Universiteit
Leiden
The Netherlands

Blood and Biomarkers in Huntington's Disease

Mastrokolias, A.

Citation

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

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

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

Cover Page

Universiteit Leiden

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

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

Blood and Biomarkers in Huntington's Disease

Anastasios Mastrokolias

Financial support for all research described in this thesis was provided by the Centre for Medical Systems Biology (CMSB), a joint activity of Leiden University Medical Center and Leiden University aiming to improve diagnosis, therapy and prevention of common and rare diseases.

Financial support provided by the CMSB for the printing and publication of this thesis is gratefully acknowledged.

Cover theme: STRING protein association network
Cover design: Anastasios Mastrokolias

PhD Thesis, Leiden University Medical Center, Leiden 2017
ISBN: 978-94-6332-195-2

Printed by: GVO drukkers & vormgevers B.V.

© Anastasios Mastrokolias, Athens

Copyright of the published chapters is held by the publisher of the journal in which the work appeared. All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means without permission of the copyright owner.

Blood and Biomarkers in Huntington's Disease

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 7 september 2017
klokke 11:15 uur

door

Anastasios Mastrokolias
geboren te Athene
in 1979

Promotor:
Prof. dr. G.J.B. van Ommen

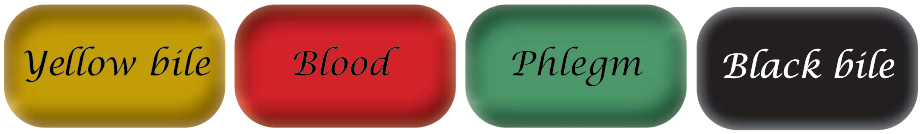
Copromotor:
Dr. W.M.C. van Roon-Mom

Overige leden:
Prof. dr. J.T. den Dunnen
Prof. dr. O.C. Meijer
Prof. dr H.H. Kampinga - University of Groningen
Dr. P.A.C. 't Hoen
Dr. J.C. Dorsman - VU Medical Center

This thesis is dedicated to
Nina, Elias and Katerina

Humoralism was an old medical philosophy according to which health and disease result from the imbalance between the following four bodily fluids or humors.

Hippocrates of Kos - 460 B.C. and Galen of Pergamum - 130 A.D.



Even though the four humors theory was much later displaced by modern medicine the word humor persists until today as a euphemism for being ... "in good humor".

Table of Contents

Chapter 1. Introduction

1.1 Huntington's Disease background
1.1.1 Huntington's disease discovery and epidemiology 12
1.1.2 Huntington's disease genetics 12
1.1.3 Huntington's disease motor symptoms 14
1.1.4 Huntington's disease cognitive and psychiatric symptoms 15

1.2 Huntington's disease neuropathology
1.2.1 Huntingtin protein structure and function 16
1.2.2 Mutant huntingtin aggregates and inclusions 18
1.2.3 Huntington's disease brain pathology 19
1.2.4 Huntington's disease peripheral pathology 20

1.3 Huntington's disease mechanisms of cytotoxicity
1.3.1 Proteasomal deregulation 21
1.3.2 Transcriptional deregulation 22
1.3.3 Mitochondrial activity and energy metabolism deregulation 24
1.3.4 Synaptic dysfunction and excitotoxicity 25
1.3.5 Autophagy 26

1.4 Huntington's disease therapeutic strategies
1.4.1 Transcriptional activity restoration strategies 27
1.4.2 Mutant huntingtin lowering strategies 28
1.4.3 Mutant huntingtin clearance and anti-aggregation strategies 30
1.4.4 Mitochondrial and synaptic function restoration strategies 31
1.4.5 Other strategies 32

1.5 Biomarkers
1.5.1 What is a biomarker ? 34
1.5.2 Biomarkers in Huntington's disease - needs and challenges 36
1.5.3 Blood based biomarkers for Huntington's disease 38
1.5.4 State of Huntington's disease clinical and imaging biomarkers 40
1.5.5 Other molecular and biochemical fluid biomarkers 42
1.6 Reference List 44

Chapter 2. Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA. BMC Genomics 2012, 13: 28 79
- Introduction 81
- Materials and Methods 83

- Results 87
- Discussion 96
- Reference List 99
- Supplementary Material 102

Chapter 3. Huntington's disease biomarker progression profile identified by transcriptome sequencing in peripheral blood. European Journal of Human Genetics (2015) 23, 1349–1356 105
- Introduction 107
- Materials and Methods 108
- Results 111
- Discussion 121
- Reference List 126
- Supplementary Material 131

Chapter 4. Integration of targeted metabolomics and transcriptomics identifies deregulation of phosphatidylcholine metabolism in Huntington's disease serum samples. Metabolomics (2016) 12: 137 137
- Introduction 140
- Materials and Methods 142
- Results 146
- Discussion 155
- Reference List 161
- Supplementary Material 169

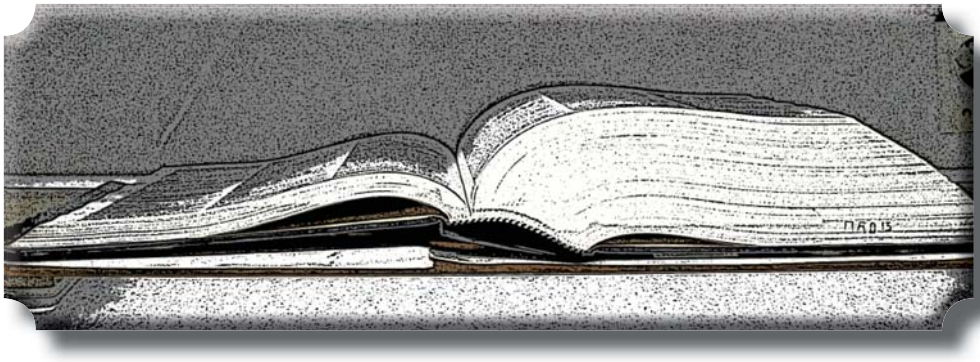
Chapter 5. General discussion

5.1 Summary 174
5.1.1 Globin Reduction Findings 174
5.1.2 DeepSAGE gene expression findings 177
5.1.3 Metabolomics and integration with transcriptomics findings 179
5.2 Limitations and challenges 181
5.3 Future directions 183
5.4 Epilogue 185
5.5 Reference List 188

Thesis summary 193
Samenvatting 197
Publication List 201
Curriculum vitae 202
Acknowledgements 203

Chapter 1

General Introduction



1.1 Huntington's Disease

1.1.1 Huntington's disease discovery and epidemiology

Huntington's disease is an autosomal dominant neurodegenerative disease that manifests through motor and psychiatric symptoms. Huntington's disease was first described and named after George Huntington in 1872; who in his paper "On Chorea" described the major aspects of the disease [1]. However, it was not until 1983 when the disease was first mapped to a restriction fragment length polymorphism marker (RFLP) (D4S10/G8) [2] on chromosome 4 and it took 10 years more (1993) before it was linked to "interesting transcript 15" (IT15) following which the HD gene was identified and isolated [3]. Since current knowledge about the human genome was unavailable at the time, the long search for the identification of the disease gene was further limited by the telomeric location of the RFLP marker, the physical distance as well as uncertain recombination frequencies between the marker and the disease locus [4,5]. HD population prevalence exhibits significant geographic differences depending on local, past migration patterns and ethnicity. The disease is most prominent in Western European populations with a rate of occurrence of 5-10 in 100.000, a prevalence of approximately 1 in 10.000 in Caucasian populations [6,7] and is least frequent in Japanese, Chinese, Finnish and African descent individuals [8]. However, as suggested by Evans et al., the disease prevalence has risen significantly in the last decades and it was found that for UK the number of HD patients was 12.3 per 100.000 [9]. Localized geographic areas exhibit very low as well as very high prevalence with the most notable example being the population of Lake Maracaibo in Venezuela which was also originally used to identify the disease gene [10].

1.1.2 Huntington's disease genetics

Huntington's disease is caused by an expanded CAG repeat in the HTT gene. Individuals that carry a CAG repeat higher than 35 are the ones that will most likely develop the disease symptoms at some point during their lifetime [11]. People with 36-39 repeats are known to carry a reduced penetrance allele and they may develop the disease rather later in life (40% chance asymptomatic at age 65) [12]. For most mutation carriers disease

onset takes place between 35 and 50 years old [13]. However, symptom onset can occur at any time during the carrier's lifetime and is inversely correlated to the CAG expansion length. As a result, people with 40 repeats or more will develop the disease and when the repeat exceeds 50 symptoms may develop as early as childhood. The elongated CAG tract results in a mutant huntingtin protein with an elongated polyglutamine tract that is associated with protein aggregation and toxic gain-of-function properties [14]. During meiotic transmission, the HD repeat is unstable in more than 80 % of the cases (increases and decreases) while the largest increases are seen in paternal transmissions [15]. This meiotic instability leads to genetic anticipation of the disorder (increasing intergenerational disease severity) but also to the generation of de-novo HD mutation carriers from unaffected parents; especially from parents which exhibited an increased glutamine repeat number and within the intermediate repeat area of the disease (27-35 CAGs) [16,17]. Nevertheless, the HD gene mutation only explains about 50-77 % of the variance in the age of symptom onset and scientists are continuously looking for genetic as well as environmental modifiers of disease onset [18,19]. Several genetic modifiers have been identified that further influence the trinucleotide repeats instability as well as the age of onset of the disease. Such genetic modifiers influence the manifestation of disease phenotypes and partially explain the variability and degree to which these symptoms develop [18,20,21]. Apart from genetic modifiers, there are also specific factors that lead to the unstable transmission of the, disease causing, trinucleotide repeat expansion and include toxic oxidation which is known to cause progressive age-dependent expansions [22,23] as well as population specific haplotypes made up of sets of single nucleotide polymorphisms [24]. On the other hand, genetic modifiers that influence the disease onset include functional polymorphisms in key genes involved in neuronal processes, the processing of huntingtin itself, gene transcription but also metabolism. Examples of affected genes include the a) kainate receptor *GluR6* [25] b) glutamate receptor, ionotropic, N-methyl-d-aspartate – *GRIN2A/2B* [26] c) gene transcription factor *TP53* [27] d) the Peroxisome Proliferator-Activated Receptor Gamma, Coactivator 1 Alpha (*PPARGC1A*) transcription factor that is involved in regulating energy metabolism genes [28,29]. Finally, in a recent very large genome wide association study two different chromosomal loci were discovered (chromosomes 8 and 15) that exhibited independent effects of disease onset acceleration and/or delay [30]. The functional variants within

the above two chromosomal loci are thought to affect genes involved in DNA synthesis and repair, mitochondrial energetics and oxidative stress (*FAN1, MTMR10, RRM2B, UBR5*). The presence of all the above genetic differences in HD mutation carriers shows that detailed genotypic information will be required in order to efficiently address the manifested disease symptom variability.

1.1.3 Huntington's disease motor symptoms

One of the most prominent features of HD is the movement disorder known as chorea [31]. Chorea is characterized by hyperkinetic dance like movements and ranges from random, exaggerated gestures and expressions to continuous and violent movements. Huntington's disease patients are also characterized by non-choreatic, hypokinetic movements such as bradykinesia and dystonia. Hypokinetic movements are usually more prominent during middle and late stages of the disease while, on the other hand, hyperkinetic movements are usually present at early to middle stages of the disease. Specifically, as the functional capacity of patients declines, the presence of chorea lessens and the presence of dystonia is intensified [32]. Additional motor symptoms include the slowness of movements (bradykinesia), rigidity and clumsiness. Aside from chroreatic or dystonic motor symptoms patients also exhibit skeletal muscle wasting and weight loss (see 1.2.4) [33]. The above two observations take place irrespective of the fact that HD patients receive a steady caloric intake [34]. HD patients further exhibit ocular abnormalities such as the inability to suppress reflexive saccades to a new optical stimulus, and are characterized by delayed initiation of voluntary saccades (reflexive and voluntary eye control respectively) [35]. These ocular dysfunctions can be in turn linked to potentially affected brain areas since it has been shown that saccadic eye movement control is partially controlled by the hippocampus, the inferior parietal lobule and the prefrontal cortex [36,37]. Finally, in subsequent and more advanced stages of the disease the patients are exhibiting dysarthria (speech problems), dysphagia (swallowing difficulties) and balance disturbances, with the latter two problems constituting potential death causes by aspiration pneumonia or serious injury respectively [38-41].

1.1.4 Huntington's disease cognitive and psychiatric symptoms

HD is characterized by cognitive and psychiatric symptoms such as depression, irritability, apathy, anxiety and dementia. The above symptoms are usually evident well before the onset of any motor symptoms and get significantly worse over time. Nevertheless, psychiatric symptoms do not always correlate with cognitive and motor aspects of the disease and specific symptoms are sometimes more highly associated with certain disease stages, such as the manifestation of depression at the same time or after onset of motor symptoms [42,43]. Previous studies have identified that depression is a significant component of the prognostic signature of HD and that such affective disorders can appear as early as 20 years before choreatic and cognitive symptom onset [44]. Another study of depression in HD has reported that over 40 % percent of HD patients had significant symptoms of depression and over 50 % had previously received treatment for depression [45]. These numbers are significantly higher than the prevalence of depression in the general population (8-12 %) [46]. The exact reason behind the high number of reports of depression in HD is not clear. However, it has been hypothesized that apart from potential perturbed biological pathways this could be also due to the adjustment of these patients to the idea that they suffer from a disabling, terminally ill disorder [47,48]. Moreover, anxiety and irritability have also been reported in HD at a prevalence varying from 34-61% and 38-73% respectively [49] which is also much higher than the reported lifetime prevalence of any other anxiety disorder as reported from the World Mental Health surveys[50]. HD patients are likewise often characterized by apathy and specifically reduced energy and lack of drive. It has been proposed that in HD and other neuropsychiatric diseases, apathy is distinct from depression and associated with disinhibition and aberrant motor behavior, in contrast to anxiety and irritability which characterize depression [51]. Finally, obsessive and compulsive symptoms as well as psychotic symptoms have also been reported in HD mutation carriers, even though at a much lower prevalence (10-52 % and 3-11% respectively) [43,49].

1.2 Huntington's disease neuropathology

Huntington's disease is characterized by atrophy of the cerebral cortex and the striatum

and HD patients are known to lose up to 40% of their brain volume compared to healthy individuals [52,53]. Neurodegeneration affects mostly the medium spiny projection neurons of the striatum while the interneurons are relatively spared [54]. The neurodegeneration mechanism of the mutant huntingtin protein however appears to be complex and neuronal death does not always correlate to huntingtin nuclear inclusions [55]. One of the most used methods to classify HD neuropathological severity is the Vonsattel ascending severity order grading system (Grades 0-4) that was formulated using postmortem brain specimens in order to identify the microscopic and macroscopic changes taking place in mutation carrier brain tissue [56]. Vonsattel et al. reported a caudate nucleus neuronal loss of 50-95% from grades 1-4 and a big increase of astrocytic population during grades 2-4. The authors also reported that the earliest changes were found in the medial paraventricular areas of the caudate nucleus and dorsal part of the putamen as well as the fact that anatomical changes were usually lagging behind the clinical symptoms something that has also been confirmed by later studies [57,58]. Apart from the neuronal degeneration pattern that is evident in the striatum of HD patients, other areas are affected too such as the substantia nigra, globus pallidus, thalamus and cerebellum [59]. The multiplicity and the variable degree of degeneration of the affected brain areas constitute the basis for the diverse clinical phenotype and disease progression rates the patients exhibit. The latter assumption has been supported by imaging studies, which have revealed a widespread brain tissue degeneration and significant volume reductions in almost all brain areas; occurring even in early stage patients [60-63].

1.2.1 Huntingtin protein structure and function

Wild type Huntingtin (Htt) protein is 348 kDa, expressed ubiquitously and its highest levels are found in the brain and the testis [3,64-66]. The exact 3D and molecular structure of the protein has not been yet achieved due to the high molecular weight of the protein and the problems this causes in X-ray crystallography. Currently, in an ongoing attempt to identify its protein structure huntingtin has been transferred to the international space station in an effort to take advantage of extraterrestrial weightlessness condition and microgravity crystallization techniques (NASA mission experiment CASIS PCG HDPCG-1). Structurally, the HD protein is characterized by a CAG repeat and

by a proline rich domain which are located in the protein N terminus [3]. The presence of similar homopolymeric glutamine and proline stretches in many transcription factors has lead scientists to the conclusion that huntingtin is also one such protein with similar functional properties [67-69]. Apart from its distinctive N-terminal structure the protein possesses several well described consensus motifs that are required for many protein-protein interactions – which is another characteristic of Htt [70]. Such patterns, sequences or motifs include: a) Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A (PP2A), and the yeast kinase TOR1 - (HEAT) - repeats [71] b) proteolytic cleavage sites that allow the fragmentation of the protein through the function of caspases 2,3,6 and 7 [72,73] c) nuclear localization sequences [74] and d) posttranslational modification sites such as those for its phosphorylation and ubiquitination [75-78]. The wild type protein was initially considered to be found only in the cytoplasm and associated with the Golgi apparatus, the microtubules and the endoplasmic reticulum [79-81]. However, an early study by Hoogeveen et al. and later one by De Rooij et al. showed that huntingtin was also present in the nucleus [82,83]. This discovery further emphasized important protein roles such as transcriptional control, a role supported by the structural features and motifs that Htt possesses, like poluglutamine stretches and leucine zipper motifs. Huntingtin is also found in clathrin coated vesicles and it has been proposed that it interacts with proteins involved in cytoskeletal transport, receptor mediated endocytosis as well synaptic recycling [84,85]. Even though the highest huntingtin levels are found in the central nervous systems [86-88] the exact function of the wild type protein is still not completely understood. Nonetheless early research in the field into the function of the protein has shown that huntingtin is required during gastrulation in early embryogenesis and has a vital role in the functioning of the basal ganglia and neuronal survival [89,90]. Additionally, further studies have demonstrated that huntingtin may possess many further roles in many key cellular pathways. Several roles have been proposed for huntingtin: possessing an anti-apoptotic role and protecting from excitotoxicity [91,92], regulating the production of neurotrophic factors such as brain derived neurotrophic factor (BDNF) [93,94], participating in or disrupting axonal transport and vesicle trafficking [95,96] and regulating synaptic function through partner proteins such as PSD-95 [97,98].

1.2.2 Mutant huntingtin aggregates and inclusions

Soon after the discovery of the HD causal mutation the presence of neuronal aggregates and inclusions of mutant huntingtin was also established as one of the main disease pathological hallmarks. The designation of the aggregates as the main disease hallmark was made possible through a series of successful experiments that used mouse and human studies to initially locate neuronal intranuclear inclusions and dystrophic neurites in HD brain [99] and later identify their role in toxicity, or protection therefrom, and in disease neurological phenotypes [100,101]. Even though the mechanisms through which Htt inclusions or Htt fragment inclusions lead to toxicity and neurodegeneration have not been completely elucidated. Some of the hypotheses advanced include interference and sequestration of essential cellular proteins and transcription factors, proteasomal component interference and functional impairment [102,103], apoptosis [104], mitochondrial function damage [104,105] and alterations in cytoskeletal and synaptic function [106]. The factors that determine the capacity of the mutant Htt to aggregate include the length of the polyglutamine expansion and the length of huntingtin polyglutamine fragments [107], the specific conformation of the mutant protein as well as the sequences that flank the polyglutamine stretch such as the poly-proline motif located downstream of the CAGs [108]. Miller et al. have suggested that the rate-limiting step for the formation of inclusion bodies (nuclear or cytoplasmic aggregates) are the different conformational changes in monomeric Htt exon 1 fragments [109]. This suggests that only modest reductions in the levels of the mutant proteins or even their relative conformation ratios can have a profound effect in aggregate formation and toxicity [109,110]. Another significant factor correlated with the increased aggregate formation is the mutant protein proteolytic cleavage and the subsequent formation of N-terminal fragments. Specifically, the function of proteases such as caspase and calpain have been shown to increase the formation of aggregates and neuronal susceptibility to cell death [111-113]. Independently of all the above factors influencing aggregate formation, the propensity of the mutant huntingtin to aggregate and cause disease pathology could additionally depend on chaperones [114,115], posttranslational modification enzymes [116,117] and can be influenced by the specific neuronal cell type properties and the mutant huntingtin expression levels and dynamics in particular cell types such as primary

neurons, astrocytes, oligodendrocytes and microglia [118-120].

1.2.3 Huntington's disease brain pathology

Postmortem examinations of HD carriers have shown that the brain tissue changes are mostly evident in advanced stages [121] and are characterized by tissue losses in the cerebral cortex and thalamus (reductions 21-29%) as well as the caudate and putamen (57% and 64% tissue loss respectively). Even though these large tissue losses are more visible in later disease stages, recent studies have shown that such brain atrophy and white matter disorganization events can even take place before the onset of disease clinical symptoms [122]. For example, medium spiny neurons make up to 95% of the striatal neuronal cell population and contain a plethora of metabotropic, glutamate and dopamine receptors [123,124]. These receptors have been the subject of intense research as to the possible mechanisms through which they can cause or contribute to neuronal cell death (see 1.3.4). In the majority of HD cases the binding capacities of these receptors is reduced and this capacity reduction has been correlated to neuronal cell death [125,126]. Nonetheless it is not completely clear yet if this receptor dysfunction leads to cell death or vice versa [126]. On the other hand, some cell types such as the striatal interneurons, the medium-sized calretinin-positive interneurons and the neuropeptide Y/somatostatin containing neurons are relatively spared, even in more advanced disease stages [127]. The reason that these cell types are less susceptible to the disease is not yet fully understood. It has been postulated however that this preservation can be related to the distribution of the cells in the corresponding brain areas as well as their relative position to excitatory neuroreceptors responsible for cell over-excitation and subsequent cell death [128]. Finally, it has been shown that the cellular and neurochemical changes that take place in HD patient brains are accompanied by gliosis [129-131]. Gliosis is characterized as an increase in the population as well as the activation of astrocytes, microglia and oligodendrocytes [128]. Microglial activation has been found widespread through the cortex, the pallidum and striatum and it has been even found to be present in pre-symptomatic carriers [132,133].

1.2.4 Huntington's disease peripheral pathology

Even though Huntington's disease has been traditionally considered as a brain disorder it is characterized by additional peripheral symptoms such as weight loss and skeletal muscle wasting, glucose intolerance, heart and gastrointestinal problems as well as testicular atrophy [134-138]. An increasing number of studies points to the fact that these peripheral symptoms are not solely or even partially due to neurodegeneration of brain cell structures and brain factors but are a result of the mutant protein expression in the specific peripheral tissue cells and occur in parallel with the central nervous system [107,139]. Notably intracellular mutant protein inclusions as well as peripheral skeletal muscle cell dysfunction have been shown to be present even when the peripheral cell types are isolated [140,141]. Weight loss is one of the most common symptoms and especially in symptomatic HD individuals. Several studies have indicated that this weight loss is not only associated with nervous system damage but is also a result of an increased metabolic rate in these patients [142,143]. Moreover, it has also been shown that weight loss does not always follow symptom onset and a multifactorial system could govern the start of weight changes in symptomatic and pre-symptomatic patients [136] Furthermore according to a different study skeletal muscle wasting can also be attributed to expression of the mutant huntingtin protein in the myocytes and a transition from slow-twitch to fast-twitch muscle fibers [144]. In addition, and using mouse models, it was shown that mutant huntingtin affects gene expression patterns in myocytes and that these expression changes are in parallel with early pathological and behavioral changes [144,145]. Other human as well as mouse studies have shown that HD mutation carriers are more likely to develop osteoporosis, diabetes, pancreatic dysfunction and impaired glucose tolerance [146-148]. Finally, the role of the immune system has been the focus of several studies and it has been shown that in HD blood is characterized by increased concentrations of proinflammatory and chemotactic cytokines such as interleukins 6 and 8, altered inflammatory signaling and a general overactivation [149-151]. This deregulation of the immune system has been attributed to expression of the mutant Htt protein in the immune system cell types affecting cell signaling and cytokine release and migration [152]. Most importantly such an immune system over activation does not have an isolated cell type effect but appears to influence other prominent peripheral degeneration events described

above such as muscle wasting and gene expression deregulation [153,154].

1.3 Huntington's disease mechanisms of cytotoxicity

Even though HD is a monogenic disorder the disease pathology is very multifaceted. The cytotoxic function of the aggregated mutant Htt protein can exert its effect through a plethora of processes as described above (see 1.2.2). In addition, it has been shown that mutant huntingtin cleavage products (especially N-terminal fragments) exacerbate the disease pathology. They are more toxic than full length huntingtin and increase the number of intranuclear inclusions [55,155]. Huntingtin is proteolytically cleaved by caspases like caspase 3 and 6 and it has been shown that inhibiting caspase activity can reduce toxicity and aggregate formation (see 1.4.5). It has been proposed that increased caspase activation and substrate cleavage, caused in HD by intranuclear Htt accumulation and release of mitochondrial cytochrome c [156], can also be the cause behind phenotypic similarities across other neurodegenerative diseases such as Alzheimer's disease [55,157,158]. In the following paragraphs, previously mentioned and additional mechanisms of cytotoxicity are described in more detail.

1.3.1 Proteasomal deregulation

The ubiquitin-proteasomal system is one of the main mechanisms of cellular protein degradation in the nervous system and is involved in processes such as neuronal transmission and neuronal plasticity [159]. In HD, it has been shown both in humans and animal models that the mutant huntingtin protein is targeted for proteolysis but is resistant to its processing [99]. Studies that used HTT exon1 transfection polyglutamine aggregates labelled with antibodies against parts of the ubiquitin-proteasomal system demonstrated co-localization with the protein inclusions as well as altered localization that could potentially adversely affect the proteasomal system function [160]. The reduced proteasomal ability to cleave the mutant protein as well as other polyglutamine proteins is attributed to the greatly expanded length of glutamine repeats of the mutant protein. As a result, these proteins are getting "stuck" in the 20S subunit of the proteasome and do not allow for the subsequent entry of the remaining catalytic substrates and other

proteins [161]. Additionally, using fluorescence recovery after photo-bleaching (FRAP) optical technology it was shown that the proteasome was irreversibly trapped within the N-terminal mutant huntingtin aggregates both in vivo and in vitro [162]. Even though the results of the previous study did not exclude the possibility of partial proteasome activity or Htt partial degradation – over time this delay in huntingtin or other substrate processing could have deleterious and adverse effects for cellular viability. Moreover, it should be noted that the results concerning the inability of the proteasomal system to degrade the mutant expanded proteins are often contradictory. For example, and in contrast to previous studies, Rousseau et al. have suggested that the aggregation is not a result of the main 20 S proteasomal subunit failure but rather of 19S regulatory particles (Rpt4/Rpt6) flanking the proteasomal catalytic core [163]. Additional data that further diffuse the role of proteasomal dysfunction in HD have shown that it is only the filamentous mutant Htt aggregates that have the potential to inhibit proteasomal function and not the formation of inclusion bodies (IBs) [164]. Despite of the opposing study results described above about the role of proteasome in HD, this proteolytic complex constitutes a very interesting therapeutic target since even if its function is shown not to be affected its activity may be targeted for additional stimulation and enhanced clearance of the polyglutamine fragments [165].

1.3.2 Transcriptional deregulation

The wild type huntingtin protein has also been shown to have a role in transcription regulation and to interact with many other transcription factors and regulatory proteins[166]. On the other hand, during disease it has been confirmed that Htt mutant protein and protein fragment aggregates are involved in the sequestration of proteins involved in transcription, irregular subcellular localization of CAG containing transcripts, alternative splicing; all events that lead to genome wide transcriptional deregulation in HD patients [167]. Specifically, it has been established that important cellular transcription factors such as TATA binding protein, CBP, SP1 and p53 are recruited to these aggregates [168-170]. In two such studies it was shown that Sp1, a well-known wild type Htt interactor, exhibits reduced DNA binding (caused by mutant Htt) and that genes regulated by Sp1 exhibited decreased expression [145,171] . Gene expression microarray

studies in human and animal models have shown that the functional types of those genes whose expression is affected by mutant huntingtin include neuronal structure genes, signaling receptors, axonal receptors and stress response genes [172,173]. In a recent study by Labadorf et al. it was also shown that HD brain deregulated genes were enriched for immune response and neuroinflammatory genes as well as homeotic genes (Hox – Homeobox genes) – a set of genes that has not yet been widely investigated in the light of HD [174]. Furthermore, it has been suggested that these mRNA changes were not attributable to cell loss only and exhibited a distinct disease related brain regional pattern, with gene expression similarities between those expressed in the HD caudate and the motor cortex [175]. Other important factors and regulators that have been implicated in HD transcriptional deregulation and whose function is influenced by the mutant protein include the neuronal factor transcriptional repressor *REST* [176], the peroxisome proliferator-activated receptor gamma (*PPARγ*) [177] and the heat shock protein transcription regulator *NF-Y* [178].

Another major transcriptional mechanism disrupted in HD, which has potentially even greater influence than individual transcriptional factors deregulation, is chromatin remodeling. Chromatin remodeling is regulated by epigenetic processes such as histone posttranslational modifications and DNA methylation [179]. Studies from HD models have reported that mutant huntingtin also interferes with the function of enzymes involved in chromatin remodeling such as histone acetyltransferases, histone methyltransferase, ubiquitin ligases and transglutaminases [180]. Specifically, it has been shown that inhibition of histone deacetylases by inhibitors such us sodium butyrate and phenylbutyrate reversed the mutant protein related neurodegeneration [181,182]. Histone methylation is an epigenetic mechanism used to activate or repress transcription however it has been demonstrated that in HD histone methylation related proteins such as ERG-associated protein with SET domain (ESET) and its activator Sp1 are also deregulated in HD [183] Post-translational modifications as the ones performed by transglutaminases have also been shown to exhibit increased activity while their inhibition with cystamine in HD models ameliorated symptoms and increased the expression of neuroprotective genes[184]. Much attention has been given lately to the role of miRNAs in HD and their gene expression regulatory roles like for miR-9 and the neuronal gene expression RE1-Si-

lencing Transcription Factor (REST) [185,186]. Finally, other ways in which chromatin remodeling is affected in HD and normal function of the histone core proteins is disrupted are monoubiquitilation and post-translation modifications such as phosphorylation of serine, threonine and tyrosine residues [172,187].

1.3.3 Mitochondrial activity and energy metabolism deregulation

Mitochondria are well known regulators of cellular health and are involved in cellular energy production, metabolism as well as apoptosis [188]. The mitochondria are another cellular compartment that is thought to be disrupted in HD and eventually also contributing to neuronal cell death. A lot of studies both in animal models and humans have described many different aspects of mitochondrial dysfunction in HD. The different mechanisms of HD mitochondrial dysfunction can be briefly categorized into : a) mitochondrial membrane destabilization b) reduction of the mobility of the mitochondria c) reactive oxygen species (ROS) overproduction with a concomitant Ca2+ cellular vulnerability through the deregulation of the mitochondrial permeability pore [105]. More specifically, studies have shown that HD mitochondria exhibit increased CAG repeat length correlated membrane depolarization [189], reduced membrane potential requiring lower calcium depolarization loads [190] and reduced CA2+ loading capacity as well as evidence that CA2+ dependent pathways lead to oxidative phosphorylation impairment of HD mitochondria [191]. It has been demonstrated that full length mutant Htt blocks mitochondrial movement and this effect correlates with increased mutant protein expression in the cytosol [192]. The same study showed that the mitochondria accumulated in areas of aggregates while their trafficking was unchanged at neuronal areas that lacked aggregates. Reactive oxygen species, produced by the mitochondria during oxidative phosphorylation, have shown to be increased in HD. This has been attributed to the presence of the mutant protein [193] and microglial over-activation which typically occurs after brain injury or immunological stimuli and has been shown to correlate with disease severity [194,195]. Moreover it has been shown that dysfunctional mitochondrial metabolism in HD patients leads to diminished complex II, III and IV activity in the caudate nucleus [196], decreased glucose metabolism and inhibition

of the metabolic regulator and transcriptional coactivator PGC-1alpha [197-199], as well as increased lactate production in the cortex and the basal ganglia [200,201]. Finally, the deregulation of neurotransmitters such as dopamine and glutamate present at the striatum and their disordered interactions with the mitochondrial system could also account for the preferential vulnerability of these brain areas to the mutant Htt protein which involves but is not limited to processes such as the reduction of cellular ATP stores, deregulation of intracellular Ca2+ levels, increased synaptic dopamine levels and impaired glutamate uptake due to increased production of ROS [202,203].

Despite the fact that so many studies have provided important clues about the deregulation of individual genes and complexes of mitochondrial metabolism it is still unclear to what extend these mitochondrial defects have a causal effect in the disease or if they are just the result of long term disease toxicity [204]. Finally, for any one of the aforementioned genes or molecules to be targeted in therapeutic approaches e.g. to equilibrate their levels, their role in energetic pathways, and their interactors, have to be extensively investigated to avoid unwanted non-specific detrimental effects.

1.3.4 Synaptic dysfunction and excitotoxicity

Several studies have produced strong evidence that HD neuronal cells and especially those that make up the tissues that are first affected in the disease such as the striatum, are especially vulnerable to excitotoxic events and are characterized by aberrant synaptic and extra-synaptic functioning. HD model studies have shown that pre-symptomatic mice exhibit striatal dopamine signaling deficiencies and that dopamine receptor expression gradually declined with age, indicative of the potential role of the dopamine system in HD neuritic malformation and subsequent disease pathology [205,206]. The uptake of another important neurotransmitter, glutamate, has been also shown to be reduced in the prefrontal cortex of HD patients [207]. This reduction in glutamate uptake is usually also accompanied by a reduction in glutamate transporter 1 (*GLT1/EAAT2*) in corticostriatal regions [208,209]. The exact mechanism by which mutant huntingtin interferes with synaptic functioning and the deregulation of neurotransmitter levels is not fully understood yet. Nevertheless, HD model studies have shown that potential mecha-

nisms could include the Htt mutant fragment transcriptional deregulation of glutamate transporters that are mainly expressed by astrocytes that envelop synapses [210,211] and/or an inability to sustain the bioenergetics/glycolytic requirements for the proper functioning of the enzymes responsible for glutamate uptake reactions (GLT-1, GAPDH). These enzymes in turn cause a neurotoxic release of glutamate to the extracellular medium [212,213]. Finally, other complementary studies that have provided clues to the potential mechanisms involved in such excitotoxic cell death have reported that: a) full length mutant huntingtin disturbs neuronal Ca2+ signaling and makes medium spiny neurons sensitive to glutamate-induced apoptotic events [214] b) anti-apoptotic and pro-apoptotic proteins such as the cell death regulator Bcl-2 and the transcription factor p53 are selectively deregulated in striatal neurons and make these neurons more vulnerable to excitotoxic cell death [215] c) changes in the levels or removal of BDNF or its tyrosine kinase receptor TrkB depletes the pro-survival signals that these molecules provide and inhibits striatal cellular functionality [216] and d) reductions in endocannabinoids and mutant Htt mediated disruption of their transcription which could in turn influence neurotransmitter release, synapse plasticity and the ability of striatal cells to respond to excitotoxicity challenges [217-219].

1.3.5 Autophagy

Increasing evidence points to the fact that the wild type Htt has a vital and active role in autophagy and is involved in autophagosome transport and biogenesis processes [220,221]. For this reason, several important clinical autophagy modulation clinical trials are under development that aim to enhance the ability of selective or general autophagic pathways to increase substrate recruitment, lysosomal activity and clear toxic proteins [222,223]. Autophagy is a cellular, protein clearance system that is divided into three main forms microautophagy, chaperone mediated autophagy and macroautophagy [224]. This classification serves to describe the way the cargo is transported to the lysosomes, and degraded by hydrolases. In macroautophagy the mechanism of toxic protein degradation involves autophagosome vesicle formation, its maturation and finally its fusion with the lysosomes [225,226]. In HD, it has been reported that the macroautophagic vacuoles are inefficient in engulfing their cytosolic loads and this in turn leads

to a reduction in their turnover, an increase in their total number in HD cells and can also be linked to the increased levels of protein aggregates and stored lipids [227]. Additionally, an important regulator of the macroautophagic process, the serine/threonine kinase mammalian target of rapamycin (mTOR) has also been reported to be inhibited and this inhibition has been attributed to the presence of the mutant Htt protein [228]. A separate study has also shown that the accumulation of mutant Htt can also stimulate the endosomal-lysosomal system and this in turn can lead to autophagic cell death [229]. Studies also support the notion that the expanded protein loses its ability to reversibly bind to the endoplasmic reticulum, the Golgi complex, maintain the wild type protein cellular localization properties and regulate stress induced autophagy [230-232] all of which have been associated with neuronal dysfunction and neuronal cell death in HD [233]. In chaperone mediate autophagy (CMA) the substrates are selectively chosen for destruction and delivered to the lysosomal membrane [234]. This is in contrast to the macroautophagy process where whole areas of the cytosol are engulfed and targeted to the lysosomes. In HD models, it has been shown that there is a compensatory regulatory mechanism where the CMA autophagic process activity is increased in response to a compromised macroautophagic pathway [235]. This increased activity is mediated through the upregulation of lysosome-associated protein type 2 a (LAMP-2A) and lysosomal HSC70 both of which proteins have been shown to be directly involved in the clearance of Htt [236]. However, with time the CMA autophagic pathway also gradually declines in its ability to compensate for the dysfunction of the macroautophagic system and as a result there is a generalized failure of autophagic clearance. In microautophagy the lysosomal membrane envaginates the material destined for destruction and internalized by the lysosomal lumen, however its molecular components are still not well understood in mammals [237,238].

1.4 Huntington's disease therapeutic strategies

1.4.1 Transcriptional activity restoration strategies

One of the potential methods to correct for the transcriptional deregulation observed in HD is through the inhibition of histone deacetylase activity. Histone deacetylase inhib-

itors such phenylbutyrate and sodium butyrate have been previously used in an attempt to improve transcriptional dysfunction by the relaxation of DNA conformation, target HD specific pathological events such as caspase mediated cell death [239,240] or even improve HD phenotype through common neurological mechanisms such as improving neuronal BDNF trafficking [241]. In studies where histone deacetylase inhibition was used in HD models, it improved motor performance, appearance and body weight of the mouse models. It has been shown that deacetylase inhibition also reverses mutant huntingtin incited expression and causes a potentially beneficial change in the expression of a specific set of genes associated with the immune system and cell cycle – death, in the cerebellum, the cortex and the striatum of HD mice [242]. Apart from the inhibition of genome-wide transcriptional regulators such as histone deacetylate inhibitors, the transcriptional regulation of individual genes important for the HD has also been suggested as an alternative approach. Examples of such genes are the cAMP response element-binding (*CREB*) and CREB co-activator CBP, brain-derived neurotrophic factor, *REST* as well as the zinc finger transcription factor *SP1*. CREB signaling function has been shown to be reduced in HD and several studies have attempted to increase CREB levels in an effort to improve neuronal survival [243]. Giampa et al. showed that in a rat HD model phosphodiesterase type 10 inhibition increased CREB levels in spiny neurons and provided a neuroprotective effect [244]. Finally, in order to control pathological transcriptional deregulation, several compounds that affect transcription have been used such as anthracycline antibiotics [245]. Stack et al. showed that anthracycline administration in R6/2 HD mice improved nucleosome dynamics and the disease neurological phenotype, shifting the balance between histone methylation and acetylation, while Ryu et al. demonstrated that mithracyclin suppressed *SP1, SP3* related transcription in neurons through the suppression of the ESET (H3-K9) histone methyltransferase and increased HD mouse model survival rates [246,247].

1.4.2 Mutant huntingtin lowering strategies

Silencing of mutant huntingtin by RNA interference notably using siRNAs directed towards the mutant mRNA has become one of the most promising HD therapeutic trials [248]. By blocking the expression of the mutant protein the gain of function effects of

the mutant Htt can be potentially reversed [249]. However, in this strategy it is important to prevent the silencing of the wild type protein as well as other non–specific silencing events [250,251]. In this way, the beneficial effects of silencing can be maximized and patient safety is ensured by avoiding toxic side-effects of non-specific silencing, notably, of wild type Htt that has been associated with perinatal lethality and motor and cognitive defects in specific HD mouse models [252,253]. Even though the first attempts at silencing the mutant protein were performed using RNA interference (RNAi) technologies [254] subsequent studies showed that delivery and efficacy could be improved through the use of single stranded small inhibitory RNAs (siRNAs) and by using chemical modifications and design improvements of the administered molecules [255]. Moreover, additional experiments that have also shown great promise have been targeting the silencing of the mutant protein through the use of antisense oligonucleotides (AONs). The main difference between AONs and RNAi is that while AONs are delivered as single strand molecules RNAis are delivered as delivered as duplexes. Furthermore, RNAis are consequently engaged by the Argonaute protein family as the catalytic components of the RNA-induced silencing complex (RISC) lead only one of the two strands to eventually bind the complementary RNA [256]. Even though RNAi and AONs recognize their targets through Watson and Crick base pairing rules additional key differences that separate the two technologies are that compared to AONs RNAi works through a well-defined RNA degradative pathway (Drosha/Dicer), their targeting efficiency towards the mutant protein, the delivery efficiency to the targeted tissue and their half-life within the administered tissue cells [257,258]. Furthermore, since the silencing of wild type huntingtin function is undesirable due to its vital function in a plethora of cellular, neurological and neurodevelopmental processes [259] a lot of effort has been put into the allele selective silencing of the mutant protein, something that can be done through the targeting of single nucleotide polymorphisms [260,261]. Since the first silencing studies in which a significant reduction of mHtt was achieved through the administration of siRNA and anti-Htt constructs [262,263] a lot of preclinical HD model, Htt lowering studies have already been performed [264]. Some of the most recent, promising studies include those that incorporate the administration of cholesterol-conjugated Htt-siRNA in mice in order to increase tissue uptake [265], the allele specific knockdown of mHtt with S-constrained-ethyl (cET) motif AONs [266] and the AONs RNase-H-mediated

degradation of huntingtin through administration in HD mice cerebrospinal fluid [267].

1.4.3 Mutant huntingtin clearance and anti-aggregation strategies

In order to address the issue of the expanded mutant huntingtin resulting to a toxic gain of function a potential therapeutic strategy has been to identify modulators that possess anti-aggregational roles and that can lead to mutant protein aggregation and inclusion body reduction. A good example of experiments that attempt to reduce the accumulation of aggregated mutant proteins is through the use of chaperone protein member families. Chaperone proteins such as HSP40 and HSP70 have been known to promote refolding and degradation of misfolded proteins [268,269]. In a similar approach Labbadia and colleagues showed that induction of the heat shock response (HSR) via the heat shock factor 1 transcription factor (HSF1) and inhibition of chaperone HSP90 improved HD related phenotypes in the R6/2 mouse model [270]. A subsequent study from the same group and in HD transgenic mouse crossbreeds overexpressing the neuronal chaperone HSJ1a (DNAJB2a) reduced mutant huntingtin aggregation and enhanced solubility [271]. Nonetheless, some of the above ameliorative effects, of the activation and overexpression of the chaperone machinery, have been transient and it has been suggested that synergistic combination therapies might be required for these therapeutic strategies to be more effective [272].

Other anti-aggregation approaches include those that use the administration of various, nontoxic, soluble disaccharides that act to stabilize the partially unfolded mutant huntingtin and prevent disease inclusion formation. One such compound, trehalose has been used in HD as well as other polyglutamine and neurodegenerative disease models and has shown to prevent neurodegeneration, enhance autophagy and alleviate disease mediated pathology [273-275]. Finally, much evidence suggests that cellular huntingtin clearance through induction of autophagy can also provide therapeutic benefits [276-279]. Apart from potential small molecule regulators of autophagy like trehalose and as described above, other approaches to regulate autophagic degradation of mutant huntingtin include: a) post-translational modifications of the mutant protein such phos-

phorylation of specific N-terminal serine residues [280,281] or acetylation of lysine 444 (K444) [282] b) autophagy induction through mTOR inhibition with rapamycin [283] or other mTOR dependent pathways and molecules [284,285] c) through other mTOR independent autophagy inducers such as L-type Ca2+ channel antagonists, K+ATP channel openers and other pharmacophores [286-288] .

1.4.4 Mitochondrial and synaptic function restoration strategies

Since mutant huntingtin has been shown to interfere with mitochondrial energy production and metabolism [289,290] it has been hypothesized that the administration of compounds and molecules that could replenish the energy cellular banks and improve the energy deficient profile in patients may also represent a likely therapeutic pathway. One of the molecules that has been proposed to have the potential to rectify the energy deficit in HD patients is creatine [291]. Creatine is synthesized in the kidney and the liver, taken up by tissues such as brain, heart and the skeletal muscle through a sodium dependent transporter [292] and together with phosphocreatine serve as short term energy banks when energetic demands rapidly increase [293]. Creatine has been shown to improve motor performance and reduce brain atrophy as well reverse mitochondrial dysfunction in transgenic mouse models of HD [294,295]. Additionally, a 16-week clinical trial of creatine administration showed that creatine was well tolerated and reduced oxidative stress levels in HD patients [296]. However even though creatine was also shown in a subsequent study [297] to improve cortical thinning, other aspects of the disease such as cognition were not affected. Therefore, considering the potential side effects of long term creatine administration, especially in presymptomatic individuals, further studies will be required to prove the applicability of this compound in HD. Another corresponding molecule, with neuroprotective and antioxidant properties, is the electron transport chain cofactor coenzyme q10 (CoQ10) which is involved in aerobic respiration and ATP energy production. CoQ10 has in some cases been combined with creatine administration to provide additional neuroprotection in HD as well as Parkinson's disease mouse models [298]. Multiple HD transgenic mouse models have demonstrated the neuroprotective properties of CoQ10 and shown that it can protect dopaminergic neurons, reduce

mitochondrial toxins as well as improve mouse model rotarod performance in combinatorial, preclinical trials [299,300]. It should be noted however that the results of clinical studies of coenzyme Q10 [301] have not been so promising as HD mouse model studies and considering the low treatment benefit compared to patient effort and energy requirement, further clinical trials of this molecule have been currently halted.

Along with mitochondrial restoration strategies much attention has been given to the restoration of cortico-striatal synaptic function in an effort to prevent selective vulnerability and neurodegeneration. The restoration of synaptic function has been the aim of many preclinical and clinical trials to regulate the release of neurotransmitters such as dopamine and glutamate, the functioning of neuroreceptors such as the N-methyl-D-aspartate receptor (NMDAR) and the expression levels and the release of neuromodulators and growth factors such as the endocannabinoids and brain derived neurotrophic factor [302]. In particular, experiments on transgenic R6/2 mouse models showed that dampening of the aberrant glutamate uptake by activation of metabotropic glutamate auto-receptors, resulted in increased model survival and a significant rescue in striatal and cortical neuronal loss [303]. Another attempt to prevent striatal loss and atrophy in HD as well as other neurodegenerative diseases has been through the overexpression of BDNF [304]. In HD, it has been shown that the increase of BDNF levels in YAC128 mouse models has resulted in the normalization of dopamine D2 receptor expression while striatal cell loss and motor dysfunction have also been reduced [305]. Finally, additional synaptic rescue schemes include the modulation of synaptic transmission via endocannabinoid signaling [306,307] and administration of neurotransmission inhibitory drugs as riluzole, memantine and the dopamine pathway inhibitor tetrabenzine [308-311].

1.4.5 Other strategies

Cell replacement therapy is another favorable approach that can potentially delay disease progression by replenishing the cells that have been lost due to the pathological effects of the mutant Htt protein. To date, cell therapy approaches have been through the use of stem cells, neural progenitor cells as well as fetal tissue cells [312]. El-Akabawy et al.

showed that the implantation of human neural stem cells in the R6/2 mouse model had suboptimal results due to low cell survival and differentiation. However, a different study by Rossignol et al. showed that co-transplantation of adult mesenchymal and neural stem cells in a transgenic rat model prolongs the survival of the neural stem cells and creates a more favorable transplantation microenvironment [313,314]. Transplantation of fetal striatal grafting has also been shown to improve the motor as well as the cognitive decline of HD patients. Nonetheless the grafts exhibited poor survival over time and conversely the grafting beneficial effects diminishing over time [315,316]. The poor survival of the grafts in the host has been attributed to poor vascularization. Trophic support can potentially be improved as grafting techniques and protocols are continuously improved [317,318]. Nevertheless, achieving the delicate balance of the type of stem cells to use (MSCs or iPSCs), during which HD disease stage (early, middle or late) and the specific neurodegeneration events these cells will target, are the greatest challenges of future stem cell trials and the determinant success factor for this therapeutic method [152].

To conclude the list of therapeutic approaches it has been suggested that targeting caspase dependent cleavage events is another prospective approach for the prevention of neuropathological features in HD and other neurodegenerative disorders [319]. According to the "toxic fragment hypothesis" the presence of multiple proteolytic sites on the Htt protein, in combination with the aberrant activity of proteases such as caspases and calpains may lead to numerous, variable length polyglutamine containing peptides that can be toxic. The identification of the specific combinations of enzymes and sites causing the most toxic peptides has been the focus of several research groups. Using YAC HD mouse models it has been shown that mice expressing mutant huntingtin resistant to caspase 6 cleavage, were protected against neurotoxic events and did not develop striatal neurodegeneration [320]. For this purpose, drugs like minocycline have been used in preclinical trials and in an attempt to inhibit the production of huntingtin toxic fragments by caspase cleavage [321,322]. Furthermore, initial clinical trials of minocycline use have exhibited stabilization of motor function and a significant amelioration of psychiatric symptoms in HD patients [323]. Even though the role of caspases is still not completely understood and the experimental results are not always consistent with the initial results from the caspase 6 resistant YAC mice [324], new molecules are constantly being

developed against caspase-6 aiming to increase the therapeutic potential of this method [325,326]. Yet increasing evidence points at important non-apoptotic roles of caspases such as caspase-6. Furthermore, alternative, complementary proteolytic pathways exist, which continue to give rise to the HD proteolytic toxic peptides (even after ablation or downregulation of caspase-6 function). This means that research in this direction should proceed with caution as to the molecules used and the extent to which the function of these proteases can and should be modified.

1.5 Biomarkers

1.5.1 What is a biomarker?

A biomarker is a physiological, molecular, biochemical or pharmacological measurement that can be obtained from the body of a patient and that can provide a good indication of disease prediction/progression or treatment efficacy in that patient [327]. Additionally, and according to the US Food and Drug administration (FDA) pharmacogenomics guidance (http://www.fda.gov/drugs) a biomarker should be measured in an analytical test system with clearly established performance characteristics and for which there is an established scientific framework that elucidates the significance of the biomarker test results. A good biomarker should also have the following properties: a) be able to detect disease features with high sensitivity (able to detect the condition when it is present) and high specificity (able to rule out a condition when it is absent), especially when compared to diseases with similar phenotypes and b) be easily, reliably and robustly detected, preferably in an easily collectable tissue (such as peripheral blood) and with minimum distress to the patient [328]. Finally, a good biomarker is judged by its positive predictive power (proportion of patients with a correct positive diagnosis over the total number of individuals who have been called to have the disease, true and false) and its negative predictive value (proportion of patients with a negative test and a correct diagnosis over the total number of people that have been called to not have the disease, including the false negatives).

Based on their disease specific properties and their relationship to disease, pathology

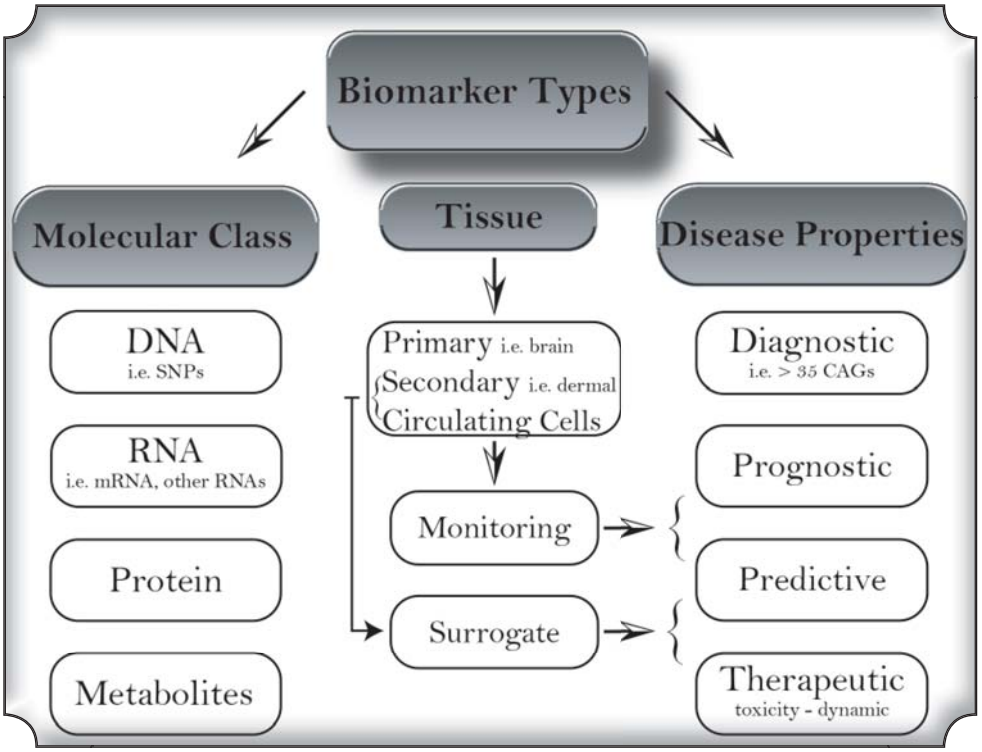


Figure.1 Biomarker type classification based on: a) the molecular subclass from which the identified biomarker originates b) biomarker tissue relative to disease central pathology tissue and c) biomarker provided information relative to disease risk (diagnostic), disease progression (prognostic), therapeutic potential (predictive) and safety or effectiveness of therapeutic approach.

biomarkers can be classified into: a) biomarkers which predict patient survival (prognostic), b) biomarkers which follow disease progression (monitoring biomarkers), c) biomarkers which can be used to evaluate drug safety and efficacy (pharmacodynamic), d) biomarkers which identify individuals likely to benefit from a treatment (predictive) and e) biomarkers which can predict the outcome given the response to therapy (surrogate biomarkers– see Figure 1) [329]. Prognostic and diagnostic biomarkers are grouped together and are otherwise called type 0 category biomarkers while therapeutic and surrogate biomarkers have been defined as type 1 and type 2 biomarkers respectively [330]. A prognostic biomarker can provide information about the outcome of the disease in

specific population groups however it cannot provide information about the patient's response to a treatment such as those provided by predictive biomarkers. On the other hand, pharmacodynamic markers can provide information about how well the treatment works (proof of principle) or can be used to optimize the dose of the administered treatment. Surrogate biomarkers are measurements that can potentially replace clinical endpoints and this is performed in a less labor intensive or expensive way than the standard and established clinical endpoint measurements methods for the specific disease. A biomarker can be an intrinsic measurement such as those provided by brain or other tissue imaging techniques, biomolecular and fluid sample tests or they can also be extrinsic. For example, an extrinsic surrogate biomarker can be cigarette smoking to the clinical end point of lung cancer. Moreover, biomarkers can be divided to the level to which they take place in a disease (molecule, cell etc.) or even according to their relationship to disease pathology (primary or secondary pathology) [331]. A good biomarker should also undergo validation and qualification [332]. Validation refers to the evaluation of the biomarker's sensitivity, specificity and reproducibility while on the other hand biomarker qualification refers to acquisition of the evidence linking a biomarker to a specific clinical endpoint [333]. Even though a biomarker can be linked, validated and qualified to a clinical endpoint (surrogate biomarker) this may not always be associated with disease causality [334]. Specifically, such a biomarker may be associated with the clinical endpoint but could nonetheless have nothing to do with the disease tissue pathology or the molecular pathways that lead to the disease specific endpoint. Such a biomarker could for example be a measurement in the peripheral blood of neurodegenerative disease patients not correlated to brain degeneration and its associated molecular pathways and yet correlated to specific disease endpoints. Finally, it should be emphasized that apart from the disease specific, drug and disease endpoint biomarker relationships, vital properties of a biomarker should be the detailed reporting of laboratory, experimental, clinical and statistical information that will allow for the unambiguous replication and reconstitution of the initial biomarker study results [330,335].

1.5.2 Biomarkers in Huntington's disease - needs and challenges

Different biomarkers can be developed for different clinical purposes. No diagnostic bio-

markers are needed for HD as diagnosis is done by the genetic identification of the number of CAG repeats. Thus, HD biomarker development is focusing on the areas of pharmacodynamics, therapeutic monitoring and disease progression diagnosis. Additionally, the ability to identify potential HD mutation carriers provides a period of several years to decades that can be used for presymptomatic therapeutic intervention. This is further emphasized by the fact that even though individuals exhibit no distinct phenotypic features during this preclinical, presymptomatic disease stage, subtle subclinical molecular and neurodegenerative events are already taking place [336-339]. Another reason that identification of good biomarkers is a top priority for HD research is the recent development of many potential treatments and therapeutic molecules (see 1.4). These therapeutic strategies have been partially validated with the help of various HD models. However, unless effective and standardized biomarkers have been described it will be very difficult to objectively describe any ameliorative outcomes of potential treatments when patients progress to the clinical trial stages. Such biomarkers should ideally describe the disease progression or even the pathology reversal equally well in early and presymptomatic patients, considering that these patients would benefit the most from therapeutic intervention. Nevertheless, even though high standards are required for the discovery of sensitive biomarkers, the multiplicity of pathological features in HD dictate that more than one type of biomarkers will be required to accurately describe the disease progression and the clinical trial benefits. Such a panel of biomarkers would be taking advantage of clinical progression tests, technological developments in imaging technologies and refinement of biochemical markers already under development. For Huntington's disease as well as other neuropsychiatric disorders, one of the main problems in studying the actual disease pathology is that it is hardly possible to collect tissue samples from brain areas of interest or obtain representative patient sample numbers to achieve the desired statistical power for a conclusive result [340,341]. Additionally, postmortem samples, which are sometimes also used to study the disease, are usually of suboptimal quality, either due to postmortem delay or due to pathological events affecting the tissue for several years, like in Huntington's disease [342-344]. Finally, development of qualified (see 1.5.1) and informative pharmacodynamic, monitoring and surrogate biomarkers will not only aid in the development of clinical trials and therapeutic agents but also assist researchers to get a better understanding of the disease pathology itself.

1.5.3 Blood based biomarkers for Huntington's disease

As described above HD patients are often irritable and exhibit choreatic symptoms and psychiatric disturbances (see 1.1.3/1.1.4). This makes it even more challenging to perform biopsies and secondary pathology tissue samplings and examinations in symptomatic patients. At the same time the ease with which peripheral blood can be collected repeatedly over a period of time, makes it a very attractive tissue for use in preclinical, clinical and biomarker discovery studies. Furthermore, huntingtin is also expressed in blood and, as such, surrogate biomarker changes could be identified in blood independently of potential causality to central disease pathology. The changes found in blood could reflect the effect of the mutant huntingtin itself within the tissue or secondary pathological events from other tissues, such as muscle. For example, it has been suggested that in neurodegenerative diseases the blood-brain barrier might be compromised, potentially exacerbating of brain pathology and causing leakage of proteins or other molecules [345]. Finally, profiling of gene expression in HD patients has shown that in whole blood and caudate samples transcription of large genomic regions is regulated in a synchronized fashion [346]. However, independently of the pathological mechanisms or other events traced by blood biomarkers they will need to be robustly associated with well-established HD clinical endpoints.

Blood consists mainly of plasma (55%), red blood cells (45%) and leukocytes and platelets (last two together ~ 1%). Most blood HD studies have been performed either using plasma and serum to examine the free proteins and molecules in these liquids; or using the leukocytes to examine cellular organization and disease related processes. The blood component or molecule under examination is often chosen based on the suspected deregulated disease pathway. By following one such hypothesis-driven approach and to study oxidative damage and dysfunction of the mitochondria in HD, serum levels of 8-hydroxy-2'-deoxyguanosine (8OH2'dG), an oxidative DNA damage indicator, have been investigated [296,347,348]. Its levels have been reported to be elevated in HD and other neurodegenerative diseases and HD mouse models overexpressing a hydrolase that break it down have conferred a protection against HD-like symptoms [349,350]. Additionally, its pharmacodynamic potential was tested by comparing its levels in HD

individuals before and after CoQ10 treatment [347] (for role of Coq10 in HD see 1.4.4). Nevertheless, the results have not always been consistent with the initial positive findings and it has been proposed that 8-hydroxy-2'-deoxyguanosine has only limited potential at assessing clinical disease features [351-353]. A more direct approach of disease specific biomarker identification involves the measurement of the mutant huntingtin itself in circulating T cells, B cells and monocytes [354,355]. The results of the study performed by Weiss et al. showed that mHtt leucocyte levels were lower in presymptomatic patients when compared to symptomatic and were also associated with disease burden and caudate atrophy scores. The quantification of mHtt itself as a biomarker obtained from peripheral readings such as blood or other peripheral fluids such as cerebrospinal fluid (CSF) can be very important since it can identify the effect of mutant protein lowering therapeutic strategies (pharmacodynamic) (see 1.4.2) ; can be indicative or associated with peripheral inflammation and immune system deregulations (see 1.2.4) (monitoring – predictive biomarker properties); and it can provide mechanistic links and clues to disease pathology. For example, the latter can be achieved by measuring mutant protein relationships in blood and other peripheral tissues, such as the relationship of soluble and aggregated mutant huntingtin levels [356]. Since it is not yet clear to which extent the interplay between aggregated and soluble mutant proteins forms influences neuronal cellular toxicity, the delineation of the relative amounts of the above two species or conformers can provide additional clues to the disease pathology and constitute on its own a type of biomarker. Other individual molecule and compound findings that have been described in HD blood include increased cytokine levels (mainly IL-6, IL-8 and TNF-a), kynurenine : tryptophan ratios as markers of microglial activation and persistent inflammation [357,358] and lower levels of BDNF, which is associated with striatal neuron survival signals, in HD patients serum and it association with longer CAG repeats [359].

Apart from specific hypothesis driven approaches, as illustrated above, peripheral blood has also been used in non-hypothesis driven efforts to identify peripheral biomarkers. Such non-hypothesis driven methodologies take advantage of state of the art high-throughput techniques such as mass spectrometry and genome wide gene expression profiling platforms to identify additional peripheral molecular biomarkers, among others metabolites and mRNAs [360,361]. To identify changes in mRNA expression levels in HD scientists

have mainly used whole peripheral blood and isolated leukocyte samples of patients as well as animal models [362-366]. Even though the overlap between studies result has not been very good for the mRNA changes identified, the per study findings have demonstrated that blood holds potential in being used as a source for disease progression and therapeutic efficacy. For example, Hu et.al. have demonstrated that dynamic regulator of chromatin plasticity H2A histone family (*H2AFY*) is overexpressed both in the blood and frontal cortex of HD patients, that this association precedes clinical symptoms and that the levels of this marker respond to a neurodegeneration suppressor, thus implying pharmacodynamic potential [364]. The above biomarker also contains many of the properties of an ideal biomarkers as it is validated in HD mouse models, changing according to disease severity and found early in the disease allowing for a wider therapeutic intervention window. Moreover, it has a known mechanism in disease pathology i.e. chromatin regulation (for the role of chromatin regulation in HD see 1.3.2) [367,368]. Apart from gene expression platforms metabolomics and proteomics technologies have been widely used to also identify such biomarkers in HD blood. Notably metabolomics could be a promising approach as deregulation of metabolism is a known HD pathology. Metabolomics and proteomics studies have also provided much information about the changes in patient blood samples [141,369-371]. For instance, proteomic and metabolomic changes in HD blood include altered levels of brain derived metabolites such as 24S-hydroxycholesterol in pre-manifest carriers, associated with metabolically active neuronal cells [372]; the presence of a catabolic phenotype in early HD patients, associated with aliphatic amino acids and changes in fatty acid breakdown markers like glycerol and malonate [373] and identification of neuroinflammation activation proteins in plasma [138]. Together all of the above results show that blood has great potential to provide prodromal, monitoring as well as pharmacodynamic HD biomarkers.

1.5.4 State of Huntington's disease clinical and imaging biomarkers

The identification of clinical biomarkers involves a series of tests and examinations that have the purpose of measuring cognitive and neuropsychiatric condition and motor function performance of mutation carriers. One of the best known of these tests

and the standard method of assessment of clinical disease progression is the Unified Huntington's Disease Rating Scale (UHDRS) [374]. This score comprises of different sub-scores that assess the patient's motor performance with tests such as oculomotor and finger tapping performance, tongue control and force, grip and posture control but also tests assessing their cognitive and functional condition. The advantages of such clinical measures are that they are usually inexpensive, standardized, relative quick and noninvasive. On the other hand, the disadvantages are that they are subjective and display inter and intra-rater variability [375]. However, despite potential disadvantages of some clinical markers there have been organized efforts such as the TRACK-HD and PREDICT-HD consortia that aim to improve the objectivity of these clinical measures as well as providing higher longitudinal change sensitivity and cross sectional differences in presymptomatic patients [376-380]. For example, 36-month observational data from the TRACK-HD study showed that several quantitative motor and cognitive tasks showed increased rates of decline. Furthermore, psychiatric markers such as apathy also exhibited significant increases [379] while the PREDICT-HD study reported that symptoms such as depression, anxiety and obsessive-compulsiveness were present more in HD mutation carriers [381].

A group of HD biomarkers that have produced some of the most significant results during the last years have been collected with the help of (neuro)imaging technologies. These technologies have the ability to describe brain loss changes as well as brain morphological changes in HD mutation carriers. The advantages of imaging techniques are that they non-invasive and that the data produced are in general well reproducible [382]. On the other hand, imaging techniques are usually quite expensive, especially when compared to clinical scoring systems. An additional disadvantage is that these examinations are sometimes difficult to perform due to the structural design of the imagers and the fact they can cause claustrophobia or are even unsuitable for patients that exhibit highly choreatic movements [376]. One of the most used of these technologies is magnetic resonance imaging. MRI studies have demonstrated how cortical degeneration is related to the clinical symptoms and how the cortical changes develop in different patients. One such prominent study by Rosas et al. showed that in HD subjects thinning of the cortical ribbons takes place from the posterior to anterior regions, that these changes take

place early in the disease and could be accountable for the disease heterogeneity [383]. In a subsequent study these patterns of cortical thinning were associated with different motor phenotypes, while functional and cognitive scores were closely related to specific cortical areas [384]. Other promising imaging techniques include functional MRI, PET, diffusor tensor imaging and magnetic resonance spectroscopy. These techniques can offer information about brain area dynamic changes, functional interconnectivity and microstructural degeneration [385,386]. Notably data from the IMAGE-HD study, using quantitative susceptibility mapping and multimodal magnetic resonance imaging, has shown that presymptomatic patient brain areas exhibited a higher iron content and that caudate volume loss could separate presymptomatic HD mutation carriers from controls respectively [387,388]. When it becomes possible to combine complementary information from different imaging types in longitudinal studies, this will greatly improve the potential and the sensitivity of imaging as a biomarker source. However, such an effort is not a small task and it will require greater standardization as well as the integration of different classes of professionals such as doctors, clinical researchers and statisticians [389].

1.5.5 Other molecular and biochemical fluid biomarkers

Aside from blood and the relative advantages it offers compared to other tissues, cerebrospinal fluid (CSF), muscle, saliva and urine have also provided useful and informative "wet" biomarkers for HD. The use of CSF is becoming increasingly popular in monitoring disease progression in other major neurodegenerative disorders such as Alzheimer's disease and Parkinson's disease [369,390-392]. CSF advantages include the fact that it is also relatively accessible and that it interacts with central nervous system tissues and is "open" to brain derived proteins which can potentially mirror the pathological changes in the brain tissue. Wild et al. were able to reliably quantify mHtt at femtomolar sensitivity in mutation carriers' CSF and also showed that higher mHtt load was correlated to reduced cognitive function and increased motor dysfunction in HD patients [393]. In a separate study, Tan et al. used postmortem samples to show that CSF possesses an mHtt and tissue specific property called cell-based aggregation (CBA) seeding, that seeds mutant huntingtin aggregation. This provides a mechanistic explanation for disease pa-

thology and can also act as a potential molecular biomarker for HD [394]. This CSF aggregate seeding propensity could be used to track disease progression (i.e. constituting a monitoring biomarker) by measuring CBA activity in a cell based assay. Saliva is another tissue that has been used to identify disease biomarkers, notably to study neuroendocrine regulation disturbances such as the hypothalamic-pituitary-adrenal (HPA) axis hyperactivity and depression found in HD carriers. Cortisol concentrations obtained from HD mutation carriers' saliva samples were increased, indicative of HPA axis dysfunction and also exhibited a correlation with motor and functional scores [395]. It should be noted however that the association of HPA axis activity with HD disease stages (pre-symptomatic vs (motor)symptomatic) and psychiatric symptoms (depression, suicidality) is not yet clear and further longitudinal studies are required for such saliva measurements to constitute potential biomarkers [396,397]. Additional examples of the use of other accessible fluids and tissues in biomarker identification include the reports of progressively increased cortisol levels in patient urine samples [398], the reduction of muscle contractility markers and a switch from fast to slow twitch muscle types in skeletal muscle of patients [399]. Finally, the function of the autonomic nervous system (ANS) in Huntington's disease has also been investigated to explore how this is affected in patients [400]. When standardized ANS system tests (sympathetic skin responses, blood pressure, heart rate variability etc.) were performed in patients, autonomic symptoms were found to be related to HD functional disability and depression and autonomic deregulation correlated with central nervous system degeneration tests and scores, such as the symbol digit modalities test and the UHDRS motor score [401,402]. These results support the notion that autonomic functional assessment could present a complementary non-invasive measure of disease progression.

Reference List

1. Huntington G: On chorea. George Huntington, M.D. J Neuropsychiatry Clin Neurosci 2003, 15: 109-112.
2. Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE et al.: A polymorphic DNA marker genetically linked to Huntington's disease. Nature 1983, 306: 234-238.
3. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. Cell 1993, 72: 971-983.
4. MacDonald ME, Anderson MA, Gilliam TC, Tranejaerg L, Carpenter NJ, Magenis E et al.: A somatic cell hybrid panel for localizing DNA segments near the Huntington's disease gene. Genomics 1987, 1: 29-34.
5. Gilliam TC, Tanzi RE, Haines JL, Bonner TI, Faryniarz AG, Hobbs WJ et al.: Localization of the Huntington's disease gene to a small segment of chromosome 4 flanked by D4S10 and the telomere. Cell 1987, 50: 565-571.
6. Gusella JF, MacDonald ME: Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. Nat Rev Neurosci 2000, 1: 109-115.
7. Warby SC, Visscher H, Collins JA, Doty CN, Carter C, Butland SL et al.: HTT haplotypes contribute to differences in Huntington disease prevalence between Europe and East Asia. Eur J Hum Genet 2011, 19: 561-566.
8. Walker FO: Huntington's disease. Lancet 2007, 369: 218-228.
9. Evans SJ, Douglas I, Rawlins MD, Wexler NS, Tabrizi SJ, Smeeth L: Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. J Neurol Neurosurg Psychiatry 2013, 84: 1156-1160.
10. Conneally PM: Huntington disease: genetics and epidemiology. Am J Hum Genet 1984, 36: 506-526.
11. Novak MJ, Tabrizi SJ: Huntington's disease. BMJ 2010, 340: c3109.
12. Quarrell OW, Rigby AS, Barron L, Crow Y, Dalton A, Dennis N et al.: Reduced penetrance alleles for Huntington's disease: a multi-centre direct observational study. J Med Genet 2007, 44: e68.
13. Myers RH: Huntington's disease genetics. NeuroRx 2004, 1: 255-262.

14. Williams AJ, Paulson HL: Polyglutamine neurodegeneration: protein misfolding revisited. Trends Neurosci 2008, 31: 521-528.
15. Duyao M, Ambrose C, Myers R, Novelletto A, Persichetti F, Frontali M et al.: Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat Genet 1993, 4: 387-392.
16. Ridley RM, Frith CD, Crow TJ, Conneally PM: Anticipation in Huntington's disease is inherited through the male line but may originate in the female. J Med Genet 1988, 25: 589-595.
17. Semaka A, Collins JA, Hayden MR: Unstable familial transmissions of Huntington disease alleles with 27-35 CAG repeats (intermediate alleles). Am J Med Genet B Neuropsychiatr Genet 2010, 153B: 314-320.
18. Gusella JF, MacDonald ME: Huntington's disease: the case for genetic modifiers. Genome Med 2009, 1: 80.
19. Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, Adam S et al.: The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nat Genet 1993, 4: 398-403.
20. Gusella JF, MacDonald ME, Lee JM: Genetic modifiers of Huntington's disease. Mov Disord 2014, 29: 1359-1365.
21. Reiner A, Dragatsis I, Dietrich P: Genetics and neuropathology of Huntington's disease. Int Rev Neurobiol 2011, 98: 325-372.
22. Kovtun IV, Liu Y, Bjoras M, Klungland A, Wilson SH, McMurray CT: OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. Nature 2007, 447: 447-452.
23. Jonson I, Ougland R, Klungland A, Larsen E: Oxidative stress causes DNA triplet expansion in Huntington's disease mouse embryonic stem cells. Stem Cell Res 2013, 11: 1264-1271.
24. Warby SC, Montpetit A, Hayden AR, Carroll JB, Butland SL, Visscher H et al.: CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup. Am J Hum Genet 2009, 84: 351-366.
25. MacDonald ME, Vonsattel JP, Shrinidhi J, Couropmitree NN, Cupples LA, Bird ED et al.: Evidence for the GluR6 gene associated with younger onset age of Huntington's disease. Neurology 1999, 53: 1330-1332.

26. Arning L, Kraus PH, Valentin S, Saft C, Andrich J, Epplen JT: NR2A and NR2B receptor gene variations modify age at onset in Huntington disease. Neurogenetics 2005, 6: 25-28.
27. Chattopadhyay B, Baksi K, Mukhopadhyay S, Bhattacharyya NP: Modulation of age at onset of Huntington disease patients by variations in TP53 and human caspase activated DNase (hCAD) genes. Neurosci Lett 2005, 374: 81-86.
28. Weydt P, Soyal SM, Gellera C, Didonato S, Weidinger C, Oberkofler H et al.: The gene coding for PGC-1alpha modifies age at onset in Huntington's Disease. Mol Neurodegener 2009, 4: 3.
29. Taherzadeh-Fard E, Saft C, Andrich J, Wieczorek S, Arning L: PGC-1alpha as modifier of onset age in Huntington disease. Mol Neurodegener 2009, 4: 10.
30. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. Cell 2015, 162: 516-526.
31. Huntington G: On chorea. George Huntington, M.D. J Neuropsychiatry Clin Neurosci 2003, 15: 109-112.
32. Feigin A, Kieburtz K, Bordwell K, Como P, Steinberg K, Sotack J et al.: Functional decline in Huntington's disease. Mov Disord 1995, 10: 211-214.
33. Zielonka D, Piotrowska I, Marcinkowski JT, Mielcarek M: Skeletal muscle pathology in Huntington's disease. Front Physiol 2014, 5: 380.
34. Kirkwood SC, Su JL, Conneally P, Foroud T: Progression of symptoms in the early and middle stages of Huntington disease. Arch Neurol 2001, 58: 273-278.
35. Lasker AG, Zee DS: Ocular motor abnormalities in Huntington's disease. Vision Res 1997, 37: 3639-3645.
36. Pierrot-Deseilligny C, Rivaud S, Gaymard B, Muri R, Vermersch AI: Cortical control of saccades. Ann Neurol 1995, 37: 557-567.
37. Patel SS, Jankovic J, Hood AJ, Jeter CB, Sereno AB: Reflexive and volitional saccades: biomarkers of Huntington disease severity and progression. J Neurol Sci 2012, 313: 35-41.
38. Heemskerk AW, Verbist BM, Marinus J, Heijnen B, Sjogren EV, Roos RA: The Huntington's Disease Dysphagia Scale. Mov Disord 2014, 29: 1312-1316.
39. Andrich JE, Wobben M, Klotz P, Goetze O, Saft C: Upper gastrointestinal findings in Huntington's disease: patients suffer but do not complain. J Neural Transm (Vi-

enna) 2009, 116: 1607-1611.

40. Williams S, Heron L, France K, Mulrooney P, Edmondston SJ: Huntington's Disease: Characteristics of Fallers. Physiother Res Int 2014.
41. Busse ME, Wiles CM, Rosser AE: Mobility and falls in people with Huntington's disease. J Neurol Neurosurg Psychiatry 2009, 80: 88-90.
42. Paulsen JS, Ready RE, Hamilton JM, Mega MS, Cummings JL: Neuropsychiatric aspects of Huntington's disease. J Neurol Neurosurg Psychiatry 2001, 71: 310-314.
43. Craufurd D, Thompson JC, Snowden JS: Behavioral changes in Huntington Disease. Neuropsychiatry Neuropsychol Behav Neurol 2001, 14: 219-226.
44. Folstein S, Abbott MH, Chase GA, Jensen BA, Folstein MF: The association of affective disorder with Huntington's disease in a case series and in families. Psychol Med 1983, 13: 537-542.
45. Paulsen JS, Nehl C, Hoth KF, Kanz JE, Benjamin M, Conybeare R et al.: Depression and stages of Huntington's disease. J Neuropsychiatry Clin Neurosci 2005, 17: 496-502.
46. Alonso J, Ferrer M, Romera B, Vilagut G, Angermeyer M, Bernert S et al.: The European Study of the Epidemiology of Mental Disorders (ESEMeD/MHEDEA 2000) project: rationale and methods. Int J Methods Psychiatr Res 2002, 11: 55-67.
47. Robins Wahlin TB: To know or not to know: a review of behaviour and suicidal ideation in preclinical Huntington's disease. Patient Educ Couns 2007, 65: 279-287.
48. van der Meer LB, van DE, Giltay EJ, Tibben A: Do Attachment Style and Emotion Regulation Strategies Indicate Distress in Predictive Testing? J Genet Couns 2015, 24: 862-871.
49. van DE, Kingma EM, van der Mast RC: Psychopathology in verified Huntington's disease gene carriers. J Neuropsychiatry Clin Neurosci 2007, 19: 441-448.
50. Kessler RC, Aguilar-Gaxiola S, Alonso J, Chatterji S, Lee S, Ormel J et al.: The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. Epidemiol Psichiatr Soc 2009, 18: 23-33.
51. Levy ML, Cummings JL, Fairbanks LA, Masterman D, Miller BL, Craig AH et al.: Apathy is not depression. J Neuropsychiatry Clin Neurosci 1998, 10: 314-319.
52. Macdonald V, Halliday G: Pyramidal cell loss in motor cortices in Huntington's disease. Neurobiol Dis 2002, 10: 378-386.

53. Rosas HD, Salat DH, Lee SY, Zaleta AK, Pappu V, Fischl B et al.: Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity. Brain 2008, 131: 1057-1068.
54. Reiner A, Albin RL, Anderson KD, D'Amato CJ, Penney JB, Young AB: Differential loss of striatal projection neurons in Huntington disease. Proc Natl Acad Sci U S A 1988, 85: 5733-5737.
55. Saudou F, Finkbeiner S, Devys D, Greenberg ME: Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. Cell 1998, 95: 55-66.
56. Vonsattel JP, Myers RH, Stevens TJ, Ferrante RJ, Bird ED, Richardson EP, Jr.: Neuropathological classification of Huntington's disease. J Neuropathol Exp Neurol 1985, 44: 559-577.
57. Aylward EH, Brandt J, Codori AM, Mangus RS, Barta PE, Harris GJ: Reduced basal ganglia volume associated with the gene for Huntington's disease in asymptomatic at-risk persons. Neurology 1994, 44: 823-828.
58. Aylward EH, Sparks BF, Field KM, Yallapragada V, Shpritz BD, Rosenblatt A et al.: Onset and rate of striatal atrophy in preclinical Huntington disease. Neurology 2004, 63: 66-72.
59. Glass M, Dragunow M, Faull RL: The pattern of neurodegeneration in Huntington's disease: a comparative study of cannabinoid, dopamine, adenosine and GABA(A) receptor alterations in the human basal ganglia in Huntington's disease. Neuroscience 2000, 97: 505-519.
60. Rosas HD, Koroshetz WJ, Chen YI, Skeuse C, Vangel M, Cudkowicz ME et al.: Evidence for more widespread cerebral pathology in early HD: an MRI-based morphometric analysis. Neurology 2003, 60: 1615-1620.
61. Squitieri F, Cannella M, Simonelli M, Sassone J, Martino T, Venditti E et al.: Distinct brain volume changes correlating with clinical stage, disease progression rate, mutation size, and age at onset prediction as early biomarkers of brain atrophy in Huntington's disease. CNS Neurosci Ther 2009, 15: 1-11.
62. Wild EJ, Henley SM, Hobbs NZ, Frost C, MacManus DG, Barker RA et al.: Rate and acceleration of whole-brain atrophy in premanifest and early Huntington's disease. Mov Disord 2010, 25: 888-895.

63. Rosas HD, Salat DH, Lee SY, Zaleta AK, Pappu V, Fischl B et al.: Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity. Brain 2008, 131: 1057-1068.
64. Sharp AH, Loev SJ, Schilling G, Li SH, Li XJ, Bao J et al.: Widespread expression of Huntington's disease gene (IT15) protein product. Neuron 1995, 14: 1065-1074.
65. DiFiglia M, Sapp E, Chase K, Schwarz C, Meloni A, Young C et al.: Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. Neuron 1995, 14: 1075-1081.
66. Gutekunst CA, Levey AI, Heilman CJ, Whaley WL, Yi H, Nash NR et al.: Identification and localization of huntingtin in brain and human lymphoblastoid cell lines with anti-fusion protein antibodies. Proc Natl Acad Sci U S A 1995, 92: 8710-8714.
67. Gerber HP, Seipel K, Georgiev O, Hofferer M, Hug M, Rusconi S et al.: Transcriptional activation modulated by homopolymeric glutamine and proline stretches. Science 1994, 263: 808-811.
68. Benn CL, Sun T, Sadri-Vakili G, McFarland KN, DiRocco DP, Yohrling GJ et al.: Huntingtin modulates transcription, occupies gene promoters in vivo, and binds directly to DNA in a polyglutamine-dependent manner. J Neurosci 2008, 28: 10720-10733.
69. Kaltenbach LS, Romero E, Becklin RR, Chettier R, Bell R, Phansalkar A et al.: Huntingtin interacting proteins are genetic modifiers of neurodegeneration. PLoS Genet 2007, 3: e82.
70. Zuccato C, Valenza M, Cattaneo E: Molecular mechanisms and potential therapeutical targets in Huntington's disease. Physiol Rev 2010, 90: 905-981.
71. Tartari M, Gissi C, Lo S, V, Zuccato C, Picardi E, Pesole G et al.: Phylogenetic comparison of huntingtin homologues reveals the appearance of a primitive polyQ in sea urchin. Mol Biol Evol 2008, 25: 330-338.
72. Wellington CL, Ellerby LM, Gutekunst CA, Rogers D, Warby S, Graham RK et al.: Caspase cleavage of mutant huntingtin precedes neurodegeneration in Huntington's disease. J Neurosci 2002, 22: 7862-7872.
73. Hermel E, Gafni J, Propp SS, Leavitt BR, Wellington CL, Young JE et al.: Specific caspase interactions and amplification are involved in selective neuronal vulnerability in Huntington's disease. Cell Death Differ 2004, 11: 424-438.

74. Xia J, Lee DH, Taylor J, Vandelft M, Truant R: Huntingtin contains a highly conserved nuclear export signal. Hum Mol Genet 2003, 12: 1393-1403.
75. Bhat KP, Yan S, Wang CE, Li S, Li XJ: Differential ubiquitination and degradation of huntingtin fragments modulated by ubiquitin-protein ligase E3A. Proc Natl Acad Sci U S A 2014, 111: 5706-5711.
76. Atkin G, Paulson H: Ubiquitin pathways in neurodegenerative disease. Front Mol Neurosci 2014, 7: 63.
77. Dong G, Callegari E, Gloeckner CJ, Ueffing M, Wang H: Mass spectrometric identification of novel posttranslational modification sites in Huntingtin. Proteomics 2012, 12: 2060-2064.
78. Havel LS, Wang CE, Wade B, Huang B, Li S, Li XJ: Preferential accumulation of N-terminal mutant huntingtin in the nuclei of striatal neurons is regulated by phosphorylation. Hum Mol Genet 2011, 20: 1424-1437.
79. Wang Y, Lin F, Qin ZH: The role of post-translational modifications of huntingtin in the pathogenesis of Huntington's disease. Neurosci Bull 2010, 26: 153-162.
80. DiFiglia M, Sapp E, Chase K, Schwarz C, Meloni A, Young C et al.: Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. Neuron 1995, 14: 1075-1081.
81. Bhide PG, Day M, Sapp E, Schwarz C, Sheth A, Kim J et al.: Expression of normal and mutant huntingtin in the developing brain. J Neurosci 1996, 16: 5523-5535.
82. Hoogeveen AT, Willemsen R, Meyer N, de Rooij KE, Roos RA, van Ommen GJ et al.: Characterization and localization of the Huntington disease gene product. Hum Mol Genet 1993, 2: 2069-2073.
83. De Rooij KE, Dorsman JC, Smoor MA, Den Dunnen JT, Van Ommen GJ: Subcellular localization of the Huntington's disease gene product in cell lines by immunofluorescence and biochemical subcellular fractionation. Hum Mol Genet 1996, 5: 1093-1099.
84. Qin ZH, Wang Y, Sapp E, Cuiffo B, Wanker E, Hayden MR et al.: Huntingtin bodies sequester vesicle-associated proteins by a polyproline-dependent interaction. J Neurosci 2004, 24: 269-281.
85. Li SH, Li XJ: Huntingtin-protein interactions and the pathogenesis of Huntington's disease. Trends Genet 2004, 20: 146-154.

86. Gutekunst CA, Levey AI, Heilman CJ, Whaley WL, Yi H, Nash NR et al.: Identification and localization of huntingtin in brain and human lymphoblastoid cell lines with anti-fusion protein antibodies. Proc Natl Acad Sci U S A 1995, 92: 8710-8714.
87. Sapp E, Schwarz C, Chase K, Bhide PG, Young AB, Penney J et al.: Huntingtin localization in brains of normal and Huntington's disease patients. Ann Neurol 1997, 42: 604-612.
88. Hoogeveen AT, Willemsen R, Meyer N, de Rooij KE, Roos RA, van Ommen GJ et al.: Characterization and localization of the Huntington disease gene product. Hum Mol Genet 1993, 2: 2069-2073.
89. Nasir J, Floresco SB, O'Kusky JR, Diewert VM, Richman JM, Zeisler J et al.: Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. Cell 1995, 81: 811-823.
90. Reiner A, Dragatsis I, Zeitlin S, Goldowitz D: Wild-type huntingtin plays a role in brain development and neuronal survival. Mol Neurobiol 2003, 28: 259-276.
91. Rigamonti D, Bauer JH, De-Fraja C, Conti L, Sipione S, Sciorati C et al.: Wild-type huntingtin protects from apoptosis upstream of caspase-3. J Neurosci 2000, 20: 3705-3713.
92. Leavitt BR, van Raamsdonk JM, Shehadeh J, Fernandes H, Murphy Z, Graham RK et al.: Wild-type huntingtin protects neurons from excitotoxicity. J Neurochem 2006, 96: 1121-1129.
93. Zuccato C, Ciammola A, Rigamonti D, Leavitt BR, Goffredo D, Conti L et al.: Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. Science 2001, 293: 493-498.
94. Gauthier LR, Charrin BC, Borrell-Pages M, Dompierre JP, Rangone H, Cordelieres FP et al.: Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. Cell 2004, 118: 127-138.
95. Her LS, Goldstein LS: Enhanced sensitivity of striatal neurons to axonal transport defects induced by mutant huntingtin. J Neurosci 2008, 28: 13662-13672.
96. Gunawardena S, Her LS, Brusch RG, Laymon RA, Niesman IR, Gordesky-Gold B et al.: Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila. Neuron 2003, 40: 25-40.
97. Parsons MP, Kang R, Buren C, Dau A, Southwell AL, Doty CN et al.: Bidirectional

control of postsynaptic density-95 (PSD-95) clustering by Huntingtin. J Biol Chem 2014, 289: 3518-3528.

98. Twelvetrees AE, Yuen EY, Arancibia-Carcamo IL, MacAskill AF, Rostaing P, Lumb MJ et al.: Delivery of GABAARs to synapses is mediated by HAP1-KIF5 and disrupted by mutant huntingtin. Neuron 2010, 65: 53-65.
99. DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP et al.: Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science 1997, 277: 1990-1993.
100. Mangiarini L, Sathasivam K, Seller M, Cozens B, Harper A, Hetherington C et al.: Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. Cell 1996, 87: 493-506.
101. Davies SW, Turmaine M, Cozens BA, DiFiglia M, Sharp AH, Ross CA et al.: Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. Cell 1997, 90: 537-548.
102. Steffan JS, Kazantsev A, Spasic-Boskovic O, Greenwald M, Zhu YZ, Gohler H et al.: The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proc Natl Acad Sci U S A 2000, 97: 6763-6768.
103. Donaldson KM, Li W, Ching KA, Batalov S, Tsai CC, Joazeiro CA: Ubiquitin-mediated sequestration of normal cellular proteins into polyglutamine aggregates. Proc Natl Acad Sci U S A 2003, 100: 8892-8897.
104. Hickey MA, Chesselet MF: Apoptosis in Huntington's disease. Prog Neuropsychopharmacol Biol Psychiatry 2003, 27: 255-265.
105. Damiano M, Galvan L, Deglon N, Brouillet E: Mitochondria in Huntington's disease. Biochim Biophys Acta 2010, 1802: 52-61.
106. DiProspero NA, Chen EY, Charles V, Plomann M, Kordower JH, Tagle DA: Early changes in Huntington's disease patient brains involve alterations in cytoskeletal and synaptic elements. J Neurocytol 2004, 33: 517-533.
107. Legleiter J, Mitchell E, Lotz GP, Sapp E, Ng C, DiFiglia M et al.: Mutant huntingtin fragments form oligomers in a polyglutamine length-dependent manner in vitro and in vivo. J Biol Chem 2010, 285: 14777-14790.
108. Arrasate M, Finkbeiner S: Protein aggregates in Huntington's disease. Exp Neurol 2012, 238: 1-11.

109. Miller J, Arrasate M, Shaby BA, Mitra S, Masliah E, Finkbeiner S: Quantitative relationships between huntingtin levels, polyglutamine length, inclusion body formation, and neuronal death provide novel insight into huntington's disease molecular pathogenesis. J Neurosci 2010, 30: 10541-10550.

110. Colby DW, Cassady JP, Lin GC, Ingram VM, Wittrup KD: Stochastic kinetics of intracellular huntingtin aggregate formation. Nat Chem Biol 2006, 2: 319-323.

111. Martindale D, Hackam A, Wieczorek A, Ellerby L, Wellington C, McCutcheon K et al.: Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. Nat Genet 1998, 18: 150-154.

112. Hermel E, Gafni J, Propp SS, Leavitt BR, Wellington CL, Young JE et al.: Specific caspase interactions and amplification are involved in selective neuronal vulnerability in Huntington's disease. Cell Death Differ 2004, 11: 424-438.

113. Bizat N, Hermel JM, Boyer F, Jacquard C, Creminon C, Ouary S et al.: Calpain is a major cell death effector in selective striatal degeneration induced in vivo by 3-nitropropionate: implications for Huntington's disease. J Neurosci 2003, 23: 5020-5030.

114. Tam S, Spiess C, Auyeung W, Joachimiak L, Chen B, Poirier MA et al.: The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation. Nat Struct Mol Biol 2009, 16: 1279-1285.

115. Kitamura A, Kubota H, Pack CG, Matsumoto G, Hirayama S, Takahashi Y et al.: Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. Nat Cell Biol 2006, 8: 1163-1170.

116. Aiken CT, Steffan JS, Guerrero CM, Khashwji H, Lukacsovich T, Simmons D et al.: Phosphorylation of threonine 3: implications for Huntingtin aggregation and neurotoxicity. J Biol Chem 2009, 284: 29427-29436.

117. Wang Y, Lin F, Qin ZH: The role of post-translational modifications of huntingtin in the pathogenesis of Huntington's disease. Neurosci Bull 2010, 26: 153-162.

118. Bates G: Huntingtin aggregation and toxicity in Huntington's disease. Lancet 2003, 361: 1642-1644.

119. Tagawa K, Hoshino M, Okuda T, Ueda H, Hayashi H, Engemann S et al.: Distinct aggregation and cell death patterns among different types of primary neurons induced by mutant huntingtin protein. J Neurochem 2004, 89: 974-987.

120. Jansen AH, van HM, Op den Kelder IC, Meier RT, de Ruiter AA, Schut MH et al.: Frequency of nuclear mutant huntingtin inclusion formation in neurons and glia is cell-type-specific. Glia 2016.

121. de la Monte SM, Vonsattel JP, Richardson EP, Jr.: Morphometric demonstration of atrophic changes in the cerebral cortex, white matter, and neostriatum in Huntington's disease. J Neuropathol Exp Neurol 1988, 47: 516-525.

122. Reading SA, Yassa MA, Bakker A, Dziorny AC, Gourley LM, Yallapragada V et al.: Regional white matter change in pre-symptomatic Huntington's disease: a diffusion tensor imaging study. Psychiatry Res 2005, 140: 55-62.

123. Kuppenbender KD, Standaert DG, Feuerstein TJ, Penney JB, Jr., Young AB, Landwehrmeyer GB: Expression of NMDA receptor subunit mRNAs in neurochemically identified projection and interneurons in the human striatum. J Comp Neurol 2000, 419: 407-421.

124. Joyce JN, Lexow N, Bird E, Winokur A: Organization of dopamine D1 and D2 receptors in human striatum: receptor autoradiographic studies in Huntington's disease and schizophrenia. Synapse 1988, 2: 546-557.

125. Tippett LJ, Waldvogel HJ, Thomas SJ, Hogg VM, van Roon-Mom W, Synek BJ et al.: Striosomes and mood dysfunction in Huntington's disease. Brain 2007, 130: 206-221.

126. Glass M, Dragunow M, Faull RL: The pattern of neurodegeneration in Huntington's disease: a comparative study of cannabinoid, dopamine, adenosine and GABA(A) receptor alterations in the human basal ganglia in Huntington's disease. Neuroscience 2000, 97: 505-519.

127. Cicchetti F, Prensa L, Wu Y, Parent A: Chemical anatomy of striatal interneurons in normal individuals and in patients with Huntington's disease. Brain Res Brain Res Rev 2000, 34: 80-101.

128. Waldvogel HJ, Kim EH, Tippett LJ, Vonsattel JP, Faull RL: The Neuropathology of Huntington's Disease. Curr Top Behav Neurosci 2015, 22: 33-80.

129. Selkoe DJ, Salazar FJ, Abraham C, Kosik KS: Huntington's disease: changes in striatal proteins reflect astrocytic gliosis. Brain Res 1982, 245: 117-125.

130. Reiner A, Del MN, Deng YP, Meade CA, Sun Z, Goldowitz D: R6/2 neurons with intranuclear inclusions survive for prolonged periods in the brains of chimeric

mice. J Comp Neurol 2007, 505: 603-629.

131. Yu ZX, Li SH, Evans J, Pillarisetti A, Li H, Li XJ: Mutant huntingtin causes context-dependent neurodegeneration in mice with Huntington's disease. J Neurosci 2003, 23: 2193-2202.

132. Pavese N, Gerhard A, Tai YF, Ho AK, Turkheimer F, Barker RA et al.: Microglial activation correlates with severity in Huntington disease: a clinical and PET study. Neurology 2006, 66: 1638-1643.

133. Politis M, Su P, Piccini P: Imaging of microglia in patients with neurodegenerative disorders. Front Pharmacol 2012, 3: 96.

134. van der Burg JM, Bjorkqvist M, Brundin P: Beyond the brain: widespread pathology in Huntington's disease. Lancet Neurol 2009, 8: 765-774.

135. Van Raamsdonk JM, Murphy Z, Selva DM, Hamidizadeh R, Pearson J, Petersen A et al.: Testicular degeneration in Huntington disease. Neurobiol Dis 2007, 26: 512-520.

136. Hamilton JM, Wolfson T, Peavy GM, Jacobson MW, Corey-Bloom J: Rate and correlates of weight change in Huntington's disease. J Neurol Neurosurg Psychiatry 2004, 75: 209-212.

137. Robbins AO, Ho AK, Barker RA: Weight changes in Huntington's disease. Eur J Neurol 2006, 13: e7.

138. Dalrymple A, Wild EJ, Joubert R, Sathasivam K, Bjorkqvist M, Petersen A et al.: Proteomic profiling of plasma in Huntington's disease reveals neuroinflammatory activation and biomarker candidates. J Proteome Res 2007, 6: 2833-2840.

139. Bjorkqvist M, Wild EJ, Thiele J, Silvestroni A, Andre R, Lahiri N et al.: A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington's disease. J Exp Med 2008, 205: 1869-1877.

140. Sathasivam K, Hobbs C, Turmaine M, Mangiarini L, Mahal A, Bertaux F et al.: Formation of polyglutamine inclusions in non-CNS tissue. Hum Mol Genet 1999, 8: 813-822.

141. Bjorkqvist M, Wild EJ, Thiele J, Silvestroni A, Andre R, Lahiri N et al.: A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington's disease. J Exp Med 2008, 205: 1869-1877.

142. Goodman AO, Murgatroyd PR, Medina-Gomez G, Wood NI, Finer N, Vidal-

Puig AJ et al.: The metabolic profile of early Huntington's disease--a combined human and transgenic mouse study. Exp Neurol 2008, 210: 691-698.

143. van der Burg JM, Bacos K, Wood NI, Lindqvist A, Wierup N, Woodman B et al.: Increased metabolism in the R6/2 mouse model of Huntington's disease. Neurobiol Dis 2008, 29: 41-51.

144. Strand AD, Aragaki AK, Shaw D, Bird T, Holton J, Turner C et al.: Gene expression in Huntington's disease skeletal muscle: a potential biomarker. Hum Mol Genet 2005, 14: 1863-1876.

145. Luthi-Carter R, Hanson SA, Strand AD, Bergstrom DA, Chun W, Peters NL et al.: Dysregulation of gene expression in the R6/2 model of polyglutamine disease: parallel changes in muscle and brain. Hum Mol Genet 2002, 11: 1911-1926.

146. Bjorkqvist M, Fex M, Renstrom E, Wierup N, Petersen A, Gil J et al.: The R6/2 transgenic mouse model of Huntington's disease develops diabetes due to deficient beta-cell mass and exocytosis. Hum Mol Genet 2005, 14: 565-574.

147. Andreassen OA, Dedeoglu A, Stanojevic V, Hughes DB, Browne SE, Leech CA et al.: Huntington's disease of the endocrine pancreas: insulin deficiency and diabetes mellitus due to impaired insulin gene expression. Neurobiol Dis 2002, 11: 410-424.

148. Goodman AO, Barker RA: Body composition in premanifest Huntington's disease reveals lower bone density compared to controls. PLoS Curr 2011, 3: RRN1214.

149. Silvestroni A, Faull RL, Strand AD, Moller T: Distinct neuroinflammatory profile in post-mortem human Huntington's disease. Neuroreport 2009, 20: 1098-1103.

150. Ellrichmann G, Reick C, Saft C, Linker RA: The role of the immune system in Huntington's disease. Clin Dev Immunol 2013, 2013: 541259.

151. Crotti A, Glass CK: The choreography of neuroinflammation in Huntington's disease. Trends Immunol 2015, 36: 364-373.

152. Kwan W, Trager U, Davalos D, Chou A, Bouchard J, Andre R et al.: Mutant huntingtin impairs immune cell migration in Huntington disease. J Clin Invest 2012, 122: 4737-4747.

153. Langen RC, Schols AM, Kelders MC, Wouters EF, Janssen-Heininger YM: Inflammatory cytokines inhibit myogenic differentiation through activation of nuclear factor-kappaB. FASEB J 2001, 15: 1169-1180.

154. Meyer SU, Krebs S, Thirion C, Blum H, Krause S, Pfaffl MW: Tumor Necrosis Factor Alpha and Insulin-Like Growth Factor 1 Induced Modifications of the Gene Expression Kinetics of Differentiating Skeletal Muscle Cells. PLoS One 2015, 10: e0139520.

155. Li H, Li SH, Johnston H, Shelbourne PF, Li XJ: Amino-terminal fragments of mutant huntingtin show selective accumulation in striatal neurons and synaptic toxicity. Nat Genet 2000, 25: 385-389.

156. Li SH, Lam S, Cheng AL, Li XJ: Intranuclear huntingtin increases the expression of caspase-1 and induces apoptosis. Hum Mol Genet 2000, 9: 2859-2867.

157. Wong BK, Ehrnhoefer DE, Graham RK, Martin DD, Ladha S, Uribe V et al.: Partial rescue of some features of Huntington Disease in the genetic absence of caspase-6 in YAC128 mice. Neurobiol Dis 2015, 76: 24-36.

158. Cusack CL, Swahari V, Hampton HW, Michael RJ, Deshmukh M: Distinct pathways mediate axon degeneration during apoptosis and axon-specific pruning. Nat Commun 2013, 4: 1876.

159. Fonseca R, Vabulas RM, Hartl FU, Bonhoeffer T, Nagerl UV: A balance of protein synthesis and proteasome-dependent degradation determines the maintenance of LTP. Neuron 2006, 52: 239-245.

160. Wyttenbach A, Carmichael J, Swartz J, Furlong RA, Narain Y, Rankin J et al.: Effects of heat shock, heat shock protein 40 (HDJ-2), and proteasome inhibition on protein aggregation in cellular models of Huntington's disease. Proc Natl Acad Sci U S A 2000, 97: 2898-2903.

161. Venkatraman P, Wetzel R, Tanaka M, Nukina N, Goldberg AL: Eukaryotic proteasomes cannot digest polyglutamine sequences and release them during degradation of polyglutamine-containing proteins. Mol Cell 2004, 14: 95-104.

162. Holmberg CI, Staniszewski KE, Mensah KN, Matouschek A, Morimoto RI: Inefficient degradation of truncated polyglutamine proteins by the proteasome. EMBO J 2004, 23: 4307-4318.

163. Rousseau E, Kojima R, Hoffner G, Djian P, Bertolotti A: Misfolding of proteins with a polyglutamine expansion is facilitated by proteasomal chaperones. J Biol Chem 2009, 284: 1917-1929.

164. Diaz-Hernandez M, Valera AG, Moran MA, Gomez-Ramos P, Alvarez-Caste-

lao B, Castano JG et al.: Inhibition of 26S proteasome activity by huntingtin filaments but not inclusion bodies isolated from mouse and human brain. J Neurochem 2006, 98: 1585-1596.

165. Schipper-Krom S, Juenemann K, Reits EA: The Ubiquitin-Proteasome System in Huntington's Disease: Are Proteasomes Impaired, Initiators of Disease, or Coming to the Rescue? Biochem Res Int 2012, 2012: 837015.

166. Harjes P, Wanker EE: The hunt for huntingtin function: interaction partners tell many different stories. Trends Biochem Sci 2003, 28: 425-433.

167. Marti E: RNA toxicity induced by expanded CAG repeats in Huntington's disease. Brain Pathol 2016.

168. Dunah AW, Jeong H, Griffin A, Kim YM, Standaert DG, Hersch SM et al.: Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. Science 2002, 296: 2238-2243.

169. Steffan JS, Kazantsev A, Spasic-Boskovic O, Greenwald M, Zhu YZ, Gohler H et al.: The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proc Natl Acad Sci U S A 2000, 97: 6763-6768.

170. Obrietan K, Hoyt KR: CRE-mediated transcription is increased in Huntington's disease transgenic mice. J Neurosci 2004, 24: 791-796.

171. Dunah AW, Jeong H, Griffin A, Kim YM, Standaert DG, Hersch SM et al.: Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. Science 2002, 296: 2238-2243.

172. Moumne L, Betuing S, Caboche J: Multiple Aspects of Gene Dysregulation in Huntington's Disease. Front Neurol 2013, 4: 127.

173. Luthi-Carter R, Strand A, Peters NL, Solano SM, Hollingsworth ZR, Menon AS et al.: Decreased expression of striatal signaling genes in a mouse model of Huntington's disease. Hum Mol Genet 2000, 9: 1259-1271.

174. Labadorf A, Hoss AG, Lagomarsino V, Latourelle JC, Hadzi TC, Bregu J et al.: RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. PLoS One 2015, 10: e0143563.

175. Hodges A, Strand AD, Aragaki AK, Kuhn A, Sengstag T, Hughes G et al.: Regional and cellular gene expression changes in human Huntington's disease brain.

Hum Mol Genet 2006, 15: 965-977.

176. Zuccato C, Belyaev N, Conforti P, Ooi L, Tartari M, Papadimou E et al.: Widespread disruption of repressor element-1 silencing transcription factor/neuron-restrictive silencer factor occupancy at its target genes in Huntington's disease. J Neurosci 2007, 27: 6972-6983.

177. Chiang MC, Chen CM, Lee MR, Chen HW, Chen HM, Wu YS et al.: Modulation of energy deficiency in Huntington's disease via activation of the peroxisome proliferator-activated receptor gamma. Hum Mol Genet 2010, 19: 4043-4058.

178. Yamanaka T, Miyazaki H, Oyama F, Kurosawa M, Washizu C, Doi H et al.: Mutant Huntingtin reduces HSP70 expression through the sequestration of NF-Y transcription factor. EMBO J 2008, 27: 827-839.

179. Gapp K, Woldemichael BT, Bohacek J, Mansuy IM: Epigenetic regulation in neurodevelopment and neurodegenerative diseases. Neuroscience 2014, 264: 99-111.

180. Lee J, Hwang YJ, Kim KY, Kowall NW, Ryu H: Epigenetic mechanisms of neurodegeneration in Huntington's disease. Neurotherapeutics 2013, 10: 664-676.

181. Steffan JS, Bodai L, Pallos J, Poelman M, McCampbell A, Apostol BL et al.: Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. Nature 2001, 413: 739-743.

182. Ferrante RJ, Kubilus JK, Lee J, Ryu H, Beesen A, Zucker B et al.: Histone deacetylase inhibition by sodium butyrate chemotherapy ameliorates the neurodegenerative phenotype in Huntington's disease mice. J Neurosci 2003, 23: 9418-9427.

183. Ryu H, Lee J, Hagerty SW, Soh BY, McAlpin SE, Cormier KA et al.: ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease. Proc Natl Acad Sci U S A 2006, 103: 19176-19181.

184. Karpuj MV, Becher MW, Springer JE, Chabas D, Youssef S, Pedotti R et al.: Prolonged survival and decreased abnormal movements in transgenic model of Huntington disease, with administration of the transglutaminase inhibitor cystamine. Nat Med 2002, 8: 143-149.

185. Marti E, Pantano L, Banez-Coronel M, Llorens F, Minones-Moyano E, Porta S et al.: A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. Nucleic Acids Res 2010, 38: 7219-

7235.

186. Packer AN, Xing Y, Harper SQ, Jones L, Davidson BL: The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. J Neurosci 2008, 28: 14341-14346.

187. Seredenina T, Luthi-Carter R: What have we learned from gene expression profiles in Huntington's disease? Neurobiol Dis 2012, 45: 83-98.

188. Chan DC: Mitochondria: dynamic organelles in disease, aging, and development. Cell 2006, 125: 1241-1252.

189. Sawa A, Wiegand GW, Cooper J, Margolis RL, Sharp AH, Lawler JF, Jr. et al.: Increased apoptosis of Huntington disease lymphoblasts associated with repeat length-dependent mitochondrial depolarization. Nat Med 1999, 5: 1194-1198.

190. Panov AV, Gutekunst CA, Leavitt BR, Hayden MR, Burke JR, Strittmatter WJ et al.: Early mitochondrial calcium defects in Huntington's disease are a direct effect of polyglutamines. Nat Neurosci 2002, 5: 731-736.

191. Gellerich FN, Gizatullina Z, Nguyen HP, Trumbeckaite S, Vielhaber S, Seppet E et al.: Impaired regulation of brain mitochondria by extramitochondrial Ca2+ in transgenic Huntington disease rats. J Biol Chem 2008, 283: 30715-30724.

192. Chang DT, Rintoul GL, Pandipati S, Reynolds IJ: Mutant huntingtin aggregates impair mitochondrial movement and trafficking in cortical neurons. Neurobiol Dis 2006, 22: 388-400.

193. Wyttenbach A, Sauvageot O, Carmichael J, Diaz-Latoud C, Arrigo AP, Rubinsztein DC: Heat shock protein 27 prevents cellular polyglutamine toxicity and suppresses the increase of reactive oxygen species caused by huntingtin. Hum Mol Genet 2002, 11: 1137-1151.

194. Block ML, Zecca L, Hong JS: Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. Nat Rev Neurosci 2007, 8: 57-69.

195. Pavese N, Gerhard A, Tai YF, Ho AK, Turkheimer F, Barker RA et al.: Microglial activation correlates with severity in Huntington disease: a clinical and PET study. Neurology 2006, 66: 1638-1643.

196. Gu M, Gash MT, Mann VM, Javoy-Agid F, Cooper JM, Schapira AH: Mitochondrial defect in Huntington's disease caudate nucleus. Ann Neurol 1996, 39: 385-389.

197. Kuhl DE, Phelps ME, Markham CH, Metter EJ, Riege WH, Winter J: Cerebral metabolism and atrophy in Huntington's disease determined by 18FDG and computed tomographic scan. Ann Neurol 1982, 12: 425-434.

198. Antonini A, Leenders KL, Spiegel R, Meier D, Vontobel P, Weigell-Weber M et al.: Striatal glucose metabolism and dopamine D2 receptor binding in asymptomatic gene carriers and patients with Huntington's disease. Brain 1996, 119 (Pt 6): 2085-2095.

199. Cui L, Jeong H, Borovecki F, Parkhurst CN, Tanese N, Krainc D: Transcriptional repression of PGC-1alpha by mutant huntingtin leads to mitochondrial dysfunction and neurodegeneration. Cell 2006, 127: 59-69.

200. Jenkins BG, Koroshetz WJ, Beal MF, Rosen BR: Evidence for impairment of energy metabolism in vivo in Huntington's disease using localized 1H NMR spectroscopy. Neurology 1993, 43: 2689-2695.

201. Costa V, Scorrano L: Shaping the role of mitochondria in the pathogenesis of Huntington's disease. EMBO J 2012, 31: 1853-1864.

202. Zeron MM, Chen N, Moshaver A, Lee AT, Wellington CL, Hayden MR et al.: Mutant huntingtin enhances excitotoxic cell death. Mol Cell Neurosci 2001, 17: 41-53.

203. Jakel RJ, Maragos WF: Neuronal cell death in Huntington's disease: a potential role for dopamine. Trends Neurosci 2000, 23: 239-245.

204. Polyzos AA, McMurray CT: The chicken or the egg: mitochondrial dysfunction as a cause or consequence of toxicity in Huntington's disease. Mech Ageing Dev 2016.

205. Cummings DM, Milnerwood AJ, Dallerac GM, Waights V, Brown JY, Vatsavayai SC et al.: Aberrant cortical synaptic plasticity and dopaminergic dysfunction in a mouse model of Huntington's disease. Hum Mol Genet 2006, 15: 2856-2868.

206. Bibb JA, Yan Z, Svenningsson P, Snyder GL, Pieribone VA, Horiuchi A et al.: Severe deficiencies in dopamine signaling in presymptomatic Huntington's disease mice. Proc Natl Acad Sci U S A 2000, 97: 6809-6814.

207. Hassel B, Tessler S, Faull RL, Emson PC: Glutamate uptake is reduced in prefrontal cortex in Huntington's disease. Neurochem Res 2008, 33: 232-237.

208. Estrada-Sanchez AM, Rebec GV: Corticostriatal dysfunction and glutamate

transporter 1 (GLT1) in Huntington's disease: interactions between neurons and astrocytes. Basal Ganglia 2012, 2: 57-66.

209. Miller BR, Dorner JL, Bunner KD, Gaither TW, Klein EL, Barton SJ et al.: Up-regulation of GLT1 reverses the deficit in cortically evoked striatal ascorbate efflux in the R6/2 mouse model of Huntington's disease. J Neurochem 2012, 121: 629-638.

210. Lievens JC, Rival T, Iche M, Chneiweiss H, Birman S: Expanded polyglutamine peptides disrupt EGF receptor signaling and glutamate transporter expression in Drosophila. Hum Mol Genet 2005, 14: 713-724.

211. Behrens PF, Franz P, Woodman B, Lindenberg KS, Landwehrmeyer GB: Impaired glutamate transport and glutamate-glutamine cycling: downstream effects of the Huntington mutation. Brain 2002, 125: 1908-1922.

212. Estrada-Sanchez AM, Montiel T, Massieu L: Glycolysis inhibition decreases the levels of glutamate transporters and enhances glutamate neurotoxicity in the R6/2 Huntington's disease mice. Neurochem Res 2010, 35: 1156-1163.

213. Camacho A, Montiel T, Massieu L: The anion channel blocker, 4,4'-dinitrostilbene-2,2'-disulfonic acid prevents neuronal death and excitatory amino acid release during glycolysis inhibition in the hippocampus in vivo. Neuroscience 2006, 142: 1005-1017.

214. Zhang H, Li Q, Graham RK, Slow E, Hayden MR, Bezprozvanny I: Full length mutant huntingtin is required for altered Ca2+ signaling and apoptosis of striatal neurons in the YAC mouse model of Huntington's disease. Neurobiol Dis 2008, 31: 80-88.

215. Liang ZQ, Wang XX, Wang Y, Chuang DM, DiFiglia M, Chase TN et al.: Susceptibility of striatal neurons to excitotoxic injury correlates with basal levels of Bcl-2 and the induction of P53 and c-Myc immunoreactivity. Neurobiol Dis 2005, 20: 562-573.

216. Li Y, Yui D, Luikart BW, McKay RM, Li Y, Rubenstein JL et al.: Conditional ablation of brain-derived neurotrophic factor-TrkB signaling impairs striatal neuron development. Proc Natl Acad Sci U S A 2012, 109: 15491-15496.

217. Van LK, Casteels C, Dhollander I, Goffin K, Grachev I, Bormans G et al.: Widespread decrease of type 1 cannabinoid receptor availability in Huntington dis-

ease in vivo. J Nucl Med 2010, 51: 1413-1417.

218. Horne EA, Coy J, Swinney K, Fung S, Cherry AE, Marrs WR et al.: Down-regulation of cannabinoid receptor 1 from neuropeptide Y interneurons in the basal ganglia of patients with Huntington's disease and mouse models. Eur J Neurosci 2013, 37: 429-440.

219. Sepers MD, Raymond LA: Mechanisms of synaptic dysfunction and excitotoxicity in Huntington's disease. Drug Discov Today 2014, 19: 990-996.

220. Martin DD, Ladha S, Ehrnhoefer DE, Hayden MR: Autophagy in Huntington disease and huntingtin in autophagy. Trends Neurosci 2015, 38: 26-35.

221. Rubinsztein DC, Codogno P, Levine B: Autophagy modulation as a potential therapeutic target for diverse diseases. Nat Rev Drug Discov 2012, 11: 709-730.

222. Wild EJ, Tabrizi SJ: Targets for future clinical trials in Huntington's disease: what's in the pipeline? Mov Disord 2014, 29: 1434-1445.

223. Levine B, Packer M, Codogno P: Development of autophagy inducers in clinical medicine. J Clin Invest 2015, 125: 14-24.

224. Cortes CJ, La Spada AR: The many faces of autophagy dysfunction in Huntington's disease: from mechanism to therapy. Drug Discov Today 2014, 19: 963-971.

225. Wong E, Cuervo AM: Autophagy gone awry in neurodegenerative diseases. Nat Neurosci 2010, 13: 805-811.

226. Nixon RA: The role of autophagy in neurodegenerative disease. Nat Med 2013, 19: 983-997.

227. Martinez-Vicente M, Talloczy Z, Wong E, Tang G, Koga H, Kaushik S et al.: Cargo recognition failure is responsible for inefficient autophagy in Huntington's disease. Nat Neurosci 2010, 13: 567-576.

228. Sarkar S, Ravikumar B, Floto RA, Rubinsztein DC: Rapamycin and mTOR-independent autophagy inducers ameliorate toxicity of polyglutamine-expanded huntingtin and related proteinopathies. Cell Death Differ 2009, 16: 46-56.

229. Kegel KB, Kim M, Sapp E, McIntyre C, Castano JG, Aronin N et al.: Huntingtin expression stimulates endosomal-lysosomal activity, endosome tubulation, and autophagy. J Neurosci 2000, 20: 7268-7278.

230. Atwal RS, Truant R: A stress sensitive ER membrane-association domain in Huntingtin protein defines a potential role for Huntingtin in the regulation of auto-

phagy. Autophagy 2008, 4: 91-93.

231. Lee H, Noh JY, Oh Y, Kim Y, Chang JW, Chung CW et al.: IRE1 plays an essential role in ER stress-mediated aggregation of mutant huntingtin via the inhibition of autophagy flux. Hum Mol Genet 2012, 21: 101-114.

232. Reijonen S, Putkonen N, Norremolle A, Lindholm D, Korhonen L: Inhibition of endoplasmic reticulum stress counteracts neuronal cell death and protein aggregation caused by N-terminal mutant huntingtin proteins. Exp Cell Res 2008, 314: 950-960.

233. Vidal R, Caballero B, Couve A, Hetz C: Converging pathways in the occurrence of endoplasmic reticulum (ER) stress in Huntington's disease. Curr Mol Med 2011, 11: 1-12.

234. Cuervo AM: Chaperone-mediated autophagy: selectivity pays off. Trends Endocrinol Metab 2010, 21: 142-150.

235. Koga H, Martinez-Vicente M, Arias E, Kaushik S, Sulzer D, Cuervo AM: Constitutive upregulation of chaperone-mediated autophagy in Huntington's disease. J Neurosci 2011, 31: 18492-18505.

236. Qi L, Zhang XD, Wu JC, Lin F, Wang J, DiFiglia M et al.: The role of chaperone-mediated autophagy in huntingtin degradation. PLoS One 2012, 7: e46834.

237. Kaushik S, Singh R, Cuervo AM: Autophagic pathways and metabolic stress. Diabetes Obes Metab 2010, 12 Suppl 2: 4-14.

238. Ahlberg J, Glaumann H: Uptake--microautophagy--and degradation of exogenous proteins by isolated rat liver lysosomes. Effects of pH, ATP, and inhibitors of proteolysis. Exp Mol Pathol 1985, 42: 78-88.

239. Ferrante RJ, Kubilus JK, Lee J, Ryu H, Beesen A, Zucker B et al.: Histone deacetylase inhibition by sodium butyrate chemotherapy ameliorates the neurodegenerative phenotype in Huntington's disease mice. J Neurosci 2003, 23: 9418-9427.

240. Gardian G, Browne SE, Choi DK, Klivenyi P, Gregorio J, Kubilus JK et al.: Neuroprotective effects of phenylbutyrate in the N171-82Q transgenic mouse model of Huntington's disease. J Biol Chem 2005, 280: 556-563.

241. Xu X, Kozikowski AP, Pozzo-Miller L: A selective histone deacetylase-6 inhibitor improves BDNF trafficking in hippocampal neurons from Mecp2 knockout mice: implications for Rett syndrome. Front Cell Neurosci 2014, 8: 68.

242. Thomas EA, Coppola G, Desplats PA, Tang B, Soragni E, Burnett R et al.: The HDAC inhibitor 4b ameliorates the disease phenotype and transcriptional abnormalities in Huntington's disease transgenic mice. Proc Natl Acad Sci U S A 2008, 105: 15564-15569.

243. Gines S, Seong IS, Fossale E, Ivanova E, Trettel F, Gusella JF et al.: Specific progressive cAMP reduction implicates energy deficit in presymptomatic Huntington's disease knock-in mice. Hum Mol Genet 2003, 12: 497-508.

244. Giampa C, Patassini S, Borreca A, Laurenti D, Marullo F, Bernardi G et al.: Phosphodiesterase 10 inhibition reduces striatal excitotoxicity in the quinolinic acid model of Huntington's disease. Neurobiol Dis 2009, 34: 450-456.

245. Kumar A, Kumar SS, Kumar V, Kumar D, Agarwal S, Rana MK: Huntington's disease: an update of therapeutic strategies. Gene 2015, 556: 91-97.

246. Stack EC, Del Signore SJ, Luthi-Carter R, Soh BY, Goldstein DR, Matson S et al.: Modulation of nucleosome dynamics in Huntington's disease. Hum Mol Genet 2007, 16: 1164-1175.

247. Ryu H, Lee J, Hagerty SW, Soh BY, McAlpin SE, Cormier KA et al.: ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease. Proc Natl Acad Sci U S A 2006, 103: 19176-19181.

248. Choong CJ, Baba K, Mochizuki H: Gene therapy for neurological disorders. Expert Opin Biol Ther 2016, 16: 143-159.

249. Keiser MS, Kordasiewicz HB, McBride JL: Gene suppression strategies for dominantly inherited neurodegenerative diseases: lessons from Huntington's disease and spinocerebellar ataxia. Hum Mol Genet 2016, 25: R53-R64.

250. Drouet V, Perrin V, Hassig R, Dufour N, Auregan G, Alves S et al.: Sustained effects of nonallele-specific Huntingtin silencing. Ann Neurol 2009, 65: 276-285.

251. Zhang Y, Friedlander RM: Using non-coding small RNAs to develop therapies for Huntington's disease. Gene Ther 2011, 18: 1139-1149.

252. White JK, Auerbach W, Duyao MP, Vonsattel JP, Gusella JF, Joyner AL et al.: Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. Nat Genet 1997, 17: 404-410.

253. Nasir J, Floresco SB, O'Kusky JR, Diewert VM, Richman JM, Zeisler J et al.: Targeted disruption of the Huntington's disease gene results in embryonic lethality

and behavioral and morphological changes in heterozygotes. Cell 1995, 81: 811-823.

254. Harper SQ, Staber PD, He X, Eliason SL, Martins IH, Mao Q et al.: RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. Proc Natl Acad Sci U S A 2005, 102: 5820-5825.

255. Yu D, Pendergraff H, Liu J, Kordasiewicz HB, Cleveland DW, Swayze EE et al.: Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant huntingtin expression. Cell 2012, 150: 895-908.

256. Watts JK, Corey DR: Silencing disease genes in the laboratory and the clinic. J Pathol 2012, 226: 365-379.

257. Geary RS, Norris D, Yu R, Bennett CF: Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. Adv Drug Deliv Rev 2015, 87: 46-51.

258. McClorey G, Wood MJ: An overview of the clinical application of antisense oligonucleotides for RNA-targeting therapies. Curr Opin Pharmacol 2015, 24: 52-58.

259. Dragatsis I, Levine MS, Zeitlin S: Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. Nat Genet 2000, 26: 300-306.

260. Pfister EL, Kennington L, Straubhaar J, Wagh S, Liu W, DiFiglia M et al.: Five siRNAs targeting three SNPs may provide therapy for three-quarters of Huntington's disease patients. Curr Biol 2009, 19: 774-778.

261. Evers MM, Toonen LJ, van Roon-Mom WM: Antisense oligonucleotides in therapy for neurodegenerative disorders. Adv Drug Deliv Rev 2015, 87: 90-103.

262. Chen ZJ, Kren BT, Wong PY, Low WC, Steer CJ: Sleeping Beauty-mediated down-regulation of huntingtin expression by RNA interference. Biochem Biophys Res Commun 2005, 329: 646-652.

263. Harper SQ, Staber PD, He X, Eliason SL, Martins IH, Mao Q et al.: RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. Proc Natl Acad Sci U S A 2005, 102: 5820-5825.

264. Godinho BM, Malhotra M, O'Driscoll CM, Cryan JF: Delivering a disease-modifying treatment for Huntington's disease. Drug Discov Today 2015, 20: 50-64.

265. DiFiglia M, Sena-Esteves M, Chase K, Sapp E, Pfister E, Sass M et al.: Therapeutic silencing of mutant huntingtin with siRNA attenuates striatal and corti-

cal neuropathology and behavioral deficits. Proc Natl Acad Sci U S A 2007, 104: 17204-17209.

266. Carroll JB, Warby SC, Southwell AL, Doty CN, Greenlee S, Skotte N et al.: Potent and selective antisense oligonucleotides targeting single-nucleotide polymorphisms in the Huntington disease gene / allele-specific silencing of mutant huntingtin. Mol Ther 2011, 19: 2178-2185.

267. Kordasiewicz HB, Stanek LM, Wancewicz EV, Mazur C, McAlonis MM, Pytel KA et al.: Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. Neuron 2012, 74: 1031-1044.

268. Labbadia J, Novoselov SS, Bett JS, Weiss A, Paganetti P, Bates GP et al.: Suppression of protein aggregation by chaperone modification of high molecular weight complexes. Brain 2012, 135: 1180-1196.

269. Kampinga HH, Craig EA: The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat Rev Mol Cell Biol 2010, 11: 579-592.

270. Labbadia J, Cunliffe H, Weiss A, Katsyuba E, Sathasivam K, Seredenina T et al.: Altered chromatin architecture underlies progressive impairment of the heat shock response in mouse models of Huntington disease. J Clin Invest 2011, 121: 3306-3319.

271. Labbadia J, Novoselov SS, Bett JS, Weiss A, Paganetti P, Bates GP et al.: Suppression of protein aggregation by chaperone modification of high molecular weight complexes. Brain 2012, 135: 1180-1196.

272. Jackrel ME, Shorter J: Shock and awe: unleashing the heat shock response to treat Huntington disease. J Clin Invest 2011, 121: 2972-2975.

273. Tanaka M, Machida Y, Niu S, Ikeda T, Jana NR, Doi H et al.: Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease. Nat Med 2004, 10: 148-154.

274. Sarkar S, Rubinsztein DC: Small molecule enhancers of autophagy for neurodegenerative diseases. Mol Biosyst 2008, 4: 895-901.

275. Sarkar S, Rubinsztein DC: Huntington's disease: degradation of mutant huntingtin by autophagy. FEBS J 2008, 275: 4263-4270.

276. Lin F, Qin ZH: Degradation of misfolded proteins by autophagy: is it a strategy for Huntington's disease treatment? J Huntingtons Dis 2013, 2: 149-157.

277. Harris H, Rubinsztein DC: Control of autophagy as a therapy for neurodegenerative disease. Nat Rev Neurol 2012, 8: 108-117.

278. Martin DD, Ladha S, Ehrnhoefer DE, Hayden MR: Autophagy in Huntington disease and huntingtin in autophagy. Trends Neurosci 2015, 38: 26-35.

279. Krainc D: Huntington's disease: tagged for clearance. Nat Med 2010, 16: 32-33.

280. Thompson LM, Aiken CT, Kaltenbach LS, Agrawal N, Illes K, Khoshnan A et al.: IKK phosphorylates Huntingtin and targets it for degradation by the proteasome and lysosome. J Cell Biol 2009, 187: 1083-1099.

281. Gu X, Greiner ER, Mishra R, Kodali R, Osmand A, Finkbeiner S et al.: Serines 13 and 16 are critical determinants of full-length human mutant huntingtin induced disease pathogenesis in HD mice. Neuron 2009, 64: 828-840.

282. Jeong H, Then F, Melia TJ, Jr., Mazzulli JR, Cui L, Savas JN et al.: Acetylation targets mutant huntingtin to autophagosomes for degradation. Cell 2009, 137: 60-72.

283. Ravikumar B, Vacher C, Berger Z, Davies JE, Luo S, Oroz LG et al.: Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. Nat Genet 2004, 36: 585-595.

284. Ravikumar B, Stewart A, Kita H, Kato K, Duden R, Rubinsztein DC: Raised intracellular glucose concentrations reduce aggregation and cell death caused by mutant huntingtin exon 1 by decreasing mTOR phosphorylation and inducing autophagy. Hum Mol Genet 2003, 12: 985-994.

285. Lee JH, Tecedor L, Chen YH, Monteys AM, Sowada MJ, Thompson LM et al.: Reinstating aberrant mTORC1 activity in Huntington's disease mice improves disease phenotypes. Neuron 2015, 85: 303-315.

286. Sarkar S, Ravikumar B, Floto RA, Rubinsztein DC: Rapamycin and mTOR-independent autophagy inducers ameliorate toxicity of polyglutamine-expanded huntingtin and related proteinopathies. Cell Death Differ 2009, 16: 46-56.

287. Tsvetkov AS, Miller J, Arrasate M, Wong JS, Pleiss MA, Finkbeiner S: A small-molecule scaffold induces autophagy in primary neurons and protects against toxicity in a Huntington disease model. Proc Natl Acad Sci U S A 2010, 107: 16982-16987.

288. Williams A, Sarkar S, Cuddon P, Ttofi EK, Saiki S, Siddiqi FH et al.: Novel

targets for Huntington's disease in an mTOR-independent autophagy pathway. Nat Chem Biol 2008, 4: 295-305.

289. Shirendeb UP, Calkins MJ, Manczak M, Anekonda V, Dufour B, McBride JL et al.: Mutant huntingtin's interaction with mitochondrial protein Drp1 impairs mitochondrial biogenesis and causes defective axonal transport and synaptic degeneration in Huntington's disease. Hum Mol Genet 2012, 21: 406-420.

290. Goodman AO, Murgatroyd PR, Medina-Gomez G, Wood NI, Finer N, Vidal-Puig AJ et al.: The metabolic profile of early Huntington's disease--a combined human and transgenic mouse study. Exp Neurol 2008, 210: 691-698.

291. Beal MF: Neuroprotective effects of creatine. Amino Acids 2011, 40: 1305-1313.

292. Sora I, Richman J, Santoro G, Wei H, Wang Y, Vanderah T et al.: The cloning and expression of a human creatine transporter. Biochem Biophys Res Commun 1994, 204: 419-427.

293. Wallimann T, Wyss M, Brdiczka D, Nicolay K, Eppenberger HM: Intracellular compartmentation, structure and function of creatine kinase isoenzymes in tissues with high and fluctuating energy demands: the 'phosphocreatine circuit' for cellular energy homeostasis. Biochem J 1992, 281 (Pt 1): 21-40.

294. Ferrante RJ, Andreassen OA, Jenkins BG, Dedeoglu A, Kuemmerle S, Kubilus JK et al.: Neuroprotective effects of creatine in a transgenic mouse model of Huntington's disease. J Neurosci 2000, 20: 4389-4397.

295. Dedeoglu A, Kubilus JK, Yang L, Ferrante KL, Hersch SM, Beal MF et al.: Creatine therapy provides neuroprotection after onset of clinical symptoms in Huntington's disease transgenic mice. J Neurochem 2003, 85: 1359-1367.

296. Hersch SM, Gevorkian S, Marder K, Moskowitz C, Feigin A, Cox M et al.: Creatine in Huntington disease is safe, tolerable, bioavailable in brain and reduces serum 8OH2'dG. Neurology 2006, 66: 250-252.

297. Rosas HD, Doros G, Gevorkian S, Malarick K, Reuter M, Coutu JP et al.: PRECREST: a phase II prevention and biomarker trial of creatine in at-risk Huntington disease. Neurology 2014, 82: 850-857.

298. Yang L, Calingasan NY, Wille EJ, Cormier K, Smith K, Ferrante RJ et al.: Combination therapy with coenzyme Q10 and creatine produces additive neuro-

protective effects in models of Parkinson's and Huntington's diseases. J Neurochem 2009, 109: 1427-1439.

299. Beal MF, Matthews RT, Tieleman A, Shults CW: Coenzyme Q10 attenuates the 1-methyl-4-phenyl-1,2,3, tetrahydropyridine (MPTP) induced loss of striatal dopamine and dopaminergic axons in aged mice. Brain Res 1998, 783: 109-114.

300. Stack EC, Smith KM, Ryu H, Cormier K, Chen M, Hagerty SW et al.: Combination therapy using minocycline and coenzyme Q10 in R6/2 transgenic Huntington's disease mice. Biochim Biophys Acta 2006, 1762: 373-380.

301. Hyson HC, Kieburtz K, Shoulson I, McDermott M, Ravina B, de Blieck EA et al.: Safety and tolerability of high-dosage coenzyme Q10 in Huntington's disease and healthy subjects. Mov Disord 2010, 25: 1924-1928.

302. Sepers MD, Raymond LA: Mechanisms of synaptic dysfunction and excitotoxicity in Huntington's disease. Drug Discov Today 2014, 19: 990-996.

303. Reiner A, Lafferty DC, Wang HB, Del MN, Deng YP: The group 2 metabotropic glutamate receptor agonist LY379268 rescues neuronal, neurochemical and motor abnormalities in R6/2 Huntington's disease mice. Neurobiol Dis 2012, 47: 75-91.

304. Lu B, Nagappan G, Guan X, Nathan PJ, Wren P: BDNF-based synaptic repair as a disease-modifying strategy for neurodegenerative diseases. Nat Rev Neurosci 2013, 14: 401-416.

305. Xie Y, Hayden MR, Xu B: BDNF overexpression in the forebrain rescues Huntington's disease phenotypes in YAC128 mice. J Neurosci 2010, 30: 14708-14718.

306. Navarrete M, Araque A: Endocannabinoids potentiate synaptic transmission through stimulation of astrocytes. Neuron 2010, 68: 113-126.

307. Micale V, Di M, V, Sulcova A, Wotjak CT, Drago F: Endocannabinoid system and mood disorders: priming a target for new therapies. Pharmacol Ther 2013, 138: 18-37.

308. Schiefer J, Landwehrmeyer GB, Luesse HG, Sprunken A, Puls C, Milkereit A et al.: Riluzole prolongs survival time and alters nuclear inclusion formation in a transgenic mouse model of Huntington's disease. Mov Disord 2002, 17: 748-757.

309. Landwehrmeyer GB, Dubois B, de Yebenes JG, Kremer B, Gaus W, Kraus PH et al.: Riluzole in Huntington's disease: a 3-year, randomized controlled study. Ann

Neurol 2007, 62: 262-272.

310. Lee ST, Chu K, Park JE, Kang L, Ko SY, Jung KH et al.: Memantine reduces striatal cell death with decreasing calpain level in 3-nitropropionic model of Huntington's disease. Brain Res 2006, 1118: 199-207.

311. Huntington Study Group: Tetrabenazine as antichorea therapy in Huntington disease: a randomized controlled trial. Neurology 2006, 66: 366-372.

312. Chen Y, Carter RL, Cho IK, Chan AW: Cell-based therapies for Huntington's disease. Drug Discov Today 2014, 19: 980-984.

313. Rossignol J, Fink K, Davis K, Clerc S, Crane A, Matchynski J et al.: Transplants of adult mesenchymal and neural stem cells provide neuroprotection and behavioral sparing in a transgenic rat model of Huntington's disease. Stem Cells 2014, 32: 500-509.

314. El-Akabawy G, Rattray I, Johansson SM, Gale R, Bates G, Modo M: Implantation of undifferentiated and pre-differentiated human neural stem cells in the R6/2 transgenic mouse model of Huntington's disease. BMC Neurosci 2012, 13: 97.

315. Cisbani G, Freeman TB, Soulet D, Saint-Pierre M, Gagnon D, Parent M et al.: Striatal allografts in patients with Huntington's disease: impact of diminished astrocytes and vascularization on graft viability. Brain 2013, 136: 433-443.

316. Paganini M, Biggeri A, Romoli AM, Mechi C, Ghelli E, Berti V et al.: Fetal striatal grafting slows motor and cognitive decline of Huntington's disease. J Neurol Neurosurg Psychiatry 2014, 85: 974-981.

317. Gallina P, Paganini M, Biggeri A, Marini M, Romoli A, Sarchielli E et al.: Human striatum remodelling after neurotransplantation in Huntington's disease. Stereotact Funct Neurosurg 2014, 92: 211-217.

318. Bachoud-Levi AC, Gaura V, Brugieres P, Lefaucheur JP, Boisse MF, Maison P et al.: Effect of fetal neural transplants in patients with Huntington's disease 6 years after surgery: a long-term follow-up study. Lancet Neurol 2006, 5: 303-309.

319. Wang XJ, Cao Q, Zhang Y, Su XD: Activation and regulation of caspase-6 and its role in neurodegenerative diseases. Annu Rev Pharmacol Toxicol 2015, 55: 553-572.

320. Graham RK, Deng Y, Slow EJ, Haigh B, Bissada N, Lu G et al.: Cleavage at the caspase-6 site is required for neuronal dysfunction and degeneration due to mutant

huntingtin. Cell 2006, 125: 1179-1191.

321. Hersch S, Fink K, Vonsattel JP, Friedlander RM: Minocycline is protective in a mouse model of Huntington's disease. Ann Neurol 2003, 54: 841-843.

322. Wang X, Zhu S, Drozda M, Zhang W, Stavrovskaya IG, Cattaneo E et al.: Minocycline inhibits caspase-independent and -dependent mitochondrial cell death pathways in models of Huntington's disease. Proc Natl Acad Sci U S A 2003, 100: 10483-10487.

323. Bonelli RM, Hodl AK, Hofmann P, Kapfhammer HP: Neuroprotection in Huntington's disease: a 2-year study on minocycline. Int Clin Psychopharmacol 2004, 19: 337-342.

324. Landles C, Weiss A, Franklin S, Howland D, Bates G: Caspase-6 does not contribute to the proteolysis of mutant huntingtin in the HdhQ150 knock-in mouse model of Huntington's disease. PLoS Curr 2012, 4: e4fd085bfc9973.

325. Varma H, Voisine C, DeMarco CT, Cattaneo E, Lo DC, Hart AC et al.: Selective inhibitors of death in mutant huntingtin cells. Nat Chem Biol 2007, 3: 99-100.

326. Chu W, Rothfuss J, Chu Y, Zhou D, Mach RH: Synthesis and in vitro evaluation of sulfonamide isatin Michael acceptors as small molecule inhibitors of caspase-6. J Med Chem 2009, 52: 2188-2191.

327. Dunckley T, Coon KD, Stephan DA: Discovery and development of biomarkers of neurological disease. Drug Discov Today 2005, 10: 326-334.

328. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001, 69: 89-95.

329. de GA, Watson S, Ellis LM, Rodon J, Tabernero J, de GA et al.: Pragmatic issues in biomarker evaluation for targeted therapies in cancer. Nat Rev Clin Oncol 2015, 12: 197-212.

330. Azuaje F, Devaux Y, Wagner D: Challenges and standards in reporting diagnostic and prognostic biomarker studies. Clin Transl Sci 2009, 2: 156-161.

331. Aronson JK: Research priorities in biomarkers and surrogate end-points. Br J Clin Pharmacol 2012, 73: 900-907.

332. Chau CH, Rixe O, McLeod H, Figg WD: Validation of analytic methods for biomarkers used in drug development. Clin Cancer Res 2008, 14: 5967-5976.

333. Wagner JA: Overview of biomarkers and surrogate endpoints in drug develop-

ment. Dis Markers 2002, 18: 41-46.

334. Strimbu K, Tavel JA: What are biomarkers? Curr Opin HIV AIDS 2010, 5: 463-466.

335. Beam CA: Statistical considerations when analyzing biomarker data. Clin Immunol 2015, 161: 31-36.

336. Siemers E, Foroud T, Bill DJ, Sorbel J, Norton JA, Jr., Hodes ME et al.: Motor changes in presymptomatic Huntington disease gene carriers. Arch Neurol 1996, 53: 487-492.

337. Reading SA, Dziorny AC, Peroutka LA, Schreiber M, Gourley LM, Yallapragada V et al.: Functional brain changes in presymptomatic Huntington's disease. Ann Neurol 2004, 55: 879-883.

338. Harrington DL, Rubinov M, Durgerian S, Mourany L, Reece C, Koenig K et al.: Network topology and functional connectivity disturbances precede the onset of Huntington's disease. Brain 2015, 138: 2332-2346.

339. Morris JC, Price JL: Pathologic correlates of nondemented aging, mild cognitive impairment, and early-stage Alzheimer's disease. J Mol Neurosci 2001, 17: 101-118.

340. Warren JD, Schott JM, Fox NC, Thom M, Revesz T, Holton JL et al.: Brain biopsy in dementia. Brain 2005, 128: 2016-2025.

341. Schott JM, Reiniger L, Thom M, Holton JL, Grieve J, Brandner S et al.: Brain biopsy in dementia: clinical indications and diagnostic approach. Acta Neuropathol 2010, 120: 327-341.

342. Ravid R, Van Zwieten EJ, Swaab DF: Brain banking and the human hypothalamus--factors to match for, pitfalls and potentials. Prog Brain Res 1992, 93: 83-95.

343. Webster MJ: Tissue preparation and banking. Prog Brain Res 2006, 158: 3-14.

344. Hynd MR, Lewohl JM, Scott HL, Dodd PR: Biochemical and molecular studies using human autopsy brain tissue. J Neurochem 2003, 85: 543-562.

345. Carvey PM, Hendey B, Monahan AJ: The blood-brain barrier in neurodegenerative disease: a rhetorical perspective. J Neurochem 2009, 111: 291-314.

346. Anderson AN, Roncaroli F, Hodges A, Deprez M, Turkheimer FE: Chromosomal profiles of gene expression in Huntington's disease. Brain 2008, 131: 381-388.

347. Biglan KM, Dorsey ER, Evans RV, Ross CA, Hersch S, Shoulson I et al.: Plas-

ma 8-hydroxy-2'-deoxyguanosine Levels in Huntington Disease and Healthy Controls Treated with Coenzyme Q10. J Huntingtons Dis 2012, 1: 65-69.

348. Long JD, Matson WR, Juhl AR, Leavitt BR, Paulsen JS: 8OHdG as a marker for Huntington disease progression. Neurobiol Dis 2012, 46: 625-634.

349. Bogdanov M, Brown RH, Matson W, Smart R, Hayden D, O'Donnell H et al.: Increased oxidative damage to DNA in ALS patients. Free Radic Biol Med 2000, 29: 652-658.

350. De LG, Russo MT, Degan P, Tiveron C, Zijno A, Meccia E et al.: A role for oxidized DNA precursors in Huntington's disease-like striatal neurodegeneration. PLoS Genet 2008, 4: e1000266.

351. Ciancarelli I, De AD, Di MC, Di SC, Pistarini C, D'Orazio N et al.: Peripheral biomarkers of oxidative stress and their limited potential in evaluation of clinical features of Huntington's patients. Biomarkers 2014, 19: 452-456.

352. Borowsky B, Warner J, Leavitt BR, Tabrizi SJ, Roos RA, Durr A et al.: 8OHdG is not a biomarker for Huntington disease state or progression. Neurology 2013, 80: 1934-1941.

353. Killoran A, Biglan KM: 8-OHdG: its (limited) potential as a biomarker for Huntington's disease. Biomark Med 2012, 6: 777-780.

354. Massai L, Petricca L, Magnoni L, Rovetini L, Haider S, Andre R et al.: Development of an ELISA assay for the quantification of soluble huntingtin in human blood cells. BMC Biochem 2013, 14: 34.

355. Weiss A, Trager U, Wild EJ, Grueninger S, Farmer R, Landles C et al.: Mutant huntingtin fragmentation in immune cells tracks Huntington's disease progression. J Clin Invest 2012, 122: 3731-3736.

356. Baldo B, Paganetti P, Grueninger S, Marcellin D, Kaltenbach LS, Lo DC et al.: TR-FRET-based duplex immunoassay reveals an inverse correlation of soluble and aggregated mutant huntingtin in huntington's disease. Chem Biol 2012, 19: 264-275.

357. Stoy N, Mackay GM, Forrest CM, Christofides J, Egerton M, Stone TW et al.: Tryptophan metabolism and oxidative stress in patients with Huntington's disease. J Neurochem 2005, 93: 611-623.

358. Politis M, Lahiri N, Niccolini F, Su P, Wu K, Giannetti P et al.: Increased central microglial activation associated with peripheral cytokine levels in premanifest

Huntington's disease gene carriers. Neurobiol Dis 2015, 83: 115-121.

359. Ciammola A, Sassone J, Cannella M, Calza S, Poletti B, Frati L et al.: Low brain-derived neurotrophic factor (BDNF) levels in serum of Huntington's disease patients. Am J Med Genet B Neuropsychiatr Genet 2007, 144B: 574-577.

360. Seredenina T, Luthi-Carter R: What have we learned from gene expression profiles in Huntington's disease? Neurobiol Dis 2012, 45: 83-98.

361. Botas A, Campbell HM, Han X, Maletic-Savatic M: Metabolomics of neurodegenerative diseases. Int Rev Neurobiol 2015, 122: 53-80.

362. Runne H, Kuhn A, Wild EJ, Pratyaksha W, Kristiansen M, Isaacs JD et al.: Analysis of potential transcriptomic biomarkers for Huntington's disease in peripheral blood. Proc Natl Acad Sci U S A 2007, 104: 14424-14429.

363. Borovecki F, Lovrecic L, Zhou J, Jeong H, Then F, Rosas HD et al.: Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. Proc Natl Acad Sci U S A 2005, 102: 11023-11028.

364. Hu Y, Chopra V, Chopra R, Locascio JJ, Liao Z, Ding H et al.: Transcriptional modulator H2A histone family, member Y (H2AFY) marks Huntington disease activity in man and mouse. Proc Natl Acad Sci U S A 2011, 108: 17141-17146.

365. Chang KH, Chen YC, Wu YR, Lee WF, Chen CM: Downregulation of genes involved in metabolism and oxidative stress in the peripheral leukocytes of Huntington's disease patients. PLoS One 2012, 7: e46492.

366. Kocerha J, Liu Y, Willoughby D, Chidamparam K, Benito J, Nelson K et al.: Longitudinal transcriptomic dysregulation in the peripheral blood of transgenic Huntington's disease monkeys. BMC Neurosci 2013, 14: 88.

367. Reilmann R: Does chromatin modulation provide the first wet biomarker for Huntington's disease? Mov Disord 2012, 27: 473.

368. Ehrlich ME, Gandy S: Chromatin plasticity and the pathogenesis of Huntington disease. Proc Natl Acad Sci U S A 2011, 108: 16867-16868.

369. Kim J, Amante DJ, Moody JP, Edgerly CK, Bordiuk OL, Smith K et al.: Reduced creatine kinase as a central and peripheral biomarker in Huntington's disease. Biochim Biophys Acta 2010, 1802: 673-681.

370. Chang KH, Wu YR, Chen YC, Chen CM: Plasma inflammatory biomarkers for Huntington's disease patients and mouse model. Brain Behav Immun 2015, 44:

121-127.

371. Silajdzic E, Rezeli M, Vegvari A, Lahiri N, Andre R, Magnusson-Lind A et al.: A critical evaluation of inflammatory markers in Huntington's Disease plasma. J Huntingtons Dis 2013, 2: 125-134.

372. Leoni V, Mariotti C, Tabrizi SJ, Valenza M, Wild EJ, Henley SM et al.: Plasma 24S-hydroxycholesterol and caudate MRI in pre-manifest and early Huntington's disease. Brain 2008, 131: 2851-2859.

373. Underwood BR, Broadhurst D, Dunn WB, Ellis DI, Michell AW, Vacher C et al.: Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. Brain 2006, 129: 877-886.

374. Shoulson I, Fahn S: Huntington disease: clinical care and evaluation. Neurology 1979, 29: 1-3.

375. Henley SM, Bates GP, Tabrizi SJ: Biomarkers for neurodegenerative diseases. Curr Opin Neurol 2005, 18: 698-705.

376. Andre R, Scahill RI, Haider S, Tabrizi SJ: Biomarker development for Huntington's disease. Drug Discov Today 2014, 19: 972-979.

377. Tabrizi SJ, Langbehn DR, Leavitt BR, Roos RA, Durr A, Craufurd D et al.: Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. Lancet Neurol 2009, 8: 791-801.

378. Tabrizi SJ, Scahill RI, Durr A, Roos RA, Leavitt BR, Jones R et al.: Biological and clinical changes in premanifest and early stage Huntington's disease in the TRACK-HD study: the 12-month longitudinal analysis. Lancet Neurol 2011, 10: 31-42.

379. Tabrizi SJ, Scahill RI, Owen G, Durr A, Leavitt BR, Roos RA et al.: Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. Lancet Neurol 2013, 12: 637-649.

380. Paulsen JS, Langbehn DR, Stout JC, Aylward E, Ross CA, Nance M et al.: Detection of Huntington's disease decades before diagnosis: the Predict-HD study. J Neurol Neurosurg Psychiatry 2008, 79: 874-880.

381. Duff K, Paulsen JS, Beglinger LJ, Langbehn DR, Stout JC: Psychiatric symp-

toms in Huntington's disease before diagnosis: the predict-HD study. Biol Psychiatry 2007, 62: 1341-1346.

382. Weir DW, Sturrock A, Leavitt BR: Development of biomarkers for Huntington's disease. Lancet Neurol 2011, 10: 573-590.

383. Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH et al.: Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology 2002, 58: 695-701.

384. Rosas HD, Salat DH, Lee SY, Zaleta AK, Pappu V, Fischl B et al.: Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity. Brain 2008, 131: 1057-1068.

385. Georgiou-Karistianis N: A peek inside the Huntington's brain: will functional imaging take us one step closer in solving the puzzle? Exp Neurol 2009, 220: 5-8.

386. Poudel GR, Stout JC, Dominguez DJ, Churchyard A, Chua P, Egan GF et al.: Longitudinal change in white matter microstructure in Huntington's disease: The IMAGE-HD study. Neurobiol Dis 2015, 74: 406-412.

387. Dominguez DJ, Ng AC, Poudel G, Stout JC, Churchyard A, Chua P et al.: Iron accumulation in the basal ganglia in Huntington's disease: cross-sectional data from the IMAGE-HD study. J Neurol Neurosurg Psychiatry 2016, 87: 545-549.

388. JF D, Stout JC, Poudel G, Churchyard A, Chua P, Egan GF et al.: Multimodal imaging biomarkers in premanifest and early Huntington's disease: 30-month IMAGE-HD data. Br J Psychiatry 2016, 208: 571-578.

389. Puopolo M, Pocchiari M: Need to improve clinical trials in rare neurodegenerative disorders. Ann Ist Super Sanita 2011, 47: 55-59.

390. Kang JH, Irwin DJ, Chen-Plotkin AS, Siderowf A, Caspell C, Coffey CS et al.: Association of cerebrospinal fluid beta-amyloid 1-42, T-tau, P-tau181, and alpha-synuclein levels with clinical features of drug-naive patients with early Parkinson disease. JAMA Neurol 2013, 70: 1277-1287.

391. Mollenhauer B, Locascio JJ, Schulz-Schaeffer W, Sixel-Doring F, Trenkwalder C, Schlossmacher MG: alpha-Synuclein and tau concentrations in cerebrospinal fluid of patients presenting with parkinsonism: a cohort study. Lancet Neurol 2011, 10: 230-240.

392. Blennow K, Zetterberg H, Fagan AM: Fluid biomarkers in Alzheimer disease.

Cold Spring Harb Perspect Med 2012, 2: a006221.

393. Wild EJ, Boggio R, Langbehn D, Robertson N, Haider S, Miller JR et al.: Quantification of mutant huntingtin protein in cerebrospinal fluid from Huntington's disease patients. J Clin Invest 2015, 125: 1979-1986.

394. Tan Z, Dai W, van Erp TG, Overman J, Demuro A, Digman MA et al.: Huntington's disease cerebrospinal fluid seeds aggregation of mutant huntingtin. Mol Psychiatry 2015, 20: 1286-1293.

395. Aziz NA, Pijl H, Frolich M, van der Graaf AW, Roelfsema F, Roos RA: Increased hypothalamic-pituitary-adrenal axis activity in Huntington's disease. J Clin Endocrinol Metab 2009, 94: 1223-1228.

396. Shirbin CA, Chua P, Churchyard A, Lowndes G, Hannan AJ, Pang TY et al.: Cortisol and depression in pre-diagnosed and early stage Huntington's disease. Psychoneuroendocrinology 2013, 38: 2439-2447.

397. Hubers AA, van der Mast RC, Pereira AM, Roos RA, Veen LJ, Cobbaert CM et al.: Hypothalamic-pituitary-adrenal axis functioning in Huntington's disease and its association with depressive symptoms and suicidality. J Neuroendocrinol 2015, 27: 234-244.

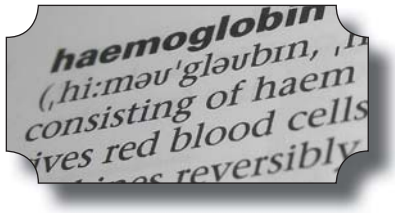
398. Bjorkqvist M, Petersen A, Bacos K, Isaacs J, Norlen P, Gil J et al.: Progressive alterations in the hypothalamic-pituitary-adrenal axis in the R6/2 transgenic mouse model of Huntington's disease. Hum Mol Genet 2006, 15: 1713-1721.

399. Strand AD, Aragaki AK, Shaw D, Bird T, Holton J, Turner C et al.: Gene expression in Huntington's disease skeletal muscle: a potential biomarker. Hum Mol Genet 2005, 14: 1863-1876.

400. Andrich J, Schmitz T, Saft C, Postert T, Kraus P, Epplen JT et al.: Autonomic nervous system function in Huntington's disease. J Neurol Neurosurg Psychiatry 2002, 72: 726-731.

401. Kobal J, Melik Z, Cankar K, Strucl M: Cognitive and autonomic dysfunction in presymptomatic and early Huntington's disease. J Neurol 2014, 261: 1119-1125.

402. Aziz NA, Anguelova GV, Marinus J, van Dijk JG, Roos RA: Autonomic symptoms in patients and pre-manifest mutation carriers of Huntington's disease. Eur J Neurol 2010, 17: 1068-1074.



Chapter 2

Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA

Anastasios Mastrokolias[1], Johan T den Dunnen[1,2], GertJan B van Ommen[1], Peter AC 't Hoen[1], Willeke MC van Roon-Mom[1]

1.Center for Human and Clinical Genetics, Leiden University Medical Center, Einthovenweg 20, 2333ZC, Leiden, The Netherlands
2.Leiden Genome Technology Center, Leiden University Medical Center, Einthovenweg 20, 2333ZC Leiden, The Netherlands

BMC Genomics (2012) 13:28

Abstract

Background

Transcriptome analysis is of great interest in clinical research, where significant differences between individuals can be translated into biomarkers of disease. Although next generation sequencing provides robust, comparable and highly informative expression profiling data, with several million of tags per blood sample, reticulocyte globin transcripts can constitute up to 76% of total mRNA compromising the detection of low abundant transcripts. We have removed globin transcripts from 6 human whole blood RNA samples with a human globin reduction kit and compared them with the same non-reduced samples using deep Serial Analysis of Gene Expression.

Results

Globin tags comprised 52-76% of total tags in our samples. Out of 21,633 genes only 87 genes were detected at significantly lower levels in the globin reduced samples. In contrast, 11,338 genes were detected at significantly higher levels in the globin reduced samples. Removing globin transcripts allowed us to also identify 2112 genes that could not be detected in the non-globin reduced samples, with roles in cell surface receptor signal transduction, G-protein coupled receptor protein signalling pathways and neurological processes.

Conclusions

The reduction of globin transcripts in whole blood samples constitutes a reproducible and reliable method that can enrich data obtained from next generation sequencing-based expression profiling.

Introduction

Transcriptomics technologies have successfully been used for biomarker discovery and the study of physiological and pathophysiological mechanisms [1]. Recently, advances in sequencing technology have allowed for direct identification of transcript specific sequences (tags) that are digitally counted, and the analysis of differences in gene expression with unprecedented accuracy [2]. One specific application of next generation gene expression analysis is DeepSAGE, a tag sequencing method on the Illumina high-throughput sequencing platform that is analogous to LongSAGE [3]. Such sequencing-based technologies offer distinct advantages over expression micro-arrays, such as a higher dynamic range, the increased power for detection of low abundance transcripts, and the detection of novel transcripts and transcript variants [4,5]. Furthermore, next generation sequencing technologies show less variation between different study sites than microarray technology and are not content-limited [6].

Transcriptome analysis of peripheral blood is of great interest for clinical research; where significant differences between samples obtained in a minimally invasive and cost-effective manner can be translated into gene signatures of disease stage, drug response and toxicity [7]. Blood comes into contact with almost every tissue and organ of the human body and due to its cellular composition it can reflect both physiological and pathogenic stimuli. Gene expression differences in peripheral whole blood have been used to determine signatures related to acute myeloid leukaemia [8] , but also in neuropsychiatric disorders and Huntington's disease where significant correlation was found between blood and brain gene expression [9,10]. Furthermore disease signatures can be robust across tissues and experiments and in a large meta-analysis study performed by Dudley and colleagues, it was demonstrated that gene expression profiles in various tissues from the same disease were more similar than gene expression profiles from identical tissues from different diseases [11]. Furthermore, 80% of genes expressed in peripheral blood cells are shared with other important tissues [12]. These results suggest that transcriptome analysis of the blood is often informative even when the originating pathology stems from a different tissue.

Blood is composed of three main cell types. The main component is red blood cells (95%), including the progenitor erythrocytes called reticulocytes, followed by platelets (5%) and white blood cells (<1%). While white blood cells make up the minority, they are the most informative and in the past, many studies have been performed on the isolated white blood cell fraction. However, white blood cell isolation kits can induce a technical bias that can confound the original expression profile of the samples and delayed sample processing can affect gene expression profiles [13,14]. This means that white blood cell separation protocols require quick and accurate processing, which can be difficult in a clinical setting. The expression profile of a sample is better preserved by a whole blood collection method, like the PAX gene blood RNA system [15]. Another advantage of this system is that samples can be frozen for up to 2 years without affecting the expression profile [16]. However, cell sorting and counting is not possible because blood cells are lysed directly after collection. Therefore, abundant transcripts in abundant cell types may conceal the more interesting less abundant transcripts from less abundant cell types. Particularly, the presence of globin transcripts originating from reticulocytes in whole blood samples may limit the sensitivity of gene expression profiling experiments [14], since globin transcripts can constitute up to 70% of the total whole blood mRNA population [17]. While in microarray experiments the presence of globin transcripts may reduce the amount of fluorescent label available for other transcripts but otherwise just results in a saturated spot on the microarray, the high abundance of globin transcripts is more of a concern in sequencing-based expression profiling studies. Since measuring absolute abundance, globin transcripts will be sequenced over and over again, while limiting the coverage of other transcripts.

To deal with the abundance of globin transcripts in whole blood mRNA, several globin reduction protocols have been successfully used in gene expression studies [18-21]. The removal of alpha and beta globin mRNA can be achieved by selective hybridization of biotinylated globin sequence specific oligos with the globin transcripts and depleting them from the total mRNA population through magnetic beads [17]. Another approach developed by Affymetrix and PreAnalytiX utilizes 3' specific peptide nucleid acids (PNAs) to inhibit reverse transcription during cDNA synthesis. Alternatively, a high abundance transcript depletion protocol adopts the properties of the Kamtchaka crab

duplex specific nuclease to selectively reduce the most abundant RNA transcripts [22].

Our aim in the present study is to investigate the effect of globin transcript reduction on the specificity and sensitivity of digital gene expression profiling (deep SAGE).

Materials and Methods

Blood collection & RNA isolation

Samples from six healthy individuals of Caucasian decent (4 females, 2 males) aged from 39 to 70 years old were collected after informed consent and with ethical approval. Whole blood was drawn into PAX gene tubes (Qiagen, Venlo, The Netherlands) and inverted 10 times. Samples were allowed to equilibrate at room temperature for 2 hours, placed at -20°C overnight and stored at -80°C until further processing. Before total RNA extraction, samples were thawed overnight at 4°C and total RNA was isolated using the PAX RNA isolation kit following the manufacturer's instructions, including DNAse treatment. RNA quality and the RIN values were determined using the RNA Nano Lab-Chip assay (Agilent Technologies, Santa Clara, CA, USA). The concentration of each sample was validated using the Nanodrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

Globin reduction

From each sample, 1.5 μg of total RNA was treated using the GLOBINclear™ Human Kit (Ambion, Austin, TX, USA) according the manufacturer's instructions. The kit contains 10 globin sequence specific biotinylated oligonucleotides (http://www.freepatentsonline.com/y2006/0257902.html). Transcripts hybridized to the oligonucleotides are removed through streptavidin magnetic beads. Concentration and quality of the samples were checked as described above.

Quantitative real-time PCR

cDNA was synthesized from 1µg of total RNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Mannheim, Germany) with Random Hexamer primers at 50°C for 1 hour. The primer sequences designed to target genes of interest and the reference gene *ACTB* are described in Additional file 4. The qPCR was performed using 1µl of 4x diluted cDNA, 2x FastStart Universal SYBR Green Master mix (Roche), 2.5 pmol forward primer,2.5 pmol reverse primer and PCR grade water to a total volume of 10µl. The reaction was performed using the LightCycler 480 (Roche) with initial denaturation 10 min. at 95°C, followed by 45 cycles of 10 sec. denaturation at 95°C, 30 sec. annealing at 60°C and 20 sec. elongation at 72°C and final elongation was performed 5 min. at 72°C. This was followed by melting curve analysis from 60°C to 98°C with a ramp rate of 0.02°C per sec. The primer efficiencies and relative expression of the transcript levels was determined using LinRegPCR_v11.3[29].

SAGE library production

SAGE libraries were produced according to the Illumina protocol [6]. In short, after hybridization of 1µg total RNA to polydT magnetic beads (Dynabeads, Invitrogen Life Technologies, Carlsbad, CA, USA), first and second strand synthesis was performed. The beads attached to the double stranded DNA were digested with NlaIII restriction endonuclease to produce short double stranded constructs starting at the most 3' CATG of the transcript. After ligation of the GEX adapter 1, the construct was digested with MmeI to create a 21 base pair fragment downstream of the GEX adapter 1. This fragment was then ligated to GEX adapter 2 to complete the cassette. The adaptor ligated constructs were amplified by 15 cycles of PCR and the products were loaded on 6% Novex TBE precast acrylamide gels (Invitrogen). The 96 base pair band corresponding to the NlaIII construct was excised and purified from the gel slice using the soak and crush method followed by ethanol precipitation. Sample quality was checked on a DNA 1000 Lab-on-a-Chip (Agilent). Sequencing was performed at the Leiden Genome Technology Center on Illumina GA2 sequencer (Illumina, San Diego, CA, USA). Purified samples were diluted to 10nM and loaded on a single lane of the flow cell where, after

cluster amplification, samples were put through an ultra short 18 cycle sequencing run.

Sequence processing

Illumina Pipeline Software version 1.5 was used for data sequence processing. The FASTQ files were analysed using the open source GAPSS_B(v2) pipeline (www.lgtc.nl/GAPSS). All sequences were trimmed to 17 base pairs to remove the first lower quality base pair from the 3' end of the sequences. After trimming, the NlaIII recognition site (CATG) was added to the 5' end of the sequence to create the complete 21/22-mer nucleotide sequences. Sequences were aligned using the Bowtie short read aligner (version 0.12.7) against the UCSC hg19 reference genome, allowing for a maximum of one mismatch and a maximum of two possible positions in the genome (options: -k 1 -m 2 -n 1 --best --strata –solexa1.3-quals).

A custom Perl script was used to create reference region files from the SAGE region files that were composed of the overlapping tags from all samples. A second Perl script was then used to link all individual region files to the reference region file, reporting the number of tags in each individual region of the reference region file. Finally, the reference region file was annotated with transcript information using BioMart (Ensembl build 60). For all reported downstream statistical and biological analysis only sequences aligned to known exons were used. Analyses were performed at the gene level, and in case of multiple SAGE tags per gene, e.g. as a consequence of alternative polyadenylation, tags were summed. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE33701 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33701).

Statistical analysis

To identify genes detected at lower or higher levels, data files containing the count expression data were analyzed using the edgeR package (version 2.0.5) [30,31] in R (version 2.12.0). Data was normalized creating libraries of equal size (11 million tags). To determine differences in detection levels between the two groups an exact test for the

negative binomial distribution was used. The *P* values were adjusted for multiple testing using the Benjamini-Hochberg approach for controlling the false discovery rate (FDR) and genes were considered to be different between reduced and non-reduced samples when the *P* value was less than 0.01. For pathway enrichment analysis, the online tool DAVID was used. To identify gene biotypes, Ensembl genes IDs were uploaded to the BioMart site and supplied gene biotype attributes (Ensembl build 60, Homo sapiens genome reference GRCh37.p2). To examine if transcripts are uniquely and consistently identified in the globin reduced or non-reduced samples only, we applied a tag count threshold of 5 tags or more and this threshold had to be met in 5 out of the 6 samples in the respective group.

Results

Globin transcripts can successfully be reduced in whole blood samples

To investigate whether the RNA quality was affected by globin reduction, the RNA Integrity Number (RIN) value was determined for each sample (summarized in Table 1) with a RIN value of 10 representing the highest RNA quality. On average, the RIN value was slightly lower in the globin-reduced samples compared to the non-reduced samples indicating a minor reduction in RNA quality. Additionally, we measured RNA concentrations before and after globin reduction and observed an RNA yield loss of 5-9 %. General sequencing and alignment statistics are given in Table 1. To verify if globin reduction was successful, the percentage of globin transcripts was determined in the reduced and non-reduced samples. The most abundant transcripts in the non-reduced samples corresponded to globin alpha1 (*HBA1*), globin beta (*HBB*) and globin alpha2 (*HBA2*) (Figure 1). The percentage of globin transcripts constituted on average 60% (52-76%) of total reported alignments and for the reduced samples this dropped to 0.1% - 0.4 % (Table 1), demonstrating the efficiency of the globin reduction protocol. The two transcripts

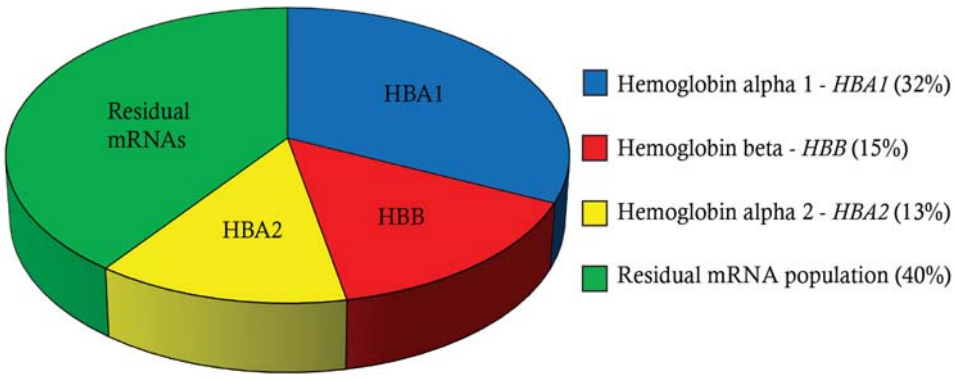


Figure 1. Transcript abundance in the non-globin reduced samples. Pie chart showing the most abundant transcripts in whole blood (average of 6 samples). Hemoglobin alpha 1 (*HBA1*), hemoglobin beta (*HBB*) and hemoglobin alpha 2 (*HBA2*) comprise the majority of transcripts.

that were most abundant after the globin transcripts, lysozyme (*LYZ*) and ribosomal protein (*RPLP2*) transcripts were detected at higher levels in the globin-reduced samples. In order to validate the next generation sequencing results we performed qPCR for HBA and *HBB* for 4 non reduced and reduced samples and the results confirmed the next generation sequencing results (see Additional file 1). To examine the reproducibility of globin reduction we performed a logarithmic transformation on the count data from the 6 globin reduced samples and calculated the correlation. Correlation values ranged from 0.88 to 0.94, showing good reproducibility after globin reduction (Additional file 2).

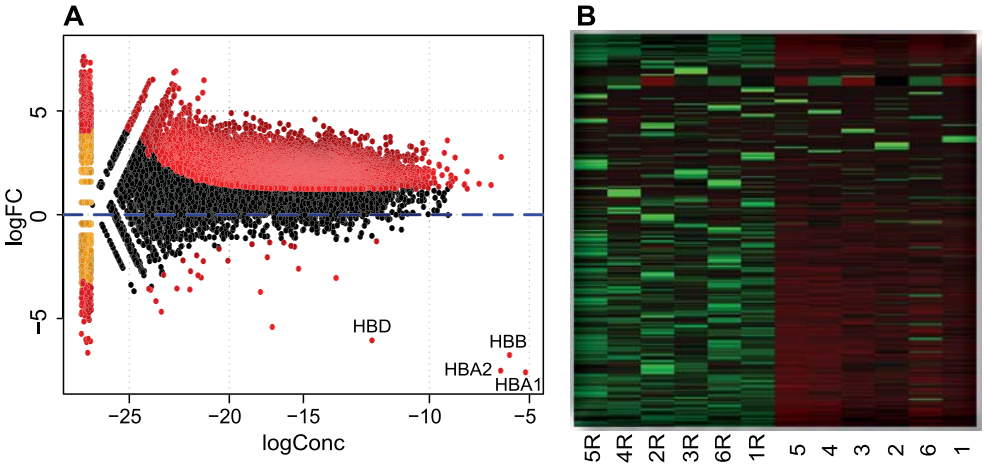


Figure 2. Differential expression of non-globin reduced versus globin reduced whole blood samples **A.** MA-plot (log concentration on x-axis, log fold-change on y-axis) showing that the majority of differentially expressed genes (P value < 0.01, indicated in red) are up-regulated while the 4 globin genes showed the strongest reduction. Genes with zero expression in one of the groups are shown on the left end of the plot. *HBD*: hemoglobin delta, *HBB*: hemoglobin beta, *HBA1*: hemoglobin alpha 1, *HBA2*: hemoglobin alpha 2. CXorf25 has a logConc of -18.79 and a logFC of -14.3 and is not shown on this graph. **B.** Heatmap of normalised expression values of 21633 genes across the 6 non reduced and 6 reduced (R) samples where green depicts high expression and red low expression values.

Samples	Non- globin reduced						Globin-reduced					
Number	1	2	3	4	5	6	1	2	3	4	5	6
RIN	8.1	7.7	7.1	6.9	7.3	8.8	7.6	7.2	6.7	6.9	7	8.2
Tag numbers	$29.2 \cdot 10^6$	$13.2 \cdot 10^6$	$32.0 \cdot 10^6$	$11.6 \cdot 10^6$	$13.6 \cdot 10^6$	$10.1 \cdot 10^6$	$16.5 \cdot 10^6$	$28.7 \cdot 10^6$	$15.9 \cdot 10^6$	$15.3 \cdot 10^6$	$15.9 \cdot 10^6$	$14.9 \cdot 10^6$
Non-aligned	1.27%	5.63%	1.43%	13.80%	5.64%	1.92%	3.64%	2.95%	3.37%	3.29%	3.45%	3.38%
Aligned	89.9%	84.4%	89.9%	80.1%	89.1%	91%	78.3%	78.9%	78.4%	78.9%	77.6%	77.5%
Globin tags	60.1%	52.1%	57.2%	69.6%	76%	71%	0.3%	0.4%	0.4%	0.2%	0.4%	0.1%

Table 1. Sequence Statistics and RNA qualities. Quality of RNA depicted by RNA integrity numbers (RIN) for the 6 non-globin reduced and globin reduced samples. The number of tags, percentages of aligned and non-aligned sequences, and relative number of globin tags are shown as percentage of the total number of aligned tags.

Gene Description	Gene name	logConc[1]	logFC[2]	*P*-Value	FDR
Pseudogene	CXorf25	-18.79	-14.3	2.10e-54	3.68e-50
Globin Alpha1	HBA1	-5.19	-7.59	7.48e-34	6.53e-30
Globin Alpha2	HBA2	-6.43	-7.51	2.12e-33	1.23e-29
Globin Beta	HBB	-5.99	-6.76	3.82e-29	1.66e-25
Globin Delta	HBD	-12.87	-6.05	4.37e-25	1.52e-21
APX/Shroom Family	SHROOM4	-17.84	-5.37	6.66e-20	1.93e-16
Ataxia/Rad3	ATR	-15.71	4.57	6.80e-17	1.69e-13
Pseudogene	RP11-431J24.4	-21.08	6.1	9.66e-17	2.10e-13
Trub sudouridine synthase	TRUB1	-19.14	4.83	6.43e-16	1.24e-12
Pseudogene	CTD-2165H16.1	-18.82	4.71	1.48e-15	2.58e-12

Table 2. Top ten most differentially expressed transcripts. The 4 most abundant globin transcripts are in the top five of differentially down-regulated transcripts ([1] logConc - Overall concentration for a tag across the two groups compared, [2] logFC - The log-fold change difference for the counts between the groups compared).

Differential gene expression differences between globin reduced and non-reduced samples

To investigate if more genes could be identified after globin reduction we used the edgeR Bioconductor statistical R package to examine differential gene expression between the 6 globin reduced and 6 non-reduced samples. We identified 21,633 uniquely expressed genes in the 12 samples and 11,633 out of these 21,633 genes were detected at significantly higher levels (FDR 1%) in the globin reduced samples (Figure 2). There were 87 genes detected at significantly lower levels in the globin reduced samples. The transcripts that were the most significantly decreased (p-value) were the pseudogene *CXorf25* and the 4 globin transcripts globin delta (*HBD*), hemoglobin beta (*HBB*), hemoglobin alpha 1 (*HBA1*), hemoglobin alpha 2 (*HBA2*) (see Table 2). Furthermore, 82 other non-globin transcripts were significantly reduced by the globin reduction procedure (Additional File 3).

To examine why 83 non-globin transcripts were reduced by the globin reduction procedure, we examined possible non-target specific hybridization of the globin oligos. Using the sequences of the ten biotinylated oligonucleotides from the globin reduction kit we performed a blastn search against the human mRNA reference genome but only the 4 globin transcripts that were found to be most significantly decreased in our globin-reduced samples showed significant hits in blastn. Moreover, oligonucleotide sequences were analyzed with RNAhybrid, an online tool for determining the minimum free hybridization energy between a long and a short RNA [23] but no conclusive evidence was found for non-specific hybridization between oligonucleotides from the globin reduction kit and the 83 non-globin transcripts that were reduced in our globin-reduced samples. Finally we wanted to investigate if any of the 83 transcripts could have been co-captured with the globin transcripts. Hence we investigated whether there was sequence homology between the 83 top reduced transcripts and the globin transcripts but there was no indication of co-capturing based on sequence homology. Furthermore, more than half the transcripts reported as non-specifically detected at lower levels were derived from pseudogenes or expressed at very low levels.

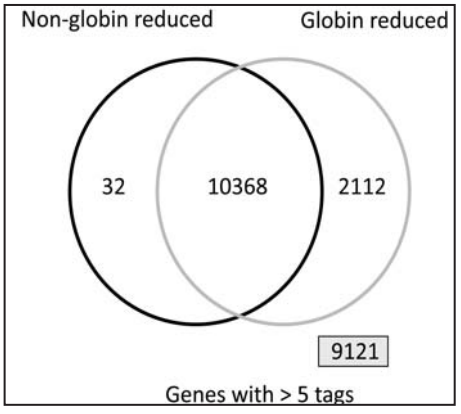


Figure 3. Venn diagram displaying the number of transcripts in at least 5 out of 6 samples at a level of 5 tags or more. Reducing the globin transcripts increases the number of transcripts that can be reliably detected (2112 genes). 9121 genes were not consistently expressed in either group.

Transcripts detected in globin reduced samples only

From the total number of 21,633 transcripts discovered in our SAGE experiments, 10,368 were expressed above the threshold of at least 5 tags in 5 out of the 6 samples in both the globin-reduced and non-reduced group. We identified 2,112 transcripts in the globin-reduced samples that were not expressed above the threshold in the non-re-

duced samples and 32 transcripts in the non-globin reduced samples that could not be detected in the globin-reduced samples (see Figure 3). Additionally, 9,121 transcripts were detected in globin-reduced samples only but did not reach the threshold (low abundance transcripts). Increasing the threshold tag number resulted in a higher number of genes detected by the globin reduced samples while the number of genes only detected in non-reduced samples remained approximately the same. Furthermore increasing the threshold resulted in less genes identified by both groups.

Functional enrichment and gene biotype analysis

For the 2,112 transcripts identified in the globin reduced samples only, we performed gene ontology (GO) term enrichment analysis using the online DAVID Bioinformatics Resources 6.7 [24] (see Table 3). The GO terms that were most significantly enriched were "metal and ion binding" and "DNA binding", the latter including genes encoding for proteins with roles in transcription and regulation of transcription. We also performed the same enrichment analysis for the 9,121 transcripts that did not reach the

GO:	Term	Count	Gene %	P value
MF	Ion binding	383	23.7	7.90e-6
MF	Cation binding	380	23.6	4.30e-6
MF	Metal ion binding	379	16.4	2.20e-6
MF	Transition metal ion binding	272	16.9	9.80e-7
MF	Zinc ion binding	229	14.2	3.50e-6
MF	DNA binding	214	13.3	9.50e-4
BP	Regulation of transcription	214	13.3	8.40e-2
CC	Intracellular non membrane bound organelle	210	13.0	1.30e-2
BP	Transcription	183	11.3	2.50e-2
CC	Plasma membrane part	176	10.9	3.70e-2

Table 3. GO term enrichment analysis using DAVID. Top 10 Gene Ontology terms found enriched by the DAVID online analysis tool in the 2,112 transcripts uniquely found in the globin reduced samples. *P*-value reflects the *P*-value from the hypergeometric test. CC-Cellular Component, BP-Biological process, MF-Molecular function.

more stringent threshold. This group of genes also showed significant enrichment for metal and ion binding proteins. Additionally, we observed enrichment in transcripts coding for cell surface receptors signal transduction proteins, G-protein coupled receptors and proteins involved in neurological processes. Since these transcripts are present in low copy numbers in our samples, these transcripts would benefit from globin reduction and even greater depth of sequencing.

There are around 20,000 protein-coding genes that together comprise only 2% of the human genome [25]. However, a large fraction of transcripts map outside known protein coding genes [26] and it is becoming clear that non-protein-coding RNAs play critical roles as transcriptional, post-transcriptional and chromatin-modifying regulators [27]. As such, we were interested in characterizing the gene biotype of the 21,633 transcripts identified in our gene expression analysis (see Figure 4). We analyzed the transcripts

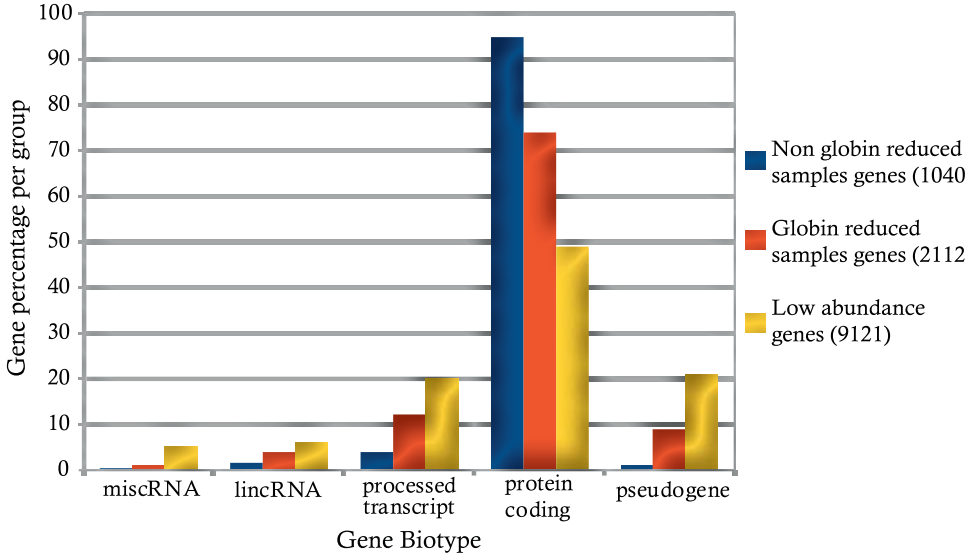


Figure 4. Gene biotype distribution in transcripts with and without globin reduction. Percentage of identified transcripts in the non-globin reduced and globin-reduced samples; low abundance transcripts consist of transcripts detected at levels too low to be assigned to any of the two previous groups. The 9121 low abundance transcripts contain relatively more processed transcripts, pseudogenes, large intergenic non coding RNAs (lincRNAs) and miscellaneous RNAs (miscRNAs).We statistically compared the distribution of the number of transcripts over the different categories with a chi-square test and all pairwise comparisons were highly significant ($p<2.2e-16$).

according to the groups defined in Figure 3. The majority of transcripts from all three groups were protein coding transcripts. However, there were subtle differences in the identified transcripts. The 2112 transcripts that were uniquely identified in the globin-reduced samples contained slightly elevated numbers of pseudogenes, processed transcripts and large intergenic non-coding RNAs (lincRNAs). The 9121 transcripts whose tag numbers were too low to be consistently ascribed to either of the experimental groups contained relatively more pseudogenes.

Sequencing depth vs. globin reduction - Genes discovered

The previous results showed that by reducing the number of globin transcripts before sequencing, the number of different transcripts that can be identified increases. Next, we investigated how much additional sequencing is required to achieve this effect of globin reduction just by sequencing non-globin reduced samples at a greater depth. For this purpose we compared results from non-globin reduced RNA samples that were sequenced at high and low sequencing depth. At two times higher sequencing depth, the reported alignments as well as the number of identified transcripts for each of the samples increased to the level of the globin reduced samples (Table 4). The number of 13,000

Group	Non-reduced				Reduced	
Samples	Sample1 Low	Sample2 Low	Sample1 High	Sample2 High	Sample1	Sample2
Total number of sequences	$7.8 \cdot 10^6$	$11.6 \cdot 10^6$	$29.2 \cdot 10^6$	$32.0 \cdot 10^6$	$16.5 \cdot 10^6$	$15.9 \cdot 10^6$
Reported Alignments	$4.7 \cdot 10^6$	$5.5 \cdot 10^6$	$23.6 \cdot 10^6$	$25.6 \cdot 10^6$	$9.8 \cdot 10^6$	$9.4 \cdot 10^6$
Transcripts Detected	9,823	9,003	13,085	12,738	13,302	12,831

Table 4. Number of transcripts identified across different sequencing depths. Two non-globin-reduced samples were sequenced twice, at low and high sequencing depths, and compared to the sequencing results from the same samples with reduction of the globin transcripts.

detected transcripts in globin-reduced samples is achieved at a sequencing depth of ~15 million reads while the same number of transcripts is detected at a sequencing depth of ~25 million reads in the non-globin reduced samples. *In silico* analysis was performed to determine at which sampling rate the threshold of at least 5 tags in 5 out the 6 samples was reached. The difference in detected transcripts between the reduced and non-reduced samples at any given sequencing depth remained equal, indicating that even at a very large numbers of reads (>60 million reads) whole blood samples are not sequenced to saturation (see Figure 5).

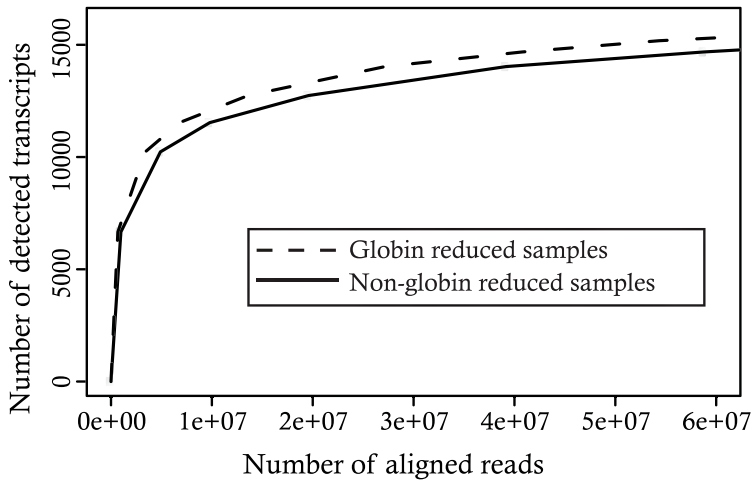


Figure 5. Number of transcripts detected at increased sampling rate. *In silico* analysis using the sum of aligned reads (total number of reads ~60 million) across each group (reduced, non-reduced). A random subset of these aligned reads is then used to find the number of detected transcripts across the two sample groups over the 5 tag, 5 out of 6 samples threshold. The plot shows that at similar sequencing depths, there is a consistently higher number of transcripts detected in the globin reduced samples compared to the non-globin reduced samples.

Discussion

We performed globin transcript reduction on RNA from human peripheral blood to investigate if this improved the sensitivity of SAGE digital gene expression profiling. In the non-reduced samples, globin percentages were highly variable between samples and ranged from 52% to 76%. These numbers match previous reports for globin transcript abundances in blood [17]. The globin reduction process with biotinylated oligonucleotides complementary to globin transcripts was successful since there was a >99.6% reduction in globin transcripts. This was consistent with previous studies [28]. We observed that after globin reduction there was a slight decrease in total RNA quality as well as an RNA yield loss of 5-9%. The slight reduction in RNA quality and yield had no effect on the quality of the data. From the 21,633 transcripts that were detected across all 12 samples, 11,633 transcripts were detected at significantly higher levels in the 6 globin reduced samples, while only 87 transcripts were reported as detected at significantly lower levels. We can not explain why there were transcripts, other than the globin transcripts, found to be expressed at lower levels in the globin reduced samples. *In silico* analysis showed that this was not likely due to cross hybridization with the globin transcripts or with the globin oligonucleotides from the globin reduction kit.

We detected robust expression of 2112 transcripts in globin reduced samples that could not be detected in non-globin reduced samples. This is similar to what has been previously reported using classical microarray gene expression platforms [17]. The important advantage of more detectable transcripts is that there will be increased statistical power to detect differences between samples. Furthermore, this increased number of detected transcripts after globin reduction provides more information about low abundance transcripts that could give new insights into the underlying disease mechanism.

The reduction of globin transcripts provides a practical and reproducible method for improving the number of transcripts detected in human peripheral whole blood samples, and reduces required sequencing capacity by a factor two. DeepSAGE is a technique that detects short tags of 21-22 base pairs representing specific transcripts. For this reason the data complexity is lower compared to other digital gene expression profiling

techniques such as whole transcriptome sequencing (RNA-seq). For RNA-seq, where the aim is to also identify all the transcript isoforms, saturation is definitely not reached at 25 million reads, which would make globin transcript reduction in blood even more advantageous for RNA-seq.

Conclusions

The reduction of globin transcripts together with whole blood collection methods, such as the PAX system, constitutes a reproducible and reliable method that increases the number of transcripts detected in next generation sequencing-based gene expression profiling. This will increase the statistical power to detect disease relevant signatures in patient-control studies.

Authors' contributions

AM carried out the sample processing, sequencing, statistical analysis, and drafted the manuscript. JTDD participated in the design of the study and helped to draft the manuscript. GJBVO participated in the design of the study and helped to draft the manuscript PACTH participated in the design of the study, the statistical analysis and helped to draft the manuscript. WMCVRM conceived the study, and participated in its design and coordination, and helped to draft the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

The blood samples were kindly provided by Prof Dr R.C. van der Mast and Dr E. van Duijn at the Psychiatry Department, LUMC, The Netherlands. This research was partly funded by the European Commission 7th Framework Program, Project no. 261123, GEUVADIS, Project no. 201413, ENGAGE, by the Centre for Medical Systems Biology within the framework of the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research and Dutch Centre for Biomedical Genetics.

Reference List

1. Le-Niculescu H, Kurian SM, Yehyawi N, Dike C, Patel SD, Edenberg HJ et al.: Identifying blood biomarkers for mood disorders using convergent functional genomics. Mol Psychiatry 2009, 14: 156-174.
2. Shendure J, Ji H: Next-generation DNA sequencing. Nat Biotechnol 2008, 26: 1135-1145.
3. Nielsen KL, Hogh AL, Emmersen J: DeepSAGE digital transcriptomics with high sensitivity, simple experimental protocol and multiplexing of samples. Nucleic Acids Res 2006, 34: e133.
4. Feng L, Liu H, Liu Y, Lu Z, Guo G, Guo S et al.: Power of deep sequencing and agilent microarray for gene expression profiling study. Mol Biotechnol 2010, 45: 101-110.
5. van Iterson M., 't Hoen PA, Pedotti P, Hooiveld GJ, Den Dunnen JT, Van Ommen GJ et al.: Relative power and sample size analysis on gene expression profiling data. BMC Genomics 2009, 10: 439.
6. 't Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX et al.: Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Res 2008.
7. Pahl A: Gene expression profiling using RNA extracted from whole blood: technologies and clinical applications. Expert Rev Mol Diagn 2005, 5: 43-52.
8. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani, Boer JM et al.: Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 2004, 350: 1617-1628.
9. Borovecki F, Lovrecic L, Zhou J, Jeong H, Then F, Rosas HD et al.: Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. PNAS 2005, 102: 11023-11028.
10. Sullivan PF, Fan C, Perou CM: Evaluating the comparability of gene expression in blood and brain. Am J Med Genet B Neuropsychiatr Genet 2006, 141B: 261-268.
11. Dudley JT, Tibshirani R, Deshpande T, Butte AJ: Disease signatures are robust across tissues and experiments. Mol Syst Biol 2009, 5: 307.

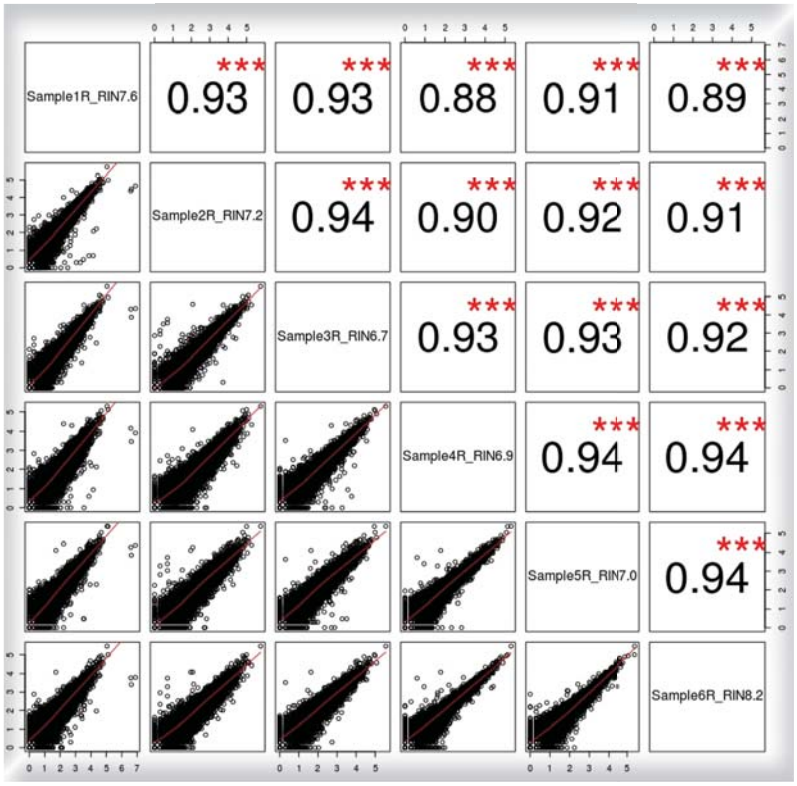
12. Liew CC, Ma J, Tang HC, Zheng R, Dempsey AA: The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. J Lab Clin Med 2006, 147: 126-132.
13. Muller MC, Merx K, Weisser A, Kreil S, Lahaye T, Hehlmann R et al.: Improvement of molecular monitoring of residual disease in leukemias by bedside RNA stabilization. Leukemia 2002, 16: 2395-2399.
14. Debey S, Schoenbeck U, Hellmich M, Gathof BS, Pillai R, Zander T et al.: Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. Pharmacogenomics J 2004, 4: 193-207.
15. Rainen L, Oelmueller U, Jurgensen S, Wyrich R, Ballas C, Schram J et al.: Stabilization of mRNA expression in whole blood samples. Clin Chem 2002, 48: 1883-1890.
16. Ovstebo R, Lande K, Kierulf P, Haug KB: Quantification of relative changes in specific mRNAs from frozen whole blood - methodological considerations and clinical implications. Clin Chem Lab Med 2007, 45: 171-176.
17. Field LA, Jordan RM, Hadix JA, Dunn MA, Shriver CD, Ellsworth RE et al.: Functional identity of genes detectable in expression profiling assays following globin mRNA reduction of peripheral blood samples. Clin Biochem 2007, 40: 499-502.
18. Vartanian K, Slottke R, Johnstone T, Casale A, Planck SR, Choi D et al.: Gene expression profiling of whole blood: comparison of target preparation methods for accurate and reproducible microarray analysis. BMC Genomics 2009, 10: 2.
19. Wright C, Bergstrom D, Dai HY, Marton M, Morris M, Tokiwa G et al.: Characterization of globin RNA interference in gene expression profiling of whole-blood samples. Clin Chem 2008, 54: 396-405.
20. Tian Z, Palmer N, Schmid P, Yao H, Galdzicki M, Berger B et al.: A Practical Platform for Blood Biomarker Study by Using Global Gene Expression Profiling of Peripheral Whole Blood. PLoS ONE 2009, 4.
21. Debey S, Zander T, Brors B, Popov A, Eils R, Schultze JL: A highly standardized, robust, and cost-effective method for genome-wide transcriptome analysis of peripheral blood applicable to large-scale clinical trials. Genomics 2006, 87: 653-664.
22. Bogdanova EA, Shagin DA, Lukyanov SA: Normalization of full-length enriched cDNA. Mol Biosyst 2008, 4: 205-212.

23. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R: Fast and effective prediction of microRNA/target duplexes. RNA 2004, 10: 1507-1517.
24. Huang dW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009, 4: 44-57.
25. Human Genome Sequencing Consortium: Finishing the euchromatic sequence of the human genome. Nature 2004, 431: 931-945.
26. Wright MW, Bruford EA: Naming 'junk': human non-protein coding RNA (ncRNA) gene nomenclature. Hum Genomics 2011, 5: 90-98.
27. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS: Non-coding RNAs: regulators of disease. J Pathol 2010, 220: 126-139.
28. Liu J, Walter E, Stenger D, Thach D: Effects of globin mRNA reduction methods on gene expression profiles from whole blood. J Mol Diagn 2006, 8: 551-558.
29. Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB et al.: Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research 2009, 37.
30. Robinson MD, Smyth GK: Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics 2008, 9: 321-332.
31. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010, 26: 139-140.

Supplementary material

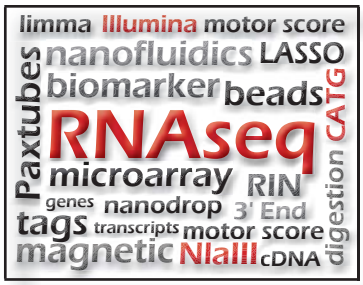
Additional file 1. qPCR results showing relative (**A**) and absolute (**B**) hemoglobin alpha (HBA) and hemoglobin beta (HBB) levels before (Globin) and after globin reduction (Globin Reduced). Results representthe biological average of 4 samples.

Additional file 2. Correlation matrix for the globin reduced samples. Upper panels represent correlation values of log transformed data with asterisks depicting significance. Lower panels contain scatter plots of log transformed globin reduced samples.



Additional file 4. qPCR primer Sequences. Primer sequences used for qPCR validation of HBA, HBB transcript levels.

HBA - FWD	ACTCTTCTGGTCCCCACAGA	75bp product
HBA - REV	GCCTTGACGTTGGTCTTGTC	
HBB - FWD	TTGAGTCCTTTGGGGATCTG	75bp product
HBB - REV	CTTTCTTGCCATGAGCCTTC	
BACTIN - FWD	GGCATCCTCACCCTGAAGTA	82bp product
BACTIN - REV	AGGTGTGGTGCCAGATTTTC	



Chapter 3

Huntington's disease biomarker progression profile identified by transcriptome sequencing in peripheral blood

Anastasios Mastrokolias[1] , Yavuz Ariyurek[2], Jelle J. Goeman[3,4], Erik van Duijn[5,6] Raymund A.C. Roos[7], Roos C. van der Mast[5], GertJan B. van Ommen[1] ,Johan T. den Dunnen[1,2], Peter A.C. 't Hoen[1], Willeke M.C. van Roon-Mom[1]

1.Department of Human Genetics, Leiden University Medical Center,The Netherlands
2.Leiden Genome Technology Center, Leiden University Medical Center, The Netherlands
3.Department of Medical Statistics and Bioinformatics, Leiden University Medical Center,The Netherlands
4.Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands
5.Department of Psychiatry, Leiden University Medical Center, The Netherlands
6.Center for Mental Health Care Delfland, Delft, The Netherlands
7.Department of Neurology, Leiden University Medical Centre,The Netherlands

European Journal of Human Genetics (2015) 23, 1349–1356

Abstract

With several therapeutic approaches in development for Huntington's disease, there is a need for easily accessible biomarkers to monitor disease progression and therapy response. We performed next generation sequencing-based transcriptome analysis, of total RNA from peripheral blood of 91 mutation carriers (27 presymptomatic and, 64 symptomatic) and 33 controls. Transcriptome analysis by DeepSAGE identified 167 genes significantly associated with clinical total motor score in Huntington's disease patients. Relative to previous studies this yielded both novel genes, and confirmed previously identified genes, such as *H2AFY*, an overlap in results which has proven difficult in the past. Pathway analysis showed enrichment of genes of the immune system and of target genes of miRNAs which are downregulated in Huntington's disease models. Using a highly parallelized microfluidics array chip (Fluidigm) we validated 12 of the top 20 significant genes in our discovery cohort and 7 in a second independent cohort. The five genes (*PROK2*, *ZNF238*, *AQP9*, *CYSTM1*, and *ANXA3*) which were validated independently in both cohorts present a candidate biomarker panel for stage determination and therapeutic readout in Huntington's disease. Finally, we suggest a first, empiric formula predicting total motor score from the expression levels of our biomarker panel. Our data support the view that peripheral blood is a useful source to identify biomarkers for Huntington's disease and monitor disease progression in future clinical trials.

Introduction

Huntington's disease (HD) is a heritable neurodegenerative disorder that manifests itself through cognitive, psychiatric and motor symptoms. The pathology is caused by an expanded CAG repeat in the *HTT* gene, resulting in a mutant huntingtin protein. Patients also develop peripheral pathology [1] and increasing evidence indicate that peripheral inflammation has a role as a disease progression modulator [2]. HD brain tissue is characterized by mutant protein aggregate formation and neuronal cell loss with transcriptional deregulation as a prominent feature [3,4]. Several mechanisms have been implicated in this deregulation such as: histone modifications, transcription factor impairment and aberrant miRNA expression [5]. For HD clinical trials it is important to identify disease progression biomarkers. Longitudinal studies have shown that imaging biomarkers and clinical measures provide valuable information [6]. However clinical measures can be subject to inter-rater variability and imaging is expensive. A biomarker should be able to identify changes before clinical symptoms, should be easily obtained and respond well to disease-modifying interventions. As it is impossible to measure molecular biomarkers in brain, the use of more accessible tissues has been proposed, such as blood. Leukocytes involved in immune system regulation, make blood an ideal source for identifying HD events such as peripheral inflammation. In addition, as huntingtin is ubiquitously expressed, mutant huntingtin-specific changes could also be reflected by gene expression changes in blood. Several studies have identified HD blood mRNA changes using microarray technology but it has proven difficult to validate these across studies [7-9].

Advances in next generation sequencing offer new inroads to study the transcriptome. The digital nature of next-generation sequencing allows for accurate quantification of unknown transcripts, low- and high-abundance transcripts. Sequence-based methods allow the measurement of known as well as unknown transcripts, thus obviating the past limitation to the microarray content. In addition, sequence-based methods are more precise than microarrays and more robust across experiments because of much greater depth and the absence of the background signal and cross hybridization issues that were associated with microarrays [10]. One such method, the 3' digital expression profiling (DGE/DeepSAGE) creates 21 base pair sequences (tags) near the 3' ends of polyade-

nylated mRNAs [11] and uniquely identifies transcripts using these tags. Thus by counting the matching transcripts one can estimate differences in gene expression between samples across a large dynamic range. In comparison with full-length RNA sequencing, DeepSAGE has the advantage of comprehensive coverage of all (transcribed) genes at great depth, at the cost of not detecting different splice variants. In this study we investigated the suitability of blood to identify HD transcriptomic biomarkers, validated the outcome in an independent cohort and derived a first empiric panel of biomarkers capable of predicting HD motor scores. Finally, we examined whether patient gene expression profiles could provide information about HD affected biological pathways.

Material and Methods

Cohort assessment and characteristics

Peripheral blood from 33 controls, 27 presymptomatic mutation carriers and 64 symptomatic mutation carriers were collected for the discovery cohort and independent validation cohort from 12 symptomatic mutation carriers and 11 controls. Collection was done with IRB approval and after informed consent. All subjects were examined by an experienced neurologist using the motor section of the Unified Huntington's Disease Rating Scale (UHDRS) as described previously [12]. All the controls were free of known medical conditions. Age considered for the analysis was the age at the time of blood collection. For a detailed summary of the study cohort's average age, gender composition, UHDRS TMS and HD progression total functional capacity scores (TFC) see Supplementary Table S3.

RNA isolation and DeepSAGE library production

RNA isolation and cDNA library production was performed as described previously [13]. In short total RNA was extracted from PAXgene blood tubes (Qiagen, Venlo, The Netherlands) and 1 μg of total RNA was used to synthesize double-stranded cDNA constructs for next-generation sequencing.

Sequence processing

Illumina GA Pipeline (version 1.5.1) was used for data sequence processing. The FASTQ files were analyzed using the open source GAPSS_B pipeline (http://www.lgtc.nl/GAPSS) as described previously[13]. In addition, a custom Perl script was used to obtain gene annotations from BioMart and a custom python script was used to count the tags in each Ensembl gene using the SAM output files from "bowtie". To avoid batch effects, the samples were randomized during RNA isolation and DeepSAGE sample preparation. To identify potential sample swaps and contaminations all samples were checked for the correct expression of *XIST* and *RPS4Y1* gender specific genes. Batch effects were assessed using multidimensional scaling (MDS) plots for gender, sequencing flow cell and disease stage and by using the edgeR bioconductor package for RNA-Seq. The sequencing gene expression data used for this study have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE51799.

Fluidigm RT-qPCR

cDNA synthesis was performed using 1 µg of total RNA from each blood sample and using random hexamer primers with the Transcriptor First Strand cDNA synthesis kit (Roche). cDNA was diluted 4 times and 1.25 µl of each sample was preamplified using 2.5µl of 2x Taqman pre-amplification master mix (Applied Biosystems) and 1.25 µl of the primer pool (0.2pmol each primer/µl). The preamplifications were performed using a 10 minute 95 °C denaturation step and 14 cycles of 15 seconds at 95° C and 4 minutes at 60° C. The preamplified reactions were diluted 5x times in H_20. Five microliters from a sample mix containing preamplified cDNA and amplification Master mix (20mM $MgCl_2$, 10mM dNTPs, FastStart Taq polymerase, DNA binding Dye loading reagent, 50x ROX, 20x Evagreen) was loaded into each sample inlet of the 48.48 dynamic array chip (Fluidigm Corporation, San Francisco) and 5 µL from an assay mix containing DA assay loading reagent as well as forward and reverse primers (10pmol/µl) were loaded into each detector inlet. The chip was then placed on the NanoFlexTM 4-IFC Controller for loading and mixing. After loading the chip was placed on the BioMark™ Real-Time PCR System using a cycling program of 10 min at 95°C followed by 40 cycles of 95°C for 15 sec and 60°C for 30 sec and 72 °C for 30 sec. Data were analyzed using the Bio-

Mark Gene Expression Data Analysis software to obtain Ct values and/or ΔCt values. Fluidigm data were corrected for differences in input RNA using the geometric mean of 3 reference genes *ACTB*, *HPRT* and *RPL22*.The array accommodated reactions for all 48 validation samples and 23 genes in duplicate (duplicate values were averaged).

Statistical analysis

All DeepSAGE downstream analyses were performed at the gene level, and in case of multiple SAGE tags per gene, for example, as a consequence of alternative polyadenylation, tags were summarized. All the tag counts for a certain gene across all 124 samples were summarized. Low abundance genes with <124 tags were removed as were the top 3 overabundant genes (*HBA1*, *HBA2* and *HBB*). Gene expression analysis was performed using the "limma" package and the "voom" function for RNA-seq data and by applying linear modeling and empirical Bayes statistics [14]. The model tested gene expression as a function of the subject's total motor score (TMS), while accounting for gender, age and relative cell content (measured by the ratio of hemoglobin tags versus total aligned tags per sample) as confounders. Fluidigm expression analysis was performed using the linear modeling function in R and by testing the individual Δct expression values against the subject's TMS, while accounting for gender and age. Global test pathway analysis was performed using the same model as was used for the DeepSAGE analysis. For GO pathway analysis the top *P* value pathways that consisted of a minimum of 10 genes were reported. For IPA analysis the top 250 DeepSAGE genes were used (*P* value <0.001) For TMS prediction a linear regression model with a "lasso" penalty was fitted using the R package "penalized", optimizing the lasso tuning parameter using leave-one-out cross-validation [15].The effects of age and gender were not penalized.

Results

Gene expression analysis

Samples were sequenced at an average library size of 23.5 million tags. Alignment to the human genome resulted in an average library size of 20.4 million tags with at least one reported alignment (87.1%). A detailed description of the sequenced samples RNA integrity numbers (RIN) and sequence alignment characteristics can be seen in Supplementary Table S4. After removal of very low abundance genes we could reliably detect a total of 16.657 genes. To find HD-specific stage or progression biomarkers the DeepSAGE gene expression data were modeled as a function of the individual UHDRS total motor score (TMS); while accounting for gender, age and the percentage of hemoglobin tags (a proxy for the reticulocyte content) as confounders. The TRACK HD study has shown that in presymptomatic HD gene carriers the motor score scale (0-124) is a strong predictor of subsequent clinical conversion [6]. Our HD group consisted of 27 presymptomatic (TMS=2.4±1.8) and 64 symptomatic (TMS=37.4±24.3). After linear modeling,

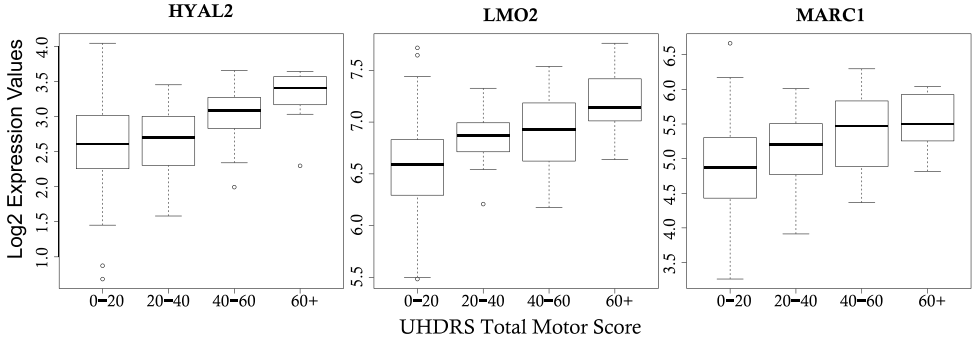


Figure 1. Boxplots of the DeepSAGE expression values for the top 3 upregulated genes discovered from linear modeling with TMS and for all 124 samples. The plot confirmed our linear modelling analysis and demonstrated a gradual increase in gene expression across the different total motor score groups.

a total of 167 genes significantly associated with motor score at an adjusted *P* value of 0.05 or less, suggesting that these constitute potential disease stage biomarkers. Of these 167 genes, 99 were positively associated with motor score and upregulated in HD sam-

ples compared with controls, while 68 were negatively associated and downregulated. The top 10 upregulated and top 10 downregulated genes are shown in Table 1. When we grouped the samples based on TMS we could confirm our linear modeling results. Boxplots for the top 3 upregulated genes showed a gradual increase in gene expression with increasing TMS (Figure 1). A full list of all the genes significantly associated with TMS as well as with total functional capacity score (TFC) disease staging are provided in Supplementary Tables S5 and S6 respectively (available online). Reassuringly in the TFC based analysis, 60% of the genes were the same as the TMS-based significant genes. Among the top TMS *P* value significant genes were genes involved in regulation of circadian rhythm such as prokineticin 2 (*PROK2*), genes associated with motor learning such as protein tyrosine phosphatase non receptor 4 (*PTPN4*) and genes implicated in the development of the brain cortex such as G protein coupled receptor 56 (*GPR56*) [16-18]. The genes with the biggest expression change but lacking statistical significance were the small nuclear RNA host gene 9 (*SNGH9*) and the major histocompatibility complex class II DQ alpha1 gene (*HLA-DQA1*). *HLA-DQA1* has been previously reported as a candidate RNA biomarker in human lymphocyte microarray data from HD patients, ranking among the top most changed genes [9]. The highest expressed significant gene was S100 calcium binding protein A9 (*S100A9*) with a log2 expression value of 11.7 while the lowest expressed significant gene was sperm acrosome associated 3 (*SPACA3*) with a value of -2.8, indicative of the high dynamic range of the sequencing platform ($2^{11.7 - (-2.8)}$ =23170 fold).

Using EBI Gene Expression Atlas (http://www.ebi.ac.uk/gxa/) and literature searches, we found that 40 of the 167 genes had been previously reported as differentially expressed in at least one HD gene expression study with the same direction in expression change. These included mechanistic target of rapamycin (*MTOR*), a potential target for therapy in HD, H2A histone family member Y (*H2AFY*), a gene whose transcript levels have been recently reported to mark HD activity in human and mouse, CDC-like kinase 3 (*CLK3*) another gene from the top 99 genes from the previous study, and aquaporin 9 (*AQP9*), a gene that has been described as a potential biomarker in blood [8,9,19].

Table 1. DeepSAGE top 10 upregulated (coefficient +) and downregulated (coefficient -) genes in HD blood samples.

Gene	Description	Coeff.[a]	Expression[b]	Adj.P-Val	Protein Function
HYAL2	Hyaluronoglucosaminidase 2	+0.4	2.6	1.0E-03	Hydrolyzes hyaluronic acid
LMO2	LIM domain only 2	+0.3	6.6	1.0E-03	Yolk sac hematopoiesis
MARC1	Mitochondrial amidoxime reducing C1	+0.4	5.0	5.0E-03	N-hydroxylate prodrug conversion
NT5DC2	5'-Nucleotidase domain containing 2	+0.4	2.8	9.0E-03	Hydrolase and metal ion binding
RNF135	Ring finger protein 135	+0.3	5.8	9.0E-03	DDX58 Ubiquitination ~ IFN-β
PROK2	Prokineticin 2	+0.5	7.9	1.0E-02	Circadian clock – GI contraction
RPN1	Ribophorin I	+0.3	5.5	1.0E-02	26S Proteasome ubiquitin binding
CYSTM1	Cysteine-rich transmembrane module 1	+0.4	6.0	1.0E-02	Stress tolerance
VCAN	Versican	+0.3	8.2	1.6E-02	Intercellular signal.-Binds hyal. acid
NCF4	Neutrophil cytosolic factor 4	+0.3	8.9	1.8E-02	NADPH-oxidase component
ARL4C	ADP-Ribosylation factor-like 4C	-0.3	8.2	1.0E-03	Microtubule vesicular transport
TMEM109	Transmembrane protein 109 (Mg23)	-0.3	7.0	6.0E-03	UVC αB-Crystallin protection
MACF1	Microtubule-actin crosslinking factor 1	-0.2	7.2	6.0E-03	Actin-microtubule stabilization
MDN1	Midasin homolog	-0.2	5.3	7.0E-03	AAA-ATPase(dynein)
PTPN4	Protein tyrosine phosphatase NR type 4	-0.3	5.1	9.0E-03	Glutamate receptor signaling
PRF1	Perforin 1	-0.4	9.5	1.0E-02	Cytolysis
CD3G	CD3g Molecule gamma	-0.3	7.5	1.0E-02	CD3 complex signal transduction
NMT2	N-Myristoyltransferase 2	-0.3	3.4	1.0E-02	N-terminal Myristoylation
KLRD1	Killer cell lectin receptor subfamily D1	-0.4	6.1	1.0E-02	Recognition of MHC classI HLA-E
GPR56	G Protein-coupled receptor 56	-0.4	7.3	1.0E-02	Brain cortical patterning

[a] Coefficients of gene expression change per motor score unit multiplied by average motor score. [b] Average $\log_2$ gene expression levels. Protein function based on Genecards.

Global test pathway analysis

To elucidate affected biological pathways in HD blood that were associated with TMS, we used the "Global test" bioconductor package [20]. We included KEGG pathways, GO terms and predicted / validated target genes of miRNAs (Broad-GSEA). In the KEGG pathway analysis (see Supplementary Table S7) we found terms frequently reported in HD and neurodegenerative disorder pathway analyses such as neuroactive ligand receptor interaction, amyotrophic lateral sclerosis and long term depression. We also found less common terms such as the pentose phosphate pathway (PPP), Jak-STAT signaling and type II diabetes mellitus. The genes that contributed most to PPP were glucose phosphate isomerase (GPI), aldolase A (*ALDOA*), phosphogluconolactonase (*PGLS*) and transketolase-like 1 (*TKTL1*) an enzyme linking PPP with the glycolytic pathway. Mitochondria-associated metabolic dysfunction and increased glycolytic rate have been previously associated with HD [21]. The Jak-STAT pathway, a common signaling pathway used by many cytokines was characterized by the upregulation of serine-threonine protein kinase (*AKT1*), suppressor of cytokine signaling 3 (*SOCS3*), son of sevenless homolog 2 (*SOS2*) and interferon-alpha/beta receptor beta chain (*IFNAR2*). Finally, for diabetes for which an increased frequency in HD patients has been previously described, the most significant genes were *MTOR* and protein kinase C delta (*PRKCD*) [22]. In the GO analysis we identified terms such as NADP binding, positive regulation of interleukin 6 production and response to cholesterol. The most significant genes for NADP binding were neuronal nitric oxide synthase 1 (*NOS1*), flavin containing monooxygenase 4 (*FMO4*) and homocysteine methyltransferase reductase (*MTRR*). The deregulation of genes linked to response to cholesterol could also be important since cholesterol biosynthesis has been shown to be impaired in HD cells, while Leoni et al. have demonstrated that 24OHC, a brain cholesterol turnover marker, correlated with disease progression [23]. All the genes reported for response to cholesterol can be seen in a Global test covariate plot in Supplementary Figure S5A. This result was also in agreement with Chou et al. who showed that the mutant Htt protein suppresses the secretion of *CCL5* [24]. The analysis for enrichment of target genes of miRNAs showed enrichment of miR-138 and miR-218 targets. These miRNAs were found downregulated in YAC128 and R6/2 HD mouse models [25]. For the miR-138 and miR-218 target genes, a separate

enrichment analysis, using DAVID (http://david.abcc.ncifcrf.gov), showed that terms enriched specifically for miR-138 target genes were histone modification and axon guidance while terms enriched specifically for miR-218 target genes were ubiquitin like conjugation, proto-oncogene and mental retardation. Other potentially interesting miRNAs that were identified previously were miR-18a, miR-504, miR-337 and miR-492 [26],[27]. To further validate our "Global test" pathway analysis results and obtain a better visual representation of the interconnections of the genes involved in the above biological processes we also analyzed our data through the use of the Ingenuity Pathway Analysis (IPA) (Ingenuity® Systems, www.ingenuity.com). Top diseases and functions reported by IPA network analysis were nervous system development, skeletal and muscular disorders but also immune cell trafficking and inflammatory response (Supplementary Table S8). The gene network plot for the genes and molecules involved in the IPA network 6 and for skeletal and muscular disorders, connective tissue disorder and cancer are shown in Supplementary Figure S5B. Interestingly this gene plot interconnected terms such as histones, 26s proteasome, pro-inflammatory cytokines, Hsp70 and insulin; all of which have previously been implicated in HD. Canonical pathway analysis using IPA further confirmed our initial "Global test" results since common pathways reported were those of diabetes mellitus, Toll-like receptor and T cell receptor signaling. Finally, upstream regulators from our top genes were reported to be IL-2, IL-6 and IL-12 (complex) by IPA analysis, which was also in good correlation with the Global test analysis.

Validation

To validate the DeepSAGE gene expression results, we performed nanoliter RT-qPCR using the Fluidigm Biomark microfluidics chip [28], using 25 samples from the original discovery cohort as technical validation, supplemented with 23 patient and control samples as a biological validation in an independent cohort. Twenty genes in total, all from our DeepSAGE list of 167 significantly differentially expressed genes, were examined; the top 12 based on *P* value, 6 further down the 167 gene list based on differential expression in previous HD studies (*H2AFY*, *AQP9*, *ANXA3*, *RGS14*, *ZNF238*, *NOL3*) and another 2 genes based on possible involvement in HD pathology (*CEBPA*, *TAF15*) [3,8,9,19,29-31]. Fluidigm data were analyzed using a linear model as a function of TMS, while accounting for gender and age. In the basic validation cohort, 12 out of

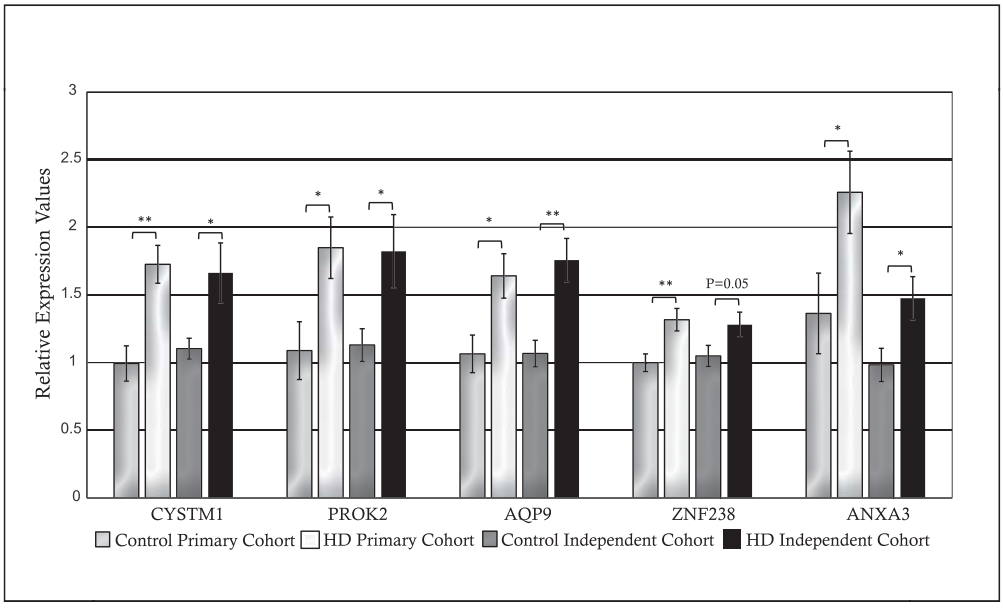


Figure 2. Relative expression of the most significant Fluidigm RT-qPCR genes across the two independent cohorts for controls and HD patients. Asterisks represent statistical significance from a student's t-test (* P<0.05, ** P<0.01). Error bars represent SEM values.

the 20 genes tested were significantly associated with TMS, while in the independent validation cohort, 7 out of the 20 genes were significant (see Table 2). Most other genes, while not reaching significance, showed trends in the same direction as in the discovery cohort. Five of the 20 genes (*PROK2*, *ZNF238*, *AQP9*, *CYSTM1* and *ANXA3*) were the most robust and significantly associated with TMS in both the discovery and the independent cohort. The intergroup relative expression levels of these 5 genes across HD versus control samples, irrespective of TMS, can also be seen in Figure 2. Finally, when the linear modeling analysis was performed on all Fluidigm samples (n=48), we were able to validate 12 of the 20 genes tested (see rightmost column of Table 2).

Biomarker motor score prediction

To evaluate which panel of genes would optimally predict TMS, we fitted a linear regression model with a "lasso" penalty using the Fluidigm expression data, age and gender as predictors and TMS as the response. The gene expression values of three genes

Table 2. Fluidigm RT-qPCR technical and biological validation results of DeepSAGE genes

Gene	Description	Discovery Cohort (n=25)		Independent Cohort (n=23)		All Samples (n=48)
		P value	Coeff.*	P value	Coeff.	P value
CYSTM1	Cysteine-rich transmembrane module 1	6.0E-03	0.5	2.0E-03	0.5	1.0E-04
PROK2	Prokineticin 2	1.0E-02	1.0	2.0E-03	0.7	1.0E-03
AQP9	Aquaporin 9	2.0E-03	0.5	6.0E-05	0.7	2.0E-05
ZNF238	Zinc finger protein 238	2.0E-02	0.3	8.0E-03	0.5	1.0E-03
ANXA3	Annexin 3	4.0E-02	0.5	6.0E-03	0.7	7.0E-04
RNF135	Ring finger protein 135	7.0E-02	0.2	7.0E-03	0.2	6.0E-03
LMO2	LIM domain only 2	3.0E-02	0.2	7.0E-02	0.2	6.0E-03
ARL4C	ADP-ribosylation factor like 4	3.0E-02	-0.3	9.0E-01	0.02	7.0E-03
TMEM109	Transmembrane protein 109	4.0E-02	-0.2	5.0E-01	-0.04	1.0E-02
CEBPA	CCAAT/enhancer binding A	2.0E-02	0.4	1.0E-01	0.2	1.0E-02
MACF1	Microtubule-actin crosslinking F1	2.0E-02	-0.3	6.0E-01	-0.04	1.0E-02
PTPN4	Protein tyrosine phosphatase NR 4	1.0E-02	-0.35	6.0E-01	-0.04	3.0E-02
MARC1	Mitochondrial amidoxime reducing C1	2.0E-02	0.5	1.0E-01	0.4	6.0E-02
H2AFY	H2A histone family, member Y	1.0E-01	0.15	1.0E-01	0.15	1.0E-01
HYAL2	Hyaluronoglucosaminidase 2	1.0E-01	0.20	4.0E-02	0.2	1.0E-01
NOL3	Nucleolar protein 3	2.0E-01	0.15	8.0E-02	0.2	2.0E-01
MDN1	Midasin homolog	2.0E-01	-0.15	9.0E-01	NC**	1.0E-01
NT5DC2	5'-Nucleotidase domain containing 2	5.0E-01	0.1	2.0E-01	0.4	3.0E-01
RGS14	Regulator of G-protein signaling 14	1.0E-01	0.2	5.0E-01	0.2	3.0E-01
TAF15	TATA box - associated factor	1.0E-01	-0.1	2.0E-01	0.15	2.0E-01

*Coeff. = Coefficients of gene expression change per motor score unit multiplied by group average motor score. **NC = No change.

(*AQP9*, *ANXA3* and *ARL4C*), together with age and gender, were the best predictors of TMS. The last gene (*ARL4C*) was non-significantly downregulated in HD blood and specifically served the purpose of enlarging the "biomarker chip" set towards tolerance for smaller individual gene changes, providing additional informativeness. The results of the cross-validated prediction analysis can be seen in Figure 3. The prediction model performed better for earlier disease stages (Stage I, II), while it was less accurate for later stages (Stage III-V) and especially for patients with a motor score of 50 points and over. Only one patient was assigned a predicted TMS greater than 50 points (patient no.29). This patient was the oldest HD carrier (>70 years). We also observed that for one patient the blood-based signature indicated a higher predicted motor score compared to the clinical motor score. This could be explained by the fact that this patient had a much lower TFC score (TFC=4) compared to other patients with similar motor score, indicative of a more advanced disease stage. Finally, the control sample with the highest clinical motor score (control no.4) was our oldest control sample (69 years) and also received a higher predicted score. When we plotted the DeepSAGE gene expression levels of these 3 genes across the controls, the presymptomatic carriers and the different HD TFC-based disease stages we could confirm that for *ANXA3* and *AQP9* there was an increase in gene expression even in the presymptomatic stage. For *ARL4C*, contrary to *ANXA3* and *AQP9* there was a decrease in gene expression, the expression changes were more prominent in the more advanced disease stages and hence provided complementary information to the other two genes (Figure 4). On the basis of this analysis, we formulated a TMS predictive equation to measure the disease stage based on gene expression of the 3 genes (see below Figure 4).

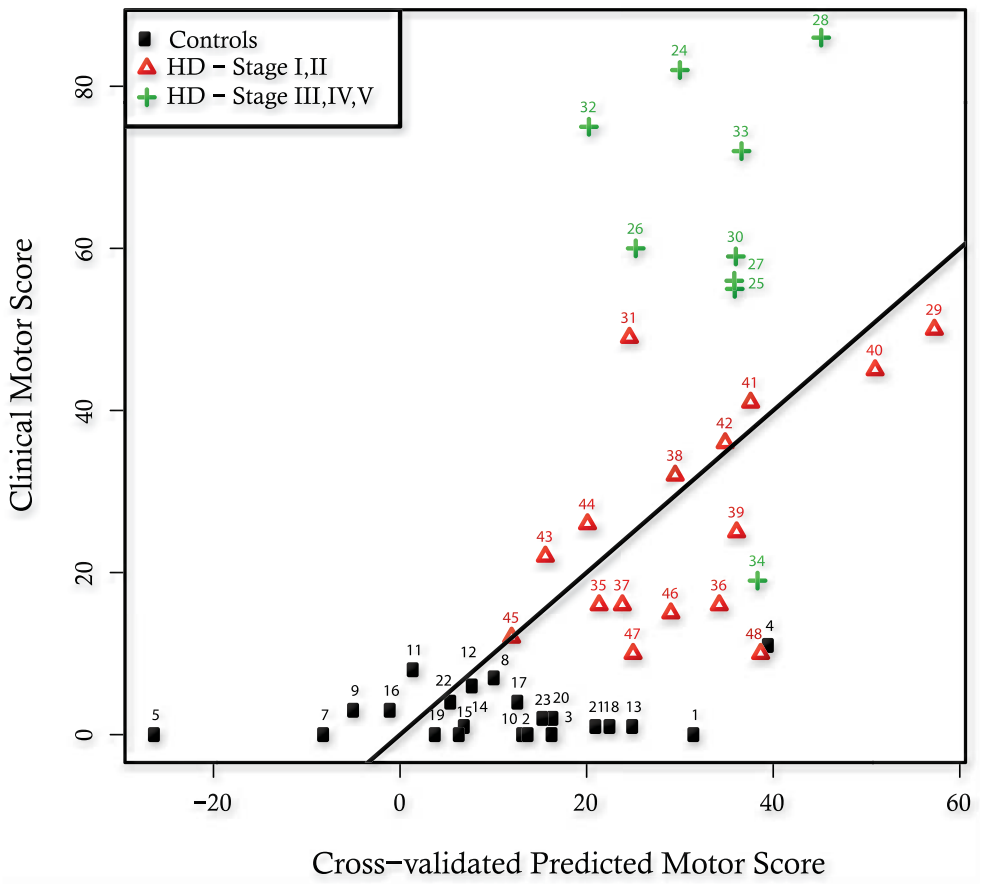
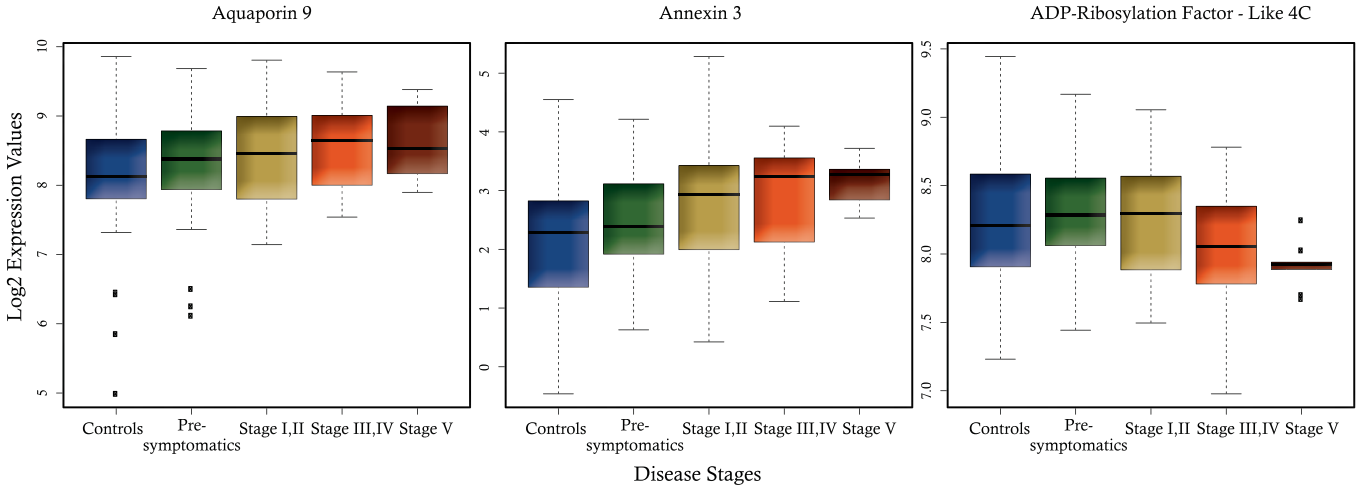


Figure 3. Plot of clinical TMS against cross-validated predicted TMS based on Fluidigm RT-qPCR gene expression data. The cross-validated motor score is predicted for each subject by a model trained on a data set in which the subject itself was left out. Stage classification was based on total functional capacity (TFC) scores (Stage I, II = TFC score 7-13/ Stage III-V=TFC score 0-7).

Figure 4. DeepSAGE gene expression levels for the best three TMS predictor genes as these were reported from the LASSO algorithm prediction analysis and across controls, presymptomatics and different HD symptomatic stages.



Total motor score predictive equation:

$$Ptms = \frac{(-1.476 * Gender + 0.964 * Age - 8.868 * X_{ARL4C} + 3.801 * X_{ANXA3} + 14.159 * X_{AQP9})}{-33.59}$$

Where Ptms is the gene expression predicted TMS, gender is 0 or 1 for male or female and X the respective, housekeeping gene corrected, ΔCt RT-qPCR value

Discussion

To date, thousands of disease biomarkers have been published while <100 have been validated in independent cohorts [32]. This inability to validate disease biomarkers has been attributed to the lack of large enough study cohorts as well as standardization in sample collection and storage [33]. For HD, validation has been even more challenging as the disease presents itself through a variety of symptoms and progression rates. For these reasons we performed gene expression profiling, taking advantage of the sensitivity of next-generation sequencing and Fluidigm technologies, and our experience in standardized blood collection and sample analysis[34,35]. Using the UHDRS TMS as a clinical parameter we identified a set of 167 genes differentially expressed in HD blood. Furthermore, we validated our findings by a targeted approach, using an entirely different technology. Technical validation (in the same cohort) confirmed 12/20 of the discovered genes and biological validation (in a different cohort) confirmed 7/20 of the discovered genes in a different cohort. Our discovery and validation cohorts (n=124 and n=48) are to our knowledge among the largest to have been used in HD gene expression studies. In contrast to previous studies, we have selected a sizable group of 20 genes for validation in duplicate (~2300 reactions). Indeed, the very fact that so many of the top 20 discovered genes can be validated argues in favor of the robustness of the discovery approach. Genes with more variation or smaller changes in principle are more difficult to validate in a small cohort. Yet, we should stress that these biomarkers presently constitute a candidate biomarker set that requires further validation in other HD cohorts before further used in a clinical setting.

The Fluidigm qPCR analysis yielded a panel of 5 genes (*PROK2*, *ZNF238*, *AQP9*, *CYSTM1* and *ANXA3*) as a potential HD biomarker set and this was validated in both the original cohort and an independent validation cohort. *PROK2* is expressed in the suprachiasmatic nucleus (SCN) and has been proposed to have a role in the regulation of circadian rhythms [17]. Circadian rhythm alterations have been shown to correlate with cognitive impairment in HD [36] and in HD models pharmacological imposition of sleep slows cognitive decline and reverses deregulation of *PROK2* [37]. As a blood marker of HD progression *PROK2* is very promising, since this could also be reflecting

brain changes. *ZNF238* is a transcriptional repressor involved in brain development and myogenesis [38] and increasing evidence suggests that gene repression mechanisms are associated with HD [39,40]. This is in agreement with the reported involvement in HD of *H2AFY*, which is also involved in transcriptional repression and further studies link HD with SP1, another zinc finger protein [41]. Aquaporins are water selective channels with possible roles in the nervous system and expression levels were upregulated after brain injury [42]. The presence of *AQP9* in blood could represent peripheral or central inflammatory events since a recent gene expression study showed that the mRNA levels of *AQP9* and four other genes can discriminate patients with chronic inflammation from controls [43]. *CYSTM1* is a relatively unknown gene and bioinformatics analysis has demonstrated a role in stress response and confer tolerance to heavy metals such as cadmium and copper [44]. Finally *ANXA3* was upregulated in two neuronal injury models [45],[46]. It is important to note that the levels of annexin *ANXA1* have also been found upregulated in a previous gene expression study in HD blood [8].

Our pathway analysis showed a wide range of processes changed in HD. The most prominent terms pointed towards the involvement of the immune system. It has been suggested by previous studies that pro-inflammatory cytokines such as IL-6, IL-8, and TNF-α can be used as peripheral HD biomarkers [47,48]. Other terms such as diabetes mellitus could also be interesting since mouse models of HD can develop diabetes mellitus [49], and it was shown that type II diabetes exhibits common features with other neurodegenerative disorders [50]. Finally, we discovered enrichment of target genes of miRNAS (miR-138/218) previously reported to be downregulated in HD models. This warrants further investigation as miR-9 was found to be downregulated in human HD brain samples and target complexes, such as REST, that regulates neuronal gene expression in non-neuronal tissues [51]. A disadvantage of whole blood may be considered its cellular heterogeneity. The more informative white blood cells comprise a small percentage of the total cell population, while 95% of blood consists of red blood cells, with hemoglobin transcript percentages, ranging from 30-90%. This could well account for the fact that until now less sensitive techniques failed to replicate results between different HD blood microarray studies [52]. For the same reason, in the past, most expression studies used isolated peripheral blood mononuclear cells. However, it

is not always possible to process samples directly after collection and preparation delays have been shown to induce biases [53]. In the present study, taking advantage of the digital nature of sequencing, we identified differentially expressed genes across a wide dynamic range, with high sensitivity, directly from whole blood. This provides a clearer image of the transcriptional alterations in HD; although biomarkers with higher expression will be more useful and easier to detect with less sensitive routine techniques. Our motor score prediction analysis showed that the gene expression predictive power was stronger for early-stage and weaker for later-stage patients. While this could be explained by the increasing impact of generalized tissue degeneration in late disease stages, the increased reliability in earlier stages is in fact of major benefit as, notably in this early phase, robust therapeutic readouts are challenging. Furthermore, previous gene expression studies have found small individual gene expression changes in HD blood. In the future and for a potential "biomarker chip" to survive further validation, a larger group of genes may be required that will better allow for variation in individual gene expression changes. For this reason, we used the predictive capabilities of the LASSO algorithm to see which genes would jointly perform most optimally in UHDRS TMS prediction. The formula we have derived links a small set of easily definable gene expression levels to the UHRDS Total Motor Score, and is thus a promising candidate biomarker set to monitor disease state, progression and putative therapeutic effect of interventions. Taking into account the great symptomatic variability in HD patients, different sets of biomarkers can be further trained and optimized, depending on the disease stage that is most prominent in the group of patients included in each study.

Considering the complexity of HD most likely a collection of biomarkers will best define disease progression and response to therapy. The biomarker changes found in this study monitor disease progression in blood and may be relatively independent of the changes taking place in brain. Such biomarkers, if validated clinically, may be useful as surrogate markers to test the effectiveness of therapeutic strategies even when they may not have a robust relationship with actual clinical endpoints [54].

Owing to the design of our study, comparing various HD stages with unaffected controls, we cannot exclude that the detected changes might also (partly) track progression of other neurodegenerative diseases. Thus, prior to putative diagnostic application, this

needs to be further assessed. However, this does not reduce the potential differentiating significance of this biomarker panel for prognostic application in a known (pre)symptomatic HD carrier setting.

In conclusion, we describe the development of a panel of candidate HD biomarkers which can be easily measured by transcript analysis of whole blood and which may have application in disease staging and the monitoring of therapeutic effectiveness. Longitudinal and cross sectional studies in additional cohorts will be needed to further validate this panel before its application in the clinic. Finally, the assessment of the disease relevance of the genes involved may well contribute to finding new HD therapeutic targets.

Acknowledgements

This work was supported by the European Commission 7th Framework Program, [Project no. 261123]; GEUVADIS, [Project no. 201413]; ENGAGE by the Centre for Medical Systems Biology within the framework of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research and Dutch Centre for Biomedical Genetics and the European Community's Seventh Framework Programme (FP7/2007-2013) [grant n° 2012-305121] "Integrated European -omics research project for diagnosis and therapy in rare neuromuscular and neurodegenerative diseases (NEUROMICS)".The authors thank Rolf Vossen (Department of Human Genetics, LUMC) for his assistance with the Fluidigm experiments, Judith Boer (Laboratory of Pediatric Oncology/ Hematology, Erasmus Medical Centre) for her assistance with IPA and Marloes Hubers (Department of Psychiatry, LUMC) for her assistance with the clinical data.

Conflicts of Interest

The authors declare that they have not conflict of interests in relation to the work described.

Reference List

1. van der Burg JM, Bjorkqvist M, Brundin P: Beyond the brain: widespread pathology in Huntington's disease. Lancet Neurol 2009, 8: 765-774.
2. Trager U, Tabrizi SJ: Peripheral inflammation in neurodegeneration. J Mol Med (Berl) 2013.
3. Hodges A, Strand AD, Aragaki AK, Kuhn A, Sengstag T, Hughes G et al.: Regional and cellular gene expression changes in human Huntington's disease brain. Hum Mol Genet 2006, 15: 965-977.
4. Runne H, Regulier E, Kuhn A, Zala D, Gokce O, Perrin V et al.: Dysregulation of gene expression in primary neuron models of Huntington's disease shows that polyglutamine-related effects on the striatal transcriptome may not be dependent on brain circuitry. J Neurosci 2008, 28: 9723-9731.
5. Luthi-Carter R, Cha JHJ: Mechanisms of transcriptional dysregulation in Huntington's disease. Clinical Neuroscience Research 2003, 3: 165-177.
6. Tabrizi SJ, Scahill RI, Owen G, Durr A, Leavitt BR, Roos RA et al.: Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. Lancet Neurol 2013.
7. Lovrecic L, Kastrin A, Kobal J, Pirtosek Z, Krainc D, Peterlin B: Gene Expression Changes in Blood as a Putative Biomarker for Huntington's Disease. Mov Disord 2009, 24: 2277-2281.
8. Borovecki F, Lovrecic L, Zhou J, Jeong H, Then F, Rosas HD et al.: Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. Proc Natl Acad Sci U S A 2005, 102: 11023-11028.
9. Runne H, Kuhn A, Wild EJ, Pratyaksha W, Kristiansen M, Isaacs JD et al.: Analysis of potential transcriptomic biomarkers for Huntington's disease in peripheral blood. Proc Natl Acad Sci U S A 2007, 104: 14424-14429.
10. Hoen PAC, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RHAM, de Menezes RX et al.: Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Res 2008, 36.
11. Nielsen KL, Hogh AL, Emmersen J: DeepSAGE - digital transcriptomics with high

sensitivity, simple experimental protocol and multiplexing of samples. Nucleic Acids Res 2006, 34.

12. van DE, Kingma EM, Timman R, Zitman FG, Tibben A, Roos RA et al.: Cross-sectional study on prevalences of psychiatric disorders in mutation carriers of Huntington's disease compared with mutation-negative first-degree relatives. J Clin Psychiatry 2008, 69: 1804-1810.
13. Mastrokolias A, den Dunnen JT, van Ommen GB, 'T Hoen PAC, van Roon-Mom WMC: Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA. Bmc Genomics 2012, 13.
14. Law CW, Chen Y, Shi W, Smyth GK: voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014, 15: R29.
15. Goeman JJ: L1 penalized estimation in the Cox proportional hazards model. Biom J 2010, 52: 70-84.
16. Li SH, Jin ZH, Koirala S, Bu LH, Xu L, Hynes RO et al.: GPR56 regulates pial basement membrane integrity and cortical lamination. Journal of Neuroscience 2008, 28: 5817-5826.
17. Cheng MY, Bullock CM, Li CY, Lee AG, Bermak JC, Belluzzi J et al.: Prokineticin 2 transmits the behavioural circadian rhythm of the suprachiasmatic nucleus. Nature 2002, 417: 405-410.
18. Kina S, Tezuka T, Kusakawa S, Kishimoto Y, Kakizawa S, Hashimoto K et al.: Involvement of protein-tyrosine phosphatase PTPMEG in motor learning and cerebellar long-term depression. Eur J Neurosci 2007, 26: 2269-2278.
19. Hu Y, Chopra V, Chopra R, Locascio JJ, Liao ZX, Ding HL et al.: Transcriptional modulator H2A histone family, member Y (H2AFY) marks Huntington disease activity in man and mouse. Proc Natl Acad Sci U S A 2011, 108: 17141-17146.
20. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC: A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004, 20: 93-99.
21. Ferreira IL, Cunha-Oliveira T, Nascimento MV, Ribeiro M, Proenca MT, Januario C et al.: Bioenergetic dysfunction in Huntington's disease human cybrids. Exp Neurol 2011, 231: 127-134.
22. Podolsky S, Leopold NA, Sax DS: Increased frequency of diabetes mellitus in pa-

tients with Huntington's chorea. Lancet 1972, 1: 1356-1358.
23. Leoni V, Mariotti C, Nanetti L, Salvatore E, Squitieri F, Bentivoglio AR et al.: Whole body cholesterol metabolism is impaired in Huntington's disease. Neurosci Lett 2011, 494: 245-249.
24. Chou SY, Weng JY, Lai HL, Liao F, Sun SH, Tu PH et al.: Expanded-polyglutamine huntingtin protein suppresses the secretion and production of a chemokine (CCL5/RANTES) by astrocytes. Journal of Neuroscience 2008, 28: 3277-3290.
25. Lee ST, Chu K, Im WS, Yoon HJ, Im JY, Park JE et al.: Altered microRNA regulation in Huntington's disease models. Exp Neurol 2011, 227: 172-179.
26. Huang WH, Li MD: Differential Allelic Expression of Dopamine D1 Receptor Gene (DRD1) Is Modulated by microRNA miR-504. Biol Psychiatry 2009, 65: 702-705.
27. Roshan R, Ghosh T, Scaria V, Pillai B: MicroRNAs: novel therapeutic targets in neurodegenerative diseases. Drug Discov Today 2009, 14: 1123-1129.
28. Spurgeon SL, Jones RC, Ramakrishnan R: High Throughput Gene Expression Measurement with Real Time PCR in a Microfluidic Dynamic Array. Plos One 2008, 3.
29. van Roon-Mom WMC, Pepers BA, 'T Hoen PAC, Verwijmeren CACM, den Dunnen JT, Dorsman JC et al.: Mutant huntingtin activates Nrf2-responsive genes and impairs dopamine synthesis in a PC12 model of Huntington's disease. BMC Mol Biol 2008, 9.
30. Chiang MC, Chen HM, Lee YH, Chang HH, Wu YC, Soong BW et al.: Dysregulation of C/EBP alpha by mutant Huntingtin causes the urea cycle deficiency in Huntington's disease. Hum Mol Genet 2007, 16: 483-498.
31. King OD, Gitler AD, Shorter J: The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res 2012, 1462: 61-80.
32. Poste G: Bring on the biomarkers. Nature 2011, 469: 156-157.
33. Ransohoff DF, Gourlay ML: Sources of Bias in Specimens for Research About Molecular Markers for Cancer. Journal of Clinical Oncology 2010, 28: 698-704.
34. 't Hoen PA, Friedlander MR, Almlof J, Sammeth M, Pulyakhina I, Anvar SY et al.: Reproducibility of high-throughput mRNA and small RNA sequencing across laboratories. Nat Biotechnol 2013, 31: 1015-1022.
35. Lappalainen T, Sammeth M, Friedlander MR, 't Hoen PA, Monlong J, Rivas MA

et al.: Transcriptome and genome sequencing uncovers functional variation in humans. Nature 2013, 501: 506-511.

36. Aziz NA, Anguelova GV, Marinus J, Lammers GJ, Roos RAC: Sleep and circadian rhythm alterations correlate with depression and cognitive impairment in Huntington's disease. Parkinsonism & Related Disorders 2010, 16: 345-350.
37. Pallier PN, Maywood ES, Zheng ZG, Chesham JE, Inyushkin AN, Dyball R et al.: Pharmacological imposition of sleep slows cognitive decline and reverses dysregulation of circadian gene expression in a transgenic mouse model of huntington's disease. Journal of Neuroscience 2007, 27: 7869-7878.
38. Hirai S, Miwa A, Ohtaka-Maruyama C, Kasai M, Okabe S, Hata Y et al.: RP58 controls neuron and astrocyte differentiation by downregulating the expression of Id1-4 genes in the developing cortex. EMBO J 2012, 31: 1190-1202.
39. Zhai W, Jeong H, Cui L, Krainc D, Tjian R: In vitro analysis of huntingtin-mediated transcriptional repression reveals multiple transcription factor targets. Cell 2005, 123: 1241-1253.
40. Zuccato C, Belyaev N, Conforti P, Ooi L, Tartari M, Papadimou E et al.: Widespread disruption of repressor element-1 silencing transcription factor/neuron-restrictive silencer factor occupancy at its target genes in Huntington's disease. J Neurosci 2007, 27: 6972-6983.
41. Qiu ZH, Norflus F, Singh B, Swindell MK, Buzescu R, Bejarano M et al.: Sp1 is up-regulated in cellular and transgenic models of Huntington disease, and its reduction is neuroprotective. J Biol Chem 2006, 281: 16672-16680.
42. Shenaq M, Kassem H, Peng CY, Schafer S, Ding JY, Fredrickson V et al.: Neuronal damage and functional deficits are ameliorated by inhibition of aquaporin and HIF1 alpha after traumatic brain injury (TBI). J Neurol Sci 2012, 323: 134-140.
43. Mesko B, Poliska S, Szegedi A, Szekanecz Z, Palatka K, Papp M et al.: Peripheral blood gene expression patterns discriminate among chronic inflammatory diseases and healthy controls and identify novel targets. Bmc Medical Genomics 2010, 3.
44. Venancio TM, Aravind L: CYSTM, a novel cysteine-rich transmembrane module with a role in stress tolerance across eukaryotes. Bioinformatics 2010, 26: 149-152.
45. Konishi H, Namikawa K, Kiyama H: Annexin III implicated in the microglial response to motor nerve injury. Glia 2006, 53: 723-732.

46. Chong KWY, Chen MJ, Koay ESC, Wong BS, Lee AYW, Russo-Marie F et al.: Annexin A3 is associated with cell death in lactacystin-mediated neuronal injury. Neurosci Lett 2010, 485: 129-133.
47. Bjorkqvist M, Wild EJ, Thiele J, Silvestroni A, Andre R, Lahiri N et al.: A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington's disease. J Exp Med 2008, 205: 1869-1877.
48. Dalrymple A, Wild EJ, Joubert R, Sathasivam K, Bjorkqvist M, Petersen A et al.: Proteomic profiling of plasma in Huntington's disease reveals neuroinflammatory activation and biomarker candidates. Journal of Proteome Research 2007, 6: 2833-2840.
49. Bjorkqvist M, Fex M, Renstrom E, Wierup N, Petersen A, Gil J et al.: The R6/2 transgenic mouse model of Huntington's disease develops diabetes due to deficient beta-cell mass and exocytosis. Hum Mol Genet 2005, 14: 565-574.
50. Han WP, Li C: Linking type 2 diabetes and Alzheimer's disease. Proc Natl Acad Sci U S A 2010, 107: 6557-6558.
51. Packer AN, Xing Y, Harper SQ, Jones L, Davidson BL: The Bifunctional microRNA miR-9/miR-9*Regulates REST and CoREST and Is Downregulated in Huntington's Disease. Journal of Neuroscience 2008, 28: 14341-14346.
52. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA et al.: Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci U S A 2003, 100: 1896-1901.
53. Debey S, Schoenbeck U, Hellmich M, Gathof BS, Pillai R, Zander T et al.: Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. Pharmacogenom J 2004, 4: 193-207.
54. Cohn JN: Introduction to surrogate markers. Circulation 2004, 109: IV20-IV21.

Supplementary Table S3. Cohort characteristics of samples used for DeepSAGE gene expression profiling and Fluidigm RT-qPCR validation analysis.

Groups	HD state	Gender	n	Age Mean (Range)	TMS[a] Mean (Range)	TFC[b] Mean (Range)
DeepSAGE Primary Cohort n=124	Controls n=33	Female	18	44 (25-68)	1.9 (0-11)	12.7 (11-13)
		Male	15	39 (23-58)	2 (0-12)	12.9 (12-13)
	Presymptomatics n=27	Female	17	41 (26-57)	2.6 (0-5)	11.7 (5-13)
		Male	10	43 (31-55)	2.1 (0-5)	12.8 (11-13)
	Symptomatics n=64	Female	32	52 (22-76)	40.5 (7-107)	5.9 (0-13)
		Male	32	55 (32-74)	34.2 (7-86)	6.9 (0-13)
Fluidigm Discovery Cohort n=25	Controls n=11	Female	6	46 (25-68)	3.3 (0-11)	12.5 (11-13)
		Male	5	42 (35-55)	2.2 (0-8)	12.8 (12-13)
	Carriers n=14	Female	6	46 (22-54)	46.5 (13-82)	5.8 (2-12)
		Male	8	62 (37-73)	53.6 (16-86)	4.9 (0-13)
Fluidigm Independent Cohort n=23	Controls n=12	Female	6	46 (38-55)	1.8 (0-4)	13
		Male	6	48 (35-58)	2.3 (0-6)	13
	Carriers n=11	Female	7	52 (33-61)	26.6 (10-45)	9.9 (7-13)
		Male	4	52 (40-61)	22.0 (10-41)	10.3 (7-13)

[a] Total Motor Score (0-124) of the Unified Huntington's Disease Rating Scale, with higher scores indicating more motor symptoms.[b] Total Functional Capacity of the Unified Huntington's Disease Rating Scale, with lower scores indicating worse functional capacity

Supplementary Table S4. Sample RNA and sequencing quality characteristics for all 124 samples used for DeepSAGE gene expression profiling.

Average RNA Quality (Range 1-10)	8.4 (7 - 9.6)
Average Number of Total Reads Sequenced	23.5 Million (11 – 37)
Average Percentage of Total Reads Aligned [a]	87.1% (20.4 M) (73 – 97 %)
Average Percentage of Total Reads Supressed [b]	9.37 % (2.2 M) (2 – 17 %)
Average Percentage of Total Reads Failed To Align	3.54 % (0.8 M) (0.6 - 15 %)

[a] At least one reported alignment.

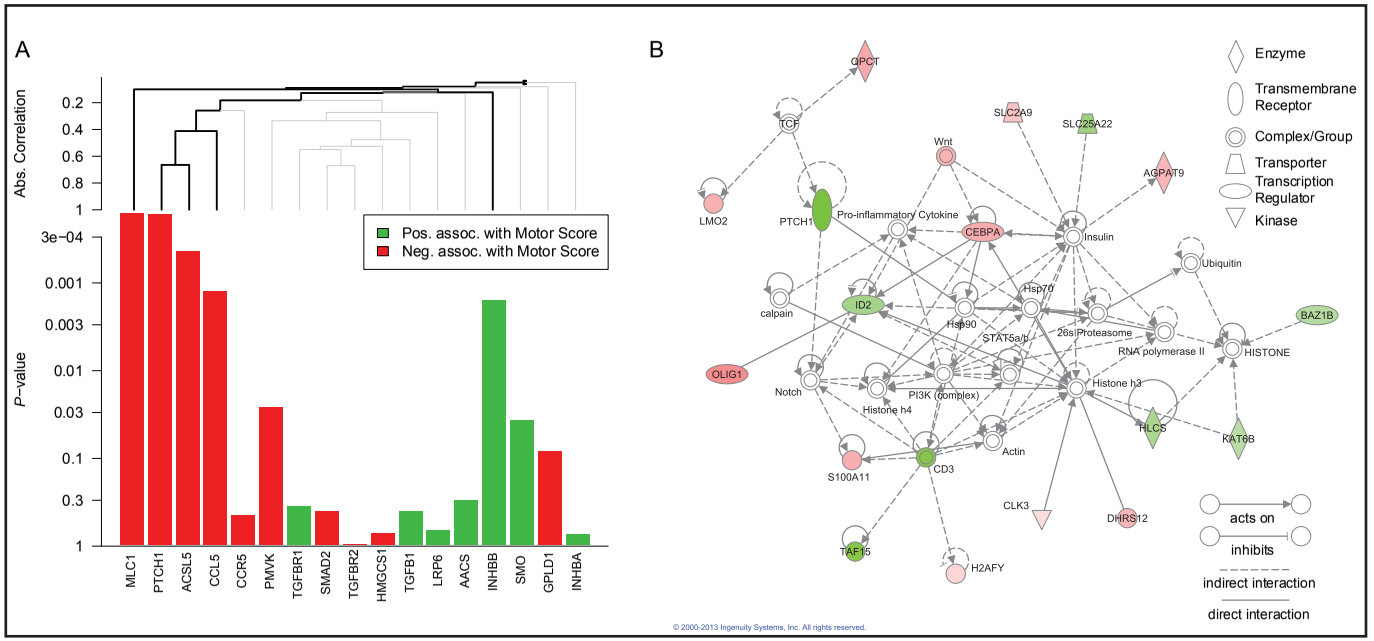
[b] Bowtie –m option - suppress all alignments for a particular read or pair if more than 2 reportable alignments exist for it.

Supplementary Table S7. Global test pathway analysis of DeepSAGE gene expression data for KEGG pathways, GO terms and MicroRNA target sets.

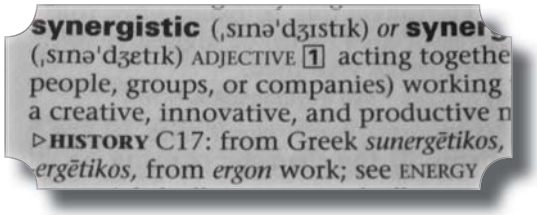
KEGG Pathways	Genes	GO Terms	Genes	Micro-RNAs	Genes
Toll-like receptor signaling	82	NADP binding	35	MiR-504	71
Pentose phosphate	22	Striated muscle cell differentiation	119	MiR-138	191
Osteoclast differentiation	115	Lung alveolus development	25	MiR-337	128
Jak-STAT signaling	102	Pancreas development	42	MiR-218	305
Amyotrophic lateral sclerosis	39	Cellular response to light stimulus	28	MiR-522	138
Neuroactive ligand receptor interaction	97	Response to cholesterol	17	MiR-492	51
Pathways in cancer	266	Positive regulation of IL-6 production	25	MiR-520D	273
Type II diabetes mellitus	36	Negative regulation of IL-12 production	10	MiR-498	92
Long-term depression	50	Cellular carbohydrate biosynthetic process	52	MiR-28	84
Chagas disease	90	Glucose catabolic process	62	MiR-129	151
Focal adhesion	154	Regulation of heart contraction	54	MiR-513	104
Cytokine-cytokine receptor interaction	171	Positive regulation of neurological system	26	MiR18A MiR-18B	104

Supplementary Table S8. Ingenuity pathway analysis top disease and functions for the top DeepSAGE differentially expressed genes.

ID	Top Diseases and Functions	Score	Focus Molecules
1	Cell Death and Survival, Nervous System Development and Function, Tissue Development	33	21
2	Cell-To-Cell Signaling and Interaction, Hematological System Development, Immune Cell Trafficking	30	21
3	Hematological System Development and Function, Inflammatory Response, Tissue Morphology	29	18
4	Infectious Disease, Renal and Urological Disease, Lipid Metabolism	28	18
5	Cell Cycle, Cell Morphology, Cell-To-Cell Signaling and Interaction	27	17
6	Cancer, Connective Tissue Disorders, Skeletal and Muscular Disorders	26	17
7	Carbohydrate Metabolism, Small Molecule Biochemistry, Immunological Disease	25	16
8	Hereditary Disorder, Renal and Urological Disease, Carbohydrate Metabolism	24	16
9	Cell Cycle, Skeletal and Muscular System Development and Function, Cell Death and Survival	24	16
10	Auditory Disease, Immunological Disease, Cellular Movement	21	14



Supplementary Figure S5. A. Covariate plot analysis from global test for the GO-term "response to cholesterol". All genes found in the DeepSAGE analysis and for the above term are shown on the x axis, while the P-values for these individual genes are plotted on the y-axis (most significant are represented by highest bars, red are negatively associated with TMS, green are positively associated with TMS). The dendrogram reflects the correlation in expression of these genes across all samples, with bold lines indicating branches (groups of genes) that are significant **B.** Molecules reported from IPA top diseases and functions network analysis and for the terms skeletal and muscular disorders, connective tissue disorders and cancer (colored genes represent genes discovered from DeepSAGE sequencing, red=upregulated genes, green=downregulated genes).



Chapter 4

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

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

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

Metabolomics (2016) 12: 137

Abstract

Introduction

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

Objectives

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

Methods

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

Results

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

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

Conclusion

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

Introduction

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

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

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

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

Materials and methods

Metabolite measurements

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

Serum collection

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

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

Quality controls

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

Statistical analysis

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

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

Integration of metabolomics with transcriptomics analysis

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

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

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

Results

The metabolomic dataset

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

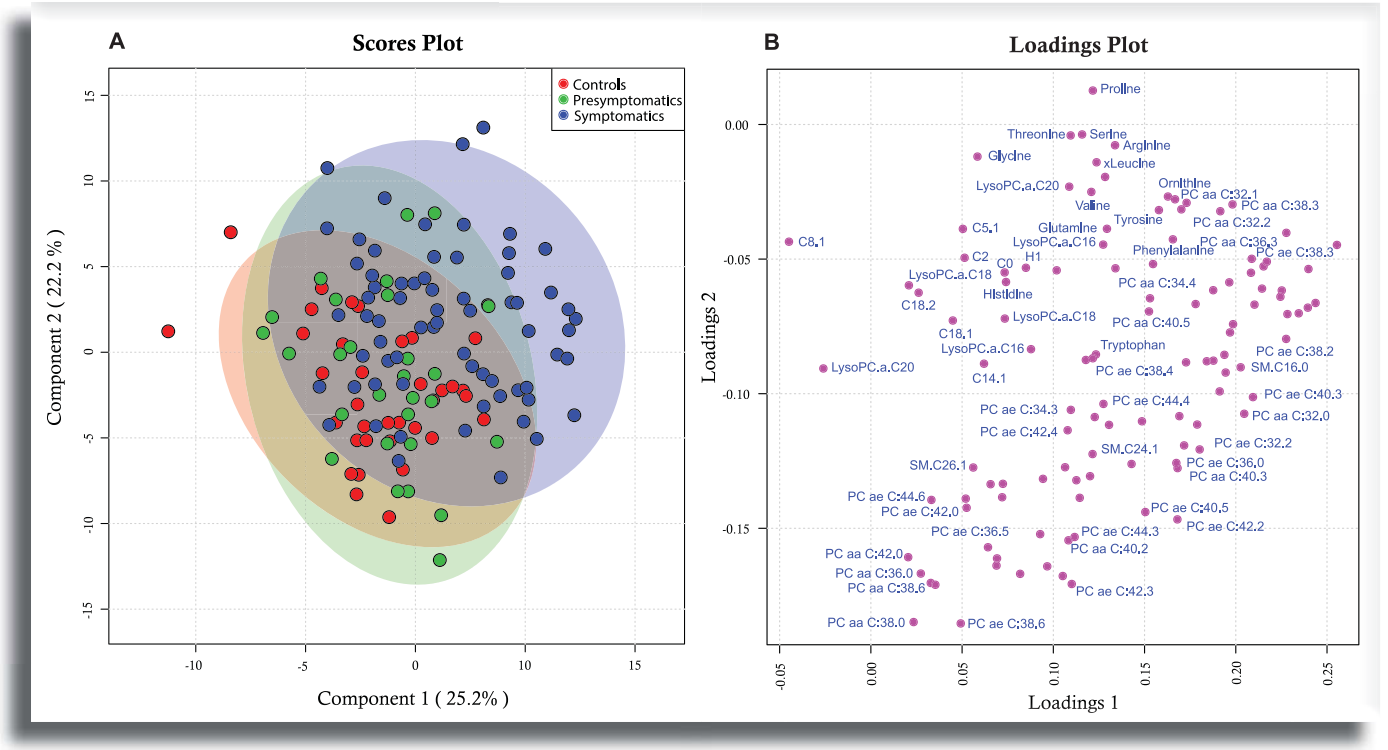


Figure 1. Diagram representing principal component analysis performed using Metaboanalyst v3.0 **A.** Principal component analysis results represent HD presymptomatic, symptomatic patients and control group separation based on the top 2 most significantly contributing components. Colored circles represent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

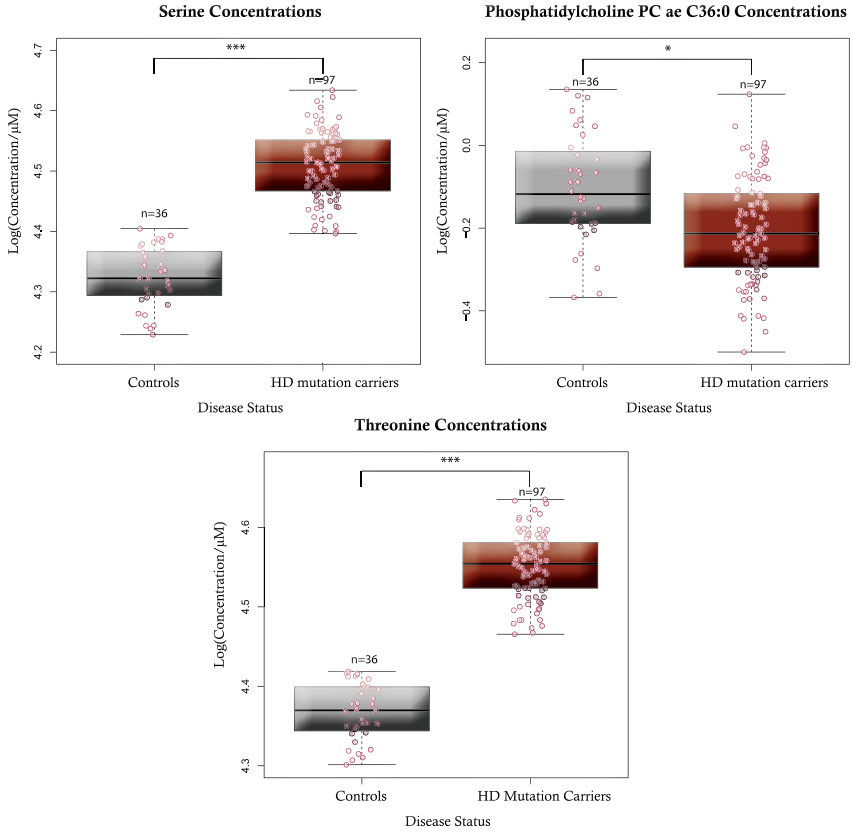
Metabolite ID	HD Mutation Carriers vs Controls Analysis	P Value	Concentration Change
Ser	**Serine**	**3.08E-05**	**Higher in HD**
PC ae C36:0	**Phosphatidylcholine acyl-alkyl C 36:0**	**6.61E-05**	**Lower in HD**
Thr	**Threonine**	**1.02E-03**	**Higher in HD**
PC ae C42:0	Phosphatidylcholine acyl-alkyl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

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

[2] Higher in earlier symptomatics

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

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



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

Association of metabolite pair ratios with disease status and severity

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

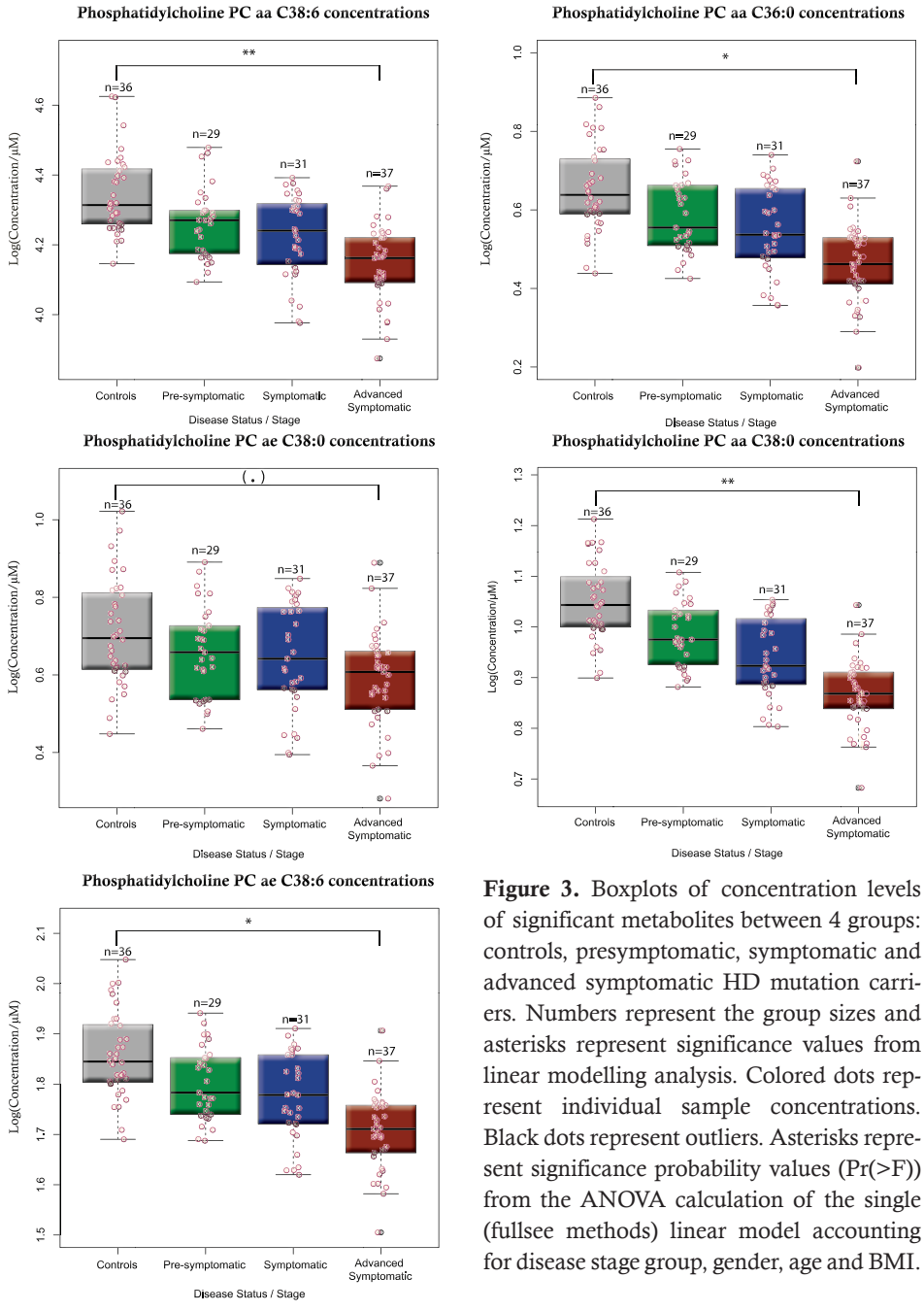


Figure 3. Boxplots of concentration levels of significant metabolites between 4 groups: controls, presymptomatic, symptomatic and advanced symptomatic HD mutation carriers. Numbers represent the group sizes and asterisks represent significance values from linear modelling analysis. Colored dots represent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

Integration of metabolomics with transcriptomics

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

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

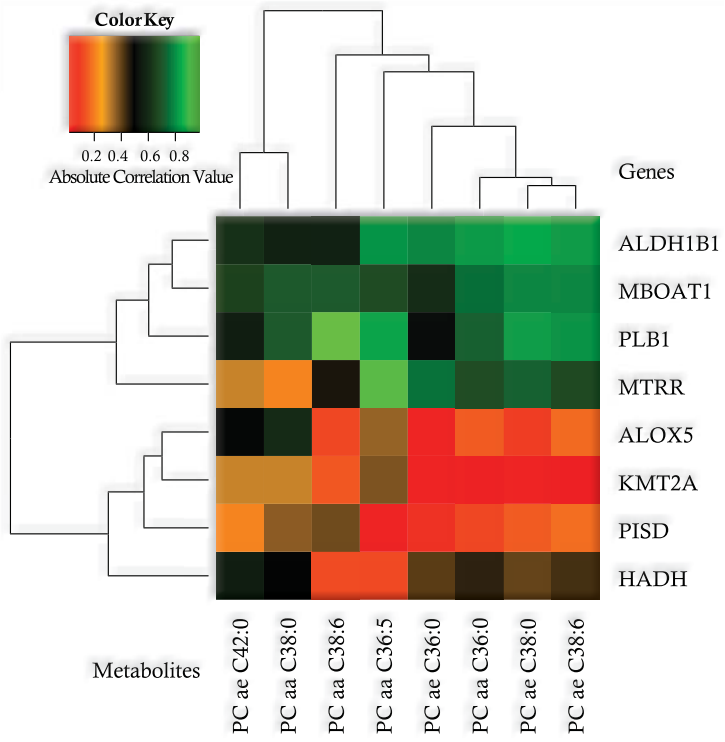


Figure 4. Heatmap of correlation values between gene expression levels and phosphatidylcholines metabolite concentrations. The selected genes shown here are genes identified using our previous gene expression data and that participate in phosphatidylcholine KEGG and BioCyc pathways and reactions. Phosphatidylcholines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

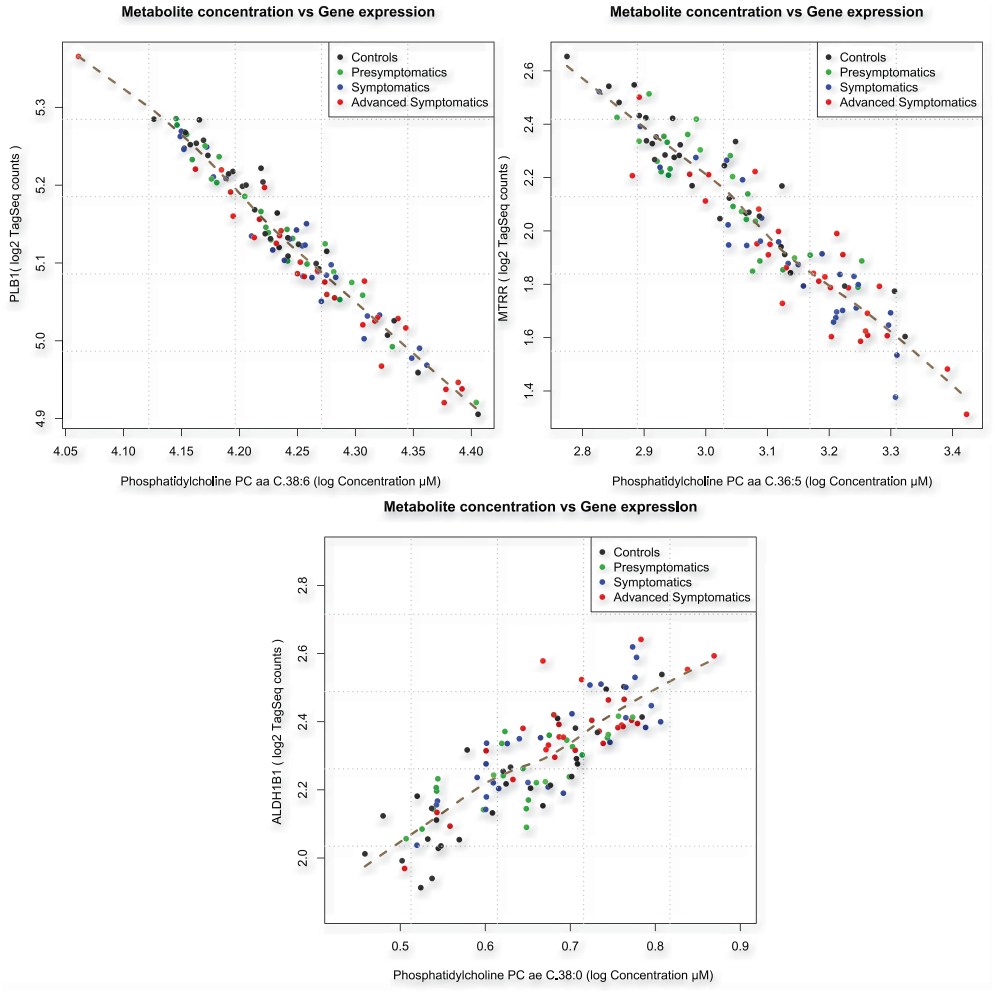


Figure 5. Plots of selected phosphatidylcholine metabolites versus their corresponding gene that participates in phosphatidylcholine pathways or reaction. The 3 plots represent the most highly correlated metabolite to gene pairs from the integration of metabolites to gene expression data analysis. Different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

Discussion

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

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

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

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

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

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

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

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

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

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

Acknowledgments

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

Compliance with Ethical Standards

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

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

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

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

Reference List

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

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

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

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

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

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

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

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

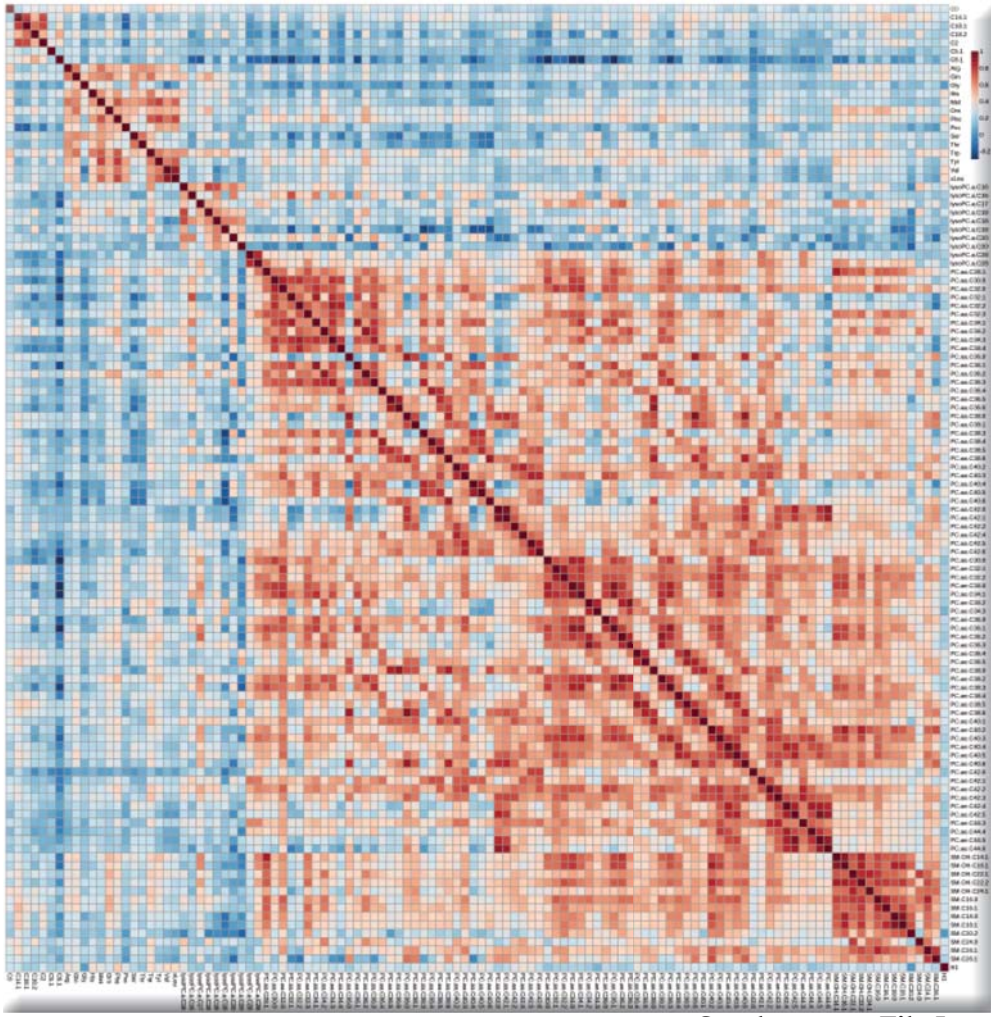
Supplementary Material

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

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

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

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



Supplementary File 5. ▸

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

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

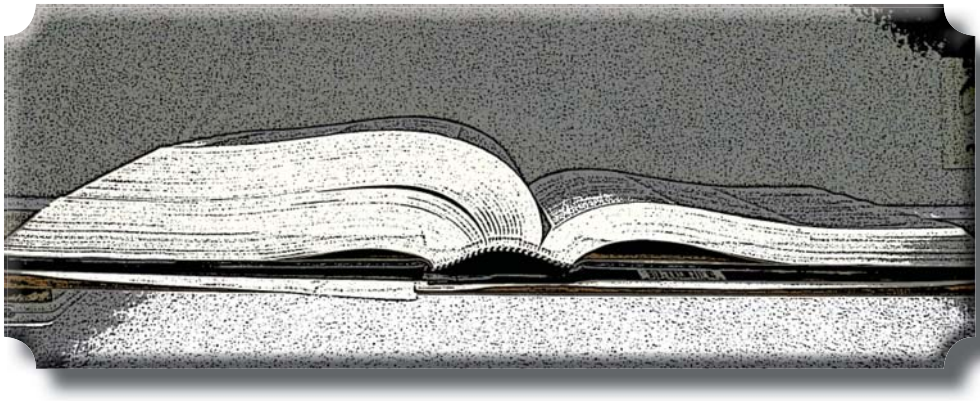
[1] Median concentration lower in earlier symptomatics (group 3) and relative to pre-symptomatics (group2).

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

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

Chapter 5

Discussion



5. Discussion

5.1 Summary

The aim of my thesis was to study the suitability of peripheral whole blood to identify disease pathology related events in HD patients. In Chapter 2 I explored the dynamic range of next generation sequencing-based transcriptomics by depleting the overabundant hemoglobin transcripts present in peripheral whole blood, which limit the dynamic range of this experimental approach. In addition, I focused on technical aspects of using whole blood as a tissue. This is particularly important considering the problems in reproducibility and lack of inter-study validation that previous HD studies have reported until now [1]. In Chapters 3 and 4 I performed cross sectional transcriptomics and metabolomics studies using samples from the same cohort of patients. The experiments performed were based on the well described notion that the expanded mutant huntingtin protein causes both genome wide transcriptional deregulation and metabolic dysfunction and failure in energy metabolism. The aim was to discover novel transcriptomic and metabolomic blood biomarkers of disease progression, validate previous ones and to identify novel or elucidate previous molecular pathways perturbed in HD. I did this by analyzing whole blood with state of the art technologies such as Illumina RNA-seq gene expression tag profiling and the highly standardized tandem mass spectrometry-based metabolomics platform of Biocrates. Subsequently we examined whether these technologies can produce useful results for disease stage and disease pathology event identification. Finally, I investigated if the results from the two different omics platforms could be combined as part of a systems biology approach that could provide further depth and an additional layer of disease knowledge relative to those that the individual study results could offer.

5.1.1 Globin Reduction Findings

In Chapter 2 I examined the effect of depleting peripheral whole blood of its highly abundant, hemoglobin transcripts (hemoglobin alpha and beta) and that constitute the majority of transcripts per sample library size during next generation sequencing (NGS)

transcriptional profiling. I compared gene expression in 6 hemoglobin reduced whole blood mRNA samples with their 6 non-reduced counterparts. The gene expression analysis of the non-reduced samples showed that hemoglobin transcripts (α and β) ranged from 52-76 % of the total transcript population while the residual mRNAs constituted, on average, only 40% of the samples' total number of transcripts. However, genome-wide gene expression profiling studies in blood aim mainly at studying this residual population, made up of the comprehensive set of leukocyte transcripts, rather than studying events in the non-nucleated mature erythrocytes. This is because, apart from providing potential biomarkers, the leukocyte mRNAs can also reflect possible functional aspects of disease pathology or highlight disease modifiers.

Our results showed that depletion of the hemoglobin transcripts can provide additional sequencing reads for the remaining blood transcripts. I found that approximately 54% of the total number of genes discovered were significantly overexpressed in the globin reduced samples at a false discovery rate of 1%. On the other hand, only around 4 transcripts of other genes per mille (87/21633) were inadvertently reduced in the treated samples. This indicates that the potentially unwanted side effects of the globin depletion on gene discovery rate is very small. Moreover, we could not identify sequence specific properties (using BLASTN) particular to the few genes that were lost which would robustly explain the systematic loss of these genes. By studying the composition of the genes present in low copy numbers, I discovered that the functional biotype of these genes included cell surface receptors, signal transduction proteins and G-protein coupled receptors. Moreover, the relative percentage of low abundance genes was higher for gene classes like processed transcripts, lincRNAs and miscellaneous RNAs, when compared to the more typical gene class of protein coding transcripts. It has been postulated that transcripts that are currently considered as non-coding or non-functional could also be called "currently unannotated". Indeed, when the function of these transcripts is discovered this will also have implications to the already discovered protein coding transcript functions [2]. Low abundance genes such as the ones reported above can be also important for HD and other neurodegenerative diseases since they are often found to be differentially regulated in gene expression studies. Thus, it is vital that their levels can be measured with high accuracy and at sufficient depth. However, analytical protocols and

techniques have not been sufficiently developed for these genes to be easily quantified. This means that these are often overlooked and substituted by other candidate transcripts that are more highly expressed and easily detected. For example, Labadorf et al., using RNA sequence analysis of human brain samples, showed that the most significantly differentially expressed genes in HD were Hox and other homeotic genes [3]. However, they also showed that the relative abundance of almost all of those Hox genes was low, around 10 % or less of the median normalized tag count average of all genes. Therefore, even though the levels of such genes could be strongly associated with disease it would not be possible to routinely and robustly quantify (or even detect) them unless a very high sequencing depth is employed. The latter situation is complicated even more when studies are performed in a tissue such as blood that is dominated by a few overabundant globin transcripts. For this reason and in order to determine how the hemoglobin over-abundance effects the detection of all the transcripts/genes present in blood I performed an in-silico analysis of gene expression sequencing depth versus the total number of detectable transcripts. This analysis showed that for a given number of detected genes (in this case approx. 13.000 genes) non-globin reduced samples require almost twice as many sequencing reads to detect the same number of genes found in globin depleted samples. As NGS costs continuously drop, sequencing cost savings could best be used to increase the study sample size, the within study sample replication numbers or to perform separate replication experiments.

Finally, at the end of the second chapter I demonstrated that for the same number of sequencing reads a higher number of genes can be detected in globin reduced samples even at sequencing depths as high as 60 million reads. This means that even a great and costly increase in sequencing depth of blood samples would still be insufficient to detect all the genes present in these samples when hemoglobin transcripts are present (at 60 million reads). Additionally, the fact that even at such great sequencing depths not all the genes or transcripts are identified can cause inaccuracies in gene expression studies. For example, studies claiming that a specific set of genes is exclusively expressed (differential expression) in a disease state, have to provide strong evidence that the transcriptome of interest has been extensively sequenced and have sufficient statistical power and sample size [4]. Even though advanced statistical methods have been designed to estimate and

normalize the sample library sizes; such as the "Trimmed Mean of M-values" (TMM), these methods are at best very good approximations of the actual library sizes and are applicable only under the condition that exactly the same amount of starting RNA has been used [5]. Moreover, even the most current and widely used sequencing-based gene expression algorithms, such as edgeR [6], estimate the sample gene frequencies based on the total number of genes found in the sample. Therefore, if a relatively high number of low abundance genes, like transcription factors or others known to be deregulated in HD, are not robustly sequenced (or not detected at all), these genes will have a low chance to be correctly estimated and appropriately associated with disease state. Consequently, globin reduced samples, yielding twice as many genes than their non-reduced counterparts, further contribute to obtaining a better representation of the transcriptome and a more accurate estimation of low abundance transcripts. This is of the utmost importance for gene expression studies and transcriptional biomarker studies, since the transcriptional capacity of mammalian genomes, in contrast to lower level eukaryotes, is immense - consisting of thousands of genes and numerous isoforms that are transcribed in an interleaved fashion across the whole genome [2,7].

5.1.2 DeepSAGE gene expression findings

In the second chapter of this thesis I used peripheral whole blood samples of HD patients and controls to perform a cross sectional genome wide gene expression profiling study. Until recently gene expression profiling studies were performed using hybridization-fluorescence-based microarrays. However, these platforms suffer from several limitations like cross-hybridization, low sensitivity and exclusive focus on the genes present on the array. This is not the case with NGS-based genome-wide transcriptional profiling. In addition, this approach provides better sensitivity, discovers novel transcripts and better quantifies low and high abundance transcripts. We collected 150 patient and control samples in a systematic and reproducible manner using Paxgene tubes, which preserve the level and quality of the RNA samples at venipuncture. Since it was not feasible to use the globin reduction approach described above in our current HD cohort study, we used a sequencing protocol (DeepSAGE tag profiling) that only sequences a small part (tag) of the 3 'UTR, which uniquely characterizes each mRNA transcript. Thus, we

could achieve a much higher sequencing depth than with standard RNAseq. To identify transcriptomic changes that would function as potential disease progression/severity biomarkers we performed a linear modeling analysis using as, our main covariate the UHDRS total motor score (TMS) and which describes muscle and motor degeneration in HD patients. Despite the fact that TMS, is characterized interrater variability and other disadvantages, we were the first to replicate results for several genes reported as potential blood biomarkers by previous HD gene expression studies. Furthermore, when we performed the linear modelling analysis using the total functional capacity score as an alternative main covariate, a large proportion of the top differentially expressed genes (≈ 60%) were the same as the ones discovered by the total motor score analysis reflecting the good correlation between TMS and TFC in HD patients. Moreover, we used a high-throughput nanoliter RT-qPCR platform to validate a selected number of 20 genes in the original discovery cohort and in a second independent cohort. We were able to validate 5 of the 20 genes that were significantly associated with TMS in both cohorts.

Furthermore, we produced a gene expression based algorithm using the "Lasso" approach (Least Absolute Shrinkage and Selection Operator) that could be used to predict the motor symptom progression in patients. This uses the expression levels of a specific set of genes, age and gender. This analysis showed that gene expression could be used with good results to predict motor symptom pathology, especially for the earlier stage symptomatic patients. For the later stage patients (higher TMS) the scores predicted by the blood gene expression were mostly significantly lower than their corresponding clinical motor scores. One reason for this could be the reported plateau effect of disease progression and specifically of chorea in more advanced HD patients [8]. We note that this gene expression-based motor score predicting formula may well reflect, besides previous neurodegeneration, also degenerative events from other tissues that blood comes into contact with. This has been previously also suggested by Hodges et.al. [9] who noted as a potential limitation in HD brain studies that changes in susceptible cells are inextricably mixed with effects of previous neurodegeneration. For our blood study this could also explain why the motor score algorithm predicted lower motor scores for the most advanced patients. Furthermore, this analysis demonstrated that even genes (such as *ARL4C* – see Chapter 2.3.3), which do not achieve statistical significance in statistical

or validation analyses may still play a role in a more accurate representation of disease progression. Here we have used the "Lasso" approach to show that a panel of genes can be trained to describe disease stage or severity. Similar to using a combination of different types of biomarkers (molecular, imaging etc.) to better define overall disease progression in a multifaceted disease like HD, the combined use of several blood biomarkers may best describe disease progression and/or severity. In our dataset, the TFC score and the TMS score exhibited high correlation for most, but not all HD patients. This could be caused by the fact that some HD patients exhibit severe psychiatric symptoms affecting everyday independent functioning at an early stage, while developing motor symptoms only later in live. For these patients, most likely a different profile of gene expression biomarkers will best reflect their progression. Thus, different groups of gene expression biomarkers could be formed by a core of robust and/or highly expressed, statistically significant genes, but also contain other genes which represent more phenotype-specific information.

Finally, while there is an urgent need for the identification of panels of biomarkers with high performance characteristics such as robust expression and sensitivity, knowledge about the biological functions and pathway modules that these biomarkers are involved in will also be very important. Such functional information will provide further specificity and clues about the mode of action and effectivity of potential therapeutic agents. For this reason, I performed different pathway analyses and identified disease related biological terms and pathways. These can be used to choose biologically relevant genes that will be the most appropriate for chosen therapeutic approaches and the most informative regarding the underlying processes that are altered during disease progression.

5.1.3 Metabolomics and integration with transcriptomics findings.

In recent years, several systems biology studies have combined the use of genomics, transcriptomics and proteomics data to identify biomarkers for therapeutic clinical trials. The field of metabolomics has also made great strides to extend such multidimensional system-level approaches and produced some promising results in the last decade [10,11].

However, few metabolomics studies have been performed in comparison with DNA, RNA and protein studies, the three biological classes following the central dogma of biology. In our study, we applied targeted metabolomics in serum samples of HD patients and controls. This targeted approach is important given the heterogeneity of the peripheral whole blood, which contains several overabundant proteins that hamper detection and identification of the less abundant metabolites [12]. We could identify a total of 10 metabolites that either separated controls from mutation carriers or tracked disease progression across clearly defined disease stages. Our results are in agreement with the results of Tsang et al. who showed phosphatidylcholine to be decreased but in the frontal cortex of the 3-NP HD rat model. [13]. Some of our results have been partially validated at the individual metabolite level by Cheng et al. who, similarly using targeted metabolomics, also reported phosphatidylcholine acyl alkyl C:36 (PC ae C36:0) to be deregulated in the blood plasma fraction of HD patients [14]. The fact that one of the biomarkers identified in the current thesis could be validated in an independent cohort and in a different laboratory setting warrants additional validation experiments to identify, in detail, the extent to which this phosphatidylcholine level can track disease progression.

Finally, we combined our previously published gene expression data from the same cohort with the metabolomics data, to explore what could cause the concentrations of these metabolites to change in the patients' serum. To perform the above integration, we used correlation analysis, workflow based exploitation of databases and open public platforms like WikiPathways. Approaches like WikiPathways enable researchers to take advantage of the expert knowledge of their colleagues, who continuously update, curate and enrich biological pathway information based on novel information from their research. Our analyses identified a set of 8 genes, 4 of which exhibited high correlation with most of the metabolites changed in disease. The most prominent biological terms that were associated with these highly-correlated genes were deregulation of the folate cycle and homocysteine metabolism (*MTRR* gene) but also potential deregulation of phospholipase activity – based on the levels of PLB1 gene. These results are supported by several HD metabolomics papers concerning the deregulation of genes and metabolites involved in antioxidant defenses and neurodegeneration. Notably, Andrich et al. reported increased plasma total homocysteine in treated HD patients and two further stud-

ies reported deregulation of enzymes involved in the regulation of the levels of cysteine [15-17]. The exact mechanism by which the levels of these amino acids and expressed genes are connected is not clear, but our results stress the need for further studies. Finally, our findings underscore the value of such system-biology approaches to enhance our knowledge on disease pathology. Genes such as the ones found by the above integrative approaches could have easily been overlooked and replaced with slightly more robust or more significant genes which might nonetheless be less meaningful biologically, disease specific or functionally relevant. Indeed, small fold changes in gene expression or metabolite concentration levels identified by the previous individual studies can become far more significant when the data from the separate experiments are aggregated into a single multidimensional biological readout, highlighting much more dynamic shifts in the overall flux through relevant biological pathways.

5.2 Limitations and challenges

The advent of next generation and high throughput "-omics" technologies in the recent years has provided the unique opportunity to generate a plethora of data about patient genomic variants, mRNA and miRNA expression changes as well as the concentration levels of proteins and metabolites in various patient tissues. Nonetheless, despite these unique opportunities the clinical utility is still limited compared to the volume and breadth of the originally produced data. Several study design limitations as well as technical and biological challenges hinder the fields of personalized genomics and systems biology from achieving its goals and obtaining robust, reproducible results. In the following paragraphs, such limitations relative to the studies published in this thesis are discussed and critically analyzed.

The use of peripheral blood for the discovery of biomarkers and surrogate pathologies and endpoints has proven to be a double-edged sword in HD. Despite the advantages of this tissue, researchers have not yet been able to fully harness the potential of peripheral blood as a source of clinical biomarkers. This is mainly due to the fact the different blood cell types are also sensitive to numerous preanalytical aspects of the study design, such as circadian rhythm and time of sample collection, the use of different sample collection

and storage protocols, patient nutritional status or other lifestyle factors that can affect cellular expression levels, such as medication, alcohol consumption or smoking. Therefore, even though the cohort used in our study was subjected to very strict age and gender stratification, the specificity of the study would have benefited from additional information regarding sample collection time, fasting state and additional information regarding medication and lifestyle. The Paxgene tube collection method preserves the samples as whole blood lysates, and as discussed earlier this results in the overabundance of hemoglobin transcripts which in turn limits the experimental dynamic range and consequently the statistical power of the results. Even if a depletion approach as described in Chapter 2 is followed, one needs to consider additional variation and reproducibility issues resulting from hemoglobin reduction, like decreased RNA quality and increased inter-sample variability, but also whether post-depletion sample concentrations will be adequate for secondary validation studies. However, considering the continuous development of increasingly sensitive state of the art platforms, enabling the progressive lowering of required sample amounts, it is this author's opinion that RNA quality is the more critical factor. It is our experience that whole blood RNA extracts exhibit lower post RNA purification quality, notably relative to cell lysates, and it is well known that RNA quality influences gene expression levels detectable. Even though in Chapter 2 we showed that RNA quality degradation of the depletion protocol is relatively small, this decrease can still cause the exclusion of a good number of available samples in, subsequent quality control because of missing data, notably when their initial quality is already low.

Our metabolomics approach had as a unique and important characteristic that large numbers of reproducibly collected HD samples were analyzed by a well-known and standardized targeted metabolomics platform. Furthermore, linking metabolomics with state of the art gene expression data showed promising correlations between genes and metabolites, both at the functional and the expression level. Metabolites can act both as feed forward and feedback reaction modulators and our linear modeling and gene expression correlation analyses may reveal only a limited amount of the underlying disease pathology. Furthermore, our study could not quantify metabolites whose concentration levels were below the lower limit of quantification and/or the limit of detection. The metabolite group below the limit of detection were mostly carnitines. Nonetheless, from

previous studies it is likely that the vast majority of the undetected metabolites is due to the tissue type (serum) and platform used rather than the absence of the metabolites in the disease stage [18]. Finally, results from cross-sectional studies are sometimes limited because they only capture a snapshot of the patients' biological fluids during a single sampling point, a static representation of the patients' disease stage at the time of venipuncture. Thus, two patients that share an almost identical biofluid profile at the time of blood sampling might have a very different disease course previous to this point. This applies particularly well for HD since it is known that the disease manifests itself though a multitude of symptoms and with highly variable progression rates. Therefore, such studies cannot be used alone to derive meaningful therapeutic endpoints. Even though they can provide important knowledge about disease bio-pathology, any conclusion(s) emerging from such studies will require additional validation in longitudinal studies.

5.3 Future directions

The future of biomarker discovery will arguably revolve around the use of systems biology approaches, integrating genomics, transcriptomics, proteomics and metabolomics to monitor human health and disease as well as the efficacy of therapeutic strategies. Such multi-omics approaches will benefit from further developments in analytical and/or clinical technologies which provide observable characteristics – known as endophenotypes - to extract highly accurate and informative molecular biomarkers. Potential endophenotypes are for instance brain atrophy scores produced by MRI experiments, psychopathology and rates of cognitive decline. The combination of these disparate data types with molecular measurements will require the collaboration of many disciplines to interconnect the data generated by clinicians and basic investigators. Independently of which type of biomarker might become more easily adopted or be the most informative, there is a widespread consensus that for a multifaceted disease like HD a single biomarker will not be sufficient. Rather, a battery of biomarkers will be required to fully characterize the disease progression and evaluate the therapeutic potential of future treatments. In the following paragraphs, additional future directions are described, both based on the studies performed in this thesis and more generalized future directions of blood based study designs based on information from the HD research community and field experts.

For personalized genomics and systems biology to make the transition from future aims to current and realistic tools in the hands of researchers, the simultaneous processing of a much larger number of samples is required, offering greater statistical and discovery power. The collection of enough informative samples can be achieved either through a centralized approach, where all samples are preprocessed together, or by meta-analysis of data collected in several smaller decentralized laboratories. Irrespective of which of the last of the two approaches proves the most efficient, for the results to be meaningful and informative, preprocessing variation needs to be kept to an absolute minimum. It has been demonstrated by Schurman et al. that several technical and experimental variation factors such as technical platform parameters, sample collection parameters and batch wise RNA isolation and amplification influence the gene expression results.

Therefore, considering the gene expression protocol used in this study we emphasize the need for simpler and easier to perform protocols to avoid batch effects. Even though the DeepSAGE protocol employed in the current study offers unique advantages such as increased sequencing depth and improved dynamic range, the subsequent use of restriction enzyme and polymerase reactions to produce the required sequencing "tags" make it a demanding and laborious protocol. As the NGS platforms evolve we may expect that sample processing and sequencing construct generation protocols will also be further optimized i.e. by sample processing in a 96-well format which is already being applied. These improvements allow for the easy and simultaneous processing of a larger number of samples, eliminating difficult to spot batch effects. Using more, and more uniformly treated samples facilitates the greater statistical power needed in a multifaceted, highly variable disease like HD and eliminate inter-lab assay performance and variability. In addition to the use of uniform sample collection, storage and analytical protocols, further efforts should be intensely advocated to select a narrower group of high quality clinical markers and progression scores that can best define disease pathology. Initiatives to this end are being taken by well-respected researchers and authorities in the HD field such as the TRACK-HD consortium [19]. These carefully chosen scores and grades can then be more widely and more uniformly adopted as potential covariates in the various statistical analyses employed by the different studies and research groups. This will assist

in reducing the gap in reproducibility and common findings between various studies, which even though similar in nature, nonetheless have used very different cohort stratification designs or clinical sub-scores to interpret the biological data produced. This would respond to the goal set by the European Huntington's Disease Network (EHDN) to standardize clinical scores. Technologies such as flow sorting and cytometry, that capture and sequence the transcriptome of a specific number and type of blood cell types, will provide more detailed blood transcriptome data and should overcome the hurdles posed by overabundant hemoglobin transcripts. In addition, more advanced statistical models could be applied to identify additional perturbed pathways since disease relevant associations between metabolite and gene expression are not always explained by 1:1 stoichiometric relationships. One such example of additional bioinformatics approaches is the use of artificial neural networks (AANs). These are inspired by biological neural networks and employ tunable numerical parameters or weights which can be trained on the data and used to approximate non-linear relationships between an input and an output molecule or disease state [20]. Finally, even though targeted and untargeted metabolomics approaches have their own limitations and advantages, in the future they could both be used in a complementary manner to reveal additional and more comprehensive metabolite disturbances.

5.4 Epilogue

The role of blood as a pathology and biomarker discovery tissue in HD and other neurological disorders has been the topic of many heated debates and stalemated scientific discussions. Irrespective of the extent to which blood and its biomarkers will be proven to associate with central HD brain pathology, it is this author's opinion that in the future, blood biomarkers will irrefutably have an important role in disease progression and therapeutic monitoring. Even if these blood indices are more distal and less specific than more proximal biomarkers, this drawback is more than cancelled by the ability to repeatedly and continuously collect disease signature measurements, in an easy and reproducible manner, across a large number of individuals, providing an otherwise unobtainable sensitivity for the monitoring of clinical and therapeutic trials. Considering the multiplicity of HD pathology and its symptoms it is already well established and

widely accepted that blood alone will not be sufficient to describe all disease aspects. Therefore, in such a context, the high sensitivity measurements of a biomarker obtained in a casual repeated fashion can overshadow the potential loss in specificity. Although continuous technological improvements of platforms and protocols will provide further breakthroughs, a great deal can be gained from additional standardization of sample collection/storage and the communal adoption of the same standardized clinical scores and endophenotypes for the inter-laboratory statistical analyses of extracted data. The latter will be especially useful in inductive biomarker identification approaches where the unbiased identification of a large number of molecules is performed through the use of state of the art high throughput technologies like the ones described in this thesis

To truly estimate whether blood can provide the required disease progression and therapeutic efficacy information, it is necessary to account for additional sources that contribute to biological variation in this tissue. This thesis has analyzed several technical and experimental tissue specific sources of variation. However, other genetic regulatory sources of human gene expression variation could have effects on interpreting expression differences, such as epigenetic regulation or expression quantitative trait loci (eQTLs). For instance, recently von Schimmelmann et al. showed that deficiency of Polycomb repressive complex 2 (PRC2), an epigenetic gene regulator, causes transcriptional changes and a temporal type of PRC2 dependent gene de-repression that leads to a progressive and fatal neurodegeneration reminiscent of HD [21]. Furthermore, Sudarsanam et al. showed that variations in genomic regulatory features such as quantitative trait nucleotides (QTNs) found in Saccharomyces cerevisiae transcription factor genes, produced variation in gene expression (49%) which however was not predictive of the phenotypic variation (98% sporulation efficiency) that could be explained by the QTNs [22]. Integrative genomic and QTL mapping approaches in Arabidopsis lines have shown that genetic variation in gene expression is not always evidenced by phenotypic variation [23]. Finally, human cortical gene expression analysis has also shown that inter-individual cortical expression was correlated with genome wide genotypic variation. As shown by the above studies in Saccharomyces and Arabidopsis, phenotypic changes are often masked or delayed by transcription network buffering interactions. All these studies lead to the conclusion that assessing the impact of the individual, natural and genome wide

expression variation across different tissues will be an important requirement for correctly defining pathological thresholds or biomarker level changes in disease. Such expression variations or buffering networks need to be better studied before we attempt to correlate tissue gene expressions signatures such as those between blood and brain across individuals with a different genotypic background. This kind of information will go hand in hand with disease specific genomic studies such as the one recently performed by the GeM-HD consortium that has identified genetic modifiers of HD and improve our understanding of genome to phenome correlates. All of these considerations emphasize the need to further promote the design of multi-level systems biology approaches since only the detailed mapping of all disease processes, from the genome to the metabolome, will be the most effective for developing therapeutic insights. Individualized disease specific information will not only allow for the finer stratification of patient cohorts but will also allow us to better interpret the results identified by the individual segments of the central dogma of molecular biology and produce a more personally tailored, accurate and effective therapeutic approach. Increasing evidence has led to the suggestion that Huntington's disease could be an "unfortunate by-product" of the evolution toward a highly-sophisticated brain which required an ever-lengthening HD gene [24]. However, we have yet to explain why after a certain number of CAG repeats the gene product becomes toxic and what mechanisms allow carriers to (variably) not develop symptoms. Therefore, the identification of genetic or environmental features that will provide or strengthen phenotypic buffering might not only protect present mutation carriers but could also protect non-carriers from a future evolutionary perishing.

Reference List

1. Seredenina T, Luthi-Carter R: What have we learned from gene expression profiles in Huntington's disease? Neurobiol Dis 2012, 45: 83-98.
2. Kapranov P, Willingham AT, Gingeras TR: Genome-wide transcription and the implications for genomic organization. Nat Rev Genet 2007, 8: 413-423.
3. Labadorf A, Hoss AG, Lagomarsino V, Latourelle JC, Hadzi TC, Bregu J et al.: RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. PLoS One 2015, 10: e0143563.
4. Garcia-Ortega LF, Martinez O: How Many Genes Are Expressed in a Transcriptome? Estimation and Results for RNA-Seq. PLoS One 2015, 10: e0130262.
5. Finotello F, Di CB: Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. Brief Funct Genomics 2015, 14: 130-142.
6. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010, 26: 139-140.
7. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP: Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet 2014, 15: 121-132.
8. Dorsey ER, Beck CA, Darwin K, Nichols P, Brocht AF, Biglan KM et al.: Natural history of Huntington disease. JAMA Neurol 2013, 70: 1520-1530.
9. Hodges A, Strand AD, Aragaki AK, Kuhn A, Sengstag T, Hughes G et al.: Regional and cellular gene expression changes in human Huntington's disease brain. Hum Mol Genet 2006, 15: 965-977.
10. Zhang A, Sun H, Yan G, Wang P, Wang X: Metabolomics for Biomarker Discovery: Moving to the Clinic. Biomed Res Int 2015, 2015: 354671.
11. Sethi S, Brietzke E: Omics-Based Biomarkers: Application of Metabolomics in Neuropsychiatric Disorders. Int J Neuropsychopharmacol 2016, 19: yv096.
12. Anderson NL, Anderson NG: The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 2002, 1: 845-867.
13. Tsang TM, Haselden JN, Holmes E: Metabonomic characterization of the 3-nitropropionic acid rat model of Huntington's disease. Neurochem Res 2009, 34: 1261-

1271.

14. Cheng ML, Chang KH, Wu YR, Chen CM: Metabolic disturbances in plasma as biomarkers for Huntington's disease. J Nutr Biochem 2016, 31: 38-44.
15. Andrich J, Saft C, Arz A, Schneider B, Agelink MW, Kraus PH et al.: Hyperhomocysteinaemia in treated patients with Huntington's disease homocysteine in HD. Mov Disord 2004, 19: 226-228.
16. Sbodio JI, Snyder SH, Paul BD: Transcriptional control of amino acid homeostasis is disrupted in Huntington's disease. Proc Natl Acad Sci U S A 2016, 113: 8843-8848.
17. Paul BD, Sbodio JI, Xu R, Vandiver MS, Cha JY, Snowman AM et al.: Cystathionine gamma-lyase deficiency mediates neurodegeneration in Huntington's disease. Nature 2014, 509: 96-100.
18. Draisma HH, Beekman M, Pool R, van Ommen GJ, Adamski J, Prehn C et al.: Familial resemblance for serum metabolite concentrations. Twin Res Hum Genet 2013, 16: 948-961.
19. Tabrizi SJ, Scahill RI, Owen G, Durr A, Leavitt BR, Roos RA et al.: Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. Lancet Neurol 2013, 12: 637-649.
20. Zhang Z, Yu Y, Xu F, Berchuck A, van Haaften-Day C, Havrilesky LJ et al.: Combining multiple serum tumor markers improves detection of stage I epithelial ovarian cancer. Gynecol Oncol 2007, 107: 526-531.
21. von SM, Feinberg PA, Sullivan JM, Ku SM, Badimon A, Duff MK et al.: Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration. Nat Neurosci 2016.
22. Sudarsanam P, Cohen BA: Single nucleotide variants in transcription factors associate more tightly with phenotype than with gene expression. PLoS Genet 2014, 10: e1004325.
23. Fu J, Keurentjes JJ, Bouwmeester H, America T, Verstappen FW, Ward JL et al.: System-wide molecular evidence for phenotypic buffering in Arabidopsis. Nat Genet 2009, 41: 166-167.
24. Zuccato C, Cattaneo E: The Huntington's Paradox. Sci Am 2016, 315: 56-61.

Summary

Samenvatting

Publication list

Curriculum vitae

Acknowledgements

Summary

Even though Huntington's disease (HD) and its most prominent symptoms were described in the middle of the 19th century, it was not until the last decades of the 20th century that a systematic analysis of the biology and genetics of the disease started. Great efforts by the scientific and patient community have raised disease awareness, reduced the disease stigma, allowed for a better and earlier recognition of potential symptoms and offered perspective for therapy development. The disease is known to exhibit geographic prevalence differences. The most recent data suggest that disease prevalence in Australia, North America and Western Europe has increased by approximately 15-20 % per decade for the second half of the 20th century. This increase accentuates the need to care for these patients and to identify symptomatic treatments and therapeutic approaches for the disease.

The work described in this thesis presents part of a framework that can be used to extract detailed disease biological information from peripheral tissue. This framework is based on the central dogma of biology "DNA to RNA to protein" and on a systems biology approach that aims to produce synergetic data whose disease pathological, prognostic and predictive value is greater than the sum of the individual experiment results. HD patients are often characterized by a multifaceted clinical profile, consisting of several symptoms and variable disease progression rates. Therefore, a systems approach such as the one described above is expected to be the most effective in identifying potential treatments and predictive biomarkers that will be most informative for the different patient subpopulations.

In Chapter 1, a detailed report of the disease background and symptoms is provided as well as current potential therapeutic approaches. In the last part of the chapter a thorough description is presented of the ideal properties of a disease biomarker. These properties are discussed in the context of the idiomorphic character of HD symptoms and, most importantly, the development of a blood biomarker reflecting brain pathology.

In Chapter 2 we investigate the depletion of the overabundant hemoglobin transcripts from next generation sequencing (NGS) gene expression peripheral blood samples. Even though technological breakthroughs such as NGS offer increased sensitivity and dynamic range the disproportionate number of hemoglobins in whole blood samples reduces the sequencing depth potential and the identification of the less abundant blood transcripts. We showed that by removing the hemoglobins, other types of transcripts such as lincRNAs, misc-RNAs and processed transcripts are increasingly and more reliably detected. To our knowledge this study was the first to describe the effects of such a depletion process on RNA sequencing data and especially in human blood samples. The various low abundance transcripts often play potentially important roles in disease in general which is also the case for Huntington's disease [1].

Chapter 3 describes the results of our whole genome gene expression profiling study employing NGS in peripheral whole blood samples. To date, our study constitutes the largest peripheral whole blood gene expression study; including such a substantial number of pre-symptomatic patients. Additionally, this study was the first HD study to have used next generation sequencing-based gene expression profiling, in contrast to hybridization microarrays which were previously the most commonly used platform. Our results confirmed for the first time blood gene expression changes found in previous studies, thus highlighting the potential of NGS to obtain disease relevant biomarkers. We showed that by employing appropriate statistical approaches, genes other than the most statistically significant ones can be used to establish disease progression biomarker panels. Such biomarker panels can consist of a combination of genes which, in addition to mere statistical significance, may provide complementary disease-specific information. These panels can potentially be refined to best reflect the pathological state in the different severity groups.

In Chapter 4 we performed an advanced integration of high throughput, large cohort data from blood and from more than one biological data classes. We integrated our targeted metabolomics results with our previously described NGS transcriptomics data from the same patients and controls. Metabolomics can reflect changes occurring downstream of the gene expression level and its deregulation that takes place in HD. Further-

more, targeted metabolomics has the advantage of enabling robust, reproducible and repeated measuring of the discovered changes in potential future longitudinal studies. By this multidisciplinary approach and employing community-driven bioinformatics research tools, we could link the changes in the HD patients' metabolite concentrations with specific subsets of the top deregulated genes of our gene expression dataset. Our integrative effort constitutes one of the first multi-level "omics" approaches, in the future to be enriched by additional biological data classes, thus providing a much-needed, more holistic description of HD disease pathology.

Finally, in Chapter 5 a critical evaluation was given of all the data produced and the conclusions drawn in the previous chapters. The knowledge gained by each of the above phases was assessed in the context of the novelty and applicability of the methodological approaches and platforms employed. In the penultimate part of this chapter the limitations of each of the above experiments were scrutinized with respect to the inherent technical limitations of the study designs and the analytic tools used. Finally, potential future directions and advancements associated with the study design, analytical components and bioinformatics tools used in each individual study were outlined.

References

1. Labadorf A, Hoss AG, Lagomarsino V, Latourelle JC, Hadzi TC, Bregu J et al.: RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. PLoS One 2015, 10: e0143563.

Samenvatting

Ook al werd de ziekte van Huntington (ZvH) met de belangrijkste symptomen reeds beschreven in het midden van de 19e eeuw, het duurde tot de laatste decennia van de 20e eeuw voor er een systematische analyse van de onderliggende biologie en genetica tot stand kwam. Enorme inspanningen van wetenschappers, artsen en patiënten hebben ervoor gezorgd dat de bekendheid van de ziekte is toegenomen, het taboe rond de ziekte is verminderd, er een betere en vroegere herkenning van potentiële symptomen kwam, en er nieuwe therapieën ontwikkeld werden. Bekend is dat de prevalentie van de ZvH kan verschillen per geografische locatie en recente data suggereren dat de prevalentie van de ZvH in Australië, Noord Amerika en West Europa over de laatste helft van de 20e eeuw is toegenomen met 15-20% per decennium. Deze toename in prevalentie benadrukt nog eens de noodzaak voor goede zorg voor deze patiënten en om symptomatische behandelingsmethodes te ontwikkelen en nieuwe therapeutische toepassingen.

Het werk beschreven in dit proefschrift vertegenwoordigt een onderdeel van een kader dat gebruikt kan worden om ziekte specifieke biologische informatie te halen uit perifeer weefsel. Dit kader is gebaseerd op het centrale dogma uit de biologie DNA -> RNA -> eiwit en een systeembiologische benadering die tot doel heeft om synergistische data te produceren waarbij de som van de pathologische en voorspellende waardes meer is dan die van de resultaten van de individuele experimenten. Sinds patiënten met de ZvH vaak een veelzijdig klinisch profiel hebben dat bestaat uit verschillende symptomen en een variabele progressie van de ziekte, ligt het voor de hand dat een benadering zoals hierboven beschreven effectiever zal zijn in het identificeren van potentiële therapieën en voorspellende biomarkers die het beste passen bij de verschillende patiëntenpopulaties.

In **hoofdstuk 1** wordt een gedetailleerde beschrijving gegeven van de symptomen van de ZvH en de huidige potentiële therapieën. Het laatste deel van dit hoofdstuk geeft een beschrijving van hoe de ideale ziekte-biomarker eruit zou moeten zien. Deze eigenschappen worden besproken in de context van het idiomorfe karakter van de ZvH symptomen, en voornamelijk in relatie tot de potentiële surrogaat eigenschappen die

nodig zijn zodat een bloed-biomarker een weerspiegeling is van de pathologie die zich afspeelt in de hersenen.

In **hoofdstuk 2** onderzoeken we de depletie van de hoog tot expressie komende hemoglo- bine transcripten op de next generation sequencing genexpressieanalyse van bloed. Voor zover we weten was ons artikel het eerste welke het effect beschreef van deze depletie op RNA sequencing data en specifiek in humane bloedmonsters. Ook al leveren technologische ontwikkelingen zoals next generation sequencing een toegenomen sensitiviteit en dynamisch bereik, de onevenredig hoge expressie van hemoglobines in totaal bloed vermindert het sequencing potentieel en de identificatie van transcripten die laag tot ex- pressie komen. We hebben aangetoond dat door het verwijderen van de hemoglobines, andere transcripten zoals lincRNAs, miscRNAs en processed transcripten in meerdere mate en consequenter gedetecteerd kunnen worden. Dit soort transcripten die laag tot expressie komen spelen vaak een belangrijke rol in ziekten in het algemeen maar ook specifiek in de ZvH [1]. Dus laten we zien dat reductie van hemoglobine transcripten in totaal bloed het totaal aantal gedetecteerde transcripten toe doet nemen en daardoor ook de statistische kracht om relevante ziektesignaturen te ontdekken.

Hoofdstuk 3 beschrijft de resultaten van onze genoomwijde genexpressie studie met next generation sequencing in totaal bloed. Tot op heden is onze studie de grootste genexpressiestudie in perifeer bloed van ZvH patiënten, met een substantieel aantal presymptomatische patiënten. Verder was onze studie de eerste die gebruik maakte van next generation sequencing genexpressie profilering waarbij in het verleden voornameljk hybridisatie microarrays gebruikt werden. Onze resultaten bevestigen voor het eerst eerder gepubliceerde data, wat het potentieel benadrukt dat next generation sequencing biomarkers kan identificeren die relevant zijn voor de ZvH. We tonen aan dat door het toepassen van relevante statistische methodes, genen die niet het meest significant verschillen in expressie, gebruikt kunnen worden voor een biomarker panel voor ziekteprogressie. Dit biomarker panel kan bestaan uit een combinatie van genen die complementaire informatie bevatten die specifiek zijn voor de ziekte en niet alleen zijn geselecteerd op hun significante expressie. Een dergelijk genenpanel kan zo ontworpen worden dat het een betere weerspiegeling is van de pathologie in de verschillende groepen patiënten.

In **hoofdstuk 4** laten we een unieke integratie zien van verschillende soorten data uit bloed van ons cohort ZvH patiënten en controles: metabolic data en de hiervoor beschreven next generation genexpressie data. Metabolomics kan in potentie veranderingen laten zien die plaatsvinden downstream van genexpressieregulatie in de ZvH. Door een getargete metabolomics aanpak konden we op een robuuste en reproduceerbare manier veranderingen meten die later in longitudinale studies toegepast kunnen worden. Door een multidisciplinaire aanpak en het gebruik van community-driven bioinformatische methoden, konden we de genexpressie data en de metabolomics data van ons ZvH cohort met elkaar verbinden. Een dergelijke integratie van data is een van de eerste multi-level "omics" die in de toekomst nog verder verrijkt kan worden met additionele biologische datasets en zodoende een meer holistische beschrijving van het ZvH fenotype tot stand kan brengen.

Ten slotte wordt in **hoofdstuk 5** een kritische evaluatie gegeven van alle geproduceerde data en de conclusies die hieruit getrokken zijn. Welke kennis er is opgedaan in de hier- boven beschreven hoofdstukken in de context van onze methodologie en opzet van de studies. In het één na laatste deel van dit hoofdstuk worden de beperkingen van ieder hoofdstuk beschreven in relatie tot de inherente limitaties van de technische bestuderingsopzet en de analyseplatforms die gebruikt zijn. Tot slot worden potentiële toekomstige onderzoeksmogelijkheden en vernieuwingen die geassocieerd zijn met de onderzoeks-opzet, bioinformatische analyse en analytische componenten van iedere individuele studie besproken.

Referenties

1. Labadorf A, Hoss AG, Lagomarsino V, Latourelle JC, Hadzi TC, Bregu J et al.: RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. PLoS One 2015, 10: e0143563.

Publication List

1. **Mastrokolias A**, den Dunnen JT, van Ommen GB, 't Hoen PA, van Roon-Mom WM: Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA. BMC Genomics 2012, 13: 28.

2. Zhernakova DV, de Klerk E, Westra HJ, **Mastrokolias A**, Amini S, Ariyurek Y et al.: DeepSAGE reveals genetic variants associated with alternative polyadenylation and expression of coding and non-coding transcripts. PLoS Genet 2013, 9: e1003594.

3. **Mastrokolias A**, Ariyurek Y, Goeman JJ, van DE, Roos RA, van der Mast RC et al.: Huntington's disease biomarker progression profile identified by transcriptome sequencing in peripheral blood. Eur J Hum Genet 2015, 23: 1349-1356.

4. **Mastrokolias A**, Pool R, Mina E, Hettne KM, van DE, van der Mast RC et al.: Integration of targeted metabolomics and transcriptomics identifies deregulation of phosphatidylcholine metabolism in Huntington's disease peripheral blood samples. Metabolomics 2016, 12: 137.

Curriculum Vitae

Anastasios Mastrokolias was born in 1979 in Athens, Hellas. He obtained his National Lyceum Apolytirion Diploma in 1997 from Ionidios School of Piraeus. In 2001 he joined the Hellenic Navy Force where he served his 19-month national military service at the Salamis Naval Base and at the Athens Navy Hospital. In 2003 he got accepted at King's College London where he studied Molecular Genetics and from where he graduated in 2006 with an upper second class bachelor's degree. During his Bachelor of Science research project Anastasios worked on the mutational analysis of lysosomal pathways using Dictyostelium Discoideum. In 2006 he got admitted to Imperial College London where he obtained his Human Molecular Genetics master's degree diploma, also with an upper second class classification. During his Master of Science program Anastasios undertook a 6-month research project with leading scientists at the University College of London Institute of Ophthalmology where he worked on candidate gene studies based on X Linked retinal degeneration. In 2008 Anastasios worked as a postgraduate trainee under an FP7 educational training scholarship at the Foundation of Research and Technology at Heraklion, Crete in Greece where he received training on wet lab and bioinformatics aspects of cDNA microarray analyses. Finally, from May of 2009 until the December of 2013 Anastasios completed the main part of his PhD on Huntington's disease at the department of Human and Clinical Genetics of the Leiden University Medical Center and under the supervision of Prof. GertJan van Ommen and Dr. Willeke van Roon-Mom.

Acknowledgements

I would like to thank my colleagues from the polyglutamine disorders research group Menno, Melvin, Lodewijk and Barry for helping me with all technical and academic aspects of my PhD and making the long busy days at work pass more pleasantly. I would also like to thank them for teaching me how to appreciate different brands of quality Dutch beer and introducing me to the secrets of the very interesting Dutch culture. To avoid forgetting someone I would like to collectively thank all the colleagues and students from the groups of Human Genetics, Bionformatics, Medical Statistics and the LGTC, with whom at one point or another I was involved in fruitful research discussions or having had pleasant afternoons and nights out.

I would like to specifically thank Roos van der Mast, Erik van Duijn and Marloes Hubers for supporting me but also providing me with crucial information about HD that made my bioinformatics and genetics analyses more meaningful and fertile.

Finally I would like to thank the members of my promotion/supervision committee: Willeke for being a great supervisor and teaching me many things, Peter-Bram for his overall support and for his continuous help with all my individual projects, Professor van Ommen for his involvement in the design of my project and sharing with me valuable lessons from his great research experience and last but not least Johan den Dunnen and Yuri van der Burgt for their aid with next generation sequencing and metabolomics respectively.