Longitudinal expression changes are weak correlates of disease progression in Huntington’s disease

Running title: Blood mRNAs longitudinally in Huntington’s

Christopher T. Mitchell¹, Irina Krier², Jamshid Arjomand³,⁴, Beth Borowsky³,⁵, Sarah J. Tabrizi⁶, Blair R. Leavitt⁷, TRACK-HD Investigators⁸ and Ruth Luthi-Carter¹,²,⁹

¹ University of Leicester, Leicester LE1 7RH, United Kingdom
² École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
³ CHDI Foundation, Princeton, NJ 08540, USA
⁴ Current address: FSHD Society, 450 Bedford Street, Lexington, MA 02420, USA
⁵ Current address: Novartis Pharmaceuticals, East Hanover, NJ 07936, USA
⁶ UCL Department of Neurodegenerative disease and Huntington’s disease centre, UCL Queen Square Institute of Neurology, Dementia Research Institute at UCL, London WC1N 3BG, United Kingdom
⁷ Centre for Molecular Medicine and Therapeutics, Department of Medical Genetics, University of British Columbia, Vancouver, BC 75Z 4H4, Canada
⁸ See Appendix
⁹ Corresponding author: relc3@leicester.ac.uk; University of Leicester, Leicester LE1 7RH, United Kingdom

© The Author(s) (2020). Published by Oxford University Press on behalf of the Guarantors of Brain. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

https://mc.manuscriptcentral.com/braincom
Abstract

Huntington’s disease is a severe but slowly progressive hereditary illness for which only symptomatic treatments are presently available. Clinical measures of disease progression are somewhat subjective and may require years to detect significant change. There is a clear need to identify more sensitive, objective and consistent measures to detect disease progression in Huntington’s disease clinical trials. Whereas Huntington’s disease demonstrates a robust and consistent gene expression signature in the brain, previous studies of blood cell RNAs have lacked concordance with clinical disease stage. Here we utilised longitudinally collected samples from a well-characterised cohort of control, Huntington’s disease-at-risk and Huntington’s disease subjects to evaluate the possible correlation of gene expression and disease status within individuals. We interrogated these data in both cross-sectional and longitudinal analyses. A number of changes in gene expression showed consistency within the present study and as compared to previous reports in the literature. The magnitude of the mean disease effect over two years’ time was small, however, and did not track closely with motor symptom progression over the same time period. We therefore conclude that while blood-derived gene expression indicators can be of value in understanding Huntington’s disease pathogenesis, they are insufficiently sensitive to be of use as state biomarkers.

Keywords: Neurodegenerative disease; Biomarker; Expression profiling; Microarray.

Abbreviations: FDR, False Discovery Rate; HTT, Huntingtin; UHDRS, Unified Huntington’s Disease Rating Scale; TMS, Total Motor Score; SAGE, Serial Analysis of Gene Expression; GSEA, Gene Set Enrichment Analysis; EST, Expressed Sequence Tag.
**Introduction**

Huntington’s disease is a chronic, incurable and ultimately fatal neurodegenerative disease affecting an estimated 12 individuals per 100,000 in the UK (Evans *et al*., 2013), and approximately 6 per 100,000 in the USA, Europe and Australia (Pringsheim *et al*., 2012). It is caused by an expanded CAG trinucleotide repeat within the *huntingtin (HTT)* gene inherited in an autosomal dominant manner (Huntington’s Disease Collaborative Research Group, 1993). This gives rise to an extension of a polyglutamine tract near the N-terminus of the huntingtin protein (HTT), which is expressed ubiquitously throughout the body. Huntington’s disease pathogenesis is complex, involving multiple cellular mechanisms including protein aggregation, mitochondrial dysfunction, and transcriptional dysregulation (Luthi-Carter, 2007).

Mutant HTT (mHTT) has a progressive and ravaging effect on the central nervous system (CNS), to the greatest extent in the striatum (Ross and Tabrizi, 2011). Cell death and dysfunction in Huntington’s disease brain regions give rise to a range of motor signs and symptoms, including chorea, alongside cognitive, behavioural and psychiatric disturbances (Walker, 2007). Current treatments rely on symptomatic relief, although several promising experimental treatments are being developed that show potential to slow the disease (Dickey and La Spada, 2018).

Having an established trait marker, the extended CAG repeat, makes Huntington’s disease unusual among neurodegenerative disorders. Presymptomatic predictive genetic testing can confirm whether an individual is carrying this mutation, providing an opportunity to study Huntington’s disease development before onset and the possibility of presymptomatic treatment (Paulsen *et al*., 2008). However, measuring presymptomatic Huntington’s disease
progression is difficult; insidious non-motor changes often precede more apparent extrapyramidal signs, although these effects can occur in either order (Tippett et al., 2017).

Estimating the prospective age of disease onset from the length of the individual’s CAG repeat can serve as a crude proxy for disease state before diagnosis (Langbehn et al., 2004) as age of onset is inversely correlated with the length of CAG expansion (with mitigating effects of interruption by a CAA codon (Wright et al., 2019)). This approach nonetheless has very limited precision, since the CAG repeat length only accounts for a maximum of 60% of variation in disease onset (Wexler et al., 2004), and has an even lower effect in individuals carrying disease alleles with fewer than 50 repeats, which comprise the majority of Huntington’s disease patients. Remaining heritable variation in age of onset may be accounted for in part by mutations in DNA maintenance genes (Genetic Modifiers of Huntington’s Disease Consortium, 2019).

In manifest Huntington’s disease, measurement of disease progression is also complicated by extensive variation of the clinical phenotype and the slow rate of progression. The Unified Huntington’s Disease Rating Scale (UHDRS), a combined clinical assessment of motor function, functional capability, cognition, and behaviour (Huntington Study Group, 1996) is the predominant method of tracking disease severity over time. Observational scales are subjective, however, and therefore prone to bias and inter-rater variability (Henley et al., 2005). More sensitive, robust and objective methods of tracking Huntington’s disease progression are therefore urgently needed to make the most of longitudinal clinical studies and especially to assess the potential benefits of new therapies in clinical trials.
Several different approaches have been taken to develop Huntington’s disease biomarkers that can monitor Huntington’s disease progression, both before and after disease manifestation. For these to be useful they should be highly reproducible, show little variability, and show quantitative, sensitive association with disease progression (Weir et al., 2011). Such biomarkers would considerably reduce the number of participants needed for clinical trials and aid the identification of effective treatments (Paulsen et al., 2006). Furthermore, the search for biomarkers could give further insight into the pathology of Huntington’s disease and potentially reveal novel targets for treatment.

Transcriptional dysregulation is a hallmark of Huntington’s disease, thereby making RNA measurements an attractive area for potential biomarker development. Over almost two decades, genome-wide expression techniques have built a comprehensive picture of the transcriptional alterations mHTT provokes in the brain through multiple cellular mechanisms (reviewed in Seredenina and Luthi-Carter, 2012). Extensive changes in mRNA expression across human Huntington’s disease brain regions has been associated with vulnerability to degeneration, establishing close links between transcriptional dysregulation and Huntington’s disease pathology (Hodges et al., 2006). Several studies have identified parallel changes in Huntington’s disease mouse models (Luthi-Carter et al., 2000; Kuhn et al., 2007) which show progression over time and correlation with losses in motor function and behavioural disturbances (Luthi-Carter et al., 2002a; Kuhn et al., 2007; Giles et al., 2012). Not all expression changes contribute to the disease process; some changes appear to have no pathological effect (Rudinskiy et al., 2009), while others are compensatory (Seredenina et al., 2011).
The transcriptional regulatory pathways known to be altered by mHTT are involved in gene expression in multiple tissues throughout the body in model systems (Sugars and Rubinsztein, 2003). Parallel gene expression changes have been observed in R6/2 mouse and human Huntington’s disease muscle (Luthi-Carter, et al., 2002b; Strand et al., 2005) and correlate with muscle atrophy. Expression changes have also been observed in Huntington’s disease mouse liver and pancreas (Andreassen et al., 2002; Chiang et al., 2011), and brown and white adipose tissues (Weydt et al., 2006; Phan et al., 2009), including genes related to energy expenditure (McCourt et al., 2016) that could contribute to weight loss. Data from model systems provide evidence that mHTT accumulation in peripheral tissues may cause functional abnormalities independently of neurodegeneration, including muscle strength, osteoporosis, cardiac failure and impaired glucose tolerance (van der Burg et al., 2009). Transcriptional changes have also been described in human Huntington’s disease fibroblasts (Marchina et al., 2014).

Peripheral blood, being a renewable, easy-to-collect tissue, is an attractive substrate for biomarker development. Several exploratory studies have reported transcriptomic changes in Huntington’s disease blood and highlighted specific RNAs as potential biomarkers. These have used both genome-wide (microarray, DeepSAGE) and targeted reverse transcription real-time quantitative PCR (RT-qPCR) gene expression analyses (Borovecki et al., 2005; Runne et al., 2007; Lovrecic et al., 2010; Hu et al., 2011; Cesca et al., 2015; Mastrokolias et al., 2015). Despite the number of potential biomarkers identified, there has been little overlap between the reported signatures across these studies. Results of focused attempts to cross-validate results have also been largely negative (Runne et al., 2007; Lovrecic et al., 2009; Mastrokolias et al., 2015; Zadel et al., 2018).
Additional studies have taken a somewhat different approach by considering transcriptional changes within the contexts of identified etiologic processes in Huntington’s disease (Krzysztoń-Russjan et al., 2013; Träger et al., 2014; Miller et al., 2016). If differential expression of genes identified by these studies can be replicated, they may also warrant consideration as potential biomarkers. Lack of concordance across these studies has hindered the implementation of RNA-based Huntington’s disease biomarkers. Discrepancies have been attributed to technological limitations, methodological inconsistencies, and human blood sample variability (Rodrigues et al., 2018).

The development of blood-based state biomarkers for Huntington’s disease has also been limited by the lack of longitudinal studies. On the positive side, sequential analyses of samples from individual subjects might minimise effects of patient-to-patient variability (Tabrizi et al., 2009). On the other hand, cross-sectional studies may maximise detection of RNA changes in Huntington’s disease versus control subjects because they reflect a long and cumulative process of mHTT exposure (Weir et al., 2011). To be useful, however, longitudinal measures need to detect changes that occur in a timeframe relevant for clinical trials, e.g. 1-3 years.

It has also been suggested that independence of neurological and peripheral Huntington’s disease pathologies might fundamentally limit the utility of using transcriptional changes in the blood to track the progression of Huntington’s disease (Zadel et al., 2018). However, previous claims of parallels between transcriptional dysregulation in the blood and brain have challenged this notion (Mastrokoli et al., 2015; Mina et al., 2016). One study purported to show similarities between Huntington’s disease blood and brain expression signatures through gene set enrichment analysis (GSEA) (Hensman Moss et al., 2017). However, none of the individual genes within these sets was significantly differentially expressed in Huntington’s
disease compared to controls after appropriate false discovery rate (FDR) correction, raising
caveats in these conclusions. Questions therefore remain over the nature and degree of
similarities between neurological and peripheral transcriptional dysregulation, and whether
these changes progress at rates that allow blood mRNA to be used as a reliable predictor of
neurological disease progression.

The present study aimed to evaluate potential mRNA state biomarkers for Huntington’s disease
in human blood samples collected in the TRACK-HD study (Tabrizi et al., 2009). TRACK-
HD was a large multi-site observational study in which quantitative MRI, cognitive, motor,
and neuropsychiatric assessments were collected in more than 350 subjects, comprising
individuals with an Huntington’s disease diagnosis, individuals at risk for Huntington’s disease
(carrying an expanded HTT allele) and control subjects. The high standardisation of these
measures across sites and the relatively long 24-month follow up time with individual subjects
provided an opportunity to compare established clinical indicators with additional phenotypic
measures. We therefore undertook transcriptomic analyses in this well characterised subject
population.

We performed both cross-sectional and longitudinal microarray gene expression analyses of
TRACK-HD blood samples from Huntington’s disease subjects in early clinical stages of
disease, presymptomatic carriers of the Huntington’s disease mutation, and controls, collected
from the same set of individuals two years apart (see also Cohort below). These data were
analysed cross-sectionally at both timepoints and also used to construct a longitudinal analysis
of gene expression and disease progression. Both of these analyses were then interrogated to
identify novel candidate biomarker genes, and to test progressive differential expression of
genes identified in previous studies, using publicly available datasets (Borovecki et al., 2005;
Runne et al., 2007; Lovrecic et al., 2010; Hu et al., 2011; Krzysztoń-Russjan et al., 2013; Träger et al., 2014; Cesca et al., 2015; Mastrokolias et al., 2015; Miller et al., 2016; Zadel et al., 2018), as well as an independent TRACK-HD sample cohort. Finally, we sought to shed new light on, or to reproduce previous evidence of, molecular mechanisms of Huntington’s disease pathogenesis identified using GSEA (Hensman Moss et al., 2017). We confirm that blood RNA changes may reflect small Huntington’s disease-related effects, but find that such measures are likely to be unsatisfactory as clinical state biomarkers of Huntington’s disease.
Materials and methods

Study cohort

Whole blood samples were collected from participants in the Track-HD in London and Vancouver (Supplementary Tables 1 and 2) into PAXgene Blood RNA tubes (Qiagen). Presymptomatic subjects (UHDRS Total Motor Score (TMS) ≤5) were further stratified based on their estimated time to diagnosis based on their current age and their estimated age of onset (calculated using the repeat length of their expanded HTT allele) according to (Langbehn et al., 2004). Subjects were then classified as farther (>10.8 years) from predicted onset (preHD A) or nearer (<10.8 years) to predicted onset (preHD B) (Supplementary Table 3). Diagnosed early symptomatic Huntington’s disease participants were classified by the Total Functional Capacity (TFC) score of the UHDRS into Stage 1 (zHD1, TFC = 11-13) and Stage 2 (zHD2, TFC = 7-10) disease subgroups (Tabrizi et al., 2009, 2012). The subset of samples represented balanced numbers of preHD and HD cases at each stage. The study was approved by the local ethics committees and written informed consent was obtained from each subject according to the Declaration of Helsinki.

Microarray analyses

DNA microarray analyses were conducted using Affymetrix Human Genome U133 Plus 2.0 Arrays. Expression values were normalised using the RMA algorithm included in the Limma package (Bioconductor). A linear regression model was fitted and adjusted to remove the effects of sample origin (London or Vancouver), gender and age. P-values were adjusted using the False Discovery Rate (FDR) method (Benjamini and Hochberg, 1995).

Cross-sectional comparisons of microarray gene expression in Huntington’s disease versus control samples at each time-point were performed using R. Cross-sectional microarray
analyses of differential expression were performed in Year 1 samples of 37 diagnosed Huntington’s disease subjects (18 zHD1, 19 zHD2) versus 24 controls. Year 3 samples were tested in the same manner and included 32 diagnosed HD subjects (16 zHD1, 16 zHD2) and 20 controls. The Year 3 samples were collected from a subset of participants whose samples had been analysed in Year 1. We also conducted cross-sectional analyses including all presymptomatic samples (data not shown), but did not draw conclusions from these analyses, as no differential expression could be detected.

**Analyses of gene expression vs disease state**

Cohorts for longitudinal analyses comparing microarray expression and Total Motor Score (TMS) were selected by comparing mean TMS scores of sample groups to the control group using Dunnett’s multiple comparisons test of an ordinary one-way ANOVA, performed using Prism 7 (GraphPad; SupplementaryTable 3). On this basis, we excluded preHD A samples, as the median TMS scores of these groups were not significantly different from that of the control group.

Longitudinal analyses included samples from 15 presymptomatic Huntington’s disease gene carriers (all preHD B) and 32 diagnosed Huntington’s disease subjects (16 zHD1, 16 zHD2). An additional longitudinal analysis of gene expression was performed in samples from a control cohort of 20 individuals for sake of comparison; as expected, this group of individuals did not show a change in TMS over the two-year period.

Microarray gene expression data was analysed and normalised using R, as described above. Expression values were unlogged and the change in gene expression (ΔExpression) was calculated for each individual across two years. Change in Total Motor Score (ΔTMS) was
plotted against ΔExpression and fitted with a linear regression model. Pearson’s correlation coefficient was calculated and the \( p \)-values were also adjusted using FDR. A cross-sectional analysis of relationship between gene expression and TMS was performed at each individual time-point (Year 1, Year 3). A longitudinal gene expression analysis was also performed on the control samples (without consideration of ΔTMS).

**Pathway analyses**

Gene set enrichment analysis (GSEA) (Broad Institute; Subramanian *et al.*, 2005) was performed on expression data of 97 Year 1 samples, implementing the analysis parameters of Hensman Moss *et al.* (2017). This included a control group of 24 samples and a Huntington’s disease group of 73 samples (including all four presymptomatic and symptomatic Huntington’s disease groups). Genes were ranked by t test of differential expression between the Huntington’s disease and control groups. We then tested 5,519 overlapping pathway gene sets obtained from the Gene Ontology (GO), REACTOME, Kyoto Encyclopedia of Genes and Genomes (KEGG), NCI Pathway Interaction Database (PID), and BioCarta databases, focusing testing on sets containing between 3 and 500 genes. The GSEA method calculated enrichment scores for each set reflecting representation of that pathway at the top and bottom of the ranked gene list (i.e. the most probably differentially expressed genes in Huntington’s disease, both increased and decreased). To estimate the statistical significance of the GSEA enrichment scores obtained, we used a sample label permutation strategy to estimate the null distribution (as recommended by Mooney and Wilmot, 2015). This involved permuting the phenotype labels 1,000 times and generating enrichment scores and \( p \) values in the same manner. Finally, \( q \) values were calculated from the \( p \) value distributions. \( q \) values were calculated separately for up- and down-regulated gene sets.
Data availability

The NCBI Gene Expression Omnibus accession number for the microarray data reported in this paper is GSE 135589.
Results

Microarray differential expression analyses in Huntington’s disease versus control subjects from the TRACK-HD study

The Track-HD study provided an opportunity to obtain a considerable number of samples collected in a highly standardised manner from a clinically well-characterised patient cohort (Tabrizi et al., 2009). We reasoned that analyses of TRACK-HD samples therefore maximised the chances of discovering or validating blood RNA biomarkers for Huntington’s disease progression. To this end, microarray gene expression analyses were conducted from Year 1 total blood RNA samples from 37 manifest Huntington’s disease subjects, 35 presymptomatic Huntington’s disease gene carriers, and 24 controls (Tabrizi et al., 2009) (see also Materials and Methods and Supplementary Tables 1 and 2 for further details). As most of these subjects would carry on participating in the study to its completion, we also had the opportunity to perform longitudinal analyses of Year 3 samples (see below).

In order to highlight the largest potential disease-related changes, we first conducted analyses of samples from symptomatic Huntington’s disease subjects versus controls. For the Year 1 microarray data, these comprised samples from 37 Huntington’s disease subjects and 24 controls, while Year 3 data comprised 32 of Huntington’s disease subjects and 20 controls. A linear regression model was fitted to analyse the relationship between expression and condition. The model also included gender, age and sample collection site (London or Vancouver) to avoid confounding effects.

Analyses of Year 1 samples detected 2,575 probesets of 54,675 that reported differential expression at cutoff of nominal p < 0.05 between the control and Huntington’s disease groups (1,125 increased and 1,450 decreased in Huntington’s disease (Fig. 1 and Supplementary Table
Supplementary Table 4 also lists the 20 probesets with most significant differential expression and the functions of proteins encoded by the genes they represent, which include gene expression (RORC, ZFP42) and involvement in cell structure (IQCA1, DYNC1I2). This number of changes does not convincingly exceed chance levels, however. Also consistent with the null hypothesis, multiple testing correction by a well-established method assigned all FDR p values (q values) to be greater than 0.05.

Because most participants remained in the study for its duration, analyses of RNA samples from most of the same subjects were analysed again after a two-year interval (TrackHD study Year 3). In the Year 3 analysis, 3,261 probesets showed a potentially significant change of expression between control and Huntington’s disease subjects (by the criterion of nominal \( p < 0.05 \); 1,962 up-regulated, 1,299 down-regulated (Fig. 1 and Supplementary Table 5)), but still no individual probesets showed differential expression that could be considered significant after correction for multiple testing.

The 20 probesets with the lowest p values in comparisons between Huntington’s disease and control samples are also listed in Supplementary Table 5; these include probes for genes whose products are involved in gene regulation (FOXK2, MRGBP), are active as serine/threonine kinases (SGK494, AKT3, ATM) and comprise another glutamine-rich protein (QRICH1). Despite the apparent low level of differential expression signal, there was a small overlap of differential gene expression signatures between the two time-points (Table 1). Of the probesets meeting criteria of nominal \( p < 0.05 \) in the Year 1 data, 299 meet the same criteria for differential expression in Year 3, with 270 of these the changing in the same direction. The high ratio of concordant versus discordant changes across two separate experimental analyses suggests that these positive results are not due to technical variability.
Focused analyses of blood RNA differential expression measures identified by previous Huntington’s disease studies

A disease state biomarker should ideally be universally applicable, such that it is detectable in Huntington’s disease versus control samples in different laboratories and independent sample cohorts. Replication of potential blood RNA biomarker signatures in Huntington’s disease has been generally inconsistent in this regard. We nonetheless examined whether RNA changes identified as potential biomarkers in previous studies were represented in our analyses of differential expression between Huntington’s disease subjects and controls (Supplementary Table 6). This included a panel of 12 potential biomarker genes identified through microarray analysis of whole blood by (Borovecki et al., 2005) and 2 further genes discovered through machine learning analysis of the same microarray data (Lovrecic et al., 2010). It also examined the expression of the histone gene H2AFY identified by (Hu et al., 2011) and subsequently analysed in (Mastrokolias et al., 2015). Five additional differentially expressed blood RNAs identified by (Mastrokolias et al., 2015) using DeepSAGE were also evaluated. Finally, we tested 7 other RNAs showing differential expression by microarray analyses of Huntington’s disease and control blood samples published recently by (Zadel et al., 2018).

We also considered the results of five additional RNA studies that focused on putative mechanisms of disease. One RNA was included from a previous study of inflammatory processes in lymphocytes (Runne et al., 2007). Two further candidates were chosen from a study that considered Ca\textsuperscript{2+} homeostasis and neuroprotection using mononuclear blood cells (Cesca et al., 2015). Six genes were identified based on a study of metabolic- and cellular stress-related RNA measures in whole blood compared to other clinical, anthropometric and biochemical parameters (Krzysztoń-Russjan et al., 2013). Larger gene sets were also considered based on data from Huntington’s disease monocytes: these comprised the 20 most
significantly dysregulated genes from a qPCR study of the NFκB pathway (Träger et al., 2014) and the top 20 from a full RNA-Seq analysis (Miller et al., 2016).

All of these candidate genes had been previously subject to technical validation using qPCR in their respective initial reports. The majority of genes from earlier studies had been selected for independent validation in additional cohorts, with limited success. To the best of our knowledge the present study is the first to attempt to reproduce the positive biomarker study results of Cesca et al. (2015) and Krzysztoń-Russjan et al. (2013).

In total, 70 unique transcripts identified in previous studies were tested in our two cross-sectional differential expression analyses. In the Year 1 Huntington’s disease versus control analysis, 8 of these showed statistically significant differential expression ($p < 0.05$, without FDR correction). However, only two of these transcripts (SAP30 and ENSA) were dysregulated in the same direction as reported in their original studies. The protein encoded by SAP30 is a subunit of the histone deacetylase complex, noteworthy due to the well-established interference by mHTT of the regulatory machinery of gene expression. ENSA, which has been suggested as a candidate gene for insulin dependent diabetes mellitus (Héron et al., 1999), encodes the endogenous ligand of the ABCC8 regulatory subunit of $K_{ATP}$ channels. The Year 3 analysis showed even less concordance: 6 genes were significantly dysregulated, yet only LTBR (which encodes the lymphotoxin-β receptor) showed a change in the same direction as previously reported. The candidate biomarker panel identified by Krzysztoń-Russjan and colleagues (2013) showed a high proportion of statistically significant transcripts in both years; however all of these were dysregulated in the opposite direction to that which was originally reported, which suggests that these are genes that show a large degree of heterogeneity of expression (or detection) between individuals, independent of Huntington’s disease status.
Global longitudinal analysis of gene expression vs Total Motor Score

It is essential for a disease state biomarker to show a clear relationship to disease progression. A robust way of assessing this is to identify and evaluate potential biomarkers is to follow this potential relationship in a single cohort longitudinally. This also provides a means of testing biomarkers that largely circumvents the interference of inter-individual variability.

Along with blood RNA microarray data, the Track-HD study collected clinical measurements including Total Motor Score (TMS), the component of UHDRS that assesses 31 motor symptoms of Huntington’s disease; the higher the score, the more severe the disease state (Huntington Study Group, 1996). In order to focus our analysis on individuals with significantly progressive disease, the cohort for the longitudinal analysis was selected through analysis of the mean TMS score change across two years of each sample group (Supplementary Table 3). As such, the analysis included 15 presymptomatic HTT mutation carriers and 32 diagnosed Huntington’s disease subjects. In this analysis, ΔTMS was used as a scale of disease progression across the two years and was plotted against ΔExpression in each individual. For each microarray probeset, a linear regression analysis was performed to model the slope and statistical significance of the relationship between the two variables (Fig. 2). Pearson’s correlation coefficient was used to determine the degree to which they were related.

The 20 probesets that showed most significant correlation are displayed in Table 2. 4,273 probesets, out of a total of 54,675, showed a statistically significant correlation ($p < 0.05$) (Supplementary Table 7). 2,962 of these showed positive correlation and 1,311 an inverse relationship. Furthermore, 394 probesets showed very highly significant correlation ($p < 0.005$). Notably, however, when multiple testing correction was applied, the significance of
association for all probesets was \( p > 0.05 \). The most significant 20 probesets included genes that encoded proteins involved in mitochondrial maintenance (\( HSPA9 \)), transcriptional regulation (\( NACC1 \)), and cell signalling pathways (\( LYPD6 \)).

To further assess the genes identified for their ability to detect Huntington’s disease progression, a comparative longitudinal analysis was performed in a group of 20 controls from the Track-HD study. It would be expected that expression changes associated with Huntington’s disease progression would not be seen in a control group. However, in this analysis, 5,392 probesets showed very highly statistically significant (adj. \( p < 0.005 \)) differential expression. These included 319 of the 4,273 statistically significant probesets identified in the Huntington’s disease correlation analysis. This result suggests that some of the signal observed might be due to technical variation (e.g. due to microarray batch processing).

The utility of a potential biomarker would be increased if it displayed a clear association with disease state at cross-sectional time-points, to inform clinical trial recruitment, for example. Therefore, cross-sectional analyses of the relationship between TMS and expression were also conducted. These included the sample groups with a significantly different mean TMS than the control group (at both timepoints) (Supplementary Table 3). In the Year 1 samples, 3,991 probesets showed statistically significant (\( p < 0.05 \)) association with TMS, including 467 of the changes identified in the longitudinal correlation analysis (Supplementary Table 8). At Year 3, the number of probesets with a statistically significant (\( p < 0.05 \)) relationship between TMS and expression was 1,601. Of these, 255 probesets had been also identified as significant in our longitudinal profiling (Supplementary Table 9).

**Evaluation of gene signatures from previous studies in our longitudinal analysis**
Along with a lack of replication of potential Huntington’s disease blood biomarkers across independent cohorts, longitudinal assessments of these potential biomarkers have not previously been carried out. Therefore, candidate genes from three cross-sectional studies were evaluated in our longitudinal analysis cohort to provide further evaluation of their suitability as Huntington’s disease biomarkers (Supplementary Table 10). It would be expected that if a gene is differentially expressed between controls and Huntington’s disease in these studies, it would be correlated with disease progression in the same direction.

Of the same 70 transcripts described above, the expression of only four changed at a rate that could be statistically significantly associated with disease progression over the two year interval ($p < 0.05$ without FDR correction). These are plotted in Fig. 3, and Table 3 summarises their functions and correlation statistics.

ROCK1, which encodes a protein serine/threonine kinase (Borovecki et al., 2005), showed the strongest correlation of any of the previously identified candidate biomarkers in our longitudinal analysis. To the contrary, however, it was not detected as differentially expressed in our cross-sectional differential gene expression analyses (nominal $p > 0.05$), where it showed a trend in the opposite direction.

SAP30 also showed an upregulation over the two year interval that was significantly associated with disease progression. This corroborates the findings of (Zadel et al., 2018) and was also detected in the cross-sectional analysis our Year 1 samples.

Likewise, LTBR showed significant positive correlation with TMS progression, consistent with a previous study of NFκB pathway disturbances in Huntington’s disease (Träger et al., 2014).
LTBR also showed upregulation (nominal p<0.05) in our Year 3 differential expression analysis.

The other gene showing significant correlation with disease progression was IL23A, which encodes an inflammatory cytokine. However, IL23A was reported to be upregulated in the original paper (Miller et al., 2016), but showed negative association with disease progression in our analysis.

**Functional pathway analyses of a putative Huntington’s disease signature**

Several studies have interrogated specific functional pathways represented in gene expression signatures from Huntington’s disease blood using gene set enrichment analysis (GSEA). In their analysis of Huntington’s disease monocytes, Miller et al. (2016) demonstrated significant enrichment of upregulated genes relating to inflammation, immunity, and intracellular signalling, and downregulated genes relating to vacuolar and catabolic functions.

Hensman Moss et al. (2017) conducted a GSEA of differentially expressed genes from a cohort of 186 Huntington’s disease and preHD subjects and 49 control blood samples, a subset of which were also analysed in the present study. They reported significant associations of RNA measures with TMS that represented pathways similar to those reported by Miller et al. (2016). However, as in our study, they did not observe statistically significant differential expression in individual genes between Huntington’s disease subjects and controls.

We first conducted GSEA analyses of differentially expressed genes across our entire Year 1 cohort. Of the 5,519 gene sets tested, none were found to be statistically significantly overrepresented (q value cutoff of 0.05) (Supplementary Tables 11 and 12). We then repeated
the analysis with the exclusion of samples in the presymptomatic groups (as we did not observe statistically significant difference in mean TMS scores in this group compared to controls). Again, we did not observe any significant enrichment of pathways (Supplementary Tables 13 and 14). We discuss possible biological and technical explanations for these discrepancies below.
Discussion

In this study we aimed to identify a transcriptional signature in whole blood that could be used as a Huntington’s disease state biomarker, and to evaluate candidate RNA biomarkers identified in previous studies. These aims were pursued through differential expression analyses of blood samples from manifest Huntington’s disease subjects, presymptomatic subjects carrying the Huntington’s disease mutation and control subjects. Both cross-sectional and longitudinal analyses were applied and correlated with UHDRS measures. To our knowledge, this is the first time a longitudinal analysis in a single Huntington’s disease cohort has been used to try to identify novel transcriptional Huntington’s disease biomarkers and to test previously reported candidates.

The high level of variability in blood RNA expression may be a substantial barrier to the clinical application of gene expression changes as clinical biomarkers. Here, differential expression in Huntington’s disease subjects versus controls was investigated using multiple linear regression modelling, to attempt to assign potential confounding effects of collection site, gender and age \textit{a priori}. The p-values discussed here are not adjusted for multiple testing, because adjustments using the standard FDR method resulted in p-values greater than 0.05 and therefore would not have aided us in discriminating top candidates. This likely reflects the disease effect sizes being very small, particularly considering the available sample numbers and other variables (Ideker \textit{et al.}, 2011). Quality control measures of the microarray data showed low technical variability, which was thus ruled out as a confounding factor.

Furthermore, our pathway analyses did not show convincing evidence of dysregulation of any particular functional cellular or molecular pathways. This does not exclude the possibility of functional pathway dysregulation (Hensman Moss \textit{et al.}, 2017), but differences from previous
approaches and their caveats need to be considered. Here, we randomised the disease status assignments of the samples to evaluate the background signal in our analyses. These maintain the data structure of the defined background, thereby circumventing confounding effects due to tissue-specific gene expression. For example, in blood, signatures are dominated by immune cell and reticulocyte RNAs, which would typically represent specific biological pathways compared to randomly sampled gene sets. In such a case, false-positive assignments of disease-related differential gene expression to pathways of immune cell signalling and inflammation become more likely.

We also evaluated the relationship between differential expression attributed to Huntington’s disease and continuous quantitative measures of disease status (UHDRS TMS) using linear correlation. The slopes of the resultant lines were also low, further supporting a small Huntington’s disease effect size. The small effect size could however reflect the proportion of blood cell subtypes susceptible to Huntington’s disease-related changes (Björkqvist et al., 2008; Wild et al., 2011).

We nonetheless entertained the possibility that small transcriptional changes in peripheral blood cells might be informative as state biomarkers in individual Huntington’s disease patients. However, longitudinal analyses of the 70 genes identified by previous cross-sectional studies (not including those from Hensman Moss et al., 2016, whose cohort overlapped with our own) detected only three that showed significant association with disease state in the same direction (ROCK1, SAP30, LTBR). The very minimal overlap of Huntington’s disease-associated gene expression changes from previous studies in our 2-year longitudinal analyses in does not bode well for their suitability as biomarkers. Again, the most likely explanation is
that the small Huntington’s disease effect size is insufficient compared to the inherent biological and technical variability of blood RNA differential expression measures.

The functions of the proteins encoded by genes with the most reproducible differential expression in Huntington’s disease samples are nonetheless consistent with idea that mHTT has pathological effects in blood. For example, ROCK1 is a member of the Rho kinase pathway and is involved in cell motility and cytoskeleton organisation. This pathway has previously been documented to be dysregulated in Huntington’s disease blood, post-mortem brain samples and mouse models (Narayanan et al., 2016).

SAP30 is a component of the histone deacetylase complex, and its dysregulation may relate to changes in the post-translational modification of histone proteins, which are known to occur in Huntington’s disease and Huntington’s disease model systems (Glajch and Sadri-Vakili, 2015; Naia et al., 2017). Likewise, dysregulation of LTBR, a receptor involved in cytokine release, contributes to the wider evidence of dysregulated inflammation in Huntington’s disease (Hensman Moss et al., 2017; Miller et al., 2016). Clearly, these changes and their functional consequences could help understand Huntington’s disease pathology in peripheral tissues, and how this differs from, or is similar to, mHTT’s effects in the brain (Weir et al., 2011).

Intra- and inter-individual variation and the potential for uncorrelated peripheral and central pathologies seem to be at the root of difficulties discerning a temporal association between blood gene expression and the neurological symptoms of Huntington’s disease. In contrast, other biomarker approaches may prove more useful. These include neuroimaging (Rees et al., 2013), alternative measures of motor function (e.g. speeded tapping, tongue force and eye movements) (Tabrizi et al., 2009) and other molecular analyses, including measurements of...
neurofilament proteins (Johnson et al., 2018; Byrne et al., 2018), HTT protein (Wild et al., 2015; Byrne et al., 2018), oxidative stress markers, inflammatory proteins, and endocrine markers (Björkqvist et al., 2008; Wild et al., 2011; Scahill et al., 2012).

In conclusion, changes in mRNA expression in whole blood showed limited association with Huntington’s disease pathology and progression. These changes may reflect interesting pathological or compensatory roles within blood cells, but their effect sizes in whole blood are very small, thereby constraining their deployment as state biomarkers for Huntington’s disease.
Acknowledgements

We thank Flav Giorgini, Karishma Joshi and Christiane Reick for critical reading of the manuscript.
We also acknowledge the invaluable support of the TRACK-HD participants, their family members, and the TRACK-HD staff.

Funding

This study was funded primarily by the CHDI Foundation. Salary support was also provided by the University of Leicester and the École Polytechnique Fédérale de Lausanne.

Competing Interests

The authors declare no competing interests.
Appendix

**TRACK-HD Investigators:** N. Arran, E. Axelson, E. Bardinet, N. Bechtel, J. Callaghan, J.C. Campbell, M. Campbell, D.M. Cash, A. Coleman, D. Craufurd, R. Dar Santos, J. Decolongon, E.M. Dumas, A. Durr, N.C. Fox, E. Frajman, C. Frost, S.L. Hicks, N. Hobbs, A. Hoffman, C. Jauffret, H. Johnson, R. Jones, C. Jurgens, D. Justo, S. Keenan, C. Kennard, P. Kraus, N. Lahiri, B. Landwehrmeier, D.R. Langbehn, S. Lee, S. Lehericy, C. Marelli, C. Milchman, W. Monaco, K. Nigaud, R. Ordidge, A. O’Regan, G. Owen, T. Pepple, S. Queller, J. Read, R. Reilmann, R.A.C. Roos, H.D. Rosas, M. Say, R.I. Scahill, J.C. Stout, A. Sturrock, E.P. ’t Hart, A.J. Tobin, R. Valabrègue, S.J.A. van den Bogaard, J. van der Grond, C. Wang, K. Whitlock, E. Wild, M.-N. Witjes-Ane.
References

Andreassen OA, Dedegolu 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. Neurobiology of Disease 2002; 11: 410–24.

Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995; 57: 289–300.

Björkvist M, Wild EJ, Thiele J, Silvestroni A, Andre R, Magnusson A, et al. A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington’s disease. Journal of Experimental Medicine 2008; 205: 1869–77.

Borovecki F, Lovrecic L, Zhou J, Jeong H, TF, Rosas HD, et al. Genome-wide expression profiling of human blood reveals biomarkers for Huntington’s disease. Proceedings of the National Academy of Sciences of the United States of America 2005; 102: 11023–28.

Byrne LM, Rodrigues FB, Johnson EB, De Vita E, Blennow K, Sechill R, et al. Cerebrospinal fluid neurogranin and TREM2 in Huntington’s disease. Scientific Reports 2018; 8: 1–7.

Cesca F, Bregant E, Peterlin B, Zadel M, Dubsky de Wittenau G, Siciliano G, et al. Evaluating the SERCA2 and VEGF mRNAs as potential molecular biomarkers of the onset and progression in Huntington’s disease. PLoS ONE 2015; 10: e0125259.

Chen C-M, Wu Y-R, Cheng M-L, Liu J-L, Lee Y-M, Lee P-W et al. The dysfunction of hepatic transcriptional factors in mice with Huntington’s Disease. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease 2011; 1812: 1111–20.

Dickey AS and La Spada AR. Therapy development in Huntington disease: From current strategies to emerging opportunities. [Review]. American Journal of Medical Genetics Part A 2018; 176: 842–61.

Evans SJW, Douglas I, Rawlins MD, Wexler NS, Tabrizi SJ, Smeeth L et al. Prevalence of adult Huntington’s disease in the UK based on diagnoses recorded in general practice records. Journal of Neurology Neurosurgery and Psychiatry 2013; 84: 1156–60.

Genetic Modifiers of Huntington’s Disease (GeM-HD) Consortium. CAG repeat not polyglutamine length determines timing of Huntington’s disease onset. Cell 2019; 178: 887-900.

Giles PJ, Elliston LA, Higgs G, Brooks SP, Dunnett SB, Jones L. Longitudinal analysis of gene expression and behaviour in the HdhQ150 mouse model of Huntington’s disease. Brain Research Bulletin 2012; 88: 88199–209.

Glajch KE and Sadri-Vakili G. Epigenetic mechanisms involved in Huntington’s disease pathogenesis. [Review]. Journal of Huntington’s Disease 2015; 4: 1–15.

Heemskerk A-W and Roos RAC. Aspiration pneumonia and death in Huntington’s disease. PLoS Currents 2012; 4: RRN1293.
Henley SM, Bates GP, Tabrizi SJ. Biomarkers for neurodegenerative diseases. [Review]. Current Opinion in Neurology 2005; 18: 698-705.

Hensman Moss DJ, Henley SM, Bates GP, Tabrizi SJ, et al. Huntington’s disease blood and brain show a common gene expression pattern and share an immune signature with Alzheimer’s disease. Scientific Reports 2017; 7: 44849.

Héron L, Virsolvy A, Apiou F, Cam AL, Bataille D. Isolation, characterization and chromosomal localization of the human ENSA gene that encodes alpha-endosulfine a regulator of beta-cell K (ATP) channels. Diabetes 1999; 48: 1873–76.

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.

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. Proceedings of the National Academy of Sciences of the United States of America 2011; 108: 17141–46.

Huntington Study Group. Unified Huntington’s disease rating scale: Reliability and consistency. Movement Disorders 1996; 11: 136–42.

Huntington’s Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington’s disease chromosomes. Cell 1993; 72: 971–83.

Ideker T, Dutkowski J, Hood L. Boosting signal-to-noise in complex biology: Prior knowledge is power. [Review]. Cell 2011; 144: 860–63.

Johnson EB, Byrne LM, Gregory S, Rodrigues FB, Blennow K, Durr A, et al. Neurofilament light protein in blood predicts regional atrophy in Huntington disease. Neurology 2018; 90: e717–23.

Krzysztoń-Russjan J, Zielonka D, Jackiewicz J, Kuśmirek S, Bubko I, Klimberg A, et al. A study of molecular changes relating to energy metabolism and cellular stress in people with Huntington’s disease: looking for biomarkers. Journal of Bioenergetics and Biomembranes 2013; 45: 71–85.

Kuhn A, Goldstein DR, Hodges A, Strand AD, Sengstag T, Kooperberg C, et al. Mutant huntingtin’s effects on striatal gene expression in mice recapitulate changes observed in human Huntington’s disease brain and do not differ with mutant huntingtin length or wild-type huntingtin dosage. Human Molecular Genetics 2007; 16: 1845–61.

Langbehn DR, Brinkman RR, Falush D, Paulsen JS, Hayden MR. A new model for prediction of the age of onset and penetrance for Huntington’s disease based on CAG length. Clinical Genetics 2004; 65: 267–77.

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. Movement Disorders: Official Journal of the Movement Disorder Society 2009; 24: 2277–81.
Lovrecic L, Slavkov I, Džeroski S, Peterlin B. ADP-ribosylation factor guanine nucleotide-exchange factor 2 (ARFGEF2): A new potential biomarker in Huntington’s disease. Journal of International Medical Research 2010; 38: 1653–62.

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. Human Molecular Genetics 2000; 9: 1259–71.

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. Human Molecular Genetics 2002a; 11: 1911–26.

Luthi-Carter R, Strand AD, Hanson SA, Kooperberg C, Schilling G, La Spada AR, et al. Polyglutamine and transcription: gene expression changes shared by DRPLA and Huntington’s disease mouse models reveal context-independent effects. Human Molecular Genetics 2002b; 11: 1927–37.

Luthi-Carter R. Huntington’s and other polyglutamine diseases: many effects of single gene mutations. [Review]. Drug Discovery Today: Disease Mechanisms 2007; 4: 111–19.

Marchina E, Misasi S, Bozzato A, Ferraboli S, Agosti C, Rozzini L, et al. Gene expression profile in fibroblasts of Huntington’s disease patients and controls. Journal of the Neurological Sciences 2014; 337: 42–46.

Mastrokolias A, Ariyurek Y, Goeman JJ, van Duijn E, Roos RAC, 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; 23: 1349–56.

McCourt AC, Jakobsson L, Larsson S, Holm C, Piel S, Elmér E, Björkqvist M. White adipose tissue browning in the R6/2 mouse model of Huntington’s disease. PLoS ONE 2016; 11: e0159870.

Miller JRC, Lo KK, Andre R, Hensman Moss DJ, Träger U, Stone TC, et al. RNA-Seq of Huntington’s disease patient myeloid cells reveals innate transcriptional dysregulation associated with proinflammatory pathway activation. Human Molecular Genetics 2016; 25: 2893–2904.

Mina E, van Roon-Mom W, Hettne K, van Zwet E, Goeman J, Neri C, et al. Common disease signatures from gene expression analysis in Huntington’s disease human blood and brain. Orphanet Journal of Rare Diseases 2016; 11: 97.

Mooney MA and Wilmot B. Gene Set Analysis: A step-by-step guide. American Journal of Medical Genetics Part B Neuropsychiatric Genetics 2015; 168: 517–27.

Naia L, Cunha-Oliveira T, Rodrigues J, Rosenstock TR, Oliveira A, Ribeiro M, et al. Histone deacetylase inhibitors protect against pyruvate dehydrogenase dysfunction in Huntington’s disease. The Journal of Neuroscience 2017; 37: 2776–94.

Narayanan KL, Chopra V, Rosas HD, Malarick K, Hersch S. Rho kinase pathway alterations in the brain and leukocytes in Huntington’s disease. Molecular Neurobiology 2016; 53: 2132–40.
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. Journal of Neurology, Neurosurgery, and Psychiatry 2008; 79: 874–880.

Paulsen JS, Langbehn DR, Stout JC, Aylward E, Ross CA, Nance M, et al. Preparing for preventive clinical trials: The Predict-HD study. Archives of Neurology 2006; 63: 883–90.

Phan J, Hickey MA, Zhang P, Chesselet M-F, Reue K. Adipose tissue dysfunction tracks disease progression in two Huntington’s disease mouse models. Human Molecular Genetics 2009; 18: 1006–16.

Pringsheim T, Wiltshire K, Day L, Dykeman J, Steeves T, Jette N. The incidence and prevalence of Huntington’s disease: a systematic review and meta-analysis. Movement Disorders 2012; 27: 1083–91.

Rees EM, Seahill RI, Hobbs NZ. Longitudinal neuroimaging biomarkers in Huntington’s Disease. [Review]. Journal of Huntington’s Disease 2013; 2: 21–39

Rodrigues FB, Byrne LM, Wild EJ. Biofluid biomarkers in Huntington’s disease. In: Precious SV, Rosser AE, Dunnett SB, editors. Huntington’s disease methods in molecular biology. New York: Springer; 2018. p. 329–96.

Ross CA and Tabrizi SJ. Huntington’s disease: from molecular pathogenesis to clinical treatment. [Review]. Lancet Neurology 2011; 10: 83–98.

Rudinskiy N, Kaneko YA, Beesen AA, Gokce O, Régulier E, Déglon N, et al. Diminished hippocalcin expression in Huntington’s disease brain does not account for increased striatal neuron vulnerability as assessed in primary neurons. Journal of Neurochemistry 2009; 111: 460–72.

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. Proceedings of the National Academy of Sciences United States of America 2007; 104: 14424–29.

Seahill RI, Wild EJ, Tabrizi SJ. Biomarkers for Huntington’s disease: an update. Expert Opinion on Medical Diagnostics 2012; 6: 371–75.

Seredenina T, Gokce O, Luthi-Carter R. Decreased striatal RGS2 expression is neuroprotective in Huntington’s disease (HD) and exemplifies a compensatory aspect of HD-induced gene regulation. PLOS One 2011; 6: e22231.

Seredenina T and Luthi-Carter R. What have we learned from gene expression profiles in Huntington’s disease? [Review]. Neurobiology of Disease 2012; 45: 83–98.

Stout JC, Paulsen JS, Queller S, Solomon AC, Whitlock KB, Campbell JC, et al. Neurocognitive signs in prodromal Huntington disease. Neuropsychology 2011; 25: 1–14.

Strand AD, Aragaki AK, Shaw D, Bird T, Holton J, Turner C, Tapscott SJ, et al. Gene expression in Huntington’s disease skeletal muscle: a potential biomarker. Human Molecular Genetics 2005; 14: 1863–76.
Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles Proceedings of the National Academy of Sciences of the United States of America 2005; 102: 15545–50.

Sugars KL and Rubinsztein DC. Transcriptional abnormalities in Huntington disease. [Review]. Trends in Genetics 2003; 19: 233–38.

Tabrizi SJ, Langbehn DR, Leavitt BR, Roos RAC, 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 Neurology 2009; 8: 791–801.

Tabrizi SJ, Reilmann R, Roos RAC, Durr A, Leavitt B, Owen G, et al. Potential endpoints for clinical trials in premanifest and early Huntington’s disease in the TRACK-HD study: analysis of 24 month observational data. Lancet Neurology 2012; 11: 42–53.

Tippett LJ, Waldvogel HJ, Snell RG, Vonsattel J-P, Young AB, Faull RLM. The complexity of clinical Huntington’s disease: Developments in molecular genetics neuropathology and neuroimaging biomarkers. [Review]. Advances in Neurobiology 2017; 15: 129–61.

Träger U, Andre R, Lahiri N, Magnusson-Lind A, Weiss A, Grueninger S, et al. HTT-lowering reverses Huntington’s disease immune dysfunction caused by NFκB pathway dysregulation. Brain 2014; 137: 819–33.

tvan der Burg JM, Björkqvist M, Brundin P. Beyond the brain: widespread pathology in Huntington’s disease. [Review]. Lancet Neurology 2009; 8: 765–74.

Walker FO. Huntington’s disease. [Review]. Seminars in Neurology 2007; 27: 143–150.

Weir DW, Sturrock A, Leavitt BR. Development of biomarkers for Huntington’s disease. [Review]. The Lancet Neurology 2011; 10: 573–90.

Wexler NS, Lorimer J, Porter J, Gomez F, Moskowitz C, Shackell E, et al. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington’s disease age of onset. Proceedings of the National Academy of Sciences of the United States of America 2004; 101: 3498–3503.

Weydt P, Pineda VV, Torrence AE, Libby RT, Satterfield TF, Lazarowski ER, et al. Thermoregulatory and metabolic defects in Huntington’s disease transgenic mice implicate PGC-1α in Huntington’s disease neurodegeneration. Cell Metabolism 2006; 4: 349–62.

Wild EJ, Boggio R, Langbehn D, Robertson N, Haider S, Miller JRC, et al. Abnormal peripheral chemokine profile in Huntington’s disease. PLoS Currents 2011; 3: RRN1231.

Wild EJ, Magnusson A, Lahiri N, Krus U, Orth M, Tabrizi SJ, et al. Quantification of mutant huntingtin protein in cerebrospinal fluid from Huntington’s disease patients. The Journal of Clinical Investigation 2015; 125: 1979–86.

Wright GEB, Collins JA, Kay C, McDonald C, Dolzhenko E, Xia Q, et al. Length of uninterrupted CAG, independent of polyglutamine size, results in increased somatic instability, hastening onset of Huntington disease. American Journal of Human Genetics 2019; 104: 1116-1126.
Zadel M, Maver A, Kovanda A, Peterlin B. Transcriptomic biomarkers for Huntington’s disease: Are gene expression signatures in whole blood reliable biomarkers? OMICS: A Journal of Integrative Biology 2018; 22: 283–94.

Zielonka D, Piotrowska I, Marcinkowski JT, Mielcarek M. Skeletal muscle pathology in Huntington’s disease. [Review]. Frontiers in Physiology 2014; 5: 380.
Figure legends

Figure 1. Gene expression of four probesets that showed the strongest evidence of differential expression in Huntington’s disease subjects compared to controls. Asterisks indicate nominal p < 0.05. Adjusted p values > 0.05 for all genes. Error bars represent SEM. Expression signals are unlogged for display purposes.

Figure 2. The four probesets with the strongest correlation between disease progression and differential expression over two years. These analyses are based on data from preHD B (n = 15; mean age at year 3 (y) ± SD: 39.9 ± 8.9 y; mean TMS change ± SD: 4.67 ± 4.47) and diagnosed Huntington’s disease subjects zHD stage 1 (n = 16; 50.0 ± 9.3 y; 6.56 ± 6.74) and zHD stage 2 (n = 16; 54.5 ± 6.6; 7.13 ± 6.16). Correlation calculated using Pearson’s correlation coefficient (r). When the significance of correlation was calculated and multiple testing correction performed, adjusted p > 0.05 in all probesets. (A) AQR: r = -0.61; nominal p = 4.3 × 10^{-6}. (B) LMNA: r = +0.58; nominal p = 1.6 × 10^{-5}. (C) ELAVL3: r = +0.58; nominal p = 2.3 × 10^{-5}. (D) TMEM5: r = -0.56; nominal p = 4.9 × 10^{-5}.

Figure 3. Probesets representing genes found to be dysregulated in previously transcriptomic Huntington’s disease blood studies that showed significant association between disease progression and differential expression in our longitudinal analysis. These analyses are based on data from preHD B (n = 15; mean age at year 3 (y) ± SD: 39.9 ± 8.9 y; mean TMS change ± SD: 4.67 ± 4.47) and diagnosed Huntington’s disease subjects zHD stage 1 (n = 16; 50.0 ± 9.3 y; 6.56 ± 6.74) and zHD stage 2 (n = 16; 54.5 ± 6.6; 7.13 ± 6.16). Correlation calculated using Pearson’s correlation coefficient (r). When the significance of correlation was calculated and multiple testing correction performed, adjusted p > 0.05 in all probesets. (A) ROCK1: r = +0.36; nominal p = 0.013. (B) SAP30: r = +0.35; nominal p = 0.017. (C) IL23A: r = -0.33; nominal p = 0.022. (D) LTBR: r = +0.29; nominal p = 0.049.
Table 1. Top 5 most significant probesets from each year showing significant differential expression at the other time point. $P$ values have not been adjusted for multiple testing. EST stands for Expressed Sequence Tag. Protein function is inferred from GeneCards Human Gene Database.

| Gene symbol | Probeset     | Protein function                          | Year 1 | Year 3 |
|-------------|--------------|-------------------------------------------|--------|--------|
|             |              | logFC | $p$     | logFC | $p$     |
| HSA277841   | 1560855_at   | -0.34 | 1.5E-04 | 0.24  | 0.038   |
| MARCH8      | 231933_at    | -0.57 | 8.7E-04 | -0.36 | 0.046   |
| TRIM58      | 215047_at    | -0.59 | 8.8E-04 | -0.44 | 0.013   |
| TOMM34      | 201870_at    | 0.23  | 9.1E-04 | 0.25  | 5.5E-03 |
| LIMS1       | 207198_s_at  | -0.25 | 1.3E-03 | -0.18 | 0.027   |
| SESN3       | 235683_at    | -0.41 | 0.042   | -0.68 | 2.7E-04 |
| ELL2        | 214446_at    | -0.44 | 8.6E-03 | -0.55 | 3.0E-04 |
| CYP19A1     | 239459_s_at  | -0.27 | 0.038   | -0.58 | 7.3E-04 |
| CCDC15      | 220466_at    | -0.33 | 0.046   | -0.49 | 7.3E-04 |
| N4BP2L2     | 214753_at    | 0.19  | 3.7E-03 | 0.20  | 7.4E-04 |
Table 2. Top 20 probesets with strongest correlation longitudinal analysis of gene expression and TMS progression. These analyses are based on data from 15 preHD B and 32 diagnosed Huntington’s disease subjects. Correlations were calculated using Pearson's correlation coefficient. $P$ values have not been adjusted. EST stands for Expressed Sequence Tag and signifies microarray tags for which the gene they represent is currently unknown. Protein function from GeneCards Human Gene Database.

| Gene symbol | Probeset | Protein function                                     | Correlation | $p$ Value |
|-------------|----------|------------------------------------------------------|-------------|-----------|
| AQR         | 212584_at| Intron-binding spliceosomal factor                   | -0.61       | 4.3E-06   |
| LMNA        | 214213_x_at| Component of nuclear lamina                           | 0.58        | 1.6E-05   |
| ELAVL3      | 206338_at| Cell differentiation                                  | 0.58        | 2.2E-05   |
| TMEM5       | 204808_s_at| Transmembrane protein                                 | -0.56       | 5.0E-05   |
| HSPA9       | 200690_at| Stress response, mitochondrial maintenance            | -0.55       | 5.4E-05   |
| (EST)       | 242526_at| (Unknown)                                             | -0.55       | 6.6E-05   |
| NACC1       | 235047_x_at| Apoptosis, transcriptional regulation                 | 0.55        | 7.2E-05   |
| PRX         | 220024_s_at| Axon ensheathment                                     | 0.54        | 8.2E-05   |
| LINC00691   | 1561059_a_at| (Non-protein coding)                                 | 0.54        | 1.1E-04   |
| IGFBP1      | 237989_at| IGF binding protein                                   | 0.53        | 1.1E-04   |
| (EST)       | 239312_at| (Unknown)                                             | 0.53        | 1.2E-04   |
| (EST)       | 220129_at| (Unknown)                                             | 0.53        | 1.4E-04   |
| MGC16025    | 1553747_at| (Non-protein coding)                                 | 0.53        | 1.5E-04   |
| LYPD6       | 227763_at| Interacts with nAChR subunits                          | 0.52        | 1.6E-04   |
| MAPRE3      | 203841_x_at| Regulation of microtubule cytoskeleton                | 0.52        | 1.8E-04   |
| TSGA10      | 223838_at| (Unknown)                                             | -0.52       | 1.9E-04   |
| (EST)       | 237213_at| (Unknown)                                             | -0.52       | 2.0E-04   |
| CCDC150     | 1553459_at| (Unknown)                                             | 0.52        | 2.1E-04   |
| (EST)       | 1557505_a_at| (Unknown)                                           | 0.51        | 2.2E-04   |
| CNTN1       | 1554784_at| Neuronal cell adhesion                                | 0.51        | 2.5E-04   |

https://mc.manuscriptcentral.com/braincom
Table 3. Functions of four probesets representing genes identified in previous studies as dysregulated in Huntington’s disease blood and their association with disease progression calculated in our longitudinal analysis of gene expression and TMS. Samples included in the analysis were those of 15 preHD B individuals and 32 diagnosed Huntington’s disease subjects. Correlation calculated using Pearson's correlation coefficient. P values have not been adjusted. Protein function from GeneCards Human Gene Database.

| Gene symbol | Probeset | Protein function                                      | Correlation | p Value |
|-------------|----------|-------------------------------------------------------|-------------|---------|
| ROCK1       | 235854_x_at | Regulation of actomyosin cytoskeleton                 | 0.36        | 0.013   |
| SAP30       | 204899_s_at | Histone deacetylase complex subunit                   | 0.35        | 0.017   |
| IL23A       | 220054_at  | Innate and adaptive immunity                         | -0.33       | 0.022   |
| LTBR        | 243400_x_at | Lymphoid development, lipid metabolism, apoptosis     | 0.29        | 0.049   |
Neuropathological changes

Blood RNA levels are poor correlates of clinical progression in Huntington’s disease.

No longitudinal changes in RNA levels were detectable in TRACK-HD subjects over a two-year interval.
Huntington’s disease is a severe hereditary neurodegenerative disease. Clinical measures of disease are somewhat subjective, and therefore state biomarkers could be of great benefit. Mitchell et al. profiled blood RNAs in well-characterized subjects for prospective changes over a two-year interval; these proved insufficiently sensitive to be useful in quantifying disease progression.