ORIGINAL ARTICLE

Serum protein profiles of patients with lung cancer of different histological types

Rui Hong Yang,1 Rui Fen Tian,2 Qiao Li Ren,2 Hong Ying Chui,2 Su Tang Guo,1 Xu Dong Zhang3 and Xia Song2

Departments of 1 Molecular Biology and 2 Pulmonary Oncology, Shanxi Cancer Hospital and Institute, Affiliated Cancer Hospital of Shanxi Medical University, Taiyuan, China, and 3 School of Medicine and Public Health, University of Newcastle, Callaghan, New South Wales, Australia

Abstract

Aims: To compare serum protein expression profiles between lung cancer patients and healthy individuals, and to examine whether there are differences in serum protein expression profiles among patients with lung cancers of different histological types and whether the characteristic expression of serum proteins may assist in differential diagnosis of various subtypes of lung cancers.

Methods: Blood samples were collected from 123 lung cancer patients before commencement of treatment who attended Shanxi Cancer Hospital, China, between 2008 and 2013. Blood samples from 60 healthy individuals were also collected in the same period. Serum protein expression profiles were analyzed using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry. The differences in the serum protein spectrums of lung cancer patients with different histological subtypes were analyzed by one-way Analysis of Variance and receiver operating characteristic curves.

Results: A cluster of 48 protein mass-to-change ratio (M/Z) peaks was differentially expressed between sera of lung cancer patients and healthy individuals. The M/Z 1205, 4673, 1429 and 4279 peaks were differentially expressed among patients with lung squamous cell carcinomas, adenocarcinomas and small-cell lung carcinomas.

Conclusion: These results reinforce the notion that profiling of serum proteins may be of diagnostic value in lung cancer, and suggest that the differences in serum protein profiles may be useful in differential diagnosis of lung cancers of varying histological subtypes.

Key words: biomarkers, lung cancer, proteomics

INTRODUCTION

Lung cancer is one of the most commonly diagnosed malignancies and the leading cause of cancer-related mortality worldwide.1–3 However, despite recent advances in understanding molecular mechanisms responsible for the pathogenesis of lung cancer and the development of molecularly targeted approaches in the treatment of the disease, there has been no overall improvement in prognosis of patients with lung cancers. In addition, the incidence of lung cancer has been constantly increasing in many parts of the world, in particular, in developing countries.4–8
Lung cancers are histologically classified into non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC), with NSCLCs subdivided into squamous cell carcinoma (SCC), large-cell carcinoma and adenocarcinoma (AC). Nevertheless, it is now understood that lung cancer is a highly heterogeneous disease with a wide variety of oncogenic anomalies driving the pathogenesis in individual patients, such as mutations in the epidermal growth factor receptor gene (EGFR) and rearrangement of anaplastic lymphoma kinase (ALK). This has led to molecular subclassification of lung cancer that is instructive for personalized management of the patients. However, accurately subtyping lung cancers histologically and molecularly before treatment is not always achievable primarily due to limited availability of tissue samples.

On the other hand, liquid biomarkers such as serum carcinoembryonic antigen (CEA), cytokeratin-19 fragments (CYFR-211) and neuron-specific enolase (NSE) may assist in differentiating the diagnosis, but their specificity and practical values remain to be further evaluated.

Characteristic protein profiles determined by proteomic approaches have been increasingly investigated as biomarkers of diagnosis, prognosis and responses to treatment in cancer patients. Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) is a high throughput technique that allows obtaining protein profiles from several complex biological samples with minimal requirements for purification and separation in a rapid and efficient way. Small amount of samples such as body fluids or tissue lysates is directly applied onto biochips that are available with different chromatographic surfaces. Selectively retained proteins are then directly analyzed by laser desorption and ionization. The result is a mass spectrum composed of the mass to charge (M/Z) ratio and intensities of the bound peptide/protein. Comparison of protein profiles according to patient phenotypes generated by SELDI-TOF-MS has led to identification of isolated or clustered peaks characteristic of pathologic conditions. This approach has been used for discovery of new biomarkers, as exemplified in patients with pancreatic, ovarian and liver cancers.

In this study, we have compared serum protein expression profiles among patients with lung cancer of different histological subtypes by use of SELDI-TOF-MS. We report here that a cluster of 48 protein peaks is variably expressed between sera of lung cancer patients and those of healthy individuals, whereas a subset of four protein peaks is differentially expressed among sera of patients with SCLCs, SCCs and ACs. These results reinforce the notion that profiling of serum proteins may be of diagnostic value in lung cancer, and suggest that the differences in serum protein profiles may be useful in differential diagnosis of lung cancers of varying histological subtypes.

METHODS

Sample preparation

Blood samples were collected from lung cancer patients who attended Shanxi Cancer Hospital and Institute, Taiyuan, China, between 2008 and 2013 before surgery. Blood samples from healthy volunteers (HVs) were also collected in the same period. Studies on human samples were reviewed and approved by the Human Ethics Committee of Shanxi Cancer Hospital and Institute. Diagnosis of lung cancer was confirmed by two independent pathologists at the Department of Anatomic Pathology of Shanxi Cancer Hospital. Approximately 2 mL whole blood was collected by phlebotomy into 4 mL vacuum tube (red top tube, Greiner Bio-One, Kremsmunster, Austria). Blood was centrifuged (1000 × g at 4 °C) for 5 min twice. The serum was collected and labeled with a unique number and transferred into four Eppendorf Micro Test tubes (0.5 mL) and stored at −80 °C till analysis. Serum samples from a total of 123 lung cancer patients were collected. The histopathological and clinical characteristics of the patients are shown in Table 1. Serum samples from 60 healthy individuals (30 males and 30 females) who attended the hospital for regular physical examination were used as normal controls. These individuals had no cancer history or acute/chronic inflammatory disease.

SELDI-TOF-MS

We chose weak cationic exchange (CM-10) chip arrays and Cu-metal binding (IMAC-30) chip arrays (Bio-Rad Laboratories, Inc. Hercules, California, USA) for pilot experiments and found that CM-10 chips captured more proteins than IMAC-30 chips. CM-10 chips were therefore used for further studies. Briefly, 5 μL of each serum sample was denatured by addition of 10 μL U9 solution (9M urea, 2% CHAPS and 150 mM Tris-HCl, pH 9). The array spots were washed three times by loading buffer 200 μL (50 mM NaAC, pH 4) for 5 min each time. The denatured sera were then diluted by 190 μL loading buffer. And 200 μL diluted sera was added on the spot of chip array and incubated 2 h at room temperature, followed by two washes with loading buffer (5 min each time) and one wash with distilled water (High Performance
Liquid Chromatography grade). The air-dried arrays were spotted with 0.8 μL sinapinic acid matrix (SPA, saturated with 50% ACN +0.5TFA) (Sigma-Aldrich, St. Louis, MO, USA). The array was placed into the SELDI instrument and read the information of each spot. Each sample was tested in triplicate.

### Data collection and analysis

Serum protein profiling was generated by SELDI-TOF-MS in a PBS-IIc ProteinChip reader (Ciphergen Biosystems, Fremont, CA) according to an automated data collection protocol with the following settings: laser intensity 200, detector sensitivity 8, and molecular mass range M/Z 1000–20000. The data were analyzed with Ciphergen Express software version 3.0 (Ciphergen Biosystems). The software was calibrated using the all in one protein standard kit (Bio-Rad Laboratories, Inc. USA). All spectrums were normalized by total ion current and baseline subtraction. When the signal-to-noise ratio >5 and requirement peaks in full the spectra >25%, the spectrum was effective. Data from triplicate spectra were combined prior to further analysis. Spectrums were divided into three groups according to their pathological diagnosis. Protein peaks were clustered by Biomarker Wizard software version 3.1 to derive a mean, standard deviation (SD) and P value, by comparing peak intensity between different groups.

### Reproducibility analysis of protein spectra

We detected the 60 serum samples from healthy individuals and found two types of spectrum patterns present in all spectrums (Fig. 1). There were apparently different clusters of peaks with M/Z ranging from 9000 to 95000 between the two types (type 1 or normal 1 vs type 2 or normal 2). Type 1 cluster appeared in each serum samples of 26 healthy individual, whereas type 2 cluster appeared in serum samples of the other 34 healthy individuals. Two protein peaks M/Z 2742 and 8690 randomly selected were used to calculate the coefficient of variance (CV). The CV values of intensity for the protein peaks M/Z 2742 and 8690 were 3.98% and 11.14% in type1, whereas the CV values of intensity for the protein peaks M/Z 2742 and 8690 in type 2 were 5.05% and 8.75%, respectively, indicating that these spectrums were highly reproducible. Pooled serum from 60 healthy individuals served as control (QC serum) over the course of the study. All spectrums were calibrated with two protein peaks M/Z 2742 and 8690 of QC serum.

### Statistical analysis

Statistical analysis was carried out using Biomarker Wizard Software 3.1. The discriminatory power for the valuable protein peaks was characterized by receiver operating characteristic (ROC) area under the curve (AUC) (SPSS 17.0) (SPSS Inc, Chicago, IL, USA). One-way
Table 2  Differential expressed peaks among squamous cell lung carcinoma, lung adenocarcinoma and small-cell lung cancer

| M/Z  | P value | Mean ± SD SCC | Mean ± SD AC | Mean ± SD SCLC |
|------|---------|---------------|--------------|----------------|
| 1205 | 0.002   | 8.39 ± 6.98   | 6.89 ± 3.97  | 5.55 ± 2.99    |
| 4673 | 0.006   | 6.26 ± 2.06   | 4.54 ± 1.09  | 2.59 ± 2.39    |
| 1429 | 0.007   | 5.45 ± 3.56   | 3.95 ± 2.23  | 4.92 ± 3.05    |
| 4279 | 0.009   | 21.43 ± 17.68 | 22.30 ± 8.25 | 29.12 ± 20.35 |

Abbreviations: M/Z: mass-to-charge ratio; SD: standard deviation.

ANOVA was used to compare the difference between two experimental groups. A P value less than 0.01 was considered statistically significant.

RESULTS

Comparison of serum protein profiles between lung cancer patients and healthy individuals

We compared serum protein profiles of 123 lung cancer patients with those of 60 healthy individuals, and found that 48 protein peaks were differentially expressed between the two groups. Among them, 12 peaks were expressed at decreased levels in lung cancer patients compared to healthy controls, whereas the other 36 peaks displayed higher abundance in lung cancer patients.

Comparison of serum protein profiles among patients with lung cancers of different histological subtypes

Of the 48 protein peaks that were differentially expressed between lung cancer patients and healthy individuals, a subset of four peaks (M/Z 1205, 4673, 1429 and 4279) was variably expressed among patients with lung cancers of different histological subtypes (SCC vs AC vs SCLC) (P < 0.01) (Table 2). The peaks M/Z 1205, 4673 and 1429 were all elevated, whereas the peak M/Z 4279 was decreased in patients with SCCs compared with those with the other two types of lung cancers (Table 2). Patients with SCLCs had lowest levels of the protein peaks M/Z 1205 and 4673, whereas patients with ACs exhibited intermediate levels of all the four protein peaks (Table 2). Representative spectra of protein peaks M/Z 1205 and 4279 from patients with lung cancers of the three different histological subtypes are shown in Figure 2.

ROC analysis

We conducted ROC analysis to determine the sensitivity and specificity of the four serum protein peaks (M/Z 1205, 4673, 1429 and 4279) in differentiation of the
three pathological subtypes of lung cancer (SCC vs AC vs SCLC). The AUCs of the four peaks were 0.84, 0.80, 0.82 and 0.18, respectively, for discrimination between patients with SCCs and those with SCLCs; 0.82, 0.72, 0.80 and 0.41, respectively, for discrimination between patients with SCCs and those with ACs; 0.53, 0.41, 0.58 and 0.87, respectively, for discrimination between patients with ACs and those with SCLCs (Fig. 3a–c).

We further conducted ROC analysis for combination of M/Z 1205 and M/Z 4279 and found that AUC was 0.91 for discrimination between patients with SCCs and those with SCLCs.

DISCUSSION

The treatment and outcome of lung cancer is dictated by the histological and molecular profiles of the tumor. Advances in understanding genetic and molecular characteristics of lung cancers have led to identifications of novel biomarkers expressed by lung cancer cells that assist in histological diagnosis of the disease. For example, immunohistochemistry staining of AE1/AE3, TTF-1, P63, CGA, CK7, CK20 and CD56 is now widely applied in practice. However, it is not uncommon that staining positivity of these markers is not sufficiently high to differentiate lung cancers of different subtypes. In addition, accurately defining the histological subtype and molecular profiles before commencement of treatment is often precluded by the limited availability of tissue samples.

Increasing studies have focused on identifying biomarkers from liquid biopsies such as blood samples. This is largely attributed to the development of novel technologies that provide simple, less time-consuming approaches with high sensitivity and specificity. Among them is SELDI-TOF-MS that has widely been applied in the field of biomarker discovery by profiling proteomic patterns of various diseases. However, reproducibility of SELDI-TOF-MS has been a concern, as different researchers may choose varying molecular ranges and using different chips and experimental procedures. In this study, we employed the commonly used chip and experimental process and identified a group of 48 protein peaks that were differentially expressed between patients with lung cancers and healthy individuals. Nevertheless, a number of peaks that we identified, for example, the peaks M/Z 1205 and 4279, were not recognized previously. This is conceivably due to the different experimental procedures used in the studies. Moreover, the varying approaches in data analysis may also contribute to the difference. This reinforces the urgent need for development of unified protocols for biomarker discovery using SELDI-TOF-MS.

Figure 3 ROC curves analysis of protein peaks at M/Z 1205, 4673, 1429 and 4279 for differentiating the three pathological subtypes of lung cancer. ROC curves were generated for SCCs versus SCLCs (a), SCCs versus ACs (b) and ACs versus SCLCs (c).
cancers with different histological subtypes by profiling serum protein patterns using proteomic approaches. In this study, we employed SELDI-TOF-MS to compare the serum protein profiles of patients with lung cancers of different histological subtypes (SCC vs AC vs SCLC), and found that a subset of four protein peaks (M/Z 1205, 4673, 1429 and 4279) was differentially expressed. While the protein peak M/Z 4279 was significantly higher in patients with SCLCs but lower in those with SCCs, the peak M/Z 1205 is lower in patients with SCLCs but higher in those with SCCs. Patients with ACs displayed intermediate levels of these two peaks. ROC analysis showed that the differences in the protein peaks M/Z1205, 4673, 1429, and 4279 may be sufficient to differentiate SCCs from SCLCs individually, whereas the discriminatory power was further improved by combination of M/Z 1205 and M/Z 4279. The difference in the peaks M/Z 1205 and 1429 appeared better to discriminate SCCs from ACs. In addition, the peak M/Z 4279 may differentiate ACs from SCLCs. Nevertheless, these results need to be further verified with prospective studies in larger cohort of patients.

In summary, we have shown in this study that SELDI-TOF-MS provides a rapid approach to assist in determining the histological subtypes of lung cancer using blood samples from patients. This may be particularly useful when tumors are located in difficult anatomical sites and/or the patients are in poor physical conditions that precludes biopsy before treatment. However, the protein identities of the identified peaks are unknown. Purification of serum proteins remains a technique challenge in the field of proteomics. Combination of Nano Aquity Ultra Performance Liquid Chromatography and LTQ Orbitrap may shed light in this regard.

ACKNOWLEDGMENTS

This study was supported by the scientific and technological project of Shanxi province (scholarship council of China (2013-133)). The authors thanked the pathology department of Shanxi Cancer Hospital and Institute to provide the data of pathology.

REFERENCES

1 Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013 (Jan); 63 (1): 11–30.
2 Ettinger DS. Ten years of progress in non-small cell lung cancer. J Natl Compr Canc Netw 2012 (Mar); 10 (3): 292–5.
3 Ahmed K, Emran AA, Jesmin T, Mukti RF, Rahman MZ, Ahmed F. Early detection of lung cancer risk using data mining. Asian Pac J Clin Oncol 2013; 14 (1): 595–8.
4 Normanno N, Rachiglio AM, Roma C et al. Molecular diagnostics and personalized medicine in oncology: challenges and opportunities. J Cell Biochem. 2013 (Mar); 114 (3): 514–24.
5 McCormack VA, Boffetta P. Today’s lifestyles, tomorrow’s cancers: trends in lifestyle risk factors for cancer in low- and middle-income countries. Ann Oncol 2011 (Nov); 22 (11): 2349–57.
6 Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev 2010 (Aug); 19 (8): 1893–907.
7 Wen C, Dehnel T. China wrestles with lung cancer. Lancet Oncol 2011; 12: 15.
8 She J, Yang P, Hong Q, Bai C. Lung cancer in China: challenges and interventions. Chest 2013 (Apr); 143 (4): 1117–26.
9 Travis WD. Pathology of lung cancer. Clin Chest Med 2011; 32: 669–92.
10 Mok TS, Wu YL, Thongprasert S et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009; 361: 947–57.
11 Camidge DR, Bang YJ, Kwak EL et al. Activity and safety of crizotinib in patients with ALK-positive non-small-cell lung cancer: updated results from a phase 1 study. Lancet Oncol 2012; 13: 1011–9.
12 Scagliotti GV, Parikh P, van Pawel J et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008; 26: 3543–51.
13 Yamagishi T, Shimizu K, Ochi N et al. Histological comparison between preoperative and surgical specimens of non-small cell lung cancer for distinguishing between “squamous” and “non-squamous” cell carcinoma. Diagn Pathol 2014; 29 (9): 103.
14 Chu XY, Hou XB, Song WA, Xue ZQ, Wang B, Zhang LB. Diagnostic values of SCC, CEA, Cyfra21-1 and NSE for lung cancer in patients with suspicious pulmonary masses: a single center analysis. Cancer Biol Ther 2011 (Jun 15); 11 (12): 995–1000.
15 Meng Z, Veenstra TD. Targeted mass spectrometry approaches for protein biomarker verification. J Proteomics 2011 (Nov 18); 74 (12): 2650–9.
16 Ye J, Liu H, Kirmiz C, Lebrilla CB, Rocke DM. On the analysis of glycomics mass spectrometry data via the regularized area under the ROC curve. BMC Bioinformatics 2007 (Dec 12); 8: 477.
17 Wilson GD, Geddes TJ, Pruett BL et al. SELDI-TOF-MS serum profiling reveals predictors of cardiac MRI changes in marathon runners. Int J Proteomics 2012; 2012: 679301.
18 Qian JY, Mou SH, Liu CB. SELDI-TOF MS combined with magnetic beads for detecting serum protein biomarkers and establishment of a boosting decision tree model for diagnosis of pancreatic cancer. Asian Pac J Cancer Prev 2012; 13 (5): 1911–5.

19 Wegdam W, Moerland PD, Meijer D et al. A critical assessment of SELDI-TOF-MS for biomarker discovery in serum and tissue of patients with an ovarian mass. Proteome Sci 2012 (Jul 23); 10 (1): 45.

20 Cao XL, Li H, Yu XL et al. Predicting early intrahepatic recurrence of hepatocellular carcinoma after microwave ablation using SELDI-TOF proteomic signature. PLoS One 2013 (Dec 13); 8 (12): e82448.

21 Travis WD, Brambilla E, Noguchi M et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multi-disciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011; 6: 244–85.

22 Navani N, Brown JM, Nankivell M et al. Suitability of endobronchial ultrasound-guided transbronchial needle aspiration specimens for subtyping and genotyping of non-small cell lung cancer: a multicenter study of 774 patients. Am J Respir Crit Care Med 2012; 185: 1316–22.

23 Wang XY, Yu H, Linnoila RI et al. Masushil as a potential therapeutic target and diagnostic marker for lung cancer. Oncotarget 2013 (May); 4 (5): 739–50.

24 Marchetti A, Ardizzoni A, Papotti M et al. Recommendations for the analysis of ALK gene rearrangements in non-small-cell lung cancer: a consensus of Italian Association of Medical Oncology and Italian Society of Pathology and Cytopathology. J Thorac Oncol 2013 (May); 8 (3): 352–8.

25 Broggini M, Garassino MC, Damia G. Evaluation of safety and efficacy of tivantinib in the treatment of inoperable or recurrent non-small-cell lung cancer. Cancer Manage Res 2013; 5: 15–20.

26 Travis WD, Brambilla E, Riely GJ. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. J Clin Oncol 2013; 31: 992–1001.

27 Vollmer E, Schultz H, Stellmacher F et al. Tumors in the lung–morphologic features and the challenge of integrating biomarker signatures into diagnostics. Rom J Morphol Embryol 2010; 51 (4): 607–14.

28 Ulivi P, Mercatanti L, Casoni GL et al. Multiple marker detection in peripheral blood for NSCLC diagnosis. PLoS One 2013; 8(2): e57401.

29 Wang XM, Li J, Yan MX et al. Integrative analyses identify osteopontin, LAMB3 and ITGB1 as critical pro-metastatic genes for lung cancer. PLoS One 2013; 8 (2): e55714.

30 Liu K, Zhuang X, Mai Z. p73 expression is associated with cellular chemosensitivity in human non-small cell lung cancer cell lines. Oncol Lett 2013 (Feb); 5 (2): 583–7.

31 Cairns D, Perkins D, Stanley A et al. Integrated multi-level quality control for proteomic profiling studies using mass spectrometry. BMC Bioinformatics 2008; 4 (9): 519.

32 Simsek C, Sonmez O, Keyf AI et al. Importance of serum SELDI-TOF-MS analysis in the diagnosis of early lung cancer. Asian Pac J Cancer Prev 2013; 14 (3): 2037–42.

33 Song QB, Hu WG, Wang P et al. Identification of serum biomarkers for lung cancer using magnetic bead-based SELDI-TOF-MS. Acta Pharmacol Sin 2011 (Dec); 32 (12): 1537–42.

34 Jiang M, Gu G, Ni B et al. Detection of serum protein biomarkers by surface enhanced laser desorption/ionization in patients with adenocarcinoma of the lung. Asia Pac J Clin Oncol 2014 (Jun); 10 (2): e7–12.

35 Selvaraju S, Rassi ZE. Liquid-phase-based separation systems for depletion, prefractionation and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis—an update covering the period 2008–2011. Electrophoresis 2012 (Jan); 33 (1): 74–88.

36 Zhang J, Guan S, Sun J et al. Characterization and profiling of phenolic amides from Cortex Lycii by ultra-high performance liquid chromatography coupled with LTQ-Orbitrap mass spectrometry. Anal Bioanal Chem 2015 (Jan); 407 (2): 581–95.