

Proteomic Analysis and Identification of Lung Cancer Associated Proteins by Two-dimensional Gel Electrophoresis and Matrix-assisted Laser Desorption Ionization Time of Flight Mass Spectrometry

Huizhen Zhang¹, Hongxiang Guo², Qingtang Fan¹, Yiming Wu¹

1 College of Public Health, Zhengzhou University, Zhengzhou 450052, China

2 College of Life Sciences, Henan Agriculture University, Zhengzhou 450052, China

Abstract objective To find the cancer associated proteins in cancerous lung tissues, and further identify new molecular markers for diagnosis and prognosis of lung carcinoma, as well as develop new therapeutic targets of the disease. **Methods** The soluble proteins from matched pairs of lung cancer tissues were analyzed for their two-dimensional electrophoresis (2DE) profiles. With the help of special image analysis software, up-regulated protein spots in two-dimensional electrophoresis gels of cancerous lung tissues were found compared with normal lung tissues. These spots were cut off from Coomassie Brilliant Blue staining gels, digested in gel with L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-trypsin and then identified by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF-MS). **Results** Tumor-specific protein spots were identified, including Apolipoprotein A-I precursor (Apo-AI), Peptidyl-prolyl cis-trans isomerase A, Calgranulin B (MRP-14), Calgiz-zarin (S100C protein), Ras-related protein Rab-14, apolipoprotein E. Among these 6 proteins, the potential significance of the differential expressions is discussed. **Conclusion** These findings demonstrate that differential expression analysis of proteomes may be useful for the development of new molecular markers for diagnosis and prognosis of lung cancer.

Key words proteome; two-dimensional electrophoresis (2DE); matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS); lung cancer, associated protein

Lung cancer is the leading cause of cancer death worldwide, and the number of lung cancer patients has been increased with the exposure to the environmental risk factors^[1-4]. Prevention of lung cancer is a serious worldwide challenge. It is clear that the successful prevention of lung cancer will depend on the reduction of risk factors along with better methods of screening and early detection. Tumor markers are widely used for

screening, diagnosis, staging, prognosis, monitoring response to treatment, and detection of recurrent disease^[5]. The analysis of proteins overexpressed in lung cancer, and making the proteins serve as tumor markers, has been the subject of extensive research.

Although many insights into the molecular pathology of lung tumors have been achieved, additional information is critical to our understanding of the development and progression of these tumors as well as to aid in early diagnosis. The most commonly evaluated markers include neuron-specific enolase, carcinoembryonic antigen, cytokeratin 19 fragments (CYFRA21-1), squamous cell carcinoma antigen, cancer antigen CA 125, and tissue polypeptide antigen^[6].

The detection of new candidate markers is complex because of the known heterogeneity of lung cancers. Methods have been developed to identify tumor associ-

*Grant item: the national "211 Project" and the National Natural Science Foundation of China (No.30371695).

Correspondence to: Huizhen Zhang, MD

Tel: 371-677-81461

Email:Huizhenzhang@zzu.edu.cn

Huizhen2@hotmail.com

ated antigens such as molecular cloning in system or using a biochemical strategy based on the extraction of antigenic peptides bound to major histocompatibility complex class 1 molecules from tumor cells. These methods have allowed the recognition of certain human tumor antigens [7]. However, no evidence has been obtained indicating that the detection of these markers precedes clinical diagnosis of lung cancer.

Proteome is defined as the total proteins expressed by a genome and the comparative analyses of proteomes from cancer cells or tissues promise the discovery of new biomarkers for early detection of cancers^[8,9]. In this proteomics approach, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and mass spectrometry have been the most important technologies for the separation and identification of proteins respectively^[10,11]. 2D-PAGE is a powerful research technique, which makes it possible to simultaneously examine hundreds of polypeptides in a tissue sample. It has been widely used for the detection and identification of potential tumor markers^[12]. Using 2D-PAGE and matrix assisted laser desorption/ionization-time of flight-mass spectrometry (MALDI-TOF-MS), we tried to identify the biomarkers of lung cancer by the comparative analysis of proteomes from human normal lung tissues and cancerous lung tissues, help elucidate the molecular mechanism of cellular events associated with cancer progression, such as cellular signaling^[13,14].

MATERIALS AND METHODS

Reagents

Immobiline DryStrips (pH 3-10, 17 cm), DryStrip coverfluids, immobilized pH gradient (IPG) buffer, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), bromophenol blue, agarose, acrylamide, Tris-base, glycine, sodium dodecyl sulfate (SDS), N,N,N',N'-tetramethylethylenediamine (TEMED), and Coomassie brilliant blue R-250, Dithiothreitol (DTT), iodoacetamide, acetic acid, Trypsin (sequencing grade), Trifluoroacetic acid. The remaining chemicals were of analytical grade. All buffers were prepared with Milli-Q water.

Tissues and extract preparation

Lung cancer tissues and corresponding adjacent non-cancerous lung tissues were obtained with informed consent from patients at the First Affiliated Hospital of Zhengzhou University. Cancer samples were obtained from the "core" part of the tumor to avoid the adjacent noncancerous tissue. For each of the normal tissues, surface epithelium was procured selectively by dissection with special care for minimal contamination of non-epithelial cells, and samples were immediately snap-frozen in liquid nitrogen. They were classified histologically according to Lauren's classification after hematoxylin and eosin staining.

Fragments of normal and malignant tissues were sharp dissected and homogenized with a homogenizer in 2mL fresh lysis buffer [7mol/L urea, 2mol/L thiourea, 40mmol/L Tris, 40g/L 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 100mmol/L dithiothreitol, 0.5mmol/L PMSF, 0.5mmol/L EDTA, 2%(V/V)NP-40, 2%(V/V)Bio-Lyte 3/10, 1%(V/V)TritonX-100], then placed into tubes on ice for dealing with ultrasonic for 10 min. The mixture was centrifuged at 1 000r/min for 5min to remove tissue and cell debris, then centrifuged in a tabletop ultracentrifuge at 12000g for 30min at 4°C. The supernatant was pipetted off and stored at -80°C until use. These were used as the 2DE samples for the soluble fraction. Protein concentration of 2-DE samples was estimated according to a commercial Bradford reagent. BSA was used as standard.

2D electrophoresis and image analysis

First-dimension iso-electric focusing (IEF) was carried out on a ProteanIEF cell (Bio-Rad) using precast 17cm pH 3-10 IPG gelstrips (Bio-Rad). 300µg total protein for silver staining gels (1000µg for Coomassie Brilliant Blue staining gels) was mixed with rehydration solution (8 M urea, 2% CHAPS, 50 mM DTT, 0.2% Bio-Lyte 3/10 ampholyte, and 0.001% bromophenol blue) to a total volume of 500µL. Rehydration and IEF were carried out as follows: 12h of passive rehydration, IEF at 500 V for 30 min, 5000 V for 3h, and 10,000 V for 8h. The current was limited to 70µA per gel strip. All IEF steps were carried out at 17°C. After IEF sepa-

ration, the gel strips were immediately equilibrated using two steps in equilibrium buffer containing 50 mM Tris-HCl (pH 8.8), 6 M urea, 30% glycerol, and 2% SDS. At the first step, 2% DTT (W/V) was included in the equilibrium buffer. 2.5% iodoacetamide (W/V) was added at the second step. The second dimension separation (13% T, 2.7% C) was carried out using Tris-glycine buffer containing 1g/L SDS on a Protean II xi 2-D cell (Bio-Rad) according to the following procedure: 5mA/gel for the initial 0.5h and 18mA/gel thereafter and a temperature of 4°C, until the bromophenol blue front reached the bottom of the gel.

For silver staining, gels were immersed in ethanol: acetic acid: water(35:7:58) for 1.5h, followed by washed twice in deionized water for 20 min. Gels were pre-treated for 1 min in a solution of 0.2g/L Na₂S₂O₃ and followed by 3 of 1-min washes in deionized water. Proteins were stained in a solution containing 2g/L AgNO₃ and 0.075% formalin (37g/L formaldehyde in water) for 20 min, and washed twice in deionized water for 1 min. Subsequently, gels were developed in a solution of 0.6g/L formaldehyde, 20g/L Na₂S₂O₃ and 0.004 g/L Na₂S₂O₃. When the desired intensity was attained, the developer was discarded and reaction was stopped by 10g/L EDTA-Na₂. For Coomassie Brilliant Blue staining of gels, gels were equilibrated in a solution containing 500ml/L methanol, 50 ml/L acetic acid and 25g/L Coomassie Brilliant Blue R-250. Gels were rinsed in 300ml/L ethanol containing 70ml/L acetic acid.

To account for experimental variation, three batches of total proteins extracted from the experimental and control sample respectively, were subjected to 2-D electrophoresis and replicate gels were simultaneously run three times. Protein patterns in the gels after staining were recorded as digitalized images using a high-resolution scanner. Gel image matching was done with PDQuest software (version 6.2, Bio-Rad, Richmond, CA).

In-gel digestion

Protein spots on Coomassie blue stained gel were performed essentially as described. After the completion of staining, the gel slab was washed twice with water for

10 min. The spots of interest were excised with a scalpel and put into 1.5mL micro-tubes. The particles were washed twice with water and then twice with water/acetonitrile (1:1) for 15 min. The solvent volumes were about twice that of the gel. Liquid was removed, acetonitrile was added to the gel particles and the mixture was left for 2h. After that, liquid was removed and the particles were rehydrated in 25mmol/L NH₄HCO₃ for 5 min. Acetonitrile was added to produce a 1:1 mixture of 25mmol/L NH₄HCO₃/acetonitrile and the mixture was incubated for 15 min. All liquid was removed. Gel particles were dried in a vacuum centrifuge, reswelled in 10mmol/L dithiothreitol and 25mmol/L NH₄HCO₃, and incubated for 30 min at 56°C to reduce the peptides. After chillness of tubes to room temperature and removal of the liquid, 55mmol/L iodoacetamide in 25mmol/L NH₄HCO₃ was added. The tubes were incubated for 30 min at room temperature in the dark to S-alkylate the peptides. Then iodoacetamide solution was removed, the particles were washed with 25mmol/L NH₄HCO₃ and acetonitrile, dried in a vacuum centrifuge, rehydrated in digestion buffer containing 50mmol/L NH₄HCO₃ and 12.5ng/L trypsin (TPCK-treated, proteomics grade, Sigma, USA), incubated for 8 h~12h at 37°C. After digestion, 25mmol/L NH₄HCO₃ was added, and the tube was incubated for 15 min. Acetonitrile was added and the tube was incubated for another 15 min. The supernatant was recovered, and the extraction was repeated twice with 50g/L TFA/acetonitrile (1:1). The three extracts were pooled and dried in a vacuum centrifuge.

MALDI-TOF-MS identification of peptide mixtures and database searching

The peptide mixtures solubilized with matrix solution CCA (α-cyano-4-hydroxycinnamic acid) was spotted on the target and dried. Dried spots were analyzed in an REFLEX-III (Bluker) MALDI-TOF mass spectrometer. The spectrometer was run in positive ion mode and in reflector mode with the setting: accelerating voltage, 20 kV; grid voltage, 76%; guide wire voltage, 0.01%; and a delay time of 150ns. The low mass gate was set at 500m/z. Protein identification was carried out by peptide mass fingerprinting with searching

the protein databases of Swiss-Prot/TrEMBL (<http://www.expasy.ch/tools/peptident.html>) and Mascot (<http://www.matrixscience.com/>). The following search parameters were applied: a mass tolerance of 50ppm and one incomplete cleavage were allowed; acetylation of the N-terminus, alkylation of lysine by carboxyamidomethylation were considered as possible modifications.

RESULTS

Protein expression maps of paired samples and image analysis

Two-dimensional gel electrophoresis maps were constructed for human normal lung tissues and cancerous lung tissues proteins in Fig.1. We made separate maps for the pH range 3~10. The second-dimensional gel electrophoresis was performed with 12.5% SDS-PAGE. For each sample, over 800 protein spots were resolved in a 2-DE gel (20cm×20cm) with silver-stained by computer-aided image analysis, and over 500 protein spots were resolved with Coomassie Brilliant Blue-stained. Figure 1A (from cancer tissues) contains a total of 530 spots, whereas Figure 1B (from normal tissues) contains a total of 560 spots. A total of 515 spots from cancer tissues could be matched to those from normal tissues. In total, we were able to identify can-

cer-specific spots in 2DE gels. The positions of the identified proteins are shown in Figure 1A. All of the identified spots could be considered as abundant proteins because of Coomassie Blue staining.

In order to evaluate the reproducibility of the soluble protein preparations and to quantify the protein extracts for 2DE gel analysis, mini 2DE gels (7cm, pH3-10) were used at first. To ensure quality and reproducibility of results, three gels per sample were processed simultaneously with the same power supply and analyzed using PDQuest 2-D software (version 7.1, Bio-Rad, Richmond, CA). Protein extracts prepared from tissues were compared in this way and found to be highly reproducible. Protein patterns were analyzed using PDQuest software. Qualitative and quantitative comparisons were made between replicate groups, comprising the three gels for each sample. Quantitation is based on the peak intensity and area of Gaussian-fitted spots, allowing more accurate quantitation than summation of pixel intensities. Any differential spots that were not present or absent in all three gels within a replicate group were excluded.

To evaluate the potential for spurious differences that are not stage-specific, we also compared replicate groups in which gels were arbitrarily assigned to one of three groups, each containing one gel from each stage. No differences were detected between these groups, which suggested that the differentially represented spots

Table 1 MALDI-TOF-MS identification of overexpression proteins in lung cancer tissues

Spot*	AC**	Matched Peptide	Top score	Theoretic Mr	Sequence*** coverage(%)	Name and description
1	P02647	20	179	30759	42	ApolipoproteinA-I precursor (Apo-AI)
2	gi4557325	8	66	36246	34	apolipoprotein E precursor
3	P35287	9	66	23912	49	Ras-related protein Rab-14
4	P06702	6	60	13234	52	Calgranulin B (MRP-14)
5	P05092	10	62	17870	47	Peptidyl-prolyl cis-trans isomerase A
6	P31949	16	110	11733	54	(CyclophilinA) Calcizzarin (S100C protein)

* The numbering and lettering corresponding to the two-dimensional gel electrophoresis image in Fig. 1.

**Accession number in NCBIInr, SWISS-PROT.

***Percentage of the protein sequence covered by the matched peptides.

are genuinely sample-specific.

Identification of overexpression proteins using peptide mass fingerprinting

Overexpression proteins spots were excised from the 2-D gels of cancerous lung tissues and subjected to trypsin digestion and MALDI mass spectrometry. 6 spots were identified successfully. The criteria used to accept identifications including the extent of sequence coverage, the number of peptides matched, the probability score, and whether human protein appeared as the top candidates in the first pass search where no restriction was applied to the species of origin. The data for the protein spot 1 as an example are shown in Fig.2. Fig. 2A shows the MALDI-TOF MS peptide mass fingerprint spectrum of trypsin-digested protein spot 1. Fig. 2B is the probability based mowse score. Score is $-10^{\times} \text{Log}(P)$, where P is the probability that the observed match is a random event. Protein scores greater than 53 are significant ($P < 0.05$). The score of two proteins is significant, we choose the protein with the highest score. Fig.2C lists the matching peptides. Twenty peptides were matched with ApolipoproteinA-I, accession-numbered as P02647 in the protein database SWISS-PROT. These 20 peptides are indicated in the

protein sequence shown in fig.2D. The results of identification are summarized in the table 1. These identified protein spots are numbered as shown in Fig.1.

DISCUSSION

Proteomic study means the analysis of the entire protein complement expressed by a genome [15]. Classically, proteome analysis consists of three steps: Two-dimensional gel electrophoresis (2-DE) used for proteins profiling, mass spectrometry (MS) for critical confirmation of the molecular weights and bioinformatics for protein identification. Comparative proteome analysis is a good strategy to discover proteins that undergo changes in expression level and may underlie the differences of phenotype [16]. 2-DE with its recent developments has been seen as an ideal tool for proteome analyses. Immobilized pH-gradient (IPG) strips are used in the first-dimensional gel electrophoresis to provide a basis for reproducible separation according to proteins isoelectric points. Advanced computer graphics and image analysis systems have offered a possibility for quantitative detection of protein spots, in addition to edition and storage of the 2-DE images. Proteome comparison between tissues in the different situations can cast a light

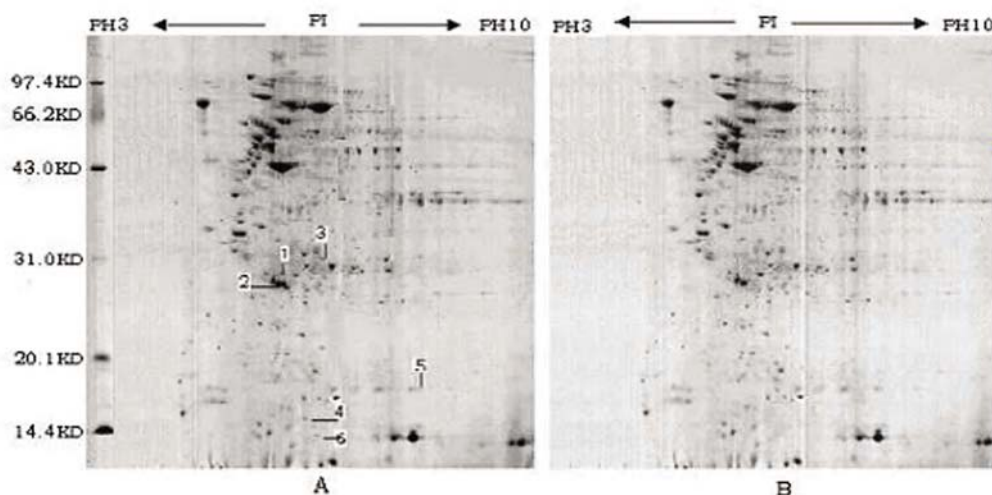


Fig. 1 2-D gel images of protein expression in lung cancer tissue and normal lung tissue

A: was from lung cancer. B: from normal lung sample of the same patient.

Proteins were separated on pH 3-10 linear IPG strip in the first dimension and 125 g/L SDS-PAGE in the second dimension. All labeled spots were tumor-specific.

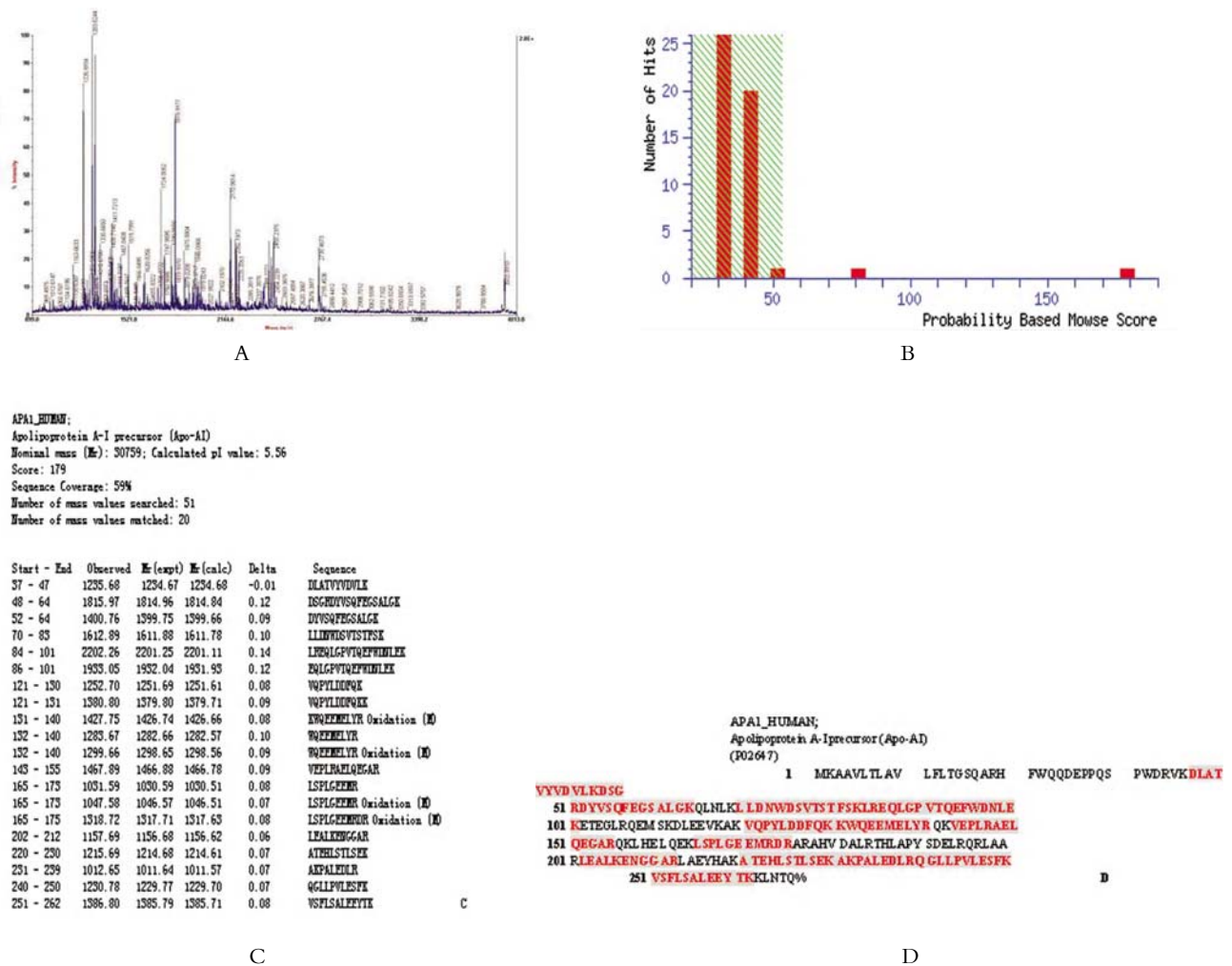


Fig. 2 Analysis of spot NO.1 from human lung 2-DE map (pH 3-10) by MALDI-TOF MS.

A MALDI-TOF MS peptide mass fingerprint spectrum obtained from crude peptide mixture after in-gel tryptic digest of spot NO.1.

B Probability based mouse score

C The list of matching peptides between the experimental and theoretical values.

D The sequence of Apolipoprotein A-I identified. The matched peptides are shadowed in the sequence.

on some protein spots, which are differently expressed in quantity under different surroundings. Nowadays, mass spectrometry such as matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF-MS) [17], can routinely identify and characterize proteins. Peptide mass fingerprinting identifies a protein based on the molecular weights of its peptides obtained by MS after digestion by trypsin [18]. The developed bioinformatics allows the identification of most proteins [19].

In this paper, we reported comparative proteome studies between normal lung tissues and cancerous lung tissues. The pH 3-10 range in the first-dimensional gel

electrophoresis separates the soluble proteins in the lung tissues. All paired samples, which would be compared each other, were run together at the same conditions. This avoids any artificial differences between samples due to the different running conditions.

The 2-DE patterns of normal lung tissues and cancerous lung tissues were compared each other with the help of the software PDQest, and different expressional protein spots were found. These protein spots have a consistent increase in all 8 independent paired samples. Over-expressed protein spots were excised from the gel and digested with trypsin. Molecular weights of the tryptic peptides were determined by MALDI-TOF-

MS. Obtained protein scores are significant ($P < 0.05$) in the protein database search. 19%~72% coverage of protein sequences was obtained in the analyses of MALDI-TOF MS.

Number 4 and number 6, were identified as Calgranulin B (S100A9, MRP14) and Calgizzarin (S100A11, S100C protein), two members of the S100 family. S100 proteins comprise a family of 10~14 kDa EF-hand-containing calcium binding proteins that function to transmit calcium-dependent cell regulatory signals, which are specifically expressed in a variety of tissues and cell types. It is thought that they are primarily involved in Ca^{2+} -mediated signal transduction. Changes in intracellular calcium levels alter the structure and function of these proteins^[20]. They have been frequently described in association with cell growth and differentiation, with cell cycle regulation and with several human diseases such as cancer and skin disorders. Studies suggest that these members of the S-100 protein family have a role in cytoskeletal changes seen in various skin diseases. Calgizzarin (S100A11) is a calcium-binding protein implicated in a variety of biologic functions such as proliferation and differentiation in cancer. Calgizzarin, a new tumor marker protein in nasopharyngeal carcinoma (NPC)^[21], plays a significant role in tumor suppression and it is involved in prostate cancer development and progression. Calgranulin B represents a novel molecular parameters of the early events of inflammatory reactions that reveal interesting aspects for the pathomechanism of chronic inflammatory reactions^[22]. Tumor specific protein markers were identified in colon tumors as calgranulin B that is also upregulated in colorectal cancer^[23, 24]. Calgizzarin and Calgranulin B present in lung cancer has not been documented.

Apolipoprotein A-I (apoA-I) is the major protein constituent of high density lipoproteins (HDL) and lymph hylomicrons. In human, proapoA-I is synthesized as a precursor protein, preproapoA-I, of 267 amino acids and is thought to occupy a surface position on the lipoprotein. ApoA-I activates lecithin-cholesterol acyltransferase, which is the cholesterol-esterifying enzyme of plasma involved in the production of mature circulating HDL^[25]. Apolipoprotein E (apoE), a protein with three common isoforms, has a large impact on

longevity and successful aging. Impairments in cognitive performance have been observed in aged apolipoprotein (apoE)-deficient mice^[26]. In this study, apolipoprotein (apo)A-I and Apolipoprotein E are found to be up-regulated in lung cancer.

Cyclophilins (CyPs) are a large class of highly conserved ubiquitous peptidyl-prolyl cis-trans isomerases. CyPs have also been identified as a specific receptor for the immunosuppressive drug cyclosporin A and are involved in a variety of biological functions. Cyclophilin A (CypA) is a cytosolic protein that has many biological functions including immune modulation, cell growth, tumorigenesis, and vascular disease. The objective of this study was to determine the effect of CypA on cell proliferation and several gene expressions in human endothelial cells and vascular smooth muscle cells^[27]. Cyclophilin A (CyPA) is overexpressed in non-small cell lung cancer (NSCLC), but it was not found that CyPA is of prognostic significance^[28]. This paper reported Cyclophilin A (CypA) overexpressed in lung cancer, but its function of a biomarker for lung cancer need further study to validate.

In conclusion, the differences of the proteins between normal and cancerous lung tissues are complex. To affirm proteins studying above be cancer-associated of lung cancer and to serve for the further basic and clinical investigation, we need do some other works such as enlarging sample, purifying and analyzing these differential proteins, and study the actions during the multistage process of lung cancer.

ACKNOWLEDGEMENTS

This work was supported by the key disciplines granted from the national "211 Project" and the National Natural Science Foundation of China (No. 30371695). This work was carried out in Henan Key Laboratory for Molecular Medicine.

REFERENCES

1. Greenlee RT, Murray T, Bolden S, *et al.* Cancer statistics. *CA Cancer J Clin*, 2000, 50: 27-33.
2. Kamp DW, Graceffa P, Pryor WA, *et al.* The role of free radicals in asbestos-induced diseases. *Free Radical Biol Med*,

- 1999, 12: 293–315.
3. Weitzman SA, Graceffa P. Asbestos catalyzes hydroxyl and superoxide radical generation from hydrogen peroxide. *Arch Biochem Biophys*, 1984, 228: 373–376.
 4. Bailey–Wilson JE, Pugh EW, Wiest JS, *et al.* Lung cancer: genetic epidemiology. In: Bertino JR, editor. *Encyclopedia of cancer*. San Diego, CA, Academic Press: 995–1003.
 5. Lindblom A, Liljegren A. Tumor markers in malignancies. *Br Med J*, 2000, 320: 424–427.
 6. Stieber P, Aronsson AC, Bialk P, *et al.* Tumor markers in lung cancer: EGTM recommendations. *Anti-cancer Res*, 1999, 19: 2817–1819.
 7. Lawrie LC, Fothergill JE, Murray GI. Spot the differences: proteomics in cancer research. *Lancet Oncol*, 2001, 2: 270–277.
 8. Simpson RJ, Dorow DS. Cancer proteomics: signaling networks to tumor markers. *Trends Biotechnol*, 2001, 19 (10 Suppl): S40–S48.
 9. Bichsel VE, Liotta LA, Petricoin EF 3rd. Cancer proteomics: from biomarker discovery to signal pathway profiling. *Cancer J*, 2001, 7: 69–78.
 10. Gorg A, Obermaier C, Boguth G, *et al.* The current state two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*, 2000, 21: 1037–1053.
 11. Lahm HW, Langen H. Mass spectrometry: for the identification of proteins separated by gels. *Electrophoresis*, 2000, 21: 2105–2114.
 12. Okuzawa K, Franzen B, Lindholm J, *et al.* Characterization of gene expression in clinical lung cancer materials by two-dimensional polyacrylamide gel electrophoresis. *Electrophoresis*, 1994, 15: 382–390.
 13. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A*, 2000, 97: 9390–9395.
 14. Ni XG, Zhao P, Liu Y, *et al.* Application of proteomic approach for solid tumor marker discovery. *Aizheng [Chinese journal of cancer]*, 2003, 22: 664–667.
 15. Wilkins MR, Sanchez JC, Gooley AA, *et al.* Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev*, 1996, 13: 19–50.
 16. O'Farrel PZ. High-resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, 1975, 250: 4007–4021.
 17. Nordhoff E, Egelhofer V, Giavalisco P, *et al.* Large-gel two-dimensional electrophoresis-matrix-assisted laser desorption/ionization time of flight mass spectrometry: an analytical challenge for studying complex protein mixtures. *Electrophoresis*, 2001, 22: 2844–2855.
 18. Mann M, Hojrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol Mass Spectrom*, 1993, 22: 338–345.
 19. Tripathi KK. Bioinformatics: the foundation of present and future biotechnology. *Curr Sci*, 2000, 79: 570–575.
 20. Ruse M, Lambert A, Robinson N, *et al.* S100A7, S100A10, and S100A11 are transglutaminase substrates. *Biochemistry*, 2001, 40: 3167–3173.
 21. Fung LF, Lo AK, Yuen PW, *et al.* Differential gene expression in nasopharyngeal carcinoma cells. *Life Sci*, 2000, 67: 923–936.
 22. Sorg C. The calcium binding proteins MRP8 and MRP14 in acute and chronic inflammation. *Behring Inst Mitt*, 1992, 91: 126–37.
 23. Chaurand P, DaGue BB, Pearsall RS, *et al.* Profiling proteins from azoxymethane-induced colon tumors at the molecular level by matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics*, 2001, 1: 1320–1326.
 24. Jungblut PR, Zimny-Arndt U, Zeindl-Eberhart E, *et al.* Proteomics in human disease: cancer, heart and infectious diseases. *Electrophoresis*, 1999, 20: 2100–2110.
 25. Ghiselli G, Gotto AM, Tanenbaum S, *et al.* Proapolipoprotein A-I conversion kinetics in vivo in human and in rat. *Proc Natl Acad Sci U S A*, 1985, 82: 874–878.
 26. Law A, Gauthier S, Quirion R. Alteration of nitric oxide synthase activity in young and aged apolipoprotein E-deficient mice. *Neurobiol Aging*, 2003, 24: 187–190.
 27. Huang LL, Zhao XM, Huang CQ, *et al.* Structure of recombinant human cyclophilin J, a novel member of the cyclophilin family. *Acta Crystallogr D Biol Crystallogr*, 2005, 61: 316–321.
 28. Howard BA, Zheng Z, Campa MJ, *et al.* Translating biomarkers into clinical practice: prognostic implications of cyclophilin A and macrophage migratory inhibitory factor identified from protein expression profiles in non-small cell lung cancer. *Lung Cancer*, 2004, 46: 313–323.