# Protein Profile of Human Lung Squamous Carcinoma Cell Line NCI-H226[1]

HAO ZHANG[*], NA LI[*], YUE CHEN[*], LING-YUN HUANG[*], YI-CHING WANG[#], GANG FANG[#], DA-CHENG HE[*,2], AND XUE-YUAN XIAO[*,2]

[*]*Key Laboratory of Cell Proliferation and Regulation of Ministry of Education, Universities' Confederated Institute of Proteomics, Beijing Normal University, Beijing 100875, China;* [#]*Department of Life Sciences, National Taiwan Normal Universities, Sec. 4 Tingchou Rd, Taipei 116, Taiwan*

**Objective** To construct a database of human lung squamous carcinoma cell line NCI-H226 and to facilitate discovery of novel subtypes markers of lung cancer. **Method** Proteomic technique was used to analyze human lung squamous carcinoma cell line NCI-H226. The proteins of the NCI-H226 cells were separated by two-dimensional gel electrophoresis and identified by mass spectrometry. **Results** The results showed that a good reproducibility of the 2-D gel pattern was attained. The position deviation of matched spots among three 2-D gels was 1.95±0.53 mm in the isoelectric focusing direction, and 1.73±0.45 mm in the sodium dodecyl sulfate-polyacrylamide gel electrophoresis direction. One hundred and twenty-seven proteins, including enzymes, signal transduction proteins, structure proteins, transport proteins, *etc.* were characterized, of which, 29 identified proteins in NCI-H226 cells were reported for the first time to be involved in lung cancer carcinogenesis. **Conclusion** The information obtained from this study could provide some valuable clues for further study on the carcinogenetic mechanism of different types of lung cancer, and may help us to discover some potential subtype-specific biomarkers of lung cancer.

**Key words**: Lung squamous carcinoma; NCI-H226 cell line; Proteomics; Two-dimensional datapase

## INTRODUCTION

Lung cancer is the leading cause of cancer deaths in both men and women in Canada, the United States, and China. Lung cancer is usually divided into two groups accounting for about 90% of all cases. These two types are called small cell lung cancer and non-small cell lung cancer including squamous carcinoma, adenocarcinoma and large-cell carcinoma. Squamous carcinoma is the second most common type, constituting about 30% of all types of lung cancer[1]. Therefore, it is extremely important for us to understand the pathogenesis and development of squamous carcinoma and thus to further identify potential subtype markers of lung cancer by exploring its protein profile. Although some insights into the proteome of lung cancer have been achieved, additional information about protein profile of this type of cancer will be critical for understanding its carcinogenesis and development, and is also helpful in making differential diagnosis of subtypes of lung cancer[2]. More recently, different proteomic analysis of lung adenocarcinoma and normal lung tissues was performed using 2D-PAGE and MALDI-TOF-MS. Nine proteins have been found to have significant over-expression in lung adenocarcinoma. Among them, UCHL1 (PGP9.5) was identified to up-regulate at the same time in A549 lung adenocarcinoma cell line by using 2D Western blot analysis[3]. UCHL1 belongs to the ubiquitin carbosyl-terminal hydrolase family that is a part of the cellular proteolytic pathway regulating many cellular processes, including cell cycle progression and cell death. UCHL1 is widely expressed in neuronal tissues at all stages of differentiation and is highly expressed in lung cancer tissue and could be a candidate marker for non-small-cell lung cancer[3]. Some previous work suggested that cell lines in outline might retain a partial property of the original tumor tissue and might act as a relatively simple and useful model for tumor

[1]This work was supported by the National Natural Science Foundation of China (Grant No.30370712) and Beijing Key Project (Grant No. 7051002).

[2]Correspondence: Dr. Xue-Yuan XIAO and Dr. Da-Cheng HE, Beijing Normal University, Key Laboratory of Cell Proliferation and Regulation of Ministry of Education, 19th Xinjiekouwai St. 100875, Beijing China. Fax: 8610-58805042. E-mail: xyxiao@bnu.edu.cn or dhe@bnu.edu.cn

Biographical note of the first author: Hao ZHANG, male, born in 1981, Ph. D candidate, majoring in cell biology and proteomics.

0895-3988/2007
CN 11-2816/Q

proteomic analysis. Recently, a 2D database for lung squamous carcinoma tissue containing thirty-four proteins has been constructed and is publicly available[4], but a more comprehensive profile of this type of lung cancer has not been seen yet.

In this study, a well-documented human lung squamous carcinoma cell line NCI-H226 was selected, and its protein profile was analyzed using 2-DE. The proteins of NCI-H226 cells were separated by two-dimensional gel electrophoresis and identified by MALDI-MS and MS/MS analysis following in-gel digestion with trypsin. The reference map for NCI-H226 cells was constructed and contained one hundred and twenty-seven proteins. The results presented here should be able to facilitate further studies on the carcinogenetic mechanism of lung squamous carcinoma and help us to discover potential subtype-specific biomarkers, and detect novel drug targets in pharmacotherapy for lung squamous carcinoma.

## MATERIALS AND METHODS

### *Materials*

Immobilized pH-gradient (IPG) strips and IPG buffers were purchased from Amersham Pharmacia Biotechnology (Uppsala, Sweden). Acrylamide, CHAPS, urea and DTT were obtained from Sigma (St, Louis, Mo, USA). The other reagents for the polyacrylamide gel preparation were purchased from Bio-Rad Labs (Hercules, CA, USA). IPG strips were frozen at -20℃ and the other reagents were kept at 4℃.

### *Sample Preparation*

NCI-H226 cell line (gift of Dr. Yi-Qing Wang and Dr. Gang Fang, National Taiwan Normal Universities, Taipei, Taiwan) was grown in RPMI 1640 supplemented with 10% bovine serum at 37℃ and 5% $CO_2$. The total cell extracts were prepared by scraping cells on ice in 500 μL of lysis buffer containing 8 mol/L urea, 4% CHAPS, 40 mmol/L Tris-base, 65 mmol/L DTT, 1% Pharmalyte (pH 3-10), 5 mmol/L PMSF, and 5 mmol/L $Na_3VO_4$, 1 mg/mL Aprotinin, and 1 mg/mL Leupeptin. After 1 hour at 4℃, the extracts were centrifuged at 1 3000 rpm for 1 hour at 4℃. The supernatant was collected and dispensed, then stored in 100 μL aliquots at -80℃. Protein concentrations were determined using the Coomassie blue method[5].

### *Two-dimensional Gel Electrophoresis*

Two dimensional gel electrophoresis were performed as reported[6]. IEF was performed using IPG strips (pH 3-10 L, 18mm×3mm×0.5mm) on IPG phorisoelectric focusing cell (Amersham Biosciences). After IEF, the IPG strips were immediately equilibrated for 2×15 minutes with gentle shaking in 10 mL equilibration solutions. The second dimensional separation was performed in 12% SDS polyacrylamide gels. The gels (180mm×200mm× 1.5 mm) were run at 40 mA per gel, in an Ettan DALT six apparatus (Amersham Biosciences). After 2-DE, the gels were stained with Coomassie brilliant blue R350 (Amersham Biosciences) according to the manufacturer's instructions. Molecular masses were determined by running standard protein markers and pI values were used as given by the supplier of the IPG strips. Electronic images of the gels were obtained using the Magicscan densitometer (resolution: 300 dpi). The images were stored as both TIFF and JPEG formats.

### *MALDI-MS and MS/MS Analyses*

MS and MS/MS analyses were performed using Qstar Pulser I Quadrupole time-flight mass spectrometer (Applied Biosciences/MDS Sciex Toronto, Canada). Spots were excised from the gels by sucking with pipette tips and placed into 0.5 mL siliconized-tubes. One protein-free gel piece was treated in parallel as negative control. Each spot was destained with 100 μL of 50% acetonitrile in 25 mmol/L $NH_4HCO_3$ and dried in a vacuum centrifuge. The dried gel pieces were incubated in the digestion solution containing 25 mmol/L $NH_4HCO_3$ and 0.1g/L trypsin (Progema, Mannheim, Germany) for 12 h at 37℃. The tryptic peptide mixture was extracted and purified with Millipore ZIPTIP C18 column. The purified tryptic peptide mixture was mixed with matrix, consisting of a saturated solution of 2, 5-dithydroxybenzoic acid in 50% acetonitrile and 0.5% trifluoroacetie. A volume (1.5 μL) of the mixture containing 2, 5-dithydroxybenzoic acid matrix was loaded on a stainless steel plate and air-dried. An accelerating voltage of 20 kv was used. The acquired mass spectra were corrected using two trypsin autodigestion fragments (Mr 842.51 and 2211.10 Da) as internal calibrants. MS and MS/MS data were used to identify each protein by searching the SWISS-PROT database with the MASCOT search engine (http://www.matrixscience.com). The searching parameters were set up as follows: the mass tolerance was ±0.3Da, the number of missed cleavage site was allowed up to 1, modifications were carbamidomethylation (fixed) and oxidation of methionine (variable), the peptide ion was $[M+H]^+$, and the monoisotopic masses were used.

## RESULTS

### *Two Dimensional Electrophoretic Analysis*

The proteins of the NCI-H226 cells were analyzed by 2-D electrophoresis. The 2-D electrophoretic separation was performed on pH 3-10 linear IPG strips and the spots were visualized after having been stained with Coomassie blue. Fig.1 shows a representative analysis of NCI-H226 proteins on a pH 3-10 IPG gel. On each gel 1.0 mg of total proteins was applied. A high-resolution and reproducibility of 2-D gel pattern of NCI-H226 cells was attained (Fig. 1). The position deviation of matched spots among three 2-D gels was 1.95±0.53 mm in the isoelectric focusing direction, and 1.73±0.45 mm in the sodium dodecyl sulfate-polyacrylamide gel electrophoresis direction. A total of ca. 900 protein spots were detected in the 2-D map (Fig. 1). The most abundant proteins were tubulin chains (P07437), glyceraldehyde 3-phosphate dehydrogenase (P04406), alpha enalose (P06733) and nucleophosmin (P06748), while the less abundant ones were mainly enzymes, and structural and heat shock proteins.

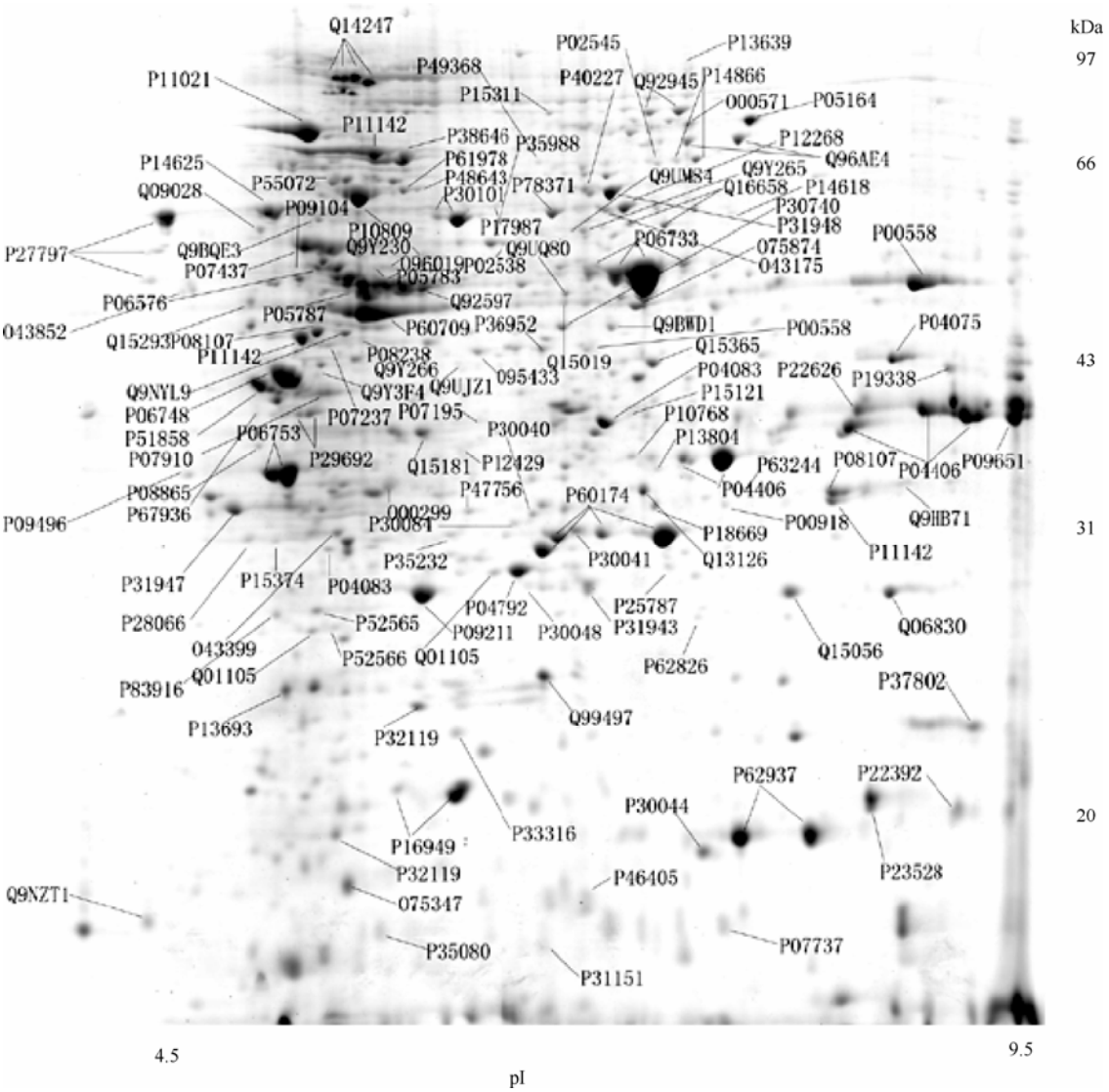


FIG. 1. Two dimensional electrophoretic analysis of NCI_H226 cell line total proteins. The proteins were separated on a pH 3-10 linear IPG strip, followed by a 12% SDS-polyacrylamide gel. The gel was stained with Coomassie blue. The spots were analyzed by MALDI-MS and MS/MS. The proteins characterized are designated with the accession number of the SWISSPROT database. The identities assigned are listed in Tables 1 and 2.

*Protein Identification*

Proteins were identified by MALDI-MS and MS/MS, following in-gel digestion with trypsin. About 500 spots, excised from three gels, were analyzed. The peptide masses were matched with the theoretical peptide masses of human proteins or with all known proteins from all species, if not successful. Figure 2 contains a representative MALDI- MS and MS/MS analysis for spot P06753 (tropomyosin alpha 3 chain). One hundred and twenty-seven proteins were identified and 98 identified proteins according to their known function are listed in Table 1 and the remaining 29 proteins which were reported for the first time to be involved in human squamous cell lung cancer and even in lung cancer are listed in Table 2. The more detailed information about these identified proteins could be obtained from authors.

TABLE 1

Classificaiton of Identified Proteins According to Their Known Functions

| SWISS-PROT Accession Number | Protein Name | SWISS-PROT Accession Number | Protein Name |
|---|---|---|---|
| Cell Structure and Mobility Proteins | | P06576 | ATP Synthase beta Chain, Mitochondrial Precursor |
| O96019 | Actin-like Protein 6A | P06733 | alpha Enolase |
| P02538 | Keratin, Type II Cytoskeletal 6A | P07195 | L-lactate Dehydrogenase B Chain |
| P05783 | Keratin, Type I Cytoskeletal 18 | P09104 | Gamma Enolase |
| P05787 | Keratin, Type II Cytoskeletal 8 | P10768 | Esterase D |
| P06753 | Tropomyosin alpha 3 Chain | P12268 | Inosine-5'-monophosphate Dehydrogenase 2 |
| P07437 | Tubulin beta-2 Chain | P14618 | Pyruvate Kinase, Isozymes M1/M2 |
| P07737 | Profilin-1 | P18669 | Phosphoglycerate Mutase 1 |
| P15311 | Ezrin | P22392 | Nucleoside Diphosphate Kinase B |
| P23528 | Cofilin, Non-muscle Isoform | P30084 | Enoyl-CoA hydratase, Mitochondrial Precursor |
| P35080 | Profilin-2 | P31939 | Bifunctional Purine Biosynthesis Protein PURH |
| P35232 | Prohibitin | P33316 | dUTP Pyrophosphatase |
| P37802 | Transgelin-2 | P60174 | Triosephosphate Isomerase |
| P47756 | F-actin Capping Protein beta-subunit | Q15181 | Inorganic Pyrophosphatase |
| P60709 | Actin, Cytoplasmic 1 | Q9BWD1 | Acetyl-CoA Acetyltransferase, Cytosolic |
| P67936 | Tropomyosin alpha 4 Chain | Chaperones | |
| Q16658 | Fascin | P04792 | Heat shock 27 kDa Protein |
| Q9BQE3 | Tubulin alpha-6 Chain | P07237 | Protein Disulfide-isomerase Precursor |
| Transcription and Translation | | P08107 | Heat Shock 70 kDa Protein 1 |
| P07910 | Heterogeneous Nuclear Ribonucleoproteins C1/C2 | P08238 | Heat Shock Protein HSP 90-beta |
| P08865 | 40S Ribosomal Protein SA | P10809 | 60 kDa Heat Shock Protein, Mitochondrial Precursor |
| P13639 | Elongation Factor 2 | P11021 | 78 kDa Glucose-regulated Protein Precursor |
| P14866 | Heterogeneous Nuclear Ribonucleoprotein L | P11142 | Heat Shock Cognate 71 kDa Protein |
| P19338 | Nucleolin | P14625 | Endoplasmin Precursor |
| P22626 | Heterogeneous Nuclear Ribonucleoproteins A2/B1 | P15374 | Ubiquitin Carboxyl-terminal Hydrolase Isozyme L3 |
| P29692 | Elongation Factor 1-delta | P17987 | T-complex Protein 1, alpha Subuni |
| P31943 | Heterogeneous Nuclear Ribonucleoprotein H | P25787 | Proteasome Subunit alpha Type |
| P55072 | Transitional Endoplasmic Reticulum ATPase | P28066 | Proteasome Subunit alpha Type 5 |
| P61978 | Heterogeneous Nuclear Ribonucleoprotein K | P30101 | Protein Disulfide-isomerase A3 Precursor |
| P63241 | Eukaryotic Translation Initiation Factor 5A | P35998 | 26S Protease Regulatory Subunit 7 |
| P83916 | Chromobox Protein Homolog 1 | P38646 | Stress-70 Protein, Mitochondrial Precursor |
| Q15056 | Eukaryotic Translation Initiation Factor 4H | P40227 | T-complex Protein 1, Zeta Subunit |
| Q9UQ80 | Proliferation-associated Protein 2G4 | P48643 | T-complex Protein 1, Epsilon Subunit |

(To be continued)

(Continued)

| SWISS-PROT Accession Number | Protein Name | SWISS-PROT Accession Number | Protein Name |
|---|---|---|---|
| Cellular Signal Proteins | | P49368 | T-complex Protein 1, Gamma Subunit |
| O00299 | Chloride Intracellular Channel Protein 1 | P62937 | Peptidyl-prolyl Cis-trans Isomerase A |
| P04083 | Annexin A1 | P78371 | T-complex Protein 1, beta Subunit |
| P09651 | Heterogeneous Nuclear Ribonucleoprotein A1 | Antioxidant proteins | |
| P12429 | Annexin A3 | P09211 | Glutathione S-transferase P |
| P30740 | Leukocyte Elastase Inhibitor | P13804 | Electron Transfer Flavoprotein alpha-subunit, Mitochondrial Precursor |
| P31947 | 14-3-3 Protein Sigma | P30041 | Peroxiredoxin 6 |
| P36952 | Maspin Precursor | P30044 | Peroxiredoxin 5, Mitochondrial Precursor |
| P51858 | Hepatoma-derived Growth Factor | P30048 | Thioredoxin-dependent Peroxide Reductase, Mitochondrial Precursor |
| P52565 | Rho GDP-dissociation Inhibitor 1 | P32119 | Peroxiredoxin 2 |
| P52566 | Rho GDP-dissociation Inhibitor 2 | Q06830 | Peroxiredoxin 1 |
| P63244 | Guanine Nucleotide-binding Protein beta-subunit 2-like 1 | Q99497 | DJ-1 Protein |
| Q15019 | Septin 2 | Others | |
| Intermediary Metabolism | | O43852 | Calumenin Precursor |
| O43175 | D-3-phosphoglycerate Dehydrogenase | P09496 | Clathrin Light Chain A |
| O75874 | Isocitrate Dehydrogenase | P13693 | Translationally Controlled Tumor Protein |
| P04075 | Fructose-bisphosphate Aldolase A | P27797 | Calreticulin Precursor |
| P04406 | Glyceraldehyde-3-phosphate Dehydrogenase | P30040 | Endoplasmic Reticulum Protein ERp29 precursor |
| P05164 | Myeloperoxidase Precursor | Q09028 | Chromatin Assembly Factor 1 Subunit C |
| | | Q15365 | Poly(rC)-binding Protein 1 |

TABLE 2

Identified Proteins in NCI-H26 Cell Line

| SWISS-PROT Accession Number | Protein Name | SWISS-PROT Accession Number | Protein Name |
|---|---|---|---|
| O00571 | DEAD-box Protein 3 | Q13126 | S-methyl-5-thioadenosine Phosphorylase |
| O43399 | Tumor Protein D54 | Q14247 | Src Substrate Cortactin |
| O75347 | Tubulin-specific Chaperone A | Q15293 | Reticulocalbin 1 Precursor |
| O95433 | Activator of 90 kDa Heat Shock Protein ATPase Homolog | Q92597 | NDRG1 Protein |
| P00558 | Phosphoglycerate Kinase 1 | Q92945 | Far Upstream Element Binding Protein 2 |
| P00918 | Carbonic Anhydrase II | Q96AE4 | Far Upstream Element Binding Protein 1 |
| P02545 | Lamin A/C | Q9HB71 | Calcyclin-binding Protein |
| P06748 | Nucleophosmin | Q9NYL9 | Ubiquitous Tropomodulin |
| P15121 | Aldose Reductase | Q9NZT1 | Calmodulin-like Protein 5 |
| P16949 | Stathmin | Q9UMS4 | PRP19/PSO4 Homolog |
| P31151 | S100 Calcium-binding Protein A7 | Q9Y230 | 48-kDa TATA Box-binding Protein-interacting Protein |
| P31948 | Stress-induced-phosphoprotein 1 | Q9Y265 | RuvB-like 1 |
| P46405 | 40S Ribosomal Protein S12 | Q9Y266 | Nuclear Migration Protein NudC |
| P62826 | GTP-binding Nuclear Protein Ran | Q9Y3F4 | Serine-threonine Kinase Receptor-associated Protein |
| Q01105 | SET Protein | | |

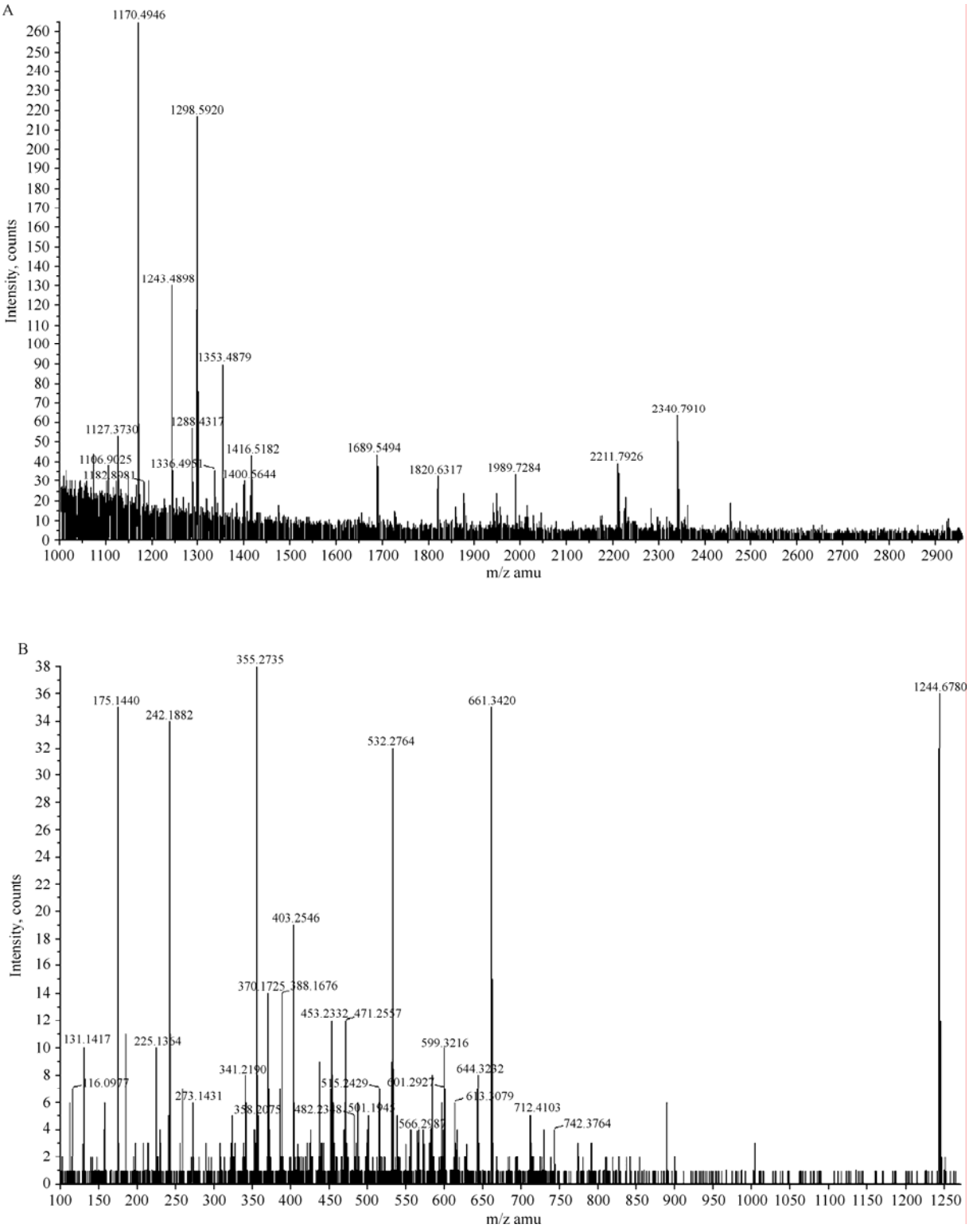


FIG. 2. MALDI-MS spectrum of spot P06753 (A) and MS/MS mass spectrum of a tryptic peptide of spot P06753 (B).

Identity could be assigned to about 70% of the analyzed spots. The number of matched peptides ranged from 4 to 20. The matched ratio (the number of matched peptides to the number of input peptides) ranged from 0.25 to 1.0. The sequence coverage of the PMF data ranged from 11% to 71%. Most (95%) coverage was $\geqslant$15%, and 86% of the coverage was $\geqslant$20%. On average, the matches had 10 matches covering 35% of the sequence, and were much higher than a minimum of four peptides, which are statistically necessary to qualify as a confident match[7]. For about 50% of the unidentified spots, good MS was collected, but no identity could be assigned. This might be due to spots overlapping, or to multiple post-translational modifications of the proteins. For the remaining 50% of the unidentified spots most of which had low intensity, insufficient peptides for protein identification were found.

About 87% of the identified proteins had masses between 20 kD and 60 kD. No protein smaller than 10 kD was identified. In general, low and high molecular mass proteins were underrepresented in Table 1. For large proteins, this might be due to limitations of the IPG strips (large proteins entering the strips with low efficiency). For small proteins, this might be due to limitations of staining methods (small proteins absorb colored substances less efficiently) and the detection limit of the gels (the lower mass detection limit of the gels was about 8 kD). The theoretical pI values of the proteins characterized varied between 4 and 9. No protein with pI value below 4 was detected, most likely due to detection limit of the IPG strips (the lower pI detection limit was about 3.5).

Many identified proteins showing heterogeneity were represented by more than one spot, most of which were structural proteins and high abundance enzymes. The multiple spots might partially result from consequence of different splicing, processing and post-translational modification, which would affect a protein's electrophoretic migration during IEF and SDS-PAGE. Heterogeneity might also result from artifacts of the 2-D electrophoresis. For most of the observed heterogeneities, neither the origin nor the biological significance was known.

### *Subcellular Location and Protein Function*

The subcellular location of the characterized proteins was analyzed using annotations in the SWISS-PRORT database. About 26% of the identified proteins had no location assigned, and approximately 43% of the proteins were cytoplasmic. The other proteins were classified into several groups, including nuclear, mitochondrial, endoplasmic reticulum, extracellular proteins, *etc*. (Fig. 3).

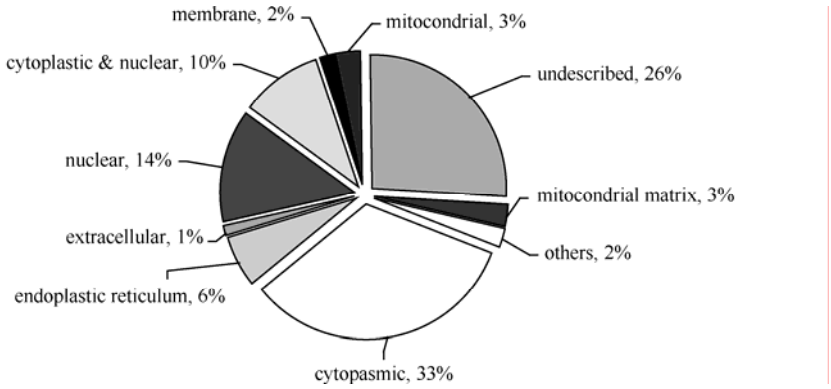


FIG. 3. Locations of the NCI-H226 cell line proteins. The proteins identified in the study were classified into location groups according to the location description in the SWISS-PROT.

## DISCUSSION

Proteomic technologies are used to separate and identify proteins. Up to now, no single technology can achieve these purposes. To get the best results, alternative technologies should be selected and in many cases they should be used in combination. In our previous works, SELDI-TOF combined with bioinformatics was used to analyze the serum samples from several hundreds of lung cancer patients. Fifteen biomarkers which could discriminate the lung cancer patients from normal controls were screened out and based on them, a classification tree was constructed[8]. The SELDI technology can be used with much less amounts of sample, compared to the other proteomic technologies. Moreover, it almost

naturally results in an integrating and quantitative biomarker pattern instead of separating pieces of data. However, it is very difficult, if not impossible, to identify the candidate biomarkers directly when the differential proteins are discovered on chips.

In the present study, a squamous carcinoma cell line, NCI-H226, was analyzed by 2-DE to obtain the protein profile. One hundred and twenty-seven proteins were identified by MALDI-MS and MS/MS analysis (Fig. 1). Such a profile could be later used to compare with a normal lung cell protein profile and may provide some information about cancer markers, and it also provides some function-related data. Most of the identified proteins were classified as metabolism-related, signaling, cytoskeleton, transport, chaperone, and nucleic acid binding proteins. Up to now, this kind of protein profiles is generated from a lot of individual laboratories respectively, and with time progressing it tends to form a comprehensive profile *via* comparison to each other, thus greatly facilitating latter works.

Li *et al.*[4] used a similar 2-DE and MALDI-TOF-MS technique to compare the protein profiles of human lung squamous carcinoma tissue and paired surrounding normal bronchial epithelial tissue, and found that 27 proteins were over-expressed in cancer tissues. In comparison of our data with Li's report, it is noticed that three proteins, fascin, nucleophosmin, and DJ-1 which were annotated to be over-expressed in lung cancer tissues[9-11] were also detected in our tested NCI-H226 cells. It has been reported, for example, that fascin (Q 16658), an actin binding protein, is marginally with strong immunostaining in lung squamous cell carcinoma and upregulated in invasive and aggressive NSCLC[15]. DJ-1 (Q99497) is a novel oncogene in cooperation with activated H-ras oncogene in conferring tumorigenic properties on normal cells in culture and over-expressed in lung cancer[9-10]. While as nucleophosmin (P06748) could reportedly contribute to p53 inactivation and tumor progression[11]. Some other proteins in Table 1 which may also correlate to various carcinomas include Prohibitin (P35232), Ezrin (P15311), and hepatoma-derived growth factor (P51858). Putting together, some proteins found in cancer cell line but not detectable in corresponding types of cancer tissues may be due to the following reasons. First, these proteins are less abundant proteins and enriched in the homogeneous cell line. Second, cell lines when cultured *in vitro* usually could not represent exact situation as they grow in the tissue. In fact, they are more or less de-differentiated and more active in proliferation. So cell lines may express or up-regulate some proteins which are not detectable *in vivo* tissues by 2-DE analysis. Third, for proteomic study, the differences among individuals could be significant when the amount of cases is small. Some proteins found in cell lines cannot be detected in some individuals. Nevertheless, the information obtained from this study could still provide some valuable clues for further study on the carcinogenetic mechanism of different types of lung cancer, and also might help us to discover some potential subtype-specific biomarkers of lung cancers. In addition, it is worthy to notice that 29 proteins listed in Table 2 have never been reported to relate with lung cancer and were first detected in NCI-H226 cells. Among them, RUVBL1 gene, located at 3q21, a region with frequent rearrangements in different types of leukemia and solid tumors, is involved in repair and/or transcription and may be essential for activation of transcription of certain essential genes[3], although the precise mechanism underlying remains to be determined. The other four proteins including DEAD-boxing protein 3, LaminA/C, Nucleophosmin and 40S ribosomal protein S12 are up-regulated in proliferative human colonic intestinal epithelial cells when compared with the same type of differentiated cells. The results indicate that these four proteins are the positive factors involved in cell proliferation[12] and may be related with carcinogenesis.

The detection of hydrophobic and less abundant proteins has been a common problem for 2D analysis. Among the one hundred and twenty-seven identified proteins, the membrane proteins and less abundant proteins are underrepresented in our list. Using liquid chromatography / tandem spectrometric (LC/MS/MS) techniques, a great number of membrane proteins can be identified. An increasing number of less abundant proteins through enriching steps by chromatography and electrophoretic techniques can be detected[13-15]. Besides, using SELDI technology and special kinds of surface-modificated chips, the hydrophobic or high-charged differential proteins could be detected effectively.

In summary, we have constructed a 2D database for the NCI-H226 cell line. The database consists of 127 proteins, obtained from MALDI-MS and MS/MS analysis of about 500 spots which were taken from three different 2D gels. This database is one of the largest 2D databases for lung cancer proteomes. Further analysis of the proteins, which are only detectable in NCI-H226 cells and have no known functions, may contribute to a better understanding of the development and pathogenesis of lung cancers.

## REFERENCES

1. Travis W D, Travis L B, Devesa S S (1995). Lung cancer. *Cancer* **75**, 191-102.
2. Oh J M, Brichory F, Puravs E, *et al.* (2001). A database of protein expression in lung cancer. *Proteomics* **1**, 1303-1319.

3. Chen G., Gharib T G, Huang C C, *et al.* (2002). Proteomic Analysis of Lung Adenocarcinoma: Identification a Highly Expressed Set of proteins in Tumor. *Clinical Cancer Research* **8**, 2298-2305.
4. Li C, Chen Z, Xiao Z, *et al.* (2003). Comparative proteomics analysis of human lung squamous carcinoma. *Biochemical and Biophysical Research Communications* **309**, 253-260.
5. Bradford M M (1976). A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-754.
6. Görg A, Obermaier C, Boguth G, *et al.* (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**,1037-1053.
7. Pappin DJC, Hojrup P and Bleasby A J (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* **3**, 327-332.
8. Xiao X Y, Liu D H, Tang Y, *et al.* (2004). Development of proteomic patterns for detecting lung cancer. *Dis Markers* **19**(1), 33-39.
9. Pelosi G, Pastorino U, Pasini F, *et al.* (2003). Independent prognostic value of fascin immunoreactivity in stage I nonsmall cell lung cancer. *Br J Cancer* **88**(4), 537-547.
10. Maiguel D A, Jones L, Chakravarty D, *et al.* (2004). Nucleophosmin Sets a Threshold for p53 Response to UV Radiation. *Molecular and Cellular Biology* **24**(9), 3703-3711.
11. Sekito A, Taira T, Niki T, *et al.* (2005). Stimulation of transforming activity of DJ-1 by Abstrakt, a DJ-1 binding protein. *Int J Oncol* **26**(3), 685-689.
12. Turck N, Richert S, Gendry P, *et al.* (2004). Proteomic analysis of nuclear proteins from proliferative and differentiated human colonic intestinal epithelial cells. *Proteomics* **4**, 93-105
13. Fountoulakis M, Takács MF, Takács B (1999). Enrichment of low-copy-number gene products by hydrophobic interaction chromatography. *J Chromatogr A* **833**, 157-168.
14. Righetti P G, Castagna A, Herbert B (2001). Prefractionation techniques in proteome analysis. *Anal Chem* **73**, 320A-326A.
15. Righetti P G, Castagna A, Herbert B, *et al.* (2003). Prefractionation techniques in proteome analysis. *Proteomics* **3**, 1397-1407.

(*Received January* 2, 2006 *Accepted August* 10, 2006)