Metabolic signatures of Huntington's disease (HD): 1H NMR analysis of the polar metabolome in post mortem human brain

Graham, S. F., Kumar, P. K., Bjorndahl, T., Han, B., Yilmaz, A., Sherman, E., ... Green, B. D. (2016). Metabolic signatures of Huntington's disease (HD): 1H NMR analysis of the polar metabolome in post mortem human brain. BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR BASIS OF DISEASE, 1862(9), 1675-1684. DOI: 10.1016/j.bbadis.2016.06.007

Published in:
BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR BASIS OF DISEASE

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Metabolic signatures of Huntington's disease (HD): ^{1}H NMR analysis of the polar metabolome in post mortem human brain.

Stewart F. Graham[1]*, Praveen K. Kumar[1], Trent Bjorndahl[2], BeomSoo Han[2], Ali Yilmaz[1], Eric Sherman[3], Ray O. Bahado-Singh[1], David Wishart[2], David Mann[4] and Brian D. Green[5].

1. Beaumont Research Institute, Beaumont Health, 3811 W. 13 Mile Road, Royal Oak, MI 48073.
2. Departments of Biological and Computing Sciences, University of Alberta, Edmonton, AB Canada
3. University of Michigan, Ann Arbour MI.
4. Institute of Brain Behavior and Mental Health, University of Manchester, UK
5. Advanced Asset Technology Centre, Institute for Global Food Security, Queen's University Belfast, Stranmillis Road, Belfast, BT9 5AG, UK.

*Corresponding Author

Tel: +1 248-551-2038; Fax: +1 248-551-2947; e-mail: stewart.graham@beaumont.edu

Abbreviations: ^{1}H NMR- Proton nuclear magnetic resonance; AD – Alzheimer's disease; AUC – area under the curve; AUROC – area under the receiver operating curve; BBB – blood brain barrier; BCAA - branched-chain amino acid; CNS - central nervous system; CSF – cerebral spinal fluid; DSS - Sodium 2,2-dimethyl-2-silapentane-5-sulfonate; FDR – false discovery rate; fMRI – functional magnetic resonance imaging; GC-Tof-MS – gas chromatography time of flight mass spectrometry; HD – Huntington's disease; IGF-1 – insulin like growth factor; MRI - magnetic resonance imaging; MAS-NMR – magic angle spinning NMR; MRS - magnetic resonance spectroscopy; ROC – Receiver operating characteristic.

Abstract

Huntington's disease (HD) is an autosomal neurodegenerative disorder affecting approximately 5-10 persons per 100,000 worldwide. The pathophysiology of HD is not fully understood but the age of onset is known to be highly dependent on the number of CAG triplet repeats in the huntingtin gene. Using ^{1}H NMR spectroscopy this study biochemically profiled 39 brain metabolites in post-mortem striatum (n=14) and frontal lobe (n=14) from HD sufferers and controls (n=28). Striatum metabolites were more perturbed with 15 significantly affected in HD cases, compared with only 4 in frontal lobe ($P<0.05$; $q<0.3$). The metabolite which changed most overall was urea which decreased 3.25-fold in striatum ($P<0.01$). Four metabolites were consistently affected in both brain regions. These included the neurotransmitter precursors tyrosine and L-phenylalanine which were significantly depleted by 1.55-1.58-fold and 1.48-1.54-fold in striatum and frontal lobe, respectively ($P=0.02$-0.03). They also included L-leucine which was reduced 1.54-1.69-fold ($P=0.04$-0.09) and *myo*-inositol which was increased 1.26-1.37-fold ($P<0.01$). Logistic regression analyses performed with MetaboAnalyst demonstrated that data obtained from striatum produced models which were profoundly more sensitive and specific than those produced from frontal lobe. The brain metabolite changes uncovered in this first ^{1}H NMR investigation of human HD offer new insights into the disease pathophysiology. Further investigations of striatal metabolite disturbances are clearly warranted.

Keywords: Huntington's disease; metabolomics; ^{1}H NMR; brain; metabolites.

Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder caused by the extension of a CAG repeat at exon 1 of chromosome 4 (4p63) and is clinically characterized by chorea and dystonia, cognitive decline and behavioural changes [1-6]. It affects 30,000 US citizens (1 in every 10,000) and it is estimated that an additional 150,000-200,000 are at greater risk because they have at least one parent with HD [7]. The appearance of symptoms is inversely correlated to the number of CAG repeats, which is also an influential factor in determining the age of HD onset (it is responsible for ~50-70% of the variance) [8]. The unaffected range is $(CAG)_{6-35}$ repeats, alleles of $(CAG)_{>40}$ are considered fully penetrant and these individuals carry a 100 % lifetime risk of developing HD and CAG repeat size with alleles of $(CAG)_{>60}$ causes juvenile onset. Although HD can present itself at any age, the age of onset is typically 40-45 years with death typically occurring 15-20 years after the initial manifestation [3,9-11]. Currently there is no neuroprotective therapy [3] or ultimate "cure" for this debilitating neurodegenerative disease [5,8].

There are major knowledge gaps regarding the underlying biomolecular mechanisms of HD [2,11]. However, there is some evidence that mechanisms contributing to HD pathogenesis include: polyglutamine aggregation and misfolding [12], oxidative stress and mitochondrial dysfunction [9], misregulation of energy expenditure [10], transcriptional deregulation [13,14], excitotoxicity [15,16] and dopamine toxicity [17,18]. Despite research advances in the last two decades there has been no meaningful progress in medical treatments for HD. Few drugs are available for HD treatment and these offer only symptomatic relief (of chorea only) [2]. The most promising research to date has been with co-enzyme Q10 (currently in Phase 3 clinical trial (n=608 participants); 2Care, The Huntington Study Group) which acts in part to enhance mitochondrial anti-oxidative and

free radical scavenging mechanisms [8]. The aim would be to target new treatments to pre-manifest patients as the discovery that changes due to HD happen many years prior to diagnosable onset. Significant technological advances now make it possible to measure, screen and identify thousands of potential biomarkers in biosamples. There is an unprecedented opportunity to identify reliable "state" biomarkers of pre-manifest HD progression that can be used as outcome measures in preventative clinical trials.

Most biomarker research in HD has concentrated on identifying clinical and neuroimaging biomarkers of disease. Clinical biomarkers are standardised clinical tests and rating scales that measure the progression of various characteristics of the HD phenotype, such as cognition and motor deterioration [3]. Data from the full PREDICT-HD study reported that a standardised cognitive tasks (n=51) demonstrate psychomotor processing, emotion recognition and working memory to be very sensitive when differentiating individuals according to time to predicted HD onset [3,19]. However clinical biomarkers are limited when differentiating between symptomatic improvement and progression of the disease [20]. Additionally they seldom provide any information pertaining to the fundamental disease mechanisms or disease pathogenesis, emphasizing the need for additional non-clinical biomarkers [3].

Very few studies have investigated the potential of metabolomics methodologies to discover novel biochemical biomarkers for HD. A range of studies have demonstrated the utility of metabolomic profiling techniques in accurately distinguishing neurodegenerative diseases from healthy controls [9,10,21-30]. Indeed, it has been successful in identifying plasma biomarker panels for the clinical diagnosis of Alzheimer's disease (AD) in individuals with amnestic mild cognitive impairment [31]. However there is a significant paucity of reliable "state" biomarkers which accurately discriminates pre-manifest HD from manifest HD. The majority of HD

metabolomics experiments have been conducted with rodent models which mimic some of the pathology of human HD. For instance Tsang et al., (2005) used proton nuclear magnetic resonance (^{1}H NMR) and magic angle spinning NMR (MAS-NMR) to discriminate between R6/2 HD transgenic mice and wild-type controls. In this study they analysed skeletal tissue, post-mortem (PM) brain, serum and urine from mice aged 4, 8 and 12 weeks. They highlighted metabolite differences and potential pathways that may be affected [11]. Underwood et al., (2006) applied GC-Tof-MS metabolite profiling techniques to serum samples from human HD patients (prodromal) and a transgenic mouse model in their search for biomarkers. They identified 1275 metabolite peaks but none of their predictive models reached statistical significance. However, they did find that fatty acid β-oxidation and nucleic acid breakdown were commonly affected in human and murine models [10]. Verwaest et al., (2011) applied ^{1}H NMR metabolomics to study the difference between transgenic mice and WT-control litter mates using CSF and serum. They produced multivariate models which distinguished between transgenic mice and WT controls with 84.9% and 72.73% predictive power for serum and CSF, respectively. In addition they produced support vector machine models; one of which was capable of differentiating between transgenic mice and WT controls with a receiver operating characteristics (ROC) value of 0.71 for serum. Unfortunately no significant differences where observed within the SVM model created from CSF data, but the study did suggest that mitochondrial energy dysfunction occurs in HD [9]. Chang et al., (2011) applied GC-Tof-MS metabolomic profiling to the plasma and brain tissue of the 3-NP early stage HD rat model (proposed as a model of pre-manifest HD). They produced predictive models which weakly differentiating transgenic mice from WT-controls with 52.4% and 30.2 % accuracy for brain and plasma, respectively [29]. Having reviewed the current literature (described above) we can conclude that progress in this research field has been

hampered by a lack of studies involving human HD specimens. Therefore, we undertook ^{1}H NMR biochemical profiling of the polar metabolome of post-mortem human brain from two different regions (frontal lobe and striatum) from HD patients and aged-matched control subjects. The aim was to identify novel CNS biomarkers of HD, and also to discover previously unknown fronto-striatal perturbations associated with the onset of HD.

Materials and Methods

Samples

Brain tissue specimens (frontal lobe and striatum) were obtained from post-mortem HD cases (n=14) and also from control subjects (n=14) with no apparent Huntington's pathology. All HD cases showed a moderately to severely atrophied corpus striatum consistent with grades 2 or 3. Exact CAG repeat numbers were not available; the clinical diagnosis of HD was confirmed by genetic testing in all cases, except cases BBN_3211 and BBN_6070. Diagnosis of HD in these instances was made by the presence of ubiquitinated/p62 positive intra-nuclear inclusions within cortical and striatal neurons. All other HD cases also demonstrated such inclusions. None were observed in the control cases. Details such as Vonsattel grading, age, gender, race and post-mortem delay can be found in Supplementary Table 1. Tissues were obtained from the University of Manchester Brain and Tissue Bank.

Sample Preparation

Frozen tissue samples (~5 g) were lyophilized (Christ Freeze Dryer, IMA Life, USA) and milled to a fine powder (Freezer/Mill 6870, Spex Sample Prep, USA) and 50 mg (±0.5 mg) was added to 500 µL of 50% methanol/water in a 2 mL sterile Eppendorf tube. The samples were mixed for 10 min, sonicated for 20 min and the protein removed by centrifugation at 13,000 *g* at 4 °C for 20 min [21,22]. Supernatants were collected, dried under vacuum using a Savant DNA Speedvac (Thermo Scientific, USA) and reconstituted in 285 µl of 50 mM sodium phosphate buffer (pH 7.0), 30 µl of Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) and 35 µl of D_2O [32]. 200 µl of sample was transferred to a 3 mm Bruker NMR tube for NMR analysis. All samples were housed at 4°C in a thermostatically controlled SampleJet autosampler (Bruker-Biospin, USA).

Prior to analysis by NMR, samples were heated to room temperature over a 3 min before being transferred to the magnet.

NMR Analysis

All ^{1}H-NMR experiments were recorded at 300.0 (±0.05) K on a Bruker Avance III HD 600 MHz spectrometer (Bruker-Biospin, USA) operating at 600.13 MHz equipped with a 5 mm TCI cryoprobe. Using a randomized running order 1D ^{1}H-NMR spectra were acquired using a pulse sequence developed by Ravanbakhsh et al. (2015) [33]. Two hundred and fifty six transients were acquired. Chemical shifts (d) are reported in parts per million (ppm) of the operating frequency. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm, concentration 500 µM) and for quantification, all ^{1}H-NMR spectra were processed and analysed using an in-house version of the Bayesil NMR automation software [33]. Bayesil is a web based system that automatically identifies and quantifies metabolites based on a library of pure compounds.

Statistical Analysis

For the comparisons of each NMR metabolite, a Student's t-test was performed. In the case of non-normal distributions, p-values were calculated based on the Mann-Whitney U test. P-values < 0.05 were considered statistically significant. Bonferroni corrected p-values ($p=0.05$/number of metabolites) were used to correct for multiple comparisons. Multivariate statistical analysis was used to determine if a predictive model could be produced based on the concentrations of the identified metabolites to differentiate between the two brain regions (frontal lobe and striatum) and the controls. Data were log-transformed and Pareto-scaled prior to using Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA). The variable importance in projection (VIP) plot that ranks the metabolites in order of their

importance to a predictive model was generated. The greater the score on the x-axis is the greater the significance of that metabolite to the generated PLS-DA model. The PLS-DA models were subsequently subjected to permutation testing (2,000 iterations) to establish whether the observed discrimination between the groups was statistically significant (p-value < 0.05).

Logistic regression analysis was performed with the generalized log-transformed data. The stepwise variable selection was also utilized for optimizing all the model components. Furthermore, a k-fold cross-validation (CV) technique was used to ensure that the logistic regression models were robust [41]. In k-fold CV, the entire sample data is randomly divided into k equal sized subsets. Of the k subsets, only one subset is used as the validation data for testing the model, and the remaining (k-1) subsets are used as training set to generate the model. This results in predictive biomarker predictive models that are both robust and optimal.

To determine the performance of each generated model, the area under the receiver operating characteristics curve (AUROC or AUC) was calculated with sensitivity and specificity using previously described techniques [34]. A receiver operator characteristic (ROC) curve is plotted with sensitivity values on the Y-axis and the corresponding FPR (1-specificity) on the X-axis. The area under the ROC curve (AUC) indicates the accuracy of a test for correctly distinguishing cases from controls. An AUC=1 indicates perfect discrimination. The 95 % CI for the AUC curves were also calculated. All these analyses were performed using the MetaboAnalyst software [35-38].

Pathway Analysis

Metabolites that were found to be significantly different (p-value<0.05) between HD and controls were analyzed using the pathway topology search tool in Metaboanalyst (v 3.0) [35-37].

The pathway library chosen was for *Homo sapiens* (human) and all compounds in selected pathways were used when referencing the specific metabolome. Fisher's exact test was applied to perform over-representation analysis and "relative betweeness centrality" was chosen for the pathway topology testing. Pathways that had both a Holm adjusted p-value <0.05 and FDR p-value <0.05 were considered to be altered due to HD.

Results

Figure 1 displays a labelled ^{1}H NMR spectrum of the extract taken from the frontal lobe of PM human brain of a HD sufferer. In total 39 metabolites were accurately identified and quantified using ^{1}H NMR. Figure 2a and 2c display the PLS-DA scores plots for both frontal lobe and striatum from controls vs. HD sufferers, respectively. The scores plots show good separation between the two sample sets. Extracts from striatum produced better models with greater separation compared with frontal lobe extracts. Figures 2b and 2d display the variable importance in projection (VIP) plots for both the frontal lobe and striatum discriminant models, respectively. These two VIP plots rank the metabolites in order of their importance to the predictive models. The greater the height on the x-axis the greater the significance of that metabolite to that particular model. In addition following 2000 rounds of permutation testing the probability of each model being statistically significant ($p<0.05$) were $p=0.025$ and $p=0.0004$ for the frontal lobe and striatum, respectively. With these p-values, extracts from the striatum were found to produce better models with increased separation in comparison with frontal lobe extracts.

Table 1 lists the results of the univariate analysis of all the metabolites recorded and quantified in the frontal lobe of controls and HD sufferers. Of the 39 metabolites only 4 reached statistical significance ($p<0.05$; $q<0.3$). These include tyrosine, l-phenylalanine, *myo*-inositol and l-leucine. Using the concentration data a pathway analysis was completed and it found that a total of 5 pathways were disrupted as a result of HD. However when stringent thresholds (Holm Adjusted p-value<0.05; $q<0.05$) were placed on these results only 3 pathways were found to be affected. These included inositol phosphate metabolism, galactose metabolism and ascorbate and aldarate metabolism (Supplementary Table 2).

Table 2 lists the results of the univariate analysis of all the metabolites recorded and quantified in striatum from control and HD sufferers. Of the 39 metabolites 15 were recorded as being significantly different ($p<0.05$; $q<0.3$) between HD and controls. As for the frontal lobe a pathway analysis was performed to determine which biochemical pathways are affected as a direct result of HD in the striatum. A total of 7 pathways were detected as being disrupted as a result of HD; however this number was reduced to 4 when we applied the same thresholds for significance as previously ($P<0.05$; $q<0.05$). The remaining pathways include: Inositol phosphate metabolism, galactose metabolism, glyoxylate and dicarboxylate metabolism and ascorbate and aldarate metabolism (Supplementary Table 3).

PLS-DA analysis was conducted to determine differences between the two brain regions. Figure 3a displays the results of the PLS-DA analysis between control subjects taken from frontal lobe and striatum. As is evident complete separation was achieved ($p=0.003$ following permutation testing; 2,000 repeats). Figure 3c shows the PLS-DA scores plot from HD sufferers taken from both the frontal lobe and striatum ($p<0.001$ following permutation testing; 2,000 repeats). Here also there was complete separation of the two brain regions. Figures 3b and 3d display the VIP plots for the control and HD regions, respectively. As may have been expected different metabolites were responsible for the variation between regions collected from controls and the regions collected from HD patients. Supplementary Table 4 displays the results of the univariate analysis for frontal lobe extracts vs. striatum extracts from controls. Of the 39 metabolites a total of 6 metabolites were deemed to be statistically different ($p<0.05$; $q<0.3$) between the two regions. Supplementary Table 5 displays the univariate analysis for frontal lobe extracts vs. striatum extracts from HD cases. Of the 39 metabolites identified and quantified only 7 were found to be at statistically significantly different concentrations between the two regions.

Figure 4a shows the results of ROC analysis undertaken using the concentrations of tyrosine, L-leucine and L-phenylalanine for control vs. HD data acquired from the frontal lobe extracts. During the process of selecting the best performing model it was deemed necessary to exclude *myo*-inositol from the PLS-DA model. An AUC of 0.752 (0.539-0.92) was achieved and following permutation testing (1000 repeats) a p-value of 0.108 was achieved demonstrating that the model does not reach significance when analysed using the PLS-DA algorithm. Figure 4b, displays the results of the logistic regression model ROC analysis following 10 fold cross validations. After performing the stepwise variable selection with the significantly different metabolites ($p < 0.05$), a logistic regression algorithm was created using the concentrations of tyrosine alone in the frontal extracts. The formula for the logistic regression algorithm is as follows:

The formal equation of the logistic regression model is written as $\text{logit}(\pi) = \beta_0 + \beta_0 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$, where π is the probability of the proportion of HD case in a group, and X_i is the metabolite concentrations as k covariates. $\text{logit}(\pi) = -0.006 - 1.439$ Tyrosine, where 0.43 is the threshold. The performance values for both the ROC analysis and the logistic regression algorithm are available as Tables 3a and 3b, respectively.

Figure 4c displays the results of the ROC analysis calculated using the concentrations of urea, valine, tyrosine and 4-aminobutyrate for control vs HD data acquired using striatum extracts. An AUC of 0.917 (CI: 0.75-1) was calculated and following permutation testing (1000 repeats) for the ROC analysis, the PLS-DA model reached significance (p=0.015). Figure 4d displays the ROC results of the logistic regression model created using the concentrations of tyrosine and urea with a 10-fold cross validation applied. The logistic regression algorithm for these two metabolite concentrations in striatum is:

The formal equation of logistic regression model is written as $\text{logit}(\pi) = -0.594 - 3.068$ Tyrosine $- 2.421$ Urea, where a threshold is 0.43 is applied. The performance values for both the ROC analysis and logistic regression algorithm are available in table 4a and 4b, respectively.

Pathway analysis showed which biochemical pathways differed between the two regions under for Control cases and for HD patients. For control subjects only Butanoate metabolism was found to be significantly different (Holm adjusted P=0.006; fdr=0.006) between the two regions. However, 7 biochemical pathways were found to be affected across the two regions in HD brains, and of these only glycoxylate and dicarboxylate metabolism reached statistical significance (Holm adjusted p=0.0002; q=0.0002). The results are summarised as Supplementary Table 6.

Discussion

This is the first ^{1}H NMR based metabolomic investigation of human HD brain. We analysed specimens collected from two brain regions: the frontal lobe and the striatum. The decision to focus on these regions was based on the fact that fronto-striatal circuitry dysfunction is a recognized clinical feature of HD [39]. Using the ^{1}H NMR data acquired we confidently identified and quantified 39 metabolites in both brain regions. Metabolite concentration data led to the development of two multivariate discriminant models which accurately differentiated between the control and HD specimens for both the frontal lobe and striatum ($p=0.003$ and $p<0.001$ following permutation testing, respectively; 2,000 repeats).

Importantly, we observed different metabolites to be significantly affected in the two brain regions. This heterogeneity demonstrates that the biochemistry in these two anatomical regions is perturbed to different extents after the onset of HD. Analysis of frontal lobe showed that the concentrations of four metabolites were significantly perturbed in HD compared with control extracts. In this instance we have based significance on $p<0.05$ and $q<0.3$ due to the small sample number, and also the small number of metabolites identified and quantified in this study. The metabolites which significantly differed between controls and HD sufferers in the frontal lobe included: tyrosine, L-phenylalanine, *myo*-inositol and L-leucine (Table 1).

Analysis of striatum revealed that 15 metabolites were significantly perturbed in HD compared with control extracts (Table 2). These include: 4-aminobutyrate, glycine, formate, L-glutamic acid, tyrosine, L-phenylalanine, aspartate, inosine, *myo*-inositol, taurine, urea, uracil, L-leucine, valine and niacinamide (Table 2). Only four metabolites tyrosine, L-phenylalanine, *myo*-inositol and L-leucine differed across both brain regions. There was however a degree of consistency in the responses we observed. Tyrosine, L-phenylalanine and L-leucine were always

decreased in HD brain and *myo*-inositol was consistently increased. Tyrosine is an essential amino acid which readily crosses the blood-brain barrier (BBB). Tyrosine is a precursor for the biosynthesis of the neurotransmitters of the sympathetic nervous system (i.e. dopamine, norepinephrine and epinephrine). L-phenylalanine is a precursor of tyrosine, therefore making L-phenylalanine also precursor of these catecholamine neurotransmitters. Hyperactivity of the sympathetic nervous system has been reported in HD sufferers[40], and it could be speculated that increased neurotransmitter biosynthesis leads to depletion of tyrosine and L-phenylalanine precursors in both the frontal lobe and striatum.

L-leucine is one of three essential branched chain amino acids (BCAA) which regulates protein synthesis by activating mTor (mammalian target of rapamycin)[41], increases reutilization of amino acids and reduces protein breakdown (www.HMDB.ca). The results here corroborate earlier findings that L-leucine is lower in HD patients[42,43]. The strong correlation between essential amino acids and IGF-1 has been extensively described [44,45,46]. Interestingly Mochel et al., report a correlation between low BCAA levels and IGF-1[43] with IGF-1 known to activate the serine-threonine Akt pathway to which huntingtin is a substrate[47]. The potential disease relevance is that decreased activation of this particular biochemical pathway has been linked to the neuronal toxicity resulting from the reduced phosphorylation of the mutated huntingtin protein[43].

Myo-inositol is a cyclic polyalcohol playing an important second messenger role (inositol phosphates) in the cell (www.HMDB.ca). *Myo*-inositol is considered to be a strong a marker/indicator of gliosis[48] with the prototypical biochemical change being the increase in glia fibrillary acid protein. This change is common to all forms of brain injury characterized as an increase in astrocyte cell body and its processes [49]. Here, *myo*-inositol is significantly elevated in

HD brain[50] which can be directly correlated to an increase in gliosis as the disease progresses across the frontal-striatal circuits.

Using the acquired data we examined brain region-specific differences in the metabolome of the frontal lobe and the striatum. There was complete separation of frontal lobe and striatum (figures 3a and 3c) both when the controls were selected or the HD cases. However, the top ranking measured metabolites (in terms of VIP scores) differed substantially for the control and HD scores plots (Figures 3b and 3d, respectively). Univariate statistical examination of control frontal lobe and striatum found six metabolites to significantly ($p<0.05$; $q<0.3$) differ between the two brain regions (Supplementary Table 4). These included: 4-aminobutyrate, ethanolamine, inosine, homocitrulline, N-acetylaspartic acid (NAA) and niacinamide. Similarly, for HD cases a total of seven metabolites significantly differed ($p<0.05$; $q<0.3$) between the two brain regions (Supplementary Table 4). These included: formate, L-glutamic acid, ethanolamine, *myo*-inositol, succinate, homocitrulline and NAA. We then eliminated those metabolites which significantly differed between controls and HD. This enabled us to focus entirely on those fronto-stratial metabolite changes which were impacted by HD pathology. These metabolites were formate, *myo*-inositol and succinate which were all increased in striatum, and also L-glutamic acid which was increased in the frontal lobe. Formate or formic acid is an intermediary metabolite under normal metabolic conditions. It plays a role in metabolic acidosis and inhibiting cytochrome oxidase activity (terminal electron in the electron transport chain) leading to cell death by depleting ATP reserves and producing reactive oxygen species. Succinate (the anion of succinic acid) is a component of the citric acid cycle which is capable of donating electrons to the electron transfer chain (www.HMDB.ca). Higher concentrations of both succinate and formate in the striatum could suggest that both the citric acid cycle and electron transfer chain are

significantly reduced in the frontal lobe. Higher concentrations of *myo*-inositol could indicate that gliosis is more prevalent in the striatum than the frontal lobe. L-glutamic acid (glutamate) is the most abundant fast excitatory neurotransmitter in the mammalian nervous system. Lower glutamate levels in the striatum would seem to fit with the higher levels of *myo*-inositol and increased gliosis. Following the synaptic release of neurotransmitters, glia cells restrict diffusion and inactivate and recycle a variety of neurotransmitters to include: glutamate, GABA and catecholamines [51,52]. Decreased concentration of glutamate in the striatum suggests glial damage has occurred. Taken together with the increased concentrations of *myo*-inositol this suggests that there is more gliosis in the striatum than the frontal lobe.

One of the main findings of this study is that HD pathology has a much greater effect on biochemical perturbations in the striatum than it does in the frontal lobe. In addition to the multivariate analysis we undertook logistic regression analyses and this demonstrated that the data obtained from the striatum extracts produced logistic regression models with increased sensitivity and specificity following a 10-fold cross validation (Table 3 and 4; frontal lobe and striatum, respectively). In addition, frontal lobe data (unlike the striatum) produced a PLS-DA model which failed to reach any statistical significance ($p=0.108$) following 1,000 permutation tests. The PLS-DA model created with striatum data did reach statistical significance following cross validation (1000 permutation tests; $p=0.015$) which enabled us to be confident that the algorithm developed using the striatum data is accurate for the identification of controls from HD sufferers based on the concentrations of tyrosine and urea evident from the equation.

Conclusions

This first ^{1}H NMR metabolomics investigation examined how HD affects two regions of the human brain and identified a number of biochemical changes. The metabolite data produces statistical models that accurately discriminate between the striatum of control subjects and HD patients. Metabolites identified here could be considered potential biomarkers for detecting and monitoring HD and perhaps *in vivo* magnetic resonance spectroscopy methodologies could be employed here. The major metabolites which were significantly affected ($p<0.05$) in the frontal lobe of HD specimens were L-leucine, myo-inositol, L-phenylalanine and tyrosine. Those metabolite concentrations significantly different ($p<0.05$) in the striatum of HD specimens were: 4-aminobutyrate, aspartate, formate, L-glutamic acid, glycine, inosine, L-leucine, niacinamide, myo-inositol, L-phenylalanine, taurine, tyrosine, uracil, urea and valine. This study demonstrates the suitability and potential power of applying NMR-based metabolomics protocols for the study of HD. Indeed more metabolomic methodologies could be focused on HD. For instance, there are rodent models of HD available which could allow a longitudinal examination of the changes occurring in blood and brain metabolome. This approach has recently been demonstrated for Alzheimer's disease-like pathology [53]. Also, the use of larger sample sizes and different sample types could enable researchers to identify the earliest signs of disease (pre-manifest HD). In the long-term the discovery of metabolite biomarkers could improve patient stratification and in turn improve clinical trial outcomes which could aid in the development of disease-modifying therapies for HD.

Acknowledgements: The work of the Manchester Brain Bank is supported by Alzheimer's Research UK and Alzheimer's Society through their funding of the Brains for Dementia Research (BDR) Program. Manchester Brain Bank also receives Service Support costs from Medical Research Council. We would also like to thank Mr. Muhammad Bin Nasaruddin from Queen's University, Belfast for accurately preparing the PM brain for analysis.

References

1 Reilmann, R., Rolf, L. H. & Lange, H. W. Decreased plasma alanine and isoleucine in Huntington's disease. *Acta neurologica Scandinavica* **91**, 222-224 (1995).

2 Walker, F. O. Huntington's disease: the road to progress. *The Lancet. Neurology* **12**, 624-625, doi:10.1016/s1474-4422(13)70105-4 (2013).

3 Weir, D. W., Sturrock, A. & Leavitt, B. R. Development of biomarkers for Huntington's disease. *The Lancet. Neurology* **10**, 573-590, doi:10.1016/s1474-4422(11)70070-9 (2011).

4 Dorsey, E. R. *et al.* Natural history of Huntington disease. *JAMA neurology* **70**, 1520-1530, doi:10.1001/jamaneurol.2013.4408 (2013).

5 Beste, C. *et al.* A novel cognitive-neurophysiological state biomarker in premanifest Huntington's disease validated on longitudinal data. *Scientific reports* **3**, 1797, doi:10.1038/srep01797 (2013).

6 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* **72**, 971-983 (1993).

7 Association, A. s. *Huntington's Disease*, <http://www.alz.org/dementia/huntingtons-disease-symptoms.asp> (2012).

8 Ross, C. A. & Shoulson, I. Huntington disease: pathogenesis, biomarkers, and approaches to experimental therapeutics. *Parkinsonism & related disorders* **15 Suppl 3**, S135-138, doi:10.1016/s1353-8020(09)70800-4 (2009).

9 Verwaest, K. A. *et al.* (1)H NMR based metabolomics of CSF and blood serum: a metabolic profile for a transgenic rat model of Huntington disease. *Biochimica et biophysica acta* **1812**, 1371-1379, doi:10.1016/j.bbadis.2011.08.001 (2011).

10 Underwood, B. R. *et al.* Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. *Brain : a journal of neurology* **129**, 877-886, doi:10.1093/brain/awl027 (2006).

11 Tsang, T. M. *et al.* Metabolic characterization of the R6/2 transgenic mouse model of Huntington's disease by high-resolution MAS 1H NMR spectroscopy. *Journal of proteome research* **5**, 483-492, doi:10.1021/pr050244o (2006).

12 Bates, G. Huntingtin aggregation and toxicity in Huntington's disease. *Lancet* **361**, 1642-1644, doi:10.1016/s0140-6736(03)13304-1 (2003).

13 Zucker, B. *et al.* Transcriptional dysregulation in striatal projection- and interneurons in a mouse model of Huntington's disease: neuronal selectivity and potential neuroprotective role of HAP1. *Human molecular genetics* **14**, 179-189, doi:10.1093/hmg/ddi014 (2005).

14 Landles, C. & Bates, G. P. Huntingtin and the molecular pathogenesis of Huntington's disease. Fourth in molecular medicine review series. *EMBO reports* **5**, 958-963, doi:10.1038/sj.embor.7400250 (2004).

15 Zeron, M. M. *et al.* Increased sensitivity to N-methyl-D-aspartate receptor-mediated excitotoxicity in a mouse model of Huntington's disease. *Neuron* **33**, 849-860 (2002).

16 Zeron, M. M. *et al.* Mutant huntingtin enhances excitotoxic cell death. *Molecular and cellular neurosciences* **17**, 41-53, doi:10.1006/mcne.2000.0909 (2001).

17 Cyr, M., Sotnikova, T. D., Gainetdinov, R. R. & Caron, M. G. Dopamine enhances motor and neuropathological consequences of polyglutamine expanded huntingtin. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **20**, 2541-2543, doi:10.1096/fj.06-6533fje (2006).

18 Charvin, D., Vanhoutte, P., Pages, C., Borrelli, E. & Caboche, J. Unraveling a role for dopamine in Huntington's disease: the dual role of reactive oxygen species and D2 receptor stimulation.

Proceedings of the National Academy of Sciences of the United States of America **102**, 12218-12223, doi:10.1073/pnas.0502698102 (2005).

19 Stout, J. C. *et al.* Neurocognitive signs in prodromal Huntington disease. *Neuropsychology* **25**, 1-14, doi:10.1037/a0020937 (2011).

20 Henley, S. M., Bates, G. P. & Tabrizi, S. J. Biomarkers for neurodegenerative diseases. *Current opinion in neurology* **18**, 698-705 (2005).

21 Graham, S., Holscher, C. & Green, B. Metabolic signatures of human Alzheimer's disease (AD): 1H NMR analysis of the polar metabolome of post-mortem brain tissue. *Metabolomics* **10**, 744-753, doi:10.1007/s11306-013-0610-1 (2014).

22 Graham, S., Holscher, C., McClean, P., Elliott, C. & Green, B. 1H NMR metabolomics investigation of an Alzheimer's disease (AD) mouse model pinpoints important biochemical disturbances in brain and plasma. *Metabolomics* **9**, 974-983, doi:10.1007/s11306-013-0516-y (2013).

23 Graham, S. F. *et al.* Untargeted Metabolomic Analysis of Human Plasma Indicates Differentially Affected Polyamine and L-Arginine Metabolism in Mild Cognitive Impairment Subjects Converting to Alzheimer's Disease. *PLoS ONE* **10**, e0119452, doi:10.1371/journal.pone.0119452 (2015).

24 Graham, S. F. *et al.* Investigation of the human brain metabolome to identify potential markers for early diagnosis and therapeutic targets of Alzheimer's disease. *Analytical chemistry* **85**, 1803-1811, doi:10.1021/ac303163f (2013).

25 Kaddurah-Daouk, R. *et al.* Alterations in metabolic pathways and networks in Alzheimer's disease. *Translational psychiatry* **3**, e244, doi:10.1038/tp.2013.18 (2013).

26 Han, X. *et al.* Metabolomics in early Alzheimer's disease: identification of altered plasma sphingolipidome using shotgun lipidomics. *PLoS One* **6**, e21643, doi:10.1371/journal.pone.0021643 (2011).

27 Inoue, K. *et al.* Metabolic profiling of Alzheimer's disease brains. *Scientific reports* **3**, 2364, doi:10.1038/srep02364 (2013).

28 Motsinger-Reif, A. A. *et al.* Comparing metabolomic and pathologic biomarkers alone and in combination for discriminating Alzheimer's disease from normal cognitive aging. *Acta neuropathologica communications* **1**, 28, doi:10.1186/2051-5960-1-28 (2013).

29 Chang, K. L. *et al.* Metabolic profiling of 3-nitropropionic acid early-stage Huntington's disease rat model using gas chromatography time-of-flight mass spectrometry. *Journal of proteome research* **10**, 2079-2087, doi:10.1021/pr2000336 (2011).

30 Tkac, I. *et al.* Homeostatic adaptations in brain energy metabolism in mouse models of Huntington disease. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* **32**, 1977-1988, doi:10.1038/jcbfm.2012.104 (2012).

31 Mapstone, M. *et al.* Plasma phospholipids identify antecedent memory impairment in older adults. *Nature medicine* **20**, 415-418, doi:10.1038/nm.3466 (2014).

32 Mercier, P., Lewis, M., Chang, D., Baker, D. & Wishart, D. Towards automatic metabolomic profiling of high-resolution one-dimensional proton NMR spectra. *J Biomol NMR* **49**, 307-323, doi:10.1007/s10858-011-9480-x (2011).

33 Ravanbakhsh, S. *et al.* Accurate, fully-automated NMR spectral profiling for metabolomics. *PLoS One* **10**, e0124219, doi:10.1371/journal.pone.0124219 (2015).

34 Xia, J., Broadhurst, D. I., Wilson, M. & Wishart, D. S. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics* **9**, 280-299, doi:10.1007/s11306-012-0482-9 (2013).

35 Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D. & Wishart, D. S. MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. *Nucleic Acids Res* **40**, W127-133, doi:10.1093/nar/gks374 (2012).

36 Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res* **37**, W652-660, doi:10.1093/nar/gkp356 (2009).

37 Xia, J., Sinelnikov, I. V., Han, B. & Wishart, D. S. MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res* **43**, W251-W257, doi:10.1093/nar/gkv380 (2015).

38 Wishart, D. S. Computational approaches to metabolomics. *Methods Mol Biol* **593**, 283-313, doi:10.1007/978-1-60327-194-3_14 (2010).

39 Tabrizi, S. J. *et al.* Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. *The Lancet. Neurology* **11**, 42-53, doi:10.1016/s1474-4422(11)70263-0 (2012).

40 Kobal, J., Meglic, B., Mesec, A. & Peterlin, B. Early sympathetic hyperactivity in Huntington's disease. *European journal of neurology* **11**, 842-848, doi:10.1111/j.1468-1331.2004.00894.x (2004).

41 Flier, J. S. Regulating Energy Balance: The Substrate Strikes Back. *Science* **312**, 861-864, doi:10.1126/science.1127971 (2006).

42 Mochel, F., Benaich, S., Rabier, D. & Durr, A. VAlidation of plasma branched chain amino acids as biomarkers in huntington disease. *Archives of Neurology* **68**, 264-271, doi:10.1001/archneurol.2010.358 (2011).

43 Mochel, F. *et al.* Early Energy Deficit in Huntington Disease: Identification of a Plasma Biomarker Traceable during Disease Progression. *PLoS ONE* **2**, e647, doi:10.1371/journal.pone.0000647 (2007).

44 Straus, D. S. & Takemoto, C. D. Regulation of albumin mRNA in H4 rat hepatoma cells by the availability of essential amino acids. *Biochimica et biophysica acta* **972**, 33-36 (1988).

45 Harp, J. B., Goldstein, S. & Phillips, L. S. Nutrition and somatomedin. XXIII. Molecular regulation of IGF-I by amino acid availability in cultured hepatocytes. *Diabetes* **40**, 95-101 (1991).

46 Thissen, J. P., Ketelslegers, J. M. & Underwood, L. E. Nutritional regulation of the insulin-like growth factors. *Endocrine reviews* **15**, 80-101, doi:10.1210/edrv-15-1-80 (1994).

47 Humbert, S. *et al.* The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt. *Developmental cell* **2**, 831-837 (2002).

48 Ross, B. D. *et al.* In vivo MR spectroscopy of human dementia. *Neuroimaging clinics of North America* **8**, 809-822 (1998).

49 Wu, V. W. & Schwartz, J. P. Cell culture models for reactive gliosis: New perspectives. *Journal of Neuroscience Research* **51**, 675-681, doi:10.1002/(SICI)1097-4547(19980315)51:6<675::AID-JNR2>3.0.CO;2-8 (1998).

50 Sturrock, A. *et al.* A longitudinal study of magnetic resonance spectroscopy Huntington's disease biomarkers. *Movement disorders : official journal of the Movement Disorder Society* **30**, 393-401, doi:10.1002/mds.26118 (2015).

51 Ortinski, P. I. *et al.* Selective induction of astrocytic gliosis generates deficits in neuronal inhibition. *Nature neuroscience* **13**, 584-591, doi:10.1038/nn.2535 (2010).

52 Bak, L. K., Schousboe, A. & Waagepetersen, H. S. The glutamate/GABA-glutamine cycle: aspects of transport, neurotransmitter homeostasis and ammonia transfer. *J Neurochem* **98**, 641-653, doi:10.1111/j.1471-4159.2006.03913.x (2006).

53 Pan, X. *et al.* Alzheimer's disease-like pathology has transient effects on the brain and blood metabolome. *Neurobiol Aging* **38**, 151-163, doi:10.1016/j.neurobiolaging.2015.11.014 (2016).

Figure Legends

Figure 1. NMR spectroscopy of Huntington Disease (HD) brain extract (Frontal lobe). Typical 1D ^{1}H NMR spectrum of a polar extract taken from HD striatum with identified metabolites labelled in the aliphatic (a) and aromatic regions of the spectrum (b). 1, 1-Methylhistidine; 2, Adenine; 3, Acetic Acid; 4, Ascorbic Acid; 5, Creatine; 6, Glycerophosphocholine; 7, Choline; 8, 4-Aminobutyrate; 9, Glycine; 10, Formate; 11, L-Glutamic Acid; 12, Ethanolamine; 13, Hypozxanthine; 14, Tyrosine; 15, L-Phenylalanine; 16, L-Alanine; 17, L-Threonine; 18, Isoleucine; 19, L-Lactic Acid; 20, Aspartate; 21, Anserine; 22, Inosine; 23, Myo-inositol; 24, Taurine; 25, Succinate; 26, Urea; 27, Uracil; 28, 3-Hydroxybutyric acid; 29, Adenosine triphosphate; 30, L-Glutamine; 31, Homocitrulline; 32, L-Leucine; 33, N-Acetylaspartic Acid; 34, Valine; 35, Niacinamide; 36, Phosphorylcholine; 37, Isobutyric acid; 38, Propylene glycol; 39, Glutathione-Oxidised.

Figure 2. Multivariate comparisons of control and HD. (a) the PLS-DA (showing the separation between groups) scores plot of control (n=14; blue dots) vs. HD (n=14; red crosses) data from the frontal lobe; (b) the VIP plot (showing the metabolites most important for classifying groups) for the frontal lobe data; (c) the PLS-DA scores plot of control (n=14; blue dots) vs. HD (n=14; red crosses) data from the striatum region of the brain; (d) the VIP plot for the striatum region data.

Figure 3. Multivariate comparisons of striatum and frontal lobe. (a) the PLS-DA scores plot of frontal lobe (n=14; blue dots) vs. the striatum region (n=14; red crosses) from the control cases; (b) the VIP plot for the control data; (c) the PLS-DA scores plot of frontal lobe (n=14;

blue circles) vs. striatum region (n=14; red dots) from the HD cases; (d) the VIP plot for the HD cases.

Figure 4. Receiver operating characteristics (ROC) curve analysis of metabolite data. (a) the ROC analysis illustrates the performance of metabolites as biomarkers discriminating control vs. HD in the frontal lobe. AUC: 0.752 (CI: 0.539-0.92); (b) logistic regression ROC analysis of control vs. HD of frontal lobe data following 10-fold cross validations. AUC: 0.745 (CI: 0.558-0.931); (c) the ROC analysis for control vs. HD data acquired from striatum. AUC: 0.917 (CI: 0.75-1.00); (b) logistic regression ROC analysis of control vs. HD from striatum following 10-fold cross validations. AUC: 0.838 (CI: 0.673-1.00).