Comparative Proteomics Analysis of Barrett Metaplasia and Esophageal Adenocarcinoma Using Two-dimensional Liquid Mass Mapping*

Jia Zhao‡, Andrew C. Chang§¶, Chen Li‡, Kerby A. Shedden||, Dafydd G. Thomas**, David E. Misek§, Arun Prasad Manoharan‡‡, Thomas J. Giordano¶**, David G. Beer§¶, and David M. Lubman‡¶§§

Esophageal adenocarcinoma, currently the seventh leading cause of cancer-related death, has been associated with the presence of Barrett metaplasia. The malignant potential of Barrett metaplasia is evidenced by ultimate progression of this condition to invasive adenocarcinoma. We utilized liquid phase separation of proteins with chromatofocusing in the first dimension and nonporous reverse phase HPLC in the second dimension followed by ESI-TOF mass spectrometry to identify proteins differentially expressed in six Barrett metaplasia samples as compared with six esophageal adenocarcinoma samples; all six Barrett samples were obtained from the identical six patients from whom we obtained the esophageal adenocarcinoma tissue. Approximately 300 protein bands were detected by mass mappings, and 38 differentially expressed proteins were identified by μLC-MS/MS. The false positive rates of the peptide identifications were evaluated by reversed database searching. Among the proteins that were identified, Rho GDP dissociation inhibitor 2, α-enolase, Lamin A/C, and nucleoside-diphosphate kinase A were demonstrated to be up-regulated in both mRNA and protein expression in esophageal adenocarcinomas relative to Barrett metaplasia. Candidate proteins were examined at the mRNA level using high density oligonucleotide microarrays. The cellular expression patterns were verified in both esophageal adenocarcinomas and in Barrett metaplasia by immunohistochemistry. These differentially expressed proteins may have utility as useful candidate markers of esophageal adenocarcinoma. Molecular & Cellular Proteomics 6:987–999, 2007.

Esophageal adenocarcinoma is increasing rapidly in Western countries and is currently the seventh leading cause of cancer-related death (1). Esophageal adenocarcinoma has been associated with the presence of Barrett metaplasia, a condition in which the normal squamous epithelium of the esophagus is replaced by columnar epithelium. The malignant potential of this condition is evidenced by the progression of Barrett metaplasia to low grade dysplasia, high grade dysplasia, and ultimately to invasive adenocarcinoma. The risk of developing adenocarcinoma is 30–125 times higher in people who have Barrett metaplasia than people who do not. The prognosis of patients with esophageal adenocarcinoma remains poor with overall 5-year survival rates of only 5–15% (1). Unfortunately patients often present with regionally advanced disease (2). Given the poor prognosis associated with esophageal adenocarcinoma, it is imperative to improve our understanding of the tumorigenesis and the factors associated with increased risk. It is possible that therapeutic targets or protein markers can be identified that will ultimately facilitate improved patient survival.

Proteomics technologies have been used for the identification of candidate markers for early cancer detection (3). The global analysis of protein expression complements genomics analyses. For example, proteomics analysis may provide further insight into post-translational modifications affecting cellular function that otherwise could not be identified by genomics analysis. It is important to identify changes in global protein expression to identify specific proteins that are involved in cancer-related processes. We and others have demonstrated that two-dimensional (2-D)1 liquid mass mapping can be used for quantitative and comparative proteomics analyses (4–6). Ion intensity-based quantitative approaches have progressively gained more popularity as mass spectrometry performance has improved significantly. 2-D fractionation techniques before mass detection simplify the complex proteome. Because mass is a unique tag for intact proteins, this method avoids the problem in quantitation induced by incomplete separation. In addition, liquid phase analysis allows easy interface to mass spectrometry analysis.

In this study, we evaluated protein expression differences to identify markers of disease progression of Barrett metapla-

1 The abbreviations used are: 2-D, two-dimensional; NPS-RP, nonporous reverse phase; CF, chromatofocusing; OG, octyl β-D-glucopyranoside; bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; RhoGDI, Rho GDP dissociation inhibitor.
Proteomics of Esophageal Adenocarcinoma

Changes that are associated with progression of Barrett metaplasia have utility for analysis of cellular protein expression, and the intact molecular weights were obtained after deconvolution of multiple charged peaks. Proteins were quantified based on individual molecular weight intensity and assembly of a protein mass map. Hierarchical clustering of the mass maps correctly segregated each of the 12 samples by histological type into Barrett metaplasia or esophageal adenocarcinoma. Differentially expressed proteins were identified by LC-MS/MS analysis by NPS-RP HPLC/ESI-TOF-MS and LC-MS/MS protein identification.

**EXPERIMENTAL PROCEDURES**

**Tissue Samples**—Patients seen by the Section of General Thoracic Surgery at the University of Michigan Hospital between May 1994 and July 2004 for resection of esophageal adenocarcinoma were evaluated for inclusion in this study. Patients receiving previous chemotherapy or radiation treatment were excluded. Consent was obtained from all patients, and the project was approved by the local Institutional Review Board. Medical records were reviewed, and data were coded to protect patient confidentiality. Tumors and adjacent Barrett metaplasia tissue were collected immediately at the time of surgery and briefly vortexed prior to agitation for 20 h at 37 °C. Hematoxylin-stained 5-μm frozen sections were reviewed by a board-certified pathologist (T. J. Giordano) with expertise in gastrointestinal pathology for tumor cellularity (adenocarcinomas) or mucosa (Barrett metaplasia). Specimens were excluded if tumor cellularity was less than 80% or if extensive lymphocytic infiltration or fibrosis was present.

**Cell Lysis and Buffer Exchange**—Barrett metaplasia or adenocarcinoma tissue samples were quickly thawed and immediately lysed with 2 ml of lysis buffer consisting of 7.5 M urea, 2.5 M thiourea, 4% n-octyl β-D-glucopyranoside (OG), 10 mM Tris(2-carboxyethyl)phosphine, 10% (v/v) glycerol, 50 mM Tris, and 40 μl of protease inhibitor solution (10-ng tablet in 1 ml of PBS buffer; Roche Applied Science). Samples were homogenized mechanically, vortexed frequently over a period of 1 h at room temperature, and then centrifuged at 30,000 rpm for 70 min at 4 °C. After the supernatant was collected, buffer exchange was conducted against the chromatofocusing start buffer using a PD-10 G-25 column (Amersham Biosciences). Protein concentration in each sample was quantified by the Bradford assay.

**Chromatofocusing**—Chromatofocusing (CF) separation was performed on an HPCF-1D column (250 × 2.1 mm) (Beckman-Coulter, Fullerton, CA) using a Beckman HPLC pump. The column was equilibrated with start buffer containing 25 mM bis-Tris propane, 6 M urea, and 1% OG adjusted to pH 7.4 with saturated iminodiacetic acid solution. After equilibration, 4.5 mg of total protein was loaded onto the CF column. Elution was achieved at pH 4.0 with elution buffer consisting of 10% (v/v) Polybuffer 74 (Amersham Biosciences), 6 M urea, and 1% OG at a flow rate of 0.2 ml/min. A linear pH gradient was generated at the outlet of the column that resulted in proteins eluting off the column according to their pl. Separation was monitored at 280 nm using a UV detector (Beckman-Coulter). The pH of the eluent was measured on line by a postdetector pH electrode/ cell (Lazar Research Laboratories, Los Angeles, CA) with low dead volume. Fractions were collected at every 0.2 pH unit from pH 7.0 to 4.0 and split for further analysis by NPS-RP HPLC/ESI-TOF-MS and LC-MS/MS/protein identification.

**NPS-RP HPLC and On-line ESI-TOF MS**—Fractions obtained from CF were subjected to NPS-RP HPLC separation using an ODSIII-E (4.6 × 33-mm) column (Beckman-Coulter) packed with 1.5-μm nonporous silica. The reverse phase separation was performed at 0.5 ml/min, and a postcolumn splitter was used so that the 200 μl/min flow was directed into an orthogonal acceleration ESI-TOF MS system (LCT; Micromass/Waters, Milford, MA). Formic acid (0.5%) was added after the splitter through a T-connector using a syringe pump. The other 0.3 ml/min flow was monitored at 214 nm using a 166 Model UV detector (Beckman-Coulter), and peaks were collected by an automated fraction collector (Model SC 100; Beckman-Coulter) controlled by in-house designed DOS-based software. To enhance the speed, resolution, and reproducibility of the separation, the reverse phase column was heated to 60 °C using a column heater (Model 7971; Jones Chromatography). Both mobile phase A (water) and mobile phase B (ACN) contained 0.1% (v/v) TFA. The gradient profile was as follows: 5–15% B in 1 min, 15–25% B in 2 min, 25–31% B in 3 min, 31–41% B in 10 min, 41–47% B in 3 min, 47–67% B in 4 min, and 67–100% B in 1 min. The capillary voltage for electrospray was set at 3200 V, the sample cone was set at 35 V, the extraction cone was set at 3 V, and the reflection lens was set at 750 V. Desolvation was accelerated by maintaining the desolvation temperature at 330 °C and the source temperature at 130 °C. The desolvation gas flow was 650–800 liters/h. 1 μg of bovine insulin was introduced to each sample as an internal standard. The intact molecular weight was obtained by deconvoluting the combined ESI spectra with MaxEnt1 software.

**Data Analysis and Clustering**—A mass map was generated by integrating the pl, M, and protein intensity of five CF fractions ranging from pH 4.6 to 5.6 into one single image using DeltaVue software (7). Maps with normalized protein value for Barrett tissue and the corresponding esophageal adenocarcinoma tissue from the same patient are shown at either side with a differential map of these two samples shown in the middle for comparison. The significance of protein expression differences was determined by χ² analysis.

After the mass maps of the pH 4.6–5.6 fractions from all 12 tissue samples were normalized based on insulin intensity and actin levels, the correlation of the mass maps was then visualized as a dendrogram by average-linked hierarchical clustering, which utilizes the mean distance between all possible pairs of entities of the two clusters. Pairs of samples are joined sooner if they have greater correlation. The length and the subdivision of the dendrogram branches reflect the relatedness of the 12 tissue samples based upon the expression of the individual proteins.

**Tryptic Digestion**—Protein fractions of interest from NPS-RP HPLC were concentrated down to ~20 μl with a SpeedVac concentrator (Labconco Corp., Kansas City, MO) operating at 60 °C. 20 μl of 100 mM ammonium bicarbonate with 5 mM DTT (Sigma-Aldrich) was then mixed with each concentrated sample to obtain a pH value of about 7.8. 0.5 μl of tosylphenylalanyl chloromethyl ketone-modified sequencing grade porcine trypsin (Promega, Madison, WI) was added and briefly vortexed prior to agitation for 20 h at 37 °C.

**Protein Identification by Capillary LC-MS/MS**—Digested peptide
mixture were separated by a capillary RP column (C18, 0.3 × 50 mm) (Michrom Bioresources, Auburn, CA) on Paradigm MG4 micropumps (Michrom Bioresources) with a flow rate of 5 μl/min. Both solvent A (water) and solvent B (ACN) contained 0.3% formic acid. The gradient was initiated at 5% solvent B, ramped to 60% solvent B in 25 min, and then finally ramped to 90% in 5 min. The resolved peptides were analyzed on an LTQ mass spectrometer with an ESI ion source (Thermo, San Jose, CA). The capillary temperature was 175 °C, spray voltage was 4.2 kV, and capillary voltage was 30 V. The normalized collision energy was set at 35% for MS/MS.

MS/MS spectra were searched using SEQUEST against the Swiss-Prot (8) human protein database. One miscleavage was allowed. Methionine oxidation was considered as a variable modification during the SEQUEST analysis. Peptide tolerance was set at 1.5 Da for DTA file search. The search results were filtered according to specific criteria: the Xcorr cutoffs were set as 1.9, 2.5, and 3.5 for 1+, 2+, and 3+ charged peptides, respectively, and the ΔCn cutoff was set as 0.1. Positive protein identification was validated by the Trans-Proteomics Pipeline. This software includes both the PeptideProphet and ProteinProphet programs that were developed by Keller et al. (9).

The MS/MS data from all the UV peak fractions collected from reverse phase separation were searched independently against the normal and reversed human International Protein Index (IPI) protein database (41,216 protein entries; ncrp.nl.gov/data/) using the SEQUEST algorithm (Thermo Finnigan, San Jose, CA) for evaluation of the false positive rate (10). The reversed protein database was created by reversing the order of amino acid sequences for each protein (the carboxyl terminus becomes the amino terminus and vice versa). The searching parameters were the same as above. The searching results were filtered according to specific criteria: the Xcorr cutoffs were set as 1.9, 2.5, and 3.5 for 1+, 2+, and 3+ charged peptides, respectively, and the ΔCn cutoff was set as 0.1.

**Analysis of mRNA Using Oligonucleotide Arrays**—For correlation of protein and RNA gene expression patterns, a range of non-dysplastic and dysplastic Barrett samples and esophageal adenocarcinomas were examined by high density oligonucleotide microarrays as described previously (11, 12). We compared mRNA expression of 46 samples, including nine Barrett metaplasia, seven samples with a mixture of Barrett metaplasia and low grade dysplasia, eight low grade dysplasia, seven high grade dysplasia, and 15 esophageal adenocarcinomas. Total RNA was isolated from 30 esophageal samples using TRizol (Invitrogen) and purified with RNeasy spin columns (Qiagen, Valencia, CA) according to the manufacturers’ instructions. RNA quality was assessed by 1% agarose gel electrophoresis and A260/A280 ratios. RNA quality was reassessed with the Agilent Bioanalyzer (Agilent Technologies, Palo Alto, CA) at intermediate steps after double-stranded cDNA and cRNA synthesis. Four samples were excluded due to insufficient quantity of RNA (<10 μg), cDNA synthesis, cRNA amplification, hybridization, and washing of the HG-U133A GeneChips (Affymetrix, Santa Clara, CA) were performed by the University of Michigan Cancer Center Microarray Core according to the manufacturer’s instructions.

To normalize the microarray data, a summary statistic was calculated using the 11 probe pairs for each gene and the robust multichip average method (13) as implemented in the Affymetrix library of Biocductor (version 1.3), which provides background adjustment, quantile normalization, and summarization. mRNA expression values for each sample were then compared with the mean expression value for the nine Barrett metaplasia samples. -Fold change greater than 2.0 was considered significant (14).

**Tissue Microarray and Immunohistochemistry**—A tissue microarray was created, as described previously (15), with formalin-fixed, paraffin-embedded tissues from 70 patients including 64 tumors, eight lymph node metastases, 11 dysplastic Barrett mucosa, 11 non-dysplastic Barrett metaplasia samples, and normal esophagus. 4-μm sections were transferred to poly(l-lysine)-coated slides, deparaffinized with xylene, rehydrated with a graded series of alcohols, and then rinsed with PBS. Antigen retrieval was performed using microwave pretreatment for 15 min in 0.01 M citrate buffer, pH 6.0. The slides were then blocked with 5% normal goat serum in PBS for 20 min, and endogenous peroxidase activity was quenched with 0.5% hydrogen peroxide in PBS. The slides were incubated with either a mouse anti-EN01 antibody at a 1:1000 dilution (Abnova Corp., New Taipei, Taiwan), rabbit anti-ARHGD1B antibody at a 1:50 dilution (Abnova Corp.), or anti-Lamin A/C (Chemicon International, Temecula, CA) at a 1:50 dilution for 1 h in 1.5% normal goat serum. After washing with PBS, the slides were incubated with a 1:1000 dilution of biotinylated goat anti-mouse antibody (Southern Biotechnology Associates, Birmingham, AL) or anti-rabbit antibody (Vector Laboratories Inc., Burlingame, CA) for 1 h and visualized using an avidin-biotin complex detection kit (VECTASTAIN ABC-Go kit; Vector Laboratories Inc.). A 0.1% solution of 3,3'-diaminobenzidine (Vector Laboratories Inc.) served as the chromagen. The slides were then counterstained with hematoxylin. Each sample was scored using a scale of 0, 1, 2, or 3 corresponding to absent, light, moderate, or intense staining. Samples without tumor or esophageal mucosa were excluded from analysis.

**RESULTS**

**Liquid Phase 2-D Separation and On-line ESI-TOF MS Analysis**—Tissue lysates containing 4.5 mg of protein each from the six premalignant Barrett metaplasia tissues and six esophageal adenocarcinomas (both from the same set of six patients to eliminate confounding genetic polymorphisms) were resolved individually in the first dimension using CF, which combines the high capacity of ion-exchange chromatography with the high resolution of IEF. The proteins elute off the column sequentially according to their pI. Fig. 1a shows the CF chromatogram of a representative esophageal adenocarcinoma extract measured at 280 nm. Approximately 60 min was required for protein to elute off the column linearly from pH 7 to 4. Fractions were collected at 0.2 pH intervals with most protein eluting out between pH 5.8 and 4.4.

The fractions of interest (pH 4.6–5.6) were further separated by NPS-RP HPLC at a flow rate of 0.5 ml/min. A postcolumn splitter was used so that 0.2 ml/min flow was directed into an orthogonal acceleration ESI-TOF MS instrument. The other 0.3 ml/min flow was directed through a UV detector at 214 nm, and the fractions were collected by UV peak for protein identification. Fig. 1b shows the UV chromatogram of an esophageal adenocarcinoma CF fraction (pH 5.2–5.0) at 214 nm. On-line ESI-TOF MS was used to obtain intact protein molecular weights. As shown in Fig. 1c, a series of multiple charged ion peaks in the combined spectrum having the same retention time as the circled UV peak in Fig. 1b were deconvoluted, and an intact mass was obtained (Fig. 1d). The theoretical mass of the protein, nucleoside-diphosphate kinase A, is 17,149 Da, and the experimental mass is 17,150 Da, which suggests a mass accuracy of 58 ppm. The abundance of each protein is indicated by peak area of the deconvoluted peak of the intact protein. Combined spectra of a selected ion
chromatogram were generated for each protein to obtain a composite of that present in each CF fraction. 1 μg of pure insulin was also added to each CF fraction as shown in Fig. 1b, and the deconvoluted peak area was used as internal standard. The flow rate was 0.5 ml/min. A postcolumn splitter was used so that 200 μl/min flow was directed into an orthogonal acceleration ESI-TOF MS instrument. Formic acid (0.5%) was added after the splitter through a T-connector using a syringe pump. The other 0.3 ml/min flow was monitored at 214 nm, and the fractions were collected by UV peak. c, combined multiple charged ion spectrum detected by ESI-TOF MS of intact protein in the circled peak in b. d, intact Mr was obtained after deconvolution. e, a representative 2-D mass map of Barrett metaplasia (red, left), differential mass map (center), and a mass map of esophageal adenocarcinoma (green, right) in pl range 4.6–5.6 and mass range 5–75 kDa are displayed.

2-D Mass Map Generation and Comparison—A mass map was created by integrating the Mr, pl, and the relative abundance of all proteins in all CF fractions into one single image using DeltaVue software. As shown in Fig. 1e, a mass map was generated by integrating the pl, Mr, and abundance of five CF fractions from pH 4.6 to 5.6. The map for Barrett tissue is shown on the left, and the map for esophageal adenocarcinomas from the same patient is shown on the right. The y axis represents intact Mr, and the x axis indicates the pH range of the CF fractions. The abundance of each protein is indicated by the intensity of each band on this map.

To test the dynamic range and ion abundance variance of our approach, different amounts of bovine pancreatic Ribonuclease A were spiked into 30 μg of whole cell lysate of an ovarian cancer cell line. The theoretical mass of Ribonuclease A is 13,690 Da with the actual mass detected as 13,686 Da. The correlation of protein abundance and its corresponding molecular weight ion intensity is shown in Fig. 2. As indicated, the Mr ion intensities increased linearly with the protein abundance at 1:1 ratio when the protein abundance fell in the range of 0.001–2.0 μg. Above 2 μg, the ions were saturated; thus the intensity will not increase linearly. No ion signals were detected when the protein level fell below 0.001 μg. This might be due to limitations in instrument sensitivity and/or data processing software. From the ion intensities, the individual protein abundance in a CF fraction usually fell in the 0.01–1 level. This result indicates that the dynamic range is ~2000, thereby facilitating our ability to semiquantify protein abundance in complex protein mixtures.
As shown in Table I, the standard deviation of total ion intensities of the six individual samples of esophageal tissues is 15.3% of the averaged value. For specific ions, the variance percentage in the six individual samples is 13.2% for the band at 26,867 with an averaged intensity of 2277. For proteins with higher abundance, the variance percentage is 8.2% for the band at 32,716 and 5.7% for the band at 54,318 with an averaged intensity of 8301 and 48,638, respectively. As indicated, lower ion variance was achieved for the higher abundant proteins. We also examined the ion variance by repeating the analysis of the same amount of whole cell lysate from the same ovarian cancer cells (OVCA 432 cell line) five times using LC-ESI-TOF. As shown in Table II, the averaged ion variance is 5.1% for the \( M_r \) ions with intensities higher than 10,000 and 9.2% for intensities lower than 10,000. For the total ion intensities, the variance is 5.8%. The ion variance for the cell line sample was lower than that of the tissue samples because it was not affected by individual variance. The mass maps of three individual experiments are shown in Supplemental Fig. 1. Very similar patterns were observed especially with the high intensity \( M_r \) bands, thus demonstrating good reproducibility of the approach.

There are \( \sim 300 \) bands present for each sample with the dynamic range of observable protein banding patterns being approximately 3 orders of magnitude. A differential map is shown in the middle of Fig. 1e that indicates the differentially expressed proteins between the Barrett metaplasia and the corresponding adenocarcinoma sample from the same patient. Both bovine insulin and actin were used for normalization. Insulin was injected with each sample as an internal control and thus to correct for experimental variation. The total intensity of the actin isoforms, which is shown in the \( M_r \) range of 41,600–41,900, was then used for normalization of sample amounts and for the protein assay to load similar amounts of protein from each sample.

**Clustering**—The mass maps of all pH 4.6–5.6 fractions from the 12 tissue samples were normalized based on insulin intensity and actin levels. Correlation of the mass maps was assessed using hierarchical clustering. Fig. 3a shows a cluster map.
illustrates the average linkage clustering of six Barrett tissue samples and the corresponding esophageal adenocarcinoma samples. All 12 tissue samples segregated into two major clusters according to their histological type, confirming that differences in protein expression may be used to characterize the neoplastic progression from Barrett metaplasia to adenocarcinoma.

**Average Linkage**

![Hierarchical clustering analysis](image)

**Fig. 3.** Hierarchical clustering analysis. 12 tissue samples are grouped based on similarities in their protein expression in a mass map using a hierarchical clustering analysis technique. t, esophageal adenocarcinoma; p, Barrett metaplasia.

Identification of Differentially Expressed Proteins by Capillary LC-MS/MS—Protein bands demonstrating differential expression between Barrett metaplasia and esophageal adenocarcinoma were of most interest. For these proteins, the eluents from the UV peak collection of the reverse phase separation having the same retention time were selected and subjected to tryptic digestion. The proteins were identified by peptide sequencing utilizing capillary LC-MS/MS. Confident identification was indicated if matches were obtained for both the $M_r$ and retention time of the intact protein. Fig. 4, a and b, shows two MS/MS spectra of nucleoside-diphosphate kinase A (P15531) corresponding to the digest of the circled UV peak in Fig. 1b and having the same retention time of the circled band in the mass map shown in Fig. 1e. The theoretical mass of the identified protein is 17,149 Da, matching the experimental mass of the 17,150 Da protein band in the differential map. As indicated in the mass map, this protein demonstrated increased expression in the adenocarcinoma tumor samples. For some proteins, differences between the experimental $M_r$/pI and the theoretical value were noted. This may be due to post-translational modifications such as truncation and/or protein phosphorylation (16). Some high abundance proteins were also present in more than one CF fraction. For example, the pI of ATP synthase D chain is 5.2, but it was also detected from pH 5.0 to 5.6 as shown in Fig. 5. To quantify these high

**Fig. 4.** MS/MS spectrum of two peptides from nucleoside-diphosphate kinase A (P15531), which is the circled band in Fig. 1b, using μLC-MS/MS. Fragments of peptide YMHSGPVVAM-VWEGLNVK at 1058.32 (2+) (A) and VMLGETNPADSKPTIR at 893.29 (2+) (B) are shown.
abundance proteins accurately, the peak area of each protein in each CF fraction was added together to estimate the total amount of the protein in each tissue sample. We identified 38 proteins differentially expressed in the six differential protein mass maps that were analyzed. Protein identities, abundance, and the significance of differential expression are listed in Table III and displayed graphically in Fig. 5. The peptide sequences identified for each protein are listed in the supplemental table. It should be noted, however, that as most proteins are modified they may have unpredictable intact masses. Furthermore many proteins may not ionize effectively as an intact molecule, whereas their tryptic peptides do ionize. Thus, assignment of protein identifications to specific intact masses should be considered tentative, pending additional verification.

Positive protein identification was validated by the TransProteomics Pipeline, which includes both PeptideProphet and ProteinProphet software (9). PeptideProphet automatically validates peptide assignment to MS/MS spectra made by a database search program such as SEQUEST. For each dataset, it calculates the distribution of search scores and peptide properties among correct and incorrect peptides and uses those distributions to compute for each obtained peptide sequence a probability that it is correct. ProteinProphet takes the peptides and search results and statistically validates the identifications at the protein level. For each differentially expressed protein, the probability that the database search result is correct, as determined by PeptideProphet and ProteinProphet, is shown in the supplemental table. The sensitivity and error rate are affected by the probability threshold, which is for distinguishing the correct and incorrect identifications. Nesvizhskii et al. (17) had reported previously that \( p \geq 0.7 \) has a sensitivity level of 94% and a false positive error rate of 1.2%. A typical graph of estimated sensitivity and error for different thresholds is shown in Supplemental Fig. 2. The functions are not very smooth for these data due to the small number of proteins in each dataset. Because the analyzed fractions are collected according to UV peak, which corre-
| No. | Acc. no. | M_p (pl) (theo) | Barrett Tumor | mRNA | Gene expression | No. | No. | p-value | symbol | mRNA Prot/mRNA |
|-----|---------|-----------------|---------------|-------|----------------|-----|-----|---------|---------|---------------|
| 1   | P05709  | 10.17/35.3      | -             |       |               | 1   | 1   | 0.054  | S100A6  | 1/10002/11001 |
| 2   | Q99467  | 13.36/41.9      | -             |       |               | 2   | 2   | 0.025  | U2B2    | 1/11002/11001 |
| 3   | Q99468  | 13.12/42.5      | -             |       |               | 3   | 3   | 0.025  | E2M       | 1/11002/11001 |
| 4   | P09224  | 11.74/51.3      | -             |       |               | 4   | 4   | 0.017  | U2A2    | 1/11001/11001 |
| 5   | P09251  | 11.67/50.4      | -             |       |               | 5   | 5   | 0.032  | U2F       | 1/11001/11001 |
| 6   | P09280  | 11.72/51.0      | -             |       |               | 6   | 6   | 0.003  | U2G       | 1/11001/11001 |

**Table III**

Differentially expressed proteins between Barrett metaplasia and esophageal adenocarcinoma tissue lysates and their frequency among 12 samples.

**Abbreviations:** Acc, accession; theo, theoretical; ID, identity; LCT, ESI-TOF MS; Prot, protein; ATL, adult T-cell leukemia; snRNA, small nuclear RNA; E2, ubiquitin carrier protein; N, no, Y, yes.
The probability of most selected proteins is 1, which means that the identifications of these differentially expressed proteins are expected to be correct. All of the proteins reported in Table III have a probability higher than 0.95.

The use of the reversed protein database for assessing the false positive rates of peptide identifications resulting from SEQUEST searching has been reported previously for yeast proteome, human plasma, and other samples (10, 18–20). Peptides passing the filtering parameters derived from the reversed database (Nrev) are defined as a false positive. The peptides identified from the normal database (Nnor) contain both true and false positives. Therefore, by dividing the number of peptides found from the reversed database by the number of identified peptides from normal databases according to the formula Nrev/Nnor, the overall false positive rate can be estimated as reported by Qian et al. (10). The false positive rate for all 38 fractions was calculated and listed corresponding to the identified differentially expressed proteins in the supplemental table. The average false positive rate of the peptide identification for these fractions is 0.065, which is reasonable with this Xcorr and Cn cutoff.

Expression of mRNA and Localization of Candidate Proteins—Proteins demonstrating altered expression in the progression from metaplasia to adenocarcinoma may play an important role in the esophageal adenocarcinoma tumorigenesis. We evaluated the mRNA expression levels for the selected proteins as determined by high density Affymetrix oligonucleotide microarrays. Several proteins identified in Table III (notably Rho GDP dissociation inhibitor 2 (RhoGDI-2; ARHGDIB), -enolase (ENO1), Lamin A/C (LMNA), and nucleoside-diphosphate kinase A (NME1)) demonstrated increased mean mRNA expression in esophageal adenocarcinomas relative to Barrett metaplasia with individual tumors showing higher expression (Fig. 6). RhoGDI-2 plays an essential role in control of a variety of cellular functions through interactions with Rho family GTPases, and it is overexpressed on the protein level in several tumor types, including ovarian (21) and breast carcinomas (22). Additionally, overexpression of RhoGDI is associated with increased resistance of cancer cells to the induction of drug-induced apoptosis (23). As shown in Table III, our result indicates that both RhoGDI-1 (P52565) and RhoGDI-2 (P52566) expression is increased in esophageal cancer and that Rho GTPases might participate in apoptosis of these cancers. By immunohistochemistry, using an antibody to RhoGDI-2, we observed relatively low protein expression in Barrett metaplasia or in dysplastic Barrett mucosa but much higher levels of cytoplasmic expression in esophageal adenocarcinomas (Fig. 7, a–d).

α-Enolase (P06733) was increased at the protein and mRNA level in esophageal adenocarcinomas (Table III and Fig. 6c) and was found by immunohistochemistry to be more...
highly expressed in the cytoplasm of a subset of tumors (Fig. 7) as compared with Barrett metaplasia with or without dysplasia (Fig. 7). Overexpression of α-enolase has been reported previously in cervical (24) and colon (25) cancer cell lines and in breast cancer (26). This gene product likely reflects the increased glycolysis in cancer cells. Although attempts to examine nucleoside-diphosphate kinase (NME1) by immunochemistry were not successful we found both protein and mRNA expression (Table III and Fig. 6a) to be increased in esophageal tumors. As shown in Table III and Fig. 6a we observed increased NME1 in esophageal adenocarcinomas. NME1 is involved in the synthesis of nucleoside
triphosphates and may regulate signal transduction by complexing with G proteins, causing activation/inactivation of developmental pathways. In our analyses the phosphorylated form of NME1 was not detected. One plausible reason may be that phosphorylated NME1 has a reactive intermediate having a rapid turnover and may be lost during analysis. This protein appears to have distinct, if not opposite, roles in different tumors. It is reported as reduced in tumors with increased metastatic potential (27) but increased in aggressive neuroblastoma (28). This protein was found to have increased expression in the esophageal adenocarcinomas relative to Barrett metaplasia, and its expression might be related to increased cell proliferation in these cancers.

Similarly Lamin A was found to be increased (Table III) in esophageal adenocarcinoma relative to Barrett metaplasia (Fig. 6d). A specific antibody directed against just Lamin A was not available. However, analysis of esophageal tissues using Lamin A/C antibodies demonstrated a very strong perinuclear localization that showed variable expression in individual cells in the Barrett mucosa but generally increased and more uniform expression in the esophageal adenocarcinomas (Fig. 7, e–h). The increase in the expression of this protein may reflect increased neoplastic cell proliferation.

Tropomyosin 1 (P09493) and tropomyosin 2 (P07951) were found to have decreased expression in esophageal adenocarcinomas, consistent with reports of down-regulation of high M<sub>r</sub> tropomyosin isoforms observed in breast cancer (29) and transformed cells (30). Multiple tropomyosin protein isoforms are known to bind actin filaments in muscle and non-muscle human cells (31). The down-regulation of key cytoskeletal proteins such as tropomyosin 1 might reflect or contribute to the formation of aberrant cytoskeleton and account at least in part for the metastatic potential of cancer cells (32).

Finally, we identified several additional interesting candidates showing differential protein expression but with no significantly different mRNA expression. Calgranulin B (S100A9, P06702), a member of the large family of S100 calcium-binding proteins, has been reported as overexpressed in poorly differentiated adenocarcinomas (33). In our study, both Calgranulin B and Calgranulin A (S100A8, P05109) were detected in esophageal adenocarcinoma samples but not in Barrett metaplasia by liquid separation mass mapping. The mRNA levels for their associated genes were not demonstrably increased in esophageal adenocarcinoma, demonstrating the potential for these combined techniques to identify proteins subject to post-translational rather than transcriptional regulation in these cancers.
regulation. Conversely it is known that both Calgranulins A and B are expressed by inflammatory cells; thus increased protein expression found in the tumors may be reflective of the inflammatory nature of esophageal adenocarcinomas.

DISCUSSION

We utilized liquid phase separation of proteins incorporating chromatofocusing in the first dimension and NPS-RP HPLC in the second dimension followed by ESI-TOF mass spectrometry to identify multiple proteins differentially expressed in the progression from Barrett metaplasia to esophageal adenocarcinoma. In our analysis, we used six Barrett metaplasia samples and six esophageal adenocarcinoma samples; all six Barrett samples were obtained from the identical six patients from whom we obtained the esophageal adenocarcinoma tissue, thereby eliminating possible genetic polymorphisms between preneoplastic and neoplastic tissues as a confounding variable. This increased expression was verified at the messenger level and further confirmed by immunohistochemistry for several of these proteins, notably Rho GDP dissociation inhibitor 2 and α-enolase. This approach of candidate identification with subsequent verification has strong potential for identifying candidate protein markers and mechanisms involved in the tumorigenesis of esophageal adenocarcinoma. In contrast to large scale analysis of transcriptional regulation, this combined approach has the additional capability to identify post-translationally modified proteins.

We identified several such candidates, including Calgranulin B, whose mRNA levels did not differ significantly in adenocarcinoma compared with Barrett metaplasia samples.

Furthermore with our clustering analysis we observed that both metastatic and neoplastic tissues segregated into two large groups, recapitulating their designated histology. This is reflective of the differential protein expression between Barrett metaplasia and esophageal adenocarcinoma that we demonstrated for several individual proteins. Although our findings require extensive validation in larger independent sets of tissue samples, this high throughput approach may have broad potential for tumor identification and classification as we have demonstrated previously in other tissue types (34) and as others have suggested for a variety of proteomics strategies (35, 36).

Currently detection of esophageal cancer, especially adenocarcinoma arising in Barrett esophagus, relies on histological examination of multiple endoscopic biopsies obtained at regularly spaced intervals in a segment of abnormal-appearing esophagus. Given the poor prognosis associated with this disease, a better understanding of the pathogenesis of the disease is essential. Utilization of 2-D mass maps with corresponding protein identification can provide detailed analyses of cellular protein expression changes that are associated with esophageal adenocarcinoma progression. These identified proteins may have utility as candidate markers of progression from Barrett metaplasia to esophageal adenocarcinoma.

Acknowledgment—We thank Alexey Nesvizhskii for help on the Protein/PeptideProphet software.

* This work was supported in part by the NCI, National Institutes of Health Grant CA71606 (to D. G. B.) and National Institutes of Health Grants R01CA100101 (to D. M. L.), R01CA90503 (to D. M. L.), and R01GM49500 (to D. M. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

§§ To whom correspondence should be addressed: Dept. of Surgery, University of Michigan, Ann Arbor, MI 48109. Tel.: 734-647-8834; Fax: 734-615-2088; E-mail: dmlubman@umich.edu.

REFERENCES

1. Jemal, A., Thomas, A., Murray, T., and Thun, M. (2002) Cancer statistics, 2002. CA Cancer J. Clin. 52, 23–47
2. Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., Smigal, C., and Thun, M. J. (2006) Cancer statistics, 2006. CA Cancer J. Clin. 56, 106–130
3. Tyers, M., and Mann, M. (2003) From genomics to proteomics. Nature 422, 193–197
4. Lubman, D. M., Kachman, M. T., Wang, H., Gong, S., Yan, F., Hamler, R. L., O’Neill, K. A., Zhu, K., Buchanan, N. S., and Barder, T. J. (2002) Two-dimensional liquid separations-mass mapping of proteins from human cancer cell lysates. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 782, 183–196
5. Hamler, R. L., Zhu, K., Buchanan, N. S., Kreunin, P., Kachman, M. T., Miller, F. R., and Lubman, D. M. (2004) A two-dimensional liquid-phase separation method coupled with mass spectrometry for proteomic studies of breast cancer and biomarker identification. Proteomics 4, 562–577
6. Zhao, J., Zhu, K., Lubman, D. M., Miller, F. R., Shekhar, M. P., Gerard, B., and Barder, T. J. (2006) Proteomic analysis of estrogen response of premalignant human breast cells using a 2-D liquid separation/mass mapping technique. Proteomics 6, 3847–3861
7. Yan, F., Sreekumar, A., Laxman, B., Chinmeyan, A. M., Lubman, D. M., and Barder, T. J. (2003) Protein microarrays using liquid phase fractionation of cell lysates. Proteomics 3, 1228–1235
8. Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., Martin, J. M., Michoud, K., O’Donovan, C., Phan, I., Pilbout, S., and Schneider, M. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucelic Acids Res. 31, 365–570
9. Kelleher, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. 74, 5383–5392
10. Qian, W. J., Liu, T., Monroe, M. E., Strittmatter, E. F., Jacobs, J. M., Kangas, L. J., Petritis, K., Camp, D. G., III, and Smith, R. D. (2005) Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. J. Proteome Res. 4, 53–62
11. Lin, J., Raoof, D. A., Thomas, D. G., Greenson, J. K., Giordano, T. J., Robinson, G. S., Bourner, M. J., Bauer, C. T., Orringer, M. B., and Beer, D. G. (2004) α-type amino acid transporter-1 overexpression and melanophan sensitivity in Barrett’s adenocarcinoma. Neoplasia 6, 74–84
12. Miller, C. T., Lin, L., Casper, A. M., Lim, J., Thomas, D. G., Oriinger, M. B., Chang, A. C., Chambers, A. F., Giordano, T. J., Glover, T. W., and Beer, D. G. (2006) Genomic amplification of MET with boundaries within fragile site FRA7G and upregulation of MET pathways in esophageal adenocarcinoma. Oncogene 25, 409–418
13. Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003) Summaries of Affymetrix GeneChips probe level data. Nucleic Acids Res. 31, e15
14. Risinger, J. I., Laghi, A., Madaras, G. V., Jazayeri, A., Aprilekovic, O., Patterson, T., Berchuck, A., and Barrett, J. C. (2003) Microarray analysis reveals distinct gene expression profiles among different histologic types of endometrial cancer. Cancer Res. 63, 6–11
15. Kononen, J., Budendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tu
mor specimens. Nat. Med. 4, 844–847
16. Zhu, K., Zhao, J., Lubman, D. M., Miller, F. R., and Barder, T. J. (2005) Protein pl shifts due to posttranslational modifications in the separation and characterization of proteins. Anal. Chem. 77, 2745–2755
17. Nevzihski, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75, 4646–4656
18. Cargile, B. J., Bundy, J. L., and Stephenson, J. L., Jr. (2004) Potential for false positive identifications from large databases through tandem mass spectrometry. J. Proteome Res. 3, 1082–1085
19. Fiol, T., Rahamim, K., Radulovic, D., Cox, B., Rossant, J., and Emili, A. (2003) PRISM, a generic large scale proteomic investigation strategy for mammals. Mol. Cell. Proteomics 2, 96–106
20. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J. Proteome Res. 2, 43–50
21. Jones, M. B., Krutzsch, H., Shu, H., Zhao, Y., Liotta, L. A., Kohn, E. C., and Petriconi, E. F., III (2002) Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. Proteomics 2, 76–84
22. Fritz, G., Just, I., and Kaina, B. (1999) Rho GTPases are over-expressed in human tumors. Int. J. Cancer 81, 682–687
23. Zhang, B., Zhang, Y., Dagher, M. C., and Shacter, E. (2005) Rho GDP dissociation inhibitor protects cancer cells against drug-induced apoptosis. Cancer Res. 65, 6054–6062
24. Lee, K. A., Shim, J. H., Kho, C. W., Park, S. G., Park, B. C., Kim, J. W., Lim, J. S., Choe, Y. K., Paik, S. G., and Yoon, D. Y. (2004) Protein profiling and identification of modulators regulated by the E7 oncogene in the C33A cell line by proteomics and genomics. Proteomics 4, 839–848
25. Sterum, R., Gaspari, M., Dommels, Y., Ouatas, T., Pluk, H., Jespersen, S., Vogels, J., Verhoeckx, K., Groten, J., and van Ommen, B. (2003) Proteome analysis reveals novel proteins associated with proliferation and differentiation of the colorectal cancer cell line Caco-2. Biochim. Biophys. Acta 1650, 73–91
26. Durany, N., Joseph, J., Jimenez, O. M., Climent, F., Fernandez, P. L., Rivera, F., and Carreras, J. (2000) Phosphoglycerate mutase, 2,3-bisphosphoglycerate phosphatase, creatine kinase and enolase activity and isoenzymes in breast carcinoma. Br. J. Cancer 82, 20–27
27. Rosengard, A. M., Krutzsch, H. C., Shearn, A., Biggs, J. R., Barker, E., Margulies, I. M., King, C. R., Liotta, L. A., and Steeg, P. S. (1989) Reduced Nm23/Awd protein in tumour metastasis and aberrant Drosophila development. Nature 342, 177–180
28. Halil, N., Keim, D. R., Melhem, R. F., Zhu, X. X., Eckerskorn, C., Brodeur, G. M., Reynolds, C. P., Seeger, R. C., Lottsche, F., Strahler, J. R., and Hanash, S. (1991) High levels of p19/nm23 protein in neuroblastoma are associated with advanced stage disease and with N-myc gene amplification. J. Clin. Investig. 88, 341–345
29. Raval, G. N., Bharadwaj, S., Levine, E. A., Willingham, M. C., Geary, R. L., Kute, T., and Prasad, G. L. (2003) Loss of expression of tropomyosin-1, a novel class II tumor suppressor that induces anoikis, in primary breast tumors. Oncogene 22, 6194–6203
30. Hendricks, M., and Weintraub, H. (1981) Tropomyosin is decreased in transformed cells. Proc. Natl. Acad. Sci. U. S. A. 78, 5633–5637
31. Giometti, C. S., and Anderson, N. L. (1984) Tropomyosin heterogeneity in human cells. J. Biol. Chem. 259, 14113–14120
32. Bharadwaj, S., Thanawala, R., Bon, G., Falcioni, R., and Prasad, G. L. (2005) Resensitization of breast cancer cells to anoikis by Tropomyosin-1: role of Rho kinase-dependent cytoskeleton and adhesion. Onco- gene 24, 8291–8303
33. Arai, K., Teratani, T., Kuruto-Niwa, R., Yamada, T., and Nozawa, R. (2004) S100A9 expression in invasive ductal carcinoma of the breast: S100A9 expression in adenocarcinoma is closely associated with poor tumour differentiation. Eur. J. Cancer 40, 1179–1187
34. Wang, Y., Wu, R., Cho, K. R., Shedden, K. A., Barder, T. J., and Lubman, D. M. (2005) Classification of cancer cell lines using an automated two-dimensional liquid mapping method with hierarchical clustering techniques. Mol. Cell. Proteomics 5, 43–52
35. Aliyama, A. A., Fanzen, B., Hagman, A., Dysvik, B., Roblick, U. J., Becker, S., Moberger, B., Auer, G., and Linder, S. (2002) Molecular classification of borderline ovarian tumors using hierarchical cluster analysis of protein expression profiles. Int. J. Cancer 98, 895–899
36. Yanagisawa, K., Shyr, Y., Xu, B. J., Massion, P. P., Larsen, P. H., White, B. C., Roberts, J. R., Edgerton, M., Gonzalez, A., Nadaf, S., Moore, J. H., Caprioli, R. M., and Carbone, D. P. (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet 362, 433–439

Molecular & Cellular Proteomics 6.6 999

Proteomics of Esophageal Adenocarcinoma