Proteomic Analysis of Lung Adenocarcinoma: Identification of a Highly Expressed Set of Proteins in Tumors¹

Guoan Chen, Tarek G. Gharib, Chiang-Ching Huang, Dafydd G. Thomas, Kerby A. Shedden, Jeremy M. G. Taylor, Sharon L. R. Kardia, David E. Misek, Thomas J. Giordano, Mark D. Iannettoni, Mark B. Orringer, Samir M. Hanash, and David G. Beer²


ABSTRACT

Purpose: The goal of this study was to identify potential protein markers in lung adenocarcinomas.

Experimental Design: A series of 93 lung adenocarcinomas (64 stage I and 29 stage III) and 10 uninvolved lung samples were examined for quantitative differences in protein expression using two-dimensional PAGE. Candidate proteins were identified using matrix-assisted laser desorption/ionization mass spectrometry or peptide sequencing. The levels of the individual isoforms of nine proteins found to be overexpressed in the lung tumors were examined. Potential mechanisms for overexpression were examined by comparing mRNA expression levels, assessed using oligonucleotide arrays, to the protein values in the same samples.

Results: Antioxidant enzyme AOE372, ATP synthase subunit d (ATP5D), β1,4-galactosyltransferase, cytosolic inorganic pyrophosphatase, glucose-regulated Mₘ 58,000 protein, glutathione-S-transferase Mₘ 58,000 protein, glutathione-S-transferase Mₘ 58,000 protein, ATP synthase subunit d, triosephosphate isomerase, and ubiquitin thiolesterase (UCHL1) were identified as being significantly overexpressed in lung adenocarcinomas. The expression of these proteins was increased from 1.4- to 10.6-fold as compared with uninvolved lung tissue. The expression of the individual protein isoforms was correlated with 10 clinicopathological variables as well as with each gene’s mRNA level in the same sample. Both isoforms of glucose-regulated Mₘ 58,000 protein were found to be significantly correlated with their mRNA expression profiles (P < 0.05), indicating that increased transcription likely underlies the increased expression of these proteins.

Conclusions: Two-dimensional PAGE and mass spectrometry can identify proteins showing increased expression in lung adenocarcinoma. The association of specific isoforms of these proteins with clinical variables and understanding the regulation of their expression will aid in determination of their potential use as biomarkers in this cancer.

INTRODUCTION

Adenocarcinomas constitute a biologically heterogeneous group of lung tumors and are now the most common type of lung cancer (1). Although many insights into the molecular pathology of lung tumors have been achieved, additional information is critical to both our understanding of the development and progression of these tumors as well as to aid in early diagnosis. The analysis of genes overexpressed in lung cancer, and that they may serve as tumor markers, has been the subject of extensive research. The most commonly evaluated markers include neuron-specific enolase, carcinoembryonic antigen, cytokeratin 19 fragments (CYFRA 21-1), squamous cell carcinoma antigen, cancer antigen CA 125, and tissue polypeptide antigen (2). Although the analysis of multiple biological markers may be more informative than the use of a single marker (3), very few markers have been accepted for routine clinical use, either because of conflicting reports or because associations are insufficient for formulating clinical treatment plans (4). The detection of new candidate markers is complex because of the known heterogeneity of lung cancers.

2D-PAGE is a powerful research technique, which makes it possible to simultaneously examine hundreds of polypeptides in a tissue sample. It has been widely used for the detection and identification of potential tumor markers (5). This study analyzed 93 lung adenocarcinomas and 10 uninvolved lung samples for protein expression using 2D-PAGE. Analysis software were used to obtain quantitative measures for individual protein spots. Proteins of interest were identified using MALDI-MS or peptide sequencing. Associations between the proteins that were overexpressed in the lung adenocarcinomas and clinicopathological features of the tumor were determined. Evaluation of the same tumor samples for mRNA expression using oligonucleotide
arrays was used to examine the mechanisms underlying the expression profiles identified by proteomic analysis.

**MATERIALS AND METHODS**

**Tissue Specimens and Preparation.** All lung tumors and adjacent normal lung tissue were obtained at the time of surgery at the University of Michigan Hospital from May 1991 to July 2000. Consent was received from all patients, and the project was approved by the Institutional Review Board. Patients’ medical records were reviewed, and patient identifiers coded to protect confidentiality. The samples were transported to the laboratory in DMEM (Life Technologies, Inc., Gaithersburg, MD) on ice. Sixty-four stage I lung adenocarcinomas, 29 stage III lung adenocarcinomas, and 10 uninvolved lung tissue samples were examined. A portion of each sample was embedded in OCT (Miles Scientific, Naperville, IL), frozen in isopentane, cooled with liquid nitrogen for cryostat sectioning, and stored at −80°C. Hematoxylin-stained cryostat sections (5 μm), prepared from tumor pieces to be used for protein or mRNA isolation, were evaluated by a study pathologist (T. J. G.), as well as compared with H&E-stained sections made from paraffin blocks of the same tumors. The same selected region of the tumor was used for protein and mRNA isolation. Specimens were excluded if there was: (a) unclear or mixed histology (e.g., adenosquamous); (b) tumor cellularity <70%; (c) potential metastatic origin as indicated by previous tumor history; (d) extensive lymphocytic infiltration or fibrosis; or (e) the patient had experienced chemotheray or radiotherapy. Tumors were histopathologically divided into two categories: bronchial-derived, if they exhibited invasive features with architectural destruction, or bronchioloalveolar, if they exhibited preservation of the lung architecture.

**2D-PAGE and Protein Quantification.** Analytical 2D-PAGE was performed as described previously (6). After separation, the protein spots were visualized by a photochemical silver-based staining technique (7). Each gel was scanned using a Kodak CCD camera. Spot detection was accomplished by using Bio Image Visage System software (Bioimage Corp., Ann Arbor, MI). Each gel generated 1600–2200 detectable spots, of which 820 spots were selected for quantitative measurement. The integrated intensity, which is the value of each spot, was calculated as the measured absorbance units multiplied by square millimeters. The protein spots from each gel were matched to the 820 spots on a “master” gel (8) to allow identification of identical polypeptides between each gel. A total of 250 spots were chosen as ubiquitously expressed reference spots to allow adjustment for subtle variation in protein loading and gel staining. Each of the 820 spots was then mathematically adjusted in relation to the reference spots (9). The resulting data can be accessed with common spreadsheet software.

**Mass Spectrometry and Polypeptide Sequencing.** Some of the polypeptides included in the analysis had been identified prior to this study on the basis of sequencing (10). For MALDI-MS, the protein spots identified for analysis were cut from preparative 2D gels using extracts from A549 lung adenocarcinoma cells (American Type Culture Collection, Rockville, MD). The run parameters were the same as those used for the analytical 2D gels. Identification of proteins was performed by trypsin digestion, followed by MALDI-MS. This allowed for a “fingerprint” to be created for each protein spot based on the molecular weight of the trypsin-digested products. The masses were compared with the known trypsin digest databases using the MS-FIT database. The results were given as probability matches to known tryptic digest patterns established by multiple databases.

**Affymetrix Oligonucleotide Microarrays.** Total RNA was isolated from 76 of the tumors and 9 of the normal lung samples using Trisol reagent (Life Technologies, Inc.). The resulting RNA was subjected to further purification using RNeasy spin columns (Qiagen, Inc., Valencia, CA) and used to generate cRNA probes. All protocols used for mRNA reverse transcription, second strand synthesis, production of cDNA and cRNA amplification, hybridization, and washing conditions for the 6800 gene HuGeneFL oligonucleotide arrays were as provided by the manufacturer (Affymetrix, Santa Clara, CA). The arrays were scanned using a GeneArray scanner with data analysis performed using GeneChip 4.0 software (Affymetrix). Details of data trimming and normalization are described elsewhere (11).

**K-ras Mutational Status.** Genomic DNA was isolated from each tumor sample, and 50 ng were subjected to PCR amplification using the primers that encompass codons 12 and 13 of the K-ras gene. The sequences of forward and reverse primers are 5′-TATAAGGCTGCTGAAAATGACT-3′ and 5′-CCTGACACGATAATATGC-3′, respectively. Two ng of purified PCR products containing the exon 1 of the K-ras gene were then subjected to thermal cycle sequencing with an internal nested primer (5′-AGGCCTGCTGAAAATGACT-3′) and resolved in 8% urea PAGE gels, dried, and exposed to PhosphorImage screens and visualized using a PhosphorImage scanner (Molecular Dynamics, Sunnyvale, CA). The mutations were determined by comparing each tumor DNA sequence of K-ras 12th and 13th codon to its wild-type sequence GGTGGC.

**2D Western Blotting.** Protein extracts of A549 lung adenocarcinoma cells were run on 2D gels using the identical conditions as used for the analytical 2D gels. The separated proteins were transferred onto polyvinylidene fluoride membranes and incubated for 2 h at room temperature with a blocking buffer consisting of TBST (Tris-buffered saline, 0.01% Tween 20) and 5% nonfat dry milk. Individual membranes were washed and incubated with anti-Erp57 (GRP58) mouse monoclonal antibody (SPA-725, 0.6 μg/ml; StressGen Biotechnologies Corp., Ann Arbor, MI) each generated 1600–2200 detectable spots, of which 820 spots were selected for quantitative measurement. The integrated intensity, which is the value of each spot, was calculated as the measured absorbance units multiplied by square millimeters. The protein spots from each gel were matched to the 820 spots on a “master” gel (8) to allow identification of identical polypeptides between each gel. A total of 250 spots were chosen as ubiquitously expressed reference spots to allow adjustment for subtle variation in protein loading and gel staining. Each of the 820 spots was then mathematically adjusted in relation to the reference spots (9). The resulting data can be accessed with common spreadsheet software.

**Mass Spectrometry and Polypeptide Sequencing.** Some of the polypeptides included in the analysis had been identified prior to this study on the basis of sequencing (10). For MALDI-MS, the protein spots identified for analysis were cut from preparative 2D gels using extracts from A549 lung adenocarcinoma cells (American Type Culture Collection, Rockville, MD). The run parameters were the same as those used for the analytical 2D gels. Identification of proteins was performed by trypsin digestion, followed by MALDI-MS. This allowed for a “fingerprint” to be created for each protein spot based on the molecular weight of the trypsin-digested products. The masses were compared with the known trypsin digest databases using the MS-FIT database. The results were given as probability matches to known tryptic digest patterns established by multiple databases.

**Affymetrix Oligonucleotide Microarrays.** Total RNA was isolated from 76 of the tumors and 9 of the normal lung samples using Trisol reagent (Life Technologies, Inc.). The resulting RNA was subjected to further purification using RNeasy spin columns (Qiagen, Inc., Valencia, CA) and used to generate cRNA probes. All protocols used for mRNA reverse transcription, second strand synthesis, production of cDNA and cRNA amplification, hybridization, and washing conditions for the 6800 gene HuGeneFL oligonucleotide arrays were as provided by the manufacturer (Affymetrix, Santa Clara, CA). The arrays were scanned using a GeneArray scanner with data analysis performed using GeneChip 4.0 software (Affymetrix). Details of data trimming and normalization are described elsewhere (11).

**K-ras Mutational Status.** Genomic DNA was isolated from each tumor sample, and 50 ng were subjected to PCR amplification using the primers that encompass codons 12 and 13 of the K-ras gene. The sequences of forward and reverse primers are 5′-TATAAGGCTGCTGAAAATGACT-3′ and 5′-CCTGACACGATAATATGC-3′, respectively. Two ng of purified PCR products containing the exon 1 of the K-ras gene were then subjected to thermal cycle sequencing with an internal nested primer (5′-AGGCCTGCTGAAAATGACT-3′) and resolved in 8% urea PAGE gels, dried, and exposed to PhosphorImage screens and visualized using a PhosphorImage scanner (Molecular Dynamics, Sunnyvale, CA). The mutations were determined by comparing each tumor DNA sequence of K-ras 12th and 13th codon to its wild-type sequence GGTGGC.

**2D Western Blotting.** Protein extracts of A549 lung adenocarcinoma cells were run on 2D gels using the identical conditions as used for the analytical 2D gels. The separated proteins were transferred onto polyvinylidene fluoride membranes and incubated for 2 h at room temperature with a blocking buffer consisting of TBST (Tris-buffered saline, 0.01% Tween 20) and 5% nonfat dry milk. Individual membranes were washed and incubated with anti-Erp57 (GRP58) mouse monoclonal antibody (SPA-725, 0.6 μg/ml; StressGen Biotechnologies Corp., Ann Arbor, MI) each generated 1600–2200 detectable spots, of which 820 spots were selected for quantitative measurement. The integrated intensity, which is the value of each spot, was calculated as the measured absorbance units multiplied by square millimeters. The protein spots from each gel were matched to the 820 spots on a “master” gel (8) to allow identification of identical polypeptides between each gel. A total of 250 spots were chosen as ubiquitously expressed reference spots to allow adjustment for subtle variation in protein loading and gel staining. Each of the 820 spots was then mathematically adjusted in relation to the reference spots (9). The resulting data can be accessed with common spreadsheet software.

**Mass Spectrometry and Polypeptide Sequencing.** Some of the polypeptides included in the analysis had been identified prior to this study on the basis of sequencing (10). For MALDI-MS, the protein spots identified for analysis were cut from preparative 2D gels using extracts from A549 lung adenocarcinoma cells (American Type Culture Collection, Rockville, MD). The run parameters were the same as those used for the analytical 2D gels. Identification of proteins was performed by trypsin digestion, followed by MALDI-MS. This allowed for a “fingerprint” to be created for each protein spot based on the molecular weight of the trypsin-digested products. The masses were compared with the known trypsin digest databases using the MS-FIT database. The results were given as probability matches to known tryptic digest patterns established by multiple databases.
of Kononen et al. (12) and used the best representative morphological areas of the tumors in this study. Deparaffinized sections of the pulmonary adenocarcinoma tissue microarray were microwaved after pretreatment in citric acid to retrieve antigenicity. The sections were incubated with blocking solution containing PBS and 1% bovine serum albumin for 60 min at room temperature. The antibodies examined included the anti-UCHL1 (PGP9.5) rabbit polyclonal antibody, anti-Erp57 (GRP58) mouse monoclonal antibody (used for frozen tissue sections), and anti-p53 mouse monoclonal antibody (1.0 μg/ml; Dako Corp., Carpinteria, CA). The sections were incubated with primary antibodies overnight at 4°C. The immunocomplex was visualized by the immunoglobulin enzyme bridge technique using a Vector ABC-peroxidase kit (Vector Laboratories, Burlingame, CA) with 3,3′-diaminobenzidine tetrachloride as a substrate. The sections were lightly counterstained with hematoxylin.

**mRNA in Situ Hybridization of Tissue Microarrays.**

Because of the lack of availability of an antibody, the presence and cellular abundance of mRNA for TPI was determined using ISH with the tissue microarrays. Biotinylated oligonucleotides (*GCCCCATTAGTCACTTTGTAGC) and (*CAGAGGGACTCGGAGTAATCG) were synthesized using the published TPI mRNA sequence (13). ISH was carried out as described previously (14). The sections were deparaffinized, hydrated, and washed with PBS for 10 min. Next, they were treated with proteinase K (20 μg/ml in PBS) for 10 min at 37°C, followed by treatment with 0.2 M HCl. Postfixation was performed using 4% (w/v) paraformaldehyde. The sections were then hybridized overnight at 42°C with the labeled oligonucleotides. After hybridization, the sections were washed using increased stringency with SSC (0.15 M NaCl, 15 mM sodium citrate, pH 7.2) and then subjected to immunostaining with a mouse anti-biotin monoclonal antibody. After amplification of the biotin-antibody complex with biotin-labeled horse anti-mouse IgG and alkaline phosphatase-labeled streptavidin, the sites of alkaline phosphatase were visualized with nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate as reported previously (14). For each run, hybridization with an antisense probe and without the antisense probe (control) was used. ISH of β-actin (a gift from Dr. Sakiyama, Chiba, Japan) was performed on tissue microarray sections of all of the lung tissue specimens to assess their mRNA integrity. The sections were lightly counterstained with nuclear fast red.

**Statistical Analysis.**

S-plus software (Insightful Corp., Seattle, WA) was used to investigate the specific features of protein expression. t tests were used to identify the differences in mean values between comparison groups. The relationship between the levels of protein and mRNA expression was examined using the Spearman correlation coefficient statistical method (15). P < 0.05 (two-sided) was considered statistically significant.

**RESULTS**

By comparing protein expression levels among 93 lung adenocarcinomas and 10 uninvolved lung samples, nine different enzyme proteins were identified using 2D-PAGE and MALDI-MS or peptide sequencing. These proteins were all significantly increased in the lung adenocarcinomas (a 1.4–10.6-fold increase). They included the antioxidant enzyme AOE372, ATP5D, B4GALT, cytosolic inorganic pyrophosphatase (PPase), GRP58, GSTM4, P4HB, TPI, and ubiquitin thiolesterase (UCHL1; Fig. 1 and Table 1). Some of these proteins were identified as having multiple isoforms. The proteins in general exhibited increased expression of all individual
isoforms relative to normal lung, except for P4HB, which demonstrated one isoform that was significantly overexpressed in lung adenocarcinomas, and one isoform that was unchanged relative to uninvolved lung tissue.

The frequency of protein expression in the individual tumor and normal samples was determined using the value of the normal lung, with the mean ± 2 SD as the cutoff value. Proteins that were significantly increased in lung adenocarcinomas were detected at this level in 35.5–96.8% of the tumor samples (Fig. 2). Only three proteins (or one isoform) were detected at this level in 10% or fewer of the normal lung samples.

The protein expression values of the nine tumor-associated enzyme proteins were examined for potential correlation with clinicopathological variables including: tumor stage, tumor classification, tumor differentiation, angiolymphatic invasion, lymphocytic response, P53 nuclear protein accumulation, K-ras 12th/13th codon mutation, and smoking status (Table 2). AOE372 was found to be overexpressed in poorly differentiated tumors relative to moderately differentiated tumors (P = 0.04). PPase was increased in bronchial-derived adenocarcinomas (P = 0.02) and in patients with a positive smoking history (P = 0.04). GRP58 (isoform 353) was increased in tumors with K-ras mutations (P = 0.04). P4HB (isoform 320) was decreased in tumors having a positive lymphocytic response (P = 0.03). TPI

### Table 1. Protein expression in lung adenocarcinomas and normal lung tissues

<table>
<thead>
<tr>
<th>Protein spot no.</th>
<th>Protein name</th>
<th>Tumor mean ± SD (n = 93)</th>
<th>Normal mean ± SD (n = 10)</th>
<th>Fold change</th>
<th>T/N*</th>
<th>P SI vs. SIII</th>
<th>P SI vs. N</th>
<th>P SIII vs. N</th>
<th>P T vs. N</th>
<th>Identification methods</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1193</td>
<td>AOE372</td>
<td>0.208 ± 0.218</td>
<td>0.020 ± 0.039</td>
<td>10.6</td>
<td>0.140</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td>MALDI-MS</td>
<td>This study</td>
</tr>
<tr>
<td>1354</td>
<td>ATP5D</td>
<td>1.155 ± 0.418</td>
<td>0.663 ± 0.210</td>
<td>1.7</td>
<td>0.848</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td>MALDI-MS</td>
<td>This study</td>
</tr>
<tr>
<td>823</td>
<td>B4GALT</td>
<td>0.503 ± 0.229</td>
<td>0.351 ± 0.067</td>
<td>1.4</td>
<td>0.827</td>
<td>0.0001</td>
<td>0.0017</td>
<td>&lt;0.0001</td>
<td></td>
<td>MALDI-MS</td>
<td>This study</td>
</tr>
<tr>
<td>902</td>
<td>PPase</td>
<td>0.272 ± 0.503</td>
<td>0.036 ± 0.039</td>
<td>7.6</td>
<td>0.222</td>
<td>0.0001</td>
<td>0.0076</td>
<td>&lt;0.0001</td>
<td></td>
<td>MALDI-MS</td>
<td>This study</td>
</tr>
<tr>
<td>350</td>
<td>GRP58</td>
<td>0.161 ± 0.087</td>
<td>0.055 ± 0.045</td>
<td>2.9</td>
<td>0.547</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>353</td>
<td>GRP58</td>
<td>0.527 ± 0.230</td>
<td>0.240 ± 0.113</td>
<td>2.2</td>
<td>0.996</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>1138</td>
<td>GSTM4</td>
<td>0.874 ± 0.356</td>
<td>0.219 ± 0.065</td>
<td>4.0</td>
<td>0.919</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>320</td>
<td>P4HB</td>
<td>0.072 ± 0.069</td>
<td>0.038 ± 0.024</td>
<td>1.9</td>
<td>0.418</td>
<td>0.0094</td>
<td>0.0213</td>
<td>0.0029</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>321</td>
<td>P4HB</td>
<td>0.096 ± 0.088</td>
<td>0.127 ± 0.057</td>
<td>0.8</td>
<td>0.115</td>
<td>0.0503</td>
<td>0.8796</td>
<td>0.1495</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>1161</td>
<td>TPI</td>
<td>1.276 ± 0.817</td>
<td>0.587 ± 0.184</td>
<td>2.2</td>
<td>0.013</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>1210</td>
<td>TPI</td>
<td>0.278 ± 0.199</td>
<td>0.130 ± 0.105</td>
<td>2.1</td>
<td>0.162</td>
<td>0.0051</td>
<td>0.0023</td>
<td>0.0014</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>1213</td>
<td>TPI</td>
<td>0.307 ± 0.227</td>
<td>0.116 ± 0.175</td>
<td>2.7</td>
<td>0.218</td>
<td>0.0149</td>
<td>0.0028</td>
<td>0.0074</td>
<td></td>
<td>Sequencing</td>
<td>(10)</td>
</tr>
<tr>
<td>1242</td>
<td>UCHL1</td>
<td>0.333 ± 0.471</td>
<td>0.097 ± 0.089</td>
<td>3.5</td>
<td>0.160</td>
<td>0.0010</td>
<td>0.0075</td>
<td>&lt;0.0001</td>
<td></td>
<td>MALDI-MS</td>
<td>(22)</td>
</tr>
<tr>
<td>1246</td>
<td>UCHL1</td>
<td>0.597 ± 0.399</td>
<td>0.294 ± 0.061</td>
<td>2.0</td>
<td>0.528</td>
<td>&lt;0.0001</td>
<td>0.0005</td>
<td>&lt;0.0001</td>
<td></td>
<td>MALDI-MS</td>
<td>(22)</td>
</tr>
</tbody>
</table>

* T, tumor; N, normal; SI, stage I; SIII, stage III. T test was used.
Table 2. Protein changed in other clinicopathological variables and correlated with mRNA

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Spot no.</th>
<th>Classification</th>
<th>Smoking status</th>
<th>Lymphocytic response</th>
<th>Protein ratio in BD</th>
<th>Protein ratio in smoking</th>
<th>mRNA ratio in BD</th>
<th>mRNA ratio in smoking</th>
<th>Cooperated with mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>P+ (n = 36)</td>
<td>P+ (n = 40)</td>
<td>P- (n = 22)</td>
<td>P- (n = 22)</td>
<td>P+ (n = 47)</td>
<td>P- (n = 47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P+ vs. P-</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>n vs. M</td>
<td>P</td>
<td>M</td>
<td>W</td>
<td>P</td>
<td>M</td>
<td>W</td>
</tr>
<tr>
<td>ras mutation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4GALT</td>
<td>0.5909</td>
<td>0.8247</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
</tr>
<tr>
<td>ATP5D</td>
<td>0.6114</td>
<td>0.8142</td>
<td>0.6973</td>
<td>0.6973</td>
<td>0.6973</td>
<td>0.6973</td>
<td>0.6973</td>
<td>0.6973</td>
<td>0.6973</td>
</tr>
<tr>
<td>PPase</td>
<td>0.5909</td>
<td>0.8247</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
<td>0.3611</td>
</tr>
<tr>
<td>GRP58</td>
<td>0.7038</td>
<td>0.4229</td>
<td>0.4887</td>
<td>0.4887</td>
<td>0.4887</td>
<td>0.4887</td>
<td>0.4887</td>
<td>0.4887</td>
<td>0.4887</td>
</tr>
<tr>
<td>GRP58</td>
<td>0.3221</td>
<td>0.2654</td>
<td>0.2049</td>
<td>0.2049</td>
<td>0.2049</td>
<td>0.2049</td>
<td>0.2049</td>
<td>0.2049</td>
<td>0.2049</td>
</tr>
<tr>
<td>P4HB</td>
<td>0.5382</td>
<td>0.1792</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
</tr>
<tr>
<td>P4HB</td>
<td>0.5382</td>
<td>0.1792</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
<td>0.8651</td>
</tr>
<tr>
<td>TPI</td>
<td>0.7829</td>
<td>0.9021</td>
<td>0.3075</td>
<td>0.3075</td>
<td>0.3075</td>
<td>0.3075</td>
<td>0.3075</td>
<td>0.3075</td>
<td>0.3075</td>
</tr>
<tr>
<td>GSTM4</td>
<td>0.8729</td>
<td>0.2249</td>
<td>0.5482</td>
<td>0.5482</td>
<td>0.5482</td>
<td>0.5482</td>
<td>0.5482</td>
<td>0.5482</td>
<td>0.5482</td>
</tr>
<tr>
<td>UCHL1</td>
<td>0.6724</td>
<td>0.4158</td>
<td>0.8322</td>
<td>0.8322</td>
<td>0.8322</td>
<td>0.8322</td>
<td>0.8322</td>
<td>0.8322</td>
<td>0.8322</td>
</tr>
</tbody>
</table>

The correlation coefficients of the protein expression values within the same tumor samples were calculated using the formula: \( r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \), where \( x \) and \( y \) represent the protein expression values of the same sample in BD and smoking, respectively, and \( \bar{x} \) and \( \bar{y} \) represent their respective means.

Note: BA, bronchioloalveolar; BD, bronchial-derived; the values in parentheses are the sample number; P, poor; M, moderate; W, well differentiated.

To examine whether the changes in protein expression may be attributable to transcriptional or other mechanisms of regulation, a comparison of the mRNA expression values and the protein expression values within the same tumor samples was made. Table 2 shows the correlation coefficients for the proteins for which probes for the corresponding genes were also present on the oligonucleotide arrays. Both GRP58 isoforms were significantly correlated with their respective mRNA levels in these tumors (\( P < 0.05 \)). This suggests that the increase in GRP58 protein expression in these tumors is associated with a corresponding increase in its mRNA, thus reflecting transcriptional regulation. No statistically significant correlation of the other protein isoforms with their respective mRNA expression was observed. The relative mRNA expression of these genes between lung adenocarcinomas and uninvolved normal lung tissues is shown in Table 3. The levels of mRNA for AOE372, GRP58, P4HB, TPI, and UCHL1 were found to be significantly increased in lung adenocarcinomas relative to normal lung (\( P < 0.005 \)); however, ATP5D, B4GALT, PPase, and GSTM4 were not.

2D Western blot analysis of A549 lung adenocarcinoma cell lines using an antibody to UCHL1 (PGP9.5) revealed four immunoreactive protein spots (Fig. 3A). Two of the spots (1242 and 1246) were also identified using MALDI-MS. The other two spots were not quantified in the primary lung tumors because most of the gels demonstrated multiple overlapping patterns in that region. Spot 1246 was the predominant isoform identified in the lung tumors, and a high level of expression of UCHL1 was present in 61.3% of these tumors (Table 1 and Fig. 2). The immunohistochemical analysis of UCHL1 in tissue microarrays using the same antibody demonstrated abundant cytoplasmic staining in the lung tumor cells and very low levels in normal lung tissue (Fig. 4, A and B).

2D Western blot analysis of GRP58 confirmed the two spots (353 and 350) identified by peptide sequencing from 2D gels (Fig. 3B). The level of GRP58 isoform (353) was found to be 3.2-fold higher than the other isoform (350); however, both demonstrated the same frequency (52.7%) of expression in the tumor samples (Table 1 and Fig. 2). Immunohistochemical analysis of GRP58 was performed on frozen lung adenocarcinoma tissue samples because of the lack of antibody reactivity using formalin-fixed tissue. GRP58 staining was abundant within the cytoplasm of the tumor cells, with lower levels of staining detected in normal lung from the same patients (Fig. 4, C and D).

TPI mRNA analysis of tumor arrays using ISH indicated a higher level of mRNA expression in the cytoplasm of the tumor cells as compared with normal lung tissue (Fig. 4, E and F). This finding is consistent with the increased TPI mRNA in lung tumors relative to normal lung tissue determined using the oligonucleotide arrays (Table 3; \( P < 0.0001 \)).
DISCUSSION

Previous studies have shown increased expression of specific enzyme proteins in lung cancer or elevated levels in the serum of lung cancer patients. These include neuron-specific enolase, lactate dehydrogenase, 5'-nucleotide phosphodiesterase, thymidine kinase, hexokinase, sialyltransferase, uridine kinase, glucose-6-phosphate dehydrogenase, and ceruloplasmin (16, 17). The increased expression of these proteins may reflect the overall changes in cellular metabolism and growth rate that occur during malignancy (18). The aim of this study was to identify proteins demonstrating increased expression in lung adenocarcinomas and to assess potential changes in the expression of individual protein isoforms, reflecting posttranslational patterns that may correlate with their clinicopathological features.

The candidate tumor markers in this study were found using quantitative assessment with 2D-PAGE gels and mass spectrometry, and a number were increased in lung adenocarcinomas (1.4–10.6-fold) as compared with normal lung tissue. Among these, AOE372, PPase, GSTM4, and UCHL1 were increased 10.6-, 7.6-, 4.0- and 3.5-fold, respectively. The frequency of elevated expression in lung adenocarcinomas was found to range from 35.5 to 96.8% among the 93 tumors examined. GSTM4 was the most consistently overexpressed protein, being present in 96.8% of the tumors. Correlations were observed between overexpression of some proteins and specific clinicopathological variables, including tumor differentiation (AOE372 and TPI), tumor subhistology (PPase), a positive smoking history (PPase, TPI and UCHL1), and a positive lymphocytic response (P4HB). Additional studies will determine the potential implications of these findings.

Oligonucleotide microarray analysis make it possible to survey the expression of thousands of genes in parallel, with applications in expression monitoring, polymorphism analysis, and sequencing (19). To date, however, there have been only a limited number of studies in which both DNA microarrays and proteomic technologies have been compared by applying them to the same sample to describe potential mechanisms of regulation of protein expression (20). Both isoforms of the GRP58 were significantly correlated with their respective mRNA levels. GRP58 mRNA levels were also significantly increased in lung adenocarcinomas as compared with normal lung tissue. This suggests that the increase in protein expression in these tumors reflects increased GRP58 transcription. The lack of correlation observed among the other proteins with their mRNA levels may reflect other posttranslational regulation differences.

In terms of biological processes, the proteins identified in this study can be divided into two main groups. One group includes the enzymes involved in energy-related pathways: ATP5D (energy generation), PPase (phosphate metabolism), TPI (glycolysis), B4GALT (lactose biosynthesis), and UCHL1 (protein degradation). The elevated expression of these enzymes may relate to the increased requirements of both energy and protein synthetic and degradation pathways in rapidly growing tumors. The other group of proteins may be involved in antigen recognition or detoxification reactions or related pathways and include: AOE372 (oxidoreductase), P4HB (oxidoreductase), GRP58 (protein disulfide isomerase), and GSTM4 (stress response). These enzymes may help clear the toxic byproducts resulting from elevated metabolism in these tumors.

As a method of validation of the 2D-PAGE protein expression data, 2D Western blot and immunohistochemistry were used with the same antibodies. The specific isoforms identified by 2D Western blot of UCHL1 and GRP58 were revealed as the same spots identified by MS. Furthermore, the expression of UCHL1 and GRP58 were substantially elevated within the cytoplasm of the tumor cells. UCHL1 (PGP9.5) belongs to the ubiquitin carboxyl-terminal hydrolase (UCH) family that is a part of the cellular proteolytic pathway that regulates many cellular processes, including cell cycle progression and cell death (21). PGP9.5 is widely expressed in neuronal tissues at all stages of differentiation, and more recently, it has been found that PGP9.5 mRNA and protein are highly expressed in lung cancer tissues (55%), independent of neuronal differentiation (21), and thus similar to our findings. Recently, our group demonstrated that autoantibodies to PGP9.5 are detectable in the sera from 9 of 64 patients with lung cancer (22).

Some of the other proteins examined in this study have shown a prior relationship to cancer. GRP58, isoform of protein disulfide isomerase is present in lung tissue (23), and increased following the v-src transformation of normal rat kidney cells and NIH3T3 cells (24). Interestingly, GRP58 (isoform 353) expression was significantly increased in tumors containing the

Table 3

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Tumor mean (n = 76)</th>
<th>Normal mean (n = 9)</th>
<th>Fold change T/N</th>
<th>P (t test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOE372</td>
<td>1112</td>
<td>680</td>
<td>1.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>ATP5D</td>
<td>513</td>
<td>557</td>
<td>0.9</td>
<td>0.1184</td>
</tr>
<tr>
<td>B4GALT</td>
<td>488</td>
<td>479</td>
<td>1.0</td>
<td>0.6874</td>
</tr>
<tr>
<td>PPase</td>
<td>67</td>
<td>48</td>
<td>1.4</td>
<td>0.0902</td>
</tr>
<tr>
<td>GSTM4</td>
<td>447</td>
<td>484</td>
<td>0.9</td>
<td>0.5394</td>
</tr>
<tr>
<td>GRP58</td>
<td>1313</td>
<td>1035</td>
<td>1.3</td>
<td>0.0049</td>
</tr>
<tr>
<td>P4HB</td>
<td>3522</td>
<td>1664</td>
<td>2.1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TPI</td>
<td>5788</td>
<td>3285</td>
<td>1.8</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>UCHL1</td>
<td>406</td>
<td>90</td>
<td>4.5</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* T, tumor; N, normal; boldface indicates P < 0.05.
12th/13th codon K-ras gene mutations. TPI, a key component of the glycolytic pathway that converts dihydroxyacetone phosphate to glyceraldehyde 3-phosphate, shows increased expression in bladder and colon carcinomas (25). Elevated expression of TPI mRNA was present in the lung tumors, and TPI (isoform 1161) was significantly higher in stage III tumors relative to stage I tumors ($P < 0.02$).

Because of differential protein processing and posttranslational modifications, multiple protein isoforms of individual proteins can be identified on 2D gels (20). In this study, TPI, UCHL1, GRP58, and P4HB express two or three different isoforms, respectively, in lung tumors (Fig. 1). These isoforms likely reflect phosphorylation or other modifications and interestingly were found to correlate with different clinicopathological variables. For example, the levels of P4HB (isoform 320) were highly elevated in lung tumors but reduced in patients showing a positive lymphocytic response. These findings suggest that the different protein isoforms may be correlated with specific features or functions of lung tumors. Thus, they may provide useful diagnostic information and further indicate the need to identify the specific modifications underlying these specific protein isoforms.

ACKNOWLEDGMENTS

We thank Dr. Lin Lin for the K-ras mutational analyses and Rork D. Kuick, Eric Puravs, Robert Hinderer, Melissa C. Krause, and Christopher Wood for their assistance in this study.

REFERENCES


3. Schneider, P. M., Praeuer, H. W., Stoeltzing, O., Boehm, J., Manning, J., Metzger, R., Fink, U., Wegerer, S., Hoelscher, A. H., and Roth,


Proteomic Analysis of Lung Adenocarcinoma: Identification of a Highly Expressed Set of Proteins in Tumors


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/8/7/2298

Cited articles
This article cites 22 articles, 6 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/8/7/2298.full.html#ref-list-1

Citing articles
This article has been cited by 25 HighWire-hosted articles. Access the articles at:
/content/8/7/2298.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.