Protein clusters associated with carcinogenesis, histological differentiation and nodal metastasis in esophageal cancer

Hiromitsu Hatakeyama1, 4, Tadashi Kondo1, Kiyonaga Fujii3, Yukihiro Nakanishi2, Hoichi Kato3, Satoshi Fukuda4 and Setsuo Hirohashi1

1 Proteome Bioinformatics Project, National Cancer Center Research Institute, Tokyo, Japan
2 Pathology Division, National Cancer Center Research Institute, Tokyo, Japan
3 Esophageal Surgery Division, Department of Surgery, National Cancer Center Hospital, Tokyo, Japan
4 Department of Otolaryngology, Head and Neck Surgery, Hokkaido University Graduate School of Medicine, Hokkaido, Japan

We examined the proteomic background of esophageal cancer. We used laser microdissection to obtain tumor tissues from 72 esophageal squamous cell carcinoma cases and adjacent normal tissues in 57 of these cases. The 2D-DIGE generated quantitative expression profiles with 1730 protein spots. Based on the intensity of the protein spots, unsupervised classification distinguished the tumor tissues from their normal counterparts, and subdivided the tumor tissues according to their histological differentiation. We identified 498 protein spots with altered intensity in the tumor tissues, which protein identification by LC-MS/MS showed to correspond to 217 gene products. We also found 41 protein spots that were associated with nodal metastasis, and identified 33 proteins corresponding to the spots, including cancer-associated proteins such as alpha-actinin 4, hnRNP K, periplakin, squamous cell carcinoma antigen 1 and NudC. The identified cancer-associated proteins have been previously reported to be individually involved in a range of cancer types, and our study observed them collectively in a single type of malignancy, esophageal cancer. As the identified proteins are involved in important biological processes such as cytoskeletal/structural organization, transportation, chaperon, oxidoreduction, transcription and signal transduction, they may function in a coordinate manner in carcinogenesis and tumor progression of esophageal cancer.

Keywords:
2D-DIGE / Bioinformatics / Carcinogenesis / Esophageal cancer / Lymph node metastasis

1 Introduction

Esophageal cancer is the eighth most common cancer [1] and the sixth leading cause of cancer death worldwide [2]. Despite the use of modern surgical techniques in combination with radio- and chemotherapy, early recurrence is common and the overall 5-year survival rate remains below approximately 40% [3–5]. Poor prognosis of esophageal cancer is attributed
to extensive local invasion and frequent regional lymph node metastasis even at initial diagnosis, however, the mechanisms of esophageal cancer progression remain largely obscure. Previous investigations reported that aberrant regulation of PGP9.5 [6], epidermal growth factor receptor [7], p21 [8] and stromelysin [9] is associated with short postoperative survival. However, as multiple genetic and epigenetic events have been observed in cancer progression, the development of complex malignant phenotypes is unlikely to be solely attributable to any single gene. Thus, comprehensive and integrative studies that will link the individual molecular aberrations observed is required to broaden our understanding of esophageal carcinogenesis and subsequently lead to the development of better treatment options for esophageal cancer patients.

Global mRNA expression studies conducted using array-based methodologies identified the gene clusters responsible for the carcinogenesis and progression of esophageal cancer [10–12]. However, practical biomarkers for novel therapeutic strategies and genes to act as molecular targets have not been identified yet, indicating some of the limitations of this approach. The proteins contained in the cells are the functional translations of the genome and directly control the malignant phenotypes of tumor cells. However, studying DNA sequences and measuring the amount of RNA do not predict post-translational aberrations resulting from phosphorylation, glycosylation or proteolysis that occur in cancer progression. In addition, global comparison between transcriptome and proteome revealed that the expression level of mRNAs is not always parallel with that of corresponding proteins [13-15].

Previous studies using proteomic approaches detected proteins that were aberrantly regulated in esophageal cancer tissues [16-18]. In this study, we were able to detect candidate biomarkers for early diagnosis and monitoring cancer progression by improving the existing proteomics strategies.

In this report, we used laser microdissection to recover tumor cells and the matched neighboring normal epithelial cells from the surgical specimens of 72 esophageal cancer cases, and subjected the recovered cells to proteomic analysis using 2D-DIGE. The 2D-DIGE has previously been applied to laser-microdissected esophageal tumor tissues [19, 20]. However, the fluorescent dyes used to label the proteins had low sensitivity, equivalent to that of silver staining, and the number of samples examined in these studies was not sufficient to allow meaningful bioinformatics analyses. Here, we used highly sensitive fluorescent dyes for 2D-DIGE, integrated proteomic and clinico-pathological information using bioinformatics methods on a large clinical sample set, and found protein clusters associated with carcinogenesis, histological differentiation and lymph node metastasis of esophageal cancer. We identified the proteins included in the clusters by MS, and validated the proteomic results using specific antibodies. The functional classification and chromosomal localization of the identified proteins were also examined. This is the first detailed, comprehensive and quantitative proteomic study on esophageal cancer using a large clinical sample set.

2 Materials and methods

2.1 Patients

Tumor cells and their adjacent normal mucosal cells were collected from 72 of 185 cases of esophageal squamous cell carcinoma, which were surgically resected in 1998 and 1999 at the National Cancer Center Hospital. Two or three tissue fragments, less than 10 mm³ in volume, were grossly obtained from the 72 cases. Matched normal mucosal tissues were also obtained from 57 of these cases. The resected tissues were snap-frozen in liquid nitrogen and stored at -80°C until use. All 72 enrolled cases were newly diagnosed as squamous cell carcinoma of the esophagus. The patients did not receive anticancer treatment prior to surgery and the cases were followed up for at least five years after surgery. The summary of the clinicopathological data of the cases analyzed is shown in Table 1, while individual case information is given in Table 2.

Table 1. Clinicopathological data of 72 esophageal cancer cases analyzed

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location</th>
<th>Cervical</th>
<th>Upper thoracic</th>
<th>Middle thoracic</th>
<th>Lower thoracic</th>
<th>Abdominal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>8</td>
<td>40</td>
<td>18</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Histological differentiation</th>
<th>Well differentiated</th>
<th>Moderately differentiated</th>
<th>Poorly differentiated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>16</td>
<td>29</td>
<td>27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prognosis (5-year survival)</th>
<th>No evidence of disease</th>
<th>Dead of disease</th>
<th>Dead of another disease</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>39</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2. TNM classification of 72 esophageal tumors analyzed

<table>
<thead>
<tr>
<th>Lymph node metastasis</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth of invasion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>pT2</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>pT3</td>
<td>41</td>
<td>12</td>
<td>53</td>
</tr>
<tr>
<td>pT4</td>
<td>6</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>55</td>
<td>17</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage II A</th>
<th>Stage II B</th>
<th>Stage III</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>8</td>
<td>51</td>
</tr>
</tbody>
</table>

a) Tumors were staged according to the WHO Tumor-Node-Metastasis classification scheme [21].
tion is available in Supplementary Table 1. The tumors were staged based on WHO guidelines [21] and the results are summarized in Table 2. This study was approved by the ethics committee of the National Cancer Center and written informed consent was obtained from the patients.

Based on the WHO classification scheme for esophageal tumors [22], tumors in this study were well differentiated in 16 (22%) cases, moderately differentiated in 29 (40%), and poorly differentiated in 27 (38%) (Table 1) [22]. Based on the same scheme, the depth of invasion was variable, although most cases were pT3 stage (n = 53, 74%). Lymph node involvement was observed in 55 cases (76%) (Table 2). Tumors with distant metastases were not included in this study. As histological differentiation is a significant variable in predicting overall survival [23] and proteins affecting differentiation may be prognostic biomarker candidates, particular emphasis was placed on data analysis in relation to the degree of tumor differentiation.

2.2 Laser microdissection and protein extraction

We specifically recovered tumor cell populations to the exclusion of non-cancerous cells using laser microdissection for the subsequent proteomic study as in our previous report [24]. In brief, the frozen tissues were embedded in OCT compound (Sakura Finetechical, Tokyo, Japan). 8-μm thick sections were prepared using a cryostat (Leica CM 3050S, Leica Microsystems, Wetzlar, Germany) and stained with hematoxylin and eosin to confirm the pathological diagnosis. Eight-micron thick neighboring sections were mounted on a thin supporting polyethylene membrane pretreated with a tissue-adhesive solution (0.1% poly-L-lysine, Sigma Aldrich, St. Louis, MO). The sectioned tissues were routinely stained with Mayer Hematoxylin; all staining procedures were performed on ice. Mayer Hematoxylin-stained sections were subjected to laser microdissection (Leica Laser Microdissection version 3.1.0.0, Leica Microsystems). Tumor cells were recovered from non-necrotic tissues using a pulsed ultraviolet laser beam, avoiding sample contamination with infiltrating inflammatory cells, stromal cells and vascular components (Fig. 1). Corresponding morphologically normal esophageal epithelium samples, located at least 5 cm away from the cancerous tissues, were also obtained from 57 of the 72 patients. Hematoxylin and eosin-stained sections were examined to confirm the diagnosis and following sections were stained only with hematoxylin for the proteomic study, as eosin staining has been found to hinder 2D-DIGE [24]. Protein corresponding to 1 mm² of microdissected area, recorded during microdissection, was recovered from hematoxylin-stained tissues for each 2D-DIGE gel.

Figure 1. Representative histology of normal esophageal epithelium and esophageal tumor with well, moderate, and poor differentiation. Tissue sections were stained with hematoxylin and eosin for histological observation and hematoxylin alone for the proteomic study. The microscopic appearance of the tissues before and after laser microdissection is demonstrated.
2.3 Preparation of fluorescence-labeled protein samples

The microdissected tissues were immediately treated with urea lysis buffer, consisting of 6 M urea, 2 M thiourea, 3% CHAPS and 1% Triton X-100. The total area of microdissected tissue was 1 mm² per 2D-DIGE gel [24]. Protein labeling was carried out as in our previous report [25]. In brief, microdissected tissues with an area of 3 mm² were incubated with 50 µL of the urea lysis buffer containing 40 mM Tris-HCl (pH 8.0). The protein samples were then reduced with 8 nmol tris-(2-carboxyethyl) phosphine hydrochloride (TCEP; Sigma Aldrich) at 37°C for 60 min and were fluorescence labeled by incubation with 12 nmol of Cy5 (CyDye DIGE Fluor saturation dye, GE Healthcare Biosciences, Uppsala, Sweden) at 37°C for 30 min. The labeling reaction was terminated by addition of urea lysis buffer containing DTT and Pharmalyte (pH 4–7, GE Healthcare Biosciences) so that their final concentrations were 65 mM and 2.0%, respectively.

We created an internal control sample by mixing a small portion of an individual protein samples. The protein concentration was measured with a Protein Assay Kit (Bio-Rad Laboratories, Hercules, CA). Five micrograms of the internal control sample was incubated with 2 nmol TCEP at 37°C for 60 min and labeled with 3 nmol of Cy3 dye (CyDye DIGE Fluor saturation dye, GE Healthcare Biosciences) at 37°C for 30 min. After terminating the labeling reaction, individual Cy5-labeled samples corresponding to 1 mm² of laser-microdissected area were mixed with 5 µg of Cy3-labeled internal control sample and urea lysis buffer containing 35 mM DTT and 1.0% Pharmalyte (GE Healthcare Biosciences) was added to a final volume of 420 µL per sample.

2.4 2D-PAGE and image acquisition

Protein expression profiles were created as in our previous report, with some modifications [25]. In brief, for first dimension separation, IPG gels (pI range 4 to 7, 24 cm length) were rehydrated with a mixture of the Cy3-labeled internal control sample with the Cy5-labeled individual samples at room temperature overnight. IEF was performed using Multiphor II (GE Healthcare Amersham Biosciences) at 20°C. Following equilibration in a buffer containing 6 M urea, 2% SDS, 50 mM Tris-HCl pH 8.8, 30% glycerol and 32 mM DTT, IPG gels were transferred, in batches of 12, onto 12.5% homogenous polyacrylamide gels and embedded in agarose between low-fluorescent glass plates. Proteins were then subjected to a second dimension separation at 17 W for 15 h at 20°C on a 40-cm long SDS-PAGE gel using a vertical electrophoresis apparatus equipped with a cooling system (Bio-Craft, Tokyo, Japan). For preparative purposes, 200 µg of proteins were first labeled with a single fluorescent dye (CyDye DIGE Fluor saturation dye) and subjected to 2D-PAGE. Gels were scanned at the appropriate wavelengths for Cy3 and Cy5 with Typhoon Trio (GE Healthcare Biosciences) to obtain the images of labeled proteins. Spot detection, quantification, and standardization of spot intensities were carried out using the DeCyder 5.0 software (GE Healthcare Biosciences).

Figure 2A demonstrates the protocol of the 2D-DIGE experiments. Both Cy3- and Cy5-images were generated from single gels by laser scan. All Cy3-images contained all spots that were detected on the Cy5-images because they represented the common internal control sample, which was a mixture of all individual samples. Thus, the gel-to-gel variations were canceled, because the Cy5 to Cy3-intensity ratio was standardized for every spot and every gel.

To assess the reproducibility of our 2D-DIGE system, we examined the similarity of the protein expression profiles of identical tissue samples (Fig. 2B). Microdissection was performed twice from the same tumor tissue of case No. 50 (Supplementary Table 1) and protein samples were independently prepared. Protein expression profiles were created from duplicate gels for each protein sample using 2D-DIGE.
A pair wise comparison revealed that more than 98% spots were scattered within twofold differences and the correlation value was high ($r = 0.9406$). Visual inspection of the gel images revealed that the protein spots that showed higher than twofold differences between the profiles were shadowed by the streaking of the other spots or located near the end of the gel (data not shown). Such spots were filtered out during the bioinformatics analyses.

2.5 Data analysis

As previously [25], we used bioinformatics to link quantitative proteomic data with clinicopathological parameters in order to identify the fraction of the proteome most relevant to esophageal cancer progression. In brief, standardized spot intensities were exported to Expressionist (GeneData, Basel, Switzerland), a data mining program. The standardized spot intensities were averaged between the duplicate gels and analyzed using scatter plotting, self-organizing map (SOM) [26, 27], hierarchical clustering and principal component analysis. Survival curves were calculated by the Kaplan-Meier method [28] and differences in survival probabilities were examined with the log-rank test.

2.6 Protein identification by MS

In-gel digestion and MS protein identification were described as in our previous report [29]. In brief, the target protein spots were recovered from the gels using an automated spot recovering machine, ProHunter (AsOne, Osaka, Japan) into a 96-well PCR plate. Gel plugs were washed with methanol, ammonium bicarbonate, and ACN three times, dried and treated with TPCK-treated trypsin overnight. The peptides were extracted from the gel by treating the gel with ACN.

The tryptic peptides resulting from the in-gel digestion were subjected to analysis by nano-scale microcapillary RP LC-ESI MS/MS. Paradigm MS4 HPLC dual solvent delivery system (Michrom BioResource, Auburn, CA) for micro-flow HPLC, an HTS PAL auto sampler (CTC Analytics, Zwingen, Switzerland) and a Finnigan LTQ linear IT mass spectrometer (Thermo Electron, San Jose, CA) equipped with a nanoESI (NSI) source (AMR, Tokyo, Japan) were used for protein identification as described [30]. Digested peptide mixtures were separated on a microcapillary RP Magic C18 column (3 μm, 200 Å, 50×0.2 mm i.d.; Michrom). Peptides were eluted through 10 to 80% linear gradient buffer B (10% water and 0.1% formic acid in ACN v/v) in buffer A (2% ACN and 0.1% formic acid in water v/v) over 10 min. The effluent solvent from the HPLC was placed into the mass spectrometer through an NSI needle at a flow rate of 1.0–1.2 μL/min (FortisTip; OmniSeparo-TJ, Hyogo, Japan). The voltage was 1.8 kV and the capillary was heated to 200°C. No sheath or auxiliary gas was used. The mass spectrometer was operated in a data-dependent acquisition mode in which MS acquisition with a mass range of m/z 450–1800 would automatically switch to MS/MS acquisition under the automated control of the Xcalibur software (version 2.0, Thermo Electron). The full MS scan was acquired by the following MS/MS experiments with an isolation width of m/z 2.0; the activation amplitude parameter was set at 35%, on the three most abundant ions detected in the survey scan. Data were acquired with dynamic mass-exclusion windows that had a 30-s exclusion duration and exclusion mass widths of –1.0 and +2.0 Da.

Raw data were converted to dta format (peak list file) using ExtractMS version 2.11 (ThermoElectron), the software supplied with the instrument, with peptide mass range set at 450 to 600 Da and MS/MS minimum peak list set at 25, prior to launching MASCOT searches. MASCOT (version 2.1, Matrix Science, London, UK) searches were performed against Homo sapiens subsets of the sequences in the Swiss-Prot (12867 sequences in the Sprot_47.8 fasta file) and NCBI (131447 sequences in the NCBI_nr_20050422 fasta file) non-redundant protein sequence databases. The following search parameters were used in all MASCOT searches: tolerance of two missed trypsin cleavages, variable modification on the methionine residue (oxidation, +16 Da), and a maximum error tolerance of ±2.0 Da in the MS data and ±1.0 Da in the MS/MS data. Protein hits with more than two significant matched peptides with the distinct sequences ($p < 0.05$, which with our search parameters equals a MASCOT ions score of 35 or more for the Swiss-Prot database and 42 or more for the NCBI database) were statistically considered to estimate the confidence of protein identifications. In addition, the MS/MS spectra of the identified peptides were manually inspected.

2.7 Western blotting

Proteins were separated by SDS-PAGE on a 10–20% polyacrylamide gradient gel and transferred onto an NC membrane. The differential expression of the identified proteins was monitored using antibodies against cytokeratin 14 (1:500, Neo Markers, Fremont, CA), periplakin (1:200, Santa Cruz Biotechnology, Santa Cruz, CA), annexin I (1:5000, BD Bioscience, San Jose, CA), squamous cell carcinoma antigen 1/2 (SCCA1/2) (1:200, Santa Cruz Biotechnology), calgelulin B (1:200, Santa Cruz Biotechnology), HSP60 (1:5000, BD Biosciences), and beta-actin (1:1000, Abcam, Cambridge, CB, UK) as controls. The secondary antibodies against mouse IgG, rabbit IgG (both GE Healthcare Biosciences) and goat IgG (Santa Cruz) were used for the antibodies against cytokeratin 14, annexin I, HSP60 and beta-actin; against SCCA1/2 and calgelulin B; and against periplakin, respectively. Antibody-antigen complexes were visualized with an ECL system (GE Healthcare Biosciences) using LAS 1000 (Fuji Film, Tokyo, Japan). The intensity of the protein bands was quantified and the relative intensity for the examined proteins was calculated by standardizing the intensity of the beta-actin bands on the same membrane.
2.8 Chromosomal location of the genes corresponding to the identified proteins

The chromosome location of the identified proteins was studied by searching the NCBI database using a database search software, Annotation Tracker (GE Healthcare Biosciences).

3 Results

3.1 Hierarchical clustering analysis of tumor samples based on their protein expression profiles

Seventy-two tumor tissues and 57 normal tissues were classified by hierarchical clustering analysis according to their protein expression profiles (Fig. 3). Protein spots that appeared in at least 80% of the Cy3-images were used for the analysis. The numbers of spots analyzed per sample ranged between 1414 and 1730. Based on the overall similarity of protein expression, the samples were divided into two groups: tumor tissues (Tree I) and normal epithelial tissues (Tree II) (Fig. 3). The tumor samples in Tree I were further grouped reflecting their histological differentiation; all but one well-differentiated tumors were located in one branch together with the moderately differentiated tumors, and all poorly differentiated tumors were distinguished from the well-differentiated tumors. The presence or absence of lymph node metastases did not appear to be associated with the proteomic classification of tumors (Tree I) or normal epithelial tissues (Tree II): both were grouped independently of their lymph node metastasis status. The anatomic site was not associated with the proteomic profile classification either. Enlarged trees with the patient ID and the number of spots in the individual samples are demonstrated in Supplementary Figs 1 and 2, respectively. These observations suggested that the proteomic profiles most dominantly reflected the malignant transformation, and secondly the histological differentiation.

The protein spots were also clustered according to their expression level across the 129 samples (72 tumor tissues and 57 normal counterparts). Of 1730 protein spots identified, the protein spots in categories A (544 spots, 31%) and C (705 spots, 41%) showed decreased or increased intensity, respectively, in many of the tumor tissues compared with their normal tissues. There did not seem to be any obvious or consistent differences between the protein expression levels of the normal and the tumor tissues in category B (481 spots, 28%). Selection of spots with different signal intensity levels was later achieved, however, taking into account statistical significance.

3.2 Comparison of the protein expression profiles

We examined the similarity of the protein expression profiles of the 129 laser-microdissected tissue samples. Pair wise correlation coefficients across all samples were performed and the results are summarized in the correlation matrix (Fig. 4A). The correlation matrix demonstrated that samples in the same category, either normal epithelial tissues or tumor tissues, had similar protein expression profiles. By visual inspection, normal epithelial tissues showed more homogeneous protein expression profiles compared with
Figure 4. (A) Correlation matrix summarizing the overall similarity of the expression profiles of unselected protein spots for all sample pairs. The red and green colors indicate that the protein expression in the paired samples had high and low similarity, respectively. Samples in the same category, either normal epithelial tissues or tumor tissues, had similar protein expression profiles. (B) Principal component analysis grouped the tissue samples based on unselected protein spots. The normal tissues and the status of histological differentiation of tumor tissues are color-coded as indicated in the panel. Normal epithelial tissues were distinguished from tumor tissues, and well-differentiated tumors from poorly differentiated ones. The distances between the normal epithelial tissue samples were smaller than the distances among the tumor tissue samples, suggesting that the protein expression profiles of the normal tissues may be more homogeneous than the profiles of the tumor tissues. The patient ID and the number of spots used in the study are shown in Supplementary Fig. 2.

Tumor tissues. Indeed, the average correlation coefficient between the normal tissues was higher than that between the tumor tissues (0.54 and 0.43, respectively). In contrast, the average correlation coefficient of pairs of normal and tumor tissues was low (r = -0.088). The differences in the protein expression profiles between the tumor tissues with different histological differentiation, and those with and without lymph node metastases were less obvious in this study.

The similarity of protein expression profiles was also examined by non-hierarchical classification. Principal component analysis defines the directions of maximum variance between the samples and represents the samples in a multi-dimensional space constructed by the resulting dimensions. The 129 samples were represented in the 3-D space made by the first three major directions generated (Fig. 4B). In the principal component analysis, the normal epithelial tissues were distinguished from the tumor tissues, and the tumor tissues with well-differentiated histology were separated from those with poorly differentiated histology. In addition, the distances between the normal epithelial tissue samples were smaller than those among the tumor samples, suggesting that the protein expression profiles of the normal tissues may be more homogeneous than those of the tumor tissues. The tumor tissues were not divided according to the status of lymph node metastasis or the anatomical site of origin in the multidimensional space (data not shown).

Taken together, the greatest differences were, predictably, observed between the proteomic profile of normal and tumor cells. Histological differentiation appeared to be the second most dominant factor affecting the proteome. Although differences in the lymph node metastasis status affected patient survival significantly, they were not reflected in the overall proteomic profile or individual proteomic characteristics of the cases examined in this study.

3.3 Proteins differentially expressed between tissue groups

Hierarchical clustering analysis suggested the presence of protein spots whose intensity seemed to be different between normal epithelial tissues and tumor tissues (Fig. 3). To identify the proteins associated with carcinogenesis and histological differentiation, we compared protein expression levels between sample groups. We selected spots that corresponded to proteins whose average expression level showed more than twofold differences between sample groups that were statistically significant (Wilcoxon test, p-value <0.01). The comparison between normal and tumor tissue groups resulted in the identification of 338 such proteins, while the comparison of the protein expression profiles of normal tissues with those of highly-, moderately- and poorly differentiated tumors resulted in the selection of 326, 316, and 389 spots, respectively. Many protein spots appeared repeatedly in the comparisons, and as a total, 498 distinct proteins were selected. The number of proteins differentially expressed, including the overlapping ones, is summarized in Supplementary Fig. 3. Of the 498 spots, 221 had more and 227 had less intensity in the tumor tissues compared with normal tissues.
The relations between the fold differences and the number of spots is summarized in Fig. 5A. Regarding the spots whose intensity was increased in the tumor tissue samples, the differences were mostly less than threefold. In contrast, many of the spots with lower intensity in the tumor tissues showed more than threefold difference compared with the normal tissues.

The pair wise similarity of the expression profiles of the 498 selected protein spots was examined in all samples, and the results are summarized in Fig. 5B. The normal and tumor tissues shared similar expression patterns with each other as a whole, the average r value being 0.65 and 0.49, respectively. In contrast, the average r value was only −0.14 in the pair of normal and tumor tissues, clearly showing that the expression pattern of the 498 selected protein spots is significantly different between normal and tumor tissues.

Similarly, principal component analysis based on the expression profiles of the 498 spots also distinguished between the normal and tumor tissues in the multidimensional space (Fig. 5C). Within tumor tissues, the well-differentiated tissues were separated from the poorly differentiated ones.

### 3.4 Self-organizing map of laser-microdissected tissues using 498 selected protein spots

We constructed a self-organizing map that demonstrated that the average spot intensity of the up- or down-regulated proteins, which were categorized in cluster A and B, respectively, was constant in all tumor tissues without obvious correlation with histological differentiation (Figs. 6A and B, upper panel). However, visual inspection of the heat-map suggested that the intensity of some of the 498 spots correlates with histological differentiation (Fig. 6A and B, lower panel). To further examine this point, we created separate self-organizing maps for the 221 and 277 spots with increased and decreased intensity, respectively (Fig. 6C and D, upper panel). We found that the up-regulated protein
Figure 6. Self-organizing map of the protein spots. The plots in the upper panels of (A) and (B) demonstrate the average profile of the 221 up-regulated (cluster A) and the 277 down-regulated proteins (cluster B) in tumors. Average profiles of the sub-cluster of the up- and down-regulated proteins in tumors are shown in clusters C and D in the upper panel of (C), and in clusters E and F in the upper panel of (D), respectively. Individual spot intensity of the individual samples is demonstrated as heat-maps in the lower panels of (C) and (D). Rows in the heat-maps represent protein spots, while columns represent individual samples. Tissue types are color-coded. The averaged expression levels of the proteins in cluster A and B show constant up- and down-regulation in tumors without obvious correlation with histological differentiation (A and B). The subclusters, clusters C and F, show the histology-dependent regulation.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

spots were divided into two subgroups. Seventy-two spots in cluster C showed notably higher expression levels in the well-differentiated tumor tissues compared with moderately and poorly differentiated ones (Fig. 6C, upper panel). The expression levels of the 72 spots in the individual samples are demonstrated in the heat-map panels (Fig. 6C, lower left panel). In contrast, the average level of the 149 spots in cluster D was constant in the tumor tissues irrespective of histological differentiation (Fig. 6C, upper panel and lower right panel). Proteins down-regulated in tumors were also divided into two clusters, clusters E and F, reflecting the correlation between expression level and histological differentiation. Spot intensity in cluster E was consistently lower in all tumors. Spot intensity in cluster F was higher in well- and moderately differentiated tumors (Fig. 6D, upper panel). Spot intensity in the individual samples is demonstrated in the heat-map panels (Fig. 6D, lower right panel). The localization of the spots in clusters C, D, E and F on representative 2-D gels is shown in Supplementary Figs. 4, 5, 6 and 7, respectively.
The 3-D views of selected spots belonging to clusters C, D, E, and F and varying based on tissue histology are shown in Fig. 7. The 3-D images derived from the Cy5-images. As the intensity of the Cy5-image was standardized based on that of the Cy3-image in the identical gel before statistical analysis, this 3-D view does not always reflect the results of the expression study precisely. However, the different intensities of the spots across the samples were obvious even in the Cy5-images.

### 3.5 Proteins associated with lymph node metastasis

We examined whether the number of lymph node metastases is associated with patient outcome in our sample set. A previous study correlated nodal metastasis with dismal prognosis in esophageal cancer [31]. The frequency of lymph node metastases in this study is summarized in Table 3, and the data concerning the individual patients are described in Supplementary Table 1. Kaplan-Meier analysis showed that the patients without lymph node metastasis had a significantly higher survival rate compared with the patients with more than five lymph node metastases ($p = 0.003$, Fig. 8). In contrast, patients with one to five lymph node metastases did not show significant differences in terms of survival rate compared with patients without lymph node metastases (Fig. 8). Thus, we assumed that the spots that had distinct intensity between the tumors without nodal metastases and those with more than five lymph node metastases would correspond to proteins that may be tumor marker candidates to predict patient survival, and we examined these differences in more detail. We found 41 such protein spots that had significantly different intensity between the two groups (Wilcoxon test, $p < 0.01$). The localization of these spots on a representative 2-D gel is shown in Supplementary Fig. 8. The 41 spots corresponded to 32 up- and 9 down-regulated proteins in the lymph node positive groups. Expression patterns of selected spots in all tumors are shown in Fig. 9A. By visual inspection, the intensity of the 32 spots corresponding to up-regulated proteins appeared to correlate with the number of lymph node metastases as a whole (cluster G, Fig. 9A). In contrast, the intensity of the 9 spots corresponding to down-regulated proteins was consistent among tumors with four or less lymph node metastases (cluster H, Figs. 9A and B). In principal component analysis, tumor tissues were distinguished from normal tissues on the basis of the expression pattern of the 41 protein spots (Fig. 9C), suggesting that the expression pattern of these proteins changes during the course of carcinogenesis.

### 3.6 Protein identification by MS and confirmation of the identification using specific antibodies

MS protein identification revealed that the 498 spots showing different intensity between normal and cancer tissues corresponded to proteins generated from 217 distinct genes, and the 41 spots associated with nodal metastasis corresponded to the protein products of 33 distinct genes. The results of identification and data supporting the protein identification are shown in Supplementary Tables 2 and 3.

**Figure 7.** Screen shots using the DeCyder software in BVA mode showing 3-D views of the Cy5-images of representative protein spots in clusters C, D, E and F. Cluster names correspond to those in Fig. 6. The spot numbers correspond to those in Supplementary Figs. 4–7 and Supplementary Tables 2 and 3. The proteins corresponding to the spots were later identified by MS and are demonstrated in the left side of the panels.
We validated the results of 2D-DIGE using specific antibodies. We selected the proteins that showed a representative expression pattern for each cluster. The proteins were separated by SDS-PAGE, transferred onto a membrane and incubated with specific antibodies. The intensity of each band was quantified and standardized by that of beta-actin in the same membrane (Fig. 10). A number of different protein spots for cytokeratin 14 were observed in clusters C and D, and cytokeratin 14 expression in Western blotting was similar to that in cluster C (Fig. 10A). This observation may reflect the fact that a higher number of cytokeratin 14 spots with similar molecular weight was present in cluster C (12 spots) than cluster D (2 spots) (Supplementary Table 2). Western blotting showed that periplakin, annexin I and SCCA1/2 expressions were lower in tumor tissues (Fig. 10A). These results were consistent with those obtained by 2D-DIGE; these proteins were included in cluster E, in which proteins were consistently down-regulated in tumor

![Figure 8](image1)

**Figure 8.** Survival curves of esophageal cancer patients in our sample set, subgrouped according to the number of lymph node metastases present. Survival of the patients with more than five lymph node metastases was significantly shorter than that of patients without lymph node metastases. Survival curves were calculated by the Kaplan–Meier method and statistical differences were calculated by log-rank test.

![Figure 9](image2)

**Figure 9.** Protein spots associated with lymph node metastasis. Average profile of the 41 protein spots the intensity of which was statistically significantly different between tumors without lymph node metastases and those with more than five lymph node metastases are demonstrated (A). Spots with increased or decreased intensity in the tumor tissues compared with the normal tissues are separately demonstrated as clusters G and H, respectively. The intensity of the 41 protein spots in the individual samples is shown in the heat-map (B). Principal component analysis grouped all tissue samples according to the expression level of the 41 protein spots (C). The localization of the protein spots on a representative 2-D gel is shown in Supplementary Fig. 8.

<table>
<thead>
<tr>
<th>Number of lymph node metastases</th>
<th>Number of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>4 or 5</td>
<td>8</td>
</tr>
<tr>
<td>More than 5</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 3. Frequency of lymph node metastasis in the esophageal cancer cases studied
Figure 10. Western blotting results for selected proteins for each cluster validated the 2D-DIGE results. The intensity of each band was quantified and standardized by that of beta-actin in the same membrane (upper panel). The intensity of the bands was standardized by that of beta-actin and compared between the samples (lower panel). (A) The expression levels of selected proteins belonging to clusters C, D, E and F were examined. (B) The expression level of HSP 60 in cluster G was monitored. The color codes correspond to those in Fig. 3. Western blotting successfully validated the results of 2D-DIGE.

3.7 Functional characterization of the proteins identified and chromosomal localization of the corresponding genes

We classified the identified proteins based on their function according to their classification in Gene Ontology and the literature curation (Fig. 11). The proteins frequently observed in clusters A and B were categorized as cytoskeletal/structural proteins, transporters, chaperones/heat shock proteins, proteins in the signal transduction pathway, and proteins involved in proteolysis. The transporter proteins appeared more frequently in cluster A than in cluster B. Proteins involved in proteolysis processes were preferentially down-regulated in the tumor tissues.
Figure 11. Functional classification of the 217 proteins identified corresponding to the 498 protein spots detected. Proteins function is based on their classification in Gene Ontology and the literature curation. Proteins up- \((n = 113)\) and down- \((n = 104)\) regulated in tumor tissues are shown in panels (A) and (B), respectively. The number of proteins in each group and the percentage are indicated in the parentheses. Categories containing only one protein were combined in the ‘miscellaneous’ group. The classification is shown in more detail in Supplementary Table 2.

We investigated the chromosomal localization of the genes corresponding to the identified 223 proteins by searching the NCBI database (Supplementary Table 2). We plotted the number of the identified proteins as a function of chromosome location in Supplementary Fig. 8, and found that a relatively large number of the genes identified were located on chromosomes 1, 12, and 17. We also found that only one gene, L-plastin, derived from chromosome 13, and no gene from the Y chromosome, although multiple genes were located on all the other chromosomes.

4 Discussion

In this study, all cases were treated in a consistent manner by a single hospital-based surgeon and no patient received preoperative therapy. In addition, all tumors were staged pathologically according to strict clinicopathological criteria, and complete follow-up and outcomes data were available for all patients. The tumors studied were matched with normal epithelial tissues from the same patients, avoiding differences due to genetic variations. Moreover, the cells were collected using laser microdissection to avoid contamination from other tissue components. The use of this advantageous sample set allowed us to investigate proteomic differences correlating with the clinicopathological parameters.

We first conducted unsupervised classification of the tissue samples to explore which clinicopathological characters most dominantly corresponded to the proteomic data. The proteome of tumors was obviously distinct from that of normal tissues, and corresponded to the histological differentiation of the tumors. The presence or otherwise of nodal metastases and the anatomic site did not seem to affect the global protein expression pattern. These observations may be consistent with our current understanding of cancer biology. During the course of carcinogenesis, a small number of genetic and epigenetic alterations may cause chromosomal instability and drastic changes in entire regions of the ge-
nomic. Indeed, global genomic studies using array comparative genomic hybridization technique showed that amplifications might concern entire chromosomal regions in esophageal cancer [32–34]. The proteomic differences between normal and tumor tissues might reflect these genomic alterations. Tumor tissues were separated according to their differentiation status on the basis of their proteomic pattern, because histological differentiation is one of the most dominant factors affecting the morphology and behavior of esophageal tumors. Although the lymph node metastasis status and the tumor location are also clinically important features, tumors could not be distinguished into groups based on these features with histological observation alone. Thus, the overall proteomic profiles of tissues may reflect their major phenotypes, as they appear histologically. Proteomic studies may have the potential to identify the mechanisms underlying histopathological differentiation in tumors.

Another interesting observation in the unsupervised study was that the degree of heterogeneity of protein expression profiles depended on the tissue types; the protein expression profiles of the normal tissues were more homogeneous compared with those of the tumor tissues. The proteomic pattern of the normal cells was similar between the different samples, probably because the global protein expression in the normal epithelial cells is strictly regulated. In contrast, the proteomic pattern of the tumor tissues was less homogenous than the normal tissues, given the fact that tumor cells accumulate alterations that are random and/or advantageous for survival during the transformation process. This hypothesis is supported by the observation that the DNA copy number and mRNA expression pattern were more homogenous in the normal mucosa compared with that in colorectal tumor tissues [35].

Of 1730 protein spots studied, 498 showed statistically significant differences in intensity between normal and tumor tissues. MS study revealed that 113 and 104 distinct proteins corresponded to spots with increased and decreased intensity in the tumor tissues, respectively. The proteins in clusters A, C and D, that were up-regulated in the tumor tissues, included several cancer-associated proteins such as alpha-actinin 4 [36], heterogeneous nuclear ribonucleoprotein K (hnRNP K) [37–39], RNA-binding protein FUS [40], N-myc downstream-regulated gene 1 protein [41], sorcin [42, 43], DNA replication licensing factor MCM7 [44], N-myc interactor [45], cathepsin D [46], annexin A5 [47], annexin A1 [48], breast carcinoma amplified sequence 2 [49], tumor protein D54 [50], tumor protein D52 [51], DJ-1 [52], transgelin-2 [53], galectin-1 [54], and nucleophosmin [55]. The proteins in clusters B, E and F, that were down-regulated in the tumor tissues, included proteins known to be associated with cancer, such as periplakin [16], gelsolin [56], caldesmon [57], src-substrate cortactin [58], squamous cell carcinoma antigen 1 [59], annexin A1 [48], hepatoma-derived growth factor [60], calgranulin B [61], and thioredoxin [62]. These proteins are involved in a variety of biological processes including signal transduction, apoptosis, RNA processing, apoptosis, transcription, DNA replication, multi-drug resistance, vesicle trafficking, cytoskeletal organization, and intracellular redox regulation. Notably, although these proteins have been reported to be aberrantly expressed individually in various cancer types, our comprehensive study enabled a panoramic view of their expression in a single type of cancer, namely esophageal cancer. Although these proteins may function in an aberrant but coordinate manner in esophageal cancer, it is presently difficult to map their functional network. Our findings may, therefore, be used as a basis for biomarker development. Proteins found to be up-regulated in esophageal cancer in our study include prognostic markers that are established in other types of cancer; these were sorcin, implicated in acute myeloid leukemia [42], MCM7 in neuroblastoma [63], cathepsin D in endometroid adenocarcinoma [64], breast carcinoma [65], glioma [66], and colon cancer [67], breast carcinoma amplified sequence 2 in breast cancer [68], gelsolin in lung cancer [69], and galectin-1 in lung cancer [70]. Sorcin is also implicated in multidrug-resistance in leukemia cells [71]. These results prompt future studies aimed at further elucidating the prognostic relevance or otherwise of the above proteins, individually or in combination, to develop improved prognostic biomarkers, which in turn may lead to the establishment of, better therapeutic strategies for esophageal cancer patients.

Lymph node metastasis is one of the critical parameters affecting prognosis. In our study, the number of lymph node metastases correlated with patient survival (Fig. 8). We found 41 protein spots the intensity of which was statistically different between the patients without nodal metastasis and those with more than five lymph node metastases. Although only 4 of the 41 protein spots had significantly different intensity between the normal and tumor tissue groups, principal component analysis distinguished the cancer tissues from the normal ones on the basis of the intensity of the 41 protein spots. These observations suggest that the overall expression pattern of this group of proteins becomes aberrant during the course of carcinogenesis, and changes further during cancer progression. Molecular targeting therapy to revert the aberrant expression of these proteins to their normal status may be effective in preventing lymph node metastasis.

The 41 protein spots associated with nodal metastasis corresponded to 33 distinct proteins. Of these 33 proteins, 10 were identified as corresponding to spots that showed increased or decreased intensity in the tumor tissues compared with their normal counterparts. However, apart from spots 791 and 1584, the protein spots whose intensity differed between normal and tumor tissues were different from those associated with nodal metastasis for the same protein. For example, the intensity of spot 1084, identified as corresponding to gelsolin, was diminished in tumors with more than five lymph node metastases. Spot 1084 did not show significant differences in intensity between normal and tumor tissues, but the intensity of the other gelsolin spots, spots 1223 and 1230, was decreased in the tumor tissues, such as periplakin [16], gelsolin [56], caldesmon [57], src-substrate cortactin [58], squamous cell carcinoma antigen 1 [59], annexin A1 [48], hepatoma-derived growth factor [60], calgranulin B [61], and thioredoxin [62]. These proteins are involved in a variety of biological processes including signal transduction, apoptosis, RNA processing, apoptosis, transcription, DNA replication, multi-drug resistance, vesicle trafficking, cytoskeletal organization, and intracellular redox regulation. Notably, although these proteins have been reported to be aberrantly expressed individually in various cancer types, our comprehensive study enabled a panoramic view of their expression in a single type of cancer, namely esophageal cancer. Although these proteins may function in an aberrant but coordinate manner in esophageal cancer, it is presently difficult to map their functional network. Our findings may, therefore, be used as a basis for biomarker development. Proteins found to be up-regulated in esophageal cancer in our study include prognostic markers that are established in other types of cancer; these were sorcin, implicated in acute myeloid leukemia [42], MCM7 in neuroblastoma [63], cathepsin D in endometroid adenocarcinoma [64], breast carcinoma [65], glioma [66], and colon cancer [67], breast carcinoma amplified sequence 2 in breast cancer [68], gelsolin in lung cancer [69], and galectin-1 in lung cancer [70]. Sorcin is also implicated in multidrug-resistance in leukemia cells [71]. These results prompt future studies aimed at further elucidating the prognostic relevance or otherwise of the above proteins, individually or in combination, to develop improved prognostic biomarkers, which in turn may lead to the establishment of, better therapeutic strategies for esophageal cancer patients.

Lymph node metastasis is one of the critical parameters affecting prognosis. In our study, the number of lymph node metastases correlated with patient survival (Fig. 8). We found 41 protein spots the intensity of which was statistically different between the patients without nodal metastasis and those with more than five lymph node metastases. Although only 4 of the 41 protein spots had significantly different intensity between the normal and tumor tissue groups, principal component analysis distinguished the cancer tissues from the normal ones on the basis of the intensity of the 41 protein spots. These observations suggest that the overall expression pattern of this group of proteins becomes aberrant during the course of carcinogenesis, and changes further during cancer progression. Molecular targeting therapy to revert the aberrant expression of these proteins to their normal status may be effective in preventing lymph node metastasis.

The 41 protein spots associated with nodal metastasis corresponded to 33 distinct proteins. Of these 33 proteins, 10 were identified as corresponding to spots that showed increased or decreased intensity in the tumor tissues compared with their normal counterparts. However, apart from spots 791 and 1584, the protein spots whose intensity differed between normal and tumor tissues were different from those associated with nodal metastasis for the same protein. For example, the intensity of spot 1084, identified as corresponding to gelsolin, was diminished in tumors with more than five lymph node metastases. Spot 1084 did not show significant differences in intensity between normal and tumor tissues, but the intensity of the other gelsolin spots, spots 1223 and 1230, was decreased in the tumor tissues, such as periplakin [16], gelsolin [56], caldesmon [57], src-substrate cortactin [58], squamous cell carcinoma antigen 1 [59], annexin A1 [48], hepatoma-derived growth factor [60], calgranulin B [61], and thioredoxin [62]. These proteins are involved in a variety of biological processes including signal transduction, apoptosis, RNA processing, apoptosis, transcription, DNA replication, multi-drug resistance, vesicle trafficking, cytoskeletal organization, and intracellular redox regulation. Notably, although these proteins have been reported to be aberrantly expressed individually in various cancer types, our comprehensive study enabled a panoramic view of their expression in a single type of cancer, namely esophageal cancer. Although these proteins may function in an aberrant but coordinate manner in esophageal cancer, it is presently difficult to map their functional network. Our findings may, therefore, be used as a basis for biomarker development. Proteins found to be up-regulated in esophageal cancer in our study include prognostic markers that are established in other types of cancer; these were sorcin, implicated in acute myeloid leukemia [42], MCM7 in neuroblastoma [63], cathepsin D in endometroid adenocarcinoma [64], breast carcinoma [65], glioma [66], and colon cancer [67], breast carcinoma amplified sequence 2 in breast cancer [68], gelsolin in lung cancer [69], and galectin-1 in lung cancer [70]. Sorcin is also implicated in multidrug-resistance in leukemia cells [71]. These results prompt future studies aimed at further elucidating the prognostic relevance or otherwise of the above proteins, individually or in combination, to develop improved prognostic biomarkers, which in turn may lead to the establishment of, better therapeutic strategies for esophageal cancer patients.
tissues. Diminished expression of gelsolin has been observed in many types of cancer [56, 72, 73] and considered as a negative prognostic factor [69]. Our observation suggests that identical proteins may contribute to carcinogenesis and cancer progression in varying ways depending on the status of PTM.

Our study revealed that tumors with nodal metastases had higher amount of hnRNP K than those without lymph node metastases. The expression of hnRNP K was also up-regulated in esophageal cancer tissues compared with the normal tissues. The overexpression of hnRNP K has been linked to a range of cancers including breast cancer [38], hepatocellular carcinoma [39], and colorectal cancer [37]. In breast cancer, grade III tumors contained more hnRNP K protein than grade II tumors, suggesting that hnRNP K is involved in the mechanisms of metastasis [38]. The hnRNP K functions as a transcriptional coactivator of p53 [74] and regulates the expression of genes involved in mitogenic responses [75, 76]. Therefore, hnRNP K overexpression may cause various genes to be aberrantly expressed and subsequently contribute to the malignant transformation of normal cells. Previous studies using comparative genomic hybridization revealed that gains of 5p15 correlated with distant organ metastasis in esophageal cancer [77]. As hnRNP K is located in this region, it may be one of the target genes for amplification in esophageal cancer. We also found an inverse correlation between NudC expression levels and nodal metastasis. NudC is a microtubules-associating protein and functions in mitosis and cytokinesis [78]. Lin et al. [79] demonstrated that adenovirus expressing NudC inhibited the growth of prostate tumors by blocking cell division in a mouse xenograft model. Both these observations suggest that NudC may potentially be used as a molecular therapy target in esophageal cancer.

Functional classification of the proteins detected in our study demonstrated that more than 60% of the up- and down-regulated proteins in esophageal cancer were categorized as cytoskeletal/structural proteins, transporter proteins, proteins involved in signal transduction, chaperone/heat shock proteins, and proteins controlling proteolysis. Proteins involved in proteolysis were observed more frequently in the down-regulated protein group than in the up-regulated one. These proteins included both proteases and protease inhibitors, and others functioning in the ubiquitin-proteasome pathway. We found that three serine protease inhibitors such as SCCA1, maspin, and serpin B13 were down-regulated in esophageal cancer. All these serpins are located on 18q21 and were down-regulated in head and neck squamous cell carcinoma because of loss of heterozygosity (LOH) [80, 81]. In vitro experiments revealed that maspin [82] and SCCA1 [83] are involved in suppression of tumor growth and invasion in head and neck tumors. Taken together, our observations suggest that decreased expression of maspin and SCCA1 in esophageal cancer may promote cancer cell growth, probably due to genomic alterations. The other 17 proteins, which function in proteolysis processes and were aberrantly expressed in tumors in our study, may also play an important role in tumor progression.

Cancer is a disease of the DNA and all proteomic changes are directly or indirectly attributable to genetic and epigenetic lesions. In this proteomic study, although we found possible association of the changes with reported chromosomal abnormalities in a limited number of proteins, the correlation of proteomic and genomic alterations was largely obscure. Of 223 genes, we found that only 1 gene was coded on chromosome 13 and no gene on chromosome Y. It is unlikely that 2D-DIGE allows the observation of gene products only from chromosomes other than chromosome 13 and Y, as 2D-DIGE identifies proteins with reduced cysteine residues with some expression level. We also noted many products of genes on chromosomes 1, 12 and 17. To understand the molecular background of these observations, we may have to identify the proteins for all spots on the 2-D gel used and examine the chromosomal distribution of the genes. Genomic factors including DNA copy number, somatic mutations, single-nucleotide-polymorphisms, and DNA methylation may influence the mRNA and protein expression levels, and a more comprehensive study may be able to explain the mechanisms that lead to proteomic aberrations at the DNA level.

In conclusion, we combined the application of laser microdissection, 2D-DIGE and MS to a large set of well-characterized cases to construct the proteomic profile of esophageal cancer and identified 240 proteins the expression level of which was associated with carcinogenesis, histological differentiation and the number of lymph node metastases. Notably, proteins previously reported to be associated with cancer individually in different types of tumors, were shown to be aberrantly expressed in a single type of malignancy, esophageal cancer. Integrated and comprehensive studies on such proteins will be beneficial for biomarker development and molecular targeting therapy for esophageal cancer. Further studies that will correlate proteomic and genomic alterations will give us clues to understand the mechanisms controlling the overall protein expression pattern in tumor cells, and such understanding will likely lead to novel therapeutic strategies.

This work was supported by a grant from the Ministry of Health, Labour and Welfare and by the Program for Promotion of Fundamental Studies in the National Cancer Institute of Biomedical Innovation of Japan. Hiromitsu Hatakeyama is a recipient of Research Resident Fellowship from the Foundation for Promotion of Cancer Research (Japan) within the framework of the 3rd Term Comprehensive 10-Year Strategy for Cancer Control.
5 References


