Data in Brief 3 (2015) 103–107

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Data for a comprehensive map and functional annotation of the human cerebrospinal fluid proteome

Yang Zhang [a,1], Zhengguang Guo [b,1], Lili Zou [b], Yehong Yang [b], Liwei Zhang [a], Nan Ji [a], Chen Shao [c], Yajie Wang [d,e,*], Wei Sun [b,**]

[a] Department of Neurosurgery/China National Clinical Research Center for Neurological Diseases, Beijing Tiantan Hospital, Capital Medical University, 6 Tian Tan Xi Li, Beijing 100050, China
[b] Core Facility of Instrument, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, 5 Dong Dan San Tiao, Beijing 100005, China
[c] National Key Laboratory of Medical Molecular Biology, Department of Physiology and Pathophysiology, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, 5 Dong Dan San Tiao, Beijing 100005, China
[d] Core Laboratory for Clinical Medical Research, Beijing Tiantan Hospital, Capital Medical University, 6 Tian Tan Xi Li, Beijing 100050, China
[e] Department of Clinical Laboratory Diagnosis, Beijing Tiantan Hospital, Capital Medical University, 6 Tian Tan Xi Li, Beijing 100050, China

ARTICLE INFO

Article history:
Received 27 January 2015
Received in revised form
3 February 2015
Accepted 9 February 2015
Available online 20 February 2015

ABSTRACT

Knowledge about the normal human cerebrospinal fluid (CSF) proteome serves as a baseline reference for CSF biomarker discovery and provides insight into CSF physiology. In this study, high-pH reverse-phase liquid chromatography (hp-RPLC) was first integrated with a TripleTOF 5600 mass spectrometer to comprehensively profile the normal CSF proteome. A total of 49,836 unique peptides and 3256 non-redundant proteins were identified. To obtain high-confidence results, 2513 proteins with at least 2 unique peptides were further selected as bona fide CSF proteins. Nearly 30% of the identified CSF proteins have not been previously reported in the normal CSF proteome. More than 25% of the CSF proteins were components of CNS cell microenvironments, and

DOI of original article: http://dx.doi.org/10.1016/j.jprot.2015.01.017
* Corresponding author at: Core Laboratory for Clinical Medical Research, Beijing Tiantan Hospital, Capital Medical University, 6 Tian Tan Xi Li, Beijing 100050, China.
** Corresponding author.
E-mail addresses: zhangyang8025@yeah.net (Y. Zhang), gzg0625@sina.com (Z. Guo), tiantanwyj@aliyun.com (Y. Wang), sunwei1018@hotmail.com (W. Sun).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.dib.2015.02.004

network analyses indicated their roles in the pathogenesis of neurological diseases. The top canonical pathway in which the CSF proteins participated was axon guidance signaling. More than one-third of the CSF proteins (788 proteins) were related to neurological diseases, and these proteins constitute potential CSF biomarker candidates. The mapping results can be freely downloaded at http://122.70.220.102:8088/csf/, which can be used to navigate the CSF proteome. For more information about the data, please refer to the related original article [1], which has been recently accepted by Journal of Proteomics.

Subject area	*Biology*
More specific subject area	*Proteomics*
Type of data	*Tables*
How data was acquired	*Instruments including Waters Acquity nano-UPLC system, AB SCIEX Triple TOF* 5600 *system*
Data format	*Analyzed*
Experimental factors	*Protein samples were reduced with* 10 mM *DTT, alkylated with* 55 mM *iodoacetamide, digested using sequencing-grade modified trypsin.*
Experimental features	*CSF pooled from* 14 *individuals* (7 *women and* 7 *men*) *was subjected to the depletion of* 14 *high-abundance proteins with an immunoaffinity column. The flow-through proteins, bound proteins, and original proteins were collected separately, digested, and then separated into* 30 *fractions each by high-pH RPLC. Total* 90 *fractions were subjected to nano-RPLC-MS/MS analysis.*
Data source location	*Beijing, China*
Data accessibility	*The data in the data in brief, in the related original article* [1] *and at* http://122.70.220.102:8088/csf/ *can be freely downloaded.*

Value of the data

- This study identified the largest high-confidence dataset of the human CSF proteome.
- Some CSF proteins' abundances are quantified by the iBAQ method.
- High proportion of the CSF proteins is microenvironment components of CNS cells.
- High proportion of the CSF proteins participate in neurocyte connectivity.
- A large part of CSF proteins are biomarker candidates of neurological diseases.

1. Data, experimental design, materials and methods

1.1. Data

Table 1 lists all the CSF proteins with at least 2 unique peptide identifications. Table 2 lists the proteins and their abundances, which were quantified by the iBAQ [2] method. Table 3 lists the CSF proteins that participate in the axon guidance signaling pathway. Table 4 lists the CSF proteins involved in neurological diseases.