CANCER GENOMICS & PROTEOMICS *3*: 355-368 (2006)

Proteomic Analysis of the MCF7 Breast Cancer Cell Line

JULIE HARDOUIN*, LUDOVIC CANELLE*, CÉLINE VLIEGHE,
JEAN-PAUL LASSERRE, MICHEL CARON and RAYMONDE JOUBERT-CARON

Laboratory of Protein Biochemistry and Proteomics, UMR CNRS 7033 (BioMoCeTi), UFR SMBH, Paris13 University, 93017 Bobigny Cedex, France

Abstract. *Background: The MCF7 breast cancer cell line is a cellular model for breast cancer studies and marker discovery. Therefore, a better knowledge of its proteome is a prerequisite for a more efficient use of this model. Materials and Methods: Proteins expressed during the exponential growth phase of MCF7 cells were analyzed and mapped using two-dimensional gel electrophoresis and mass spectrometry. Results: From the spots excised from preparative gels of whole-cell extracts, a subset of 368 different polypeptides, corresponding to 249 different proteins, was identified. These polypeptides were positioned on a silver-stained gel to construct a reference map. Conclusion: The data allowed the construction of the most extensive reference map for MCF7 published to date, with 189 novel proteins, which had not been previously listed on maps, and are now accessible on World 2D-PAGE database, providing a basis for further studies on MCF7.*

Breast cancer is the second leading cause of cancer death in women (1, 2). The MCF7 cell line is commonly used as a cellular model for breast cancer studies and marker discovery (3-7), hence information on the MCF7 proteome is a prerequisite for a better understanding of this model. Proteomics has become a popular approach for the analysis, identification and characterization of proteins, combining large-scale protein separation with mass spectrometry and bioinformatics (8-10). Two-dimensional gel electrophoresis (2-DE) (11) remains the most frequently used approach for global protein analysis and the identification of protein isoforms, because it allows the separation and comparison of several thousand proteins on a single gel (12). Consequently, a number of dedicated 2-DE image databases have been constructed, providing maps for different cells and tissues (13, 14). However, few up-to-date informative maps dedicated to breast tissue or breast cancer cell lines are available on the web (http://www.bio-mol.unisi.it/2d/2d.html). This study used a global proteomic approach based on a combination of 2-DE protein separation, image analysis and mass spectrometry analysis (MALDI-TOF) for an extensive identification of 368 polypeptides, corresponding to 249 proteins and 56 isoforms. The use of bioinformatics tools provided useful information on functional pathways and the cellular localization of the identified proteins.

*Both authors contributed equally to this work.

Correspondence to: Raymonde Joubert-Caron, UMR CNRS 7033, Protein Biochemistry and Proteomics Laboratory, UFR SMBH Léonard de Vinci, 74 rue Marcel Cachin, F-93017 Bobigny Cedex, France. Tel: +33 148 387 754; Fax: +33 148 387 313, e-mail: rm_caron@smbh.univ-paris13.fr

Key Words: MCF7 cell line, proteome, breast cancer, mass spectrometry, 2-DE databases.

Materials and Methods

Two-dimensional gel electrophoresis of MCF7 protein extract. The cell pellet of MCF7 cells grown to confluence, prepared as described previously (15, 16), was resuspended at 0˚C in lysis buffer (25 mM Tris, 10 mM EDTA, 7 M urea, 2 M thiourea, 5% v/v glycerol, 0.33% v/v CHAPS, 0.35% v/v Triton X100, 0.35% w/v sulfobetaine 3-10, 10% v/v isopropanol, 12.5% v/v isobutanol, 100 mM DTT, 1 mM orthovanadate and protein inhibitor cocktail) (17). The lysate was then centrifuged twice at 150,000 xg for 25 min at 4˚C.

For 2-DE, isoelectric focusing (IEF) (15, 16) was carried out on 17 cm IPG strips in a linear pH gradient (pH 5-8) using the Protean IEF Cell system (Bio-Rad, Richmond, CA, USA). Analytical-run IPG strips were rehydrated with 150 μg of protein in rehydration medium and 2% v/v carrier ampholyte mixture (Pharmalytes 5-8, GE Healthcare, Uppsala, Sweden). Focalization was complete at 90,000 Vhs. Preparative IPG strips were loaded with 500 μg of protein and 2% v/v carrier ampholyte by in-gel passive rehydration. IEF was then performed until a total of 195,000 Vhs was reached. After focusing, analytical and preparative gel strip equilibration occurred in 6 M urea, 60 mM SDS, 65 mM DTT, 30% v/v glycerol and 0.05 M Tris-HCl pH 8.6 for 15 min, then for a further 20 min with 53 mM iodoacetamide instead of DTT. The two-dimensional gels (18 cm x 20 cm x 1 mm, 8-18.5 polyacrylamide) were run in a Protean Plus Dodeca cell from Bio-Rad. Analytical gels were stained with silver nitrate as described previously (16), while preparative gels were stained with Colloidal

Coomassie blue. Analysis of protein abundance and experimental determination of *p*I and M*r* were done using Image Master Platinum 5 software (GE Healthcare, Uppsala, Sweden).

Mass spectrometric analysis of MCF7 protein spots. Individual spots were excised from a preparative gel using the Proteineer SP spot picker (Bruker Daltonique, Wissembourg, France). In-gel trypsin digestion was carried out using the GE Healthcare Ettan Digester, as described previously (18). A sandwich spotting method was used for peptide mass fingerprinting (PMF), with α-cyano-4-hydroxy-cinnamic acid (HCCA) as matrix, on a Biflex IV (Bruker Daltonique) instrument (16, 18). Using the Mascot Server software package with the Mascot Daemon client application (Matrix Science Ltd, London, UK), mass spectra of tryptic digest peptides were sought in batch mode against a downloaded copy of the biweekly updated UniProtKB/Swiss-Prot database (http://www.expasy.ch/sprot/download.html). The key parameters used were: taxonomy: *Homo sapiens* (human); fixed modification: carbamidomethyl (C); variable modification: methionine, monoisotopic; peptide mass tolerance: 0.2 Da; peptide charge state: +1; maximum missed cleavages: 1. Functional annotation of the set of identified proteins was performed using the FatiGO tool available on Babelomics (http://babelomics.bioinfo.cipf.es/index.html).

Results and Discussion

Protein identification and 2-DE map. A total of 727 picked spots were subjected to trypsinolysis and the resulting peptide fragments subjected to PMF identification using MALDI-TOF MS. Three hundred and sixty eight spots, showing a large range of molecular weights, *p*I and spot intensity, were successfully identified (Table I). The identification rate was approximately 47% (368 out of 727 spots). This rate was lower than that usually obtained in our mapping experiments (14, 19). This is probably related to the use of robotic spot excision instead of the manual picking previously used in our lab. Manual picking favors the excision of large spots, which are more visible on the gels, while automation allows systematic picking independently of the size of the spot.

It is well known that accurate identification *via* PMF is favored by a sufficient number of peptides for matching. Thus, the relationship between the rate of identification and the protein abundance (%Vol) for the 727 spots subjected to trypsinolysis was examined (Table II). The correlation coefficient (r^2) between these two parameters was 0.968 (Figure 1). The rate of successful identification was 92.3% for the spots with a %Vol >1, while it was less than 48% for the spots with a %Vol ranging from 0.05% to 0.09%. These results confirmed that the amount of protein found in a given spot is crucial for its subsequent identification.

A comprehensive 2-DE map was annotated with the spot ID of proteins listed in Table I (Figure 2). Correlating protein expression profile with cell physiology requires an estimation of the relative abundance of each protein. Thus, to provide a base for further studies on MCF7, the relative abundance of spots determined on the map is reported in Table I and illustrated in Figure 3. The values were <0.1% for most of the spots detected on MCF7 gel (1039/1317).

Protein isoforms. The 368 spots identified corresponded to 249 distinct proteins. Of these, 56 proteins (176 spots) were present in at least 2 isoforms (Table III). These data are informative, but probably incomplete because not all the isoforms were excised and identified for each protein resolved in the gel. Generally, the isoforms affected the *p*I value more than the molecular weight, suggesting the presence of modifications affecting the charge of the polypeptides. The presence of numerous isoforms for the same protein highlights the complexity of the MCF7 proteome. For instance, the protein P05787 (Keratin, type II cytoskeletal 8) was found in at least in 21 spots. The molecular weights of these spots ranged from 42 to 57 kDa, with most spots being found as trains at 51 kDa. Some of these spots were distributed along a large streak and may be a result of a problem of focalization of this abundant protein. But it is tempting to explain many isoforms by PTMs. Using the information available in the Swiss-Prot database in conjunction with careful examination of the experimental spectra, we were able to analyze PTMs. For instance, HNRH1_HUMAN (P31943) was identified in seven spots with *p*I ranging from 6.36 to 7.04 and M*r* from 47,600 to 52,600. Annotations for PTMs on this protein in Swiss-Prot indicate the possibility of an acetylation of MET1, as well as phosphorylation on SER22, THR99, SER103, TYR245 and 305. The mass spectra obtained for the seven digested spots were comparable. However, phosphorylation of THR 99 and SER 103 were deduced from the observation of the pertinent ions in the spectrum (Figure 4). Significant information on PTMs can be directly deduced from MALDI-TOF spectra, as exemplified in the protein CSDE1_HUMAN (O75534). The interpretation of the spectra of the spots ID 607 (*p*I 6.54/M*r* 88438) and ID 608 (*p*I 6.61/M*r* 88438) indicated that the MET signal was cleaved. The sequence begins, therefore, with a serine. From the presence of the ion at m/z 2608.22, it can be deduced that for spot ID 608, SER1 is acetylated, while this acetylation was not observed for spot ID 607.

In addition to PTMs, isoforms are produced by the replacement of an amino acid. This occurred in the case of the BLVRB_HUMAN protein (P30043), identified in two spots ID 1729 (*p*I 7.45/Mr25106) and 1766 (*p*I 7.68/M*r* 23346), respectively. A variant amino acid at position 45 is annotated in Swiss-Prot. ARG can be replaced by GLN. Experimentally, spot 1766 contains an ARG, whereas spot 1729 contains GLN residue in position 45. These two amino acids differ in the charge of their side chains, which may explain the different *p*I observed on 2-DE gels for this protein.

Table I. *Features of the proteins identified in 2-DE patterns of MCF7 cells.*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
298	5.67	137829	0.057	Q9Y4L1	OXRP_HUM.	87	16	
462	6.60	109215	0.031	P14735	IDE_HUM.	56	15	
484	6.38	107306	0.042	P36776	LONM_HUM.	174	27	
499	6.26	102138	0.108	Q14697	GANAB_HUM.	223	37	
500	6.31	100708	0.234	Q14697	GANAB_HUM.	243	32	
507	6.37	101420	0.058	Q14697	GANAB_HUM.	199	29	
527	5.68	98808	0.045	P55072	TERA_HUM.	61	13	
607	6.54	88438	0.026	O75534	CSDE1_HUM.	66	14	
608	6.61	88438	0.030	O75534	CSDE1_HUM.	160	25	
625	6.42	87647	0.047	Q16891	IMMT_HUM.	83	18	
669	6.75	82548	0.021	Q96RP9	EFG1_HUM.	68	17	
674	6.44	82301	0.012	P26368	U2AF2_HUM.	63	19	
675	5.93	81809	0.029	Q96ED9	HOOK2_HUM.	186	33	
690	6.44	79871	0.011	P52888	MEPD_HUM.	146	29	
703	6.68	78448	0.017	P15311	EZRI_HUM.	75	14	1
704	6.02	78213	0.044	P49959	MRE11_HUM.	56	15	
713	5.84	77050	0.038	P13798	ACPH_HUM.	65	17	
728	5.35	74805	0.352	P11021	GR78_HUM.	123	33	1, 2
730	5.44	74415	0.703	P11021	GR78_HUM.	220	37	1, 2
745	7.68	75000	0.015	O60506	HNRPQ_HUM.	66	17	
766	6.00	71560	0.558	P38646	GR75_HUM.	228	26	1, 2
767	6.09	71374	0.774	P38646	GR75_HUM.	114	25	1, 2
799	5.74	69718	0.181	P11142	HSP7C_HUM.	155	32	1
800	5.82	69536	0.172	P11142	HS7C_HUM.	106	29	1
801	5.90	69900	0.215	P11142	HS7C_HUM.	151	41	1
822	6.05	68101	0.533	P08107	HS71_HUM.	365*	14	1
828	6.88	68457	0.041	P31040	DHSA_HUM.	165	24	
830	5.90	67746	0.287	P08107	HS71_HUM.	130	10	1
841	6.48	67570	0.063	P09960	LKHA4_HUM.	170	28	
846	6.39	67218	0.045	P09960	LKHA4_HUM.	84	33	
850	6.72	67218	0.020	Q8TAT6	NPL4_HUM.	76	12	
869	5.66	66347	0.055	P20700	LAM1_HUM.	120	31	1
870	5.69	66347	0.177	P20700	LAM1_HUM.	241	49	1
873	7.06	65830	0.088	O75083	WDR1_HUM.	212	46	
877	5.48	65317	0.339	Q9BWX0	Q9BWX0_HUM.	148	23	
882	6.09	64639	0.038	Q9BR76	COR1B_HUM.	73	20	
882	6.09	64639	0.038	Q13283	G3BP_HUM.	66	22	
883	6.15	64639	0.026	Q9BR76	COR1B_HUM.	152	23	
884	6.19	64639	0.055	P13797	PLST_HUM.	94	20	
886	7.57	63802	0.216	P14866	ROL_HUM.	107	16	1
891	7.46	63968	0.126	P14866	HNRPL_HUM.	98	20	1
893	6.39	64303	0.006	P48163	MAOX_HUM.	144	23	
895	6.94	63636	0.077	P02545	LAMA_HUM.	136	36	
897	7.20	63802	0.052	P02545	LAMA_HUM.	102	23	1
899	7.31	63802	0.100	P02545	LAMA_HUM.	84	10	1
901	6.68	63636	0.041	P49368	TCPG_HUM.	118	26	1
902	6.89	63636	0.061	P02545	LAMA_HUM.	73	21	1
903	7.35	63802	0.086	P14866	HNRPL_HUM.	77	26	1
905	6.84	63470	0.083	P02545	LAMA_HUM.	91	19	1
910	6.79	62484	0.148	P49368	TCPG_HUM.	173	43	1
916	6.11	62159	0.039	Q99829	CPNE1_HUM.	91	19	
920	6.83	62322	0.054	P48444	COPD_HUM.	106	27	
924	7.07	61997	0.066	P35520	CBS_HUM.	79	22	
927	5.65	61194	0.080	P61978	HNRPK_HUM.	74	18	1
929	7.35	61514	0.013	P54577	SYYC_HUM.	94	25	
933	7.22	61194	0.064	P08243	ASNS_HUM.	115	22	
936	6.51	60876	0.123	P17987	TCPA_HUM.	217	44	

Table I. *continued*

Table I. *continued*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
937	7.13	61035	0.032	P11413	G6PD_HUM.	125	32	1
940	5.70	60559	0.219	P61978	HNRPK_HUM.	118	38	1
941	7.61	61035	0.028	O60701	UGDH_HUM.	155	27	
944	6.04	60396	0.169	P61978	HNRPK_HUM.	129	37	1
947	5.67	60396	0.115	P61978	HNRPK_HUM.	93	28	1
950	6.29	60559	0.023	P13674	P4HA1_HUM.	66	17	
951	6.39	60559	0.057	P17987	TCPA_HUM.	118	22	
955	7.40	60233	0.048	P30038	AL4A1_HUM.	133	27	1
957	5.64	60233	0.045	P61978	HNRPK_HUM.	130	29	1
959	7.57	59426	0.036	P04040	CATA_HUM.	148	33	1
965	7.46	59586	0.020	P04040	CATA_HUM.	126	24	1
967	5.69	58157	0.260	P10809	CH60_HUM.	213	50	1, 2
970	6.05	58947	0.163	P48643	TCPE_HUM.	176	50	1
972	6.14	59106	0.048	O15460	P4HA2_HUM.	98	26	
977	6.47	58788	0.053	Q13177	PAK2_HUM.	78	17	
980	6.38	58629	0.025	P13674	P4HA1_HUM.	206	39	
982	6.88	58471	0.035	P49591	SYS_HUM.	222	46	
983	7.28	58314	0.094	Q9HCC0	MCCC2_HUM.	223	39	
986	7.22	58314	0.026	Q9HCC0	MCCC2_HUM.	102	22	
987	6.25	57844	0.056	P05787	K2C8_HUM.	188	42	1, 2
988	6.77	58157	0.017	P49591	SYS_HUM.	64	14	
990	7.14	58157	0.023	Q9HCC0	MCCC2_HUM.	167	28	
1003	6.04	57377	0.100	P50990	TCPQ_HUM.	224	39	
1007	5.81	58947	0.701	P10809	CH60_HUM.	135	41	1, 2
1008	6.73	57377	0.025	Q16881	TRXR1_HUM.	125	27	
1009	6.81	57223	0.027	Q16881	TRXR1_HUM.	114	24	
1011	6.67	57068	0.013	P40123	CAP2_HUM.	57	18	
1017	6.55	56761	0.020	P62191	PRS4_HUM.	188	37	
1018	7.38	56001	0.275	P31948	STIP1_HUM.	82	31	
1020	5.49	56608	0.029	Q02818	NUCB1_HUM.	126	33	
1022	7.24	56608	0.105	P09622	DLDH_HUM.	79	18	
1029	7.10	55250	0.245	O43175	SERA_HUM.	157	39	1
1031	7.28	55699	0.095	P11413	G6PD_HUM.	198	46	1
1034	6.91	55549	0.099	Q13217	DNJC3_HUM.	155	39	
1036	6.82	55250	0.203	Q9UMS4	PRP19_HUM.	106	27	1
1037	7.15	55101	0.177	P06733	ENOA_HUM.	962*	45	1
1038	6.04	55549	0.023	P05787	K2C8_HUM.	119	35	1, 2
1039	7.19	54953	0.181	P11413	G6PD_HUM.	245	55	1
1041	6.72	55250	0.055	O43175	SERA_HUM.	133	34	
1042	7.05	54953	0.107	P11413	G6PD_HUM.	116	35	1
1042	7.05	54953	0.107	P12268	IMDH2_HUM.	94	21	
1049	5.93	54510	0.122	Q02790	FKBP4_HUM.	157	37	
1050	6.01	54363	0.112	Q02790	FKBP4_HUM.	166	39	
1056	6.76	53634	0.266	P78371	TCPB_HUM.	224	45	1
1059	5.55	53345	0.533	P68363	TBAK_HUM.	230	48	
1062	6.34	53490	0.423	P30101	PDIA3_HUM.	222	51	1, 2
1064	6.66	53634	0.141	P78371	TCPB_HUM.	220	42	1
1065	7.29	53924	0.120	P11413	G6PD_HUM.	174	40	1
1068	7.78	54070	0.071	Q01518	CAP1_HUM.	75	20	
1072	6.54	53634	0.169	P23381	SYW_HUM.	159	34	
1073	6.24	51646	1.496	P05787	K2C8_HUM.	138	28	1, 2
1076	6.42	53779	0.125	P30101	PDIA3_HUM.	230	53	1, 2
1077	6.99	53779	0.085	P28838	AMPL_HUM.	148	39	
1078	7.14	53490	0.184	P28838	AMPL_HUM.	133	38	
1079	7.32	53779	0.034	P49419	AL7A1_HUM.	101	23	
1082	7.05	53779	0.043	P31943	HNRH1_HUM.	123	40	
1087	7.49	53202	0.053	P43490	NAMPT_HUM.	134	24	

Table I. *continued*

Table I. *continued*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
1088	6.58	52773	0.133	Q13228	SBP1_HUM.	142	41	
1089	6.40	52915	0.023	P61978	HNRPK_HUM.	72	18	1
1094	6.98	52347	0.124	Q9Y265	RUVB1_HUM.	228	47	1
1097	6.13	51646	0.475	P05787	K2C8_HUM.	254	48	1, 2
1098	5.28	51229	0.950	P05787	K2C8_HUM.	138	26	1, 2
1100	6.84	52206	0.150	Q9Y265	RUVB1_HUM.	227	50	1
1104	7.43	51925	0.083	P11172	PYR5_HUM.	115	28	
1106	7.62	52347	0.028	P43490	NAMPT_HUM.	121	31	
1112	7.25	52065	0.028	Q9NVA2	SEP11_HUM.	125	30	
1119	6.35	51646	0.190	P05787	K2C8_HUM.	233	44	1, 2
1122	7.36	51368	0.059	Q8TDM6	DLG5_HUM.	68	5	
1123	7.51	51368	0.132	P00367	DHE3_HUM.	200	31	1, 2
1125	6.65	51506	0.051	Q9UJV9	DDX41_HUM.	108	22	
1125	6.65	51506	0.051	P31943	HNRH1_HUM.	100	36	
1128	7.18	51368	0.018	Q9NVA2	SEP11_HUM.	125	30	
1134	6.45	50679	0.111	P31943	HNRH1_HUM.	110	42	
1134	6.45	50679	0.111	P05787	K2C8_HUM.	57	17	1, 2
1135	6.62	50816	0.065	P36957	ODO2_HUM.	81	20	
1138	6.36	50543	0.090	P31943	HNRH1_HUM.	112	40	
1144	6.54	49542	0.047	Q07960	RHG01_HUM.	128	35	
1145	6.97	49656	0.016	P11182	ODB2_HUM.	88	20	
1147	5.45	48752	0.683	P06576	ATPB_HUM.	158	36	1
1151	7.43	49202	0.498	P06733	ENOA_HUM.	171	47	1
1152	6.18	49429	0.081	Q9Y230	RUVB2_HUM.	256	64	
1157	7.20	48752	0.300	P06733	ENOA_HUM.	72	20	1
1160	5.53	48417	0.424	P06576	ATPB_HUM.	201	46	1
1162	5.88	48864	0.041	P49005	DPOD2_HUM.	57	20	
1168	6.99	48528	0.076	Q05048	CSTF1_HUM.	149	46	
1172	6.82	48528	0.038	P50395	GDIB_HUM.	181	54	
1173	5.63	48305	0.202	Q15084	PDIA6_HUM.	117	34	
1176	5.93	48195	0.084	Q9BYG5	PAR6B_HUM.	65	25	
1176	5.93	48195	0.084	P05787	K2C8_HUM.	62	18	1, 2
1178	5.84	47863	0.160	P05787	K2C8_HUM.	264	56	1, 2
1179	6.24	48417	0.038	P61158	ARP3_HUM.	184	62	
1179	6.24	48417	0.038	P05787	K2C8_HUM.	109	43	1, 2
1181	6.77	48084	0.054	P05455	LA_HUM.	63	28	
1191	5.77	47425	0.097	Q8NBS9	TXND5_HUM.	100	31	
1193	6.57	47644	0.047	O75439	MPPB_HUM.	237	47	
1199	5.69	47208	0.068	Q8NBS9	TXND5_HUM.	88	32	
1200	6.54	47534	0.020	P31943	HNRH1_HUM.	150	38	
1202	6.24	47316	0.068	Q9UJZ1	STML2_HUM.	103	31	
1208	6.06	46883	0.171	P31930	UCR1_HUM.	164	8	1, 2
1209	6.13	47316	0.012	O14745	NHERF_HUM.	63	34	1
1213	5.46	47099	0.086	P80303	NUCB2_HUM.	67	32	
1221	5.86	46241	0.102	P05787	K2C8_HUM.	109	58	1, 2
1223	6.88	46561	0.132	Q9UQ80	P2G4_HUM.	138	1	1
1235	6.94	46241	0.149	P26641	EF1G_HUM.	56	20	
1251	5.90	45503	0.046	O96019	ACL6A_HUM.	62	19	
1258	5.83	45191	0.118	P52597	HNRPF_HUM.	133	44	
1259	7.05	45399	0.014	P55263	ADK_HUM.	90	17	
1264	6.89	45295	0.011	P31153	METK_HUM.	119	28	
1269	6.78	44777	0.054	P20073	ANXA7_HUM.	176	37	
1276	7.35	44367	0.044	P11310	ACADM_HUM.	104	30	
1278	6.62	43860	0.334	P49585	PCY1A_HUM.	105	36	
1280	6.85	44164	0.063	P36507	MP2K2_HUM.	58	19	
1281	6.97	44164	0.030	P61163	ACTZ_HUM.	149	48	
1282	7.41	44164	0.096	P08559	ODPA_HUM.	65	8	

Table I. *continued*

Table I. *continued*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
1284	5.69	43860	0.251	P05787	K2C8_HUM.	115	35	1, 2
1284	5.69	43860	0.251	P05783	K1CR_HUM.	89	27	1, 2
1285	5.98	43259	1.244	P60709	ACTB_HUM.	63	19	1
1286	5.37	43860	0.184	P05783	K1CR_HUM.	235	45	1, 2
1292	6.06	43559	0.284	P63621	ACTG_HUM.	113	37	
1294	6.38	43961	0.079	Q07889	SOS1_HUM.	58	12	
1296	7.28	43359	0.185	O75874	IDHC_HUM.	156	31	
1297	7.48	43559	0.070	P53582	AMPM1_HUM.	109	37	
1300	7.04	43459	0.076	P55263	ADK_HUM.	143	32	
1301	7.35	43559	0.031	O75874	IDHC_HUM.	140	28	
1306	6.67	43160	0.093	P23526	SAHH_HUM.	158	31	
1308	7.10	43259	0.057	O75874	IDHC_HUM.	98	29	
1309	6.86	43259	0.017	P55263	ADK_HUM.	65	16	
1311	5.57	42667	0.452	P05787	K2C8_HUM.	131	45	1, 2
1316	6.43	42667	0.061	Q03154	ACY1_HUM.	193	53	
1317	7.67	42471	0.208	P62333	PRS10_HUM.	179	48	
1322	5.72	42962	0.759	P05783	K1CR_HUM.	281	69	1, 2
1324	7.38	42569	0.021	P09972	ALDOC_HUM.	114	39	
1327	6.33	42471	0.029	P60709	ACTB_HUM.	63	19	1
1330	7.15	42083	0.017	P16930	FAAA_HUM.	136	42	
1331	6.72	41986	0.013	O75821	IF34_HUM.	67	21	
1332	6.82	41986	0.058	Q15019	SEPT2_HUM.	126	45	
1333	5.45	40938	1.171	P08727	K1CS_HUM.	267	70	1
1338	6.22	41697	0.040	P49903	SPS1_HUM.	69	29	
1339	7.60	41411	0.126	O96008	TOM40_HUM.	108	40	
1340	6.99	41316	0.132	Q15366	PCBP2_HUM.	80	33	
1341	6.13	41506	0.021	Q9UNM6	PSD13_HUM.	192	43	
1343	6.63	41506	0.018	P30740	ILEU_HUM.	127	31	
1344	5.33	40938	0.495	P08727	K1C19_HUM.	325	73	1
1346	6.20	41411	0.036	P60709	ACTB_HUM.	63	19	1
1347	5.24	40938	0.459	P08727	K1C19_HUM.	218	49	1
1348	6.00	41221	0.091	O95433	AHSA1_HUM.	95	33	
1349	6.08	41316	0.028	P60709	ACTB_HUM.	99	34	1
1350	6.15	41221	0.044	Q9BT78	CSN4_HUM.	161	52	
1352	5.87	41126	0.076	Q96I99	SUCB2_HUM.	153	40	
1356	7.23	40938	0.101	P28482	MK01_HUM.	83	28	
1358	7.32	40938	0.108	Q9BXW7	CECR5_HUM.	143	40	
1360	7.46	40938	0.083	P17174	AATC_HUM.	51	6	
1361	7.27	40750	0.091	O43684	BUB3_HUM.	117	37	1
1364	7.37	40470	0.262	Q15365	PCBP1_HUM.	122	56	
1365	5.67	40470	0.146	Q9NYL9	TMOD3_HUM.	108	29	
1372	6.97	40377	0.044	Q15365	PCBP1_HUM.	122	56	
1377	7.63	40008	0.052	P35908	K22E_HUM.	54	11	
1379	6.68	39916	0.090	P51570	GALK1_HUM.	132	41	
1381	5.74	40100	0.013	Q9UBE0	ULE1A_HUM.	138	42	
1386	6.05	39916	0.036	Q6FI81	CPIN1_HUM.	64	26	
1395	5.81	39641	0.059	Q04323	U33K_HUM.	71	36	
1398	7.12	39641	0.085	Q92890	UFD1_HUM.	111	31	
1401	5.67	39460	0.025	P61962	WDR68_HUM.	103	36	
1406	6.18	39369	0.048	O95861	BPNT1_HUM.	115	38	1
1407	6.40	39188	0.052	P06132	DCUP_HUM.	105	33	
1408	6.60	39369	0.034	Q9UBS4	DNJBB_HUM.	55	27	
1413	6.89	39188	0.013	P82650	RT22_HUM.	154	40	
1419	7.14	38919	0.041	P14550	AK1A1_HUM.	142	43	1
1427	7.64	38563	0.066	O95299	NUDM_HUM.	99	25	
1429	6.82	38386	0.231	P53004	BIEA_HUM.	63	39	
1430	7.20	38474	0.135	P09467	F16P1_HUM.	212	47	

Table I. *continued*

Table I. *continued*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
1433	6.35	38386	0.066	P53365	ARFP2_HUM.	93	35	
1445	7.36	37947	0.041	P51665	PSD7_HUM.	108	36	
1452	6.20	37773	0.030	O15266	SHOX_HUM.	55	20	
1454	6.36	37514	0.085	P50213	IDH3A_HUM.	87	29	
1457	6.60	37514	0.024	P62136	PP1A_HUM.	79	24	
1462	6.02	37085	0.060	Q13347	IF32_HUM.	53	19	
1469	7.41	36651	0.025	Q9Y2S7	PDIP2_HUM.	110	35	
1474	6.59	36651	0.047	P62136	PP1A_HUM.	233	40	
1484	7.22	36049	0.088	O00151	PDLI1_HUM.	126	47	
1485	6.49	36134	0.045	P50226	ST1A2_HUM.	56	20	
1486	7.09	36049	0.123	P31942	HNRH3_HUM.	104	39	
1491	6.52	36049	0.053	P62140	PP1B_HUM.	212	46	
1492	6.80	35878	0.056	P36873	PP1G_HUM.	194	41	
1494	7.28	36049	0.071	Q9NYK5	RM39_HUM.	116	39	
1498	6.87	35878	0.069	O00170	AIP_HUM.	73	26	
1504	6.25	35288	0.197	P05388	RLA0_HUM.	122	43	
1515	7.76	34873	0.154	P52895	AK1C2_HUM.	143	51	
1531	6.00	34138	0.061	P52907	CAZA1_HUM.	113	47	1
1555	6.00	32713	0.055	P62879	GBB2_HUM.	81	29	
1563	6.73	32405	0.048	Q9H2U2	IPYR2_HUM.	132	51	
1576	7.38	31646	0.051	P45880	VDAC2_HUM.	98	29	1
1584	6.99	31422	0.007	Q00169	PIPNA_HUM.	83	36	
1594	7.26	30687	0.153	P10768	ESTD_HUM.	160	53	
1599	7.34	30614	0.021	P46926	GNPI_HUM.	128	61	
1601	6.37	30470	0.040	P09525	ANXA4_HUM.	85	37	1
1607	7.22	30326	0.019	P11802	CDK4_HUM.	131	48	
1610	7.01	30254	0.121	Q13011	ECH1_HUM.	81	28	
1611	7.05	30111	0.028	Q13011	ECH1_HUM.	80	28	
1620	6.34	29476	0.100	P61289	PSME3_HUM.	71	42	
1622	6.96	29476	0.099	P25786	PSA1_HUM.	88	40	
1629	5.66	29060	0.176	O00299	CLIC1_HUM.	87	34	1
1637	5.96	28632	0.105	P46777	RL5_HUM.	59	21	
1643	6.93	28390	0.023	P48556	PSD8_HUM.	59	22	
1644	5.83	28091	0.281	P07339	CATD_HUM.	117	37	2
1646	6.74	28210	0.084	P07951	TPM2_HUM.	70	28	
1649	6.36	28032	0.195	Q06323	PSME1_HUM.	142	50	1
1657	6.11	27502	0.361	P35232	PHB_HUM.	167	55	1
1669	6.71	27212	0.072	P42126	D3D2_HUM.	55	21	
1672	6.97	27154	0.038	P18669	PGAM1_HUM.	74	37	1
1686	6.66	26360	0.308	P30084	ECHM_HUM.	122	38	1, 2
1689	6.59	26138	0.909	P04792	HS27_HUM.	117	37	1
1691	7.40	26138	0.460	P60174	TPIS_HUM.	57	5	1
1694	7.00	26249	0.070	P60900	PSA6_HUM.	92	40	1
1697	5.80	26193	0.016	P35232	PHB_HUM.	122	52	1
1698	6.94	26138	0.272	O75838	KIP2_HUM.	58	34	
1700	6.29	25808	0.563	P04792	HSPB1_HUM.	74	30	1
1706	6.22	25808	0.193	Q13162	PRDX4_HUM.	55	25	
1707	6.83	25808	0.220	P60174	TPIS_HUM.	111	40	1
1709	5.24	25808	0.056	P29312	1433Z_HUM.	141	52	
1720	5.75	25374	0.108	P21266	GSTM3_HUM.	134	48	
1726	7.60	25106	0.294	P62826	RAN_HUM.	93	43	
1729	7.45	25106	0.074	P30043	BLVRB_HUM.	57	38	1
1734	5.77	24852	0.073	P05091	DHAM_HUM.	217	17	
1735	7.42	24558	0.174	P25787	PSA2_HUM.	70	38	
1738	6.57	24631	0.031	P36957	ODO2_HUM.	57	17	
1751	5.39	23626	0.362	Q16206	COVA1_HUM.	94	25	
1754	5.62	23767	0.023	P61086	UBC1_HUM.	75	34	

Table I. *continued*

Table I. *continued*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
1760	7.58	23346	0.044	P04179	SODM_HUM.	78	36	1, 2
1762	6.57	23277	0.274	P30048	PRDX3_HUM.	55	21	1, 2
1766	7.68	23346	0.018	P30043	BLVRB_HUM.	77	49	1
1772	7.36	22662	0.006	Q9Y3C8	UFC1_HUM.	70	32	
1786	6.12	21226	0.742	P32119	PRDX2_HUM.	168	65	1
1792	5.76	21100	0.092	Q96T22	CASP8_HUM.	104	25	
1806	7.11	19688	0.012	Q9UBQ0	VPS29_HUM.	67	32	
1810	5.80	19600	0.022	O75947	ATP5H_HUM.	69	50	
1831	6.42	17914	0.220	P15531	NDKA_HUM.	79	42	
1835	6.74	17754	0.016	P04792	HSPB1_HUM.	61	32	1
1837	6.91	17321	0.042	P23528	COF1_HUM.	82	50	1
1838	6.76	17321	0.034	P24666	PPAC_HUM.	88	47	
1849	7.29	16411	0.167	P62937	PPIA_HUM.	78	43	
1853	5.54	16228	0.321	P63241	IF5A_HUM.	78	31	
1855	7.68	16337	0.139	P30044	PRDX5_HUM.	145	56	1
1865	6.51	15584	0.147	P61088	UBE2N_HUM.	74	33	
1866	7.13	15584	0.060	Q96DB5	FA82B_HUM.	62	23	
1877	6.87	15068	0.084	P31151	S107_HUM.	72	36	
1883	6.16	14486	0.076	P29373	RABP2_HUM.	134	51	
1887	7.13	14112	0.229	P49773	HINT1_HUM.	58	40	
1889	7.29	13989	0.132	P10606	COX5B_HUM.	54	24	1
1905	7.62	12877	0.084	P04080	CYTB_HUM.	76	55	
1907	5.26	12655	0.220	Q96HG2	Q96HG2_HUM.	81	36	1
1917	6.78	12010	0.040	Q96FQ6	S10AG_HUM.	61	38	1
1923	6.57	11298	0.320	P31949	S10AB_HUM.	161	27	1
1924	7.48	11397	0.224	P05109	S108_HUM.	57	40	1
2133	5.83	107306	0.048	P49588	SYA_HUM.	84	15	
2161	6.36	35039	0.015	O00764	PDXK_HUM.	62	22	1
2163	6.06	51368	0.463	P05787	K2C8_HUM.	255	58	1, 2
2181	5.81	42864	0.458	P05783	K1CR_HUM.	243	73	1, 2
2182	5.89	43061	0.255	P05783	K1CR_HUM.	207	44	1, 2
2190	6.52	43459	0.136	Q9UL16	CCD19_HUM.	131	36	
2194	5.93	51785	0.183	P05787	K2C8_HUM.	168	36	1, 2
2196	5.69	51646	0.306	P05787	K2C8_HUM.	137	25	1, 2
2202	7.00	44572	0.058	P53367	ARFP1_HUM.	172	55	
2207	5.69	42864	0.022	P60709	ACTB_HUM.	137	45	1
2216	5.97	67746	0.353	P08107	HSP71_HUM.	161	28	1
2219	5.89	60396	0.266	P61978	HNRPK_HUM.	96	28	1
2220	5.92	57532	0.175	P54578	UBP14_HUM.	118	25	
2221	5.93	59909	0.035	P10809	CH60_HUM.	109	33	1, 2
2222	5.89	51785	0.086	P05787	K2C8_HUM.	196	45	1, 2
2223	5.83	51925	0.072	P05787	K2C8_HUM.	154	41	1, 2
2226	5.77	51785	0.140	P05787	K2C8_HUM.	189	45	1, 2
2236	6.47	26083	0.244	P04792	HSPB1_HUM.	142	37	1
2240	6.31	79871	0.002	O60568	PLOD3_HUM.	193	36	
2241	6.27	58157	0.016	O75131	CPNE3_HUM.	116	19	
2242	6.94	25644	0.072	P00492	HPRT_HUM.	94	38	
2242	6.94	25644	0.072	P04792	HSPB1_HUM.	64	39	1
2243	6.97	63470	0.069	P31939	PUR9_HUM.	118	28	
2243	6.97	63470	0.069	P02545	LAMA_HUM.	88	28	1
2244	6.99	60233	0.004	P40227	TCPZ_HUM.	69	16	1
2245	6.99	61675	0.159	P40227	TCPZ_HUM.	128	36	1
2246	5.76	21932	0.024	P02794	FRIH_HUM.	62	28	
2247	6.24	53924	0.175	P05787	K2C8_HUM.	130	35	1, 2
2248	5.85	46991	0.044	P05787	K2C8_HUM.	228	54	1, 2
2249	7.03	56152	0.137	P11413	G6PD_HUM.	188	43	1
2252	7.24	55250	0.058	P12268	IMDH2_HUM.	219	47	

Table I. *continued*

Table I. *continued*

ID on map	*p*I	M*r* gel (Da)	Abundance (%Vol)	UniProtKB/Swiss Prot accession number	Abbr. name	Mascot score	Coverage (%)	Entry in 2D databases
2253	7.57	60396	0.016	P30038	AL4A1_HUM.	120	31	1
2255	7.72	16301	0.070	P30044	PRDX5_HUM.	108	27	1
2256	7.07	63802	0.169	P31939	PUR9_HUM.	198	40	
2257	6.91	35793	0.017	P31942	HNRH3_HUM.	71	25	
2258	6.51	49885	0.061	P31943	HNRH1_HUM.	54	20	
2259	6.43	51925	0.073	P31943	HNRH1_HUM.	57	19	
2260	6.54	35878	0.011	P37837	TALDO_HUM.	87	27	
2262	6.67	58629	0.012	P48382	RFX5_HUM.	56	14	
2263	6.73	63305	0.069	P49368	TCPG_HUM.	63	13	1
2264	6.16	32251	0.018	P50224	ST1A3_HUM.	54	20	
2265	6.71	40192	0.043	Q9UBS4	DNJBB_HUM.	92	33	
2266	5.75	41602	0.192	P60709	ACTB_HUM.	69	28	1
2267	5.67	42179	0.157	P60709	ACTB_HUM.	64	28	1
2268	5.94	60876	0.020	P61978	HNRPK_HUM.	90	25	1
2269	6.94	54805	0.186	P98170	BIRC4_HUM.	77	30	
2271	5.81	56456	0.099	Q12874	SF3A3_HUM.	129	30	
2272	6.38	43259	0.036	Q13148	TADBP_HUM.	54	14	
2275	6.26	99701	0.041	Q14697	GANAB_HUM.	99	17	
2278	7.82	24852	0.008	Q99714	HCD2_HUM.	78	38	
2279	6.82	39733	0.016	Q9BQ04	RBM30_HUM.	59	22	1
2280	5.17	43860	0.050	Q9BS26	TXND4_HUM.	117	36	
2281	6.38	42667	0.009	Q9P2R7	SUCB1_HUM.	74	21	
2283	6.73	38298	0.017	Q9UBS4	DNJBB_HUM.	67	24	
2284	6.47	50000	0.014	Q9UKS6	PACN3_HUM.	95	35	
2285	6.76	55699	0.002	Q9UMS4	PRP19_HUM.	77	17	1
2286	6.52	51646	0.027	P36957	ODO2_HUM.	57	17	

Protein identification using MALDI-TOF MS (except proteins analyzed with LC-MS/MS, *in Mascot score column). Protein scores >54 are significant at $p<0.05$. %Vol of a spot = ((O. D.*Area)/Sum of the volume of all spots detected on the gel) *100. Entry in databases: (1)Swiss 2D-PAGE and (2) Siena 2D-PAGE.

Table II. *Distribution of spots according to protein abundance.*

Class[1]	% Vol[2]	Number of spots picked	Number of identifications
1	>1	13	12
2	0.999-0.500	16	12
3	0.499-0.200	54	35
4	0.199-0.150	51	33
5	0.149-0.100	79	43
6	0.099-0.050	195	94
7	0.049-0.019	259	96
8	<0.0189	60	19

[1]Classes were used to draw the diagrams shown in Figure 1 and 3.
[2]The %Vol represents a normalized value related to the protein abundance in a given spot and was calculated on the 2-DE gel used to build the MCF7 map.

Table III. *List of proteins identified as multiple isoforms.*

Number of isoforms	Number of affected proteins	AC N° in UniProtKB/Swiss Prot
2	34	O43175, O75534, P04040, P06576, P09960, P11021, P12268, P13674, P17987, P20700, P28838, P30038, P30043, P30044, P30101, P31939, P31942, P35232, P38646, P40227, P43490, P49591, P60174, P62136, P78371, Q02790, Q13011, Q15365, Q16881, Q8NBS9, Q9BR76, Q9NVA2, Q9UMS4, Q9Y265
3	12	O75874, P06733, P08107, P08727, P10809, P11142, P14866, P36957, P49368, P55263, Q9HCC0, Q9UBS4
4	1	Q14697,
5	2	P04792, P05783
6	2	P02545, P11413
7	2	P31943, P60709
8	1	P61978
21	1	P05787

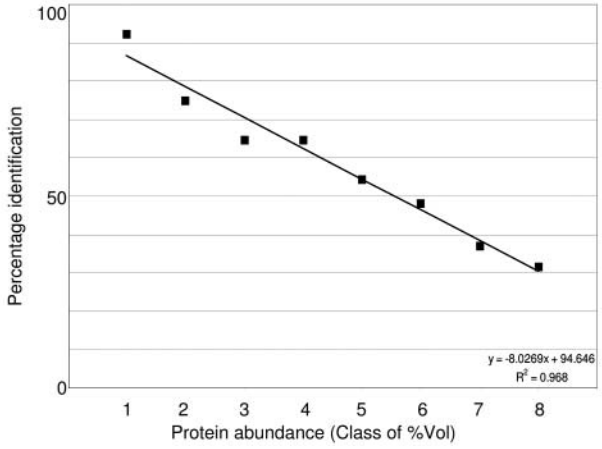


Figure 1. *Scatter plot of the relation between protein abundance and rate of successful identification. Protein abundance was calculated in %Vol. The values of %Vol for all the spots detected on the gel were distributed in 8 classes as reported in Table II.*

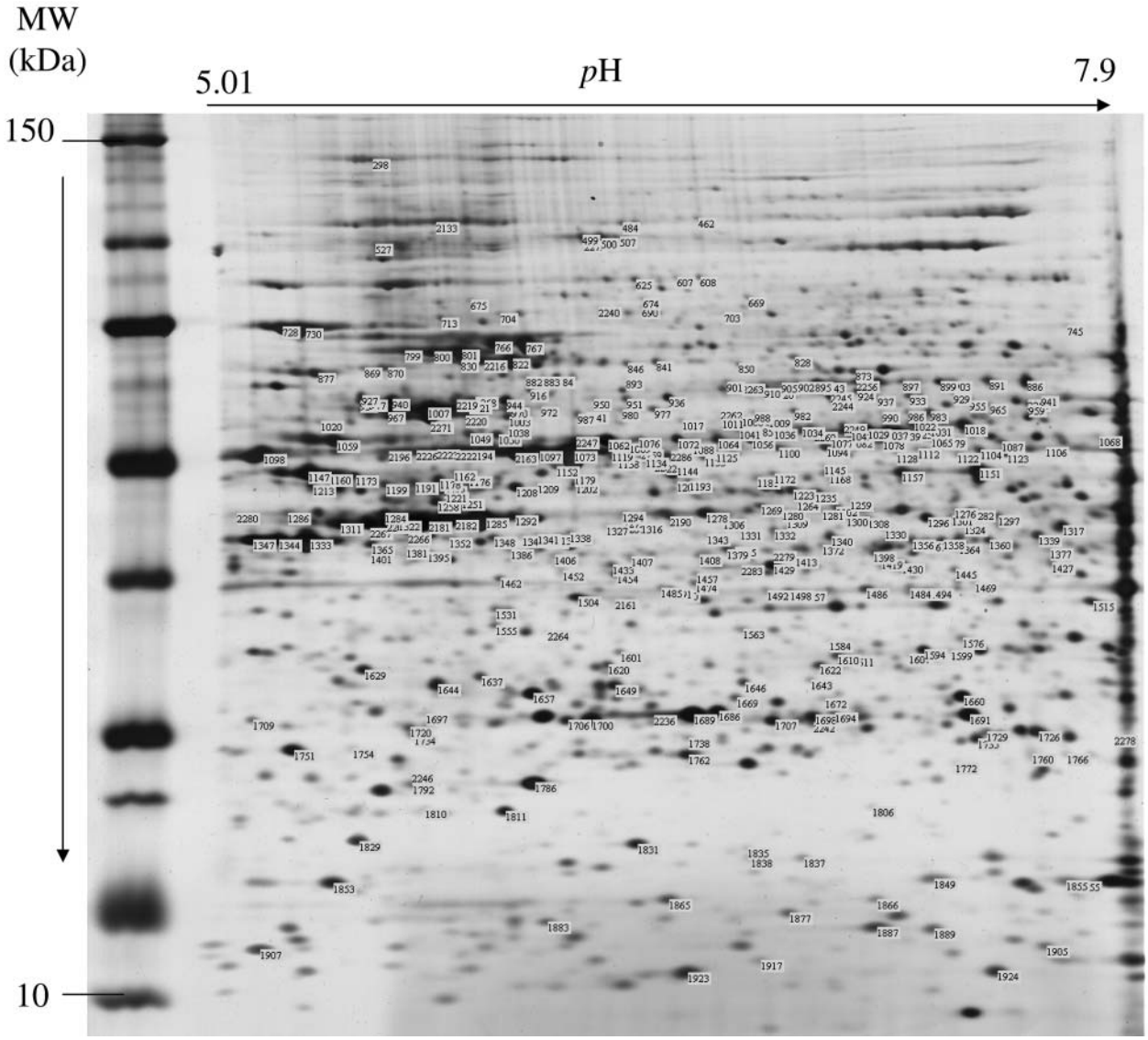


Figure 2. *Annotated 2-DE map of MCF7. Cells were collected during the exponential growth phase. Whole cells (~ 2x10^8) were lysed in 1 ml of extraction buffer. 150 μg of the cell extract were loaded on 5-8 IPG gradient strip (17 cm-length). The labels show the spot ID in relation to the identifications listed in Table I.*

Table IV. *Classification of identified proteins in biological and molecular processes according to the Gene Ontology (GO) terms.*

Biological processes			Molecular processes	
	Number of genes	% for each process	Number of genes	% for each process
Cellular physiological process	183	90.59	Protein binding	70	33.82
Metabolism	167	82.67	Hydrolase activity	48	23.19
Localization	43	21.29	Nucleotide binding	48	23.19
Regulation of physiological process	35	17.33	Ion binding	42	20.29
Regulation of cellular process	34	16.83	Oxidoreductase activity	36	17.39
Cell communication	24	11.88	Nucleic acid binding	30	14.49
Response to stress	23	11.39	Transferase activity	24	11.59
Response to abiotic stimulus	13	6.44	Electron transporter	17	8.21
Negative regulation of process	12	5.94	Ligase activity	13	6.28
Physiological process	11	5.45	Lipid binding	7	3.38
Morphogenesis	9	4.46	Isomerase activity	7	3.38
Response to biotic stimulus	9	4.46	Lyase activity	7	3.38
Positive regulation of process	6	2.97	Structural constituent of cytoskeleton	6	2.9
Death	6	2.97	Ion transporter activity	6	2.9
Response to external stimulus	5	2.48	Enzyme activator activity	6	2.9
Locomotion	5	2.48	Helicase activity	4	1.93
Regulation of enzyme activity	4	1.98	Carrier activity	4	1.93
System development	4	1.98	Cofactor binding	4	1.93
Organ development	4	1.98	Structural constituent of ribosome	3	1.45
Response to endogenous stimulus	3	1.49	Small protein activating enzyme activity	3	1.45
Regulation of growth	3	1.49	Receptor binding	3	1.45
Cell growth	3	1.49	GTPase regulator activity	3	1.45
Homeostasis	3	1.49	Channel or pore class transporter activity	3	1.45
Tissue development	2	0.99	Transcriptional activator activity	3	1.45
Behavior	2	0.99	Translation factor activity, nucleic acid binding	3	1.45
Cell differentiation	1	0.5	Enzyme inhibitor activity	2	0.97
Cell adhesion	1	0.5	Peroxidase activity	2	0.97
			Chromatin binding	2	0.97
			Transcription factor activity	2	0.97
			Receptor signaling protein	2	0.97
			Transcription cofactor activity	2	0.97
			ATPase activity, movement of substances	2	0.97
			Selenium binding	2	0.97
			Vitamin binding	1	0.48
			Isoprenoid binding	1	0.48
			Transcriptional repressor activity	1	0.48
			Small protein conjugating enzyme activity	1	0.48
			Caspase regulator activity	1	0.48
			Chaperone activator activity	1	0.48
			Lipid transporter activity	1	0.48
			Drug binding	1	0.48
			Protein transporter activity	1	0.48
			Organic acid transporter activity	1	0.48

Functional classification and localization. Mining a proteome currently needs a lot of time and effort in seeking all the available information. This is further hampered by the wide variation in the use of vocabulary that inhibits effective searching. The classification of identified proteins in biological processes and molecular functions based on the Gene Ontology (GO) terms facilitates annotation. According to GO, a biological process is a series of events accomplished by one or more ordered assemblies of molecular functions. As it is often difficult to distinguish between a biological process and a molecular function, the general rule is that a process must have more than one distinct step. FatiGo (20) was used for MCF7 protein classification. Among the 249 identified proteins, 6 genes

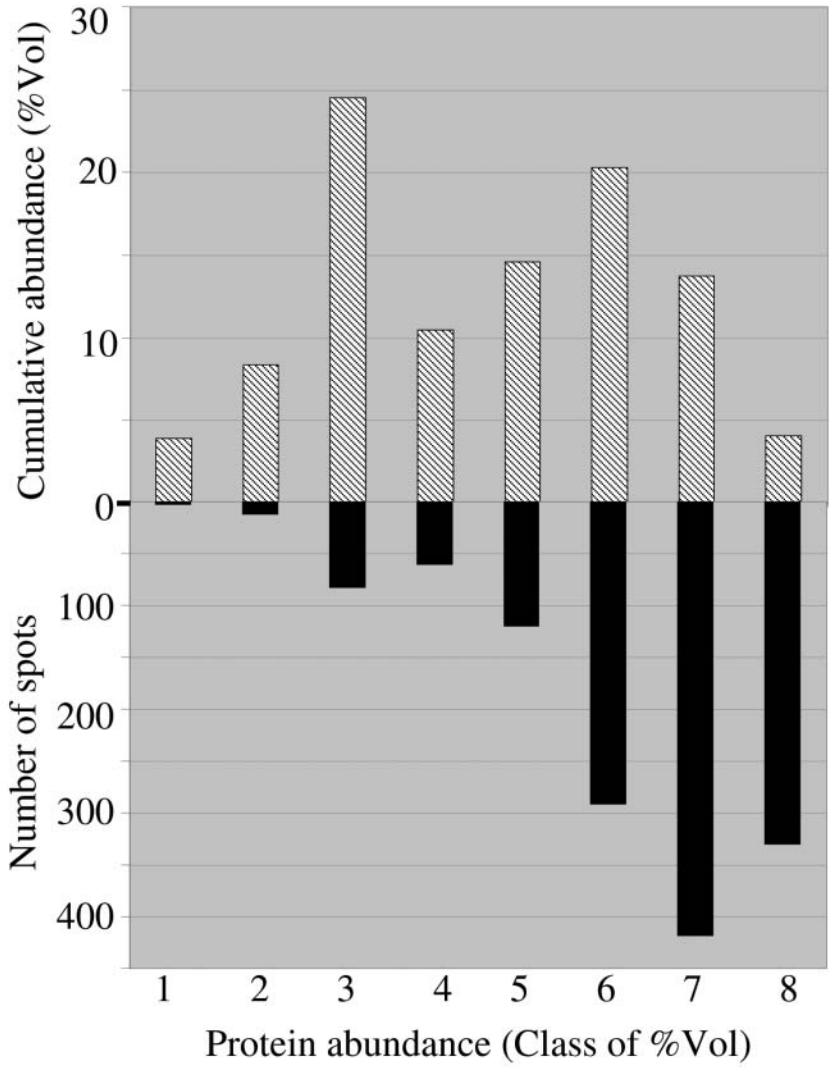


Figure 3. *Histograms illustrating the relative dynamics of the MCF7 protein abundance after 2-DE separation. Protein abundance (%Vol) was evaluated by the image analysis software. The signal intensity (O. D.) measured for each spot on the gel was below the saturation threshold. Classes of percentage of volume are shown in Table II. The volume values were <0.1 for most of the spots (1039/1317).*

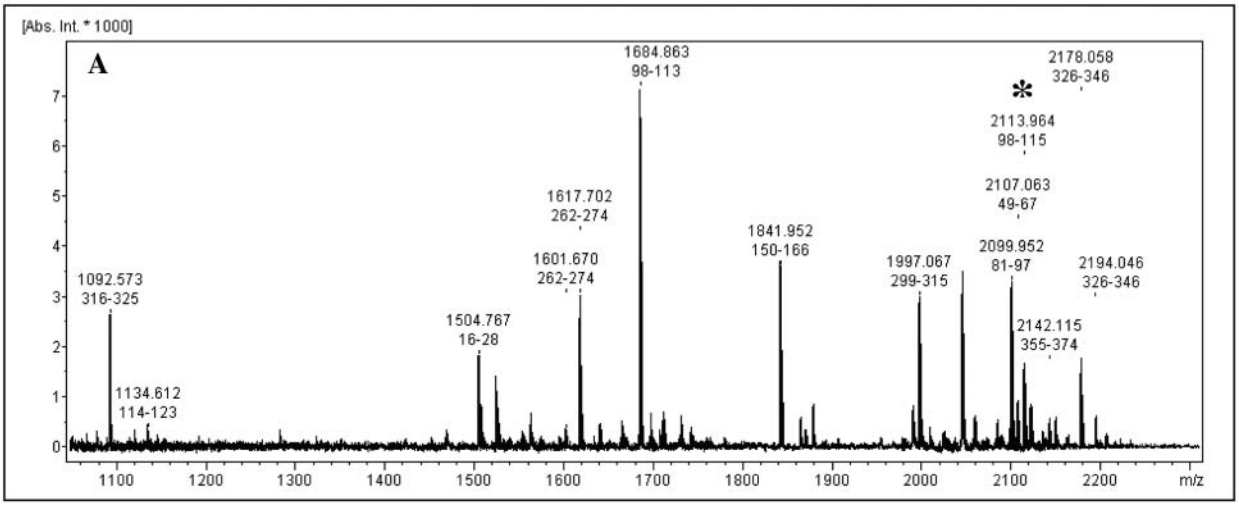
were not annotated in databases. The analysis was then carried out on 243 genes. The distribution of MCF7 proteins in various biological and molecular processes is reported in Table IV. As expected 183 proteins (91%) are involved in cellular physiology: 167 are directly implicated in metabolism, 43 participate in targeting (GO term: localization), and 24 are implicated in cell communication.

Conclusion

The generation of a comprehensive MCF7 2-DE reference map represents a base for studies of the MCF7 proteome. As illustrated in this paper, the use of bioinformatics tools allows a deeper analysis of the proteins identified in a proteome, not only for differential studies but also for cell profiling. Our results allowed the construction of the most extensive reference map of MCF7 published so far, with 189 novel proteins, which have never been listed on maps, and are now accessible on World 2D-PAGE. This 2-DE map will be available on our website at: http://www-smbh.univ-paris13.fr/lbtp/ Biochemistry/Biochimie/bque.htm.

Acknowledgements

This work was supported by the Ministère des Finances et de l'Industrie (NODDICCAP contract), by the European Union (European Regional Development Fund) and by the Association de la Recherche contre le Cancer.



B

m/z	Start-End	Sequence
1504.76	16-28	GLPWSCSADEVQR
2107.06	49-67	EGRPSGEAFVELESEDEVK
2099.95	81-97	YVEVFKSNNVEMDWVLK
1684.86	98-113	HTGPNSPDTANDGFVR
1134.61	114-123	LRGLPFGCSK
1841.95	150-166	STGEAFVQFASQEIAEK
1601.67	262-274	DLNYCFSGMSDHR
1617.70	262-274	DLNYCFSGmSDHR
2113.96*	275-293	H**pT**GPN**pS**PDTANDGFVRLR
1997.06	299-315	ATENDIYNFFSPLNPVR
2178.05	326-346	VTGEADVEFATHEDAVAAMSK
2194.04	326-346	VTGEADVEFATHEDAVAAmSK
1092.57	316-325	VHIEIGPDGR
2142.11	355-374	YVELFLNSTAGASGGAYEHR

Figure 4. *A: MALDI mass spectrum of HNRH1_HUMAN (P31943) identified from the spot ID 1138. B: Table summarizing the tryptic peptides identified on the mass spectrum. M: oxidized methionine, pS: phosphorylated serine, pT: phosphorylated threonine; *peptide contains phosphorylated SER 99 and THR 103.*

References

1 Hortobagyi GN, de la Garza Salazar J, Pritchard K, Amadori D, Haidinger R, Hudis CA, Khaled H, Liu MC, Martin M, Namer M, O'Shaughnessy JA, Shen ZZ and Albain KS: The global breast cancer burden: variations in epidemiology and survival. Clin Breast Cancer *6*: 391-401, 2005.

2 Jemal A, Tiwari RC, Murray T, Ghafoor A, Samuels A, Ward E, Feuer EJ and Thun MJ: Cancer statistics, 2004. CA Cancer J Clin *54*: 8-29, 2004.

3 Wang D, Jensen RH, Williams KE and Pallavicini MG: Differential protein expression in MCF7 breast cancer cells transfected with *ErbB2*, neomycin resistance and luciferase plus yellow fluorescent protein. Proteomics *4*: 2175-2183, 2004.

4 Chopin V, Slomianny C, Hondermarck H and Le Bourhis X: Synergistic induction of apoptosis in breast cancer cells by cotreatment with butyrate and TNF-alpha, TRAIL, or anti-Fas agonist antibody involves enhancement of death receptors' signaling and requires P21(waf1). Exp Cell Res *298*: 560-573, 2004.

5 Gehrmann ML, Fenselau C and Hathout Y: Highly altered protein expression profile in the adriamycin resistant MCF-7 cell line. J Proteome Res *3*: 403-409, 2004.

6 Gehrmann ML, Hathout Y and Fenselau C: Evaluation of metabolic labeling for comparative proteomics in breast cancer cells. J Proteome Res *3*: 1063-1068, 2004.

7 Hathout Y, Gehrmann ML, Chertov A and Fenselau C: Proteomic phenotyping: metastatic and invasive breast cancer. Cancer Lett *210*: 245-253, 2004.

8 Englbrecht CC and Facius A: Bioinformatics challenges in proteomics. Comb Chem High Throughput Screen *8*: 705-715, 2005.
9 Righetti PG, Castagna A, Antonucci F, Piubelli C, Cecconi D, Campostrini N, Zanusso G and Monaco S: The proteome: Anno Domini 2002. Clin Chem Lab Med *41*: 425-438, 2003.
10 Thiede B, Hohenwarter W, Krah A, Mattow J, Schmid M, Schmidt F and Jungblut PR: Peptide mass fingerprinting. Methods *35*: 237-247, 2005.
11 Gorg A, Weiss W and Dunn MJ: Current two-dimensional electrophoresis technology for proteomics. Proteomics *4*: 3665-3685, 2004.
12 Pietrogrande MC, Marchetti N, Dondi F and Righetti PG: Decoding 2D-PAGE complex maps: Relevance to proteomics. J Chromatogr B Analyt Technol Biomed Life Sci *833*: 51-62, 2006.
13 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.
14 Poirier F, Pontet M, Labas V, le Caer JP, Sghiouar-Imam N, Raphael M, Caron M and Joubert-Caron R: Two-dimensional database of a Burkitt lymphoma cell line (DG 75) proteins: protein pattern changes following treatment with 5'-azycytidine. Electrophoresis *22*: 1867-1877, 2001.
15 Joubert-Caron R and Caron M: Proteome analysis in the study of lymphoma cells. Mass Spectrom Rev *24*: 455-468, 2005.
16 Pionneau C, Canelle L, Bousquet J, Hardouin J, Bigeard J, Caron M and Joubert-Caron R: Proteomic analysis of membrane-associated proteins from the breast cancer cell line MCF7. Cancer Genom Proteom *2*: 199-208, 2005.
17 Canelle L, Bousquet J, Pionneau C, Deneux L, Imam-Sghiouar N, Caron M and Joubert-Caron R: An efficient proteomics-based approach for the screening of autoantibodies. J Immunol Methods *299*: 77-89, 2005.
18 Canelle L, Pionneau C, Marie A, Bousquet J, Bigeard J, Lutomski D, Kadri T, Caron M and Joubert-Caron R: Automating proteome analysis: improvements in throughput, quality and accuracy of protein identification by peptide mass fingerprinting. Rapid Commun Mass Spectrom *18*: 2785-2794, 2004.
19 Caron M, Imam-Sghiouar N, Poirier F, Le Caer JP, Labas V and Joubert-Caron R: Proteomic map and database of lymphoblastoid proteins. J Chromatogr B Analyt Technol Biomed Life Sci *771*: 197-209, 2002.
20 Al-Shahrour F, Minguez P, Vaquerizas JM, Conde L and Dopazo J: Babelomics: a suite of web-tools for functional annotation and analysis of group of genes in high-throughput experiments. Nucleic Acids Res (Web Server issue) *33*: 460-464, 2005.

Received November 14, 2006
Accepted November 26, 2006