Published online 3 December 2013 Nucleic Acids Research, 2014, Vol. 42, Database issue D959–D965
doi:10.1093/nar/gkt1251

Plasma Proteome Database as a resource for proteomics research: 2014 update

Vishalakshi Nanjappa[1,2], Joji Kurian Thomas[1,2], Arivusudar Marimuthu[1], Babylakshmi Muthusamy[1,3], Aneesha Radhakrishnan[1,4], Rakesh Sharma[5], Aafaque Ahmad Khan[1], Lavanya Balakrishnan[1,6], Nandini A. Sahasrabuddhe[1], Satwant Kumar[1], Binit Nitinbhai Jhaveri[7], Kaushal Vinaykumar Sheth[7], Ramesh Kumar Khatana[8], Patrick G. Shaw[9], Srinivas Manda Srikanth[1,3], Premendu P. Mathur[4], Subramanian Shankar[10], Dindagur Nagaraja[11], Rita Christopher[5], Suresh Mathivanan[12], Rajesh Raju[1], Ravi Sirdeshmukh[1], Aditi Chatterjee[1,4], Richard J. Simpson[12], H. C. Harsha[1], Akhilesh Pandey[13,14,15,16,*] and T. S. Keshava Prasad[1,2,3,*]

[1]Institute of Bioinformatics, International Technology Park, Bangalore 560 066, Karnataka, India, [2]Amrita School of Biotechnology, Amrita University, Kollam 690 525, Kerala, India, [3]Centre of Excellence in Bioinformatics, School of Life Sciences, Pondicherry University, Puducherry 605 014, India, [4]Department of Biochemistry and Molecular Biology, Pondicherry University, Puducherry 605014, India, [5]Department of Neurochemistry, National Institute of Mental Health and Neurosciences, Bangalore 560 022, Karnataka, India, [6]Department of Biotechnology, Kuvempu University, Shankaraghatta 577 451, Karnataka, India, [7]Government Medical College, Bhavnagar 364 001, Gujarat, India, [8]Mahatma Gandhi Institute of Medical Sciences, Sevagram, Wardha 442 012, Maharashtra, India, [9]The Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA, [10]Department of Internal Medicine, Armed Forces Medical College, Pune 411 040, Maharashtra, India, [11]Department of Neurology, National Institute of Mental Health and Neurosciences, Bangalore 560 022, Karnataka, India, [12]Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3084, Australia, [13]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD 21205, USA, [14]Department of Biological Chemistry, Johns Hopkins University, Baltimore, MD 21205, USA, [15]Department of Oncology, Johns Hopkins University, Baltimore, MD 21205, USA and [16]Department of Pathology, Johns Hopkins University, Baltimore, MD 21205, USA

Received August 16, 2013; Revised November 10, 2013; Accepted November 11, 2013

ABSTRACT

Plasma Proteome Database (PPD; http://www.plasmaproteomedatabase.org/) was initially described in the year 2005 as a part of Human Proteome Organization's (HUPO's) pilot initiative on Human Plasma Proteome Project. Since then, improvements in proteomic technologies and increased throughput have led to identification of a large number of novel plasma proteins. To keep up with this increase in data, we have significantly enriched the proteomic information in PPD. This database currently contains information on 10 546 proteins detected in serum/plasma of which 3784 have been reported in two or more studies. The latest version of the database also incorporates mass spectrometry-derived data including experimentally verified proteotypic peptides used for multiple reaction monitoring assays. Other novel features include published plasma/serum concentrations for 1278 proteins along with a separate category of plasma-derived extracellular vesicle proteins. As plasma proteins have become a major

*To whom correspondence should be addressed. Tel: +410 502 6662; Fax: +410 502 7544; Email: pandey@jhmi.edu
Correspondence may also be addressed to T. S. Keshava Prasad. Tel: +080 28416140; Fax: +28416132; Email: keshav@ibioinformatics.org
Present address:
Dindagur Nagaraja, Dharwad Institute of Mental Health and Neuro Sciences (DIMHANS), Dharwad, Karnataka, India.

© The Author(s) 2013. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024

PREOMICS EXHIBIT 1014
Page 1 of 7

thrust in the field of biomarkers, we have enabled a batch-based query designated Plasma Proteome Explorer, which will permit the users in screening a list of proteins or peptides against known plasma proteins to assess novelty of their data set. We believe that PPD will facilitate both clinical and basic research by serving as a comprehensive reference of plasma proteins in humans and accelerate biomarker discovery and translation efforts.

INTRODUCTION

Plasma proteome represents an important subproteome, as it harbors proteins secreted from almost all tissues (1,2). In addition to classical blood proteins, plasma contains proteins secreted by various cells, glands and tissues along with proteins derived from commensal and infectious organisms and parasites residing inside the body (1). The plasma proteome comprises 22 highly abundant proteins including albumin, immunoglobulins, transferrin and haptoglobin, which make up 99% of total protein abundance in plasma. The remaining fraction is composed of proteins of much lower abundance including proteolytically cleaved protein fragments (3). This wide dynamic range of protein abundance, greater than 10 orders of magnitude, makes plasma proteome a challenging proteome to analyze. Plasma proteins also undergo various types of post-translational modifications such as glycosylation, which add to its complexity (4,5). An additional level of heterogeneity is brought about by inter- and intra-individual variation due to age, gender and genetic factors. All these factors make plasma the most complex and diverse body fluid, and its analysis poses challenge for the proteomic community.

Despite the analysis of plasma not being straightforward from an analytical standpoint, it is the most investigated body fluid in clinical diagnostics. Components of plasma including circulating tumor cells, cell-free RNA and DNA, metabolites, electrolytes and proteins are all considered as molecular markers for early detection of diseases, disease monitoring and prognosis (6–9). When compared with other body fluids such as cerebrospinal fluid, gastric juice, bile and synovial fluid, plasma is more readily accessible and requires a simple collection procedure (1,10,11). Thus, detection and quantitation of endogenous or foreign antigens or antibodies directed against these antigens in plasma could be used to determine the physiological and pathological states of an individual (4). Several plasma or serum proteins have already been identified as potential biomarkers of diseases including cardiovascular diseases, autoimmune diseases, infectious diseases and neurological disorders (12–19).

Owing to the importance of plasma proteins, several proteomic efforts have been carried out to explore human plasma proteins. In 2002, Anderson and Anderson compiled immunoassay- and 2D gel electrophoresis-based investigations of plasma proteome and reported the presence of 289 proteins (1). Adkins *et al.* used two different separation techniques followed by mass spectrometry (MS) for characterization of proteins from depleted serum and reported 490 proteins (20). Human Proteome Organization (HUPO) initiated a pilot phase of a major community initiative—the Human Plasma Proteome Project (HPPP)—in 2002 to determine the human plasma or serum protein constituents (21). With the involvement of 35 laboratories across the globe, this led to the identification of 3020 plasma proteins (22). As a part of this initiative, we developed a web-based resource called the Plasma Proteome Database (PPD) (23). Recent incorporation of depletion strategies to remove high-abundance proteins and multiple fractionation strategies coupled to LC-MS/MS approaches and high-resolution MS have resulted in a substantial increase in the number of proteins identified from plasma. For example, a study by Liu *et al.* coupled two different fractionation methods to MS and identified 9087 proteins in plasma, which is the largest data set on plasma proteins reported thus far (24). Farrah *et al.* have analyzed raw MS/MS data from plasma/serum, submitted to the Proteomics Identifications Database (PRIDE) and Human Plasma PeptideAtlas using Trans-Proteomic Pipeline and reported a set of high-confidence 1929 proteins (25–27). We systematically documented these newly described human plasma proteins and made them available for the biomedical community through the updated version of PPD.

RESULTS

Documentation of plasma proteins in PPD

Plasma proteome occupies an important niche at the interface of proteomics, diagnostics and medicine. To enable sharing of data across laboratories involved in HPPP, we developed the PPD as a web-based resource for plasma proteins (23). To keep up with exponential increase in plasma proteome data published in recent times, we systematically documented newly described human plasma proteins and curated information pertaining to them in PPD. For this, as a first step, we carried out a PubMed-based literature by searching for groups that have extensively contributed to plasma or serum proteomics. Additional PubMed searches were carried out to fetch scientific articles by using keywords pertaining to plasma proteins. The information from articles was annotated in PPD by experienced researchers at the Institute of Bioinformatics (IOB) along with clinical investigators from collaborating clinical centers in India. Currently, PPD contains 10 546 proteins linked to 509 scientific articles. However, as experimental data were obtained from diverse experimental platforms and published in the form of peer-reviewed research articles, we did not attempt any *post hoc* determination of any false positives in the database. However, as an alternative, to increase the confidence of identification of proteins in plasma or serum, we have included the number of articles reporting the detection of each protein in plasma or serum. This is listed as 'Total number of studies' under 'Experimental evidence' in the molecule page of respective protein as shown in Figure 1. In this database, the detection of 12 proteins in plasma is supported by >50 publications and

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024

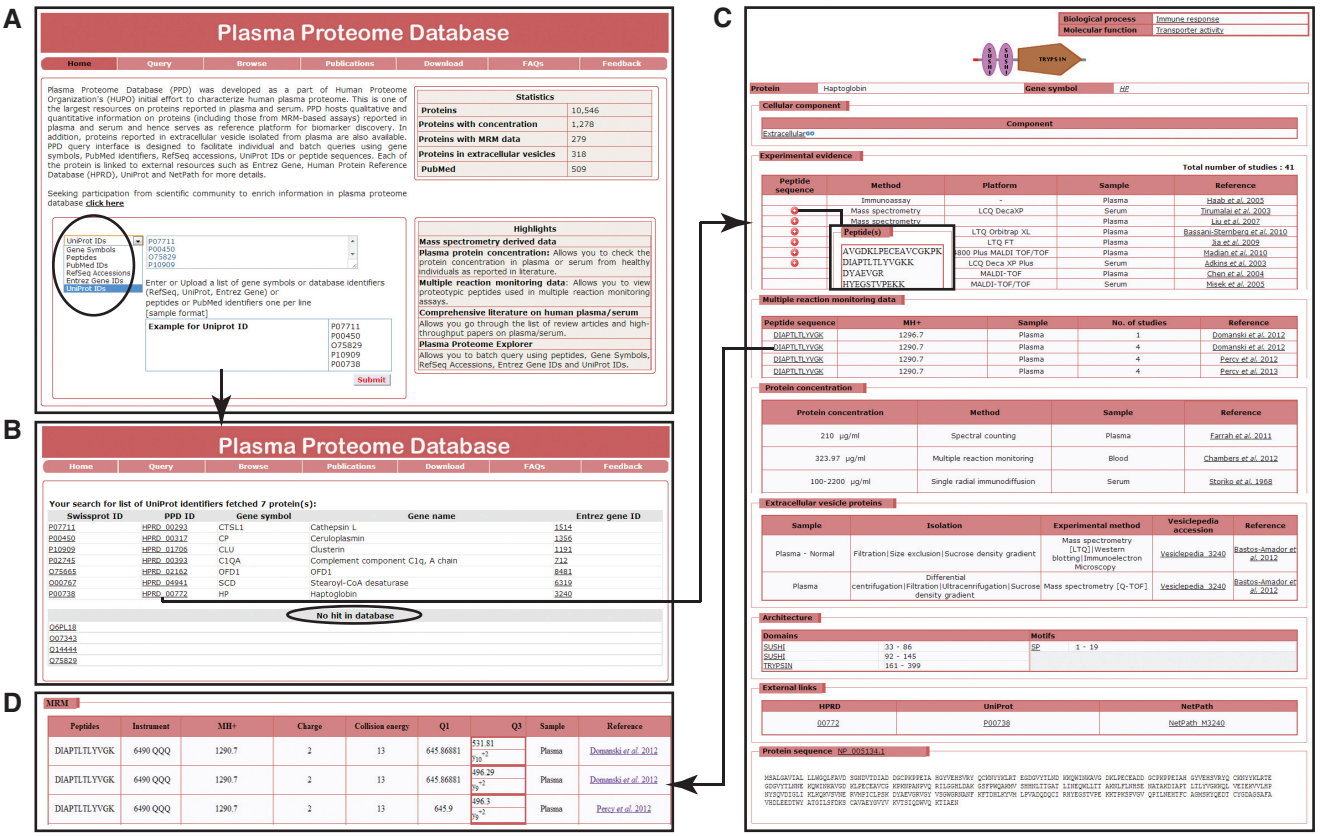

Figure 1. Data analysis workflow using Plasma Proteome Explorer and a screenshot of a 'Molecule Page' for Haptoglobin in PPD. When queried using UniProt IDs (as shown in **A**), the plasma proteome explorer displays two different type of results (shown in **B**). These correspond to (i) proteins present in PPD, which are hyperlinked to their corresponding molecule pages and (ii) proteins not present in PPD, which are linked to an external database, UniProt. Clicking the molecule leads the user to the respective molecule page (shown in **C**). The graphical representation shows domains and motifs found in the protein. The molecule page also displays the alternate localization of protein, the associated biological process and molecular function of the protein. In addition, the plasma concentration reported in healthy individuals along with corresponding PubMed identifiers and any other information regarding presence in plasma EVs is displayed. Further, MRM data is provided with information on proteotypic peptides, peptide m/z values, charge, collision energy, transitions determined, type of fragment ions and the mass spectrometer used along with a link to the corresponding publication (shown in **D**).

that of 167 proteins is supported by >10 publications. Of the remaining 10 367 proteins, 2199 are supported by two publications, whereas 6762 proteins are supported by a single publication. A histogram depicting the distribution of proteins with corresponding number of publications is shown Figure 2. In the current update, we have included peptide data for proteins identified from MS-based studies, which is one of the newly added features of PPD. A comparison of previous and current versions of PPD provided in Table 1 shows a 3-fold increase of protein information in this 2014 update. From our analysis, we observed that 4668 and 1972 proteins were unique to PPD when compared with data from other publicly available data resources on plasma or serum proteins—the Sys-BodyFluid and Healthy Human Individual's Integrated Plasma Proteome (HIP2), respectively (28,29). The data present in the Sys-BodyFluid database were annotated from 15 articles, whereas the HIP2 database contained plasma proteome data cataloged from 4 articles. In contrast to other databases that provide plasma/serum protein catalogs generated from a limited number of studies, PPD includes data annotated from 509 peer-reviewed publications.

In addition to documentation of plasma proteins and peptide information, for each protein in PPD, we have provided external links to Human Protein Reference Database (HPRD) (30), NetPath (31), Entrez Gene (32) and UniProt (33) (Figure 1). For the benefit of users, a collection of important published research articles and reviews describing high-throughput investigations on plasma/serum proteome has been provided. We are seeking participation of the scientific community by inviting investigators to submit relevant articles for inclusion in PPD and have created a new portal in PPD to submit such articles. To facilitate engagement with the community, there is a 'feedback' option, which allows users to submit their comments or suggestions. Finally, we have provided an option to download PPD data in XML and Microsoft Excel formats (http://www.plasmaproteomedatabase.org/download).

Plasma concentration of proteins

An important step in plasma biomarker discovery is the relative quantification of proteins in healthy and disease conditions. Determination of relative and absolute

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024

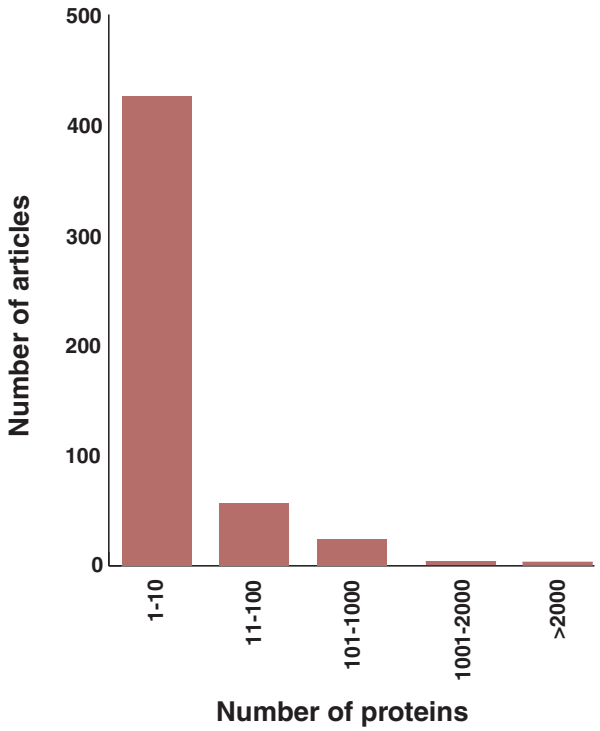

Figure 2. Distribution of proteins annotated in PPD according to number of corresponding publications. The histogram shows that a high number of proteins (8793 proteins) in plasma have been reported only in a single study. Of the 509 articles annotated, 426 were found to support the presence of ≤10 proteins in plasma or serum.

Table 1. A comparison of data in the current version of PPD with the initial database in 2005

Data Features	PPD-2005	PPD-2014	Number of articles annotated in the current update
Total number of proteins	3778	10 546	509
Total number of proteins with plasma levels	Not annotated	1278	276
Total number of proteins derived from plasma extracellular vesicles	Not annotated	318	9
Total number of proteins with MRM data	Not included	279	24

concentration of plasma proteins is helpful for monitoring disease predisposition, onset, diagnosis and progression. Several studies have been carried out in this direction. An early effort was reported by Anderson and Anderson in 2002, where they cataloged plasma concentration range for 70 proteins from published literature (1). Subsequently in 2005 and 2008, their group documented plasma levels for 177 and 211 proteins, respectively, which were reported to be associated with cardiovascular disease, stroke and cancer based on literature evidence (34,35). Recently, Farrah *et al*. provided an estimate for plasma levels for 1200 proteins based on spectral counting (26). We have incorporated plasma protein levels in this update of PPD and documented plasma protein concentrations for 1278 proteins from 276 articles. Availability of this information will be useful for both clinicians and researchers alike. To document plasma protein levels, we searched PubMed using Gene symbols OR Protein name OR alternate name(s) of the protein AND Plasma level* OR Serum level* as queries. We engaged four medical students (S.K., B.N.J., K.V.S. and R.K.K, who are also co-authors) to document the plasma protein levels in normal individuals from the literature. The plasma protein concentration data in PPD can be browsed as shown in Figure 1 along with the method used for the measurement in each case. We also developed a web-based submission portal (http://www.plasmaproteomedatabase.org/protein_concentration) for submission of data pertaining to protein concentration in plasma/serum by the clinical/analytical community. The researchers can submit the information obtained from their study through this portal, and the information will be processed in PPD along with the source of the information. The 'Plasma protein concentration data' file contains a list of plasma proteins sorted based on their concentration in plasma/serum. This file can be downloaded at http://www.plasmaproteomedatabase.org/download.

Incorporation of data from multiple reaction monitoring experiments

Traditionally, enzyme-linked immunosorbent assay and western blotting have been used for protein identification and quantitation in plasma. The success of these conventional methods relies heavily on the availability of specific antibodies against each protein. These methods also provide limited throughput, as each assay can monitor one protein at a time. Recent advances in MS have enabled simultaneous identification and quantitation of hundreds of proteins in plasma in a single experiment (36–38). Targeted approaches like multiple reaction monitoring (MRM) have enabled quantitation of plasma proteins with high specificity as well as increased throughput (36,38).

MRM assays use multiple parameters to identify and quantify specific protein/s in complex samples. Proteotypic peptides that uniquely represent a protein are first selected, and assays are developed, which take into account the retention time of the peptides, their m/z values and measurement of specific fragment ions (39). This multi-step process provides superior specificity and allows for accurate quantification of protein levels over a wide dynamic range of abundance (36,40). MRM analysis has thus revolutionized biomarker research and is being widely used to determine altered protein levels across various disease phenotypes. A compilation of proteotypic peptides of plasma proteins that generate good signals in MS experiments along with their transitions will be immensely beneficial in developing MRM assays. In PPD, we have annotated 591 reported peptides corresponding to 279 proteins for MRM analysis. This includes details of precursor m/z, fragment ion m/z, collision energy, charge

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024

of the precursor ion and instrument used for the analysis (Figure 1).

Proteins identified from extracellular vesicles in plasma

Extracellular vesicles (EVs) are membranous sacs released by both normal and diseased cell types and are also reported in body fluids such as blood, urine, saliva and synovial fluid (41–44). EVs include ectosomes, exosomes and apoptotic bodies. They facilitate communication between cells by transfer of membrane and cytosolic proteins and are also shown to play role in mediating immune response and antigen presentation (44). EVs are shown to be present in plasma from healthy and certain cancer patients including glioblastoma and colon cancer (45,46). Therefore, identification and characterization of EVs in plasma should aid in disease diagnosis and discovery of biomarkers. Considering the importance of EVs, we have curated and included plasma EV proteins in PPD. There are 318 EV proteins annotated in PPD and each is linked to Exocarta and Vesiclepedia, both manually curated resources for proteins, messenger RNA, micro RNA and lipids reported from EVs including exosomes (42,43).

Plasma Proteome Explorer–a batch query interface in PPD

We have enabled a batch query utility called 'Plasma Proteome Explorer', which allows users to query the plasma proteome by providing a list of peptides, gene symbols, Entrez Gene IDs, RefSeq accessions or UniProt IDs. Data analysis workflow of plasma proteome explorer is represented in Figure 1. The results obtained are displayed in a web page with following features—PPD ID, which are hyperlinked to PPD 'molecule page', gene symbol and Entrez Gene ID.

Plasma proteome explorer will enable biologists with limited bioinformatics skills to compare their own experimental data set with that of plasma proteome. In a single step, users can differentiate between novel plasma proteins, novel MS-derived peptides detected from plasma, known plasma proteins and peptides. This will enable researchers in biomarker development to screen candidate molecules that can be detected in plasma, from a larger set of candidate protein biomarkers identified in the discovery phase to pursue further for validation. We believe that the plasma proteome explorer will occupy the intermediate niche between 'biomarker discovery' and 'biomarker validation' phases. We have enabled a separate query tab, which allows querying of proteins based on protein name, gene symbol, Entrez Gene ID, UniProt ID, sample type, experimental method and MS-based platform. In addition, search by protein sequence can be carried out in PPD. The 'Browse' option allows the user to pursue proteins categorized on the basis of their protein name, biological process, molecular function, proteins with information on MRM data and plasma exosomal proteins.

CONCLUSIONS

The main goal of the PPD is to foster research in the area of clinical proteomics, especially the discovery of plasma biomarkers, to diagnose and monitor human diseases. We have introduced a number of new features to expand the utility of this database, particularly in the field of biomarker discovery and validation. Plasma proteome explorer will prove useful for selecting secreted candidate biomarkers from a larger set of differentially expressed genes/proteins in any disease or perturbed biological conditions. As the database develops, we foresee the need for the incorporation of information on disease association and availability of reagents such as enzyme-linked immunosorbent assay kits. We anticipate that PPD will provide a platform to bring together the communities of biomedical researchers engaged in biomarker discovery and validation along with proteomic researchers developing analytical solutions to overcome problems in plasma proteome research. We foresee that PPD will serve as a centralized reference repository for such a community approach.

ACKNOWLEDGEMENTS

The authors thank the Department of Biotechnology (DBT), Government of India, for research support to the Institute of Bioinformatics, Bangalore. B.M. and A.R. are recipients of Senior Research Fellowship from the Council of Scientific and Industrial Research (CSIR), New Delhi, India. R.S is a Research Associate supported by DBT. S.M.S. is a recipient of Senior Research Fellowship from University Grants Commission (UGC), India. H.C.H. is a Wellcome Trust/DBT India Alliance Early Career Fellow. T.S.K.P. is a recipient of the research grant on 'Development of Infrastructure and a Computational Framework for Analysis of Proteomic Data' from DBT. A.P. is supported by a contract HHSN268201000032C from the National Heart Lung and Blood Institute and an NIH Roadmap grant 'Technology Center for Networks and Pathways' [U54GM103520].

FUNDING

Funding for open access charge: Waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Anderson,N.L. and Anderson,N.G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell Proteomics*, **1**, 845–867.
2. Zhang,H., Liu,A.Y., Loriaux,P., Wollscheid,B., Zhou,Y., Watts,J.D. and Aebersold,R. (2007) Mass spectrometric detection of tissue proteins in plasma. *Mol. Cell Proteomics*, **6**, 64–71.
3. Tirumalai,R.S., Chan,K.C., Prieto,D.A., Issaq,H.J., Conrads,T.P. and Veenstra,T.D. (2003) Characterization of the low molecular weight human serum proteome. *Mol. Cell Proteomics*, **2**, 1096–1103.

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024

4. Ping,P., Vondriska,T.M., Creighton,C.J., Gandhi,T.K., Yang,Z., Menon,R., Kwon,M.S., Cho,S.Y., Drwal,G., Kellmann,M. *et al.* (2005) A functional annotation of subproteomes in human plasma. *Proteomics*, **5**, 3506–3519.
5. Yang,Z., Hancock,W.S., Chew,T.R. and Bonilla,L. (2005) A study of glycoproteins in human serum and plasma reference standards (HUPO) using multilectin affinity chromatography coupled with RPLC-MS/MS. *Proteomics*, **5**, 3353–3366.
6. Cristofanilli,M., Budd,G.T., Ellis,M.J., Stopeck,A., Matera,J., Miller,M.C., Reuben,J.M., Doyle,G.V., Allard,W.J., Terstappen,L.W. *et al.* (2004) Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.*, **351**, 781–791.
7. Hanash,S.M., Pitteri,S.J. and Faca,V.M. (2008) Mining the plasma proteome for cancer biomarkers. *Nature*, **452**, 571–579.
8. Kopreski,M.S., Benko,F.A., Kwak,L.W. and Gocke,C.D. (1999) Detection of tumor messenger RNA in the serum of patients with malignant melanoma. *Clin. Cancer Res.*, **5**, 1961–1965.
9. Wang,B.G., Huang,H.Y., Chen,Y.C., Bristow,R.E., Kassauei,K., Cheng,C.C., Roden,R., Sokoll,L.J., Chan,D.W. and Shih Ie,M. (2003) Increased plasma DNA integrity in cancer patients. *Cancer Res.*, **63**, 3966–3968.
10. Lathrop,J.T., Anderson,N.L., Anderson,N.G. and Hammond,D.J. (2003) Therapeutic potential of the plasma proteome. *Curr. Opin. Mol. Ther.*, **5**, 250–257.
11. Rosenblatt,K.P., Bryant-Greenwood,P., Killian,J.K., Mehta,A., Geho,D., Espina,V., Petricoin,E.F. III and Liotta,L.A. (2004) Serum proteomics in cancer diagnosis and management. *Annu. Rev. Med.*, **55**, 97–112.
12. Agranoff,D., Fernandez-Reyes,D., Papadopoulos,M.C., Rojas,S.A., Herbster,M., Loosemore,A., Tarelli,E., Sheldon,J., Schwenk,A., Pollok,R. *et al.* (2006) Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. *Lancet*, **368**, 1012–1021.
13. Bauer,J.W., Baechler,E.C., Petri,M., Batliwalla,F.M., Crawford,D., Ortmann,W.A., Espe,K.J., Li,W., Patel,D.D., Gregersen,P.K. *et al.* (2006) Elevated serum levels of interferon-regulated chemokines are biomarkers for active human systemic lupus erythematosus. *PLoS Med.*, **3**, e491.
14. Berhane,B.T., Zong,C., Liem,D.A., Huang,A., Le,S., Edmondson,R.D., Jones,R.C., Qiao,X., Whitelegge,J.P., Ping,P. *et al.* (2005) Cardiovascular-related proteins identified in human plasma by the HUPO plasma proteome project pilot phase. *Proteomics*, **5**, 3520–3530.
15. He,Q.Y., Lau,G.K., Zhou,Y., Yuen,S.T., Lin,M.C., Kung,H.F. and Chiu,J.F. (2003) Serum biomarkers of hepatitis B virus infected liver inflammation: a proteomic study. *Proteomics*, **3**, 666–674.
16. Kim,H.J., Cho,E.H., Yoo,J.H., Kim,P.K., Shin,J.S., Kim,M.R. and Kim,C.W. (2007) Proteome analysis of serum from type 2 diabetics with nephropathy. *J. Proteome Res.*, **6**, 735–743.
17. Liao,H., Wu,J., Kuhn,E., Chin,W., Chang,B., Jones,M.D., O'Neil,S., Clauser,K.R., Karl,J., Hasler,F. *et al.* (2004) Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. *Arthritis Rheum.*, **50**, 3792–3803.
18. Sawai,S., Umemura,H., Mori,M., Satoh,M., Hayakawa,S., Kodera,Y., Tomonaga,T., Kuwabara,S. and Nomura,F. (2010) Serum levels of complement C4 fragments correlate with disease activity in multiple sclerosis: proteomic analysis. *J. Neuroimmunol.*, **218**, 112–115.
19. Zimmermann-Ivol,C.G., Burkhard,P.R., Le Floch-Rohr,J., Allard,L., Hochstrasser,D.F. and Sanchez,J.C. (2004) Fatty acid binding protein as a serum marker for the early diagnosis of stroke: a pilot study. *Mol. Cell Proteomics*, **3**, 66–72.
20. Adkins,J.N., Varnum,S.M., Auberry,K.J., Moore,R.J., Angell,N.H., Smith,R.D., Springer,D.L. and Pounds,J.G. (2002) Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell Proteomics*, **1**, 947–955.
21. Omenn,G.S. (2004) The Human Proteome Organization Plasma Proteome Project pilot phase: reference specimens, technology platform comparisons, and standardized data submissions and analyses. *Proteomics*, **4**, 1235–1240.
22. Omenn,G.S., States,D.J., Adamski,M., Blackwell,T.W., Menon,R., Hermjakob,H., Apweiler,R., Haab,B.B., Simpson,R.J., Eddes,J.S. *et al.* (2005) Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics*, **5**, 3226–3245.
23. Muthusamy,B., Hanumanthu,G., Suresh,S., Rekha,B., Srinivas,D., Karthick,L., Vrushabendra,B.M., Sharma,S., Mishra,G., Chatterjee,P. *et al.* (2005) Plasma Proteome Database as a resource for proteomics research. *Proteomics*, **5**, 3531–3536.
24. Liu,X., Valentine,S.J., Plasencia,M.D., Trimpin,S., Naylor,S. and Clemmer,D.E. (2007) Mapping the human plasma proteome by SCX-LC-IMS-MS. *J. Am. Soc. Mass Spectrom.*, **18**, 1249–1264.
25. Farrah,T., Deutsch,E.W. and Aebersold,R. (2011) Using the human plasma PeptideAtlas to study human plasma proteins. *Methods Mol. Biol.*, **728**, 349–374.
26. Farrah,T., Deutsch,E.W., Omenn,G.S., Campbell,D.S., Sun,Z., Bletz,J.A., Mallick,P., Katz,J.E., Malmstrom,J., Ossola,R. *et al.* (2011) A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Mol. Cell Proteomics*, **10**, M110 006353.
27. Vizcaino,J.A., Cote,R.G., Csordas,A., Dianes,J.A., Fabregat,A., Foster,J.M., Griss,J., Alpi,E., Birim,M., Contell,J. *et al.* (2013) The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.*, **41**, D1063–D1069.
28. Li,S.J., Peng,M., Li,H., Liu,B.S., Wang,C., Wu,J.R., Li,Y.X. and Zeng,R. (2009) Sys-BodyFluid: a systematical database for human body fluid proteome research. *Nucleic Acids Res.*, **37**, D907–D912.
29. Saha,S., Harrison,S.H., Shen,C., Tang,H., Radivojac,P., Arnold,R.J., Zhang,X. and Chen,J.Y. (2008) HIP2: an online database of human plasma proteins from healthy individuals. *BMC Med. Genomics*, **1**, 12.
30. Prasad,T.S.K., Goel,R., Kandasamy,K., Keerthikumar,S., Kumar,S., Mathivanan,S., Telikicherla,D., Raju,R., Shafreen,B., Venugopal,A. *et al.* (2009) Human protein reference database—2009 update. *Nucleic Acids Res.*, **37**, D767–D772.
31. Kandasamy,K., Mohan,S.S., Raju,R., Keerthikumar,S., Kumar,G.S., Venugopal,A.K., Telikicherla,D., Navarro,J.D., Mathivanan,S., Pecquet,C. *et al.* (2010) NetPath: a public resource of curated signal transduction pathways. *Genome Biol.*, **11**, R3.
32. Maglott,D., Ostell,J., Pruitt,K.D. and Tatusova,T. (2011) Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.*, **39**, D52–D57.
33. UniProt Consortium. (2013) Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.*, **41**, D43–D47.
34. Anderson,L. (2005) Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J. Physiol.*, **563**, 23–60.
35. Polanski,M. and Anderson,N.L. (2007) A list of candidate cancer biomarkers for targeted proteomics. *Biomark. Insights*, **1**, 1–48.
36. Domanski,D., Percy,A.J., Yang,J., Chambers,A.G., Hill,J.S., Freue,G.V. and Borchers,C.H. (2012) MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. *Proteomics*, **12**, 1222–1243.
37. Huttenhain,R., Soste,M., Selevsek,N., Rost,H., Sethi,A., Carapito,C., Farrah,T., Deutsch,E.W., Kusebauch,U., Moritz,R.L. *et al.* (2012) Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Sci. Transl. Med.*, **4**, 142ra194.
38. Percy,A.J., Chambers,A.G., Yang,J. and Borchers,C.H. (2013) Multiplexed MRM-based quantitation of candidate cancer biomarker proteins in undepleted and non-enriched human plasma. *Proteomics*, **13**, 2202–2215.
39. Lange,V., Picotti,P., Domon,B. and Aebersold,R. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.*, **4**, 222.
40. Stahl-Zeng,J., Lange,V., Ossola,R., Eckhardt,K., Krek,W., Aebersold,R. and Domon,B. (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol. Cell Proteomics*, **6**, 1809–1817.

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024

41. Caby,M.P., Lankar,D., Vincendeau-Scherrer,C., Raposo,G. and Bonnerot,C. (2005) Exosomal-like vesicles are present in human blood plasma. *Int. Immunol.*, **17**, 879–887.
42. Kalra,H., Simpson,R.J., Ji,H., Aikawa,E., Altevogt,P., Askenase,P., Bond,V.C., Borras,F.E., Breakefield,X., Budnik,V. *et al.* (2012) Vesiclepedia: a compendium for extracellular vesicles with continuous community annotation. *PLoS Biol.*, **10**, e1001450.
43. Mathivanan,S., Fahner,C.J., Reid,G.E. and Simpson,R.J. (2012) ExoCarta 2012: database of exosomal proteins, RNA and lipids. *Nucleic Acids Res.*, **40**, D1241–D1244.
44. Simpson,R.J., Lim,J.W., Moritz,R.L. and Mathivanan,S. (2009) Exosomes: proteomic insights and diagnostic potential. *Expert Rev. Proteomics*, **6**, 267–283.
45. Ji,H., Greening,D.W., Barnes,T.W., Lim,J.W., Tauro,B.J., Rai,A., Xu,R., Adda,C., Mathivanan,S., Zhao,W. *et al.* (2013) Proteome profiling of exosomes derived from human primary and metastatic colorectal cancer cells reveal differential expression of key metastatic factors and signal transduction components. *Proteomics*, **13**, 1672–1686.
46. Shao,H., Chung,J., Balaj,L., Charest,A., Bigner,D.D., Carter,B.S., Hochberg,F.H., Breakefield,X.O., Weissleder,R. and Lee,H. (2012) Protein typing of circulating microvesicles allows real-time monitoring of glioblastoma therapy. *Nat. Med.*, **18**, 1835–1840.

Downloaded from https://academic.oup.com/nar/article/42/D1/D959/1064084 by guest on 28 September 2024