Human Proteinpedia as a Resource for Clinical Proteomics*

Suresh Mathivanan‡§ and Akhilesh Pandey¶

Clinical proteomics is an emerging field that deals with the use of proteomic technologies for medical applications. With a major objective of identifying proteins involved in pathological processes and as potential biomarkers, this field is already gaining momentum. Consequently, clinical proteomics data are being generated at a rapid pace, although mechanisms of sharing such data with the biomedical community lag far behind. Most of these data are either provided as supplementary information through journal web sites or directly made available by the authors through their own web resources. Integration of these data within a single resource that displays information in the context of individual proteins is likely to enhance the use of proteomic data in biomedical research. Human Proteinpedia is one such portal that unifies human proteomic data under a single banner. The goal of this resource is to ultimately capture and integrate all proteomic data obtained from individual studies on normal and diseased tissues. We anticipate that harnessing of these data will help prioritize experiments related to protein targets and also permit meta-analysis to uncover molecular signatures of disease. Finally, we encourage all biomedical investigators to maximize dissemination of their valuable proteomic data to rest of the community by active participation in existing repositories such as Human Proteinpedia. Molecular & Cellular Proteomics 7: 2038-2047, 2008.

Advancements in proteomics and its clinical applications have led researchers to exploit them to discover protein markers for cancer diagnosis, interrogate key components of signaling pathways, capture protein-protein interactions, dissect organellar proteomes, identify post-translational modifications and to catalog protein expression and subcellular localization profiles (1, 2). Clinical proteomics deals with the application of proteomic technologies to help decipher the changes that occur in cells, tissues, and organs under diseased conditions. With the increase in the use of recent high-throughput technologies such as mass spectrometry, data generation far outstrips the pace of data storage and dissemination. Data once generated can always be revisited and queried in new or different ways that could even lead to potential breakthroughs in terms of identifying diagnostic markers or therapeutic targets. Although proteomic data can be submitted to public repositories, this is neither popular nor mandated, even for published data. Given the high experimental and labor costs in addition to the precious nature of the data, it is imperative that there are concerted community efforts to capturing such data and making them available in formats that would be most useful to biomedical researchers.

Cancer Biomarkers and Disease Proteomics — The potential of mass spectrometry to identify proteins in samples in a high throughput (3) manner with reduced sample requirements have made mass spectrometry an ideal tool to be deployed in clinical proteomics (4). Thus, use of proteomics for identification of cancer biomarkers for diagnostic, prognostic, or therapeutic applications is of substantial interest. In this regard, quantitative analysis of protein expression in normal and cancer tissues to identify proteins overexpressed in cancers has already been successfully reported by a number of groups (5–10). Because it has already been demonstrated that early diagnosis of breast, colorectal, and cervical cancers through screening approaches can lead to a reduction in mortality rates (11), there is sufficient justification for aggressive pursuit of novel biomarkers for early detection of all cancers.

In addition to the search for biomarkers, it is also of interest to identify proteomic changes that occur in diseases to gain insights into their pathogenesis. Such proteomic changes could include alterations in abundance of proteins or their post-translational modifications or subcellular localization, among others (12). In the future, it may even be possible to diagnose a particular disease condition from organ-specific proteomic signatures present in serum. For this, we must first systematically obtain proteomic data from individual organs. Such data can be archived, and meta-analysis can be carried out to decipher the signatures, as was recently reported for head and neck and colon cancers (13).

Is Proteomics Synonymous with Mass Spectrometry?—The routine use of mass spectrometers to identify a multitude of proteins in a high-throughput fashion has led to a situation where the terms "proteomics" and "mass spectrometry" are sometimes used interchangeably. A number of repositories have been developed that only accept data from mass spec-

From the ‡Institute of Bioinformatics, International Tech Park, Bangalore 560 066, India, §Department of Biotechnology, Kuvempu University, Shankaraghatta, Karnataka, India, and ¶McKusick-Nathans Institute of Genetic Medicine and the Departments of Biological Chemistry, Pathology, and Oncology, Johns Hopkins University, Baltimore, Maryland 21205

Received, May 15, 2008, and in revised form, June 23, 2008

Published, MCP Papers in Press, June 23, 2008, DOI 10.1074/ mcp.R800008-MCP200

^{© 2008} by The American Society for Biochemistry and Molecular Biology, Inc. This paper is available on line at http://www.mcponline.org

trometry experiments. However, proteomics includes a broad array of techniques that are still in common use including Western blotting, immunohistochemistry, yeast two-hybrid, peptide and protein microarrays, x-ray crystallography, NMR spectroscopy, fluorescence microscopy, and flow cytometry. Among these techniques, antibody-based methods are especially used in the oncology field for diagnosis and classification of cancers (14). HUPO¹ Antibody Initiative (15) was initiated to accelerate the production and use of validated antibodies against human proteins (16). With the availability of a large number of antibodies, assays such as immunohistochemistry and enzyme-linked immunosorbent assay can be used for biomarker validation. Therefore, it is important to remember the clinical platforms that are relevant to oncology research when proteomic platforms are being discussed.

Genomic Versus Proteomic Data-In the case of genomic data, the International Nucleotide Sequence Consortium has already established a working principle according to which any sequence data that is submitted to any one of the 3 members, GenBank (17), European Molecular Biology Laboratory (EMBL) (18), or DNA Data Bank of Japan (DDBJ) (19), will automatically be reflected in the other data bases. Further, all sequences submitted to these data bases are freely available to the public without any restrictions. This method of data sharing has been in practice for over 20 years now. Further, if a manuscript contains novel sequences, submission of the nucleotide sequences to any one of the three major nucleotide sequence data bases prior to publication is mandatory. In fact, manuscripts are accepted subject to the condition that a unique data base accession number assigned by these data bases will be provided by the authors before publication.

Unlike genomic data, however, proteomic data is diverse with a multitude of experimental platforms and data types with the result that there are no general working principles for data submission that apply to all types of proteomic data. However, for specific data types such as mass spectrometry data, specific guidelines are beginning to emerge (20) although they are not universally adopted at the current time.

Data in Centralized Repositories Versus Supplementary Information—Given the current size of most proteomic data sets, the authors are often unable to accommodate them in the body of the article. Most of them end up publishing the majority of such data as supplementary information either at the web site of the journal or on their own web site (21). However, there are a number of disadvantages of submitting data as supplementary information instead of contributing them to centralized repositories as listed. 1) Most scientific articles are not freely available and preclude many scientists from accessing published articles. Even if the supplementary information is provided freely by the journals, it would be of no use without the original article that is only accessible by a fraction of the scientific community. 2) Data added as supplementary information might not be easily accessible, as most are in pdf or word document formats and cannot be searched readily. 3) The supplementary data provided by the authors generally does not follow a specific format. This makes it difficult to combine independent data sets for data mining or meta-analysis purposes. 4) Retrieving information on a specific gene from supplementary information is not a trivial task because the nomenclature system is often decided by the authors. 5) Supplementary information is most often limited to the web space provided by the journals and large raw mass spectrometry data (in the gigabyte range) are mostly left out.

On the contrary, data contributed to centralized repositories can be downloaded freely, is more searchable, and is often constrained so that common standard formats are used. Moreover, it is possible for information from diverse research articles to be integrated and presented to the user at the context of the protein or a biological pattern as is done in the case of Human Proteinpedia. With the recent advancements in semantic web (22) and data base interoperability (23), it will become even more fruitful for the scientific community to contribute their data to centralized repositories for optimal utilization of data.

Standardization and Vocabulary Issues in Proteomic Data-Gene nomenclature is regulated by human genome organization, whereas naming of proteins is largely left to individual investigators. This is unfortunate because even literature searches are based on text and not sequences, which makes it almost impossible to retrieve the published literature on any given protein in a comprehensive fashion. Some features of proteins are beginning to be standardized using controlled vocabularies such as eVOC (24) for describing tissue expression, Gene Ontology (25) for cellular component, molecular function, and biological process, while RESID (26) and Proteomics Standard Initiative-Molecular Interaction (27) vocabularies are available for post-translational modifications and protein-protein interactions, respectively. Proteomics Standard Initiative-Mass Spectrometry (PSI-MS) vocabularies are used to standardize mass spectrometry-based experimental annotations. Nevertheless, even though these controlled vocabularies are available, they are by no means in common use as major data bases themselves do not always adhere to the available vocabularies (28).

A Need for Unified Information about Proteins – Some of the most popular public repositories store information about specific aspects of proteins. For instance, Protein Data Bank (PDB) (29) is an archive of structural data of biological macromolecules. PRoteomics IDEntifications PRIDE data base (30) and PeptideAtlas (31) are some of the leading mass spectrometry-based data repositories. HPRD (32), IntAct (33), Mint (34), BioGrid (35), and data base of interacting proteins

¹ The abbreviations used are: HUPO, human proteome organization; SMEK1, suppressor of mek1; FGL2, fibrinogen like 2; HCC, hepatocellular carcinoma; HPRD, human protein reference data base.

Ним Home

Experiment Principal In Title E-mail Address Country Lab URL Data submit Title E-mail Published/ Journal nar Published/ Sample sou Sample sou Source org Labeling te Protease u Is the samp Reduction of Mass spect

Tumor p K.TQ 1. R.LG 3. 4.

		man Prote	(III)						1 3	a con a	
	Referenc	e Databa	se								
	-	You		Query >> Tumor pr	rotein D52 like 2						
	1 Query		Tumor p	protein D52	like 2						
	Browse								Molecular Class	Unclassified	
	🧆 Blast								Molecular Function	Molecular function unknown	
	FAQs								Biological Process	Cell growth and/or maintena	ince
			Isoform 1	1					biological i roccas	Con grown and or manicena	
	Downlo	ad									
	1.00										
	Human Proteinpe	odla									
							-	_	-		
	Pathw	avs									
	Phospho Finder	Motif									GO TO: Isoform_1
	Become a Molecule A										
	Molecule A	uthority	SUMMA		DISEASES	PTMs & SUBSTRATES	EXTERNAL LINKS				
	(General								
	1.00		Gene Sym	nbol: <u>TPD</u>	<u>52L2</u>	Molecular Weight ()a):	24854	Gene Map Locus:	20q13.2-q13	3.3
	1										
	faces.		Localiza	ation							
			Primary					Alternate			
								Alternate			
	- 200		Human	n Proteinpedia							
	100 Contra						L	ocalization			
			Otoplas								
			Cytopias	sm 60							
	1.4		Cytopias	<u>sm</u> 60							
				sm co Is and Motifs	1			Expres	sion		
					1			Expres	sion	Site of Expression	
	OTEINPED		Domain	is and Motifs]					Site of Expression	
	COTEINPED		Domain	is and Motifs	ions				n Proteinpedia		
Query L	Login to contribute d	ata Download o	Domain	is and Motifs	ions				n Proteinpedia	Site of Expression	
Query L	Login to contribute d	ata Download o	Domain	is and Motifs	ions			Huma		Site of Expression Disease Tissue	Cell Line
Query L	Login to contribute d esited supporting this a nor protein D52 like 2.(oplasm	ata Download o	Domain	is and Motifs	ions			Huma <u>B Cell</u>	n Proteinpedia	Site of Expression	<u>293T</u>
ataset depo Tun lization Cyto lass spectron	Login to contribute d esited supporting this a nor protein D52 like 2 (oplasm	ata Download (motation (PD52L2)	ata HPRD	is and Motifs				Huma <u>B Cell</u> Brain	n Proteinpedia	Site of Expression Disease Tissue	293T HeLa
ataset depo Tun lization Cyto ass spectron pe on	Login to contribute d esited supporting this a nor protein D52 like 2 (oplasm	Anta Download (motation PD5212) Mass spectrome Platelet subprot	ata HPRD	es and Motifs				Huma <u>B Cell</u>	n Proteinpedia	Site of Expression Disease Tissue	<u>293T</u>
e e e e e e e e e e e e e e e e e e e	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	Anta Download (motation PD5212) Mass spectrome Platelet subprot	Ata HPRD	es and Motifs		eef compariments, such as ment	ranes, cytosol, nucleus and	B Cell Brain Liver	n Proteinpedia	Site of Expression Disease Tissue	293T HeLa
e e e e e e e e e e e e e e e e e e e	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	Mass spectrome Platelet subprot Commercially a cytoskeleton. Pl Dr. Gerard Capy Assistant Profes	Experimental I Prome analysis onitable kits were us telefer proteome, acti py sy r	es and Motifs		ent compartments, such as ment	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue	293T HeLa
e e e e e e e e e e e e e e e e e e e	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	Mass spectrom PIDS2L2) Mass spectrom Platelet subpro Commercially a cytopkaletach Dr. Gerard Cap Assistant Profee gerard.cagney@	Domain: Domain: Dependent of the second Dependent of the second of th	es and Motifs		ent comparisons, such as ment of compared.	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue	293T HeLa
ataset depo Tun ization Cyto ass spectron e on escription	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	Mass spectrome Platelet subprot Commercially a cytoskeleton. Pl Dr. Gerard Capy Assistant Profes	Domain: Domain: Dependent of the second Dependent of the second of th	es and Motifs		onf compariments, such as ment	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue	293T HeLa
taset depo Tun zation Cyte ess spectron escription igator's Nat	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	hta Download o metation Pip5212) Mass spectrum Platele subpro Commercially a Dr. Gerard Cap Platele subpro Commercially a Dr. Gerard Cap Platele subpro Commercially a Dr. Gerard Cap Commercially a Dr. Gerard Cap Sector Platfin University Colle Beffeld, Dublin Treland http://protoco	Domain: ato HPRD Performental I yr ato HPRD ato HPR	Protein Annotati		ent compartments, such as mend	ranes, cylosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
taset depo Tun zation Cyte ess spectron escription igator's Nat	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	Na Download of enotation PPOSICIO Platelet subprol Crytoskelston, PL Dr. Gerard Cog Assistant Profe gerard cogneyy Conway Institut Breard Day Interview Colling Interview Col	Domain: ato HPRO Paperimental 1 rome manylo alable kits were us alable kits were us al	Protein Annotati		onf compared., such as ment	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue	293T HeLa
taset depo Tun zation Cyto ass spectron e on escription igator's National by	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	Mass spectrome prostical metation Prostical Platelet subprot cytoskelston, P Dr. Gerard Cap Assistant Prote gerard capacy Convey Institut University Cole Hathew Salilyon BioInformatics i matthew.salilyon	Domain: ato HPRO Paperimental 1 rome manylo alable kits were us alable kits were us alable the protons, alable were used in the second second rome and the second second second rome and the second second second second second second rome a	Protein Annotati		ert compartments, such as ment of compared.	ranes, cytosol, nucleas and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
taset depo Tun zation Cyto ass spectron e on escription igator's National by	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	hte Download e motation PDS2L3 Mass spectrom Platelet subprot Commercially a cytoskeleton. Pl Dr. Gerard Capey Assistant Profe gerard.capey Comvey Institut Belfield, Dublin Treland http://protoca Mathew Sulliva BioInformatics	Domain: ato HPRO Paperimental 1 rome manylo alable kits were us alable kits were us alable the protons, alable were used in the second second rome and the second second second rome and the second second second second second second rome a	Protein Annotati		ert compartments, such as ment	ranes, cytosol, nucleas and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
uery L ntaset depo ization Cyto ass spectron	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	No Download of motation PDSID	Domain: ato HPRO Paperimental 1 rome manylo alable kits were us alable kits were us alable the protons, alable were used in the second second rome and the second second second rome and the second second second second second second rome a	Protein Annotati Protein Annotati Details - HuPA 00007 sed for crude fractionati vivated and resting states		ent compariments, such as ment	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
tterry I I taset depo I I I zation Cyty ass spectron e on escription igator's Nar by ablished	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	No Download d metation Platele shiprof PD522.2	Domain: ato HPRD Paperimental 1 77 ato HPRD atable kits were su- atable kits were su- station of the su- t substation of the substation of the sub- station of the substation of the substation of the sub- station of the substation of the substation of the sub- station of the substation of the substation of the sub- station of the substation of	Protein Annotati Protein Annotati Details - HuPA 00007 sed for crude fractionati vivated and resting states		ent compartments, such as ment	ranes, cylosof, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
taset depo Tun zation Cytv ess spectron escription igator's Nar bby bbilished	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	to Download of motation PDSLDD PDSLDD Mass spectrom Platelet subpro- Commercially a cross-spectrom Platelet subpro- commercially a cross-spectrom but- devector se	Domain: ato HPRD Paperimental 1 77 mase analytic allable kits were su- allable kits were su- re obsilin allable kits were su- allable kits were su- allable kits were su- re obsilin allable kits were su- allable kits were su- station of the super- supersupersupersupersupersuper- supersupersupersuper- allable kits were supersuper- supersupersupersuper- supersupersupersuper- supersupersupersupersuper- supersupersupersupersupersuper- supersupersupersupersuper- supersupersupersuper- supersupersupersupersupersuper- supersupersupersupersuper- supersupersupersupersuper- supersupersupersupersupersupersupersuper	Protein Annotati Protein Annotati Details - HuPA 00007 sed for crude fractionati vivated and resting states		ent compartments, such as ment	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
tery 1 te	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	No Download d medation PB522.2.3 Mass spectrom Platele subject Commercially a Dr. Gerard Cap Corr, Gerard Cap Corr, Gerard Cap Corr, Carard Cap Corr, Cap Corr, Carard Cap Corr, Corr, Cor	Domain: ato HPRD Paperimental 1 77 mase analytic allable kits were su- allable kits were su- re obsilin allable kits were su- allable kits were su- allable kits were su- re obsilin allable kits were su- allable kits were su- station of the super- supersupersupersupersupersuper- supersupersupersuper- allable kits were supersuper- supersupersupersuper- supersupersupersuper- supersupersupersupersuper- supersupersupersupersupersuper- supersupersupersuper- supersupersupersupersuper- supersupersupersupersuper- supersupersupersupersupersuper- supersupersupersupersuper- supersupersupersupersupersupersupersuper	Protein Annotati Protein Annotati Details - HuPA 00007 sed for crude fractionati vivated and resting states		orf compared. such as ment	ranes, cytosol, nucleus and	Huma B Cell Brain Liver Ovary	n Proteinpedia	Site of Expression Disease Tissue <u>Ovarian Cancer</u>	293T HeLa
tery I a taset dopo I un attain Cyte ss spectrors n ssscription blished n ue comminged	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	No Download d	Domain: Domain: Deperimental 1 Province analysis ata HPRD Province analysis ata sea analysis at	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		ent compartments, such as ment and compared.	ranes, cytosol, nucleas and	B Cell Brain Liber Ovary Platelet	n Proteinpedia	Site of Expression Disease Tissue Divatian cancer Comments	293T HeLa
tery A testet daps. tantion Cytic ss spectrons n sscription blished highlight daps. n n tester tester teste tester tester tester tester tester test	Login to contribute d ssited supporting this at mor protein DS2 like 2 (1 oplasm imetry	No Download d	Domain: Dom	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		onf compariments, such as ment	ranes, cytosol, nucleus and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2 <u>231</u> HeL <u>3</u> K-562
tery and term of the second se	Login to contribute d wind supporting this a merry relation 52 like 2 (metry me	No Download Passage Common Passage Common	Domain: Dom	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		ent compartments, such as ment	ranes, cylosol, nucleus and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Divatian cancer Comments	2 <u>231</u> HeL3 K-562
utery 1 and topo Taster topo Tasterto Tasterto Tasterto Taster topo Taster top	Login to contribute d wind supporting this a merry like 2 (metry me atabase searching (MS) atabase searching (MS)	Mass spectrome Platele solycov PD522.2 Mass spectrome Platele solycov Cytokeleton, Pl Dr. Gererd Capy Assistant Profe gerand.caparoy University Cole Balifield Mathem BioInformatics Mathew Salibos BioInformatics Mathew Salibos Mathew Salibos Mathew Salibos Not applicable Tissue: Plate Cell Ine; Homo sapicable Tissue: Plate Cell Ine; Homo sapicable Types No No No No No Salifield Mathematics Homo sapicable Tissue: Plate Cell Ine; Homo Salifield No No So De Salibos	Domain: Dom	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		nt compared. such as ment	ranes, cytosol, success and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2 <u>231</u> HeL <u>3</u> K-562
utery 21 and 25	Login to contribute d wind supporting this a margeretic D2 like 2 (metry	No Download d medation Pab521 2 Mass spectrome Pab522 2 Mass spectrome Pab522 2 Masses spectrome Pabeles Dr. Gererd Cap Assistant Profe gerand-agnetic Compared and the Compared and the pabeles Compared and the pabeles Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan Mathew Saillyan BioInformatics Mathew Saillyan Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan BioInformatics Mathew Saillyan Mathew Saillya	Domain: Dom	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		ert compartments, such as ment	ranes, cytosol, nucleas and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2 <u>231</u> HeL <u>3</u> K-562
nerry 20 Tasset for a second s	Login to contribute d wind supporting this a margeretic D2 like 2 (metry	No Download d	Domain: Dom	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		ent compariments, such as ment	ranes, cytosol, nucleus and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2231 Hela K-562
enery of a case of energy of a case of energy of a case of energy of a case	Login to contribute d satural supporting this a marg potentia 052 like 2 (oplass metry stabase searching (HS; atabase searching (HS; ng	No Download d modation Pabola States Pabola Stat	Experimental T Experimental T Try Try Service analysis salable kills were as salable	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		ort compartments, such as ment	ranes, cylosod, nucleus and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2 <u>231</u> HeL <u>3</u> K-562
An and a second	Login to contribute d wind supporting this a margeretic D2 like 2 (metry	No Download d modation Pabola States Pabola Stat	ata Deperimental 1 September 1	Details Huffs 00007 Details Huffs 00007 sed for crude fractional invited and resting states html		ent compartments, such as ment and compared.	ranes, cytosol, nucleus and	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2231 Hela K-562
arry of the second seco	Login to contribute 4 wind supporting this a more protein 52 like 2 (possion metry me atabase searching (HS, ng (TP252L2) (s localized)	Download	ata Deperimental 1 September 1	ns and Motifs Protein Annotati Details HurA 00007 sed for crude fractionati html 11301 11301	to a of cells hato differ		MS/MS Spectrum InDA_514065	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2231 Hela K-562
the second secon	Login to contribute 4 wind supporting this a more protein 52 like 2 (possion metry me atabase searching (HS, ng (TP252L2) (s localized)	No Download Packard States Spectrame Plateles subjects Packard States Spectrame Plateles subjects Dr. Gereard Capa Dr. Gereard Capa Dr. Gereard Capa Common Spectrame Native Sulliva Biolinformatics Interlead, Duble Interlead, Duble I	ata Deperimental 1 September 1	Is and Motifs Protein Annotati Datails - HuPA_00007 sed for closed http://datails.and/focused http://	on of calls into differences of the second sec	Sequence Mentifier	K/MS Spectrum	B Call Brain Lorr Platele	n Proteinpedia	Site of Expression Disease Tissue Ovarian cancer Comments	2231 Hela K-562

Fig. 1. Display of expression and subcellular localization of tumor protein D52-like 2 in HPRD molecule page. Molecule page of tumor protein D52-like 2 in HPRD is displayed. Almost all information for this protein is derived from community annotations through Human Proteinpedia including subcellular localization and expression in tissues, cell lines, and diseases. The annotated data shows that this molecule is expressed in B cells, brain, liver, ovary, and platelets. It is also expressed in ovarian cancer and in several cell lines (293T, HeLa, and K-562). Clicking on any of these hyperlinked terms opens a pop-up window (e.g. cytoplasm or platelet, as shown), which provides additional experimental data and details about the contributing laboratory as well as any publications. For example, the window on the left shows peptide identification data, peptide scores, precursor mass, charge state, and sequence identifiers from this unpublished study. If available, the MS/MS spectra are hyperlinked to another window as shown in the right lower part that allows the users to manually inspect the data.

otei	n				ATT O C	<u>_</u>
base	e					263
You are	e at: <u>HPRD</u> >> (Query >> Tumor protein D52 like 2				
	Tumor p	rotein D52 like 2				
				_	Molecular Class Unclassified	
					Molecular Function Molecular function unkno	wn
					Biological Process Cell growth and/or mainte	enance
	lsoform 1					
	ALTERNAT		s & SUBSTRATES			GO TO
F	SUMMAR		TERACTIONS EXTER	RNAL LINKS		
	Protein I	nteractions				
	PROTEIN II	NTERACTORS				
		Name of Interact	or		Experiment Type	1
	Mal T cell d	fifferentiation protein 2			In Vitro ; Yeast 2 Hybrid	C
	Tumor prot	tein D52-like 1			Yeast 2 Hybrid	(
		tein D52 like 2			Yeast 2 Hybrid	C
	Tumor prot				Yeast 2 Hybrid	(
	Tumor prof	tein D52-like 3			In Vitro ; Yeast 2 Hybrid	E
	Human	Proteinpedia				
		Name of Interactor			Experiment Type	Туре
	IKKE			In Vivo		Direct
			V			
		Home Query Login to contribute data	Download data HPRD	Protein Annotation	5	
	1 e	Home Query Login to contribute data		Protein Annotation	5	
		xperimental dataset deposited supporting this annotation		Protein Annotation	5)	
	1.	xperimental dataset deposited supporting this annotation		Protein Annotation		
		xperimental dataset deposited supporting this annotation Platform: Mass spectrometry Experiment type	Experimental D Mass spectrometry Large-scale mapping of protein-prote	etails - HuPA_00144 in interactions by mass sp	ectrometry	
		perimental dataset deposited supporting this annotation Platform: Mass spectrometry Experiment type Short description	Experimental D Mass spectrometry Large-scale mapping of protein-prote	etails - HuPA_00144 in interactions by mass sp	ectrometry	
		Platform: Mass spectrometry Platform: Mass spectrometry Experiment type Short description Experimental description	Experimental C Mass spectrometry Large-scale mapping of protein-prote 338 bait proteins were selected bas tagged versions of these proteins for 6 46/3 interactions between 2233 d 6 46/3 interactions between 2233 d	etails - HuPA_00144 in interactions by mass sp		_
		Platform: Nass spectrometry Platform: Nass spectrometry Experiment Type Short description Experimental description Platform: Name Title	Experimental D Mass spectrometry Large-scale mapping of protein-prote 328 balt protein were selected bas Lapped versions, of these proteins to 6463 interactions between 3225 Ur. Daniel Figuys Canada Research Chair in Proteomics	etails - HuPA_00144 in interactions by mass sp d on known or suspected liowed by LC-ESI-HS/MS is the filtered out using empl is the typoteins.	ectrometry	
		Patform: Mass spectrometry Patform: Mass spectrometry Experiment type Short description Principal Investigator's Name Title c-mail	Experimental D Hass spectrometry Large - colt megaling of protein-prote 238 balt proteins were selected has tapped versions of these proteins of False proteins and redendaris this w larged versions of these proteins of the protein second selection of the D. Daniel Flagers, Southerne, a	etails - HuPA_00144 in interactions by mass sp d on known or suspected liowed by IC-ESI-HS/MS en filtered out using empl Stinct proteins. and Systems Biology	ectrometry	
		Platform: Mass spectrometry Platform: Mass spectrometry Cooperiment type Short description Principal Investigator's Name Title Email Address	Experimental D Mass spectrometry Large-scale mapping of proteins prote 338 balt proteins were selected lass tapped versions of these proteins of of 6463 literactions between 2233 d Dr. Daniel Figory Canada Research Chair In Proteomics dignosybuitteme.ca Case Center for Proteomics/Departm Ottome your KL 1463	etails - HuPA_00144 in interactions by mass sp d on known or suspected liowed by IC-ESI-HS/MS en filtered out using empl Stinct proteins. and Systems Biology	ectrometry	
		Patform: Mass spectrometry Patform: Mass spectrometry Experiment type Short description Principal Investigator's Name Title c-mail	Experimental D Hass spectrometry Large - colt megaling of protein-prote 238 balt proteins were selected has tapped versions of these proteins of False proteins and redendaris this w larged versions of these proteins of the protein second selection of the D. Daniel Flagers, Southerne, a	etails - HuPA_00144 in interactions by mass sp d on known or suspected liowed by IC-ESI-HS/MS en filtered out using empl Stinct proteins. and Systems Biology	ectrometry	
			Experimental D Hass spectrometry Large-scale mapping of protein-prote 23B half proteins were selected has probability of the selected has probability of the selected has probability of the selected has probability of the selected has difgory questions.ca Case Coeffer (or broteins/Chapman Biology Case Coeffer (or broteins/Chapman Biology Case Coeffer (or broteins/Chapman Biology Case Coeffer (or broteins/Chapman Biology Case) of the selected has different of the selected has differe	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
		Platform: Mass spectrometry Platform: Mass spectrometry Experiment type Short description Experiment description Frincipal Investigator's Name Title Frinal Address County Data submitted by	Experimental IT Mass spectrometry Large-scale mapping of proteins prote 338 half proteins were selected basi proteins and redentiant have of 64.63 interactions between 223.54 Dr. Daniel Florger Dr. Daniel Florger Dr. Daniel Florger Dr. Bar (Strateger Cargory Daniel and redentiant have Cargory Daniel and redentiant have Cargory Daniel of Articles Cargory Daniel (Strateger Ottawa Institute of Systems Biology Ones Conference Proteinos/Cheganiel Cargory Daniel (Strateger Cargory Daniel (Strateger Cargory Daniel (Strateger Cargory Daniel (Strateger Ottawa Institute of Systems Biology Ottawa Dr. Rob M Wrwing Assistant Professor & Director of Biol rob erwingScats.edu	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
		Platform: Nass spectrometry Platform: Nass spectrometry Platform: Nass spectrometry platform: Nass spectry platform: Nass spectry platform: Nass spectrometr	Experimental IT Mass spectrometry Large-scale mapping of proteins prote 338 half proteins were selected basi proteins and redentiant have of 64.63 interactions between 223.54 Dr. Daniel Florger Dr. Daniel Florger Dr. Daniel Florger Dr. Bar (Strateger Cargory Daniel and redentiant have Cargory Daniel and redentiant have Cargory Daniel of Articles Cargory Daniel (Strateger Ottawa Institute of Systems Biology Ones Conference Proteinos/Cheganiel Cargory Daniel (Strateger Cargory Daniel (Strateger Cargory Daniel (Strateger Cargory Daniel (Strateger Ottawa Institute of Systems Biology Ottawa Dr. Rob M Wrwing Assistant Professor & Director of Biol rob erwingScats.edu	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Mass spectrometry Large-scale mapping of proteins prote 23B kalf proteins were selected lass frake positives and relatediant bits price positives and relatediant bits price positives and relatediant bits price positives and relations of the price positive and relations of the Canada Research Chair in Proteomics Classicatar Performations (Page) Canada Dr. Rob H Weing Assistant Performs A Director of Iloi robusingGrass-du Published understant	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Mass spectrometry Large-scale mapping of proteins prote 33B kalf proteins were selected lass frake positives and relatediant bits processing and relationship they have been been been been been been been chandle Research Chair in Proteomics difgory gloutenew.ca Caasc dense for proteomics/Department Ottom on on ULL 14 Section Chandle Chair Chair Chair Charles and Chair Charlow Charles and Charles and Charles Charles and Charles and Charles Charles and Charles and Charles and Charles and Charles and Charles and Charles and Charles and Charles and Char	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Nass spectrometry Large-scale mapping of proteins prote 338 balt proteins were selected base independent of the selected base independent of the selected base of 64.63 literactions between 225.54 On Daniel Flogra. Canada Benesarch Charl in Proteomics differed base of the selected base of the selected base of the selected differed base of the selected base of the selected base of the selected differed base of the selected base of the selected base of the selected differed base of the selected base of the selected base of the selected differed base of the selected base of the selected base of the selected differed base of the selected base of the selected base of the selected base of the selected base of the selected base of the selected base of the selected base of the difference of the selected base of the selected base of the selected base of the selected base of the difference of the selected base of t	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Nass spectrometry Large-scale mapping of proteins prote 338 balf proteins were selected lass fapped versions of these proteins of of 64.63 literactions between 225.54 On Daniel Flagory Canada Besearch Chail in Proteomics dipro-Southane 225.54 Dr. Balf Projections and the dipro-Southane 225.54 Canada Dr. Rob H (Wolg Assistant Professor & Director of Biol Assistant Professor & Director of Biol Southane (BH 1995) Tissues : HET2933 Cell line:	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Mass spectrometry Large-scale mapping of proteins prote 33B kalf proteins were selected has processing of the selected has processing of the selected has processing of the selected has processing of the selected has differed proteins and readenant has chands Research Andre in Proteonics/ Department Chands of Selected has differed proteins (Change and Selected To Rob H Wring Assistant Professor & Director of Inter- published manager bar basistant professor & Director of Inter- basistant professor & Director & Directo	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Hass spectrometry Large-scale mapping of protein-prote 33B half proteins were selected has frage-scale mapping of protein-prote 33B half proteins were selected has for character for the selected has dispersively and the selected has dispersively and the dispersively and the selected has dispersively and the selected has dispersively and the selected has dispersively and the selected has dispersively and the selected has dispersively and the selected has dis	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Nass spectrometry Large-scale mapping of proteins prote 338 balf proteins were selected lass fapped versions of these proteins of of 64.63 literactions between 225.9 d Dr. Daniel Flagys Canada Besearch Chair la Proteonics dignsy-fauetane action Cate Center for Proteonics Distribution Cate Center for Proteonics Distribution Distribution Center for Center for Center Distribution Center for Center for Center Trypsin Instrument: LCQ Decs	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Nass spectrometry Large-scale mapping of proteins prote 338 balf proteins were selected lass fapped versions of these proteins of of 64.63 literactions between 225.9 d Dr. Daniel Flagos, Canada Besearch Chail in Proteomics dipro-youthown 225.9 d Dr. Balf Proteins for the proteomics dipro-youthown 225.9 d Dr. Rob H Wolg Assistant Professor & Director of Biol Assistant Professor & Director B	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Mass spectrometry Large-scale mapping of proteins prote 23B half proteins were selected has free positive and reduction the ver- price positive and reduction the ver- price positive and reduction the ver- class center for proteins(2) Department diago youtherw.ca Canada Beserver Alon in Proteonics (2) Casa Center for Proteins(2) Department Ottow of Section 100 of the case Center for Proteins(2) Department diago youtherw.ca Casa Center for Proteins(2) Department diago youtherw.ca Casa Center for Proteins(2) Department diago youtherw.ca Case Center for Proteins(2) Department diago youtherw.ca Case Center for Proteins(2) Department biology 12333301 Taseus: III: Traysia Cell line: Homo splies [Izonemy: Section] New Yes	etails - HuPA_00344 interactions by mass go d on known or suspected for the suspected for the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the suspect of the su	ectrometry	
			Experimental IC Insessectionentry Large-scale mapping of proteins protein Targe-scale mapping of proteins protein Targe-scale mapping of proteins protein Targe-scale mapping of proteins proteins Targe-scale mapping of 64.51 interaction targe scale of the scale of t	netalis - HuirA, DO144 in Interactions by mass sp do halown or xappedie static approximation of the state of the state approximation of the state of the state of conseties ent of Genetics	ectrometry	
			Experimental C Hass spectrometry Large-scale mapping of proteines for package devisions of these proteines for package packings and relationship to the package packing and relationship to the characteristic of the proteomics of the characteristic of the characteristic of the characteristic of the proteomics of the characteristic of the proteomics of the characteristic of the characteristic of the characteristic of the characteristic of the characteristic of the characteristic of the ch	netalis - HuirA, DO144 in Interactions by mass sp do halown or xappedie static approximation of the state of the state approximation of the state of the state of conseties ent of Genetics	ectrometry	
			Experimental IC Mass spectrometry Large-scale mapping of proteins prote 230 Bal (proteins process) process process process and relationship (proteins) process process process) process process process digosydoutawa.ca Canada Research Chair in Proteomics/ digosydoutawa.ca Canada Research Chair in Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Departmet Sociatar Proteomics/ Departmet Sociatar Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Sociatar Pr	netalis - HuirA, DO144 in Interactions by mass sp do halown or xappedie static approximation of the state of the state approximation of the state of the state of conseties ent of Genetics	ectrometry	
			Experimental IC Mass spectrometry Large-scale mapping of proteins prote 230 Bal (proteins process) process process process and relationship (proteins) process process process) process process process digosydoutawa.ca Canada Research Chair in Proteomics/ digosydoutawa.ca Canada Research Chair in Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Departmet Sociatar Proteomics/ Departmet Sociatar Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Discourse (Proteomics/ Departmet Sociatar Proteomics/ Sociatar Pr	netalis - HuirA, DO144 in Interactions by mass sp do halown or xappedie static approximation of the state of the state approximation of the state of the state of conseties ent of Genetics	ectrometry	
			Experimental IC Assessectometry Large-scale maging of proteins were selected has Targe-scale to the target of target of the target of target o	netalis - HuirA, DO144 in Interactions by mass sp do halown or xappedie static approximation of the state of the state approximation of the state of the state of conseties ent of Genetics	ectrometry	

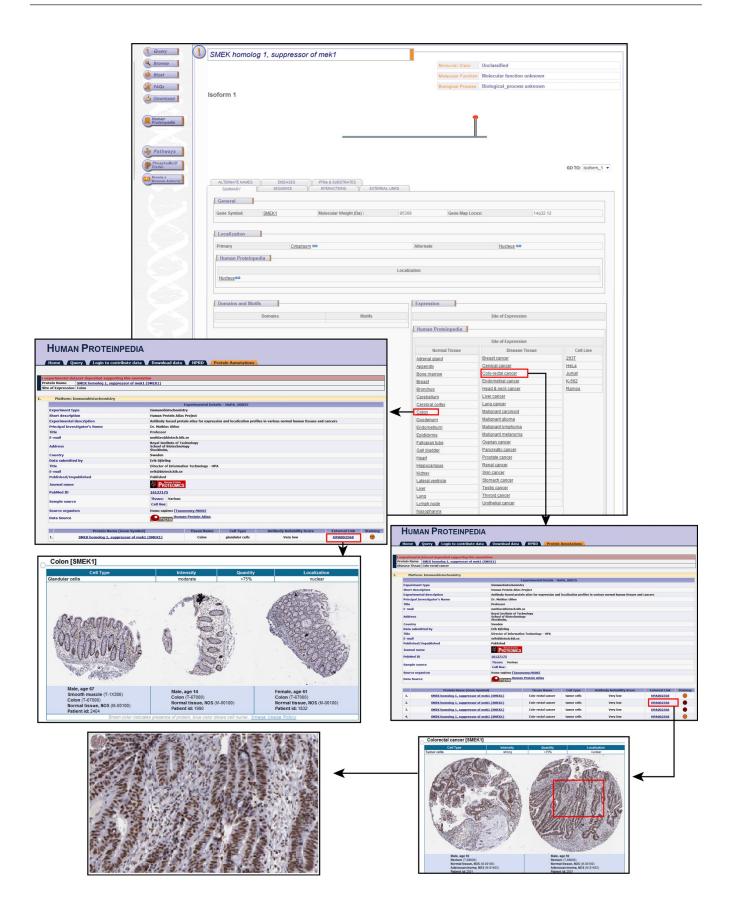
Fig. 2. Display of post-translational modifications and protein interactors for tumor protein D52-like 2. *a*, the molecule page for tumor protein D52-like 2 is shown with several interacting proteins manually annotated in HPRD and one protein, I-Kappa-B Kinase-Epsilon, based on data contributed to Human Proteinpedia from a mass spectrometry experiment. The experimental details along with information about the contributing laboratory are also shown. *b*, no curated post-translational modifications exist for this protein in HPRD. However, the Human Proteinpedia tab shows that there are two phosphorylation sites that have been contributed based on a published study. The *lower panel* provides a description of the experiment, phosphopeptides identified, and the peptide score.

b

You are at HPRD >> Query >> Tumor protein D52 like 2 Tumor protein D52 like 2	
You are at <u>HPRD</u> >> <u>Query</u> >> Tumor protein D52 like 2	
1 Tumor protein D52 like 2	1
	Molecular Class Unclassified
	Molecular Function Molecular function unknown
1	Biological Process Biological_process unknown
Isoform 1	biological Frocess biological_process anknown
	GO TO: Isofo
SUMMARY SEQUENCE INTERACTIONS	~
ALTERNATE NAMES DISEASES PTMs & SUBSTRATES EXTERNAL LINKS	
PTMs	Substrates
Residue Type Site Upstream	m Enzymes Title Residue Type Si
include inpe site opsitied	
Human Proteinpedia	
Desidue Tons Dia U	and an
Residue Type Site Up S Phosphorylation 189 189	pstream Enzymes
T Phosphorylation 196	
HUMAN PROTEINPEDIA	
Platform: Mass spectrometry	
Experimental Details - HuPA_00389	
Experiment type Mass spectrometry Short description 293T phosphoproteome analysis - ETD - Lys-c	
Experiment type Mass spectrometry Short description 2927 theophopreteem anylysis - LTD - Lys-c Experimental description Calls trasted with plosphatase inhibitors, tractor tractor were used.	sctionated using reversed phase chromatography. Three different proteases sing fitanium dioxide.
Experiment type Ress spectrometry Short description P3217 phospharporene analysis - ETD - Lys - c Experimental description Calls treated with phospharpore in the phospharpore in the phospharpore is Principal Treating of the phospharpore is Dr. Abilities Head with	actionated using reversed phase chromatography. Three different proteases
Experiment type Hess spectrometry Short description 2037 [Josobydownolesse analysis - ETD - Lys - c Experimental description Crist treated with phosphatese lubilityme, Tasket proteins were for principal truestigator's Name Principal Investigator's Name Dr. Abilities Nearly Title Associate Professor Title Associate Professor	vcloaated using reversed phase chromatography. Three different proteases sing thankon disuble.
Experiment type Mass spectrometry Short description 2021 phospharporesem analysis - ETD - Lys-c Experimental description Calls treated with phospharporesem analysis - ETD - Lys-c Principal provides analysis Calls treated with phospharporesem analysis - ETD - Lys-c Principal provides analysis Calls treated with phospharporesem analysis - ETD - Lys-c Principal provides analysis Or. Addless Phonely Title Associate Professor E-mail pandrog Planuedon Address Johns topaking and provides	sclonated using reversed phase chromatography. Three different proteases sing Itlanium dioxide.
Experiment type Mass spectrometry Short description 2937 plosybapyoreone analysis - ETD - Lys -c Experimental description CBI treated with plosphatese inhibitors, Instate more for were used. Digitated proteins were enriched for phosphapeptides un Principal Investigator's Name Dr. Abales.P andery Title Associate Professor E-read E-read pasteryspin.deu Dates inspin.deu Address Dates inspin.deu Dates inspin.deu Address Dates inspin.deu Dates inspin.deu Country USA USA Dates inspin.deu	xclionated using reversed phase chromatography. Three different proteases
Experiment type Mass spectrometry Short description 2037 Josphaphoretone analysis - ETD - Lys - c Experimental description Cells treated with phosphatese inhibitors, Instart proteins were for over used, Dipated proteins were end, of the phosphape phase Principal Investigator's Name Dr. Ahilies Pradey Title Associate Proteins Associate Proteins E- mail pader phase Datas isolation University of the phosphape phase Address Address Datas isolation University of Datas submitted by USA Lab URL http://randeykab.gm.lim.edu Data submitted by Dr. iterrit Notas	sclausted using reversed phase chromatography. Three different proteases
Experiment Type Ness spectrametry Short decorption 2021 Responsemental Approximation approximati approximati approximation approximation approximation approximat	cclusated using reversed phase chromatography. Three different proteases
Experiment Type Mess Spectrometry Short docsription 2031 Thosphaperseem analysis - ETD - Lys - C Experimental description Click treated with phosphates leakhblers, Tated processing and the phosphaperseem analysis - ETD - Lys - C Precided Treated with phosphates leakhblers, Tated processing and the phosphaperseem analysis - ETD - Lys - C Click treated with phosphates leakhblers, Tated processing and the phosphase photos of the phosphates leakhblers, Tated processing and the phosphaperseem analysis - ETD - Lys - C Precided Treated State Processing and the phosphates leakhblers, Tated phosphate	ccloaded using reversed phase chromatography. Three different proteases
Experiment type Mass spectrometry Short description 2021 plosphoretone analysis - ITD - Lys - c Experimental description Call treated with phosphates inhibitors, instar processing with the phosphates inhibitors, instar phosphotes inhibitore inhibitors, instar phosphotes inhibitors, instar	schanned uning reversed phase chromatography. Three different protoases for this have divide.
Experiment type Mass spectrometry Short description 2021 plosphoreneous analysis - ITD - Lys <	sclausted using reversed phase chromatography. Three different proteases
Experiment type Mess sectoratory Short description 2017 psolyaphysicseem analysis - ITD - Lys <-	sclausted using reversed phase chromatography. Three different proteases
Experiment type Mass spectrometry Short description 2021 polyaphysicsee analysis - ETD - Lys <	sclosated using reversed phase chromatography. Three different proteases
Experiment type Notes spectrometry Short description 2021 polephysicrones analysis - ETD - Lys <	schanned uning recorded phase chromatography. Three different protoases
Experiment type Mass spectrometry Stort decryption 2021 Decryptoresen analysis - ITD - Lys <	sclausted using reversed phase chromatography. Three different proteases one thanken disate.
Experiment type Mass spectrometry Short description 2021 Jobs/provensem analysis - ETD - Lys <	sclausted using reversed phase chromatography. Three different proteases one thanken decade.
Experiment type Mass spectrometry Short description 2017 plosphoresem analysis - tTD - tys - c Experimental description Cell treated with phospharse inhibitory. Unside the phosphare inhibitory. Unside th	schanned using recorded phase chromatography. Three different proteases often instance divisio.
Experiment type Mass spectrometry Short description 2021 piophysicates mithibure. Trait of profiles with the prophetices mithibure. Trait of profiles with the prophetices with the prophetic	schlansnel uning revensed phase chromatography. Three different proteases
Experiment type Mass spectrometry Stort docuption 2021 produptoresen analysis - ITD - Lys <	scelarated using revenued phase chromatography. Three different proteases tigs that an dease.
Experiment type Mass spectrometry Stort occupient 2021 polyaphysicrosen analysis (TD - Lys < Experimental description Control Control Principal Investigator's Name Control Principal Investigator's Name Stort end with polyaphysicrosen analysis (TD - Lys <	sclausted using reversed phase chromatography. Three different proteases of the final decade.
Experiment Type Mass spectrometry Bind occupition 2031 Topologioarroeme analysis - ETD - Lys <- Experimental description Cell treated with phosphates with Bindra, Bater profiles were for obsphates with the phosphates with Bindra, Bater phosphates were project provided treated with phosphates with Bindra, Bater phosphates were project provided treated with phosphates were bindra, Bater phosphates with the project provided treated with phosphates were bindra, Bater phosphates with the project provided treated with phosphates were bindra, Bater phosphates with the project provided treated were bindra, Bater phosphates with the project phosphates were bindra, Bater phosphates with the phosphates were bindra, Bater phosphates with the phosphates were bindra, Bater phosphates with the provided phosphates were bindra, Bater phosphates with the phosphates were bindra, Bater phosphates with the phosphates were bindra, Bater phosphates were bindra, Bater phosphates were bindra, Bater phosphates were bindra, Bater phosphates were b	scriptaned using recorded phase chromatography. Three different proteases

Fig. 2-continued

(36) are some of data bases capturing protein-protein interaction data. LifeDB (37) catalogs subcellular localization, whereas Human Protein Atlas (38) archives immunohistochemistry data. These data bases were designed to either collect or accommodate data only from specific experiment types; very few archive data from multiple platforms. Thus, it is currently impossible for a researcher to view all of these data stored in these specialized data bases in one location. Further, there is a lack of mechanisms to automatically exchange most proteomic data types between repositories



without substantial manipulation and, in most instance, manual intervention or curation.

In developing a resource for housing proteomic data including that from clinical proteomics, two major issues should be considered. The first is that the data should be shared regardless of the size of the dataset (*i.e.* it is not just high-throughput data that are worth sharing; data from individual experiments is often even more valuable and should not be ignored). Second, there should be a central portal where the available data is compiled and displayed in the context of a gene/ protein. The latter feature would permit users to construct complex queries such as "what are the post-translational modifications on my protein of interest, its interacting proteins, its subcellular localization, and if it is overexpressed in cancers". Such queries cannot be made in any of the existing proteomic repositories although some provide links to other data bases for certain data types.

Human Proteinpedia as a Portal for Basic and Clinical Information about Proteins-Human Proteinpedia (39) is a community portal for sharing human proteomic data that is developed with the active participation of more than 70 laboratories around the world. It allows researchers to share their human proteomic data in a manner that is somewhat similar to that of Wikipedia. However, experimental evidence is mandatory for inclusion of data in Human Proteinpediaand; the contributions are always linked to the investigator and the laboratory. Annotations pertaining to post-translational modifications, expression in cell lines or tissues, protein-protein interactions, enzyme substrate, and subcellular localization can be submitted. Human Proteinpedia includes data from diseases such as cancers thereby allowing the biomedical community to take a system's view of the disease proteome. Moreover, it can accommodate data from multiple experimental platforms such as yeast two-hybrid screens, peptide/protein arrays, immunohistochemistry, Western blots, mass spectrometry, co-immunoprecipitation, and fluorescence microscopy.

Thus, Human Proteinpedia represents an early attempt to unify human proteomic data under a single resource. An important feature of Human Proteinpedia is that it displays the data in the context of proteins that are annotated in HPRD, a literature curated data base for human proteins (32). An example of tumor protein D52-like 2, which is an uncharacterized protein, will illustrate how Human Proteinpedia can not only handle the complex query described above but provide meaningful answers that otherwise might be difficult to find or derive. Fig. 1 shows the expression of tumor protein D52-like 2 in normal tissues, diseases, and in cell lines along with its subcellular localization. These are all based on data submitted by the community, and the name of the contributing laboratory is clearly displayed when a user clicks on a link (the figure shows the link from the term "cytoplasm" and "platelet"). In addition, in this case, we would not know that this protein is expressed in ovarian cancer without the data contributed by the community. Similarly, Fig. 2*a* shows that tumor protein D52-like 2 interacts with I-Kappa-B Kinase-Epsilon, a kinase that phosphorylates IkappaB- α , based on a large-scale protein interaction mapping experiment. Finally, Fig. 2*b* shows that this protein is phosphorylated on serine and threonine residues with links to the primary data that can be explored by the users.

Likewise, Fig. 3 shows the molecule page of suppressor of mek1 (SMEK1) in HPRD. The molecule is unclassified and its site of expression in normal human tissues is also unknown in the literature. However, annotations contributed by the scientific community through Human Proteinpedia reveal the site of expression of SMEK1 in normal and disease tissue as well as cell lines (Fig. 3). These annotations reveal that SMEK1 is moderately expressed in glandular cells of normal colon tissue while being strongly expressed in tumor cells of colorectal cancer tissue. Fig. 4 shows the expression of an extracellular matrix protein, fibrinogen like 2 (FGL2), in hepatocellular carcinoma (HCC). This protein, similar to fibrinogen β and γ , was not previously reported to be involved in HCC. However, it is shown to be expressed in HCC by immunohistochemistry as well as by Western blotting (Fig. 4). Given the fact that early diagnosis will improve prognosis, it is important to pursue such overexpressed molecules, which could turn out to be potential biomarkers.

Human Proteinpedia have several advantages over other proteomic resources with respect to clinical proteomic data. Human Proteinpedia incorporates data from multiple experimental platforms, whereas most of the centralized repositories accumulate data from one or two experimental platforms. Given the advantages of each proteomic platform, integration of clinical data produced from all of them under a single banner was lacking. However, Human Proteinpedia displays such clinical information along with the literature-curated data in the context of a protein molecule. With gaining popularity, we expect that even more diverse clinical studies will be integrated and it will be possible to extract biologically meaningful patterns of molecules expressed in particular disease conditions. Further, such data could drive planning of new clinical studies.

Conclusions and Outlook—To systematically take advantage of the explosion in proteomic data, it must be captured efficiently for the explicit purpose of sharing with the community. In this regard, the researchers should pursue depositing

Fig. 3. SMEK1 expression in colon and colorectal cancer. SMEK1 molecule page is shown with links to the Human Proteinpedia page indicating expression in colon and colorectal cancer (*highlighted*), among other sites and diseases. Links from colon displays the experiment description and the information of the contributing group. Human protein atlas links are provided from the Human Proteinpedia page, which indicate moderate expression of SMEK1 in the glandular cells of normal colon tissue. A hyperlink from colorectal cancer again leads to the same resource, which reveals strong expression of SMEK1 in the tumor cells in colorectal cancer tissue.

Human P Reference Date	Prote abas	in Ge re at: <u>HPRD</u> >> <u>Query</u> >> Fibrin	nogen like 2			43		
Query		Fibrinogen like 2	2			I		
Browse						Molecular Class	Extracellular matrix protein	
🧆 Blast						Molecular Function	Extracellular matrix structural c	onstituent
FAQs						Biological Process	Cell growth and/or maintenance	9
Download Download						FBG		
Pathways PhosphoMolif		ALTERNATE NAMES	DISEASES	PTMs & SUBSTRATES	EXTERNAL LINKS			
		General						
Become a "Molecule Authority"			FGL2	Molecular Weight (Da) :	50228	Gene Map Locu	s: 7q11.23	
		Localization						
		Primary	Extracellular	60		Alternate		
1.2		Domains and Motifs	-			Expression		
Aller Aller			mains	Mo	tifs	Lymphocyte	Site of Expression	
		FBG 208 - 435		<u>SP</u> 1 - 15 <u>CC</u> 73 - 165		Ovary		
						Small intestine		
						Human Proteinpedia		
							Site of Expression	
						Normal Tissue	Disease Tissue	Cell Line

isease Ti	ime Fibrinogen like 2 (FGL2) issue Liver cancer					
	Side Liver cancer					
Plat						
	form: Immunohistochemistry					
				tal Details - HuPA_0002	20	
	ment type		istochemistry			
	lescription		llular carcinoma analysi			
	mental description			as performed using norma	al and cancer liver tissues with antibody sp	ecific for the respective proteins.
Journa	ed/Unpublished	Unpublish Not applic				
PubMe		Not applic				
		Tissue:				
Sample	source		tissue: Liver cancer			
Source	organism		iens [Taxonomy:9606	a		
	- gampin		in the second second second			
Fibrin	ogen like 2 (FGL2) is expressed in Li	iver cancer				
	Protein Name (Gene Symbol)	Tissue Name	Species of	Antibody type	Antibody source	Primary antibody dilution
	Protein Name (Gene Symbol)	lissue name	primary antibody	Antibody type	Antibody source	Primary andbody dilute
					Santa Cruz Biotech, #sc-30869	1:100
1. Showin	Fibrinogen like 2 (FGL2) g 1 to 3 of 3 images	Liver cancer	Goat	Polyclonal		
			Goat	Polycional		
Showin		Liver cancer	Goat	Potycional		
Showin	g 1 to 3 of 3 images			Potycional		
Showin	g 1 to 3 of 3 images	Liver cancer	Experiment			
Plat Experies Short of	e 1 to 3 of 3 images form: Western blot detection ment type lescription	Western I Liver can	Experiment bioting cer analysis	tal Details - HuPA_0063	4	
Plat Experin Short of Experin	g 1 to 3 of 3 images	Western I Liver can Western I	Experiment kotting ser analysis kotting was performed	tal Details - HuPA_0063		respective proteins.
Showin Plat Experin Short of Experin Publish	e 1 to 3 of 3 images	Western I Liver can Western I Unpublish	Experiment location cer analysis location ded	tal Details - HuPA_0063	4	respective proteins.
Plat Experin Short of Experin Journa	g 1 to 3 of 3 images	Western Liver can Western Urspublis Not appli	Experiment lotting cer analysis lotting vas performed adie	tal Details - HuPA_0063	4	respective proteins.
Plat Experin Short of Experin Publish	g 1 to 3 of 3 images	Western I Liver can Western I Unpublish Not appli	Experiment lotting cer analysis lotting vas performed adie	tal Details - HuPA_0063	4	respective proteins.
Plat Experin Short of Experin Publish Journa Publish	g 1 to 3 of 3 images	Western Liver can Urgutalia Not appli Tissue:	Experiment locality l	tal Details - HuPA_0063	4	respective proteins.
Plat Plat Experint Short of Experint Publist Journa Publist Sample	e 1 to 3 of 3 images	Western 1 Liver can Western 1 Compatibil not applic Tissue Disease	Experiment botting det able able able tissue: Liver cancer	tal Details - HuPA_DOG3 using normal and cancer	4	respective proteins.
Plat Plat Experint Short of Experint Publist Journa Publist Sample	e 1 to 3 of 3 images	Western 1 Liver can Western 1 Compatibil not applic Tissue Disease	Experiment locality l	tal Details - HuPA_DOG3 using normal and cancer	4	respective proteins.
Plat Experin Short of Experin Publish Journa PubMe Sample Source	e 1 to 3 of 3 images	Western Western Urpublish Not appli Not applish Sicease Homo sap	Experiment botting det able able able tissue: Liver cancer	tal Details - HuPA_DOG3 using normal and cancer	4	respective proteins.
Plat Experin Short of Experin Publish Journa Pubbe Sample Source	e 1 to 3 of 3 images	Western I Liver can Western I Viewstern Viewstern Disease Hono sap	Experiment lotting cer analysis lobiting was performed ad able lable tissue: Liver cancer less [Taxonomy:5000	nal Details - HuPA, D063 using normal and cancer	16	
Plat Experin Short of Experin Publish Journa PubMe Sample Source	e 1 to 3 of 3 images	Western Western Urpublish Not appli Not applish Sicease Homo sap	Experiment botting det able able able tissue: Liver cancer	tal Details - HuPA_DOG3 using normal and cancer	4	respective proteins.

Fig. 4. FGL2 expression in hepatocellular carcinoma. The molecule page of FGL2, a secreted protein, is shown. Based on unpublished data submitted to Human Proteinpedia, there are two entries based on two different experimental platforms showing that it is expressed in HCC². Immunohistochemical staining shows that it is expressed in HCC; this is accompanied by information about the antibody used. The second entry shows that it is overexpressed based on Western blot analysis.

² R. Chaerkady and A. Pandey, unpublished data.

their data to any of the public repositories. In addition, the peer-reviewed journals should actively encourage the authors to submit their data to such proteomic repositories as proposed recently by Nature Biotechnology (40) and Nature Methods (41). Human Proteinpedia allows referees of submitted manuscripts to access the data anonymously if the authors have submitted the data prior to publication for this purpose.

To capture the proteomic data that has already been generated, our team at the Institute of Bioinformatics is scanning through the published issues to date in all of the major proteomic journals including Molecular & Cellular Proteomics, Proteomics, and Journal of Proteome Research for possible inclusion in Human Proteinpedia. The corresponding authors of the relevant articles are being contacted and requested to contribute the data. Those who volunteer work with the team so that the data submission is as simple and painless as possible for the contributor. In addition, the team will obtain data that is not present in Human Proteinpedia from other public proteomic repositories on a regular basis and integrate them with the existing information.

Cancer Genome Anatomy Project (42) aims to catalog the gene expression profiles of normal, precancer, and cancer tissue samples. The goal of this initiative is to improve detection, diagnosis, and treatment of patients through worldwide collaboration. While this project is mainly targeted toward genomic and transcriptomic analysis, future plans that include analyses of cancer proteomes are almost certain. Genomic analysis alone cannot predict the various proteomic alterations in cancers and a better understanding of these alterations will impact detection, diagnosis, and treatment. With additional initiatives being announced to dissect various aspects of the human proteome, including a recent one by HUPO, the need for a portal that allows effective sharing of data effectively among scientists is almost a prerequisite. We anticipate that Human Proteinpedia will be one such portal.

The day when biologists will have a single integrated portal to view data from genomics, transcriptomics, and proteomics data might not be too far off. An initial step to unify the human proteomic data has been taken with the development of Human Proteinpedia. However, this would not have been possible without the enthusiastic participation of the proteomics community. We hope that investigators will continue to share their data to maintain the momentum and anticipate that more and more laboratories will join. Future goals include the addition of protein structure information, and efforts are already on to allow users to view proteomic information submitted to Human Proteinpedia at the genomic level by mapping the peptides onto the genome. We anticipate that the availability of such data will spur the development of additional "omics" tools and newer bioinformatics approaches for harvesting the information provided by the datasets.

* This work was supported, in whole or in part, by National Institutes of Health Grant U54 RR020839 (Roadmap Initiative for Technology Centers for Networks and Pathways). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

|| To whom correspondence should be addressed. E-mail: pandey@jhmi.edu.

REFERENCES

- Pandey, A., and Mann, M. (2000) Proteomics to study genes and genomes. Nature 405, 837–846
- Nakamura, K., Aebersold, R., Bairoch, A., Dunn, M., Celis, J., Hanash, S., Hochstrasser, D., Humphrey-Smith, I., James, P., Klose, J., LaBaer, J., Langen, H., Mann, M., Parekh, R., Patterson, S., Pearce, C., Poepstorff, P., Simpson, R. J., Tomlinson, I., Tsugita, A., and Yates, J. (2004) From genome to proteome-aim of human proteomics. *Seikagaku* 76, 1271–1274
- Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. Nature 422, 198–207
- Cravatt, B. F., Simon, G. M., and Yates, J. R., 3rd (2007) The biological impact of mass spectrometry-based proteomics. *Nature* 450, 991–1000
- Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* **19**, 946–951
- Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **73**, 2836–2842
- Everley, P. A., Krijgsveld, J., Zetter, B. R., and Gygi, S. P. (2004) Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell. Proteomics* 3, 729–735
- Gronborg, M., Bunkenborg, J., Kristiansen, T. Z., Jensen, O. N., Yeo, C. J., Hruban, R. H., Maitra, A., Goggins, M. G., and Pandey, A. (2004) Comprehensive proteomic analysis of human pancreatic juice. *J. Proteome Res.* 3, 1042–1055
- Gagne, J. P., Ethier, C., Gagne, P., Mercier, G., Bonicalzi, M. E., Mes-Masson, A. M., Droit, A., Winstall, E., Isabelle, M., and Poirier, G. G. (2007) Comparative proteome analysis of human epithelial ovarian cancer. *Proteome Sci.* 5, 16
- Crnogorac-Jurcevic, T., Gangeswaran, R., Bhakta, V., Capurso, G., Lattimore, S., Akada, M., Sunamura, M., Prime, W., Campbell, F., Brentnall, T. A., Costello, E., Neoptolemos, J., and Lemoine, N. R. (2005) Proteomic analysis of chronic pancreatitis and pancreatic adenocarcinoma. *Gastroenterology* **129**, 1454–1463
- 11. Cuzick, J. (1999) Screening for cancer: future potential. *Eur. J. Cancer* **35**, 685–692
- Kocher, T., and Superti-Furga, G. (2007) Mass spectrometry-based functional proteomics: from molecular machines to protein networks. *Nat. Methods* 4, 807–815
- Muller, U., Ernst, G., Melle, C., Guthke, R., and von Eggeling, F. (2006) Convergence of the proteomic pattern in cancer. *Bioinformatics* 22, 1293–1296
- Simon, R., and Sauter, G. (2003) Tissue microarray (TMA) applications: implications for molecular medicine. *Expert Rev. Mol. Med.* 5, 1–12
- Haab, B. B., Paulovich, A. G., Anderson, N. L., Clark, A. M., Downing, G. J., Hermjakob, H., Labaer, J., and Uhlen, M. (2006) A reagent resource to identify proteins and peptides of interest for the cancer community: a workshop report. *Mol. Cell. Proteomics* 5, 1996–2007
- Uhlen, M. (2007) Mapping the human proteome using antibodies. *Mol. Cell.* Proteomics 6, 1455–1456
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2008) *GenBank. Nucleic Acids Res.* 36, D25–D30
- Cochrane, G., Akhtar, R., Aldebert, P., Althorpe, N., Baldwin, A., Bates, K., Bhattacharyya, S., Bonfield, J., Bower, L., Browne, P., Castro, M., Cox, T., Demiralp, F., Eberhardt, R., Faruque, N., Hoad, G., Jang, M., Kulikova, T., Labarga, A., Leinonen, R., Leonard, S., Lin, Q., Lopez, R., Lorenc, D., McWilliam, H., Mukherjee, G., Nardone, F., Plaister, S., Robinson, S., Sobhany, S., Vaughan, R., Wu, D., Zhu, W., Apweiler, R., Hubbard, T., and Birney, E. (2008) Priorities for nucleotide trace, sequence and annotation data capture at the Ensembl Trace Archive and the EMBL nucleotide sequence database. *Nucleic Acids Res.* 36, D5–D12

- Sugawara, H., Ogasawara, O., Okubo, K., Gojobori, T., and Tateno, Y. (2008) DDBJ with new system and face. *Nucleic Acids Res.* 36, D22–D24
- Bradshaw, R. A., Burlingame, A. L., Carr, S., and Aebersold, R. (2006) Reporting protein identification data: the next generation of guidelines. *Mol. Cell. Proteomics* 5, 787–788
- Santos, C., Blake, J., and States, D. J. (2005) Supplementary data need to be kept in public repositories. *Nature* 438, 738
- Cannata, N., Schroder, M., Marangoni, R., and Romano, P. (2008) A Semantic Web for bioinformatics: goals, tools, systems, applications. *BMC Bioinformatics* 9, suppl. 4, S1
- Marenco, L., Nadkarni, P., Martone, M., and Gupta, A. (2007) Interoperability across neuroscience databases. *Methods Mol. Biol.* 401, 23–36
- Kelso, J., Visagie, J., Theiler, G., Christoffels, A., Bardien, S., Smedley, D., Otgaar, D., Greyling, G., Jongeneel, C. V., McCarthy, M. I., Hide, T., and Hide, W. (2003) eVOC: a controlled vocabulary for unifying gene expression data. *Genome Res.* **13**, 1222–1230
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29
- Garavelli, J. S. (2003) The RESID Database of Protein Modifications: 2003 developments. *Nucleic Acids Res.* 31, 499–501
- Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Ceol, A., Moore, S., Orchard, S., Sarkans, U., von Mering, C., Roechert, B., Poux, S., Jung, E., Mersch, H., Kersey, P., Lappe, M., Li, Y., Zeng, R., Rana, D., Nikolski, M., Husi, H., Brun, C., Shanker, K., Grant, S. G., Sander, C., Bork, P., Zhu, W., Pandey, A., Brazma, A., Jacq, B., Vidal, M., Sherman, D., Legrain, P., Cesareni, G., Xenarios, I., Eisenberg, D., Steipe, B., Hogue, C., and Apweiler, R. (2004) The HUPO PSI's molecular interaction format–a community standard for the representation of protein interaction data. *Nat. Biotechnol.* 22, 177–183
- Mathivanan, S., Periaswamy, B., Gandhi, T. K., Kandasamy, K., Suresh, S., Mohmood, R., Ramachandra, Y. L., and Pandey, A. (2006) An evaluation of human protein-protein interaction data in the public domain. *BMC Bioinformatics* 7, suppl. 5, S19
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank. *Nucleic Acids Res.* 28, 235–242
- Martens, L., Hermjakob, H., Jones, P., Adamski, M., Taylor, C., States, D., Gevaert, K., Vandekerckhove, J., and Apweiler, R. (2005) PRIDE: the proteomics identifications database. *Proteomics* 5, 3537–3545
- Desiere, F., Deutsch, E. W., Nesvizhskii, A. I., Mallick, P., King, N. L., Eng, J. K., Aderem, A., Boyle, R., Brunner, E., Donohoe, S., Fausto, N., Hafen, E., Hood, L., Katze, M. G., Kennedy, K. A., Kregenow, F., Lee, H., Lin, B., Martin, D., Ranish, J. A., Rawlings, D. J., Samelson, L. E., Shiio, Y., Watts, J. D., Wollscheid, B., Wright, M. E., Yan, W., Yang, L., Yi, E. C., Zhang, H., and Aebersold, R. (2005) Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry. *Genome Biol.* 6, R9
- Peri, S., Navarro, J. D., Amanchy, R., Kristiansen, T. Z., Jonnalagadda, C. K., Surendranath, V., Niranjan, V., Muthusamy, B., Gandhi, T. K., Gronborg, M., Ibarrola, N., Deshpande, N., Shanker, K., Shivashankar, H. N., Rashmi, B. P., Ramya, M. A., Zhao, Z., Chandrika, K. N., Padma, N., Harsha, H. C., Yatish, A. J., Kavitha, M. P., Menezes, M., Choudhury, D. R., Suresh, S., Ghosh, N., Saravana, R., Chandran, S., Krishna, S., Joy, M., Anand, S. K., Madavan, V., Joseph, A., Wong, G. W., Schiemann, W. P., Constantinescu, S. N., Huang, L., Khosravi-Far, R., Steen, H., Tewari, M., Ghaffari, S., Blobe, G. C., Dang, C. V., Garcia, J. G., Pevsner, J., Jensen, O. N., Roepstorff, P., Deshpande, K. S., Chinnaiyan, A. M., Hamosh, A., Chakravarti, A., and Pandey, A. (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* **13**, 2363–2371
- Kerrien, S., Alam-Faruque, Y., Aranda, B., Bancarz, I., Bridge, A., Derow, C., Dimmer, E., Feuermann, M., Friedrichsen, A., Huntley, R., Kohler, C., Khadake, J., Leroy, C., Liban, A., Lieftink, C., Montecchi-Palazzi, L.,

Orchard, S., Risse, J., Robbe, K., Roechert, B., Thorneycroft, D., Zhang, Y., Apweiler, R., and Hermjakob, H. (2007) IntAct-open source resource for molecular interaction data. *Nucleic Acids Res.* **35**, D561–D565

- Chatr-aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L., and Cesareni, G. (2007) MINT: the molecular INTeraction database. *Nucleic Acids Res.* 35, D572–D574
- Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., Breitkreutz, A., and Tyers, M. (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34, D535–D539
- Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., and Eisenberg, D. (2004) The database of interacting proteins: 2004 update. *Nucleic Acids Res.* 32, D449–D451
- Mehrle, A., Rosenfelder, H., Schupp, I., del Val, C., Arlt, D., Hahne, F., Bechtel, S., Simpson, J., Hofmann, O., Hide, W., Glatting, K. H., Huber, W., Pepperkok, R., Poustka, A., and Wiemann, S. (2006) The LIFEdb database in 2006. *Nucleic Acids Res.* 34, D415–D418
- Uhlen, M., Bjorling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergstrom, K., Brumer, H., Cerjan, D., Ekstrom, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., Forsberg, M., Bjorklund, M. G., Gumbel, K., Halimi, A., Hallin, I., Hamsten, C., Hansson, M., Hedhammar, M., Hercules, G., Kampf, C., Larsson, K., Lindskog, M., Lodewyckx, W., Lund, J., Lundeberg, J., Magnusson, K., Malm, E., Nilsson, P., Odling, J., Oksvold, P., Olsson, I., Oster, E., Ottosson, J., Paavilainen, L., Persson, A., Rimini, R., Rockberg, J., Runeson, M., Sivertsson, A., Skollermo, A., Steen, J., Stenvall, M., Sterky, F., Stromberg, S., Sundberg, M., Tegel, H., Tourle, S., Wahlund, E., Walden, A., Wan, J., Wernerus, H., Westberg, J., Wester, K., Wrethagen, U., Xu, L. L., Hober, S., and Ponten, F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4**, 1920–1932
- 39. Mathivanan, S., Ahmed, M., Ahn, N. G., Alexandre, H., Amanchy, R., Andrews, P. C., Bader, J. S., Balgley, B. M., Bantscheff, M., Bennett, K. L., Bjorling, E., Blagoev, B., Bose, R., Brahmachari, S. K., Burlingame, A. S., Bustelo, X. R., Cagney, G., Cantin, G. T., Cardasis, H. L., Celis, J. E., Chaerkady, R., Chu, F., Cole, P. A., Costello, C. E., Cotter, R. J., Crockett, D., DeLany, J. P., De Marzo, A. M., DeSouza, L. V., Deutsch, E. W., Dransfield, E., Drewes, G., Droit, A., Dunn, M. J., Elenitoba-Johnson, K., Ewing, R. M., Van Eyk, J., Faca, V., Falkner, J., Fang, X., Fenselau, C., Figeys, D., Gagne, P., Gelfi, C., Gevaert, K., Gimble, J. M., Gnad, F., Goel, R., Gromov, P., Hanash, S. M., Hancock, W. S., Harsha, H. C., Hart, G., Hays, F., He, F., Hebbar, P., Helsens, K., Hermeking, H., Hide, W., Hjerno, K., Hochstrasser, D. F., Hofmann, O., Horn, D. M., Hruban, R. H., Ibarrola, N., James, P., Jensen, O. N., Jensen, P. H., Jung, P., Kandasamy, K., Kheterpal, I., Kikuno, R. F., Korf, U., Korner, R., Kuster, B., Kwon, M. S., Lee, H. J., Lee, Y. J., Lefevre, M., Lehvaslaiho, M., Lescuyer, P., Levander, F., Lim, M. S., Lobke, C., Loo, J. A., Mann, M., Martens, L., Martinez-Heredia, J., McComb, M., McRedmond, J., Mehrle, A., Menon, R., Miller, C. A., Mischak, H., Mohan, S. S., Mohmood, R., Molina, H., Moran, M. F., Morgan, J. D., Moritz, R., Morzel, M., Muddiman, D. C., Nalli, A., Navarro, J. D., Neubert, T. A., Ohara, O., Oliva, R., Omenn, G. S., Oyama, M., Paik, Y. K., Pennington, K., Pepperkok, R., Periaswamy, B., Petricoin, E. F., Poirier, G. G., Prasad, T. S., Purvine, S. O., Rahiman, B. A., Ramachandran, P., Ramachandra, Y. L., Rice, R. H., Rick, J., Ronnholm, R. H., Salonen, J., Sanchez, J. C., Sayd, T., Seshi, B., Shankari, K., Sheng, S. J., Shetty, V., Shivakumar, K., Simpson, R. J., Sirdeshmukh, R., Siu, K. W., Smith, J. C., Smith, R. D., States, D. J., Sugano, S., Sullivan, M., Superti-Furga, G., Takatalo, M., Thongboonkerd, V., Trinidad, J. C., Uhlen, M., Vandekerckhove, J., Vasilescu, J., Veenstra, T. D., Vidal-Taboada, J. M., Vihinen, M., Wait, R., Wang, X., Wiemann, S., Wu, B., Xu, T., Yates, J. R., Zhong, J., Zhou, M., Zhu, Y., Zurbig, P., and Pandey, A. (2008) Human proteinpedia enables sharing of human protein data. Nat. Biotechnol. 26, 164-167
- 40. Editorial (2007) Democratizing proteomics data. Nat. Biotechnol. 25, 262
- 41. Editorial (2008) Thou shalt share your data. Nat. Methods 5, 209
- Strausberg, R. L., Buetow, K. H., Emmert-Buck, M. R., and Klausner, R. D. (2000) The cancer genome anatomy project: building an annotated gene index. *Trends Genet.* 16, 103–106