Database, 2021, **2021(0)**, 1–14
DOI: https://doi.org/10.1093/database/baab065
Original article

HBFP: a new repository for human body fluid proteome

Dan Shao[1,2,3], Lan Huang[1], Yan Wang[1,*], Xueteng Cui[2], Yufei Li[2], Yao Wang[1], Qin Ma[4], Wei Du[1] and Juan Cui[3,*]

[1]Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, 2699 Qianjin Street, Changchun 130012, China
[2]Department of Computer Science and Technology, Changchun University, 6543 Weixing Road, Changchun 130022, China
[3]Department of Computer Science and Engineering, University of Nebraska-Lincoln, 122E Avery Hall, 1144 T St., Lincoln, NE 68588, USA
[4]Department of Biomedical Informatics, College of Medicine, The Ohio State University, 310G Lincoln tower, 1800 cannon drive, Columbus, OH 43210, USA

*Corresponding author: Tel: +86-0431-85168752; Fax: +86-0431-85168752; Email: wy6868@jlu.edu.cn
Correspondence may also be addressed to Juan Cui. Tel: +1-(402) 472-5023; Fax: +1-(402) 472-7767; Email: jcui@unl.edu

Citation details: Shao, D., Huang, L., Wang, Y. *et al.* HBFP: a new repository for human body fluid proteome. *Database* (2021) Vol. 2021: article ID baab065; DOI: https://doi.org/10.1093/database/baab065

Abstract

Body fluid proteome has been intensively studied as a primary source for disease biomarker discovery. Using advanced proteomics technologies, early research success has resulted in increasingly accumulated proteins detected in different body fluids, among which many are promising biomarkers. However, despite a handful of small-scale and specific data resources, current research is clearly lacking effort compiling published body fluid proteins into a centralized and sustainable repository that can provide users with systematic analytic tools. In this study, we developed a new database of human body fluid proteome (HBFP) that focuses on experimentally validated proteome in 17 types of human body fluids. The current database archives 11 827 unique proteins reported by 164 scientific publications, with a maximal false discovery rate of 0.01 on both the peptide and protein levels since 2001, and enables users to query, analyze and download protein entries with respect to each body fluid. Three unique features of this new system include the following: (i) the protein annotation page includes detailed abundance information based on relative qualitative measures of peptides reported in the original references, (ii) a new score is calculated on each reported protein to indicate the discovery confidence and (iii) HBFP catalogs 7354 proteins with at least two non-nested uniquely mapping peptides of nine amino acids according to the Human Proteome Project Data Interpretation Guidelines, while the remaining 4473 proteins have more than two unique peptides without given sequence information. As an important resource for human protein secretome, we anticipate that this new HBFP database can be a powerful tool that facilitates research in clinical proteomics and biomarker discovery.

Database URL: https://bmbl.bmi.osumc.edu/HBFP/

Background

Human body fluids are thought to be rich resources of disease-associated proteins that are secreted or leaked from pathological tissues across the body, many of which are commonly obtainable through non-invasive procedures (1, 2). Driven by these factors, research interests have soared a few decades ago toward biomarker discovery by examining body fluid proteomes. It is highly plausible that empowered by innovative high-throughput technologies, modern proteomic studies have successfully identified a large number of proteins in various body fluids such as plasma, serum, saliva and urine (3).

With great effort by a few large consortiums, several community-based proteomic databases have been developed in the past decades. For example, in 2002, the international Human Proteome Organization initiated the Human Plasma Proteome Project and reported human plasma and serum protein constituents in its online databases (4). Another similar database, named Plasma Proteome Database, archived more than 10 000 proteins detected in human blood (5). Additionally, the Proteomics Identifications database (6) and Human Plasma PeptideAtlas (7) report a total of 3509 high-confidence plasma proteins. More recently, the extracellular vesicles community also reports new proteins identified in exosomes in multiple different resources including blood and breast milk, e.g. in ExoCarta (8). Additionally, the global Human Proteome Project (HPP) announces a set of mass spectrometry (MS) data interpretation guidelines that are presented to the broader research community (9).

Our team has recently conducted a systematical assessment of human proteome identified using quantitative proteomics tools such as MS and computational predictive models, as documented in a recent review article (10). To expand this effort, we developed a new human body fluid proteome (HBFP) database to organize 11 827 unique proteins reported in 164 scientific articles since 2001, which has a maximal false discovery rate (FDR) of 0.01 on both the peptide and protein levels. Until today, this database

Received 12 April 2021; Revised 23 September 2021; Accepted 28 September 2021

stores information about proteins from 17 types of body fluids including plasma/serum, saliva, urine, cerebrospinal fluid (CSF), seminal fluid (SF), amniotic fluid, tear fluid, bronchoalveolar lavage fluid (BALF), milk, synovial fluid, nipple aspirate fluid, cervical-vaginal fluid, pleural effusion, sputum, exhaled breath condensate, pancreatic juice and sweat.

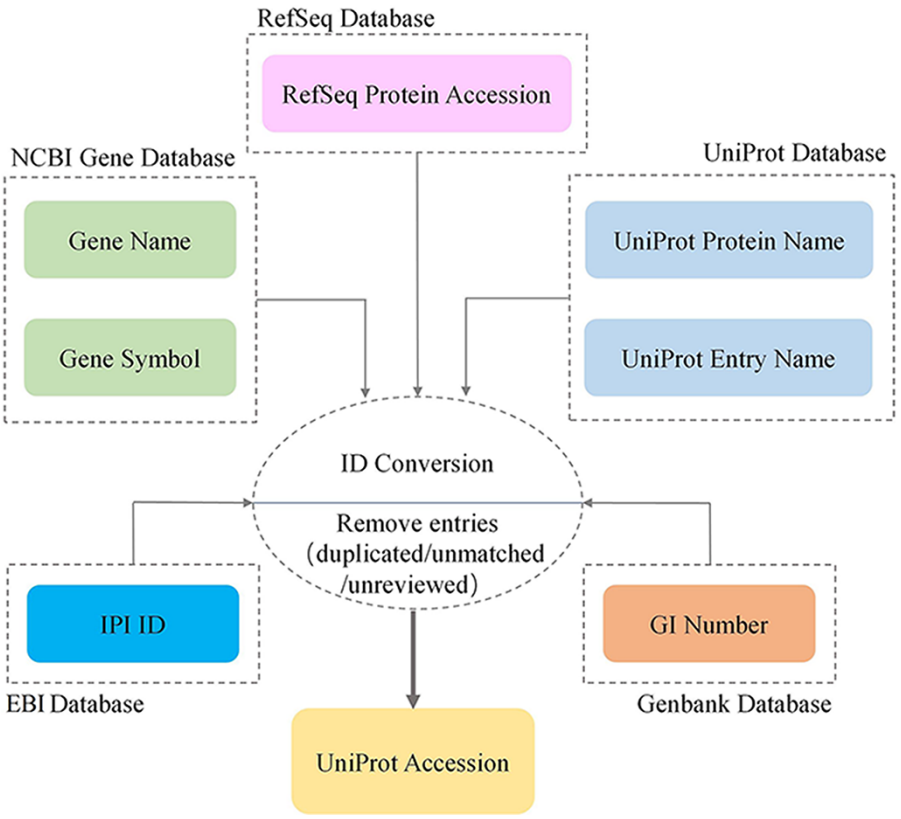


Figure 1. Workflow of protein identifier conversion.

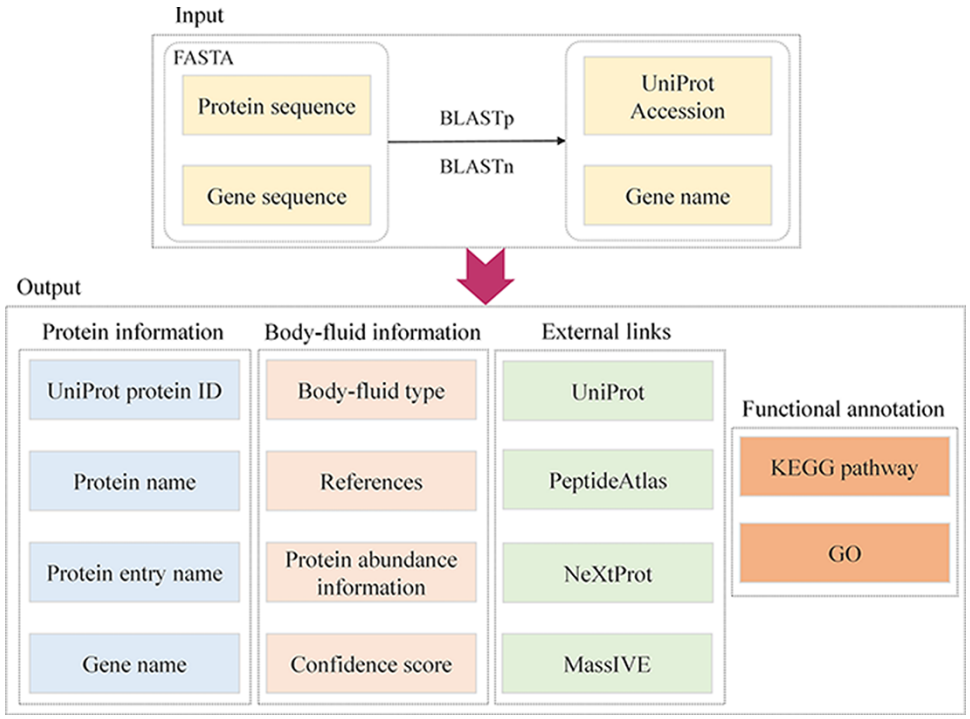


Figure 2. Construction workflow and utilities of querying page.

For each protein entry, description about protein secretion information, literature source, abundances, confidence and functional annotation is provided. This database system also provides users easy access to data visualization and download and functional analysis based on Gene Ontology (GO) and pathways.

Database content and design

Protein entries

We have manually collected proteins reported in 17 types of body fluids by carefully reviewing 164 scientific references published since 2001 based on a PubMed search with FDR $\leq$1% on both the peptide and protein levels.

In the HBFP database, each protein is assigned with a unique identifier of UniProtKB/Swiss-Prot accession (UniProt release 2020_06) (11). Since different identifiers have been mixed used in the referenced studies, we first used conversion tools at BioDBnet (https://biodbnet-abcc.ncifcrf.gov/) (12) and UniProt (https://www.UniProt.org/) to confidently convert different identifiers to UniProt accession numbers. The common identifiers involved in this study include International Protein Index ID [hosted at European Bioinformatics Institute (EBI) (closed in 2011)], GI number (from Genbank database), RefSeq protein accession (from RefSeq database), Gene name/symbol (from NCBI Gene database) and UniProt protein/entry name (from UniProt database). The ID conversion process is shown in Figure 1. During the conversion, poorly curated proteins with ambiguous identifiers were eliminated. For examples, many International Protein Index ID links to unclearly described instances that cannot be mapped to a UniProt entry are excluded.

Database utilities

The interface of the HBFP database is constructed by PHP, while the database system is based on MySQL. The main contents of the current database include query and browse pages described as follows.

Table 1. Overall statistics

Body fluid types		Statistics		References
		Number of protein entries	Number of references	
1	Plasma/serum	5790	38	(18–55)
2	Saliva	2758	21	(19, 56–75)
3	Urine	7330	23	(19, 76–97)
4	CSF	4364	12	(19, 90, 98–107)
5	SF	4084	5	(108–112)
6	Amniotic fluid	3025	6	(19, 113–117)
7	Tear fluid (TF)	1882	11	(118–128)
8	BALF	3434	6	(41, 129–133)
9	Milk	2457	14	(134–147)
10	Synovial fluid	1637	7	(148–154)
11	Nipple aspirate fluid	1734	5	(155–159)
12	Cervical–vaginal fluid	949	4	(160–163)
13	Pleural effusion	1519	3	(164–166)
14	Sputum	1809	3	(167–169)
15	Exhaled breath condensate	351	5	(170–174)
16	Pancreatic juice	702	4	(175–178)
17	Sweat	1244	3	(179–181)
Total (non-redundant)		11 827	164	

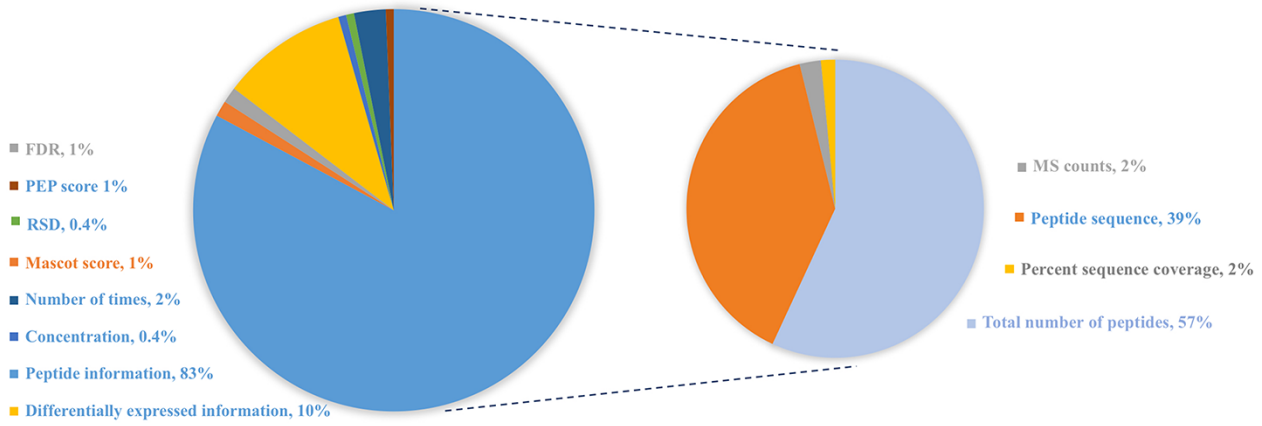


Figure 3. Distribution of protein abundance methods in HBFP database based on a number of original quantitative analysis methods from the original literature studies. Note that the sum of protein abundance is not 100% since not all of the literature studies provide quantitative analysis information.

Querying page

As one of the most important functions, the querying page allows users to search for body fluid proteins based on different types of input including protein ID, gene name, and protein or gene sequence. When given a FASTA input, BLASTp or BLASTn is used to translate sequence input to the best-match protein entry. The top hit (the highest bit score) from the BLAST search is considered the best match of the query. Figure 2 illustrates the workflow and content of querying page.

The annotation of each protein contains the following information:

- Protein ID/name/entry name
- Gene name
- Associated body fluid type along with indicated discovery confidence (explained in the next section)
- References and protein abundance information where the protein is reported
- External links to public databases including UniProt, PeptideAtlas and NeXtProt (13), MassIVE (14)
- Functional annotation based on the KEGG pathway (15) and GO (16)

Browsing page

This page provides an overview list of proteins associated with 17 types of body fluids and links to view and download selected proteins.

Database highlights

Data statistics

When determining the inclusion of reported proteins, we applied the following criteria for credibility of the MS evidence. First, for papers that issued peptide sequence details, we remapped all those peptide sequences to neXtProt (release 2021-02-15) using the neXtProt peptide uniqueness checker to remove unreliable matches (17). Specifically, we applied guideline #15 of HPP Guidelines 2.1 (9) to include proteins that contain at least two non-nested uniquely mapping peptides of nine amino acids into the HBFP database. According to this criterion, 7354 proteins were confirmed confidently. Another 4473 proteins were also included as they were not explicitly provided with peptide sequence information but have more than two unique peptides.

The overall statistics about the protein entries and references in terms of each body fluid are summarized in Table 1. The current HBFP database contains 11 827 distinct proteins from 17 types of body fluids. Note that urine exceeds all other body fluids in terms of protein counts while blood is at the second rank. All data are made publicly available in the HBFP and via links at https://bmbl.bmi.osumc.edu/HBFP/.

Protein abundance

In order to provide users experimental evidence from the original study, this database also displays relatively abundant information from the corresponding literature studies. General proteomics approaches using MS identify proteins by matching identified peptides against predefined protein sequence databases. The qualitative measures of protein reported in the original reference include the following: (i) peptide information: most of cited studies provide explicit information about peptide sequence, the total number of peptides, MS counts or the percent sequence coverage; (ii) differential expression information including fold change (positive value demonstrates up-regulated expression and negative

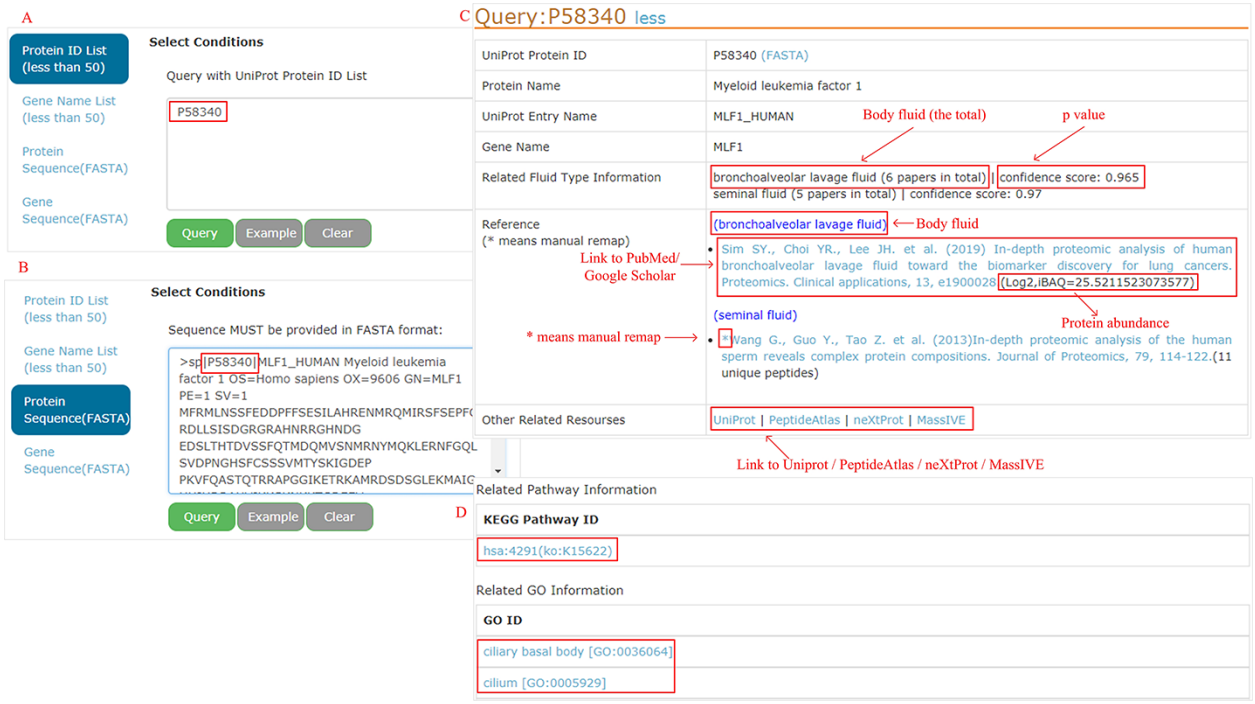


Figure 4. Example of query response with input as 'P58340' in the protein ID and protein sequence box.

value indicates down-regulated expression), up- or down-regulated expression in case vs. control or (normalized) spectral counts and (iii) other statistical information including FDR, relative standard deviation and the number of times across different samples or experiments, as shown in Figure 3.

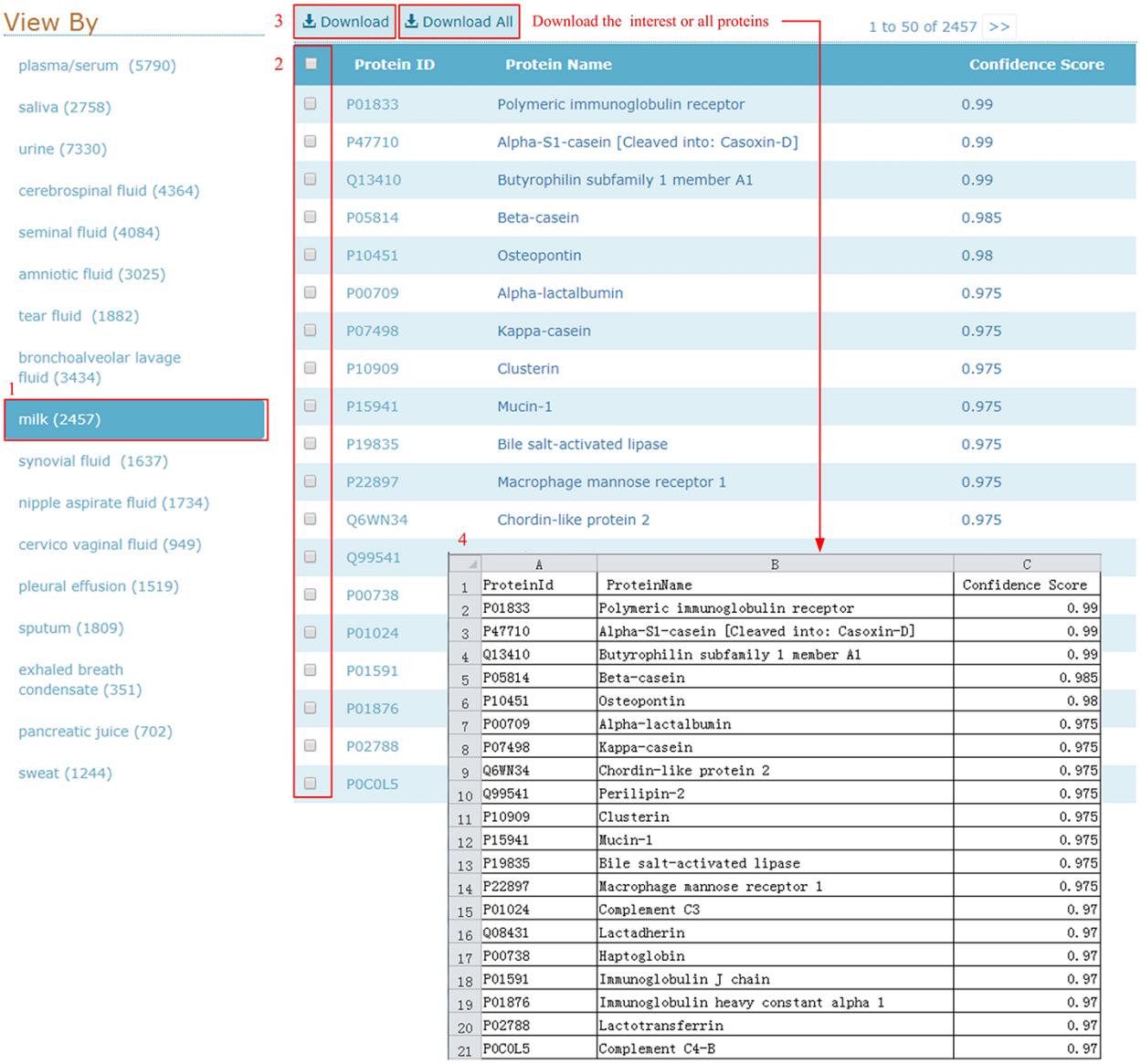
Confidence score

In the HBFP database, to evaluate the confidence level of each discovered protein in each body fluid, a new statistical measure is calculated based on Guideline # 9 of HPP guidelines 2.1 for the combined datasets. It is a well-known phenomenon that when taking N datasets with a substantial FDR and piling them all together, the overall FDR increases with the number of datasets. For example, for plasma, there are 38 papers with plasma protein lists, each with a substantial FDR ($\leq$1%). It is probably a conservative estimate to suppose that the FDR of such a combined result is $1\% + 0.5\% \times (N$ datasets$-1)$ (9). It means that 50% of the correct identifications overlap and none of the incorrect ones does, so the resulting FDR is added in a 0.5% increment. Meanwhile, the confidence level of protein in the combined datasets is also reduced. Otherwise, considering the overlap of the true positives, the larger the number of datasets in which a protein is associated with a specific fluid, the more reliable this protein is. In the end, a confidence score C is calculated as follows:

$$C_{i,j} = A_i + 0.5\% \times (M_j - 1) \quad (1)$$

$$A_i = 1 - FDR_i \quad (2)$$

$$FDR_i = 1\% + 0.5\% \times (N_i - 1) \quad (3)$$



	A	B	C
1	ProteinId	ProteinName	Confidence Score
2	P01833	Polymeric immunoglobulin receptor	0.99
3	P47710	Alpha-S1-casein [Cleaved into: Casoxin-D]	0.99
4	Q13410	Butyrophilin subfamily 1 member A1	0.99
5	P05814	Beta-casein	0.985
6	P10451	Osteopontin	0.98
7	P00709	Alpha-lactalbumin	0.975
8	P07498	Kappa-casein	0.975
9	Q6WN34	Chordin-like protein 2	0.975
10	Q99541	Perilipin-2	0.975
11	P10909	Clusterin	0.975
12	P15941	Mucin-1	0.975
13	P19835	Bile salt-activated lipase	0.975
14	P22897	Macrophage mannose receptor 1	0.975
15	P01024	Complement C3	0.97
16	Q08431	Lactadherin	0.97
17	P00738	Haptoglobin	0.97
18	P01591	Immunoglobulin J chain	0.97
19	P01876	Immunoglobulin heavy constant alpha 1	0.97
20	P02788	Lactotransferrin	0.97
21	P0C0L5	Complement C4-B	0.97

Figure 5. Download illustration where user can choose the body fluid name and download the proteins of interest or all proteins.

where N_i is the number of relevant literature studies (datasets) of a specific fluid i; FDR_i represents the overall FDR of multiple datasets in body fluid i; A_i means the confidence level of proteins in the combined datasets of body fluid i and M_j refers to the number of literature studies in which a protein j is identified in body fluid i.

For example, there are 38 literature studies related to blood in the HBFP, so $N_i = 38$, $FDR_i = 0.195$ and $A_i = 0.805$. The protein O14791 is identified in blood by 19 independent studies, i.e. $M_j = 19$. As a result, the calculated $C_{i,j}$ score for O14791 in blood is 0.895. Meanwhile, protein Q9UJV9 only is identified in one paper for blood, so $M_j = 1$ and $C_{i,j} = A_i = 0.805$. It means that protein Q9UJV9 maintains only the confidence level in the combined datasets of blood. Specifically, protein P01833 is identified in milk by 14 studies, and a total of 14 literature studies on milk are included in the HBFP, so protein P01833 maintains the original confidence level, i.e. 0.99. The larger the C score, the higher the confidence that a protein reported in that fluid will be. Note that this score can only be compared within the same type of body fluid.

Database applications

Data access

The website can be accessed through https://bmbl.bmi.osumc.edu/HBFP/.

Query

All proteins can be easily accessed by searching protein ID, gene name, protein sequence (FASTA) or gene sequence (FASTA) (<50 items per query) (Figure 4A and B as an example). A BLAST (182) is performed locally to find the best match when the sequence FASTA format is given. For each protein, detailed information is displayed (Figure 4C). Users can connect directly to the PubMed or Google Scholar to view the original study through the provided links. Four databases (UniProt, PeptideAtlas, NeXtProt and MassIVE) are cross-linked for additional protein annotation, while the KEGG pathway and GO are focused on the functional aspects (Figure 4D).

Download

HBFP allows users to browse the entire protein list in each body fluid, where the proteins are ordered based on descending confidence scores. Users can check and download all entries of the selected body fluid type in one go, as shown in Figure 5.

Demo of comparative analysis using the HBFP database

Body fluid analysis

Many proteins in the HBFP database have a broad distribution in terms of body fluid types. An internal comparative analysis across different fluids can provide further information regarding the specificity of a proposed marker protein. Of 11 827 identified proteins, 66.8% are identified in at least two body fluids (Figure 6). A total of 93 proteins (0.79%) are shared among all analyzed body fluids, which may indicate that these proteins are essential for various life activities (Table 2).

Venn diagram and GO annotation

To take a closer look at this comparison, we focused on five body fluids that have the most protein counts, including blood, urine, CSF, SF) and BALF. An interesting discovery is that urine shares large numbers of common proteins with other fluids (Figure 7). A total of 4109, 3212, 2990 and 2950 proteins overlapped between the plasma and the other

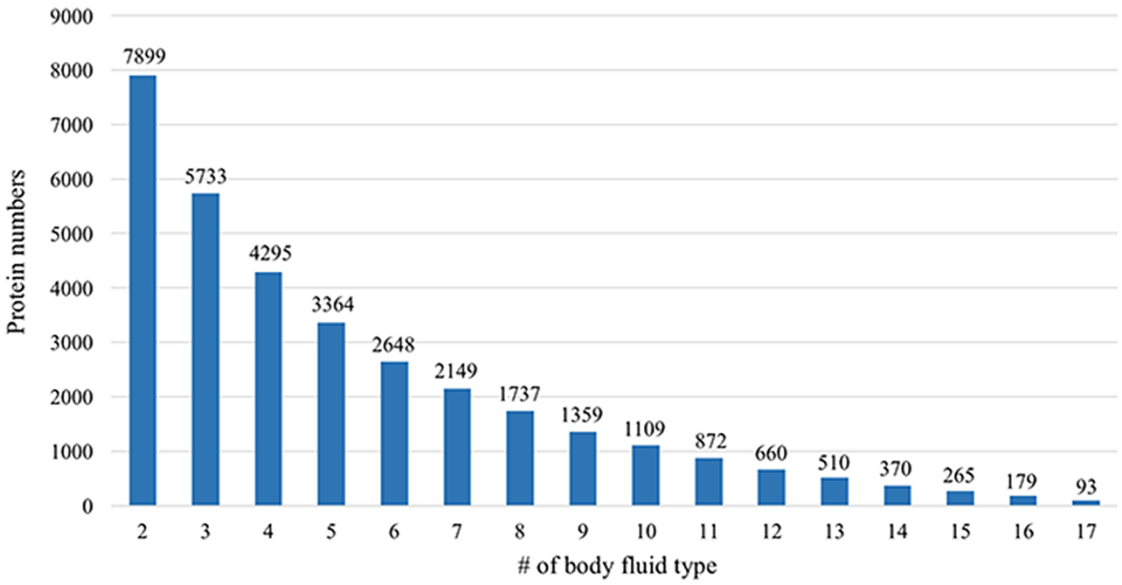


Figure 6. Comparative analysis across different body fluids. Seven thousand eight hundred and ninety-nine (7899) proteins are presented in at least two body fluids and 5733 proteins existed in at least three body fluids. Only 93 proteins exist in all 17 body fluids.

Table 2. List of 93 proteins shared among all 17 body fluids

	UniProt accession number	Protein name	Gene name
1	P11021	Endoplasmic reticulum chaperone BiP	HSPA5
2	P55072	Transitional endoplasmic reticulum ATPase	VCP
3	P13647	Keratin, type II cytoskeletal 5	KRT5
4	O00299	Chloride intracellular channel protein 1	CLIC1
5	P02787	Serotransferrin	TF
6	P22314	Ubiquitin-like modifier-activating enzyme 1	UBA1
7	P13645	Keratin, type I cytoskeletal 10	KRT10
8	P02533	Keratin, type I cytoskeletal 14	KRT14
9	P07237	Protein disulfide-isomerase	P4HB
10	P06576	ATP synthase subunit beta, mitochondrial	ATP5F1B
11	P30041	Peroxiredoxin-6	PRDX6
12	P63104	14-3-3 protein zeta/delta	YWHAZ
13	P62258	14-3-3 protein epsilon	YWHAE
14	P14923	Junction plakoglobin	JUP
15	P04040	Catalase	CAT
16	P01834	Immunoglobulin kappa constant	IGKC
17	P06702	Protein S100-A9	S100A9
18	P52209	6-Phosphogluconate dehydrogenase, decarboxylating	PGD
19	P18669	Phosphoglycerate mutase 1	PGAM1
20	P14618	Pyruvate kinase PKM	PKM
21	P61981	14-3-3 protein gamma	YWHAG
22	P07384	Calpain-1 catalytic subunit	CAPN1
23	P50395	Rab GDP dissociation inhibitor beta	GDI2
24	Q00610	Clathrin heavy chain 1	CLTC
25	P26641	Elongation factor 1-gamma	EEF1G
26	P32119	Peroxiredoxin-2	PRDX2
27	P19971	Thymidine phosphorylase	TYMP
28	P26038	Moesin	MSN
29	P40121	Macrophage-capping protein	CAPG
30	P35754	Glutaredoxin-1	GLRX
31	P01009	Alpha-1-antitrypsin	SERPINA1
32	P01860	Immunoglobulin heavy constant gamma 3	IGHG3
33	P06753	Tropomyosin alpha-3 chain	TPM3
34	P68871	Hemoglobin subunit beta	HBB
35	P62805	Histone H4	H4C1
36	P30086	Phosphatidylethanolamine-binding protein 1	PEBP1
37	P35579	Myosin-9	MYH9
38	P01023	Alpha-2-macroglobulin	A2M
39	Q06830	Peroxiredoxin-1	PRDX1
40	P02042	Hemoglobin subunit delta	HBD
41	P07737	Profilin-1	PFN1
42	P80188	Neutrophil gelatinase-associated lipocalin	LCN2
43	P02679	Fibrinogen gamma chain	FGG
44	P40925	Malate dehydrogenase, cytoplasmic	MDH1
45	P08758	Annexin A5	ANXA5
46	P46940	Ras GTPase-activating-like protein IQGAP1	IQGAP1
47	P01833	Polymeric immunoglobulin receptor	PIGR
48	P31949	Protein S100-A11	S100A11
49	P04792	Heat shock protein beta-1	HSPB1
50	P07339	Cathepsin D	CTSD
51	P01857	Immunoglobulin heavy constant gamma 1	IGHG1
52	P06733	Alpha-enolase	ENO1
53	P23284	Peptidyl-prolyl cis-trans isomerase B	PPIB
54	P02647	Apolipoprotein A-I	APOA1
55	O43707	Alpha-actinin-4	ACTN4
56	P30740	Leukocyte elastase inhibitor	SERPINB1
57	Q16610	Extracellular matrix protein 1	ECM1
58	P60709	Actin, cytoplasmic 1	ACTB
59	P15924	Desmoplakin	DSP
60	P62937	Peptidyl-prolyl cis-trans isomerase A	PPIA
61	P17931	Galectin-3	LGALS3
62	P00491	Purine nucleoside phosphorylase	PNP
63	P04080	Cystatin-B	CSTB
64	P02788	Lactotransferrin	LTF
65	P13639	Elongation factor 2	EEF2
66	P35527	Keratin, type I cytoskeletal 9	KRT9
67	P06396	Gelsolin	GSN
68	P59998	Actin-related protein 2/3 complex subunit 4	ARPC4

(continued)

Table 2. (Continued)

	UniProt accession number	Protein name	Gene name
69	P25311	Zinc-alpha-2-glycoprotein	AZGP1
70	P02768	Albumin	ALB
71	P61160	Actin-related protein 2	ACTR2
72	P04406	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH
73	P60174	Triosephosphate isomerase	TPI1
74	P18206	Vinculin	VCL
75	P08670	Vimentin	VIM
76	P10599	Thioredoxin	TXN
77	P11142	Heat shock cognate 71 kDa protein	HSPA8
78	P01011	Alpha-1-antichymotrypsin	SERPINA3
79	P04075	Fructose-bisphosphate aldolase A	ALDOA
80	P04264	Keratin, type II cytoskeletal 1	KRT1
81	P37837	Transaldolase	TALDO1
82	P35908	Keratin, type II cytoskeletal 2 epidermal	KRT2
83	P02545	Prelamin-A/C	LMNA
84	P69905	Hemoglobin subunit alpha	HBA1
85	P07900	Heat shock protein HSP 90-alpha	HSP90AA1
86	P29401	Transketolase	TKT
87	P00558	Phosphoglycerate kinase 1	PGK1
88	P00338	L-lactate dehydrogenase A chain	LDHA
89	P01861	Immunoglobulin heavy constant gamma 4	IGHG4
90	P05109	Protein S100-A8	S100A8
91	P04083	Annexin A1	ANXA1
92	P01024	Complement C3	C3
93	P09211	Glutathione S-transferase P	GSTP1

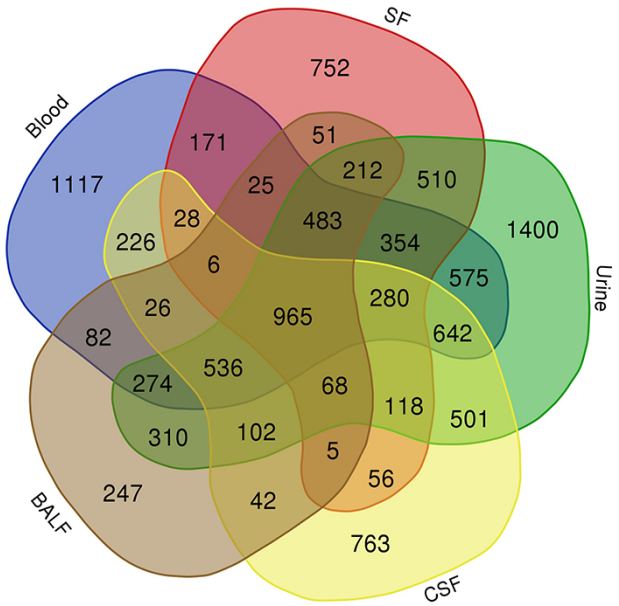


Figure 7. Venn diagram showing the common proteins among five body fluids (blood, urine, CSF, SF and BALF) that have the most number of proteins in the HBFP.

four body fluids (blood, CSF, SF and BALF, respectively). There are 965 proteins commonly detected in all five body fluids. The functional analysis using the BiNGO tool (183) in Cytoscape (184), reflecting information about cellular localization, molecular function and biological process of these proteins (Figure 8).

Conclusions

The new HBFP database developed in this study represents the first of its kind as a comprehensive reference resource of HBFP. All data are available through an open-access user-friendly Web platform. All protein entries were manually curated, which can be easily traced back to the original literature. Users can query and download proteins of interest to verify discovery in their own study or conduct an *in silico* analysis on human secretomes. We currently schedule a regular update every 6 months. The future plan is to include computationally identified proteins using statistical and machine learning approaches (185–191). In the past decade, many computational studies have revealed unique strengths in overcoming challenges in profiling-based proteomics research in terms of discovering new protein bioavailability and functions. Those computationally predicted proteins can serve as a secondary resource for biomarker discovery. In summary, by providing a wealth of information and functional analysis, we believe the HBFP database can be an excellent tool for the research community to explore human proteome in various body fluids.

Funding

National Natural Science Foundation of China (no. 62072212); Development Project of Jilin Province of China (nos 20200401083GX, 2020LY500L06 and 2020C003); Guangdong Key Project for Applied Fundamental Research (grant 2018KZDXM076); Jilin Province Key Laboratory of Big Data Intelligent Computing (no. 20180622002JC).

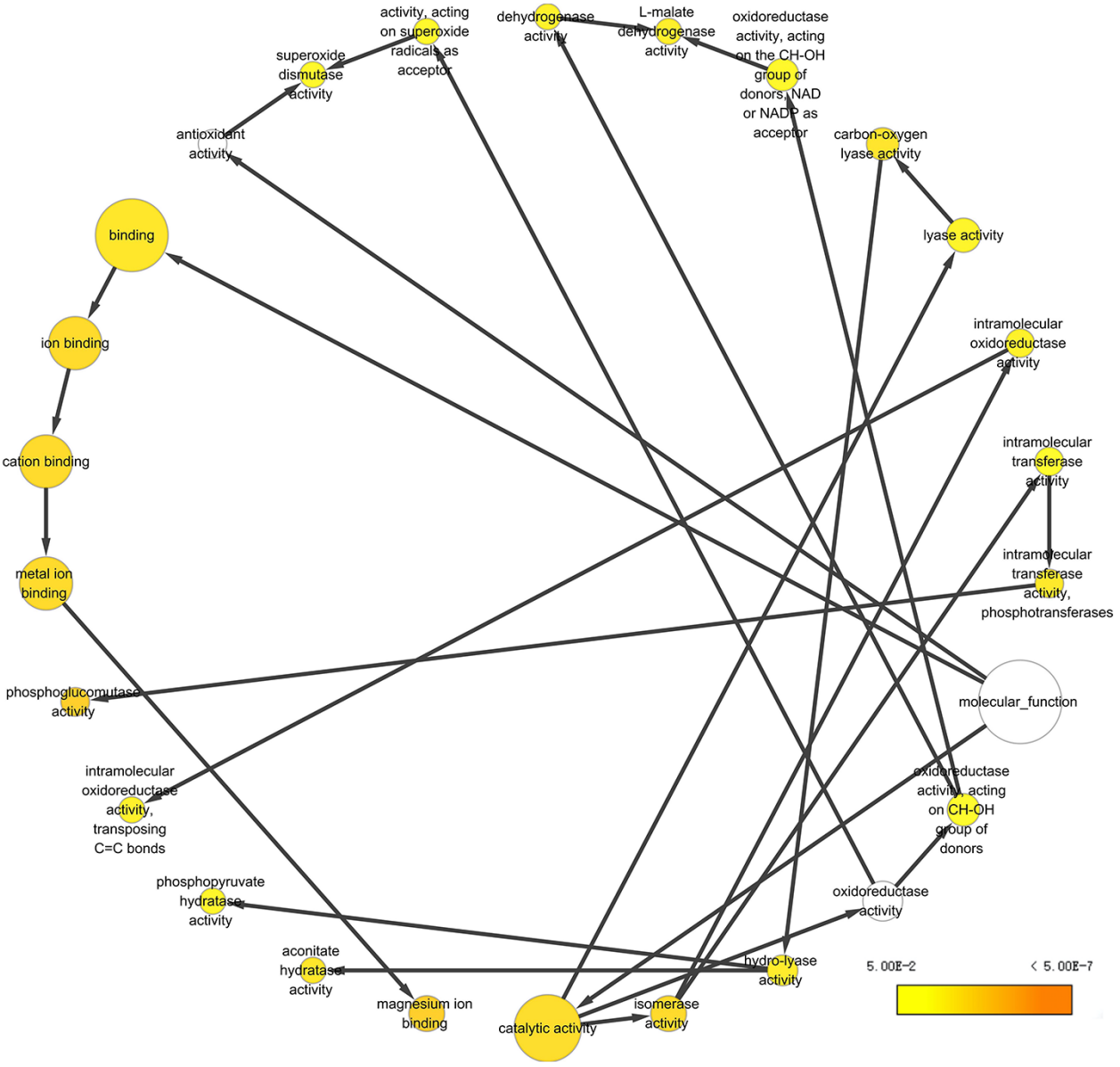


Figure 8. Example of GO annotation based on the 965 proteins common in five body fluids.

Conflict of interest

The authors declare that they have no competing interests.

References

1. Anderson,N.L. (2010) The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin. Chem.*, **56**, 177–185.
2. Lathrop,J., Anderson,N., Anderson,N. *et al.* (2003) Therapeutic potential of the plasma proteome. *Curr. Opin. Mol. Ther.*, **5**, 250–257.
3. Hu,S., Loo,J. and Wong,D. (2006) Human body fluid proteome analysis. *Proteomics*, **6**, 6326–6353.
4. Omenn,G.S., States,D.J., Adamski,M. *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.
5. Nanjappa,V., Thomas,J.K., Marimuthu,A. *et al.* (2014) Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic Acids Res.*, **42**, D959–D965.
6. Vizcaíno,J., Côté,R., Reisinger,F. *et al.* (2010) The Proteomics Identifications database: 2010 update. *Nucleic Acids Res.*, **38**, D736–D742.
7. Schwenk,J.M., Omenn,G.S., Sun,Z. *et al.* (2017) The human plasma proteome draft of 2017: building on the Human Plasma PeptideAtlas from mass spectrometry and complementary assays. *J. Proteome Res.*, **16**, 4299–4310.
8. Keerthikumar,S., Chisanga,D., Ariyaratne,D. *et al.* (2016) ExoCarta: a web-based compendium of exosomal cargo. *J. Mol. Biol.*, **428**, 688–692.
9. Deutsch,E.W., Overall,C.M., Eyk,J.V. *et al.* (2016) Human Proteome Project mass spectrometry data interpretation guidelines 2.1. *J. Proteome Res.*, **15**, 3961–3970.
10. Huang,L., Shao,D., Wang,Y. *et al.* (2021) Human body-fluid proteome: quantitative profiling and computational prediction. *Brief. Bioinformatics*, **22**, 315–333.

11. UniProt Consortium (2015) UniProt: a hub for protein information. *Nucleic Acids Res.*, **43**, D204–D212.
12. Mudunuri,U., Che,A., Yi,M. *et al.* (2009) bioDBnet: the biological database network. *Bioinformatics*, **25**, 555–556.
13. Lydie,L., Ghislaine,A.P., Aurore,B. *et al.* (2012) neXtProt: a knowledge platform for human proteins. *Nucleic Acids Res.*, **40**, D76–D83.
14. Wang,M., Wang,J., Carver,J. *et al.* (2018) Assembling the community-scale discoverable human proteome. *Cell Syst.*, **7**, 412–421.
15. Minoru,K., Michihiro,A., Susumu,G. *et al.* (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res.*, **36**, D480–D484.
16. Carbon,S., Ireland,A., Mungall,C.J. *et al.* (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics*, **25**, 288–289.
17. Mathieu,S., Alain,G., Daniel,T. *et al.* (2017) The neXtProt peptide uniqueness checker: a tool for the proteomics community. *Bioinformatics*, **33**, 3471–3472.
18. Zhao,Y., Chang,C., Qin,P. *et al.* (2016) Mining the human plasma proteome with three-dimensional strategies by high-resolution Quadrupole Orbitrap Mass Spectrometry. *Anal. Chim. Acta*, **904**, 65–75.
19. Zhao,M., Yang,Y., Guo,Z. *et al.* (2018) A comparative proteomics analysis of five body fluids: plasma, urine, cerebrospinal fluid, amniotic fluid and saliva. *Proteomics Clin. Appl.*, **12**, e1800008.
20. Yan,W., Apweiler,R., Balgley,B.M. *et al.* (2010) Systematic comparison of the human saliva and plasma proteomes. *Proteomics Clin. Appl.*, **3**, 116–134.
21. Moreno,S.O., Cominetti,O., Galindo,A.N. *et al.* (2017) The differential plasma proteome of obese and overweight individuals undergoing a nutritional weight loss and maintenance intervention. *Proteomics Clin. Appl.*, **12**, 1600150.
22. Li,L., Xu,Y. and Yu,C.X. (2012) Proteomic analysis of serum of women with elevated Ca-125 to differentiate malignant from benign ovarian tumors. *Asian Pac. J. Cancer Prev.*, **13**, 3265–3270.
23. Chen,L.Z., Gu,H., Li,J. *et al.* (2016) Comprehensive maternal serum proteomics identifies the cytoskeletal proteins as non-invasive biomarkers in prenatal diagnosis of congenital heart defects. *Sci. Rep.*, **6**, 19248.
24. Acosta-Martin,A.E., Panchaud,A., Chwastyniak,M. *et al.* (2011) Quantitative mass spectrometry analysis using PAcIFIC for the identification of plasma diagnostic biomarkers for abdominal aortic aneurysm. *PLoS One*, **6**, e28698.
25. Pietzner,M., Engelmann,B., Kacprowski,T. *et al.* (2017) Plasma proteome and metabolome characterization of an experimental human thyrotoxicosis model. *BMC Med.*, **15**, 6.
26. Boichenko,A.P., Govorukhina,N., Klip,H.G. *et al.* (2014) A panel of regulated proteins in serum from patients with cervical intraepithelial neoplasia and cervical cancer. *J. Proteome Res.*, **13**, 4995–5007.
27. Geyer,P.E., Kulak,N.A., Pichler,G. *et al.* (2016) Plasma proteome profiling to assess human health and disease. *Cell Syst.*, **2**, 185–195.
28. Geyer,P.E., Wewer Albrechtsen,N.J., Tyanova,S. *et al.* (2016) Proteomics reveals the effects of sustained weight loss on the human plasma proteome. *Mol. Syst. Biol.*, **12**, 901.
29. Yadav,A.K., Bhardwaj,G., Basak,T. *et al.* (2011) A systematic analysis of eluted fraction of plasma post immunoaffinity depletion: implications in biomarker discovery. *PLoS One*, **6**, e24442.
30. Liu,Z., Fan,S., Liu,H. *et al.* (2016) Enhanced detection of low-abundance human plasma proteins by integrating polyethylene glycol fractionation and immunoaffinity depletion. *PLoS One*, **11**, e0166306.
31. Limonier,F., Steendam,K.V., Waeterloos,G. *et al.* (2016) An application of mass spectrometry for quality control of biologicals: highly sensitive profiling of plasma residuals in human plasma-derived immunoglobulin. *J. Proteomics*, **152**, 312–320.
32. Bjelosevic,S., Pascovici,D., Ping,H. *et al.* (2017) Quantitative age-specific variability of plasma proteins in healthy neonates, children and adults. *Mol. Cell Proteomics*, **16**, 924–935.
33. Farrah,T., Deutsch,E.W., Omenn,G.S. *et al.* (2011) A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. *Mol. Cell. Proteomics*, **152**, 312–320.
34. Gautam,P., Nair,S.C., Ramamoorthy,K. *et al.* (2013) Analysis of human blood plasma proteome from ten healthy volunteers from Indian population. *PLoS One*, **8**, e72584.
35. Cheon,D.H., Nam,E.J., Park,K.H. *et al.* (2016) Comprehensive analysis of low-molecular-weight human plasma proteome using top-down mass spectrometry. *J. Proteome Res.*, **15**, 229–244.
36. Zhou,M., Prieto,D.A., Lucas,D.A. *et al.* (2006) Identification of the SELDI ProteinChip human serum retentate by microcapillary liquid chromatography-tandem mass spectrometry. *J. Proteome Res.*, **5**, 2207–2216.
37. Zeng,Z., Hincapie,M., Pitteri,S.J. *et al.* (2011) A proteomics platform combining depletion, multi-lectin affinity chromatography(M-LAC), and isoelectric focusing to study the breast cancer proteome. *Anal. Chem.*, **83**, 4845–4854.
38. Pan,S., Chen,R., Crispin,D.A. *et al.* (2011) Protein alterations associated with pancreatic cancer and chronic pancreatitis found in human plasma using global quantitative proteomics profiling. *J. Proteome Res.*, **10**, 2359–2376.
39. Surinova,S., Choi,M., Tao,S. *et al.* (2015) Prediction of colorectal cancer diagnosis based on circulating plasma proteins. *EMBO Mol. Med.*, **7**, 1166–1178.
40. Harel,M., Oren-Giladi,P., Kaidar-Person,O. *et al.* (2015) Proteomics of microparticles with SILAC Quantification (PROMIS-Quan): a novel proteomic method for plasma biomarker quantification. *Mol. Cell. Proteomics*, **14**, 1127–1136.
41. Carvalho,A.S., Cuco,C.M., Lavareda,C. *et al.* (2017) Bronchoalveolar lavage proteomics in patients with suspected lung cancer. *Sci. Rep.*, **7**, 42190.
42. Zhou,B., Zhou,Z., Chen,Y. *et al.* (2019) Plasma proteomics-based identification of novel biomarkers in early gastric cancer. *Clin. Biochem.*, **76**, 5–10.
43. Du,Z., Liu,X., Wei,X. *et al.* (2020) Quantitative proteomics identifies a plasma multi-protein model for detection of hepatocellular carcinoma. *Sci. Rep.*, **10**, 15552.
44. Park,J., Kim,H., Kim,S.Y. *et al.* (2020) In-depth blood proteome profiling analysis revealed distinct functional characteristics of plasma proteins between severe and non-severe COVID-19 patients. *Sci. Rep.*, **10**, 22418.
45. Garay-Baquero,D.J., White,C.H., Walker,N.F. *et al.* (2020) Comprehensive plasma proteomic profiling reveals biomarkers for active tuberculosis. *JCI Insight*, **5**, e137427.
46. Kumar,V., Ray,S., Ghantasala,S. *et al.* (2020) An integrated quantitative proteomics workflow for cancer biomarker discovery and validation in plasma. *Front. Oncol.*, **10**, 543997.
47. Geyer,P.E., Arend,F.M., Doll,S. *et al.* (2021) High-resolution serum proteome trajectories in COVID-19 reveal patient-specific seroconversion. *EMBO Mol. Med.*, **13**, e14167.
48. Messner,C.B., Demichev,V., Wendisch,D. *et al.* (2020) Ultra-high-throughput clinical proteomics reveals classifiers of COVID-19 infection. *Cell Syst.*, **11**, 11–24.e4.
49. Ming,C., Yi,L., Hla,B. *et al.* (2020) Quantitative proteomics and reverse engineer analysis identified plasma exosome derived protein markers related to osteoporosis. *J. Proteomics*, **228**, 103940.
50. Yang,T., Fu,Z., Zhang,Y. *et al.* (2020) Serum proteomics analysis of candidate predictive biomarker panel for the diagnosis of trastuzumab-based therapy resistant breast cancer. *Biomed. Pharmacother.*, **129**, 110465.

51. Dey,K.K., Wang,H., Niu,M. *et al.* (2019) Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer's disease. *Clin. Proteomics*, **16**, 16.
52. Smolarz,M., Pietrowska,M., Matysiak,N. *et al.* (2019) Proteome profiling of exosomes purified from a small amount of human serum: the problem of co-purified serum components. *Proteomes*, **7**, 18.
53. Lin,L., Zheng,J., Yu,Q. *et al.* (2017) High throughput and accurate serum proteome profiling by integrated sample preparation technology and single-run data independent mass spectrometry analysis. *J. Proteomics*, **174**, 9–16.
54. Ren,J., Zhao,G., Sun,X. *et al.* (2017) Identification of plasma biomarkers for distinguishing bipolar depression from major depressive disorder by iTRAQ-coupled LC-MS/MS and bioinformatics analysis. *Psychoneuroendocrinology*, **86**, 17–24.
55. Liu,C.W., Bramer,L., Webb-Robertson,B.J. *et al.* (2017) Temporal expression profiling of plasma proteins reveals oxidative stress in early stages of Type 1 diabetes progression. *J. Proteomics*, **172**, 100–110.
56. Rao,P.V., Reddy,A.P., Lu,X. *et al.* (2009) Proteomic identification of salivary biomarkers of type-2 diabetes. *J. Proteome Res.*, **8**, 239–245.
57. Guo,T., Rudnick,P.A., Wang,W.J. *et al.* (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.
58. Wilmarth,P.A., Riviere,M.A., Rustvold,D.L. *et al.* (2004) Two-dimensional liquid chromatography study of the human whole saliva proteome. *J. Proteome Res.*, **3**, 1017–1023.
59. Gonzalezbegne,M., Lu,B.W., Liao,L.J. *et al.* (2011) Characterization of the human submandibular/sublingual saliva glycoproteome using lectin affinity chromatography coupled to Multidimensional Protein Identification Technology. *J. Proteome Res.*, **10**, 5031–5046.
60. Sivadasan,P., Gupta,M.K., Sathe,G.J. *et al.* (2015) Data from human salivary proteome – a resource of potential biomarkers for oral cancer. *J. Proteomics*, **4**, 374–378.
61. Cho,H.R., Kim,H.S., Park,J.S. *et al.* (2017) Construction and characterization of the Korean whole saliva proteome to determine ethnic differences in human saliva proteome. *PLoS One*, **12**, e0181765.
62. Winck,F.V., Ribeiro,A.C.P., Domingues,R.R. *et al.* (2015) Insights into immune responses in oral cancer through proteomic analysis of saliva and salivary extracellular vesicles. *Sci. Rep.*, **5**, 16305.
63. Aboodi,G.M., Sima,C., Moffa,E.B. *et al.* (2016) Salivary cytoprotective proteins in inflammation and resolution during experimental gingivitis—a pilot study. *Front. Cell. Infect. Microbiol.*, **5**, 92.
64. Xie,H., Rhodus,N.L., Griffin,R.J. *et al.* (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.
65. Bandhakavi,S., Stone,M.D., Onsongo,G. *et al.* (2009) A dynamic range compression and three-dimensional peptide fractionation analysis platform expands proteome coverage and the diagnostic potential of whole saliva. *J. Proteome Res.*, **8**, 5590–5600.
66. De Jong,E.P., Xie,H.W., Onsongo,G. *et al.* (2010) Quantitative proteomics reveals myosin and actin as promising saliva biomarkers for distinguishing pre-malignant and malignant oral lesions. *PLoS One*, **5**, e11148.
67. Franco-Martínez,L., Hernández,J.M.G., Horvati,A. *et al.* (2019) Differences on salivary proteome at rest and in response to an acute exercise in men and women: a pilot study. *J. Proteomics*, **214**, 103629.
68. Contini,C., Olianas,A., Serrao,S. *et al.* (2021) Top-down proteomics of human saliva highlights anti-inflammatory, antioxidant, and antimicrobial defense responses in alzheimer disease. *Front. Neurosci.*, **15**, 668852.
69. Sembler-Mller,M.L., Belstrm,D., Locht,H. *et al.* (2020) Proteomics of saliva, plasma, and salivary gland tissue in Sjögren's syndrome and non-Sjögren patients identify novel biomarker candidates. *J. Proteomics*, **225**, 103877.
70. Xiao,X., Liu,Y., Guo,Z. *et al.* (2017) Comparative proteomic analysis of the influence of gender and acid stimulation on normal human saliva using LC/MS/MS. *Proteomics Clin. Appl.*, **11**, 1600142.
71. Sun,Y., Huo,C., Qiao,Z. *et al.* (2018) Comparative proteomic analysis of exosomes and microvesicles in human saliva for lung cancer. *J. Proteome Res.*, **17**, 1101–1107.
72. Wu,C.C., Chu,H.W., Hsu,C.W. *et al.* (2015) Saliva proteome profiling reveals potential salivary biomarkers for detection of oral cavity squamous cell carcinoma. *Proteomics*, **15**, 3394–3404.
73. Suresh,A. (2015) Human salivary proteome — a resource of potential biomarkers for oral cancer. *J. Proteomics*, **127**, 89–95.
74. Cecchettini,A., Finamore,F., Ucciferri,N. *et al.* (2019) Phenotyping multiple subsets in Sjögren's syndrome: a salivary proteomic SWATH-MS approach towards precision medicine. *Clin. Proteomics*, **16**, 26.
75. Jehmlich,N., Dinh,K., Gesell-Salazar,M. *et al.* (2013) Quantitative analysis of the intra- and inter-subject variability of the whole salivary proteome. *J. Periodont. Res.*, **48**, 392–403.
76. Castagna,A., Cecconi,D., Sennels,L. *et al.* (2005) Exploring the hidden human urinary proteome via ligand library beads. *J. Proteome Res.*, **4**, 1917–1930.
77. Alamgir,K. and Packer,N.H. (2006) Simple urinary sample preparation for proteomic analysis. *J. Proteome Res.*, **5**, 2824–2838.
78. Li,Q.R., Fan,K.X., Li,R.X. *et al.* (2010) A comprehensive and non-prefractionation on the protein level approach for the human urinary proteome: touching phosphorylation in urine. *Rapid Commun. Mass Spectrom. RCM*, **24**, 823–832.
79. Guo,Z., Wang,Z., Lu,C. *et al.* (2018) Analysis of the differential urinary protein profile in IgA nephropathy patients of Uygur ethnicity. *BMC Nephrol.*, **19**, 358.
80. Hogan,M.C., Johnson,K.L., Zenka,R.M. *et al.* (2014) Subfractionation, characterization, and in-depth proteomic analysis of glomerular membrane vesicles in human urine. *Kidney Int.*, **85**, 1225–1237.
81. Nielsen,H.H., Beck,H.C., Kristensen,L.P. *et al.* (2015) The urine proteome profile is different in neuromyelitis optica compared to multiple sclerosis: a clinical proteome study. *PLoS One*, **10**, e0139659.
82. Lin,L., Yu,Q., Zheng,J.X. *et al.* (2018) Fast quantitative urinary proteomic profiling workflow for biomarker discovery in kidney cancer. *Clin. Proteomics*, **15**, 42.
83. Zhao,M.D., Li,M.L., Yang,Y.H. *et al.* (2017) A comprehensive analysis and annotation of human normal urinary proteome. *Sci. Rep.*, **7**, 3024.
84. Onile,O.S., Calder,B., Soares,N.C. *et al.* (2017) Quantitative label-free proteomic analysis of human urine to identify novel candidate protein biomarkers for schistosomiasis. *PLoS Negl. Trop. Dis.*, **11**, e0006045.
85. Simona,P., Yunee,K., Simona,F. *et al.* (2012) Identification of prostate-enriched proteins by in-depth proteomic analyses of expressed prostatic secretions in urine. *J. Proteome Res.*, **11**, 2386–2396.
86. Adachi,J., Kumar,C., Zhang,Y.L. *et al.* (2006) The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol.*, **7**, R80.
87. Liu,X.J., Shao,C., Wei,L.L. *et al.* (2012) An individual urinary proteome analysis in normal human beings to define the minimal

sample number to represent the normal urinary proteome. *Proteome Sci.*, **10**, 70.
88. Zheng,J.H., Liu,L.G., Wang,J. *et al.* (2013) Urinary proteomic and non-prefractionation quantitative phosphoproteomic analysis during pregnancy and non-pregnancy. *BMC Genomics*, **14**, 777.
89. Marimuthu,A., O'Meally,R.N., Chaerkady,R. *et al.* (2011) A comprehensive map of the human urinary proteome. *J. Proteome Res.*, **10**, 2734–2743.
90. Guo,Z.G., Zhang,Y., Zou,L.L. *et al.* (2015) A proteomic analysis of individual and gender variations in normal human urine and cerebrospinal fluid using iTRAQ quantification. *PLoS One*, **10**, e0133270.
91. Prikryl,P., Satrapova,V., Frydlova,J. *et al.* (2020) Mass spectrometry-based proteomic exploration of the small urinary extracellular vesicles in ANCA-associated vasculitis in comparison with total urine. *J. Proteomics*, **233**, 104067.
92. Swensen,A.C., He,J., Fang,A.C. *et al.* (2021) A comprehensive urine proteome database generated from patients with various renal conditions and prostate cancer. *Front. Med.*, **8**, 548212.
93. Li,Y., Wang,Y., Liu,H. *et al.* (2021) Urine proteome of COVID-19 patients. *URINE*, **2**, 1–8.
94. Chen,R., Yi,Y., Xiao,W. *et al.* (2021) Label-free liquid chromatography–mass spectrometry proteomic analysis of urinary identification in diabetic vascular dementia in a han chinese population. *Front. Aging Neurosci.*, **13**, 619945.
95. Huo,S., Wang,H.X., Yan,M.X. *et al.* (2021) Urinary proteomic characteristics of hyperuricemia and their possible links with the occurrence of its concomitant diseases. *ACS Omega*, **6**, 9500–9508.
96. Ahn,H.S., Kim,J.H., Jeong,H. *et al.* (2020) Differential urinary proteome analysis for predicting prognosis in type 2 diabetes patients with and without renal dysfunction. *Int. J. Mol. Sci.*, **21**, 4236.
97. Chen,C.J., Chou,C.Y., Shu,K.H. *et al.* (2021) Discovery of novel protein biomarkers in urine for diagnosis of urothelial cancer using iTRAQ proteomics. *J. Proteome Res.*, **20**, 2953–2963.
98. Pan,S., Wang,Y., Quinn,J.F. *et al.* (2006) Identification of glycoproteins in human cerebrospinal fluid with a complementary proteomic approach. *J. Proteome Res.*, **5**, 2769–2779.
99. Guldbrandsen,A., Vethe,H., Farag,Y. *et al.* (2014) In-depth characterization of the cerebrospinal fluid (CSF) proteome displayed through the CSF proteome resource (CSF-PR). *Mol. Cell. Proteomics Mcp*, **13**, 3152–3163.
100. Mouton-Barbosa,E., Roux-Dalvai,F., Bouyssié,D. *et al.* (2010) In-depth exploration of cerebrospinal fluid by combining peptide ligand library treatment and label-free protein quantification. *Mol. Cell. Proteomics*, **9**, 1006–1021.
101. Schutzer,S.E., Liu,T., Natelson,B.H. *et al.* (2010) Establishing the proteome of normal human cerebrospinal fluid. *PLoS One*, **5**, e10980.
102. Borg,J., Campos,A., Diema,C. *et al.* (2011) Spectral counting assessment of protein dynamic range in cerebrospinal fluid following depletion with plasma-designed immunoaffinity columns. *Clin. Proteomics*, **8**, 6.
103. Hu,Z.Y., Zhang,H.Y., Zhang,Y. *et al.* (2014) Nanoparticle size matters in the formation of plasma protein coronas on Fe_3O_4 nanoparticles. *Colloids Surf B Biointerfaces*, **121**, 354–361.
104. Schutzer,S.E., Angel,T.E., Liu,T. *et al.* (2011) Distinct cerebrospinal fluid proteomes differentiate post-treatment lyme disease from chronic fatigue syndrome. *PLoS One*, **6**, e17287.
105. Pan,S., Zhu,D., Quinn,J.F. *et al.* (2007) A combined dataset of human cerebrospinal fluid proteins identified by multidimensional chromatography and tandem mass spectrometry. *Proteomics*, **7**, 469–473.
106. Begcevic,I., Brinc,D., Drabovich,A.P. *et al.* (2016) Identification of brain-enriched proteins in the cerebrospinal fluid proteome by LC-MS/MS profiling and mining of the Human Protein Atlas. *Clin. Proteomics*, **13**, 11.
107. Charlotte,M., Lydie,L., Antonio,N.G. *et al.* (2018) Deep dive in the proteome of human cerebrospinal fluid: a valuable data resource for biomarker discovery and missing protein identification. *J. Proteome Res.*, **17**, 4113–4126.
108. Yang,C., Guo,W.B., Zhang,W.S. *et al.* (2017) Comprehensive proteomics analysis of exosomes derived from human seminal plasma. *Andrology*, **5**, 1007–1015.
109. Ashok,A., Ahmet,A., Luna,S. *et al.* (2015) Comparative proteomic network signatures in seminal plasma of infertile men as a function of reactive oxygen species. *Clin. Proteomics*, **12**, 23.
110. Pilch,B. and Mann,M. (2006) Large-scale and high-confidence proteomic analysis of human seminal plasma. *Genome Biol.*, **7**, R40.
111. Wang,G., Guo,Y., Tao,Z. *et al.* (2013) In-depth proteomic analysis of the human sperm reveals complex protein compositions. *J. Proteomics*, **79**, 114–122.
112. Zhang,X.G., Vos,H.R., Tao,W. *et al.* (2020) Proteomic profiling of two distinct populations of extracellular vesicles isolated from human seminal plasma. *Int. J. Mol. Sci.*, **21**, 7957.
113. Lee,J., Lee,J.E., Choi,J.W. *et al.* (2020) Proteomic analysis of amniotic fluid proteins for predicting the outcome of emergency cerclage in women with cervical insufficiency. *Reprod. Sci.*, **27**, 1318–1329.
114. Cho,C.K., Smith,C.R. and Diamandis,E.P. (2010) Amniotic fluid proteome analysis from Down syndrome pregnancies for biomarker discovery. *J. Proteome Res.*, **9**, 3574–3582.
115. Liu,X., Song,Y.J., Guo,Z.G. *et al.* (2019) A comprehensive profile and inter-individual variations analysis of the human normal amniotic fluid proteome. *J. Proteomics*, **192**, 1–9.
116. Jeon,H.S., Lee,S.M., Jung,Y.M. *et al.* (2020) Proteomic biomarkers in mid-trimester amniotic fluid associated with adverse pregnancy outcomes in patients with systemic lupus erythematosus. *PLoS One*, **15**, e0235838.
117. Hong,S., Ji,E.L., Yu,M.K. *et al.* (2020) Identifying potential biomarkers related to pre-term delivery by proteomic analysis of amniotic fluid. *Sci. Rep.*, **10**, 19648.
118. Zhou,L., Zhao,S.Z., Koh,S.K. *et al.* (2012) In-depth analysis of the human tear proteome. *J. Proteomics*, **75**, 3877–3885.
119. Liu,Q., Liu,J., Ren,C. *et al.* (2017) Proteomic analysis of tears following acupuncture treatment for menopausal dry eye disease by two-dimensional nano-liquid chromatography coupled with tandem mass spectrometry. *Int. J. Nanomed.*, **12**, 1663–1671.
120. Huang,Z., Du,C.X. and Pan,X.D. (2018) The use of in-strip digestion for fast proteomic analysis on tear fluid from dry eye patients. *PLoS One*, **13**, e0200702.
121. Soria,J., Acera,A., Merayo-Lloves,J. *et al.* (2017) Tear proteome analysis in ocular surface diseases using label-free LC-MS/MS and multiplexed-microarray biomarker validation. *Rep*, **7**, 17478.
122. Nttinen,J., Mkinen,P., Aapola,U. *et al.* (2020) Early changes in tear film protein profiles after femtosecond LASIK surgery. *Clin. Proteomics*, **17**, 36.
123. Csősz,É., Boross,P., Csutak,A. *et al.* (2012) Quantitative analysis of proteins in the tear fluid of patients with diabetic retinopathy. *J. Proteomics*, **75**, 2196–2204.
124. Chen,X.L., Rao,J., Zheng,Z. *et al.* (2019) Integrated tear proteome and metabolome reveal panels of inflammatory-related molecules via key regulatory pathways in dry eye syndrome. *J. Proteome Res.*, **18**, 2321–2330.
125. Tong,L., Zhou,X.Y., Jylha,A. *et al.* (2015) Quantitation of 47 human tear proteins using high resolution multiple reaction monitoring (HR-MRM) based-mass spectrometry. *J. Proteomics*, **115**, 36–48.
126. Boerger,M., Funke,S., Leha,A. *et al.* (2019) Proteomic analysis of tear fluid reveals disease-specific patterns in patients with

Parkinson's disease – a pilot study. *Parkinsonism Relat. Disord.*, **63**, 3–9.
127. Cheung,J.K.W., Bian,J.F., Sze,Y.H. *et al.* (2021) Human tear proteome dataset in response to daily wear of water gradient contact lens using SWATH-MS approach. *Data Brief*, **36**, 107120.
128. Dor,M., Eperon,S., Lalive,P.H. *et al.* (2018) Investigation of the global protein content from healthy human tears. *Exp. Eye Res.*, **179**, 64–74.
129. Almatroodi,S.A., Mcdonald,C.F., Collins,A.L. *et al.* (2015) Quantitative proteomics of bronchoalveolar lavage fluid in lung adenocarcinoma. *Cancer Genomics Proteomics*, **12**, 39–48.
130. Sim,S.Y., Choi,Y.R., Lee,J.H. *et al.* (2019) In-depth proteomic analysis of human bronchoalveolar lavage fluid toward the biomarker discovery for lung cancers. Proteomics. *Clin. Appl.*, **13**, e1900028.
131. Foster,M.W., Thompson,J.W., Que,L.G. *et al.* (2013) Proteomic analysis of human bronchoalveolar lavage fluid after subsgemental exposure. *J. Proteome Res.*, **12**, 2194–2205.
132. Ortea,I., Rodríguez-Ariza,A., Chicano-Gálvez,E. *et al.* (2016) Discovery of potential protein biomarkers of lung adenocarcinoma in bronchoalveolar lavage fluid by SWATH MS data-independent acquisition and targeted data extraction. *J. Proteomics*, **138**, 106–114.
133. Foster,M.W., Morrison,L.D., Todd,J.L. *et al.* (2015) Quantitative proteomics of bronchoalveolar lavage fluid in idiopathic pulmonary fibrosis. *J. Proteome Res.*, **14**, 1238–1249.
134. Yang,M., Cong,M., Peng,X. *et al.* (2016) Quantitative proteomic analysis of milk fat globule membrane (MFGM) proteins in human and bovine colostrum and mature milk samples through iTRAQ labeling. *Food Funct.*, **7**, 2438–2450.
135. Liao,Y., Alvarado,R., Phinney,B. *et al.* (2011) Proteomic characterization of human milk whey proteins during a twelve-month lactation period. *J. Proteome Res.*, **10**, 1746–1754.
136. Beck,K.L., Weber,D., Phinney,B.S. *et al.* (2015) Comparative proteomics of human and macaque milk reveals species-specific nutrition during postnatal development. *J. Proteome Res.*, **14**, 2143–2157.
137. Liao,Y.L., Alvarado,R., Phinney,B. *et al.* (2011) Proteomic characterization of specific minor proteins in the human milk casein fraction. *J. Proteome Res.*, **10**, 5409–5415.
138. Zhang,Q., Cundiff,J.K., Maria,S.D. *et al.* (2013) Quantitative analysis of the human milk whey proteome reveals developing milk and mammary-gland functions across the first year of lactation. *Proteomes*, **1**, 128–158.
139. Molinari,C.E., Casadio,Y.S., Hartmann,B.T. *et al.* (2012) Proteome mapping of human skim milk proteins in term and preterm milk. *J. Proteome Res.*, **11**, 1696–1714.
140. Kim,B.J. and Dallas,D.C. (2021) Systematic examination of protein extraction, proteolytic glycopeptide enrichment and MS/MS fragmentation techniques for site-specific profiling of human milk N-glycoproteins. *Talanta*, **224**, 121811.
141. Dallas,D.C., Guerrero,A., Khaldi,N. *et al.* (2013) Extensive in vivo human milk peptidomics reveals specific proteolysis yielding protective antimicrobial peptides. *J. Proteome Res.*, **12**, 2295–2304.
142. Picariello,G., Ferranti,P., Mamone,G. *et al.* (2012) Gel-free shotgun proteomic analysis of human milk. *J. Chromatogr. A*, **1227**, 219–233.
143. Liao,Y., Alvarado,R., Phinney,B. *et al.* (2011) Proteomic characterization of human milk fat globule membrane proteins during a 12 month lactation period. *J. Proteome Res.*, **10**, 3530–3541.
144. Dayon,L., Macron,C., Lahrichi,S. *et al.* (2021) Proteomics of human milk: definition of a discovery workflow for clinical research studies. *J. Proteome Res.*, **20**, 2283–2290.
145. Goonatilleke,E., Huang,J., Xu,G. *et al.* (2019) Human milk proteins and their glycosylation exhibit quantitative dynamic variations during lactation. *J. Nutr.*, **149**, 1317–1325.
146. Zhou,Y.H., Le,Z., Yu,Z.B. *et al.* (2019) Peptidomic analysis reveals multiple protection of human breast milk on infants during different stages. *J. Cell. Physiol.*, **234**, 15510–15526.
147. Gan,J., Robinson,R.C., Wang,J. *et al.* (2018) Peptidomic profiling of human milk with LC-MS/MS reveals pH-specific proteolysis of milk proteins. *Food Chem.*, **274**, 766–774.
148. Balakrishnan,L., Nirujogi,R.S., Ahmad,S. *et al.* (2014) Proteomic analysis of human osteoarthritis synovial fluid. *Clin. Proteomics*, **11**, 6.
149. Balakrishnan,L., Bhattacharjee,M., Ahmad,S. *et al.* (2014) Differential proteomic analysis of synovial fluid from rheumatoid arthritis and osteoarthritis patients. *Clin. Proteomics*, **11**, 1.
150. Rydholm,U. (2016) Synovial fluid proteome in rheumatoid arthritis. *Acta Orthop.*, **77**, 1–11.
151. Mahendran,S.M., Keystone,E.C., Krawetz,R.J. *et al.* (2019) Elucidating the endogenous synovial fluid proteome and peptidome of inflammatory arthritis using label-free mass spectrometry. *Clin. Proteomics*, **16**, 23.
152. Birkelund,S., Bennike,T.B., Kastaniegaard,K. *et al.* (2020) Proteomic analysis of synovial fluid from rheumatic arthritis and spondyloarthritis patients. *Clin. Proteomics*, **17**, 29.
153. Foers,A.D., Dagley,L.F., Chatfield,S. *et al.* (2020) Proteomic analysis of extracellular vesicles reveals an immunogenic cargo in rheumatoid arthritis synovial fluid. *Clin. Transl. Immunol.*, **9**, e1185.
154. Lee,J.H., Jung,J.H., Kim,J. *et al.* (2020) Proteomic analysis of human synovial fluid reveals potential diagnostic biomarkers for ankylosing spondylitis. *Clin. Proteomics*, **17**, 20.
155. Brunoro,G.V., Carvalho,P.C., Ferreira,A.T. *et al.* (2015) Proteomic profiling of nipple aspirate fluid (NAF): exploring the complementarity of different peptide fractionation strategies. *J. Proteomics*, **117**, 86–94.
156. Alexander,H., Stegner,A.L., Wagner-Mann,C. *et al.* (2004) Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin. Cancer Res.*, **10**, 7500–7510.
157. Kurono,S., Kaneko,Y., Matsuura,N. *et al.* (2016) Identification of potential breast cancer markers in nipple discharge by protein profile analysis using two-dimensional nano-liquid chromatography/nanoelectrospray ionization-mass spectrometry. *Proteomics Clin. Appl.*, **10**, 605–613.
158. Shaheed,S.U., Tait,C., Kyriacou,K. *et al.* (2017) Nipple aspirate fluid - a liquid biopsy for diagnosing breast health. *Proteomics Clin. Appl.*, **11**, 1700015.
159. Pavlou,M.P., Kulasingam,V., Sauter,E.R. *et al.* (2010) Nipple aspirate fluid proteome of healthy females and patients with breast cancer. *Clin. Chem.*, **56**, 848–855.
160. Kim,Y.E., Kim,K., Han,B.O. *et al.* (2021) Quantitative proteomic profiling of Cervicovaginal fluid from pregnant women with term and preterm birth. *Proteome Sci.*, **19**, 3.
161. Muytjens,C., Yu,Y. and Diamandis,E.P. (2017) Discovery of antimicrobial peptides in cervical-vaginal fluid from healthy non-pregnant women via an integrated proteome and peptidome analysis. *Proteomics*, **17**, 1600461.
162. Federi,C.C., Valerie,W., Silvia,T. *et al.* (2013) Proteome profiles of vaginal fluids from women affected by bacterial vaginosis and healthy controls: outcomes of rifaximin treatment. *J. Antimicrob. Chemother.*, **68**, 2648–2659.
163. Starodubtseva,N.L., Brzhozovzkiy,A.G., Bugrova,A.E. *et al.* (2019) Label-free cervicovaginal fluid proteome profiling reflects the cervix neoplastic transformation. *J. Mass Spectrom.*, **54**, 693–703.
164. Hosako,M., Muto,T., Nakamura,Y. *et al.* (2012) Proteomic study of malignant pleural mesothelioma by laser microdissection and two-dimensional difference gel electrophoresis identified cathepsin D as a novel candidate for a differential diagnosis biomarker. *J. Proteomics*, **75**, 833–844.
165. Mundt,F., Johansson,H.J., Forshed,J. *et al.* (2013) Proteome screening of pleural effusions identifies galectin 1 as a diagnostic

biomarker and highlights several prognostic biomarkers for malignant mesothelioma. *Mol. Cell. Proteomics*, **13**, 701–715.
166. Park,J.O., Choi,D.Y., Choi,D.S. *et al.* (2013) Identification and characterization of proteins isolated from microvesicles derived from human lung cancer pleural effusions. *Proteomics*, **13**, 2125–2134.
167. Burg,D., Schofield,J.P.R., Brandsma,J. *et al.* (2018) Large-scale label-free quantitative mapping of the sputum proteome. *J. Proteome Res.*, **17**, 2072–2091.
168. Hailemariam,M., Yu,Y.B., Singh,H. *et al.* (2020) Protein and microbial biomarkers in sputum discern acute and latent tuberculosis in investigation of pastoral Ethiopian cohort. *Front. Cell. Infect. Microbiol.*, **11**, 595554.
169. Gray,R.D., Macgregor,G., Noble,D. *et al.* (2008) Sputum proteomics in inflammatory and suppurative respiratory diseases. *Am. J. Respir. Crit. Care Med.*, **178**, 444–452.
170. Muccilli,V., Saletti,R., Cunsolo,V. *et al.* (2015) Protein profile of exhaled breath condensate determined by high resolution mass spectrometry. *J. Pharm. Biomed. Anal.*, **105**, 134–149.
171. Hayes,S.A., Haefliger,S., Harris,B. *et al.* (2016) Exhaled breath condensate for lung cancer protein analysis: a review of methods and biomarkers. *J. Breath Res.*, **10**, 034001.
172. Lacombe,M., Marie-Desvergne,C., Combes,F. *et al.* (2018) Proteomic characterization of human exhaled breath condensate. *J. Breath Res.*, **12**, 021001.
173. Chen,D.P., Bryden,W.A. and Mcloughlin,M. (2020) A novel system for the comprehensive collection of nonvolatile molecules from human exhaled breath. *J. Breath Res.*, **15**, 016001.
174. Gade,I.L., Schultz,J.G., Cehofski,L.J. *et al.* (2020) Exhaled breath condensate in acute pulmonary embolism; a porcine study of effect of condensing temperature and feasibility of protein analysis by mass spectrometry. *J. Breath Res.*, **15**, 026005.
175. Paulo,J.A., Lee,L.S., Banks,P.A. *et al.* (2011) Difference gel electrophoresis identifies differentially expressed proteins in endoscopically collected pancreatic fluid. *Electrophoresis*, **32**, 1939–1951.
176. Paulo,J.A., Lee,L.S., Wu,B. *et al.* (2010) Identification of pancreas-specific proteins in endoscopic (ePFT) collected pancreatic fluid with mass spectrometry (GeLC-MS/MS). *Pancreas*, **39**, 889.
177. Marchegiani,G., Paulo,J.A., Sahora,K. *et al.* (2015) The proteome of postsurgical pancreatic juice. *Pancreas*, **44**, 574–582.
178. Paulo,J.A., Kadiyala,V., Gaun,A. *et al.* (2013) Analysis of endoscopic pancreatic function test (ePFT)-collected pancreatic fluid proteins precipitated via ultracentrifugation. *J. Pancreas*, **14**, 176–186.
179. Csősz,É., Emri,G., Kalló,G. *et al.* (2015) Highly abundant defense proteins in human sweat as revealed by targeted proteomics and label-free quantification mass spectrometry. *J. Eur. Acad. Dermatol. Venereol.*, **29**, 2024–2031.
180. Raiszadeh,M.M., Ross,M.M., Russo,P.S. *et al.* (2012) Proteomic analysis of eccrine sweat: implications for the discovery of schizophrenia biomarker proteins. *J. Proteome Res.*, **11**, 2127–2139.
181. Yu,Y.J., Prassas,I., Muytjens,C. *et al.* (2017) Proteomic and peptidomic analysis of human sweat with emphasis on proteolysis. *J. Proteomics*, **155**, 40–48.
182. Johnson,M., Zaretskaya,I., Raytselis,Y. *et al.* (2008) NCBI BLAST: a better web interface. *Nucleic Acids Res.*, **36**, W5–W9.
183. Maere,S., Heymans,K. and Kuiper,M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **2**, 3448–3449.
184. Shannon,P., Markiel,A., Ozier,O. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
185. Cui,J., Liu,Q., Puett,D. *et al.* (2008) Computational prediction of human proteins that can be secreted into the bloodstream. *Bioinformatics*, **24**, 2370–2375.
186. Hong,C.S., Cui,J., Ni,Z.H. *et al.* (2011) A computational method for prediction of excretory proteins and application to identification of gastric cancer markers in urine. *PLoS One*, **6**, e16875.
187. Hu,L.L., Huang,T., Cai,Y.D. *et al.* (2011) Prediction of body fluids where proteins are secreted into based on protein interaction network. *PLoS One*, **6**, e22989.
188. Liu,Q., Cui,J., Yang,Q. *et al.* (2010) In-silico prediction of blood-secretory human proteins using a ranking algorithm. *BMC Bioinform.*, **11**, 250.
189. Sun,Y., Du,W., Zhou,C. *et al.* (2015) A computational method for prediction of saliva-secretory proteins and its application to identification of head and neck cancer biomarkers for salivary diagnosis. *IEEE Trans. Nanobiosci.*, **14**, 167–174.
190. Wang,J.X., Liang,Y.C., Wang,Y. *et al.* (2013) Computational prediction of human salivary proteins from blood circulation and application to diagnostic biomarker identification. *PLoS One*, **8**, e80211.
191. Wang,Y., Du,W., Liang,Y.C. *et al.* (2016) PUEPro: a computational pipeline for prediction of urine excretory proteins. In: *2016 Advanced Data Mining and Applications (ADMA)*. Gold Coast, QLD, Australia.