



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

LC-HRMS data as a result of untargeted metabolomic profiling of human cerebrospinal fluid



Florence Mehl^a, Héctor Gallart-Ayala^a, Ioana Konz^a,
 Tony Teav^a, Aikaterini Oikonomidi^b, Gwendoline Peyratout^b,
 Vera van der Velpen^{a,b}, Julius Popp^{b,c,1}, Julijana Ivanisevic^{a,1,*}

^a *Metabolomics Platform, Faculty of Biology and Medicine, University of Lausanne, Switzerland*

^b *Old Age Psychiatry, Department of Psychiatry, Lausanne University Hospital, Switzerland*

^c *Geriatric Psychiatry, Department of Mental Health and Psychiatry, Geneva University Hospitals, Switzerland*

ARTICLE INFO

Article history:

Received 18 April 2018

Received in revised form

16 October 2018

Accepted 23 October 2018

Available online 27 October 2018

ABSTRACT

Cerebrospinal fluid (CSF) is a key body fluid that maintains the homeostasis in central nervous system (CNS). As a biofluid whose content reflects the brain metabolic activity, the CSF is analyzed in the context of neurological diseases and is rarely collected from healthy subjects. For this reason, the metabolite variation associated with general phenotypic characteristics such as gender and age have hardly ever been studied. Here we present the hydrophilic interaction liquid chromatography-high resolution mass spectrometry (HILIC-HRMS) data as a result of untargeted metabolomics analysis of a cohort of elderly cognitively healthy volunteers ($n = 32$). 146 unambiguously identified water soluble metabolites (using accurate mass, retention time and MS/MS matching against spectral libraries) were measured and their abundances across all the subjects depending on their gender are provided in this article. Data tables are available at <https://data.mendeley.com/datasets/c73xtsd4s5/1>. It's published on mendeley, the DOI is DOI:10.17632/c73xtsd4s5.1. The data presented in this article are related to the research article entitled "A global HILIC-MS approach to measure polar human cerebrospinal fluid metabolome: Exploring

DOI of original article: <https://doi.org/10.1016/j.aca.2018.04.002>

* Corresponding author.

E-mail addresses: florence.mehl@unil.ch (F. Mehl), julijana.ivanisevic@unil.ch (J. Ivanisevic).

¹ Equal contribution.

<https://doi.org/10.1016/j.dib.2018.10.113>

2352-3409/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

gender-associated variation in a cohort of elderly cognitively healthy subjects” (Gallart-Ayala et al., 2018, In press).

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications table

Subject area	Clinical study
More specific subject area	Metabolomics
Type of data	XCMS preprocessed and signal drift corrected LC–MS data
How data was acquired	Global-untargeted metabolite profiling was performed using HILIC coupled to high-resolution mass spectrometry (HRMS) operating in both positive and negative (ESI+ and ESI-) mode on 6550 iFunnel Q-TOF mass spectrometer interfaced with 1290 UHPLC system (Agilent Technologies). Data were acquired in the full <i>m/z</i> range 50–1200 in both modes.
Data format	Tables in.txt format
Experimental factors	Cerebrospinal fluid samples were collected using lumbar puncture from cognitively healthy, elderly volunteers ($n = 32$, with no history or symptoms of relevant psychiatric or neurologic disease, no cognitive impairment, and without cerebral Alzheimer’s disease pathology as indicated by CSF biomarkers). Clinical characteristics of the cohort are provided below in Experimental design section.
Experimental features	Cerebrospinal fluid samples were extracted by the addition of ice-cold MeOH:ACN (1:1). CSF extracts were analyzed with ESI-LS-MS system and processed using ProFinder (Agilent Technologies) and XCMS software. For the quality control (QC), pooled QC samples (representative of the entire sample set) were analyzed periodically (every 7 samples) throughout the run.
Data source location	Metabolomics Platform, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland
Data accessibility	Data tables are available at https://data.mendeley.com/datasets/c73xtsd4s5/1 . it’s published on mendeley, the DOI is DOI:10.17632/c73xtsd4s5.1

Value of the data

- The dataset consists of the LC-HRMS metabolomics data as a result of untargeted analyses of cerebrospinal fluid samples from a cohort of elderly cognitively healthy human subjects ($n = 32$).
- In this paper [1], metabolites were unambiguously identified by AMRT in-house database matching and MS/MS spectra matching against publically available spectral libraries (METLIN & mzCloud).
- The identified metabolites ($n = 146$) are implicated in multiple metabolic pathways including glycolysis and TCA cycle, purine and pyrimidine, amino acid and fatty acid metabolism.
- The acquired data allowed for the exploration inter-individual variability and the discovery of gender-associated differences.

1. Data

Data tables are available at <https://data.mendeley.com/datasets/c73xtsd4s5/1>. it’s published on mendeley, the DOI is DOI:10.17632/c73xtsd4s5.1. The dataset contains the XCMS processed and signal

drift corrected data. Data table is reported in.txt format and contains the information (i.e. peak areas) about all detected metabolite features across analyzed samples and pooled quality control sample. Metabolite features are assigned by the couple of specific RT and accurate m/z ratio information and annotated with CAMERA for isotopes. Peaks were detected and integrated using XCMS software (parameters are specified in Materials and Methods). Metabolites abundances were corrected for signal drift effect by fitting a locally quadratic (loess) regression model to the QC values (as described in Materials and Methods below). A table of all identified metabolites derived from this dataset is available as supplementary table in the research article entitled “A global HILIC-MS approach to measure polar human cerebrospinal fluid metabolome: Exploring gender-associated variation in a cohort of elderly cognitively healthy subjects” [1]. These data were log transformed prior to univariate statistical analysis to explore gender-associated differences.

2. Experimental design, materials and methods

2.1. Study population and CSF sample collection

Cerebrospinal fluid samples were collected using lumbar puncture from cognitively healthy (with no history or symptoms of relevant psychiatric or neurologic disease, and no cognitive impairment) volunteers ($n = 32$) at the Department of Psychiatry and the Department of Clinical Neurosciences, University Hospital in Lausanne, Switzerland (CHUV). The participants have been selected among cognitively healthy, community dwelling volunteers participating on an observational cohort study on cognitive aging and Alzheimer’s disease (AD) [2]. They were recruited by announcements and word of mouth. All participants had a comprehensive medical, neuropsychological, and psychosocial evaluation to exclude cognitive impairment, as well as brain MRI or CT scans, and venous and lumbar punctures. As a significant percentage of elderly people with normal cognition may have cerebral AD pathology, only participants with a non-AD CSF biomarker profile have been included in this study [2]. The study was approved by the cantonal ethics committee (Vaud) and informed consent was obtained from all participants [3].

Lumbar punctures were performed between 8 and 10 a.m. after an overnight fast. For lumbar puncture, a standardized technique with a 22-gauge “atraumatic” spinal needle and the subject in a sitting or lying position was applied [4]. A volume of 10–12 ml of CSF was collected in polypropylene tubes. CSF samples were centrifuged, frozen in aliquots, and stored at -80°C before further use (Table 1).

2.2. CSF metabolite extraction

Cerebrospinal fluid samples (100 μL) were extracted by the addition of 400 μL of ice-cold MeOH:ACN (1:1, v/v) to maintain MeOH:ACN:H₂O (2:2:1, v/v) ratio. The samples were then vortexed for 30 s, incubated for 1 h at -20°C , and finally centrifuged at 13,000 rpm at 4°C for 15 min. The resulting supernatant was evaporated to dryness using a speedVac. The dry extracts were then reconstituted in

Table 1

Clinical characteristics of the cohort studied defining median and inter-quartile range.

Characteristic	
Gender (male:female)	12:20
Age (median, IQR), years	65.0 (62.5,70.5)
BMI (median, IQR)	23.8 (21.8,26.1)
Smokers (non:past:active)	12:16:4
Cholesterol (median, IQR), mmol L ⁻¹	6.1 (5.5,7.0) ^a
Glucose (median, IQR), mmol L ⁻¹	5.6 (5.1,6.0) ^a
Creatinine (median, IQR), mmol L ⁻¹	76.0 (61.0,89.3)

^a Data not available for 1 subject.

100 μL of $\text{H}_2\text{O}:\text{MeOH}:\text{ACN}$ (2:1:1, v/v), sonicated for 1 min and centrifuged for 10 min at 13,000 rpm at 4 °C to remove the insoluble debris. The supernatants were transferred to HPLC vials and stored at –80 °C prior to LC-MS analysis.

2.3. Untargeted LC–HRMS analysis

Global-untargeted metabolite profiling was performed by HILIC chromatography coupled to high-resolution mass spectrometry (HRMS) operating in both positive and negative (ESI+ and ESI-) mode on 6550 iFunnel Q-TOF mass spectrometer interfaced with 1290 UHPLC system (Agilent Technologies). Samples were analyzed using two chromatographic separations: i) BEH Amide, 1.7 μm , 100 mm \times 2.1 mm I.D. column (Waters, Massachusetts, US) in positive ionization mode and, ii) SeQuant[®] ZIC-pHILIC, 5 μm , 100 mm \times 2.1 mm I.D. column (Merck, Darmstadt, Germany) with a SeQuant[®] ZIC-pHILIC, 5 μm , 20 mm \times 2.1 mm I.D. guard column (Merck, Darmstadt, Germany) in negative ionization mode. The column temperature was maintained constant at 25 °C and 30 °C in positive and negative ionization mode, respectively. The mobile phase was composed of A = 20 mM ammonium formate and 0.1% formic acid in water (pH 3.8) and B = 0.1% formic acid in 100% ACN for positive mode and A = 20 mM ammonium acetate and 20 mM ammonium hydroxide in water (pH 9.3) and B = 100% ACN for negative mode. In positive mode, the linear elution gradient from 95% B (0–1.5 min) to 45% B (17–19 min) was applied. The initial gradient conditions were restored within one minute and a 5-min post-run re-equilibration was applied to maintain the system reproducibility. In negative mode, the linear step-wise elution gradient from 90% B (0–1.5 min) to 50% B (8–11 min) to 45% B (12–15 min) was applied. The initial gradient conditions were restored within one minute and a 9-min post-run equilibration was applied to maintain the system reproducibility. The flow rates were 400 $\mu\text{L min}^{-1}$ and 300 $\mu\text{L min}^{-1}$ in positive and negative ionization mode, respectively. In both cases, the sample injection volume was 2 μL . ESI source conditions were set as follows: dry gas temperature 290 °C and flow 14 L min^{-1} , fragmentor voltage 380 V, sheath gas temperature 350 °C and flow 12 L min^{-1} , nozzle voltage 0 V, and capillary voltage +2000 V in positive mode and –2000 V in negative. The instrument was set to acquire over the full m/z range 50–1200 in both modes, with the MS acquisition rate of 2 spectra/s. In addition, targeted MS/MS analysis was performed using the inclusion list of ions of interest, in narrow isolation width ($\sim 1.3 m/z$), with a MS acquisition rate of 500 ms and a MS/MS acquisition rate of 500 ms at a collision energy (CE) of 20 V. For the quality control (QC), pooled QC samples (representative of the entire sample set) were analyzed periodically (every 7 samples) throughout the overall analytical run to assess the analytical variability and correct for the potential signal intensity drift inherent to LC-MS technique [5].

2.4. Data processing

Raw LC-MS data were converted to mzXML files (using ProteoWizard MS Convert) and uploaded to XCMS plus software (stand-alone server hosted within bioinformatics facility at UNIL) for data processing including peak detection, retention time correction, profile alignment, and isotope annotation [6–8]. Data were processed as a multi-group experiment and the parameter settings were as follows: centWave algorithm for feature detection ($\Delta m/z = 20$ ppm, minimum peak width = 5 s and maximum peak width = 30 s, S/N threshold = 10, $mzdiff = 0.01$, integration method = 1, prefilter peaks = 3, prefilter intensity = 10000, noise filter = 5000); obiwrap settings for retention time correction (profStep = 1); and parameters for chromatogram alignment, including $mzwid = 0.015$, minfrac = 0.5 and bw = 5 [9].

2.5. Signal intensity drift correction and statistical analysis

The data table containing the abundances or peak areas of all detected and identified metabolites across all samples (CSF extracts from 32 individuals) was imported to Workflow4Metabolomics [10,11] where the peak intensities were corrected for signal drift effect by fitting a locally quadratic (loess) regression model to the QC values [5,12]. The α parameter controlling the smoothing was set to 1 to avoid over fitting [11].

Acknowledgements

The present study was supported by the Fondation Pierre Mercier pour la Science and the Swiss National Science Foundation, grant: 320030L_141179.

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at <https://doi.org/10.1016/j.dib.2018.10.113>.

References

- [1] H. Gallart-Ayala, I. Konz, F. Mehl, T. Teav, A. Oikonomidi, G. Peyratout, V. van der Velpen, J. Popp, J. Ivanisevic, A global HILIC-MS approach to measure polar human cerebrospinal fluid metabolome: exploring gender-associated variation in a cohort of elderly cognitively healthy subjects, *Anal. Chim. Acta* 1037 (2018) 327–337.
- [2] J. Popp, A. Oikonomidi, D. Tautvydaite, L. Dayon, M. Bacher, E. Migliavacca, H. Henry, R. Kirkland, I. Severin, J. Wojcik, G. L. Bowman, Markers of neuroinflammation associated with Alzheimer's disease pathology in older adults, *Brain Behav. Immunity* 62 (2017) 203–211.
- [3] J. Mathys, M. Gholamrezaee, H. Henry, A. von Gunten, J. Popp, Decreasing body mass index is associated with cerebrospinal fluid markers of Alzheimer's pathology in MCI and mild dementia, *Exp. Gerontol.* (2017).
- [4] J. Popp, M. Riad, K. Freymann, F. Jessen, Diagnostic lumbar puncture performed in the outpatient setting of a memory clinic. Frequency and risk factors of post-lumbar puncture headache, *Der Nervenarzt* 78 (2007) 547–551.
- [5] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, *Nat. Protoc.* 6 (2011) 1060–1083.
- [6] H. Gowda, J. Ivanisevic, C.H. Johnson, M.E. Kurczy, H.P. Benton, D. Rinehart, T. Nguyen, J. Ray, J. Kuehl, B. Arevalo, Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses, *Anal. Chem.* 86 (2014) 6931–6939.
- [7] C.A. Smith, E.J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, *Anal. Chem.* 78 (2006).
- [8] R. Tautenhahn, G.J. Patti, D. Rinehart, G. Siuzdak, XCMS Online: a web-based platform to process untargeted metabolomic data, *Anal. Chem.* 84 (2012) 5035–5039.
- [9] G. Libiseller, M. Dvorzak, U. Kleb, E. Gander, T. Eisenberg, F. Madeo, S. Neumann, G. Trausinger, F. Sinner, T. Pieber, C. Magnes, IPO: a tool for automated optimization of XCMS parameters, *BMC Bioinform.* 16 (2015) 118.
- [10] F. Giacomoni, G. Le Corguille, M. Monsoor, M. Landi, P. Pericard, M. Petera, C. Duperier, M. Tremblay-Franco, J.F. Martin, D. Jacob, S. Goultquer, E.A. Thevenot, C. Caron, Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics, *Bioinformatics (Oxf., Engl.)* 31 (2015) 1493–1495.
- [11] Y. Guitton, M. Tremblay-Franco, G. Le Corguille, J.F. Martin, M. Petera, P. Roger-Mele, A. Delabriere, S. Goultquer, M. Monsoor, C. Duperier, C. Canlet, R. Servien, P. Tardivel, C. Caron, F. Giacomoni, E.A. Thevenot, Create, run, share, publish, and reference your LC-MS, FIA-MS, GC-MS, and NMR data analysis workflows with the Workflow4Metabolomics 3.0 Galaxy online infrastructure for metabolomics, *Int. J. Biochem. Cell Biol.* (2017).
- [12] F.M. van der Kloet, I. Bobeldijk, E.R. Verheij, R.H. Jellema, Analytical error reduction using single point calibration for accurate and precise metabolomic phenotyping, *J. Proteome Res.* 8 (2009) 5132–5141.