Mapping the Huntington's disease process using cerebrospinal fluid analysis

VALTER NIEMELÄ
Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder, caused by a CAG-repeat expansion in the HTT-gene. Today there are no disease-modifying therapies (DMTs), but several promising clinical trials are underway, including therapies that reduce mutant huntingtin expression.

Reliable biomarkers could empower such trials and guide the timing for initiation of future DMTs.

Neurofilament light (NFL) and tau, which are cerebrospinal fluid markers of neuronal death, have been implicated as markers of disease progression. Increased levels of the inflammatory marker YKL-40 have also been reported in HD.

The aim was to validate and compare the above biomarker candidates by targeted analyses, while explorative liquid chromatography-mass spectrometry (LC-MS) was used to identify new candidates. Clinically well-characterized HD patients, premanifest gene expansion carriers (pGECs), and controls were enrolled from Uppsala University Hospital in Sweden.

In contrast to tau, NFL levels differed between all three groups and NFL had stronger correlations with symptom severity compared with total-tau and phospho-tau. Longitudinally, only NFL maintained intergroup differences and rose with disease progression.

Soluble CD27, a marker of T cell-mediated inflammation, differed between all three groups, with the highest levels in manifest HD, and mostly undetectable levels in controls. YKL-40 showed a non-significant trend toward increase in manifest HD.

We applied LC-MS metabolomics and discovered a metabolite signature unique to manifest HD, with deranged tyrosine metabolism including L-DOPA, dopamine, and thyroxine. Utilizing LC-MS we also identified altered proteins in manifest HD, including proenkephalin that was decreased and associated with symptom severity.

In conclusion, NFL may be used as a pharmacodynamic marker in intervention trials. Interestingly, elevated sCD27 implies a role of adaptive immunity before disease onset, but validation is needed. YKL-40 is not suitable as an early marker in HD. The CSF metabolome constitutes a new compartment of potential biomarkers but challenges in metabolite identification should be addressed in future studies. Proenkephalin levels potentially reflect the remaining number of striatal medium spiny neurons and hold promise as a marker of disease progression.

Keywords: Neurodegeneration, Huntington's disease, Neurogenetics, Neuroinflammation, Neurofilament light, tau, Proenkephalin

Valter Niemelä, Department of Neuroscience, Box 593, Uppsala University, SE-75124 Uppsala, Sweden.

© Valter Niemelä 2019

ISSN 1651-6206
ISBN 978-91-513-0795-4
urn:nbn:se:uu:diva-393256 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-393256)
List of papers


# Table of contents

Introduction ................................................................................................ 9  
Huntington's disease – a single-gene neurodegenerative disorder ........... 9  
Disease progression, its determinants and rating scales ......................10  
Molecular pathogenesis ...........................................................................16  
Outlook for future disease-modifying therapies ..............................16  
Biomarkers for HD – why are they needed? .......................................17  
Biomarkers for HD – the current state of research .................................18

Aims .........................................................................................................19  
What is a suitable biomarker for Huntington's disease? ......................19  
Aim of individual papers .......................................................................20

Study participants and methods .............................................................21  
Study participants – the Uppsala HDCSF cohort .................................21  
Clinical assessment ..............................................................................23  
CSF sample collection ..........................................................................23  
Biochemical analyses ............................................................................23  
Enzyme-linked immunosorbent assay (ELISA) ..................................23  
Untargeted liquid chromatography-mass spectrometry (LC-MS) .......25  
Statistical analyses ................................................................................26

Results including key illustrations .........................................................27  
Paper I ..................................................................................................27  
Paper II ................................................................................................29  
Paper III................................................................................................31  
Paper IV................................................................................................35

Discussion ...............................................................................................39  
Neuronal damage markers ..................................................................39  
Neuroinflammatory markers .................................................................40  
Metabolomic characteristics ..................................................................40  
Proteomic characteristics .......................................................................42

Limitations ..............................................................................................44

Conclusions ............................................................................................45

Summary in Swedish/ Sammanfattning på svenska ..........................46
Abbreviations

Antisense oligonucleotide (ASO)
Central nervous system (CNS)
Cerebrospinal fluid (CSF)
Chitinase-3 like protein 1 (YKL-40)
False discovery rate (FDR)
Gene negative controls (controls)
Huntington's disease (HD)
Liquid-Chromatography Mass-Spectrometry (LC-MS)
Mutant huntingtin (mHTT)
Neurofilament light (NFL)
Phospho-tau (P-tau)
Presymptomatic gene expansion carriers (pGEC)
Proenkephalin (PENK)
Soluble CD27 receptor (sCD27)
Stroop color (SC)
Stroop interference (SI)
Stroop word (SWR)
Symbol digit modality test (SDMT)
Total-tau (T-tau)
Total Functional Capacity (TFC)

Special terms
Disease burden score (DBS) (see introduction for a definition)
5-year risk of onset (see introduction for a definition)
Introduction

Huntington's disease – a single-gene neurodegenerative disorder

Huntington's disease (HD) is an inheritable and severe neurodegenerative disorder. Symptoms are progressive and lead to death after about 20 years of disease. HD is an autosomal dominant disease which means the risk of passing the gene mutation down to offspring is 50%. Because no disease-modifying treatments are proven so far, the mainstay of treatment is to manage symptoms and provide support to patients, family and caretakers. This is best achieved through specialized multidisciplinary teams (1). Predictive gene testing can be provided to those above 18 years of age who wish to know their status, in accordance with established guidelines (2). However only about 1/6 choose to have the test among those with 50% risk of HD (3-5).

Although sporadic reports of the disease occurred before George Huntington's, his essay on the disease (“On Chorea” in 1872) is acknowledged for its accurate description of both clinical characteristics and inheritance pattern that remains true today (6).

The cause for HD is an expansion of the CAG-repeat segment (“polyglutamine”) expansion on exon 1 of the huntingtin (HTT) gene, which was identified in 1993 resulting from a longstanding and collaborative effort (7). The penetrance is complete with 40 or more repeats on one of the HTT alleles. Repeat lengths between 36 and 39 can cause HD but with incomplete penetrance and typically onset later in life. While the majority inherit HD from affected parents, about 6-8% of cases have no family history, resulting from meiotic expansion into the disease-causing CAG-range (new mutations) (8). Increased repeat length causes instability and higher risk of further expansion in the offspring, leading to earlier onset (anticipation).

Intermediate alleles (IA) are in the range of 27-35 repeats and are common in the general population (9). Their potential pathological significance (a suggested behavioral phenotype (10) and rare reports of HD like symptoms (11) is controversial (12) and subject to continued research. What is known is that IAs are a source for meiotic expansion and therefore a source for new HD mutations.

HD is the most common cause of inherited dementia in the developed world. The highest prevalence of HD is found in populations with predominantly European ancestry estimated at 5.7/100 000 (13), but the more recent
reports using multisource ascertainment provide substantially higher estimates as for instance 12.3 - in the UK (14) and 13.7 per 100 000 in Canada (15). Importantly, HD also affects people in non-Caucasian regions although with much lower prevalence, as demonstrated in studies in Asian countries with a pooled prevalence of 0.4 per 100 000 (13).

The national prevalence of HD in Sweden has not been calculated after the genetic test was made available (16).

Disease progression, its determinants and rating scales

The introduction of the genetic test for HD has made it possible to study pre-manifest individuals many years before predicted disease onset throughout the disease course. Larger observational studies (such as REGISTRY (17), Enroll-HD (18) and COHORT (19)) include thousands of premanifest and manifest HD patients, whereas TRACK-HD and PREDICT-HD with smaller numbers have dense imaging data aimed to identify biomarkers that are linked to disease onset. Thanks to such international efforts our understanding of the natural history of HD has been strengthened.

HD typically manifests in midlife between 30 to 50 years of age, but the onset is subject to large variation and may occur at any time. Age of onset is inversely correlated with the number of CAG-repeats which explains about 60% of the variation (20), (Fig 1). Genome wide association studies revealed that the residual variation in age of onset and rate of disease progression is largely due to genetic variation in regions other than the CAG-repeat tract. Important genetic modifiers of HD have been found in the DNA damage repair system, which is associated with modulating the expansion of the CAG repeat tract in germline and somatic cells (21).

There is a long-standing debate about how the onset should be clinically defined, but a prevailing consensus is still to judge by the presence of unequivocal extrapyramidal motor signs (22, 23). Symptoms are progressive and can be divided into three domains; movement disorder, psychiatric disorder, and cognitive impairment. Subtle symptoms emerge 5-10 years before a clinical diagnosis can be established and this period is referred to as the prodromal stage.

Which symptom emerges first, or becomes most disabling for a given patient cannot be predicted at this time and the reason for this variation is poorly understood.

The movement disorder in HD is of a complex nature featuring chorea, dystonia, bradykinesia and rigidity, impaired saccadic function and eye movements, dysphagia and motor apraxia. Chorea is a typical early feature while most patients progress into a more rigid stage later during the clinical course (19).
Cognitive decline begins in prodromal HD, and is characterized by mental slowing, reduced attention and flexibility, impaired visuospatial functions and emotional recognition (24-26).

The psychiatric symptoms associated with HD include a wide span from the commonly occurring apathy, depression, anxiety and sleep disturbances, obsessive compulsive disorders, to the less frequently observed psychosis (27). The psychiatric symptoms, except apathy, do not progress as the motor- and cognitive symptoms do.

Disease progression is relentless and in the latest stage patients often lose their ability to communicate and need assistance with all activities of daily living.

The currently most used method of rating HD symptoms is the Unified HD rating scale (UHDRS), which includes scales for motor, emotional, cognitive, behavioral and functional components (22). Onset is judged according to a set of criteria, defined as diagnostic level of confidence (DCL), ranging from 0 (no motor abnormalities suggestive of HD) to 4 (motor abnormalities ≥99% likely to be due to HD) (28).

Total motor score (UHDRS-TMS) includes scoring of eye movements, speech, chorea, and dystonia, rapid alternating movements, bradykinesia and gait (range 0-123 points). UHDRS is able to track disease progression but lacks sensitivity in the premanifest period.
Disease stages
Manifest disease can be divided into five stages (Table 1 below). These stages are based on the level of functional impairment, which is assessed with the instrument Total Functional Capacity (TFC), included in the UHDRS (29).

Table 1. Disease stages

<table>
<thead>
<tr>
<th>Stage</th>
<th>Total Functional Capacity (TFC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>11-13</td>
</tr>
<tr>
<td></td>
<td>7-10</td>
</tr>
<tr>
<td>Moderate</td>
<td>4-6</td>
</tr>
<tr>
<td>Advanced</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Disease burden score
The term disease burden score (DBS) or CAG-age-product (CAP) score, is frequently used in HD research. It is simply a function of age and number of CAG repeats on the expanded HTT-allele, calculated by the formula, \((CAG_n - 35.5) \times \text{age}\) (30). Disease burden is considered to be a measure of accumulated genetic toxicity, and can predict the HD phenotype to a large extent.

5-year risk of onset
Langbehn et al. (31) used epidemiological data to create a model for predicting the risk of disease onset in a certain window of time, based on CAG repeat number and current age alone. Previous attempts were based solely on age of onset of HD patients recorded at care centers, whereas this new strategy also included data from pre-symptomatic individuals. In this way, the model utilizes all gene expansion carriers for a more reliable prediction.
Fig. 1. CAG repeat number and age of motor onset in HD. Reprinted with permission from Holmans et al. Genetic modifiers of Mendelian disease: Huntington's disease and the trinucleotide repeat disorders. Human Molecular Genetics, 2017, Volume 26, Issue R2, Pages R83–R90.
Molecular pathogenesis

Huntingtin is expressed ubiquitously throughout the human tissues. Although some peripheral pathology can be detected (32), the most deleterious effects occur in the CNS. The normal function of huntingtin has been subject to intense research but is still elusive. HTT has many sites for interactions with other proteins. It is essential for the development of the nervous system. Homozygous knockout of \textit{HTT} in mice results in lethality before development of the nervous system (33, 34), and heterozygotes display abnormalities of the basal ganglia (34). The medium spiny neurons of the striatum are affected first and neurodegeneration spreads to the frontal lobes and later on to cortical grey matter (35, 36). There seems to be no single causative mechanism for neuronal death, but instead a pathological cascade of events mediated by the diverse effects of m\textit{HTT}. Mechanisms involve both gain of function, such as from the cleaved N-terminal fragments (containing the expanded polyQ-stretch) that interfere with transcription in the nucleus, as well as loss of function exemplified by reduced neurotrophic support (37, 38). The disruption of other systems has been reported including, mitochondrial function (39-41), axonal transport (42), and more recently the hormonal axis (43) and neuroinflammation (44, 45).

Outlook for future disease-modifying therapies

More than 25 years after the discovery of the gene causing HD, despite about 100 clinical trials of drugs and devices, we still have no disease-modifying therapies (46). The HD community has waited a long time. Nonetheless, there is now reason to be optimistic. In contrast to previous trials, the ongoing and planned trials are increasingly targeting HD specific pathological mechanisms. As HD is an autosomal dominant disorder, it is tempting to go straight to the source of the disease and reduce the production of m\textit{HTT}. The evolving techniques with the common goal to lower huntingtin (“huntingtin lowering strategies”) turn out to be quite diverse tools with different qualities and unique limitations.

Starting upstream with the case of actual gene editing for intervention at the DNA level techniques such as CRISPR/Cas9 and Zinc-finger are being developed for the purpose. Judging from cell studies and murine models it is possible, yet this field is still in its early research phase. The obvious risk with gene editing is that potential long-term effects of the intervention (a modified genotype) would lead to adverse and irreversible effects. For a review of the above therapeutic strategies see Wild et al. (47).

Another strategy for huntingtin lowering is targeting the messenger RNA of \textit{HTT}. A recent phase 1b/2a trial with HD patients investigated the safety and tolerability of an antisense oligonucleotide (ASO, IONIS HTT-Rx) that
binds and induces degradation of huntingtin mRNA transcripts. In December 2017, IONIS confirmed that their ASO was able to reduce CSF mHTT in a dose-dependent manner. This unique breakthrough caught the attention of the HD community world-wide as a sign of hope for the future. Notably, this ASO unselectively lowers both wild type- and mHTT. This type of drug acts in a way that any given HD patient could potentially benefit from treatment, but also infers a risk that adverse effects may arise from loss of function of wild type huntingtin. Therefore, the goal was limited to a 50% reduction of HTT and the drug has now entered a phase 3 trial. Ongoing are also safety trials with ASOs that act through an allele-specific manner, thus allowing greater mHTT lowering, while leaving the wild type unaffected. These ASOs however require genotyping ahead of study entry to find out if patients have the certain genetic traits (SNPs) that the ASO targets or not.

Biomarkers for HD – why are they needed?

Accurate quantitation of disease activity in any disease is crucial when assessing the therapeutic efficacy of interventions. Monitoring HD progression with clinical instruments such as the UHDRS has been the standard in trials so far. However clinical assessment is challenging due to slow progression of diverse symptoms, and also due to inter- and intra-rater variability. Trials involving presymptomatic gene expansion carriers (pGECs) would benefit from tools able to monitor “subclinical” disease activity and estimate the time of clinical onset. While biomarkers are not generally used in the clinical setting for HD today, reliable biomarkers could empower clinical trials by facilitating the assessment of therapeutic efficacy. When disease-modifying treatments arrive, important issues arise concerning the optimal timing of treatment-initiation (perhaps in the pre-symptomatic stage or prodromal stage?), and assessment of response. For pGECs it is especially important that we have markers that can objectively quantitate their subclinical neuropathological disease progression, to prolong the pre-symptomatic stage or even prevent disease-onset.

When disease-modifying treatments arrive, important issues arise concerning the timing of treatment-initiation (i.e. in the pre-symptomatic stage or prodromal stage?).
Biomarkers for HD – the current state of research

*Imaging markers*

Structural radiological correlates of disease-stage have been identified, such as rate of striatal atrophy on Magnetic Resonance Imaging (MRI) which can be detected from ten to fifteen years before disease onset (48, 49). White matter volume is also reduced long before motor onset (48, 50-54). However, such markers do not provide information on molecular pathological processes and imaging needs to be repeated. Another concern is that volumetrics may reflect factors other than disease progression, such as increase of striatal volume due to inflammation or edema. Potential functional imaging markers include fluorodeoxyglucose (FDG)-PET (55), and more recently PDE10A imaging which depicts progressive striatal changes in the pre-symptomatic stage (56, 57). The disadvantage is that nuclear imaging is complex and expensive.

*Biochemical markers*

Blood is of course an accessible source, but CSF has the advantage of proximity to the neurodegenerative process, and is still relatively accessible. CSF markers of neuronal death, neurofilament light (NFL) and total tau (t-tau), although not specific for HD, have been implicated as potential biomarkers for disease progression in HD (58-62). A number of biomarkers reflecting different aspects of immune system activation have also been reported to be abnormal in HD. Among these are IL-1-beta, IL-6, IL-8 (45), TNF-alpha, and YKL-40 (62-64). Although involved in different mechanisms, most of these markers conform to the innate immune system.

Interestingly, a single molecule assay has been developed that can measure mutant huntingtin mHTT in CSF. mHTT was detected in pGECs and manifest HD patients, in contrast to gene negative controls (60).

Despite the large number of individual biomarker candidates that have been reported, most of them have only been studied in single centers with smaller cohorts and there is a general need for validation (65, 66), especially in pre-symptomatic gene expansion carriers. In addition, a number of studies were carried out before the genetic test for HD was developed and may have included phenocopies.
Aims

What is a suitable biomarker for Huntington's disease?
We aimed to discover or validate cerebrospinal fluid markers that predict the HD phenotype with emphasis on the following aspects:

1) markers that reflect disease progression across all disease stages, from premanifest to late stage disease.

2) markers that correlate highly with HD symptom severity.

3) markers that reflect disease-specific mechanisms and potential therapeutic targets.

4) markers for use as potential surrogate end-points in clinical trials.
Aim of individual papers

Paper I To validate and compare NFL and tau as biomarkers for Huntington's disease

Paper II To investigate which immune mechanisms are involved early in HD, and explore the longitudinal profile of YKL-40, NFL, T-tau and P-tau.

Paper III To assess the metabolic profile of Huntington's disease gene expansion carriers.

Paper IV To assess the proteomic profile of Huntington's disease gene expansion carriers.
Study participants and methods

Study participants – the Uppsala HDCSF cohort

The clinical data and bio-samples that provide the fundament for this thesis were donated by the participants and stored in the Uppsala Huntington's disease CSF collection. This project enrolls manifest HD patients, premanifest gene expansion carriers (pGEC) and gene negative controls. Controls consist of spouses or other family members to HD patients including persons at risk for HD who had a negative predictive gene test result. All participants except the controls underwent clinical assessment at the visit for CSF-sample collection. This sample collection also stores matched blood samples of most participants and belongs under Uppsala biobank (registration number 827). Repeated sampling and clinical assessment were undertaken where possible at intervals of at least one year. Patients were initially included between 2012 and 2013 by JS and thereafter by VN. The number of participants varied between Papers depending on when samples were sent for analysis and how many samples had been collected at that time (Table 1).

As recruitment of controls was slow, additional sex- and age-matched neurological controls were enrolled. Their respective CSF-samples were obtained from a sample collection at the same clinic including patients who had undergone investigation for suspected neurological disease including lumbar puncture. Only samples belonging to patients where a neurological diagnosis had been ruled out were included.

*HD gene expansion carriers*

pGEC were defined as *HTT* gene expansion (>35 CAG repeats) and a diagnostic confidence level (DCL) below 4. Manifest HD was defined by *HTT* gene expansion (>35 CAG repeats) and DCL of 4. (28)
Table 2. Number of participants per study

<table>
<thead>
<tr>
<th>Category</th>
<th>Paper</th>
<th></th>
<th></th>
<th>Paper IV</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
<td></td>
<td>Uppsala</td>
<td>Validation</td>
</tr>
<tr>
<td>Manifest HD</td>
<td>12</td>
<td>14</td>
<td>13</td>
<td></td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>pGEC</td>
<td>11</td>
<td>13</td>
<td>13</td>
<td></td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Controls</td>
<td>0</td>
<td>25</td>
<td>43</td>
<td></td>
<td>38</td>
<td>23</td>
</tr>
</tbody>
</table>

HD, Huntington's disease; pGEC, premanifest gene expansion carriers; Controls, gene negative controls. Paper I included no controls for comparison. Paper II: repeated samples from 11 manifest HD patients, 9 premanifest gene expansion carriers and 3 controls were analyzed as well. In Paper III, repeated samples from 7 premanifest and 8 manifest patients were analyzed.

Controls
Paper I: no controls were included.

Paper II: 20 healthy young CSF donors without family history of HD, and 5 subjects who had undergone predictive gene testing with negative results. Three of the 5 mentioned controls donated a second CSF sample.

Paper III & IV: Paper III included 5 subjects with a family history of HD that had undergone predictive gene testing with negative results. Thirty-eight additional controls (where neurological diagnoses were ruled out) were included in each paper. Paper IV included a validation cohort of 23 additional controls and 23 manifest HD patients that were recruited from the Sahlgrenska University Hospital in Gothenburg and the Karolinska University Hospital in Stockholm.

Ethics
The studies were approved by the Regional Ethical Review Board in Uppsala (DNR 2012/274 and 2013/278). All participants provided written consent. The HD clinic at Sahlgrenska University Hospital in Gothenburg enrolled patients as part of the ethics approval (DNR 2012/274) whereas Karolinska University Hospital in Stockholm included patients after approval from the ethics review board in Stockholm (DNR 2012/2224-32/4).
Clinical assessment
For all included HD gene expansion carriers, a predefined battery of tests was used, including total motor score (TMS), symbol digit modality test (SDMT), letter verbal fluency, category verbal fluency (animals), Stroop color, Stroop word, Stroop interference and total functional capacity (TFC).

CSF sample collection
CSF collection was standardized to procedure and handling but the timing of meals and time of day varied. Atraumatic needles were preferentially used to minimize the risk of post-puncture headache. Polypropylene tubes and collection vessels were used to avoid protein adsorption. After centrifugation with 1300 G at 4 degrees Celsius for 10 minutes, the acellular proportion was pipetted off for storage at minus 70 degrees C.

Biochemical analyses
Enzyme-linked immunosorbent assay (ELISA)
Enzyme-Linked Immunosorbent Assay (ELISA) is a method that allows detection and quantification of substances (proteins, peptides, metabolites, hormones) using antibody binding, an immunoassay. ELISA was invented by Peter Perlmann and Eva Engvall in the 1970s (67) and has gained massive use ever since (68). Previous to development of immunoassays using enzyme detection the prevailing technique was the radioimmunoassay (RIA) which uses an antibody labeled with a radioactive substance. RIAs have continuously been replaced by ELISA due to both health risks and practical reasons. The principle for ELISA is simple. An antibody has specificity to bind a certain substance, its antigen. The samples are pipetted into wells of an ELISA plate. The plate wells are coated with a surface that binds proteins strongly. When samples are injected the proteins are bound in the wells. Often a primary antibody, called a capture antibody binds the protein of interest. A secondary antibody is used for detection. The detection antibody binds to the tail (other end) of the capture antibody. In a washing procedure all detection antibodies that are not bound to the complex of antigen and capture antibody are removed. Detection occurs when a substrate is added that is converted by an enzyme attached to the tail of the detection antibody.

The term sandwich ELISA is a variation of this method in which the capture antibodies are added to the surface of the wells first, and the rest of the wells’ sticky surface is blocked by a blocking buffer. When samples are loaded, only the protein of interest is bound to antibodies in the wells. After
washing, detection antibodies are added – creating a “sandwich” with the antigen in the middle. ELISA offers very high sensitivity, but its power lies mostly in the specificity of the antibodies (69). Incidents have been reported where commercial ELISA kits recognize a different antigen than the one it is marketed for (70, 71). Use of different antibodies across studies can influence the results, thus reducing the reproducibility of findings.

Five different commercial ELISA kits were used to quantify the levels of targeted proteins in Paper I and II (table 3 below). Analyses were performed at the Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal campus, except sCD27 which was analyzed at the Department of Clinical Chemistry, Uppsala University Hospital. Samples were run in a single batch, using the same set of reagents, to reduce inter-assay variation.

Table 3. ELISA kits applied:

<table>
<thead>
<tr>
<th>Protein</th>
<th>ELISA type</th>
<th>ELISA kit, manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-tau</td>
<td>Solid phase, direct</td>
<td>INNOTEST® hTAU Ag, Fujirebio Europe, Gent, Belgium.</td>
</tr>
<tr>
<td>P-tau</td>
<td>Solid phase, direct</td>
<td>INNOTEST® phosphotau[181P], Fujirebio Europe, Gent, Belgium.</td>
</tr>
<tr>
<td>NFL</td>
<td>Sandwich</td>
<td>NF-light ELISA kit, UmanDiagnostics, Umeå, Sweden.</td>
</tr>
<tr>
<td>sCD27</td>
<td>Sandwich</td>
<td>Human CD27/TNFRSF7 DY382, R&amp;D Systems, Minneapolis, MN, USA.</td>
</tr>
<tr>
<td>YKL-40</td>
<td>Sandwich</td>
<td>YKL-40, R&amp;D Systems, Minneapolis, MN, USA.</td>
</tr>
</tbody>
</table>
Untargeted liquid chromatography-mass spectrometry (LC-MS)

Both studies (Paper III and IV) utilized LC-MS to examine which molecules were present and their relative quantities between the samples. The main difference was that in paper III we studied the CSF metabolome (small molecules such as fatty acids, sugars, amino acids and lipids etc.) and in paper IV the proteome. Therefore, sample preparation differed as metabolites were extracted for paper III, and proteins (removal of the most abundant proteins, and digestion into peptides) for paper IV.

Briefly, LC-MS offers high resolution – the ability to analyze and separate a large number of compounds. The technique involves three main steps: 1) After preparation the samples are injected in Liquid-chromatography columns and travel through at different speed (retention time). Retention time depends on the affinity to the porous content in the column. 2) The next step is ionization of molecules into gas-phase. Ionization can be achieved by a process called electrospray ionization (ESI), where the molecules in solvent are sprayed through a needle tip at high voltage and temperature. 3) Entering the mass spectrometer, a magnetic field deflects the ionized molecules. The degree of deflection depends on the mass to charge (m/z) ratio and this property is finally recorded by a detector. Tandem MS was performed applying higher-energy collisional dissociation (HCD). The general mass-spectrometry method was the same in both experiments, and both used similar instruments (Orbitrap systems). Protein identities were confirmed by comparing the generated MS spectra with databases for human proteins. The identification of metabolites is less straightforward largely because hundreds or thousands of metabolites may have the same mass. To avoid this issue the MS metabolomics results were run against an in-house library of known and previously characterized metabolites. This allows reliable metabolite identification, however limited to the scope of the library.

As opposed to the relative intensity values derived from the experiments above, absolute quantitation of proenkephalin in paper IV was achieved by spiking an internal standard in each sample (AQUA peptide ELLETGDNRELC13N15, Sigma-Aldrich International GmbH).
Statistical analyses

Paper I and II: Differences in protein concentrations between groups were tested using non-parametric tests (Mann-Whitney). Associations to potential confounders like sex and age were assessed, and if there was an interaction, this variable was included as a covariant in a regression model. Correlations to clinical scores, disease burden, and 5-year risk of onset were expressed as Spearman correlations, based on the pGECs and manifest HD subjects pooled into one group. In paper I, an external statistician, was consulted, whereas in paper II all statistics were conducted by VN.

Paper III: Statistics were conducted by SH using R software. Principal component analysis (PCA) was used to identify which components and thereby which combination of metabolomic features that contributed most to separation of the groups. Partial least squares-discriminant analysis (PLS-DA) was used for the repeated samples analysis. ROC curves were plotted to visualize the yield of metabolomic features in discriminating between groups, and finally pathway analyses to identify the most impacted pathways.

Paper IV: A three-step predefined data analysis strategy was used to optimize the statistical power and validity of findings. 1) Differences in protein intensity scores between the pGEC and manifest HD group were assessed with t-tests and 2) The influence from age and sex was assessed in the controls and removed from the first analysis using a linear regression model 3) All proteins still showing differences were tested for correlations with clinical test scores using Spearman rank correlations. Analyses were carried out in SPSS and GraphPad Prism software by VN.
Results including key illustrations

Paper I
When comparing NFL and tau, we found that NFL was a superior HD biomarker in the following aspects, 1) Manifest patients had higher age-adjusted levels than the premanifest group (p = 0.003), but this was not the case for tau (p = 0.416). 2) NFL had stronger correlations with HD symptom scores (Fig. 4).
Fig. 4. Comparing correlations of NFL and tau with cognitive tests

Fig. 4. Neurofilament light (NFL) has stronger correlations with clinical scores like Stroop interference (A), and category fluency (B) compared with tau (C and D)
sCD27, a highly specific marker of intrathecal T cell-mediated inflammation, was elevated in pGECs, and even higher levels were observed in the manifest HD subjects. Controls, as expected, mostly had undetectable levels. YKL-40, reflecting astrocytosis, appears to be a feature of later disease stages, showing a non-significant trend toward elevated levels in the manifest group compared to the pGECs (Fig. 5). YKL-40, in contrast to sCD27, was closely correlated to HD symptoms, and to concentrations of the markers of neurodegeneration.

By visual interpretation of the longitudinal profiles of YKL-40, NFL, T-tau and P-tau (Fig. 6), we concluded that NFL has the most predictable dynamics, with generally and slowly increasing levels, and preserved intergroup differences over time. One subject who converted from pGEC to manifest HD between the two samples, had the highest NFL value among the pGECs at baseline, and continued rising slowly, whereas both forms of tau and YKL-40 had a clearly rising trend between samples. This is in line with NFL being a marker suitable for the early prodromal disease stage as well.
Fig. 5. Protein concentrations plotted un-adjusted by group. p-values are adjusted for age or sex, as described below. Boxes show first and third quartiles, the central bands show the median, and the whiskers show data within 1.5 IQR of the median. Red point ID indicates a subject who converted from premanifest to manifest disease during the study. (A) YKL-40 levels differed between all groups before, but not after adjustment for age. (B) Sex-adjusted sCD27 concentrations were higher in manifest HD than premanifest gene expansion carriers, and close to zero in controls. (C) Age-adjusted NFL concentrations differed between all groups. Levels of (D) T-tau and (E) P-tau were higher in manifest patients compared to premanifest before adjustment for age, but not after. The premanifest group had lower levels than controls, but the difference was significant only after adjustment for age.
Fig. 6. Longitudinal dynamics of biomarkers in Huntington's disease including YKL-40, NFL, T-tau and P-tau.

A

CSF YKL-40

B

CSF NFL

C

CSF T-tau

D

CSF P-tau

Fig. 6. Longitudinal biomarker dynamics of (A) Neurofilament light levels (NFL), (B) YKL-40 and (C) Total tau (T-tau), and (D) Phospho-tau (P-tau). Each line represents one individual, starting at the first CSF sample and protein concentration, finishing at the time of the second sample and concentration.

Paper III

The LC-MS metabolomics analysis yielded a total of 3701 metabolic features 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 principal component analysis (PCA) based on all metabolic features separated groups to a larger extent compared to PCA based only on identified metabolites. The manifest group was most clearly separated from the rest, indicating a unique metabolic signature.

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. 7a). 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. 7b). Both
models indicated segregations between the groups. The third model, comparing the premanifest HD subjects with controls was found to be non-predictive ($Q^2 < 0$).

*Altered CSF metabolites in manifest HD patients compared with controls.*

24 metabolites were significantly altered in this comparison. Ten of these were independently altered, and lumichrome, xanthine, N-acetylproline, isoleucine and L-DOPA remained significant after adjustment for multiplicity. No longitudinal differences in levels of the altered metabolites were found in manifest patients ($n=8$) on a second sample.

Comparing the two groups, five pathways were affected: aminoacyl-tRNA biosynthesis; the phenylalanine metabolism; valine, leucine and isoleucine biosynthesis; valine, leucine, isoleucine degradation and the purine metabolism. (Fig. 7c).

*Altered CSF metabolites in manifest compared with premanifest HD subjects.*

28 metabolites were significantly altered in this comparison, of which 14 were also of importance in distinguishing between manifest HD patients and controls. 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 multiplicity.

Eight pathways were significantly altered between the two groups: 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. 7d).

*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 tyrosine and phenylalanine metabolisms respectively. Additionally, tryptophan, tyrosine, phenylalanine and valine were recurrent in multiple pathways.

*Metabolite association to disease severity.*

A cluster consisting of thyroxine and dopamine from the tyrosine metabolism as well as 5-hydroxytryptophan and ophthalmate displayed positive correlations to the cognitive measurements and TFC, and negative correlations to DBS, TMS and to the 5 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. Notably, all metabolites from the tyrosine metabolism were associated to disease severity and HD symptoms, where thyroxine in particular showed significant correlations to all measures of disease severity, DBS and the 5 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 5 year risk of onset. Dopamine correlated with 5 year risk of onset as well as TMS, SC, SWR, SI and SDMT. Finally, tyrosine displayed correlations to TMS, SI, CVF and SDMT, while lumichrome only displayed a significant correlation to the 5 year risk of onset in premanifest HD subjects.
Fig. 7. Metabolomics. AUROC curves and altered pathways.

(a) Manifest HD vs. Controls

![Graph a](image)

(b) Manifest HD vs. Premanifest HD

![Graph b](image)

(c)

![Graph c](image)

(d)

![Graph d](image)

**Fig. 7.** 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.
Paper IV

In the fourth study, also utilizing LC-MS we identified proteins with altered levels in manifest HD compared with premanifest subjects. Age related effects on protein levels were estimated in the control group that completely spanned the age-distribution of all gene expansion carriers.

The significantly upregulated proteins (n=43) were in majority against the downregulated (n=10). Decreased levels of proenkephalin (PENK) and transthyretin (TTR) along with upregulated proteins (vasorin (VASN), stanniocalcin-2 (STC2), epsilon-sarcoglycan (SGCE) and complement component C7 (C7)) were all closely linked to HD symptom severity.

Validation of PENK results in a separate cohort including absolute quantitation is shown below in Fig. 8.

**Fig. 8. Proenkephalin (PENK) intergroup comparisons.**

---

**Fig 8.** Proenkephalin (PENK) intergroup comparisons. Boxes show first and third quartiles, the central bands show the median, and the whiskers show data within 1.5 IQR of the median. (A) PENK Intensity values in the Uppsala cohort are plotted unadjusted whereas the p-values indicate differences after age-adjustment by removing the age contribution from controls. (B). The validation cohort consisting of controls (n=23) and manifest HD patient samples (n=23) from Stockholm and Gothenburg. Unadjusted concentrations are plotted whereas the p-value indicates the difference after age-adjustment by including age in a regression model.
### Correlation matrix 1. Down-regulated proteins and correlations with HD symptoms

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Motor</th>
<th>Cognitive tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DBS</td>
<td>TFC</td>
<td>TMS</td>
</tr>
<tr>
<td>TTR*</td>
<td>-0.409</td>
<td>.444</td>
<td>.478</td>
</tr>
<tr>
<td>F12</td>
<td>-0.346</td>
<td>.213</td>
<td>.324</td>
</tr>
<tr>
<td>ORM2</td>
<td>-0.514</td>
<td>.327</td>
<td>.420</td>
</tr>
<tr>
<td>VTN</td>
<td>-0.452</td>
<td>.360</td>
<td>.447</td>
</tr>
<tr>
<td>HPX</td>
<td>-0.490</td>
<td>.371</td>
<td>.488</td>
</tr>
<tr>
<td>PENK*</td>
<td>-0.696</td>
<td>.652</td>
<td>.702</td>
</tr>
<tr>
<td>CFHR1</td>
<td>-0.417</td>
<td>.240</td>
<td>.381</td>
</tr>
<tr>
<td>RBP4</td>
<td>-0.377</td>
<td>.320</td>
<td>.405</td>
</tr>
<tr>
<td>AZGP1</td>
<td>-0.385</td>
<td>.270</td>
<td>.308</td>
</tr>
<tr>
<td>SOD3</td>
<td>-0.521</td>
<td>.470</td>
<td>.457</td>
</tr>
</tbody>
</table>

Significantly down-regulated proteins (*after age-adjustment) and their correlations with measures of symptom severity. Proteins are labeled with their respective gene names. Unadjusted spearman correlations (Rho-values). Statistically significant correlations have been bolded. DBS, Disease burden score; Total Motor Score; TMS, Stroop Word; S-W, Stroop Color; S-C, Stroop Interference; S-I, Category Verbal Fluency (animals); CVF, Verbal Fluency Letters; VFL, Total Functional Capacity; TFC, Symbol Digit Modalities Test; SDMT.
Correlation matrix 2. Up-regulated proteins and correlations with HD symptoms

<table>
<thead>
<tr>
<th>Protein</th>
<th>Motor</th>
<th>Cognitive tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DBS</td>
<td>TFC</td>
</tr>
<tr>
<td>GM2A</td>
<td>.222</td>
<td>- .455</td>
</tr>
<tr>
<td>OMG</td>
<td>.127</td>
<td>- .303</td>
</tr>
<tr>
<td>VASN*</td>
<td>.431</td>
<td>- .534</td>
</tr>
<tr>
<td>B4GAT1</td>
<td>.478</td>
<td>- .301</td>
</tr>
<tr>
<td>EFEMP1</td>
<td>.152</td>
<td>- .266</td>
</tr>
</tbody>
</table>
Significantly up-regulated proteins (*after age-adjustment) and their correlations with measures of symptom severity. Proteins are labeled with their respective gene names. Spearman correlations (unadjusted, Rho-values). Statistically significant correlations have been bolded. DBS, Disease Burden Score; Total Motor Score; TMS, Stroop Word; S-W, Stroop Color; S-C, Stroop Interference; S-I, Category Verbal Fluency (animals); CVF, Verbal Fluency Letters; VFL, Total Functional Capacity; TFC, Symbol Digit Modalities Test; SDMT.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Motor</th>
<th>Cognitive tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DBS</td>
<td>TFC</td>
<td>TMS</td>
</tr>
<tr>
<td>NEO1</td>
<td>.293</td>
<td>-345</td>
<td>.349</td>
</tr>
<tr>
<td>C1s</td>
<td>.214</td>
<td>.023</td>
<td>.189</td>
</tr>
<tr>
<td>ENO2</td>
<td>.555</td>
<td>-292</td>
<td>.500</td>
</tr>
<tr>
<td>PLXDC2</td>
<td>.296</td>
<td>-486</td>
<td>.484</td>
</tr>
<tr>
<td>ADAM22</td>
<td>.277</td>
<td>-057</td>
<td>.284</td>
</tr>
<tr>
<td>SEMA3G</td>
<td>.560</td>
<td>-284</td>
<td>.512</td>
</tr>
<tr>
<td>NDRG2</td>
<td>.288</td>
<td>-204</td>
<td>.357</td>
</tr>
<tr>
<td>SGCE</td>
<td>.537</td>
<td>-313</td>
<td>.535</td>
</tr>
<tr>
<td>GPC1</td>
<td>.037</td>
<td>-552</td>
<td>.508</td>
</tr>
<tr>
<td>PPIA</td>
<td>.309</td>
<td>-403</td>
<td>.450</td>
</tr>
<tr>
<td>FAT2*</td>
<td>.453</td>
<td>-256</td>
<td>.417</td>
</tr>
<tr>
<td>TYRO3</td>
<td>.165</td>
<td>-278</td>
<td>.384</td>
</tr>
<tr>
<td>STC2</td>
<td>.561</td>
<td>-505</td>
<td>.654</td>
</tr>
</tbody>
</table>
Discussion

Neuronal damage markers

The fact that correlations between NFL and clinical symptoms were higher than between NFL and disease burden score (DBS) is interesting. This excess power in prediction should reflect factors other than CAG and age, such as the impact of genetic modifiers and environment. This hypothesis was also strengthened in paper II where the somewhat higher number of participants allowed adjustment for (DBS) leaving mostly significant correlations with clinical symptoms. This shows that NFL has predictive power on the HD phenotype superior to that of DBS, a finding that is consistent with other studies (60, 72).

In addition to the current thesis findings, the work of other groups together mounts to convincing evidence that NFL is a robust marker of disease progression in HD, from early premanifest stages into advanced disease. We suggest NFL be used in clinical trials as a biomarker that can potentially be responsive to successful disease-modifying therapies. It has now been implemented as an explorative surrogate pharmacodynamic marker in the IONIS/Roche trial of HTTrx/RG6042.

Why NFL is especially well suited as a biomarker in HD

NFL is a structural component of axons which are predominantly located in white matter. The white matter degeneration begins before disease onset and follows the progression of HD as shown in the Track-HD study. This axonal loss is reflected by increasing NFL levels in CSF. Hence, NFL levels reflect the rate of axonal degeneration. Importantly, the turnover rate of NFL in CSF is notoriously slow, where increased levels from an isolated traumatic event lasts for weeks to months. HD however, being a rather slowly progressing disease poses no pressing need for markers with high temporal resolution. This may however mean that NFL cannot respond as fast as for instance mutant huntingtin levels were shown to do (in the IONIS HTTrx Phase 1b/2a trial (73)). Despite the slow turnover of NFL in CSF, it provides a tool for monitoring the rate of axonal degeneration, unlike for instance static imaging markers such as MRI that merely provide a snapshot of atrophy that has occurred at some unspecified time-point.

In contrast to another disorder with prominent white matter involvement, relapsing-remitting multiple sclerosis (RRMS), HD should not cause any
sudden increases or unpredictable fluctuations in NFL concentrations by virtue of its slowly progressive course. The longitudinal profile of NFL in HD gene expansion carriers also supports this hypothesis (paper II and Byrne et al (72)). Further the anatomical areas affected in HD are more homogenous compared with both RRMS, and amyotrophic lateral sclerosis which are other disorders where NFL measurement is applied (74). This fact explains the strong correlations observed between NFL levels and HD symptoms as demonstrated in the current thesis findings.

The case for tau as a disease progression marker is overshadowed by the performance of NFL that has often been reported concomitantly in previous publications. Tau (and P-tau, displaying almost identical levels in paper III) also has significant predictive power on the HD phenotype after adjustment for disease burden. The issue here as with many candidate markers is that it has not been proven that tau increases before motor onset, although one study reported increased levels in a pooled group of manifest and premanifest subjects (61).

Neuroinflammatory markers

Interestingly the increase of sCD27 indicates intra-thecal T-cell mediated inflammation before disease onset in HD, with even higher levels in manifest HD. The lack of associations to other HD symptoms however is unexpected. Perhaps adaptive immune activation as reflected by sCD27 is an early feature that contributes to disease onset rather than to further deterioration of neurological function. In any case, this finding needs validation in a larger cohort. YKL-40, which showed no significant intergroup differences after age-adjustment in our study, consistent with previous studies altogether shows limited promise as a biomarker of disease progression especially since it is not sensitive to progression in premanifest HD.

Metabolomic characteristics

The tyrosine metabolism, including tyrosine, thyroxine, L-DOPA and dopamine, showed the highest impact based on the pathway analysis indicating a role in disease progression. Thyroxine demonstrated reduced levels in manifest vs. premanifest HD subjects, moderate associations to all measures of disease severity and correlation with 5 year risk of onset in premanifest subjects. Reduced thyroxine levels in CSF have been reported in other neurodegenerative disorders such as Alzheimer’s disease (AD) (75) and decreased tyrosine has been reported in post mortem HD brain tissue (frontal lobe and striatum) (76). In addition, perturbed tyrosine metabolism has also been reported in AD
and Parkinson’s disease (PD) (77). Wild-type HTT has been reported to bind to the nuclear thyroid hormone receptor-α1 (TRα1) (78) providing a potential mechanism involved in 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 reduced in manifest HD patients and to various degrees associated with disease severity. Similar to thyroxine, dopamine levels also correlated with the 5 year risk of onset. These observations suggest a derangement in tyrosine metabolism and more specifically in the dopaminergic pathway in HD, even though the derangement of this system is not well understood and previous findings have been inconclusive, as reviewed by Byrne and Wild (66).

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) (77).

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 slight correlations to symptom severity. Decreased levels of these large neutral amino acids have previously been reported in post mortem brain tissue, CSF and/or serum and plasma in HD subjects, findings hypothesized to be the result of impaired energy metabolism (76, 79-82). The valine, leucine and isoleucine degradation pathway, herein found perturbed in manifest vs. premanifest HD and controls, has also been reported in PD (77).

Moreover, the purine metabolism was altered in manifest compared with premanifest HD and controls. Previously, this has also been found in AD, ALS and PD and suggested to be an alternative pathway for neurodegenerative diseases to overcome the inadequate glucose supply and the deficient energy metabolism (77). Manifest HD patients had increased levels of purine metabolites (xanthine and its precursor hypoxanthine, adenosine, adenine, cyclic AMP, and inosine). Increased adenosine levels found in the striatum of transgenic HD mice have been attributed to mitochondrial dysfunction (83).

Intriguingly, lumichrome displayed a twofold decrease in manifest HD patients compared with controls, but no difference between manifest and premanifest HD. Comparing all mHTT carriers to controls supports the finding that decreased levels of lumichrome are typical to mHTT carriers. Interestingly, lumichrome correlated with the 5 year risk of onset, 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). Riboflavin is an essential compound in the brain which has, as a majority of other findings herein, been connected to the energy metabolism.
where a riboflavin deficiency has resulted in a mitochondrial dysfunction (84). Riboflavin is also essential for monoamine oxidase (MAO) activity (85), which is involved in the metabolism of neurotransmitters such as norepinephrine, serotonin and dopamine. The changes in lumichrome levels could be linked to increased MAO-activity due to locally increased levels of dopamine (86), 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 the nervous system in HD possibly mediated by mitochondrial dysfunction, previously suggested by multiple studies (76, 87-90).

**Proteomic characteristics**

In the present study we identified CSF proteomic differences between manifest HD patients and pGECs. Because the goal was to identify markers of disease progression we highlighted proteins that also displayed correlations with symptom severity. Below we will discuss the potential markers identified and their connection to HD pathophysiology.

CSF PENK was decreased in manifest HD patients compared with pGECs, and in manifest HD compared with controls as demonstrated in the validation cohort. There was also a trend towards lower levels in pGECs compared to controls. Although this trend did not reach significance, there was an inverse correlation between PENK and the 5-year risk of onset among the pGECs. PENK levels also decreased along with the progression of HD symptoms. PENK which is expressed highly in a subset of medium spiny neurons (MSNs) of the caudate nucleus, has to our knowledge not previously been measured in the CSF of HD gene carriers. However, a PENK related peptide (Met-enkephalin-Arg6-Gly7-Leu8, (MERGL)) was reported to be down-regulated in HD (and progressive supranuclear palsy) compared to controls (91). MERGL levels seemed to be affected by antipsychotics in the former study, but we could not find differing levels of PENK in the participants using antipsychotics. Contrary from the gaba- and enkephalin producing MSNs, neuropeptide Y (NPY)-containing neurons are spared in HD and higher levels of NPY have been reported (92). The authors suggested that a ratio between NPY and Dynorphin-A or enkephalin peptides could serve as a biomarker for HD onset or disease progression. Furthermore, PENK is a stable surrogate marker for the enkephalins that due to their instability ($T_{1/2}$ 15 min) are less likely to serve as meaningful HD biomarkers. Development of PENK as a blood marker for HD may be challenging as it is also expressed in non-neuronal tissues (93) and has already proven to increase with renal- and cardiac dysfunction (94).

Strikingly low levels of TTR (fold change 0.61) were found in the manifest HD group compared with pGECs and controls, and the levels correlated
moderately with HD symptoms. Interestingly, decreased TTR has been reported in manifest HD before (95, 96) although associations to HD symptom severity were not reported. While TTR is a carrier of thyroxine in the brain it is also an amyloid binding protein and has been implicated as a possible therapeutic target in Alzheimer’s disease (AD) (97). Reduced CSF TTR in AD patients has also been described (98). Retinol binding protein 4 which is carried by TTR was also decreased in manifest HD compared with pGEC in the present study. In a recent study (99) in the same HD-cohort we found that the tyrosine metabolism was the most impacted pathway with markedly reduced levels of thyroxine in CSF of manifest HD patients compared to pGEC. The reduction of TTR could, according to the dominant hypothesis from AD, reflect a failing protective mechanism against the aggregation of toxic beta-amyloid. Since mutant huntingtin forms amyloid like aggregates (100) this mechanism warrants further study.

In agreement with previous studies in HD (44, 101) we found evidence of complement activation. C7, a component of the membrane attack complex, was elevated and correlated highly with symptoms. Elevated plasma C7 in HD patients has also been reported (44). C7 is regulated by clusterin which was also upregulated in our study, but hardly correlated significantly with any HD symptoms. Other complement components (C4A and -B), both increased with severity of symptoms, but neither met significance for intergroup difference after age-adjustment. C1s was on the contrary elevated but did not correlate with HD symptoms. Both IL-6ST and CD44 were increased but displayed only weak- to moderate correlations with symptoms. Elevated IL-6 levels in plasma and CSF in HD have indeed been reported before (45).

VASN, which was strongly associated to HD symptoms, is a hypoxia induced protein which attenuates TGF-beta by binding (102). Other proteins that were closely linked to HD symptoms include STC2 - a regulator of insulin like growth factor bioavailability (103), and the brain enriched protein SGCE which is encoded by the same gene that is linked to myoclonus dystonia syndrome (DYT11) (104).
Limitations

The main restraint was limited sample size, which can be considered small to moderate compared with similar studies. This may cause insufficient statistical power which is a common issue in the research field of neurology. The main issue is that of type I error in which a study appears negative (“p-values > 0.05”) because of limited power. In the targeted analyses our results were congruent with similar studies and evidently power increased with only a few more participants between paper I and II that strengthened the findings (for instance when correlations with clinical scores of NFL and tau survived adjustment for disease burden). Furthermore, non-hypothesis driven methods as applied here in metabolomics and proteomics can yield type II error (random findings due to multiple statistical testing). We addressed this with false discovery rate (FDR) testing in paper III and by applying several criteria for markers to be clinically relevant in paper IV and also by validating results in a separate cohort. The manifest group was deemed too small for stratification into manifest disease stages and instead correlations were tested with TFC. A potential limitation is that the sample collection was not strictly standardized to time of day or timing of meals. We cannot exclude bias from diurnal variation of markers. This is not generally an issue with markers of neuronal death (i.e. tau and NFL), but would rather be a concern for metabolomic features and reassuringly the findings were similar on analysis of repeated samples. Antidepressants and antipsychotics were more commonly used by the manifest HD patients than controls and premanifest subjects. We hypothesized that including only drug-free patients would cause a selection bias and limit external validity. Instead we repeated statistical tests after excluding patients on drugs which resulted in similar but underpowered results in the metabolomics study.
Conclusions

NFL is a suitable biomarker for disease progression in HD and may be of use as a pharmacodynamics marker in intervention studies. This finding has been consistent throughout several studies in different HD cohorts. Blood based NFL-assays may also be useful as Byrne and colleagues have demonstrated (105).

Our results regarding sCD27, suggest that T cells may contribute to neuroinflammation before motor onset in HD. sCD27, may after further study come to serve as a marker of inflammation in HD. The adaptive immune system may contribute to inflammation and neuronal death, by cell-mediated immunity. This field is largely unexplored and needs further study for clarification.

Tyrosine metabolism, including the dopaminergic species, were 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 5 year risk of onset in premanifest subjects. A large proportion of the biochemical changes suggest impaired energy metabolism in HD, where decreased levels of lumichrome were typical for all mHTT carriers. These findings contribute to characterizing the molecular pathogenesis in HD.

PENK and TTR (with decreased levels) along with VASN, STC2, SGCE and C7 (with upregulated levels) were all closely linked to HD symptom severity. The levels of PENK begin to decrease before disease onset and follow disease progression. We hypothesize that declining PENK levels reflect the degeneration of medium spiny neurons (MSNs) that produce PENK, and that assays for PENK may serve as a surrogate marker for the state of MSNs in HD.
Huntingtons sjukdom är en allvarlig och ärtlig neuropsykiatrisk sjukdom. Den som bär på anlaget insjuknar i genomsnitt kring 40 års ålder och får oftast psykiatriska symtom samt framskridande intellektuell och rörelsefunktionshinder (inkusive karakteristiska ofrivilliga rörelser, sk. chorea). Även om vi vet att sjukdomen framskrider över tid, är det viktigt att hitta pålitliga mätmetoder för sjukdomsaktiviteten, så kallade sjukdomsmarkörer. Sjukdomsmarkörer som mäts i kroppsvätska från patienter kan förhoppningsvis användas för uppskattning av effekten av bromsläkemedel i kommande studier.

I denna avhandling som består av fyra delarundersöktes sjukdomsmarkörer i ryggmärgsvätska hos en grupp personer med anlaget för Huntingtons sjukdom. Deltagarna (vars antal varierade något mellan studierna) var i olika sjukdomarveld varit en förutsättning för att kunna se om ämnena som mättes i ryggmärgsvätska hade koncentrationer som hängde ihop med svårighetsgrad på symtomen. En andel av personerna med yngre genomsnittsålder var friskmärgsålder (11-13 st) medan en andra grupp bestod av insjuknade patienter (12-14 st). Även personer utan anlaget för Huntingtons sjukdom inkluderades i studierna (43st) för att utgöra en jämförelsegrupp gentemot anlagsbärarna.

Olika mätmetoder användes i delstudierna (I-IV). Immunbaserade metoder användes för att mäta koncentrationer av på förhand utvalda markörer. Resultaten bekräftade att nivåerna av ett tidigare föreslaget protein, neurofilament light (NFL) självständigt förutsäger svårighetsgraden av symtomen och sjukdomsstadier. NFL fanns i låga nivåer hos friska, något högre nivåer hos symtomfria anlagsbäraren och högst nivåer mättes hos de insjuknade. Fördeken med NFL var att det verkar kunna mäta den diskreta och tidiga sjukdomsaktivitet som föregå debutsymtomen. Ett protein som reflekterar immunsystemets aktivering, löslig CD27 receptor (sCD27), var förhöjt enligt samma mönster som NFL men utan tydliga samband med de enskilda symtomens svårighetsgrad. sCD27 föreslås studeras ytterligare före några slutgiltiga slutsatser dras om dess användbarhet.

En kemisk metod, masspektrometri, användes för att identifiera nya substanser, proteiner och andra mindre molekyler. Intressant nog identifierades både tidigare föreslagna proteiner och några nya vars nivåer skiljer sig mellan sjukdomens stadier. Proenkephalin som produceras av en nervcellpopulation...
som är särskilt utsatt vid Huntingtons sjukdom var lågt och nära associerat till
symtomens svårighetsgrad. Vi kunde dessutom bekräfta att proenkephalin var
lägt i en separat grupp huntingtonpatienter från sjukhus i Stockholm och Gö-
teborg. Utöver proteinstudierna gjordes även en kartläggning av mindre mo-
lekyler (s.k. metaboliter) där man kunde se en särskild molekylär signatur för
den insjuknade subgruppen. Resultaten pekade på en störning av centrala ni-
våer av vissa ämnen (aminosyror, sköldkörtelhormon och signalsubstanser) –
samtliga kopplade till dopaminsystemet som är av betydelse för koordinat-
onen av kroppens rörelser. Dessa fynd är nya och behöver bekräftas i ytterli-
gare studier, helst med fler studiedeltagare.
Acknowledgments

Most importantly, thanks to all the brave patients who gave their time and provided samples.

Anne-Marie Landtblom, for your patience and intense encouragement.

Dag Nyholm, for always being the most reasonable person in the room with high scientific rigor combined with a great sense of humor.

Jimmy Sundblom, for introducing me to the field of HD, giving me the relevant projects for the thesis, for continuous hands-on advice and support, and being a constant source of inspiration.

Joachim Burman, for your excellent support as my clinical supervisor, and for great advice in scientific writing.

Stephanie Herman, for a great collaboration, impressive statistical knowledge, and many nice moments working on metabolomics. Also thanks to Kim Kultima, for this successful collaboration.

Anna Shevchenko and Jonas Bergquist, for excellent work in the proteomics facility and introducing me to the LC-MS method.

Kristina Becanovic, for stringent monitoring of data and many nice moments working together.

Everyone from the neurology outpatient clinic who helped me with sample collection. No one is forgotten.

The clinical HD team, without you there would be no steady stream of patients coming to our clinic. Thanks for procuring the data for this thesis together.

HD collaborators in Sweden, especially Per Svenningsson, Martin Paucar, Radu Constantinescu and Maria Kneider for our inspiring collaboration on biosamples.
My parents, Pekka and Birgitta, who unintentionally inspired me to work with neurology.

Emily, my love, for your patience when I talk about the thesis and being my in-house English language expert.


38. Ferrer I, Goutan E Fau - Marin C, Marin C Fau - Rey MJ, Rey Mj Fau - Ribalta T, Ribalta T. Brain-derived neurotrophic factor in Huntington disease. (0006-8993 (Print)).
43. Petersen A, Gabery S. Hypothalamic and Limbic System Changes in Huntington's Disease. (1879-6400 (Electronic)).


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)