A systems-level “misunderstanding”: the plasma metabolome in Huntington’s disease

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1002/acn3.214

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:21462478

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
A systems-level “misunderstanding”: the plasma metabolome in Huntington’s disease

Herminia D. Rosas, Gheorghe Doros, Swati Bhasin, Beena Thomas, Sona Gevorkian, Keith Malarick, Wayne Matson & Steven M. Hersch

1Department of Neurology, Boston, Massachusetts
2Center for Neuro-imaging of Aging and Neurodegenerative Diseases, Boston, Massachusetts
3Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, Massachusetts
4Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts
5Department of Biostatistics, School of Public Health, Boston University, Boston, Massachusetts
6Edith Nourse Rogers Memorial Veterans Hospital, Bedford, Massachusetts
7MassGeneral Institutes for Neurodegenerative Disease, Laboratory of Neurodegeneration and Neurotherapeutics, Boston University, Boston, Massachusetts

Correspondence
Herminia D. Rosas, Center for Neuro-imaging of Aging and Neurodegenerative Diseases, 149 13th Street Room 2275, Charlestown, MA 02129. Tel: (617) 726-0658; Fax (617) 724-1227; E-mail: rosas@helix.mgh.harvard.edu

Abstract

Objective: Huntington’s disease (HD) is a rare neurodegenerative disease caused by the expansion of an N-terminal repeat in the huntingtin protein. The protein is expressed in all cells in the body; hence, peripheral tissues, such as blood, may recapitulate processes in the brain. The plasma metabolome may provide a window into active processes that influence brain health and a unique opportunity to noninvasively identify processes that may contribute to neurodegeneration. Alterations in metabolic pathways in brain have been shown to profoundly impact HD. Therefore, identification and quantification of critical metabolomic perturbations could provide novel biomarkers for disease onset and disease progression. Methods: We analyzed the plasma metabolomic profiles from 52 premanifest (PHD), 102 early symptomatic HD, and 140 healthy controls (NC) using liquid chromatography coupled with a highly sensitive electrochemical detection platform. Results: Alterations in tryptophan, tyrosine, purine, and antioxidant pathways were identified, including many related to energetic and oxidative stress and derived from the gut microbiome. Multivariate statistical modeling demonstrated mutually distinct metabolomic profiles, suggesting that the processes that determine onset were likely distinct from those that determine progression. Gut microbiome-derived metabolites particularly differentiated the PHD metabolome, while the symptomatic HD metabolome was increasingly influenced by metabolites that may reflect mutant huntingtin toxicity and neurodegeneration. Interpretation: Understanding the complex changes in the delicate balance of the metabolome and the gut microbiome in HD, and how they relate to disease onset, progression, and phenotypic variability in HD are critical questions for future research.

Introduction

Huntington’s disease (HD) is an autosomal dominant inherited neurodegenerative disorder characterized by progressive motor, psychiatric, cognitive, and metabolic dysfunction. HD is caused by the abnormal expansion of a polymorphic triplet (CAG) repeat in the N-terminus of the Huntington gene leading to an excessive and toxic polyglutamine sequence in the huntingtin protein. The mutant huntingtin protein is expressed ubiquitously throughout the body but causes its greatest harm to neurons, especially in the striatum and cerebral cortex, where dysfunction and neurodegeneration cause the most consequential clinical symptoms of the disease. Aberrant interactions between mutant huntingtin, or its proteolytic fragments, and many other proteins, as well as down-
stream effects have been identified, which collectively play roles in neurodegeneration and which have become therapeutic targets for disease modification. Because HD is highly variable and slowly progressive clinically, there is an urgent need for useful biomarkers to help detect disease activity, monitor progression, and assess the pharmacodynamic effects and potential efficacy of experimental therapies. Since blood is easily and repeatedly accessible clinically and since its collection and processing is readily standardized, we have sought to discover markers of HD in blood that could be useful clinically.

Metabolomics is a global approach to understanding metabolic pathways and metabolic networks, including the precursors and products of all cellular biochemical pathways. The metabolome reflects dynamic interactions between the genome, transcriptome, proteome, and environment and provides information about the chemical “state” at a particular time. Metabolomic profiling has tremendous potential to provide critical information about “when” a system is perturbed, information about “which” specific molecular pathways might be implicated, and about “how” profiles change with disease. These are all difficult questions that remain largely unanswered in HD; identifying affected pathways could provide markers of disease onset or progression and may represent pathogenic pathways that could be targets for treatment and provide pharmacodynamic markers of potential treatments. As the huntingtin protein is present ubiquitously, analyzing the plasma metabolome is a less invasive way of investigating biochemical changes taking place in the presence of the mutant protein that may reflect centrally acting processes.

We therefore applied a targeted approach to metabolomic profiling to identify global biochemical changes in HD in plasma samples derived from a cohort of premanifest subjects (PHD), early symptomatic HD patients (HD), and age- and gender-matched healthy controls (NC). We used high-performance liquid chromatography coupled with highly sensitive electrochemical detection to profile plasma metabolites and focused on tryptophan, tyrosine, and purine pathway constituents. These biochemical pathways have been previously implicated as relevant to neurodegeneration in HD, and may reflect cellular events involving mutant huntingtin, oxidative stress, inflammation, mitochondrial dysfunction, synaptic dysfunction, and cell death.

Materials and Methods

Patients and sample processing

Blood samples were collected prospectively from 140 healthy controls (NC, F:M 68:72; age 50.8 ± 8.8), 102 patients with early symptomatic HD (HD, F:M 58:44; age 47 ± 8.8; CAG repeat 44.6 ± 2.9), and 52 subjects known to carry the trinucleotide expansion but who were without clinical symptoms (premanifest) of HD (PHD, F:M 33:19; age 43 ± 9.3; CAG 42.2 ± 2.0) at the MGH HD Center of Excellence as part of the REVEAL-HD translational biomarker program. A detailed history was obtained for each subject, including age, medications, and total functional capacity assessment. Procedures were explained and consent obtained according to the Declaration of Helsinki (BMJ 1991; 302:1194). Study protocols were approved by the Partners Human Research Committee.

Blood was collected by venipuncture into tubes containing ethylenediaminetetraacetic acid as an anticoagulant and kept on ice until centrifugation, which occurred within 3 h of collection, first at 1000 g for 10 min to remove red blood cells, and then at 15,800 g for 20 min. The plasma was aliquoted into 500 µL aliquots and subsequently frozen at −80°C until analyzed.

Sample preparation and metabolomics analysis

Plasma samples were prepared for analysis using extraction/precipitation of 125 µL of plasma with 500 µL of acetonitrile/0.4% acetic acid, centrifugation, centrifugal evaporation of the supernatant, and reconstitution in the Liquid Chromatography Electrochemical Array running buffer for injection. Samples were analyzed in a blinded fashion using an established well-validated Liquid Chromatography Electrochemical Array long gradient method. This method has been used previously for defining markers of caloric restriction, signatures of neurodegenerative diseases and response to therapy in depression. Samples were randomized and run along with a standard containing 60 metabolites of interest. Duplicate samples were interleaved for quality control. All chromatograms were time normalized (stretched) to the chromatogram of a sample pool run in the middle of the run (sequence sets of 14) to 0.5 sec for major peaks and 1.5–2 sec for minor peaks. Data were then response normalized to the average pool value for each sequence.

Statistical analyses

Although ~1300 metabolites are resolved on the Liquid Chromatography Electrochemical Array platform, we restricted this initial analysis to the 29 known metabolites related to the pathways of interest (Table S1).

Bivariate analyses of individual compounds were carried out using t-tests and analysis of variance models to test differences in mean metabolite levels between gene-expanded subjects (including both PHD and early HD) as compared to age- and gender-matched healthy controls as...
well as between the PHD and early HD groups. Differences amongst the groups in product/substrate ratios within the pathways as potential indicators of altered enzymatic processes were assessed by calculating their respective Pearson correlations. Correlations between metabolites were obtained by calculating their respective Pearson’s coefficients and are reported along with unadjusted $P$-values based on two-sided Student’s $t$-tests. Heat maps and network graphs are used to graphically represent the correlations within each group.

To further evaluate differences between the metabolomic profile of subjects comprising each of the groups we employed a number of multivariate techniques. The discrimination performance (measured by area under receiver operating curve, ROC) of the metabolite combinations was evaluated using well-validated receiver operator curve methodology. Models were validated internally using bootstrap techniques. The calibration of a prediction model measures how well the predicted probabilities agree with the actual observed outcome. The Hosmer–Lemeshow test was used to evaluate the calibration of the models.

Composite metabolomic profiles were used to construct partial least square discriminant analysis models, k-nearest neighbor hierarchical cluster analysis (including k-NN1 and k-NN3) to evaluate the categorical separation of the three study groups and to identify metabolites that contributed most to discriminating amongst groups. Correlation frequency distribution analysis was used to assess the degree to which relationships among compounds discriminated among the groups. Scoring using correct frequency discriminant analysis of the ratios underlying these correlations among and within pathways asks the question “to what extent do the metabolic relationships and sources among and within these pathways separate one state from another?” This approach differs fundamentally from the information obtained from the degree of separation given by individual compound concentrations using hierarchical cluster analysis, partial least square analysis, or linear discriminant analysis (LDA), for example. It relates both to the underlying enzymatic processes driving metabolic transitions and to the relative efficacy of sources of those compounds.

Results

Participant demographics

The demographic characteristics of the groups were compared; there were no significant differences with respect to age or gender amongst the three groups. The mean CAG repeat length for the mutant Huntington allele in the PHD subjects and symptomatic HD patients did not differ. Nine NC subjects, ten PHD subjects and 31 symptomatic HD patients reported taking selective serotonin reuptake inhibitors (SSRIs) of these, nine NC, seven PHD subjects, and 17 symptomatic HD patients had evidence of plasma serotonin re-uptake inhibitor metabolites at the time of sample acquisition. No other medications (e.g., antihistamines, analgesics, antihypertensives) were used with a frequency or regularity that permitted analysis or modeling and were considered essentially as part of the inherent variability in any human study.

Individual and pathway metabolite analyses to evaluate differences between NC, PHD and early symptomatic HD subjects

We used several multivariate analyses to determine the extent to which the entire data determined the category of an individual and the specific metabolite and relationship amongst them most responsible for the categorization. Table 1 provides a summary of the mean, and standard deviations, metabolite concentrations for each group and comparisons between NC versus PHD subjects, NC versus HD patients, and PHD versus HD patients including both the uncorrected and FDR corrected significance levels. Several metabolites from the tyrosine, tryptophan, purine, and tocopheral pathways demonstrated significant differences between groups, the changes were not necessarily linear between PHD and HD groups; in some cases, metabolites were higher in the PHD and lower in the HD group (or vice versa) as compared to NC, suggesting dynamic changes through the disease progression process. While several of the individual compounds studied did not reach significance when corrected for multiple comparisons, taken together, they compose unique descriptive metabolotypes for PHD and HD subjects.

Tryptophan pathway

Serotonin was reduced in both PHD and HD groups as compared to NC. In contrast, N-acetylserotonin was increased in both groups. 5-hydroxytryptophan, 3-hydroxyanthranilic acid, and kynurenine were significantly reduced in the HD group but were not significantly different in PHD as compared to NC groups; however, differences between PHD and HD groups were significant except for kynurenine. In contrast, 5-hydroxyindoleacetate and N-methyltryptamine were significantly reduced in the PHD group but not significantly different between the HD and NC groups; differences between PHD and HD groups were significant for 5-hydroxyindoleacetate. The tryptophan metabolite, indole-3-propionic acid was significantly reduced in both the PHD and HD groups. There were no significant differences amongst groups in the concentra-
tions of indole-3-acetic acid, 3-hydroxykynurenine or in tryptophan itself.

**Tyrosine pathway**

Tyrosine was increased in the PHD and symptomatic HD groups compared to NC group but the increase was significantly only in the PHD group. Homogentisic acid was significantly increased in the PHD group, but was not significantly different in the HD group. Although 2-hydroxyphenylacetate was higher in the PHD group compared to the NC group, the difference was not significant; in contrast, 2-hydroxyphenylacetate was significantly reduced in symptomatic HD patients. 4-hydroxyphenyllactic acid was significantly reduced in HD in comparison to both NC and PHD subjects. Homovanillic acid was significantly higher in HD but reduced in PHD subjects. There were no differences in the concentrations of 4-hydroxyphenylacetate or vanillylmandelic acid concentrations amongst the groups.

**Purine pathway**

Xanthine was significantly reduced in both HD and PHD groups as compared to the NC group; there was no significant difference between the HD and PHD groups. Xanthosine was significantly lower in the HD group compared to the NC group. Guanosine, hypoxanthine, and paraxanthine were not significantly different between groups. Urate concentrations were significantly lower in HD as compared to the PHD group.

**Tocopherols**

The tyrosine pathway-related antioxidant gamma tocopherol was significantly increased in HD; alpha tocopheral was increased in both HD and PHD subjects.

---

**Table 1. Metabolite concentrations (ng/ml) in controls (NC), premanifest (PHD) and early symptomatic Huntington’s disease (HD).**

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Metabolite</th>
<th>Mean (SD)</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptophan</td>
<td>Tryptophan (TRP)</td>
<td>4225.2 (872.3)</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>Serotonin (5HT)</td>
<td>17.5 (30.8)</td>
<td>0.006**</td>
</tr>
<tr>
<td></td>
<td>N-acetylseryotonin (NASHT)</td>
<td>12.1 (7.31)</td>
<td>&lt;0.001**</td>
</tr>
<tr>
<td></td>
<td>5-hydroxytryptophan (5HTP)</td>
<td>4 (3.2)</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>5-hydroxyindoleacetic acid (5HIAA)</td>
<td>18 (10.2)</td>
<td>0.014**</td>
</tr>
<tr>
<td></td>
<td>Homogentisic acid (HGA)</td>
<td>227 (144.1)</td>
<td>0.243</td>
</tr>
<tr>
<td></td>
<td>Kynurinine (KYN)</td>
<td>149.1 (38.6)</td>
<td>0.537</td>
</tr>
<tr>
<td></td>
<td>3-hydroxyanthranilic acid (3OHAN)</td>
<td>3.3 (2.7)</td>
<td>0.284</td>
</tr>
<tr>
<td></td>
<td>3-hydroxykynurenine (3OHKY)</td>
<td>6.2 (3.8)</td>
<td>0.381</td>
</tr>
<tr>
<td></td>
<td>Indole-3-acetic acid (I3AA)</td>
<td>227 (144.1)</td>
<td>0.081*</td>
</tr>
<tr>
<td></td>
<td>Tyrosine (TYR)</td>
<td>3570.8 (1128.5)</td>
<td>0.023**</td>
</tr>
<tr>
<td></td>
<td>Homovanillic acid (HVA)</td>
<td>21.6 (9.8)</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>Homogentisic acid (HGA)</td>
<td>2.8 (2.45)</td>
<td>0.045**</td>
</tr>
<tr>
<td></td>
<td>Vanillylmandelic acid (VMA)</td>
<td>23.6 (11.04)</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>2-hydroxyphenylacetate (2HPAC)</td>
<td>104.6 (102.3)</td>
<td>0.048**</td>
</tr>
<tr>
<td></td>
<td>3-hydroxyphenylacetate (3HPAC)</td>
<td>124.2 (150.8)</td>
<td>0.862</td>
</tr>
<tr>
<td></td>
<td>4-hydroxyphenylacetate (4HPAC)</td>
<td>79.2 (110.1)</td>
<td>0.429</td>
</tr>
<tr>
<td></td>
<td>4-hydroxyphenyllactic acid (4HPLA)</td>
<td>155 (60.8)</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Purine</td>
<td>Hypoxanthine (HX)</td>
<td>285.3 (399.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Guanosine (GR)</td>
<td>56.9 (51.67)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Xanthine (XAN)</td>
<td>122.6 (61.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Xanthosine (XANTH)</td>
<td>57.7 (19.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urate (URATE)</td>
<td>40,099.7 (1001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Paraxanthine (PARAXAN)</td>
<td>68.9 (60.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antioxidant</td>
<td>Gamma tocopherol (GTOCO)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alpha tocopherol (ATOCO)</td>
<td>2699 (1562.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methionine (MET)</td>
<td>1409.1 (666.5)</td>
</tr>
</tbody>
</table>

*P < 0.1, **P < 0.05, FDR corrected.
Methionine

Methionine was significantly increased in PHD subjects only; this difference was significant as compared to both the NC and HD groups.

Multivariate analysis of the NC, PHD, and HD groups

Receiver operator curves

To determine whether the metabolite profiles discriminate amongst groups, we used Receiver Operator Curves. The area under the curve (AUC) provides the discriminating power of a given variable, or group of variables to differentiate between NC and PHD subjects, between NC and HD patients and between PHD subjects and HD patients (Fig. 1).

The AUC values for sensitivity and specificity (index corrected) for the entire set of metabolites were between 74% and 80%. Finally, to validate the models and obtain an unbiased estimate of future model performance we used the Bootstrap method proposed by Efron. The bootstrap adjusted AUC values are provided in Table 2, further supporting group differences.

Partial least squares-discriminant analysis

We used projection on latent structures discriminate analysis models with one out testing to define the level of categorical separation. Strong group separation was achieved between the groups; PHD versus NC; symptomatic HD versus NC, and PHD versus HD (Fig. 2).

Comparative discriminant models were developed using LDA, hierarchical cluster analysis (kNN1 and kNN3) and LDA; the relationships amongst compounds as discriminators among groups were evaluated using correlation frequency distribution analysis (shown in Fig. 3 as a heat map and in Fig. 4 as a network graph.)

The correlation frequency distribution analysis of these correlations is shown in Figure 5.

The correct classification rates among the three categories NC, PHD subjects, and HD patients are summarized in Table 3, for all discriminant models tested. We validated the performance of each model using one out scoring. The training and validation sets were iterated until each subject was included in a validation set.

Table 4 provides the metabolite ratios that have the strongest effect on discrimination between group pairs. In particular, \( N\)-acetylsertotonin/xanthine, \( N\)-acetylsertotonin/indole lactic-3-acid, 4-hydroxyphenylactic acid/5-hydroxyindoleacetate, \( N\)-acetylsertotonin/kynurenine, 5-hydroxyindoleacetate/N-acetylsertotonin, 5-hydroxytryptophan/
homovanillic acid, indole-3-propionic acid/homovanillic acid, N-acetylserotonin/xanthosine, 4-hydroxyphenyllactic acid/homovanillic acid; indole-3-propionic acid/N-acetylserotonin, urate/homovanillic acid, indole-3-propionic acid/N-acetylserotonin, homovanillic acid/methionine.

Controlling for drug effects

In order to determine if concomitant SSRI’s might have an impact or explain reduced serotonin, as has been previously reported, or might exert an effect on other metabolites in the tyrosine pathway, we independently evaluated the metabolomics profiles of those participants who were on SSRI’s (nine NC, ten PHD subjects, and 31 HD patients). No subject was on Haldol. We were able to identify peaks corresponding to known metabolites for the parent drug in all NC, 65% of PHD subjects who reported taking these medications and in 60% of symptomatic HD patients who were prescribed SSRI’s, suggesting that some of these individuals were not actually taking these medications. We found a reduction in the mean concentration of serotonin (8.68 vs. 6.98 ng/mL) in SSRI users, which was not significant ($P = 0.09$). Removing those individuals from the original analyses did not significantly change the results (Fig. 6).
Discussion

Although the cardinal symptoms of HD are due to neurodegeneration in the central nervous system, mutant huntingtin protein is expressed throughout all tissues. The plasma metabolome integrates the effects of neurodegeneration, as leaked into the plasma; the systemic effects of the widespread expression of the mutant protein; and the global effects of disease on systemic metabolism. We identified distinct alterations in tyrosine, tryptophan, and purine pathways in prodromal and early HD, many related to oxidative stress and cellular metabolism. Selective dysregulation of some pathways and paradoxically increased regulation of other pathways suggests complex alterations in the feedback control of the underlying enzymes, proteins, or genes in HD. Remarkably, the normal control, PHD, and early HD plasma metabolomes were mutually distinct rather than differing along a continuum suggesting differing influences during the prodromal and symptomatic stages of the disease. Many of the metabolites differentiating the control from the PHD and HD metabolomes are highly linked to the gut microbiota, suggesting that the HD genetic mutation favors a distinct microbiome or “enterotype”. Enteric mutant huntingtin or systemic effects of HD on the gut could influence energy homeostasis, the supply of vitamins and metabolites, and neuroimmune function and impact the clinical expression of HD. Our data also suggests that the PHD metabolome is more influenced by the gut microbiome than the HD metabolome, perhaps due to increasing effects of mutant huntingtin (mthtt) toxicity and neurodegeneration.
We found indole-3-propionic acid, a product of tryptophan metabolism through indole-3-pyruvic acid and indole-3-lactic acid, to be significantly reduced in pre-manifest subjects and in HD. Indole-3-propionic acid is a potent hydroxyl radical scavenger normally found in the plasma and cerebrospinal fluid (CSF) that is produced exclusively by the commensal gut bacteria *Clostridium sporogenes*. Its primary metabolite, kynuric acid, was undetectable at levels of pg/mL even after concentrating plasma up to a factor of 20 arguing against increased metabolism. Consumption by covalent binding to oxidized protein residues is also inconsistent with the modest increase of bound indole-3-propionic acid reductions we have observed and suggests an HD microbiome that discriminates against *Clostridium sporogenes*. Interestingly, indole-3-propionic acid, is protective against A-Beta-mediated toxicity in vitro and iron-mediated oxidative damage and is under development as a possible treatment for Alzheimer’s disease. As levels of indole-3-propionic acid, and other circulating antioxidants diminish in HD, the level of the oxidative damage to proteins likely increases.

Quinolinic acid-induced excitotoxic striatal lesions led to the hypothesis that the tryptophan pathway could play an important role in HD pathophysiology. Tryptophan is converted to, kynurenine, which is reduced in HD patients, and thence to kynurenic acid, a glutamate receptor antagonist and endogenous neuroprotectant. Its metabolite, 3-hydroxyanthranilic acid, is reduced in HD, as is the 3-hydroxyanthranilic/anthranilate ratio, as previously reported. This reduction in plasma 3-hydroxyanthranilic acid could be due to modulation of 3-hydroxyanthranilic acid oxidase in favor of reducing the downstream production of quinolinic acid. If brain levels of 3-hydroxyanthranilic acid are also low, kynurenine 3-monoxygenase inhibitors, a potential treatment for HD, may simply shift the production and metabolism of 3-hydroxyanthranilic acid to other pathways and may not have a major effect on quinolinic acid levels. Interestingly, 3-hydroxyanthranilic levels were not correlated with its precursor 3-hydroxykynurenine in any group, suggesting alternate routes in its production.

In controls, kynurenine correlated with its precursor tryptophan and with the subsequent pathway from kynurenine through 3-hydroxykynurenine and 3-hydroxyanthranilic acid. With phenoconversion, kynurenine became more strongly correlated with tryptophan and with its product, 3-hydroxykynurenine (and not 3-hydroxyanthranilic acid) in HD. This suggests that the pathway through 3-hydroxyanthranilic acid to quinolinate, and eventually NAD/NADH, becomes more dominant with symptom onset. Although the kynurenine branch product kynurenic acid was not measured in this study, the shift of regulation is consistent with progres-
sive loss of protection from kynurenic acid and with previous findings of lower levels in HD. Alternatively, 3-hydroxyanthranilic acid could arise from the tryptophan pathway leading through indole-3-acetic acid to anthranilate and to 3-hydroxyanthranilic acid, either through enzymatic processes or direct hydroxyl radical attack on anthranilate. The strong correlation between tryptophan and indole-3-acetic acid in HD patients but not in healthy controls supports this.

Other metabolites comprising the tryptophan pathway were also altered. Serotonin was markedly reduced in both PHD and HD and its primary metabolite 5-hydroxyindoleacetate was significantly reduced in PHD subjects. Since plasma serotonin levels reflect release from enterochromaffin cells in the gut wall, selective uptake by platelets, and metabolism to 5-hydroxyindoleacetate by the liver, the reductions observed are certainly peripheral in origin. Despite the reduction in serotonin, N-acetyl serotonin, a neuroprotective antioxidant metabolite,36 and precursor for melatonin, was significantly increased in both PHD and HD. Reduced metabolism, would be consistent with reports of reduced melatonin release in brain HD,37 even though serum levels of melatonin are not altered in HD.38,39 A more likely explanation is increased production of N-acetyltryptophan in the gut with direct conversion to N-acetyltserotonin. These findings, which were not explained by SSRI’s, may be relevant clinically for HD given the importance of serotonin levels for mood, cognition, appetite, and gastrointestinal motility.

Tyrosine pathway

Alterations in the gut microflora could also underlie the tyrosine elevations that were observed via excess production from the essential dietary amino acid phenylalanine. 2-hydroxyphenylacetaet, which was increased in HD; 3-hydroxyphenylacetaet, which was increased in HD; and 4-hydroxyphenylacetaet, which was reduced in HD, are polyphenol metabolites of phenylacetaet, which is derived from dietary phenylalanine in the gut.40 These polyphenols are largely formed nonenzymatically from direct hydroxyl radical attack on the free amino acid or on residues in proteins that are subsequently degraded. Interestingly, the gut microflora forms these compounds through the microbial styrene degradation pathway by which 2-hydroxyphenylacetaet and 3-hydroxyphenylacetaet are converted to homogentisic acid, which we also find to be increased in premanifest plasma. 4-hydroxyphenyllactic acid, which was reduced progressively from PHD to HD, is a tyrosine metabolite formed by gut microbes, such as lactobacilli41 or consumed in dietary plant material.42 Interestingly, 4-hydroxyphenyllactic acid decreases reactive oxygen species production in both mitochondria and neutrophils and like indole-3-propionic acid, is an antioxidant supplied by the gut microbiome.43 Its diminution in HD could further indicate an HD/gut interaction or a change in mitochondrial status in HD.

Tyrosine metabolism in the gut may also relate to the increases in alpha tocopherol we observed in both premanifest and symptomatic groups and to increases in
gamma tocopherol (vitamin E) we observed in symptomatic HD. The tocopherols are absorbed by the intestine into the lymphatic system and released into the plasma by the liver after consumption in the diet. While most vitamin E in humans originates from ingested plant material in the form of dietary vegetable oil, it is also synthesized by microbes in a pathway through tyrosine and homogentisic acid. We hypothesize that the Huntington enterotype that leads to an increase in tyrosine and homogentisic acid also increases production and absorption of the tocopherols. Further research will be necessary to determine whether an HD enterotype significantly modulates the onset and course and thus might also be a therapeutic target as has been suggested for other disorders.

More widespread alterations in the tyrosine pathway were suggested. HVA, the primary metabolite of dopamine was increased in HD. A significant percentage of the HVA in plasma is believed to originate from the central nervous system. Higher HVA levels in CSF have been correlated with psychotic symptoms in schizophrenia while lower levels have been correlated with depression in HD CSF levels have been reported to be unchanged, so higher levels of HVA in the plasma may represent reduced metabolism in the periphery. Although reduced dopamine beta hydroxylase activity could also explain the increase, there are several reports that plasma dopamine beta hydroxylase activity is normal in HD.

Purine pathway
Alterations in the purine pathway were also significant. While its precursors hypoxanthine and guanosine were unchanged and xanthosine trended down in HD, xanthine was markedly reduced in both PHD and HD groups. In contrast, there were no significant differences between controls and PHD or HD in its metabolite, the systemic antioxidant urate (although there were significant differences between HD and PHD groups). Perhaps urate oxidase activity is increased to maintain levels of urate in HD, resulting in precursor depletion. The purine pathway is also closely linked to cellular energetics and mitochondrial function through adenosine metabolism and ADP/ATP dynamics, which are altered in HD transgenic mouse models. Mitochondrial dysfunction would spill ADP into the production of hypoxanthine, decoupling it from its immediate product xanthine, consistent with the elevated hypoxanthine/xanthine ratio we observed in the premalignant group. Both xanthine and hypoxanthine are normally about 10-fold higher in ventricular CSF than in plasma, so alterations in purine metabolism in brain could also effect peripheral levels. Lower plasma levels of urate have been associated with Parkinson’s disease while higher levels may be associated with slower progression of HD. Inosine treatment to raise urate levels is currently being tested in clinical trials for Parkinson disease. Levels of paraxanthine, the principle metabolite of caffeine were not altered amongst groups.

We observed an increase in plasma methionine in premanifest subjects. This is consistent with other suggestions of altered one-carbon metabolism, such as the increases in DNA and histone methylation. Since there is normally an equilibrium between methionine and cysteine, this finding could also be connected to alterations in L-cysteine levels that have been reported in HD and related to oxidative mechanisms of pathogenesis.

Limitations
This work is based on one of the largest plasma data sets available for HD and is based on the use of several models and while cross-validation and boot-strap methods were used in an effort to validate the models, a follow-up study, using an independent sample set is warranted.

Conclusion
HD is a nervous system disorder that demonstrates complex brain-body perturbations, including dysregulation of bioenergetic pathways and of the gut microbiome. Collectively, our findings identify specific and selective alterations of the plasma metabolome that could provide useful markers of prodromal and symptomatic HD patients; confirm that HD involves important metabolic pathways; demonstrate that mthtt affects systemic as well as central nervous system biochemistry; and identifies significant interactions between HD and the gut that could influence disease onset and progression and have important implications for treatment.

To determine the predictive potential of the compounds, these models were evaluated using leave-one-out cross-validation to assess the predictive performance of these models and to further refine the variables included in prediction analyses. By using cross-validation, we were able to estimate a prediction error from the withheld data used in the validation process. While such an analysis is not as powerful in assessing the true predictive performance of a model, it does provide a reasonable estimate of the predictive performance.

Because of the high dimensional and sparse nature of the data, to try to assess whether the resulting models were better than would be expected by chance, we per-
formed permutation testing to ascribe statistical significance to the resulting models. The best correlation coefficient and AUC for each permuted data set were recorded, and an empirical distribution of model fit statistics was generated across the permuted data sets. The values from the real data analysis were compared to the empirical distribution to generate an empirical P-value.

We also used a network approach to evaluate differences amongst groups. Categorizing disease based on the correlations in the metabolomics network is a different strategy representing a different biological phenomenon from the use of ROCs of partial least squares discriminant analysis (receiver operator curve or partial least squares discriminant analysis), which rely on the levels of compounds in the data set. Correlation networks reflect the aggregate influence of the underlying enzymes and sources of input among categories and the information for discriminating categories is inherent in those correlations.

Receiver operator curve, k-nearest neighbor hierarchical cluster analysis, and LDA analyses were performed using R (R Core Team [2013]). Partial least squares discriminant analyses were performed using Matlab. Correlation frequency distribution analysis was performed using an in house program and approach.60

Acknowledgments
This work was supported by the National Institutes of Health (NS042861, NS058793, AT000613, and FD003359 to H. D. R.; NS058793 to G. D.; NS058793, NS042861, AT000613, and FD003359 to S. M. H.)

Author Contributions
H. D. R. was responsible for the study design, conception and protocol development, clinical assessments, data analyses, and interpretation and authoring the manuscript. G. D. was the study statistician responsible for the data analysis and interpretation and contributed to authoring and reviewing the manuscript and generation of figures. S. B. was involved in the sample preparation, analysis, metabolomic analysis, figure preparation modeling, and reviewing of this manuscript. B. T. was involved in the sample preparation and reviewing of this manuscript. S. G. was involved in the study design, data management, and reviewing of the data. K. M. was involved in the protocol development, data analysis, and reviewing of this manuscript. W. M. was responsible for the metabolomics analysis and contributed to the editing of this manuscript. S. M. H. was the Principal investigator and was responsible for obtaining funding, study design and conception, data interpretation, and played a key role in authoring this manuscript. Other: We are very grateful to the individuals who participated in this study, who so generously contributed their time and energy to this work, and without whom it would not have been possible. We thank the individuals who helped recruit participants for this study, including: Susan Maya, Caleb Dresser, Jessica Meyer, Puja Turakhia, Rachel Goldstein, Lauren Woo, Alex Bender, Angela Hu, and Jennifer Lee and who helped with sample processing, including Samantha Matson.

Conflict of Interest
None declared.

References


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** List of known compounds quantified by the LCECA (Liquid Chromatography Electrochemical Array) platform.