

SCIENTIFIC REPORTS



OPEN

Alterations in the tyrosine and phenylalanine pathways revealed by biochemical profiling in cerebrospinal fluid of Huntington's disease subjects

Stephanie Herman^{1,2}, Valter Niemelä³, Payam Emami Khoonsari¹, Jimmy Sundblom⁴, Joachim Burman³, Anne-Marie Landtblom³, Ola Spjuth², Dag Nyholm³ & Kim Kultima¹

Huntington's disease (HD) is a severe neurological disease leading to psychiatric symptoms, motor impairment and cognitive decline. The disease is caused by a CAG expansion in the huntingtin (HTT) gene, but how this translates into the clinical phenotype of HD remains elusive. Using liquid chromatography mass spectrometry, we analyzed the metabolome of cerebrospinal fluid (CSF) from premanifest and manifest HD subjects as well as control subjects. Inter-group differences revealed that the tyrosine metabolism, including tyrosine, thyroxine, L-DOPA and dopamine, was significantly altered in manifest compared with premanifest HD. These metabolites demonstrated moderate to strong associations to measures of disease severity and symptoms. Thyroxine and dopamine also correlated with the five year risk of onset in premanifest HD subjects. The phenylalanine and the purine metabolisms were also significantly altered, but associated less to disease severity. Decreased levels of lumichrome were commonly found in mutated HTT carriers and the levels correlated with the five year risk of disease onset in premanifest carriers. These biochemical findings demonstrates that the CSF metabolome can be used to characterize molecular pathogenesis occurring in HD, which may be essential for future development of novel HD therapies.

Huntington's disease (HD) is a severe neurodegenerative disease characterized by psychiatric symptoms, motor impairment and cognitive decline. The disease is caused by a polyglutamine (encoded by the nucleotide triplet CAG) expansion on exon 1 of the huntingtin (HTT) gene, which is dominantly inherited¹. CAG expansion does not lead to disease until a threshold is reached. At 36 repeats, there is a risk for development of HD, usually presenting late in life. Individuals with 40 repeats or more will invariably develop the disease. A very large number of CAG repeats (60 or more) usually leads to onset before 20 years of age.

Although the molecular mechanisms remain elusive, HD pathogenesis have been proposed to be associated with a combination of toxic gain of function of the mutant HTT (mHTT) protein and loss of function of the wild-type HTT protein, untranslated RNA toxicity², rising instability and oxidative damage of genomic DNA³ and CAG repeat associated non-ATG translation⁴, where the impaired systems include mitochondrial function, axonal transport, hormone secretion and neurotrophic support^{5,6}. Currently there are no therapies for HD, which typically is fatal within twenty years of onset⁷.

The cerebrospinal fluid (CSF) metabolome is a valuable source to monitor biochemical changes in the central nervous system (CNS) occurring in the events of neurological diseases⁸ and it represents an unexplored compartment in understanding the effects of the HTT mutation in humans. Using metabolomics based methods, disturbed energy metabolism was reported in CSF of premanifest HD transgenic rats compared with wild-type⁹ and in human plasma of HD patients¹⁰. Also in *post mortem* HD brain tissue (frontal lobe and striatum), changes

¹Department of Medical Sciences, Clinical Chemistry, Uppsala University, Uppsala, Sweden. ²Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden. ³Department of Neuroscience, Neurology, Uppsala University, Uppsala, Sweden. ⁴Department of Neuroscience, Neurosurgery, Uppsala University, Uppsala, Sweden. Correspondence and requests for materials should be addressed to K.K. (email: Kim.Kultima@medsci.uu.se)

in energy metabolism as well as phospholipid metabolism was reported^{11,12}. CSF is the closest biological compartment to the CNS (except for the brain itself) and should therefore reflect the pathological changes occurring in the brain more accurately than i.e. plasma or serum¹³.

The human CSF metabolome of mHTT carriers has to the best of our knowledge not been investigated before. To fill this gap, we devised a mass spectrometry (MS) based metabolomics approach to investigate changes in the CSF metabolome of mHTT carriers in comparison to controls, as well as differences between the metabolomes of premanifest and manifest HD subjects. To enable monitoring over time, repeated samples were collected for a subset of the premanifest and manifest HD subjects. Finally, the metabolites underlying the differences were connected to biochemical pathways and to measures of disease severity and symptoms.

Methods

Ethical approval. The study was approved by the Regional Ethical Review Board in Uppsala, Sweden (Dnr 2012/274 and 2013/278). All participants provided written informed consent before any samples were collected.

Enrollment of study participants. Thirteen manifest and 13 premanifest mHTT carriers, and 42 controls were enrolled at Uppsala University Hospital. Premanifest HD was defined as individuals with a mHTT CAG expansion (≥ 36 repeats) and a diagnostic confidence level below four, whereas manifest HD was defined by a similar CAG expansion and diagnostic confidence level of four¹⁴. Eight of the manifest and seven of the premanifest mHTT carriers donated a second sample at follow-up.

Clinical assessment. All mHTT carriers were assessed using the total motor score (TMS), total functional capacity (TFC)¹⁵ as well as the cognitive performance measures: Stroop interference (SI), Stroop color (SC), Stroop word reading (SWR)¹⁶, category verbal fluency (CVF), verbal fluency letters (VFL) and the symbol digit modalities test (SDMT). The number of CAG repeats and age were used to calculate the disease burden score (DBS) for all mHTT carriers and the five year risk of onset for the premanifest HD subjects¹⁷.

CSF handling and storage. Samples were collected and stored in accordance with the guidelines for CSF collection, formed by the BioMS-eu network¹⁸. The control samples were collected by lumbar punctures made in routine health care. Patients were asked to donate 3 mL for research purposes. If consent was given the samples were brought to the laboratory within 30 minutes and then entered a semi-automatic process of spinning, aliquoting and freezing. The samples were centrifuged at 250 g for five minutes, stored in polypropylene tubes in aliquots of 240 μ L at -80°C until analyzed.

Metabolite extraction. CSF samples were thawed on ice and 100 μ L was mixed with 410 μ L ice-cold methanol (MeOH) supplemented with internal standards at a final concentration of 0.25 μ M. The samples were vortexed for 15 s and incubated at -20°C for 30 minutes, followed by centrifugation at 20400 g for 12 minutes at 4°C . The supernatants were transferred to fresh Eppendorf tubes which were dried using a centrifugal vacuum concentrator (overnight).

Upon analysis the dried samples were reconstituted in 100 μ L 5% MeOH, 0.1% formic acid and 94.9% deionized MilliQ water. 10 μ L of each sample was pooled to create a quality control (QC) sample to be injected repeatedly to monitor the MS performance by ensuring that the internal standards were present and in correct magnitude.

Mass spectrometry analysis. The samples were injected twice for duplicates, in a constrained randomized order. A QC injection was done every 16th sample and blank injection every 8th. Finally, a 2-fold serial dilution series ranging from 0.5 to 32.0 μ L QC was injected. In addition, MS/MS analysis was performed on pooled samples containing all samples within the three study groups (premanifest HD, manifest HD and controls) and on the global pool (the QC sample). The analysis was performed on a Thermo Ultimate 3000 HPLC and Thermo Q-Exactive Orbitrap mass spectrometer. Detailed description of the analysis can be found in the Supplementary information.

Quantification and quality assessment. The acquired raw data was converted to an open source format (mzML). Peaks were centroided by *msconvert* from ProteoWizard¹⁹ and preprocessed using the following pipeline within the KNIME platform²⁰: the peak-picked data was quantified by *FeatureFinderMetabo*²¹ and features were linked across samples using *featureLinkerUnlabelledQT*²², allowing 10 s retention time tolerance and 5 ppm mass deviation (performed irrespective of charge state across the samples). The non-default parameters used can be found as Supplementary Table S1.

The quantified data was loaded into the statistical software environment R v3.4.0²³. Blank filtering was performed according to our previously introduced pipeline^{8,24}. A quality filtering procedure was performed using the dilution series keeping only spectral features with at least an absolute Pearson correlation of 0.3 with the injection volume. To stabilize the variance, the intensity values were replaced by its \log_2 value and potential sample outliers were detected and removed by calculating the total ion count (TIC) of each sample. Samples with a TIC less than 60% of the average TIC were seen as outliers and removed from the study. One sample replicate was removed. To further reduce the intensity decay over run time, the spectral features were normalized using normalize cyclic loess from the R-package *limma*²⁵. The in-between-replicate correlation was calculated (minimum replicate correlation achieved was 0.92) and the replicates averaged. Finally, the coverage within metabolites were required to be at least 75% for inclusion and the remaining missing values were replaced by its average feature value.

Metabolite identification. All metabolic features with a 75% coverage were matched against an *in house* library of characterized metabolites using a 15 ppm mass tolerance and a 20 s time window. Metabolites (identified metabolic features) of interest were manually curated on MS/MS level when available. Identities confirmed by m/z and elution time of the pure standards and by MS/MS fragmentation pattern were depicted as verified on validation level 2. Identities confirmed only by m/z and elution time of the pure standards were depicted as verified on validation level 1. Identities of metabolites with available MS/MS fragmentation that did not match the authentic standard were rejected.

Statistical analysis. *Post hoc* comparisons were performed for age, number of CAG repeats, gender, DBS, the measures of disease severity and time intervals between repeated sampling for premanifest and manifest HD subjects in the first and second sampling sets, respectively. For variables not significantly different from normal distribution, Welch's *t*-test was used, otherwise the non-parametric Mann-Whitney test was used. The Shapiro-Wilks Normality test was used to assess normality of variables. For the categorical variable gender, a Chi-squared test was used. To compare symptom severity before and after follow-up, paired *t*-tests were performed comparing the HD symptom metrics including DBS at the first and second sampling occasions. A *p*-value lower than 0.05 was considered statistically significant.

In order to eliminate age as a confounder, age was corrected for using linear detrending based on only controls. The controls had been selected to completely span the age distribution of all mHTT carriers. Metabolic features with a significant age dependence (*p*-value < 0.05) were corrected by fitting a linear regression model (R function *lm*) for each metabolic feature in the controls, with age as the explanatory variable. The age coefficient was extracted from the model and used to correct the metabolic levels in all individuals²⁶.

Principal component analysis (PCA) was performed using the function *pca* from the R package *mixOmics*²⁷ using all metabolic features. To specifically target inter-group differences, supervised multivariate analysis using partial least square discriminant analysis (PLS-DA) was performed on identified metabolites. The data was scaled (zero mean, unit variance) and three PLS-DA models were trained on the first sampling set, comparing two groups at a time using the R package *ropls*²⁸. The most significant variables were obtained using the "Variable Importance in the Projection" (VIP). To assess model performance the quality metrics R^2 and Q^2 were extracted and the area under the receiver operating characteristic (AUROC) curve was computed using the R-package *pROC*²⁹. To ensure reproducibility, VIP scores and AUROC values including the receiver operating characteristic (ROC) curves were collected through a 5-fold cross-validation (CV) repeated ten times, where the number of components were fixed to one. Briefly, the 5-fold CV divides the data into five balanced groups using stratified sampling. Four of these groups are used for training the model, while the fifth is used for validation and performance estimation. The procedure is repeated five times, so that each group may act as validation set. Variables with an average VIP score above 1.0 were seen as significantly altered and subjected to a Welch's *t*-test, accounting for unequal variance and unequal sample size. The *p*-values were adjusted for multiple comparisons using false discovery rate (FDR). A FDR value < 0.1 was considered significant. Finally, paired *t*-tests were done comparing the longitudinal first and second samplings of premanifest and manifest HD subjects respectively. A *p*-value < 0.05 was considered statistically significant.

Altered metabolites (VIP > 1.0) with an available HMDB code from each model were separately subjected to a pathway analysis using MetaboAnalyst³⁰ and the *Homo sapiens* pathway library. The significance levels were based on the hypergeometric test and the relative-betweenness centrality was used to compute the pathway impact. A FDR value < 0.1 was considered statistically significant.

Finally, Spearman's rank correlation analyses were performed between all altered metabolites and the measures of disease severity, including DBS and the five year risk of onset. A *p*-value < 0.05 was considered statistically significant. A clustering analysis was performed on the correlation values of the metabolites using the R function *hclust* (default parameters). The Euclidean distance was used as measure of similarity and computed using the R function *dist*.

Results

Research subjects characteristics. As expected, the manifest HD patients were significantly older compared with the premanifest mHTT carriers, had higher DBS and TMS, lower TFC as well as lower cognitive performance, whereas no significant difference in gender distribution or the number of CAG repeats were present (Table 1, Fig. 1a). No significant differences were neither found in gender distribution of the premanifest and manifest HD subjects compared with the control subjects.

A second CSF sample was collected from seven premanifest and eight manifest HD subjects on average 20 months (range: 10.4–44.6) after the first sample was taken, showing that the premanifest and manifest HD subjects had higher DBS at follow-up. The premanifest subjects also displayed an increased five years risk of onset. No other significant changes were found.

One of the premanifest subjects and six of the manifest patients were on antidepressants. Two of the manifest patients were on antipsychotics. Six controls were on antidepressants (Table 1). At follow-up, one of the premanifest HD subjects had developed manifest HD and started taking antidepressants. In addition, one premanifest and one manifest HD patient had started taking antidepressants.

The CSF metabolome distinguished mHTT carriers from controls. In total, 3701 metabolic features were matched across all samples with a 75% coverage. Out of these, 94 metabolites were successfully identified using an *in house* library. In total, 22% of all metabolic features and 28(30%) of the identified metabolites were age dependent and therefore corrected for age in all subjects.

A PCA based on all metabolic features revealed that the premanifest and manifest HD groups were separated from each other over PC1 and from the controls over PC2 that explained ten and eight percentage of the variance

Cohort	Controls	Premanifest HD	Manifest HD		
<i>n</i>	42	13	13		
Female/Male	27/15	6/7	5/8		
On medication, <i>n</i>	6	1	8		
Age*, median(range)	44(20–74)	34(19–56)	51(30–72)		
CAG, median(range)	n/a	42(40–54)	43(39–49)		
5 year risk of onset, % median(range)	n/a	23(1–46)	n/a		
DBS*, mean(±SD)	n/a	265(±73.7)	385(±62.9)		
TFC*, mean(±SD)	n/a	13(±0.0)	9.5(±3.50)		
TMS*, median(range)	n/a	1(0–4)	29(8–64)		
CVF*, mean(±SD)	n/a	22(±3.7)	12(±5.3)		
VFL*, mean(±SD)	n/a	42(±15.3)	17(±10.1)		
SDMT*, mean(±SD)	n/a	48(±12.5)	20(±9.94)		
SI*, median(range)	n/a	47(38–62)	22(6–36)		
SWR* mean(±SD)	n/a	88(±15.7)	48(±22.2)		
SC*, mean(±SD)	n/a	69(±11.2)	33(±14.9)		
Follow-up		Premanifest HD	Manifest HD		
<i>n</i>		7	8		
Female/Male		3/4	3/5		
On medication, <i>n</i>		3	5		
Transitioned, <i>n</i>		1	n/a		
Sampling interval in months, median(range)		12(11.2–44.6)	28(10.4–44.2)		
Sampling^a		1st	2nd	1st	2nd
5 year risk of onset [§] , % median(range)		14(1–30)	17.5(2–33)	n/a	n/a
DBS [§] , mean(±SD)		254(±69.2)	267(±69.0)	380(±79.6)	396(±86.0)
TFC, mean(±SD)		13(±0.0)	13(±0.0)	9.8(±3.20)	8.0(±3.93)
TMS, median(range)		1.5(0–4)	1.5(0–5)	28(8–64)	32.5(17–60)
CVF, mean(±SD)		22(±3.11)	19(±7.28)	13(±3.31)	12(±7.95)
VFL, mean(±SD)		38(±12.8)	40(±11.1)	16.7(±11.3)	16.9(±12.5)
SDMT, mean(±SD)		46(±17.0)	46(±14.3)	24(±3.99)	24(±11.6)
SI, median(range)		45.5(38–50)	46(34–57)	22(6–31)	19(11–28)
SWR, mean(±SD)		85(±19.6)	87(±28.5)	50(±17.4)	51(±11.9)
SC, mean(±SD)		72(±14.2)	66(±13.6)	36(±10.4)	33(±9.25)

Table 1. Clinical and demographic data including follow-up data of the mHTT carriers. Only antidepressants and antipsychotics have been included in the number of subjects on medication. *A significant difference between premanifest and manifest HD subjects. [§]A significant difference between the 1st and 2nd samplings for both premanifest and manifest HD subjects. ^aThe phenoconverter was excluded in comparisons between 1st and 2nd samplings. DBS: disease burden score; TFC: total functional capacity; 5yrisk: the five year risk of onset; TMS: total motor score; SC: Stroop color; SWR: Stroop word reading; SI: Stroop interference; CVF: category verbal fluency; VFL: verbal fluency letters; SDMT: symbol digit modalities test.

each (Fig. 1b). When including PC4, explaining seven percent variance, the manifest HD patients were clearly separated from the rest, indicating a specific biochemical signature of this group (Fig. 1c,d). We also performed a PCA limited to the 94 identified metabolites, which separated the groups to a lesser degree (see Supplementary Fig. S1).

In order to extract altered metabolites, PLS-DA models were used to compare premanifest HD, manifest HD and controls. The first model (manifest HD vs. controls) achieved quality metrics of $R^2 = 0.76$; $p < 0.05$, $Q^2 = 0.30$; $p < 0.05$ and an average AUROC of $0.72(\pm 0.141)$, Fig. 2a. The second model (manifest vs. premanifest HD) achieved quality metrics of $R^2 = 0.54$; $p < 0.85$, $Q^2 = 0.16$; $p < 0.15$ and an average AUROC of $0.76(\pm 0.173)$, Fig. 2b. Both indicating segregations between the groups. The third model, comparing the premanifest HD subjects with controls were found to be non-predictive ($Q^2 < 0$).

Altered CSF metabolites in manifest HD patients compared with controls. Comparing manifest HD patients with controls, 24 metabolites were significantly altered, Table 2. The univariate tests demonstrated that ten of these were independently altered, of which lumichrome, xanthine, N-acetylproline, isoleucine and L-DOPA remained significant after correcting for multiple comparisons. Xanthine, O-succinyl-homoserine, N-acylproline, glutarylcarnitine, tyrosine, 5-methylcytosine had been corrected for a significant age dependence. In the eight manifest patients, from which a second longitudinal sample was collected, no differences were found between first and follow-up for any of the altered metabolites.

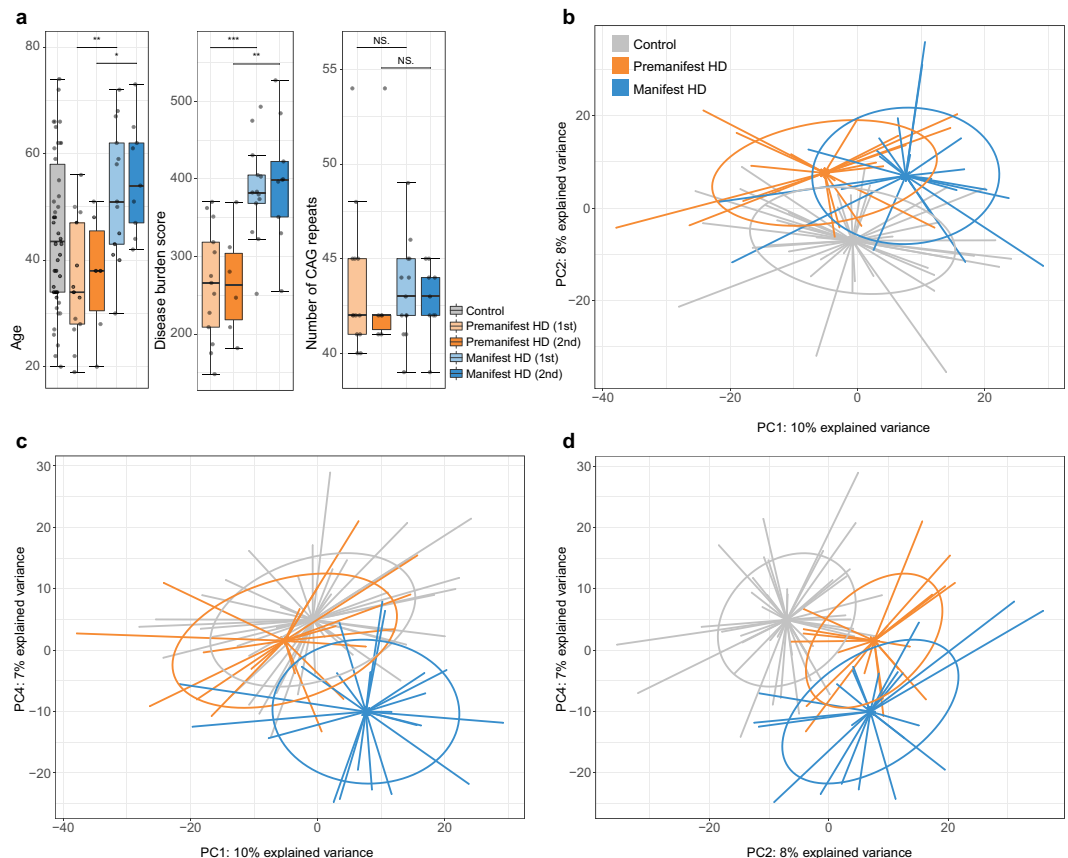


Figure 1. (a) The age, disease burden score and CAG expansion distributions in the study cohort stratified on diagnostic groups and sampling occasion. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. (b–d) Age-adjusted principal component analysis (PCA) on all metabolic features. Separation between all study groups can be seen in (b) PC1 vs. PC2 and (d) PC2 vs. PC4 whereas a clear separation of the manifest HD patients can be seen in (c) PC1 vs. PC4.

The pathway analysis revealed five biochemical pathways that were affected in manifest HD compared with controls: aminoacyl-tRNA biosynthesis; the phenylalanine metabolism; valine, leucine and isoleucine biosynthesis; valine, leucine, isoleucine degradation and the purine metabolism, where the phenylalanine metabolism showed the highest impact (Fig. 2c, Table 3). A complete table of the pathway analysis can be found as Supplementary Table S2.

Altered CSF metabolites in manifest compared with premanifest HD subjects. Comparing manifest HD patients with premanifest subjects revealed 28 metabolites as significantly altered, Table 4. Out of these, 14 were also of importance in distinguishing between manifest HD patients and controls. The univariate tests demonstrated that 11 metabolites were perturbed, of which L-DOPA, xanthine, ophthalmate, creatinine, tyrosine, 5-hydroxytryptophan, adenosine and phenylalanine remained significant after correcting for multiple comparisons. In addition to those previously mentioned, 4-acetamidobutanoate and S-adenosylhomocysteine had been corrected for an age dependence.

There was a significant longitudinal decrease in the levels of O-acetylcarnitine in the manifest HD patients. Significantly altered levels of ophthalmate, phenylalanine, 4-quinolinecarboxylate and N,N,N-trimethyllysine were found in the six premanifest HD subjects (excluding the phenoconverting patient) where a longitudinal sample had been collected.

Using pathway analysis, eight biochemical pathways were significantly perturbed between manifest and premanifest HD subjects: aminoacyl-tRNA biosynthesis; the phenylalanine, tyrosine and tryptophan biosynthesis; valine, leucine and isoleucine biosynthesis; the tyrosine metabolism; the nitrogen metabolism; valine, leucine and isoleucine degradation; the phenylalanine metabolism and the purine metabolism (Fig. 2d, Table 5), where tyrosine metabolism and phenylalanine metabolism showed the highest impacts. A complete table of the pathway analysis can be found as Supplementary Table S3.

Metabolite-to-pathway linkages for all altered metabolites. Disregarding pathway impacts, investigating only linkages between all altered metabolites and biochemical pathways revealed that six metabolites have been connected to purine metabolism and aminoacyl-tRNA biosynthesis respectively, and four metabolites to

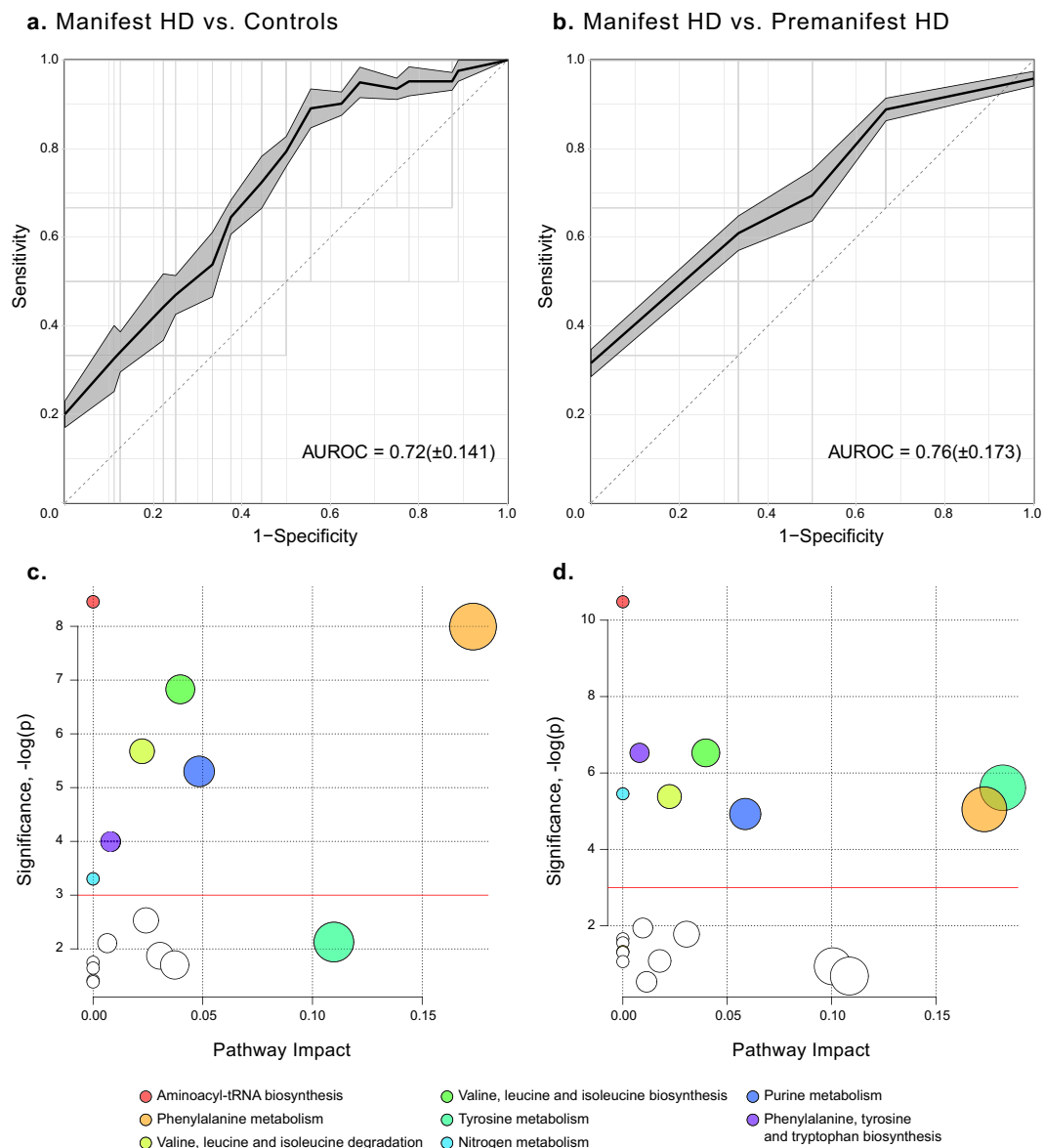


Figure 2. Average ROC curves with corresponding average and standard deviation of the AUROC for the PLS-DA models comparing (a) manifest HD patients with controls and (b) manifest HD patients with premanifest HD subjects. The shadowed areas indicate the standard error of the mean of the sensitivity and 1-specificity. (c,d) Pathway analyses of the altered metabolites between manifest HD compared with (c) controls and (d) premanifest HD. The size of the node indicates the pathway impact (similar to the x-axis) computed by the relative-betweenness centrality and the color corresponds to the corresponding pathway. Pathways that were non-significant in both comparisons have been colored white. The red lines indicate the significance level of a p -value = 0.05.

tyrosine and phenylalanine metabolisms respectively (Fig. 3). Additionally, tryptophan, tyrosine, phenylalanine and valine were recurrent in multiple pathways.

Metabolite association to disease severity. To investigate the effects of disease progression, the associations between CSF levels of altered metabolites and measures of disease severity were investigated in all mHTT carriers and to the five year risk of onset in premanifest HD subjects. The hierarchical clustering revealed a cluster consisting of thyroxine and dopamine from the tyrosine metabolism as well as 5-hydroxytryptophan and ophthalmate that displayed positive correlations to the cognitive measurements and TFC, and negative correlations to DBS, TMS and to the five year risk of onset. Another cluster consisting of creatinine, xanthine, adenine and adenosine from the purine metabolism as well as phenylacetate from the phenylalanine metabolism correlated positively with DBS and TMS and negatively with the cognitive measurements as well as TFC (Fig. 4). Notably, all metabolites from the tyrosine metabolism depicted associations to disease severity and HD symptoms, where

Metabolite	HMDB	VIP mean(95% CI)	log ₂ FC HD-C	p-value	FDR	Validation level
Lumichrome	—	2.56(2.47, 2.64)	-1.02	6.3 × 10⁻³	0.047	2
Xanthine*	HMDB00292	2.33(2.28, 2.38)	0.28	3.9 × 10⁻³	0.047	2
O-succinyl-homoserine*	YMDB00917	1.94(1.83, 2.05)	-0.38	0.043	0.101	1
N-acetylproline*	—	1.90(1.81, 1.99)	0.25	6.0 × 10⁻³	0.047	1
Phenylacetate	HMDB00209	1.81(1.72, 1.90)	0.20	0.030	0.100	2
Isoleucine	HMDB00172	1.81(1.72, 1.89)	-0.51	7.8 × 10⁻³	0.047	2
L-DOPA	HMDB00181	1.77(1.69, 1.86)	-0.34	0.015	0.070	1
Leucine	HMDB00687	1.76(1.66, 1.86)	-0.23	0.033	0.100	2
Corticosterone	HMDB01547	1.52(1.40, 1.64)	-0.70	0.066	0.101	1
Ophthalmate	HMDB05765	1.44(1.35, 1.53)	-0.34	0.065	0.101	1
Salicylate	HMDB01895	1.42(1.33, 1.51)	-0.46	0.031	0.100	1
Phenylalanine	HMDB00159	1.42(1.31, 1.52)	-0.12	0.087	0.116	2
Deoxyuridine	HMDB00012	1.36(1.27, 1.46)	-0.20	0.075	0.106	1
Valine	HMDB00883	1.33(1.24, 1.43)	-0.16	0.059	0.101	2
Creatinine	HMDB00562	1.33(1.23, 1.42)	0.16	0.053	0.101	2
Inosine	HMDB00195	1.32(1.22, 1.43)	0.14	0.147	0.160	2
Cyclic AMP	HMDB00058	1.31(1.19, 1.42)	0.21	0.106	0.127	1
Hypoxanthine	HMDB00157	1.29(1.23, 1.36)	0.20	0.059	0.101	2
Glutarylcarntine*	HMDB13130	1.24(1.16, 1.33)	-0.24	0.114	0.130	2
Phosphocreatine	HMDB01511	1.16(1.08, 1.25)	0.27	0.067	0.101	1
Aldosterone/Cortisone	—/HMDB02802	1.13(1.07, 1.19)	-0.36	0.044	0.101	1
Tyrosine*	HMDB00158	1.13(1.04, 1.21)	-0.11	0.095	0.120	2
5-methylcytosine*	HMDB02894	1.11(1.01, 1.22)	0.11	0.163	0.170	1
1-methyladenosine	HMDB03331	1.02(0.91, 1.13)	-0.45	0.198	0.198	1

Table 2. Altered metabolites with an average VIP score greater than 1.0 from the PLS-DA comparing manifest HD patients (HD) with controls (C). A positive log₂ fold change (FC) HD-C indicates an increase in manifest HD patients compared with controls and vice versa. p-values < 0.05 and FDR values < 0.1 have been bolded. *Metabolites with a demonstrated age dependence that have been corrected. Identities confirmed by m/z and elution time of the internal standards and by MS/MS fragmentation pattern (validation level 2). Identities confirmed by m/z and elution time of the internal standards (validation level 1).

Pathway	Coverage	p-value	FDR	Impact
Aminoacyl-tRNA biosynthesis	5/75	2.1 × 10⁻⁴	0.014	0.0
Phenylalanine metabolism	4/45	3.4 × 10⁻⁴	0.014	0.173
Valine, leucine and isoleucine biosynthesis	3/27	1.1 × 10⁻³	0.029	0.040
Valine, leucine and isoleucine degradation	3/40	3.4 × 10⁻³	0.069	0.022
Purine metabolism	4/92	5.0 × 10⁻³	0.080	0.048
Phenylalanine, tyrosine and tryptophan metabolism	2/27	0.018	0.246	0.008
Nitrogen metabolism	2/39	0.037	0.420	0.0

Table 3. Results from the pathway analysis based on the altered metabolites in manifest HD patients compared with controls. p-values < 0.05 and FDR values < 0.1 have been bolded.

thyroxine in particular showed significant correlations to all measures of disease severity, DBS and the five year risk of onset. The correlation values can be found in Supplementary Table S4.

The CSF levels of creatinine, ophthalmate and 5-hydroxytryptophan correlated with all measures of disease severity and DBS, but not to the five year risk of onset. L-DOPA correlated with all cognitive measurements as well as TFC and TMS, whereas xanthine correlated with all measures of disease severity except for SI and the five year risk of onset. Dopamine displayed significant correlations to the five year risk of onset as well as to TMS, SC, SWR, SI and SDMT. Finally, tyrosine displayed significant correlations to TMS, SI, CVF and SDMT, while lumichrome only displayed a significant correlation to the five year risk of onset in premanifest HD subjects.

Discussion

There is an unmet clinical need for understanding the pathological events occurring in HD and objective markers to monitor disease progression and treatment response are needed. Using CSF as an explorative source, we found multiple biological pathways including tyrosine metabolism, phenylalanine metabolism and purine metabolism that were deranged in manifest HD patients compared with subjects with premanifest HD and controls. In

Metabolite	HMDB	VIP mean (95% CI)	log ₂ FC HD-pHD	p-value	FDR	log ₂ FC HD ₂ -HD ₁	log ₂ FC pHD ₂ -pHD ₁	Validation level
L-DOPA	HMDB00181	2.14 (2.08, 2.20)	-0.53	2.5 × 10⁻³	0.041	n.s.	n.s.	1
Xanthine*	HMDB00292	2.10 (2.05, 2.14)	0.29	3.8 × 10⁻³	0.041	n.s.	n.s.	2
Ophthalmate	HMDB05765	2.05 (1.98, 2.13)	-0.81	4.4 × 10⁻³	0.041	n.s.	-0.44	1
Creatinine	HMDB00562	1.81 (1.75, 1.87)	0.23	0.013	0.082	n.s.	n.s.	2
Tyrosine*	HMDB00158	1.78 (1.70, 1.86)	-0.22	0.015	0.082	n.s.	n.s.	2
5-hydroxytryptophan	HMDB00472	1.74 (1.65, 1.84)	-0.61	0.018	0.084	n.s.	n.s.	1
Adenosine	HMDB00050	1.71 (1.63, 1.78)	0.49	0.021	0.084	n.s.	n.s.	1
Phenylalanine	HMDB00159	1.64 (1.59, 1.69)	-0.18	0.027	0.096	n.s.	0.04	2
Phenylacetate	HMDB00209	1.55 (1.46, 1.65)	0.25	0.034	0.107	n.s.	n.s.	2
DHEAS	—	1.53 (1.47, 1.59)	-0.99	0.040	0.112	n.s.	n.s.	2
Thyroxine	HMDB00248	1.49 (1.40, 1.58)	-0.37	0.045	0.114	n.s.	n.s.	1
Glutarylcarntine*	HMDB13130	1.38 (1.30, 1.46)	-0.33	0.063	0.143	n.s.	n.s.	2
O-succinyl-homoserine*	YMDB00917	1.36 (1.27, 1.46)	-0.39	0.071	0.143	n.s.	n.s.	1
Adenine	HMDB00034	1.35 (1.24, 1.45)	0.25	0.080	0.143	n.s.	n.s.	1
Isoleucine	HMDB00172	1.34 (1.25, 1.43)	-0.32	0.079	0.143	n.s.	n.s.	2
Aldosterone/Cortisone	—/HMDB02802	1.31 (1.21, 1.42)	-0.27	0.082	0.143	n.s.	n.s.	1
4-quinolinecarboxylate	—	1.26 (1.17, 1.36)	-0.22	0.097	0.159	n.s.	-0.25	1
O-acetylcarnitine	HMDB00201	1.20 (1.12, 1.28)	-0.25	0.108	0.164	-0.22	n.s.	1
4-acetamidobutanoate*	HMDB03681	1.18 (1.06, 1.30)	0.19	0.117	0.164	n.s.	n.s.	2
Hypoxanthine	HMDB00157	1.17 (1.09, 1.26)	0.18	0.117	0.164	n.s.	n.s.	2
N,N,N-trimethyllysine	HMDB01325	1.17 (1.07, 1.28)	0.21	0.123	0.164	n.s.	0.18	1
Leucine	HMDB00687	1.13 (1.04, 1.22)	-0.16	0.136	0.166	n.s.	n.s.	2
Dopamine	HMDB00073	1.13 (1.01, 1.25)	-0.28	0.137	0.166	n.s.	n.s.	1
N-acetylalanine	HMDB00766	1.08 (0.98, 1.17)	0.14	0.143	0.166	n.s.	n.s.	1
Tryptophan [M-H]	HMDB00929	1.07 (0.99, 1.16)	-0.19	0.154	0.172	n.s.	n.s.	2
Valine	HMDB00883	1.03 (0.94, 1.13)	-0.11	0.181	0.180	n.s.	n.s.	2
Tryptophan [M + H]	HMDB00929	1.03 (0.94, 1.13)	-0.15	0.170	0.180	n.s.	n.s.	2
S-adenosylhomocysteine*	HMDB00939	1.03 (0.92, 1.14)	0.20	0.175	0.180	n.s.	n.s.	1

Table 4. Altered metabolites with an average VIP score greater than 1.0 from the PLS-DA comparing premanifest (pHD) and manifest (HD) subjects. A positive log₂ fold change (FC) HD-pHD indicates an increase in manifest HD patients compared with premanifest subjects and vice versa. A positive log₂ FC HD₂-HD₁ indicates an increase over time in manifest HD patients and similar for log₂ FC pHD₂-pHD₁ in premanifest HD subjects. Non-significant FC ($p < 0.05$) have been marked with n.s. p -values < 0.05 and FDR values < 0.1 have been bolded. *Metabolites with a demonstrated age dependence that have been corrected. Identities confirmed by m/z and elution time of the internal standards and by MS/MS fragmentation pattern (validation level 2). Identities confirmed by m/z and elution time of the internal standards (validation level 1).

Pathway	Coverage	p-value	FDR	Impact
Aminoacyl-tRNA biosynthesis	6/75	2.8 × 10⁻⁵	2.2 × 10⁻³	0.0
Phenylalanine, tyrosine and tryptophan biosynthesis	3/27	1.5 × 10⁻³	0.039	0.008
Valine, leucine and isoleucine biosynthesis	3/27	1.5 × 10⁻³	0.039	0.040
Tyrosine metabolism	4/76	3.7 × 10⁻³	0.061	0.182
Nitrogen metabolism	3/39	4.3 × 10⁻³	0.061	0.0
Valine, leucine and isoleucine degradation	3/40	4.6 × 10⁻³	0.061	0.022
Phenylalanine metabolism	3/45	6.4 × 10⁻³	0.073	0.173
Purine metabolism	4/92	7.3 × 10⁻³	0.073	0.059

Table 5. Results from the pathway analysis based on the altered metabolites in manifest HD patients compared with premanifest HD subjects. p -values < 0.05 and FDR values < 0.1 have been bolded.

addition, the measures of disease severity enabled us to connect these changes to HD symptoms and the five year risk of onset in premanifest HD.

The tyrosine metabolism, including tyrosine, thyroxine, L-DOPA and dopamine, showed the highest impact based on the pathway analysis and may thus play a central role in disease progression. Thyroxine demonstrated reduced levels in manifest compared with premanifest HD subjects, moderate to strong associations to all measures of disease severity and lower thyroxine levels were associated to an increased five year risk of onset in premanifest HD subject. Importantly, thyroxine was not associated to age in the control group ($r = 0.03$,

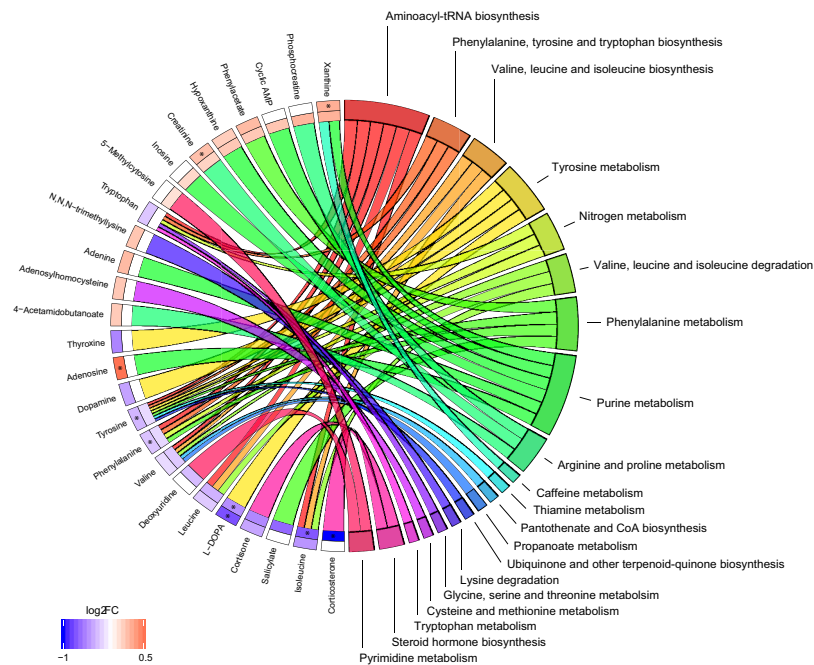


Figure 3. Linkages between altered metabolites and biochemical pathways. The altered metabolites have been linked with pathways as color-coded ribbons. Blue-to-red coding next to the metabolites depicts the magnitude of the \log_2 fold change (FC), where the inner layer represents FC in manifest compared with premanifest HD (HD-pHD) and the outer in comparison with controls (HD-C). Significant FC with an FDR < 0.1 have been marked with ‘*’.

p-value = 0.84), thus eliminating age as a potential confounder³¹. In serum, the results for thyroxine levels in mHTT carriers have been inconclusive, reporting no significant change between premanifest and manifest HD³² as well as increased levels in HD compared with controls³³. Reduced thyroxine levels in CSF have been reported in other neurodegenerative disorders such as Alzheimer’s disease (AD)³⁴ and tyrosine has been reported decreased in *post mortem* HD brain tissue (frontal lobe and striatum)¹². In addition, tyrosine metabolism has as a common denominator been found to be perturbed in both AD and Parkinson’s disease (PD)³⁵. Thyroid hormones, including thyroxine, are synthesized and released by the thyroid gland and then transported into the CNS³⁶. They are mainly involved in the regulation of energy homeostasis and are important for normal brain function. Previously it has been shown that the wild-type HTT protein binds to thyroid hormone receptor- α 1 (TR α 1)³⁷, a nuclear receptor for the deiodinated product of thyroxine (T₄), triiodothyronine (T₃), which may be a reason for the altered tyrosine metabolism.

Dopamine and L-DOPA are part of the dopaminergic pathway in tyrosine metabolism where dopamine is produced by its precursor L-DOPA, which further is synthesized from tyrosine and phenylalanine, all of which were shown to be reduced in manifest HD patients and were to various degrees associated with disease severity. Similarly to thyroxine, dopamine levels also correlated with the five year risk of onset. These observations suggest a derangement in tyrosine metabolism and more specifically in the dopaminergic pathway in HD, even though the role is still not well understood and previous findings have been inconclusive, as reviewed by Byrne and Wild³⁸.

The phenylalanine metabolism also achieved high impact and significance in the pathway analyses, where phenylalanine, salicylate and tyrosine were reduced and phenylacetate increased in manifest compared with premanifest HD and/or controls. The phenylalanine metabolism has a significant role in neurotransmitter regulation and has previously been reported to be affected in both AD and amyotrophic lateral sclerosis (ALS)³⁵.

The amino acids phenylalanine, valine, isoleucine and leucine are part of the aminoacyl-tRNA biosynthesis and were all reduced in manifest HD compared with premanifest and controls, but showed only moderate correlations to measures of disease severity. Decreased levels of these large neutral amino acids have previously been reported in *post mortem* brain tissue, CSF and/or serum/plasma in HD subjects on multiple occasions and suggested to be a result of a deficiency in the energy metabolism^{10,12,39–42}. Similarly, the valine, leucine and isoleucine degradation pathway, herein perturbed in manifest compared with premanifest HD and controls, has previously been found to be affected in PD³⁵.

Moreover, the purine metabolism was found to be significantly altered in manifest compared with premanifest HD and controls. Previously, this has also been found in AD, ALS and PD and therefore suggested to be an alternative pathway for neurodegenerative diseases to overcome the inadequate glucose supply and the deficient energy metabolism³⁵. Xanthine and its precursor hypoxanthine was found to be increased in manifest compared with premanifest HD and controls, adenosine and adenine was increased in manifest compared with premanifest HD, and cyclic AMP and inosine was raised in manifest HD compared with controls, which are all part of the

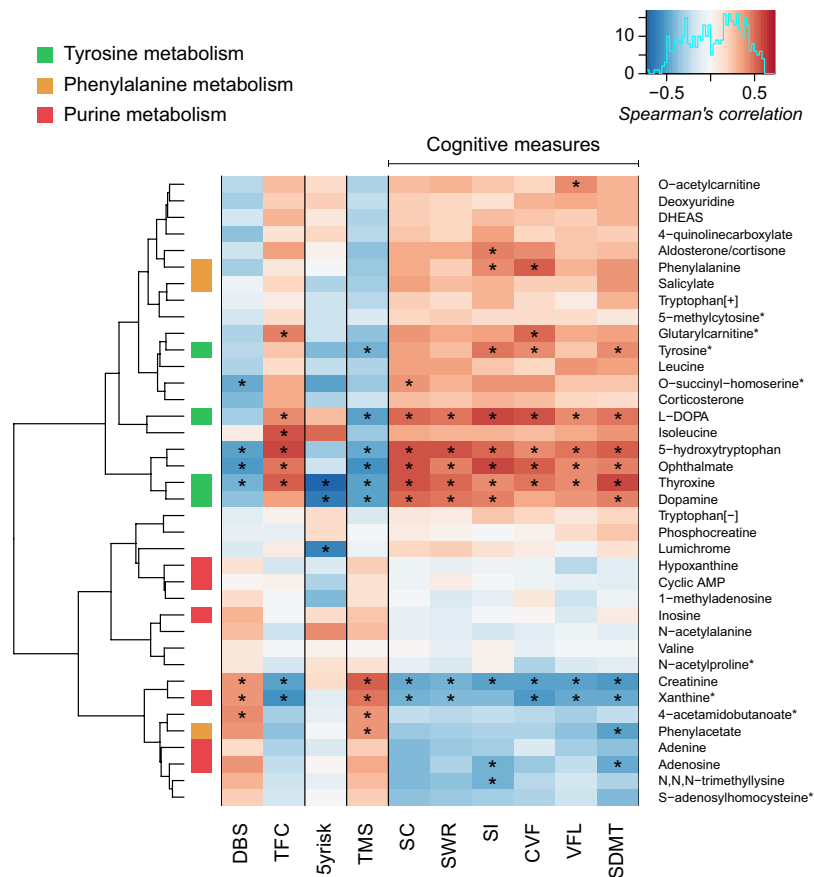


Figure 4. Associations between the biochemical changes and HD symptoms, where a “*” indicates a corresponding p -value < 0.05. Hierarchical clustering has been performed on the correlation values where metabolites from the tyrosine metabolism, phenylalanine metabolism as well as the purine metabolism have been marked. DBS: disease burden score; TFC: total functional capacity; 5yrisk: the five year risk of onset; TMS: total motor score; SC: Stroop color; SWR: Stroop word reading; SI: Stroop interference; CVF: category verbal fluency; VFL: verbal fluency letters; SDMT: symbol digit modalities test. *Metabolites with a demonstrated age dependence that have been corrected.

purine metabolism. Increased adenosine levels have previously been demonstrated in the striatum of transgenic HD mice and suggested to be a consequence of the mitochondrial dysfunction⁴³, whereas the remaining compounds are novel findings in HD.

Intriguingly, lumichrome displayed a twofold decrease in manifest HD patients compared with controls, but no difference between manifest and premanifest HD (see Supplementary Fig. S2). Comparing mHTT carriers (manifest and premanifest HD subjects) to controls supports the finding that decreased levels of lumichrome are typical to mHTT carriers (mHTT carriers- $C \log_2 FC = -0.87$, p -value < 0.0003, AUROC = 0.77). Interestingly, lumichrome was correlated with the five year risk of onset in premanifest HD subjects (no correlation to ageing was found, $r = 0.05$, p -value = 0.76), but did not correlate with any other measure of disease severity. Lumichrome is part of the riboflavin metabolism where it is in strict equilibrium with the amount of riboflavin (vitamin B₂), which is an essential compound in the brain that is transported from the blood through the CSF into the brain. Riboflavin has, as a majority of other findings herein, been connected to the energy metabolism where a riboflavin deficiency has resulted in a mitochondrial dysfunction⁴⁴. Riboflavin is also essential for monoamine oxidase (MAO) activity⁴⁵, which in its different isoforms is needed for metabolism of several important neurotransmitters such as norepinephrine, serotonin and dopamine. The changes in lumichrome levels might be directly connected to increased MAO-activity due to locally increased levels of dopamine⁴⁶, but this finding may also indicate a deficiency in the energy metabolism before disease onset.

Taken together, these observations support the hypotheses of a negative energy balance in HD and mitochondrial dysfunction, previously supported by multiple studies^{9,11,47–49}.

The most prominent limitation of the study was the limited number of mHTT carriers, which is a natural restraint of a single center HD study. This limitation suppresses the statistical power and confidence of our findings. Facing this limitation, the pathway analysis was advantageous as it couples statistical testing to molecular functioning, showing significance at a higher level⁵⁰. To strengthen our findings further the metabolites showing altered levels were also associated to measures of disease severity, enabled by the well-characterized demographic data of the mHTT carriers. Moreover, to ensure reproducibility and extraction of variation true to the global HD

population, profound cross-validation was performed to estimate more accurate AUROC values and VIP scores of the metabolites in the PLS-DA models. Finally, very few of the altered metabolites were found to change significantly over time. The limited number of mHTT carriers where a second sample was collected and the relative short time interval between samplings may be a reason to this.

Aside from the limited sample size, the study contained a number of other restraints. The natural age difference between premanifest and manifest HD subjects with a similar number of CAG repeats could be eliminated by age effect estimation and correction using the age matched controls. By such, we assumed that the mHTT carriers were depicting an accelerated effect caused by a combination of aging and neurodegeneration. To prevent removal of neurodegenerative alterations that are confounding with age and are of interest, we estimated the “normal” age effect in solely controls.

Moreover, the administration of drugs to some but not all subjects, could affect the results. Excluding patients on antidepressants and antipsychotics, univariate analyses of the altered metabolites between manifest and premanifest HD, as well as manifest HD compared with controls demonstrated similar changes but with less significant power, implying that the findings were not treatment specific (see Supplementary Table S5). However, as only the use of antidepressants and antipsychotics were considered, effects of other medications cannot be ruled out. Information regarding menopausal status of female patients as well as estrogen replacement therapy were neither collected. Trends of menopausal status in the age of onset have been indicated⁵¹. However, stratification on presumed menopausal status (females <50 versus ≥50 years of age), resulted in only minority groups of presumed post-menopausal women ($n_{HD} = 3$, $n_{preHD} = 1$, $n_{control} = 10$). As such, Chi-squared tests comparing the distributions of menopausal status between diagnostic groups resulted in no significant differences (p -value > 0.05). But it cannot be ruled out that there may be metabolite expressions that are affected by menopausal status. Moreover, the inter-group *post hoc* comparisons could rule out gender, number of CAG repeats and sampling intervals as confounding factors.

Finally, one of the major challenges in metabolomics studies is metabolite identification. Identification by an *in house* library has the advantage of high reliability but the restraint of being limited to the composed compounds. The PCA model generated on the full metabolome suggested global biochemical differences between the three groups. But when limiting the data to metabolites with an identity, the segregations were not as clear, indicating that there are other biochemical alterations of importance that have not been fully revealed herein. The PLS-DA models extracted 24 and 28 metabolites each, comparing manifest HD with controls and manifest with premanifest HD, where 14 metabolites were in common. A majority (15 and 18 respectively) of the altered metabolites were found to be decreased in manifest compared with premanifest HD and controls, in agreement to what has previously been found in the proteome of HD CSF¹³.

In conclusion, we found tyrosine metabolism, including the dopaminergic species, to be altered in manifest HD patients. All corresponding metabolites from the tyrosine metabolism demonstrated moderate to strong associations to measures of disease severity and symptoms. Thyroxine showed particular promise as it was associated to all measures including the five year risk of onset in premanifest HD subjects. Moreover, a large proportion of the biochemical changes supported previous findings of an impaired energy metabolism in HD, where decreased levels of lumichrome were typical for all mHTT carriers. These findings are important steps towards characterizing the molecular pathogenesis in HD and support the hypothesis that the CSF metabolome may be used to explore changes caused by the HTT mutation, which may be essential for future development of novel disease-modifying therapies for HD.

Data Availability

The mass spectrometry data generated/analyzed in current study is available on the MetaboLights database, study identifier: MTBLS749.

References

1. 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).
2. Li, L.-B., Yu, Z., Teng, X. & Bonini, N. M. RNA toxicity is a component of ataxin-3 degeneration in *Drosophila*. *Nature* **453**, 1107–1111 (2008).
3. Kovtun, I. V. *et al.* OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature* **447**, 447–452 (2007).
4. Pearson, C. E. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic entities! *PLoS Genet.* **7**, e1002018 (2011).
5. Ross, C. A. *et al.* Huntington disease: natural history, biomarkers and prospects for therapeutics. *Nat. Rev. Neurol.* **10**, 204–216 (2014).
6. Saudou, F. & Humbert, S. The Biology of Huntingtin. *Neuron* **89**, 910–926 (2016).
7. Rinaldi, C. *et al.* Predictors of survival in a Huntington's disease population from southern Italy. *Can. J. Neurol. Sci.* **39**, 48–51 (2012).
8. Herman, S. *et al.* Integration of magnetic resonance imaging and protein and metabolite CSF measurements to enable early diagnosis of secondary progressive multiple sclerosis. *Theranostics* **8**, 4477–4490 (2018).
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. *Biochim. Biophys. Acta* **1812**, 1371–1379 (2011).
10. Cheng, M.-L., Chang, K.-H., Wu, Y.-R. & Chen, C.-M. Metabolic disturbances in plasma as biomarkers for Huntington's disease. *J. Nutr. Biochem.* **31**, 38–44 (2016).
11. Graham, S. F. *et al.* Targeted biochemical profiling of brain from Huntington's disease patients reveals novel metabolic pathways of interest. *Biochim. Biophys. Acta* **1864**, 2430–2437 (2018).
12. Graham, S. F. *et al.* Metabolic signatures of Huntington's disease (HD): (1)H NMR analysis of the polar metabolome in post-mortem human brain. *Biochim. Biophys. Acta* **1862**, 1675–1684 (2016).
13. Fang, Q. *et al.* Brain-specific proteins decline in the cerebrospinal fluid of humans with Huntington disease. *Mol. Cell. Proteomics* **8**, 451–466 (2009).
14. Reilmann, R., Leavitt, B. R. & Ross, C. A. Diagnostic criteria for Huntington's disease based on natural history. *Mov. Disord.* **29**, 1335–1341 (2014).

15. Shoulson, I. & Fahn, S. Huntington disease: clinical care and evaluation. *Neurology* **29**, 1–3 (1979).
16. Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group. *Mov. Disord.* **11**, 136–142 (1996).
17. Langbehn, D. R. *et al.* A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length. *Clin. Genet.* **65**, 267–277 (2004).
18. Teunissen, C. E. *et al.* A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. *Neurology* **73**, 1914–1922 (2009).
19. Chambers, M. C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
20. Berthold, M. R. *et al.* KNIME: The Konstanz Information Miner. In *Studies in Classification, Data Analysis, and Knowledge Organization* 319–326 (2008).
21. Kenar, E. *et al.* Automated label-free quantification of metabolites from liquid chromatography-mass spectrometry data. *Mol. Cell. Proteomics* **13**, 348–359 (2014).
22. Weisser, H. *et al.* An automated pipeline for high-throughput label-free quantitative proteomics. *J. Proteome Res.* **12**, 1628–1644 (2013).
23. Website. Available at: R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>. (Accessed: 28th March 2018).
24. Herman, S. *et al.* Mass spectrometry based metabolomics for *in vitro* systems pharmacology: pitfalls, challenges, and computational solutions. *Metabolomics* **13**, 79 (2017).
25. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
26. Falahati, F. *et al.* The Effect of Age Correction on Multivariate Classification in Alzheimer's Disease, with a Focus on the Characteristics of Incorrectly and Correctly Classified Subjects. *Brain Topogr.* **29**, 296–307 (2016).
27. Rohart, F., Gautier, B., Singh, A. & Cao, K.-A. L. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **13**, e1005752 (2017).
28. Thévenot, E. A., Roux, A., Xu, Y., Ezan, E. & Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **14**, 3322–3335 (2015).
29. Robin, X. *et al.* pROC: an open-source package for R and S to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).
30. Xia, J., Sinelnikov, I. V., Han, B. & Wishart, D. S. MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Res.* **43**, W251–7 (2015).
31. Chaleckis, R., Murakami, I., Takada, J., Kondoh, H. & Yanagida, M. Individual variability in human blood metabolites identifies age-related differences. *Proc. Natl. Acad. Sci. USA* **113**, 4252–4259 (2016).
32. Kalliolia, E. *et al.* A 24-Hour Study of the Hypothalamo-Pituitary Axes in Huntington's Disease. *PLoS One* **10**, e0138848 (2015).
33. Aziz, N. A., Pijl, H., Frölich, M., Roelfsema, F. & Roos, R. A. C. Altered thyrotropic and lactotropic axes regulation in Huntington's disease. *Clin. Endocrinol.* **73**, 540–545 (2010).
34. Johansson, P. *et al.* Reduced cerebrospinal fluid level of thyroxine in patients with Alzheimer's disease. *Psychoneuroendocrinology* **38**, 1058–1066 (2013).
35. Kori, M., Aydın, B., Unal, S., Arga, K. Y. & Kazan, D. Metabolic Biomarkers and Neurodegeneration: A Pathway Enrichment Analysis of Alzheimer's Disease, Parkinson's Disease, and Amyotrophic Lateral Sclerosis. *OMICS* **20**, 645–661 (2016).
36. Dratman, M. B., Crutchfield, F. L., Gordon, J. T. & Jennings, A. S. Iodothyronine homeostasis in rat brain during hypo- and hyperthyroidism. *Am. J. Physiol.* **245**, E185–93 (1983).
37. Futter, M. *et al.* Wild-type but not mutant huntingtin modulates the transcriptional activity of liver X receptors. *J. Med. Genet.* **46**, 438–446 (2009).
38. Byrne, L. M. & Wild, E. J. Cerebrospinal Fluid Biomarkers for Huntington's Disease. *J. Huntingtons Dis.* **5**, 1–13 (2016).
39. Oepen, G., Cramer, H., Bernasconi, R. & Martin, P. Huntington's disease - imbalance of free amino acids in the cerebrospinal fluid of patients and offspring at-risk. *Arch. Psychiatr. Nervenkr.* **231**, 131–140 (1982).
40. Mochel, F. *et al.* Early energy deficit in Huntington disease: identification of a plasma biomarker traceable during disease progression. *PLoS One* **2**, e647 (2007).
41. Underwood, B. R. *et al.* Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. *Brain* **129**, 877–886 (2006).
42. Mochel, F., Benaich, S., Rabier, D. & Durr, A. Validation of plasma branched chain amino acids as biomarkers in Huntington disease. *Arch. Neurol.* **68**, 265–267 (2011).
43. Gianfriddo, M., Melani, A., Turchi, D., Giovannini, M. G. & Pedata, F. Adenosine and glutamate extracellular concentrations and mitogen-activated protein kinases in the striatum of Huntington transgenic mice. Selective antagonism of adenosine A2A receptors reduces transmitter outflow. *Neurobiol. Dis.* **17**, 77–88 (2004).
44. Depeint, F., Bruce, W. R., Shangari, N., Mehta, R. & O'Brien, P. J. Mitochondrial function and toxicity: role of the B vitamin family on mitochondrial energy metabolism. *Chem. Biol. Interact.* **163**, 94–112 (2006).
45. Bonnefil, V., Castiglione, C. M., Cawthon, R. M. & Breakefield, X. O. Effect of riboflavin on monoamine oxidase activity in cultured neuroblastoma cells. *Cell. Mol. Neurobiol.* **1**, 351–359 (1981).
46. Chen, J. Y., Wang, E. A., Cepeda, C. & Levine, M. S. Dopamine imbalance in Huntington's disease: a mechanism for the lack of behavioral flexibility. *Front. Neurosci.* **7**, 114 (2013).
47. Gu, M. *et al.* Mitochondrial defect in Huntington's disease caudate nucleus. *Ann. Neurol.* **39**, 385–389 (1996).
48. Brennan, W. A. Jr., Bird, E. D. & Aprille, J. R. Regional mitochondrial respiratory activity in Huntington's disease brain. *J. Neurochem.* **44**, 1948–1950 (1985).
49. Goodman, A. O. G. *et al.* The metabolic profile of early Huntington's disease—a combined human and transgenic mouse study. *Exp. Neurol.* **210**, 691–698 (2008).
50. García-Campos, M. A., Espinal-Enríquez, J. & Hernández-Lemus, E. Pathway Analysis: State of the Art. *Front. Physiol.* **6**, 383 (2015).
51. Arning, L. *et al.* NR2A and NR2B receptor gene variations modify age at onset in Huntington disease in a sex-specific manner. *Hum. Genet.* **122**, 175–182 (2007).

Acknowledgements

We wish to thank all participants who donated their samples. Per-Johan Jakobsson is acknowledged for contributing with metabolite standards. We greatly acknowledge the financial support by the Åke Wiberg Foundation and The European Commission's Horizon 2020 Programme funded under grant agreement number 654241 (PhenoMeNal). The Swedish state support for clinical research (ALF) also funded the research, including research time for V.N. J.S. received grants from the Åland cultural foundation for HD research.

Author Contributions

S.H., V.N., J.S., A.M.L., D.N. and K.K. conceptualized and designed the study. S.H. performed the metabolomics analysis. V.N. and J.S. collected and curated the clinical data. S.H. and P.E.K. performed the data analysis. J.S., J.B., A.M.L., O.S., D.N. and K.K. supervised the work and provided resources to the study. S.H., V.N. and K.K. wrote the manuscript. All authors contributed to the data interpretation and reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-40186-5>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019