

This is the **accepted version** of the following article: Pont, L., Benavente, F., Jaumot, J., Tauler, R., Alberch, J., Ginés, S., Barbosa, J. and Sanz-Nebot, V. (2016), Metabolic profiling for the identification of Huntington biomarkers by on-line solid-phase extraction capillary electrophoresis mass spectrometry combined with advanced data analysis tools. *ELECTROPHORESIS*, 37: 795–808. doi:10.1002/elps.201500378, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1002/elps.201500378/abstract>

Information of the Journal in which the present paper is published:

- WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Electrophoresis, 2016, 37, 795–808
- DOI 10.1002/elps.201500378

# **METABOLIC PROFILING FOR THE IDENTIFICATION OF HUNTINGTON BIOMARKERS BY ON-LINE SOLID-PHASE EXTRACTION CAPILLARY ELECTROPHORESIS MASS SPECTROMETRY COMBINED WITH ADVANCED DATA ANALYSIS TOOLS**

Laura Pont<sup>a</sup>, Fernando Benavente<sup>a\*</sup>, Joaquim Jaumot<sup>b</sup>, Romà Tauler<sup>b</sup>, Jordi Alberch<sup>c,d,e</sup>,

Silvia Ginés<sup>c,d,e</sup>, José Barbosa<sup>a</sup>, Victoria Sanz-Nebot<sup>a</sup>

<sup>a</sup>Departament de Química Analítica, Facultat de Química, Universitat de Barcelona, Barcelona, Spain.

<sup>b</sup>Department of Environmental Chemistry, IDAEA-CSIC, Jordi Girona 18-26, 08034 Barcelona, Spain.

<sup>c</sup>Departament de Biologia Cel·lular, Immunologia i Neurociències, Facultat de Medicina, Universitat de Barcelona, Spain.

<sup>d</sup>Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.

<sup>e</sup>Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain.

\*Corresponding author: [fbenavente@ub.edu](mailto:fbenavente@ub.edu) (F. Benavente, PhD)

Tel: (+34) 934039778, Fax: (+34) 934021233

**NONSTANDARD ABBREVIATIONS:** HD: Huntington's disease; MCR-ALS: multivariate curve resolution alternating least squares; PLS-DA: partial least squares discriminant analysis; SPE-CE-MS: solid-phase extraction capillary electrophoresis mass spectrometry; wt: wild type.

**KEYWORDS:** biomarker / Huntington / metabolomics / multivariate data analysis / SPE-CE-MS

**TOTAL WORDS:** 7620

## **ABSTRACT**

In this work, an untargeted metabolomic approach based on sensitive analysis by on-line solid-phase extraction capillary electrophoresis mass spectrometry (SPE-CE-MS) in combination with multivariate data analysis is proposed as an efficient method for the identification of biomarkers of Huntington's disease (HD) progression in plasma. For this purpose, plasma samples from wild type (wt) and HD (R6/1) mice of different ages (8, 12 and 30 weeks), were analysed by C<sub>18</sub>-SPE-CE-MS in order to obtain the characteristic electrophoretic profiles of low molecular mass compounds. Then, multivariate curve resolution alternating least squares (MCR-ALS) was applied to the multiple full scan MS data sets. This strategy permitted the resolution of a large number of metabolites being characterised by their electrophoretic peaks and their corresponding mass spectra. A total number of 29 compounds were relevant to discriminate between wt and HD plasma samples, as well as to follow-up the HD progression. The intracellular signalling was found to be the most affected metabolic pathway in HD mice after 12 weeks of birth, when mice already showed motor coordination deficiencies and cognitive decline. This fact agreed with the atrophy and dysfunction of specific neurons, loss of several types of receptors and changed expression of neurotransmitters.

## 1. Introduction

Huntington's disease (HD) is an inherited neurodegenerative disorder, which is characterised by progressive motor and cognitive disturbances. HD is caused by an expansion of the cytosine-adenine-guanine (CAG) repeat in the exon 1 of the huntingtin gene (*HTT*), which encodes a stretch of glutamines in the huntingtin protein [1–8]. Although the *HTT* gene is ubiquitously expressed as the huntingtin protein in most tissues, HD pathology has primarily been located to the basal ganglia and to the neocortex. The pathology involves atrophy and dysfunction of specific neurons, loss of several types of receptors, changed expression of neurotransmitters and key proteins, as well as formation of ubiquitin positive aggregates [1–8]. HD is a fatal disease, and the median interval between clinical diagnosis and death is typically given as 15 to 20 years [2,4,6,8].

By use of predictive genetic testing, it is possible to identify individuals who carry the *HTT* gene defect before the onset of symptoms, providing a unique window of opportunity for intervention aimed at preventing or delaying disease onset [4,7]. However, without robust and practical measures of disease progression, the efficacy of therapeutic interventions in this premanifest HD cannot be readily assessed. Neuroimaging and biochemical biomarkers are being investigated for their potential in clinical use and their value in the development of future treatments [4,7]. Modern neuroimaging techniques such as magnetic resonance imaging (MRI) enable high-quality images of brain structure and function to be obtained [9,10]. However, metabolites that can be quantified in biofluids, such as blood or urine, are appealing due to the improved selectivity, the minimal requirement for patient involvement,

opportunity for rapid bulk processing of specimens, availability of reliable assays and possibility of carrying out multiple analyses on a single sample [5,7,11].

Metabolomics aims to obtain a comprehensive coverage of low molecular mass compounds from biological systems [12–14]. Metabolomics studies can be approached using targeted or untargeted analysis [15–17]. In targeted analysis, a specified list of metabolites is analysed. In contrast, untargeted analysis requires comprehensive metabolite measurements. Furthermore, it can implicate previously unrecognised metabolites or pathways with a unique phenotype and, therefore, is a powerful platform to elucidate novel biomarkers and gain insight into disease pathogenesis.

Different techniques are currently used for untargeted metabolomics, including nuclear magnetic resonance (NMR), gas chromatography mass spectrometry (GC-MS), liquid chromatography mass spectrometry (LC-MS) and capillary electrophoresis mass spectrometry (CE-MS) [17–21]. For the first time to our knowledge, the use of on-line C<sub>18</sub> solid-phase extraction capillary electrophoresis mass spectrometry (C<sub>18</sub>-SPE-CE-MS) is proposed as an alternative sensitive method for metabolomic studies of biological fluids, which are complex diluted samples. CE is a versatile, high-performance separation technique with many desirable characteristics such as instrumental simplicity, full automation, high efficiency, low consumption of sample and reagents and reduced analysis times. However, like many other microanalytical techniques, it has poor concentration sensitivity for most analytes, from low molecular mass compounds to biopolymers such as proteins [22,23]. Several strategies have been proposed to improve CE sensitivity. Today, SPE-CE is becoming widely recognised as a powerful approach that overcomes this major drawback [22–27]. In SPE-CE, a

microcartridge placed inside and near the inlet of the separation capillary contains an appropriate extraction sorbent (in our case, C<sub>18</sub>). This sorbent selectively retains the target analyte, enabling large volumes of sample to be introduced ( $\approx$ 50-100  $\mu$ L). The retained analyte is eluted in a small volume of an appropriate solution ( $\approx$ 25-50 nL), which results in sample clean-up and concentration enhancement with minimum sample handling before separation and detection, for example, by on-line mass spectrometry (SPE-CE-MS) [22–27].

Chemometric methods play a crucial role in data processing, exploration and classification of the massive data sets generated in metabolomic studies [28–34]. If the goal of the study is the compound detection, the use of resolution methods such as multivariate curve resolution alternating least squares (MCR-ALS) can be an excellent alternative. MCR-ALS can resolve overlapped electrophoretic/chromatographic peaks from the collected data and provide the separation profiles and mass spectra of the constituents in the analysed samples. This approach allows overcoming problems such as retention time shifts, background noise contributions, and differences in signal-to-noise (S/N) ratios among different injections. Several published articles focus on the application of MCR-ALS to solve similar problems in LC-MS [32,33] and GC-MS [34]; but only a few studies have been previously reported combining CE-MS and MCR-ALS in metabolomic applications [35].

In this paper, we evaluate the capacity of SPE-CE-MS combined with advanced multivariate data analysis to preconcentrate, separate, detect and identify low molecular mass metabolites in plasma samples from wild type (wt) and HD (R6/1) mice of different ages (8, 12 and 30 weeks). A comparison between the different untargeted

metabolomic profiles allows us to propose novel potential biomarker candidates involved in the progression of Huntington's disease, which could be useful for prediction of disease onset or response to treatment.

## **2. Materials and methods**

### **2.1 Chemicals and reagents**

All the chemicals used in the preparation of buffers and solutions were of analytical reagent grade. Acetonitrile (99.9%), methanol (99.9%), 2-propanol ( $\geq 99.9\%$ ), formic acid (HFor) (99.0%), glacial acetic acid (HAc), ammonia (25%) and sodium hydroxide ( $\geq 99.0\%$ , pellets) were purchased from Merck (Darmstadt, Germany). Methionine enkephalin (Met,  $\geq 97.0\%$ ) and endomorphin 1 (End 1,  $\geq 98.0\%$ ) were provided by Sigma (St. Louis, MO, USA). Dynorphin A (1-7) (Dyn A,  $\geq 98.0\%$ ) was supplied by Bachem (Bubendorff, Switzerland). Polyethylene glycol (PEG) 8,000  $M_r$  (~50% in water) was purchased from Fluka (Buchs, Switzerland). Water with a conductivity value lower than 0.05  $\mu\text{S}/\text{cm}$  was obtained using a Milli-Q water purification system (Millipore, Molsheim, France).

### **2.2 Electrolyte solutions, sheath liquid and standard solutions**

Aqueous standard solutions (2500  $\mu\text{g}\cdot\text{mL}^{-1}$ ) of Dyn A, End 1 and Met peptides were prepared and stored in a freezer at  $-20^\circ\text{C}$  when not in use. A 10  $\text{ng}\cdot\text{mL}^{-1}$  standard mixture of the three peptides was prepared and analysed at the beginning and at the end of each SPE-CE-MS sequence, in order to check the proper functioning of the on-line SPE microcartridges. The background electrolyte (BGE) contained 50 mM of HAc and 50 mM of HFor and was adjusted to pH 3.50 with ammonia. The sheath liquid solution

consisted of a hydroorganic mixture of 60:40 v/v 2-propanol:water with 0.05% v/v of HFor. All solutions were passed through a 0.45  $\mu$ m nylon filter (MSI, Westboro, MA, USA) before analysis and were stored at 4°C when not in use. The sheath liquid was degassed for 10 min by sonication before use.

### **2.3 Mice blood plasma and sample preparation**

Plasma samples from male wild type mice (wt) and R6/1 transgenic mice (B6CBA background) expressing exon 1 of mutant huntingtin with 145 repeats (HD, R6/1) of different ages (8, 12 and 30 weeks, early, middle and late disease stage, respectively), were kindly supplied by the Department of Cellular Biology, Immunology and Neurosciences (Faculty of Medicine, University of Barcelona) [36]. Blood from mice was collected by cardiac puncture in standard clinical vials and placed on ice. Plasma was separated from the blood cells, pooled, deposited into polyethylene tubes and frozen at -20°C. It is worth mentioning that due to the small amount of blood that was possible to extract from a single mouse (between 1 and 2 mL), each set of samples corresponded to the combination of the plasma obtained from 4 or 5 mice. All animal procedures were approved by the CEEA committee of the University of Barcelona and were in accordance with the European Communities Council Directive (2010/63/EU).

The sample pretreatment used for the analysis of low molecular mass compounds in plasma samples was described elsewhere [22,37]. The off-line double step pretreatment of plasma samples consisted of protein precipitation with cold acetonitrile (plasma:acetonitrile, 200  $\mu$ L:1200  $\mu$ L) followed by centrifugal filtration with 10000  $M_r$



cut-off cellulose acetate filters (Amicon® Ultra-0.5, Millipore). Centrifugal filters were passivated before the first use with 5% v/v of PEG in water [37].

## **2.4 Apparatus and procedures**

pH measurements were made with a Crison 2002 potentiometer and a Crison electrode 52-03 (Crison Instruments, Barcelona, Spain). Centrifugal filtration was carried out in a cooled Rotanta 460 centrifuge (Hettich Zentrifugen, Tuttlingen, Germany) for centrifugation at controlled temperature (25°C).

### **2.4.1 On-line solid-phase extraction capillary electrophoresis mass spectrometry**

The construction of the microcartridge or analyte concentrator for C<sub>18</sub>-SPE-CE-MS was carried out as described elsewhere [22,37]. All fused silica capillaries were supplied by Polymicro Technologies (Phoenix, AZ, USA). The microcartridge (7 mm L<sub>T</sub> x 250 µm id x 360 µm od) was inserted inside the separation capillary (72 cm L<sub>T</sub> x 75 µm id x 360 µm od), at 7.5 cm from the inlet, using two plastic sleeves. Previously, it was filled with the sorbent found in C<sub>18</sub> Sep-pak cartridges (Waters, Milford, MA, USA). The sorbent particles were retained in the microcartridge between two frits (0.1 cm).

All capillary rinses were performed at high pressure (930 mbar). New separation capillaries were flushed with 1.0 M NaOH (20 min) and water (15 min) before inserting the microcartridge. This activation procedure was performed off-line to avoid the unnecessary entrance of NaOH into the MS system. Once inserted the microcartridge, the SPE-CE-MS capillaries were first conditioned by consecutive flushes of water (1

min), methanol (1 min), water (1 min) and BGE (3 min) at 930 mbar. Standard peptide mixture (Dyn A, End 1 and Met) or mice plasma samples were then introduced at 930 mbar for 10 min (approximately 60  $\mu\text{L}$  using the Hagen-Poiseuille equation [38]). A final rinse with the BGE (2 min at 930 mbar) eliminated non-retained molecules and equilibrated the capillary before the elution. Retained compounds were eluted by injecting a solution of 60:40 v/v methanol:water with 50 mM HAc and 50 mM HFor at 50 mbar for 10 s (approximately 50 nL [38]). Separation was carried out at 25°C by applying a voltage of 17 kV (normal polarity, cathode in the outlet). Between runs, the capillary was rinsed for 2 min with water and 2 min with acetonitrile, in order to avoid carry-over between consecutive analyses. In general, the different plasma samples (i.e. 8wt, 12wt, 30wt and 8HD, 12HD and 30HD) were analysed in triplicate (with the exception of 12wt, 12HD and 30HD, for which only two replicates were analysed due to the small volume of plasma sample available). Each series of replicate analyses was performed in a new SPE-CE-MS capillary due to the limited durability of the SPE microcartridges ( $\approx 10$  analyses) because of the complexity of the plasma matrix and the limited selectivity of the  $\text{C}_{18}$  sorbent. After these analyses, the extraction efficiency decreased and the microcartridge was packed until it was completely clogged [39]. At the beginning and at the end of each sequence, a 10  $\text{ng}\cdot\text{mL}^{-1}$  standard peptide mixture was analysed as a quality control of the system.

All CE-MS experiments were performed in an HP<sup>3D</sup> CE system coupled with an orthogonal G1603A sheath-flow interface to a 6220 oa-TOF LC/MS spectrometer (Agilent Technologies, Waldbronn, Germany). The sheath liquid was delivered at a flow rate of 3.3  $\mu\text{L}\cdot\text{min}^{-1}$  by a KD Scientific 100 series infusion pump (Holliston, MA, USA). ChemStation C.01.06 software (Agilent Technologies) was used for CE control

and separation data acquisition (e.g. voltage, temperature and current), and was run in combination with MassHunter B.04.00 workstation software (Agilent Technologies) for control of the mass spectrometer and MS data acquisition.

The TOF mass spectrometer was operated under optimum conditions in positive mode using the following parameters: capillary voltage 4000 V, drying gas temperature 200°C, drying gas flow rate 4 L·min<sup>-1</sup>, nebulizer gas 7 psig, fragmentor voltage 215 V, skimmer voltage 60 V, OCT 1 RF Vpp voltage 300 V. Data were collected in profile at 1 spectrum/s between 40 and 1250 m/z, with the mass range set to high resolution mode (4 GHz). A standard tune and an external mass calibration were performed daily at the beginning of the day following the manufacturer instructions using the typical LC-MS sprayer and ESI-L tuning mix (Agilent Technologies).

## **2.5 Data analysis**

SPE-CE-MS data was analysed by a combination of advanced chemometric tools to evaluate the most significant metabolic changes involved in HD. Figure 1 shows a summary of the data analysis workflow, which is explained in detail in this section.

### **2.5.1. Data pre-processing of data set**

First, SPE-CE-MS raw data was converted to .txt format by using the ProteoWizard software [40] and, then, imported into the MATLAB environment (The Mathworks Inc. Natick, MA, USA) using in-house made routines. During this import process, MS information was compressed to 0.01 Da/e resolution. Every sample provided a data

matrix with 2490 rows (migration times, 40 minutes of electrophoretic run) and 121001 columns (m/z values, from 40 to 1250) (see Figure 1A). An automatic weighted least squares baseline correction was applied before to the MCR-ALS analysis.

### 2.5.2. Full scan MS data arrangement and MCR-ALS analysis

MCR-ALS is a chemometric method especially useful to analyse multicomponent systems with strongly overlapping contributions, such as those present in CE separations, where the electrophoretic behaviour of metabolites is rather similar [41]. In the case of SPE-CE-MS, full scan MS data matrix **D** contains the experimental mass spectra at all retention times in their rows and the electropherograms at all m/z values in their columns. MCR-ALS analysis of the data matrix **D**, following a bilinear model, gives two factor matrices, **C** and **S<sup>T</sup>**, as in Eq. 1:

$$\mathbf{D} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad (1)$$

where matrix **C** contains the electrophoretic profiles of the resolved contributions (components), matrix **S<sup>T</sup>** contains the corresponding mass spectra of the resolved contributions, and matrix **E** contains the residuals unexplained by the model.

The different samples can be simultaneously analysed and compared by MCR-ALS using a column-wise augmented data matrix configuration (see matrix **D<sub>aug</sub>** in Eq. 2 and Figure 1B), following the strategy described in the work of Ortiz-Villanueva [35]:

$$\mathbf{D}_{\text{aug}} = \begin{bmatrix} \mathbf{D}_1 \\ \vdots \\ \mathbf{D}_{15} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_{15} \end{bmatrix} \mathbf{S}^T + \begin{bmatrix} \mathbf{E}_1 \\ \vdots \\ \mathbf{E}_{15} \end{bmatrix} = \mathbf{C}_{\text{aug}}\mathbf{S}^T + \mathbf{E}_{\text{aug}} \quad (2)$$

This approach allowed obtaining a common matrix of the mass spectra of the resolved components (**S<sup>T</sup>**) for all samples, and a set of matrices describing the resolved

electrophoretic profiles ( $\mathbf{C}_{\text{aug}}$ ) in every sample. These electrophoretic peaks resolved in matrix  $\mathbf{C}_{\text{aug}}$  are allowed to vary in position (shifts) and shape among samples because the only requirement for a proper resolution is that the resolved spectra are the same for the common constituents in the different samples [42]. This aspect is especially useful in the case of CE data where migration shifts among samples occur and, hence, the alignment of electrophoretic peaks before analysis is not needed.

In this study, the electropherograms were partitioned in two time windows corresponding to the two regions with the most intense peaks (selected regions, depending on the sample, varied approximately from 10 to 25 min and from 30 to 40 min, respectively, Figure 2). Then, the resulting data matrices were further reduced in their  $m/z$  mode dimension in 30 different  $m/z$  ranges ( $m/z$  widths for reduction were 20, 50 and 100  $m/z$  in the  $m/z$  ranges from 40-400, 400-800 and 800-1250  $m/z$ , respectively) (Figure 1B) [35].

MCR-ALS analysis was carried out following standard procedures for the determination of the number of components (SVD, [43]) and initial estimates (SIMPLISMA, [44]). ALS optimization was performed under non-negativity constraints for electrophoretic ( $\mathbf{C}_{\text{aug}}$ ) and spectral ( $\mathbf{S}^T$ ) profiles, and spectral normalization (equal height) [45,46].

### **2.5.3. Detection and identification of potential metabolites**

For every resolved MCR-ALS component, electropherogram (peak) profiles of the six sample sets (i.e. 8wt, 12wt, 30wt and 8HD, 12HD and 30HD) were compared. Only resolved components of  $\mathbf{C}_{\text{aug}}$  that showed S/N ratios higher than 10% of the abundance of the most intense component were selected. Next, their corresponding mass spectra

profiles ( $S^T$ ) were used to identify the  $m/z$  values causing the differentiation between wt and HD plasma samples at 8, 12 and 30 weeks. Finally, peak areas of these candidate  $m/z$  values were recovered from the full scan SPE-CE-MS data using the MassHunter workstation software, taking as a reference the  $m/z$  value and the migration time of the MCR-ALS resolved components (Figure 1C). Areas were finally normalised considering the peak area corresponding to a compound present in all the samples that was not discriminant between control and HD samples ( $m/z$  of 72.9858, in the first time window).

These areas were used to build a data matrix containing the area of each candidate (feature) in every sample. This data matrix was autoscaled in order to give equal weighting to all candidates in the measured samples. Finally, partial least squares discriminant analysis (PLS-DA) models were applied to the autoscaled data matrix to evaluate sample discrimination and to identify the most important features. There are numerous methods for feature selection when considering PLS models. In this work, the variable importance in the projection (VIP) method was used [47], because it is one of the preferred methods to deal with metabolomic data due to its ability for handling multicollinear data [48]. For each model, VIP scores estimate the importance of each feature in the projection. Only features with a VIP score over a particular threshold (usually  $VIP=1$ ) are considered important and selected for further analysis. In all the cases, leave-one-out cross-validation was used to assess the performance of the built models. Thereafter, the accurate experimental molecular mass values of the finally VIP selected metabolites were searched in on-line databases resources, such as METLIN Metabolite Database [49] and Human Metabolome Database (HMDB) [50]. A small error from the calculated (theoretical) molecular mass ( $M_r$ ) was used to evaluate the

accuracy of possible molecular formulas ( $E_r \leq 20$  ppm,  $\left| M_{r \text{ experimental}} - M_{r \text{ theoretical}} \right| / M_{r \text{ theoretical}} * 10^6$ ). Finally, the list of the tentatively identified metabolites was used to investigate the possible metabolic pathways and mechanisms involved in HD according to the KEGG database [51] (Figure 1C).

#### **2.5.4. Software**

Most of the calculations and data analysis were performed under MATLAB R2013a (The Mathworks Inc. Natick, MA, USA). PLS Toolbox 7.3.1 (Eigenvector Research Inc., Wenatchee, WA, USA) was used for PLS-DA and VIP calculations; and MCR-ALS toolbox [42] was used for resolution of electrophoretic and mass spectral metabolite profiles from full MS scan augmented data matrices.

### **3. Results and discussion**

#### **3.1 Analysis of mice plasma by C<sub>18</sub>-SPE-CE-MS**

Untargeted metabolomics analysis requires a comprehensive coverage of low molecular mass compounds from biological samples. However, very often sample amount limitations, matrix complexity and metabolite concentration preclude direct analysis with CE-MS. With the aim of solving these issues, plasma samples from wt (control) and HD mice were analysed by C<sub>18</sub>-SPE-CE-MS in order to preconcentrate, separate, detect and identify low molecular mass compounds and establish significant differences between the global metabolite profiles from different groups of samples. In order to evaluate HD progression in individuals at the premanifest motor stage of the disease, plasma samples from wt and HD mice were analysed at 8, 12 and 30 weeks of age. In

HD mice, these samples corresponded to early (asymptomatic), middle (symptomatic) and late (terminal) disease stage mice, respectively, although this classification is only based on motor coordination deficiencies [52].

The applied C<sub>18</sub>-SPE-CE-MS method in positive ESI mode was developed for the analysis of peptides in human plasma in previous works [22,37], but preliminary experiments showed that it was also useful to obtain a rich fingerprint of low molecular mass compounds in mouse plasma. As shown in those studies, all the plasma samples were subjected to an off-line sample pretreatment before C<sub>18</sub>-SPE-CE-MS in order to prevent microcartridge saturation due to the limited selectivity of the C<sub>18</sub> sorbent. A double step pretreatment based on solvent precipitation and centrifugal filtration with M<sub>r</sub> cut-off filters was applied to eliminate salts and high molecular mass compounds (i.e. proteins). This pretreatment allowed excellent recoveries for low molecular mass opioid peptides (>70%) [37]. Furthermore, LODs were improved by C<sub>18</sub>-SPE-CE-MS between 1000 and 10000 times compared to CE-MS, depending on the peptides and the sample [39]. Figure 2 shows the total ion electropherograms (TIEs) obtained for the mice plasma samples by C<sub>18</sub>-SPE-CE-MS. As can be observed, separation resolution is not high because of the complexity of the sample. All the electropherograms present a characteristic profile with two time regions with the most intense peaks (approximately at 10-25 min and 30-40 min, respectively), and advanced chemometrics methods are necessary for high throughput and reliable comparison between the different sets of plasma samples.

### **3.2 MCR-ALS analysis and detection of the most relevant metabolites**



MCR-ALS was applied using a column-wise augmented data matrix containing simultaneously the information of the 15 samples (wt and HD, both at 8, 12 and 30 weeks) and allowed the resolution of the electropherogram profiles and corresponding mass spectra of the plasma metabolites.

MCR-ALS analysis was performed separately on column-wise augmented data matrices of different  $m/z$  ranges (at the resolution of 0.01 Da/e), corresponding to the two selected time windows. A total number of 60 column-wise augmented matrices (two time windows  $\times$  30  $m/z$  intervals) were separately analysed. The number of components selected was related to the number of electrophoretic peaks, despite the fact that some of these resolved components could be due to contributions such as solvent background or instrumental noise. In most of the cases, MCR-ALS models showed an explained variance ( $R^2$ ) of almost 100%. The electropherogram profiles for the resolved MCR-ALS components in the six sample sets (i.e. 8wt, 12wt, 30wt and 8HD, 12HD and 30HD) were compared and only resolved components of  $\mathbf{C}_{\text{aug}}$  that showed S/N ratios higher than 10% of the abundance of the most intense component were finally selected (in order to remove contributions such as solvent background or instrumental noise). The mass spectra of these components (from  $\mathbf{S}^T$ ) were used to identify the  $m/z$  values causing the discrimination between samples. After the resolution and analysis of the 60 augmented data matrices, a total number of 74 features were detected. Finally, peak areas of these candidate  $m/z$  values were recovered from the full scan raw  $\text{C}_{18}$ -SPE-CE-MS data using the MassHunter workstation software, taking as a reference the  $m/z$  value and the migration time of features obtained from the MCR-ALS resolved components.

PLS-DA was then applied to identify the most important metabolites responsible for the sample discrimination considering the raw peak areas for the selected 74 candidate metabolites. In order to identify potential Huntington biomarkers which could be useful to discriminate between wt and HD samples, as well as to follow-up the HD progression, three different PLS-DA models were built. Figure 3 shows the PLS-DA scores plot for the mice plasma samples taking into account the three mentioned models. As can be observed in Figure 3A, the first PLS-DA model was applied to discriminate between control and HD samples. This model permitted us to propose possible biomarkers involved in HD. Two latent variables (LVs) explained 34% and 89% of the X and Y variances, respectively. The second PLS-DA model was applied to differentiate between wt samples of different ages and identify metabolites involved in aging of healthy controls. In order to improve the reliability of the PLS-DA model due to the limited amount of samples, a two-class model was used, which presented at least 3 samples in each class (8 weeks and 12-30 weeks). These two sets of samples were also the best option to differentiate later between aging and early HD progression. A PLS-DA model with two latent variables (LVs) explained 47% of the X-variance and the 99% of the Y-variance (Figure 3B). Finally, the third PLS-DA model was applied to distinguish between HD samples of different ages and identify possible biomarkers which could be useful to follow-up the disease progression. Again, the same two sets of samples were defined (8 weeks and 12-30 weeks). In this case, two latent variables (LVs) explained 52% and 98% of the X and Y variances, respectively (see Figure 3C). All PLS-DA models allowed class discrimination and the detection of the most relevant components for the differentiation of the samples. It is worth mentioning that HD samples of 12 and 30 weeks were slightly separated in the scores plot (Figure 3C), whereas this separation was not observed for wt samples (Figure 3B). Anyway, a 3 class

PLS-DA model was not recommended because of the limited amount of samples. VIP scores values higher than 1 were used as a feature selection tool in order to choose only the most relevant candidate metabolites for each PLS-DA model (33 out of 74).

### **3.3. Tentative metabolite identification and biological meaning**

The most contributing metabolites to sample discrimination (33) were tentatively identified, taking advantage of the highly accurate experimental molecular mass values provided by the oa-TOF mass spectrometer. Only 29 features of the total of 33 were tentatively identified with an error lower or equal to 20 ppm (Table 1) (the 4 non-identified features were discarded for further discussion). As can be observed in Table 1, there were some ambiguities on the metabolite identities because this tentative identification was solely based on the agreement between the experimental and the theoretical molecular mass values. For example, in some cases several isobaric metabolites were proposed for a certain molecular formula and experimental molecular mass value (i.e. ID 2, 3, 4, 6, 7, 9, 13, 15, 16, 17, 18, 19, 29). In other cases, it was not possible to differentiate between metabolites with very close theoretical molecular mass values because  $E_r \leq 20$  ppm (i.e. ID 6 ( $E_r=7$  and 14 ppm), ID 8 ( $E_r=16$  and 18 ppm), ID 9 ( $E_r=2$  and 15 ppm), ID 10 ( $E_r=13$  and 17 ppm), and ID 25 ( $E_r=7$  and 16 ppm)). In the future, analysis of standard samples and MS/MS measurements for structure characterisation would be necessary to improve reliability of these identity assignments. The Venn diagrams that appear as insets in Figures 4A and 4B show the relations between the identified metabolites that explain HD progression and aging of healthy controls. As can be observed, 7 metabolites (4+3) were useful to specifically explain HD progression (HD set, ID 1, 10, 28, 3, 13, 15 and 24 in Table 1 and Figure 4).

Similarly, 8 metabolites (8+0, ID 5, 6, 14, 16, 20, 21, 22 and 23 in Table 1 and Figure 4) were useful to specifically explain aging of healthy controls (wt set). The concentration trends of these specific metabolites were varied (Figures 4A and 4B), some of them decrease, while others increase after 12 weeks of birth. Finally, there were 8 metabolites (7+1, ID 29, 4, 11, 26, 7, 8, 9 and 19 in Table 1 and Figure 4) that were explaining both progression of HD and aging. 4 of them showed a clear different concentration trend in HD and wt plasma samples, but for the other 4 metabolites the trend was similar, indicating that differences were found on their absolute concentration (e.g., ID 8 normalised areas in HD and wt plasma samples were 466.4072 and 431.2187, respectively). With regard to differentiation in general of wt and HD samples (wt/HD set), there were 6 metabolites that were useful to specifically distinguish between control and HD samples, 4 downregulated and 2 upregulated in HD samples, as shown in Figure 4C (ID 2, 12, 18, 25, 17 and 27 in Table 1). For metabolites explaining also HD progression and/or aging (ID 10, 15, 1 and 11 in Table 1 and Figure 4C), the concentration trends were varied (2 were downregulated and 2 upregulated).

The identified metabolites were searched against different on-line databases to identify the potential metabolic pathways that could be involved in HD pathology. Different metabolic pathways were found to be related to 13 of the 29 identified metabolites (see Table 2). It is well-known that HD could affect different metabolic pathways. Huntingtin is ubiquitously expressed and, in addition to neurological features, the peripheral phenotype of HD could include weight loss, energy disturbances and alteration of endocrine function.

As it is shown in Figure 4A, concentrations of phenylalanyl-arginine and arginyl-phenylalanine (ID 13, Table 1) were found increased in HD mice after 12 weeks of birth (Figure 4A). These metabolites, which were specific to explain HD progression, are incomplete breakdown products of protein digestion or protein catabolism known to have physiological or cell-signalling effects (Table 2) [53]. Similarly, prostaglandins, thromboxanes, lipoxins and leukotrienes (ID 15, Table 1) were found upregulated after 12 weeks (Figure 4A), but downregulated when all HD samples were compared to all controls (Figure 4C), thus indicating a change of trend after 12 weeks. These metabolites are related with regulation of inflammatory processes and signalling pathways, mainly the arachidonic acid metabolism, the neuroactive ligand-receptor interaction, the serotonergic synapse, the cAMP signalling pathway and the oxytocin signalling pathway (see Table 2). The arachidonic acid metabolism has been also related with the synthesis of cytochromes involved in the mitochondrial oxidative phosphorylation, and altered mitochondrial function has been associated to HD [54,55]. Furthermore, cAMP levels have been found reduced in the striatum of several HD mouse models [56], while the oxytocin signalling pathway has been related with changes in the hypothalamic and limbic systems that take place at HD early stages [57]. Concentration of L-urobilinogen (ID 24, Table 1), which is related with the porphyrin metabolism (Table 2), was also found increased after 12 weeks of birth (Figure 4A). All these changes in 12 weeks-old HD mice suggest an onset on specific neuronal dysfunction, altered expression of several types of receptors and changed expression of neurotransmitters and key proteins. Unbalanced activity within these pathways provides a potential mechanism for many of the pathological phenotypes associated with HD, such as transcriptional dysregulation, inflammation and ultimately neurodegeneration [58–60].

With regard to metabolites explaining both progression of HD and aging (Figure 4A and 4B), gangliosides (ID 29, Table 1), which are cell plasma membrane components that modulate cell signal transduction events, showed a different concentration trend on HD progression compared to aging. Gangliosides levels decreased after 12 weeks of birth in HD progression (Figure 4A), while increased in controls (Figure 4 B). Decreased ganglioside concentration has been also found in the cerebellum of R6/1 (HD) mice at 35-40 weeks [61], and in fibroblasts, cortex and striatum of YAC128 mice [62]. Similarly, L-hexanoylcarnitine levels (ID 11, Table 1), which decreased with aging in healthy controls (Figure 4B), were found to increase with HD progression (Figure 4A), and also when all HD samples were compared to all controls (Figure 4C), suggesting that the disease involves disturbances in energy production, which are characterised by production and excretion of unusual acylcarnitines [63]. Concentration of PC(14:1(9Z)/14:1(9Z)) (ID 26, Table 1), which is related with signalling pathways (the arachidonic acid metabolism and the retrograde endocannabinoid signalling), the glycerophospholipid metabolism and the linoleic acid metabolism, was also found increased with HD progression and decreased with wt aging (Table 2, Figure 4A and 4B). In contrast, changes on the concentration trend with HD progression or aging of (-)-epinephrine and normetanephrine (ID 7, Table 1), which are metabolites related with tyrosine metabolism and signalling pathways (cAMP signalling pathway, adrenergic signalling in cardiomyocytes and neuroactive ligand-receptor interaction) (Table 2) were not observed (Figure 4A and 4B). These metabolites were found decreased after 12 weeks in HD progression and aging (Figure 4A and 4B). The same trend was observed for vanlyglycol and phosphorylcholine (ID 8, Table 1, Figures 4A and 4B), which are related with the tyrosine and the glycerophospholipid metabolisms,

respectively (Table 2). Finally, metabolites with ID 9 (Table 1), presented again a decreasing trend in both HD progression and wt aging (Figures 4A and 4B). In this case, 3-indolebutyric acid is related with the tryptophan metabolism, while the other metabolites are incomplete products of protein digestion or protein catabolism associated with cell signalling effects (Table 2) [64].

With regard to metabolites explaining only wt aging, dimethylbenzimidazole (ID 5, Table 1), which is related with the riboflavin and porphyrin metabolisms (Table 2), was found reduced after 12 weeks of birth (Figure 4B). The same concentration trend was observed for 18-hydroxycorticosterone and cortisol (ID 16, Table 1, Figure 4B), which are metabolites associated with the steroid hormone biosynthesis (Table 2).

Comparing all HD samples with all controls, concentration levels of m-cresol and p-cresol (ID 2, Table 1), which are involved in protein digestion and absorption, as well as in degradation of aromatic compounds, were found downregulated in HD samples (Figure 4C). The same concentration trend was observed for histidinyl-histidine (ID 12, Table 1, Figure 4C), an incomplete breakdown product of protein digestion or catabolism with cell signalling effects [65,66].

#### **4. Concluding remarks**

An optimised sample pretreatment was applied to wild type and R6/1 mice plasma samples (of 8, 12 and 30 weeks) prior to the analysis by C<sub>18</sub>-SPE-CE-MS. The proposed methodology demonstrated to be suitable to ensure a reliable and comprehensive metabolite profiling of the plasma samples. The combination of MCR-ALS with other

chemometric tools, such as PLS-DA, allowed the comprehensive analysis of the C<sub>18</sub>-SPE-CE-MS metabolomic data, resolving electrophoretic peaks and mass spectra of a large number of metabolites. Finally, a list of potential metabolites useful to discriminate between control and HD plasma samples, as well as to follow-up the HD progression, were tentatively identified, and the most affected metabolic pathways were discussed. Although different pathways were found altered in HD, the intracellular signalling was observed to be the most affected, especially after 12 weeks of birth, thus suggesting that the pathology involves dysfunction of specific neurons, altered expression of several types of receptors and changed expression of neuro-transmitters. In addition, although some of the identified metabolites have been previously described in the striatum of R6/1 (HD) mice or other rat models, attempts to find such biomarkers in plasma have hitherto been unsuccessful. In the present work, we propose direct brain-striatal metabolites as good biomarkers that can be found in periphery (plasma samples). Therefore, we provide a window of opportunity for prediction of disease onset, evaluation of HD early progression or response to treatment.

## **Acknowledgements**

Laura Pont acknowledges the Spanish Ministry of Economy and Competitiveness for a FPI fellowship. This study was supported by a grant from the Spanish Ministry of Education and Science (CTQ2011-27130). Part of the study was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 320737. We also thank Josep Maria Marimón for the blood sample collection.



The authors have no conflict of interest to declare.

## References

- [1] Mastrokolias, A., Ariyurek, Y., Goeman, J. J., van Duijn, E., Roos, R. A., van der Mast, R. C., van Ommen, G. B., den Dunnen, J. T., 't Hoen, P. A., van Roon-Mom, W. M., *Eur. J. Hum. Genet.* 2015, 1–8.
- [2] Zielonka, D., Mielcarek, M., Landwehrmeyer, G. B., *Parkinsonism Relat. Disord.* 2015, 21, 169–178.
- [3] van den Bogaard, S. J., Dumas, E. M., Teeuwisse, W. M., Kan, H. E., Webb, A., van Buchem, M. A., Roos, R. A., van der Grond, J., *J. Huntingtons. Dis.* 2014, 3, 377–386.
- [4] Kim, S. D., Fung, V. S., *Mov. Disord.* 2014, 27, 477–483.
- [5] Andre, R., Scahill, R. I., Haider, S., Tabrizi, S. J., *Drug Discov. Today* 2014, 19, 972–979.
- [6] Ross, C. A., Aylward, E. H., Wild, E. J., Langbehn, D. R., Long, J. D., Warner, J. H., Scahill, R. I., Leavitt, B. R., Stout, J. C., Paulsen, J. S., Reilmann, R., Unschuld, P. G., Wexler, A., Margolis, R. L., Tabrizi, S. J., *Nat. Rev. Neurol.* 2014, 10, 204–216.
- [7] Weir, D. W., Sturrock, A., Leavitt, B. R., *Lancet. Neurol.* 2011, 10, 573–590.
- [8] Ross, C. A., Shoulson, I., *Parkinsonism Relat. Disord.* 2009, 15S3, S135–S138.
- [9] Niccolini, F., Politis, M., *World J. Radiol.* 2014, 6, 301–312.
- [10] Rees, E. M., Scahill, R. I., Hobbs, N. Z., *J. Huntingtons. Dis.* 2013, 2, 21–39.
- [11] Krzysztoń-Russjan, J., Zielonka, D., Jackiewicz, J., Kuśmirek, S., Bubko, I., Klimberg, A., Marcinkowski, J. T., Anuszevska, E. L., *J. Bioenerg. Biomembr.* 2013, 45, 71–85.

- [12] Hirayama, A., Wakayama, M., Soga, T., *Trends Anal. Chem.* 2014, *61*, 215–222.
- [13] Villas-Bôas, S. G., Mas, S., Akesson, M., Smedsgaard, J., Nielsen, J., *Mass Spectrom. Rev.* 2005, *24*, 613–46.
- [14] Ramautar, R., Shyti, R., Schoenmaker, B., De Groote, L., Derks, R. J., Ferrari, M. D., Van Den Maagdenberg, A. M., Deelder, A. M., Mayboroda, O. A., *Anal. Bioanal. Chem.* 2012, *404*, 2895–2900.
- [15] Meier, F., Garrard, K. P., Muddiman, D. C., *Rapid Commun. Mass Spectrom.* 2014, *28*, 2461–2470.
- [16] Bignardi, C., Cavazza, A., Corradini, C., Salvadeo, P., *J. Chromatogr. A* 2014, *1372*, 133–144.
- [17] Rao, J. U., Engelke, U. F., Sweep, F. C., Pacak, K., Kusters, B., Goudswaard, A. G., Hermus, A. R., Mensenkamp, A. R., Eisenhofer, G., Qin, N., Richter, S., Kunst, H. P., Timmers, H. J., Wevers, R. A., *J. Clin. Endocrinol. Metab.* 2014, 2014–2138.
- [18] Gertsman, I., Gangoit, J. A., Barshop, B. A., *Metabolomics* 2014, *10*, 312–323.
- [19] Arbulu, M., Sampedro, M. C., Gómez-Caballero, A., Goicolea, M. A., Barrio, R. J., *Anal. Chim. Acta* 2014, *858*, 32–41.
- [20] Wehrens, R., Weingart, G., Mattivi, F., *J. Chromatogr. B* 2014, *966*, 109–116.
- [21] Carty, D. M., Siwy, J., Brennand, J. E., Zürlig, P., Mullen, W., Franke, J., McCulloch, J. W., North, R. A., Chappell, L. C., Mischak, H., Poston, L., Dominiczak, A. F., Delles, C., *Hypertension* 2011, *57*, 561–569.
- [22] Benavente, F., Medina-Casanellas, S., Barbosa, J., Sanz-Nebot, V., *J. Sep. Sci.* 2010, *33*, 1294–1304.
- [23] Ramautar, R., de Jong, G. J., Somsen, G. W., *Electrophoresis* 2015, in press.
- [24] Medina-Casanellas, S., Tak, Y. H., Benavente, F., Sanz-Nebot, V., Sastre

- Toraño, J., Somsen, G. W., de Jong, G. J., *Electrophoresis* 2014, 35, 2996–3002.
- [25] Guzman, N. A., Blanc, T., Phillips, T. M., *Electrophoresis* 2008, 29, 3259–3278.
- [26] Guzman, N. A., Phillips, T. M., *Electrophoresis* 2011, 32, 1565–1578.
- [27] Breadmore, M. C., Tubaon, R. M., Shallan, A. I., Phung, S. C., Keyon, A. S., Gstoettenmayr, D., Prapatpong, P., Alhusban, A. A., Ranjbar, L., See, H. H., Dawod, M., Quirino, J. P., *Electrophoresis* 2015, 36, 36–61.
- [28] Madsen, R., Lundstedt, T., Trygg, J., *Anal. Chim. Acta* 2010, 659, 23–33.
- [29] Kotłowska, A., *Drug Dev. Res.* 2014, 75, 283–290.
- [30] González-Domínguez, R., García, A., García-Barrera, T., Barbas, C., Gómez-Ariza, J. L., *Electrophoresis* 2014, 35, 3321–3330.
- [31] Tseng, Y. J., Kuo, C. T., Wang, S. Y., Liao, H. W., Chen, G. Y., Ku, Y. L., Shao, W. C., Kuo, C. H., *Electrophoresis* 2013, 34, 2918–2927.
- [32] Farrés, M., Piña, B., Tauler, R., *Metabolomics* 2015, 11, 210–224.
- [33] Siano, G. G., Pérez, I. S., García, M. D., Galera, M. M., Goicoechea, H. C., *Talanta* 2011, 85, 264–275.
- [34] Parastar, H., Jalali-Heravi, M., Sereshti, H., Mani-Varnosfaderani, A., *J. Chromatogr. A* 2012, 1251, 176–187.
- [35] Ortiz-Villanueva, E., Jaumot, J., Benavente, F., Piña, B., Sanz-Nebot, V., Tauler, R., *Electrophoresis* 2015, in press.
- [36] Brito, V., Giralt, A., Enriquez-barreto, L., Puigdel·lívol, M., Suelves, N., Zamora-moratalla, A., Ballesteros, J. J., Martín, E. D., Dominguez-iturza, N., Morales, M., Alberch, J., Ginés, S., *J. Clin. Invest.* 2014, 124, 4411–4428.
- [37] Pont, L., Benavente, F., Barbosa, J., Sanz-Nebot, V., *J. Sep. Sci.* 2013, 36, 3896–3902.
- [38] Heiger, D., *High Performance Capillary Electrophoresis. An Introduction*,

*Agilent Technologies.*

- [39] Ortiz-Villanueva, E., Benavente, F., Giménez, E., Yilmaz, F., Sanz-Nebot, V., *Anal. Chim. Acta* 2014, 846, 51–59.
- [40] Kessner, D., Chambers, M., Burke, R., Agus, D., Mallick, P., *Bioinformatics* 2008, 24, 2534–2536.
- [41] Saurina, J., *Chemometric Methods in Capillary Electrophoresis*.
- [42] Jaumot, J., Gargallo, R., de Juan, A., Tauler, R., *Chemom. Intell Lab* 2005, 76, 101–110.
- [43] Golub, G., Solna, K., Dooren, P. V., *SIAM J. Matrix Anal. Appl.* 2000, 22, 1–19.
- [44] Windig, W., Guilment, J., *Anal. Chem.* 1991, 63, 1425–1432.
- [45] Tauler, R., Smilde, A., Kowalski, B., *J. Chemom.* 1995, 9, 31–58.
- [46] de Juan, A., Jaumot, J., Tauler, R., *Anal. Methods* 2014, 6, 4964–4976.
- [47] Wold, S., Sjöström, M., Eriksson, L., *Chemom. Intell Lab* 2001, 58, 109–130.
- [48] Palermo, G., Piraino, P., Zucht, H. D., *Adv. Appl. Bioinforma. Chem.* 2009, 2, 57–70.
- [49] Smith, C. A., Maille, G. O., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., Custodio, D. E., Abagyan, R., Siuzdak, G., *Ther. Drug Monit.* 2005, 27, 747–751.
- [50] Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De Souza, A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazyrova, A., Shaykhutdinov, R., Li, L., Vogel, H. J., Forsythe, I., *Nucleic Acids Res.* 2009, 37, 603–610.

- [51] Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., *Nucleic Acids Res.* 2012, *40*, 109–114.
- [52] Anglada-Huguet, M., Xifró, X., Giralt, A., Zamora-Moratalla, A., Martín, E. D., Alberch, J., *Mol. Neurobiol.* 2014, *49*, 784–795.
- [53] Schug, K. A., Lindner, W., *J. Am. Soc. Mass Spectrom.* 2004, *15*, 840–847.
- [54] Wang, H., Lim, P. J., Karbowski, M., Monteiro, M. J., *Hum. Mol. Genet.* 2009, *18*, 737–752.
- [55] Shirendeb, U., Reddy, A. P., Manczak, M., Calkins, M. J., Mao, P., Tagle, D. A., Reddy, P. H., *Hum. Mol. Genet.* 2011, *20*, 1438–1455.
- [56] Gines, S., Seong, I. S., Fossale, E., Ivanova, E., Trettel, F., Gusella, J. F., Wheeler, V. C., Persichetti, F., MacDonald, M. E., *Hum. Mol. Genet.* 2003, *12*, 497–508.
- [57] Petersén, A., Gabery, S., *J. Hunt.* 2012, *1*, 5–16.
- [58] Bowles, K. R., Jones, L., *J. Huntingtons. Dis.* 2014, *3*, 89–123.
- [59] Carrasco, E., Casper, D., Werner, P., *J. Neurosci. Res.* 2007, *85*, 3109–3117.
- [60] Gabery, S., Halliday, G., Kirik, D., Englund, E., Petersén, A., *Neuropathol. Appl. Neurobiol.* 2015.
- [61] Denny, C. A., Desplats, P. A., Thomas, E. A., Seyfried, T. N., *J. Neurochem.* 2010, *115*, 748–758.
- [62] Maglione, V., Marchi, P., Di Pardo, A., Lingrell, S., Horkey, M., Tidmarsh, E., Sipione, S., *J. Neurosci.* 2010, *30*, 4072–4080.
- [63] Chang, K. L., New, L. S., Mal, M., Goh, C. W., Aw, C. C., Browne, E. R., Chan, E. C., *J. Proteome Res.* 2011, *10*, 2079–87.
- [64] Kee, A. J., Smith, R. C., Gross, A. S., Madsen, D. C., Rowe, B., *Metabolism* 1994, *43*, 1373–1378.

- [65] Kardys, I., de Maat, M. P., Klaver, C. C., Despriet, D. D., Uitterlinden, A. G., Hofman, A., de Jong, P. T., Witteman, J. C., *Am. J. Cardiol.* 2007, *100*, 646–648.
- [66] Alí-Torres, J., Rodríguez-Santiago, L., Sodupe, M., Rauk, A., *J. Phys. Chem. A.* 2011, *115*, 12523–12530.

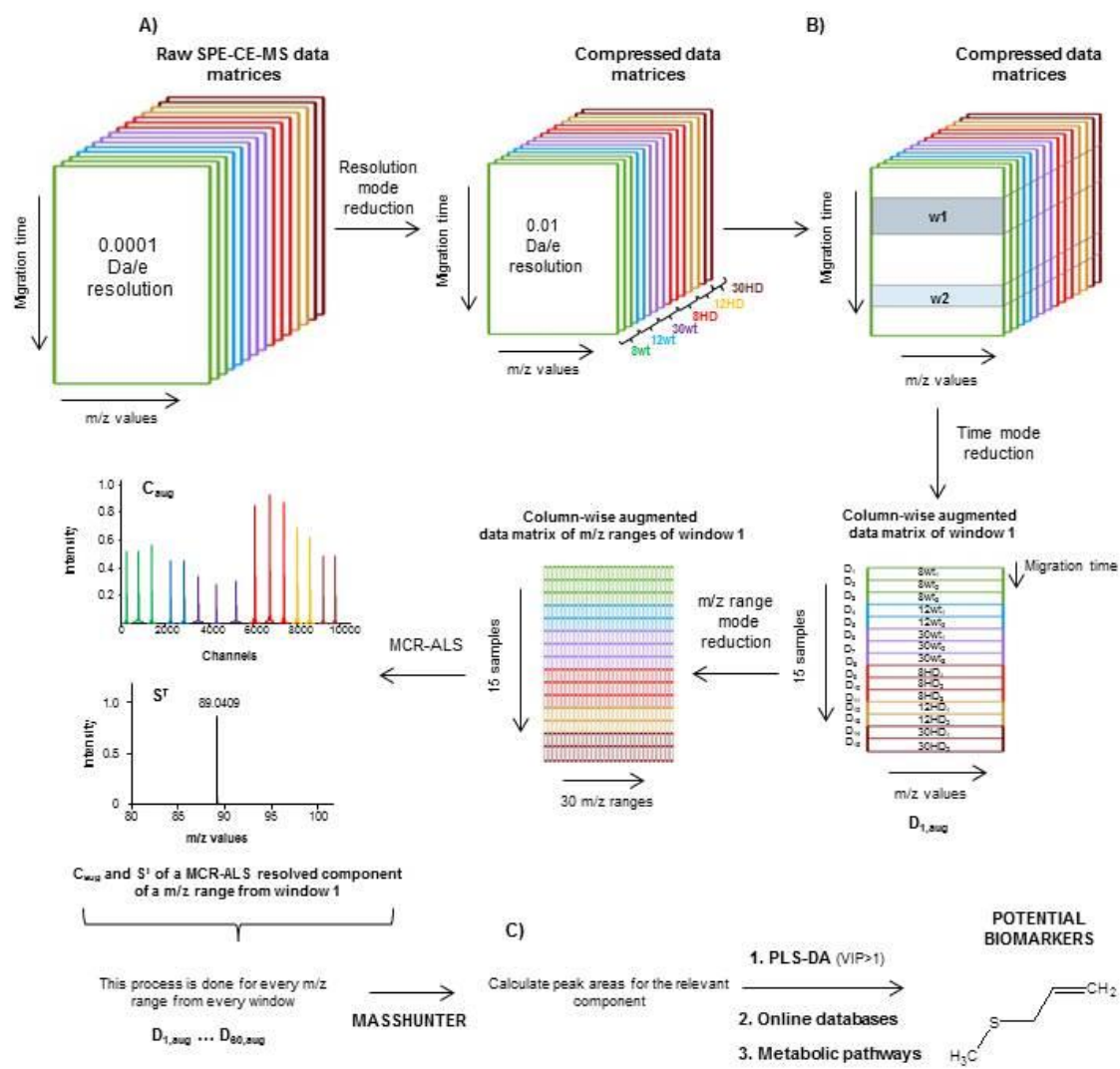
## Figure legends

**Figure 1.** Workflow of C<sub>18</sub>-SPE-CE-MS data analysis: (A) Pre-processing of data set, (B) data arrangement and MCR-ALS analysis in order to detect metabolites, and (C) tentative identification of relevant metabolites and metabolic pathways.

**Figure 2.** Total ion electropherograms (TIEs) obtained by C<sub>18</sub>-SPE-CE-MS for (i) wt and (ii) HD plasma samples at (A) 8, (B) 12 and (C) 30 weeks.

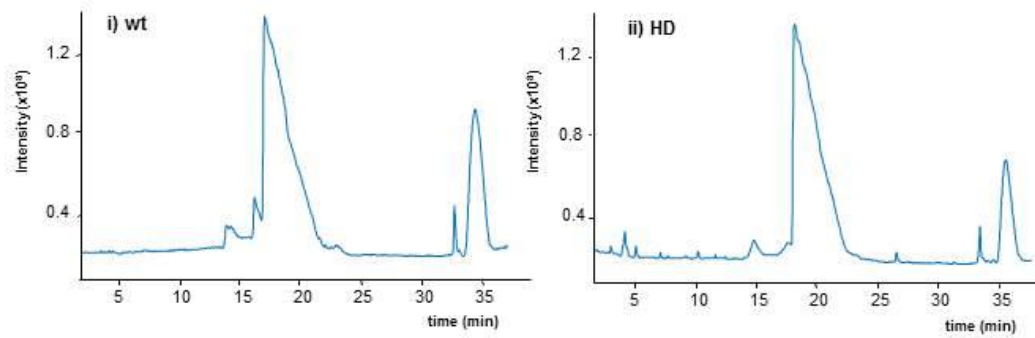
**Figure 3.** PLS-DA scores plot for the mice samples in order to differentiate between (A) all wt and HD plasma samples, (B) wt samples at 8, and from 12 to 30 weeks, and (C) HD samples at 8, and from 12 to 30 weeks. Every sample was analysed in triplicate (with the exception of 12wt, 12HD and 30HD, for which only two replicates were analysed due to the small volume of plasma sample provided).

**Figure 4.** Bar graphs with the folding trends for the identified metabolites (Table 1) that were used to explain (A) HD progression (set HD), (B) aging of healthy controls (set wt) and (C) differences between wt and HD plasma samples (set wt/HD). The % of Abundance of each metabolite was calculated normalising to the metabolite presenting the highest abundance. (\*See Table 2 for the related metabolic pathways).

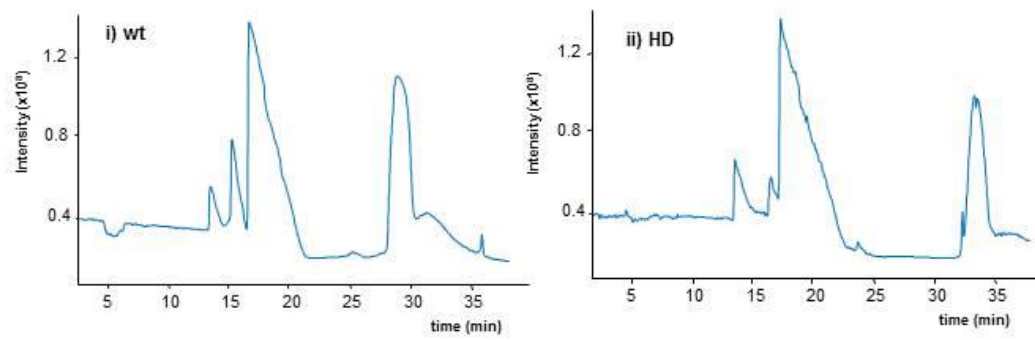


**Figure 1**

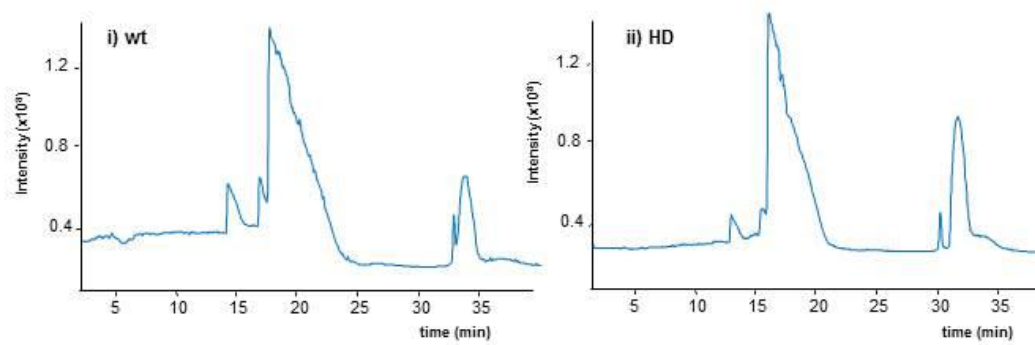
**A) 8 weeks**



**B) 12 weeks**

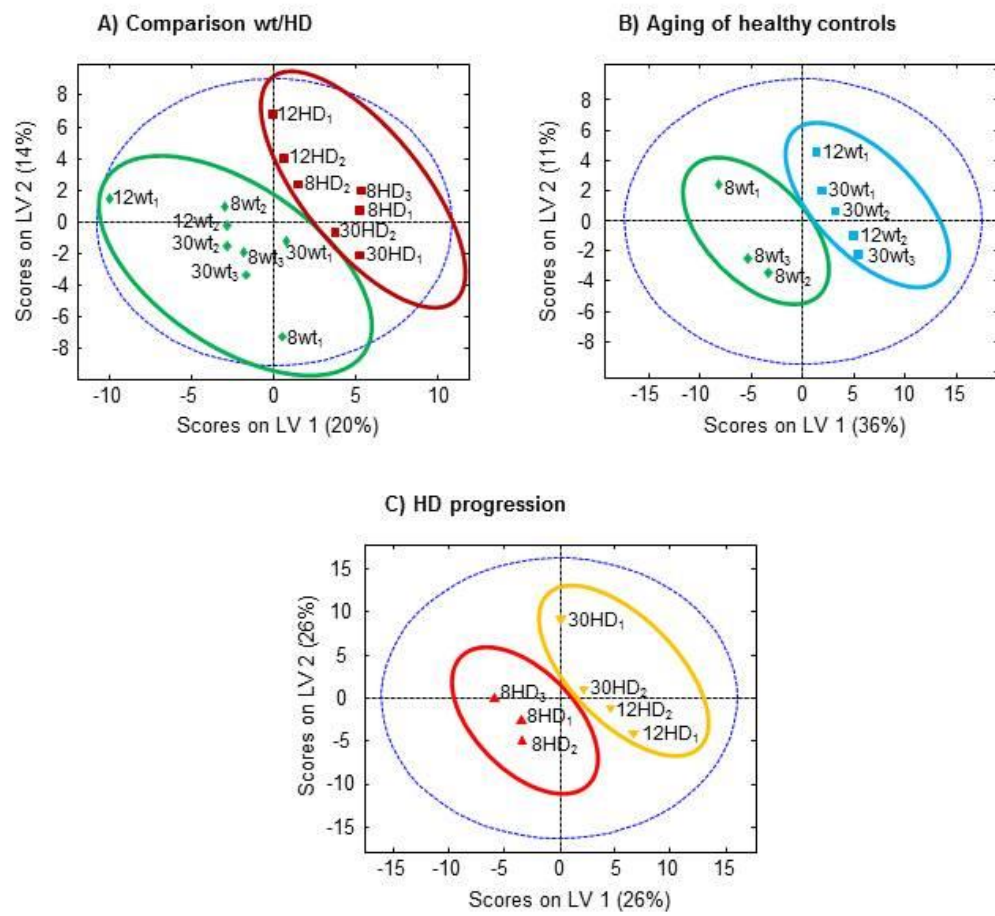


**C) 30 weeks**



**Figure 2**





**Figure 3**



