British Journal of Cancer (2008) 98, 596–603

www.bjcancer.com

Proteomic analysis of stage I primary lung adenocarcinoma aimed at individualisation of postoperative therapy

J Maeda[1], T Hirano*[,1], A Ogiwara[2,3], S Akimoto[2,3], T Kawakami[2], Y Fukui[4], T Oka[4], Y Gong[1], R Guo[1], H Inada[1], K Nawa[1], M Kojika[1], Y Suga[1], T Ohira[1], K Mukai[5] and H Kato[1]

[1]Department of Surgery, Tokyo Medical University, 6-7-1 Nishi-shinjuku, Shinjuku-ku, Tokyo 160-0023, Japan; [2]Clinical Proteome Center, Tokyo Medical University, 2-6-1 Nishi-shinjuku, Shinjuku-ku, Tokyo 163-0217, Japan; [3]Medical ProteoScope Co. Ltd, 2-6-1 Nishi-shinjuku, Shinjuku-ku, Tokyo 163-0217, Japan; [4]Taiho Pharmaceutical Co. Ltd, Tokyo, Japan; [5]Department of Pathology, Tokyo Medical University, 6-7-1 Nishi-shinjuku, Shinjuku-ku, Tokyo 160-0023, Japan

Although postoperative adjuvant chemotherapy (PAC) with uracil–tegafur significantly improves the prognosis of patients with stage I lung adenocarcinoma, subset analysis has revealed that only 11.5% of patients with stage IB derive actual benefit from such therapy. Therefore, it is extremely important to identify patients for whom adjuvant chemotherapy will be beneficial. We performed comprehensive protein analysis of 24 surgically resected specimens of stage I adenocarcinoma using liquid chromatography-tandem mass spectrometry (LC-MS/MS), followed by bioinformatical investigations to identify protein molecules. Furthermore, we carried out immunohistochemical studies of 90 adenocarcinoma specimens to validate the results of LC-MS/MS. We detected two kinds of protein molecules (myosin IIA and vimentin) by LC-MS/MS. We confirmed their immunohistochemical expression and distribution, and evaluated the relationship between the expression of these proteins and prognosis after adjuvant chemotherapy. Patients with no expression of either myosin IIA or vimentin showed a significantly better outcome regardless of PAC using uracil–tegafur. However, we were unable to select responders to uracil–tegafur using these proteins. Cases of adenocarcinoma lacking expression of either myosin IIA or vimentin show a good outcome without PAC, and therefore do not require such treatment.
British Journal of Cancer (2008) **98,** 596–603. doi:10.1038/sj.bjc.6604197 www.bjcancer.com
Published online 22 January 2008

Keywords: myosin IIA; vimentin; postoperative adjuvant chemotherapy; responder to uracil–tegafur; stage I lung adenocarcinoma

Death due to lung cancer is still increasing in most industrialised countries, including Japan, despite improvement of various diagnostic and therapeutic modalities. Even though the opportunities to detect lung cancer at an early stage are increasing, approximately 60 000 patients with lung cancer die every year in Japan, usually due to distant metastasis. Distant metastasis, including intrapulmonary metastasis, frequently occurs in patients with advanced-stage non-small cell lung cancer (NSCLC) who undergo only surgical resection, because in such cases micrometastases probably exist at the time of surgery. The concept of postoperative adjuvant chemotherapy (PAC) for control of micrometastasis does not conflict with the improved prognosis of NSCLC patients. However, the efficacy of PAC in patients after complete resection of NSCLC was a matter of controversy in the 1990s. Even as recently as 2003, the efficacy of PAC could not be demonstrated (Scagliotti *et al*, 2003). In 2004, however, some studies demonstrate a beneficial effect of PAC (Arriagada *et al*, 2004; Strauss *et al*, 2004; Winton *et al*, 2005). We have also reported that PAC with oral uracil-tegafur (DPD Inhibitory Fluoropyrimidine, Taiho Pharmaceutical Co. Ltd, Tokyo, Japan) provided better survival than surgical treatment alone in patients with stage I adenocarcinoma of the lung (Kato *et al*, 2004). The combination of uracil and tegafur (also referred to as UFT) at a molar ratio of 4 : 1 is an oral anticancer agent with good absorption in the small intestine (Fujii *et al*, 1979). Tegafur is a prodrug that is gradually converted into fluorouracil in the liver by the cytochrome P-450 enzyme system. Uracil enhances the serum concentration of fluorouracil by competitive inhibition of dihydropyrimidine dehydrogenase, the enzyme responsible for fluorouracil catabolism (Ikenaka *et al*, 1979). Oral uracil-tegafur generates a higher maximal plasma level of fluorouracil than protracted intravenous infusion of fluorouracil at a dose that is equimolar to the amount of tegafur in uracil-tegafur (Ho *et al*, 1998).

Even though PAC with uracil-tegafur has significantly improved the prognosis of patients with stage I primary lung adenocarcinoma, even subset analysis of stage IB has revealed that 11.5% of patients actually derive some benefit from the treatment (Kato *et al*, 2004). Nonresponders to uracil-tegafur, including relapse-free patients without any adjuvant therapy, gain no benefit from PAC. In this context, it is important to establish biomarkers for prediction of responders to uracil-tegafur, and/or for favourable prognosis without the use of PAC.

Clarification of the entire human genome is one of the most significant events in the history of bioscience, and has accelerated the comprehensive analysis of human genes and their protein

*Correspondence: Dr T Hirano; E-mail: thirano@tokyo-med.ac.jp
Received 10 August 2007; revised 22 November 2007; accepted 18 December 2007; published online 22 January 2008

products. Many biologists recognise the importance of protein analysis, because proteins play central role various cellular functions. However, in cancer research, there has been a tendency for most researchers to avoid investigation of cancer-related proteins, because their structures are more complicated than those of the genes. Nevertheless, techniques for the comprehensive analysis of proteins have improved greatly in recent years. The concept of comprehensive protein analysis has been established, and the new research field of proteomics has been developed. One of the main purposes of clinical proteomics in the field of oncology is the development of new therapeutic strategies for cancer, centred on individualised therapy. In this study, we attempted to identify biomarkers for the selection of responders to uracil–tegafur and nonresponders, including relapse-free patients without any requirement for PAC, using clinical proteomics methodology.

MATERIALS AND METHODS

Institutional review board approval for this investigation

The institutional review board approved the use of proteomics analysis to explore biomarkers for selection of responders to oral uracil–tegafur (294/323/480/702).

Surgical samples of stage I lung adenocarcinoma for mass spectrometry

After obtaining written informed consent, lung cancer tissues were obtained from patients with pathologically confirmed stage I adenocarcinoma resected at Tokyo Medical University Hospital between 1995 and 2001. Tissues were kept frozen at −80°C until use. We collected 11 lung adenocarcinoma specimens from 11 patients who subsequently underwent PAC using uracil–tegafur for more than 2 years. In 5 of these 11 patients, recurrent lesions were detected within 2 years after surgery (U1R1), and the remaining 6 were confirmed to be disease-free for 5 years after surgery (U1R0). Furthermore, 13 specimens of lung cancer were collected from patients receiving no adjuvant therapy after surgery. In 6 of these 13 patients, recurrence was recognised within 2 years after surgery (U0R1), but no recurrent lesions were detected during 5 years after surgery in the other 7 (U0R0).

Protein extraction

The surgically resected materials were suspended and homogenised in PBS supplemented with a protease inhibitor cocktail (Roche Diagnostics Inc., Basel, Switzerland) at 4°C. The cell lysate was then fractionated by ultracentrifugation (52 000 g, 4°C, 20 min). The resulting pellet containing plasma membranes from the cells, was solubilised in PBS containing 5% SDS with continuous ultrasonication. The resulting solution was taken as the insoluble fraction, whereas the supernatant from the ultracentrifugation, containing mainly cytosolic proteins, was taken as the soluble fraction. The total protein concentrations of both fractions were measured (Lowry *et al*, 1951) using bovine serum albumin as a standard.

Protein condensation with SDS-PAGE

We added 150 pmol egg white lysozyme (Sigma-Aldrich Inc., St Louis, MO, USA) to an aliquot containing 75 μg protein from each fraction, then dried it under vacuum. The mixture was then solubilised in sample buffer (Laemmli, 1970) with gentle stirring at 37°C for 1 h. A two-third volume of the solution containing 50 μg sample protein and 100 pmol lysozyme was subjected to SDS-PAGE on 12.5% polyacrylamide gel 1-mm-thick. SDS-PAGE was carried out at a constant current of 20 A until the bromophenol blue marker passed the boundary between the stacking and separation gels. In this 'halfway' running, most proteins remained stacked in a small area of approximately 2 mm in height between the gel boundary and the blue marker. After electrophoresis, this small gel area was excised from the gel slab, and the proteins were fixed in the gel slice with an excess volume of aqueous solution containing 40% methanol and 10% acetic acid.

In-gel tryptic digestion of protein

The gel slice was subjected to an in-gel tryptic digestion process (Shevchenko *et al*, 1996), with minor modifications. Briefly, after *S*-carboxyamidomethylation of Cys residues with iodoacetamide, the gel slice was incubated in a small volume of 50 mM ammonium bicarbonate buffer solution containing 1 μg of trypsin (Promega Co., Madison, WI, USA). The resulting peptides were extracted from the gel matrix, and dried under vacuum.

Liquid chromatography-tandem mass spectrometry

The peptide mixture (1 μg) was analysed using a liquid chromatography-tandem mass spectrometry (LC-MS/MS) system in a fully automated manner (Kawakami *et al*, 2005). Briefly, reversed-phase peptide separation was performed on a C18 capillary LC column (Michrom BioResources Inc., Auburn, CA, USA) at a flow rate of 1 μl/min. For gasification of the protonated peptides, the LC effluent was directly interfaced with an electrospray ionisation (ESI) source in a positive ion mode modified on a Finnigan LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) (Schwartz *et al*, 2002). The ESI used a Teflon-coated spray needle (20 μm id, AMR Inc., Tokyo, Japan). The ESI-MS/MS operation and continuous data acquisition of full MS scan and subsequent three MS/MS scans were carried out on an Xcalibur system controller (Thermo Fisher Scientific).

Semiquantitative analysis

All full MS data were investigated using an i-OPAL semiquantitative LC-MS data analysis system (i-OPAL algorithm: Patent no. WO 2004/090526 AI). First, the signal intensity of the full MS scan was normalised so that the total signal intensity of each sample became the same value. Several standard signals derived either from the injected egg white lysozyme or from sample intrinsic common proteins were selected as i-OPAL alignment markers. The i-OPAL alignment programme was used to align the nonlinearly fluctuating LC retention time axis of all LC-MS data to finally generate a single combined LC-MS data set for the soluble and the insoluble fractions, respectively. Analysis of variance (ANOVA) was applied for each peak signal in the final combined LC-MS data set to select candidate marker signals whose intensity differed significantly in a particular patient group. ANOVA was carried out using a Spotfire DecisionSite package.

Database searches

All MS/MS data were investigated using the Mascot search engine (Matrix Science Ltd., London, UK, http://www.matrixscience.com) against the *Homo sapiens* (human) subset of the Swiss-Prot and the RefSeq protein sequence databases. The database searches were performed allowing for fixed modification of cysteine residues (*S*-carbamidomethylation, +57.0 Da) and variable modification of methionine residues (oxidation, +16.0 Da), peptide mass tolerance ±2.0 Da and fragment *m/z* tolerance ±0.8.

Surgical specimens of stage I lung adenocarcinoma for immunohistochemical staining

Sample set A for confirmation of LC-MS semiquantitative results To confirm the semiquantitative results of LC-MS, 23

Table 1 Clinical features of lung adenocarcinoma cases subjected to LC-MS/MS

Characteristic	(%)
Age (year)	
Median	65.0
Range	32–78
Gender	
Male	19 cases (79.2%)
Female	5 cases (20.8%)
Pathological stage	
IA	10 cases (41.7%)
IB	14 cases (58.3%)
Presence of recurrence	
(+)	11 cases (45.8%)
(−)	13 cases (54.2%)
PAC	
(+)	11 cases (45.8%)
(−)	13 cases (54.2%)

PAC = postoperative adjuvant chemotherapy with oral uracil–tegafur.

Table 2 Clinical features of lung adenocarcinoma cases as revealed by immunohistochemical staining

Characteristic	(%)
Age (year)	
Median	64.8
Range	45–82
Gender	
Male	54 cases (60.0%)
Female	36 cases (40.0%)
Pathological stage	
IA	33 cases (36.7%)
IB	57 cases (63.3%)
Existence of recurrence	
(+)	41 cases (45.6%)
(−)	49 cases (54.4%)
PAC with urasil-tegafur	
(+)	51 cases (56.7%)
(−)	39 cases (43.3%)

PAC = postoperative adjuvant chemotherapy.

formalin-fixed, paraffin-embedded specimens derived from the same cases as those used for LC-MS analysis were collected for immunohistochemical investigation. As one formalin-fixed specimen had already been exhausted for the previous investigations, the remaining 23 specimens were investigated.

Sample set B for validation To validate the expression of the protein molecules on lung adenocarcinoma cells, 90 formalin-fixed, paraffin-embedded specimens from patients with lung adenocarcinoma, resected at Tokyo Medical University Hospital between 1995 and 2001, were used. All the patients had undergone curative resection of lung cancer, and after surgery, a pathologically definitive diagnosis of stage I adenocarcinoma had been obtained. We evaluated recurrence after surgery using chest roentgenography and serum tumour markers (CEA, CA19-9 and SLX) every 3 months and computed tomography of the head and body, and bone scintigraphy, every 6 months. When it was difficult to evaluate roentgenographically whether the lesion was recurrent or not, either cytological or pathological examinations were performed to obtain a definitive diagnosis (Table 1).

Of the 90 patients, 51 underwent PAC using uracil–tegafur. These 51 cases included 24 recurrences (U1R1) within 5 years and 27 cases without recurrence (U1R0) within 5 years after surgical treatment. The remaining 39 patients did not receive any adjuvant chemotherapy. These 39 patients included 17 with recurrence (U0R1) and 22 without recurrence (U0R0). The clinicopathological backgrounds of the 90 patients with lung adenocarcinoma are summarised in Table 2.

Immunohistochemical staining of surgically resected specimens of stage I lung adenocarcinoma

Four-micrometer-thick tissue sections were prepared from formalin-fixed, paraffin-embedded surgical specimens and collected on glass slides. The sections were stained immunohistochemically by the ABC method using either anti-myosin IIA mouse monoclonal antibody (clone ab24762, abcam, Cambridge, CB4 0FW, UK) (diluted 1:500) or anti-vimentin antibody (Dako Cytomation, Denmark A/S) (diluted 1:100) as the first antibody. After deparaffinisation, specimens were treated with 0.01% trypsin and an autoclave antigen retrieval system (Barbareschi *et al*, 1994). Sequentially, after inhibition of endogenous peroxidase activity with 0.5% hydrogen peroxide and incubation with 2% normal swine serum, the first antibody was applied. Biotinylated anti-mouse immunoglobulin (Vector Laboratories Inc., Burlingam, CA, USA) was applied as the second antibody (diluted 1:200), followed by application with avidin–biotin peroxidase complex (Vector Laboratories Inc.) (diluted 1:100). The specimens were reacted with 0.06% 3,3'-diaminobenzidine tetrahydrochloride and 0.03% hydrogen peroxide in Tris-buffered saline to visualise the positive areas. Meyer's haematoxylin was used for counterstaining.

Evaluation of immunohistochemically stained preparations

Cells showing cytoplasmic staining were evaluated as positive. For myosin IIA immunohistochemical staining, we evaluated a case as positive when more than 50% of the cells were stained. Also, for vimentin immunostaining, cases in which more than 25% of the cells were stained were evaluated as positive. For both kinds of staining, normal alveolar epithelium served as an internal negative control.

Statistical analysis

Statistical analysis was carried out using the SPSS program. Statistical significance of the relationship between recurrence and immunohistochemical reactivity was evaluated using χ^2 test. Disease-free survival curves were calculated from the day of surgery using the Kaplan–Meier method, and the significance of differences in survival rates between the patient groups was calculated by the log-rank test. In all statistical analyses, a P-value of <0.05 was taken to indicate a statistically significant difference.

RESULTS

LC-MS data analysis

After i-OPAL alignment and peak detection, we obtained 13 136 signal peaks from the soluble fraction and 14 984 peaks from the insoluble fraction. Using Spotfire, we restricted the candidate signal peaks on the basis of the following conditions:

(1) A Mascot search result with a score equal to or more than 50.
(2) An ANOVA P-value equal to or less than 1×10^{-5} (for the soluble fraction) or 1×10^{-6} (for the insoluble fraction).

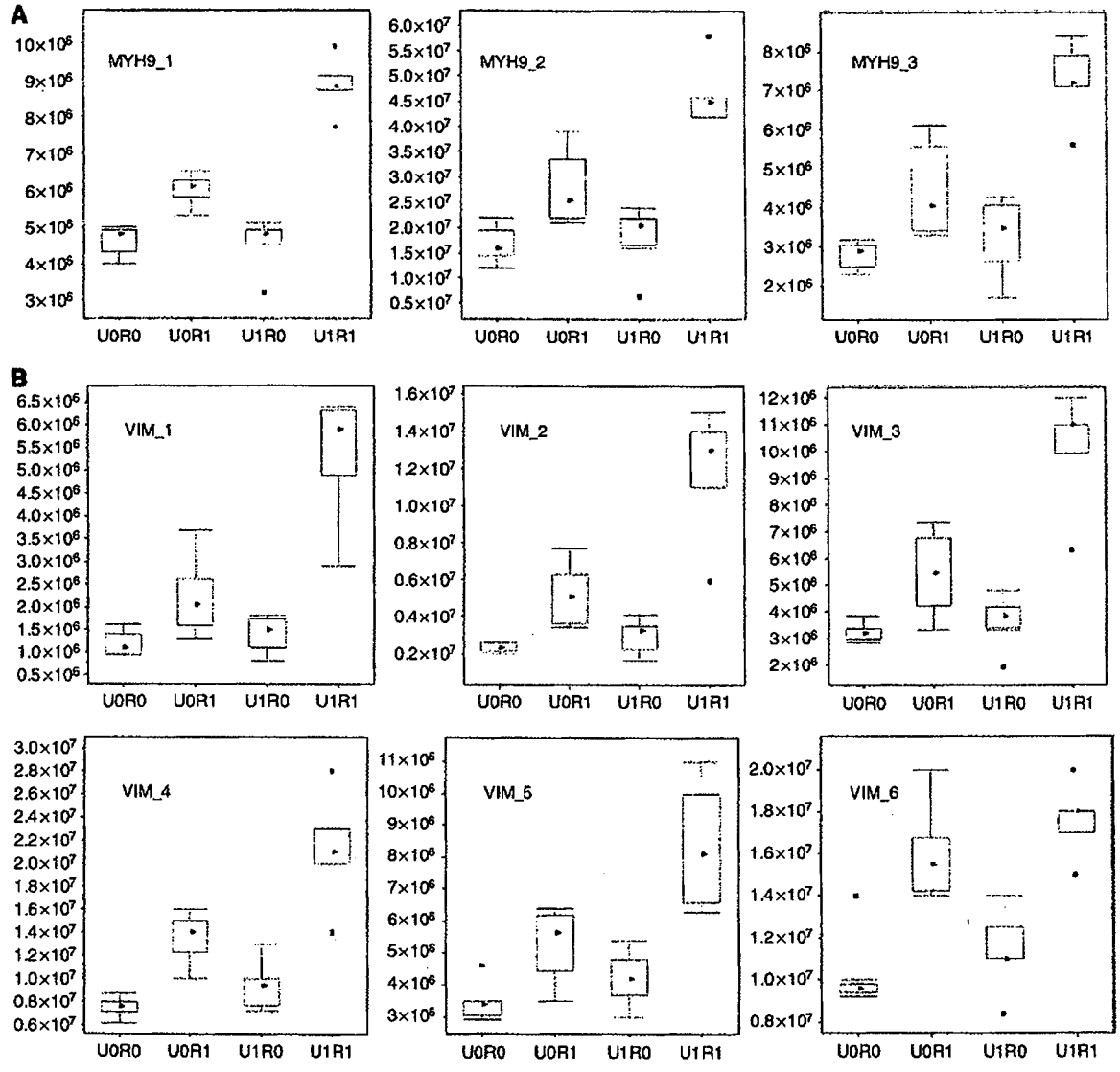

Figure 1 Comparison of the intensity of peptide signals originating from the same protein molecule in each group detected by LC-MS. The vertical axis indicates normalised signal intensity measured by LC-MS. In each box plot, the upper and lower sides of the box represent the upper and the lower quartile values (Q3/Q1), and the upper and lower horizontal bars outside the box indicate the upper and the lower adjacent values (UAV/LAV). Note that UAV is the largest observation value that is less than or equal to $Q3+1.5\times(Q3-Q1)$, and LAV is the smallest observation greater than or equal to $Q1-1.5\times(Q3-Q1)$. Black triangle marks represent the median values, and black square marks represent outliers. U0R0: patients without PAC showing no recurrence within 5 years after surgery. U0R1: patients without PAC in showing recurrence within 5 years after surgery. U1R0: patients who received PAC with uracil–tegafur and showed no recurrence within 5 years after surgery. U1R1: patients who received PAC with uracil–tegafur and showed recurrence within 5 years after surgery. (**A**) These three peptide signals were shown by MS/MS to have originated from myosin IIA. There was a significant difference between the U1R1 and the other groups ($P<9.7\times10^{-7}$). (**B**) These six peptide signals were shown by MS/MS to have originated from vimentin. There was also a significant difference between the U1R1 and the other groups ($P<8.3\times10^{-6}$).

As the peptide compositions of the soluble and the insoluble fractions differed, we applied different criteria to obtain approximately the same number of candidate signals. As a result, we were able to restrict the number of candidate signals to 23 and 28 for the soluble and insoluble fractions, respectively. From the restricted candidate signals, we selected several myosin IIA and vimentin signals as final candidate biomarker signals, because these two candidate biomarkers were identified by more than one distinct peptide sequence, and almost all of these signals had similar patterns of intensity (Figure 1A and B; Table 3).

Table 3 lists the amino-acid sequences from the selected candidate biomarker signals described above. These sequences were identified from MS/MS data using Mascot software.

Figure 1A shows the distribution of the signal intensity of several peptide ions derived from myosin IIA, and Figure 1B shows the signal intensity distribution of vimentin-derived peptide ions. For most signals, the intensity for group U1R1 patients showed patterns that differed significantly (i.e., were markedly higher) from those of the other patient groups.

Immunohistochemical staining of myosin IIA and vimentin

Representative staining of myosin IIA and vimentin is shown in Figure 2, and a summary of the immunohistochemical data is presented in Table 4. Cytoplasmic staining was observed in cases

Table 3 Amino-acid sequences from the selected peptide ion signals

Name	Fraction	Sequence
Myosin, heavy polypeptide 9, non-muscle		
MYH9_1	Insoluble	IRELESQISELQEDLESER
MYH9_2	Insoluble	KANLQIDQINTDLNLER
MYH9_3	Insoluble	HEMPPHIYAITDTAYR
Vimentin		
VIM_1	Insoluble	ETNLDSLPLVDTHSK
VIM_2	Insoluble	NLQEAEEWYK
VIM_3	Insoluble	LGDLYEEEMR
VIM_4	Insoluble	LLQDSVDFSLADAINTEFK
VIM_5	Soluble	SGDAAIVDMVPGKPMCVESFSDYPPLGR
VIM_6	Soluble	ILTVEDHYYEGGIGEAVSSAVVGEPGITVTHLAVNR

positive for myosin IIA and vimentin. We evaluated cases in which more than 50% of the cells showed immunohistochemical reactivity for myosin IIA, considered as overexpressing (positive). We also evaluated overexpressing (positive) cases in which more than 25% of the cells showed immunohistochemical reactivity for vimentin. On the basis of these criteria, we evaluated sample sets A and B.

Immunohistochemical evaluation of sample set A (Table 4A): all patients with cancers lacking expression of both myosin IIA and vimentin showed relapse-free survival at 5 years. On the other hand, all patients with cancers showing positive expression of both myosin IIA and vimentin suffered disease recurrence.

Immunohistochemical evaluation of sample set B (Table 4B): among 90 cases, 75 (83.3%) showed overexpression of myosin IIA, and 48 (53.3%) showed overexpression of vimentin. There was no

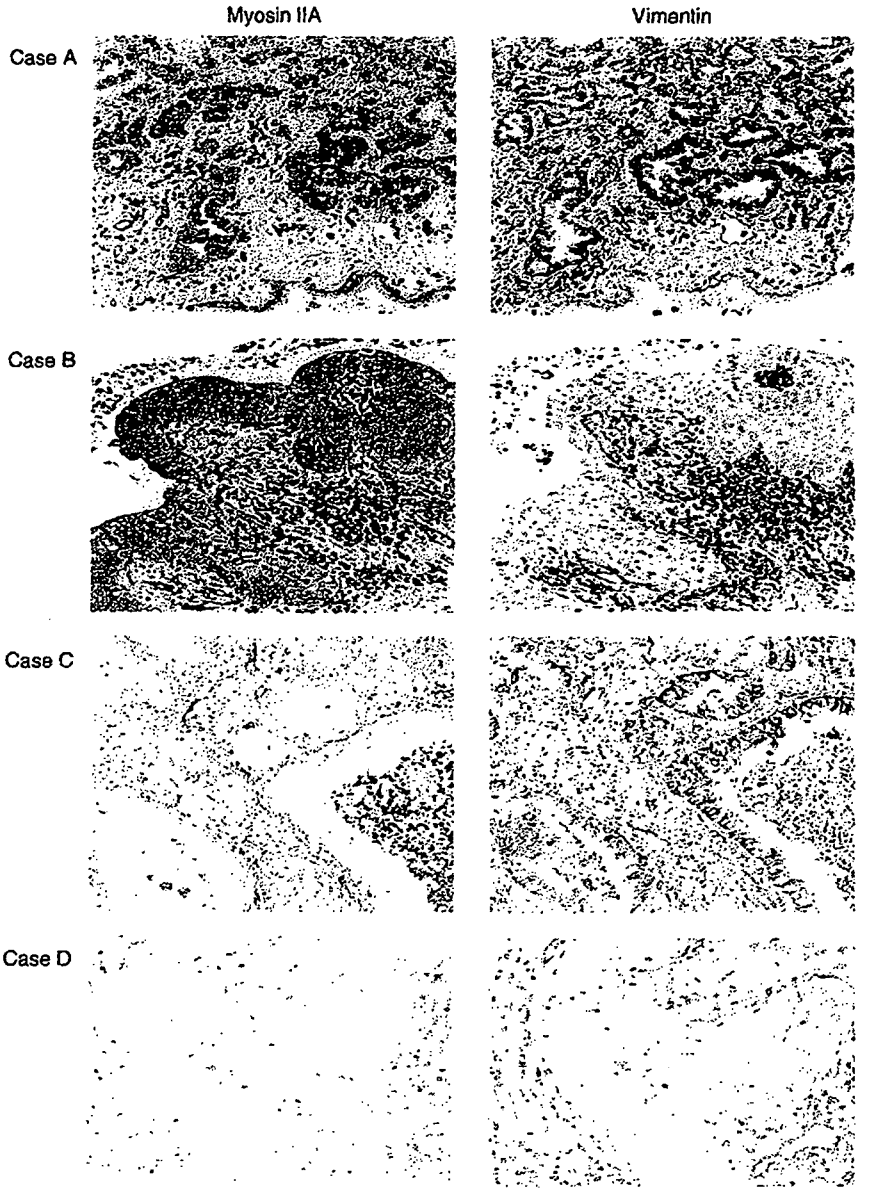

Figure 2 Immunohistochemical reactivity of representative cases using anti-myosin IIA antibody, ab24762 (abcam, Cambridge, UK) and anti-vimentin antibody (Dako Cytomation, Denmark A/S). Case A showed positive cytoplasmic staining for both myosin IIA and vimentin. Case B showed positive cytoplasmic staining for myosin IIA and negative cytoplasmic staining for vimentin. Case C showed negative cytoplasmic staining for myosin IIA and positive cytoplasmic staining for vimentin. Case D showed negative cytoplasmic staining for both myosin IIA and vimentin.

Table 4 Relationship between PAC, recurrence and immunohistochemical reactivity for myosin IIA and vimentin

	M(−)V(−)	M(−)V(+) or M(+)V(−)	M(+)V(+)	ND
(A) Sample set A (n = 24) derived from the same cases as those subjected to LC-MS/MS				
U0R0	5	2	0	0
U1R0	2	3	0	1
U0R1	0	2	4	0
U1R1	0	0	5	0
(B) Sample set B (n = 90) for validation by immunohistochemical analysis				
U0R0	6	11	5	0
U1R0	3	12	12	0
U0R1	0	6	11	0
U1R1	0	10	14	0

M = expression of myosin IIA; ND = not done; U0R0 = patients without PAC showing no recurrence within 5 years after surgery; U0R1 = patients without PAC showing recurrence within 5 years after surgery; U1R0 = patients who received PAC with uracil–tegafur and showed no recurrence within 5 years after surgery; U1R1 = patients who received PAC with uracil–tegafur and showed recurrence within 5 years after surgery; V = expression of vimentin. *Statistically significant difference between U0R0 and U1R1 was detected ($P = 0.008$).

relationship between the immunohistochemical reactivities of myosin IIA and vimentin. All nine patients whose cancers lacked immunohistochemical reactivity for both myosin IIA and vimentin showed relapse-free survival at 5 years. Among cases that were immunohistochemically negative for both myosin IIA and vimentin, we recognised a statistically significant difference between U1R1 and U0R0 ($P = 0.008$), but there were no significant differences between U1R1 and the other groups.

Disease-free survival and coexpression of myosin IIA and vimentin in sample set B

The non-relapse survival curves of cases with/without PAC are shown in Figure 3A and B. Irrespective of whether patients had undergone PAC or not, the non-relapse survival rate of cases lacking expression of both myosin IIA and vimentin was 100%. Among patients who had not undergone PAC, there was a statistically significant difference between cases lacking expression of both myosin IIA and vimentin and cases that were positive for both ($P = 0.011$) (Figure 3A). Among the patients who received PAC, there was no statistically significant difference in this respect (Figure 3B). When the cases showing positive expression of both myosin IIA and vimentin were evaluated, we recognised a 5-year-survival rate benefit of approximately 19% in patients who had undergone PAC with uracil–tegafur, but there was no statistically significant difference in this respect between patients who had and who had not received PAC.

When we evaluated the non-relapse survival curves of all the studied cases, there were statistically significant differences between cases negative for myosin IIA and vimentin expression and cases that were positive for both ($P = 0.006$), and between cases positive for either myosin IIA or vimentin and cases that were negative for both ($P = 0.029$; Figure 4).

DISCUSSION

Lung cancer is the leading cause of cancer death in Japan, and its incidence is still increasing. Even if surgical resection involving either lobectomy or pneumonectomy accompanied by lymph node dissection is performed at a relatively early stage, distant metastasis often occurs within a few years. More than 20% of patients with stage I NSCLC suffer recurrence caused by distant

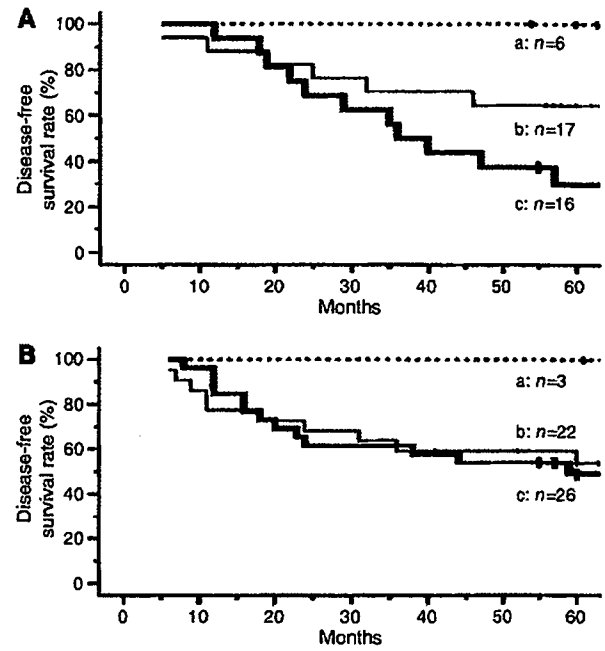

Figure 3 Kaplan–Meier curves for disease-free survival after complete resection in patients with stage I lung adenocarcinoma who received PAC with uracil–tegafur (**A**), or did not receive any PAC (**B**). a: Cases lacking both myosin IIA and vimentin expression (non-relapse survival rate at 5 years: 100% in panels **A** and **B**). b: Cases negative for myosin IIA expression and positive for vimentin expression, or positive for myosin IIA and negative for vimentin expression (non-relapse survival rate at 5 years: 64.7% in panel **A** and 53.7% in panel **B**). c: Cases positive for both myosin IIA and vimentin expression (non-relapse survival rate at 5 years: 30.0% in panel **A** and 49.0% in panel **B**). In patients who did not receive adjuvant chemotherapy, there was a statistically significant difference in disease-free survival between those who were negative and those who were positive for both proteins (a–c: $P = 0.011$). No significant difference in this respect was recognised in patients who received PAC with uracil–tegafur.

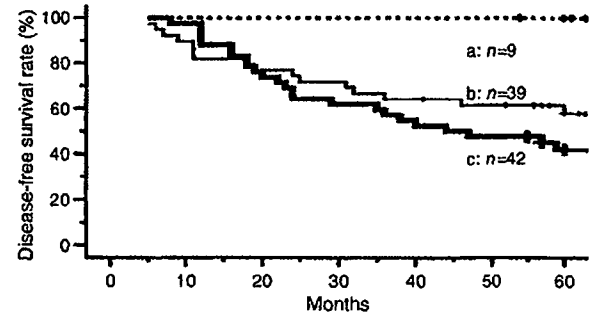

Figure 4 Kaplan–Meier curves for disease-free survival after complete resection in patients with stage I lung adenocarcinoma. a: Cases lacking both myosin IIA and vimentin expression (non-relapse survival rate at 5 years: 100%). b: Cases negative for myosin IIA and positive for vimentin, or positive for myosin IIA and negative for vimentin (non-relapse survival rate at 5 years: 58.0%). c: Cases positive for both myosin IIA and vimentin (non-relapse survival rate at 5 years: 42.0%). Group a: showed significantly higher survival than group b, and significantly higher survival than group c: (a–b: $P = 0.029$; a–c: $P = 0.006$).

metastasis. Distant metastasis is the most frequent mode of recurrence in patients who undergo surgical resection of lung cancer, and it is believed that in such patients, micrometastasis is

invariably present at the time of initial treatment. If an efficient PAC regimen could be devised for total control of micrometastasis, then the prognosis of patients with lung cancer would be markedly improved. A meta-analysis conducted in the 1990s showed that PAC using platinum-based agents had no effect on the survival of patients with NSCLC, even though previous studies had suggested a 5% increase in survival at 5 years (Non-Small Cell Lung Cancer Collaborative Group, 1995). At the 2004 ASCO meeting, the results of two randomised adjuvant trials showing the efficacy of platinum-based chemotherapy – the CALGB-9633 trial (carboplatin and paclitaxcel) (Strauss *et al*, 2004) and the JBR 10 trial (cisplatin and vinorelbine) – were reported. Furthermore, a recent large-scale randomised clinical trial involving meta-analysis concluded that patients assigned to cisplatin-based PAC had a significantly higher survival rate than those assigned to post-operative observation (44.5 *vs* 40.4% at 5 years; $P<0.03$) (Arriagada *et al*, 2004). Also our previous study showed that PAC with uracil–tegafur conferred a survival benefit for patients with resected stage I adenocarcinoma of the lung (Kato *et al*, 2004). Also, meta-analysis of PAC with tegafur–uracil supported this result (Hamada *et al*, 2005). However, even though a significant difference was found in this study, the 5-year-survival rate benefit of this therapy was 11.5% for stage IB adenocarcinoma (Kato *et al*, 2004). At present, although leading lung cancer experts appear to have reached a consensus concerning the effectiveness of PAC, none of the present PAC regimens are of benefit to more than 15% of patients with NSCLC. In this context, it is very important to predict the response to PAC and to select potential responders before carrying out PAC. Therefore, we attempted to identify biomarkers of either responders or nonresponders including relapse-free patients without PAC using uracil–tegafur, for selection of patients who would benefit from PAC using proteomic analysis of surgically resected specimens of stage I lung adenocarcinoma.

As proteins play a role in both physiological and pathological functions, it is now recognised that investigation of proteins is essential to obtain an accurate grasp of cellular physiology. Recent advances in proteomic techniques, including two-dimensional polyacrylamide gel electrophoresis and MS, have brought hope that the pathogenesis of any type of malignant neoplasm will be ultimately clarified. We believe that the present concepts of proteomic analysis will prove to be extremely valuable in the field of clinical oncology, and will lead to the development of new therapeutic strategies. Liquid chromatography-tandem mass spectrometry enables simultaneous evaluation of a large number of polypeptides, and furthermore, MS/MS has made it possible to identify protein molecules by obtaining information about their amino-acid sequences. We attempted to identify proteins associated with the effectiveness of postoperative uracil–tegafur chemotherapy and the favourable prognosis of stage I adenocarcinoma, and detected two kinds of protein molecules (myosin IIA and vimentin) showing significantly high expression in the group that suffered recurrence despite administration of uracil–tegafur, in comparison with the other groups. Our semiquantitative results of LC-MS were confirmed by immunohistochemistry for myosin IIA and vimentin (Table 4A).

Nonmuscle myosin IIA is a major component of the actomyosin cytoskeleton and is generally considered to contribute to contraction of the cell posterior during migration (Ridley *et al*, 2003). However, there is still a profound lack of understanding of the exact mechanical roles of myosin IIA during cell migration. A recent clinical study of patients with NSCLC found a significant positive correlation between the expression levels of myosin light chain kinase (which activates myosin II) and the likelihood of disease recurrence and metastasis (Minamiya *et al*, 2005), indicating that myosin IIA activation could be a factor contributing to metastasis. A key role for myosin IIA in cancer cell metastasis has been further suggested, indirectly, by a number of published studies focusing on the small calcium-binding protein, metastasin-1. This protein is upregulated in many metastasis cell lines, and when overexpressed enhances metastatic behaviour (Davies *et al*, 1993). A major cellular target of metastasin-1 seems to be myosin IIA (Garrett *et al*, 2006). Although studies of metastasin-1 suggest critical roles for myosin IIA in metastasis, it remains completely unknown how myosin IIA contributes to metastasis, and which isoforms are important for this process.

Vimentin is the most ubiquitous intermediate filament protein and the first to be expressed during cell differentiation. All primitive cell types express vimentin, but in most nonmesenchymal cells, it is replaced by other intermediate filament proteins during differentiation. Vimentin is expressed in a wide variety of mesenchymal cell types (fibroblasts, endothelial cells, etc), and also in a number of other cell types derived from mesoderm, mesothelium and ovarian granulose cells. Epithelial–mesenchymal transition is a key mechanism operating in the normal development of multicellular organisms. During this process, epithelial cells progressively acquire a reversible or irreversible mesenchymal phenotype that is essential for organogenesis (Thiery, 2002). Morphogenetic epithelial–mesenchymal transition is aberrantly recapitulated during tumorigenesis in a variety of epithelial cancers, including those of the thyroid, liver, kidney, prostate, breast and lung (Arias, 2001; Thiery, 2002). The common signature of this process involves disruption of normal epithelial integrity, with loss of morphological features including polarised epithelia, and partial or total gain of mesenchymal markers with progressive acquisition of a motile and invasive phenotype (Islam *et al*, 1996). In addition to a disrupted epithelial morphology, dysregulation of adhesion and junctional molecules and aberrant expression of *N*-cadherin, epithelial–mesenchymal transition involves *de novo* expression of other mesenchymal markers, such as fibronectin and vimentin in epithelial cells. Aberrant expression of vimentin in tumours and transformed cell lines has been correlated with increased motility, invasive behaviour and poor prognosis (Gilles *et al*, 1996; Hendrix *et al*, 1997). Recently, it was reported that the presence of vimentin-positive tumour cells mainly in fibrotic areas is consistent with other studies that have shown a correlation between tumour fibrosis and epithelial–mesenchymal transition (Blanco *et al*, 2004).

These two molecules identified by proteomic analysis might reflect the cellular functions of metastasis and the mechanism of recurrence of malignant neoplasms. We attempted to validate the results of LC-MS/MS using immunohistochemistry of an additional sample set (sample set B: 90 surgically resected lung cancer specimens) with monoclonal antibodies against the two proteins. The results showed that cases lacking expression of the two proteins had a good prognosis, irrespective of whether the patients had undergone PAC. Therefore, these two proteins appear to be potentially useful biomarkers for the selection of patients who do not require PAC. In the cases positive for both of these proteins, the 5-year-survival benefit was approximately 19% in patients with adenocarcinoma who underwent PAC with uracil–tegafur. However, there was no significant difference between patients who did, and did not, undergo PAC with uracil–tegafur. Therefore, in this investigation, we failed to select patients who might benefit from this adjuvant chemotherapy. A larger-scale investigation is therefore needed to establish suitable biomarkers for the selection of patients who might benefit from PAC with uracil–tegafur, because a few per cent of patients with stage I adenocarcinoma do obtain such a benefit.

Individualised chemotherapy for lung cancer patients is currently attracting attention, because the efficacy of systemic chemotherapy using any single agent is less than 30%. Therefore, it is extremely important to select patients who might benefit from chemotherapy. Until an ideal chemotherapy agent is established, we propose that rather than focusing only on improving the efficacy of chemotherapy regimens, we should also make efforts to identify patients who will show a good response to regimens that

are already established. This proposal is justified only on the basis of evidence-based medicine. In this situation, positive indicators for the effectiveness of PAC with uracil-tegafur are needed. However, in this study, we were unable to detect novel biomarkers for selection of good responders. The two protein molecules detected in this proteomic analysis were biomarkers indicative of good prognosis.

The ultimate purpose of clinical proteomics is to improve diagnostic procedures including the exact evaluation of biological characteristics of tumour cells and to understand the molecular pathogenesis of cancers to devise novel therapeutic strategies. We believe that proteomic analysis will become an integral tool for investigation of tumour biology. We conclude that negative expression of both myosin IIA and vimentin is an indicator of good prognosis for stage I lung adenocarcinoma without the need for PAC.

ACKNOWLEDGEMENTS

This study was supported by grants from the Japan Society for the Promotion of Science (18390382). We thank Mrs Mami Murakami for her experimental expertise and statistical analysis.

REFERENCES

Arias AM (2001) Epithelial mesenchymal interactions in cancer and development. *Cell* 105: 425–431

Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon JP, Vansteenkiste J (2004) Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* 350: 351–360

Barbareschi M, Girlanda S, Mauri MF, Forti S, Eccher C, Nauri FA, Togni R, Dalla Palma P, Doglioni C (1994) Quantitative growth fraction evaluation with MIB1and Ki67 antibodies in breast carcinomas. *Am J Clin Pathol* 102: 171–175

Blanco D, Vicent S, Elizegi E, Pino I, Fraga MF, Esteller M, Saffiotti U, Lecanda F, Montuenga LM (2004) Altered expression of adhesion molecules and epithelial-mesenchymal transition in silica-induced rat lung carcinogenesis. *Lab Invest* 84: 999–1012

Davies BR, Davies MP, Gibbs FE, Barraclough R, Rudland PS (1993) Induction of the metastatic phenotype by transfection of a benign rat mammary epithelial cell line with the gene for p9Ka, a rat calcium-binding protein, but not with the oncogene EJ-ras-1. *Oncogene* 8: 999–1008

Fujii S, Kitano S, Ikenaka K, Shirasaka T (1979) Effect of coadministration of uracil or cytosine on the anti-tumor activity of clinical doses of 1-(2-tetrahydrofuryl)-5 fluorouracil and level of 5 - fluorouracil in rodents. *Gann* 70: 209–214

Garrett SC, Varney KM, Weber DJ, Bresnick AR (2006) S100A4, a mediator of metastasis. *J Biol Chem* 281: 677–680

Gilles C, Polette M, Piette J, Delvigne AC, Thompson EW, Foidart JM, Birembaut P (1996) Vimentin expression in cervical carcinomas: association with invasive and migratory potential. *J Pathol* 180: 175–180

Hamada C, Tanaka F, Ohta M, Fujimura S, Kodama K, Imaizumi M, Wada H (2005) Meta-analysis of postoperative adjuvant chemotherapy with tegafur-uracil in non-small-cell lung cancer. *J Clin Oncol* 23: 4999–5006

Hendrix MJ, Seftor EA, Seftor RE, Trevor KT (1997) Experimental co-expression of vimentin and keratin intermediate filaments in human breast cancer cells results in phenotypic interconversion and increased invasive behavior. *Am J Pathol* 150: 483–495

Ho DH, Pazdur R, Covington W, Brown N, Huo YY, Lassere Y, Kuritani J (1998) Comparison of 5-fluorouracil pharmacokinetics in patients receiving continuous 5 -fluorouracil infusion and oral uracil plus N1-(2'-tetrahydrofuryl)-5-fluorouracil. *Clin Cancer Res* 4: 2085–2088

Ikenaka K, Shirasaka T, Kitano S, Fujii S (1979) Effect of uracil on metabolism of 5-fluorouracil *in vitro*. *Gann* 70: 353–359

Islam S, Carey TE, Wolf GT, Wheelock MJ, Johnson KR (1996) Expression of *N*-cadherin by human squamous carcinoma cells induces a scattered fibroblastic phenotype with disrupted cell–cell adhesion. *J Cell Biol* 135: 1643–1654

Kato H, Ichinose Y, Ohta M, Hata E, Tsubota N, Tada H, Watanabe Y, Wada H, Tsuboi M, Hamajima N, Ohta M (2004) A randomized trial of adjuvant chemotherapy with uracil-tegafur for adenocarcinoma of the lung. *N Engl J Med* 350: 1713–1721

Kawakami T, Tateishi K, Yamano Y, Ishikawa T, Kuroki K, Nishimura T (2005) Protein identification from product ion spectra of peptides validated by correlation between measured and predicted elution times in liquid chromatography/mass spectrometry. *Proteomics* 5: 856–864

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* 193: 265–275

Minamiya Y, Nakagawa T, Saito H, Matsuzaki I, Taguchi K, Ito M, Ogawa J (2005) Increased expression of myosin light chain kinase mRNA is related to metastasis in non-small cell lung cancer. *Tumour Biol* 26: 153–157

Non-Small Cell Lung Cancer Collaborative Group (1995) Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomized clinical trials. *BMJ* 311: 899–909

Ridley AJ, Schwartz MA, Burridge K, Firtel RA, Ginsberg MH, Borisy G, Parsons JT, Horwitz AR (2003) Cell migration: integrating signals from front to back. *Science* 302: 1704–1709

Scagliotti GV, Fossati R, Torri V, Crino L, Giaccone G, Silvano G, Martelli M, Clerici M, Cognetti F, Tonato M (2003) Randomized study of adjuvant chemotherapy for completely resected stage I, II, or IIIA non-small-cell lung cancer. *J Nat Cancer Inst* 95: 1453–1461

Schwartz JC, Senko MW, Syka JE (2002) A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom* 13: 659–669

Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68: 850–858

Strauss GM, Herndon J, Maddaus MA, Johnstone DW, Johnson EA, Watson DM, Sugarbaker DJ, Schilsky RL, Green MR (2004) Randomized clinical trial of adjuvant chemotherapy with paclitaxel and carboplatin following resection in stage IB non-small cell lung cancer (NSCLC): report of cancer and leukaemia Group B (CALGB) Protocol 9633. *Proc Am Soc Clin Oncol* 23: 621

Thiery JP (2002) Epithelial–mesenchymal transitions in tumour progression. *Nat Rev Cancer* 2: 442–454

Winton T, Livingston R, Johnson D, Rigas J, Johnston M, Butts C, Cormier Y, Goss G, Inculet R, Vallieres E, Fry W, Bethune D, Ayoub J, Ding K, Seymour L, Graham B, Tsao MS, Gandara D, Kesler K, Demmy T, Shepherd F (2005) Vinorelbine plus cisplatin *vs* observation in resected non-small-cell lung cancer. *N Engl J Med* 352: 2589–2597

Proteomic Signature Corresponding to the Response to Gefitinib (Iressa, ZD1839), an Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor in Lung Adenocarcinoma

Tetsuya Okano,[1,6] Tadashi Kondo,[1] Kiyonaga Fujii,[2] Toshihide Nishimura,[2] Toshimi Takano,[4] Yuichiro Ohe,[4] Koji Tsuta,[5] Yoshihiro Matsuno,[5] Akihiko Gemma,[6] Harbumi Kato,[2,3] Shoji Kudoh,[6] and Setsuo Hirohashi[1]

Abstract **Purpose:** We aimed to identify candidate proteins for tumor markers to predict the response to gefitinib treatment.

Experimental Design: We did two-dimensional difference gel electrophoresis to create the protein expression profile of lung adenocarcinoma tissues from patients who showed a different response to gefitinib treatment. We used a support vector machine algorithm to select the proteins that best distinguished 31 responders from 16 nonresponders. The prediction performance of the selected spots was validated by an external sample set, including six responders and eight nonresponders. The results were validated using specific antibodies.

Results: We selected nine proteins that distinguish responders from nonresponders. The predictive performance of the nine proteins was validated examining an additional six responders and eight nonresponders, resulting in positive and negative predictive values of 100% (six of six) and 87.5% (seven of eight), respectively. The differential expression of one of the nine proteins, heart-type fatty acid–binding protein, was successfully validated by ELISA. We also identified 12 proteins as a signature to distinguish tumors based on their *epidermal growth factor receptor* gene mutation status.

Conclusions: Study of these proteins may contribute to the development of personalized therapy for lung cancer patients.

Non-small cell lung carcinoma (NSCLC) accounts for ~85% of lung cancer cases (1). Biomarker(s) that predict the response to gefitinib (Iressa; AstraZeneca, Macclesfield, United Kingdom), an epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor, may help to improve the choice of therapeutic strategy in patients with NSCLC. Gefitinib improves NSCLC-related symptoms and quality of life in some patients with advanced NSCLC who do not respond to platinum-based chemotherapy. However, the response rate for gefitinib remains <20% in patients with NSCLC (2–4), and treatment with gefitinib is associated with serious adverse effects, such as severe acute interstitial pneumonia in 5.4% of the patients who received the treatment (5, 6). Thus, it is imperative to select appropriate patients for treatment with gefitinib and exclude patients in whom gefitinib is unlikely to exhibit any clinical benefit. Women, patients who have never smoked, patients with adenocarcinoma, and East Asians are major subgroups of responders (3, 4, 6–8). Recently, gain-of-function somatic mutation in the tyrosine kinase domain of the EGFR has been correlated with the response to gefitinib (9, 10). However, other studies have revealed that correction of the phenotype arising from EGFR mutation may not account for all of the clinical benefits of gefitinib (11, 12), and both preclinical and clinical studies have reported that the efficacy of gefitinib is independent of EGFR expression level (11, 13–15). Although molecular features of the *EGFR* gene, including mutation and high copy number, (16, 17) are associated with response to gefitinib, other molecular markers in the tumor, such as HER2 overexpression (18), Akt phosphorylation (19), and other EGFR downstream molecules (20), also correlate with response. These observations suggest a role for unknown, but important, factors in gefitinib sensitivity. Identification and elucidation of such factors will improve existing therapeutic protocols and contribute to further understanding of the mechanisms of gefitinib sensitivity.

Authors' Affiliations: [1]Proteome Bioinformatics Project, National Cancer Center Research Institute; [2]Clinical Proteome Center and [3]Department of Surgery, Tokyo Medical University; [4]Department of Internal Medicine and [5]Clinical Laboratory Division, National Cancer Center Hospital; and [6]The Fourth Internal Department of Medicine, Nippon Medical School, Tokyo, Japan

Received 7/7/06; revised 11/17/06; accepted 11/29/06.

Grant support: 'Third-Term Comprehensive Control Research for Cancer' conducted by the Ministry of Health, Labor, and Welfare and by the Program for Promotion of Fundamental Studies in Health Sciences in the National Institute of Biomedical Innovation of Japan. Tetsuya Okano is the recipient of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research (Japan).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Current address for K. Fujii: Proteome Bioinformatics Project, National Cancer Center Research Institute, Tokyo, Japan.

Requests for reprints: Tadashi Kondo, Proteome Bioinformatics Project, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. Phone: 81-3-3542-2511, ext. 3004; Fax: 81-3-3457-5298; E-mail: takondo@gan2.res.ncc.go.jp.

doi:10.1158/1078-0432.CCR-06-1654

To identify the gene products correlated with the efficacy of gefitinib, genome-wide screening was done recently for NSCLC. A global mRNA expression study using DNA microarrays and biopsy samples identified 51 genes associated with the sensitivity to gefitinib and established a numerical scoring system to predict the response (21). This expression study also led to the establishment of ELISA assays for the identified gene products in serum. Preclinical studies involving mRNA profiling of NSCLC xenografts resulted in the identification of a set of genes that were differentially expressed between tumors that were sensitive and insensitive to gefitinib treatment (22, 23). These studies will lead to the identification of novel biomarkers to predict the response to gefitinib treatment. However, mRNA expression does not necessarily correlate with protein level, and posttranslational modifications, such as phosphorylation, cannot be predicted from the amount of RNA or from the DNA sequence (24). With this background, comprehensive expression studies at the protein level, an approach called proteomics, have been conducted in patients with lung cancer to develop biomarkers that predict clinical outcomes (25). However, no global protein expression study has yet been done on the mechanism of response to gefitinib.

To identify the proteomic signature for sensitivity to gefitinib and to use that signature as a tumor marker to predict the response to gefitinib, we analyzed global protein expression levels in lung adenocarcinoma tissues for whom we have detailed information on *EGFR* gene status. The surgical specimens were obtained at the time of surgery from patients who subsequently had recurrence and received gefitinib monotherapy. We then used two-dimensional difference gel electrophoresis (2D-DIGE) covering ~2,000 proteins to identify a set of proteins of which expression was associated with sensitivity to gefitinib and with EGFR mutation. The predictive performance of the protein set was validated with an independent data set and compared with that of EGFR mutation.

Materials and Methods

Patients and tissue samples. We examined tumor tissues from patients who relapsed after surgery and received gefitinib monotherapy. Two hundred seventy-nine patients who received gefitinib at the National Cancer Center Hospital from July 2002 to December 2004 were evaluated for inclusion in this study. Ninety-two patients relapsed after surgical resection of primary NSCLC and started to receive monotherapy with gefitinib 250 mg/d for 14 days (n = 92). We used tumor tissues obtained at the time of surgery and stored in vapor nitrogen. Fifteen patients were excluded from our study for the following reasons: frozen tissues were not available (n = 10) and tumor histology showed squamous cell carcinoma (n = 4) or pleomorphic carcinoma (n = 1). The histologic features of the tissues were reviewed by two board-certified pathologists (Y.M and K.T.) and diagnosis was based on the latest WHO classification of lung adenocarcinoma (8, 26–28). The tumor responses were classified into complete response (CR), partial response (PR), and progressive disease (PD) using standard bidimensional measurements (29). In this study, patients without a marked reduction of tumor size were subdivided into minor response (MR) and stable disease (SD) groups. MR was defined as a 25% decrease in the sum of the products of perpendicular diameters of all measurable lesions at any point during gefitinib treatment. SD was defined as a <25% decrease in tumor size after treatment. The clinical information is summarized in Table 1, and further information, including EGFR mutation status, is summarized in Supplementary Table S1. Consent was obtained from all patients and the protocol was approved by the institutional review board of the National Cancer Center.

To identify the proteins associated with response to gefitinib, we compared the protein expression profiles of responders (CR and PR) and nonresponders (PD). Of 77 samples available, the effects of gefitinib treatment were not examined for six cases because the treatment was not completed. These six samples were excluded from this study. We constructed two sample sets in the following way (Table 2): a training sample set comprising 31 responders (2 CRs + 29 PRs) and 16 nonresponders (16 PDs) and a test set comprising six responders (6 PRs) and 8 nonresponders (8 PDs) from whom samples were obtained between June and December 2004 (Table 2). As no significant differences were observed between CRs and PRs (Supplementary Fig. S1A), we grouped CRs and PRs together in the responder group.

Protein extraction and protein expression profiling. The frozen tumor tissues were crushed to frozen powder with a Multi-Beads Shocker (Yasui-kikai, Osaka, Japan) under cooling with liquid nitrogen. The frozen powder was then treated with urea lysis buffer (7 mol/L urea, 2 mol/L thiourea, 3% CHAPS, 1% Triton X-100) for 30 min on ice. After centrifugation at 15,000 rpm for 30 min, the supernatant was recovered as cellular protein for the protein expression study.

Protein samples were labeled with CyDye DIGE Fluor saturation dye (GE Healthcare Amersham Biosciences, Uppsala, Sweden) according to

Table 1. Patient characteristics

	No. patients	%
Gender		
Female	33	43
Male	44	57
Age (y)		
Median (range)	62.2 (32-80)	—
Histologic type		
Adenocarcinoma		100
Papillary/acinar/ bronchioloalveolar/solid	30/16/9/6	49/26/15/10
Smoking history*		
Never smokers	37	48
Former smokers	12	16
Current smokers	28	36
ECOG performance status†		
0/1/2/3	24/39/9/5	31/51/12/6
Prior chemotherapy		
Yes	30	39
No	47	61
Response to gefinitib		
CR/PR/MR/SD/PD/NE	2/35/2/8/24/6	3/45/3/10/31/8
EGFR gene status		
Mutation L858R	18	23.4
DEL‡	18	23.4
G719§	2	2.6
Wild-type	35	45.4
Unknown	4	5.2

Abbbreviation: NE, not evaluated.
*Never-smokers: those who had never had a smoking habit; former smokers: those who had stopped smoking at least 1 yr before diagnosis; and current smokers: active smokers at diagnosis of NSCLC or those who had stopped smoking less than 1 yr before diagnosis.
†ECOG performance status was monitored according to the previous report (44).
‡Deletional mutations in exon 19.
§G719S and G719C.

Table 2. Training and test sets to develop the classifier for the response to gefitinib

	Training set			Test set		
	Responders, $n = 31$ (%)	Nonresponders, $n = 16$ (%)	P	Responders, $n = 6$ (%)	Nonresponders, $n = 8$ (%)	P
Age						
Mean ± SD	64.0 ± 8.9	60.5 ± 12.0	0.330	57.5 ± 12.8	62.8 ± 6.1	0.386
Gender						
Male	17 (55)	9 (56)	0.927	3 (50)	5 (62.5)	0.640
Female	14 (45)	7 (44)		3 (50)	3 (37.5)	
Smoking history						
Never smokers	17 (55)	9 (56)	0.286	4 (67)	4 (50)	0.054
Former smokers	7 (22.5)	1 (6)		2 (33)	0 (0)	
Current smokers	7 (22.5)	6 (38)		0 (0)	4 (50)	
EGFR gene status						
Mutation	27 (87)	1 (6)	<0.001	4 (66)	0 (0)	0.006
Wild type	3 (10)	13 (81)		1 (17)	8 (100)	
Unknown	1 (3)	2 (13)		1 (17)	0 (0)	
Prior chemotherapy						
(+)	12 (39)	5 (31)	0.614	6 (100)	0 (22)	<0.001
(−)	19 (61)	11 (69)		0 (0)	8 (100)	
Performance status						
0	11 (35.5)	6 (37.5)	0.945	2 (33)	1 (12.5)	0.347
1	11 (35.5)	10 (62.5)		4 (67)	7 (87.5)	
2	6 (19)	0 (0)		0 (0)	0 (0)	
3	3 (10)	0 (0)		0 (0)	0 (0)	

our previous report (30). We prepared an internal control consisting of a mixture of small portions of all protein samples obtained before May 2004 (31). The internal control sample and the individual experimental samples were labeled with Cy3 and Cy5 CyDye DIGE Fluor saturation dyes, respectively. Five micrograms of Cy3- or Cy5-labeled protein were mixed and coseparated by two-dimensional PAGE. The first-dimension separation was achieved on an Immobiline pH gradient gel (isoelectric point range, 4-7; 24 cm length) with a Multiphor II (GE Healthcare Amersham Biosciences). The second-dimension separation was done with an EttanDalt II (GE Healthcare Amersham Biosciences) with a 9% to 15% gradient polyacrylamide gel. After electrophoresis, the gels were scanned at appropriate wavelengths for Cy3 and Cy5 (Supplementary Fig. S2A). The ratio between Cy5 and Cy3 intensity was calculated for all protein spots in identical gels by the use of DeCyder software (GE Healthcare Amersham Biosciences; ref. 31). The standardized spot intensities were then logarithmically transformed and subjected to a data-mining package (Impressionist; GeneData, Basel, Switzerland). We ran triplicate gels for each sample and calculated the averaged standardized spot intensity.

To assess the reproducibility of the proteomic data with the internal control in our analyses, we generated triplicate protein profiles from identical samples (case 9; Supplementary Table S1) and compared the standardized intensity of the paired spots (Supplementary Fig. S2B). Scattergrams with 1,980, 1,646, and 1,873 spots showed that the intensities of 1,916 (93.7%), 1,599 (94.7%), and 1,770 (94.5%) spots, respectively, were scattered within a 2-fold difference, and the correlation values were also high (r values > 0.93; Supplementary Fig. S2B).

Data analysis. A bioinformatic approach based on a support vector machine (SVM) algorithm and a leave-one-out cross-validation was used to identify proteins of which expression was associated with tumor characteristics, including therapeutic response to gefitinib and the presence of EGFR mutation (32).

Protein identification. Proteins corresponding to the protein spots of interest were identified by mass spectrometry (30). The proteins were recovered in a gel plug by using an automated spot collector (SpotPicker; GE Healthcare Amersham Biosciences) and digested with sequence grade trypsin (Promega, Madison, WI; ref. 30). Trypsin digests were applied to liquid chromatography coupled with tandem mass spectrometry (LTQ, Thermo, Waltham, MA). A database search against Swiss-Prot was done with Mascot software. Patients with a Mascot score of 35 or more were used for protein identification. When multiple proteins were identified in a single spot, the proteins with the highest number of peptides were considered as those corresponding to the spot.

Mutations in the EGFR gene. EGFR mutations in the samples obtained between July 2002 and May 2004 were examined as described in our previous report (8). Analysis of samples obtained between June 2004 and December 2004 was done by high-resolution melting analysis with a LightCycler HR-1 system (Idaho Technology Inc., Salt Lake City, UT).

ELISA. The expression level of heart-type fatty acid-binding protein (H-FABP) in protein samples from 55 lung adenocarcinoma patients (2 CRs, 28 PRs, 6 SDs, 1 MR, and 18 PDs) was measured in a clinical laboratory (SRL, Tokyo, Japan) with a commercially available ELISA kit (MARKIT-M H-FABP, Dainippon Pharmaceutical, Tokyo, Japan) according to the manufacturer's instructions (Supplementary Table S1). All these 55 samples were included in a 2D-DIGE analysis set in this study.

Results

Proteomic signature for the response to gefitinib. We first selected 1,685 protein spots that appeared in at least 80% of the images of Cy3-labeled internal control. We further selected 87 protein spots that showed different intensities between responder and nonresponder groups ($P < 0.05$, Wilcoxon test). Although potentially resulting in a loss of information, this trimming process decreased the possibility that the classifier would be significantly influenced by irrelevant expression data. We selected protein sets for which expression was associated with response to gefitinib by using a SVM algorithm. Accuracy, plotted as a function of spot number, was constant until the number of spots decreased to less than nine, showing that accurate classification did not require all protein spots (Fig. 1A). The location on the two-dimensional map is shown for the selected nine spots (Fig. 1B; Supplementary Fig. S3).

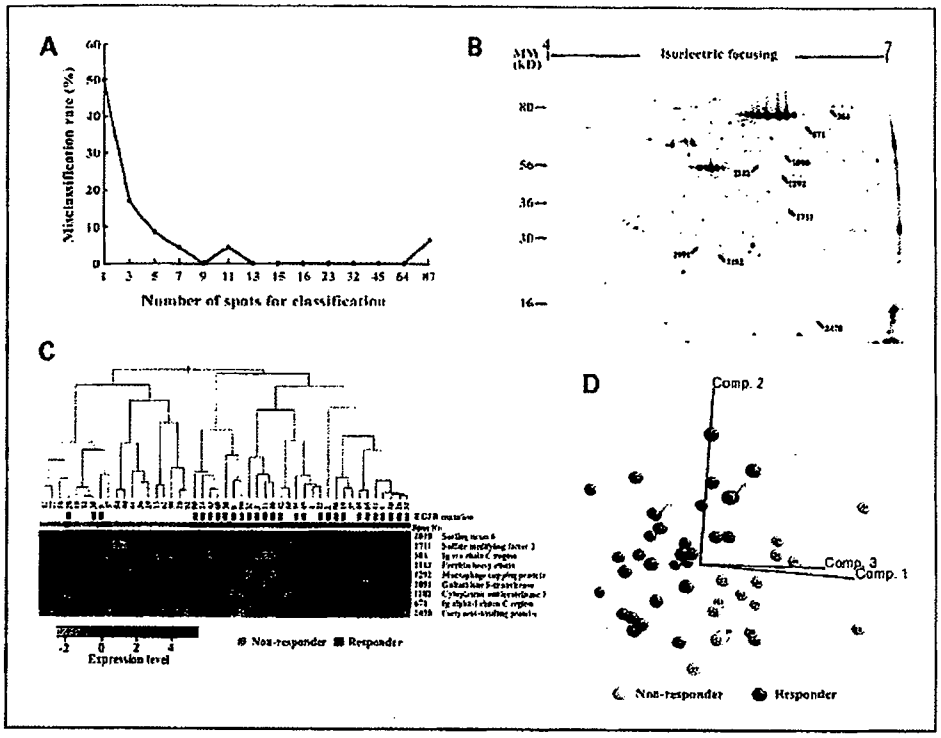

Fig. 1. Data-mining procedure to develop the prediction model for the response to gefitinib. *A*, a spot ranking method selected a few protein spots by which the cumulative error rate of a leave-one-out cross-validation became minimal. The spot ranking method indicated that the error rate was minimal when the prediction model was constructed by a particular nine protein spots. *B*, localization of the selected nine protein spots on the two-dimensional map. An enlarged two-dimensional image is shown in Supplementary Fig. S2. *C*, hierarchical clustering analysis of the samples in the learning set using the selected nine protein spots. Black bars, the presence of EGFR mutations within exons 18 to 21. *D*, principal component analysis of the samples in the learning set using the selected nine protein spots. Comp.1, 2, and 3, the first component 1, 2, and 3, respectively.

Mass spectrometry revealed that these nine spots corresponded to nine gene products (Table 3). Overall similarity of the selected spots is shown in Supplementary Fig. S1B and C. As the responder group in the training set consisted mainly of PRs, the obtained proteomic signature would presumably be more reflective of PR than CR.

The classification performance of the selected nine protein spots was validated by unsupervised classification. Hierarchical clustering showed that all tumor samples in the training set, except for cases 5, 20, and 37, were grouped according to their sensitivity to gefitinib based on the expression pattern of the nine proteins (Fig. 1C). In principal component analysis, all 47 samples seemed to be separated into two groups, although the border between these groups was not clear (Fig. 1D). Although hierarchical clustering and principal component analysis are crude methods of validation of classification, the results obtained using them were consistent.

To validate the predictive performance of the nine protein spots, we investigated a newly enrolled test sample set that was completely independent of the learning set. Based on the expression level of the nine protein spots, the distance of each sample from the hyperplane created by the SVM algorithm, defined as the SVM value, was calculated. The samples with a positive SVM value were grouped as responders and the samples with a negative SVM value were grouped as nonresponders. As a consequence, all training set samples were correctly classified in accordance with their clinical response to gefitinib (Fig. 2). All responders (six PRs) and seven of eight nonresponders (eight PDs) in the test set were also correctly classified. The expression pattern of the nine protein spots in the nonresponder patient (case 75) was more similar to that of the responder group, for unknown reasons. We also validated the results using the samples from patients who showed MR and SD. We found that the two patients showing MR were categorized as responders and that among the eight patients showing three SDs were classified into the responder group and five SDs into the nonresponder group. We did a leave-one-out cross validation for all 47 samples in the training set and the test set using nine protein spots with 1,000 times random permutation. All but two cases, cases 37 and 75, were correctly classified according to their status of response to the treatment. The overall misclassification error rate was 3.3%. Consequently, the model predicted the response to gefitinib in 13 of the 14 (92.8%) newly enrolled samples from the responders and nonresponders and may be useful for disease monitoring.

Proteomic signature for EGFR gene mutation. We studied the spots on the prediction for EGFR mutation. We set a training sample set, including 58 samples (34 mutation-positive samples and 24 mutation-negative samples; Supplementary Table S2). We found that the 12 protein spots showed the high correlation with the EGFR mutation (Supplementary Data; Supplementary Figs. S4-6). The classification and prediction performance of the selected 12 protein spots was successfully validated using the external validation sample set, including four mutation-positive samples and 11 mutation-negative samples (Supplementary Fig. S7). Only one protein, sulfate modifying factor 2, was shared between the signatures for the response and for the mutation (Table 3; Supplementary Table S3).

Expression of H-FABP measured by ELISA. We validated the differential expression of the identified proteins by the use of a widely available clinical assay. The expression level of H-FABP in the same tumor samples as those used in 2D-DIGE was measured with a commercially available ELISA kit intended for serum assays (Fig. 3). H-FABP expression measured by ELISA was highly correlated with that measured by 2D-DIGE (Pearson correlation, 0.76295; $P < 0.0001$). The ELISA study also showed

that the expression level of H-FABP was significantly different between the responder (PR and CR) and nonresponder (PD) groups (P = 0.0031, Mann-Whitney U test) and also between the patients with MR or SD and the nonresponder group (P = 0.0047, Mann-Whitney U test). These results indicate that up-regulation of H-FABP in tumor tissues can be monitored by routine clinical methods.

Discussion

We identified 87 protein spots of which the intensity was statistically significantly different between samples from the responder (CR and PR) and nonresponder (PD) groups in the training set. Application of a data-mining procedure allowed identification of a set of nine protein spots that accurately distinguished between responders and nonresponders. The different expression levels of these nine protein spots allowed classification of 13 of 14 of our test PR and PD cases in accordance with their clinical response to gefitinib. These protein spots classified cases showing a MR to gefitinib (MR) into the responder group. The intermediate cases, SD, were categorized into both responder and nonresponder groups. The usefulness of our findings will be validated in a larger clinical data set.

Table 3. List of proteins for the response to gefitinib

Spots no.*	Rank	Accession no.†	Identified protein†	MW (DA)‡	pI‡	Ion charge state (+)	MZ (obs)§	Mass‖	δ¶	Miss**	Mascot ions score††	Peptide sequence
384	5	Q96RP9	Ig mu chain C region	49,557	6.35	2	810.3	1,617.7	0.91	0	74	QVGSGVTTDQVQAEAK
						2	640.1	1,277.5	0.63	0	47	YAATSQVLLPSK
671	1	P01876	Ig α-1 chain C region	37,655	6.08	2	919.2	1,836.0	0.32	0	68	QEPSQGTTTFAVTSILR
						2	771.8	1,540.7	0.91	0	54	DASGVTFTWTPSSGK
1090	7	Q9UNH7	SNX 6	46,649	5.81	2	636.5	1,270.5	0.55	0	73	NLVELAELELK
						2	577.0	1,152.2	−0.33	0	39	SLVDYENANK
1182	8	P50453	Cytoplasmic antiproteinase 3	42,404	5.61	2	816.4	1,629.8	0.95	0	82	IEELLPGSSIDAETR
						2	626.6	1,249.4	1.66	0	75	AFQSLLTEVNK
						2	591.0	1,179.5	0.47	0	63	LVLVNAIYFK
						2	757.5	1,513.6	−0.56	0	47	LQEDYDMESVLR +Oxidation (M)
1292	6	P40121	Macrophage capping protein	38,518	5.88	2	633.8	1,264.4	1.18	0	85	VSDATGQMNLTK
						2	676.8	1,351.4	0.05	0	79	YQEGGVESAFHK
						2	932.1	1,861.1	1.11	0	50	MQYAPNTQVEILPQGR +Oxidation (M)
						2	659.8	1,317.3	0.23	0	41	EGNPEEDLTADK
1711	3	Q8NBJ7	Sulfatase modifying factor 2	33,857	7.78	2	792.5	1,581.7	1.32	0	112	MGNTPDSASDNLGFR
						2	779.9	1,557.6	0.15	0	95	GASWIDTADGSANHR
						2	740.0	1,477.6	0.36	0	83	LPTEEEWEFAAR
						2	613.2	1,224.4	−0.02	0	66	FLMGTNSPDSR
						2	629.9	1,256.5	1.27	0	55	SVLWWLPVEK
						2	818.0	1,633.8	0.12	1	55	RLPTEEEWEFAAR
						2	837.7	1,672.9	0.48	0	47	LEHPVLHVSWNDAR
2091	9	P09211	Glutathione *S*-transferase P	23,225	5.44	2	647.5	1,292.5	0.44	0	36	MLLADQGQSWK +Oxidation (M)
2182	4	P02794	Ferritin heavy chain	21,094	5.30	2	823.4	1,643.8	1.04	0	91	MGAPESGLAEYLFDK +Oxidation (M)
						2	648.3	1,294.5	0.03	0	53	NVNQSLLELHK
2478	2	P05413	Fatty acid–binding protein, heart	14,727	6.34	2	735.2	1,467.5	0.81	0	103	LGVEFDETTADDR
						2	798.7	1,595.7	−0.32	1	73	LGVEFDETTADDRK
						2	603.3	1,204.3	0.26	0	70	WDGQETTLVR
						2	455.0	907.0	1.04	0	67	SLGVGFATR
						2	774.7	1,546.8	0.56	0	61	QVASMTKPTTIIEK
						2	438.0	873.0	0.88	0	54	NGDILTLK
						1	889.6	889.0	−0.41	0	45	SIVTLDGGK

Abbreviation: pI, isoelectric point.
*Spot numbers refer to those in Fig. 1B (Supplementary Fig. S3).
†Accession nos. of proteins were derived from Swiss-Prot and National Center for Biotechnology Information nonredundant databases.
‡Theoretical molecular weight and isoelectric point were obtained from Swiss-Prot and the ExPASy database (http://au.expasy.org).
§Experimental m/z value.
‖Relative molecular mass calculated from the peptide sequence.
¶Difference (error) between the experimental and calculated masses.
**Number of missed cleavage sites.
††Mascot ions score (http://www.matrixscience.com/search_form_select.html).

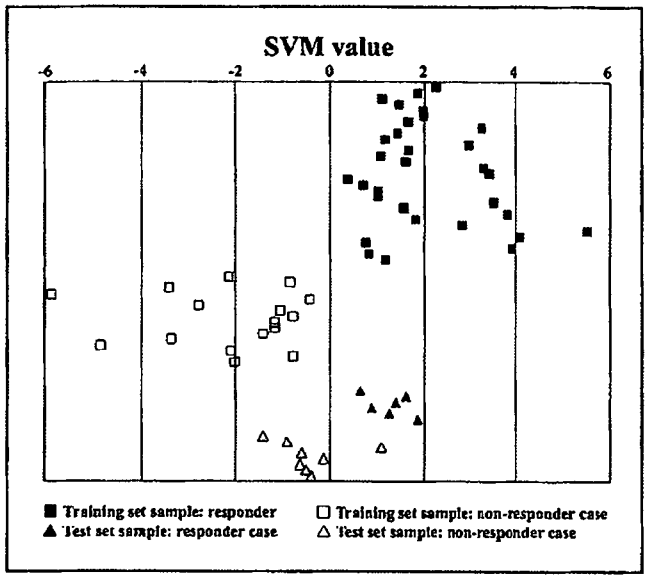

Fig. 2. Predictive performance of the nine spots was validated by examining the SVM value of each sample in the group.

We identified the proteins whose expression was correlated with response to gefitinib and found associations with the EGFR signal pathway and with the biology of lung cancer. Sorting nexin (SNX) 6 is a member of a SNX family that functions in the intracellular trafficking of plasma membrane receptors (33). SNXs form complexes with other SNXs and with plasma membrane receptors. In complexes with SNX1, SNX2, and SNX4, SNX6 interacts with the intercellular portion of the EGFR as well as with transforming growth factor-β receptor, insulin receptor, leptin receptor, and platelet-derived growth factor receptor (34). By binding to the kinase domain of the transforming growth factor-β receptor, SNX6 perturbs transforming growth factor-β signal transduction (34). The other SNX family, SNX1, decreases the expression of EGFR by activating the endosome-to-lysosome pathway with enterophilin-1 (35), although the functions of the complex of SNX6 and EGFR have not yet been reported. The functional association of SNX6 with oncogene product Pim-1, which has been implicated in the development of hematopoietic (36), gastric (37), and prostatic (38) malignancies, suggests the involvement of SNX6 in cancer biology. Kakiuchi et al. (21) reported that another SNX family member, SNX13, was correlated with the response to gefitinib in patients with NSCLC. These reports suggest that SNX6 might play an important role in signal transduction pathways that affect the phenotypes of lung cancer.

We tried to identify the proteins whose expression was associated with EGFR mutation. Because gefitinib is a specific inhibitor of EGFR and mutation of EGFR is considered to be a predictive marker for gefitinib sensitivity, we had expected some similarity between the set of proteins predicting sensitivity to gefitinib and the set of proteins reflecting EGFR mutation status. However, only sulfate modifying factor 2 was common to the two sets. Search of the PubMed database revealed no association of sulfate modifying factor 2 with the EGFR pathway and no evidence for its involvement in resistance to chemotherapy. Similarly, the other proteins correlated with EGFR mutation status had no obvious involvement in the EGFR pathway. Functional studies on these proteins will contribute to further understanding of EGF signaling in cells and to discovery of novel therapeutic targets in lung cancer.

2D-DIGE is a high-performance proteomic technology and a powerful tool to develop candidate biomarkers. However, 2D-DIGE requires expensive fluorescent dyes and well-trained operators to run the gels. Thus, routine clinical studies with multiple large-format two-dimensional gels and a 2D-DIGE protocol are unlikely to be practical. Application of our results requires a simple and cost-effective method that can be used routinely in the clinic. In addition, as we need to examine the expression of multiple proteins, a practical tool for simultaneously measuring the amount of the other proteins is required. With that in mind, we validated measurement of the differential expression of H-FABP by the use of a commercially available ELISA kit (MARLIT-M H-FABP) that is routinely used in hospitals for the early diagnosis of acute myocardial infarction using serum samples. The expression level of H-FABP in tumor tissues as monitored by the ELISA assay was highly correlated with that by 2D-DIGE, and a significant difference in H-FABP expression was observed between responders (CR + PR), minor responders (MR + SD), and nonresponders (PD). Thus, our results can provide a simple and direct method to predict the response to gefitinib.

H-FABP functions in intracellular lipid transport, storage, and metabolism. As H-FABP is highly expressed in heart and released into plasma after myocardial injury, it has been used as a plasma marker for early diagnosis of acute myocardial infarction and stroke. However, many lines of evidence also suggest an association of H-FABP with cancer biology. Higher expression of H-FABP was observed in a more tumorigenic small-cell lung cancer cell line (39) compared with its counterpart. Increased expression of H-FABP is associated with tumor aggressiveness, metastasis, and poor prognosis of gastric cancer (40). In contrast, H-FABP is known to have growth-inhibitory activity in breast cancer cells (41), and breast cancer does not express H-FABP because of gene silencing by hypermethylation (42). These observations suggest complexity in the way that H-FABP is involved in the progression of cancer. Recently, Loeffler-Ragg et al. (43) reported that another FABP family member, E-FABP, is up-regulated in gefitinib-resistant colon cancer cell lines compared with gefitinib-sensitive cell

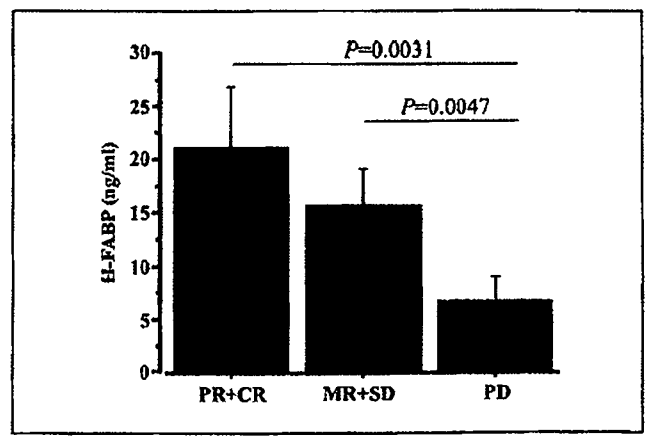

Fig. 3. ELISA assay for H-FABP. The differential expression level of H-FABP was validated by ELISA assay.

lines. Further study on the contribution of the FABP family to cancer phenotypes, including resistance to chemotherapy, will provide novel insights into cancer biology.

In conclusion, our proteomic study has identified proteins whose expression can predict the response to gefitinib in patients with recurrence of lung adenocarcinoma. Large-scale validation of the present results and functional analysis to elucidate the contribution and synergies of the identified proteins in the response to gefitinib will assist in developing novel therapeutic strategies for lung cancer.

References

1. Breathnach OS, Freidlin B, Conley B, et al. Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: sobering results. J Clin Oncol 2001;19:1734–42.
2. Herbst RS. Dose-comparative monotherapy trials of ZD1839 in previously treated non-small cell lung cancer patients. Semin Oncol 2003;30:30–8.
3. Fukuoka M, Yano S, Giaccone G, et al. Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer (The IDEAL 1 Trial). J Clin Oncol 2003; 21:2237–46.
4. Kris MG, Natale RB, Herbst RS, et al. Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. JAMA 2003;290:2149–58.
5. Inoue A, Saijo Y, Maemondo M, et al. Severe acute interstitial pneumonia and gefitinib. Lancet 2003;361: 137–9.
6. Takano T, Ohe Y, Kusumoto M, et al. Risk factors for interstitial lung disease and predictive factors for tumor response in patients with advanced non-small cell lung cancer treated with gefitinib. Lung Cancer 2004;45:93–104.
7. Miller VA, Kris MG, Shah N, et al. Bronchioloalveolar pathologic subtype and smoking history predict sensitivity to gefitinib in advanced non-small-cell lung cancer. J Clin Oncol 2004;22:1103–9.
8. Takano T, Ohe Y, Sakamoto H, et al. Epidermal growth factor receptor gene mutations and increased copy numbers predict gefitinib sensitivity in patients with recurrent non-small-cell lung cancer. J Clin Oncol 2005;23:6829–37.
9. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500.
10. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
11. Cappuzzo F, Gregorc V, Rossi E, et al. Gefitinib in pretreated non-small-cell lung cancer (NSCLC): analysis of efficacy and correlation with HER2 and epidermal growth factor receptor expression in locally advanced or metastatic NSCLC. J Clin Oncol 2003; 21:2658–63.
12. Pao W, Miller VA. Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. J Clin Oncol 2005;23:2556–68.
13. Baselga J, Rischin D, Ranson M, et al. Phase I safety, pharmacokinetic, and pharmacodynamic trial of ZD1839, a selective oral epidermal growth factor receptor tyrosine kinase inhibitor, in patients with five selected solid tumor types. J Clin Oncol 2002;20: 4292–302.
14. Moasser MM, Basso A, Averbuch SD, Rosen N. The tyrosine kinase inhibitor ZD1839 ("Iressa") inhibits HER2-driven signaling and suppresses the growth of HER2-overexpressing tumor cells. Cancer Res 2001; 61:7184–8.
15. Janmaat ML, Kruyt FA, Rodriguez JA, Giaccone G. Response to epidermal growth factor receptor inhibitors in non-small cell lung cancer cells: limited antiproliferative effects and absence of apoptosis associated with persistent activity of extracellular signal-regulated kinase or Akt kinase pathways. Clin Cancer Res 2003; 9:2316–26.
16. Cappuzzo F, Hirsch FR, Rossi E, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. J Natl Cancer Inst 2005;97:643–55.
17. Hirsch FR, Varella-Garcia M, Bunn PA, Jr., et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol 2003;21:3798–807.
18. Hirata A, Hosoi F, Miyagawa M, et al. HER2 overexpression increases sensitivity to gefitinib, an epidermal growth factor receptor tyrosine kinase inhibitor, through inhibition of HER2/HER3 heterodimer formation in lung cancer cells. Cancer Res 2005;65:4263–60.
19. Cappuzzo F, Magrini E, Ceresoli GL, et al. Akt phosphorylation and gefitinib efficacy in patients with advanced non-small-cell lung cancer. J Natl Cancer Inst 2004;96:1133–41.
20. Han SW, Hwang PG, Chung DH, et al. Epidermal growth factor receptor (EGFR) downstream molecules as response predictive markers for gefitinib (Iressa, ZD1839) in chemotherapy-resistant non-small cell lung cancer. Int J Cancer 2005;113:109–15.
21. Kakiuchi S, Daigo Y, Ishikawa N, et al. Prediction of sensitivity of advanced non-small cell lung cancers to gefitinib (Iressa, ZD1839). Hum Mol Genet 2004;13: 3029–43.
22. Zembutsu H, Ohnishi Y, Daigo Y, et al. Gene-expression profiles of human tumor xenografts in nude mice treated orally with the EGFR tyrosine kinase inhibitor ZD1839. Int J Oncol 2003;23:29–39.
23. Jain A, Tindell CA, Laux I, et al. Epithelial membrane protein-1 is a biomarker of gefitinib resistance. Proc Natl Acad Sci U S A 2005;102:11858–63.
24. Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 1999;19:1720–30.
25. Chen G, Gharib TG, Wang H, et al. Protein profiles associated with survival in lung adenocarcinoma. Proc Natl Acad Sci U S A 2003;100:13537–42.
26. Travis WD, Colby TV, Corrin B, Shimosato Y, Brambilla E. Histological typing of lung and pleural tumours. World Health Organization International Classification of Tumors. New York (NY): Springer-Verlag; 1999.
27. Colby TV, Noguchi M, Henschke C, et al. Adenocarcinoma. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. Pathology and genetics: tumors of the lung, pleura, thymus, and heart. Lyon (France): IARC; 2004. p. 35–44.
28. Ebright MI, Zakowski MF, Martin J, et al. Clinical pattern and pathologic stage but not histologic features predict outcome for bronchioloalveolar carcinoma. Ann Thorac Surg 2002;74:1640–6.
29. Green S, Weiss GR. Southwest Oncology Group standard response criteria, endpoint definitions and toxicity criteria. Invest New Drugs 1992;10:239–53.
30. Fujii K, Kondo T, Yokoo H, Yamada T, Iwatsuki K, Hirohashi S. Proteomic study of human hepatocellular carcinoma using two-dimensional difference gel electrophoresis with saturation cysteine dye. Proteomics 2005;5:1411–22.
31. Alban A, David SO, Bjorkesten L, et al. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 2003;3:36–44.
32. Brown MP, Grundy WN, Lin D, et al. Knowledge-based analysis of microarray gene expression data by using support vector machines. Proc Natl Acad Sci U S A 2000;97:262–7.
33. Worby CA, Dixon JE. Sorting out the cellular functions of sorting nexins. Nat Rev Mol Cell Biol 2002;3: 919–31.
34. Parks WT, Frank DB, Huff C, et al. Sorting nexin 6, a novel SNX, interacts with the transforming growth factor-β family of receptor serine-threonine kinases. J Biol Chem 2001;276:19332–9.
35. Pons V, Peres C, Teulie JM, et al. Enterophilin-1 interacts with focal adhesion kinase and decreases β1 integrins in intestinal Caco-2 cells. J Biol Chem 2004;279:9270–7.
36. Hammerman PS, Fox CJ, Birnbaum MJ, Thompson CB. Pim and Akt oncogenes are independent regulators of hematopoietic cell growth and survival. Blood 2005;105:4477–83.
37. Chen CN, Lin JJ, Chen JJ, et al. Gene expression profile predicts patient survival of gastric cancer after surgical resection. J Clin Oncol 2005;23: 7286–95.
38. Xie Y, Xu K, Dai B, et al. The 44 kDa Pim-1 kinase directly interacts with tyrosine kinase Etk/BMX and protects human prostate cancer cells from apoptosis induced by chemotherapeutic drugs. Oncogene 2006;25:70–8.
39. Zhang L, Cilley RE, Chinoy MR. Suppression subtractive hybridization to identify gene expressions in variant and classic small cell lung cancer cell lines. J Surg Res 2000;93:108–19.
40. Hashimoto T, Kusakabe T, Sugino T, et al. Expression of heart-type fatty acid-binding protein in human gastric carcinoma and its association with tumor aggressiveness, metastasis, and poor prognosis. Pathobiology 2004;71:267–73.
41. Huynh HT, Larsson C, Narod S, Pollak M. Tumor suppressor activity of the gene encoding mammary-derived growth inhibitor. Cancer Res 1995;55: 2225–31.
42. Huynh H, Alpert L, Pollak M. Silencing of the mammary-derived growth inhibitor (MDGI) gene in breast neoplasms is associated with epigenetic changes. Cancer Res 1996;56:4865–70.
43. Loeffler-Ragg J, Skvortsov S, Sarg B, et al. Gefitinib-responsive EGFR-positive colorectal cancers have different proteome profiles from non-responsive cell lines. Eur J Cancer 2005;41:2338–46.
44. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982;5: 649–55.

Proteomics-based identification of α-enolase as a tumor antigen in non-small lung cancer

Ping He,[1] Tetsuji Naka,[1,2,8] Satoshi Serada,[3] Minoru Fujimoto,[1,2] Toshio Tanaka,[1] Shoji Hashimoto,[4] Yoshihito Shima,[1] Tomoki Yamadori,[1] Hidekazu Suzuki,[5] Tomonori Hirashima,[5] Kaoru Matsui,[5] Hiroyuki Shiono,[6] Meinoshin Okumura,[6] Toshiro Nishida,[7] Isao Tachibana,[1] Naoko Norioka,[3] Shigemi Norioka[3] and Ichiro Kawase[1]

[1]Department of Molecular Medicine, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871; [2]Laboratory for Immune Signal, National Institute of Biomedical Innovation, 7-6-8 Saito-asagi, Ibaragi, Osaka 567-0085; [3]Functional Proteomics Group, Osaka University Graduate School of Frontier Biosciences, 1-3 Yamadaoka, Suita, Osaka 565-0871; [4]Division of Infection Control and Prevention, Osaka University Hospital, 2-2 Yamadaoka, Suita, Osaka 565-0871; [5]Department of Thoracic Malignancy, Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, 3-7-1 Habikino, Habikino, Osaka 583-8588; [6]Department of Surgery, Division of General Thoracic Surgery, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871; [7]Department of Surgery (E1), Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan

(Received December 17, 2006/Revised April 3, 2007/Accepted April 6, 2007/Online publication May 16, 2007)

Autoantibodies against tumor antigens represent one type of biomarker that may be assayed in serum for detection of cancer and monitoring of disease progression. In the present study, we used a proteomics-based approach to identify novel tumor antigens in non-small cell lung cancer (NSCLC). By combining two-dimensional electrophoresis, western blotting, mass spectrometry and enzyme-linked immunosorbent assay technology, we detected autoantibodies against α-enolase in a subset of NSCLC patients' sera. When 'Mean $OD_{\text{healthy control sera}}$ + 3 $SD_{\text{healthy control sera}}$' was used as the cut-off point, the prevalence of this autoantibody was 27.7% in patients with NSCLC (26 of 94), 1.7% in healthy control subjects (1 of 60), and not detectable in sera from 15 patients with small cell lung cancer, 18 patients with gastrointestinal cancer and nine patients with *Mycobacterium avium* complex infection of lung. Immunohistochemical staining showed that expression of α-enolase was increased in cancer tissues of NSCLC patients, and flow cytometric analysis confirmed the expression of α-enolase at the surface of cancer cells. The combined detection of autoantibodies against α-enolase, carcinoembryonic antigen and cytokeratin 19 fragment (CYFRA21-1) enhanced sensitivity for the diagnosis of NSCLC. Therefore, autoantibodies against α-enolase may constitute a promising biomarker for NSCLC. (*Cancer Sci* 2007; 98: 1234–1240)

Lung cancer is the leading cause of cancer death,[1] and NSCLC accounts for nearly 80% of lung cancer cases. There is an urgent need for a better understanding of the biological mechanisms of NSCLC as well as the identification of reliable biomarkers for its diagnosis and prognosis. To date, a number of NSCLC markers have been evaluated, including CEA, CYFRA 21-1, SCC antigen, CA125 and NSE.[2–8] Autoantibodies against several tumor antigens such as L-myc and c-myc, p53 and antineural/antinuclear antigens have also been investigated.[9–12] Recently, autoantibodies against PGP9.5, peroxiredoxin-I, annexin-I and annexin-II were identified in the sera of lung cancer patients using a proteomic approach.[13–15] However, the sensitivity and specificity of these biomarkers are not yet satisfactory and there are currently no data to support any particular method for screening for lung cancer.[16]

Autoantibodies against tumor antigens represent one type of biomarker that may be assayed in serum for detection of cancer and monitoring of disease progression. In spite of the fact that the quantity of any tumor antigen in cancer cells or in the circulation is usually very small, especially in the early stages of cancer, the body's immune response to such antigens represents a remarkable phenomenon of biological amplification of these weak signals from tumor antigens.[17] The identification of panels of tumor antigens that elicit an immune response may thus be useful for detecting potential specific biomarkers as well as for the initiation of immunotherapy against NSCLC. The aim of the present study was to identify novel candidate tumor antigens in NSCLC by means of a proteomics-based approach. One of these antigens was identified as α-enolase, and its immunogenicity was confirmed by western blotting using recombinant protein. The results obtained with enzyme-linked immunosorbent assay (ELISA) demonstrated that when 'Mean $OD_{\text{healthy control sera}}$ + 3 $SD_{\text{healthy control sera}}$' was used as the cut-off point, a humoral immune response directed against α-enolase occurred in 27.7% of NSCLC patients, but in only 1.7% of healthy control subjects. Immunohistochemical staining showed that α-enolase was overexpressed in cancer tissues of NSCLC patients. The combined detection of autoantibodies against α-enolase, CEA and CYFRA 21-1 enhanced sensitivity for NSCLC diagnosis. Therefore, autoantibodies against α-enolase may constitute a promising biomarker for NSCLC.

Materials and Methods

Subjects. Sera and tumor tissue were obtained at the time of diagnosis after informed consent had been given by the subjects. The experimental protocol was approved by the ethics committee of Osaka University. Sera from 94 patients with NSCLC, 15 patients with SCLC, 18 patients with gastrointestinal cancer (10 patients with gastric cancer, 8 patients with colon cancer) and nine patients with MAC were analyzed. In terms of TNM stages, the NSCLC patients comprised 17 cases of stage I, 14 cases of stage II, 34 cases of stage III and 29 cases of stage IV. The histological distribution of NSCLC was 73 adenocarcinoma cases and 21 SCC cases. Clinical data for the serum tumor marker CEA and CYFRA 21-1 were also collected for investigation. Sera from 60 asymptomatic healthy subjects, whose average age and sex were comparable to those of the NSCLC patient group, were used as controls.

1-DE and 2-DE. Proteins were extracted from NSCLC tumor tissues using the Complete Mammalian Proteome Extraction Kit (Calbiochem, Darmstadt, Germany). For 1-DE, extracted proteins

[8]To whom correspondence should be addressed. E-mail: tnaka@nibio.go.jp

Abbreviations: 1-DE, one-dimensional electrophoresis; 2-DE, two-dimensional electrophoresis; a.a., amino acids; CA, cancer antigen; CBB, Coomassie brilliant blue; CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment; ELISA, enzyme-linked immunosorbent assay; HRP, horseradish peroxidase; IEF, isoelectric focusing; IHC, immunohistochemical; LC, liquid chromotography; MAC, Mycobacterium avium complex; MALDI-TOF, matrix-assisted laser desorption–ionization time-of-flight; MS, mass spectrometry; NSCLC, non-small cell lung cancer; NSE, neuronspecific enolase; OD, optical density; PCR, polymerase chain reaction; PMF, peptide mass fingerprinting; PVDF, polyvinylidene difluoride; SCC, squamous cell carcinoma; SCLC, small cell lung cancer; TNM, tumour–node–metastasis.

doi: 10.1111/j.1349-7006.2007.00509.x

were resolved by using 10% Bis-Tris Criterion XT Precast gels (Bio-Rad Laboratories, Hercules, CA, USA), transferred to PVDF membranes or stained with CBB. For 2-DE, IEF was carried out using the PROTEAN IEF cell (Bio-Rad Laboratories) according to the manufacturer's instructions. Extracted proteins were reconstituted in a rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 2 mM tributyl phosphine (TBP), 0.0002% bromophenol blue (BPB), 0.2% bio-lyte ampholyte 3-10) and applied to ReadyStrip IPG strips (11 cm, pH 3–10). IEF was run for 45 000 Vh, and 2-DE was carried out using 10% Bis-Tris Criterion XT Precast gels. The gels were then stained with the Silver Stain MS Kit (Wako Pure Chemical Industries, Osaka, Japan) or used for protein transfer to PVDF membranes.

Western blotting. After blocking with 5% skim milk, the PVDF membranes were incubated with serum at a 1:100 dilution or rabbit anti-enolase antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA) at a 1:1500 dilution. The membranes were then incubated with sheep anti-human IgG or donkey anti-rabbit IgG (Amersham Biosciences UK, Buckinghamshire, UK). Membranes incubated with sheep anti-human IgG only were used as negative controls. Finally, the signals were visualized with an enhanced chemiluminescence reaction system (Perkin Elmer Life Sciences, Boston, MA, USA).

Identification of protein bands or spots. Protein bands on gels stained with CBB or protein spots on gels stained with silver, which corresponded to positive bands or spots on western blot membranes, were excised from the gel and digested with trypsin (Promega, Madison, WI, USA) according to published procedures.[18] For protein bands, the LC-MS/MS analysis was carried out using an LCQ ion trap mass spectrometer (ThermoElectron, San Jose, CA, USA) coupled on-line with Magic 2002 capillary high-performance liquid chromatography (Michrom BioResources, Auburn, CA, USA). For protein spots, all PMF spectra were obtained by using an ultraflex TOF/TOF MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). MS/MS or PMF data were then searched with Mascot software (Matrix Science, London, UK) against the NCBInr or swiss-prot databases. Protein database searching was carried out with following parameters for PMF: *Homo sapiens*, maximum of one missed, cleavage by trypsin, monoisotopic mass value, charge state of 1+, allowing a mass tolerance of 100 p.p.m., and carbamidomethyl modification of cysteine.

Preparation of recombinant protein. To prepare recombinant proteins, the human full-length α-enolase complementary DNA (1–434 a.a.) was amplified by PCR from the Hep3B cell line cDNA library using the primers: sense 5′-GTGGCTAGAAGTTCACCATG-3′, antisense 5′-TTACTTGGCCAAGGGGTTTC-3′. To map the autoepitope on α-enolase, three cDNA fragments that encode C-terminal deletion mutant proteins (α-Eno1, α-Eno2, α-Eno3) were similarly amplified. The nucleotide sequences of the primers for PCR were: α-Eno1 (1–334 a.a.) sense 5′-TGTCTATTCTCAAGATCCATGCC-3′, antisense 5′-TTACTCGTTCACGGCCTTGGC-3′; α-Eno2 (1–234 a.a.), sense 5′-TGTCTATTCTCAAGATCCATGCC-3′, antisense 5′-TTAAGCTTTCCCAATAGCAGTC-3′; and α-Eno3 (1–134 a.a.) sense 5′-TGTCTATTCTCAAGATCCATGCC-3′, antisense 5′-TTAGATGTGGCGGTACAGGGG-3′. These cDNA fragments were then subcloned into the pET-28a(+) vector (Novagen, Madison, WI, USA), resulting in expression of α-enolase or its fragments with a 6 × His tag. Recombinant proteins were produced in *Escherischia coli* BL21-CodonPlus (DE3)-RIL cells (Stratagene, La Jolla, CA, USA) and purified by affinity chromatography using Ni-NTA resin (QIAGEN, Tokyo, Japan). Recombinant human full-length or C-terminal deletion mutant α-enolase, rabbit β-enolase (Sigma, St Louis, MO, USA) and human γ-enolase (Calbiochem) were subjected to sodium dodecylsulfate–polyacrylamide gel electrophoresis, using a 4–20% precast gel, then stained with CBB or transferred to PVDF membrane and probed with anti-enolase antibody, anti-6 × His monoclonal antibody or sera as described above.

Flow cytometry. Human lung adenocarcinoma cell line A549 was maintained in RPMI-1640 medium supplemented with 10% heat-inactivated fetal bovine serum, 100 U/mL penicillin and 100 μg/mL streptomycin. Cells (10^6) were incubated with rabbit anti-enolase antibody at a 1:100 dilution and labeled with fluorescein isothiocyanate-conjugated goat anti-rabbit immunoglobulin (BD Biosciences, San Jose, CA, USA). Normal rabbit IgG was used as a control. Stained cells were analyzed using a FACS Canto cytometer (Becton-Dickinson, Mountain View, CA, USA) and the results were analyzed using FlowJo software (Tree Star, Stanford, CA, USA).

ELISA. To assess the potential of these autoantibodies as a diagnostic marker, their frequencies in the sera were determined by means of ELISA using recombinant human full-length α-enolase protein. The ELISA was carried out as published elsewhere, with modifications.[19] Briefly, each well of a Microtiter plate (MaxiSorp; Nunc A/S, Roskilde, Denmark) was coated with 1 μg of recombinant human full-length α-enolase. After blocking with 1% bovine serum albumin, all wells were incubated with human serum at a 1:500 dilution at room temperature for 1 h. To reduce the background level originating from the non-specific reactivity of sera with bacterial proteins, the sera were diluted and incubated with 100 μg/mL *E. coli* BL21-CodonPlus (DE3)-RIL cell lysate for 2 h at room temperature before incubation with coated recombinant human α-enolase. The antigen–antibody complexes were detected with 1:5000-diluted HRP-conjugate sheep anti-human IgG with TMB (Dako, Carpinteria, CA, USA) as the substrate. OD was read at 450 nm. The antibody titer was expressed by using arbitrary binding units calculated according to the formula:

$$\text{binding units of sample} = (OD_{\text{sample}}/[\text{Mean } OD_{\text{healthy control sera}} + 3\ SD_{\text{healthy control sera}}]) \times 100.$$

Based on this formula, 100 binding units was used as the cut-off point.

IHC staining. After deparaffinization, tissue sections were treated with 100% cold methanol containing freshly prepared 0.3% hydrogen peroxide for 30 min, blocked in 10% normal goat serum for 20 min and incubated with rabbit anti-enolase antibody (Santa Cruz Biotechnology) at a 1:250 dilution overnight. Incubation of parallel sections omitting the first antibody was done to generate negative controls. Staining of sections was completed with a biotin-conjugated secondary antibody, HRP-conjugated streptavidin and diaminobenzidine.

Statistical analysis. Significant differences between groups were assessed with the χ^2-test and Fisher's exact test. $P < 0.05$ was considered significant.

Results

Detection of autoantigens associated with NSCLC by 1-DE western blotting and LC-MS/MS. In order to screen for autoantibodies against cancer cells in patients with NSCLC, proteins extracted from a given patient's tumor tissue were subjected to 1-DE, transferred to membranes, and incubated with sera from the same patient or from healthy control subjects. Membranes incubated with only the secondary antibody were used as negative controls. An approximately 47-kDa band was recognized only by a subset of NSCLC patient sera, whereas no such reaction was observed with healthy control sera or negative controls (Fig. 1a). To identify this 47-kDa protein, the corresponding band on the gel stained with CBB was digested and analyzed using LC-MS/MS. The eight proteins, including α-enolase and elongation factor 1-α 1, which were identified by database searching through Mascot software, are listed in Table 1. Many of these proteins

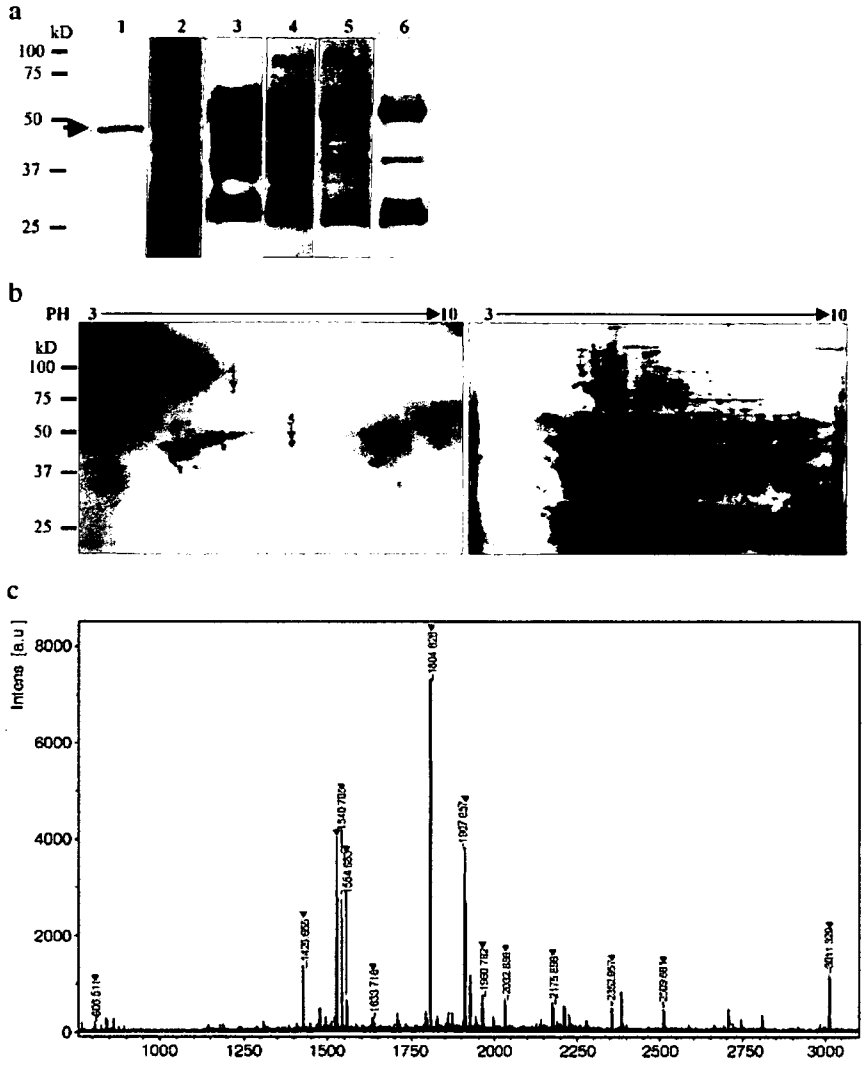

Fig. 1. (a) Screening by means of one-dimensional electrophoresis (1-DE) western blotting analysis for autoantigen associated with non-small cell lung cancer (NSCLC). Lane 1: ~47-kDa positive band (arrow), which was recognized by anti-enolase antibodies. Lanes 2, 3: 47-kDa positive band, which was recognized by NSCLC sera. Lane 4: no 47-kDa positive band was observed in a negative NSCLC case. Lane 5: no positive reaction was observed with healthy control sera. Lane 6: no positive reaction was observed in negative controls. (b) Detection by means of two-dimensional electrophoresis (2-DE) western blotting analysis of autoantigen associated with NSCLC. Left panel: Representative 2-DE western blotting analysis. Right panel: corresponding 2-DE silver-stained image. Protein spots recognized only by NSCLC sera are marked with arrows and numbers. (c) Peptide mass fingerprinting spectra of positive spot 5. For spot 5, 14 peptide masses were matched with human α-enolase by executing an NCBInr database search, yielding 52% protein sequence coverage. The matched mass peaks are marked with arrow heads.

Table 1. Mascot search results of the liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) data

Protein name	Swiss Prot Accession number	Molecular weight (Da)	pI	Score*	Peptide matched	Protein coverage (%)
Elongation factor 1-α 1	P68104	50 451	9.10	123	4	7
Cytokeratin 17	Q04695	48 230	4.97	271	13	25
α-Enolase	P06733	47 037	6.99	96	3	5
Elongation factor Tu	P49411	49 852	7.26	56	1	3
α-1-acid glycoprotein 1 precursor	P02763	23 725	4.93	160	4	18
Vimentin	P08670	53 545	5.06	115	5	9
Albumin precursor	P02768	71 317	5.92	106	4	6
Actin-like protein 3	P61158	47 797	5.61	46	2	7

*Scores > 39 indicate identity or extensive homology ($P < 0.05$). To identify the 47-kDa protein recognized only by non-small cell lung cancer patient sera, the corresponding band stained with Coomassie brilliant blue was digested and analyzed by LC-MS/MS. The eight proteins identified are listed.

are of similar molecular weight and one of them may be the autoantigen associated with NSCLC. We used western blotting with rabbit anti-enolase antibodies to confirm that the expression of α-enolase occurred at the same position as that of the 47-kDa positive band (Fig. 1a).

Autoantibodies against α-enolase present in NSCLC patient sera. To characterize autoantibodies in NSCLC sera, proteins extracted from a given patient's tumor tissue were separated by 2-DE, transferred to membranes, and incubated with sera from the same patient or from healthy control subjects. Compared with the sera of healthy control subjects, 2-DE western blotting with NSCLC patient sera showed five positive protein spots (Fig. 1b, left panel), including one spot (spot 5) that also had a molecular weight of approximately 47 kDa and a pI value of

doi: 10.1111/j.1349-7006.2007.00509.x

Table 2. Mascot search results of matrix-assisted laser desorption–ionization time-of-flight (MALDI-TOF) data

Spot no.	Protein name	Sequence coverage (%)	Molecular weight (Da)	pI
1	Chain D, myeloperoxidase	22	53 806	9.43
2	Tumor rejection antigen-1 (gp96)	17	92 696	4.76
3	Not identified			
4	α glucosidase II α subunit	15	86 236	5.71
5	α-enolase	52	47 037	6.99

To identify the immunoreactive spots in two-dimensional electrophoresis western blotting analysis recognized by non-small cell lung cancer patient sera, corresponding silver-stained spots were digested and analyzed by MALDI-TOF/mass spectrometry. 'Spot no.' corresponds to spots marked in Fig. 1b.

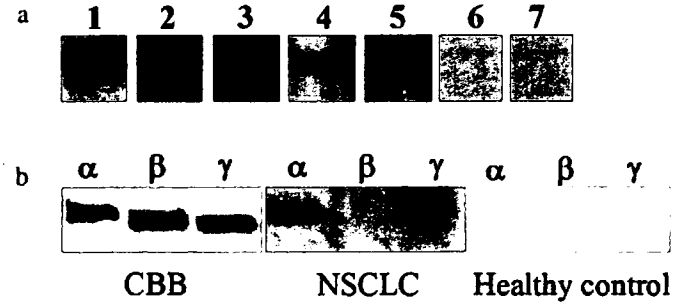

Fig. 2. (a) Western blotting analysis of recombinant human full-length α-enolase protein. Recombinant human full-length α-enolase protein was probed with rabbit anti-enolase antibodies (lane 1), with sera from non-small cell lung cancer (NSCLC) patients (lane 2–5), and with sera from healthy control subjects (lane 6–7). (b) Western blotting analysis of α, β and γ-enolase. α, β and γ-enolase protein were stained with Coomassie brilliant blue (left panel) or probed with sera from NSCLC patients (middle panel) and from healthy control subjects (right panel).

approximately 7.0. The corresponding spots on the silver-stained gel (Fig. 1b, right panel) were identified by MALDI-TOF/MS and database search. The identified proteins are summarized in Table 2. Spot 5 was recognized as α-enolase, as in the previous LC-MS/MS analysis. Its PMF spectrum is shown in Fig. 1c and the database search produced 14 peptide masses that coincided with human α-enolase, thus yielding 52% protein sequence coverage.

Next, western blotting with full-length recombinant human α-enolase protein was used to confirm and analyze the immunogenicity of α-enolase. Correct expression of the recombinant protein was verified by western blotting using rabbit anti-enolase antibodies (Fig. 2a). The recombinant proteins were then probed with sera from NSCLC patients or healthy control subjects, and positive bands were detected only in sera from the former, not from the latter (Fig. 2a). In addition, western blotting was used to determine the reactivity of NSCLC patient sera to enolase isoforms, which contain α, β and γ-enolase. The sera that were positive for autoantibodies against α-enolase reacted with neither β-enolase nor γ-enolase (Fig. 2b), whereas healthy control sera did not react with any of the enolase isoforms. This indicates the specificity of autoantibodies against α-enolase in NSCLC patient sera, and the overall results suggest that a subset of NSCLC patient sera contains autoantibodies against α-enolase.

Frequencies of autoantibodies against α-enolase in the sera were determined by means of ELISA using recombinant human full-length α-enolase protein. We tested 94 sera from patients with NSCLC, 15 from patients with SCLC, 18 from patients with gastrointestinal cancer (10 patients with gastric cancer, 8 patients with colon cancer), nine from patients with MAC infection of lung and 60 from healthy control subjects. When

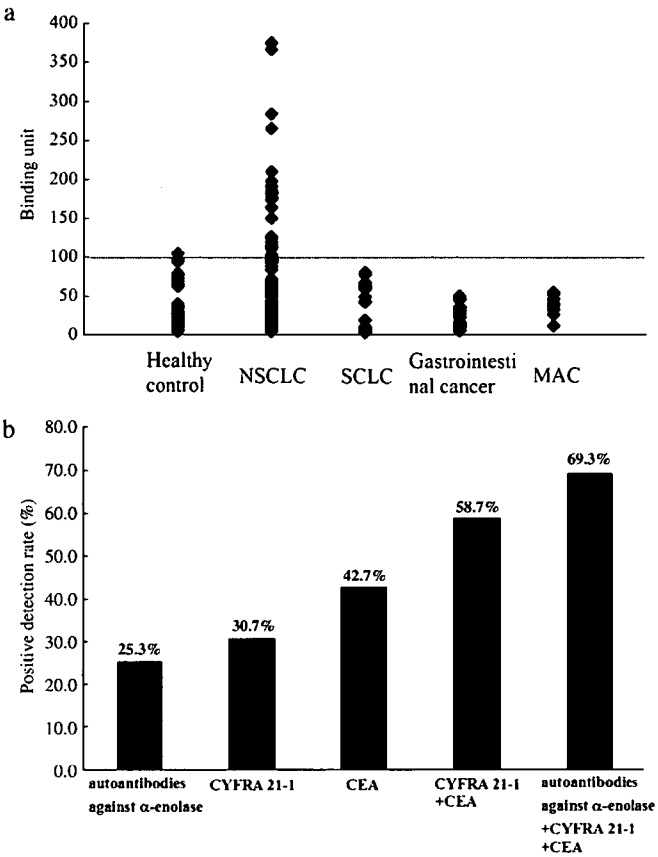

Fig. 3. (a) Prevalence of autoantibodies against α-enolase determined by enzyme-linked immunosorbent assay (ELISA). The *y*-axis denotes binding units. The solid horizontal line represents the positive cut-off limit. The prevalence of autoantibodies against α-enolase was 27.7% in patients with non-small cell lung cancer (NSCLC) (26 of 94), 1.7% in healthy control subjects (1 of 60), and not detectable in small cell lung cancer gastrointestinal cancer (10 patients with gastric cancer, 8 patients with colon cancer) and *Mycobacterium avium* complex infection of lung. (b) Positive detection rate of autoantibodies against α-enolase, cytokeratin 19 fragment (CYFRA21-1) and carcinoembryonic antigen (CEA) in NSCLC patients. Detection of CYFRA 21-1 and CEA in combination increased the positive detection rate to 58.7%. Furthermore, combined detection of autoantibodies against α-enolase, CEA and CYFRA 21-1 achieved a positive detection rate of up to 69.3%.

'Mean $OD_{\text{healthy control sera}} + 3\ SD_{\text{healthy control sera}}$' was used as the cut-off point, the prevalence of autoantibodies against α-enolase was 27.7% in patients with NSCLC (26 of 94), 1.7% in healthy control subjects (1 of 60), and not detectable in SCLC, gastrointestinal cancer or MAC infection of lung (Fig. 3a). These results

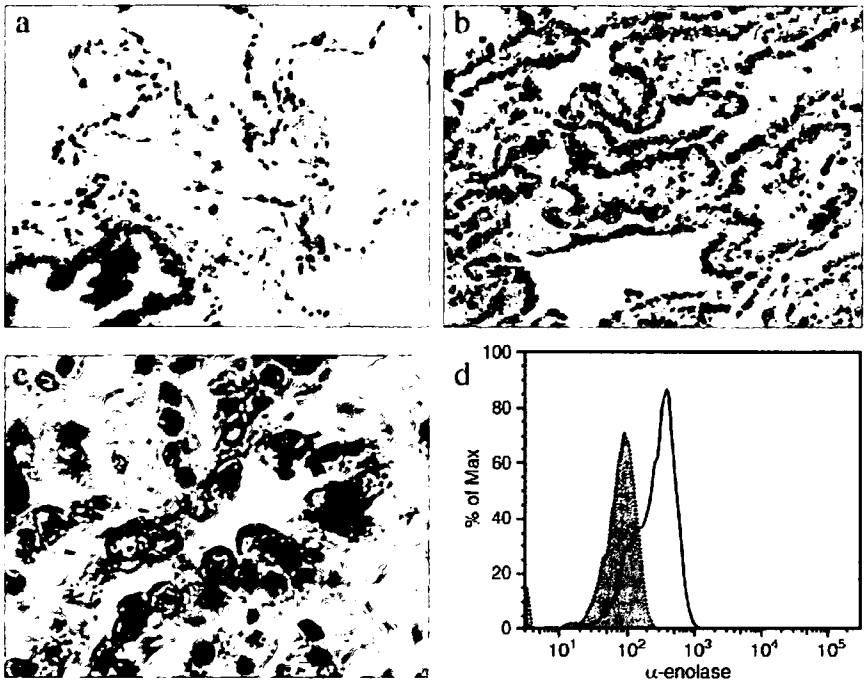

Fig. 4. Immunohistochemical and flow cytometric analysis of α-enolase. (a) Normal lung tissue (×100). (b) Lung adenocarcinoma (×100). (c) α-Enolase staining showing a mixture of cytoplasmic, and membranous immunoreactivity (×400). (d) A549 cells were stained with anti-enolase antibody, labeled with fluorescein-isothiocyanate-conjugated goat antirabbit immunoglobulin, and analyzed on a FACS Canto (open histogram). Shaded histogram indicates staining with control IgG.

showed that titers of autoantibodies against α-enolase are increased in a subset of NSCLC patients. Next, we examined the correlation between the prevalence of autoantibodies against α-enolase and clinicopathological features in NSCLC patients. Positive reactivity was detected in 21 of the 73 sera from adenocarcinoma patients (28.8%) and in five of the 21 sera from SCC patients (23.8%). There was no significant correlation between the occurrence of autoantibodies against α-enolase and pathological types ($P = 0.654$). In addition, there was a tendency for autoantibodies against α-enolase to be more prevalent in patients with advanced NSCLC cases (stage III/IV, 33.3%, 21 of 63) than in stage I/II cases (16.1%, 5 of 31), although the results of the statistical analysis suggest that the prevalence has no significant correlation with disease stage ($P = 0.08$). We also investigated the relationship between autoantibodies against α-enolase and other tumor markers (CEA, CYFRA 21-1) that have been applied to clinical practice in NSCLC patients. In a total of 94 NSCLC patients, clinical data of both CEA and CYFRA 21-1 were available for 75 patients. In these patients, 25.3% (19/75) were positive for autoantibodies against α-enolase, 42.7% (32/75) were positive for CEA, and 30.7% (23/75) were positive for CYFRA 21-1. The occurrence of autoantibodies against α-enolase didn't show a significant correlation with CEA ($P = 0.63$) or with CYFRA 21-1($P = 0.92$). Detection of CYFRA 21-1 and CEA in combination increased the positive detection rate to 58.7% (44/75). Furthermore, positive detection rate was enhanced up to 69.3% (52/75) when combined detection of autoantibodies against α-enolase, CEA and CYFRA 21-1 was used (Fig. 3b).

IHC and flow cytometric analysis of α-enolase. We used IHC staining to compare α-enolase expression in non-malignant and malignant lung tissues from 20 NSCLC patients, including 10 patients with autoantibodies against α-enolase and 10 patients without autoantibodies against α-enolase. The staining results showed that expression of α-enolase was increased in malignant lung tissue of NSCLC patients (Fig. 4a–c). Additionally, IHC staining showed not only cytoplasmic but also membranous immunoreactivity in cancer cells (Fig. 4c). Flow cytometric analysis of human lung adenocarcinoma cell line A549 also confirmed the expression of α-enolase at the surface of lung cancer cells (Fig. 4d).

Epitopes located at the N-terminal region (1–134 a.a.) of α-enolase that are recognized by autoantibodies. To locate the serological epitopes of α-enolase, full-length and C-terminal deletion mutant proteins (α-Eno1, α-Eno2, α-Eno3) were prepared. The full-length α-enolase (1–434 a.a.), α-Eno1 (1–334 a.a.), α-Eno2 (1–234 a.a.) and α-Eno3 (1–134 a.a.) recombinant proteins were clearly shown by an anti-6 × His antibody or stained with CBB, which verified their expression (Fig. 5). A commercially available rabbit anti-enolase antibody reacted only with the full-length α-enolase, α-Eno1 and α-Eno2 recombinant proteins (Fig. 5). However, sera from NSCLC patients who showed the presence of autoantibodies against α-enolase reacted with the α-Eno1, α-Eno2, and α-Eno3 recombinant proteins (Fig. 5), indicating that in NSCLC patients at least the N-terminal region of α-enolase contains epitopes.

Discussion

In the present study we used a proteomics-based screen test to identify proteins such as α-enolase and gp96 that may elicit a humoral immune response in NSCLC patients. We then confirmed that some NSCLC patients' sera contained autoantibodies against α-enolase by means of western blotting using recombinant protein. Furthermore, the results obtained with ELISA demonstrated that when 'Mean $OD_{\text{healthy control sera}} + 3\ SD_{\text{healthy control sera}}$' was used as the cut-off point, the humoral immune response directed against α-enolase occurred in 27.7% of NSCLC patients, but in only 1.7% of healthy control subjects. α-enolase is an isoenzyme of enolase, a key protein that catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate, which is the second of the two high-energy intermediates that generate ATP in glycolysis.[20] Three isoforms of enolase have been identified and are known as α, β and γ-enolase. α-enolase is present in most tissues and is predominant in early embryonic tissue, β-enolase is expressed in muscle tissue, and γ-enolase, also known as NSE, is found only in neuronal tissues.

Autoantibody responses to tumors are generally thought to be elicited in three ways. These are overexpression of specific proteins, especially on the cell surface, gene mutation or post-translational modification of proteins, which shows new epitopes

doi: 10.1111/j.1349-7006.2007.00509.x