A reference map of human nasopharyngeal squamous carcinoma proteome

FENG LI[1,2*], ZHIQIANG XIAO[1*], PENGFEI ZHANG[1], JIANLING LI[2], MAOYU LI[1], XUEPING FENG[1], YONGJUN GUAN[2] and ZHUCHU CHEN[1,2]

[1]Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital and [2]Cancer Research Institute, Central South University, Changsha, P.R. China

Received September 18, 2006; Accepted November 27, 2006

Abstract. In order to conduct a comparative proteomics study of human nasopharyngeal carcinoma (NPC) to understand the molecular mechanisms that participate in the formation of NPC, the two-dimensional gel electrophoresis (2-DE) reference map of human NPC tissue proteome was described. To provide a high level of reproducibility between gels and accurately array each protein expressed in NPC tissue proteome, the two-dimensional polyacrylamide gel electrophoresis system, modified colloidal Coomassie Brilliant Blue staining method and ImageMaster 2D Platinum image analysis software were used. The NPC 2-DE maps show that high quality and good reproducibility of the 2-DE gel pattern was attained. An average total of 1,100 protein spots were separated by 2-DE, visualized by a modified colloidal Coomassie Brilliant Blue staining method. A synthesized 2-DE reference gel was acquired after detailed analysis of the NPC 2-DE gel maps, and 216 medium to high abundant spots were identified as landmark spots of NPC 2-DE gel, which expressed on >75% of gels. To provide an unambiguous identification of the landmark spots in gels, MALDI-TOF, ESI-Q-TOF mass spectrometry and database search were used to identify the proteins expressed in NPC tissue proteome. Between the 216 landmark spots, all proteins were identified with MALDI-TOF at first, 41 of which were identified with both MALDI-TOF and ESI-Q-TOF. All identified proteins were classified in terms of their subcellular localization and physiological function with information from SWISS-PROT and NCBI websites. According to our knowledge this is the first 2-DE reference map of human NPC. This reference map will serve as a basis for further studies of human NPC and the reference map data will be used to expand the proteome database of human NPC, which can be accessed in our website http://www.xyproteomics.org/.

Correspondence to: Dr Zhuchu Chen, Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Cancer Research Institute, Central South University, #87 Xiangya Road, Changsha 410078, P.R. China
E-mail: tcbl@xysm.net

*Contributed equally

Abbreviations: NPC, nasopharyngeal carcinoma; CHCA, α-cyano-4-hydroxycinnamic acid; PMF, peptide mass fingerprint

Key words: nasopharyngeal carcinoma, proteome, reference map, two-dimensional gel electrophoresis, mass spectrometry

Introduction

Nasopharyngeal carcinoma (NPC) is one of the most common cancers in southern China and Southeast Asia. This cancer is prevalent in the southern part of China including Guangdong and Hunan Provinces (25-50/100,000) but less so among northern Chinese (3/100,000). Moreover, the Cantonese are the most frequently affected population and have an incidence rate of NPC nearly 100-fold higher than that seen in Caucasians. The marked geographic and racial differences in incidence of NPC indicate that the development of this cancer must be related to special genetic and environmental factors (1-3).

Molecular genetic studies have indicated that inactivation of tumor suppressor genes on 3p, 9p, 11q, 13q, 14q and 16q and alterations of oncogenes on chromosomes 8 and 12 are important in the pathogenesis of NPC. LOH studies also suggested that multiple tumor suppressor loci at 3p13-14.3, 3p21, and 3p21-25 were involved (4-6). In addition to genetic alteration, Epstein-Barr virus (EBV) latent infection plays an important role in pathogenesis of NPC. The detection of a single form of viral DNA in NPC cells suggests that the tumors are clonal proliferations of a single cell that was initially infected with EBV. Environmental factors including traditional diets and early exposure to salted fish have also been reported to be high risk factors with nasopharyngeal carcinoma (7-9).

However, these research topics of NPC mainly focused on a minority of interesting proteins and genes. As we all know, NPC carcinogenesis is a result of complex interaction processes of multiple genes, multiple proteins and multiple factors. The study of a minority of several proteins or genes will not be sufficient to understand the pathogenesis of NPC in a global eye (10). Therefore, a global analysis system such as proteomics is needed to analyze all parameters as far as possible that might participate in the pathogenesis of NPC (10-12).

Proteomics mapping of tissue and cell proteins is the preferred method for investigating the pathogenesis of human NPC. The differentially expressed proteins can be identified through comparing the 2-DE reference gels representing the typical pattern of human NPC with normal nasopharyngeal

epithelia. Because differentially expressed spots among the matched 2-DE gels can be varied especially when there are some inconstant differentially expressed spots, many samples are needed to confirm their validity. On the other hand, the differentially expressed spots may be artifacts that are introduced by gel warping. Thus, in order to conduct a confirmed comparative proteomics study of the NPC, a reference map of NPC proteome is needed. Several 2-DE reference maps are now available on the ExPASy website, including maps of various microorganisms and also those of human organs, cells, and fluids, as well as those of several cell lines (13,14). To our knowledge, no detailed 2-DE reference map of human NPC is currently available. To fill this gap, we constructed a 2-DE reference map of human NPC. In our laboratory, the experimental conditions and reproducibility of 2-DE have been optimized (15), and it was used in this study. A modified colloidal Coomassie Brilliant Blue staining method, which was so called Blue Silver staining, was used to visualize more protein spots in the 2-DE gels and improve compatibility with mass spectrometry (16). MALDI-TOF MS and ESI-Q-TOF tandem mass spectrometry were used at the same time to validate each other and acquire unambiguous protein identification.

On the 2-DE reference gel of human NPC we identified 216 distinct spots. The data are presented in terms of the different functional classes of proteins and their relevance to the NPC proteome. The 2-DE reference map of human NPC will offer a basis for research involved in NPC, and open the way to investigate the NPC protein changes associated with NPC pathogenesis.

Materials and methods

Instruments and software. IPGphor, Ettan DALT II System, ImageScanner (maximum resolution 9600*9600 dpi), ImageMaster 2D Platinum analysis software (Amersham Pharmacia Biotech, USA), Centrifuge (Eppendorf, Germany), Voyager-DE STR MALDI-TOF Mass Spectrometry (Applied-Biosystems, USA), ESI-Q-TOF Mass Spectrometry (Micromass, UK), Mascot Distiller and Mascot Database Search engine and Statistical Package For Social Science (SPSS for windows, version 10.01, USA) were used.

Reagent. Immobiline pH-gradient DryStrips (pH 3-10L, 24 cm), IPG buffer (pH 3-10), DryStrip cover fluids, thiourea, urea, CHAPS, dithiothreitol, Pharmalyte (pH 3-10), bromophenol blue, acrylamide/bis-acrylamide (29:1 = w:w), TEMED, low-molecular-weight marker, Tris-base, SDS and glycine were purchased from Amersham Pharmacia Biotech. Mercaptoethanol, iodoacetamide, HCl, Coomassie Brilliant Blue G-250, and 37% v/v formaldehyde were from Sigma. Acetic acid, glycerol, methanol, sodium thiosulfate, and sodium carbonate were from Fisher. PAGE (12%) resolving gel was cast with a DALT II Gel Caster. All buffers were prepared with Milli-Q water.

Tissue samples (preparation of tissue proteins). Twenty fresh nasopharyngeal carcinoma biopsy specimens were obtained from Xiangya Hospital, Central South University, Hunan Province, P.R. China, and were diagnosed by pathology. These specimens were immediately frozen in liquid nitrogen after excision and flushing out of blood and then stored at -80˚C until analysis. The tissue specimens of low-differentiated squamous cell carcinoma were selected in this study. Five tissue specimens were randomly put into a pool and a total of 85 mg nasopharyngeal tissues were ground into powder in liquid nitrogen and lysed in 400 μl lysis buffer (7 mol/l urea, 2 mol/l thiourea, 100 mM/l DTT, 4% CHAPS, 40 mM/l Tris, 2% pharmalyte, 1 mg/ml DNase I). The lysates were incubated at 4˚C for 1 h and centrifuged at 15,000 rpm for 45 min at 4˚C. The super-natant was transferred to a fresh tube and stored at -80˚C until use. The concentration of the total proteins was determined with the 2D Quantification kit (Amersham Biosciences).

2-DE. 2-DE was performed as described by the manufacturer (Amersham Biosciences) and Görg *et al* (17). Protein samples (800 μg for preparative gels) were mixed with loading buffer for IPG strips to obtain a final volume of 450 μl. First-dimension IEF was performed using IPGstrip (pH 3-10L, 240 mm x 3 mm x 0.5 mm) on an IPGphor isoelectric focusing cell. In the reswelling step of dry IPG strips, 30 v was applied at the same time. Second-dimension SDS–PAGE was performed on an Ettan DALT II system. After electrophoresis, the Blue Silver staining method was used to visualize the protein spots in the 2-DE gels.

Image analysis. The stained 2-DE gels were scanned with MagicScan software on Imagescanner. ImageMaster 2D Platinum analysis software was used for spot intensity calibration, spot detection, background subtraction, and 1-D calibration. After matching analysis of the NPC 2-DE gel maps, a synthesized NPC 2-DE reference gel was acquired. The relative volume (% Vol) of each spot was quantified by calculation of spot volume after normalization of the image using the total spot volume normalization method multiplied by 100. Two hundred and sixteen spots with medium to high relative volume and expressed on >75% gels were identified as landmark spots of the NPC 2-DE synthesized reference gel. The reproducibility of the spot position was calculated with Gorbett's method (18). Statistical analysis was carried out with SPSS for Windows 10.0 and Excel.

Protein identification. Landmark spots were excised from preparative gels using punch and transferred into a 1.5-ml siliconized Eppendorf tube. One protein-free gel piece was treated in parallel as a negative control. Proteins were in-gel digested as previously described. The gel-pieces were destained according to the EMBL Bioanalytical Research Group's in-gel digestion protocol. The dried gel-pieces were incubated in 10 μl digestion solution consisting of 40 mM/l NH_4HCO_3 in 9% acetonitrile solution and 20 μg/ml proteomics grade trypsin for 10-12 h at 37˚C. The tryptic peptide mixture was extracted and purified with a Millipore ZipTipC (18) column.

The samples were analyzed with a MALDI-TOF Mass Spectrometer (ABI) or ESI-Q-TOF mass spectrometry (Micromass). The purified tryptic peptide mixture was mixed with CCA matrix solution and vortexed gently. A volume (1 μl) of the mixture containing CCA matrix was loaded on a stainless steel plate and air-dried. The standard peptide mixture

was spotted at the same time to correct the machine. The parameters of MALDI-TOF were set up as follows: positive ion-reflector mode, 20-kV accelerating voltage, 64.5% grid voltage, 1.12 mirror voltage ratio, 337-nm N_2 laser wavelength, 3-ns pulse width, 50 laser shots, 800- to 3000-Da acquisition mass range, 100-ns delay, and $4x10^{-7}$ Torr vacuum degree.

Some mass spectra were obtained on an ESI-Q-TOF mass spectrometry fitted with a nanoelectrospray ionization source. For on-line LC/MS analysis, a Waters CapLC solvent delivery system was coupled to the mass spectrometry. The samples were loaded to a pre-column [320 μm x 50 mm, 5 μmC (18) silica beads, Waters] at 30 μl/min flow rates for concentrations and fast desalting through a Waters CapLC Autosampler, then eluted to the reversed phase column (75 μm x 150 mm, 5 μm, 100A, LC Packing) at a flow rate of 200 nl/min after flow splitting for separation. MS/MS spectra were performed in data-depended mode in which up to 4 precursor ions above an intensity threshold (7 counts/sec, cps) were selected for MS/MS analysis from each survey 'scan'. The nanoelectrospray parameters were 3,000 V for capillary voltage, 45 V for cone voltage, 80˚C for source temperature and 15 psi collision gas back pressure.

Database analysis. In peptide mass fingerprint (PMF) map database searching, Mascot Distiller was used to get the monoisotopic peak list from the raw mass spectrometry files. The monoisotopic peak list was used to search the MSDB or NCBInr database with the Mascot search engine (http://www.matrixscience.com/). The searching parameters were set up as follows: the taxonomy was selected as all entries or *Homo sapiens*; the mass tolerance was ±200 ppm; missed cleavage sites were allowed up to 1; fixed modifications were selected as carboxymethyl (cysteine); the variable modification was selected as Oxidation (methylation) or none.

In tandem mass spectrometry data database query, the PKL format file that was generated from MS/MS mass spectrometry was imported into the Mascot search engine. The searching parameters were set up as follows: the taxonomy was selected as all entries or *Homo sapiens*; the mass tolerance was ±1.0 Da; the MS/MS tolerance was ±0.5 Da; missed cleavage sites were allowed up to 1; fixed modifications were selected as carboxymethyl (cysteine); the variable modification was selected as Oxidation (methylation) or none; the data format was selected as Micromass PKL format and the Instrument was selected as ESI-QUAD-TOF.

Results and discussion

Protein preparation and separation. In NPC histological pathology, >95% of NPC belong to low differentiated nasopharyngeal squamous cell carcinoma or non-differentiated squamous cell carcinoma. The high malignancy of NPC results in tumors increasing rapidly and metastasizing at the early stage, and >75% of patients were at stage III or IV when diagnosed (2). Because of the importance of low differentiated nasopharyngeal squamous cell carcinoma in NPC, all the cases used in this study were diagnosed by pathology to be low differentiated nasopharyngeal carcinoma. Due to the especially anatomic localization and biological feature of NPC, the weight of NPC biopsy tissues usually will be no higher than 20 mg, so it is very important to recover NPC tissue powers from mortar as much as possible and increase protein extraction efficiency at the same time. To improve the protein solubility of the NPC tissue sample, chemical and physical methods were combined to process the tissues and extract proteins. Based on our previous protocol, some modifications have been made here: to improve protein solubilization, a combination of urea and thiourea was used, and pharmalyte was added to protein solubilization buffer; to remove any indissoluble material before IEF, sufficient time centrifugation (45 min x 15,000 x g, 4˚C) was used.

For the 2-DE system, a 24-cm IPGstrip combined with the vertical multigel electrophoresis system was used in this study. More protein spots can be separated by a 24-cm pH 3-10L IPGstrip than other IPGstrips because of its length. The vertical multigel electrophoresis system, which can analyze up to 12 gels at a time, can provide a higher reproducibility between 2-DE gels. 2-DE gel patterns of a human NPC proteome with high-resolution and high-reproducibility were attained (Fig. 1). An average total of 1,100 protein spots were detected in the 2-DE map shown in Fig. 1 and the average positional deviation of the matched-spots among four replicate 2-DE gels of one pool was 1.55±0.35 mm in the IEF direction, and 1.78±0.55 mm in the SDS-PAGE direction according to Gorbett's calculating method (data not shown).

2-DE map of human nasopharyngeal squamous carcinoma. To reliably identify the differentially expressed proteins in the two-dimensional difference gel electrophoresis, the landmark spots in the 2-DE gel reference maps would be very helpful for the following reasons: a) when the difference spots were low expressed proteins, it was difficult to identify them without the aims of landmark spots; b) at the current 2-DE technology condition gel distortion was difficult to fully discover, which increased the difficulty of image analysis and the need for landmark spots to correct gel distortion; c) the 2-DE image analysis software was not ideal for spot detection or gel-matching, most of the software still needed human eyes to find the matched landmark spots; d) the 2-DE maps would be insufficient for image analysis when the analyzed samples are very complex; and e) the 2-DE gel reference map provided some useful composition information of the analyzed cells. One study showed that even housekeeping proteins such as GAPDH, ß-actin and ß-tubulin varied much between a number of different established renal cancer cell lines, matched pairs of renal tumor and normal kidney lysates as well as nine different human tissues (19). Thus, a good 2-DE reference map of human nasopharyngeal squamous carcinoma with many landmark spots is essential for a reliable comparative proteomics study of different differentiated degrees of nasopharyngeal squamous carcinoma, NPC and normal nasopharyngeal epithelia. In the 2-DE maps shown in Fig. 1, Blue Silver, a modified method of colloidal Coomassie Brilliant Blue staining was used to attain the high quality NPC 2-DE gel maps, and an average of 1,100 protein spots of every gel were detected through this modified staining method by 2-DE gel analysis software ImageMaster 2D Platinum. The green box region was magnified and red arrays showed the matched landmark spots in this region. In Fig. 2, 216 spots highly expressed in 75% of gels and identified as landmark spots in

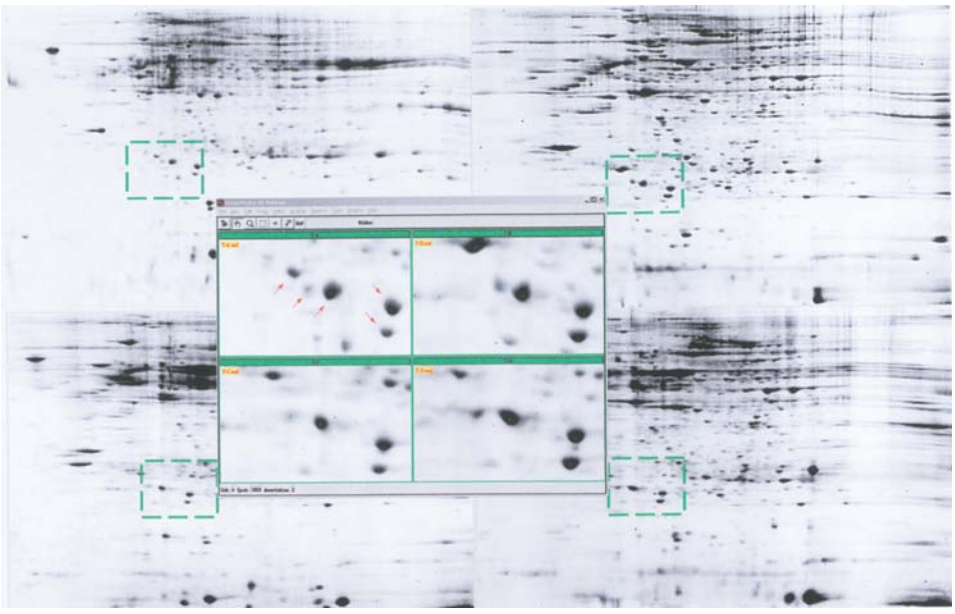

Figure 1. The 2-DE gel maps of NPC and matched landmark spots.

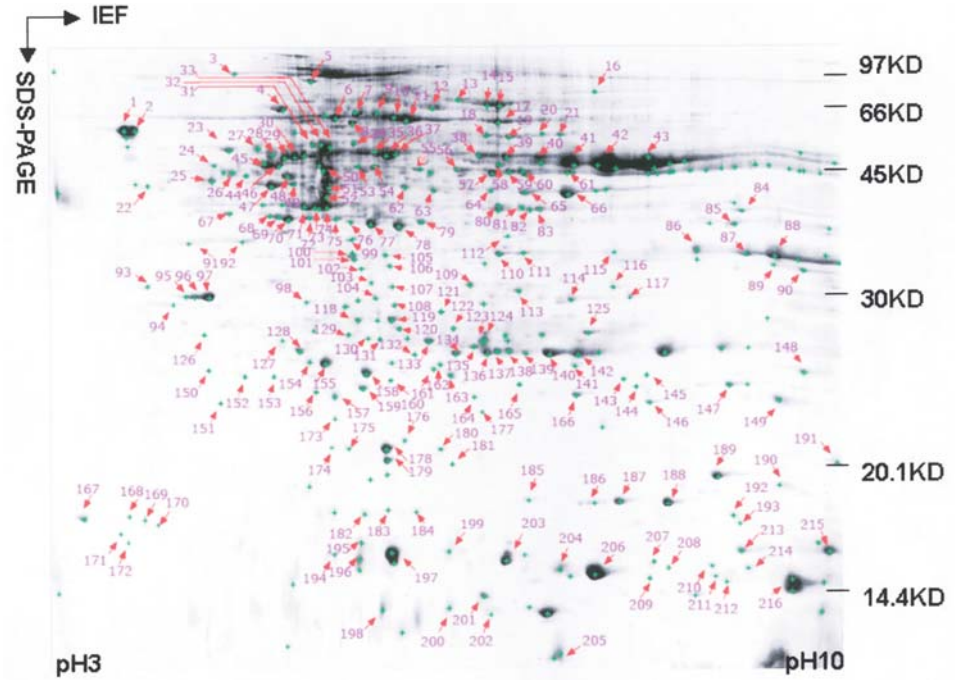

Figure 2. The synthetic 2-DE reference gel map of NPC.

the reference gel were marked with a purple number and red arrow, respectively, in the center of the spots a green plus symbol was added automatically by ImageMaster 2D Platinum.

Undoubtedly, silver staining is a more sensitive staining method than Coomassie Brilliant Blue staining as it can visualize more spots in 2-DE gels, but this method also has problems: first, the linear consistency of protein spot quantity with protein spot volume is not as good as with Coomassie Brilliant Blue staining, so many differential proteome analyses prefer classical organic dyes of the Coomassie family or use fluorescence dye; second, the compatibility of mass spectrometry is not as good as Coomassie Brilliant Blue staining, which will greatly affect the successful identification of the protein spots of the gels, especially low abundance protein spots. Blue Silver, a modification method of colloidal Coomassie Brilliant Blue with a sensitivity close to silver staining and the advantage of linear consistency and mass spectrometry compatibility should be the right choice in 2-DE gel staining.

Protein identification by MS. All spots were identified by MALDI-TOF PMF at first, 41 spots among these spots were identified by MALDI-TOF PMF and ESI-Q-TOF tandem MS/MS simultaneously. The detailed information of identified proteins is listed in Table I, which can also be accessed from

Table I. Detailed information about the identified proteins.

Spot ID	Protein name	AC	pI	MW	Score	Coverage (%)	M/Q	Localization	Function
1	Calreticulin	gil30583735	4.29	48283	188	57	17/48	Membrane	Protein process
2	Calreticulin	gil30583735	4.29	48283	182	55	17/51	Membrane	Protein process
3	Transferrin	gil37747855	6.97	79310	218	31	20/32	Extracellular	Transport
4	BiP protein	gil6470150	5.23	71002	270	50	34/69	Membrane	Energy
5	Valosin containing protein	gil30023842	5.14	89950	105	21	14/35	Membrane	Energy
6	Heat shock 70 kDa protein 9B precursor	gil12653415	6.03	73967	83	40	21/60	Cytoplasm	Cell death/defense
7	Heat shock 70 kDa protein 1B	gil4885431	5.48	70267	191	51	24/57	Cytoplasm	Cell death/defense
8	Serum albumin	gil3212456	5.62	68425	293	56	32/58	Plasma	Other
9	Serum albumin	gil3212456	5.62	68425	290	56	31/57	Plasma	Other
10	Serum albumin	gil3212456	5.62	68425	290	56	31/57	Plasma	Other
11	Serum albumin	gil3212456	5.62	68425	290	56	31/67	Plasma	Other
12	Serum albumin	gil52001697	5.63	67690	104	27	13/32	Plasma	Other
13	PRO1400	gil6650772	6.95	65304	105	30	14/49	Plasma	Transport
14	Transferrin	gil37747855	6.97	79310	326	52	34/55	Extracellular	Transport
15	PRO1400	gil6650772	6.95	65304	163	40	20/47	Plasma	Transport
16	Aconitase 2	gil47678261	7.36	86113	242	33	20/28	Mitochondria	Metabolism
17	Serum albumin	gil55669910	5.57	67174	320	57	31/60	Plasma	Other
18	Serum albumin	gil55669910	5.57	67174	320	57	31/60	Plasma	Other
19	Stress-induced-phosphoprotein 1	gil5803181	6.17	63227	183	52	23/43	Cytoplasm	Cell death/defense
20	α 1 antitrypsin	gil1942629	5.37	44280	89	26	8/22	Extracellular	Cell death/defense
21	Heterogeneous nuclear ribonucleoprotein L	gil133274	6.65	60719	69	18	9/38	Nucleus	Signal cell death
22	α 1 antitrypsin	gil1942629	5.37	44280	82	37	9/30	Extracellular	Cell death/defense
23	Prolyl 4-hydroxylase, ß subunit	gil20070125	4.76	57480	173	40	20/44	ER	Metabolism
24	Retinoblastoma binding protein 4	gil15929379	4.8	40791	69	20	7/21	Nucleus	Signal
25	Vimentin	gil340219	5.03	53738	107	36	17/58	Cytoskeleton	Structural
26	Unnamed protein product	gil28193184	5.1	35332	101	35	8/18	Unknown	Unknown
27	α 1 antitrypsin	gil1942629	5.37	44280	115	31	10/28	Extracellular	Cell death/defense
28	α 1 antitrypsin precursor	gil1942629	5.37	44280	162	37	16/47	Extracellular	Cell death/defense
29	α 1 antitrypsin precursor	gil1942629	5.37	44280	98	34	10/28	Extracellular	Cell death/defense
30	α tubulin 2	gil34740335	4.94	50804	86	42	12/35	Cytoskeleton	Structural
31	Chaperonin 60	gil6996447	5.7	61187	103	28	13/48	Cytoplasm	Protein process
32	Chaperonin	gil31542947	5.7	61187	67	10	5/8	Cytoplasm	Protein process
33	Chaperonin	gil31542947	5.7	61187	67	10	5/8	Cytoplasm	Protein process
34	Protein for IMAGE:3534054	gil14495685	6.09	43461	124	49	10/19	Unknown	Unknown
35	Protein disulfide-isomerase	gil7437388	5.98	57160	156	54	25/89	ER	Protein process
36	Protein disulfide-isomerase	gil2135267	5.98	57115	156	54	25/89	ER	Protein process
37	Protein disulfide-isomerase	gil2135267	5.98	57115	156	54	25/89	ER	Protein process
38	Fibrin ß	gil223002	7.95	51358	180	62	23/75	Plasma	Other
39	Fibrin ß	gil223002	7.95	51358	146	60	18/53	Plasma	Other
40	Fibrinogen ß chain precursor	gil7924018	8.54	56577	146	41	16/49	Plasma	Other
41	Fibrinogen ß chain precursor	gil7924018	8.54	56577	268	67	29/66	Plasma	Other
42	Fibrinogen ß chain precursor	gil7924018	8.54	56577	202	70	34/76	Plasma	Other
43	Fibrinogen ß chain precursor	gil7924018	8.54	56577	212	71	35/76	Plasma	Other
44	α 1 antitrypsin	gil1942629	5.37	44280	156	36	14/35	Extracellular	Cell death/defense
45	Tubulin ß chain	gil338695	4.75	50240	148	41	17/46	Cytoskeleton	Structural

Table I. Continued.

Spot ID	Protein name	AC	pI	MW	Score	Coverage (%)	M/Q	Localization	Function
46	Tubulin ß chain	gil338695	4.75	50240	148	41	17/46	Cytoskeleton	Structural
47	Keratin 16	gil24430192	4.99	51578	152	43	16/47	Cytoskeleton	Structural
48	ATP synthase	gil32189394	5.26	56525	139	56	18/63	Mitochondria	Energy
49	ATP synthase	gil32189394	5.26	56525	93	43	16/79	Mitochondria	Energy
50	Heterogeneous nuclear ribonucleoprotein F	gil16876910	5.38	46013	90	30	9/30	Nucleus	Signal
51	TXNDC5 protein	gil12654715	5.32	36725	98	40	11/55	ER	Unknown
52	TXNDC5 protein	gil12654715	5.32	36725	108	45	14/63	ER	Unknown
53	Keratin 8	gil33875698	5.62	55787	181	28	15/22	Cytoskeleton	Structural
54	Keratin 8	gil4504919	5.62	53671	144	57	22/79	Cytoskeleton	Structural
55	Heterogeneous nuclear ribonucleoprotein H1	gil48145673	5.79	49384	78	29	8/22	Nucleus	Signal
56	Chaperonin containing TCP1	gil5453603	6.01	57794	223	62	24/55	Cytoplasm	Protein process
57	Fibrin ß	gil223002	7.95	51358	94	37	10/33	Plasma	Other
58	Aldehyde dehydrogenase 1	gil2183299	6.3	55427	87	29	10/29	Cytoplasm	Metabolism
59	Cytosol aminopeptidase	gil12643394	6.29	53006	278	59	26/48	Cytoplasm	Structural
60	Cytosol aminopeptidase	gil12643394	6.29	53006	189	38	18/50	Cytoplasm	Structural
61	Fibrin ß	gil223002	7.95	51358	243	82	23/55	Plasma	Other
62	Proteasome 26SATPase subunit 2	gil4506209	5.71	49002	130	53	21/85	Several compartments	Metabolism
63	Fragment double-D from human fibrin	gil28373962	5.84	38081	108	33	9/23	Plasma	Other
64	ENO1 protein	gil29792061	7.01	47481	89	23	8/20	Cytoplasm	Metabolism
65	ENO1 protein	gil29792061	7.01	47481	151	50	18/50	Cytoplasm	Metabolism
66	ENO1 protein	gil29792061	7.01	47481	157	51	14/29	Cytoplasm	Metabolism
67	Keratin 19	gil1070608	5.11	44065	331	78	29/50	Cytoskeleton	Structural
68	Keratin 19	gil1070608	5.11	44065	331	78	29/50	Cytoskeleton	Structural
69	Keratin 19	gil1070608	5.11	44065	331	78	29/50	Cytoskeleton	Structural
70	Keratin 19	gil1070608	5.11	44065	331	78	29/50	Cytoskeleton	Structural
71	ß actin	gil15277503	5.55	40536	119	49	11/30	Cytoskeleton	Structural
72	Keratin 19	gil6729681	5.04	44079	78	18	7/18	Cytoskeleton	Structural
73	ß actin	gil15277503	5.55	40536	119	49	11/30	Cytoskeleton	Structural
74	ß actin	gil15277503	5.55	40536	76	37	9/39	Cytoskeleton	Structural
75	Keratin 19	gil1070608	5.11	44079	80	18	7/18	Cytoskeleton	Structural
76	Fragment double fibrin	gil28373962	5.84	38041	77	22	6/14	Plasma	Other
77	Fragment double-D from human fibrin	gil28373962	5.84	38041	376	92	29/59	Plasma	Other
78	Fragment double-D from human fibrin	gil28373962	5.84	38081	375	85	28/54	Plasma	Other
79	Fragment double-D from human fibrin	gil28373962	5.84	38081	214	62	36/60	Plasma	Other
80	Fragment double-D from human fibrin	gil28373962	5.84	38081	121	41	11/32	Plasma	Other
81	Elongation factor Tu	gil704416	7.7	49851	196	60	18/46	Nucleus	Signal
82	IDH1	gil49168486	6.53	46933	159	46	19/54	Mitochondria	Energy
83	Fragment double-D from human fibrin	gil28373962	5.84	38041	219	73	19/42	Plasma	Other
84	Phosphoglycerate kinase 1	gil4505763	8.3	44985	86	32	12/45	Cytoplasm	Signal
85	Aldolase A, fructose-bisphosphatase	gil4557305	8.49	39720	117	44	13/43	Cytoplasm	Energy
86	Glyceraldehyde-3-phosphate dehydrogenase	gil32891805	8.26	36202	84	45	10/46	Cytoplasm	Energy
87	Glyceraldehyde-3-phosphate dehydrogenase	gil32891805	8.58	36070	90	47	10/40	Cytoplasm	Energy

Table I. Continued.

Spot ID	Protein name	AC	pI	MW	Score	Coverage (%)	M/Q	Localization	Function
88	Glyceraldehyde 3-phosphate dehydrogenase	gil32891805	8.58	36070	90	47	10/39	Cytoplasm	Energy
89	HNRPA2B1 protein	gil32880197	8.97	37464	164	44	13/25	Nucleus	Signal
90	HNRPA2B1 protein	gil14043072	8.97	37464	162	44	13/26	Nucleus	Signal
91	ß tropomyosin	gil6573280	4.7	29980	84	29	11/35	Cytoskeleton	Structural
92	ß tropomyosin	gil6573280	4.7	29980	84	29	11/35	Cytoskeleton	Structural
93	Fibrin α N term fragment	gil223068	9.08	26564	134	53	13/31	Plasma	Other
94	Eukaryotic translation elongation factor 1 ß 2	gil30582997	4.5	24919	104	44	7/12	Nucleus	Signal
95	Tropomyosin 3	gil55665780	4.77	27386	160	43	17/77	Cytoskeleton	Structural
96	Tropomyosin 3	gil55665780	4.77	27386	160	43	17/44	Cytoskeleton	Structural
97	Tropomyosin 3	gil55665780	4.77	27386	160	43	17/44	Cytoskeleton	Structural
98	Chloride intracellular channel 1	gil55961618	5.09	27248	152	58	12/32	Mitochondria	Channel
99	α tubulin	gil37492	5.02	50810	82	26	8/20	Cytoskeleton	Structural
100	F-actin capping protein α-1 subunit	gil5453597	5.45	33073	103	41	7/11	Cytoskeleton	Structural
101	GTP-binding regulatory protein ß-2 chain	GBB2_HUMAN	5.6	38048	66	22	8/36	Several compartments	Signal
102	Inorganic pyrophosphatase	gil33150672	5.54	33095	147	64	12/29	Cytoplasm	Metabolism
103	Annexin A3	gil4826643	5.63	36524	131	40	11/21	Membrane	Signal
104	Proteasome activator subunit 2	gil48734793	5.44	27515	64	27	5/14	Cytoplasm	Protein process
105	Lactate dehydrogenase B	gil54696396	5.72	36769	134	51	15/55	Cytoplasm	Energy
106	Pyruvate kinase	gil35505	7.58	58411	92	31	11/34	Cytoplasm	Energy
107	Protein PP4-X	gil189617	5.65	36232	173	39	12/17	Membrane	Signal
108	PSME1	gil49456277	5.78	28862	84	41	11/41	Cytoplasm	Cell death/ defense
109	Pyruvate dehydrogenase ß subunit precursor	gil189754	6.2	39566	80	43	10/47	Cytoplasm	Energy
110	Annexin A1	gil4502101	6.57	38918	165	55	16/41	Membrane	Signal
111	PDZ and LIM domain 1	gil13994151	6.8	36604	73	41	7/29	Cytoskeleton	Channel
112	Alcohol dehydrogenase	gil1633300	6.34	36761	143	45	11	Cytoplasm	Metabolism
113	Fragment D of fibrinogen	gil49258705	7.66	36287	156	46	13/25	Plasma	Other
114	Glyceraldehyde-3-phosphate dehydrogenase	gil32891805	8.26	36201	126	51	13/43	Cytoplasm	Metabolism
115	Glyceraldehyde-3-phosphate dehydrogenase	gil32891805	8.26	36201	87	46	10/42	Cytoplasm	Metabolism
116	VDAC2 protein	gil15277577	6.81	30849	91	49	10/45	Membrane	Channel
117	Guanine nucleotide binding protein	gil21619296	7.6	35511	75	27	6/19	Cytoplasm	Signal
118	Prohibitin	gil46360168	5.57	29859	166	69	15/47	Cytoplasm	Signal
119	Proteasome activator subunit 1	gil54695544	5.78	28876	96	45	12/41	Several compartments	Protein process
120	6-phosphogluconolactonase	gil6018458	5.7	27815	63	30	6/29	Cytoplasm	Energy
121	Glutathione transferase ω 1-1	gil55925946	6.23	27833	99	28	7/11	Cytoplasm	Cell death/ defence
122	Endoplasmic reticulum protein 29 precursor	gil5803013	6.77	29032	75	26	6/13	ER	Protein process
123	Macropain subunit ι	gil296736	5.58	25078	80	56	12/36	Several compartments	Metabolism
124	IgG κ light chain	BAC01688	6.15	28989	66	40	7/17	Plasma	Other
125	PRO2044	gil6650826	6.97	30084	161	56	12/20	Plasma	Unknown
126	Proteasome endopeptidase complex ζ chain	gil88168	4.74	26579	82	48	8/39	Several compartments	Metabolism
127	rho GDP dissociation inhibitor α	gil30582607	5.02	23250	79	42	8/16	Cytoplasm	Signal

Table I. Continued.

Spot ID	Protein name	AC	pI	MW	Score	Coverage (%)	M/Q	Localization	Function
128	rho GDP dissociation inhibitor ß	gil20379030	5.1	23030	88	69	12/41	Cytoplasm	Signal
129	γ actin	gil40226101	5.5	29687	92	55	9/21	Cytoskeleton	Structural
130	Heat shock protein 27	gil662841	7.83	22313	76	28	6/19	Cytoplasm	Cell death/defense
131	Thioredoxin peroxidase	gil5453549	5.86	30749	88	54	10/28	Cytoplasm	Signal
132	GRB2	gil47496673	6.08	25246	121	45	9/16	Nucleus	Signal
133	Heat shock 27 kDa protein 1	gil54696638	5.98	22826	124	56	14/69	Cytoplasm	Cell death/defense
134	Serum albumin	gil37222202	5.25	19788	117	68	11/46	Plasma	Other
135	Peroxiredoxin 6	gil56204402	6	25133	82	45	8/30	Cytoplasm	Cell death/defense
136	Triosephosphate isomerase	gil999893	6.51	26807	69	48	10/35	Cytoplasm	Energy
137	Triosephosphate isomerase	gil999893	6.51	26807	125	68	15/38	Cytoplasm	Energy
138	Triosephosphate isomerase	gil999893	6.51	26087	98	70	15/61	Cytoplasm	Energy
139	Triosephosphate isomerase	gil999893	6.51	26807	101	72	16/61	Cytoplasm	Energy
140	Triosephosphate isomerase	gil999893	6.51	26807	119	75	17/55	Cytoplasm	Energy
141	HES1	gil1655594	8.5	28485	88	47	8/15	Nucleus	Signal
142	Triosephosphate isomerase	gil999893	6.51	26807	73	72	16/92	Cytoplasm	Energy
143	Glyceraldehyde 3-phosphate dehydrogenase	G3P2_HUMAN	8.58	36070	64	25	5/15	Cytoplasm	Energy
144	Peroxiredoxin 1	gil55959887	6.41	19135	104	53	7/16	Cytoplasm	Cell death/defense
145	Flavin reductase	gil32891807	7.13	22219	102	64	7/18	Cytoplasm	Metabolism
146	Neuropolypeptide h3	gil913159	7.42	21027	116	62	10/18	Cytoplasm	Signal
147	Peroxiredoxin 1	gil55959887	8.58	19135	130	61	9/19	Cytoplasm	Cell death/defense
148	FGA protein	gil47124990	8.57	33426	72	33	9/34	Unknown	Unknown
149	FGA protein	gil47124990	8.57	33426	94	39	13/44	Unknown	Unknown
150	ALB protein	gil27692693	5.97	48641	102	24	9/21	Plasma	Other
151	Proteasome ß subunit	gil56207312	4.92	22428	75	39	7/27	Several compartments	Protein process
152	ALB protein	gil27692693	5.97	48641	114	25	10/20	Plasma	Other
153	Glyoxalase I	gil15030212	5.12	20955	122	61	9/23	Cytoplasm	Metabolism
154	Proapolipoprotein	gil178775	5.45	28944	227	71	19/37	Plasma	Transport
155	Apolipoprotein A1	gil37499465	5.56	30759	188	66	20/55	Plasma	Transport
156	ATP synthase D chain	gil5453559	5.21	18537	95	68	10/47	Mitochondria	Energy
157	ALB protein	gil37222202	5.25	19788	169	76	14/43	Plasma	Other
158	Glutathione transferase	gil2204207	5.43	23595	119	60	11/42	Cytoplasm	Cell death/defense
159	Peroxiredoxin 2	gil61362999	5.66	22048	70	28	7/30	Cytoplasm	Cell death/defense
160	Glutathione peroxidase	gil577062	6.15	22193	85	37	6/15	Cytoplasm	Cell death/defense
161	Peroxiredoxin 3	AAH02685	7.67	28017	88	54	8/30	Cytoplasm	Cell death/defense
162	Rhoa Complexed with Gtp Analogue	gil3318980	5.07	20842	87	46	6/14	Cytoplasm	Signal
163	RNA-binding protein regulatory subunit	AAC1208	6.33	20050	168	86	15/60	ER	Signal
164	cdc42hs-Gdp Complex chain B	gil4389380	5.78	21434	147	65	10/23	Nucleus	Signal
165	Proteasome endopeptidase complex ß chain	S55040	6.51	22993	79	26	6/8	Several compartments	Protein process
166	MN-Superoxiddismutase	CAA01016	7.96	19945	111	58	10/20	Cytoplasm	Cell death/defense
167	Calmodulin 2	gil14250065	4.1	16826	76	56	9/60	Cytoplasm	Signal
168	Myosin light chain 3	gil188590	4.42	17091	64	28	5/20	Cytoskeleton	Structural
169	Myosin alkali light chain 6	gil17986264	4.51	17717	88	51	7/33	Cytoskeleton	Structural
170	Myosin alkali light chain 6	gil17986264	4.51	17717	80	51	7/35	Cytoskeleton	Structural
171	Calmodulin like 5	AAH39172	4.34	15881	96	52	5/7	Cytoplasm	Signal
172	Almodulin related-protein NB-1	CALL_HUMAN	4.3	16806	91	47	5/7	Cytoplasm	Signal
173	Sorcin	S5209	5.32	21947	78	41	8/41	Cytoplasm	Unknown
174	PACAP protein	gil18204192	5.37	21023	78	57	7/31	Unknown	Unknown

Table I. Continued.

Spot ID	Protein name	AC	pI	MW	Score	Coverage (%)	M/Q	Localization	Function
175	Haptoglobin α and ß-chain	CAA25248	6.25	42126	84	22	8/23	Plasma	Transport
176	Nucleoside-diphosphate kinase nm23-H1g	A33386	5.83	17309	88	51	7/25	Nucleus	Structural
177	Ras-related protein rap-1a rap, chain A	A34655	5.65	21040	84	35	7/17	Cytoplasm	Signal
178	Haptoglobin α and ß chain	CAA25248	6.25	42126	68	62	8/34	Plasma	Transport
179	Superoxide dismutase	CAE99479	5.7	16023	73	64	7/35	Cytoplasm	Cell death/defense
180	Haptoglobin α and ß chain	CAA25248	6.25	42126	81	20	7/19	Plasma	Transport
181	LMW-PTP	PPAC_HUMAN	6.35	18356	148	74	10/21	Cytoplasm	Signal
182	Transthyretin, chain A	2ROXA	5.33	12939	182	86	10/20	Nucleus	Signal
183	AIF-1	Q9UKS9	6.63	16792	100	57	7/16	Nucleus	Signal
184	Ubiquitin-protein ligase	JC4894	6.13	17127	77	42	6/21	Extracellular	Signal
185	Cyclophilin A	PPIA_HUMAN	7.68	17870	102	35	7/17	Cytoplasm	Structural
186	Peroxiredoxin 5	gil46015020	7.94	18327	146	69	11/34	Cytoplasm	Cell death/defense
187	Cyclophilin A	PPIA_HUMAN	7.82	18098	77	54	9/53	Cytoplasm	Structural
188	Cyclophilin A	AAH07104	7.82	18228	87	66	11/63	Cytoplasm	Structural
189	Cofilin	P23528	8.22	18371	147	67	12/33	Cytoskeleton	Structural
190	Destrin	1AK6	8.05	19794	90	36	6/10	Cytoskeleton	Structural
191	PPIB	gil48145775	9.42	23771	104	38	7/17	Cytoplasm	Structural
192	IgG Fc-Fragment, Chain A	gil28373341	7	25422	75	46	8/21	Plasma	Other
193	IGHG1 protein	gil49522738	8.35	51757	70	23	8/18	Plasma	Other
194	Calgranulin B	gil4506773	5.71	13291	101	87	7/30	Cytoplasm	Signal
195	MRP-14 Protein	CAA00999	5.55	12770	77	44	5/7	Cytoplasm	Signal
196	Calgranulin	B31848	5.71	13291	126	81	8/22	Cytoplasm	Signal
197	MRP-14 Protein	CAA00999	5.55	12770	81	84	8/48	Cytoplasm	Signal
198	Serum amyloid A1 protein precursor	gil40316910	6.28	13581	86	66	7/38	Plasma	Other
199	Unknown Protein	AAH14308	8	23322	69	34	6/25	Unknown	Unknown
200	IgG κ chain c region	A37927	5.61	11161	76	58	5/23	Plasma	Other
201	ß2 microglobulin	1HLAM	6.46	11592	63	52	5/27	Plasma	Other
202	Calgranulin A	BCHUCF	6.51	10828	112	76	8/24	Several compartments	Signal
203	Unknown Protein	AAH14308	8	23322	78	28	5/13	Unknown	Unknown
204	Hemoglobin ß chain (deoxy)	1A3NB	6.82	15872	172	91	13/46	Plasma	Transport
205	Tetraubiquitin Chain B	1TBEB	5.73	8176	74	65	5/25	Several compartments	Protein process
206	Hemoglobin ß chain (deoxy)	1A3NB	6.82	15872	163	91	14/70	Plasma	Transport
207	PRO2675	AAF69644	6.14	33466	65	22	6/17	Plasma	Unknown
208	Cytoplasmic actin	CAA23745	7.85	14207	68	70	6/37	Cytoplasm	Structural
209	Hemoglobin δ chain	HDHU	7.97	16028	163	91	13/46	Plasma	Transport
210	PRO2675	AAF69644	6.14	33466	68	22	6/15	Plasma	Unknown
211	Hemoglobin δ chain	HDHU	7.97	16028	86	71	9/63	Plasma	Transport
212	Hemoglobin α chain	1DSHA	8.07	15018	76	56	6/29	Plasma	Transport
213	Profilin, chain A	1AWIA	8.46	15014	137	24	10/37	Cytoskeleton	Structural
214	Hemoglobin δ chain	HDHU	7.97	16028	99	75	10/58	Plasma	Transport
215	Lysozyme	134L	9	15120	102	76	10/53	Cytoplasm	Cell death/defense
216	α2 globin	gil1335076	8.73	15174	115	93	10/43	Plasma	Transport

ER, endoplasmic reticulum.

our website http://www.xyproteomics.org/en/. SpotID represented the protein spot ID on the gel, protein name or AC number represented protein name or AC number in MSDB or NCBInr database. Theoretical pI, theoretical Mw,

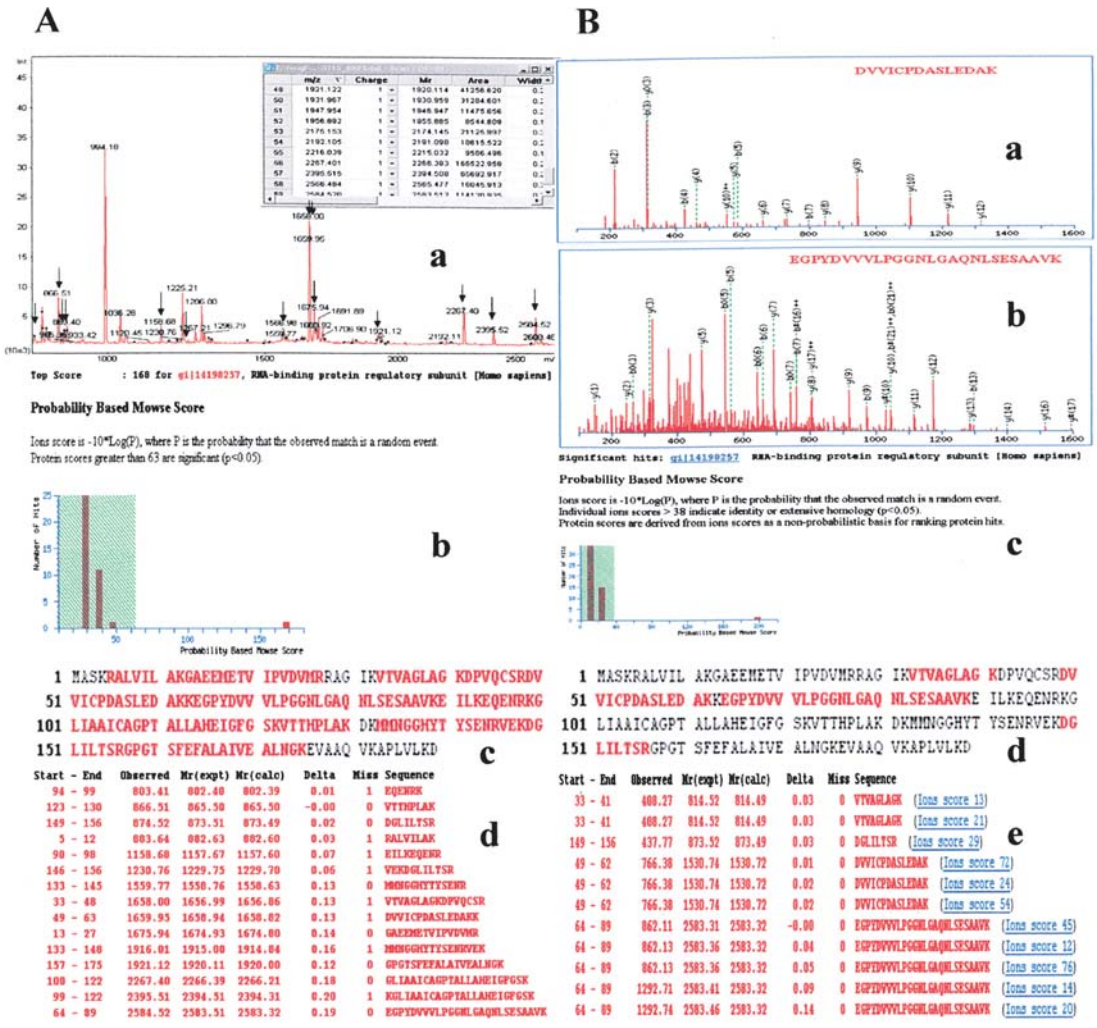

Figure 3. MALDI-TOF, Q-TOF mass spectrometry map and database query result of spot 135. (A) MALDI-TOF mass spectrometry map and database query result of spot 135. a, Monoisotopic peaks of MALDI-TOF mass spectrometry map, arrays showed the matched monoisotopic peak; b, database query result and scores; c, protein sequence and coverage sequence; d, matched peptide fragment and error between experimental mass and theoretical mass. (B) MS/MS analysis and database query result of the spot 135. a, A peptide that contains 14 amino acids; b, a peptide that contains 26 amino acids; c, database query result and its scores; d, protein sequence and its coverage sequence; e, matched peptide fragment and its sequence.

Sequence Coverage and scores of database querying are also provided. Furthermore, the number of matched query and number of input monoisotopic peaks are provided in Table I.

A representative MALDI-TOF mass spectrometry map and database query result are shown in Fig. 3A. A total of 60 monoisotopic peaks were input into MASCOT PMF database search program (http://www.matrixscience.com/cgi/search form.pl?FORMVER = 2& SEARCH = PMF). The matched peptide fragment number is 15, the database query score is 168 and the query result shows that protein is RNA-binding protein regulatory subunit. The sequence of matched peptide fragment, start-end position and the error between theoretical peptide mass and experimental peptide mass are also shown in Fig. 3A. Illustrated as this example, a high quality of PMF mapping is an essential fundament to successfully query databases. With the high quality of peptide mass fingerprint mapping, a default setting can be used to process the PMF map to get the monoisotopic peak lists without filtering the contaminated peak lists. On the other hand, it can be more easy to attain an unambiguous result with a probability based algorithm such as MASCOT than using software based simply on scoring algorithm (20). The MS/MS results of the same spot are shown in Fig. 3B, an 8-amino acid peptide and a 26-amino acid peptide were sequenced. Although the sequence coverage of the tandem mass spectrometry identified peptide fragments is lower than that acquired by MALDI-TOF mass spectrometry, the score of the tandem mass spectrometry identified result is higher than the score of PMF. Part of the coverage sequence identified by both two kinds of mass spectrometry is identical, which illuminate that the identified result is unambiguous.

Subcellular localization and functional groups of the identified proteins. The 216 identified proteins of NPC have been classified in terms of their subcellular localization (Fig. 4A) and physiological function (Fig. 4B) using information from SWISS-PROT and NCBI websites. The majority of identified proteins of NPC (Fig. 4A) are localized in the cytoplasm (34%), cytoskeleton (14%) or several major cellular compartments such as nucleus (6%) and endoplasmic reticulum (ER, 4%) or several compartments (4%). Only 4% of these proteins were membrane proteins and 3%

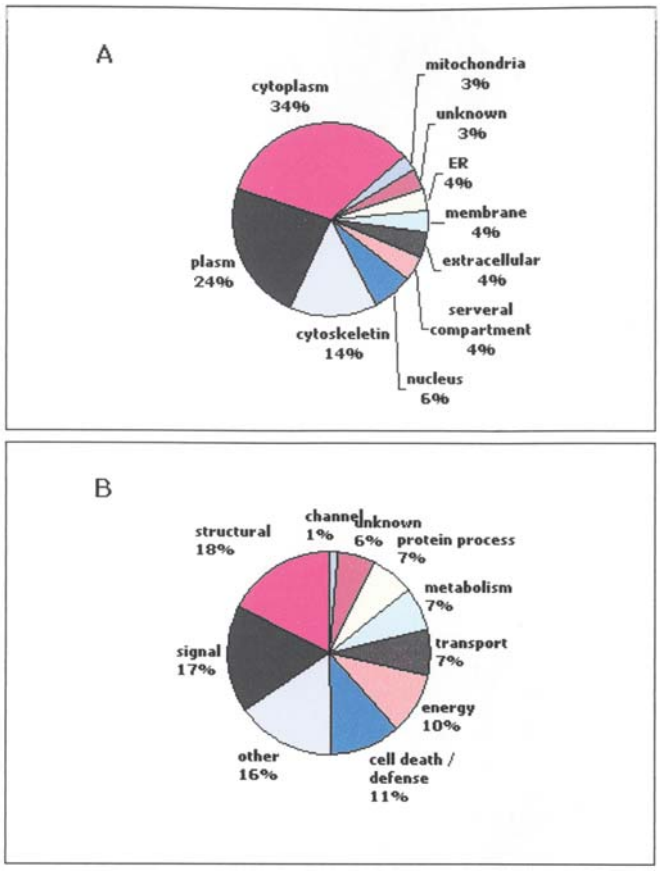

Figure 4. Pie charts representing (A) the distribution of the identified 201 proteins in NPC proteome according to their cellular localization and (B) the distribution of the 201 identified proteins in NPC proteome according to their biological function. Assignments were made on the basis of information provided on the NCBI or ExPASy websites. ER, endoplasmic reticulum.

mitochondria, which shows membrane proteins need more stringent protocol to be extracted. Moreover, a relative part of identified proteins were extracellular proteins (4%), such as vimentin that was identified in secretome. Among them, 24% identified proteins were plasma proteins, and 27 protein spots were identified as hemoglobin and fibrinogen or their isoforms, which indicates that the use of more pure nasopharyngeal carcinoma tissue would help to decrease the contamination from blood.

A large group of the identified proteins are subcellular defense proteins (11%), such as peroxiredoxin 1, 2, 3, 5, 6, the different isoforms of peroxiredoxin, which are involved in redox regulation of the cellular function and are associated with various biological processes, such as the detoxification of oxidants, cell proliferation, cell differentiation and gene expression (21). Heat shock proteins 27, 70, 9B precursor are also in this class, which are produced at a high level by cells to protect themselves against unfavorable conditions. Seven percent of the proteins identified are metabolic enzymes such as triosephosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase, α enolase, etc., which are involved in several metabolic pathways. These metabolic enzymes were also found in other expression proteome studies, which showed they were housekeeping proteins (22,23). Structural proteins are a large part of the identified proteins (18%), for example, actin, tubulin, and keratin. Several actin related proteins were also identified such as profilin, destrin and F-actin capping protein α-1 subunit. Some of these actin-binding proteins are essential for the reorganization of actin filaments as a cellular response to various growth factors, and which could play crucial roles in human disorders (24). Seventeen percent of the identified proteins are involved in signal transduction proteins, for example, phosphoglycerate kinase 1, calgranulin, and ras-related protein rap-1a. Several heterogeneous nuclear ribonucleoproteins were also identified as signal transduction proteins, such as heterogeneous nuclear ribonucleoprotein F, H1, L. More detailed information about the function and localization of the identified proteins is showed in Table I.

In conclusion, in the present study, we constructed a 2-DE reference map of human nasopharyngeal squamous carcinoma, including 216 identified proteins. To our knowledge, this is the first 2-DE reference map of human NPC. This reference map will serve as a basis for further studies of NPC, especially the proteomic comparison among different cell types of NPC, different differentiated degrees of NPC, different stages in the process of NPC, or between the NPC tissue and normal nasopharyngeal epithelial tissue. The reference map data will be used to expand the proteome database of the human NPC, which can be accessed on our website (http://www.xyproteomics.org/en). Furthermore, the mapping data also provide basic data for integrating the genomic, transcriptomic and proteomic research results to reveal the molecular mechanism of NPC and to screen tumor-related markers.

Acknowledgments

This study was supported by a grant from National 973 Project of China (2001CB510208), for Outstanding Scholars of New Era from Ministry of Education of China (2002-48), National Natural Science Foundation of China (30000028, 30240056, 30370642), and key research program from Science and Technology Committee of Hunan, P.R. China (02SSY2001-1), and key research program from Public Health Bureau of Hunan Province, P.R. China (Z02-04).

References

1. Yu MC and Yuan JM: Epidemiology of nasopharyngeal carcinoma. Semin Cancer Biol 12: 421-429, 2002.
2. Licitra L, Bernier J, Cvitkovic E, Grandi C, Spinazze S, Bruzzi P, Gatta G and Molinari R: Cancer of the nasopharynx. Crit Rev Oncol Hematol 45: 199-213, 2003.
3. Huang T, Liu Q, Huang H and Cao S: Study on genetic epidemiology of nasopharyngeal carcinoma in Guangdong, China. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 19: 134-137, 2002.
4. Lo KW and Huang DP: Genetic and epigenetic changes in nasopharyngeal carcinoma. Semin Cancer Biol 12: 451-462, 2002.
5. Cheng Y, Poulos NE, Lung ML, Hampton G, Ou B, Lerman MI and Stanbridge EJ: Functional evidence for a nasopharyngeal carcinoma tumor suppressor gene that maps at chromosome 3p21.3. Proc Natl Acad Sci USA 95: 3042-3047, 1998.
6. Chow LS, Lo KW, Kwong J, To KF, Tsang KS, Lam CW, Dammann R and Huang DP: RASSF1A is a target tumor suppressor from 3p21.3 in nasopharyngeal carcinoma. Int J Cancer 109: 839-847, 2004.
7. Raab-Traub N: Epstein-Barr virus in the pathogenesis of NPC. Semin Cancer Biol 12: 431-441, 2002.
8. Griffin BE: Epstein-Barr virus (EBV) and human disease: facts, opinions and problems. Mutat Res 462: 395-405, 2000.

9. Young LS and Murray PG: Epstein-Barr virus and oncogenesis: from latent genes to tumours. Oncogene 22: 5108-5121, 2003.
10. Baak JP, Path FR, Hermsen MA, Meijer G, Schmidt J and Janssen EA: Genomics and proteomics in cancer. Eur J Cancer 39: 1199-1215, 2003.
11. Hanash SM, Bobek MP, Rickman DS, Williams T, Rouillard JM, Kuick R and Puravs E: Integrating cancer genomics and proteomics in the post-genome era. Proteomics 2: 69-75, 2002.
12. Hanash SM: Global profiling of gene expression in cancer using genomics and proteomics. Curr Opin Mol Ther 3: 538-545, 2001.
13. Sanchez JC, Appel RD, Golaz O, Pasquali C, Ravier F, Bairoch A and Hochstrasser DF: Inside SWISS-2DPAGE database. Electrophoresis 16: 1131-1151, 1995.
14. Hoogland C, Sanchez JC, Walther D, Baujard V, Baujard O, Tonella L, Hochstrasser DF and Appel RD: Two-dimensional electrophoresis resources available from ExPASy. Electrophoresis 20: 3568-3571, 1999.
15. Li C, Chen Z, Xiao Z, Wu X, Zhan X, Zhang X, Li M, Li J, Feng X, Liang S, Chen P and Xie JY: Comparative proteomics analysis of human lung squamous carcinoma. Biochem Biophys Res Commun 309: 253-260, 2003.
16. Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, Orecchia P, Zardi L and Righetti PG: Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis 25: 1327-1333, 2004.
17. Gorg A, Obermaier C, Boguth G and Weiss W: Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12.0, longer separation distances and simplified procedures. Electrophoresis 20: 712-717, 1999.
18. Corbett JM, Dunn MJ, Posch A and Gorg A: Positional reproducibility of protein spots in two-dimensional polyacrylamide gel electrophoresis using immobilised pH gradient isoelectric focusing in the first dimension: an interlaboratory comparison. Electrophoresis 15: 1205-1211, 1994.
19. Ferguson RE, Carroll HP, Harris A, Maher ER, Selby PJ and Banks RE: Housekeeping proteins: a preliminary study illustrating some limitations as useful references in protein expression studies. Proteomics 5: 566-571, 2005.
20. Perkins DN, Pappin DJ, Creasy DM and Cottrell JS: Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20: 3551-3567, 1999.
21. Kinnula VL, Paakko P and Soini Y: Antioxidant enzymes and redox regulating thiol proteins in malignancies of human lung. FEBS Lett 569: 1-6, 2004.
22. Zhan X and Desiderio DM: A reference map of a human pituitary adenoma proteome. Proteomics 3: 699-713, 2003.
23. Dupont A, Tokarski C, Dekeyzer O, Guihot AL, Amouyel P, Rolando C and Pinet F: Two-dimensional maps and databases of the human macrophage proteome and secretome. Proteomics 4: 1761-1778, 2004.
24. Dos Remedios CG, Chhabra D, Kekic M, Dedova IV, Tsubakihara M, Berry DA and Nosworthy NJ: Actin binding proteins: regulation of cytoskeletal microfilaments. Physiol Rev 83: 433-473, 2003.