

RESEARCH HIGHLIGHT

# Biomarker discovery in human cerebrospinal fluid: the need for integrative metabolome and proteome databases

Emanuel Schwarz<sup>1</sup>, E Fuller Torrey<sup>2</sup>, Paul C Guest<sup>1\*</sup> and Sabine Bahn<sup>1,3\*</sup>

## Abstract

The number of metabolites identified in human cerebrospinal fluid (CSF) has steadily increased over the past 5 years, and in this issue of *Genome Medicine* David Wishart and colleagues provide a comprehensive update that brings the number of metabolites listed in the CSF metabolome database to 476 compounds. There is now a need for an integrative metabolome-proteome CSF database to maximize the impact of this achievement in biomedical research. Only by such efforts can we hope to unravel the complexity of molecular pathophysiological processes.

## The human cerebrospinal fluid metabolomic database

Cerebrospinal fluid (CSF) is considered to be an excellent source of molecular biomarkers of neurological function since it is in direct contact with the extracellular space of the brain. As the CSF contains metabolites and other molecules that can reflect altered brain function, systematic studies of the CSF metabolome can potentially offer new insights into central nervous system (CNS) disorders. These include neurodegenerative conditions such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis and multiple sclerosis, and neuropsychiatric diseases such as schizophrenia, depression and bipolar disorder. At the time it was first published in 2007, the CSF metabolome database [1] contained 308 detectable metabolites, with information on names, structures, concentrations, literature references and disease associations. Since this time, the database has been applied to the study of numerous CNS diseases.

## Technological advances expand the spectrum of CSF metabolites

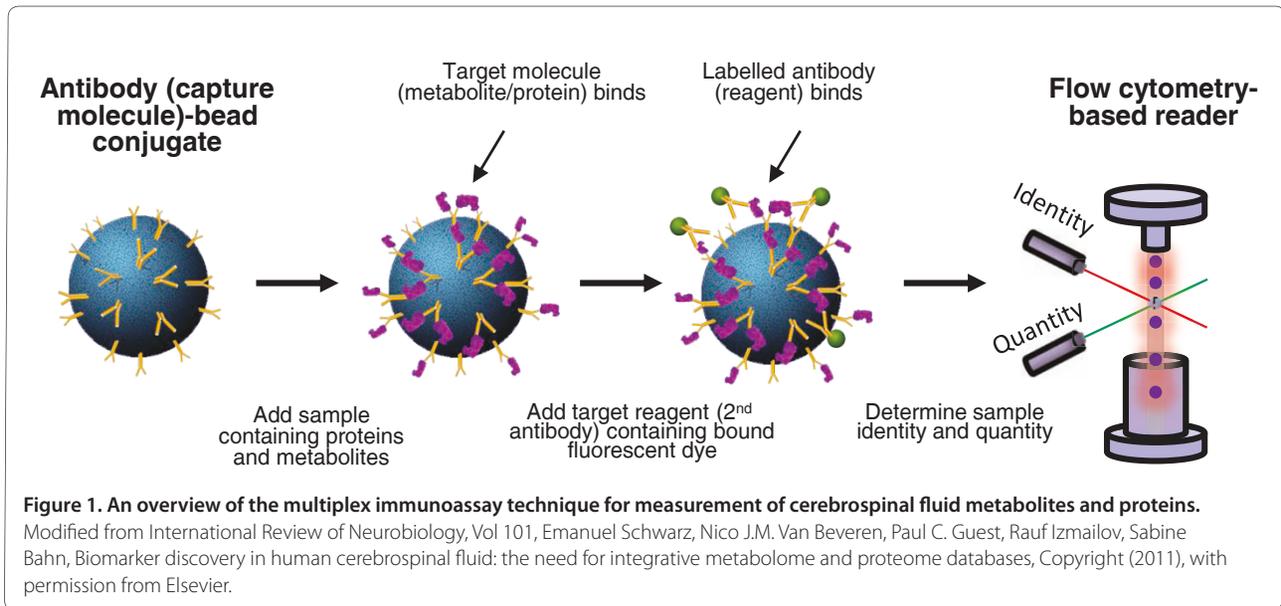
A new study from David Wishart and co-workers, published in *Genome Medicine*, represents an important work in the field of biomarker discovery [2]. The authors describe how continuing advances in analytical technologies have led to the discovery of many new CSF metabolites, allowing expansion of the CSF metabolome database. They applied five different metabolomic platforms to characterize multiple CSF samples, and this resulted in identification of new constituent metabolites and an increase in size of the database by approximately 50%. The authors now give an updated CSF metabolome database that contains detailed information of 476 compounds. Furthermore, the authors have carried out an extensive literature review for additional information on these compounds, including their concentrations and disease associations. The newly identified molecules consist of a number of metabolites (6 acylcarnitines, 13 amino acids, hexose, 42 phosphatidylcholines, 2 lysophosphatidylcholines and 14 sphingolipids) that should aid neurological studies involving changes in energy metabolism. In addition, Wishart *et al.* also identified 37 metal ions, which could be useful in studies of neurodegenerative diseases such as Alzheimer's disease, since alterations in metal ions are known to occur in this condition [3].

## Integrating metabolome and proteome profiles of CSF

Despite the comprehensiveness of this updated CSF metabolome database, there is still a need for continued developments, including integration with a CSF proteomics database. The most extensive CSF proteome characterized thus far contains 2,630 proteins [4]. However, integration of these databases will require the application of sophisticated bioinformatic approaches. Also, studies involving changes in CSF metabolites and proteins may require analyses using a single platform to rule out artifacts found as a result of cross-platform comparisons. This is important considering that proteins, metabolites

\*Correspondence: pg110@cam.ac.uk; sb209@cam.ac.uk

<sup>1</sup>Institute of Biotechnology, University of Cambridge, Cambridge CB2 1QT, UK  
Full list of author information is available at the end of the article



and other molecules are interactive as components of the same biological networks and this is manifested at multiple levels in a systems biology manner. For example, protein hormones, including insulin and VGF, regulate the levels of metabolites such as glucose, adrenocorticotropic hormone regulates production and release of the steroid hormone cortisol, and growth factor proteins such as platelet-derived growth factor can affect the action of dopaminergic and serotonergic neurotransmitters. As biomedical research evolves from traditional clinical and biological investigations to incorporate multi-omic technologies, integration of the resulting data has emerged as a critical next stage.

In line with this, several groups, including our own, have investigated both metabolic and proteomic profiles of CSF samples in normal and disease states. For example, Blanchet and colleagues [5] have shown that a fusion of proteome and metabolome data leads to higher predictive accuracy in a rat model of autoimmune encephalomyelitis. Interestingly, the molecular signature comprises transporter proteins such as hemopexin, serum albumin and serotransferrin, along with metabolites such as succinate, pantothenate and glutamine, which are involved in multiple complex biological pathways. An alternative example comes from our own group, by identification of a metabolome profile comprising glucose, acetate, alanine and glutamine that could distinguish patients with schizophrenia from control patients, with high accuracy [6]. As a complement to this metabolome study, we also profiled proteins and peptides in the same CSF samples [7]. The key alterations identified were increased levels of a VGF-derived peptide and decreased levels of the transport protein transthyretin. The two studies together provide a more complete

picture of the changes seen in patients with schizophrenia as VGF is known to affect the levels of energy-related metabolites such as glucose, and transthyretin directly interacts with and is responsible for the transport of metabolites such as the thyroid hormone thyroxine ( $T_4$ ) and the vitamin-A-related molecule retinol.

As an extension, we are now profiling CSF from patients with schizophrenia and controls using multiplex immunoassay profiling to simultaneously measure multiple proteins and metabolites (Figure 1). This multiplex platform will allow researchers to target a broader combination of metabolites. As all of these proteins interact with small molecules to exert their functions, it is clear that integration of 'omic' platforms is required for a better understanding of disease processes.

### Future perspectives in CSF metabolomics and beyond

Wishart and colleagues have made a significant contribution to our ability to unravel disease processes through their use of multiple platforms to increase the size of the metabolome database [2]. However, future profiling studies will require incorporation of assays for proteins and metabolites into single platforms to identify components that are altered in disease. This is because cross-platform comparisons are useful for cataloguing components but not for reliable quantitative studies. It is clear from the metabolome and proteome studies described above that small molecules and proteins are highly interactive in bringing about physiological effects in complex biological systems. Therefore, a more complete understanding of diseases and other biological effects requires a massive integration of technologies and statistical methods. However, there is now reason for optimism

that further technological and interdisciplinary advancements will overcome current limitations in the field to help usher biomarkers fully into the 21st century. This can be achieved either by integrating the analyses using a single platform, such as the multiplex immunoassay method described above, or by using sophisticated bioinformatic and biostatistical methods to integrate the metabolite and proteomic data acquired using different platforms.

#### Abbreviations

CNS, central nervous system; CSF, cerebrospinal fluid.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Institute of Biotechnology, University of Cambridge, Cambridge CB2 1QT, UK. <sup>2</sup>The Stanley Medical Research Institute, 8401 Connecticut Avenue Suite 200, Chevy Chase, MD 20815, USA. <sup>3</sup>Department of Neuroscience, Erasmus University Medical Centre, 3000 CA, Rotterdam, the Netherlands

Published: 30 April 2012

#### References

1. HMDB CSF Metabolome [<http://www.csfmetabolome.ca>]
2. Mandal R, Guo AC, Chaudhury K, Liu P, Yallou FS, Dong E, Aziat F, Wishart DS: An update on the human cerebrospinal fluid metabolome. *Genome Med* 2012, **4**:38.

3. Roberts BR, Ryan TM, Bush AI, Masters CL, Duce JA: The role of metallobiology and amyloid- $\beta$  peptides in Alzheimer's disease. *J Neurochem* 2012, **120**(Suppl 1):149-166.
4. Schutzer SE, Liu T, Natelson BH, Angel TE, Schepmoes AA, Purvine SO, Hixson KK, Lipton MS, Camp DG, Coyle PK, Smith RD, Bergquist J: Establishing the proteome of normal human cerebrospinal fluid. *PLoS One* 2010, **5**:e10980.
5. Blanchet L, Smolinska A, Attali A, Stoop MP, Ampt KA, van Aken H, Suidgeest E, Tuinstra T, Wijmenga SS, Luidert T, Buydens LM: Fusion of metabolomics and proteomics data for biomarkers discovery: case study on the experimental autoimmune encephalomyelitis. *BMC Bioinformatics* 2011, **12**:254.
6. Holmes E, Tsang TM, Huang JT, Leweke FM, Koethe D, Gerth CW, Nolden BM, Gross S, Schreiber D, Nicholson JK, Bahn S: Metabolic profiling of CSF: evidence that early intervention may impact on disease progression and outcome in schizophrenia. *PLoS Med* 2006, **3**:e327.
7. Huang JT, Leweke FM, Oxley D, Wang L, Harris N, Koethe D, Gerth CW, Nolden BM, Gross S, Schreiber D, Reed B, Bahn S: Disease biomarkers in cerebrospinal fluid of patients with first-onset psychosis. *PLoS Med* 2006, **3**:e428.

doi:10.1186/gm338

Cite this article as: Schwarz E, *et al.*: Biomarker discovery in human cerebrospinal fluid: the need for integrative metabolome and proteome databases. *Genome Medicine* 2012, **4**:39.