DOI 10.1002/pmic.200800777 *Proteomics* 2009, *9*, 5544–5552

RESEARCH ARTICLE

Protein profiling in human sera for identification of potential lung cancer biomarkers using antibody microarray

Min-Kyu Han[1,**], Young-Hee Oh[1*], Jimin Kang[2], Young-Pil Kim[1], Soowon Seo[2], Jhingook Kim[3], KeunChil Park[4] and Hak-Sung Kim[1,5]*

[1] Department of Biological Sciences, KAIST, Daejeon, Korea
[2] Department of Biomedical Engineering Interdisciplinary Course, Sungkyunkwan University School of Medicine, Suwon, Korea
[3] Departments of Thoracic Surgery and Pathology, Sungkyunkwan University School of Medicine, Suwon, Korea
[4] Department of Medicine, Sungkyunkwan University School of Medicine, Suwon, Korea
[5] Graduate School of Nanoscience and Technology (WCU), KAIST, Daejeon, Korea

To identify potential biomarkers of lung cancer (LC), profiling of proteins in sera obtained from healthy and LC patients was determined using an antibody microarray. Based on our previous study on mRNA expression profiles between patients with LC and healthy persons, 19 proteins of interest were selected as targets for fabrication of an antibody microarray. Antibody to each protein and five nonspecific control antibodies were spotted onto a hydrogel-coated glass slide and used for profiling of proteins in sera of LC patients in a two-color fluorescence assay. Forty-eight human sera samples were analyzed, and expression profiling of proteins were represented by the internally normalized ratio method. Six proteins were distinctly down-regulated in sera of LC patients; this observation was validated by Wilcoxon test, false discovery rate, and Western blotting. Blind test of other 32 human sera using the antibody microarray followed by hierarchical clustering analysis revealed an approximate sensitivity of 88%, specificity of 80%, and an accuracy of 84%, respectively, in classifying the sera, which supports the potential of the six identified proteins as biomarkers for the prognosis of lung cancer.

Received: October 1, 2008
Revised: September 14, 2009
Accepted: September 16, 2009

Keywords:
Antibody microarray / Biomarker / Lung Cancer / Protein arrays / Protein profiling

1 Introduction

Lung cancer (LC) is responsible for about 25% of the total cancer-related mortality, and is a leading and most rapidly increasing cause of mortality in Korea and the United States [1, 2]. About 60% of patients diagnosed in stage 1 can survive for 5 years. But, a large number of LC patients are diagnosed in the advanced stages (2–4) of LC, and their 5-year survival rate is only 10–30% [2, 3]. To improve the LC survival rate, substantial improvements are required in diagnosis; for this, identification of LC biomarkers are crucial [4, 5]. Spiral computed tomography scans are known to be effective for diagnosing LC with relatively high sensitivity. Their specificity, however, is so low that benign diseases are misinterpreted as cancer. It was reported that combining computed tomography scans with biomarker

Correspondence: Professor Hak-Sung Kim, Department of Biological Sciences, Graduate School of Nanoscience and Technology (WCU), KAIST, Daejeon, Korea
E-mail: hskim76@kaist.ac.kr
Fax: +82-42-350-2610

Abbreviations: AQP5, aquaporin 5; **ARTN,** artemin precursor; **CKB,** creatine kinase brain; **FDR,** false discovery rate; **INR,** internally normalized ratio; **LC,** lung cancer; **MCM3,** minochromosome maintenance protein; **MMP7,** matrix metalloproteinase 7; **SAA,** serum amyloid A; **TAF9,** TATA-binding protein; **TGIF2,** 5′-TG-3′ interacting protein; **VEGF,** vascular endothelial growth factor

*These authors contributed equally to this work.

**Current address: Dr. Min-Kyu Han, Technology Foresight Center, KISTEP, Seoul, Korea.

expression profiling could greatly increase the ability to distinguish LC from benign diseases [4, 6].

A disease biomarker is defined as a physical sign or measurement that originates from a specific disease-associated biomolecule, and which can be discerned in laboratory procedures [7]. Thus, biomarkers can be effectively used for diagnosis and/or prognosis of specific diseases. Of many changes caused by specific diseases, differentially expressed levels of proteins in sera between healthy persons and patients can be measured more easily than by biopsy, offering a great opportunity to use them as disease biomarkers for diagnosis. From the clinical point of view, the use of a single biomarker for diagnosis has a risk of generating a false-positive result due to links with other diseases. Therefore, an expression pattern of multiple biomarker proteins is usually preferred for disease diagnosis with sufficient sensitivity and specificity rather than the use of single biomarker [8]. To find potential disease biomarkers, protein expression profiling of sera obtained from healthy persons and patients has been investigated using various methods including MALDI-MS [9], ELISA [4], microfluidic system [10], and micro-RNA profiling [11]. However, these techniques are laborious and time-consuming. As an approach to profiling protein expression in a high-throughput manner, protein microarray technology has proved effective for a variety of applications in a simple way, and can be accomplished using small amounts of serum sample and reagents [12]. Expression of proteins in several kinds of cancers has been profiled using protein microarrays [13–15], and potential biomarkers were identified [8, 16, 17]. Thousands of proteins are present in human serum and can be used as a source of disease biomarkers [9].

In this work, in an effort to identify potential biomarkers for LC, profiling of proteins in sera obtained from healthy persons and LC patients was determined using an antibody microarray. We selected 19 proteins as targets for fabrication of antibody microarray based on our previous study on mRNA expression profiling in lung tissues of LC patients [18]. Antibodies against these selected proteins and against three previously identified proteins that are differentially expressed proteins in LC (*i.e.* matrix metalloproteinase 7 (MMP7) [19], serum amyloid A (SAA) [4] and vascular endothelial growth factor (VEGF) [20]) were spotted onto a hydrogel-coated glass slide, and used for analysis of protein profiling in sera. Additionally, human IgG and *Escherichia coli* protein OppA were employed as positive and negative controls, respectively. Using a two-color fluorescence assay, analysis of 48 human sera revealed six proteins to be distinctly down-regulated in sera from LC patients. Differential profiling between the two groups was evaluated using Wilcoxon test, the false discovery rate (FDR) [21], and Western blotting. Blind testing of other 32 human serum samples was carried out to evaluate the potential of the six identified proteins as biomarkers for the prognosis of lung cancer.

2 Materials and methods

2.1 Antibodies and reagents

Nineteen antibodies were selected based on the mRNA profiling data and commercial availability for construction of the antibody microarray (Table 1). Anti-human IgG was purchased from Sigma-Aldrich (St. Louis, MO, USA). Anti-apoptosis-antagonizing transcription factor, anti-alcohol dehydrogenase 1B, anti-centromere protein K (FKSG14), and anti-hydroxymethylglutaryl-CoA lyase (HMGCL) were from Abnova (Taipei City, Taiwan). Anti-aquaporin 5 (AQP5), anti-artemin precursor (ARTN), anti-creatine kinase brain (CKB), anti-minochromosome maintenance protein (MCM3), anti-TATA-binding protein (TFIID), and anti-5′-TG-3′ interacting protein (TGIF) were from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Anti-rho guanine nucleotide exchange factor 10 was from Orbigen (San Diego, CA, USA). Anti-ARID domain-containing protein 4A and anti-acetylcholine receptor protein subunit epsilon precursor (CHRNE) were from Aviva Antibody Corporation (San Diego, CA, USA). Anti-defensin 6 precursor (DEFA6) was from Alpha Diagnostic (San Antonio, TX, USA). Anti-eukaryotic translation initiation factor 3 subunit 8 (eIF3S8) was from Bethyl Laboratories (Montgomery, TX, USA). Anti-G protein-coupled receptor 50 and anti-arfaptin-2 (RFP2) were from Genetex (San Antonio, TX, USA). Anti-CD94 (KLRD1) was from Serotec Antibodies (Raleigh, NC, USA). Anti-mitogen-activated protein kinase kinase kinase 6 (MAP3K6) was from Abgent (San Diego, CA, USA). Anti-MMP7 was from Calbiochem (Darmstadt, Germany). Anti-SAA was from Antigenix America (Huntington Station, NY,USA). Anti-VEGF was from R&D Systems (Minneapolis, MN, USA). Anti-OppA was a home-made antibody used in our previous work [22]. Three kinds of secondary antibodies, anti-rabbit IgG, anti-mouse IgG, and anti-goat IgG, conjugated with horse radish peroxidase were obtained from Bio-Rad Laboratories (Hercules, CA, USA). Cy3-labeled secondary antibodies were from Sigma-Aldrich. Most of these antibody solutions were used as received, and the concentration of anti-OppA was 0.2 mg/mL. Reagents used were of analytical grade.

2.2 Dye-labeling of serum proteins

All sera used were provided by Samsung Medical Center (Seoul, Korea) under approval of the local Institutional Review Board and with the informed consent of all subjects. Patients with various stages of both non-small cell LC and small cell LC were included. After clotting, the sera were isolated by centrifugation for 10 min at 4°C and stored at −70°C until analysis. Serum proteins were labeled with Cy3 or Cy5 dyes for a comparative fluorescence analysis as described in our previous work [22]. Briefly, serum was diluted to 15-fold with dye-conjugation buffer (50 mM

Table 1. Target proteins and antibodies used for fabrication of the antibody microarray

No.	Gene accession no.	Abbreviation	Full protein name	Concentration (mg/mL)	Type	Source
1		human IgG	Human immunoglobulin G	1	Poly	Rabbit
2		AATF	Apoptosis-antagonizing transcription factor	0.5	Poly	Mouse
3		ADH1B	Alcohol dehydrogenase 1B	0.5	Poly	Mouse
4	NM_000362	AQP5	Aquaporin 5	0.2	Poly	Rabbit
5		ARHGEF10	Rho guanine nucleotide exchange factor 10	0.2	Poly	Rabbit
6		ARID4A	ARID domain-containing protein 4A	0.2	Poly	Rabbit
7	NM_057090	ARTN	Artemin precursor	0.2	Poly	Goat
8		CHRNE	Acetylcholine receptor protein subunit epsilon precursor	0.2	Poly	Rabbit
9	NM_001152	CKB	Creatine kinase brain	0.2	Poly	Goat
10		DEFA6	Defensin 6 precursor	0.2	Poly	Rabbit
11		EIF3S8	Eukaryotic translation initiation factor 3 subunit 8	1	Poly	Rabbit
12		FKSG14	Centromere protein K	1	Mono	Mouse
13		GPR50	G protein-coupled receptor 50	1	Poly	Rabbit
14		HMGCL	Hydroxymethylglutaryl-CoA lyase	0.2	Mono	Mouse
15		KLRD1	CD94	0.5	Mono	Mouse
16		MAP3K6	Mitogen-activated protein kinase kinase kinase 6	0.2	Poly	Rabbit
17	NM_002388	MCM3	Minochromosome maintenance protein	0.2	Poly	Goat
18		RFP2	Arfaptin-2	0.5	Poly	Rabbit
19	NM_006880	TAF9	TATA-binding protein	0.2	Poly	Rabbit
20	NM_021809	TGIF2	5′-TG-3′ interacting protein	0.2	Poly	Rabbit
21		MMP7	Matrix metalloproteinase7	0.2	Mono	Mouse
22		SAA	Serum amyloid A	1	Mono	Mouse
23		VEGF	Vascular endothelial growth factor	0.5	Mono	Mouse
24		OppA	Periplasmic oligopeptide-binding protein precursor	0.2	Poly	Rabbit

sodium carbonate buffer, pH 9.0), and incubated with either *N*-hydroxysuccinimide ester-linked Cy3 or Cy5 dye (Amersham Biosciences, Buckinghamshire, UK) in the dye-conjugation buffer for 40 min at room temperature with a gentle mixing every 10 min in the dark. The coupling reaction was stopped with addition of 1 M Tris-HCl (pH 8.0) followed by further incubation for 20 min with gentle mixing every 10 min in the dark. Dye-labeled protein solution was separated from free dyes by centrifugation at $10\,000 \times g$ for 30 min using a Microcon YM-10 apparatus (Millipore, Billerica, MA, USA). The resulting solution was further centrifuged after addition of 50 volumes of 3% skimmed milk in PBS. Recovered protein was diluted 50-fold with PBS. As a reference sample, eight healthy sera were pooled, labeled according the same procedure, and used in all experiments. All labeled-sera were stored at -70°C prior to use.

2.3 Fabrication of antibody microarray and protein profiling

The antibody microarray was constructed with a hydrogel-coated glass slide (Slide H, Schott Nexterion, Jena, Germany) using a robotic arrayer (Microsys, Cartesian Technologies, Irvine, CA, USA) equipped with CMP 3 spotting pins (Telechem International, Sunnyvale, CA, USA). The glass slides remained stable for months when stored at -20°C. As shown in Fig. 1, a single antibody microarray was composed of a set of 24 antibodies. All protein solutions for spotting were supplemented with 0.5% trehalose, and a relative humidity in the arrayer was maintained at 70%. Each antibody solution was deposited in six spots with a 2×3 format. After incubation for 1 h at room temperature, the resulting slide was treated with PBS supplemented with 1 M ethanolamine (Sigma-Aldrich) at 4°C overnight for blocking. Following a thorough washing with PBS, the microarray was incubated with dye-labeled samples in the dark for 1 h. Prior to a fluorescence analysis, the microarray was consecutively washed three times with PBST (PBS containing 0.1% Tween-20) and distilled water followed by drying with nitrogen gas. Expression profiling of proteins was carried out in duplicate.

2.4 Read out of fluorescence signal

Fluorescence intensities emitted from Cy5 and Cy3 were measured using a fluorescence scanner (GenePix Personal 4100A, Axon Instruments, Union City, CA, USA) and analyzed using GenePix Pro 4.1 software provided by the manufacturer. Both Cy5 and Cy3 intensities taken from all spots were corrected for the background intensity, and ratios of their median values were collected to calculate an average ratio for each antibody. The background fluorescence intensity was less than 100. Raw spots showing either obvious defects or a signal-to-noise ratio less than 3 were

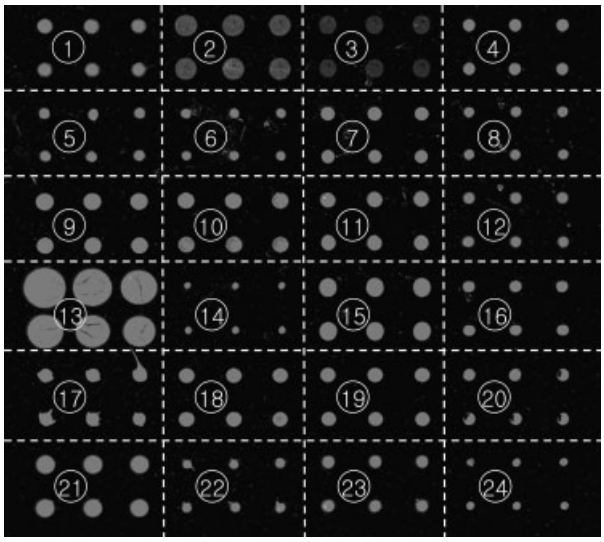

Figure 1. Layout of antibody spots in the antibody microarray. Spotted antibodies were visualized using Cy3-labeled secondary antibodies. Position of each antibody was indicated according to the number scheme presented in Table 1.

discarded prior to further analysis. Protein expression ratios between two samples were represented as internally normalized ratio (INR) as described elsewhere [23]. INR is determined using alternative labeling of two samples; Cy3-labeled reference and Cy5-labeled sample are mixed in equal amounts and applied on to the first antibody microarray, and then dye labeling is changed for the second microarray analysis. The intensity ratios of Cy5 to Cy3 are estimated on both microarrays, and INR is calculated as follows:

$$\begin{aligned} &\text{X : the Cy5/Cy3 ratio obtained Slide 1}\\ &\text{Y : the Cy5/Cy3 ratio obtained Slide 2}\\ &\text{XIY} = \frac{\text{sample}^{\text{cy5}}}{\text{reference}^{\text{cy3}}} \Big/ 00000\frac{\text{reference}^{\text{cy5}}}{\text{sample}^{\text{cy3}}}\\ &= \frac{\text{sample}^{\text{cy5}}}{\text{reference}^{\text{cy3}}} \times \frac{\text{sample}^{\text{cy3}}}{\text{reference}^{\text{cy5}}} = R_1 \times R_2 \end{aligned}$$

According to above equation, the sample-to-reference ratio, INR, is defined as square root of (X/Y). Although this method requires double microarrays and alternative dye-labeling, it can be effectively used as a universal condition for expression profiling of all proteins without external normalization factors, avoiding selecting of false-positive results by offsetting dye-labeling efficiency.

2.5 Western blotting

Serum was diluted to tenfold with PBS and SDS sample buffer. Following boiling, proteins in diluted serum were separated by SDS-PAGE. After the developed gel was equilibrated with transfer buffer (30 mM glycine, 48 mM Tris, 0.037% SDS, and 20% methanol, pH 8.3), proteins were electro-transferred from the gel onto a nitrocellulose membrane (Amersham Biosciences) at 60 V for 3 h. For immuno-blotting, the protein-transferred membrane was firstly blocked with 1% BSA in PBST (PBS containing 0.2% Tween20) and subsequently incubated with primary antibody and 1000-fold diluted secondary antibody conjugated with HRP. Antibody-bound proteins were visualized using a ECL Detection (Promega, USA).

2.6 Statistical analyses

K-mean and hierarchical clustering was performed using MultiExpriment Viewer program (TM4) (Version 4.4). Wilcoxon test was carried out to estimate a p value. To define a threshold of the differential expression, we used the FDR. When p value lower than 0.05 was applied as a criterion, an acceptable value of FDR (9.5%) was achieved. Thus, only the protein profiling result with p value < 0.05 was considered as a differential level.

3 Results and discussion

3.1 Construction of antibody microarray

Our previous study on expression profiling of mRNA in lung tissues of LC patients revealed that 47 genes were up- or down-regulated in lung tissues [18]. It is presumed that differential mRNA expression in LC tissue might affect not only the concentrations of corresponding proteins in lung tissue, but also those in serum through lysis and/or necrosis of cells or secretion [24]. As an analytical approach to profile the protein expression levels, several methods including two-dimensional gel electrophoresis and Western blotting have been widely used, but they are limited for analysis of low-abundance, large and hydrophobic proteins such as membrane proteins [25]. On the other hand, an antibody microarray offers some advantages over the conventional methods in terms of sensitivity, reagent amount, and throughput [23, 26–28].

To identify potential biomarkers of LC, we attempted an antibody microarray-based approach to investigate the expression levels of proteins in serum. Of 47 proteins encoded by differentially expressed mRNAs observed in our previous work [18], 19 proteins were selected as the target for fabrication of an antibody microarray by taking into consideration commercial availability of antibodies (Table 1). Five antibodies (no. 1 and no. 21–24) were further included as controls, and the total of 24 antibodies were used for the fabrication of the microarray. All antibodies were purchased from commercial sources except for the home-made anti-OppA. Each antibody was spotted in a 2×3 format onto a hydrogel-coated glass slide. The layout of the

spotted antibodies was visualized by probing the spots with Cy3-labeled secondary antibodies. As shown in Fig. 1, all spots displayed distinct fluorescence emissions compared with background, indicating proper immobilization of used antibodies. Fluctuation in fluorescence intensity and size of spots were observed, likely due to different buffer conditions and/or concentrations of antibodies in solution.

3.2 Protein profiling in human sera using antibody microarray

To identify potential protein biomarkers of LC, serum samples from 28 LC patients and eight healthy persons were analyzed against the pooled serum of healthy persons using the constructed antibody microarray. Individual serum and pooled serum were labeled with Cy dyes, mixed at equal volume, and applied to the antibody microarray as the test and reference samples, respectively. Comparative expression levels of proteins in sera were represented by INR, and the resulting ratios were converted into a $\log_2$ base for intuitive inspection. Following washing and drying, the fluorescence intensities from the antibody microarray was measured by fluorescence scanning. As expected, negligible signals were observed from the antibody employed as a negative control, anti-OppA (no. 24), in both samples (Fig. 2). Other antibody spots (no. 2, 3, 10, and 21) also displayed signal-to-noise ratios lower than 3 in sera from healthy persons (Fig. 2A) and LC patients (Fig. 2B). Thus, these spots were excluded from further analysis. These proteins might be expressed in sera at levels so low as to be virtually undetectable, or might exist only in lung tissue. Anti-human IgG (no. 1) employed as a positive control showed a regular circular morphology. Background intensity of the microarray was very low compared with those from the spots. Anti-KLRD1 (no. 15) displayed a strong green color due to auto-fluorescence. Spot-to-spot variations were less than 7% of coefficient of variation, and slide-to-slide variations showed coefficients of variation $<20\%$, which were almost same as those observed in our previous paper [22].

When healthy serum was applied as a test sample, fluorescence intensities and colors from the antibody spots in two microarrays were almost the same (Fig. 2A), indicating that the reference sample worked well for expression profiling of proteins in the antibody microarray. On the other hand, the use of serum of a LC patient as the test sample resulted in different colors and intensities in antibody spots (Fig. 2B). This result indicates that the antibody microarray can be effectively used for expression profiling of proteins in sera. In this work, the spots exhibiting a signal-to-noise ratio exceeding 3 were considered for further analysis.

As shown in Fig. 2C, of 16 fluorescence-producing antibodies against target proteins, six antibody spots exhibited distinct changes in the ratios of expression levels between

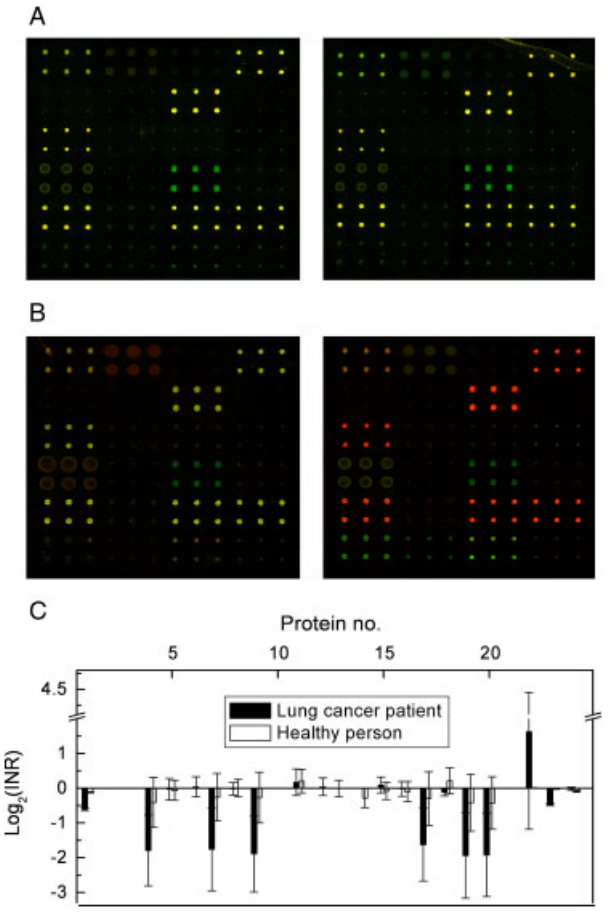

Figure 2. Representative images of antibody microarray when dye-labeled reference and test samples were mixed at equal amounts and applied. Each slide was incubated with Cy3-reference/Cy5-test sample (left) and Cy5-reference/Cy3-test sample (right). (A) Serum of a healthy person was used as test sample. (B) Serum of LC patient was employed as test sample. (C) Protein expression profiling of sera from LC patients and healthy persons. Ratios of LC patients (closed bars) and those of healthy persons (open bars) are shown as $\log_2$ values. Data represent the average values and standard deviations from duplicated experiments using four antibody microarrays.

healthy and patient sera. The six were identified to be AQP5 (no. 4), ARTN (no. 7), CKB (no. 9), MCM3 (no. 17), TAF9 (no. 19), and TGIF2 (no. 20). Interestingly, six identified proteins were found to be down-regulated in sera of LC patients. SAA (no. 22) used as a positive control showed up-regulation in LC sera as expected, which ruled out the possibility that the down-regulated ratios of six proteins in sera of LC patients might have originated from a systemic error. Previous expression profiling of mRNAs revealed that the genes encoding five down-regulated proteins also showed a similar trend in lung tissue, which strongly supports the reliability of the profiling results of proteins using an antibody microarray. In the case of MCM3, however, the opposite result was observed: up-regulation

in mRNA profiling and down-regulation by antibody microarray. This discrepancy with mRNA profiling might be due to undetectable protein levels by fluorescence-based assay or different serum sources. Two other control proteins, MMP7 (no. 21) and VEGF (no. 23), did not display any change in expression levels in sera between healthy and patients.

To further confirm the possibility that the six identified proteins could be biomarkers of LC, the protein profiles were additionally analyzed for sera from patients with benign lung disease ($n = 7$) and lymphoma ($n = 5$). In order for the identified proteins to be specific biomarkers of LC, levels of these proteins in sera of LC patients should be distinct from those of healthy persons and patients with other benign diseases. As shown in Fig. 3, the six identified proteins exhibited distinct profiles in LC patients. This was further investigated statistically using Wilcoxon test, which is non-parametric and not based on group mean values. As presented in Table 2, the p values for six proteins were estimated to be <0.05 when sera of LC patients were compared with other type of samples – healthy, benign, and lymphoma. Comparison of protein profiles between LC and lymphoma patients showed that the use of an expression pattern of biomarkers except TGIF2 and MCM3 could distinguish LC from lymphoma. On the other hand, the p values of comparison among other cases almost always exceeded 0.05. This analysis further supported validation of the experimental data. From the profiling of proteins using antibody microarray, six proteins (AQP5, ARTN, CKB, MCM3, TAF9, and TGIF2) were found to be differentially expressed in LC sera. Future studies will explore and exploit their biomarker capabilities in LC.

To our knowledge, this study is the first to identify AQP5, ARTN, and TGIF2 as potential LC biomarkers. Reports have described CKB expression and corresponding mRNA in LC-related biological samples; the results were equivocal, and expression or activity of CKB was up-regulated [29, 30] or down-regulated [24, 31]. This discrepancy might have resulted from different preparation methods and/or sources of samples. Indeed, the CKB expression may vary depending on cancer types; our study on mRNA profiling showed down-regulation of CKB in LC patients. On the contrary, mRNAs corresponding to MCM3 [32] and TFIID [33] were reported to be up-regulated in tissue, which is opposite to the results using antibody microarray. This also might be due to difference of target (mRNA *versus* protein) and sample source (tissue *versus* serum).

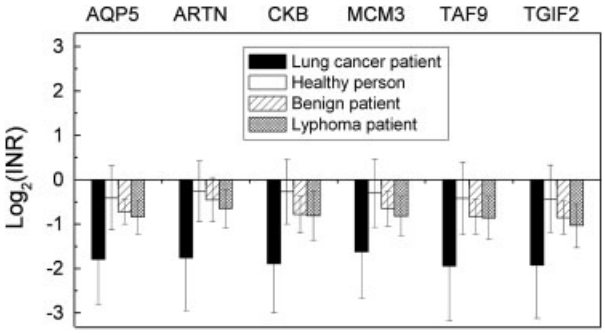

Figure 3. Profiling of six identified proteins in the sera of LC and benign disease patients against healthy persons. Log_2 values of the ratios are shown for LC patients (closed bars), healthy persons (open bars), benign lung disease patients (sparse slant bars), and lymphoma patients (dense cross-striped bars). Data represent the average values and standard deviations from duplicate experiments using four antibody microarrays.

3.3 Validation by Western blotting

Western blotting was performed to validate the antibody microarray profiling results. Of six proteins tested, five proteins could be validated by Western blotting. As expected, SAA (no. 22) as a control showed overexpression in LC sera. Their levels in sera between LC patient and healthy persons were found to be coincident with those using the antibody microarray as shown in Fig. 4.

3.4 Blind test with unknown sera

High sensitivity as well as specificity is required for reliable diagnosis of LC. To explore this, a blind test was conducted, in which new 32 serum samples including sera from LC patients and healthy persons were randomly allocated to

Table 2. Wilcoxon test for six proteins identified from the protein expression profiling by the antibody microarray

	AQP5	ARTN	CKB	TAF9	TGIF2	MCM3
LC *versus* healthy	**8.30E−04**	**7.10E−04**	**3.50E−04**	**2.10E−03**	**8.30E−04**	**2.10E−03**
LC *versus* benign	**6.70E−03**	**2.10E−03**	**3.30E−03**	**2.20E−02**	**2.00E−02**	**9.80E−03**
LC *versus* lymphoma	**2.60E−02**	**3.70E−02**	**2.00E−02**	**3.70E−02**	7.90E−02	5.20E−02
Healthy *versus* benign	4.80E−01	3.50E−01	1.70E−01	2.70E−01	2.30E−01	2.70E−01
Healthy *versus* lymphoma	2.20E−01	3.10E−01	2.80E−01	2.60E−01	8.50E−02	1.80E−01
Benign *versus* lymphoma	2.70E−01	3.20E−01	4.60E−01	4.40E−01	2.70E−01	3.20E−01

The p values between two samples were represented. The p values smaller than 0.05 were represented in bold.

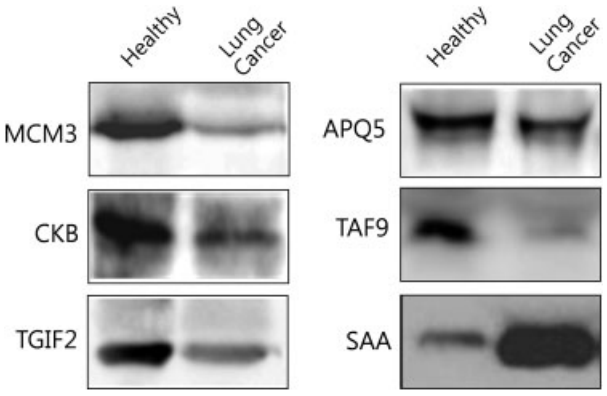

Figure 4. Western blotting of identified proteins. Pooled sera from healthy persons are shown in left lane, and right lane indicates the pooled sera from LC patients.

antibody microarray analysis without knowledge of sample origin. K-mean and hierarchical clustering was performed to classify unknown sera to cancer or healthy based on the expression profile of six identified proteins. In the clustering, INR correspond the ratio of protein levels in healthy and cancer sera. Green color indicates INR lower than 0 designating green area as cancer cluster. As shown in Fig. 5, sera from LC patients were successfully distinguished from the healthy sera. Only two LC sera were misjudged as healthy, and the remaining 15 LC sera were classified correctly, resulting in a sensitivity of 88%. Of the 15 healthy sera, 12 were designated as healthy, which corresponded to a specificity of 80%. When considering all tested sera, accuracy was estimated to be 84%. This result indicates that the six identified proteins could effectively be employed as potential biomarkers for prognosis of LC, although further studies according to histological subtypes and stages of LC are needed. When only any single protein was used for classification, median accuracy was about 69%. It is obvious that the use of multiple biomarker proteins would give rise to much higher accuracy for prognosis than that possible with a single biomarker.

4 Concluding remarks

We have demonstrated that an antibody microarray can be used for profiling of proteins in sera between LC patients and healthy persons. Based on the profiling result, six proteins were found to be differentially expressed in the sera of LC patients. Six identified proteins were further confirmed by using Wilcoxon test, Western blotting, and FDR. Blind testing of new 32 sera using the antibody microarray revealed a sensitivity and specificity of 88 and 80%, respectively, in classifying LC. Further studies will explore these promising results with the aim of exploiting their biomarker capabilities in LC.

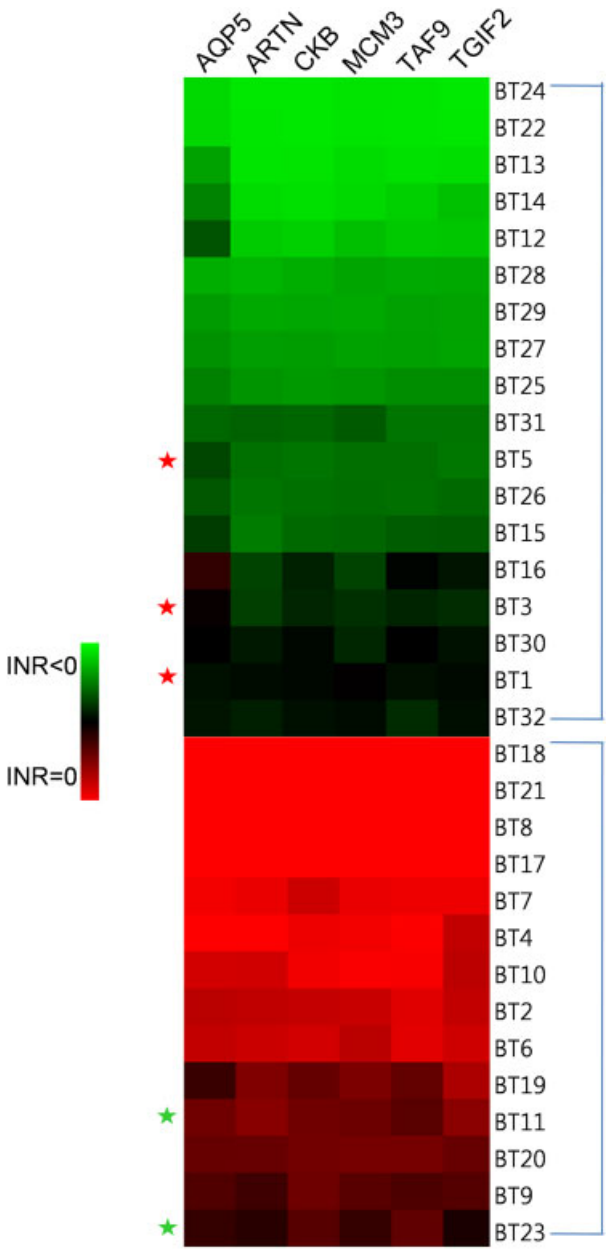

Figure 5. K-mean and hierarchical clustering of blind-tested sera. Clustering was performed according to the expression pattern of six identified proteins. Names of proteins are indicated on the top of the sure. Red color represents the INR value of zero (same expression level with healthy), and green indicates INR lower than zero (down expression level compared with healthy). Serum number was given randomly, and an asterisk indicates a false result (*i.e.* cancer serum (green asterisk) is judged as healthy or healthy serum (red asterisk) is regarded as cancerous).

This research was supported by the Korea Health 21C R&D Project (0405-MN01-0604-0007) of MIHWAF and the Nano Science & Technology Program (M10503000868-08M0300-86810), the Nano/Bio Science & Technology Program

(20090065621), Pioneer Research Program for Converging Technology (M10711300001-08M1130-00110), the Brain Korea 21 Program and WCU program (R31-2008-000-10071-0) of MEST. The authors thank Yong-Taek Kim for help regarding statistical analysis.

The authors have declared no conflict of interest.

5 References

[1] Page, G. P., Green, J. L., Lackland, D., Epidemiology of lung cancer with special reference to genetics, bioassays, women, and developing countries. *Semin. Respir. Crit. Care Med.* 2000, *21*, 365–373.

[2] Mountain, C. F., Revisions in the International System for Staging Lung Cancer. *Chest* 1997, *111*, 1710–1717.

[3] Beer, D. G., Kardia, S. L., Huang, C. C., Giordano, T. J. *et al.*, Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat. Med.* 2002, *8*, 816–824.

[4] Khan, N., Cromer, C. J., Campa, M., Patz, E. F., Jr., Clinical utility of serum amyloid A and macrophage migration inhibitory factor as serum biomarkers for the detection of nonsmall cell lung carcinoma. *Cancer* 2004, *101*, 379–384.

[5] Zhong, L., Hidalgo, G. E., Stromberg, A. J., Khattar, N. H. *et al.*, Using protein microarray as a diagnostic assay for non-small cell lung cancer. *Am. J. Respir. Crit. Care Med.* 2005, *172*, 1308–1314.

[6] Gao, W. M., Kuick, R., Orchekowski, R. P., Misek, D. E. *et al.*, Distinctive serum protein profiles involving abundant proteins in lung cancer patients based upon antibody microarray analysis. *BMC Cancer* 2005, *5*, 110.

[7] Lesko, L. J., Atkinson, A. J., Jr., Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. *Ann. Rev. Pharmacol. Toxicol.* 2001, *41*, 347–366.

[8] Orchekowski, R., Hamelinck, D., Li, L., Gliwa, E. *et al.*, Antibody microarray profiling reveals individual and combined serum proteins associated with pancreatic cancer. *Cancer Res.* 2005, *65*, 11193–11202.

[9] de Noo, M. E., Tollenaar, R. A., Ozalp, A., Kuppen, P. J. *et al.*, Reliability of human serum protein profiles generated with C8 magnetic beads assisted MALDI-TOF mass spectrometry. *Anal. Chem.* 2005, *77*, 7232–7241.

[10] Hong, J., Microfluidic systems for high throughput screening. *BioChip J.* 2008, *2*, 12–26.

[11] Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E. *et al.*, MicroRNA expression profiles classify human cancers. *Nature* 2005, *435*, 834–838.

[12] Oh, Y.-H., Kim, Y.-P., Kim, H. -S., SUMO(Small Ubiquitin-like Modifier) Chip for analysis of SUMO-conjugation to a target protein. *BioChip J.* 2007, *1*, 28–34.

[13] Zhou, H., Bouwman, K., Schotanus, M., Verweij, C. *et al.*, Two-color, rolling-circle amplification on antibody microarrays for sensitive, multiplexed serum-protein measurements. *Genome Biol.* 2004, *5*, R28.

[14] Huang, R. P., Huang, R., Fan, Y., Lin, Y., Simultaneous detection of multiple cytokines from conditioned media and patient's sera by an antibody-based protein array system. *Anal. Biochem.* 2001, *294*, 55–62.

[15] Huang, R., Lin, Y., Shi, Q., Flowers, L. *et al.*, Enhanced protein profiling arrays with ELISA-based amplification for high-throughput molecular changes of tumor patients' plasma. *Clin. Cancer Res.* 2004, *10*, 598–609.

[16] Pitha-Rowe, I., Petty, W. J., Feng, Q., Koza-Taylor, P. H. *et al.*, Microarray analyses uncover UBE1L as a candidate target gene for lung cancer chemoprevention. *Cancer Res.* 2004, *64*, 8109–8115.

[17] Tannapfel, A., Anhalt, K., Hausermann, P., Sommerer, F. *et al.*, Identification of novel proteins associated with hepatocellular carcinomas using protein microarrays. *J. Pathol.* 2003, *201*, 238–249.

[18] Choi, N., Son, D. S., Lee, J., Song, I. S. *et al.*, The signature from messenger RNA expression profiling can predict lymph node metastasis with high accuracy for non-small cell lung cancer. *J. Thorac. Oncol.* 2006, *1*, 622–628.

[19] Leinonen, T., Pirinen, R., Bohm, J., Johansson, R. *et al.*, Expression of matrix metalloproteinases 7 and 9 in non-small cell lung cancer Relation to clinicopathological factors, beta-catenin and prognosis. *Lung cancer* 2006, *51*, 313–321.

[20] Tamura, M., Oda, M., Matsumoto, I., Tsunezuka, Y. *et al.*, The combination assay with circulating vascular endothelial growth factor (VEGF)-C, matrix metalloproteinase-9, and VEGF for diagnosing lymph node metastasis in patients with non-small cell lung cancer. *Ann. Surg. Oncol.* 2004, *11*, 928–933.

[21] Breitling, R., Biological microarray interpretation: the rules of engagement. *Biochim. Biophys. Acta* 2006, *1759*, 319–327.

[22] Han, M. K., Hong, M. Y., Lee, D., Lee, D. E. *et al.*, Expression profiling of proteins in L-threonine biosynthetic pathway of *Escherichia coli* by using antibody microarray. *Proteomics* 2006, *6*, 5929–5940.

[23] Andersson, O., Kozlowski, M., Garachtchenko, T., Nikoloff, C. *et al.*, Determination of relative protein abundance by internally normalized ratio algorithm with antibody arrays. *J. Proteome Res.* 2005, *4*, 758–767.

[24] Joseph, J., Cardesa, A., Carreras, J., Creatine kinase activity and isoenzymes in lung, colon and liver carcinomas. *Br. J. Cancer* 1997, *76*, 600–605.

[25] Patton, W. F., Schulenberg, B., Steinberg, T. H., Two-dimensional gel electrophoresis; better than a poke in the ICAT? *Curr. Opin. Biotechnol.* 2002, *13*, 321–328.

[26] Madoz-Gurpide, J., Wang, H., Misek, D. E., Brichory, F. *et al.*, Protein based microarrays: a tool for probing the proteome of cancer cells and tissues. *Proteomics* 2001, *1*, 1279–1287.

[27] Sukhanov, S., Delafontaine, P., Protein chip-based microarray profiling of oxidized low density lipoprotein-treated cells. *Proteomics* 2005, *5*, 1274–1280.

[28] Kusnezow, W., Jacob, A., Walijew, A., Diehl, F. *et al.*, Antibody microarrays: an evaluation of production parameters. *Proteomics* 2003, *3*, 254–264.

[29] Coolen, R. B., Pragay, D. A., Nosanchuk, J. S., Belding, R., Elevation of brain-type creatine kinase in serum from patients with carcinoma. *Cancer* 1979, *44*, 1414–1418.

[30] Niklinski, J., Furman, M., Laudanski, J., Palynyczko, Z. *et al.*, Evaluation of carcinoembryonic antigen (CEA) and brain-type creatine kinase (CK-BB) in serum from patients with carcinoma of the lung. *Neoplasma* 1991, *38*, 129–135.

[31] Brown, D. J., McMillan, D. C., Milroy, R., The correlation between fatigue, physical function, the systemic inflammatory response, and psychological distress in patients with advanced lung cancer. *Cancer* 2005, *103*, 377–382.

[32] Ha, S. A., Shin, S. M., Namkoong, H., Lee, H. *et al.*, Cancer-associated expression of minichromosome maintenance 3 gene in several human cancers and its involvement in tumorigenesis. *Clin. Cancer Res.* 2004, *10*, 8386–8395.

[33] Wada, C., Kasai, K., Kameya, T., Ohtani, H., A general transcription initiation factor, human transcription factor IID, overexpressed in human lung and breast carcinoma and rapidly induced with serum stimulation. *Cancer Res.* 1992, *52*, 307–313.

