

Sys-BodyFluid: a systematical database for human body fluid proteome research

Su-Jun Li, Mao Peng, Hong Li, Bo-Shu Liu, Chuan Wang, Jia-Rui Wu, Yi-Xue Li* and Rong Zeng

Key Laboratory of Systems Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

Received August 13, 2008; Revised September 29, 2008; Accepted October 16, 2008

ABSTRACT

Recently, body fluids have widely become an important target for proteomic research and proteomic study has produced more and more body fluid related protein data. A database is needed to collect and analyze these proteome data. Thus, we developed this web-based body fluid proteome database Sys-BodyFluid. It contains eleven kinds of body fluid proteomes, including plasma/serum, urine, cerebrospinal fluid, saliva, bronchoalveolar lavage fluid, synovial fluid, nipple aspirate fluid, tear fluid, seminal fluid, human milk and amniotic fluid. Over 10 000 proteins are presented in the Sys-BodyFluid. Sys-BodyFluid provides the detailed protein annotations, including protein description, Gene Ontology, domain information, protein sequence and involved pathways. These proteome data can be retrieved by using protein name, protein accession number and sequence similarity. In addition, users can query between these different body fluids to get the different proteins identification information. Sys-BodyFluid database can facilitate the body fluid proteomics and disease proteomics research as a reference database. It is available at <http://www.biosino.org/bodyfluid/>.

INTRODUCTION

In the post-genome era, proteomic technology has rapidly developed to be a powerful platform for the research of human physiology. It can be applied for identifying potential novel biomarkers for prognosis, diagnosis and therapeutics (1,2). And in recent years it is shown that body

fluids have become one of the important targets for proteomics research (3). The body fluids include a wide variety of compositions like plasma/serum, urine, cerebrospinal fluid, saliva, bronchoalveolar lavage fluid, synovial fluid, nipple aspirate fluid, tear fluid, amniotic fluid and so on. Analysis of the protein composition in body fluids can help to understand human disease proteomics better. Hu *et al.*, (3) reviewed the body fluids research advances in proteome analysis and focused on its applications to human disease biomarker discovery. The importance of body fluids has also been appreciated by recent proteomics work (4). The database 'MAPU: Max-Planck Unified database of organellar, cellular, tissue and body fluid' (5) published in 2007 exhibit the close attention of the proteome researchers to the body fluids. The MAPU database stores the data from their own lab and contains several kinds of body fluids, such as urine and tear fluid. To collect more curated proteomics data in the related literatures of the body fluids and provide comprehensive protein annotation, as well as explore the relationships between the different body fluids, we constructed this database Sys-BodyFluid. Abundant proteomics data and in-depth protein annotation make Sys-BodyFluid to be a reference database for body fluid and clinical proteomics research.

DATABASE CONSTRUCTION

Sys-BodyFluid database was implemented through MySQL relational database (<http://www.mysql.com>). The web graphical user interface was constructed using JavaServer Pages technology (<http://java.sun.com/products/jsp/>). The manually curated body fluid protein data in the Sys-BodyFluid were imported to MySQL database by JAVA program. The protein annotation data were downloaded from International Protein Index (IPI) database, Gene Ontology (6), GOA database (7) and

*To whom correspondence should be addressed. Tel: +86 21 54920089; Fax: +86 21 54920143; Email: yxli@sibs.ac.cn
Correspondence may also be addressed to Rong Zeng. Tel: +86 21 54920170; Fax: +86 21 54920171; Email: zr@sibs.ac.cn

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors

© 2008 The Author(s)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.0/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. The data summary in Sys-BodyFluid database

| Body fluid name | Protein number | Paper number |
|--------------------------------------|----------------|--------------|
| Plasma/Serum (11–23) | 7748 | 13 |
| Saliva (24–31) | 2161 | 8 |
| Urine (32–40) | 1941 | 9 |
| Cerebrospinal fluid (41–46) | 1286 | 6 |
| Seminal fluid (47,48) | 916 | 2 |
| Amniotic fluid (49–51) | 899 | 3 |
| Tear (52,53) | 509 | 2 |
| Bronchoalveolar lavage fluid (54,55) | 411 | 2 |
| Milk (56,57) | 175 | 2 |
| Synovial fluid (58) | 114 | 1 |
| Nipple aspiration fluid (59,60) | 84 | 2 |
| Total | 10 138 | 50 |

KEGG (8) pathway database. Open source JAVA library named as JFreeChart (<http://www.jfree.org/jfreechart/>) distributed under LGPL was adopted to plot the image of the statistics data in the web.

DATA SOURCE AND DATABASE CONTENTS

We searched PubMed and manually curated 50 related peer-review publications published online before May 2008. The primary sequences of the proteins were retrieved by the original ID from their corresponding databases in these publications. Due to the database updates, the protein sequences reported in the literatures may have changed or depleted in the current databases. Therefore, these protein sequences were manually validated before importing into the database. Each protein was mapping to the IPI database to uniform the protein ID in Sys-BodyFluid by blasting these protein sequences against the database (Human IPI Version 3.44) (the *E*-value cutoff was set to 10^{-8} , the BLAST-HSP coverage was >0.9). Thus, each of the protein has a corresponding IPI ID in the Sys-BodyFluid database. The total unique proteins and paper numbers of the 11 kinds of body fluids in our database are summarized in Table 1. For example, there are 13 papers and 7748 proteins about the plasma/serum research in our database. Users can obtain this statistical information about the Sys-BodyFluid database in the 'DATABASE' web link in the website <http://www.biosino.org/bodyfluid>.

DATA AVAILABILITY

The Sys-BodyFluid is accessed from graphical web interface (<http://www.biosino.org/bodyfluid/>) and the data are available for download through the 'DOWNLOAD' link in the website as a text file. Users could specify their interested body fluid data to download.

DATABASE UTILITY

Sys-BodyFluid provides users the current database data statistics of different body fluids through the DATABASE link for the paper number and the unique protein number (DATABASE Link). As shown in

Figure 1, Sys-BodyFluid offers users an optimal search function, including searching by protein ID, name and sequence similarity (SEARCH link, Figure 1A). The comprehensive browse option allows users to explore comparison analysis between two or more different body fluids data (Browse link, Figure 1B). For each protein in Sys-BodyFluid, we provide detailed annotation information, including protein description, involved body fluids, paper information, domain, Gene Ontology, pathway, sequence and so on (Figure 1C). Users can choose their interested body fluid to browse or download. Web page describing the body fluid provides users particular information. Furthermore, the availability of pathway analysis will assist users to investigate the difference between body fluids through involved metabolism and signal transduction pathway (Pathway link, Figure 1D). Proteins in our database are labeled with 'red' color. The body fluid number and paper number of the proteins involved in are also showed in the web page.

RESULTS AND DISCUSSION

To get more comprehensive understanding of the relationship between body fluids, we compared the proteins composition in different body fluids. The result is shown in Figure 2A. There are 2928 proteins presented in at least two body fluids and 1359 proteins exist in at least three body fluids. Only 15 proteins exist in total 11 body fluids. For these 2928 proteins, GO annotation information were obtained and enrichment analysis was performed using BiNGO (9) and Cytoscape (10). Each node in Figure 2B represents a GO term. The node's size is scaled by protein number and node's color shows *P*-value of the enrichment analysis. The edge denotes the parent–children relationship between nodes. From this analysis, it is shown that some molecular functions like 'protein binding' and 'enzyme regulator activity' are over-presented in this dataset, as well as the biological process like 'transport' and 'secretion'. Cellular component like 'extracellular region' is significantly enriched.

Human body fluids proteome analysis is still a challenge because dynamic range and the complexity of the body fluids protein composition. It is important to construct a body fluid reference database dedicated to biomarker discovery research. Previous work like MAPU is a great effort to integrate the data from their own lab and aim to provide a 'gold standard' reference proteome database. It is still necessary to refer to other proteomic literature data. For this reason, our database Sys-BodyFluid was build as a complementary database to the MAPU and aimed to provide users more information about the body fluids accompanied by protein abundant annotations. The relationship between different body fluids was also focused in our database. Users can access this database by <http://www.biosino.org/bodyfluid>.

PERSPECTIVES

As more and more body fluid proteome data have been produced recently, it is planned to update Sys-BodyFluid

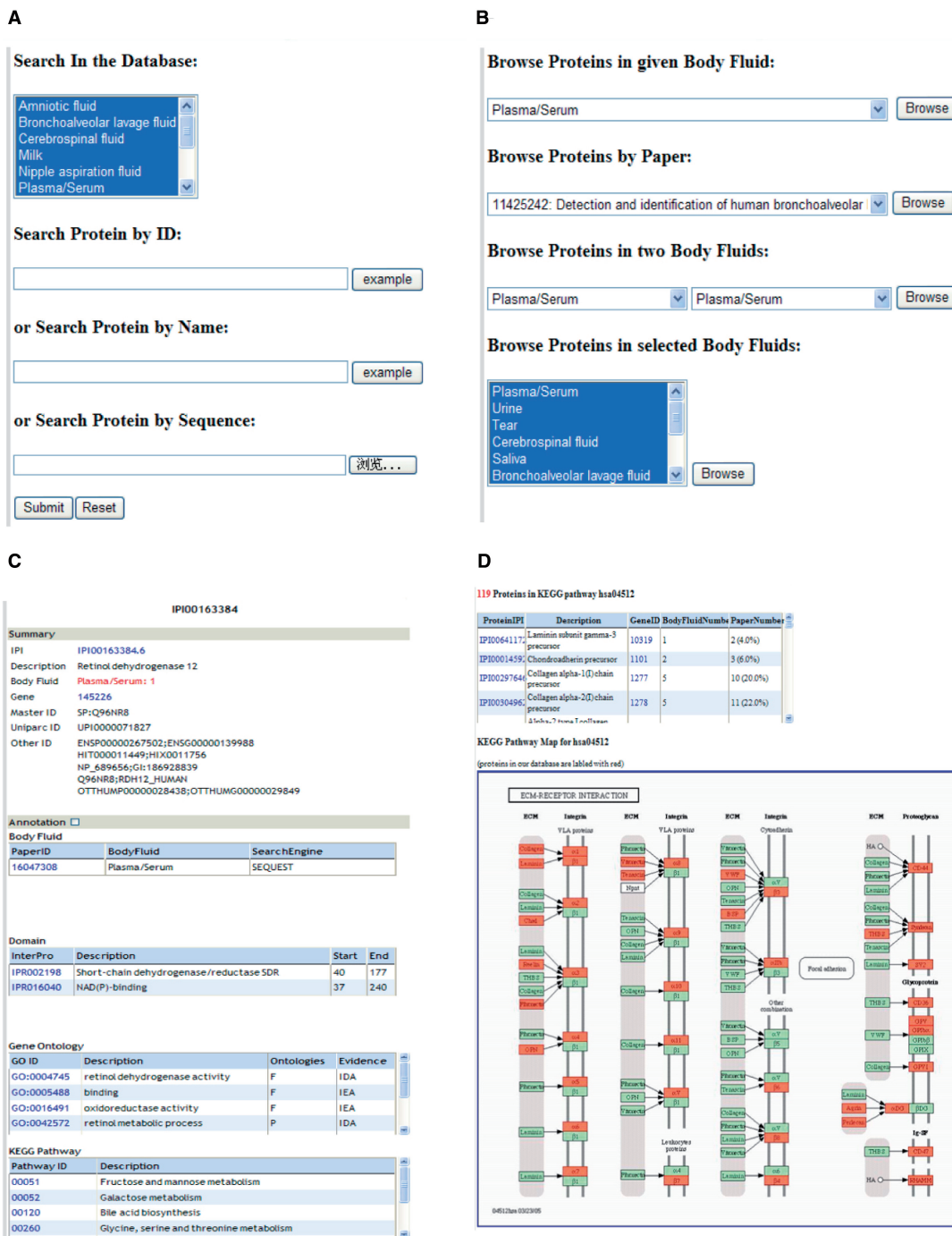


Figure 1. The web graphical user interface of Sys-BodyFluid database. (A) Search part and option. Users could search protein by protein ID, protein name and sequence similarity. (B) Browse part. Database allows user browse protein by their interested body fluid and interested paper. Protein existed in two body fluids could also be viewed and multi body fluids can be investigated. (C) Protein annotation part. There is detailed information in the database for each protein, including description, domain, Gene Ontology term, sequence and so on. (D) Pathway part. The proteins (colored by red) in different body fluids and their involved pathway are shown in pathway link. Proteins in our database are labeled with 'red' color. The body fluid number and paper number are also showed in the web page.

database every 6 months. New body fluid proteome data produced during the time will be added to our database. Furthermore, more annotation information like protein interaction data will also be included. In the future, we will collect more body fluid proteome

data in the disease proteomics research, for example, cancer and diabetes proteome data. If possible, tissue proteomics data will be also included to look into the crosstalk between the tissue protein and the body fluid protein.

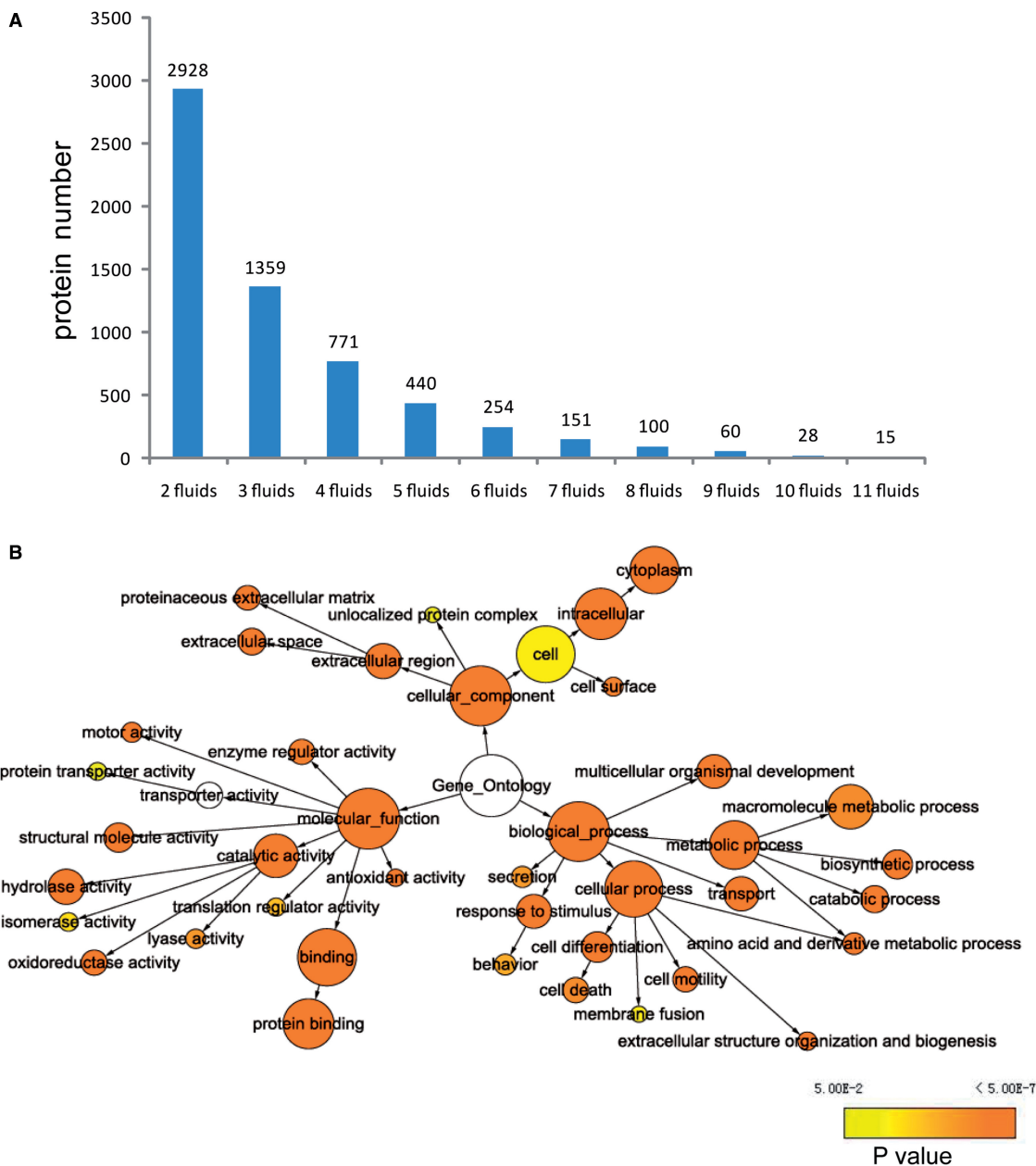


Figure 2. (A) The data comparison in different body fluids. There are 2928 proteins presented in at least two body fluids and 1359 proteins existed in at least three body fluids. Only 15 proteins exist in total 11 body fluids. (B) Gene Ontology annotation statistical analysis for the 2928 proteins existing in at least two body fluids.

FUNDING

Basic Research Foundation (2006CB910700); CAS Project (KSCX2-YW-R-106, KSCX2-YW-R-112, KGCX1-YW-13); High-technology Project (2007AA02Z334). Funding for open access charge: CAS project KSCX2-YW-R-106.

Conflict of interest statement. None declared.

REFERENCES

1. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature*, **422**, 198–207.
2. Binz, P.A., Hochstrasser, D.F. and Appel, R.D. (2003) Mass spectrometry-based proteomics: current status and potential use in clinical chemistry. *Clin. Chem. Lab. Med.*, **41**, 1540–1551.
3. Hu, S., Loo, J.A. and Wong, D.T. (2006) Human body fluid proteome analysis. *Proteomics*, **6**, 6326–6353.
4. Fusaro, V.A. and Stone, J.H. (2003) Mass spectrometry-based proteomics and analyses of serum: a primer for the clinical investigator. *Clin. Exp. Rheumatol.*, **21**, S3–S14.
5. Zhang, Y., Zhang, Y., Adachi, J., Olsen, J.V., Shi, R., de Souza, G., Pasini, E., Foster, L.J., Macek, B., Zougman, A. *et al.* (2007) MAPU: Max-Planck Unified database of organellar, cellular, tissue and body fluid proteomes. *Nucleic Acids Res.*, **35**, D771–D779.
6. 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. *et al.*

- (2000) Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
7. Camon, E., Magrane, M., Barrell, D., Lee, V., Dimmer, E., Maslen, J., Binns, D., Harte, N., Lopez, R. and Apweiler, R. (2004) The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res.*, **32**, D262–D266.
 8. Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. et al. (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res.*, **36**, D480–D484.
 9. Maere, S., Heymans, K. and Kuiper, M. (2005) BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448–3449.
 10. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
 11. Jin, W.H., Dai, J., Li, S.J., Xia, Q.C., Zou, H.F. and Zeng, R. (2005) Human plasma proteome analysis by multidimensional chromatography prefractionation and linear ion trap mass spectrometry identification. *J. Proteome Res.*, **4**, 613–619.
 12. Anderson, N.L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R.S., Conrads, T.P., Veenstra, T.D., Adkins, J.N., Pounds, J.G., Fagan, R. et al. (2004) The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol. Cell. Proteomics*, **3**, 311–326.
 13. Barnea, E., Sorkin, R., Ziv, T., Beer, I. and Admon, A. (2005) Evaluation of prefractionation methods as a preparatory step for multidimensional based chromatography of serum proteins. *Proteomics*, **5**, 3367–3375.
 14. Gong, Y., Li, X., Yang, B., Ying, W., Li, D., Zhang, Y., Dai, S., Cai, Y., Wang, J., He, F. et al. (2006) Different immunoaffinity fractionation strategies to characterize the human plasma proteome. *J. Proteome Res.*, **5**, 1379–1387.
 15. He, P., He, H.Z., Dai, J., Wang, Y., Sheng, Q.H., Zhou, L.P., Zhang, Z.S., Sun, Y.L., Liu, F., Wang, K. et al. (2005) The human plasma proteome: analysis of Chinese serum using shotgun strategy. *Proteomics*, **5**, 3442–3453.
 16. 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.
 17. 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.
 18. Sennels, L., Salek, M., Lomas, L., Boschetti, E., Righetti, P.G. and Rappsilber, J. (2007) Proteomic analysis of human blood serum using peptide library beads. *J. Proteome Res.*, **6**, 4055–4062.
 19. Tanaka, Y., Akiyama, H., Kuroda, T., Jung, G., Tanahashi, K., Sugaya, H., Utsumi, J., Kawasaki, H. and Hirano, H. (2006) A novel approach and protocol for discovering extremely low-abundance proteins in serum. *Proteomics*, **6**, 4845–4855.
 20. 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.
 21. Tu, C.J., Dai, J., Li, S.J., Sheng, Q.H., Deng, W.J., Xia, Q.C. and Zeng, R. (2005) High-sensitivity analysis of human plasma proteome by immobilized isoelectric focusing fractionation coupled to mass spectrometry identification. *J. Proteome Res.*, **4**, 1265–1273.
 22. Valentine, S.J., Plasencia, M.D., Liu, X., Krishnan, M., Naylor, S., Udseth, H.R., Smith, R.D. and Clemmer, D.E. (2006) Toward plasma proteome profiling with ion mobility-mass spectrometry. *J. Proteome Res.*, **5**, 2977–2984.
 23. Zhou, M., Prieto, D.A., Lucas, D.A., Chan, K.C., Issaq, H.J., Veenstra, T.D. and Conrads, T.P. (2006) Identification of the SELDI ProteinChip human serum retentate by microcapillary liquid chromatography-tandem mass spectrometry. *J. Proteome Res.*, **5**, 2207–2216.
 24. Denny, P., Hagen, F.K., Hardt, M., Liao, L., Yan, W., Arellano, M., Bassilian, S., Bedi, G.S., Boontheung, P., Cociorva, D. et al. (2008) The proteomes of human parotid and submandibular/sublingual gland salivas collected as the ductal secretions. *J. Proteome Res.*, **7**, 1994–2006.
 25. Fang, X., Yang, L., Wang, W., Song, T., Lee, C.S., DeVoe, D.L. and Balgley, B.M. (2007) Comparison of electrokinetics-based multidimensional separations coupled with electrospray ionization-tandem mass spectrometry for characterization of human salivary proteins. *Anal. Chem.*, **79**, 5785–5792.
 26. Guo, T., Rudnick, P.A., Wang, W., Lee, C.S., DeVoe, D.L. and Balgley, B.M. (2006) Characterization of the human salivary proteome by capillary isoelectric focusing/nanoreversed-phase liquid chromatography coupled with ESI-tandem MS. *J. Proteome Res.*, **5**, 1469–1478.
 27. Ramachandran, P., Boontheung, P., Xie, Y., Sondej, M., Wong, D.T. and Loo, J.A. (2006) Identification of N-linked glycoproteins in human saliva by glycoprotein capture and mass spectrometry. *J. Proteome Res.*, **5**, 1493–1503.
 28. Vitorino, R., Lobo, M.J., Ferrer-Correia, A.J., Dubin, J.R., Tomer, K.B., Domingues, P.M. and Amado, F.M. (2004) Identification of human whole saliva protein components using proteomics. *Proteomics*, **4**, 1109–1115.
 29. Walz, A., Stuhler, K., Wattenberg, A., Hawranke, E., Meyer, H.E., Schmalz, G., Bluggel, M. and Ruhl, S. (2006) Proteome analysis of glandular parotid and submandibular-sublingual saliva in comparison to whole human saliva by two-dimensional gel electrophoresis. *Proteomics*, **6**, 1631–1639.
 30. Wilmarth, P.A., Riviere, M.A., Rustvold, D.L., Lauten, J.D., Madden, T.E. and David, L.L. (2004) Two-dimensional liquid chromatography study of the human whole saliva proteome. *J. Proteome Res.*, **3**, 1017–1023.
 31. Xie, H., Rhodus, N.L., Griffin, R.J., Carlis, J.V. and Griffin, T.J. (2005) A catalogue of human saliva proteins identified by free flow electrophoresis-based peptide separation and tandem mass spectrometry. *Mol. Cell. Proteomics*, **4**, 1826–1830.
 32. Adachi, J., Kumar, C., Zhang, Y., Olsen, J.V. and Mann, M. (2006) The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol.*, **7**, R80.
 33. Castagna, A., Cecconi, D., Sennels, L., Rappsilber, J., Guerrier, L., Fortis, F., Boschetti, E., Lomas, L. and Righetti, P.G. (2005) Exploring the hidden human urinary proteome via ligand library beads. *J. Proteome Res.*, **4**, 1917–1930.
 34. Khan, A. and Packer, N.H. (2006) Simple urinary sample preparation for proteomic analysis. *J. Proteome Res.*, **5**, 2824–2838.
 35. Oh, J., Pyo, J.H., Jo, E.H., Hwang, S.I., Kang, S.C., Jung, J.H., Park, E.K., Kim, S.Y., Choi, J.Y. and Lim, J. (2004) Establishment of a near-standard two-dimensional human urine proteomic map. *Proteomics*, **4**, 3485–3497.
 36. Pieper, R., Gatlin, C.L., McGrath, A.M., Makusky, A.J., Mondal, M., Seonarin, M., Field, E., Schatz, C.R., Estock, M.A., Ahmed, N. et al. (2004) Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. *Proteomics*, **4**, 1159–1174.
 37. Ru, Q.C., Katenhusen, R.A., Zhu, L.A., Silberman, J., Yang, S., Orchard, T.J., Brzeski, H., Liebman, M. and Ellsworth, D.L. (2006) Proteomic profiling of human urine using multi-dimensional protein identification technology. *J. Chromatogr. A*, **1111**, 166–174.
 38. Spahr, C.S., Davis, M.T., McGinley, M.D., Robinson, J.H., Bures, E.J., Beierle, J., Mort, J., Courchesne, P.L., Chen, K., Wahl, R.C. et al. (2001) Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics*, **1**, 93–107.
 39. Sun, W., Li, F., Wu, S., Wang, X., Zheng, D., Wang, J. and Gao, Y. (2005) Human urine proteome analysis by three separation approaches. *Proteomics*, **5**, 4994–5001.
 40. Zerefos, P.G., Vougas, K., Dimitraki, P., Kossida, S., Petrolekas, A., Stravodimos, K., Giannopoulos, A., Fountoulakis, M. and Vlahou, A. (2006) Characterization of the human urine proteome by preparative electrophoresis in combination with 2-DE. *Proteomics*, **6**, 4346–4355.
 41. Ogata, Y., Charlesworth, M.C., Higgins, L., Keegan, B.M., Vernino, S. and Muddiman, D.C. (2007) Differential protein expression in male and female human lumbar cerebrospinal fluid using iTRAQ

- reagents after abundant protein depletion. *Proteomics*, **7**, 3726–3734.
42. Ogata, Y., Charlesworth, M.C. and Muddiman, D.C. (2005) Evaluation of protein depletion methods for the analysis of total-, phospho- and glycoproteins in lumbar cerebrospinal fluid. *J. Proteome Res.*, **4**, 837–845.
 43. Pan, S., Wang, Y., Quinn, J.F., Peskind, E.R., Waichunas, D., Wimberger, J.T., Jin, J., Li, J.G., Zhu, D., Pan, C. *et al.* (2006) Identification of glycoproteins in human cerebrospinal fluid with a complementary proteomic approach. *J. Proteome Res.*, **5**, 2769–2779.
 44. Wenner, B.R., Lovell, M.A. and Lynn, B.C. (2004) Proteomic analysis of human ventricular cerebrospinal fluid from neurologically normal, elderly subjects using two-dimensional LC-MS/MS. *J. Proteome Res.*, **3**, 97–103.
 45. Zhang, J., Goodlett, D.R., Peskind, E.R., Quinn, J.F., Zhou, Y., Wang, Q., Pan, C., Yi, E., Eng, J., Aebersold, R.H. *et al.* (2005) Quantitative proteomic analysis of age-related changes in human cerebrospinal fluid. *Neurobiol. Aging*, **26**, 207–227.
 46. Zougman, A., Pilch, B., Podtelejnikov, A., Kiehnopf, M., Schnabel, C., Kumar, C. and Mann, M. (2008) Integrated analysis of the cerebrospinal fluid peptidome and proteome. *J. Proteome Res.*, **7**, 386–399.
 47. Fung, K.Y., Glode, L.M., Green, S. and Duncan, M.W. (2004) A comprehensive characterization of the peptide and protein constituents of human seminal fluid. *Prostate*, **61**, 171–181.
 48. Pilch, B. and Mann, M. (2006) Large-scale and high-confidence proteomic analysis of human seminal plasma. *Genome Biol.*, **7**, R40.
 49. Cho, C.K., Shan, S.J., Winsor, E.J. and Diamandis, E.P. (2007) Proteomics analysis of human amniotic fluid. *Mol. Cell. Proteomics*, **6**, 1406–1415.
 50. Michaels, J.E., Dasari, S., Pereira, L., Reddy, A.P., Lapidus, J.A., Lu, X., Jacob, T., Thomas, A., Rodland, M., Roberts, C.T. Jr. *et al.* (2007) Comprehensive proteomic analysis of the human amniotic fluid proteome: gestational age-dependent changes. *J. Proteome Res.*, **6**, 1277–1285.
 51. Nilsson, S., Ramstrom, M., Palmblad, M., Axelsson, O. and Bergquist, J. (2004) Explorative study of the protein composition of amniotic fluid by liquid chromatography electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J. Proteome Res.*, **3**, 884–889.
 52. de Souza, G.A., Godoy, L.M. and Mann, M. (2006) Identification of 491 proteins in the tear fluid proteome reveals a large number of proteases and protease inhibitors. *Genome Biol.*, **7**, R72.
 53. Li, N., Wang, N., Zheng, J., Liu, X.M., Lever, O.W., Erickson, P.M. and Li, L. (2005) Characterization of human tear proteome using multiple proteomic analysis techniques. *J. Proteome Res.*, **4**, 2052–2061.
 54. Plymoth, A., Yang, Z., Lofdahl, C.G., Ekberg-Jansson, A., Dahlback, M., Fehniger, T.E., Marko-Varga, G. and Hancock, W.S. (2006) Rapid proteome analysis of bronchoalveolar lavage samples of lifelong smokers and never-smokers by micro-scale liquid chromatography and mass spectrometry. *Clin. Chem.*, **52**, 671–679.
 55. Sabounchi-Schutt, F., Astrom, J., Eklund, A., Grunewald, J. and Bjellqvist, B. (2001) Detection and identification of human bronchoalveolar lavage proteins using narrow-range immobilized pH gradient DryStrip and the paper bridge sample application method. *Electrophoresis*, **22**, 1851–1860.
 56. Fortunato, D., Giuffrida, M.G., Cavaletto, M., Garoffo, L.P., Dellavalle, G., Napolitano, L., Giunta, C., Fabris, C., Bertino, E., Coscia, A. *et al.* (2003) Structural proteome of human colostrum fat globule membrane proteins. *Proteomics*, **3**, 897–905.
 57. Palmer, D.J., Kelly, V.C., Smit, A.M., Kuy, S., Knight, C.G. and Cooper, G.J. (2006) Human colostrum: identification of minor proteins in the aqueous phase by proteomics. *Proteomics*, **6**, 2208–2216.
 58. Gobezie, R., Kho, A., Krastins, B., Sarracino, D.A., Thornhill, T.S., Chase, M., Millett, P.J. and Lee, D.M. (2007) High abundance synovial fluid proteome: distinct profiles in health and osteoarthritis. *Arthritis Res. Ther.*, **9**, R36.
 59. Alexander, H., Stegner, A.L., Wagner-Mann, C., Du Bois, G.C., Alexander, S. and Sauter, E.R. (2004) Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin. Cancer Res.*, **10**, 7500–7510.
 60. Varnum, S.M., Covington, C.C., Woodbury, R.L., Petritis, K., Kangas, L.J., Abdullah, M.S., Pounds, J.G., Smith, R.D. and Zangar, R.C. (2003) Proteomic characterization of nipple aspirate fluid: identification of potential biomarkers of breast cancer. *Breast Cancer Res. Treat.*, **80**, 87–97.