Comparative proteomic analysis of nucleic acid-binding proteins in ten human tumor cell lines

LEILA AFJEHI-SADAT[1], EPHREM ENGIDAWORK[1,2], IRENE SLAVC[1] and GERT LUBEC[1]

[1]Department of Pediatrics, Medical University of Vienna, Vienna, Austria; [2]Department of Pharmacology, School of Pharmacy, Addis Ababa University, Addis Ababa, Ethiopia

Received July 6, 2005; Accepted August 5, 2005

Abstract. Failure in regulation of genes involved in growth and division of cells may result in pathological conditions, particularly cancer. Regulation is exerted at various levels of the transcriptional and post-transcriptional processes involving mainly nucleic acid-binding proteins. Here, we systematically explored the proteome of ten different cell lines in search for proteins potentially serving as molecular markers and/or targets for monitoring prognostic outcome and clinical therapies. High-throughput analysis, two-dimensional electrophoresis coupled to matrix-assisted laser desorption/ionization mass spectrometry, identified 72 nucleotide-binding proteins and their interacting partners, which were differentially expressed in the cell lines investigated. Out of the 72 identified proteins, 33 of them were specifically expressed in a single cell line (for e.g., replication protein A 32 kDa, transcription intermediary factor 1, heterogeneous ribonucleoproteins). Moreover, tumor-related proteins including breast carcinoma amplified sequence 2, zinc finger proteins, chromobox protein homologs were identified in individual cell lines. The present findings demonstrate that rich protein information can be obtained by means of proteomic analysis for better understanding of oncogenesis and pathogenesis in a global way, which in turn represents the basis for the rational designs of diagnostic and therapeutic methods.

Introduction

In living systems, proteins, nucleic acids and other biomacromolecules within the cellular milieu interact with each other, thereby exhibiting various biological activities. The interaction of proteins with nucleic acids by and large is central to many cellular functions. Precisely regulated gene expression is a prerequisite for producing and maintaining the many different types of cells and tissues of a multicellular organism. Control of gene expression can be exerted at the transcription level at promoters and other DNA sequences outside the coding region. Post-transcriptional regulation is, however, the predominant regulatory mechanism for gene expression in eukaryotes. It can involve control of RNA splicing, transport, stability, localization and translation in a temporally and spatially regulated manner, allowing cells of different types or at different developmental stages to fine-tune their pattern of gene expression (1,2). These controls are mediated via interactions between regulatory DNA/RNA sequences and nucleic acid recognition motifs of a protein (3). Nucleic acid-binding proteins are modular in structure with conserved nucleic acid binding domains and auxiliary domains that function in the assembly of multiprotein complexes central to specific functions (2,4,5). Whilst some domains specifically recognize RNA or DNA, others recognize both DNA and RNA. Domains for DNA recognition include among others zinc finger, leucine zipper, helix-turn-helix, helix-loop-helix, POU and KH domains. Several motifs have also been identified for RNA-binding proteins and the list includes, but is not limited to the RNP domain (also called RNA recognition motif, RRM), the RGG box, zinc fingers, KH domain, double strand RNA-binding motif (DSRM) and the oligonucleotide/oligosaccharide binding motif (OB fold). The adaptation of a nucleic acid recognition motif to different purposes in different contexts is of importance both from functional and evolutionary perspectives.

Cancer is a disorder of cells and although it usually appears as a tumor made up of a mass of cells, the visible tumor is the end result of a whole series of changes which may have taken many years to develop. The behavior patterns of many tumors, and in some cases even the causal agents, are known, but how these agents transformed normal cells and influenced tumor cell behavior is beginning to be elucidated. The development of molecular biology opened up a major new approach to the molecular analysis of normal and tumor cells. Particularly, rapid advances in proteomic studies have greatly increased our understanding of the complex interactions of proteins within the cellular environment. The dynamic nature of the proteome of a cell or a tissue provides ample justification for studying gene expression in disease states at the protein level. There is growing interest in applying proteomics to gain insight into disease processes, develop

Correspondence to: Professor Gert Lubec, Department of Pediatrics, Division of Basic Sciences, Medical University of Vienna, Wahringer Gurtel 18-20, A-1090 Vienna, Austria
E-mail: gert.lubec@meduniwien.ac.at

Key words: tumor cell line, proteomics, nucleic acid-binding proteins, expression, markers

Table I. Identification of investigated cell lines (ATCC).

Name	ATCC no.	Cancer type
Saos-2	HTB-85	Osteosarcoma
SK-N-SH	HTB-11	Neuroblastoma
HCT116	CCL-247	Colorectal carcinoma
CaOva3	HTB-75	Adenocarcinoma of ovary
A-549	CCL-185	Alveolar cell carcinoma of lung
HL-60	CCL-240	Promyeolcytic leukemia
A-673	CRL-1598	Rhabdomyosarcoma
A-375	CRL-1619	Malignant melanoma
MCF-7	HTB-22	Adenocarcinoma of mammary gland
HeLa	CCL-2	Adenocarcinoma of cervix

new markers for diagnosis and early detection of diseases and accelerate drug development by identifying potential targets. This study reports on the identification of several nucleic acid-binding proteins and presents a comprehensive analysis of the expression of these proteins in a large panel of human tumor cell lines, forming the basis to achieve the aforementioned goals.

Materials and methods

Cell culture. Tumor cell lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA). The cell lines and their ATCC no. are given in Table I. SK-N-SH and HeLa cervix cell lines were cultured in minimum essential medium (Eagle) with 2 mM L-glutamine and Earle's basic salt solution adjusted to contain 1.5 g/l sodium bicarbonate, 0.1 mM non-essential amino acids, and 1.0 mM sodium pyruvate, with 10% fetal bovine serum. The same conditions were used to culture MCF-7 cell line except for supplementing 10% fetal bovine serum with 90% 0.01 mg/ml bovine insulin. A-673, A-375 and CaOva3 cell lines were cultured in Dulbecco's modified Eagle's medium with 4 mM L-glutamine adjusted to contain 1.5 g/l sodium bicarbonate and 4.5 g/l glucose with 10% fetal bovine serum. HCT116 and Saos-2 cell lines were cultured in McCoy's 5 medium with 1.5 mM L-glutamine and 15% fetal bovine serum.

HL-60 cell line was cultured with Iscove's modified Dulbecco's medium with 4 mM L-glutamine adjusted to contain 1.5 g/l sodium bicarbonate and 20% fetal bovine serum. A-549 cell line was cultured with Ham's F12K medium with 2 mM L-glutamine adjusted to contain 1.5 g/l sodium bicarbonate and 10% fetal bovine serum. Cell cultures were maintained in a humidified atmosphere of 5% v/v CO_2 in air at 37˚C.

Sample preparation. Harvested cells were washed three times with 10 ml phosphate-buffered saline (Gibco BRL, Gaithersburg, MD, USA) and centrifuged for 10 min at 800 x g at room temperature. The supernatant was discarded and the pellet was homogenized with 1.0 ml of sample buffer consisting of 40 mM Tris, 7 M urea (Merck, Germany), 2 M thiourea (Sigma, St. Louis, MO, USA), 4% CHAPS (3-[(3-cholamidopropyl) dimethylammonio]-1-propane-sulfonate) (Sigma, St. Louis, MO, USA), 65 mM 1,4-dithioerythritol (Merck), 1 mM EDTA (Merck) and protease inhibitors complete (Roche, Basel, Switzerland). After homogenization samples were left at room temperature for 1 h and centrifuged at 14,000 rpm for 1 h and supernatant was transferred into Ultrafree-4 centrifugal filter unit (Millipore, Bedford, MA), for desalting and concentrating proteins. Protein content of the supernatant was quantified by Bradford protein assay method (6). The standard curve was generated using bovine serum albumin and absorbance was measured at 595 nm.

Two-dimensional gel electrophoresis (2-DE). Samples prepared from each cell line were subjected to 2-DE as described elsewhere (7). Protein (1 mg) was applied on immobilized pH 3-10 non-linear gradient strips in sample cups at their basic and acidic ends. Focusing was started at 200 V and the voltage was gradually increased to 8,000 V at a rate of 4 V/min and then kept constant for a further 3 h (approximately 150,000 V/h totally). After the first dimension, strips (13 cm) were equilibrated for 15 min in a buffer containing 6 M urea, 20% glycerol, 2% sodium dodecylsulfate, 2% 1,4-dithioerythritol and then for 15 min in the same buffer containing 2.5% iodoacetamide instead of 1,4-dithioerythritol. After equilibration, strips were loaded onto 9-16% gradient sodium dodecylsulfate polyacrylamide gels for second-dimensional separation. Gels (180x200x1.5 mm) and run at 40 mA per gel. Immediately after the second dimension run, gels were fixed for 12 h in 50% methanol, containing 10% phosphoric acid and stained with Colloidal Coomassie blue (Novex, San Diego, CA) for 12 h on a rocking shaker. Molecular masses were determined by running standard protein markers (Bio-Rad), covering the range 10-250 kDa. pI values were used as given by the supplier of the immobilized pH gradient strips (Amersham Bioscience, Uppsala, Sweden). Excess of dye was washed out from the gels with distilled water and were scanned with ImageScanner (Amersham Pharmacia Biotech). Electronic images of gels were recorded using Adobe Photoshop and Microsoft PowerPoint Softwares.

Matrix-assisted laser desorption ionization mass spectrometry (MALDI/MS). Spots were excised with a spot picker (PROTEINEER sp™, Bruker Daltonics, Germany), placed into 96-well microtiter plates and in-gel digestion and sample preparation for MALDI analysis were performed by an automated procedure (PROTEINEER dp™, Bruker Daltonics) (8,9). Briefly, spots were excised and washed with 10 mM ammonium bicarbonate and 50% acetonitrile in 10 mM ammonium bicarbonate. After washing, gel plugs were shrunk by addition of acetonitrile and dried by blowing out the liquid through the pierced well bottom. The dried gel pieces were reswollen with 40 ng/μl trypsin (Roche Diagnostics, Penzberg, Germany) in enzyme buffer [consisting of 5 mM Octyl ß-D-glucopyranoside (OGP) and 10 mM ammonium bicarbonate] and incubated for 4 h at 30˚C. Peptide extraction was performed with 10 μl of 1% TFA in 5 mM OGP. Extracted peptides were directly applied onto a target (AnchorChip™, Bruker Daltonics) that was load with α-cyano-4-hydroxycinnamic acid (Bruker Daltonics) matrix thinlayer. The mass spectrometer used in this work was an Ultraflex™ TOF/TOF (Bruker Daltonics) operated in the reflector for MALDI-TOF peptide mass fingerprint (PMF) or LIFT mode for MALDI-TOF/TOF with a fully automated mode using the

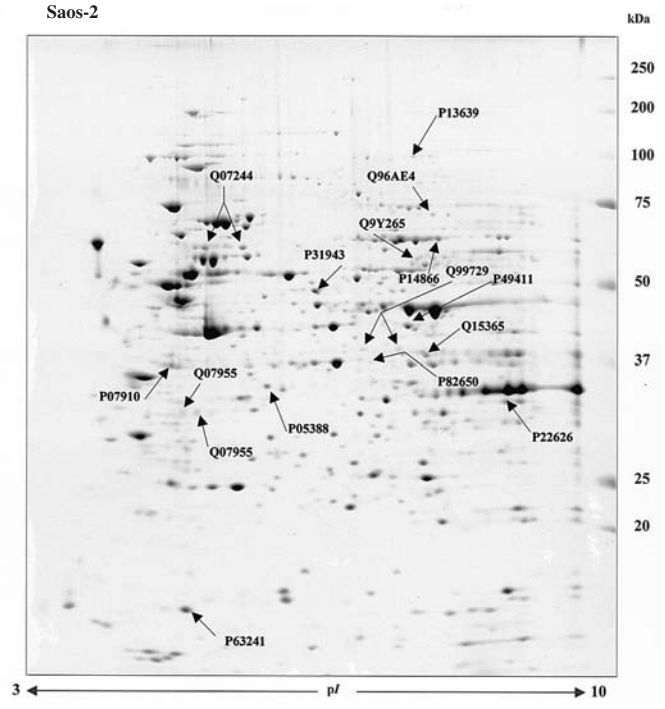

Figure 1. 2-DE gel image of human osteosarcoma (Saos-2), showing 15 proteins (18 spots). Proteins were extracted and separated on an immobilized pH 3-10 non-linear gradient strip followed by separation on a 9-16% gradient polyacrylamide gel. Gels were stained with Coomassie blue and spots were analyzed by MALDI-MS or MS/MS. Proteins are designated by their SWISS-PROT accession number. The names of the proteins are listed in Table II.

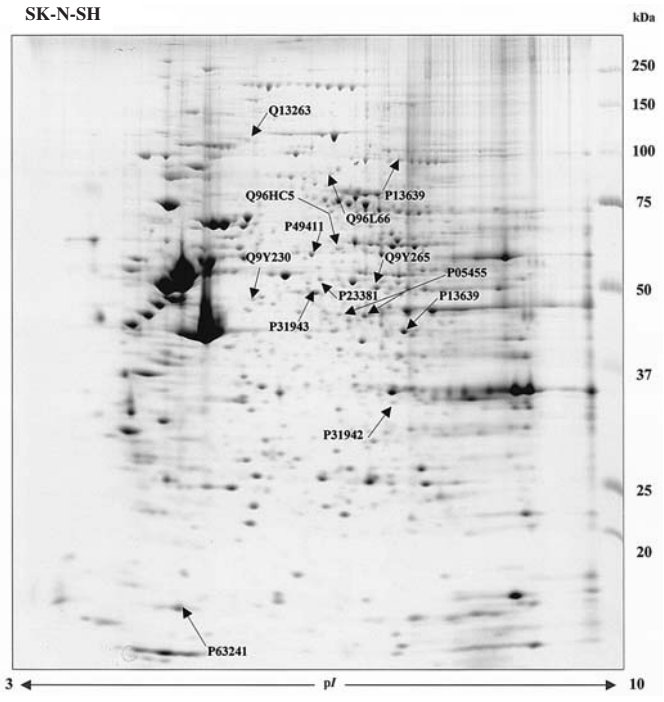

Figure 2. 2-DE gel image of human neuroblastoma cell line (SK-N-SH) showing 12 proteins (14 spots).

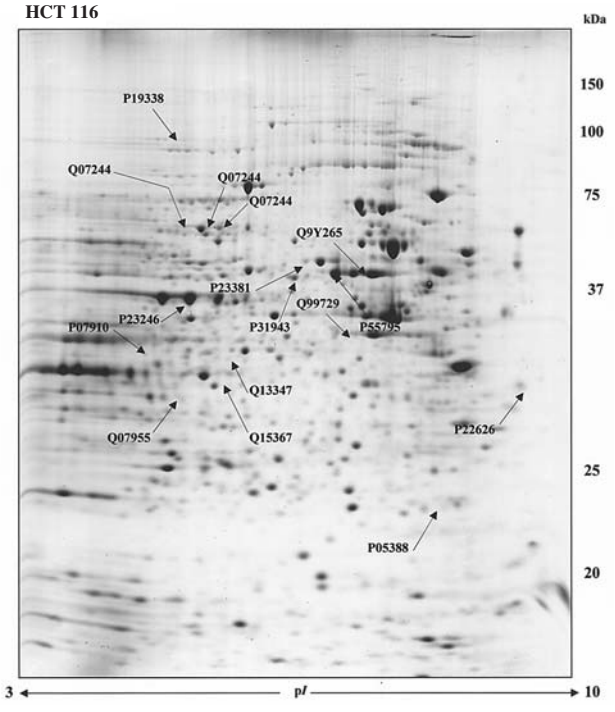

Figure 3. 2-DE gel image human colorectal carcinoma cell line (HCT116) showing 14 proteins (16 spots).

FlexControl™ software. An accelerating voltage of 25 kV was used for PMF. Calibration of the instrument was performed externally with $[M+H]^+$ ions of angiotensin I, angiotensin II, substance P, bombesin, and adrenocorticotropic hormones (clip 1-17 and clip 18-39). Each spectrum was produced by accumulating data from 200 consecutive laser shots. Those samples which were analyzed by PMF from MALDI-TOF were additionally analyzed using LIFT-TOF/TOF MS/MS from the same target. A maximum of three precursor ions per sample were chosen for MS/MS analysis. In the TOF1 stage, all ions were accelerated to 8 kV under conditions promoting metastable fragmentation. After selection of jointly migrating parent and fragment ions in a timed ion gate, ions were lifted by 19 kV to high potential energy in the LIFT cell. After further acceleration of the fragment ions in the second ion source, their masses could be simultaneously analyzed in the reflector with high sensitivity. PMF and LIFT spectra were interpreted with the Mascot software (Matrix Science Ltd., London, UK). Database searches, through Mascot, using combined PMF and MS/MS datasets were performed via BioTools 2.2 software (Bruker). A mass tolerance of 100 ppm and 2 missing cleavage sites for PMF and MS/MS tolerance of 0.5 Da and 1 missing cleavage sites for MS/MS search were allowed and oxidation of methionine residues was considered. The probability score calculated by the software was used as criterion for correct identification.

The algorithm used for determining the probability of a false positive match with a given mass spectrum is described elsewhere (10).

Results

To analyze expression profiles of different tumor cell lines, proteins were applied on 2-DE gel and visualized by staining with Coomassie blue. Representative gels of the experimental series produced following application of 1 mg total protein are shown in Figs. 1-10. Spots were analyzed by MALDI/MS

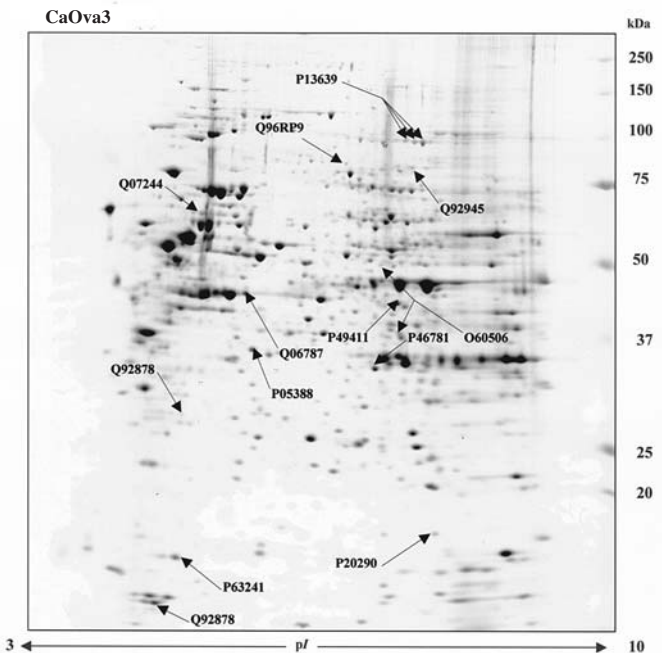

Figure 4. 2-DE gel image of human adenocarcinoma of ovary (CaOva3) showing 12 proteins (16 spots).

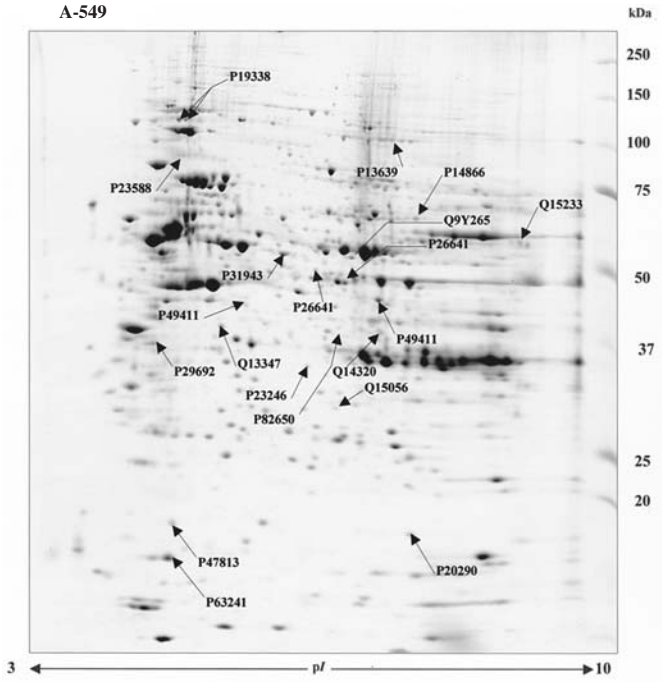

Figure 5. 2-DE gel image human lung cancer cell line (A-549) showing 17 proteins (20 spots).

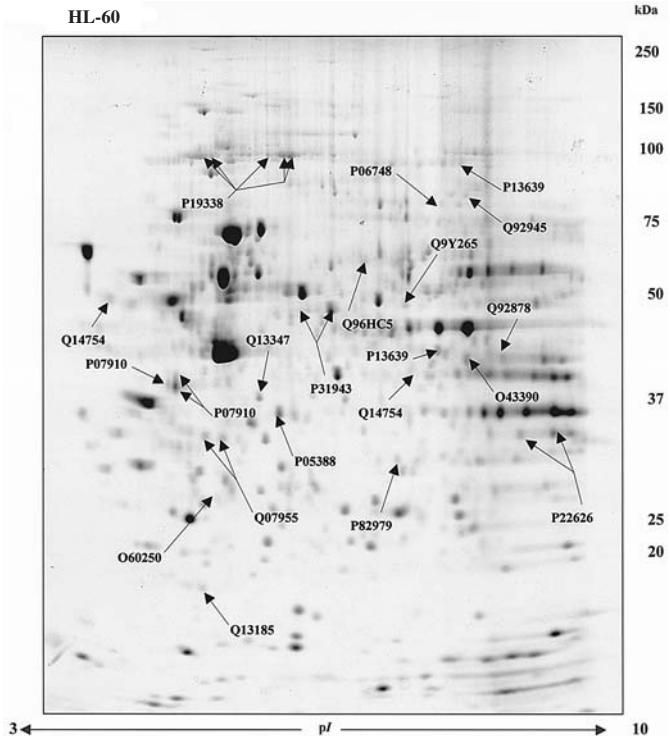

Figure 6. 2-DE gel image of human promyeloblastic leukemia cell line (HL-60) showing 18 proteins (29 spots).

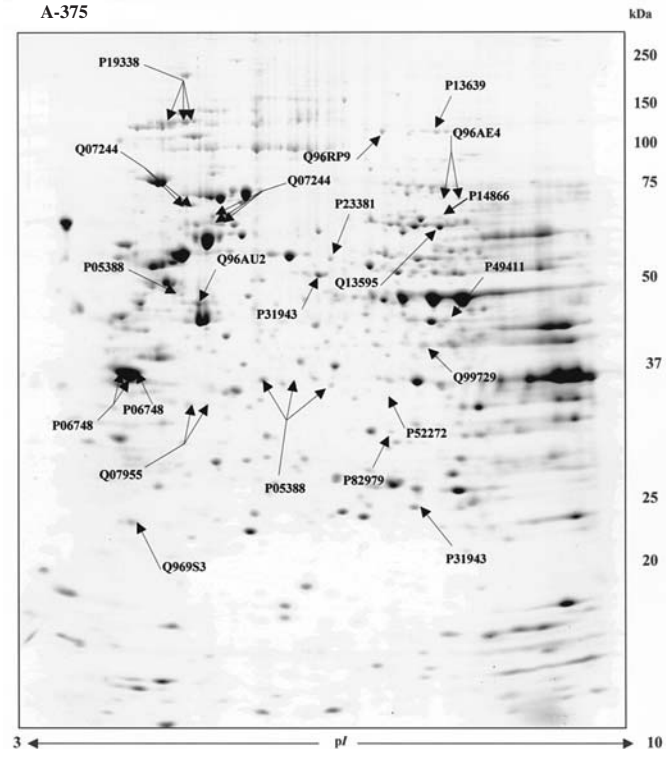

Figure 7. 2-DE gel image of human malignant melanoma cell line (A-375) showing 18 proteins (32 spots).

or MS/MS following in-gel trypsinization and identification was carried out by matching the peptide mass with the theoretical peptide mass of all proteins in the SWISS-PROT database. Internal standards were used to correct the measured peptide mass, thus reducing the windows of mass tolerance and increasing the confidence of identification. This proteomic approach revealed expression of 72 nucleic acid-binding proteins. SWISS-PROT accession numbers and names of the analyzed proteins with MALDI information are given in Table II.

Identified proteins are subgrouped into different categories based on their possible functions as DNA handling, RNA

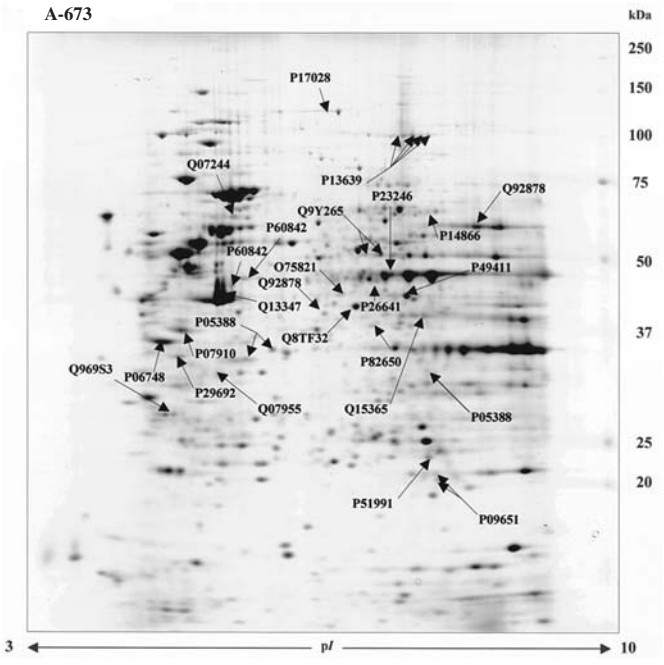

Figure 8. 2-DE gel image of human rhabdomyosarcoma cell line (A-673) showing 23 proteins (28 spots).

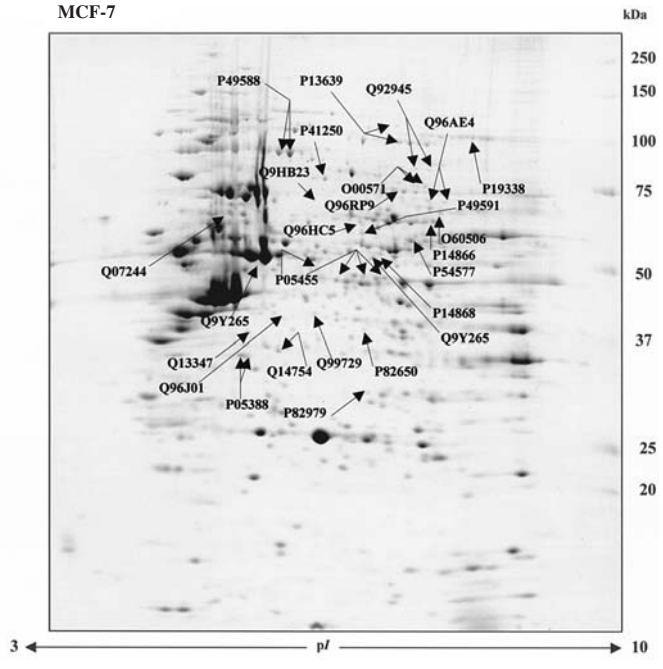

Figure 9. 2-DE gel image of human mammary gland cancer cell line (MCF-7) showing 25 proteins (33 spots).

handling and miscellaneous proteins. The results of domain analysis of nucleic acid-binding proteins are shown in Table III.

Proteomic analysis revealed variable expression of nucleic acid-binding proteins ranging from cell line-specific expression through expression limited to few cell lines to expression by multiple cell lines. About 45.8% of the total identified proteins were expressed only in one cell line, 41.7% in 2-5 cell lines, and 12.5% in 6-10 cell lines, representing tumor-specific, moderate tumor-associated and tumor-associated expression,

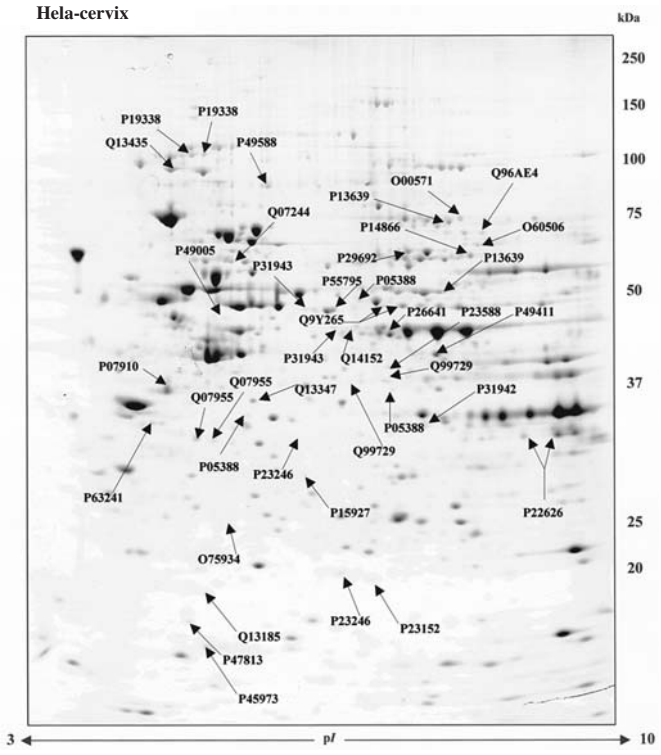

Figure 10. 2-DE gel image of human HeLa cervix cell line showing 33 proteins (42 spots).

respectively. Moreover, some proteins were represented by more than one spot. Three spots were detected for elongation factor 2 in CaOva3; three spots for heterogeneous nuclear ribonucleoprotein (hnRNP) C1/C2 in HL-60; two spots for lupus La protein in SK-N-SH, hnRNP A2/B1 in HeLa cervix, far upstream element binding protein (FAB) 1 in A-375 and MCF-7, eEF2 in SK-N-SH, and RuvB-like 1 in MCF-7. The different spots may represent different expression forms of the same protein, as post-translational modification is a common theme in protein biochemistry.

RNA handling proteins involved in splicing, transport, translation or degradation accounted for 76% and 69% of the total proteins identified and tumor-specific protein expression, respectively. RNA handling proteins involved in translation constituted the largest proportion of proteins that exhibited tumor-specific expression. This was followed by the hnRNPs that take part in pre-mRNA splicing and transport of the mature mRNA. Among the seven identified aminoacyl-tRNA synthetases, whilst six of them were expressed only in MCF-7 cell line, tryptophanyl-tRNA synthetase was detected in multiple cell lines. In stark contrast, hnRNPs expression was fairly distributed amongst cell lines. Only hnRNPs: A1, A3, F, R and M showed cell specific expression.

DNA handling proteins involved in transcription, replication or repair constituted about 18% of the total proteins identified. Not unlike the translational proteins, the transcriptional machinery proteins, notably transcription factors were the major DNA handling proteins that displayed tumor-specific expression. There was, however, little or no tumor-associated expression of DNA handling proteins, since FBP1 was the only protein that displayed expression in four cell lines. DNA handling proteins whose role in cancer is not known, such as

Table II. Mass spectroscopical identification of nucleic acid binding proteins in ten different investigated tumor cell lines (+, detected; -, undetected; each + represents one spot).

A, RNA handling proteins													
Accession no.	Protein name	TMW[a]	pI[b]	Observed tumor cell line									
				Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
O75934	Breast carcinoma amplified sequence 2	26131	5.5	-	-	-	-	-	-	-	-	-	+ OMW[c]: 26229 TS[d]: 173 PM[e]: 24
P29692	Elongation factor 1-δ	31122	4.9	-	-	-	-	+ OMW: 31217 TS: 175 PM: 20	-	-	+ OMW: 31217 TS: 169 PM: 23	-	+ OMW: 31217 TS: 95 PM: 17
P26641	Elongation factor 1-γ	49988	6.3	-	-	-	-	++ OMW: 37236 TS: 63 PM: 10	-	-	+ OMW: 37236 TS: 63 PM: 10	-	+ OMW: 50429 TS: 89 PM: 14
Q96RP9	Elongation factor G 1, mitochondrial (precursor)	83506	6.6	-	-	-	+ OMW: 84103 TS: 150 PM: 28	-	-	+ OMW: 84103 TS: 144 PM: 28	-	+ OMW: 84103 TS: 250 PM: 40	-
P49411	Elongation factor Tu, mitochondrial (precursor)	49542	7.3	+ OMW: 49852 TS: 242 PM: 34	+ OMW: 49852 TS: 130 PM: 20	-	+ OMW: 49852 TS: 183 PM: 29	++ OMW: 49852 TS: 182 PM: 28	-	+ OMW: 49852 TS: 351 PM: 45	+ OMW: 49852 TS: 204 PM: 28	-	+ OMW: 49851 TS: 369 PM: 47
P13639	Elongation factor 2	95207	6.4	+ OMW: 96115 TS: 99 PM: 20	++ OMW: 49851 TS: 130 PM: 20	-	+++ OMW: 49851 TS: 209 PM: 35	+ OMW: 96115 TS: 295 PM: 48	++ OMW: 96115 TS: 202 PM: 37	+ OMW: 96115 TS: 337 PM: 52	++++ OMW: 96115 TS: 231 PM: 38	++ OMW: 96115 TS: 290 PM: 43	++ OMW: 96115 TS: 192 PM: 30
P47813	Eukaryotic translation initiation factor 1A, X-chromosomal	16329	5.1	-	-	-	-	+ OMW: 22251 TS: 155 PM: 18	-	-	-	-	+ OMW: 16433 TS: 74 PM: 11
Q13347	Eukaryotic translation initiation factor 3 subunit 2	36502	5.4	-	-	+ OMW: 36878 TS: 159 PM: 23	-	+ OMW: 36878 TS: 102 PM: 20	+ OMW: 36878 TS: 159 PM: 23	-	+ OMW: 36878 TS: 130 PM: 21	+ OMW: 36878 TS: 257 PM: 31	+ OMW: 36878 TS: 135 PM: 21
O75821	Eukaryotic translation initiation factor 3 subunit 4	35611	5.9	-	-	-	-	-	-	-	+ OMW: 35959 TS: 89 PM: 15	-	-
Q14152	Eukaryotic translation initiation factor 3 subunit 10	166569	6.4	-	-	-	-	-	-	-	-	-	+ OMW: 29822 TS: 226 PM: 29
P60842	Eukaryotic initiation factor 4A-I	46154	5.3	-	-	-	-	-	-	-	+ OMW: 46353 TS: 159 PM: 27	-	-
P23588	Eukaryotic translation initiation factor 4B	69224	5.5	-	-	-	-	+ OMW: 69240 TS: 180 PM: 26	-	-	-	-	+ OMW: 69240 TS: 201 PM: 29
Q15056	Eukaryotic translation initiation factor 4H	27385	6.7	-	-	-	-	+ OMW: 69240 TS: 81 PM: 13	-	-	-	-	-

Table II. Continued.

A, RNA handling proteins Accession no.	Protein name	TMW	pI	Observed tumor cell line Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
P63241	Eukaryotic translation initiation factor 5A	16701	5.1	+ OMW: 16918 TS: 64 PM: 6	+ OMW: 16918 TS: 74 PM: 9	-	+ OMW: 16918 TS: 126 PM: 19	+ OMW: 16918 TS: 146 PM: 17	-	-	-	-	+ OMW 16918 TS: 130 PM: 16
Q06787	Fragile X mental retardation 1 protein	71174	7.0	-	-	-	+ OMW: 4611 TS: 67 PM: 8	-	-	-	-	-	-
P09651	Heterogeneous nuclear ribonucleoprotein A1	38715	9.3	-	-	-	-	-	-	-	++ OMW: 20946 TS: 95 PM: 13	-	-
P51991	Heterogeneous nuclear ribonucleoprotein A3	39595	9.1	-	-	-	-	-	-	-	+ OMW: 20946 TS: 111 PM: 15	-	-
Q99729	Heterogeneous nuclear ribonucleoprotein A/B	36613	9.0	++ OMW: 30683 TS: 93 PM: 13	-	+ OMW: 36041 TS: 138 PM: 17	-	-	-	+ OMW: 30683 TS: 92 PM: 13	-	+ OMW: 30683 TS: 77 PM: 12	++ OMW: 30683 TS: 105 PM: 15
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	37430	9.0	+ OMW: 36041 TS: 138 PM: 17	-	+ OMW: 36041 TS: 145 PM: 16	-	-	++ OMW: 36041 TS: 282 PM: 34	-	-	-	++ OMW: 30683 TS: 272 PM: 33
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	33688	5.0	+ OMW: 32004 TS: 82 PM: 13	-	+ OMW: 32004 TS: 76 PM: 16	-	-	+++ OMW: 32004 TS: 153 PM: 19	-	+ OMW: 32004 TS: 196 PM: 27	-	+ OMW: 32004 TS: 93 PM: 15
Q96AU2	Heterogeneous nuclear ribonucleoprotein F	45700	5.4	-	-	-	-	-	-	+ OMW: 46013 TS: 109 PM: 22	-	-	-
P31943	Heterogeneous nuclear ribonucleoprotein H	49230	5.9	+ OMW: 49848 TS: 67 PM: 16	+ OMW: 49848 TS: 70 PM: 16	+ OMW: 49848 TS: 114 PM: 22	-	+ OMW: 49848 TS: 126 PM: 22	++ OMW: 49484 TS: 155 PM: 25	++ OMW: 49484 TS: 225 PM: 31	-	-	++ OMW: 49484 TS: 186 PM: 29
P31942	Heterogeneous nuclear ribonucleoprotein H3	36926	6.4	-	+ OMW: 35273 TS: 86 PM: 19	-	-	-	-	-	-	-	+ OMW: 49484 TS: 155 PM: 25
P55795	Heterogeneous nuclear ribonucleoprotein H'	49264	5.9	-	-	+ OMW: 49517 TS: 127 PM: 26	-	-	-	-	-	-	+ OMW: 49517 TS: 179 PM: 31
Q07244	Heterogeneous nuclear ribonucleoprotein K	50976	5.4	++ OMW: 51230 TS: 78 PM: 17	-	++ OMW: 51230 TS: 97 PM: 18	+ OMW: 51230 TS: 109 PM: 16	-	-	++++ OMW: 51281 TS: 200 PM: 28	+ OMW: 51230 TS: 106 PM: 22	+ OMW: 51230 TS: 177 PM: 25	+ OMW: 51230 TS: 192 PM: 30
P14866	Heterogeneous nuclear ribonucleoprotein L	60187	6.7	+ OMW: 64617 TS: 97 PM: 23	-	-	-	+ OMW: 64617 TS: 97 PM: 23	-	+ OMW: 60719 TS: 161 PM: 30	+ OMW: 60719 TS: 139 PM: 27	+ OMW: 60719 TS: 203 PM: 33	+ OMW: 60719 TS: 235 PM: 39

Table II. Continued.

A, RNA handling proteins Accession no.	Protein name	TMW	pI	Observed tumor cell line Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
O60506	Heterogeneous nuclear ribonucleoprotein Q	69633	8.7	-	-	-	++ OMW: 62789 TS: 97 PM: 23	-	-	-	-	+ OMW: 62789 TS: 217 PM: 30	+ OMW: 62789 TS: 225 PM: 29
O43390	Heterogeneous nuclear ribonucleoprotein R	70943	8.2	-	-	-	-	-	+ OMW: 71184 TS: 61 PM: 12	-	-	-	-
P05455	Lupus La protein	46837	6.7	-	++ OMW: 46979 TS: 112 PM: 20	-	-	-	-	-	-	++++ OMW: 46979 TS: 167 PM: 25	-
P52272	Heterogeneous nuclear ribonucleoprotein M	77384	8.9	-	-	-	-	-	-	+ OMW: 73854 TS: 64 PM: 15	-	-	-
P19338	Nucleolin	76213	4.6	-	-	+ OMW: 76224 TS: 114 PM: 21	-	++ OMW: 46224 TS: 173 PM: 28	++++ OMW: 46224 TS: 80 PM: 16	+++ OMW: 76224 TS: 217 PM: 31	-	+ OMW: 76224 TS: 125 PM: 23	++ OMW: 46224 TS: 94 PM: 20
P06748	Nucleophosmin	32575	6.4	-	-	-	-	-	+ OMW: 31090 TS: 74 PM: 15	+++ OMW: 31090 TS: 197 PM: 27	+ OMW: 31090 TS: 88 PM: 15	-	-
Q14754	ORFII	83610	9.4	-	-	-	-	-	++ OMW: 84443 TS: 64 PM: 15	- -	- -	+ OMW: 84443 TS: 81 PM: 18	-
Q15365	Poly(rC)-binding protein 1	37498	6.7	+ OMW: 37987 TS: 77 PM: 17	-	-	-	-	-	-	+ OMW: 38015 TS: 75 PM: 19	-	-
P15927	Replication protein A 32 kDa subunit	29247	5.8	-	-	-	-	-	-	-	-	-	+ OMW: 29342 TS: 78 PM: 14
Q15367	Ribonucleoprotein (fragment)	40594	8.5	-	-	+ OMW: 46979 TS: 80 PM: 16	-	-	-	-	-	-	-
Q07955	Splicing factor, arginine/serine-rich 1	27613.5	10.4	+ OMW: 27711 TS: 102 PM: 11	-	+ OMW: 22560 TS: 69 PM: 8	-	-	++ OMW: 27711 TS: 210 PM: 24	++ OMW: 27711 TS: 215 PM: 23	+ OMW: 27711 TS: 137 PM: 18	-	++ OMW: 27711 TS: 227 PM: 25
P23152	Splicing factor, arginine/serine-rich 3	19330	11.6	-	-	-	-	-	-	-	-	-	+ OMW: 27711 TS: 64 PM: 7
P23246	Splicing factor, proline- and glutamine-rich	76149	9.5	-	-	+ OMW: 45277 TS: 63 PM: 13	-	+ OMW: 45277 TS: 75 PM: 13	-	-	+ OMW: 45277 TS: 76 PM: 12	-	++ OMW: 45277 TS: 81 PM: 15

Table II. Continued.

A, RNA handling proteins Accession no.	Protein name	TMW	pI	Observed tumor cell line Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
Q13435	Splicing factor 3B subunit 2	97657	5.5	-	-	-	-	-	-	-	-	-	+ OMW: 97710 TS: 96 PM: 23
Q96J01	THO complex subunit 3	38772	5.7	-	-	-	-	-	-	-	-	+ OMW: 39431 TS: 99 PM: 16	-
Q13263	Transcription intermediary factor 1-ß	88550	5.5	-	+ OMW: 90261 TS: 83 PM: 22	-	-	-	-	-	-	-	-
P20290	Transcription factor BTF3	22168	9.4	-	-	-	+ OMW: 17688 TS: 63 PM: 10	+ OMW: 17688 TS: 95 PM: 11	-	-	-	-	-
Q13595	Transformer-2 protein homolog	32689	11.3	-	-	-	-	-	-	+ OMW: 32726 TS: 61 PM: 10	-	-	-
P17028	Zinc finger protein 24	42128	5.8	-	-	-	-	-	-	-	+ OMW: 54808 TS: 63 PM: 13	-	-
Q969S3	Zinc finger protein 622	54272	5.8	-	-	-	-	-	-	+ OMW: 54808 TS: 70 PM: 12	+ OMW: 54808 TS: 63 PM: 13	-	-
Q8TF32	Zinc finger protein 431	67217	9.0	-	-	-	-	-	-	-	+ OMW: 65979 TS: 81 PM: 14	-	-

B, DNA handling proteins Accession no.	Protein name	TMW	pI	Observed tumor cell line Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
P49588	Alanyl-tRNA synthetase	106801	5.3	-	-	-	-	-	-	-	-	+ OMW: 107476 TS: 163 PM: 34	+ OMW: 107476 TS: 182 PM: 38
P14868	Aspartyl-tRNA synthetase	57136	6.1	-	-	-	-	-	-	-	-	+ OMW: 57499 TS: 174 PM: 28	-
O00571	DEAD-box protein 3, X-chromosomal	73112	6.7	-	-	-	-	-	-	-	-	++ OMW: 73466 TS: 178 PM: 31	+ OMW: 73466 TS: 252 PM: 36
P49005	DNA polymerase δ subunit 2	51289	5.4	-	-	-	-	-	-	-	-	-	+ OMW: 51884 TS: 122 PM: 21

Table II. Continued.

B, DNA handling proteins Accession no.	Protein name	TMW	pI	Observed tumor cell line Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
Q92878	DNA repair protein RAD50	153892	6.5	-	-	-	++ OMW: 51884 TS: 82 PM: 29	-	+ OMW: 154823 TS: 75 PM: 26	-	++ OMW: 139145 TS: 79 PM: 24	-	-
Q96AE4	Far upstream element binding protein 1	67473	7.2	+ OMW: 67690 TS: 78 PM: 23	-	-	-	-		++ OMW: 67690 TS: 216 PM: 41	-	++ OMW: 67690 TS: 106 PM: 23	+ OMW: 68790 TS: 155 PM: 31
Q92945	Far upstream element binding protein 2	72709	8.0	-	-	-	+ OMW: 73458 TS: 66 PM: 18	-	+ OMW: 73458 TS: 190 PM: 34	-	-	++ OMW: 73458 TS: 162 PM: 27	-
P41250	Glycyl-tRNA synthetase	83140	6.6	-	-	-	-	-	-	-	-	+ OMW: 78166 TS: 274 PM: 42	-
Q9HB23	Lysyl-tRNA synthetase	71384	6.3	-	-	-	-	-	-	-	-	+ OMW: 71795 TS: 97 PM: 20	-
Q9Y265	RuvB-like 1	50228	6.0	+ OMW: 50538 TS: 175 PM: 28	+ OMW: 50538 TS: 66 PM: 11	+ OMW: 50538 TS: 207 PM: 32	-	+ OMW: 50538 TS: 194 PM: 26	+ OMW: 50538 TS: 97 PM: 16	-	++ OMW: 50538 TS: 131 PM: 22	+ OMW: 50538 TS: 195 PM: 28	++ OMW: 50538 TS: 368 PM: 45
Q9Y230	RuvB-like 2	51156.6	5.5	-	+ OMW: 512696 TS: 199 PM: 33	-	-	-	-	-	-	-	-
P49591	Seryl-tRNA synthetase	58646	6.1	-	-	-	-	-	-	-	-	+ OMW: 59112 TS: 252 PM: 34	-
P23381	Tryptophanyl-tRNA synthetase	53165.4	5.8	-	+ OMW: 53396 TS: 84 PM: 17	+ OMW 53396 TS: 117 PM: 23	-	-	-	+ OMW: 53396 TS: 180 PM: 28	-	-	-
P54577	Tyrosyl-tRNA synthetase	59012	6.6	-	-	-	-	-	-	-	-	+ OMW: 59448 TS: 270 PM: 36	-
Q96HC5	U2 small nuclear ribonucleoprotein auxiliary factor	53120.5	9.2	-	+ OMW: 53429 TS: 81 PM: 19	-	-	-	+ OMW: 53429 TS: 106 PM: 20	-	-	+ OMW: 53429 TS: 73 PM: 14	-
Q14320	XAP-5 protein	40242	6.4	-	-	-	-	+ OMW: 40216 TS: 193 PM: 27	-	-	-	-	-

Table II. Continued.

C, Miscellaneous Accession no.	Protein name	TMW	pI	Observed tumor cell line									
				Saos-2	SK-N-SH	HCT116	CaOva3	A-549	HL-60	A-375	A-673	MCF-7	HeLa
Q15233	Non-POU domain-containing octamer-binding protein	54100	9.0	-	-	-	-	+ OMW: 54180 TS: 262 PM: 40	-	-	-	-	-
P46781	40S ribosomal protein S9	22460	10.7	-	-	-	+ OMW: 22635 TS: 67 PM: 10	-	-	-	-	-	-
P05388	60S acidic ribosomal protein P0	49204	5.1	+ OMW: 34423 TS: 127 PM: 17	-	+ OMW: 34423 TS: 90 PM: 8	+ OMW: 34423 TS: 84 PM: 14	-	+ OMW: 34423 TS: 83 PM: 12	++++ OMW: 41425 TS: 158 PM: 22	+++ OMW: 34423 TS: 197 PM: 23	++ OMW: 34423 TS: 160 PM: 18	+++ OMW: 34423 TS: 241 PM: 30
Q13185	Chromobox protein homolog 3	20824	5.2	-	-	-	-	-	+ OMW: 19879 TS: 176 PM: 19	-	-	-	+ OMW: 35273 TS: 78 PM: 11
P45973	Chromobox protein homolog 5	22225	5.7	-	-	-	-	-	-	-	-	-	+ OMW: 22251 TS: 155 PM: 18
P82650	Mitochondrial 28S ribosomal protein S22	41280	7.7	+ OMW: 41425 TS: 104 PM: 15	-	-	-	+ OMW: 41425 TS: 206 PM: 23	-	-	+ OMW: 41425 TS: 155 PM: 19	+ OMW: 41425 TS: 368 PM: 36	-
Q96L66	NRAS-related protein	79412	5.6	-	+ OMW: 89684 TS: 93 PM: 21	-	-	-	-	-	-	-	-
P82979	Nuclear protein Hcc-1	23540	6.1	-	-	-	-	-	+ OMW: 23713 TS: 95 PM: 13	+ OMW: 23713 TS: 90 PM: 12	-	+ OMW: 36136 TS: 134 PM: 23	-
O60250	Ribosomal protein L13 (fragment)	7390	10.3	-	-	-	-	-	+ OMW: 23713 TS: 95 PM: 13	-	-	-	-

Saos-2, bone osteosarcoma; SK-N-SH, brain neuroblastoma; HCT116, adenocarcinoma of colon; CaOva3, adenocarcinoma of ovary; A-549: alveolar cell carcinoma of lung; HL-60, peripheral blood, promyelocyte leukaemia; A-375, skin malignant melanoma; A-673, muscle rhabdomyosarcoma; MCF-7, breast cancer; HeLa cervix, cervix adenocarcinoma. [a]Theoretical molecular weight. [b]Theoretical isoelectric point; Mascot-search-MS-results: [c]observed molecular weight, [d]total score, [e]number of matched peptides.

THO complex subunit 3 (expressed in MCF-7) and DNA handling proteins including XAP-5 protein (expressed in A-549 cell line), both zinc finger protein 431 and zinc finger protein 24 (only in A-673) were identified.

Proteins grouped under miscellaneous include proteins that bind to DNA and RNA as well as those that may not have direct nucleic acid-binding activity, but contribute to the handling of nucleic acids indirectly for e.g., through binding to the nucleic acid-binding proteins. Such proteins accounted for 13.9% of the total identified proteins, out of which non-POU domain-containing octamer-binding protein, 40S ribosomal protein S9; chromobox protein homolog 5; NRAS-related protein and ribosomal protein L13 exhibited cell line specificity. Although proteins that bind to both DNA and RNA are grouped under miscellaneous (Table II), the poly (rc) binding protein family, including poly (rc) binding protein 1 (hnRNP E1) and hnRNP K are discussed under RNA binding proteins for purpose of simplicity.

Table III. Results of domain analysis of nucleic acid binding proteins in ten tumor cell lines.

Accession no.	Protein name	Gene name	RNA-binding motif (RRM)	DNA binding domains	Other NAB related domains/protein family	Other domains
O75934	Breast carcinoma amplified sequence 2	BCAS2	-	-	Breast carcinoma amplified sequence 2 family Pfam[a]: PF05700.1 InterPro[b]: IPR008409 spliceosome associated protein	-
P29692	Elongation factor 1-δ	EEF1D	-	-	Elongation factor 1, ß/ß'/δ chain IPR001326 PF00736	-
P26641	Elongation factor 1-γ	EEF1G	-	-	Elongation factor 1, γ chain IPR001662	Glutathione S-transferase, C-terminal-like IPR010987 Glutathione S-transferase, C-terminal IPR004046 Glutathione S-transferase, N-terminal IPR004045
Q96RP9	Elongation factor G 1, mitochondrial (precursor)	GFM1	-	-	Translation elongation factor G family IPR004540 Elongation factor G, C-terminal IPR000640 Elongation factor G, domain IV IPR005517 Elongation factor Tu, domain 2 IPR004161	Protein synthesis factor, GTP-binding IPR000795 Small GTP-binding protein domain IPR005225
P49411	Elongation factor Tu, mitochondrial (precursor)	TUFM	-	-	Translation elongation factor Tu family IPR004541 Elongation factor Tu, C-terminal domain IPR004160 Elongation factor Tu, domain 2 IPR004161	Protein synthesis factor, GTP-binding IPR000795
P47813	Eukaryotic translation initiation factor 1A, X-chromosomal	EIF1AX	-	-	Eukaryotic initiation factor 1A (eIF-1A) IPR001253 S1, IF1 type IPR006196	Nucleic acid-binding, OB-fold IPR008994
Q13347	Eukaryotic translation initiation factor 3 subunit 2	EIF3S2	-	-	WD-40 repeat IPR001680	-
O75821	Eukaryotic translation initiation factor 3 subunit 4	EIF3S4	1 RRM	-	-	-
Q14152	Eukaryotic translation initiation factor 3 subunit 10	EIF3S10	-	-	Proteasome component region PCI IPR000717	Spectrin repeat IPR002017
P60842	Eukaryotic initiation factor 4A-I	EIF4A1	1 RRM	-	DEAD/DEAH box helicase IPR001410 ATP-dependent helicase, DEAD-box IPR000629 Helicase, C-terminal IPR001650	-
Q15056	Eukaryotic translation initiation factor 4H	WBSCR1	1 RRM	-	-	-
P63241	Eukaryotic translation initiation factor 5A	EIF5A	-	-	Eukaryotic initiation factor 5A hypusine (eIF-5A) IPR001884 KOW (Kyrpides, Ouzounis, Woese) motif IPR005824 Translation protein SH3-like IPR008991	-
Q06787	Fragile X mental retardation 1 protein	FMR1	-	KH (k homology) domain KH, type 1 IPR004088	Nucleic acid-binding, OB-fold IPR008994	-

Table III. Continued.

Accession no.	Protein name	Gene name	RNA-binding motif (RRM)	DNA binding domains	Other NAB related domains/protein family	Other domains
P09651	Heterogeneous nuclear ribonucleoprotein A1	HNRPA1	2 RRM	-	-	-
P51991	Heterogeneous nuclear ribonucleoprotein A3	HNRPA3	2 RRM	-	-	-
Q99729	Heterogeneous nuclear ribonucleoprotein A/B	HNRPAB	2 RRM	-	-	-
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	HNRPA2B1	2 RRM	-	-	-
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRPC	1 RRM	-	-	-
Q96AU2	Heterogeneous nuclear ribonucleoprotein F	-	3 RRM	-	-	-
P31943	Heterogeneous nuclear ribonucleoprotein H	HNRPH1	3 RRM	-	-	-
P31942	Heterogeneous nuclear ribonucleoprotein H3	HNRPH3	2 RRM	-	-	-
P55795	Heterogeneous nuclear ribonucleoprotein H'	HNRPH2	3 RRM	-	-	-
P61978	Heterogeneous nuclear ribonucleoprotein K	HNRPK	-	3 KH domains	-	-
P14866	Heterogeneous nuclear ribonucleoprotein L	HNRPL	3 RRM	-	-	-
O60506	Heterogeneous nuclear ribonucleoprotein Q	SYNCRIP	3 RRM	-	HnRNP R and Q splicing factor family IPR006535	-
O43390	Heterogeneous nuclear ribonucleoprotein R	HNRPR	3 RRM	-	HnRNP R and Q splicing factor family IPR006535	-
P05455	Lupus La protein	SSB	1 RRM	-	RNA-binding protein Lupus La domain IPR006630	-
P52272	Heterogeneous nuclear ribonucleoprotein M	HNRPM	3 RRM	-	-	-
P19338	Nucleolin	NCL	4 RRM	-	-	-
P06748	Nucleophosmin	NPM1	-	-	Nucleoplasmin family IPR004301	-
Q14754	ORFII	-	-	-	RNA-directed DNA polymerase (reverse transcriptase) domain IPR000477	-
Q15365	Poly(rC)-binding protein 1	PCBP1	-	3 KH domains	-	-
P15927	Replication protein A 32 kDa subunit	RPA2	-	-	Nucleic acid binding, OB-fold, tRNA/helicase-type domain IPR004365	-
Q15367	Ribonucleoprotein (fragment)	-	1 RRM	-	RNA-binding protein Lupus La domain IPR006630	-
Q07955	Splicing factor, arginine/ serine-rich 1	SFRS1	2 RRM	-	-	-
P84103	Splicing factor, arginine/ serine-rich 3	SFRS3	1 RRM	-	-	-
P23246	Splicing factor, proline- and glutamine-rich	SFPQ	2 RRM	-	-	-
Q13435	Splicing factor 3B subunit 2	F3B2	-	DNA-binding SAP (after SAF-A/B, Acinus and PIAS) IPR003034	-	Protein of unknown function DUF382 IPR007180 PSP, proline-rich domain IPR006568
Q96J01	THO complex subunit 3	THOC3	-	-	-	6 WD repeat

Table III. Continued.

Accession no.	Protein name	Gene name	RNA-binding motif (RRM)	DNA binding domains	Other NAB related domains/protein family	Other domains
Q13263	Transcription intermediary factor 1-ß	TRIM28	-	-	2 B-box, C-terminal domain IPR003649 1 Bromodomain IPR001487 1 Zn-finger, B-box IPR000315 1 Zn-finger-like, PHD finger IPR001965 1 Zn-finger, RING IPR001841	-
P20290	Transcription factor BTF3	BTF3	-	-	Nascent polypeptide-associated complex NAC domain IPR002715	-
Q13595	Transformer-2 protein homolog	TRA2A	1 RRM	-	-	-
P17028	Zinc finger protein 24	ZNF24	-	-	Transcriptional regulator SCAN domain IPR003309 Zn-finger, C2H2 type IPR007087 Zn-finger, C2H2 subtype IPR007086	-
Q969S3	Zinc finger protein 622	ZNF622	-	-	Zn-finger, U1-like domain IPR003604	-
Q8TF32	Zinc finger protein 431	ZNF431	-	-	KRAB box family IPR001909 Zn-finger, C2H2 type IPR007087 Zn-finger, C2H2 subtype IPR007086	-
P49588	Alanyl-tRNA synthetase	AARS	-	-	Alanyl-tRNA synthetase, class IIc IPR002318 Alanyl-transfer RNA synthetase IPR006193	Phosphoesterase, DHHA1 domain IPR003156
P14868	Aspartyl-tRNA synthetase	DARS	-	-	Aspartyl-tRNA synthetase archea/euk type family IPR004523 nucleic acid binding, OB-fold, tRNA/helicase-type domain IPR004365 tRNA synthetase, class II (D, K and N) domain IPR004364 Aspartyl-tRNA synthetase, class IIb domain IPR002312 Aminoacyl-transfer RNA synthetase, class II domain IPR006195	-
O00571	DEAD-box protein 3, X-chromosomal	DDX3X	-	-	DEAD/DEAH box helicase domain IPR001410 ATP-dependent helicase, DEAD-box domain IPR000629 Helicase, C-terminal domain IPR001650	-
P49005	DNA polymerase δ subunit 2	POLD2	-	-	Belongs to the DNA polymerase δ/II small subunit family	-
Q92878	DNA repair protein RAD50	RAD50	-	-	Rad50 zinc hook motif IPR007517	ABC transporter related domain IPR003439 Mss4-like IPR011057
Q92945	Far upstream element binding protein 2	KHSRP	-	4 KH domains	-	-
P41250	Glycyl-tRNA synthetase	GARS	-	-	tRNA synthetase, class II (G, H, P and S) domain IPR002314 Aminoacyl-transfer RNA synthetase, class II domain IPR006195 WHEP-TRS domain IPR000738	Anticodon-binding domain IPR004154

Table III. Continued.

Accession no.	Protein name	Gene name	RNA-binding motif (RRM)	DNA binding domains	Other NAB related domains/protein family	Other domains
Q9HB23	Lysyl-tRNA synthetase	-	-	-	Nucleic acid binding, OB-fold, tRNA/helicase-type IPR004365 tRNA synthetase, class II (D, K and N) domain IPR004364 Aminoacyl-transfer RNA synthetase, class II domain IPR006195	-
Q9Y265	RuvB-like 1	RUVBL1	-	-	TBP-interacting protein 49 (TIP49), C-terminal domain IPR010339	AAA ATPase domain IPR003593 AAA ATPase, central region IPR003959
Q9Y230	RuvB-like 2	RUVBL2	-	-	TBP-interacting protein 49 (TIP49), C-terminal domain IPR010339	-
P49591	Seryl-tRNA synthetase	SARS	-	-	tRNA synthetase, class II (G, H, P and S) domain IPR002314 Aminoacyl-transfer RNA synthetase, class II domain IPR006195	
P23381	Tryptophanyl-tRNA synthetase	WARS	-	-	Aminoacyl-tRNA synthetase, class Ib domain IPR002305 WHEP-TRS domain IPR000738	-
P54577	Tyrosyl-tRNA synthetase	YARS	-	-	Aminoacyl-tRNA synthetase, class Ib domain IPR002305 t-RNA-binding region domain IPR002547	-
Q96HC5	U2 (RNU2) small nuclear RNA auxiliary factor 2, isoform b	U2AF2	RRM	-	-	-
Q14320	XAP-5 protein	XAP5	-	-	XAP5 protein family IPR007005	-
Q15233	Non-POU domain-containing octamer-binding protein	NONO	2 RRM	-	-	-
P46781	40S ribosomal protein S9	RPS9	-	-	-	RNA-binding S4 domain IPR002942
P05388	60S acidic ribosomal protein P0	RPLP0	-	-	-	Ribosomal protein L10 family IPR001790 Ribosomal protein 60S family IPR001813
Q13185	Chromobox protein homolog 3	CBX3	-	-	-	2 Chromo domain IPR000953 Chromo shadow domain IPR008251
P45973	Chromobox protein homolog 5	CBX5	-	-	-	2 Chromo domain IPR000953 Chromo shadow domain IPR008251
P82650	Mitochondrial 28S ribosomal protein S22	MRPS22	-	-	Ribosomal S22 mitochondrial 28S mitochondrion S22MT MRP-S22 C14A4.14 GK002 CG12261 PD250041	-
Q96L66	NRAS-related protein	UNR	-	Cold-shock protein, DNA-binding domain IPR002059 Cold shock protein domain IPR011129	-	Haem peroxidase, plant/ fungal/bacterial family IPR002016
P82979	Nuclear protein Hcc-1	HCC1	-	DNA-binding SAP (after SAF-A/B, Acinus and PIAS) motif IPR003034	-	-
O60250	Ribosomal protein L13 (fragment)	-	-	-	Ribosomal protein L13e family IPR001380	-

[a]Protein families database of alignments and HMMs. [b]InterPro database.

Discussion

Expression of genes, including those involved in carcinogenesis is regulated at many different levels: transcription, mRNA processing and transport, mRNA degradation, and translation of mRNA to synthesize the protein it is encoded for, and nucleic acid-binding proteins play a major role in this regard. This report explored the use of proteomics as a tool for identifying tumor-specific and/or tumor-associated expression of nucleic-acid handling proteins that could serve as diagnostic biomarkers or therapeutic targets in a wide-range of cell lines used by the scientific community. In the present comprehensive study, we hereby demonstrate cell-line specific expression of 32 proteins, which may represent the well-defined phenotype of the individual cell lineages and cell type specific functions.

RNA handling proteins. Since gene expression is regulated during further processing of pre-mRNA into mRNA that is ready to be translated in the cytosol, proteins that bind to RNA during processing may thus be involved in controlling gene expression. One such well-characterized group is the SR family of proteins, which are splicing regulatory proteins that among other properties bind to exonic splicing enhancer sequences (11). Specific phosphorylation of SR proteins is one of the key determinants regulating splicing events and antitumor drugs, such as indolocarbazoles have been shown to act by inhibiting SR phosphorylation (12). Tumor-associated expression of splicing factor SR1 and cell line-specific expression of SR3 in HeLa cell suggest that different tumors might differentially regulate each member of a protein family and this might help in defining different tumor types. Another less-well characterized group of pre-mRNA binding proteins are hnRNPs where about 30 have been identified so far and function in a staggering array of cellular activities, ranging from transcription and pre-mRNA processing in the nucleus to cytoplasmic mRNA translation and turn over (13). An accumulating body of evidence indicates that hnRNPs are implicated in tumorigenesis. In BCR/ABL leukemogenesis, hnRNP A1 and hnRNP E2 proteins are upregulated via post-translational mechanism that resulted in enhanced protein stability (14). hnRNP C1/C2, hnRNP E1 and hnRNP K (the later two bind both RNA and DNA) have been shown to bind and enhance translation at the internal ribosomal entry site element of *c-myc*, with hnRNP C1/C2 binding at the 367-411 nucleotide region and hnRNP E1 and hnRNP K at 158-211 nucleotide region (15,16). Moreover, expression of hnRNP K is up-regulated in grade III breast cancer samples and this up-regulation appears to correlate with anchorage-independence and enhanced proliferation of breast cancer cells (17). hnRNP F and hnRNP H/H', which belong to a subgroup of hnRNPs characterized by a special structure of the RRM domains that consequently have been denoted as quasi-RRM domains, are differentially regulated in a variety of tumors (11). The observation that all-trans-retinoic-acid inducing growth inhibition, differentiation, and apoptosis; down-regulated a number of hnRNPs in acute promyelocytic leukemia (18) strongly reinforces the importance of hnRNPs in cancer. Given the role of hnRNPs in tumor biology, it is not surprising to see a fair distribution of hnRNPs in the cell lines investigated. This suggests that while some hnRNPs (for e.g., A1, A3, H3, F, Q and R) could possibly serve as a specific-marker for a specific tumor, others (such as, C1/C2, A2/B1, A/B, E1, H/H', K and L) may be used as a general marker for tumor development. Lupus La protein plays a role in stabilization of nascent transcripts of RNA polymerase III and is shown to increase expression of mdm2, a negative regulator of p53, in BCR/ABL expressing cells (19). Lupus La protein expressed in three cell lines, and ribonucleoprotein (fragment), which is specifically expressed in HCT116 and has a lupus La domain, are among candidate proteins with possible modulatory role in tumorigenesis.

The process of translation is very complex and involves several translational factors that control the three distinct steps of translation i.e., initiation, elongation and termination. Regulation of gene expression at the level of translation has attracted much attention particularly in relationship to cancer following the observation that disruption by mutation or overexpression of translational factors can cause cellular transformation (20). Since then, a number of studies have been published showing that alteration in expression of, notably overexpression of eIF2 and eIF4 translational factors is often associated with human carcinogenesis (21-25). Inspite of the evidence implicating the involvement of certain translation factors in neoplastic developments, the underlying mechanisms responsible for their oncogenic potential are less well understood. Several translation factors, including seven initiation and four elongation factors have been identified in the present study. Although no direct link of the identified factors and cancer exists in the literature, there are few candidates that may be considered as relevant for tumor biology for e.g., eEF2. eEF2 is phosphorylated by eEF2-kinase, whose activity is increased in several malignancies and non-specific inhibitors of this enzyme promoted cell death (26).

Aminoacyl-tRNA synthetases catalyze activation of amino acids and their subsequent linkage to tRNAs, a process that simplifies the difficulty of formation of peptide bond between two amino acids and codon recognition. A finding that appears to be of significance is that expression of nearly all identified aminoacyl-tRNAs was restricted to MCF-7 cell line, tryptophanyl-tRNA synthetase being an exception, which was expressed in three cell lines. At this juncture, we cannot explain exactly why most aminoacyl-tRNA synthetases concentrated in one cell line. However, it seems that MCF-7 cell line largely expresses class II synthetases and all the rest class I synthetases, the exception being tyrosyl-tRNA synthetase, a class I enzyme expressed by the MCF-7 cell line. This finding suggests that class II enzymes that activate smaller amino acids are potential markers for breast cancer better than class I that activates larger and also more hydrophobic amino acids.

DNA handling proteins. Initiation, progression and maintenance of some tumors depends upon overexpression through gene amplification or inappropriate expression of normal genes, mutation in a critical region of a gene, or impaired responses to DNA damage. Transcription activator and repressor proteins that regulate gene expression by altering the rate of formation of the transcriptional apparatus, and proteins involved in the

DNA repair pathways thus have an important role in tumorigenesis. A number of DNA handling proteins, which have already been implicated in tumor biology have been identified in the investigated cell lines. Replication protein A 32 kDa, expressed only in the HeLa cell line, is a heterotrimeric protein required for DNA replication, recombination and repair, and has been found to interact with menin, a protein encoded by the tumor suppressor gene, MEN1 (27). Transcription intermediary factor ß-1, expressed only in SK-N-SH cell line, specifically interacts with the Kruppel-associated box (KRAB) domain of the human zinc finger factor and corepresses transcription (28). Although transcription intermediary factor ß-1 per se is not linked to cancer, KRAB-related domains are shown to be disrupted in human sarcomas (29). Nuclear protein Hcc-1, a protein expressed in MCF-7, is a cytokine regulated protein whose overexpression enhances cell proliferation in normal as well as cancer cells (30). Transcription factor BTF3, expressed in CaOva3 and A-549; was found to be altered in human Burkitt lymphoma cell line induced to undergo apoptosis (31). FAB1 (32) and RuvB-like 2 (33), which showed broader expression are also important in *c-myc* oncogenic transformation by binding to single stranded DNA. DNA repair protein RAD50 is one of several recombination repair genes that are required for mitotic repair of DNA damage induced by irradiation. Human RAD50 gene is shown to have two splice variants denoted as RAD50-2 and RAD50-3 (33). The RAD50-3 splice variant has a deduced open reading frame encoding a 1173 amino acid protein with a molecular size of 138 kDa dubbed as truncated RAD-50 protein. Database search (www.expasy.org/tools/sim-prot/html) revealed that truncated RAD50 protein has 100% identity to RAD50-2 from amino acid sequence 140-1312 and 99.4% identity to RAD-50 homologue hsRAD50 (RAD50-1) from amino acid sequence 146 to 1318. Truncated RAD50 lacks the N-terminal ATP-binding domain, which is important for DNA-binding activity of RAD-50 proteins (34,35). Thus, cells expressing truncated RAD50 protein may have defective DNA-repair capacity, indicating the importance of this protein in genetic instability and tumor biology. Another one is zinc finger protein 431, a protein with KRAP and zinc-finger domains and is expressed in A-673. Zinc-finger proteins bearing this domain are shown to be overexpressed in pancreatic carcinoma (37) and serve as transcriptional repressor proteins (29). XAP-5, expressed in A-549 cell line, is encoded by a gene located on Xq28 (38) and contains a XAP-5 domain whose function is uncertain and thought to bind DNA. Taken together, these findings suggest that DNA handling proteins not only be potential diagnostic markers but also treatment targets.

Miscellaneous proteins. Chromobox proteins recognize and bind to histone H3 tails methylated at Lys-9 and seem to be involved in transcriptional silencing through formation of heterochromatin-like complexes. These proteins are demonstrated to be recruited by the KRAP-ZFP-KAP1 complex in gene silencing (38). Since KRAP-like domains are implicated in cancer (28), the involvement of chromobox proteins in cancer is highly likely. Moreover, the limited expression of chromobox protein homolog 3 and 5 in HL-60 and HeLa cell may endow them to be a potential therapeutic avenue for solid tumors. Ribosome-binding protein-1 is an uncharacterized protein thought to act, by similarity, as a ribosome receptor and mediates interaction between the ribosome and the endoplasmic reticulum. The sole expression of this protein in A-375 cell line and its supposed involvement in protein synthesis and sorting calls for further investigation.

In summary, the data demonstrated cell-line specific expression of nucleic acid-binding proteins, which play an essential role in maintenance of the cellular integrity of a multicellular organism through regulation of gene expression. The proteins identified may have future clinical relevance for preventative, diagnostic and therapeutic purposes. Particularly, the protein synthetic machinery represents an attractive target for cancer therapy, since it is activated by a variety of genetic alteration during tumor progression. However, further validation of these findings with several cell lines and human tumor samples is required prior to probable use of these proteins as potential molecular markers or targets in human cancer. A handful of the structures described herein have been also reported in normal cell lines by this protein chemical technique independent of antibody availability and specificity (39-41).

Acknowledgements

We are highly indebted to Professor T. Stroebel (Medical University of Vienna, Institute of Neurology) for supplying ten tumor cell lines. We appreciate the skillful technical assistance of Dr Kiseok Lee (Medical University of Vienna, Department of Pediatrics). We acknowledge contribution of the Oesterreichische Nationalbank, Jubilaeumsfonds to Irene Slavc (no. 9187) and the Forschungsgesellschaft fuer cerebrale Tumore. E.E. is grateful for the permanent support of the Austrian Embassy Development Cooperation, Addis Ababa, Ethiopia.

References

1. Siomi H and Dreyfuss G: RNA-binding proteins as regulators of gene expression. Curr Opin Genet Dev 7: 345-353, 1997.
2. Thisted T, Lyakhov DL and Liebhaber SA: Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and alphaCP-2KL, suggest distinct modes of RNA recognition. J Biol Chem 276: 17484-17496, 2001.
3. Draper DE and Reynaldo LP: RNA binding strategies of ribosomal proteins. Nucleic Acids Res 27: 381-388, 1999.
4. Dreyfuss G, Swanson MS and Pinol-Roma S: Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trends Biochem Sci 13: 86-91, 1988.
5. Burd CG and Dreyfuss G: Conserved structures and diversity of functions of RNA-binding proteins. Science 265: 615-621, 1994.
6. Bradford MM: 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, 1976.
7. Weitzdoerfer R, Fountoulakis M and Lubec G: Reduction of actin-related protein complex 2/3 in fetal Down Syndrome brain. Biochem Biophys Res Commun 293: 836-841, 2002.
8. Yang JW, Czech T and Lubec G: Proteomic profiling of human hippocampus. Electrophoresis 25: 1169-1174, 2004.
9. Berndt P, Hobohm U and Langen H: Reliable automatic protein identification from matrix-assisted laser desorption/ionization mass spectrometric peptide fingerprints. Electrophoresis 20: 3521-3526, 1999.
10. Honore B, Baandrup U and Vorum H: Heterogeneous nuclear ribonucleoproteins F and H/H' show differential expression in normal and selected cancer tissues. Exp Cell Res 294: 199-209, 2004.

11. Pilch B, Allemand E, Facompre M, Bailly C, Riou JF, Soret J and Tazi J: Specific inhibition of serine- and arginine-rich splicing factors phosphorylation, spliceosome assembly, and splicing by the antitumor drug NB-506. Cancer Res 61: 6876-6884, 2001.
12. Krecic AM and Swanson MS: HnRNP complexes: composition, structure and function. Curr Opin Cell Biol 11: 363-371, 1999.
13. Perrotti D and Calabretta B: Post-transcriptional mechanisms in BCR/ABL leukemogenesis: role of shuttling RNA-binding proteins. Oncogene 21: 8577-8583, 2002.
14. Evans JR, Mitchell SA, Spriggs KA, Ostrowski J, Bomsztyk K, Ostarek D and Willis AE: Members of the Poly (rC) binding protein family stimulate the activity of the c-myc internal ribosome entry segment *in vitro* and *in vivo*. Oncogene 22: 8012-8020, 2003.
15. Kim JH, Paek KY, Choi K, Kim TD, Hahm B, Kim KT and Jang SK: Heterogeneous nuclear ribonucleoprotein C modulates translation of c-myc mRNA in a cell cycle phase-dependent manner. Mol Cell Biol 23: 708-720, 2003.
16. Mandal M, Vadlamudi R, Nguyen D, Wang RA, Costa L, Bagheri-Yarmand R, Mendelsohn J and Kumar R: Growth factors regulate heterogeneous nuclear ribonucleoprotein K expression and function. J Biol Chem 276: 9699-9704, 2001.
17. Harris MN, Ozpolat B, Abdi F, Gu S, Legler A, Mawuenyega KG, Tirado-Gomez M, Lopez-Berestein G and Chen X: Comparative proteomic analysis of all-trans-retinoic acid treatment reveals systematic post-transcriptional control mechanisms in acute promyelocytic leukemia. Blood 104: 1314-1323, 2004.
18. Trotta R, Vignudelli T, Candini O, Intine RV, Pecorari L, Guerzoni C, Santilli G, Byrom MW, Goldoni S, Ford LP, Caligiuri MA, Maraia RJ, Perroti D and Calabretta B: BCR/ABL activated mdm2 mRNA translation via the La antigen. Cancer Cell 3: 145-160, 2003.
19. Rhoads RE: Regulation of eukaryotic protein synthesis by initiation factors. J Biol Chem 268: 3017-3020, 1993.
20. Kerekatte V, Smiley K, Hu B, Smith A, Gelder F and De Benedetti A: The proto-oncogene/translation factor eIF4E: a survey of its expression in breast carcinomas. Int J Cancer 64: 27-31, 1995.
21. Eberle J, Krasagakis K and Orfanos CE: Translation initiation factor eIF-4A1 mRNA is consistently overexpressed in human melanoma cells *in vitro*. Int J Cancer 71: 396-401, 1997.
22. Fakuchi-Shimogori T, Ishii I, Kashiwagi K, Mashiba H, Ekimoto H and Igarashi K: Malignant transformation by overproduction of translation initiation factor eIF4G. Cancer Res 57: 5041-5044, 1997.
23. Bauer C, Brass N, Diesinger I, Kayser K, Grasser FA and Meese E: Overexpression of the eukaryotic translation initiation factor 4G-1 (eIF4G-1) in squamous cell lung carcinoma. Int J Cancer 98: 181-185, 2002.
24. Joseph P, O'Kernick CM, Othumpangat S, Lei YX, Yuan BZ and Ong TM: Expression profile of eukaryotic translation factors in human cancer tissues and cell lines. Mol Carcinog 40: 171-179, 2004.
25. Arora S, Yang JM, Kinzy TG, Utsumi R, Okamoto T, Kitayama T, Oritz PA and Hait WN: Identification and characterization of an inhibitor of eukaryotic elongation factor 2 kinase against human cancer cell lines. Cancer Res 63: 6894-6899, 2003.
26. Sukhodolets KE, Hickman AB, Agarwal SK, Sukhodolets MV, Obungu VH, Novotny EA, Crabtree JS, Chandrasekharappa SC, Collins FS, Spiegel AM, Burns AL and Marx SJ: The 32-kDa subunit of replication protein A interacts with menin, the product of the MEN1 tumor suppressor gene. Mol Cell Biol 23: 493-509, 2003.
27. Moosmann P, Georgiev O, Le Douarin B, Bourquin JP and Schaffner W: Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. Nucleic Acids Res 24: 4859-4867, 1996.
28. Lim FL, Soulez M, Koczan D, Thiesen HJ and Knight JC: A KRAB-related domain and a novel transcription repression domain in proteins encoded by SSX gene that are disrupted in human carcinoma. Oncogene 17: 2013-2018, 1998.
29. Fukuda S, Wu DW, Stark K and Pelus LM: Cloning and characterization of a proliferation-associated cytokine-inducible protein, CIP29. Biochem Biophys Res Commun 292: 593-600, 2002.
30. Brockstedt E, Otto A, Rickers A, Bommert K and Wittmann-Liebold B: Preparative high-resolution two-dimensional electrophoresis enables the identification of RNA polmerase B transcription factor 3 as an apoptosis-associated protein in the human BL60-2 Burkitt lymphoma cell line. J Protein Chem 18: 225-231, 1999.
31. Duncan R, Bazar L, Michelotti G, Tomonaga T, Krutzsch H, Avigan M and Levens D: A sequence-specific, single-strand binding protein activates the far upstream element of c-myc and defines a new DNA-binding motif. Genes Dev 8: 465-480, 1994.
32. Wood MA, McMahon SB and Cole MD: An ATPase/helicase complex is an essential cofactor for oncogenic transformation by c-Myc. Mol Cell 5: 321-330, 2000.
33. Kim KK, Shin BA, Seo KH, Kim PN, Koh JT, Kim JH and Park BR: Molecular cloning and characterization of splice variants of human RAD50 gene. Gene 235: 59-67, 1999.
34. Alani E, Padmore R and Kleckner N: Analysis of two wild-type and rad50 mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. Cell 61: 419-436, 1990.
35. Raymond WE and Kleckner N: RAD50 protein of *S. cerevisiae* exhibits ATP-dependent DNA binding. Nucleic Acids Res 21: 3851-3856, 1993.
36. Ferbus DA and Goubin G: Production and characterization of mouse monoclonal antibodies to human zinc finger OZF protein overexpressed in pancreatic carcinoma. Hybridoma 18: 431-436, 1999.
37. Chen EY, Zollo M, Mazzarella R, Ciccodicola A, Chen CN, Zuo L, Heiner C, Burough F, Repetto M, Schlessinger D and D'Urso M: Long-range sequence analysis in Xq28: thirteen known and six candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and G6PD loci. Mol Genet 5: 659-668, 1996.
38. Ryan RF, Schultz DC, Ayyanathan K, Singh PB, Friedman JR, Fredericks WJ and Rauscher FJ III: KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol Cell Biol 19: 4366-4378, 1999.
39. Afjehi-Sadat L, Krapfenbauer K, Fountoulakis M, Frischer T, Slavc I and Lubec G: Validation of hypothetical nucleic acid binding proteins in human bronchial epithelial, mesothelial, amnion, kidney and lymphocyte cell lines by proteomics. Curr Proteomics 1: 297-313, 2004.
40. Peyrl A, Krapfenbauer K, Slavc I, Strobel T and Lubec G: Proteomic characterization of the human cortical neuronal cell line HCN-2. J Chem Neuroanat 26: 171-178, 2003.
41. Yang JW, Suder P, Silberring J and Lubec G: Proteome analysis of mouse primary astrocytes. Neurochem Int (In press).