnosis of lung cancer. SELDI-TOF MS ProteinChip technology is a new technique that allows multiple serum samples obtained directly from patients to be analyzed in a relatively short time (25). It is a high-throughput approach used to generate protein expression profiles, which in combination with bioinformatics tools to extract information for biomarker discovery, has been essential in identifying novel protein biomarkers.

One of the challenges in the analysis of data generated by SELDI-TOF-MS is to reduce the false protein peaks, in which the discriminatory power is due to random variation (12). The SVM classification technique used in the study is a sophisticated machine learning method based on statistical theory. SVM can solve problems such as the generalization of the medium and small samples in pattern recognition, pattern selection, over-fitting, etc. (26–30).

In this study, we used the integrated approach of SELDI-TOF MS and SVM tools to analyze the large data of spectra. We established three protein fingerprint patterns for SCLC. First, four potential biomarkers, m/z at 4,293 Da, 4,612 Da, 6,455 Da, and 7,582 Da, were identified and constructed as Pattern 1, to distinguish SCLC from healthy controls. The specificity and the sensitivity of this pattern are 85.7% and 88.9%, respectively. We also demonstrated that Pattern 1 was significantly better than the current serum biomarker NSE at distinguishing patients with SCLC from healthy individuals because of its higher sensitivities. This result showed that the selection of a combination of multiple proteins obtained from SELDI may become a potential diagnostic approach. The peaks with m/z of 6,455 Da and 7,582 Da were highly expressed in lung cancer but weakly expressed in healthy people, so the two protein peaks remain interesting to be further investigated. Second, our study identified two potential biomarkers with m/z at 2,764 Da and 17,368 Da, and established a protein fingerprint pattern (Pattern 2) to distinguish SCLC from pneumonia with a specificity of 91.7%, and sensitivity of 88.9%. These results suggest that the two constructed protein fingerprint patterns can be used for detection and screening of SCLC. In the third pattern (Pattern 3), another three potential biomarkers were identified to distinguish SCLC from NSCLC with a sensitivity of 83.3% and a specificity of 75%. The specificity of Pattern 3 was lower than that of the other patterns. One possible reason is that the sample size in this pattern was small, and the other reason may be due to the complexity of lung cancer itself. Although lung cancer is usually divided into two different kinds, SCLC and NSCLC, according to their characteristics, there might exist some correlations between SCLC and NSCLC. Successive works need to expand sample size and further illuminate the meaning of the detected protein peaks in SCLC and NSCLC.

Of our collected SCLC serum samples, 19 patients suffered from limited disease. According to the results of the blinded test set, we found that the pattern had a sensitivity of 78.9% in the detection of limited-stage SCLC. It suggested that the pattern might be better for early detection of SCLC than other single or panel of biomarkers currently used in clinic (7,31).

In this study, we instituted various preventive measures to avoid generation of biased results caused by artifacts related to the nature of the clinical samples. All serum samples were collected and processed within the same clinical and laboratory settings. To avoid variation in the procedure, freshly collected sera were immediately aliquotted, stored at –80°C, and thawed only once. Standard protocols must be developed to minimize unwanted fluctuation, and CVs between ProteinChips must be calculated by using common peaks across different spectra. The use of paired serum samples from individual patients in the study removed most of genetic and environmental variables and made it likely that the changes in protein profile reflected the disease state more exactly. We also used QC serum to allow detection of any unusual features during the process. Such precautions led to very good reproducibility of the protein peak patterns.

In conclusion, we have shown that using SELDI-TOF MS with SVM could find new potential tumor markers for SCLC. The diagnostic models were established by potential tumor markers with SVM had high sensitivity and specificity, and they might be used for screening and early diagnosis of SCLC. More samples should be collected to validate the pattern. Furthermore, for a better understanding of SCLC, identification of the most specific biomarkers described in this study would be the keystone of further investigation.

REFERENCES

1. Manegold C, Thatcher N. Survival improvement in thoracic cancer: progress from the last decade and beyond. Lung Cancer 2007;57:S3–S5.
2. Granetzny A, Boseila A, Wagner W, et al. Surgery in the trimodality treatment of small cell lung cancer: stage-dependent survival. Eur J Cardiothorac Surg 2006;30:212–216.
3. Bharti A, Ma PC, Salgia R. Biomarker discovery in lung cancer-promises and challenges of clinical proteomics. Mass Spectrom Rev 2007;26:451–466.
4. Arita T, Kuramitsu T, Kawamura M, et al. Bronchogenic carcinoma: incidence of metastasis to normal sized lymph nodes. Thorax 1995;50:1257–1269.
5. Hatzakis KD, Froudarakis ME, Bouros D, Tzanakis N, Karkavitsas N, Siafakas NM. Prognostic value of serum tumor markers in patients with lung cancer. Respiration 2002;69:25–29.

6. Lee JH, Chang JH. Diagnostic utility of serum and pleural fluid carcinoembryonic antigen, neuron specific enolase and cytokeratin 19 fragments in patients with effusion from primary lung cancer. Chest 2005;128:2298–2303.
7. Pujol JL, Quantin X, Jacot W, Boher JM, Grenier J, Lamy PJ. Neuroendocrine and cytokeratin serum markers as prognostic determinants of small cell lung cancer. Lung Cancer 2003;39: 131–138.
8. Titulaer MK, Siccama I, Dekker LJ, et al. A database application for pre-processing, storage and comparison of mass spectra derived from patients and controls. BMC Bioinformatics 2006; 7:403.
9. Jemal A, Murray T, Samuels A, Chafoor A, Ward E, Thun MJ. Cancer statistics, 2003. CA Cancer J Clin 2003;53:5–6.
10. Engwegen JY, Gast MC, Schellens JH, Beijnen JH. Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. Trends Pharmacol Sci 2006;27:251–259.
11. Yu Y, Chen S, Wang LS, et al. Prediction of pancreatic cancer by serum biomarkers using surface-enhanced laser desorption/ionization-based decision tree classification. Oncology 2005;68:79–86.
12. Petricoin EF, Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 2004;15:24–30.
13. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359: 572–577.
14. Li J, Zhang Z, Rosenzweig J, Wang YY, Chan DW. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin Chem 2002;48:1296–1304.
15. Vlahou A, Laronga C, Wilson L, et al. A novel approach toward development of a rapid blood test for breast cancer. Clin Breast Cancer 2003;4:203–209.
16. Koopmann J, Zhang Z, White N, et al. Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. Clin Cancer Res 2004;10:860–868.
17. Chen YD, Zheng S, Yu JK, Hu X. Artificial neural networks analysis of surface-enhanced laser desorption/ionization mass spectra of serum protein pattern distinguishes colorectal cancer from healthy population. Clin Cancer Res 2004;10:8380–8385.
18. DeAngelis LM. Brain tumors. N Engl J Med 2001;344:114–123.
19. Mannes AJ, Martin BM, Yang HY, et al. Cystatin C as a cerebrospinal fluid biomarker for pain in humans. Pain 2003;102: 251–256.
20. Yip TT, Chan JW, Cho WC, et al. Protein chip array profiling analysis in patients with severe acute respiratory syndrome identified serum amyloid a protein as a biomarker potentially useful in monitoring the extent of pneumonia. Clin Chem 2005;51:47–55.
21. Zelen M. Keynote address on biostatistics and data retrieval. Cancer Chemother Rep 1973;4:31–42.
22. Liu Y. Active learning with support vector machine applied to gene expression data for cancer classification. J Chem Inf Comput Sci 2004;44:1936–1941.
23. Lin HH, Han LY, Cai CZ, Ji ZL, Chen YZ. Prediction of transporter family from protein sequence by support vector machine approach. Proteins 2006;62:218–231.
24. Nelson DI, Concha-Barrientos M, Driscoll T, et al. The global burden of selected occupational diseases and injury risks: methodology and summary. Am J Ind Med 2005;48:400–418.
25. Seibert V, Wiesner A, Buschmann T, Meuer J. Surface-enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI TOF-MS) and ProteinChip technology in proteomics research. Pathol Res Pract 2004;200:83–94.
26. Somorjai RL, Dolenko B, Baumgartner R. Class prediction and discovery using gene microarray and proteomics mass spectroscopy data: curses, caveats, cautions. Bioinformatics 2003;19: 1481–1491.
27. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics 2000;16:906–914.
28. Byvatov E, Schneider G. Support vector machine applications in bioinformatics. Appl Bioinformatics 2003;2:67–77.
29. Jorissen RN, Gilson MK. Virtual screening of molecular databases using a s support vector machine. J Chem Inf Model 2005;45:549–561.
30. Wiesner A. Detection of tumor markers with proteinchip technology. Curr Pharm Biotechnol 2004;5:45–67.
31. Liu Y, Sturgis CD, Grzybicki DM, et al. Microtubule-associated protein-2: a new sensitive and specific marker for pulmonary carcinoid tumor and small cell carcinoma. Mod Pathol 2001;14: 880–885.