Universiteit Leiden
The Netherlands

Blood and Biomarkers in Huntington's Disease

Mastrokolias, A.

Citation
Mastrokolias, A. (2017, September 7). *Blood and Biomarkers in Huntington's Disease*. Retrieved from https://hdl.handle.net/1887/54854

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/54854

Note: To cite this publication please use the final published version (if applicable).

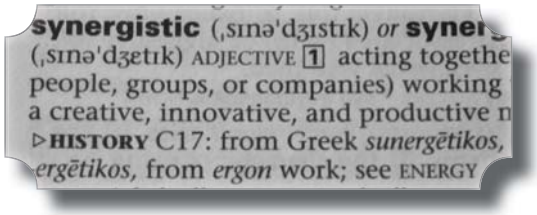
Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/54854 holds various files of this Leiden University dissertation.

Author: Mastrokolias, A.
Title: Blood and Biomarkers in Huntington's Disease
Issue Date: 2017-09-07



Chapter 4

Integration of targeted metabolomics and transcriptomics identifies deregulation of phosphatidylcholine metabolism in Huntington's disease peripheral blood samples.

Anastasios Mastrokolias[1], Rene Pool[2,3], Eleni Mina[1], Kristina M Hettne[1], Erik van Duijn[4,5], Roos C. van der Mast[4], GertJan van Ommen[1] , Peter A.C. 't Hoen[1], Cornelia Prehn[6], Jerzy Adamski [7,8], Willeke van Roon-Mom[1]

1. Department of Human Genetics, Leiden University Medical Center, The Netherlands
2. Department of Biological Psychology, Faculty of Psychology and Education, VU University Amsterdam, The Netherlands
3. The EMGO + Institute for Health and Care Research, VU University Medical Center, The Netherlands
4. Department of Psychiatry, Leiden University Medical Center, The Netherlands
5. Center for Mental Health Care Delfland, Delft, The Netherlands
6. Helmholtz Zentrum München, German Research Center for Environmental Health, München-Neuherberg, Germany
7. German Center for Diabetes Research, Neuherberg, Germany
8. Lehrstuhl für Experimentelle Genetik, Technische Universität München, Freising-Weihenstephan, Germany

Metabolomics (2016) 12: 137

Abstract

Introduction

Metabolic changes have been frequently associated with Huntington's disease (HD). At the same time peripheral blood represents a minimally invasive sampling avenue with little distress to Huntington's disease patients especially when brain or other tissue samples are difficult to collect.

Objectives

We investigated the levels of 163 metabolites in HD patient and control serum samples in order to identify disease related changes. Additionally, we integrated the metabolomics data with our previously published next generation sequencing-based gene expression data from the same patients in order to interconnect the metabolomics changes with transcriptional alterations.

Methods

This analysis was performed using targeted metabolomics and flow injection electrospray ionization tandem mass spectrometry in 133 serum samples from 97 Huntington's disease patients (29 pre-symptomatic and 68 symptomatic) and 36 controls.

Results

By comparing HD mutation carriers with controls we identified 3 metabolites significantly decreased in HD (serine and threonine and one phosphatidylcholine - PC ae C36:0) and an additional 8 phosphatidylcholines (PC aa C38:6, PC aa C36:0, PC ae C38:0, PC aa C38:0, PC ae C38:6, PC ae C42:0, PC aa C36:5 and PC ae C36:0) that exhibited a significant association with disease severity. Using workflow based exploitation of pathway databases and by integrating our

metabolomics data with our gene expression data from the same patients we identified 4 deregulated, phosphatidylcholine metabolism related, genes (ALDH1B1, MBOAT1, MTRR and PLB1) that showed significant association with the changes in metabolite concentrations.

Conclusion

Our results support the notion that phosphatidylcholine metabolism is deregulated in HD blood and that these metabolite alterations are associated with specific gene expression changes.

Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder that presents itself through motor dysfunction, psychiatric disturbances and cognitive decline. The pathology is caused by an expanded CAG repeat in the *HTT* gene, resulting in a mutant huntingtin protein [1]. A characteristic of HD is mutant protein aggregate formation and neuronal cell loss in the brain but it is also known that HD patients develop peripheral tissue symptoms such as muscle atrophy, impaired glucose tolerance and weight loss [2,3]. The mutation for HD was discovered more than 20 years ago and much is known about the underlying disease mechanisms [4]. Moreover, recent studies show that lowering mutant huntingtin protein levels using RNAi is a promising therapeutic approach that is close to clinical trials [5,6]. This highlights/prompts the need for biomarkers that track disease progression and measure clinical trial therapeutic effectiveness.

Deregulation of energy and metabolic pathways have been repeatedly implicated in HD [7-10]. Specifically, defects in lipid homeostasis have been proposed as contributors to disease onset [11-13]. Additionally, total cholesterol was found to be significantly reduced even outside the brain when human fibroblasts where cultured in lipoprotein-deprived serum [14]. Previous studies using HD transgenic models and human caudate samples have shown a deregulation of genes involved in glycosphingolipid metabolism, selected brain gangliosides as well as neutral and acidic lipids. Additionally, Wang and colleagues were able to discover metabolic hormonal plasma signatures in presymptomatic and symptomatic HD patients suggesting that in HD metabolic hormone secretion and energy regulation is affected [15]. Previous mass spectrometry studies have shown differences in the serum metabolome of transgenic HD mice and wild type controls with a similar trend in human samples implicating changes in fatty acid breakdown and certain aliphatic amino acids [16]. Consequently such approaches that use mass spectrometry metabolomics on brain as well as non-nervous system tissue constitute a promising approach for discovering novel HD metabolomics biomarkers [17]. Longitudinal studies have shown promising results in clinical and imaging HD biomarker discovery, but many of these biomarkers are either expensive or subject to inter-rater variability [18]. A good biomarker should identify changes before clinical manifestation,

should be easily obtained and should respond robustly to disease-modifying interventions. Increasingly, metabolomics technology is used in biomarker studies because it can identify intermediate biomarkers of deregulated genomic pathways [19,20]. Furthermore, metabolomics identifies changes that occur downstream of the gene expression level. This applies particularly well in HD since it is recognized that the mutant protein causes genome wide transcriptional deregulation [21,22]. The mutant huntingtin protein is ubiquitously expressed, and gene expression deregulations can be found in various HD tissues and organs. Furthermore, metabolite changes in blood may reflect changes in tissues that have been in contact with blood [23] and as it is impossible to measure molecular biomarkers in the brain, peripheral blood has been proposed as a viable alternative [24]. Nonetheless, the cellular heterogeneity of blood together with the data complexity produced by non-targeted mass-spectrometric protocols, make it difficult to quantify the levels of all metabolites simultaneously. Therefore, we have used a targeted metabolomics approach that measures the concentration of a selected group of HD relevant, key biological compounds (such as amino acids, acyl carnitines, hexoses and glycerophospholipids) in a semi-high throughput manner to identify such metabolomics markers.

The aim of this study was to detect metabolic markers of HD status and progression as well as disease deregulated metabolic pathways. Our approach was based on targeted mass-spectrometry using the Biocrates AbsoluteIDQTM p150 [25] kit to measure metabolite levels in serum from HD carriers and controls. We then tested for the association of the metabolite levels with HD mutation status, and well accepted clinical progression scores and stages such as the Unified Huntington's Disease Rating Scale (UHDRS) total motor score (TMS) and total functional capacity (TFC) score based stages. Since the integration of disparate biological data types like metabolomics and transcriptomics can provide a more complete picture of diseases we correlated our metabolomics data with our publicly available whole genome expression profiling data from the same patient cohort and investigated functional relationships between the metabolite changes and the gene expression changes.

Materials and methods

Metabolite measurements

Metabolite concentrations were determined using the targeted metabolomics kit AbsoluteIDQ™ p150 (Biocrates Life Sciences AG, Innsbruck, Austria) and flow injection electrospray ionization tandem mass spectrometry (FIA-ESI-MS/MS). A total of 163 different metabolites were quantified simultaneously by the platform in 10 µL of serum. The metabolite panel consists of 14 amino acids, Hexose (H1), free carnitine (C0), 40 acylcarnitines (Cx:y), hydroxylacylcarnitines (C(OH)x:y), and dicarboxylacylcarnitines, 15 sphingomyelins (SMx:y) and N-hydroxyacyloylsphingosyl-phosphocoline (SM(OH) x:y), 77 phosphatidylcholines (PC, aa=diacyl, ae=acyl-alkyl) and 15 lyso-phosphatidylcholines. Lipid side chains are denoted as Cx:y , where x represents the number of carbons in the side chain and y the number of double bonds. The assay procedures of the p150 kit as well as the metabolite nomenclature have been described in detail previously [25]. Sample handling was performed by a Hamilton Microlab STAR™ robot (Hamilton Bonaduz AG, Bonaduz, Switzerland) and an Ultravap nitrogen evaporator (Porvair Sciences, Leatherhead, U.K.), beside standard laboratory equipment. Mass spectrometric (MS) analyses were done on an API 4000 LC-MS/MS System (Sciex Deutschland GmbH, Darmstadt, Germany) equipped with a 1200 Series HPLC (Agilent Technologies Deutschland GmbH, Böblingen, Germany) and a HTC PAL auto sampler (CTC Analytics, Zwingen, Switzerland) controlled by the software Analyst 1.5. Data evaluation for quantification of metabolite concentrations and quality assessment was performed at the Genome Analysis Center of the Helmholtz Zentrum München using the MetIDQ™ software package, which is an integral part of the AbsoluteIDQ™ kit. Internal standards served as reference for the calculation of metabolite concentrations in µM.

Serum collection

Peripheral blood was collected from 29 presymptomatic, 68 symptomatic and 36 control, non-fasting individuals with institutional review board approval and after informed consent. For serum sample isolation, blood was collected in BD vacutainer Z tubes (no

additives) and was allowed to clot for 1 hour at room temperature. Tubes were spun at 1300 g for 10 minutes at room temperature, were aliquoted and stored at -80 °C. Detailed information about the UHDRS clinical scores and CAG repeat lengths of all the patients and controls as well as gender, age and BMI information can be found in Supplementary File 1.

Quality controls

To ensure the robustness of downstream statistical analyses, all data provided from the MetIQ software package were subjected to three quality control steps. For the first step the coefficient of variance was calculated for each experimental plate. To achieve this five aliquots of a reference plasma pool were measured on each plate together with the cohort samples. The coefficient of variance was calculated as the standard deviation to mean ratio for all five reference samples per metabolite and per experimental plate. All metabolites with a mean coefficient of variance of all plates, higher than 25 % were excluded from further analysis. All metabolites with a missing value rate larger than 5% were also excluded. In the second step any outlying data points with a value greater than mean +/- 5 SD of all measurements for this metabolite were excluded. Additionally, two Huntington disease samples were excluded due to high BMI values (outliers). For samples with less than, or equal to, three independent outlying points only the independent data points themselves were excluded. After these quality control steps 114 out of 163 metabolites and 133/138 samples remained. After the above steps, when missing values were detected these were imputed using the R package "mice". Finally, all metabolite concentrations were transformed using natural logarithm and before applying the experimental linear modeling analysis.

Statistical analysis

To identify significant differences between the HD and control samples, the statistical analysis software R (Version 3.1.2, http://www.r-project.org/) was used. After the metabolite concentrations were log-transformed, linear modelling statistical tests were applied. In specific, in the first model (disease status) the HD mutation carriers' and control

individuals' groups were coded as the main covariate and tested in a linear model using gender (categorical), age and BMI as additional (continuous) covariates. In the second model (disease severity) a four group categorical variable vector was used as the main covariate. The following groups were defined: Group 1 (n=36) – control, Group 2 (n=29) – pre-symptomatic (TMS score <=5), Group 3 (n=31) – symptomatic (TMS>5, TFC score 13-7) and Group 4 (n=37) – advanced symptomatic individuals (TMS score>5, TFC score of 0-6). For both models the final *P* value that was used to judge the validity of our findings was extracted using the ANOVA function on the two nested linear models; the reduced linear model containing only the covariates of gender, age and BMI and the full model additionally containing the main disease group categorical covariate. Since many identified metabolites showed a high degree of correlation (see Supplementary File 2), the Bonferroni method was judged too strict for multiple testing correction. Therefore, the experiment-wide significance threshold that was used was 1.34E-03. This value was calculated using the matrix spectral decomposition method and the eigenvalues of the metabolites correlation matrix (matSpDlite) [26-28]. Bar plots for all metabolites that passed quality control were created using the Platform for RIKEN Metabolomics (PRIMe) tool for Microsoft excel [29]. The PLS-DA analysis across controls, presymptomatic and symptomatic HD carriers was performed using the corresponding function of the MetaboAnalyst v.3.0 online tool for metabolomics data [30].The top, pair ratio associations for all possible metabolite pair ratios were calculated through log transformation of the ratios and the p-gain value was calculated from the individual *P* values of the ratio metabolites. In specific, p-gain is defined as the fold decrease in the *P* value of association for the pair of metabolites compared to the lower of the two *P* values for the single metabolites [31].

Integration of metabolomics with transcriptomics analysis

Gene expression data from a previously published dataset [32] were extracted using the scaled data object from the "voom" function of the "limma" package designed for RNA-seq data analysis [33]. The data for both the genes and the metabolites were regressed for the effect of age, gender and BMI. The gene expression data were regressed for cellular hemoglobin percentage (hemoglobin alpha and beta sequencing count tags) as a proxy

of the cell reticulocyte count [34]. For the extraction of the genes that were related to the metabolites the metabolic pathway databases from the Kyoto Encyclopedia of Genes and Genomes KEGG [35] (release 63) and BioCyc (version 16) were accessed for retrieving background knowledge for each metabolite [36]. Two interrogation schemes were employed: pathway scheme and reaction scheme. In a pathway scheme, for a given metabolite, all the pathways that it participated in were determined followed by the retrieval of all the genes that participated in these pathways. In a reaction scheme, given a metabolite, all the reactions that it was part of and the compounds that participated in these reactions were determined [36].

For the integrated metabolomics and transcriptomic pathway analysis, the community driven resource of curated pathways WikiPathways [37] was used to identify common pathways. The WikiPathways human pathway collection is the largest and most active collection per species. In terms of coverage of unique human genes, WikiPathways is comparable to KEGG. To investigate which metabolite-gene pathways overlapped we used all 10 significant metabolites (8 phosphatidylcholines and 2 amino acids) and the top 200 genes from our above linear modeling, as the input to the WikiPathways Web Service. Since pathway information for individual phosphatidylcholines is lacking, we also included the compounds at the phosphatidylcholine compound class level and their isomers (1,2-diacyl-sn-glycero-3-phosphocholine, alkyl,acyl-sn-glycero-3-phosphocholine and 1,2-diacyl-sn-glycero-3-phosphocholine(1+)) according to the Chemical Entities of Biological Interest ontology [38].

Results

The metabolomic dataset

Using the Biocrates p150 kit we quantified serum concentrations of 163 metabolites in 133 serum samples from 97 HD mutation carriers and 36 controls. After quality control, 114 of the initial 163 metabolites could be reliably detected and these were used for further analysis. These 114 metabolites consisted of 14 amino acids, 7 carnitines, 10 lyso-phosphatidylcholines, 69 phospatidylcholines, 5 hydroxysphingomyelins, 8 sphingomyelins and 1 hexose. Supplementary File 3 (available online) contains data distributions of all 163 metabolites, the metabolites that were excluded and their concentrations relative to the platform limit of detection (LOD) and lower limit of quantification (LLOQ). In order to analyze the group structure of the metabolomic dataset across controls, presymptomatic and symptomatic HD patients, we performed a partial least square discriminant analysis (PLS-DA). The symptomatics group (group 3 and 4 combined, TMS >5) exhibited a clear shift from the control (group 1) and the pre-symptomatic (group 2) samples. The scores plot for the symptomatics, pre-symptomatics and control groups and for the first two principal components can be seen in Figure 1A while the relative contributions and the relationships between the metabolites can be seen in the loadings plot of Figure 1B. The observed concentration levels for all 114 metabolites across all four groups can be seen in Supplementary File 4 (available online). To identify which metabolites were significantly different in HD, the concentration changes of the detected metabolites were tested using a linear regression model between HD mutation carriers versus control individuals and a linear regression model using four disease severity stage groups as described in the Material and Methods section. We identified 3 metabolites significantly changed in the HD mutation carriers vs controls analysis and 8 metabolites significantly changed in the 4 disease stage group analysis that associated with disease progression (adj.P.val < 1.34E-03) (see Table 1). In the two group analysis the amino acids serine and threonine were higher in HD mutation carriers while the phosphatidylcholine acyl-alkyl C 36:0 average level was lower (see Figure 2). In the 4 group analysis, 8 metabolites in total passed the significance threshold.

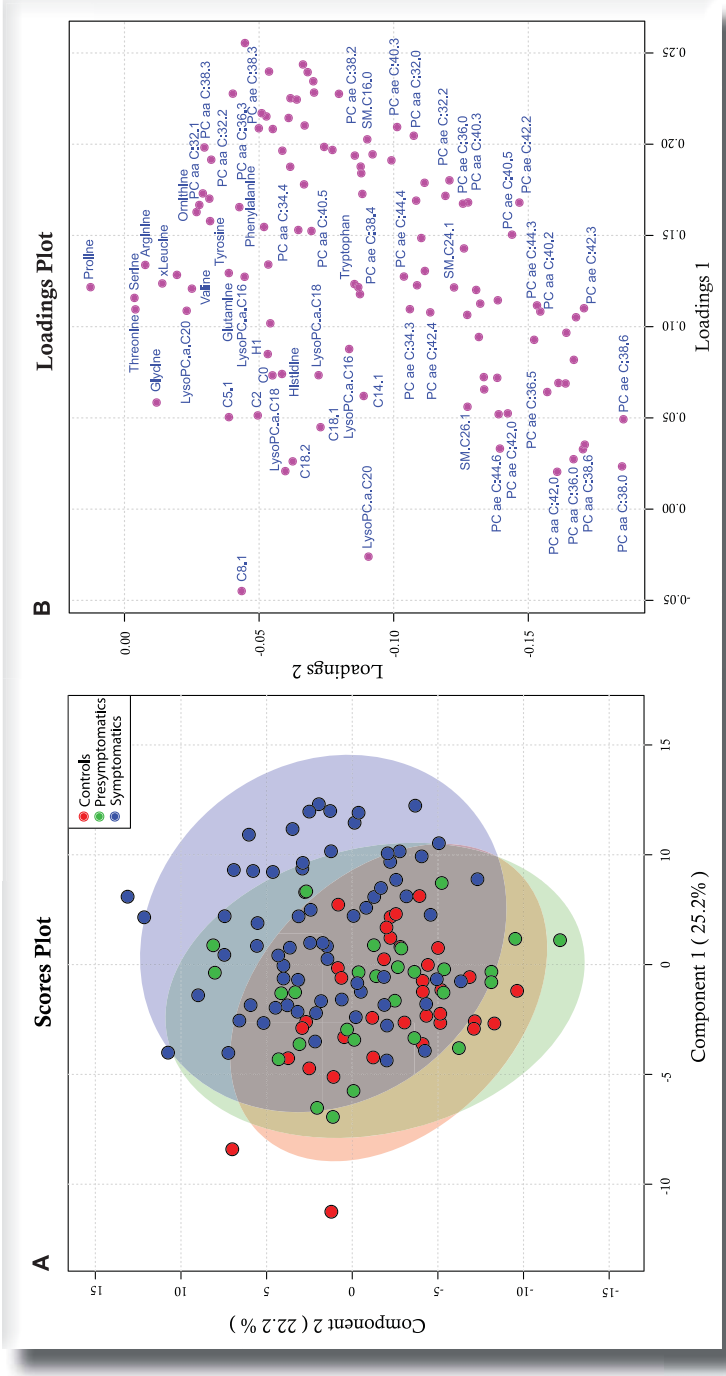


Figure 1. Diagram representing principal component analysis performed using Metaboanalyst v3.0 **A.** Principal component analysis results represent HD presymptomatic, symptomatic patients and control group separation based on the top 2 most significantly contributing components. Colored circles represent 95% confidence intervals. Colored dots represent individual samples. **B.** Loading plots of the principal components for the platform metabolites. Some metabolite names have been omitted next from their corresponding metabolite dot for figure clarity purposes.

4

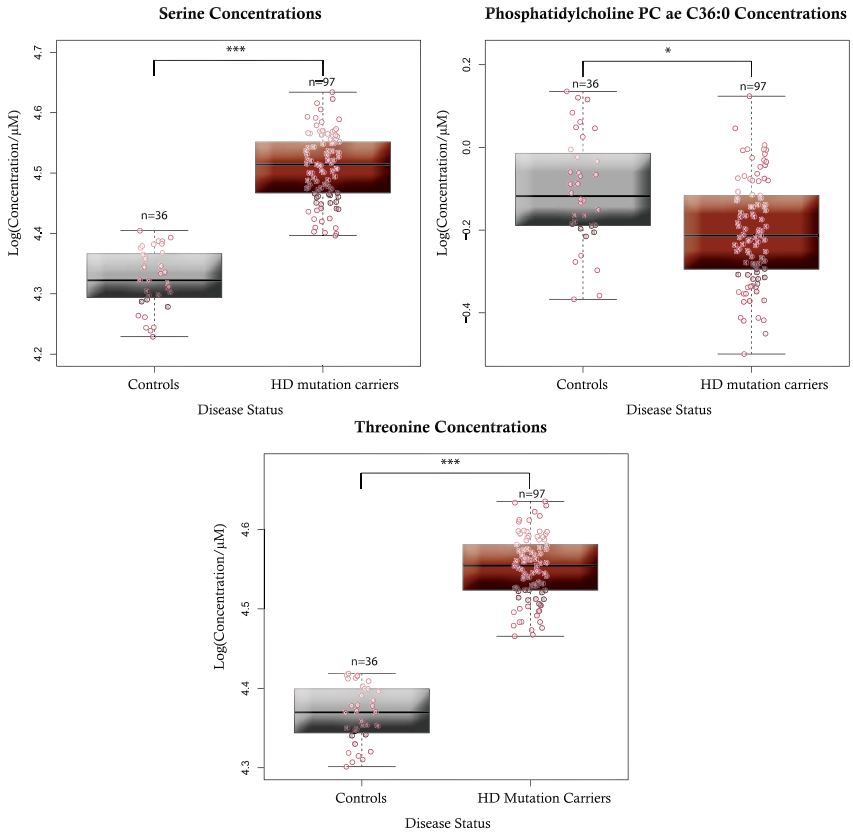
Metabolite ID	HD Mutation Carriers vs Controls Analysis	P Value	Concentration Change
Ser	**Serine**	**3.08E-05**	**Higher in HD**
PC ae C36:0	**Phosphatidylcholine acyl-alkyl C 36:0**	**6.61E-05**	**Lower in HD**
Thr	**Threonine**	**1.02E-03**	**Higher in HD**
PC ae C42:0	Phosphatidylcholine acyl-alkyl C 42:0	2.35E-03	Lower in HD
PC ae C44:3	Phosphatidylcholine acyl-alkyl C 44:3	3.72E-03	Lower in HD
PC aa C38:6	Phosphatidylcholine diacyl C 38:6	4.60E-03	Lower in HD
PC ae C38:0	Phosphatidylcholine acyl-alkyl C 38:0	8.95E-03	Lower in HD
Arg	Arginine	1.80E-02	Higher in HD
PC aa C36:0	Phosphatidylcholine diacyl C 36:6	1.99E-02	Lower in HD
PC aa C40:6	Phosphatidylcholine diacyl C 40:6	2.33E-02	Lower in HD
Metabolite ID	**Disease Progression (4) Group Analysis**	**P Value**	**Concentration Change**
PC aa C38:6	**Phosphatidylcholine diacyl C 38:6**	**1.01E-04**	**Lower in HD**
PC aa C36:0	**Phosphatidylcholine diacyl C 36:0**	**4.22E-04**	**Lower in HD**
PC ae C38:0	**Phosphatidylcholine acyl alkyl C 38:0**	**4.42E-04**	**Lower in HD**
PC aa C38:0	**Phosphatidylcholine diacyl C 38:0**	**5.48E-04**	**Lower in HD**
PC ae C38:6	**Phosphatidylcholine acyl alkyl C38:6**	**5.63E-04**	**Lower in HD**
PC ae C42:0	**Phosphatidylcholine acyl alkyl C 42:0**	**1.06E-03**	**Lower in HD[1]**
PC aa C36:5	**Phosphatidylcholine diacyl C 36:5**	**1.16E-03**	**Lower in HD[2]**
PC ae C36:0	**Phosphatidylcholine acyl alkyl C 36:0**	**1.19E-03**	**Lower in HD[1]**
PC.ae.C40:1	Phosphatidylcholine acyl alkyl C40:1	1.72E-03	Lower in HD
Ser	Serine	1.96E-03	Higher in HD

[1]Upper quartile range higher in earlier HD symptomatics (group 3) vs HD presymptomatics (group2)

[2] Higher in earlier symptomatics

◂ **Table 1.** Table of the top 10 metabolites resulting from linear modeling analysis of metabolite concentrations between HD mutation carriers and controls and accounting for disease status (top) or disease progression group (bottom) and gender, age and BMI. P values represent significance probability values (Pr(>F)) from the two-way ANOVA calculation on two (nested) linear models accounting for disease status or disease stage respectively (measurement variables) and gender, age and BMI (nominal variables). Metabolites that pass the adjusted p value threshold are highlighted in bold. Concentration changes were obtained from the fitted data of the metabolites using the full linear statistical model (see above) and disease state and stage respectively as the main covariate.

▾ **Figure 2.** Boxplots of concentration levels of metabolites that were significantly different between control individuals and HD mutations carriers. Numbers represent the group sizes and asterisks represent significance values from linear modelling analysis. Colored dots represent individual sample concentrations. Asterisks represent significance probability values (Pr(>F)) from the ANOVA calculation of the single (full – see methods) linear model accounting for disease status, gender, age and BMI. * = P-value < 0.05, *** = P-value < 0.001



4

These 8 metabolites were exclusively acyl alkyl and di-acyl phosphatidylcholines and were lower in HD versus controls and associated with increasing disease progression (for the top 5 metabolites see Figure 3). In the 4 group analysis serine was also in the top 10 metabolites but failed to pass the adjusted *P* value significance level.

Association of metabolite pair ratios with disease status and severity

Previous studies have shown that calculating the ratios of individual metabolite concentrations can reduce dataset variation. Furthermore, such metabolite ratio changes have been connected to altered enzymatic reactions and pathways, can be used as an approximation of the associated enzymatic activity [39,40] and ratios of specific pairs of metabolites have been suggested as biomarkers [41,42]. For this reason, we calculated all the pairwise ratios of the detected metabolites and tested for the association of their ratios using the same two linear models. The resulting associations were ranked according to their p-gain values. The results for the two group (HD vs controls) analysis and the four group analysis can be seen in Supplementary File 5. We observed that in the two group analysis the results were dominated by inter- phosphatidylcholine ratios, as well as ratios of phosphatidylcholines to serine and threonine. Additionally, one of the top p-gain values was that of the arginine to carnitine ratio. Most of the ratios of the sphingolipids and sphingolipids to amino acids were lower in HD carrier samples. The metabolite pair associations with controls, pre-symptomatic and the 2 symptomatic groups revealed similar results. The highest p-gain values were exhibited by inter-phosphatidylcholine ratios but also phosphatidylcholines and the hydroxy-sphingomyelin C16:1 (SM.OH.C16:1). The 4 group disease progression analysis was also characterized by the absence of any amino acids in the top p-gain analysis similar to the individual metabolite analysis. These results confirm the changes in phosphatidylcholines levels in the disease and strengthen the potential of the use of (pairs of) phosphatidylcholines as markers of disease progression since variation is reduced.

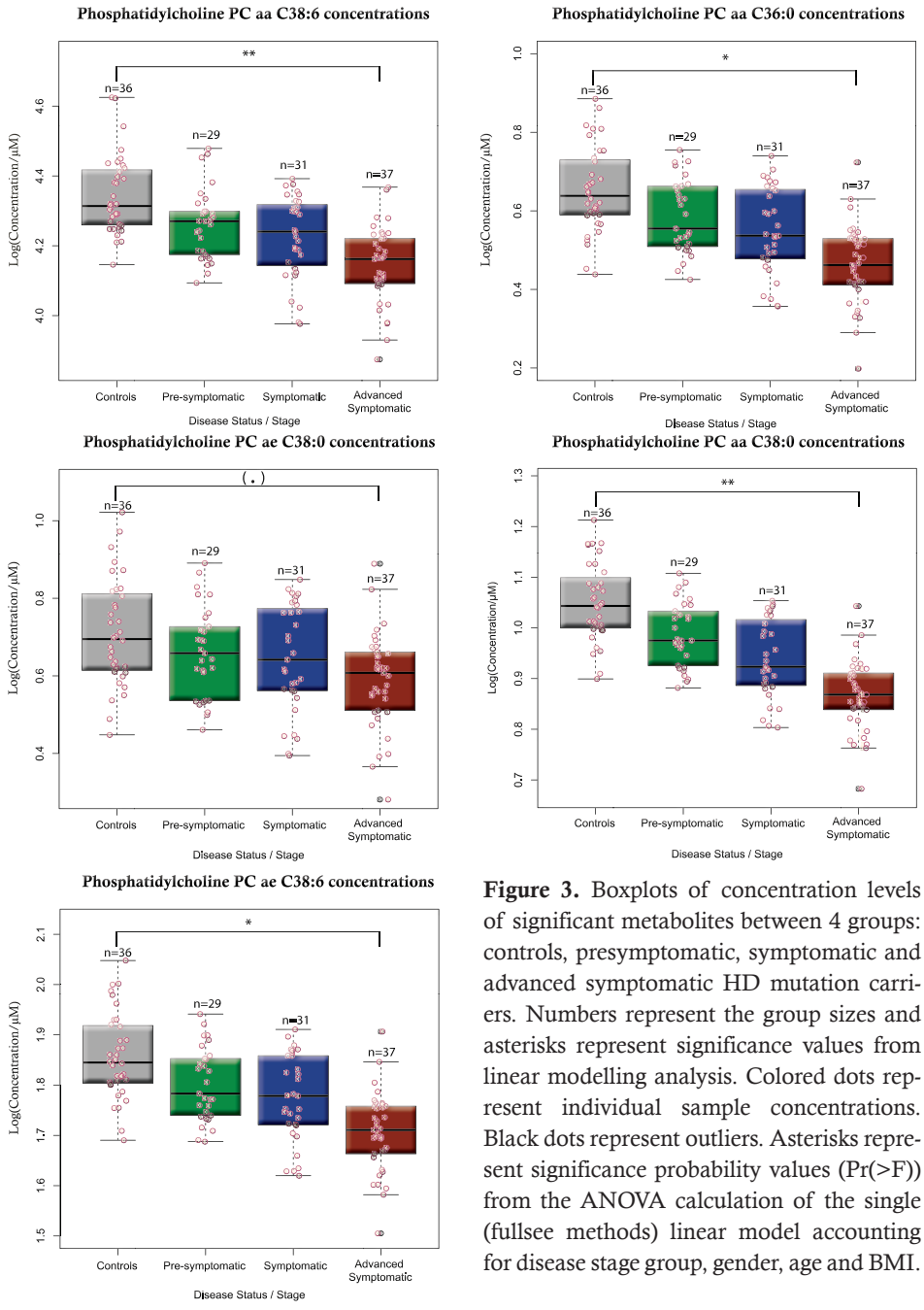


Figure 3. Boxplots of concentration levels of significant metabolites between 4 groups: controls, presymptomatic, symptomatic and advanced symptomatic HD mutation carriers. Numbers represent the group sizes and asterisks represent significance values from linear modelling analysis. Colored dots represent individual sample concentrations. Black dots represent outliers. Asterisks represent significance probability values (Pr(>F)) from the ANOVA calculation of the single (fullsee methods) linear model accounting for disease stage group, gender, age and BMI.

Integration of metabolomics with transcriptomics

To further explore potential molecular connections of the identified metabolites with HD relevant or novel disease state/progression mechanisms we combined our targeted metabolomics dataset with our previously published next-generation sequencing gene expression data from the same patient and control cohort [32]. For our initial analysis we focused on the phosphatidylcholine metabolites since these exhibited significant statistical associations with HD disease progression scores. Using the previously published methodology of Dharuri et al. [36] we extracted the genes from KEGG [43] and BioCyc [44] databases that corresponded to phosphatidylcholine related metabolic pathways. Consequently, we reanalyzed the previously published gene expression data from the same cohort, using the same linear model with the metabolomics dataset. We extracted the top 200 differentially expressed genes and compared them with phosphatidylcholine pathways related genes from the KEGG and BioCyc databases. We identified 8 genes that were present in both the differentially expressed gene list and the above databases lists. These genes were *ALOX5* (arachidonate 5-lipoxygenase), *ALDH1B1* (aldehyde dehydrogenase 1 member 1), *KMT2A* (lysine specific methyltransferase 2A), MBOAT1 (membrane bound O-acyltransferase DC1), *MTRR* (methionine synthase reductase), *PISD* (phosphatidylserine decarboxylase), *PLB1* (phospholipase B1) and *HADH* (hydroxyacyl-CoA dehydrogenase). The correlation values of each of the 8 genes with the 8 significant metabolites from our 4 group linear modeling analysis are represented in Figure 4. We observed that for the genes *MBOAT1*, *PLB1*, *ALDH1B1* and *MTRR* the correlations with the majority of the 8 metabolites were high ($r > 0.6$) while for the other 4 genes the associations were average or poor. The highest associations were observed between the genes *ALDH1B1*, *MTRR* and *PLB1* with phosphatidylcholines PC ae C.38:0, PC aa C36:5 and PC aa C38:6 (see Figure 5). Additional genes that were present both in the BioCyc and KEGG databases and our previous sequencing-based gene expression gene lists and for the amino acid serine were also *NPL* (N-acetylneuraminate pyruvate lyase), *PGLYRP1* (peptidoglycan recognition protein 1) and *TKTL1* (transketolase-like 1). Finally, for the serine and threonine metabolites we could not identify any unique common genes. It should be noted however that the above phosphatidylcholine related gene *ALDH1B1* was also present in our threonine KEGG reaction list and similarly *PISD*

and *PLB1* were also present in our serine KEGG reaction list. To expand on the above findings, we performed a second analysis using the WikiPathways Web Service and all 10 significant metabolites (8 phosphatidylcholines and 2 amino acids) from both of the above metabolomics linear models, in order to investigate further connections between potential metabolomics and transcriptomic pathways. The metabolite-gene pathways with the highest overlap of genes and metabolites we identified were glycerophospholipid biosynthesis (containing genes *PLB1*, *PISD* and Serine and 1,2-diacyl-sn-glycero-3-phosphocholine (1+)) and phase II conjugation (containing gene *MTRR* and serine and threonine), supporting the results from our first pathway analysis. All the overlapping pathways reported for the Wikipathways analysis can be seen in Supplementary File 6.

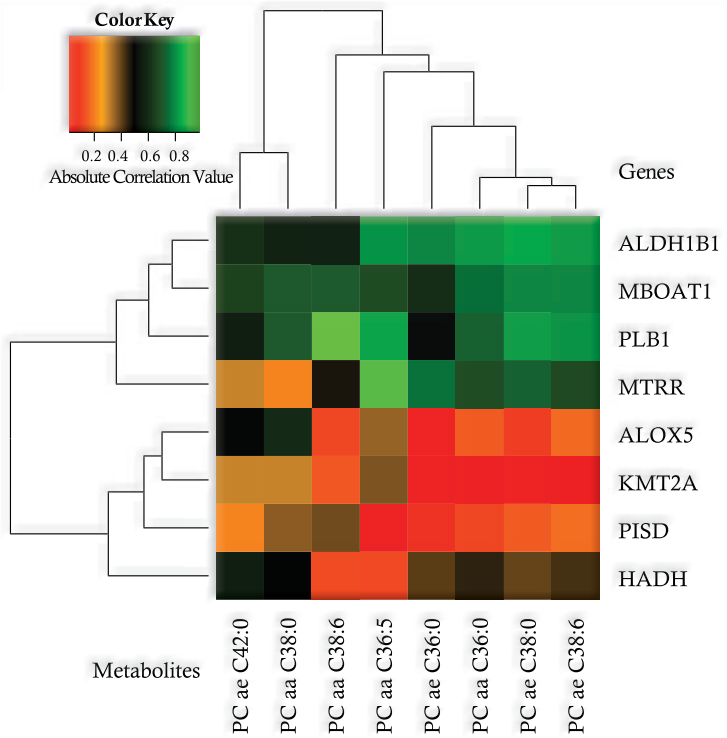


Figure 4. Heatmap of correlation values between gene expression levels and phosphatidylcholines metabolite concentrations. The selected genes shown here are genes identified using our previous gene expression data and that participate in phosphatidylcholine KEGG and BioCyc pathways and reactions. Phosphatidylcholines shown here are the statistically significant phosphatidylcholine metabolites identified from the 4 group linear modelling analysis. Color key represents absolute correlation values.

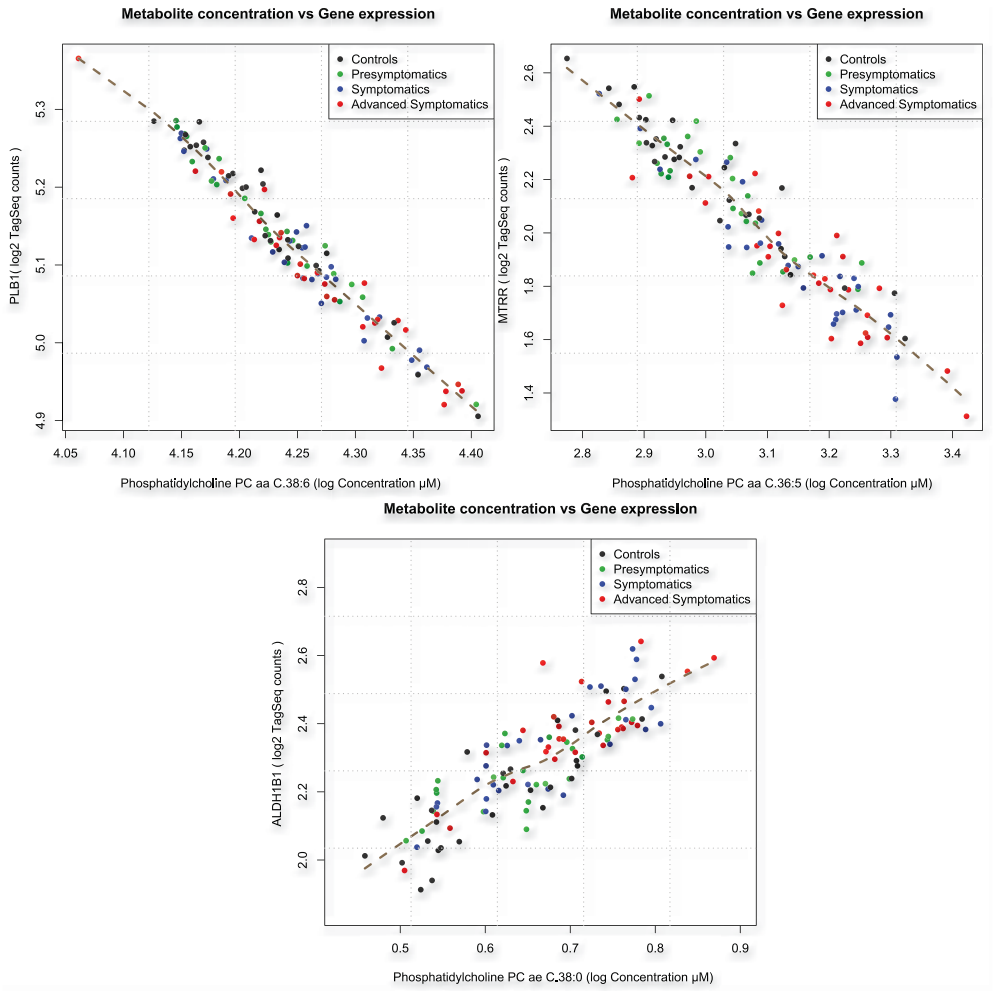


Figure 5. Plots of selected phosphatidylcholine metabolites versus their corresponding gene that participates in phosphatidylcholine pathways or reaction. The 3 plots represent the most highly correlated metabolite to gene pairs from the integration of metabolites to gene expression data analysis. Different colored dots represent individual samples from each disease state group and brown lines represent loess fit lines.

Discussion

The current study targeted approach of the Biocrates technology has been successfully applied in many cohort studies [45-47]. Comparison of this type of data with data obtained from non-targeted platforms has shown strong positive correlations for metabolites named for the same compounds. Furthermore such a comparison has shown that the results obtained are complementary and informative for future studies of comprehensive metabolomic analyses with different platforms [48,49].Using a well-defined, UHDRS-based linear model, we discovered a total of 10 metabolites whose concentrations showed significant associations with Huntington's disease state and severity stages. Eight of the 10 metabolites were phosphatidylcholines while the other two were the amino acids serine and threonine. These results are in agreement with the results of Tsang et al. that have reported a decrease of phosphatidylcholine levels in frontal cortex lipid extracts of a 3-NP treated HD rat model [50]. Phosphatidylcholine is a major membrane phospholipid and has been shown to have a role in neuronal differentiation and cell fate determination [51]. In the past, oral administration of lecithin and other choline containing dietary sources have been suggested as a replacement therapy for HD and as a potential substrate source for brain acetylcholine synthesis [52]. The current study shows an increase of serine and threonine in HD patients as shown from the HD versus controls linear modeling. Serine has an important role in the metabolism of purines and pyrimidines since it is the precursor of several other amino acids. It is also a precursor to numerous other metabolites, including sphingolipids and folate, which is the principal donor of one-carbon fragments in biosynthesis. As such, one explanation for the increased serine levels could be that in Huntington's disease these amino acids are intended for the production of phospholipids whose levels are decreasing with disease severity. Moreover, the D-serine amino acid isomer can act as a neuromodulator since it can activate NMDA receptors. NMDA receptors have been implicated in a range of processes including memory, learning and development and their excessive stimulation can be involved in a number of neurodegenerative conditions including HD [53].

The second of the two amino acids whose levels were altered in HD, threonine, is an essential amino acid and together with serine constitute the only two proteinogenic amino

acids. Threonine can be converted to pyruvate while in an intermediate step it can undergo thiolysis to produce acetyl-coA. In a less common pathway threonine can also be converted to a-ketobutyrate via serine dehydratase. It has been previously suggested that pyruvate can have a neuroprotective effect in neurological diseases by, among others, enhancing brain to blood glutamate efflux, scavenging H_2O_2 and having an anti-inflammatory action [54]. Furthermore it has been shown that pyruvate administration can have a neuroprotective effect in a quinolinic acid rat model of HD [55]. Additionally, the inability of the excess threonine to undergo thiolysis and produce acetyl-CoA could result in reduced energy production (Krebs cycle) as well as an insufficient synthesis of acetylcholine. The increased levels of threonine in the mutation carriers could therefore represent a compensatory mechanism in an attempt to produce more substrates for the generation of the above neuroprotective molecules such as pyruvate and/or the inability of the threonine metabolizing enzymes to properly process the present levels of this amino acid.

Using the pairwise combination of all the individual metabolites we discovered a series of metabolites ratios (mainly phosphatidylcholines) that changed gradually with disease severity. These consisted of the increase of inter acyl-alkyl-phosphatidylcholines ratios but also changes in the sphingomyelins to phosphatidylcholines ratios. Sphingomyelins are a group of sphingolipids found in mammalian cell membranes and especially membranes that surround nerve cell axons [56]. A decreased ratio of sphingomyelins to lipids across disease stages could indicate an increased vulnerability and damage of nerve cell axons. Even though the role of sphingolipids and gangliosides in brain damage has been investigated since the 1970s [57-59] it is not until recently that strong evidence has been presented in support of the role of gangliosides and their biosynthetic genes in autophagic and apoptotic signaling [60,61]. Additionally, the ratio of arginine to carnitine metabolites was among the top results for the two group (HD vs controls) metabolite pair analysis with modest p-gain values. Arginine is a non-essential amino acid that is also a precursor of nitric oxide (NO) a molecule involved in neurotransmission and inflammation, both of which processes are thought to be deregulated in HD [62]. It has been previously postulated that increased dietary l-arginine could accelerate motor symptom and weight loss events in HD models, through changes in cerebral blood flow and the regulation of NO and nitric oxide synthase [63,64]. Furthermore, it has been also shown

that arginine uptake by HD patients separated them in two distinct metabolic profile groups indicative of a complex and idiomorphic function of this molecule across different individuals [65]. Since the levels of arginine were already found higher in HD in the individual metabolite analysis, the increased arginine to carnitine ratio in HD patients could indicate an arginine-concomitant decrease in the levels, of the antioxidant and lipid regulator molecule, of carnitine. This is further supported by the study of Cuturic et.al, who showed that catabolism and chronic anticonvulsant administration in HD institutionalized patients predisposed to low serum carnitine and that supplementation with levocarnitine improved motor and cognitive measures in these patients [66].Finally apart from their potential roles in deregulated HD molecular pathways these ratios could also serve as potential biomarkers of disease progression since by calculating individual metabolite ratios the dataset variation is reduced and the biomarker robustness is increased.

Moreover, we integrated our previously published gene expression data with the current metabolomics dataset from the same cohort. In specific by generating bioinformatics workflow-based metabolite specific gene sets we identified a group of 8 genes that were decreased in phosphatidycholine metabolic pathways and also found deregulated in our HD patients. Three of these transcriptomics deregulated genes (*MTRR*, *PLB1* and *ALDH1B1*) exhibited especially high correlation with specific diacyl and acyl-alkyl phosphatidylcholines that were downregulated in HD in the metabolomics dataset. More specifically, *MTRR* is involved in the proper function of methionine synthase and folate metabolism [67,68]. Mutations in the *MTRR* gene are thought to be responsible for multiple disorders and especially those affected through the deregulation of the folate cycle and homocysteine metabolism [69,70]. In the past, increased levels of plasma total homocysteine have been found in HD patients and it has been hypothesized that these increased homocysteine levels are a contributing factor to neurodegeneration in these patients [71]. The second of the three genes whose expression was highly correlated with metabolite levels, *PLB1* is a membrane-associated phospholipase with phospholipase A2 activity that exhibits preferential hydrolysis at the sn-2 position of diacyl-phospholipids. A recent study by Fonteh et.al. and in Alzheimer's disease patients cerebrospinal fluid has shown that a significant increase in this phospholipase A2 activity accompanies the glycerophospholipid decrease observed in late onset AD patients [72]. This is in agree-

ment with our data since the levels of *PLB1* exhibited an inverse correlation with all 8 of the metabolites that were statistically significantly associated with HD progression. Thus in a similar fashion with the findings of Fonteh et.al the increased *PLB1* phospholipase A2 expression levels in our HD blood samples could be indicative of perturbation of membrane structures with a concomitant disruption of cellular transport and clearances processes as well as a resulting inflammation overactivation [73,74]. Finally, this integrated "-omics" analysis showed that potential pathways affected from the deregulation of the above genes and changed metabolite concentrations were glycerophospholipid biosynthesis, vitamin B12 and folate metabolism. It has been previously shown that low choline and folate levels are interrelated and that the de novo synthesis of phosphatidycholine is insufficicent to maintain choline levels when the levels of the previous two compounds are also low [75]. Low folate has been associated with cardiovascular disease, a pathology that also affects HD patients and according to some surveys is the leading cause of death in patients [76,77]. The administration of choline has been shown to reduce total plasma homocysteine levels [78], an indicative cardiovascular disease risk factor, while folate and vitamin B12 supplementation has been considered as an additional supplementation therapy for many neuropsychiatric disorders [79,80].

Novel findings from our study include the serum upregulation of serine and threonine levels as well as the inverse association of the levels of a group of 8 phospatidylcholine metabolites with disease progression. The lower level of these metabolites support the evidence found regarding altered lipid metabolism in neurodegenerative disorders as well as the use of phosphatidylcholine as a potential therapeutic avenue [81-83]. The increased amino acid level findings are in partial agreement with an older study that also identified increased serine levels but instead in the Broadmann's area 10 of HD patients [84]. On the other hand, these results are in contrast with the findings of a study by Gruber et.al. that reported decreased levels of serine and 4 more amino acids in HD mutation carriers, in plasma samples [85]. Previous studies by Mochel et.al. have identified valine, leucine and isoleucine metabolite levels to be decreased in plasma samples of HD patients versus presymptomatic and control individuals [86,87]. We could not validate this finding in our serum samples using the Biocrates platform. A possible explanation for this limitation could be the different platforms and protocols that were used to mea-

sure the metabolites. Additionally, the differences could be attributed to the different group sizes and the different UHDRS score thresholds that were used to differentiate between presymptomatic, early and mild HD patient groups. Another potential limitation or reason in regard to the disagreement of some our results with previous studies could be that our study was performed using serum samples while the previous studies were performed in plasma. In specific, using the Biocrates platform employed in the current study it has been shown that serum exhibits higher sensitivity than plasma due to the fact that metabolite concentrations are generally higher in serum samples [88,89] . An additional limitation of the study could be the potential effect of drug treatment on the metabolomics profile of the individuals used for the current study, which was not taken into account since this information was not available for all study participants (especially controls). Considering the great disease phenotypic variation and the different progression rates that characterize Huntington's disease mutation carriers our results will require further validation and refinement in even larger groups before they are used in a clinical trial setting. Such additional validation experiments can reduce the intergroup metabolite concentration overlap and clearly define the concentration thresholds that can be used to distinguish between disease progression/stages. Finally, further research would have to be performed to determine if the current metabolic changes are specific for Huntington's disease or might also partly track changes in other similar neuromuscular disorders and could therefore have additional potential diagnostic applicability.

The present study is according to our knowledge the first study that uses a targeted metabolomics approach in peripheral blood serum samples and in such a large cohort of HD patient peripheral blood samples and with so many pre-symptomatic patients. Obtaining a disease specific metabolomic profile of HD could greatly improve our understanding of the disease pathology. Additionally, these profiles can potentially be used for patient screening as well as drug safety and effectiveness assessment. This could allow for earlier diagnosis something which is very important for HD where disease progression rates and clinical evaluation scores can be highly variable. Serum samples can also be collected noninvasively allowing for longitudinal studies as well as their use both in the preclinical and clinical settings. Our findings combined with the reproducibility and standardization of platforms such as the one used in this study demonstrates the potential of metabolomics to identify disease changes as well as prospective disease progression biomarkers.

Acknowledgments

We thank Werner Römisch-Margl, Julia Scarpa and Katharina Faschinger for metabolomics measurements performed at the Helmholtz Zentrum München, Genome Analysis Center, Metabolomics Core Facility.

Compliance with Ethical Standards

Funding: This work was supported by the European Commission 7th Framework Program,[Project no. 261123]; GEUVADIS, [Project no. 201413]; ENGAGE / CMSB and the European Community's Seventh Framework Programme (FP7/2007-2013) [grant no. 2012-305121]; Integrated European -omics research project for diagnosis and therapy in rare neuromuscular and neurodegenerative diseases (NEUROMICS)'; European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 305,444 (RD-Connect), 305,121 (Neuromics)

Disclosure of potential conflicts of interest: The authors declare that they have not conflict of interests in relation to the work described.

Research involving Human Participants and/or Animals: All procedures performed were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Human peripheral blood samples were collected with institutional review approval and after informed consent from all individual participants included in the study.

Reference List

1. The Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. Cell 1993, 72: 971-983.
2. Lalic NM, Maric J, Svetel M, Jotic A, Stefanova E, Lalic K et al.: Glucose homeostasis in Huntington disease: abnormalities in insulin sensitivity and early-phase insulin secretion. Archives of Neurology 2008, 65: 476-480.
3. Zielonka D, Piotrowska I, Marcinkowski JT, Mielcarek M: Skeletal muscle pathology in Huntington's disease. Front Physiol 2014, 5: 380.
4. Ross CA, Aylward EH, Wild EJ, Langbehn DR, Long JD, Warner JH et al.: Huntington disease: natural history, biomarkers and prospects for therapeutics. Nat Rev Neurol 2014, 10: 204-216.
5. Yu D, Pendergraff H, Liu J, Kordasiewicz HB, Cleveland DW, Swayze EE et al.: Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant huntingtin expression. Cell 2012, 150: 895-908.
6. Evers MM, Pepers BA, van Deutekom JC, Mulders SA, den Dunnen JT, Aartsma-Rus A et al.: Targeting several CAG expansion diseases by a single antisense oligonucleotide. PLoS One 2011, 6: e24308.
7. Acuna AI, Esparza M, Kramm C, Beltran FA, Parra AV, Cepeda C et al.: A failure in energy metabolism and antioxidant uptake precede symptoms of Huntington's disease in mice. Nature Communications 2013, 4: 2917.
8. Mochel F, Haller RG: Energy deficit in Huntington disease: why it matters. J Clin Invest 2011, 121: 493-499.
9. Tang CC, Feigin A, Ma Y, Habeck C, Paulsen JS, Leenders KL et al.: Metabolic network as a progression biomarker of premanifest Huntington's disease. J Clin Invest 2013, 123: 4076-4088.
10. Johri A, Chandra A, Beal MF: PGC-1alpha, mitochondrial dysfunction, and Huntington's disease. Free Radic Biol Med 2013, 62: 37-46.
11. Gulati S, Liu Y, Munkacsi AB, Wilcox L, Sturley SL: Sterols and sphingolipids: dynamic duo or partners in crime? Progress in Lipid Research 2010, 49: 353-365.

12. Valenza M, Cattaneo E: Emerging roles for cholesterol in Huntington's disease. Trends in Neurosciences 2011, 34: 474-486.
13. Sipione S, Rigamonti D, Valenza M, Zuccato C, Conti L, Pritchard J et al.: Early transcriptional profiles in huntingtin-inducible striatal cells by microarray analyses. Human Molecular Genetics 2002, 11: 1953-1965.
14. Valenza M, Rigamonti D, Goffredo D, Zuccato C, Fenu S, Jamot L et al.: Dysfunction of the cholesterol biosynthetic pathway in Huntington's disease. Journal of Neuroscience 2005, 25: 9932-9939.
15. Wang R, Ross CA, Cai H, Cong WN, Daimon CM, Carlson OD et al.: Metabolic and hormonal signatures in pre-manifest and manifest Huntington's disease patients. Front Physiol 2014, 5: 231.
16. Underwood BR, Broadhurst D, Dunn WB, Ellis DI, Michell AW, Vacher C et al.: Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. Brain 2006, 129: 877-886.
17. Schnackenberg LK, Beger RD: Metabolomic biomarkers: their role in the critical path. Drug Discov Today Technol 2007, 4: 13-16.
18. Tabrizi SJ, Scahill RI, Owen G, Durr A, Leavitt BR, Roos RA et al.: Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. Lancet Neurol 2013, 12: 637-649.
19. Nishiumi S, Suzuki M, Kobayashi T, Matsubara A, Azuma T, Yoshida M: Metabolomics for biomarker discovery in gastroenterological cancer. Metabolites 2014, 4: 547-571.
20. Wang X, Zhang A, Sun H: Power of metabolomics in diagnosis and biomarker discovery of hepatocellular carcinoma. Hepatology 2013, 57: 2072-2077.
21. Hodges A, Strand AD, Aragaki AK, Kuhn A, Sengstag T, Hughes G et al.: Regional and cellular gene expression changes in human Huntington's disease brain. Human Molecular Genetics 2006, 15: 965-977.
22. Runne H, Regulier E, Kuhn A, Zala D, Gokce O, Perrin V et al.: Dysregulation of gene expression in primary neuron models of Huntington's disease shows that polyglutamine-related effects on the striatal transcriptome may not be dependent on brain circuitry. Journal of Neuroscience 2008, 28: 9723-9731.

23. Diamanti D, Mori E, Incarnato D, Malusa F, Fondelli C, Magnoni L et al.: Whole gene expression profile in blood reveals multiple pathways deregulation in R6/2 mouse model. Biomark Res 2013, 1: 28.
24. Sassone J, Colciago C, Cislaghi G, Silani V, Ciammola A: Huntington's disease: the current state of research with peripheral tissues. Experimental Neurology 2009, 219: 385-397.
25. Romisch-Margl W, Prehn C, Bogumil R, Rohring C, Suhre K, Adamski J: Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics 2012, 8: 133-142.
26. Li J, Ji L: Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb) 2005, 95: 221-227.
27. Cheverud JM: A simple correction for multiple comparisons in interval mapping genome scans. Heredity (Edinb) 2001, 87: 52-58.
28. Nyholt DR: A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. American Journal of Human Genetics 2004, 74: 765-769.
29. Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K et al.: MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods 2015.
30. Xia J, Sinelnikov IV, Han B, Wishart DS: MetaboAnalyst 3.0--making metabolomics more meaningful. Nucleic Acids Research 2015, 43: W251-W257.
31. Suhre K, Meisinger C, Doring A, Altmaier E, Belcredi P, Gieger C et al.: Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. PLoS One 2010, 5: e13953.
32. Mastrokolias A, Ariyurek Y, Goeman JJ, van DE, Roos RA, van der Mast RC et al.: Huntington's disease biomarker progression profile identified by transcriptome sequencing in peripheral blood. European Journal of Human Genetics 2015.
33. Law CW, Chen Y, Shi W, Smyth GK: voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014, 15: R29.
34. Mastrokolias A, den Dunnen JT, van Ommen GB, 't Hoen PA, van Roon-Mom WM: Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA. BMC Genomics 2012, 13: 28.

35. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M: KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Research 2012, 40: D109-D114.
36. Dharuri H, Henneman P, Demirkan A, van Klinken JB, Mook-Kanamori DO, Wang-Sattler R et al.: Automated workflow-based exploitation of pathway databases provides new insights into genetic associations of metabolite profiles. BMC Genomics 2013, 14: 865.
37. Kutmon M, Riutta A, Nunes N, Hanspers K, Willighagen EL, Bohler A et al.: WikiPathways: capturing the full diversity of pathway knowledge. Nucleic Acids Research 2015.
38. Hastings J, de MP, Dekker A, Ennis M, Harsha B, Kale N et al.: The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Research 2013, 41: D456-D463.
39. Petersen AK, Krumsiek J, Wagele B, Theis FJ, Wichmann HE, Gieger C et al.: On the hypothesis-free testing of metabolite ratios in genome-wide and metabolome-wide association studies. BMC Bioinformatics 2012, 13: 120.
40. Gieger C, Geistlinger L, Altmaier E, Hrabe de AM, Kronenberg F, Meitinger T et al.: Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet 2008, 4: e1000282.
41. Ceglarek U, Muller P, Stach B, Buhrdel P, Thiery J, Kiess W: Validation of the phenylalanine/tyrosine ratio determined by tandem mass spectrometry: sensitive newborn screening for phenylketonuria. Clin Chem Lab Med 2002, 40: 693-697.
42. Perdelli F, Cristina ML, Sartini M, Orlando P: Urinary hydroxyproline as a biomarker of effect after exposure to nitrogen dioxide. Toxicology Letters 2002, 134: 319-323.
43. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M: KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Research 2010, 38: D355-D360.
44. Romero P, Wagg J, Green ML, Kaiser D, Krummenacker M, Karp PD: Computational prediction of human metabolic pathways from the complete human genome. Genome Biol 2005, 6: R2.
45. Draisma HH, Pool R, Kobl M, Jansen R, Petersen AK, Vaarhorst AA et al.: Genome-wide association study identifies novel genetic variants contributing to varia-

tion in blood metabolite levels. Nat Commun 2015, 6: 7208.
46. Vouk K, Ribic-Pucelj M, Adamski J, Rizner TL: Altered levels of acylcarnitines, phosphatidylcholines, and sphingomyelins in peritoneal fluid from ovarian endometriosis patients. J Steroid Biochem Mol Biol 2016, 159: 60-69.
47. Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C et al.: A genome-wide perspective of genetic variation in human metabolism. Nat Genet 2010, 42: 137-141.
48. Suhre K, Meisinger C, Doring A, Altmaier E, Belcredi P, Gieger C et al.: Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. PLoS One 2010, 5: e13953.
49. Yet I, Menni C, Shin SY, Mangino M, Soranzo N, Adamski J et al.: Genetic Influences on Metabolite Levels: A Comparison across Metabolomic Platforms. PLoS One 2016, 11: e0153672.
50. Tsang TM, Haselden JN, Holmes E: Metabonomic characterization of the 3-nitropropionic acid rat model of Huntington's disease. Neurochem Res 2009, 34: 1261-1271.
51. Marcucci H, Paoletti L, Jackowski S, Banchio C: Phosphatidylcholine biosynthesis during neuronal differentiation and its role in cell fate determination. J Biol Chem 2010, 285: 25382-25393.
52. Rosenberg GS, Davis KL: The use of cholinergic precursors in neuropsychiatric diseases. Am J Clin Nutr 1982, 36: 709-720.
53. Hardingham GE, Bading H: Synaptic versus extrasynaptic NMDA receptor signalling: implications for neurodegenerative disorders. Nat Rev Neurosci 2010, 11: 682-696.
54. Zilberter Y, Gubkina O, Ivanov AI: A unique array of neuroprotective effects of pyruvate in neuropathology. Front Neurosci 2015, 9: 17.
55. Ryu JK, Kim SU, McLarnon JG: Neuroprotective effects of pyruvate in the quinolinic acid rat model of Huntington's disease. Exp Neurol 2003, 183: 700-704.
56. Slotte JP: Biological functions of sphingomyelins. Prog Lipid Res 2013, 52: 424-437.
57. Wherrett JR, Brown BL: Erythrocyte glycolipids in Huntington's chorea. Neurology 1969, 19: 489-493.
58. Higatsberger MR, Sperk G, Bernheimer H, Shannak KS, Hornykiewicz O: Striatal ganglioside levels in the rat following kainic acid lesions: comparison with Huntington's disease. Exp Brain Res 1981, 44: 93-96.

59. Heipertz R, Pilz H, Scholz W: The fatty acid composition of sphingomyelin from adult human cerebral white matter and changes in childhood, senium and unspecific brain damage. J Neurol 1977, 216: 57-65.
60. Takamura A, Higaki K, Kajimaki K, Otsuka S, Ninomiya H, Matsuda J et al.: Enhanced autophagy and mitochondrial aberrations in murine G(M1)-gangliosidosis. Biochemical and Biophysical Research Communications 2008, 367: 616-622.
61. Desplats PA, Denny CA, Kass KE, Gilmartin T, Head SR, Sutcliffe JG et al.: Glycolipid and ganglioside metabolism imbalances in Huntington's disease. Neurobiology of Disease 2007, 27: 265-277.
62. Andre R, Carty L, Tabrizi SJ: Disruption of immune cell function by mutant huntingtin in Huntington's disease pathogenesis. Curr Opin Pharmacol 2016, 26: 33-38.
63. Deckel AW, Volmer P, Weiner R, Gary KA, Covault J, Sasso D et al.: Dietary arginine alters time of symptom onset in Huntington's disease transgenic mice. Brain Res 2000, 875: 187-195.
64. Deckel AW: Nitric oxide and nitric oxide synthase in Huntington's disease. J Neurosci Res 2001, 64: 99-107.
65. Salvatore E, Rinaldi C, Tucci T, Di ML, Di SC, Savastano S et al.: Growth hormone response to arginine test differentiates between two subgroups of Huntington's disease patients. J Neurol Neurosurg Psychiatry 2011, 82: 543-547.
66. Cuturic M, Abramson RK, Moran RR, Hardin JW, Frank EM, Sellers AA: Serum carnitine levels and levocarnitine supplementation in institutionalized Huntington's disease patients. Neurol Sci 2013, 34: 93-98.
67. Wolthers KR, Lou X, Toogood HS, Leys D, Scrutton NS: Mechanism of coenzyme binding to human methionine synthase reductase revealed through the crystal structure of the FNR-like module and isothermal titration calorimetry. Biochemistry 2007, 46: 11833-11844.
68. Leclerc D, Wilson A, Dumas R, Gafuik C, Song D, Watkins D et al.: Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria. Proc Natl Acad Sci U S A 1998, 95: 3059-3064.
69. Mitchell ES, Conus N, Kaput J: B vitamin polymorphisms and behavior: evidence of associations with neurodevelopment, depression, schizophrenia, bipolar disorder and cognitive decline. Neurosci Biobehav Rev 2014, 47: 307-320.

70. Mandaviya PR, Stolk L, Heil SG: Homocysteine and DNA methylation: a review of animal and human literature. Mol Genet Metab 2014, 113: 243-252.
71. Andrich J, Saft C, Arz A, Schneider B, Agelink MW, Kraus PH et al.: Hyperhomocysteinaemia in treated patients with Huntington's disease homocysteine in HD. Mov Disord 2004, 19: 226-228.
72. Fonteh AN, Chiang J, Cipolla M, Hale J, Diallo F, Chirino A et al.: Alterations in cerebrospinal fluid glycerophospholipids and phospholipase A2 activity in Alzheimer's disease. J Lipid Res 2013, 54: 2884-2897.
73. Stephenson DT, Lemere CA, Selkoe DJ, Clemens JA: Cytosolic phospholipase A2 (cPLA2) immunoreactivity is elevated in Alzheimer's disease brain. Neurobiology of Disease 1996, 3: 51-63.
74. Sun GY, Xu J, Jensen MD, Simonyi A: Phospholipase A2 in the central nervous system: implications for neurodegenerative diseases. Journal of Lipid Research 2004, 45: 205-213.
75. Jacob RA, Jenden DJ, Allman-Farinelli MA, Swendseid ME: Folate nutriture alters choline status of women and men fed low choline diets. Journal of Nutrition 1999, 129: 712-717.
76. Abildtrup M, Shattock M: Cardiac Dysautonomia in Huntington's Disease. Journal of Huntington's disease 2013, 2: 251-261.
77. Mihm MJ, Amann DM, Schanbacher BL, Altschuld RA, Bauer JA, Hoyt KR: Cardiac dysfunction in the R6/2 mouse model of Huntington's disease. Neurobiology of Disease 2007, 25: 297-308.
78. Olthof MR, Brink EJ, Katan MB, Verhoef P: Choline supplemented as phosphatidylcholine decreases fasting and postmethionine-loading plasma homocysteine concentrations in healthy men. American Journal of Clinical Nutrition 2005, 82: 111-117.
79. Stanger O, Fowler B, Piertzik K, Huemer M, Haschke-Becher E, Semmler A et al.: Homocysteine, folate and vitamin B12 in neuropsychiatric diseases: review and treatment recommendations. Expert Rev Neurother 2009, 9: 1393-1412.
80. Kifle L, Ortiz D, Shea TB: Deprivation of folate and B12 increases neurodegeneration beyond that accompanying deprivation of either vitamin alone. J Alzheimers Dis 2009, 16: 533-540.
81. Growdon J: Use of Phosphatidylcholine in Brain Diseases: An Overview. In

Lecithin. 33 edition. Edited by Hanin I, Ansell GB. Springer US; 1987:121-136.
82. Adibhatla RM, Hatcher JF: Altered lipid metabolism in brain injury and disorders. Sub-Cellular Biochemistry 2008, 49: 241-268.
83. Adibhatla RM, Hatcher JF, Dempsey RJ: Lipids and lipidomics in brain injury and diseases. AAPS J 2006, 8: E314-E321.
84. Bonilla E, Prasad AL, Arrieta A: Huntington's disease: studies on brain free amino acids. Life Sciences 1988, 42: 1153-1158.
85. Gruber B, Klaczkow G, Jaworska M, Krzyszton-Russjan J, Anuszewska EL, Zielonka D et al.: Huntington' disease--imbalance of amino acid levels in plasma of patients and mutation carriers. Annals of Agricultural and Environmental Medicine 2013, 20: 779-783.
86. Mochel F, Charles P, Seguin F, Barritault J, Coussieu C, Perin L et al.: Early energy deficit in Huntington disease: identification of a plasma biomarker traceable during disease progression. PLoS One 2007, 2: e647.
87. Mochel F, Benaich S, Rabier D, Durr A: Validation of plasma branched chain amino acids as biomarkers in Huntington disease. Archives of Neurology 2011, 68: 265-267.
88. Yu Z, Kastenmuller G, He Y, Belcredi P, Moller G, Prehn C et al.: Differences between human plasma and serum metabolite profiles. PLoS One 2011, 6: e21230.
89. Kronenberg F, Trenkwalder E, Kronenberg MF, Konig P, Utermann G, Dieplinger H: Influence of hematocrit on the measurement of lipoproteins demonstrated by the example of lipoprotein(a). Kidney International 1998, 54: 1385-1389.

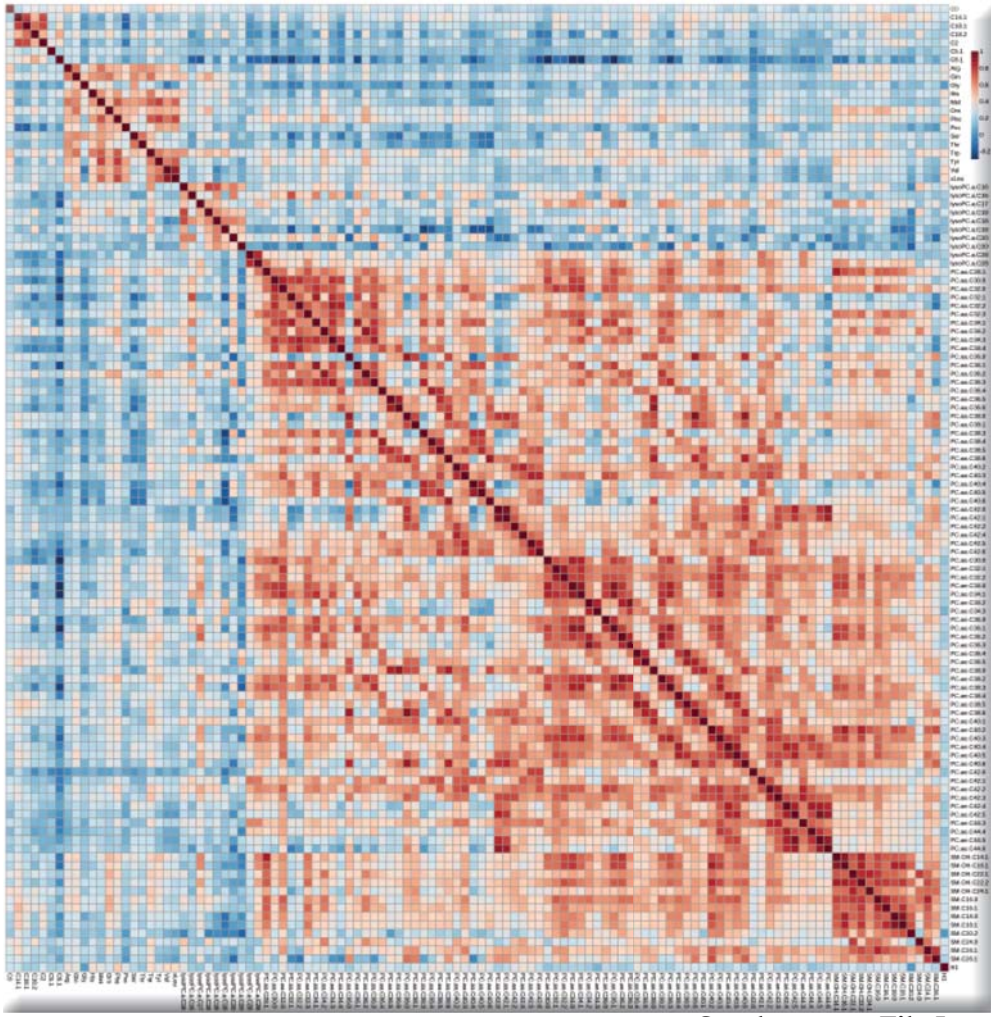
Supplementary Material

Supplementary File 1. Cohort characteristics for the peripheral blood serum samples used for Biocrates targeted metabolomics profiling.

Cohort Groups Total n = 133		Gender (n)	CAG Repeat Mean	Age Mean (Range)[1]	TMS[2] Mean (Range)	TFC[3] Mean (Range)	BMI[4] Mean (Range)
Group 1 (n=36) **Controls** (CAG ≤ 35)		Female (19)	21.1	41.6 (25-65)	1.3 (0-7)	12.7 (11-13)	24.6 (20-32)
		Male (17)	21.3	39 (23-62)	1.1 (0-8)	12.9 (12-13)	25.5 (19-31)
HD	Group 2 (n=29) **Pre-symptomatics** (TMS ≤ 5)	Female (17)	43.1	42.2 (26-60)	2.8 (0-5)	12.1 (8-13)	24.8 (20-33)
		Male (12)	41.6	44.2 (32-60)	2.3 (0-5)	12.3 (7-13)	23.9 (21-27)
	Group 3 (n=31) **Symptomatics** (TMS >5, TFC 13-7)	Female (14)	43.9	52.5 (23-64)	22.4 (8-49)	10.9 (8-13)	23.7 (18-30)
		Male (17)	43.7	51.8 (33-74)	19.1 (7-50)	10.8 (7-13)	25.9 (21-32)
	Group 4 (n=37) **Advanced Symptomatics** (TMS >5, TFC 6-0)	Female (23)	45.1	51.3 (33-76)	50.8 (7-89)	3.1 (0-6)	24.5 (17-32)
		Male (14)	44.9	54.9 (32-69)	51.3 (8-86)	2.4 (0-5)	26.4 (16-33)

[1]Range values have been rounded to the nearest whole number. [2]TMS = Total Motor Score rating (0-124) of the unified Huntington's disease rating scale (UHDRS). [3]TFC = Total Functional Capacity scale of Huntington's disease patients (13-0). [4]BMI = Body mass index

Supplementary File 2. Metabolite correlation heatmap. Color key represents correlation values (Red represents high correlation - Blue represents low correlation). Individual metabolite names are located at the bottom and at the right side of the heatmap.



Supplementary File 5. ▸

Top 10 metabolite pair values from the association analysis of metabolite pair ratios using a 2 group design matrix (controls vs all HDs) and a 4 group linear modelling design matrix (controls and 3 HD disease stages). The metabolites pairs have been sorted according to decreasing p-gain values, resulting from the p-values of the metabolite pairs relative to the smaller of the individual metabolite p-values. Concentration changes were obtained from the fitted data of the metabolite pairs ratios using the full linear statistical model (see methods) and disease state and stage respectively as the main covariate.

Metabolite pair IDs	HD Mutation Carriers vs Controls Analysis	P-gain	Concentration Change
Ser / PC ae C36:0	Serine / Phosphatidylcholine acyl alkyl	4.38E+03	Higher in HD
PC.ae.C36:0 / PC.aa.C32:0	Phosphatidylcholine acyl alkyl /diacyl	2.06E+03	Lower in HD
PC.ae.C36:0 / SM.C16:0	Phosphatidylcholine acyl alkyl / Sphingomyelin	8.62E+02	Lower in HD
PC.ae.C36:0 / Thr	Phosphatidylcholine acyl alkyl / Threonine	5.72E+02	Lower in HD
PC.ae.C44.3 / Thr	Phosphatidylcholine acyl alkyl / Threonine	4.45E+02	Lower in HD
Arg / C0	Arginine / Carnitine	4.05E+02	Higher in HD
PC.ae.C40.4:PC.ae.C44.3	Phosphatidylcholine acyl alkyl / acyl alkyl	1.96E+02	Higher in HD
PC.ae.C42.0:Thr	Phosphatidylcholine acyl alkyl / Threonine	1.83E+02	Lower in HD
PC.ae.C44.3:Ser	Phosphatidylcholine acyl alkyl / Serine	1.80E+02	Lower in HD
PC.aa.C42.1:PC.ae.C44.6	Phosphatidylcholine diacyl / acyl alkyl	1.66E+02	Lower in HD
Metabolite pair IDs	**Disease Progression (4) Group Analysis**	**P-gain**	**Concentration Change**
PC.ae.C36:0 / PC.ae.C34:1	Phosphatidylcholine acyl alkyl /acyl alkyl	1.16E+07	Lower in HD
PC.ae.C36:1 / PC.ae.C36:0	Phosphatidylcholine acyl alkyl /acyl alkyl	7.26E+04	Higher in HD
PC.ae.C40:6 / SM.OH.C16:1	Phosphatidylcholine acyl alkyl / OH Sphingomyelin	2.53E+04	Lower in HD
PC.ae.C40:3 / PC.ae.C42:3	Phosphatidylcholine acyl alkyl /acyl alkyl	2.02E+04	Higher in HD
PC.ae.C44:3 / PC.ae.C40:3	Phosphatidylcholine acyl alkyl /acyl alkyl	1.95E+04	Lower in HD
PC.ae.C36:0 / PC.ae.C38:3	Phosphatidylcholine acyl alkyl /acyl alkyl	1.73E+04	Lower in HD
SM.OH.C16:1 / PC.aa.C38:0	OH Sphingomyelin / Phosphatidylcholine diacyl	1.19E+04	Higher in HD
PC.ae.C30:0 / PC.ae.C36:0	Phosphatidylcholine acyl alkyl / acyl alkyl	1.05E+04	Higher in HD
PC.ae.C42:3 / PC.ae.C40:4	Phosphatidylcholine acyl alkyl / acyl alkyl	1.03E+04	Lower in HD
PC.ae.C42:4 / PC.ae.C42:3	Phosphatidylcholine acyl alkyl / acyl alkyl	1.02E+04	Higher in HD[1]

[1] Median concentration lower in earlier symptomatics (group 3) and relative to pre-symptomatics (group2).

Supplementary File 6. Metabolomics and transcriptomics overlapping pathways reported from the Wikipathways analysis.

Pathway ID	Pathway name	Pathway size	Hits
WP2740	Glycerophospholipid biosynthesis	101	PLB1 - PISD – Serine - 1,2-diacyl-sn-glycero-3-phosphocholine (1+)
WP1880	Phase II conjugation	139	MTRR - Serine - Threonine
WP1533	Vitamin B12 Metabolism	112	MTRR - Serine
WP2650	Arachidonic acid metabolism	102	ALOX5 - 1,2-diacyl-sn-glycero-3-phosphocholine (1+)
WP1857	Metabolism of water-soluble vitamins and cofactors	168	MTRR - Serine
WP2693	Metabolism of amino acids and derivatives	358	MTRR - Serine
WP176	Folate metabolism	132	MTRR - Serine
WP2719	Fc gamma receptor dependent phagocytosis	79	MYO10 - 1,2-diacyl-sn-glycero-3-phosphocholine (1+)
WP15	Selenium Micronutrient Network	188	ALOX5 - Serine