

# Proteomic screening of completely resected tumors in relation to survival in patients with stage I non-small cell lung cancer

FUMIHIRO OSHITA<sup>1\*</sup>, ATSUSHI MORITA<sup>5\*</sup>, HIROYUKI ITO<sup>1</sup>, YOICHI KAMEDA<sup>2</sup>,  
EIJU TSUCHIYA<sup>4</sup>, SHIGERU ASAI<sup>5</sup> and YOHEI MIYAGI<sup>3,4</sup>

Departments of <sup>1</sup>Thoracic Oncology, <sup>2</sup>Pathology and <sup>3</sup>Laboratory for Molecular Diagnostics, Kanagawa Cancer Center, Nakao 1-1-2, Asahi-ku, Yokohama 241-0815; <sup>4</sup>Division of Molecular Pathology and Genetics, Kanagawa Cancer Center Research Institute, Nakao 1-1-2, Asahi-ku, Yokohama 241-0815; <sup>5</sup>Discovery Research Laboratories, Shionogi & Co., Ltd., 12-4 Sagisu 5-chome, Fukushima-ku, Osaka 553-0002, Japan

Received March 3, 2010; Accepted May 3, 2010

DOI: 10.3892/or\_00000902

**Abstract.** To investigate the early changes in protein function that induce micro-metastasis in early-stage non-small cell lung cancer, we conducted proteomic analysis of tissue that had been completely resected. We selected sixteen patients whose tumors were pathological stage I adenocarcinoma with hobnail cell morphology. We compared their proteomic profiles between patients whose cancers recurred and did not recur after 5 years of follow-up. Proteins extracted from frozen tumor tissue were saturated by CyDye labeling and subjected to two-dimensional difference gel electrophoresis (2D-DIGE). We found approximately 2500 protein spots by DeCyder-DIA software analysis. A random forest classifier to the spot volume data sets revealed 30 extracted spots that were potentially effective marker proteins for distinguishing the recurrence and non-recurrence groups. Among them, Mann-Whitney U-test analysis showed that 15 spots were candidate marker proteins. Finally, 11 unique proteins corresponding to the 15 candidate spots were found using an improved in-gel digestion method and MALDI-TOF MS and nanoLC-ESI MS analysis. Among them, aldehyde dehydrogenase and tropomyosin were considered to have undergone protein modifications such as phosphorylation, acetylation, and glycosylation. We have thus identified some biomarkers that can predict tumor recurrence in patients with early-stage non-small cell lung cancer including hobnail-type adenocarcinoma. A further study is required to confirm the utility of these biomarkers.

## Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer death in Japan. To improve the prognosis of patients with NSCLC, attempts have been made to develop tests that will facilitate early diagnosis and treatment, and thereby decrease patient mortality. Although many chest roentgenogram-negative lung cancers can be detected by chest computed tomography, a significant number of patients with early-stage disease show aggressive tumors. Although locoregional control of NSCLC can be achieved by surgery, more than 70% of relapses in patients with stage I disease occur at distant sites (1). Thus, most patients with NSCLC have systemic disease, even at the earliest stage. Recent efforts at improving the management and outcome of patients with this disease have been directed at induction and adjuvant chemotherapy to reduce the high systemic relapse rate.

Cancer cells with genetic alterations at the early stage are considered to subsequently acquire other gene alterations, resulting in progression to locally advanced or metastatic tumors. Many genetic alterations related to cell proliferation, apoptosis, vascularization and tumor invasion have been reported to be prognostic factors in resected NSCLC. A study we conducted using immunohistochemical staining to compare the survival of 72 patients with small adenocarcinomas of the lung according to the expression of individual genes found that the overall survival of patients showing positive expression of survivin, cyclin D1 and integrin  $\beta$  was significantly worse than that of patients whose tumors did not express these genes (2). It was concluded that multiple, but not single, oncogene expression in tumor cells is an indicator of poor prognosis in patients with this type of cancer. Another study we conducted using DNA microarray also demonstrated that the expression levels of cyclin-dependent kinase 8, phosphoinositide-3-kinase, interferon regulatory factor 3 and tubulin were significantly higher in tumors of surviving patients with stage I NSCLC (3). To identify novel factors associated with prognosis in NSCLC patients, various biomarkers, including serum and tissue parameters and genetic and epigenetic factors, have been investigated using microarray, but none has been adopted for use in general clinical

---

*Correspondence to:* Dr Fumihito Oshita, Department of Thoracic Oncology, Kanagawa Cancer Center, Nakao 1-1-2, Asahi-ku, Yokohama 241-0815, Japan  
E-mail: foshita@kcch.jp

\*Contributed equally

*Key words:* lung, cancer, two-dimensional difference gel electrophoresis, lung adenocarcinoma, marker

practice (4). DNA microarray technology has had limited success in identifying gene expression profiles and biomarkers that are associated with survival in NSCLC patients (5-8) because mRNA expression is often not tightly associated with the level of protein expression and cannot identify post-translational modification of proteins. Thus, there is a need to develop new technologies for protein profiling that are applicable to clinical practice, and have high accuracy, sensitivity and throughput.

Proteomics may be defined as the large-scale characterization of proteins expressed by the genome (9). Unlike the study of a single protein or pathway, proteomic methods facilitate a systematic overview of expressed protein profiles, which, in the case of neoplasia, could ultimately improve the diagnosis, prognosis and management of patients by revealing the protein affecting overall tumor progression (10). It is becoming apparent that extensive post-translational modifications such as phosphorylation, glycosylation and proteolytic processing are common events, which present a challenge for protein analysis. These post-translational modifications can significantly alter protein function and thus the characteristics of the cell or tissue in which they occur. Thus, in the post-genome era, one of the challenges of proteomics is to understand the characteristics of tissue through knowledge of effector proteins and to apply this to clinical use. Individual cancer profiling could identify pathways that have been activated and that are suitable for chemotherapeutic strategies tailored to each individual neoplasm.

Contrary to promising expectations, the proteomic strategy has not yet yielded any biomarkers of clinical importance. Handling targets of huge diversity due to post-translational modifications of each protein, the technique seems ironically to have trouble seeking out the 'genuine article'. To overcome this situation, it is critical to adjust the clinicopathological characteristics of samples except for the specific point of interest. In the present study, therefore, we focused specifically on recurrence of stage I adenocarcinomas that had been completely resected, and followed the outcome of the patients over a 5-year period. The histology of the adenocarcinoma was also further restricted to the well differentiated type with a hobnail cell morphology, which is the most common cell type of this tumor (11). It was expected that the results would help to clarify the changes in protein function that induce micro-metastasis in early-stage NSCLC, and indicate possible new avenues for the development of drugs to suppress tumor recurrence.

## Materials and methods

This study was approved by the ethics committees of the Kanagawa Cancer Center and Shionogi Discovery Research Laboratories and performed in accordance with the guidelines of both committees.

**Materials.** CyDye DIGE Fluors (Cy3 and Cy5 for saturation-labeling), IPG buffer (pH 3-11 NL), Immobiline DryStrip (24 cm, pH 3-11 NL), acrylamide, *N,N*-methylenebisacrylamide, sodium dodecyl sulfate (SDS), *N,N,N',N'*-tetramethylethylenediamine (TEMED) and bind-silane, were obtained from GE Healthcare (Uppsala, Sweden). Urea was from ICN Biomedicals (Aurora, OH, USA). Tris-(2-carboxyethyl)

phosphine hydrochloride (TCEP) was purchased from Pierce (Rockford, IL, USA) and *N,N*-dimethylformamide (DMF) was from Sigma-Aldrich (St. Louis, MO, USA). Other chemicals were purchased from Wako Pure Chemical Industries (Osaka, Japan). Protease inhibitor mixture cocktail (Complete) and Pefabloc SC protector solution were from Roche Diagnostics GmbH (Mannheim, Germany). Lysyl endopeptidase from *Achromobacter lyticus* M497-1 (Lys-C) was purchased from Wako Pure Chemical Industries.

**Patients.** Patients with histological proven lung cancer who had undergone curative surgical resection of their lung tumors were entered into the present study. The histology was restricted to well differentiated adenocarcinoma with a hobnail cell morphology. None of the patients had received prior chemotherapy for the primary lesion. Patient without tumor recurrence for up to 5 years after tumor resection were defined as non-recurrence patients, and patients who died within 5 years after resection due to tumor recurrence and progression were defined as recurrence patients. Written informed consent for proteomic analysis of their tumor tissue was obtained from all the patients or their families.

**Tumor samples.** Specimens of tumor tissues were taken from freshly resected materials in the operating room, rinsed thoroughly with saline and snap-frozen in liquid nitrogen. The frozen tissues were stored at -80°C until use.

**Sample preparation and fluorescence labeling.** Frozen lung adenocarcinoma tissue (20 mg) was thawed and washed well with the tissue washing buffer. The tissue pellet was then homogenized in lysis buffer and subjected to 2D-DIGE. Total protein concentrations were determined with a Protein Assay kit purchased from Bio-Rad (Hemel Hempstead, UK) using bovine  $\gamma$  globulin as a standard. The fluorescence labeling for 2D-DIGE with Cy3 or Cy5 was performed as described previously (12,13). The labeled samples were stored at -80°C until use.

**2D-DIGE analysis.** Two-dimensional difference gel electrophoresis for 2D-DIGE analysis was carried out as follows. We loaded 5  $\mu$ g of the Cy5-labeled protein of analyzing sample on each gel and also loaded 5  $\mu$ g of Cy3-labeled protein which was made of mixture of all analyzed samples on another gel. The first-dimension separation was performed with Immobiline DryStrips (24 cm, pH 3-11 NL) after rehydration of the IPG strips with the labeled samples in 450  $\mu$ l of rehydration buffer, and then electric focusing was carried out. The IPG strips were then loaded and run on 12.5% polyacrylamide Laemmli gels (14) at 1 W per gel constant power at 20°C. Three replicas from each sample were made for differential analysis in order to reduce gel-to-gel variations.

Gel images were collected using a Typhoon 9400 (GE Healthcare) fluorescence gel scanner and the Cy-dye images were imported into DeCyder software version 6.5 (GE Healthcare). These images were normalized using Cy3 images as an internal control, and then we calculated average abundance changes and analyzed the differential expression of the proteins statistically across all gels. Additionally, all numerical data as normalized spot volumes were exported from the DeCyder software and utilized for further statistical analysis

 SPANDIDOS PUBLICATIONS random forests and Mann-Whitney U-test, in order to identify the protein spots that were differentially expressed between the tissue groups examined.

**In-gel digestion.** Protein spots of interest were excised by an automated spot collector (Ettan SpotPicker, GE Healthcare) from a preparative Cy3 2D-DIGE gel. The gel pieces were subjected to the modified in-gel digestion method to prepare peptide samples for identification using mass spectrometry as follows. The gel pieces were washed vigorously 5 times with 100  $\mu$ l of washing buffer (60% acetonitrile containing 100 mM ammonium bicarbonate) for 10 min, then washed additionally with 100  $\mu$ l of acetonitrile for 5 min and dried completely for 15 min using a SpeedVac Concentrator (SPD2010, Thermo Electron Corp., Waltham, MA, USA). The dried gel pieces were swollen with 20  $\mu$ l of 0.1% RapiGest SF (Waters, Milford, MA, USA) in 50 mM ammonium bicarbonate and incubated at 37°C for 10 min. After the excess solutions had been removed, the gel pieces were dried in the SpeedVac Concentrator and re-swollen in 20  $\mu$ l of 1 mM Tris-HCl (pH 8.5) containing 200 ng of Lys-C, and incubated on ice for 45 min. After the Lys-C solutions had been removed and 20  $\mu$ l of 1 mM Tris-HCl (pH 8.5) added, the enzymatic digestions were performed overnight at 37°C. Each reaction solution was transferred to a new tube, and the generated peptide fragments in each piece of gel were extracted by vortexing for 15 min with 50  $\mu$ l of 60% acetonitrile. Each extracted solution was combined with the reaction solution, followed by evaporation to complete dryness using the SpeedVac Concentrator. In order to enhance the signal intensity derived from the ionized peptides on mass spectrometric analysis, the digested peptide fragments of C-terminal lysine residues were guanidinated using a ProteomMass Guanidination kit (Sigma-Aldrich) in accordance with the manufacturer's protocol. The prepared in-gel digestion samples were stored at -30°C until analysis.

**Mass spectrometric protein identification.** The proteins corresponding to the spots of interest were identified using two mass spectrometry systems, MALDI-TOF MS and nanoLC-ESI Q-ToF MS. For MALDI-TOF MS analysis, we applied the AnchorChip technology (Bruker Daltonics, Bremen, Germany) using  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA, Bruker Daltonics) as the MALDI matrix. Preparation of the digested sample on the AnchorChip was performed by the thin-layer affinity CHCA method (15). An in-gel digestion sample was dissolved in 4  $\mu$ l of 0.2% trifluoroacetic acid (TFA) then deposited onto the CHCA matrix layer on an AnchorChip MALDI plate (600  $\mu$ m diameter). After 3 min, 4  $\mu$ l of washing buffer (10 mM ammonium phosphate, monobasic in 0.1% TFA) was added to each spot and the whole droplet was removed at once. A 1- $\mu$ l aliquot of 0.1 mg/ml CHCA in ethanol/acetone/0.1% TFA solution (6:3:1 mixture) was deposited onto the spot for recrystallization of the matrix. MALDI-TOF MS analysis of the peptide samples was performed using the UltraFlex II instrument (Bruker Daltonik) in accordance with the manufacturer's instructions. The MS and MS/MS spectra were acquired in the positive mode. The spectral data were processed and analyzed using the FlexAnalysis software package (Bruker Daltonics).

For nanoLC-ESI Q-ToF MS analysis, we used a DiNa nano-LC instrument (KYA Technologies, Tokyo, Japan) coupled with a QqTOF mass spectrometer (Q-ToF Ultima API, Micromass, Manchester, UK). In-gel digestion samples were desalted and separated using the DiNa nano-LC equipped with monolithic ODS columns (100  $\mu$ m id x 130 mm for precolumn and 50  $\mu$ m x 700 mm for analytical column, Kyoto Monotech, Kyoto, Japan) (16). The peptide fragments were eluted by a gradient flow of acetonitrile and distilled water containing 0.2% formic acid at a flow rate of 200 nl/min, then injected directly into the Q-ToF Ultima API through an online ESI nanospray emitter (PicoChip, New Objective, MA, USA). The analytical conditions of MS were: cycle time 1.10 sec; scan duration 1.00 sec; ionization mode ES<sup>+</sup>; *m/z* range 350-1300. Three precursor masses were selected for MS/MS analysis with automatically determined collision energy for the charge states and *m/z* values of the precursor ions. Data acquisition and processing were carried out using the MassLynx software package (Micromass).

MASCOT software (Matrix Science, London, UK) was used to search the mass spectral data obtained from MALDI-TOF MS and nanoLC-ESI Q-ToF MS against the NCBI database (human) for protein identification. Up to one missed cleavage with Lys-C protease was allowed and the data were searched using guanidination of lysine as a fixed modification, and oxidation of methionine, Cy3- and hydrolyzed Cy3-labeling of cysteine as variable modifications. The reliabilities of the resulting protein identification were evaluated using the following criteria: i) The same protein was assigned a top score by a Mascot search for both MALDI and ESI. ii) The resulting Mascot score was >40 and/or statistical reliability was 95%. iii) No other likely candidates for consideration were found. iv) All MS and MS/MS spectra of identified proteins were checked manually.

## Results

Among the 16 patients studied, 12 were alive without tumor recurrence and 4 had died due to lung cancer recurrence by 5 years after surgery. Their characteristics are summarized in Table I. Protein extracts were obtained from the frozen tissues. As the volume of extracted protein was limited, 2D-DIGE analysis using the saturation CyDye labeling method was used for proteomics analysis. For this, the protein extract was saturated with CyDye (Cy3 and Cy5) and then subjected to 2-dimensional electrophoresis. Protein spot images were obtained using a fluorescence scanner.

**2D-DIGE analysis of proteins from surgical samples.** The proteomic profiles of the 16 samples of resected human lung adenocarcinoma tissue were analyzed by 2D-DIGE. Prior to the differential expression analysis, the reproducibility of the present method was verified by running each sample in duplicate. The results showed that the intensity value for 99.5% of the protein spots was within a 2-fold difference, with a correlation coefficient of 0.973, thus demonstrating that the present profiling method had high reproducibility.

A representative 2D-DIGE gel image obtained from a Cy3-labeled mixture of all samples as an internal control is shown in Fig. 1. We found ~2500 protein spots on this image

Table I. Patient characteristics.

No.	Gender	Age	Smoking	Operation	Pathologic diagnosis		Recurrence	Metastatic site	Outcome
					Stage	Pathology			
1	M	60	-	2000 Dec	IB	W/D Ad	-	-	Alive
2	M	61	+	2000 Aug	IB	W/D Ad	-	-	Alive
3	M	80	+	2000 Mar	IB	W/D Ad	-	-	Alive
4	F	83	-	2001 Jul	IB	W/D Ad	-	-	Alive
5	M	64	-	2001 Jul	IB	W/D Ad	-	-	Alive
6	F	75	-	2000 Nov	IB	W/D Ad	-	-	Alive
7	F	57	-	2000 Oct	IB	W/D Ad	-	-	Alive
8	F	61	-	2001 Nov	IB	W/D Ad	-	-	Alive
9	M	70	+	2001 Mar	IB	W/D Ad	-	-	Alive
10	F	66	-	2001 Jul	IB	W/D Ad	-	-	Alive
11	F	68	-	2000 Nov	IB	W/D Ad	-	-	Alive
12	M	49	+	2000 Oct	IB	W/D Ad	-	-	Alive
13	F	46	+	2000 Sep	IA	W/D Ad	+	Brain	Dead (2005 Jun)
14	M	68	+	2001 Apr	IA	W/D Ad	+	Brain	Dead (2004 Feb)
15	M	50	+	2001 Sep	IB	W/D Ad	+	Lung	Dead (2004 Jul)
16	M	77	+	2001 Nov	IB	W/D Ad	+	Lung	Dead (2005 Mar)

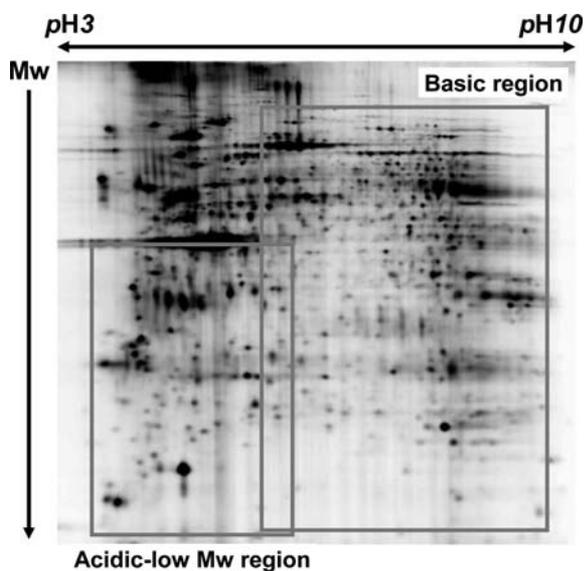


Figure 1. 2D-DIGE gel image of Cy3-labeled mixed sample used as an internal control (master image for the present analysis). The differential analysis was carried out using two cropped images, covering the basic and acidic-low Mw regions, which are indicated by squares.

by DeCyder-DIA software analysis. Among these spots, about 500 were detected in the acidic ( $pI < 6$  approximately) and high-molecular-weight (Mw) region ( $> 40$  kDa approximately) with high fluorescence intensities. Some abundant cytoskeletal proteins, such as molecules of the actin and cytochrome family, were mobilized electrophoretically in this region. In order to focus on proteins with intermediate and/or low expression in the present analysis, we cropped two images, an acidic-low Mw region and a basic region. Exclusion of the acidic-high Mw region made it possible to obtain fluorescence images with higher sensitivity because the fluores-

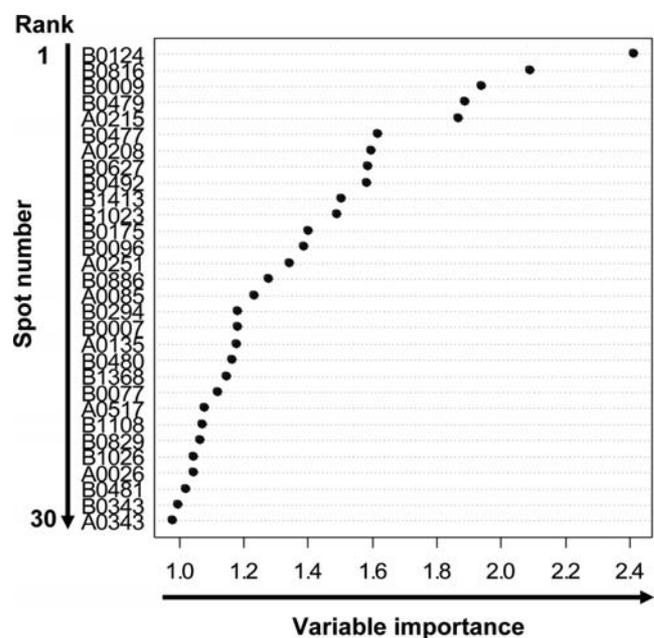


Figure 2. Candidate spots extracted using random forests (RF) classifier analysis. By this analysis, the top 30 characteristically expressed spots were extracted. Continuously performed statistical analysis using Mann-Whitney U-test demonstrated that most of these 30 spots had significant identification power ( $P < 0.05$ ).

cence range on image scanning (practically, about three to four orders of magnitude) was fixed, without inclusion of excessively abundant proteins. 2D-DIGE quantification analysis was performed by focusing on these acidic-low Mw and basic regions, for which the numbers of spots detected in the master gel were 536 and 1518, respectively (Fig. 1).

After all the images obtained from triplicate lung adenocarcinoma tissue samples had been analyzed using the



Spot no.	RF rank	Expression change (recurrence/non-recurrence)	Average ratio	P-value <sup>a</sup>
B0816	2	Up	2.4	0.006
B0479	4	Down	2.4	0.018
A0215	5	Up	1.4	0.009
B0477	6	Down	2.1	0.018
B0492	9	Down	3.7	0.034
B1413	10	Up	1.2	0.046
B1023	11	Down	1.8	0.009
B0175	12	Up	1.7	0.025
A0251	14	Up	1.5	0.009
B0886	15	Down	1.4	0.013
A0085	16	Up	1.5	0.046
B0480	20	Down	2.3	0.009
B1368	21	Up	1.3	0.025
B0077	22	Up	1.2	0.130
B0481	28	Down	2.3	0.009
B0343	29	Up	1.5	0.046

<sup>a</sup>P-values calculated by Mann-Whitney's U-test. The variables of P<0.05 considered to be significant between recurrence and non-recurrence groups of lung adenocarcinoma patients.

DeCyder-BVA module, the normalized spot volume data were exported for further statistical analysis. In order to find effective marker proteins that could correctly distinguish the recurrence and non-recurrence groups, we first fitted a random forest (RF) classifier to the spot volume data sets. The summarized result of RF statistical analysis depicted in Fig. 2 shows the top 30 extracted candidate spots with variable importance. Among the 30 spots, 8 (27%) appeared in the acidic-low Mw region and the other 22 (73%) appeared in the basic region. For these candidates, we checked their states of independence from other spots by evaluating the 2D-DIGE images with DeCyder-BVA software, and at least 16 spots were confirmed to be isolated as a single fluorescent peak. Next, in order to assess the potencies of the 16 spots selected as marker proteins, Mann-Whitney U-test analysis was performed on the spot volume of each candidate. As a result of the additional statistical analysis, 15 candidates, except for spot B0077, emerged at a P-value of <0.05 by U-test (Table II). The distribution patterns of the spot volumes for four representative candidates are shown in Fig. 3.

*Identification of differentially expressed proteins in samples from patients with recurrence.* For differential expression analysis, we used saturation fluorescence labeling for 2D-DIGE. This labeling method introduces fluorescent CyDye molecules to all cysteine residues of proteins, which leads to under-representation of the ion signals on mass spectra. In an initial study, we had clarified that identification of saturation CyDye-labeled proteins was dependent on not only ionization of the labeled peptides at the mass spectrometric analysis step, but also the efficiency of protein labeling at in-gel digestion step. In order to identify all candidates, we performed an

improved in-gel digestion method for the preparation of peptide samples, and used two MS systems for analysis, MALDI-TOF MS and nanoLC-ESI MS. The 16 statistically selected candidate spots were excised from the preparative gels using an Ettan SpotPicker. After in-gel digestion with Lys-C protease, all the prepared samples obtained from the candidate spots were analyzed using the above MS systems. The reliability of protein identification was ensured by duplicate analysis and the criteria used for the MS search. Thus our analysis resulted in the identification of 11 unique proteins corresponding to the 15 candidate spots (Table III). Five spots (B0479, B0477, B0492, B0480, B0481) aligned horizontally in the 2D-DIGE gel were identified as the same protein, aldehyde dehydrogenase, and three tropomyosin subunit components (A0085, A0215, A0251) were found in the acidic-low Mw region. Continuous depiction of these identified candidates on the master gel image revealed differences in Mw and/or pI for 11 of the 15 proteins between the values predicted on the UniProt protein database (<http://www.uniprot.org/>) and the values observed upon 2D-DIGE gel separation (Fig. 4). It was thought that these shifts were attributable to protein modifications such as phosphorylation, acetylation and glycosylation. The shifts of the Mw and pI values, and the isoforms identified from analysis of the MS/MS spectrum, are summarized additionally in Table III.

## Discussion

We have previously reported some survival-related genes in resected tumors, such as cyclin-related genes and protein kinase, characterized using DNA microarray. Unfortunately, the DNA microarray technique has limited applicability for

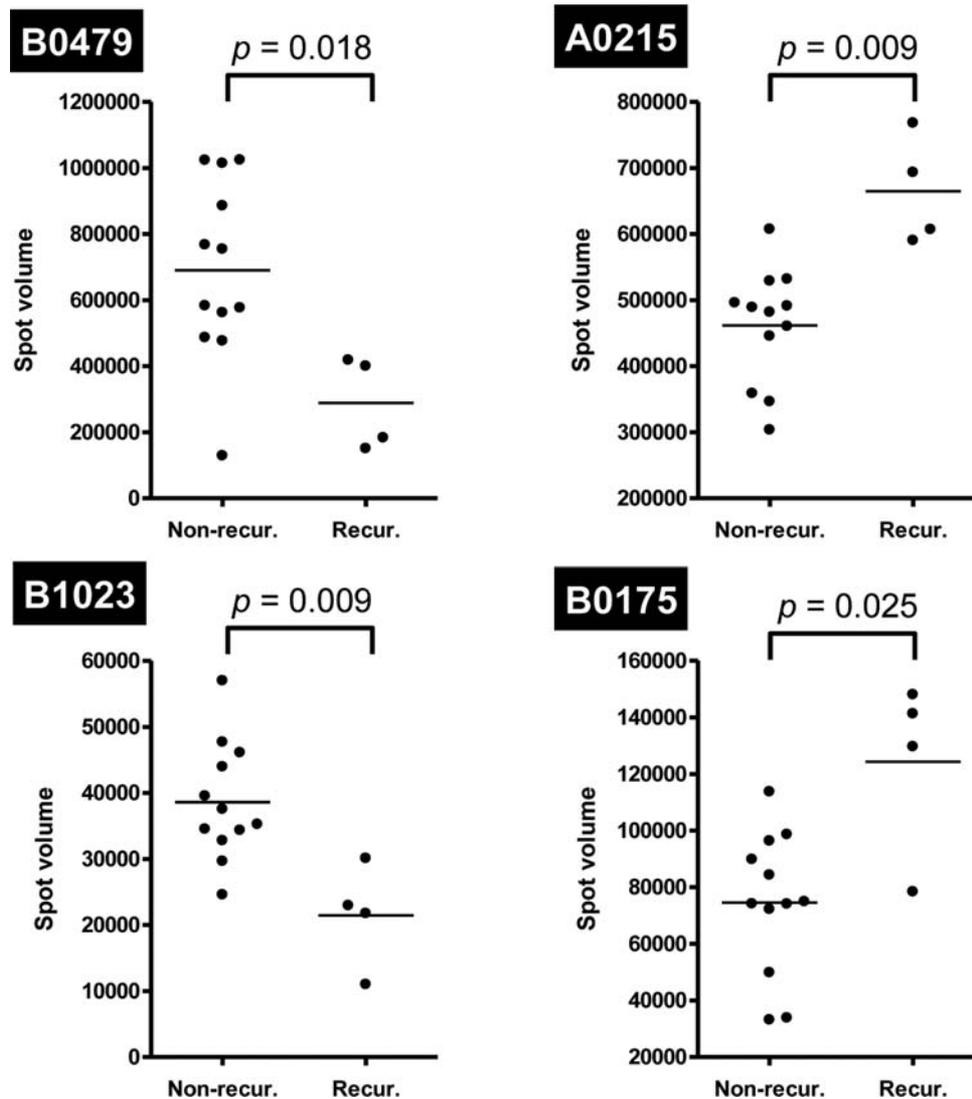


Figure 3. Normalized spot volumes of selected candidates in each individual tissue sample. Four representative spots, B0479, A0215, B1023, B0175 are shown. Each dot derived from one tissue sample was the mean of the normalized spot volume for triplicate 2D-DIGE analyses, and the horizontal bar was the mean for each respective group. Significance was evaluated using Mann-Whitney U-test.

identifying therapeutic target profiles in NSCLC. This may be because alternative splicing, post-translational modifications, compartmentalization and translocation allow many protein isoforms to be encoded by a single gene allele. Integration of both genomics and proteomics will no doubt facilitate the discovery of novel drug-target proteins and disease biomarkers. A proteomic study of 174 specimens of NSCLC using matrix-associated laser desorption-ionization mass spectrometry demonstrated a signature of 25 signals that was associated with patient survival (4). In that study, proteins corresponding to signals in the signature were identified and shown to have various cellular functions, including ribosomal protein L26-like 1, acylphosphatase and phosphoprotein enriched in astrocytes 15. The tumor samples used in that study were obtained from patients with early stage I to advanced stage III disease, and moreover the study did not take into consideration various pathological types. Our present proteomic analysis was performed on tumors of uniform pathology such as hobnail-type adenocarcinoma, and only early-stage tumors. A further advantage of our analysis was

that it compared non-recurrent and recurrent tumors with metastatic lesions at stage I for which the outcome was followed until 5 years after resection. Thus our analysis was able to clarify simply differences in the metastatic component of primary lung tumors at an early stage. Most previous studies have compared the proteomics of tumor samples and normal lung tissues, and therefore might have detected differences in not only metastatic properties but also carcinogenesis. In order to treat patients with early-stage NSCLC, the therapeutic target should be the characteristics of metastasis and not tumor progression.

Our proteomic study showed a good sensitivity of over 95% for distinguishing non-recurrent from recurrent groups. The target proteins analyzed in this study comprised about 2500 spots, and therefore genuinely prognostic proteins might have been included. Prognostic proteins detected in this analysis using samples of early-stage tumors were closely related to initiation of the invasion or adhesion stage of the metastatic process. In fact, these prognostic proteins had already been reportedly related to cancerous lesions, and to



Identified proteins that were differentially expressed between the recurrence and non-recurrence groups.

Spot no.	Protein name	Accession no. <sup>a</sup>	Gene name <sup>b</sup>	Theoretical		Cys. no.	Observed spot on 2D-DIGE gel <sup>d</sup>
				Mw <sup>c</sup> (kDa)	pI		
B0816	Annexin A2	P07355	ANXA2	42.5	7.6	4	Acidic
B0479	Aldehyde dehydrogenase	P05091	ALDH2	63.6	6.6	9	Acidic
A0215	Tropomyosin $\alpha$ -3 chain	P06753-2	TPM3	32.8	4.8	4	-
B0477	Aldehyde dehydrogenase	P05091	ALDH2	63.6	6.6	9	Acidic
B0492	Aldehyde dehydrogenase	P05091	ALDH2	63.6	6.6	9	Acidic
B1413	Peptidyl-prolyl cis-trans isomerase A	P62937	PPIA	21.1	7.8	4	Acidic
B1023	$\delta$ 3,5- $\delta$ 2,4-dienoyl-CoA isomerase	Q13011	ECH1	40.5	8.2	6	Acidic
B0175	Mitochondrial inner membrane protein	Q16891-2	IMMT	89.7	6.2	7	-
A0251	Tropomyosin $\alpha$ -4 chain	P67936	TPM4	30.8	4.7	2	-
B0886	Inorganic pyrophosphatase 2	Q9H2U2	IPYR2	45.2	7.1	9	Low Mw/Acidic
A0085	Tropomyosin $\beta$ chain	P07951-2	TPM2	35.1	4.6	1	High Mw
B0480	Aldehyde dehydrogenase	P05091	ALDH2	63.6	6.6	9	Acidic
B1368	Proteasome subunit $\beta$ type 2	P49721	PSB2	25.3	6.5	3	Basic
B0481	Aldehyde dehydrogenase	P05091	ALDH2	63.6	6.6	9	Acidic
B0343	Lamin-A/C	P02545	LMNA	79.1	6.6	5	-

<sup>a</sup>UniProt database accession number; <sup>b</sup>Gene name on NCBI database; <sup>c</sup>Calculated Mw of proteins in which all cysteine residues were labeled with Cy3; <sup>d</sup>2D-DIGE spot differences in Mw and pI compared with theoretical values.

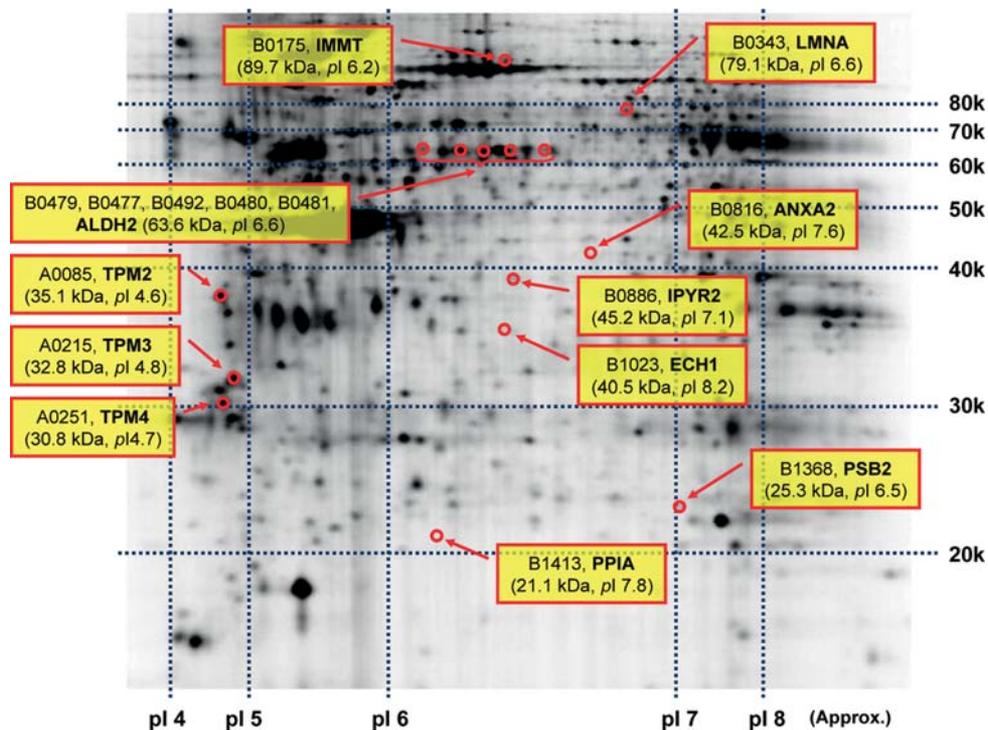


Figure 4. Identified proteins differentially expressed between the recurrence and non-recurrence groups of lung adenocarcinoma patients on the master image of the 2D-DIGE gel. The 15 spots identified in this study are circled, and each is assigned a spot number, a gene name (shown in Table III), the calculated Mw of the Cy3-labeled protein and the pI. The vertical and horizontal dotted lines show the pI and Mw values predicted from the present 2D-DIGE gel.

play a role in metastasis. Seven such prognostic proteins, annexin, tropomyosin, peptidyl-prolyl cis-trans isomerase A, mitochondrial inner membrane, proteasome, elongation and lamin, were up-regulated in patients with recurrence com-

pared to those who were cured. We also identified down-regulation of the mitochondrial enzymes aldehyde dehydrogenase, dienoyl-CoA isomerase and inorganic pyrophosphatase 2 in patients with recurrence. Annexins are calcium-

and phospholipid-binding proteins forming an evolutionarily conserved multigene family that is widely expressed in mammals (17). Dysregulation of annexins has been reported in numerous cancers, influencing patterns of cellular behavior such as proliferation, motility, invasiveness and signal pathways (18). Furthermore, numerous studies have shown that annexins are metastasis-associated prognostic factors in multiple malignancies (19-22). Tropomyosin is an elongation protein that binds to the actin subunit, and regulation of tropomyosin binding is an important step in muscle contraction. Tropomyosin is reportedly up-regulated in highly metastatic breast cancer (23) and increases the adhesion of cells to the matrix by enhancing the production of actin fibers and focal adhesions in breast cancer cells (24). Peptidyl-prolyl isomerase, a regulator of protein conformation, is necessary for the progression of human breast cancer, and the inhibitor cyclosporin A has been reported to inhibit cancer growth and invasion (25). A gene associated with high metastatic potential isolated from rat mammary adenocarcinoma has been shown to be a homologue of human elongation factor 1 subunit  $\alpha$  (26). EF-1  $\delta$  mRNA shows significantly higher expression in cancerous than in non-cancerous esophageal tissues, and its higher expression is correlated with lymph node metastasis in patients with esophageal carcinoma (27). Laminin is a major glycoprotein of the basement membrane. Both tumor cells and normal cells express a high-affinity receptor for laminin, although the expression is more pronounced in tumor cells. Metastatic invasion of the basement membrane by tumor cells involves their binding to laminin. The platelet isoform of 12-lipoxygenase is expressed in a variety of human tumors. This enzyme is able to bind to lamin A, and is involved in the cellular responses associated with tumor progression and metastasis (28). Dysregulation of the ubiquitin/proteasome system often accompanies tumorigenesis and progression. Comparative proteasome analysis of mouse lung adenocarcinoma cell lines has revealed that aberrant expression of proteasome proteins is involved in cancer development and metastasis (29).

Aldehyde dehydrogenase and inorganic pyrophosphatase were down-regulated in tumors from patients who developed recurrence. The expression of inorganic pyrophosphatase is reportedly decreased in neoplastic cells (30), similar to the findings we obtained in the present study. However, other studies have demonstrated that aldehyde dehydrogenase is highly expressed in tumors (31,32) and that stem cells in breast cancer cell lines showing positive expression of aldehyde dehydrogenase are responsible for mediating metastasis (33). In contrast, our study showed that aldehyde dehydrogenase was down-regulated in primary tumors, and was present as a transformed acidic form. Down-regulation of aldehyde dehydrogenase in mitochondria was found at 5 points to distinguish the non-recurrence and recurrence groups in our study, indicating that this minor form of the enzyme has an important role in metastasis of early-stage lung cancer.

Our prognostically relevant proteins represented forms that had been modified after translation and accounted for rather a minor proportion of each protein. For example, annexin A2 was shown to account for only 0.5% of total protein by Western blot analysis. Two other candidate prognostic proteins, aldehyde dehydrogenase and tropomyosin subunit components,

that had the potential to distinguish the non-recurrence and recurrence groups, were identified in the 2D-DIGE gel. It was thought that these shifts were attributable to protein modifications such as phosphorylation, acetylation and glycosylation. These data suggest that post-translational modification of proteins plays a very important role in the initiation of metastasis in early-stage lung adenocarcinoma. An investigation of protein glycosylation using multiplex proteomics technology has demonstrated that alteration of protein glycosylation would be of great significance for metastasis in human liver cell lines (34), and another proteomic study has also shown that post-translational modifications of high-mobility proteins are related to the metastatic potential of breast cancer cells, metastatic and non-metastatic cells showing a difference in amino acid residue acetylation and methylation (35). However, these previous studies employed cancer cell lines, and our present study is the first to have used tumor tissues from patients with lung cancer, demonstrating that post-translational modification of proteins is related to metastasis. It is unclear why no proteins related to translational modification were found, as revealed by differences in such proteins between non-recurrence and recurrence patients, although minor regulation of translationally modified prognostic proteins might have occurred. Such minor modifications of prognostic proteins may have explained the difficulty we experienced in raising antibodies against them. However, such minor modification was an early change characteristic of metastasis, and therefore might provide clues to devising new treatments for suppression of metastasis.

We conclude that up-regulation of five proteins, including annexin 2, and down-regulation of mitochondrial aldehyde dehydrogenase are indicators of poor prognosis for stage I adenocarcinoma with hobnail cell morphology. A confirmation study of these biomarkers of poor prognosis is needed with the long-term aim of allowing selection of patients with early-stage lung cancer who might benefit from chemotherapy. Further proteomic investigations are also required to examine any other biomarkers of poor prognosis in other pathological types of early-stage NSCLC.

### Acknowledgements

This study was supported in part by Kanagawa Cancer Research Fund, Kanagawa Prefectural Hospitals Cancer Research Fund and Kanagawa Health Foundation.

### References

1. Martini N and Ginsberg R: Treatment of stage I and II disease. In: Comprehensive Textbook of Thoracic Oncology. Aisner J, Arriagada R, Green M, *et al* (eds). Williams & Williams, Baltimore, MD, pp339-350, 1996.
2. Oshita F, Ito H, Ikehara M, *et al*: Prognostic impact of survivin, cyclin D1, integrin  $\beta$ 1 and VEGF in patients with small adenocarcinoma of stage I lung cancer. *Am J Clin Oncol* 27: 425-428, 2004.
3. Oshita F, Sekiyama A, Ito H, Kameda Y and Miyagi Y: Genome-wide cDNA microarray screening of genes related to survival in patients after curative resection of non-small cell lung cancer. *Oncol Rep* 16: 817-821, 2006.
4. Yanagisawa K, Tomida S, Shimada Y, Yatabe Y, Mitsudomi T and Takahashi T: A 25-signal proteomic signature and outcome for patients with resected non-small cell lung cancer. *J Natl Cancer Inst* 99: 858-867, 2007.

 SPANDIDOS JG, Kardia SL, Huang CC, *et al*: Gene-expression profiles of survival of patients with lung adenocarcinoma. *Nat Med* 8: 810-824, 2002.

6. Bhattacharjee A, Richards WG, Staunton J, *et al*: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci USA* 98: 13790-13795, 2001.
7. Garber ME, Troyanskaya OG, Schluens K, *et al*: Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci USA* 98: 13784-13789, 2001.
8. Tomida S, Koshikawa K, Yatabe Y, *et al*: Gene expression-based, individualized outcome prediction for surgically treated lung cancer patients. *Oncogene* 23: 5360-5370, 2004.
9. Hoehn GT and Suffedini AF: Proteomics. *Crit Care Med* 33: S444-S448, 2005.
10. Alessandro R, Fontana S, Kohn E and De Leo G: Proteomic strategies and their application in cancer research. *Tumori* 91: 447-455, 2005.
11. Hashimoto T, Tokuchi Y, Hayashi M, *et al*: Different subtypes of human lung adenocarcinoma caused by different etiological factors. Evidence from p53 mutational spectra. *Am J Pathol* 157: 2133-2141, 2000.
12. Shaw J, Rowlinson R, Nickson J, Stone T, Sweet A, Williams K and Tonge R: Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. *Proteomics* 3: 1181-1195, 2003.
13. Greengauz-Roberts O, Stöppler H, Nomura S, *et al*: Saturation labeling with cysteine-reactive cyanine fluorescent dyes provides increased sensitivity for protein expression profiling of laser-microdissected clinical specimens. *Proteomics* 5: 1746-1757, 2005.
14. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685, 1970.
15. Zhang X, Shi L, Shu S, *et al*: An improved method of sample preparation on AnchorChip targets for MALDI-MS and MS/MS and its application in the liver proteome project. *Proteomics* 7: 2340-2349, 2007.
16. Ishizuka N, Minakuchi H, Nakanishi K, Soga N, Nagayama H, Hosoya K and Tanaka N: Performance of a monolithic silica column in a capillary under pressure-driven and electrodriven conditions. *Anal Chem* 72: 1275-1280, 2000.
17. Gerke V and Moss SE: Annexins: from structure to function. *Physiol Rev* 82: 331-371, 2002.
18. Rescher U and Gerke V: Annexins: unique membrane binding proteins with diverse functions. *J Cell Sci* 117: 2631-2639, 2004.
19. Brichory FM, Misek DE, Yim AM, *et al*: An immune response manifested by the common occurrence of annexins I and II auto-antibodies and high circulating levels of IL-6 in lung cancer. *Proc Natl Acad Sci USA* 98: 9824-9829, 2001.
20. Wang KL, Wu TT, Resetskova E, *et al*: Expression of annexin A1 in esophageal and esophagogastric junction adenocarcinomas: association with poor outcome. *Clin Cancer Res* 12: 4598-4604, 2006.
21. Lim LH and Pervaiz S: Annexin I: the new face of an old molecule. *FASEB J* 21: 968-975, 2007.
22. Liu YF, Xiao ZQ, Li MX, *et al*: Quantitative proteome analysis reveals annexin A3 as a novel biomarker in lung adenocarcinoma. *J Pathol* 217: 54-64, 2009.
23. Li DQ, Wang L, Fei F, *et al*: Identification of breast cancer metastasis-associated proteins in an isogenic tumor metastasis model using two-dimensional gel electrophoresis and liquid chromatography-ion trap-mass spectrometry. *Proteomics* 6: 3352-3368, 2006.
24. Zheng Q, Safina A and Bakin AV: Role of high-molecular weight tropomyosins in TGF-beta-mediated control of cell motility. *Int J Cancer* 122: 78-90, 2008.
25. Zheng J, Koblinski JE, Dutson LV, Feeney YB and Clevenger CV: Prolyl isomerase cyclophilin A regulation of Janus-activated kinase 2 and the progression of human breast cancer. *Cancer Res* 68: 7769-7778, 2008.
26. Phillips SM, Bendall AJ and Ramshaw IA: Isolation of gene associated with high metastatic potential in rat mammary adenocarcinomas. *J Natl Cancer Inst* 82: 199-203, 1990.
27. Ogawa K, Utsunomiya T, Mimori K, *et al*: Clinical significance of elongation factor-1 delta mRNA expression in oesophageal carcinoma. *Br J Cancer* 91: 282-286, 2004.
28. Tang K, Finley RL Jr, Nie D and Honn KV: Identification of 12-lipoxygenase interaction with cellular proteins by yeast two-hybrid screening. *Biochemistry* 39: 3185-3191, 2000.
29. Zhang K, Wrzesinski K, Stephen JF, Larsen PM, Zhang X and Roepstorff P: Comparative proteome analysis of three mouse lung adenocarcinoma CMT cell lines with different metastatic potential by two-dimensional gel electrophoresis and mass spectrometry. *Proteomics* 8: 4932-4945, 2008.
30. Shatton JB, Williams A and Weinhouse S: Subcellular distribution of inorganic pyrophosphatase activity in various normal and neoplastic cell types. *Cancer Res* 43: 3742-3747, 1983.
31. Sreerama L and Sládek NE: Primary breast tumor levels of suspected molecular determinants of cellular sensitivity to cyclophosphamide, ifosfamide, and certain other anticancer agents as predictors of paired metastatic tumor levels of these determinants. Rational individualization of cancer chemotherapeutic regimens. *Cancer Chemother Pharmacol* 47: 255-262, 2001.
32. Marselos M and Michalopoulos G: Changes in the pattern of aldehyde dehydrogenase activity in primary and metastatic adenocarcinomas of the human colon. *Cancer Lett* 34: 27-37, 1987.
33. Charafe-Jauffret E, Ginestier C, Iovino F, *et al*: Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Res* 69: 1302-1313, 2009.
34. Zhou H, Liu Y, Chui J, *et al*: Investigation on glycosylation patterns of proteins from human liver cancer cell lines based on the multiplexed proteomics technology. *Arch Biochem Biophys* 459: 70-78, 2007.
35. Edberg DD, Bruce JE, Siems WF and Reeves R: In vivo post-translational modifications of the high mobility group A1a proteins in breast cancer cells of differing metastatic potential. *Biochemistry* 43: 11500-11515, 2004.