CMLS, Cell. Mol. Life Sci. 61 (2004) 1246–1255
1420-682X/04/101246-10
DOI 10.1007/s00018-004-4049-4

CMLS Cellular and Molecular Life Sciences

Research Article

Sequential proteome alterations during genesis and progression of colon cancer

U. J. Roblick[a, b, *], D. Hirschberg[c], J. K. Habermann[a, d], C. Palmberg[c], S. Becker[a], S. Krüger[e], M. Gustafsson[f], H.-P. Bruch[b], B. Franzén[a, g], T. Ried[d], T. Bergman[c], G. Auer[a] and H. Jörnvall[c]

[a] Department of Oncology and Pathology, Unit of Cancer Proteomics, Karolinska Institutet and Hospital, 171 76 Stockholm (Sweden), Fax: +46 8 321047, e-mail: Uwe.Roblick@cck.ki.se
[b] Department of Surgery, University of Schleswig-Holstein, Campus Lübeck Medical School (Germany)
[c] Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm (Sweden)
[d] Department of Genetics Branch, NIH Center for Cancer Research, NCI/NIH, Washington, Bethesda Maryland (USA)
[e] Institute of Pathology, University of Lübeck, Medical School, Lübeck (Germany)
[f] Gyros AB, Uppsala Science Park, 751 83 Uppsala (Sweden)
[g] Astra Zeneca AB, Preclinical R&D, 151 85 Södertälje (Sweden)

Received 2 February 2004; received after revision 16 March 2004; accepted 18 March 2004

Abstract. Changes in the proteome of colon mucosal cells accompany the transition from normal mucosa via adenoma and invasive cancer to metastatic disease. Samples from 15 patients with sporadic sigmoid cancers were analyzed. Proteins were separated by two-dimensional gel electrophoresis. Relative differences in expression levels between normal tissue, adenoma, carcinoma and metastasis were evaluated in both intra- and inter-patient comparisons. Up- and down-regulated proteins (> two fold) during development to cancer or metastasis were excised and submitted to peptide mass fingerprinting and MS/MS sequence analysis, facilitated by the use of a compact disc workstation. In total, 112 protein spots were found to be differentially regulated, of which 72 were determined as to protein identity, 46 being up-regulated toward the progression of cancer, and 26 down-regulated. Several of the identifications correlate with proteins of the cell cycle, cytoskeleton or metabolic pathways. The pattern changes now identified have the potential for design of marker panels for assistance in diagnostics and therapeutic strategies in colorectal cancer.

Key words. Proteomics; 2-D gel electrophoresis; mass spectrometry; tumor marker; colon cancer.

Colorectal carcinoma is the fourth most common malignancy in the western world (approximately one of every eighth malignant tumor in the USA). It ranks third as the cause of death from carcinoma, surpassed only by lung and prostate neoplasms in men, and lung and breast cancers in women. Multiple mutations have been suggested to lead to a progression from normal epithelium to metastatic carcinoma through hyperplastic epithelium, adenoma, carcinoma and metastasis, known as the 'adenoma-carcinoma sequence' [1]. In 65–70%, tumors are detected as advanced-stage cancer located in the sigmoid colon. As in most malignant diseases, early diagnosis and especially detection of metastases is of importance for patient prognosis. Settlement of tumor cells in the liver is the major cause of cancer-associated deaths [2]. Presently, clinical parameters combined with histopathological staging and grading are the most important diagnostic and prognostic variables. The evaluation of carcinoembryonic antigen (CEA) in serum has not fulfilled the promise of a simple test that would offer an early di-

* Corresponding author.

agnosis of colon cancer. A number of other, less well explored potential markers exist but are currently not used in routine clinical diagnosis [3–5]. Therefore, more extensive proteome tests are desirable for diagnosis, prognosis evaluation and monitoring recurrent disease.

We describe comprehensive proteome analyses of normal mucosa, adenoma, cancer and distant liver metastasic cells recovered from the same patients using two-dimensional gel electrophoresis (2-DE), image and principal component analysis (PCA), and mass spectrometry of samples purified via compact disc (CD) technology to study progression-specific protein expression profiles.

Materials and methods

Sample collection

Colon samples were obtained immediately after resection from the Lübeck University Hospital (with ethical permit). Cells were collected from the surface of non-necrotic tumor or metastasic tissue, normal mucosa and polyps by scraping with a scalpel, and were transferred into 2–5 ml ice-cold RPMI-1640 medium containing 5% fetal calf serum and 0.2 mM phenylmethylsulfonyl fluoride/0.83 mM benzamidine. This was followed by aspiration and squirting steps with a syringe and a 29-gauge needle to preferentially release the tumor cells, which are less attached to each other than the connective-tissue cells. A two-phase nylon filter (pore sizes of 250 and 100 µm) was used to catch remaining stromal components and to allow the passage of the tumor cells. These enriched cell suspensions, underlaid with 2 ml ice-cold Percoll (54.7% in PBS), were centrifuged at 1000 g for 10 min at +4 °C. Cells at the interface were collected and washed twice with PBS. This method is rapid and results in pure tumor cell populations, free from serum proteins, red blood cells, connective tissue and necrotic tissue material [6]. After preparation, each sample was quality-checked by comparison of Giemsa-stained smears with histological slides. Only samples with more than 95% tumor cells were used.

Characterization of formalin-fixed specimens

Histopathological characterization was carried out using hematoxylin-eosin-stained sections of formalin-fixed and paraffin-embedded specimens. Polyps, tumors and liver metastases were characterized according to size, lymph node status and site of metastasis. Histopathological classification followed WHO criteria (TNM classification).

Cytochemical analysis

Nuclear DNA content was assessed in all disease specimens by image cytometric analysis of Feulgen-stained cells. Samples with a single stemline in the normal diploid region (1.9c–2.1c) were classified as diploid and those with pronounced scattered DNA values and a stemline outside the diploid or tetraploid regions were classified as aneuploid [7, 8].

Protein quantification

Protein concentrations of samples were determined by addition of 25 µl concentrated assay reagent (Bio-Rad) to 1 µl solubilized sample diluted in 100 µl Milli-Q water using 96-well microplates [9]. A standard curve was constructed using different concentrations of bovine serum albumin. The plate was read using a Multiscan reader (Labsystems).

2-D gel electrophoresis

This was performed as described elsewhere [6, 10]. Before application, all samples were diluted to 500 µl containing 7 M urea, 2 M thiourea, 1% 3-(3-chloramidopropyl) dimethylammonio-1-propanesulfonate (CHAPS), 0.4% immobilized pH gradient (IPG) buffer, 0.3% dithiothreitol (DTT) and a trace of bromophenol blue. Total cell extracts were applied via active rehydration on precast IPG strips (pH 4–7, 17 cm; Bio-Rad) in sample solution for isoelectric focusing. Protein (75 µg/IPG strip) was loaded and focused in PROTEAN IEF Cells (Bio-Rad) for ~52,900 Vh overnight. After isoelectric focusing (IEF), the strips were immediately equilibrated for 2 × 15 min with 50 mM Tris-HCl, pH 8.8, in 6 M urea, 30% glycerol and 2% SDS. DTT (2%) was included in the first and iodoacetamide (2.5%) in the second equilibration step to reduce and alkylate free thiols. SDS gels with a 10–13% linear acrylamide gradient (1.5 × 200 × 230 mm) were used for the second dimension. The IEF strips were carefully applied on top of the gels and fixed in position using warm agarose (0.5%) dissolved in SDS/PAGE running buffer. Electrophoresis was carried out overnight at 42,000 Vh. For image analysis, gels were stained with silver nitrate [11, 12]. Preparative gels utilized more material (two- to tenfold more) and spots were then visualized using Coomassie Brillant Blue, Sypro Ruby (Bio-Rad) or Silver Plus staining (Bio-Rad).

Scanning, image and statistical analysis

2-D gels were scanned at 84.7 × 84.7 µm resolution using a GS-710 imaging densitometer (Bio-Rad). Data were analyzed using PDQuest software from Bio-Rad [13]. A PCA was performed with 60 gels to screen for outliers and clusters [14]. Spot data from the PDQuest gel analysis package were subsequently transferred to the Spotfire statistical software (DecisionSite for Functional Genomics; www.spotfire.com). PCA is effective for identifying discriminating features in a data set by finding two or three linear combinations of the original features that best summarize the variation in the data. If much of the variation is captured by these two or three

most significant principal components, group membership of many data points can be observed.
Using the PDQuest 7.1 software, polypeptide spots from the different stages of the disease were matched to spots in the reference pattern. Spot intensities were normalized, background was subtracted, peaks were located and the relative staining intensities were determined. Several known proteins served as landmarks to facilitate the gel matching. Spots were identified using boolean analysis and the Mann-Whitney test ($p < 0.05$). A high-level match set comparing the clinical cancer samples with two colorectal cancer cell lines (diploid HCT 116 and aneuploid Lovo cell lines) was also constructed, as were gels from two normal liver samples, recovered from the liver metastases resection specimens. Only those protein spots that indicated up- or down-regulation (> twofold) in the metastases samples were considered for identification.

In-gel digestion and CD technology for preparation of tryptic digests

Protein spots were excised manually from Coomassie-stained gels and in-gel digested [15] using a MassPREP robotic protein-handling system (Micromass). Gel pieces were destained twice with 100 µl 50 mM ammonium bicarbonate (Ambic)/50% (v/v) acetonitrile at 40°C for 10 min. Pieces containing protein were reduced by 10 mM DTT in 100 mM Ambic for 30 min, shrunk in acetonitrile, and the proteins were then alkylated with 55 mM iodoacetamide in 100 mM Ambic for 20 min. Trypsin (25 µl of a 12 ng/µl solution in 50 mM Ambic) was added and incubation was carried out for 4.5 h at 40°C. Peptides were extracted with 30 µl 5% formic acid/2% acetonitrile followed by extraction with 24 µl 2.5% formic acid/50% acetonitrile. The acetonitrile was evaporated under atmospheric pressure overnight at 10°C. For electrospray (ES) ionization MS/MS, the peptide extracts were desalted with C_{18} ZipTips (Millipore), activated and equilibrated using 10 µl 70% acetonitrile/0.1% trifluoroacetic acid (TFA) twice, 10 µl 50% acetonitrile/0.1% TFA twice, and finally 10 µl 0.1% TFA twice. The sample was loaded onto the ZipTip by pipetting 20 times and washed using 10 µl 0.1% TFA twice. The tryptic fragments were eluted with 60% acetonitrile/1% acetic acid.
Samples with proteins in low yield were analyzed using a Gyrolab MALDI SP1 Workstation (Gyros AB). In this approach, 96 microcolumns (packed with a C_{18} resin to a volume of 10 nl) are incorporated into a CD platform and used for desalting by reverse-phase chromatography. The columns are conditioned with 50% acetonitrile in water. The samples are loaded onto the columns, and solvents passed through by spinning of the disc. The wash solution (200 nl 0.1% TFA) is directed to a waste exit. Peptides are eluted from the columns using 200 nl 50% acetonitrile containing 1 mg/ml α-cyano-4-hydroxycinnamic acid matrix and 0.1% TFA. The eluate is captured in an open MALDI target area of 200 × 400 µm for solvent evaporation and the peptide/matrix crystallization. For on-CD MALDI analysis, the cut CD was accommodated in the target compartment of the MALDI instrument.

Mass spectrometry

The tryptic fragments were mass analyzed by matrix-assisted laser desorption ionization (MALDI) mass spectrometry (Voyager DE-PRO; Applied Biosystems) and, where relevant, also ES ionization quadrupole time-of-flight (Q-TOF) tandem MS (Micromass) for sequence information. Samples for MALDI analysis were mixed at a 1:1 (v/v) ratio with a saturated α-cyano-4-hydroxycinnamic acid solution in 50% acetonitrile/0.1% TFA. Database searches were carried out using the MS-Fit search program (http://prospector.ucsf.edu/). Only protein hits with three or more matching peptide masses were considered, and only if the assigned protein was human, and the theoretical pI and M_r values did not deviate excessively from the values observed in the 2-D gel separation.
Samples for the ES analysis were introduced via gold-coated nano-ES needles (Protana). A capillary voltage of 800–1000 V was applied together with a cone voltage of 40–45 V and collision energy of 4.2 eV. The sample aerosol was desolvated in a stream of nitrogen. During the collision-induced dissociation (CID), the collision energy was in the range of 15–30 eV with argon as collision gas.

Results

Samples, histopathology and ploidy analysis

Colon samples from more than one histological specimen (N for normal, P for adenoma/polyp, C for cancer/tumor, M for metastasis) were analyzed by 2-D gel electrophoresis, subsequent PDQuest evaluation and PCA from 15 patients with sporadic adenocarcinomas. In addition, tissues from single lesions were analyzed from many patients (200 in total). Data and histopathological diagnosis are provided in table 1.
The eight polyps included in the study were pedunculated tubular adenomas localized within the colon resection specimen. All polyps had a head size smaller than 1 cm diameter. Two polyps showed a high-grade dysplasia, the other six were low-grade dysplastic. Ploidy assessment showed pronounced scattered DNA values exceeding the tetraploid region in the two polyps with high-grade dysplasia, in all of the carcinomas and in the metastasis samples. These samples were classified as aneuploid. The remaining six polyps were diploid. For the normal mucosal tissue, diploidy was presumed.

Intra- and inter-tumor/metastasis heterogeneity

In two tumor (CR97T and CR151T) and two metastasis (CR69M and CR175M) samples, we took two separate

Table 1. Samples.

Sample					Patient		
	normal	adenoma	cancer	metastasis	sex	age	TNM
1			CR64T	CR65M	M	76	T3N2M1
2	CR95N		CR94T	CR96M	M	81	T3N2M1
3	CR98N		CR97T		F	78	T3N1Mx
4	CR107N	CR105P	CR106T		F	71	T3N0Mx
5	CR125N		CR124T		F	65	T3N2Mx
6	CR132N		CR131T		M	59	T3N1Mx
7	CR135N	CR133P	CR134T		M	77	T4N1Mx
8	CR139N	CR138P*	$CR137T_1$, $136T_2$		M	78	T4N1Mx
9	CR149N	CR150P	CR148T		M	60	T3N0Mx
10	CR152N		CR151T		F	71	T3N0Mx
11	CR158N		CR157T		M	76	T3N1Mx
12	CR160N		CR159T		F	75	T4N1Mx
13	CR164N	$CR162P_1$, $163P_2$*	CR161T		M	61	T3N0Mx
14	CR174N	CR173P	CR172T		M	74	T3N1Mx
15	CR211N	CR210P	CR209T		M	76	T3N1Mx

N, normal mucosa; P, adenoma/polyp; T, cancer/tumor; M, metastasis from 15 patients. For database allocation they are labeled as CR (colorectal) followed by a number and N, P, T, M. For the sample type were included and analyzed with more than one-stage samples per patient. In addition to the samples from 15 patients where more than one sample type could be obtained, metastases only were includes for 5 patients (CR21M, CR68M, CR69M, CR102M, CR175M). TNM classification according to WHO guidelines. The asterisks denote polyp material that was high-grade dysplastic and aneuploid.

preparations from different areas within the same neoplasia (designated A and B in fig. 1) to evaluate the intra-tumor and intra-metastasis heterogeneity. In these comparisons, we were able to match at least 93% of all polypeptide spots detected. The average correlation coefficient was 0.86 within the tumor and 0.80 within the metastasis samples, reflecting their close relatedness. In analogy, visual inspection showed very similar expression patterns (fig. 1). The correlation coefficients between pairs of primary tumors gave an average r value of 0.70 (0.58–0.73). For pairwise comparison of liver metastases, the average correlation coefficient was 0.57 (0.51–0.61). In the patient from whom normal tissue, primary tumor and liver metastases were available, we calculated a correlation coefficient of 0.55 between normal mucosa and the tumor, and 0.76 between the tumor and its metastases.

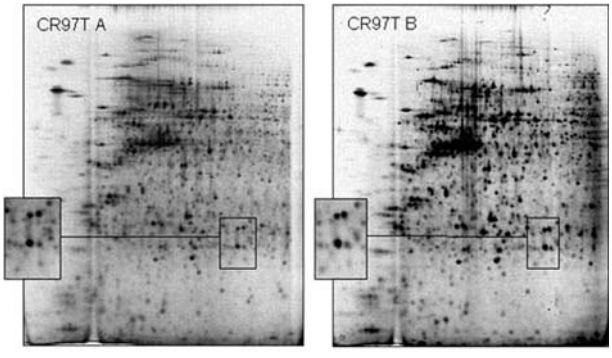

Figure 1. Comparison of two carcinoma samples (CR97T; A and B) from different locations within the same tumor. Close-up areas reflect the relatedness and relative homogeneity.

Principal component analysis

As an initial analytical step, we applied a PCA to the entire data set to identify outliers and clusters. Using spotfire statistical analysis, the relationship of the four colonic histological subtypes can then be visualized in a three-dimensional PCA plot. The PCA results showed a close clustering of the normal mucosa samples (fig. 2). In the neighborhood of the normal samples, six of the eight polyps also clustered. The two outlying polyps clustered closer to the tumor cohort, and were identified as the polyps with high-grade dysplasia and aneuploidy mentioned above (fig. 2), proving compatibility between the histopathological and PCA-proteomics grouping modes. Nevertheless, a pathological re-appraisal of those two cases did not lead to carcinoma diagnosis. The tumor samples formed a widespread, separate cluster in the three-dimensional PCA plot (fig. 2). The subset of liver metastases clustered separately, without overlap with the primary tumors.

Identification of the deviating spots in normal versus cancer samples

Using IPG strips of pH 4–7, an average of 1675 spots (range 1543–1866 spots/gel) could be resolved. The resolution of the polypeptides showed better quality in the low-molecular-mass area and toward the acidic side of the gels (fig. 1). Increased streaking and precipitation on the basic side is a well-known phenomenon observed in 2-D gel electrophoresis. The number of cells used was not determined since the analysis was not performed with cell line material. However, in parallel work with cell lines, we find about 1 million cells useful to obtain high-

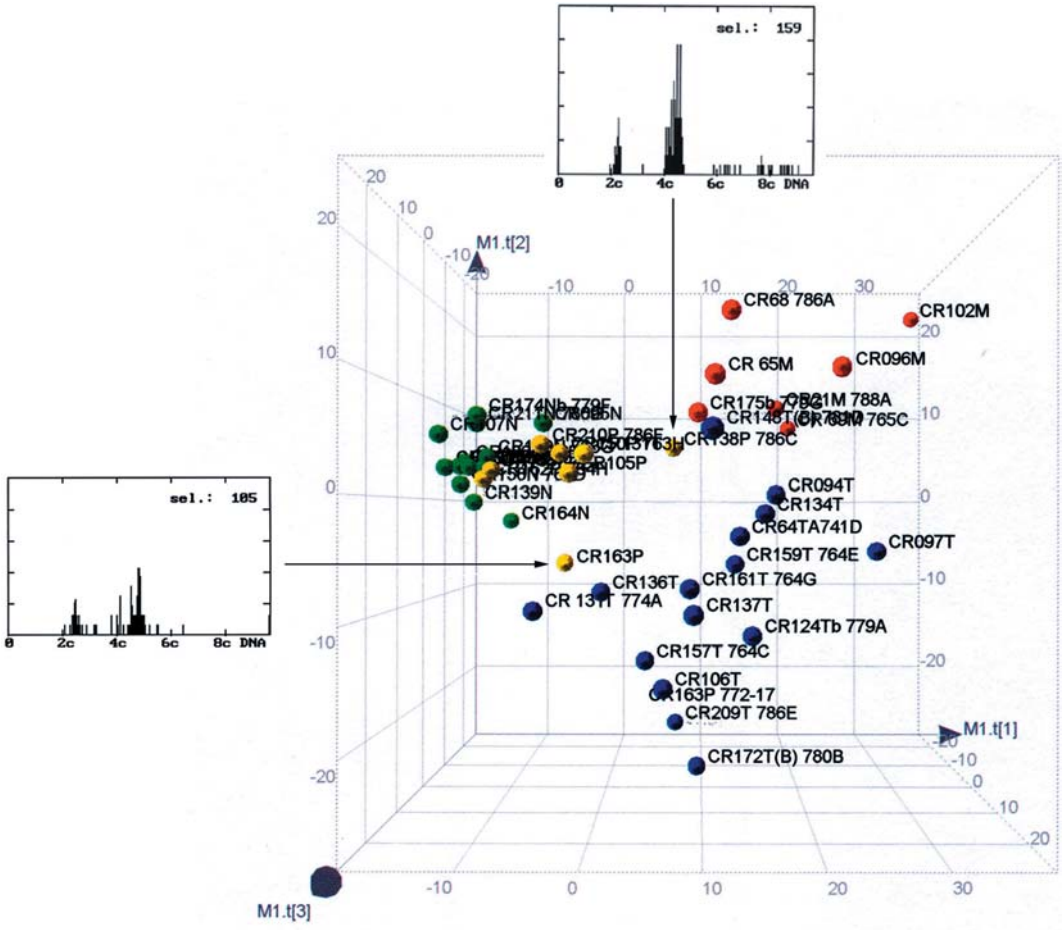

Figure 2. PCA plot of the protein expression data of all match set members, with the normal cohort (green), polyps (yellow), tumors (blue) and metastases (red). The arrows highlight two polyps that were outliers, showed aneuploidy in the DNA measurement (shown in the histograms) and clustered closer to the tumors in the three-dimensional space.

quality gels with corresponding spot numbers. Focusing on proteins that matched to all members of the match set and on polypeptides whose level was significantly up- or down-regulated from benign disease to cancer development (fig. 3), we identified 112 spots deviating by at least a twofold change in expression as judged by staining intensity in the PDQuest analysis template. Furthermore, within the carcinoma samples, we considered only those spots for further evaluation that were co-expressed in at least one of the colorectal cancer cell lines.

Of the up- and down-regulated proteins, nine high-abundant forms could be identified via gel matching toward our HCT 116 database. For identification of remaining focus, we performed preparative separations using Coomassie Brilliant Blue staining. We could detect and prepare 64 for analysis. They were cut out from the Coomassie-stained preparative gels, the corresponting proteins were in-gel digested with trypsin, analyzed by MALDI MS and, in several cases, also analyzed for sequence by tandem MS (fig. 4). All proteins identified and the proof for their identifications are summerized in table 2. Table 2 also shows the extent of sequence coverage for each spot. Identification of spot 32, determined both by MALDI and tandem MS is exemplified in figure 4. For the different preparations, three adjacent but equally expressed proteins with known identities served as internal controls and landmarks. Sixty-four samples were run on the Gyros SP-1 workstation. Of these samples, 58 proteins were identified, compared to 39 when the same sam-

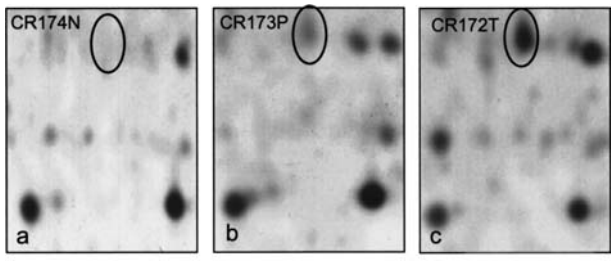

Figure 3. Intra-individual expression differences of cytokeratin 20 in patient 14. The left gel segment **a** is zoomed from the normal mucosa; **b** represents a polyp in the same patient, and **c**, the corresponding segment of the adenocarcinoma.

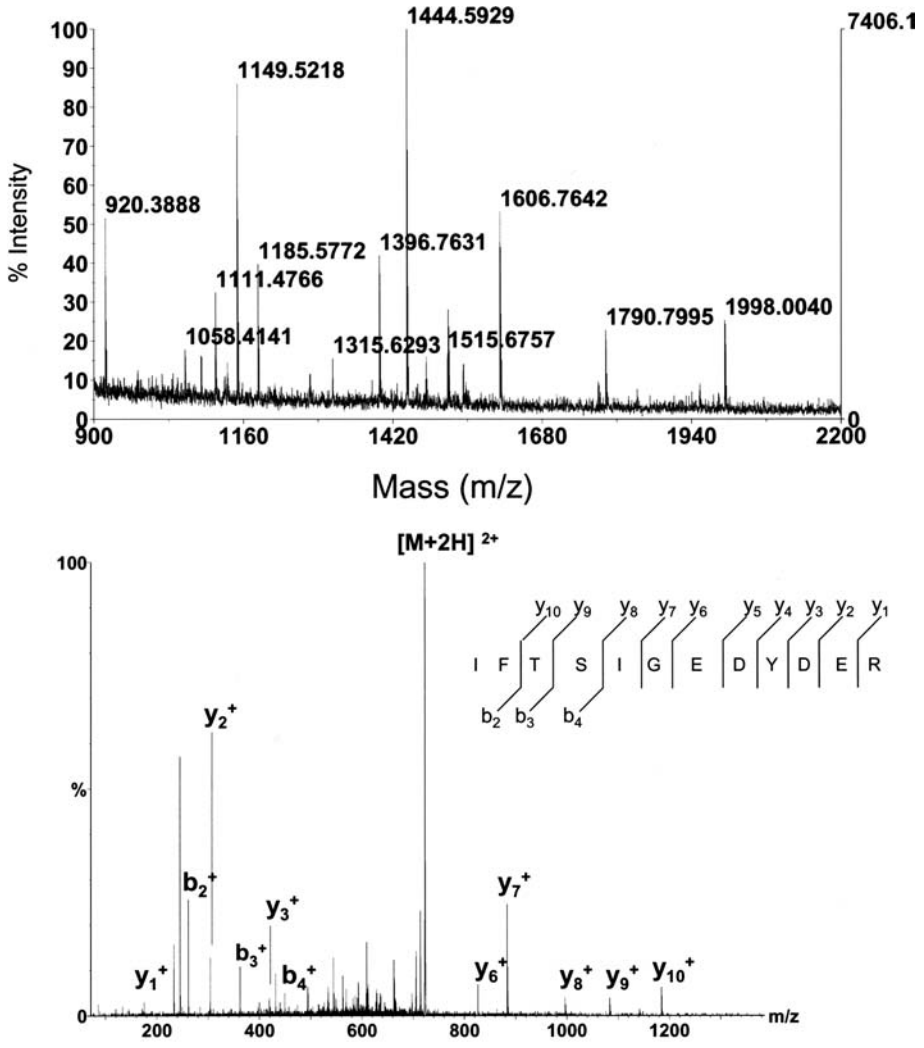

Figure 4. Mass spectrometric characterization of spot 32 (fig. 5). From MALDI MS spectrum (top) of the tryptic peptide mixture, database searches of the masses of the tryptic peptides identified the protein as prohibitin. Tandem MS via collision-induced dissociation of the doubly charged fragment with m/z 722.8 (1444.6 Da in MALDI MS) confirmed the identification from the peptide mass fingerprinting.

Table 2. Proteins identified and supporting data.

Spot	Protein	SPDB	Mass	pI	Coverage	MS/MS
1	keratin, type II cytoskeletal 7	P08729	51	5.5	18	
2	keratin, type I cytoskeletal 18	O75575	48	5.3	12	
3	tropomyosin 1 alpha chain	P09493	33	4.7	22	
4	tropomyosin 1 alpha chain	P09493	33	4.7	30	
5	reticulocalbin 1 precursor	P08729	39	4.9	29	
6	cytochrome b5	O43169	81	4.8	27	13–24
7	keratin, type I cytoskeletal 19	P08727	44	5	28	
8	keratin, type II cytoskeletal 7	P08729	51	5.5	24	
9	actin, cytoplasmic 1	P02570	42	5.3	27	
10	2′,3′-cyclic nucleotide 3′-phosphodiesterase	P09543	48	9.2	14	
11	serum albumin precursor	P02768	69	5.9	26	
12	annexin V	P08758	36	4.9	40	227–244
13	actin, cytoplasmic 2	P02571	42	5.3	25	
14	protein disulfide isomerase	P07237	57	4.8	21	
15	keratin, type I cytoskeletal 19	P08727	44	5	56	
16	endoplasmin precursor	P14625	32	4.8	24	

Table 2 (continued)

Spot	Protein	SPDB	Mass	pI	Coverage	MS/MS
17	keratin, type I cytoskeletal 19	P08727	44	5	57	
18	keratin, type I cytoskeletal 19	P08727	44	5	35	
19	keratin, type I cytoskeletal 19	P08727	44	5	36	
20	ATP synthase beta chain	P06576	57	5.3	30	
21	class I beta-tubulin	P09203	50	4.8	20	
22	class I beta-tubulin	P09203	50	4.8	20	
23	keratin, type I cytoskeletal 18	P05783	48	5.3	27	
24	vimentin	P08670	54	5.1	39	
25	ATP synthase D chain	O75947	18	5.2	51	123–143
26	keratin, type I cytoskeletal 19	P08727	44	5	56	
27	actin, cytoplasmic 1	P02570	42	5.3	26	
28	actin, cytoplasmic 1	P02570	42	5.3	20	
29	keratin, type II cytoskeletal 8	P05787	54	5.3	27	
30	endoplasmin precursor	P14625	32	4.8	18	
31	heat shock cognate 71 kDa	P11142	71	5.4	17	
32	prohibitin	P35232	30	5.6	31	106–117
33	actin, cytoplasmic 2	P02571	42	5.3	22	
34	F-actin capping protein alpha-1 subunit	P52907	33	5.4	39	38–47
35	tubulin alpha-6/1 chain	Q9BQE3	50	5	27	65–79
36	78-kDa glucose-regulated protein precursor	P11021	72	5.1	30	
37	actin, cytoplasmic 1	P02570	42	5.3	25	
38	keratin, type II cytoskeletal 8	P05787	54	5.5	35	
39	keratin, type I cytoskeletal 18	P05783	48	5.3	28	
40	keratin, type II cytoskeletal 8	P05787	54	5.5	46	
41	CYPIA/PAPS synthetase-2	O95340	70	8.2	11	
42	keratin, type II cytoskeletal 8	P05787	54	5.5	39	
43	keratin, type I cytoskeletal 20	P35900	48	5.5	49	
44	lamin B2	Q03252	68	5.3	46	
45	60-kDa heat shock protein	P10809	61	5.7	30	182–493 + 430–446
46	5-lipoxygenase	P09917	78	5.5	17	
47	actin, cytoplasmic 1	P02570	42	5.3	15	
48	putative ATP-dependent Clp protease	Q16740	30	8.3	33	215–226
49	tubulin alpha-6 chain	Q9BQE3	50	5	10	65–79
50	actin, cytoplasmic 2	P02571	42	5.3	26	
51	actin, cytoplasmic 1	P02570	42	5.3	25	
52	actin, cytoplasmic 2	P02570	42	5.3	34	
53	serum albumin precursor	P02768	69	5.9	26	
54	serum albumin precursor	P02768	69	5.9	42	
55	heat shock 27 kDa	P04792	23	6	51	28–37
56	3,2-trans-enoyl-CoA isomerase	P42126	39	8.8	21	271–283
57	T-complex protein 1, beta subunit	P78371	57	6	23	
58	DnaJ homolog subfamily B member 11	Q9UBS4	41	5.8	32	207–217
59	DnaJ homolog subfamily A member 2	Q60884	46	6.1	28	
60	hepatocellular carcinoma-associated antigen 66	Q9NYH9	70	7.2	21	
61	keratin, type II cytoskeletal 7	P08729	51	5.5	13	
62	proliferation-associated protein 2G4	Q9UQ80	44	6.1	37	333–344
63	succinate dehydrogenase	P31040	73	7.1	16	
64	elongation factor 1-delta	P29692	31	4.9		
65	heat shock cognate 71-kDa protein	P11142	71	5.4		
66	26s protease regulatory subunit	Q03527	49	5.7		
67	heat shock cognate 71-kDa protein	P11142	71	5.4		
68	glucose-regulated protein precursor	P11021	72	5.1		
69	eucaryotic translation initiation factor	O95065	50	5.8		
70	vimentin	P08670	54	5.1		
71	heat shock cognate 71-kDa protein	P11142	71	5.4		
72	ezrin (P81) cytovillin	P15311	70	6		

Material from 72 spots (shown in fig. 4) was identified as given in the first two columns. Corresponding accession number (Swissprot database, SPDB), precursor subunit mass (kDa) and precursor PI are given in the next three columns. Sixty-three identifications were by MALDI mass fingerprinting of the tryptic fragments, in which case the penultimate column gives the extent of sequence coverage (in percent). Thirteen cations were by sequence analysis using collision-induced dissociation in tandem MS (MS/MS), in which case the column gives the start and end position of each sequence determined. Spots 64–72 were identified by comparison with gel databases.

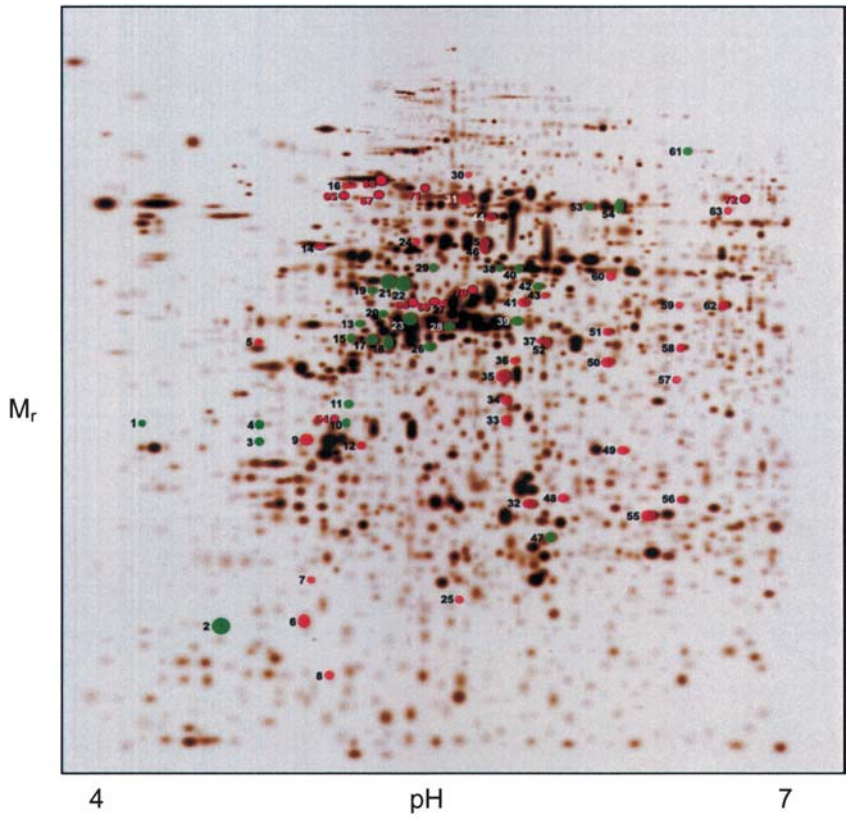

Figure 5. 2-D electrophoresis gel pattern showing all the spots identified (1–72). Spots in green define those that were overexpressed in the benign subentities. Spots in red define those that were significantly overexpressed in the malignant stages . The nine red spots with red numbers were identified by gel-to-gel matching.

ples were run manually. In total, and including all analyses, we identified no less than 72 polypeptides revealing expression differences during the progress of the disease. The identified 26 proteins with higher expression in normal and adenomatous lesions than in the primary neoplasia and metastases, and the remaining 46 proteins with an increased expression in malignant disease stages are all listed in table 2 and displayed gel-positioned in figure 5. The database entries regarding mass and pI in table 2 are given for orientation but refer to the precursor forms as listed in the database, whereas the actual positions as shown in figure 5 apply to the mature proteins identified. Differences between database and actual mass and pI values are due to modifications (such as truncations, phosphorylations and glycosylations) of the mature polypeptide chains.

Discussion

Remarkable advances in colon cancer prevention and therapy are consequences of early detection, improved therapy and a better understanding of colorectal carcinogenesis. There are several genetic events in colon cancer development that affect proteins involved in biological processes during tumor formation. By generating comprehensive proteomic expression profiles of normal mucosa, polyps, primary tumors and metastases, we traced expression changes corresponding to colon cancer development and progression at both intra- and inter-patient levels consistent with the prevailing adenoma-carcinoma paradigm of tumor progression [1]. Thus, we were able to detect the changes of 112 proteins at the proteome level and to identify 72 of those polypeptides (table 2).

We first analyzed qualitative and quantitative electrophoretic data to gain basic information concerning heterogeneity in primary colon tumors and metastases. When cell populations sampled from different areas within the same primary tumor and metastasis were analyzed, our average correlation level for the tumors was 0.86 and 0.80 in the metastases. We found a high degree of intra-tumoral homogeneity which is in accordance with studies describing average correlation coefficients of 0.85 in ovarian and 0.81 in breast carcinomas [16]. As all our tumor samples and metastases represented advanced disease stages, this homogeneity could be due to overgrowth by a dominating cell clone [17], in a manner similar to the situation at chemotherapy when response and resistance properties are related to the dominating tumor cell population [18].

PCA was used to identify outliers in our four histologically defined subgroups. The PCA data support the quality of the preparation technique, which has been validated in breast tumor samples of different histological subtypes [6, 16]. Taking into account that the background data of protein spots from more than 200 patients were processed by PCA analysis, the discrimination of the four colonic subgroups based on the protein expression data is notable. Furthermore, our polypeptide expression data reflect the closer relatedness of the benign-sample cohorts (normal mucosa, polyps) than of the malignant primary carcinomas and metastases (wider spread). Interestingly, the two dysplastic polyp samples did not conformingly cluster in their cohort and were closer in location to the malignant diseases. DNA content analysis revealed a scattered distribution pattern, characteristic for aneuploid cells, indicating an increased malignancy potential. This supports the theory that adenomas should be regarded as a potentially malignant disease entity in themselves rather than just a presymptomatic transitional lesion [19].
To identify the proteins responsible for the subentity expression groups, we searched for proteins up- or down-regulated at least twofold. This was done by evaluating a match set consisting of proteins from 15 patients, derived from normal mucosal tissue, polyps, primary carcinomas and liver metastases. We succeeded in identifying 72 of 112 consistently, deviating proteins. Of 72 identified polypeptides that showed significant expression differences in the 2-DE evaluation, 46 proteins were found to be exclusively over-expressed in tumors and metastases.
The deviating proteins identified (table 2, fig. 5) include several functional groups. One group includes cell-cycle-related proteins, e.g. elongation factor-1, proliferation-associated protein 2G4 and translation initiation factor, reflecting the proliferation status of the neoplasms. Another group includes proteins involved in metabolic pathways, e.g. 3,2-trans-enoyl-CoA isomerase, succinate dehydrogenase and 5-lipooxygenase, representing an increased metabolic activity within the malignancies [16]. A third group, with increased expression levels in the carcinoma and metastasis samples, was identified as members of the heat shock (HS) family, e.g. HS cognate 71 and HS protein (HSP) 27. HSPs are abundant intracellular polypeptides rapidly up-regulated by various stressors (e.g. temperature, infection and cancer). Their release is considered as a putative danger signal because they help to maintain the structural and metabolic integrity of cells. Several reports using therapeutic cancer vaccine based on HSPs for renal cancer and melanoma immunotherapy have been published [20–24]. A fourth group with expression changes now detected consists of structural proteins. We found a shift in cytokeratin expression from CK7, CK8, CK18 and CK19 (high in normal mucosa and polyps) to an over-expression of CK20 in the cancer samples (table 2, fig. 5). This is in accordance with immunohistochemical data where CK7 and CK20 are used as markers to distinguish primary ovarian carcinoma (CK7+/CK20–) and colonic cancer (CK7–/CK20+) metastatic to the ovary [25]. In addition, we now observed a lower expression of actin 1, actin 2 and tropomyosin 1 in the malignant cells than in the normal mucosa and the polyps. This transformation is in agreement with results in cell lines following Fos-induced deregulation of genes encoding microfilament-associated proteins [26]. A down-regulation of tropomyosin isoforms 1–3 has also been observed in breast cancer tissue [16]. Regarding the expression differences between primary carcinomas and liver metastases, we detected a significant difference for the hepatocellular-carcinoma-associated antigen 66. This 70-kDa protein is a member of a family of proteins also over-expressed in different carcinoma entities, like testis, lung, hepatocellular carcinoma, pancreas, prostate and colon cancer cell lines, but not in the corresponding normal tissues [27].
We conclude that 2-D-gel-electrophoresis-based proteomics combined with MS is an efficient tool for biomarker distinction in colon cancer material. The present study reveals a panel of well-defined markers with diagnostic potential. Although these markers will need validation before they can be used in a clinical setting, they indicate that cellular proteome patterns could be used for non-subjective diagnosis and to design adjusted individual therapies.

Acknowledgements. Grants from the Swedish Cancer Society, the Cancer Society Stockholm, the Swedish Research Council, Wallenberg Consortium North, Knut and Alice Wallenberg Foundation, the Medical University of Schleswig-Holstein, Campus Lübeck (N14-2001), Karolinska Institutet and the Swedish Institute are gratefully acknowledged.

1 Fearon E. R. and Vogelstein B. (1990) A genetic model for colorectal tumorigenesis. Cell **61:** 759–767
2 Sporn M. B. (1996) The war on cancer. Lancet **347:** 1377–1381
3 Yeatman T. J. and Chambers A. F. (2003) Osteopontin and colon cancer progression. Clin. Exp. Metastasis **20:** 85–90
4 Wallace H. M. and Caslake R. (2001) Polyamines and colon cancer. Eur. J. Gastroenterol. Hepatol. **13:** 1033–1039
5 Srivastava S., Verma M. and Henson D. E. (2001) Biomarkers for early detection of colon cancer. Clin. Cancer Res. **7:** 1118–1126
6 Franzen B., Linder S., Okuzawa K., Kato H. and Auer G. (1993) Nonenzymatic extraction of cells from clinical tumor material for analysis of gene expression by two-dimensional polyacrylamide gel electrophoresis. Electrophoresis **14:** 1045–1053
7 Auer G. U., Caspersson T. O. and Wallgren A. S. (1980) DNA content and survival in mammary carcinoma. Anal. Quant. Cytol. **2:** 161–165
8 Auer G. U., Arrhenius E., Granberg P. O. and Fox C. H. (1980) Comparison of DNA distributions in primary human breast cancers and their metastases. Eur. J. Cancer **16:** 273–278
9 Bradford M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:** 248–254

10 Okuzawa K., Franzen B., Lindholm J., Linder S., Hirano T., Bergman T. et al. (1994) Characterization of gene expression in clinical lung cancer materials by two-dimensional polyacrylamide gel electrophoresis. Electrophoresis **15:** 382–390
11 Morrissey J. H. (1981) Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. Anal. Biochem. **117:** 307–310
12 Rabilloud T., Vuillard L., Gilly C. and Lawrence J. J. (1994) Silver-staining of proteins in polyacrylamide gels: a general overview. Cell. Mol. Biol. **40:** 57–75
13 Garrels J. I. (1989) The QUEST system for quantitative analysis of two-dimensional gels. J. Biol. Chem. **264:** 5269–5282
14 Jocliffe I. T. (1986) Principal Component Analysis, Springer, New York
15 Oppermann M., Cols N., Nyman T., Helin J., Saarinen J., Byman I. et al. (2000) Identification of foetal brain proteins by two-dimensional gel electrophoresis and mass spectrometry comparison of samples from individuals with or without chromosome 21 trisomy. Eur. J. Biochem. **267:** 4713–4719
16 Franzen B., Linder S., Alaiya A. A., Eriksson E., Uruy K., Hirano T. et al. (1996) Analysis of polypeptide expression in benign and malignant human breast lesions: down-regulation of cytokeratins. Br. J. Cancer **74:** 1632–1638
17 Ponten F., Berg C., Ahmadian A., Ren Z. P., Nister M., Lundeberg J. et al. (1997) Molecular pathology in basal cell cancer with p53 as a genetic marker. Oncogene **15:** 1059–1067
18 Kern D. H. (1998) Heterogeneity of drug resistance in human breast and ovarian cancers. Cancer J. Sci. Am. **4:** 41–45
19 Villavicencio R. T. and Rex D. K. (2000) Colonic adenomas: prevalence and incidence rates, growth rates, and miss rates at colonoscopy. Semin. Gastrointest. Dis. **11:** 185–193
20 Chiosis G., Lucas B., Shtil A., Huezo H. and Rosen N. (2002) Development of a purine-scaffold novel class of Hsp90 binders that inhibit the proliferation of cancer cells and induce the degradation of Her2 tyrosine kinase. Bioorg. Med. Chem. **10:** 3555–3564
21 (2002) Cancer vaccine-antigenetics. BioDrugs **16:** 72–74
22 Nawrocki S., Murawa P., Malicki J., Kapcinska M., Gryska K., Izycki D. et al. (2000) Genetically modified tumour vaccines (GMTV) in melanoma clinical trials. Immunol. Lett. **74:** 81–86
23 Nicchitta C. V. (2003) Re-evaluating the role of heat-shock protein-peptide interactions in tumour immunity. Nat. Rev. Immunol. **3:** 427–432
24 Rakonczay Z. Jr, Takacs T., Boros I. and Lonovics J. (2003) Heat shock proteins and the pancreas. J. Cell. Physiol. **195:** 383–391
25 Berezowski K., Stastny J. F. and Kornstein M. J. (1996) Cytokeratins 7 and 20 and carcinoembryonic antigen in ovarian and colonic carcinoma. Mod. Pathol. **9:** 426–429
26 Jooss K. U. and Muller R. (1995) Deregulation of genes encoding microfilament-associated proteins during Fos-induced morphological transformation. Oncogene **10:** 603–608
27 Wang Y., Han K. J., Pang X. W., Vaughan H. A., Qu W., Dong X. Y. et al. (2002) Large scale identification of human hepatocellular carcinoma-associated antigens by autoantibodies. J. Immunol. **169:** 1102–1109