Transcriptional changes in Huntington Disease identified using genome-wide expression profiling and cross platform analysis


1) Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, Department of Medical Genetics, University of British Columbia, Vancouver, BC, V5Z 4H4, Canada.

2) Centre for High-throughput Biology and Department of Psychiatry, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

3) Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Station 15, CH1015 Lausanne, Switzerland

Address correspondence to:
Dr. Blair R. Leavitt
Centre for Molecular Medicine and Therapeutics
Department of Medical Genetics
University of British Columbia
980 West 28th Avenue
Vancouver, BC, V5Z 4H4
Ph: 604-875-3801
Fax: 604-875-3840
Email: bleavitt@cmmt.ubc.ca
Abstract

Evaluation of transcriptional changes in the striatum may be an effective approach to understanding the natural history of changes in expression contributing to the pathogenesis of Huntington disease (HD). We have performed genome-wide expression profiling of the YAC128 transgenic mouse model of HD at 12 and 24 months of age using two platforms in parallel: Affymetrix and Illumina. The data from these two powerful platforms was integrated to create a combined rank list, thereby revealing the identity of additional genes that proved to be differentially expressed between YAC128 and control mice. Using this approach we identified thirteen genes to be differentially expressed between the YAC128 and controls which were validated by qPCR in independent cohorts of animals. In addition, we analyzed additional time points relevant to disease pathology: three, six and nine months of age. Here we present data showing the evolution of changes in the expression of selected genes; *Wt1*, *Pcdh20* and *Actn2* RNA levels change as early as three months of age, whereas *Gsg1l*, *Sfmbt2*, *Acy3*, *Polr2a* and *Ppp1r9a*, RNA expression is affected later, at 12 and 24 months of age. We also analyzed the expression of these thirteen genes in human HD and control brain, thereby revealing changes in *SLC45A3*, *PCDH20*, *ACTN2*, *DDAH1* and *PPP1R9A* RNA expression. Further study of these genes may unravel novel pathways contributing to HD pathogenesis.
Introduction

Huntington Disease (HD) is a neurological disorder characterized by early selective neuronal cell death in striatum and cortex caused by mutations in the human HTT gene (1). Expansion of the CAG tract in exon 1 of the HTT gene leads to an abnormally long polyglutamine tract in huntingtin, a protein ubiquitously expressed throughout the body and brain. Neuropathological features are observed in the neostriatum, globus pallidus and cerebral cortex with a selective loss of the large pyramidal neurons in cortical layers III, IV and VI (2, 3).

Evidence derived from cellular and animal models of HD has identified a number of potential mechanisms by which polyglutamine-expanded huntingtin protein results in neuronal cell death. These include disruption of the axonal transport, mitochondrial dysfunction, imbalance in calcium homeostasis and excitotoxicity, altered proteolysis, impairment of the ubiquitin proteosome system, and transcriptional alterations (4, 5). Specific changes in mRNA expression have been described in human post-mortem brains, and in cellular- and mouse models of HD. Down-regulation of dopamine D1 and D2 receptors, enkephalin, substance P (6-8) and several striatal-enriched genes have been found altered in post-mortem caudate samples from HD cases using quantitative real time PCR analysis (qPCR)(9). Microarray studies performed on human HD brain samples (grade 0-2) have demonstrated that the caudate nucleus manifests the greatest number and magnitude of transcriptional changes, followed by motor cortex and then cerebellum, with no significant change detectable in prefrontal association cortex (10).

There is concordance of striatal-enriched genes differentially expressed in HD mouse models with gene expression changes identified in human microarray studies (11).
For example, several of the genes showing the greatest magnitude of changes, such as the cannabinoid CB1 receptor, dopamine D1 receptor and enkephalin genes, have also been identified as differentially expressed in expression profiling studies of transgenic mice (10-12). Also, a considerable number of Ca\textsuperscript{2+}-binding and Ca\textsuperscript{2+}, K\textsuperscript+ and Na\textsuperscript+ channel mRNAs are down-regulated in humans in concordance with previous observations in the R6/2 transgenic mouse model (13-15).

The YAC128 mouse model of HD expresses full-length human mutant huntingtin with 128 glutamines from a yeast artificial chromosome (YAC) (16, 17). This mouse model expresses full-length huntingtin under the endogenous human huntingtin promoter and spans both up- and down-stream regulatory regions. YAC128 mice display disease progression over time with both neuropathological and behavioral deficits similar to human disease. Initially, the YAC128 mice exhibit behavioral changes with a hyperkinetic phenotype and rotarod deficits at three months of age (18). YAC128 mice display several neuropathological changes characteristic of HD with a significant decrease in brain weight and striatal volume observed at eight months (19), and a selective neuronal loss in striatum and cortex at 12 months of age (18). The fact that YAC128 transgenic mice display early onset cognitive and motor deficits with selective neuropathology recapitulating key features of human disease makes this mouse model useful for studies into disease pathogenesis.

The majority of genome-wide expression studies performed including studies on different HD mouse models have employed the Affymetrix Gene chip platform when studying gene expression changes. More recent technologies, such as the Illumina Bead Array platform, are now also widely used for this kind of analysis. Affymetrix arrays are
produced by \textit{in situ} synthesis of 25-mer oligonucleotides with multiple probes for each gene. The Affymetrix arrays contain one-base mismatch probes as controls for cross-hybridization. In contrast, the Illumina Bead arrays are based on self-assembly of 50-mer long oligonucleotides to microbeads. Bead arrays contain 30 copies of the same oligonucleotide as an internal technical replication, which Affymetrix arrays lack. We employed both platforms to increase the power of the gene expression profile analysis of the YAC128 and controls with the ultimate goal of studying transcriptional alterations in striatum and identifying genes of importance for disease pathogenesis.

In this study we present evidence for both early and late transcriptional changes of genes in the striatum that we initially identified by genome-wide expression profiling of the YAC128 mouse model of HD using the Affymetrix and Illumina platforms. As our experiments were designed to analyze both platforms in parallel as a validation tool, we first analyzed the same striatal mRNA separately across both platforms. We identified modest transcriptional changes on the individual platforms. This prompted us to an alternative approach, to combine the results from the Affymetrix and Illumina platforms to create combined rank lists. We show that combining data sets obtained from two powerful independent platforms revealed the identity of additional genes that were differentially expressed between YAC128 and control mice at 12 and 24 months of age. To further validate our results we analyzed the top ten genes ranked from the Illumina, Affymetrix and the combined rank lists, respectively, by applying the more sensitive qPCR methodology. For genes validated to be differentially expressed we also analyzed earlier time-points: three, six and nine months old YAC128 mice and controls.
In summary, we have applied an unbiased approach to combine results from two powerful expression array platforms that identified a set of genes that showed a differential regulation between the YAC128 and controls. In total, we evaluated expression of thirteen genes at time points relevant to disease pathology. This analysis revealed several genes: Wt1, Pcdh20 and Actn2 that display transcriptional changes already at three months of age, early in the manifestation of the HD phenotype of these mice. The expression of additional genes: Gsg1l, Sfmbt2, Acy3, Polr2a and Ppp1r9a, was found to be altered only at later time points. We examined the expression of these genes in human HD caudate and controls and demonstrated transcriptional changes in SLC45A3, PCDH20, ACTN2, DDAH1 and PPP1R9A, validating the relevance of these findings in mouse to the human disease. Further study of these genes might elucidate novel disease mechanisms of importance in HD.
Results

**Genome-wide expression profiling of 24 month old YAC128 mice and controls using Affymetrix and Illumina platforms**

We initially compared the transcriptional changes in the striatum of 24 month old YAC128 and wildtype littermate control mice. YAC128 mice display a clear progression of disease with age that correlates with the degree of selective neuronal loss and neurodegeneration. There is an established correlation between levels of mutant huntingtin, disease severity and neuronal loss in YAC128 mice (20). We studied the gene expression profiles of the same mRNA using two different platforms: Illumina and Affymetrix (supplementary data, Fig. S1). YAC128 mice showed both an up- (light yellow) and down-regulation (dark brown) of genes compared to controls. However, no genes were significantly different on either platform after correction for multiple comparisons (21).

**Calculating combined ranks from the Affymetrix and Illumina profiles for 24 month old YAC128 and controls**

Gene expression studies using Affymetrix GeneChips and Illumina BeadArrays have previously been compared side by side with a high agreement for genes that were predicted to be differentially expressed (22). We therefore examined our genome-wide expression data to determine whether genes had similar ranks on both platforms, taking into account that approximately 48,000 probe sets are analyzed on the Illumina arrays and 45,000 probe sets are measured on the Affymetrix counterpart. In order to focus on the very best candidate changes, we only considered the data for the top 100 genes on
each platform (ranked by probability of differential expression). We found both concordant as well as discordant predicted changes in expression between the two platforms. Discordant predictions may be due in part to low expression levels of some genes or the fact that the specific probes employed in the two platforms assay different splice variants or transcripts of a gene (23). While our experiments were designed to analyze both platforms in parallel as a validation tool, the modest transcriptional changes observed with both platforms prompted us to attempt an alternative approach to identify differentially expressed genes in YAC128 mice compared to controls. We subsequently used the two data sets to create a combined rank list. We calculated a combined rank for each gene by averaging the individual ranks generated from the Affymetrix and Illumina platforms (supplementary data Table S1). This approach incorporates a confirmatory identification of differentially expressed genes. Furthermore, this approach helped to elute genes of interest despite small sample sizes and low magnitude of transcriptional changes in our experimental setting. However, by implementing this approach we down-weighted false-positive changes *i.e.* genes that were ranked high on one platform, but not reproduced on the other platform, would tend to be down-weighted.

**Transcriptional changes identified in the top ten genes from Affymetrix and Illumina genome-wide expression profiling**

We next decided to validate the separate findings from the Illumina and Affymetrix results using qPCR. We assessed the top ten differentially expressed genes based on the Illumina and Affymetrix results, respectively, on mRNAs from a new cohort of 24 month old YAC128 and control mice (supplementary data Fig. S1). This analysis resulted in confirmation of three genes out of ten from the Illumina results to be differentially
expressed *i.e.*, Wilm’s tumor 1 homolog (*Wt1*), DAZ interacting protein 1-like (*Dzip1l*) and Retrotransposon gag domain containing 4 (*Rgag4*), where *Wt1* was up-regulated, and both *Dzip1l* and *Rgag4* were down-regulated in the YAC128 compared to controls (supplementary data Fig. S2) (Table 1). Netrin G1 (*Ntng1*) showed a trend for up-regulation in the YAC128 mice compared to controls that was not statistically significant. The RAB21 member of the RAS oncogene family (*Rab21*), Docking protein 3 (*Dok3*), Dopa decarboxylase (*Ddc*) and Up-regulated during skeletal muscle growth 5 (*Usmg5*) were not detected as differentially expressed between YAC128 and controls in the qPCR analyses (supplementary data Fig. S2). Three different primer sets were tested for olfactory receptor 1502 (*Olfr1502*) and fibronectin type III domain containing 8 (*Fndc8*), but we were unable to amplify and quantify these genes from striatal RNA.

Similar results were obtained with four genes out of ten detected as differentially expressed based on the Affymetrix results; confirmed differentially expressed RNAs comprised *Wt1*, DNA damage inducible transcript 4-like (*Ddit4l*), Germ cell-specific gene 1-like protein (*Gsg1l*) and solute carrier family 45, member 3 (*Slc45a3*) (supplementary data Fig. S3) (Table 1). In addition to *Wt1*, which was top ranked in both the Illumina and Affymetrix rank results, *Slc45a3* was also up-regulated, while *Ddit4l* and *Gsg1l* were down-regulated in YAC128 compared to control mice. *DMPK* also showed a non-statistically significant trend towards down-regulation in YAC128 compared to controls. Deleted in lymphocytic leukemia, 7 (*Dleu7*), Spermatogenesis associated 5 (*Spata5*), Ectonucleoside triphosphate diphosphohydrolase 7 (*Entpd7*), GRP1-associated scaffold protein (*Grasp*) and Regulator of telomere elongation helicase
Identification of additional genes from the combined rank list validated as differentially expressed in YAC128

We next assessed the differential expression of the genes from the created combined rank list for 24 month old YAC128 and control striatal mRNA samples. The qPCR analysis of the top ten genes from the combined rank list showed differential expression in YAC128 compared to controls for six out of the ten genes analyzed (Table 1). Wt1, Ntng1 and Ddit4l were included in the top ten for both Affymetrix and/or Illumina and in the created combined rank list. Additional genes from the combined rank list that were shown to be differentially expressed were: Protocadherin 20 (Pcdh20), Scm-like with four MBT domains protein 2 (Sfmbt2), Aspartoacylase 3 (Acy3) and Polymerase (RNA) II (DNA directed) polypeptide A (Polr2a) (supplementary data, Fig. S4). All of these genes were significantly up-regulated in YAC128 compared to controls. There was no significant difference in gene expression for Transducer of ERBB2 (Tob2), Pogo transposable element with KRAB domain (Pogk) or HIV-1 Rev binding protein-like (Hrbl) between YAC128 and controls at 24 months of age (Table 1) (supplementary data Fig. S4).

Transcriptional alterations revealed using the combined rank strategy on 12 month old YAC128 and controls

We next studied the YAC128 mice compared to controls at the earlier 12 month time point. Both neuropathological and behavioral abnormalities are established in the transgenic mice at this time point. YAC128 mice display approximately 10-15% striatal
atrophy and 9% neuronal loss at 12 months of age (24). We used the same approach as for the 24 month old mice to calculate a combined rank by averaging the ranks obtained from the individual platforms (supplementary data Table S2). We again performed qPCR analyses of the top ten genes from the established combined rank list using a new cohort of 12 month old YAC128 and control mice. Three out of ten genes were validated to be differentially expressed; α-Actinin 2 (Actn2), Dimethylarginine dimethylaminohydrolase 1 (Ddah1) and Protein phosphatase 1, regulatory (inhibitor) subunit 9A (Ppp1r9a) (Fig. S5) (Table 1). Interestingly, all three genes were down-regulated in the YAC128 at 12 months of age. Myeloid/lymphoid leukemia 10 (Mllt10), Surfeit 4 gene (Surf4), Forkhead box J3 (Foxj3), Transportin 1 (Tnpo1), Nedd4-binding protein 1 (N4bp1) and WAP for-disulfide core domain protein 1 precursor (Wfdc1) were not differentially expressed between YAC128 and controls (supplementary data, Fig. S5). Three different primer sets were tested for Zinc finger protein 371 (Zfp371), but we were unable to amplify and quantify the expression of this gene from mouse striatal RNA.

**Time course analysis of differentially expressed genes show both early and late changes in gene expression**

To further characterize the change in expression of the genes that were identified as differentially expressed at 12 or 24 months, we studied additional time points of relevance to disease pathogenesis with regard to both early onset of behavioral abnormalities and neuropathological changes. Taken together, we analyzed the gene expression at 3, 6, 9, 12 and 24 months of age on mRNA from new cohorts of YAC128 and controls using qPCR for each target. Wt1 displayed dynamic transcriptional changes over all five time points analyzed. Initially, YAC128 mice showed an up-regulation of
*Wt1* at three months of age (p=0.029, Mann-Whitney 2-tailed U test) compared to controls (Fig. 1A). This, however, reverted to similar transcript levels expressed in YAC128 and control mice at six months. YAC128 displayed a down-regulation of *Wt1* at nine months of age (p=0.017), before a distinct two and five-fold up-regulation was observed in YAC128 compared to controls at 12 (p=0.0025) and 24 months (p=0.0012), respectively. YAC128 mice showed a continuous down-regulation of *Ddit4l* (Fig. 1B) from six to 24 months of age, except for at 12 months when YAC128 and controls displayed similar transcript levels (p=0.0079 in six month; p=0.030 in nine month; p=0.014 in 24 month YAC128 vs. controls). *Gsg1l* was only down-regulated in the YAC128 at 24 months of age (p=0.018) (Fig. 1C). In contrast, *Slc45a3* was up-regulated almost three-fold in the YAC128 mice compared to controls at 9 months of age and continuously up-regulated at later time points (p=0.029 in nine month; p=0.0087 in 12 month; p=0.015 in 24 month YAC128 vs. controls) (Fig. 1D). Transcriptional expression profiles with down-regulation in the YAC128 compared to controls were observed for *Dzip1l* (p=0.0055 in six month; p=0.0012 in 12 month; p=0.022 in 24 month YAC128 vs. controls) (Fig. 1E) and *Rgag4* (p=0.0010 in six month; p=0.0082 in 24 month YAC128 vs. controls) (Fig. 1F). For the genes obtained from the combined ranking lists, YAC128 mice showed a significant up-regulation of *Pcdh20* at all time points, except for at six months (p=0.018 at three month; p=0.017 at nine month; p=0.0012 at 12 month; p=0.035 at 24 month YAC128 vs. controls) (Fig. 2A). *Sfmbt2, Acy3 and Polr2a*, showed up-regulation at the two late-stage time points in the YAC128 compared to controls; *Sfmbt2* (p=0.0012 for both 12 and 24 month YAC128 vs. controls) (Fig. 2B); *Acy3* (p=0.0025 for 12 month; p=0.0012 for 24 month YAC128 vs. controls) (Fig. 2C); *Polr2a* (p=0.035 for
12 month; \( p=0.0047 \) for 24 month YAC128 vs. controls) (Fig. 2D). In addition, there were trends toward differential expression for several RNA targets at individual time points even though statistical significance was not reached. \( Acy3 \) (\( p=0.055 \)) (Fig. 2C) and \( Polr2a \) (\( p=0.051 \)) (Fig. 2D) showed borderline significance for up-regulation in nine month old YAC128 compared to controls. Several of the analyzed target genes displayed a significant difference in transcriptional expression at the earlier time-points. \( Actn2 \) was down-regulated in YAC128 as early as 3 months and consistently until 24 months of age (\( p=0.0040 \) in three month; \( p=0.036 \) in six month; \( p=0.0087 \) in nine month; \( p=0.0012 \) in 12 month; \( p=0.014 \) in 24 month old YAC128 vs. controls) (Fig. 3A). This was the only gene that showed a significant differential expression at every time point analyzed. \( Ddah1 \) showed a five-fold down-regulation in the YAC128 compared to controls at 24 months in addition to changes observed at earlier time points (\( p=0.025 \) for six month; \( p=0.022 \) for 12 month; \( p=0.0020 \) for 24 month old YAC128 vs. controls) (Fig. 3B). \( Ppp1r9a \) showed a down-regulation in YAC128 compared to controls at the two late-stage time points (\( p=0.0047 \) for 12 month; \( p=0.0012 \) for 24 month YAC128 vs. controls) (Fig. 3C). In general, the Affymetrix platform gave a better estimate than Illumina, of the magnitude of fold change based on our qPCR results (Table 2).

**Transcriptional alterations in HD cases are concordant with changes observed in YAC128**

We next studied the genes that were differentially expressed in mouse striatum in caudate samples from human HD and control brains. In total, we analyzed fourteen genes that were identified to be differentially expressed on the individual platforms, Affymetrix and Illumina, and the combined rank lists for 12 month and 24 month data, respectively. \( WTI \)
showed a tendency for up-regulation in HD caudate, but did not reach statistical significance (p=0.082, Mann-Whitney 2-tailed U test) (Fig. 4A). This tendency to up-regulation was concordant to observations in YAC128 compared to controls at 3, 12 and 24 months of age (Table 2). We decided to analyze DMPK expression in human HD cases, since DMPK showed borderline significance for down-regulation in 24 month old YAC128 mice compared to controls. However, in our analysis DMPK showed no transcriptional changes in human HD cases (Fig. 4C). GSG1L showed down-regulation in HD caudate with borderline significance (p=0.051) (Fig. 4D). The gene expression profile was concordant with the YAC128 profile at 24 months of age. SLC45A3 was highly up-regulated in HD caudate compared to controls (p=0.0043) (Fig. 4E), concordant with the expression profile in YAC128 mice compared to controls at 9, 12 and 24 months of age. DDIT4L (Fig. 4B), DZIP1L (Fig. 4F) and RGAG4 (Fig. 4G) did not show any significant changes in HD caudate versus controls. We also studied the additional genes that were identified as differentially expressed from the combined rank lists for the 12 and 24 month YAC128 data sets. PCDH20 was down-regulated in HD caudate compared to controls (p=0.0043) (Fig. 5A). This expression profile was discordant to the changes observed in the YAC128 and controls. SFMTB2, ACY3 and POLR2A (Fig. 5B-D) did not show any significant transcriptional changes between HD cases and controls. All three genes analyzed from the combined rank list created from the 12 month data and validated to be differentially expressed in YAC128 and controls, also showed transcriptional changes in human HD caudate samples: ACTN2 (p=0.017) (Fig. 5E), PPP1R9A (p=0.017) (Fig. 5F) and DDAHI (p=0.017) (Fig. 5G). ACTN2 and PPP1R9A showed down-regulation, while DDAHI showed up-regulation in HD case
caudate compared controls. Gene expression profiles for ACTN2 and PPP1R9A were both concordant with observations in the YAC128 data, in contrast to DDAH1 which was up-regulated in HD caudate and down-regulated in YAC128 at 6, 12 and 24 months compared to controls.
Discussion

We performed genome-wide expression profiling on striatal tissue from YAC128 and controls at 12 and 24 months of age. We utilized two powerful platforms in parallel, Illumina and Affymetrix arrays, with the ultimate goal to study transcriptional changes in striatum and to identify genes of importance for disease pathogenesis. We identified thirteen genes that were differentially expressed in YAC128 compared to controls, as confirmed in independent groups of YAC128 samples. We also used qPCR to further assess time-dependent alterations in the expression of these genes with respect to the onset and progression of disease phenotype. Most importantly, we studied these genes in HD cases and we showed transcriptional alterations consistent with the findings in YAC128. Transcriptional changes identified may derive from a variety of different mechanisms involved in disease, and may therefore lead to the identification of crucial pathways involved in HD pathogenesis.

It was first shown that long polyglutamine stretches activate transcription in vitro (25, 26) and many transcription factors contain glutamine-rich activation domains such as CREB-binding protein (CBP), TATA-box binding protein (TBP) and specificity protein 1 (Sp1). Accordingly, mutant huntingtin protein has been suggested to act directly as a transcription factor by mimicking or interfering with the actions of transcription factors that contain glutamine rich activation domains (27). More recent studies showed that huntingtin protein, both wild-type and mutant, has the ability to bind DNA without additional transcription factors (28). Mutant huntingtin, however, increased the transcriptional binding factor activity overall. In addition, mutant huntingtin displayed increased occupancy at gene promoters in vivo compared to wild-type huntingtin. Mutant
huntingtin is thus suggested to alter the transcriptional profile through modulation of the DNA confirmation and altered binding of transcriptional factors (28).

Several studies indicate that mutant huntingtin interacts with, sequesters and/or compromises the normal function of transcription factors such as Sp1 (29) (30), the nuclear receptor co-repressor (N-CoR) (31), CBP (32, 33), p53 (34, 35), TBP (36) and TAFII130 (37) or disrupts the core transcriptional machinery by interacting with the pre-initiation complex (38). Furthermore, studies have provided evidence that huntingtin may modulate chromatin structure by interfering with the acetylated and deacetylated states of histones, which led to the testing of (HDAC) inhibitors for therapeutic intervention in HD. It has also been suggested that intranuclear inclusions may non-specifically alter gene expression by reducing the association of transcription factors to DNA binding sites (39), although this was shown not to be the case in the R6/2 mouse model of HD (40).

Previously, Kuhn et al. performed a meta-analysis on seven different mouse models of HD and human post-mortem caudate (12). In this study, it was concluded that short N-terminal mouse models exhibit rapid effects and transcriptional changes similar to what is observed in human brain. Nevertheless, knock-in (CHL2$^{Q150/Q150}$, Hdh$^{Q92/Q92}$) and the YAC128 transgenic full-length huntingtin mouse models also displayed significant HD-like transcript profiles at an older age. No distinct transcriptional changes could be assigned to the differences in expression of full-length huntingtin and N-terminal huntingtin fragments (12). Polyglutamine diseases share many features including a polyQ-dependent neurodegeneration and inclusion body pathology (41, 42). Transcriptional alterations observed could thus be caused by the polyglutamine expansion alone and be independent of the huntingtin protein context (15). Moreover, the
transcriptional changes observed may be due to secondary bystander effects from affected tissue as well as surrounding tissue, and/or changes not related to disease. Regardless, transcriptional dysregulation may be an important mechanism in HD pathogenesis. Several genome-wide expression studies have been performed and it has been difficult to extrapolate what changes are relevant to disease pathogenesis. We validated that Wt1, Ddit4l, Gsg1l, Slc45a3, Dzip1l and Rgag4, which were initially identified on either the Affymetrix or Illumina platform, were differentially expressed. Remarkably, by combining the results from the Affymetrix and Illumina arrays we were able to pinpoint additional genes that were differentially expressed. These genes were Pcdh20, Sfmbt2, Acy3, Polr2a, Ddah1, Actn2 and Pppr1r9a. This is partly due to the down weighting of false positive results (reduction of Type I error) when combining the two data sets, but more notably, this approach helped in eluting genes of interest despite small sample sizes and low magnitude of transcriptional changes in our experimental setting.

In our study we showed transcriptional changes for selected genes in our YAC128 mouse model of HD already at three months of age. This indicates that the changes observed in Wt1, Pcdh20 and Actn2 might play an important role in disease pathogenesis, since neuropathological changes are not manifested at this early time point. Interestingly, Wt1 displayed transcriptional changes with some distinct fluctuations observed throughout the time-points measured, while Actn2 was consistently down-regulated in YAC128 compared to controls. In contrast, the alterations in Pppr1r9a, which show late transcriptional changes from 12 months on, may reflect a secondary response to neuronal loss and additional bystander effects from both affected as well as adjacent brain regions.
We studied the gene expression of several genes in HD cases and controls and showed transcriptional changes in *SLC45A3, PCDH20, ACTN2, DDAH1* and *PPP1R9A*. The genes showing a differential transcriptional expression profile in YAC128 and controls, but not reaching statistical significance in human samples might still be involved in HD pathogenesis. In our study, discrepancies in transcriptional expression profiles between the YAC128 mouse and human HD case data might be due to biological inter-variability in the human samples. The HD grades analyzed ranged from 1-3 in the human samples, which might reduce the power to detect differences due to sample variability. In addition, differences in the postmortem interval might have an effect on the expression of some genes together with general biological variability in the samples analyzed (43, 44).

Several of the genes we have identified to be differentially expressed are suggested to be involved in RNA regulation *i.e.* *Wt1, Dzip1l, Sfmbt2* and *Polr2a*. The gene encoding fused in sarcoma (*FUS*) was recently linked to amyotrophic lateral sclerosis (ALS) and shown to be involved in RNA regulation (45, 46). This discovery is an example of neurodegenerative disease caused by a mutation in a gene that is predominantly involved in DNA repair, and regulation of RNA transcription, splicing and transport.

Wilm’s tumor 1 homolog (*Wt1*) plays an essential role in the normal development of the urogenital system, and the human gene is mutated in a small subset of patients with Wilm's tumors (47). *Wt1* belongs to a family of zinc-finger transcription factors for which *α*-actinin 1 (*Actn1*) has been identified as an mRNA target (48). The *WT1*(+KTS) isoform is suggested to bind close to or at the start codon of *Actn1* mRNA (48). More recently, it
was shown that *WT1* associated with transcripts encoding actin-binding proteins and other cytoskeletal proteins including alpha-actinins and *Ppp1r9a*. (49). We performed a sequence comparison of the region around the transcriptional start site for murine *Actn1* and *Actn2* (data not shown). There was approximately 70% homology between these regions, suggesting *Wt1* could potentially act as a transcription factor involved in *Actn2* regulation. Additional experiments are required to study putative interactions and to confirm whether *Wt1* directly regulates *Actn2* and *Ppp1r9a*. *Wt1* mRNA was in low-abundance compared to *Actn2* in the striatum of YAC128 and control mice. We observed a significant reduction of *Actn2* mRNA levels at all time-points measured compared to levels at three months in both YAC128 and controls (data not shown). Apart from this observation there is no obvious correlation in the expression profiles of *Wt1* and *Actn2* in the YAC128 or control mice. Transcript levels in human HD caudate showed a tendency for up-regulation in HD cases compared to controls, which correlates to increased levels observed in the YAC128 mice at 3, 12 and 24 months of age. This concordance strengthens the hypothesis that *WT1* may play an important role in HD pathogenesis.

*Actn2*, was found to be down-regulated in YAC128 mice compared to controls as early as 3 months of age. This gene displays a consistent down-regulation at all time points analyzed. More importantly, this down-regulation was confirmed in human HD caudate. Our findings in YAC128 mice and HD patients is in agreement with observations made in R6/2 mice both at 6 and 12 weeks of age, in R6/1 mice and in previous analysis of human HD samples (9, 10, 50). *Actn2* is involved in cell adhesion and cytoskeletal arrangement and has been suggested to play a crucial role for the
anchoring of NMDA receptors in central neurons (51-53). It has previously been shown that Actn2 co-immunoprecipitates with the protein RGS9-2 in rat striatum, suggesting a functional relationship between RGS9-2 and α-actinin 2 in calcium-mediated inactivation of NMDA receptors (54).

\textit{Slc45a3} belongs to a family of solute carriers with over 25 members. \textit{Slc14a1}, which belongs to this same family, has previously been shown to be differentially expressed both in human HD caudate (10) as well as in the R6/2 model (55). We observed up-regulation of \textit{Slc45a3} in YAC128 at 9, 12 and 24 months of age as well as in HD caudate compared to controls. The protein is predicted to be a plasma membrane protein and has been shown to constitute the majority of erythroblast transformation–specific (ETS) family member gene fusions observed in prostate cancers. \textit{Slc45a3} gene expression is not unique to prostate tissue and may play an important role in solute regulation in striatal neurons (56).

Protocadherin 20 (\textit{Pcdh20}) belongs to the cadherin superfamily, of which most members are expressed predominantly in the central nervous system. Protocadherins are involved in Ca\textsuperscript{2+}-mediated cell-cell adhesion and are suggested to be involved in the formation and maintenance of synaptic connections. \textit{In situ} hybridization studies performed on rat brain during early postnatal stage (P3), a critical period for establishment of specific synaptic connections, provide evidence for region-dependent expression patterns in cerebral cortex of several protocadherins, including \textit{Pcdh20} (57). Data suggests that there is a correlation between regional expression patterns of different protocadherins and development of specific synaptic connections between cerebral cortex and other communicating brains regions (57). YAC128 exhibited an up-regulation of
Pcdh20 as early as 3 months of age and also at 9, 12 and 24 months, while we observed a
down-regulation of PCDH20 in the caudate of HD cases compared to controls. This data
suggests that alterations in Pcdh20 expression may lead to impairment of synaptic
connections in striatum.

Dimethylarginine dimethylaminohydrolase 1 (Ddah1), is one of two described enzymes (Ddah1 and Ddah2) that hydrolyzes methylated arginine analogues and
asymmetric dimethylarginine (ADMA) produced during cellular turnover of methylated
proteins. ADMA is an endogenous inhibitor of nitric oxide synthase (NOS), which
catalyzes NO production (58, 59). Thus, Ddah1 is a target for novel therapeutic agents
designed to modulate NO generation. Neuronal nitric-oxide synthase positive
interneurons constitute ~10% of the cells in the striatum. These interneurons persist in
HD, in contrast to the GABAergic medium spiny neurons (MSNs) which are susceptible
to neurodegeneration in the striatum. In one study it was hypothesized that NO
nitrosylates the NMDA receptors in the interneurons and render them less sensitive to
activation, while NO diffusing into the MSNs lead to toxicity when it reacts with
superoxide anion (O2-) (60, 61). Administration of NOS inhibitors in R6/2 transgenic
mice accelerated the onset of disease symptoms (62). R6/1 transgenic mice with only one
copy of nNOS showed a delayed onset of disease, while mice lacking both copies
displayed an acceleration of the disease (62). We showed decreased levels of Ddah1 in
YAC128 compared to controls which might cause decreased hydrolysis of ADMA, and
reduced inhibition of nitric oxide synthase (NOS), resulting in increased levels of NO in
the YAC128. We observed increased levels of DDAH1 in HD cases which might suggest
an up-regulation of the enzyme to compensate for reduced levels of NO production due to
increased levels of ADMA in the brain. Alternatively, increased \textit{DDAH1} levels might be a primary or secondary contributor to the pathological state of HD. We need to further investigate the relationship between \textit{DDAH1}, ADMA, NOS and NO and the potential role these molecular targets might play in HD pathogenesis.

\textit{PPP1R9A} (protein phosphatase 1, regulatory (inhibitor) subunit 9A) encodes Neurabin I (\textit{neural tissue specific F-actin-binding protein I}) whose name stems from the fact that Neurabin I has been shown to bind and inhibit the function of protein phosphatase I (\textit{PPI}) (63) (64). In human, \textit{PPP1R9A} is located in a cluster of imprinted genes and the protein has been shown to be involved in actin cytoskeleton dynamics and in synaptic formation and function (64, 65). It has been shown that \textit{Ppp1r9a} is imprinted mainly in skeletal muscle (maternally expressed), but not in brain (65). Furthermore, it has also been shown that Neurabin I is highly concentrated in the synapses of developed neurons and is involved in neurite formation (66). We showed a down-regulation of \textit{Ppp1r9a} in striatum in YAC128 mice as well as HD cases compared to controls. Down-regulation of both \textit{Ppp1r9a} and \textit{Actn2} strengthens the evidence that deficiencies in cytoskeletal dynamics play a role in HD pathogenesis.

We have identified protocadherin 20 (\textit{Pcdh20}), neurabin I (\textit{Ppp1r9a}) and actinin alpha 2 (\textit{Actn2}) to be transcriptionally altered in the striatum of the YAC128 mice and in human HD caudate. These genes are known to play a role in synaptic formation, synaptic plasticity, and cytoskeletal arrangement. Both \textit{Actn2} and \textit{Pcdh20} were altered already at 3 months in the YAC128 mice. Our results suggest that transcriptional changes of these genes reflect neuronal dysfunction of pathways involved in: axonal transport, synaptic plasticity and dendrite integrity. The hallmark of HD is selective degeneration of
vulnerable medium spiny neurons while interneurons are spared in the striatum. Prior to cell death, however, there are neuropathological changes occurring that indicate early synaptic pathology. Morphological changes, including dystrophic neurites have been described in spiny striatal and cortical pyramidal neurons in HD (67) (68, 69). HD mice have been employed to study the correlation between dysmorphic changes in neurons and behavioural abnormalities. Dysmorphic dendrites in the striatal spiny and the cortical pyramidal neurons were shown to be predictive of onset and severity of behavioural deficits in HD mice (70). Dystrophic neurites have been suggested to correlate with alterations in synaptic plasticity in HD. For example, the R6/2 mice exhibit altered synaptic plasticity, which has been suggested to contribute to the pre-symptomatic changes in cognition (71). Furthermore, both YAC46 and YAC72 HD mice exhibit early electrophysiological abnormalities indicative of altered synaptic function, including excitatory NMDA receptor activity (16, 72). It has been suggested that cell-cell interactions between cortical and striatal neurons are critical for HD pathogenesis through the study of BAC transgenic mice (73). In addition, mutant huntingtin has been shown to inhibit both fast axonal transport and elongation of neuritic processes (74). We present novel transcriptional changes in several genes that are involved in synaptic integrity and function. The finding that these genes are transcriptionally dysregulated in the YAC128 mice and human HD caudate, emphasizes the importance of these pathways in HD pathogenesis. Moreover, modulation of synaptic influences may have therapeutic potential in HD.

In our study we have identified several genes that are differentially expressed in both the YAC128 model as well as in human HD caudate. Transcriptional changes
identified in this study might reflect: 1) genes that lie within the pathogenic pathways leading to neurodegeneration; 2) genes that reflect the cellular attempts to block the disease process; and 3) genes that reflect non-specific responses to neurodegeneration. The first group of genes will provide novel insights into the specific downstream mechanisms of neurodegeneration. Knowledge of these pathways will provide new experimental starting points from which to discover novel approaches directed specifically to modify these disease-causing pathways in HD. The second group of genes represent components of pathways that are up- or down-regulated in an attempt to reduce neuronal death. These might be exploited in the future as a way to protect neurons. The profiling of YAC128 mice has illuminated several molecular changes in HD brain that have received little or no attention in previous studies.

Transcriptional changes are likely to be important in HD pathogenesis. Full-length models have less obvious transcriptional changes than HD fragment models as described previously (75). Nevertheless, in this paper we present an experimental approach that has successfully identified novel transcriptional changes in YAC128 striatum and human HD caudate that have been validated by qPCR. Additional studies can be envisaged to better understand the contribution of these genes and their potential role in HD pathogenesis. Ultimately, further characterization of the role of these genes in HD may lead to the identification of new pathways involved in HD.
Materials and Methods

Mice

Transgenic HD mice expressing human HD huntingtin with 120 CAG repeats (YAC128) and wildtype littermates were used for the described experiments (18). The mice were group housed in polystyrene cages under a normal light-dark cycle (6 am to 8 pm) in a clean facility and with free access to water and standard rodent chow. All experiments were performed in accordance with the University of British Columbia animal care committee.

RNA isolation for genome wide expression analysis using the Affymetrix and Illumina platforms

Striatal tissue was collected from 12 and 24 month old YAC128 mice and wildtype littermates. Animals were anaesthetized using Avertine and tissue was collected in RNAlater (Ambion) and stored at -80°C prior to RNA isolation. Homogenization of tissue was performed using a Fastprep Homogenizer (ThermoScientific). Total RNA from mouse striatum was extracted using the Qiagen RNeasy mini kit and eluted in 30µl RNase free water according to manufacturer’s instructions. For 12 month old mice, total RNA from four YAC128 and four wildtype littermates was extracted. For 24 month old mice, total RNA from six YAC128 and four wildtype littermates was extracted. Total RNA was quantified using the Nanodrop spectrophotometer (ThermoScientific) and the integrity of the total RNA was determined electrophoretically on the RNA Nano Assay Chip run on the Bioanalyzer 2100 (Agilent Tech).
**Illumina beadarrays**

Biotinylated antisense RNA was generated using the Illumina Totalprep RNA amplification kit (Illumina®) and hybridized to an Illumina Sentrix® Mouse-6 sample BeadChip (Illumina®). The BeadChip used contained six arrays and interrogated 47,769 transcripts derived from the National Center for Biotechnology Information (NCBI) Reference Sequence (RefSeq) database, the RIKEN® FANTOM™2 database and the Mouse Exonic Evidence Based Oligonucleotide (MEEBO) set. The BeadChips were scanned on the Bead express platform (Illumina®) and gene expression analysis was undertaken using the Illumina BeadStudio software (Illumina®).

**Affymetrix**

RNA was prepared according to the manufacturer’s two-cycle target labeling procedure (Affymetrix Inc.) with a starting amount 25 ng total RNA per sample. We used the Mouse Genome 430 2.0 Arrays (Affymetrix Inc.) for all samples. This mouse array is a single GeneChip comprising over 45,000 probe sets representing more than 34,000 well-substantiated mouse genes.

**Statistical Analysis of Illumina and Affymetrix genome-wide expression data**

Quality assessment was performed using the affy R package (76). Pre-processing of the Illumina and Affymetrix microarrays was done using Beadarray (77) and Affymetrix GeneChip® Operating Software (GCOS), respectively. In the analysis, the data was split into two based on age, 12 and 24 months, respectively. Chips were normalized using Robust Multiarray Average (RMA). Within each subset, we performed a t-test between the wild-type and transgenic YAC128 data. The genes were ranked based on p-values and...
assigned to probes using Gemma (22) and Illumina and Affymetrix provided annotations.
These ranks were subsequently used to create the combined rank list described. We
calculated the combined rank for each target by averaging the individual ranks on the
Affymetrix and the Illumina platform, creating a new gene list based on the calculated
combined rank value obtained. Corrections for multiple testing were performed using the
false discovery rate (FDR) according to Benjamini & Hochberg (21, 78).

Human material

RNA was extracted from fresh-frozen samples (stored in -80°C) of human caudate
collected with minimal postmortem interval to autopsy from six HD-gene-positive cases
and five age- and sex-matched controls as described previously (10). All samples were
carefully selected on the basis of RNA quality and the HD cases were additionally graded
according to the Vonsattel grade of disease pathology (scale 0–4) with two samples each
representing HD grade 1, 2 and 3, respectively (79). RNA was extracted from human
caudate using TRIzol (Invitrogen) followed by RNeasy column cleanup (Qiagen) using
the manufacturers’ protocols.

Relative quantification of mRNA by real-time quantitative RT-PCR

Tissue samples from striatum were sampled from naïve YAC128 and wildtype littermates
at 3, 6, 9, 12 and 24 months of age, respectively. Striatal tissue from three, six, and nine
month old samples were frozen immediately at -80°C. Striatal samples from 12 and 24
month old mice were stored in RNAlater at -80°C. Homogenization of each individual
tissue was performed using a Fastprep Homogenizer (ThermoScientific). We isolated
total RNA from homogenized tissues using the Qiagen RNeasy minikit. We carried out
reverse transcription cDNA synthesis using the Quantitect Reverse Transcription kit (Qiagen) according to manufacturer’s instructions. In total, striatal RNA was extracted from eight wildtype and five YAC128 mice for three months; nine wildtype and eight YAC128 mice for six month old; five wildtype and six YAC128 mice for nine month old; six wildtype and seven YAC128 mice for twelve month old; and seven wildtype and six YAC128 mice for 24 month old mice. Quantitative analyses of mRNA expression were performed using FastSYBR®green mix according to manufacturer’s instructions (Applied Biosystems). Amplification of cDNA was performed using the 7500 Fast Real-Time PCR System (Applied Biosystems). Primers were constructed over exon/exon boundaries to avoid amplification of contaminating genomic DNA. All primers were designed using the Primer Express software version 3.0 (Applied Biosystems). Primer sequences are available as supplementary material (Table S3 for mouse and S4 for human qPCR analysis). Relative quantification of mRNA levels was calculated using the standard curve method, with amplification of target mRNA and control genes in separate wells. Standard curves were created using ten fold serial dilutions of either mouse liver, cortex or kidney cDNA. Each sample was run in duplicate. The relative amount of mRNA in each well was calculated as the ratio between the target mRNA and a normalization factor calculated from the endogenous levels of the reference genes analyzed. In figure 1-3, each sample ratio was divided by the mean of the wildtype littermate control samples for each individual time-point. The following targets were analyzed for all five time-points according to above; Wilm’s tumor homolog 1 (Wt1), DNA-damage-inducible transcript 4-like (Ddit4), Germ cell-specific gene 1-like protein (Gsg1l), Solute carrier family 45 member 3 (Slc45a3), DAZ interacting protein 1-like
(Dzip1l), Retrotransposon gag domain containing 4 (Rgag4), Protocadherin-20 (Pcdh20), Scm-like with four MBT domains protein 2 (Sfmbt2), Aspartoacylase-2 (Aminoacylase3) (Acy3), Polymerase (RNA) II (DNA directed) polypeptide A (Polr2a), Actinin alpha-2 (Actn2), Dimethylarginine dimethylaminohydrolase 1 (Ddah1) and Protein phosphatase 1, regulatory (inhibitor) subunit 9A (Ppp1r9a). Genes that were not annotated or only described as cDNA clones at the time of analysis were not included in the qRT-PCR analysis of the top ten genes. These genes include Dio3os (Affymetrix rank list) and D830041|17Rik (Illumina rank list) (Fig. S1).

Calculation of Normalization factor for quantitative real-time PCR data

We analyzed multiple reference genes for normalization of the qPCR data; 18S ribosomal RNA (Rn18s), beta-2 microglobulin (B2m), Glyceraldehyde-3-phosphate dehydrogenase (Gapdh), actin-beta (Actb), Ribosomal protein, large, P0 (Rplp0) and Hypoxanthine guanine phosphoribosyl transferase 1 (Hprt1). We applied the GeNorm software analysis for calculation of the most accurate normalization factor for our data measurements at 3, 6, 9, 12 and 24 months, respectively (80). The normalization factor (NF) was based on geometric averaging of multiple internal control genes. Calculations were based on the average expression stability and pairwise variation analysis using the GeNorm-software (http://medgen.ugent.be/~jvdesomp/genorm/). Accurate normalization factors were calculated based on; Gapdh, Rn18s and Rplp0 for three month data; Gapdh, Actb and Hprt1 for six, nine and 24 month data; Rn18s, Actb and Hprt1 for 12 month data; ACTB, GAPDH and PGK1 (Phosphoglycerate kinase 1) for human HD and control caudate samples.
**Statistical analysis**

Mann-Whitney 2-tailed U test was performed using GraphPad Prism version 4 for statistical analysis of the mRNA expression.

**Funding**

This work was supported by the Swedish Research Council for KB, Canadian Institutes of Health Research (BRL, MRH, MAP), CHDI Foundation (BRL, MRH), the EPFL (RL-C, AK) and Michael Smith Foundation for Health Research (BRL, MAP). BRL is a Canadian Institutes of Health Research Clinician-Scientist and Michael Smith Foundation for Health Research Scholar. PP is a Canadian Institutes of Health Research New Investigator and Michael Smith Foundation for Health Research scholar. This work was supported by the National Institutes of Health [GMO76990 to PP]. MRH is a Canada Research Chair in Human Genetics.

**Acknowledgements**

The authors wish to thank Ge Lu, Catherine Carter, Nagat Bissada, Laila Al-Alwan and the Lausanne DNA array facility for technical assistance.

**Conflict of Interest Statement**

The authors declare that they have no conflicts of interest.
References


Legends to figures

Figure 1

Time course analysis of differentially expressed genes identified on the Affymetrix and Illumina platforms. We analyzed the top ten genes from the Affymetrix and Illumina results, respectively. YAC128 and wildtype littermate controls were analyzed at 3, 6, 9, 12 and 24 months of age by qPCR. We analyzed Wilm’s tumor 1 homolog (Wt1), DNA damage-inducible transcript 4 like (Ddit4l), Germ cell-specific gene 1-like protein (Gsg1l) and Solute carrier family 45, member 3 (Slc45a3). YAC128 showed an up-regulation of Wt1 transcript levels at 3 months (p=0.029), 12 months (p=0.0025) and 24 months (p=0.0012) and a down-regulation at 9 months (p=0.0173) compared to controls (A). Ddit4l was down-regulated in YAC128 compared to controls at 6 months (p=0.0079), nine months (p=0.030) and 24 months (p=0.014) (B). Gsg1l was only down-regulated in YAC128 at 24 months of age compared to controls (p=0.018) (C). Slc45a3 showed up-regulation at 9 months (p=0.029), 12 months (p=0.0087) and 24 months (p=0.0152) in YAC128 compared to controls (D). In addition to Wt1, DAZ interacting protein 1-like (Dzip1l) and Retrotransposon gag domain containing 4 (Rgag4) were identified on the Illumina platform as differentially expressed at 24 month of age. YAC128 showed down-regulation of Dzip1l at 6 months (p=0.0055), 12 months (p=0.0012) and 24 months (p=0.0221) compared to controls (E). Rgag4 was down-regulated in YAC128 compared to controls at 6 months (p=0.0010) and 24 months (p=0.0082) (F). YAC128 data was normalized to the calculated average for the wildtype controls for each individual target and time-point. The bars show the mean ± SEM for each target. Statistical analysis was performed using Mann-Whitney 2-tailed U test; *
p<0.05; ** p< 0.01; *** p<0.001. White and grey bars indicate wildtype controls and YAC128, respectively.

**Figure 2**

Time course analysis of differentially expressed genes identified using the combined rank list for 24 month old YAC128 and controls. We analyzed the top ten genes from the established combined rank list. YAC128 and wildtype littermate controls were analyzed at 3, 6, 9, 12 and 24 months of age by qPCR. In addition to *Wt1* and *Ddit4l*, we analyzed Protocadherin 20 (*Pcdh20*), Sem-like with four mbt domains 2 (*Sfmbt2*), Aspartoacylase 3 (*Acy3*) and Polymerase (RNA) II (DNA directed) polypeptide A (*Polr2a*) that were included in the top ten genes of the combined rank list for 24 month old mice. *Pcdh20* was up-regulated in YAC128 compared to controls at 3 months (p=0.0186), 9 months (p=0.0173), 12 months (p=0.0012) and 24 months (p=0.035) (A). *Sfmbt2* was up-regulated in YAC128 compared to controls at 12 months (p=0.0012) and 24 months (p=0.0012) (B). *Acy3* was up-regulated at 12 months (p=0.0025) and 24 months (p=0.0012) (C). There was a tendency for increase of *Acy3* transcript levels also at 9 months (p=0.056). *Polr2a* transcript levels were increased in YAC128 at 12 months (p=0.035) and 24 months (p=0.0047) (D). YAC128 data was normalized to the calculated average for the wildtype controls for each individual target and time-point. The bars show the mean ± SEM for each target. Statistical analysis was performed using Mann-Whitney 2-tailed U test; * p<0.05; ** p< 0.01; *** p<0.001. White and grey bars indicate wildtype controls and YAC128, respectively.
Figure 3

Time course analysis of differentially expressed genes identified using the combined rank list for 12 month old YAC128 and controls. We analyzed the top ten genes from the established combined rank list. YAC128 and wildtype littermate controls were analyzed at 3, 6, 9, 12 and 24 months of age by qPCR. We analyzed Actinin alpha 2 (Actn2), Dimethylarginine dimethylaminohydrolase 1 (Ddahl) and Protein phosphotase 1, regulatory (inhibitor) subunit 9A (Ppp1r9a). Actn2 was down-regulated in YAC128 compared to controls at all five time-points; at 3 months (p=0.0040), 6 months (p=0.036), 9 months (p=0.0087), 12 months (p=0.0012) and 24 months (p=0.014) (A). Ddahl transcript levels were lower in YAC128 compared to controls at 6 months (p=0.021), 12 months (p=0.022) and 24 months (p=0.0012) (B). Ppp1r9a was down-regulated in YAC128 compared to controls at 12 months (p=0.0047) and 24 months (p=0.0012) (C). YAC128 data was normalized to the calculated average for the wildtype controls for each individual target and time-point. The bars show the mean ± SEM for each target. Statistical analysis was performed using Mann-Whitney 2-tailed U test; * p<0.05; ** p< 0.01; *** p<0.001. White and grey bars indicate wildtype controls and YAC128, respectively.

Figure 4

Transcriptional alterations identified in human HD caudate compared to controls. We analyzed the genes identified as differentially expressed in YAC128 transgenic model when validating the results from the individual platforms: Affymetrix and Illumina. We quantified WT1, DDIT4L, DMPK, GSG1L, SLC45A3, DZIP1L and RGAG4 in human HD
caudate (n=6) and controls (n=5) by qPCR. *WT1* (A), *DDIT4L* (B) and *DZIP1L* (F) did not show differential expression, although showing trends for up-regulation of *WT1* and *DZIP1L* in HD cases compared to controls. *GSG1L* showed down-regulation in HD caudate with borderline significance (p=0.051) (D), while *SLC45A3* was up-regulated in HD cases compared to controls (p=0.0043) (E). *RGAG4* did not show any differential expression between HD cases and controls (G). The middle line of the box and whiskers plot shows the median, the top and bottom lines show the 75th and 25th percentile, respectively. The top and bottom whiskers indicate the largest and smallest values.

Statistical analysis was performed using Mann-Whitney 2-tailed U test; * p<0.05; ** p<0.01; *** p<0.001.

**Figure 5**

Transcriptional alterations identified in human HD caudate compared to controls. We analyzed the genes identified as differentially expressed in the YAC128 transgenic mice when validating results from the established 12 and 24 month combined rank lists. In addition to *WT1* and *DDIT4L*, we analyzed *PCDH20*, *SFMBT2*, *ACY3* and *POLR2A* mRNA expression in human caudate of HD cases and controls by qPCR. These genes were initially identified to be differentially expressed in the YAC128 compared to controls when validating genes from the combined rank list for 24 month old mice. *PCDH20* was down-regulated in HD cases compared to controls (p=0.0043) (A). *SFMBT2* (B), *ACY3* (C) and *POLR2A* (D) were not differentially expressed in HD cases compared to controls, although, *SFMBT2* and *ACY3* showed tendencies for up-regulation in HD cases compared to controls. We also analyzed *ACTN2*, *DDAH1* and *PPP1R9A* mRNA expression in human caudate of HD cases and controls. These genes were initially
identified to be differentially expressed in the YAC128 compared to controls when validating genes from the combined rank list for 12 month old mice. All three genes showed transcriptional changes. \textit{ACTN2} (E) and \textit{PPP1R9A} (F) were both down-regulated in HD cases compared to controls (p=0.017), while \textit{DDAH1} (G) was up-regulated in HD cases compared to controls (p=0.017). The middle line of the box and whiskers plot shows the median, the top and bottom lines show the 75\textsuperscript{th} and 25\textsuperscript{th} percentile, respectively. The top and bottom whiskers indicate the largest and smallest values. Statistical analysis was performed using Mann-Whitney 2-tailed U test; * p<0.05; ** p<0.01; *** p<0.001.
Tables

Table 1. Summary of validated genes in YAC128 and control mice using quantitative real-time PCR

<table>
<thead>
<tr>
<th></th>
<th>24 months</th>
<th>12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumina</td>
<td>Affymetrix</td>
<td>Combined rank</td>
</tr>
<tr>
<td>Wt1</td>
<td>Wt1</td>
<td>Wt1</td>
</tr>
<tr>
<td>Ntng1</td>
<td>Ddit4l</td>
<td>Ntng1</td>
</tr>
<tr>
<td>Rab21</td>
<td>DMPK</td>
<td>Tob2</td>
</tr>
<tr>
<td>Dzip11</td>
<td>Gsg11</td>
<td>Pcdh20</td>
</tr>
<tr>
<td>Dok3</td>
<td>Dleu7</td>
<td>Sfmbt2</td>
</tr>
<tr>
<td>Olfr1502</td>
<td>Spata5</td>
<td>Pogk</td>
</tr>
<tr>
<td>Fndc8</td>
<td>Entpd7</td>
<td>Acy3</td>
</tr>
<tr>
<td>Rgag4</td>
<td>Grasp</td>
<td>Hrbl</td>
</tr>
<tr>
<td>Ddc</td>
<td><strong>Sle45a3</strong></td>
<td>Polr2a</td>
</tr>
<tr>
<td>Usmg5</td>
<td>Rtel1</td>
<td>Ddit4l</td>
</tr>
<tr>
<td></td>
<td>3/10</td>
<td>4/10</td>
</tr>
</tbody>
</table>

The top ten genes from the Affymetrix, Illumina, and the combined rank lists for 12 and 24 months, were validated by qPCR in independent cohorts of YAC128 and control mice. Genes in bold were differentially expressed between YAC128 and controls. We created combined rank lists for the 12 and 24 month data by averaging the individual ranks generated from the Affymetrix and Illumina results for each target.
Table 2. Summary of expressional profiling and transcriptional alterations in YAC128 and HD cases

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>YAC128 p-value</th>
<th>YAC128 Affy 28 p-value</th>
<th>YAC128 Affy 1 Illumina p-value</th>
<th>YAC128 Illumina FC³</th>
<th>YAC128 qPCR FC³,7</th>
<th>Human Affy p-value⁵</th>
<th>Human Affy FC³,5</th>
<th>Human qPCR⁶ ↑↓</th>
<th>YAC128 vs. human HD concordance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt1¹</td>
<td>2.1E-06</td>
<td>0.76</td>
<td>1.5E-04</td>
<td>0.20</td>
<td>↑19</td>
<td>-0.12</td>
<td>p=0.08</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ddit4l¹</td>
<td>2.9E-06</td>
<td>-1.2</td>
<td>3.3E-02</td>
<td>-0.39</td>
<td>↓16,9,24</td>
<td>-1.2</td>
<td>0.23</td>
<td>7.0E-03</td>
<td>p=0.66</td>
</tr>
<tr>
<td>Gsg1l¹</td>
<td>1.5E-05</td>
<td>-0.53</td>
<td>2.8 E-01</td>
<td>-0.029</td>
<td>↓24</td>
<td>-0.79</td>
<td>-0.098</td>
<td>4.0E-03</td>
<td>p=0.052</td>
</tr>
<tr>
<td>Scl45a3¹</td>
<td>1.0E-04</td>
<td>1.1</td>
<td>9.3E-02</td>
<td>0.13</td>
<td>↑9,12,24</td>
<td>0.53</td>
<td>1.1E-02</td>
<td>0.17</td>
<td>p=0.004</td>
</tr>
<tr>
<td>Dzip1l¹</td>
<td>5.8E-01</td>
<td>0.13</td>
<td>5.4E-04</td>
<td>0.053</td>
<td>↓6,12,24</td>
<td>-0.34</td>
<td>3.6E-02</td>
<td>0.083</td>
<td>p=0.13</td>
</tr>
<tr>
<td>Rgag4²</td>
<td>3.7E-01</td>
<td>-0.30</td>
<td>1.1E-03</td>
<td>-0.16</td>
<td>↓6.24</td>
<td>-0.46</td>
<td>NA</td>
<td>NA</td>
<td>ND</td>
</tr>
<tr>
<td>Pcdh20²</td>
<td>3.0E-04</td>
<td>1.1</td>
<td>8.5E-03</td>
<td>0.11</td>
<td>↑3,9,12,24</td>
<td>0.50</td>
<td>2.5E-07</td>
<td>-0.71</td>
<td>p=0.004</td>
</tr>
<tr>
<td>Sfmbt2²</td>
<td>3.5E-03</td>
<td>0.84</td>
<td>1.4E-03</td>
<td>0.062</td>
<td>↑12,24</td>
<td>1.2</td>
<td>6.4E-01</td>
<td>-0.02</td>
<td>p=0.08</td>
</tr>
<tr>
<td>Acy3³</td>
<td>4.8E-03</td>
<td>0.96</td>
<td>1.1E-03</td>
<td>0.25</td>
<td>↑12,24</td>
<td>1.4</td>
<td>6.2E-01</td>
<td>-0.02</td>
<td>p=0.33</td>
</tr>
<tr>
<td>Polr2a¹</td>
<td>1.8E-03</td>
<td>0.53</td>
<td>2.5E-02</td>
<td>0.18</td>
<td>↑12,24</td>
<td>0.75</td>
<td>5.8E-01</td>
<td>0.03</td>
<td>p=0.13</td>
</tr>
<tr>
<td>Actn2²</td>
<td>1.0E-02</td>
<td>-0.69</td>
<td>2.8E-04</td>
<td>-0.50</td>
<td>↓3,6,9,12,24</td>
<td>-0.90</td>
<td>5.4E-11</td>
<td>-0.78</td>
<td>p=0.017</td>
</tr>
<tr>
<td>Ddahl²</td>
<td>1.6E-02</td>
<td>-0.29</td>
<td>6.3E-01</td>
<td>-0.058</td>
<td>↓6,12,24</td>
<td>-0.28</td>
<td>2.2E-06</td>
<td>0.45</td>
<td>↑p=0.017</td>
</tr>
<tr>
<td>Ppp1r9a²</td>
<td>1.8E-02</td>
<td>-1.4</td>
<td>1.0E-02</td>
<td>-0.26</td>
<td>↓12,24</td>
<td>-0.34</td>
<td>8.1E-09</td>
<td>-0.66</td>
<td>p=0.017</td>
</tr>
</tbody>
</table>

1) P-values and fold change data (FC) for Affymetrix and Illumina results are from data set comparing YAC128 and controls at 24 months of age.
2) P-values and fold change data (FC) for Affymetrix and Illumina results are from data set comparing YAC128 and controls at 12 months of age.
3) Fold change (FC) is shown as log₂ transformed values. Positive values indicate up-regulation in YAC128 compared to controls; Negative values indicate down-regulation in YAC128 compared to controls.
4) Arrows indicate up or down-regulation of significant transcriptional changes in YAC128 compared to controls at different ages indicated in months.
5) Human HD gene expression data has been previously published in Hodges et al. 2006 (10).
6) Arrows indicate up- or down-regulation of significant transcriptional changes in HD caudate samples compared to controls.
7) Fold change indicated shows the magnitude of transcriptional change (log₂ transformed) observed with qPCR for 24 month old YAC128 compared to control mice and 12 month old mice for Actn2, Ddahl and Ppp1r9a.
Figures

Figure 1

A. Wilm's tumor 1 homolog (Wt1) mRNA / NF

B. DNA damage-inducible transcript 4 like (Ddit4l) mRNA / NF

C. GSG1-like (Gsg1l) mRNA / NF

D. Solute carrier family 45, member 3 (Slc45a3) mRNA / NF

E. DAZ interacting protein 1-like (Dzip1l) mRNA / NF

F. Retrotransposon gag domain containing 4 (Rgag4) mRNA / NF
**Figure 2**

A. **Protocadherin 20**

B. **Scm-like with four mbt domains 2**

C. **Aspartoacylase 3**

D. **Polymerease (RNA) II (DNA directed) polypeptide A**

* and ** indicate statistically significant differences.
Figure 3

A. Actinin alpha 2

B. Dimethylarginine dimethylaminohydrolase 1

C. Protein phosphatase 1, regulatory (inhibitor) subunit 9A
Figure 4

A  Wilms' tumor 1 homolog

B  DNA damage-inducible transcript 4 like

C  Dystrophin myotonica-protein kinase

D  Germ cell-specific gene 1-like protein

E  Solute carrier family 45, member 3

F  DAZ interacting protein 1-like

G  Retrotransposon gag domain containing 4
Figure 5

A  Pro-protease in 20

B  Smc-like with four mbt domains 2

C  Aspartoacylase 3

D  Polymerase (RNA) II (DNA directed) polypeptide A

E  Actinin alpha 2

F  Protein phosphatase 1, regulatory (inhibitor) subunit 9A

G  Dimethylarginine dimethylaminohydrolase 1