View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by RERO DOC Digital Library

Human Molecular Genetics, 2010, Vol. 19, No. 4 **609–622**
doi:10.1093/hmg/ddp527
Advance Access published on November 23, 2009

Expression analysis of novel striatal-enriched genes in Huntington disease

Gelareh Mazarei[1], Scott J. Neal[1], Kristina Becanovic[1], Ruth Luthi-Carter[2], Elizabeth M. Simpson[1] and Blair R. Leavitt[1,*]

[1]Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada V5Z 4H4 and [2]Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Received September 1, 2009; Revised and Accepted November 18, 2009

Selective degeneration of striatal neurons is a pathologic hallmark of Huntington disease (HD). The exact mechanism(s) behind this specific neurodegeneration is still unknown. Expression studies of diseased human post-mortem brain, as well as different mouse models exhibiting striatal degeneration, have demonstrated changes in the expression of many important genes with a large proportion of changes being observed in the striatal-enriched genes. These investigations have raised questions about how enrichment of particular transcripts in the striatum can lead to its selective vulnerability to neurodegeneration. Monitoring the expression changes of striatal-enriched genes during the course of the disease may be informative about their potential involvement in selective degeneration. In this study, we analyzed a Serial Analysis of Gene Expression (SAGE) database (www.mouseatlas.org) and compared the mouse striatum to 18 other brain regions to generate a novel list of striatal-enriched transcripts. These novel striatal-enriched transcripts were subsequently evaluated for expression changes in the YAC128 mouse model of HD, and differentially expressed transcripts were further examined in human post-mortem caudate samples. We identified transcripts with altered expression in YAC128 mice, which also showed consistent expression changes in human post-mortem tissue. The identification of novel striatal-enriched genes with altered expression in HD offers new avenues of study, leading towards a better understanding of specific pathways involved in the selective degeneration of striatal neurons in HD.

INTRODUCTION

Huntington disease (HD) is an autosomal dominant neurodegenerative disorder caused by a CAG repeat expansion in exon 1 of the *HTT* gene resulting in expression of mutant huntingtin protein (HTT) with an expanded polyglutamine region. The prevalence of HD is estimated to be almost 1 in 10 000 in populations of European descent that is reduced approximately 10-fold in individuals of Asian or African ancestry. Patients with HD develop symptoms of motor, cognitive and neuropsychiatric disturbances that eventually lead to complete disability and death approximately 18–25 years following the onset of the disease.

An important hallmark of HD is the selective neuronal degeneration in the caudate and putamen (together known as the striatum), which occurs despite the ubiquitous expression of mutant HTT throughout the brain and body of HD patients. This selective degeneration can be characterized by early striatal atrophy and neuronal dysfunction (1,2) as well as 90% loss of the main population of striatal neurons, the GABAergic medium spiny neurons (MSNs), by the late stages of the disease (1). Other brain regions are also affected in HD including cortex, globus pallidus, thalamus and hypothalamus (2,3). The severity of striatal pathology is correlated with the degree of motor and cognitive impairments (2,4), suggesting a central role for striatal degeneration in HD symptomatology. However, the underlying mechanism by which mutant HTT causes selective striatal degeneration in HD is not well understood. Consequently, no effective treatment or cure exists for HD.

The selective neuronal degeneration in HD has led to many studies investigating the unique physiology of the striatum compared with other brain regions. One such study focused

*To whom correspondence should be addressed. Tel: +1 6048753801; Fax: +1 6048753840; Email: bleavitt@cmmt.ubc.ca

© *The Author 2009. Published by Oxford University Press. All rights reserved.*
For Permissions, please email: journals.permissions@oxfordjournals.org

on the expression of striatal-enriched genes (genes with higher relative expression in the striatum compared with other brain regions) in a mouse model of HD as well as in human post-mortem brain (5). Interestingly, the product of these genes has been associated with biological processes previously implicated in HD (6). Therefore, although results from brain global transcriptional analyses (7–10) have proven to be insightful, we propose that the study of striatal-enriched genes will more specifically contribute to our understanding of the unique striatal physiology and its selective susceptibility to degeneration in HD.

In this study, we analyzed the Serial Analysis of Gene Expression (SAGE) database from the Mouse Atlas of Gene Expression project (www.mouseatlas.org) (11) to compare the relative abundance of transcripts in normal striatum to that in 18 other brain regions and thus to select for striatal-enriched genes. This Mouse Atlas project includes spatially and temporally restricted gene-expression profiles throughout murine development in a total of 200 cells and tissues (11). SAGE is a relatively unbiased method for large-scale gene expression profiling as, unlike microarray methods, it does not require prior knowledge of the genes expressed. Thus, it has the potential to identify novel genes (12).

We generated a list of novel striatal-enriched genes which included annotated transcripts not previously described as striatal-enriched as well as unannotated transcripts expressed specifically in the striatum. We analyzed the expression of the candidate genes in the YAC128 mouse model of HD. YAC128 mice express mutant *HTT* under its endogenous regulatory elements from a yeast artificial chromosome (YAC) transgene. YAC128 mice exhibit progressive motor dysfunction, cognitive impairment, impaired lifespan and progressive selective degeneration of specific neurons in the striatum thereby recapitulating many aspects of the human disease (13,14). By 12 months, there are clear reductions in striatal volume, striatal neuronal counts and striatal neuronal cross-sectional area in these mice (13). Our recent studies have also identified HD-related transcriptional changes in this mouse model at 12 and 24 months using microarrays (9) (Becanovic *et al*., submitted for publication). In the present study, we employed quantitative real-time PCR as a more sensitive method to detect changes in expression. We detected significant changes in the relative abundance of 11% of our candidate striatal-enriched transcripts in this mouse model at 12 months of age. Subsequent analysis of orthologous transcripts expression in human post-mortem caudate led to the identification of the adaptor-related protein complex 1, sigma 1 subunit (*Ap1s1*) and Cd4 Antigen (*Cd4*) as genes whose expression is significantly down-regulated in both YAC128 mouse striatum and caudate from HD patients.

We propose that the study of genes with enriched striatal expression is a valid approach to identify candidate genes involved in HD pathogenesis.

RESULTS

Identification of striatal-enriched SAGE tags

To identify novel striatal-enriched transcripts, we compared the p84 SAGE library from striatum with libraries from 18 other brain regions in this database. A set of stringent selection criteria (Fig. 1 and Materials and Methods) was developed to establish a list of novel striatal-enriched genes. The algorithm used to derive our candidate gene list is depicted in Figure 1. To assess the validity of our SAGE analysis, we compiled a list of previously reported striatal-enriched genes from two publications, Desplat *et al*. (5) and de Chaldee *et al*. (15). These two studies encompass the most comprehensive list of known striatal-enriched transcripts identified to date.

Striatal enrichment/specificity of a given transcript was determined based on its relative expression in striatum when compared with 18 other brain tissues. First, a comparison was made for adult [at 84 days postnatal (dpn)] brain libraries. There were 32 216 (potential transcripts) unique tag sequences in the 84 dpn striatum library, of which 6672 tag-sequences were more abundant in the striatum than any other brain library by at least one absolute tag count (raw count in un-normalized library). This eliminated all tag-sequences that were more abundant in any other brain region than in striatum. Of these tags, 6044 were observed only once and were exclusive to the striatum (i.e. one tag count in the striatum and zero elsewhere). These 'singletons' may result from sequencing and PCR errors or may well represent bona fide low-abundance transcripts in the tissue (11). However, an un-biased and careful analysis of such a large number of tag-sequences and their corresponding transcripts were beyond the targeted scope of our analysis, and thus they were excluded from further consideration.

A *P*-value score (see Materials and Methods) representing striatal enrichment for each tag relative to other brain regions was calculated and 268 of 628 non-singleton tags were significantly enriched. The abundance of the 268 tag-sequences was also compared among all 18 brain libraries over *five* developmental time points (pre-striatum, Theiler stage 25, and at dpns 7, 35 and 84) using the mean tag count. This tested whether or not a tag-sequence was preferentially expressed in striatum throughout the development. This criterion eliminated 81 tags from further consideration, but preserved all previously documented striatal-enriched genes among the 187 remaining tags.

Independent validation of striatal enrichment

SAGE measures absolute gene expression based on the sequencing of mRNA-derived fragments. However, SAGE libraries used in this project were constructed from a single individual mouse for each developmental time point and this must be considered when performing comparisons. SAGE libraries constructed from different individual mice or mice kept in different laboratories show correlations of only 0.66–0.78, or 0.5–0.58 for counts <100, even when the same inbred strain and tissues are used (12). Thus, to further validate striatal enrichment of the selected tags, two subsequent analyses were performed. First, tag-sequences representing known transcripts but lacking striatal-enriched expression patterns in the Allen Institute for Brain Science (AIBS) *in situ* hybridization images ($n = 82$) were eliminated from the list (Fig. 1).

At this stage, we gave consideration of the mapping of the tag sequences to identified genomic loci. Twenty-three tags

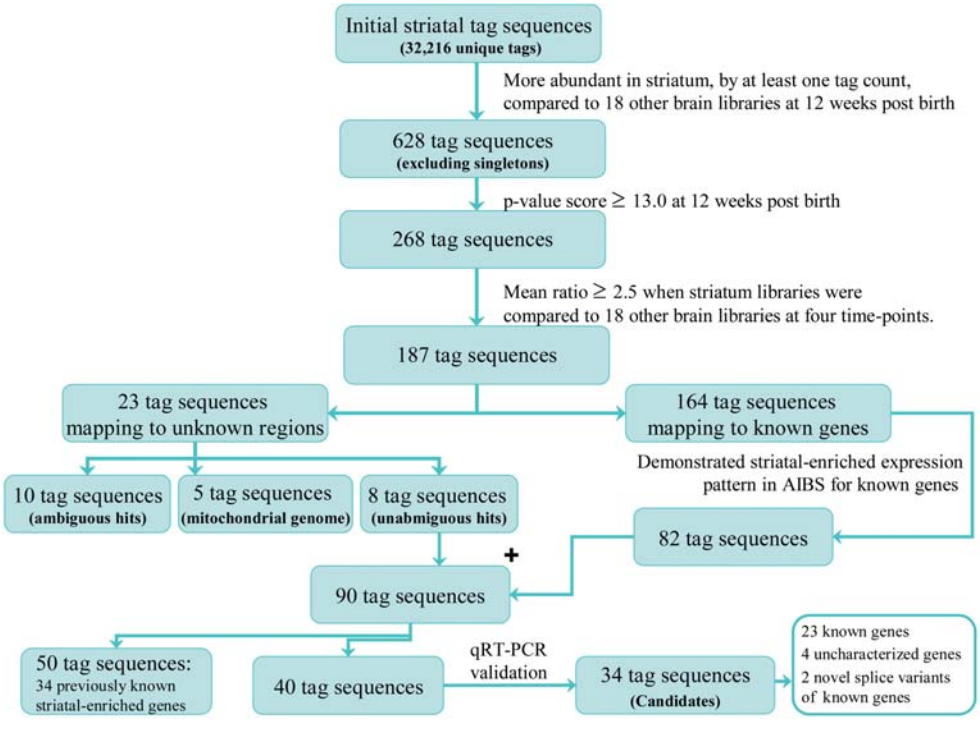

Figure 1. Selection algorithm identified novel striatal-enriched transcripts. We defined 'striatal-enriched' transcripts as transcripts with dominant patterns of expression in striatum compared to 18 other brain regions. Accordingly, selection criteria were designed to maximize detection of previously known striatal-enriched genes. Enrichment of a SAGE tag in the striatum relative to other regions was determined by using a *P*-value score ≥ 13 at postnatal day (dpn) 84 and mean ratio of ≥ 2.5 at all available time points. These two filtering steps reduced the number of tags to 187 consisting of tag-sequences mapping to both known and unknown regions of the genome. Striatal enrichment seen in Allen Institute for Brain Science (AIBS) *in situ* database images and quantitative real-time PCR were also used for further filtration of the SAGE tags. These selection criteria led to the identification of 34 novel striatal-enriched tags which corresponded to 23 known genes, 4 uncharacterized genes and 2 novel splice variants of known genes.

did not map to known genes. Fifteen of these tag-sequences either mapped to the mitochondrial genome or could not be definitely mapped to a single region in the mouse genome (ambiguous hits) using Refseq, Mammalian Gene Collection and Ensembl. Of the 90 remaining unambiguous tag-sequences, 50 mapped to 34 previously reported striatal-enriched genes (Fig. 1; Table 1) including 28/54 (~52%) of the genes reported by Desplats *et al*. (5) and 18/28 (~64%) of the genes reported by de Chaldee *et al*. (15) (Table 1).

The second validation for striatal-enriched expression was accomplished by qRT–PCR analysis. Expression analysis was conducted for 40 tag-sequences that mapped to unambiguous regions (a single sense position gene, an expressed sequence tag (EST) or an AceView predicted transcript). Eighteen of 40 transcripts (*Ablim2, Actn1, Car11, Cd4, Cyld1, Gpr83, Cacna2d3, Indo, Meis2, Ppp1r9a, Sh2d5, Smpd3*, *Tbc1d8*, *Tmem158*, predicted expansion of *Gm705*, tag downstream of *Pde7b*) were significantly enriched in wild-type mouse striatum ($n = 3$). Fifteen additional transcripts (*Ap1s1, Rgs4, Grm5, Gnao1, Gpr155, Kcnab1, Phactr1, Plxnd1, Rgs7bp, C030013G03Rik, C030007101Rik*, *BB360574*, tag downstream *Tmem16C*) demonstrated a trend toward higher striatal expression but did not reach statistical significance. Most frequently, the expression of these sequences was not significantly different between cortex and striatum, but was enriched in striatum compared with the rest of the brain. This may represent physiological similarities between these two telencephalic brain regions. Five transcripts (*Me2, Rasl10db, Sytl5, Tmod1*, tag downstream *Neb1* gene) were less abundant in striatum than one or more other brain regions (one-way ANOVA, $P \geq 0.05$). qRT–PCR results for the predicted expansion of *Gm705* (mapped to an un-annotated region of the genome) (Fig. 2A) and *Tmem158* (Fig. 2B) are examples of significant striatal-enriched expression. *Ap1s1* (Fig. 2C) expression demonstrates a trend towards striatal enrichment, and *Tmod1* (Fig. 2D) with lower expression in striatum relative to other regions. The remaining qRT–PCR results for validation of all candidate transcripts are in Supplementary Material, Figure S2.

Based on the above validation, we identified 34 novel striatal-enriched transcripts (Table 2).

Computational characterization and mapping of the novel striatal-enriched SAGE tags

Tag-to-gene mapping of the 34 novel striatal-enriched SAGE tags revealed that these tags represented 23 known genes, two uncharacterized transcripts and two potential novel splice variants of known genes (Table 2). Among the known genes, cylindromatosis (turban tumor syndrome) gene (*Cyld1*), adaptor protein complex AP-1 sigma 1 (*Ap1s1*), phosphatase and actin regulator 1 (*Phactr1*) and protein phosphatase 1, regulatory (inhibitor) subunit 9A (*Ppp1r9a*) were

Table 1. Thirty-four previously reported striatal-enriched genes obtained from our SAGE analysis

Tag	Transcript	*P*-value score*	Mean ratio[a]	Corresponding tag	ID	Transcript reference
Known transcripts						
1	*Penk1* (A)	17.0	15.2	CTGCTTTGTGCTGTACA	NM_001002927	Desplats *et al*., de Chaldee *et al*.
2	*Penk1* (B)	15.6	61.2	AGAAGGGTGGGACGCCC	NM_001002927	Desplats *et al*., de Chaldee *et al*.
3	*Ppp1r1b* (A)	17.0	14.3	TCCCTCCCTTAGTATCC	NM_144828	Desplats *et al*., de Chaldee *et al*.
4	*Ppp1r1b* (B)	14.7	14.7	CTTTTCCGGGTCTCAGA	NM_144828	Desplats *et al*., de Chaldee *et al*.
5	*Rgs9*	17.0	16.5	TGGGCCAGAAGATATCC	NM_011268	Desplats *et al*., de Chaldee *et al*.
6	*Rasd2*	17.0	17.8	TTACATATGCCTTCCTC	XM_204287	Desplats *et al*.
7	*Arpp21* (A)	17.0	7.0	GGACGGCAGTTACCTCT	NM_028755	Desplats *et al*., de Chaldee *et al*.
8	*Arpp21* (B)	17.0	6.2	TTACTTGATCTCAGACC	NM_028755	Desplats *et al*., de Chaldee *et al*.
9	*Arpp21* (C)	16.9	25.9	GAGAGTGCTTTCTTTAA	NM_028755	Desplats *et al*., de Chaldee *et al*.
10	*Rxrg*	17.0	26.6	GAAAGCAAACAAATGGA	NM_009107	Desplats *et al*.
11	*Gm705*	17.0	15.3	TGCCTCTACTCCGTCAC	AI836640	Desplats *et al*.
12	*St8sia3* (A)	16.9	5.0	TGAATGAGGAATGTGAC	NM_009182	de Chaldee *et al*.
13	*St8sia3* (B)	14.2	5.5	GGGAAGGGCTCACGAAG	NM_009182	de Chaldee *et al*.
14	*Drd1a* (A)	16.9	37.8	TGGATAGAACTGAAGGA	NM_010076	Desplats *et al*.
15	*Drd1a* (B)	16.6	26.4	CTTGAAATGGCTTTCTG	NM_010076	Desplats *et al*.
16	*Adora2a*	16.9	22.0	GACCCGAGCTGGATAGT	NM_009630	Desplats *et al*.
17	*Tmem90a* (A)	16.8	39.7	TCTCCCTGTTACTGTGG	NM_001033334	de Chaldee *et al*.
18	*Tmem90a* (B)	16.4	19.5	ATTTGGGGGACCTGGCC	NM_001033334	de Chaldee *et al*.
19	*Drd2*	16.8	16.4	CACCAAACCAATAAAAC	NM_010077	Desplats *et al*., de Chaldee *et al*.
20	*Dgkb* (A)	16.8	7.9	CACTCCCATAGGGTTTC	NM_178681	Desplats *et al*.
21	*Dgkb* (B)	16.4	12.4	GGTAACACTAATACGTC	NM_178681	Desplats *et al*.
22	*Adcy5*	16.8	11.9	GAACTCAAGGGTCTGCT	XM_993156	Desplats *et al*., de Chaldee *et al*.
23	*Gng7*	16.8	12.7	GCACTTTCCATTTCCGT	NM_010319	Desplats *et al*., de Chaldee *et al*.
24	*Tesc*	16.8	4.5	GAGACCATCGCCCTCTG	NM_021344	Desplats *et al*., de Chaldee *et al*.
25	*Scn4b*	16.6	14.7	CCCAGCGCTTGCGTCTA	NM_001013390	Desplats *et al*., de Chaldee *et al*.
26	*Tac1*	16.6	10.9	TTTCCTGTTTCGTGACT	NM_009311	Desplats *et al*., de Chaldee *et al*.
27	*pde10a* (A)	16.4	7.2	AAATATATTTAGAAATG	NM_011866	Desplats *et al*.
28	*pde10a* (B)	16.3	7.6	TCAAAGAATTAAAACAC	NM_011866	Desplats *et al*.
29	*pde10a* (C)	15.9	31.0	TATGCACAAGCACACAG	NM_011866	Desplats *et al*.
30	*pde10a* (D)	14.5	41.9	TTCATTTACTGTACAAA	NM_011866	Desplats *et al*.
31	*Gpr88*	16.3	7.2	TCTGTTGAAGATGGAAT	NM_022427	Desplats *et al*., de Chaldee *et al*.
32	*pde1b*	16.1	4.6	CTCTTGGTCTCCCAGGC	NM_053202	Desplats *et al*., de Chaldee *et al*.
33	*Rap1gap*	15.9	2.7	TATGTCAGTAAGGGTTG	XM_989545	Desplats *et al*., de Chaldee *et al*.
34	*Spock3*	15.8	5.0	TTAATTAATTGTGTAGA	NM_001040159	de Chaldee *et al*.
35	*Rasgrp2*	15.7	11.6	CATTGAGAAGAGGGCCA	NM_011242	Desplats *et al*.
36	*Cpne5* (A)	15.5	4.7	ACTTCATTGGGGAGTTC	NM_153166	Desplats *et al*.
37	*Cpne5* (B)	15.0	6.0	GTGATAGTGATGTTTTG	NM_153166	Desplats *et al*.
38	*Pcp4l1*	15.4	5.9	TTTAATAATGATGACAA	XM_992882	de Chaldee *et al*.
39	*Bcl11b*	15.3	6.9	TGTCTGCCATTTCATTT	NM_138576	Desplats *et al*.
40	*Gnal* (A)	15.0	7.5	CATAAAACCTGAGCTCA	NM_177137	Desplats *et al*.
41	*Gnal* (B)	13.5	3.6	TGCCTGGCACATTTAGA	NM_177137	Desplats *et al*.
42	*Arpp19*	14.8	3.2	ACTTTCAGTGGCCTTGT	NM_021548	Desplats *et al*.
43	*Ppp3ca*	13.6	2.5	CCCTCTGACGCCAACCT	NM_000944	Desplats *et al*., de Chaldee *et al*.
Novel splice-variants of known transcripts (AceView prediction available)						
45	*Adcy5.aSep07*	17.0	12.1	GAAATTGAAAAGTCTAT	N/A	Desplats *et al*., de Chaldee *et al*.
46	*Dgkb.aSep07*	16.6	5.2	TGTGTATCGTGTGTATT	N/A	Desplats *et al*.
47	*FoxP1.aSep07*	16.1	3.7	AGCGCATTCTTAAATGC	N/A	Desplats *et al*.
48	*Adcy5.aSep07*	14.6	18.9	GAGGCTCCTCCCCACTC	N/A	Desplats *et al*., de Chaldee *et al*.
49	*Strn.aSep07*	13.4	4.1	TCCACTCTCCCTGCTTT	N/A	Desplats *et al*.
50	*Arpp21.dSep07*	13.0	8.4	TGCTGTTTACCTATTTA	N/A	Desplats *et al*., de Chaldee *et al*.

Using selection criteria described in Figure 1, 50 SAGE tags corresponded to 34 previously reported striatal genes. More than one SAGE tag may correspond to a gene, representing its different splice variants. Genes published in Desplats *et al*. (5) and de Chaldee *et al*. (15) were used as points of reference.

*P-value score indicates relative enrichment of the SAGE tag in the striatum at 3 months.

[a]Mean ratio indicates relative enrichment of the SAGE tag in the striatum through five developmental time points.

represented by two or more enriched SAGE tags implying that two or more striatal-enriched splice variants exist for each of these genes. There are two annotated splice-variants *Cyld1*, and three for *Phactr1*. This is consistent with the number of striatal-enriched SAGE tags selected for these genes. For *Ppp1r9a*, there are three documented transcripts one of which was not striatally enriched. For *Ap1s1*, we selected two striatal-enriched tags, one of which corresponds to a splice-variant for this gene that has not previously been reported. Supplementary Material, Table S5 provides more descriptive information regarding transcript titles and molecular functions of the 34 novel striatal-enriched candidate transcripts obtained by our analysis.

Gene ontology (GO) provides standardized vocabulary to describe genes and their corresponding products. Using Expression Analysis Systematic Explorer (EASE)

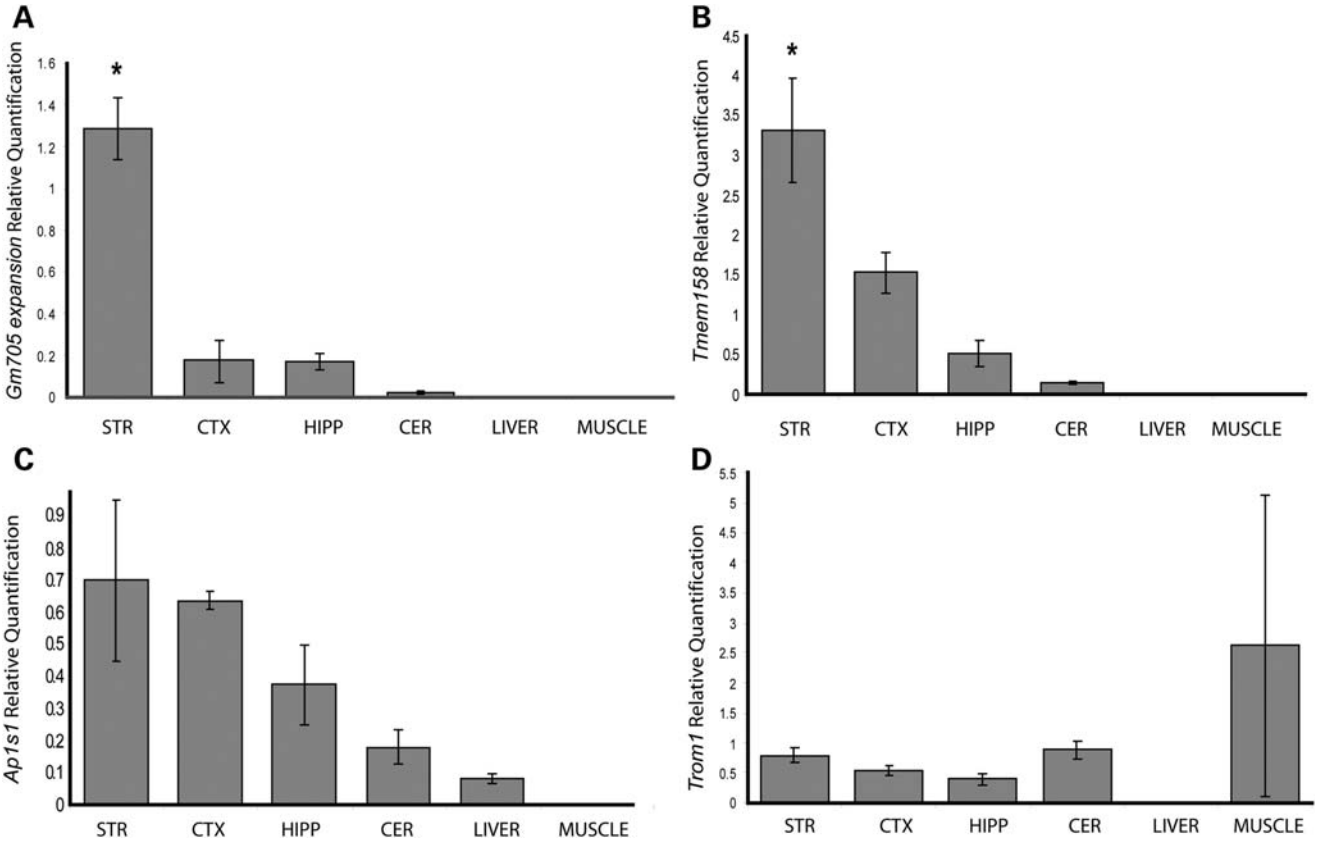

Figure 2. Examples of qRT–PCR results validating SAGE data. qRT–PCR was used for validation of 40 novel striatal-enriched transcripts obtained from SAGE and AIBS analysis. qRT–PCR experiments for these transcripts revealed three groups of candidate transcripts: (i) transcripts whose expression was shown to be significantly higher in the striatum; (ii) transcripts that show a trend towards striatal enrichment and (iii) transcripts that showed lower expression levels in the striatum compared with at least one other tissue. A cross comparison between the expression levels in striatum, cortex, hippocampus, cerebellum, liver and muscle using qRT–PCR revealed *Gm705* expansion, a previously unknown striatal-enriched transcript (**A**), and *Tmem158*, a transcript not previously known to be striatal-enriched (**B**) were significantly higher in the striatum; *Ap1s1*, a transcript selected by SAGE showed a trend towards striatal enrichment (**C**); for transcript *Trom1*, striatal-enrichment was not validated using qRT–PCR (**D**). Overall, out of 40 transcripts, 45% showed a significantly higher level in striatum, 37% showed a trend towards striatal enrichment and 17% showed lower levels of expression in striatum than at least one other tissue. Error bars represent SEM. Significant differences were determined by a one-way ANOVA analysis using a threshold value of $P < 0.05$. The qRT–PCR results for rest of the candidate transcripts are in Supplementary Material, Figure S2.

(http://david.abcc.ncifcrf.gov/), which provides statistical methods for discovering enriched biological themes within gene lists, we identified a wide range of classification throughout the list of candidate genes but no specific GO classes were over-represented in our novel striatal genes (data not shown).

The uncharacterized sequences to which our enriched tags mapped were identified as Riken transcripts *C030013G03Rik* and *C030007I01Rik*, an EST (*BB360574*), and also a genomic sequence that may contain a novel transcriptional unit (Table 2, tag no. 29). Tag-sequence no. 29 corresponds to a novel transcript on the negative strand of Chromosome 19 and may represent expansion of gene *Gm705*. There are no previously reported expression data for this region. These results further indicate that our approach was successful in the identification of novel markers of striatal expression using SAGE.

SAGE is outstanding for detecting putative new splice variants. This may include novel internal exons as well as novel alternative 3′ untranslated regions (UTRs) (12). As SAGE tags derive from the 3′ most anchoring enzyme site in a transcript, many SAGE tags are expected to be derived from the 3′-UTR which is not consistently included in gene predictions. Two of the novel striatal-enriched tag-sequences mapped to regions downstream of transmembrane protein 16C (*Tmem16C*) and phosphodiesterase 7b (*Pde7b*). AceView prediction (16) of alternate splicing exists (Table 2). Both of these predictions have demonstrated homology to the human mRNA of the corresponding genes. Similarly, in the list of the previously reported striatal-enriched transcripts, mapping of SAGE tags to AceView predictions led to identification of potential novel splice-variants for previously reported striatal-enriched genes such as adenylate cyclase 5 (*Adcy5*), diacylglycerol kinase (*Dgkb*), forkhead box P1 (*FoxP1*), Striatin (*Strn*) and cyclic AMP-regulated phosphoprotein, 21 (*Arpp21*) (Table 1).

Analysis of the candidate transcripts in the YAC128 mouse model of HD

We predicted that transcripts enriched in the striatum are relevant to the selective degeneration of this brain region in HD. Thus, expression levels of all 34 novel striatal-enriched transcripts were tested in 12-month-old YAC128 and wild-type

Table 2. Final list of novel striatal-enriched transcripts obtained using SAGE

Tag	Transcript	*P*-value score*	Mean ratio[a]	Corresponding tag	ID
Known transcripts					
1	*Tmem158*	17.0	6.3	GCCTGGTGGGCGTTTCT	NM_001002267
2	*Gnao1*	16.8	2.6	TTGGGATGTTTGTGCTG	NM_010308
3	*Phactr1* (A)	16.8	4.1	TAAATCCGAACTGCCTT	NM_198419
4	*Phactr1* (B)	14.9	14.7	CGTTCCGACTCGCTGGT	NM_198419
5	*Phactr1* (C)	13.0	25.2	GGCCTTCCGGAAATAAG	NM_198419
6	Meis2	16.6	4.9	TTAGATATTCTATGTGT	NM_131896
7	*Ap1s1* (A)	16.5	2.5	GGCCAAGTGGCTGCCTT	NM_007457
8	*Ap1s1* (B)	13.5	4.8	TGCCCTGTATGGCTGTA	NM_007457
9	*Kcnab1*	16.5	5.2	GTTAAGTTCTAAATCTT	NM_010597
10	*Car11*	16.0	3.5	TGGACGGTGGTCCTCAC	NM_009800
11	*Gpr155*	16.0	2.8	TTTCTCAATAAAATACC	XM_130346
12	*Indo*	15.9	86.8	ACTTTCCTAAGGAGTGT	NM_008324
13	*Rgs4*	15.8	6.6	AACTGCAAAGCCCTCTT	NM_009062
14	*Actn1*	15.8	2.6	GCTTCTCATTTTTCAGT	NM_134156
15	*Cacna2d3*	15.7	3.2	GAAACCAATTTAAAACT	NM_009785
16	*Ablim2*	15.4	3.5	GTCCCTTGTATAGCAAG	NM_177678
17	*Sh2d5*	15.2	2.7	GCTCTGTGTCTATGCAG	XM_357399
18	*Plxnd1*	14.9	2.7	GGCAGTGCCCATTGCCC	XM_232283
19	*Cyld (A)*	14.9	4.3	TACTCTTTCTGTATCAT	NM_173369
20	*Cyld (B)*	13.0	25.2	GATGCCAGGTTGCTAAA	NM_173369
21	*Ppp1r9a* (A)	14.9	3.1	CAGTTACAGAAGCAGAG	NM_017650
22	*Ppp1r9a* (B)	14.4	2.5	AAATAACAATTGGCCAG	NM_181595
23	*Grm5*	14.5	4.3	TTGCCAATCAATACTGG	XM_987218
24	*Tbc1d8*	14.2	5.2	CCTTTAATGTTTGTCCT	NM_018775
25	*Rgs7bp*	14.2	2.9	CAAATGGACTAACTGAC	NM_029879
26	*Smpd3*	13.9	4.6	TGCCACACTGTAGCACC	NM_021491
27	*Gpr83*	13.6	5.7	ACCAGACAGGACCCATT	NM_010287
28	*Cd4*	13.0	63.0	CTGGGGTCCTTTATGGA	NM_013488
Unknown transcripts					
29	Predicted extension of *Gm705*	14.8	21.7	TTCATCCCTCGCCCCCT	N/A
30	*C030007I01Rik cDNA*	13.9	7.4	AGCTAACATAATGGGAC	AK021055
31	*C030013G03Rik cDNA*	13.0	12.6	TTAGAAAGCAGTATTCC	AK021075
32	N/A	13.0	6.3	GATTGTGACTATTTGAA	BB360574
Novel splice-variants of known transcripts (based on AceView prediction)					
33	*Tmem16c.bSep07-unspliced*	16.5	8.9	ATGAGATTCAACTTGAA	N/A
34	*Pde7b.bSep07*	13.3	6.8	TCTATCCAGATGAACAT	N/A

Through subsequent steps of filtration, 34 novel striatal-enriched transcripts were identified and mapped to 28 transcripts representing 23 known genes not previously reported as striatal-enriched, 4 unknown transcripts and 2 novel splice-variants of known genes. Genes *Phactr1, Ap1s1, Cyld, Ppp1r9a* are represented by more than one SAGE tag corresponding to their different splice variants.

*P-value score indicates relative enrichment of the SAGE tag in the striatum at 3 months.

[a]Mean ratio indicates relative enrichment of the SAGE tag in the striatum through five developmental time points.

mice using qRT–PCR. Through this analysis, we found that three genes exhibited a significant change in their mRNA expression level. These included Cd4 antigen ($P = 0.015$, $n =$ 10 WT, 8 YAC128) and *Ap1s1* ($P = 0.003$, $n = 9$ WT, 7 YAC128) that were down-regulated in the striatum of YAC128 mice (Fig. 3A and B, respectively) and indoleamine-pyrrole 2,3 dioxygenase (*Indo*) ($P = 0.05$, $n = 10$ WT, 8 YAC128) that was up-regulated in YAC128 striatum (Fig. 3C).

We further examined the expression of these three genes in YAC128 mice at 3 months of age. Only *Indo* mRNA was significantly up-regulated at this time point ($P = 0.039$, $n = 7$ WT, 5 YAC128) (Fig. 3C). Among a selected group of previously reported striatal-enriched genes (*Ppp1rb1, Penk1, Tac1*, *Gnal, Pcp4l, Spock3, St8sia3*), we showed a significant down-regulation of *Ppp1rb* transcripts ($P = 0.029$, $n = 11$ WT, 7 YAC128) (Fig. 3D) which was consistent with reduction in *Ppp1rb1* encoded protein (DARPP-32) in YAC128 mice (17). We additionally identified, for the first time in YAC128 mice, a significant down-regulation in *Gnal* ($P = 0.02$, $n = 10$ WT, 7 YAC128) (Fig. 3E), consistent with the known expression change in human HD putamen (18). The expression level of these transcripts was further tested in 12-month YAC18 transgenic mice over-expressing normal human HTT. The results of these experiments showed no expression change of these genes between YAC18 and wild-type striatum, but demonstrate a significant expression difference between YAC18 and YAC128 striatum. These results suggest that changes in *Ap1s1, Cd4* and *Indo* are not due huntingtin dosage difference, but are likely to be a result of the expression of mutant huntingtin. The data are shown in Supplementary Material, Figure S1.

Overall, having examined all 34 candidate striatal-enriched transcripts at 12 months, we identified three genes (containing four striatal-enriched transcripts) whose levels of mRNA expression were significantly changed in the striatum of YAC128 mouse model of HD. This indicated that ~11% of the novel striatal-enriched transcripts selected through our SAGE analysis showed significantly altered expression in this HD model.

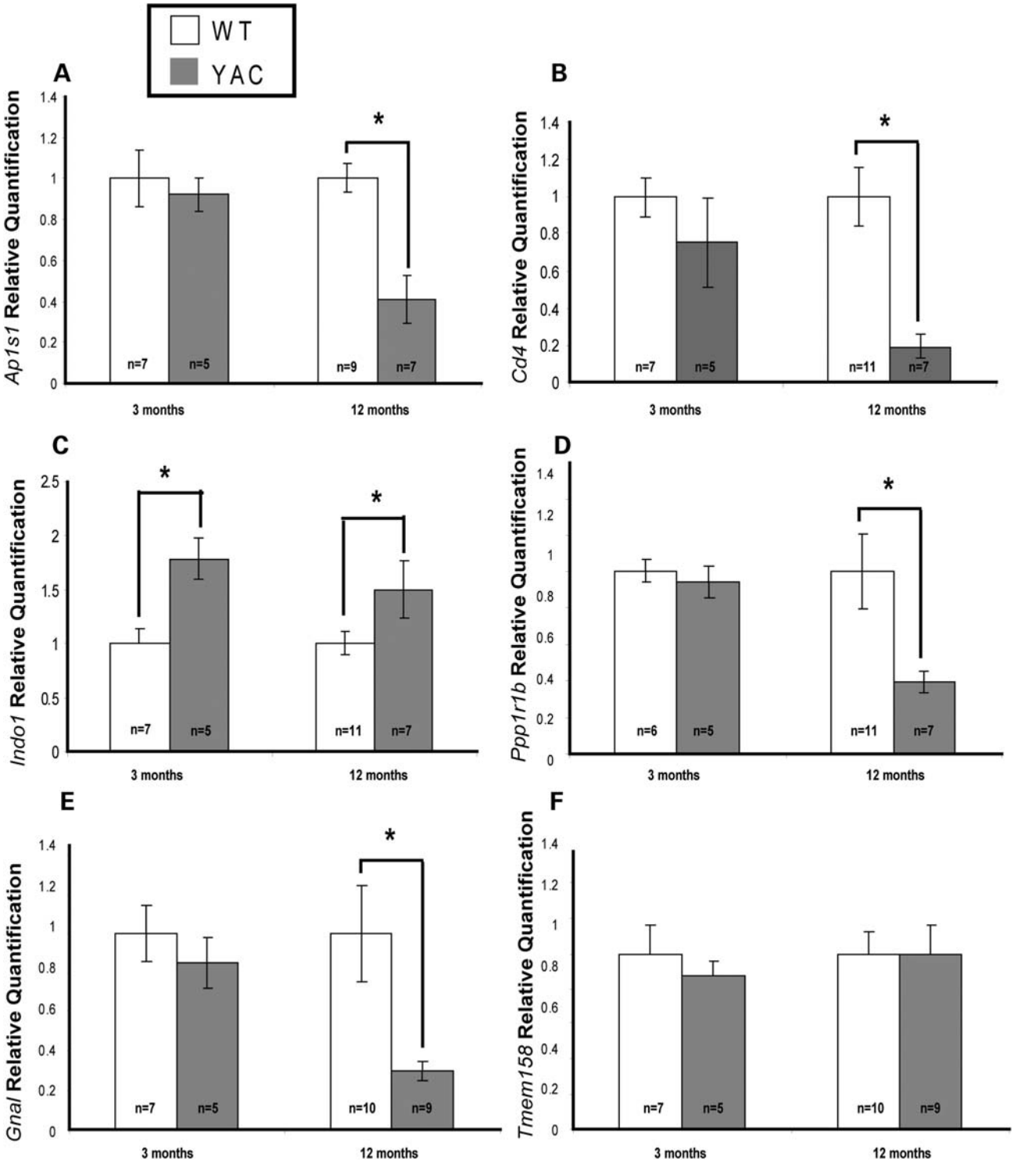

Figure 3. Five striatal-enriched genes respond to pathogenesis in the YAC128 mouse model of HD. Relative quantification of the mRNA levels of 34 novel striatal-enriched transcripts in the striatum of 12-month-old YAC128 mice revealed the greatest mRNA changes for *Ap1s1* (**A**) *Cd4* (**B**) and *Indo* (**C**). These transcripts were also tested at 3 months and only *Indo* mRNA expression exhibited a significant up-regulation at both 12 months and 3 months of age (C). As expected, well-known striatal transcripts, *Ppp1r1b* and *Gnal* showed significant changes at 12 months (**D** and **E**). *Tmem158* is an example of a striatal-enriched gene with no alteration in expression level in the YAC128 mice (**F**). Error bars represent SEM. Significant differences were determined by student's *t*-test (unpaired; two-tailed) using a threshold value of $P < 0.05$.

Analysis of mouse striatal gene expression changes in the human HD caudate

To assess the relevance of the gene expression changes we observed in the YAC128 mouse model of HD, we assayed the expression of the differentially regulated genes in caudate from HD patients and age-matched controls. Initially, *PPP1R1B* (DARPP32 mRNA) expression level was measured. This analysis was based on *PPP1R1B* previously reported altered expression levels in human HD caudate and in different mouse models of HD such as YAC128, fragment protein models and knock-in models (5,9,10,19,20). These results were reproduced in our qRT–PCR analyses of human post-mortem caudate samples where *PPP1R1B* mRNA levels were significantly reduced in HD compared with controls (Fig. 4A). qRT–PCR was also performed for human orthologs of *Cd4, Ap1s1* and *Indo*. Human *CD4* mRNA expression level was significantly increased in human HD caudate indicating an expression change opposite to what was observed in YAC128 striatum (Fig. 4B). Finally, consistent with the mouse data in YAC128, human *AP1S1* mRNA was shown, for the first time, to be significantly reduced in human post-mortem HD caudate (Fig. 4C). Despite being significantly enriched in mouse striatum, human *INDO* transcript has unexpectedly low abundance in human HD and control post-mortem tissues to date, and has not been detected in our qRT–PCR experiment. As an example of an unknown striatal-enriched transcript with un-altered levels of expression in the YAC128 mice, human ortholog of *Gm705* transcript also exhibited no transcriptional change in human HD samples.

DISCUSSION

This project identified a novel set of striatal-enriched transcripts from an existing SAGE database, compared the expression levels of these transcripts in the striatum of wild-type and YAC128 transgenic mice, and further evaluated those transcripts with altered expression in caudate samples from post-mortem HD and control brain.

The transcripts selected through our SAGE analysis are novel in two ways. (i) They have not been previously described as striatal-enriched by other studies. (ii) They are novel transcripts that were never discovered before. In either case, previous studies have failed to detect them using other techniques. SAGE allows the detection of tissue-specific low-abundance transcripts which normally remain undetectable using other techniques. SAGE also allows for the discovery of novel transcripts, without *a priori* selection bias, while other techniques such as microarrays lack this advantage. Finally, the digital nature of SAGE tags makes them easy to handle in computational analyses. In our study, 19 brain regions are easily compared. The high efficiency of this screen suggests that the MouseAtlas SAGE library database is a powerful resource to identify tissue-specific transcripts.

Confirmation of previously reported striatal-enriched genes

Although the goal of this project was to identify novel striatal-enriched genes from an existing SAGE database using a defined selection algorithm, we also assessed the selection of previously reported striatal-enriched genes at each selection step to ensure the validity of our selection process. Accordingly, while identifying novel striatal-enriched transcripts, we were also able to detect a large proportion of the striatal-enriched genes previously reported by Desplats *et al*. (5) and de Chaldee *et al*. (15). We considered four main reasons why not all of the previously reported genes met our selection criteria:

(i) Our initial selection criteria required striatal-enriched SAGE tags to be more abundant in the striatum than other brain regions at dpn 84. As an example, the previously reported striatal-enriched transcript retinoic acid receptor beta (*Rarb*) is highly expressed in striatum at dpn 7 and dpn 37, but not at dpn 84. Therefore, this transcript and other similar transcripts that are transiently enriched during striatal development are absent from our striatal-enriched candidate list.

(ii) We compared expression data for 18 different brain regions in this study and selected for SAGE tags that had higher counts in the striatum than all 18 other brain regions. Therefore, SAGE tags that manifest higher or equal abundance in at least one other brain region were eliminated. The hippocalcin (*Hpca*) SAGE tag, which was highly abundant in the striatum, showed a higher expression in CA1 and the neuronal guanine nucleotide exchange factor (*nGEF*) SAGE tag which was more abundant in visual cortex than in the striatum. These genes and other previously reported striatal-enriched genes were found to be highly expressed in the striatum, but were expressed at higher levels in at least one other brain region (e.g. synaptoporin) and based on our selection criteria, these genes are not considered striatal-enriched in this analysis.

(iii) Low abundance transcripts may not be detected by this specific SAGE analysis or may be at the singleton level, which were excluded based on our selection criteria. The SAGE tag for *Htr6* gene (5-hydroxytryptamine (serotonin) receptor 6) was excluded as a singleton.

(iv) Transcripts lacking the appropriate anchoring enzyme site (e.g. *NlaIII*) will never appear in SAGE libraries as no tag for these genes will be extracted (12). Similarly, since SAGE relies on a polyadenylated tail for tag extraction, transcripts that are not polyadenylated will be absent. Dopamine receptor 3 (*Drd3*) fails to have a valid SAGE tag since its annotated transcript lacks a poly-A tail. Therefore, despite being abundant in the striatum, many previously reported striatal-enriched transcripts were excluded from our candidate gene list due to our specific selection criteria. Table 3 summarizes why some documented striatal-enriched genes were excluded by our criteria.

In selecting our list of candidate striatal-enriched genes, we employed statistical filtration and *in situ* hybridization images from the AIBS database. The use of AIBS database was beneficial since it provided a more clear understanding of patterns of expression of annotated transcripts. The estimated correlation rate between SAGE and AIBS for known candidate transcripts was ~50% in our analysis. In a recent cDNA

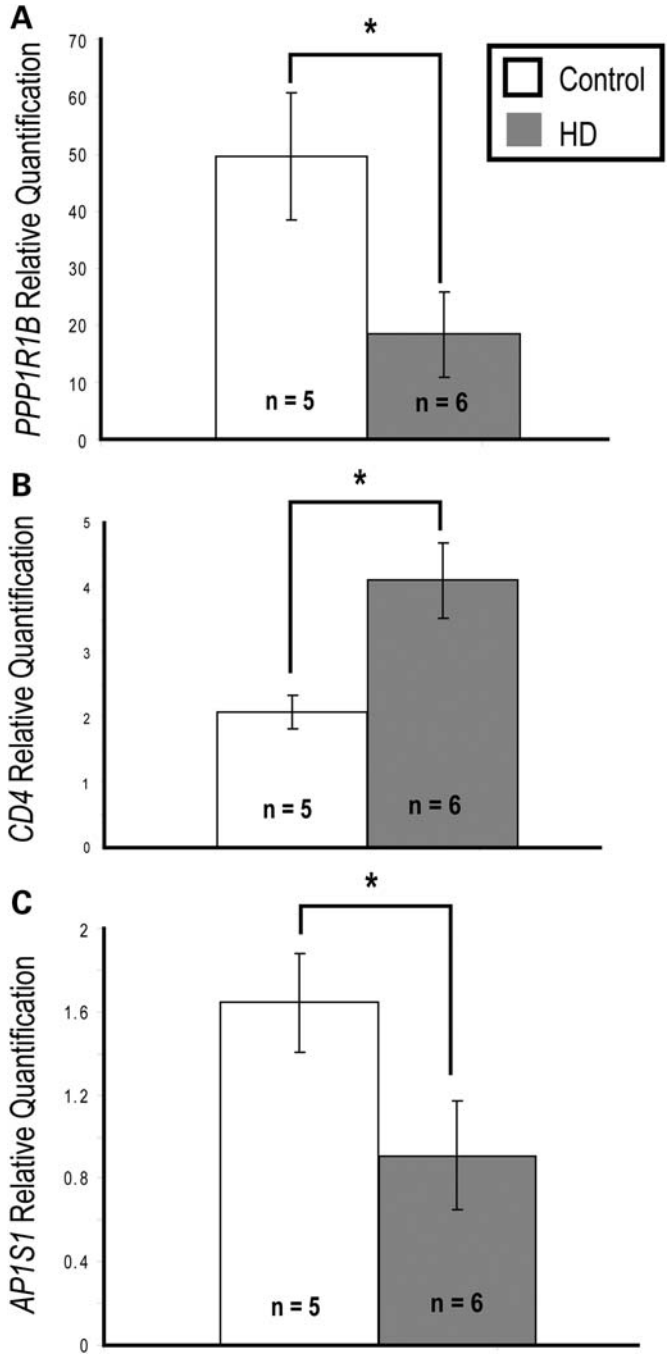

Figure 4. Alteration in expression levels of striatal-enriched transcripts in postmortem caudate of HD subjects. Transcripts that were differentially expressed in the YAC128 mouse model of HD were tested in human postmortem caudate samples from HD patients and age-matched controls. *PPP1R1B* (**A**) and *AP1S1* (**C**) showed significant down-regulation in HD subjects compared with controls, similar to expression changes observed in the YAC128 mouse model. *CD4* (**B**) was significantly up-regulated in HD subjects compared with controls, opposite to the change observed in the YAC128 mouse model. Error bars represent SEM. Significant differences were determined by student's *t*-test (unpaired; two-tailed) using a threshold value of $P < 0.05$.

microarray study by Ghate *et al*. (21), the majority of their novel striatal genes did not demonstrate clear striatal enrichment in their *in situ* images in AIBS leading to questions regarding the sensitivity of this resource. Overall, the semi-quantitative and spatially restricted expression data provided by *in situ* images was complementary to our study. However, AIBS data in isolation are not a pragmatic source for evaluating un-annotated transcripts.

Quantitative RT–PCR was used as another means to further validate selection of our candidate gene list. Unlike the Mouse Atlas of Gene Expression and AIBS data sets, multiple animals were tested, resulting in a more robust analysis of transcript expression patterns within the brain. Moreover, relative expression of all transcripts, including unknown transcribed regions, was accessible for analysis using this technique. Using qRT–PCR analysis, 18 transcripts from our candidate list (45%) were significantly enriched in striatum, and 15 transcripts (37%) showed a robust trend towards striatal enrichment. In a previous study using qRT–PCR analysis of mouse genes identified by SAGE, a correlation of 0.64 between the two methods was found, but this study only examined tags with counts greater than 45 (22). The lower correlation in our analysis could be due to the lower abundance of tags assessed (only required to be >1 count) with the majority of our candidate transcripts expressed at less than 45 counts. Similar reasoning could be used to explain the 7 out of 40 (17%) transcripts that were less abundant in striatum than at least one other tissue using qRT–PCR. These data indicate that more abundant transcripts are more likely to be validated using techniques such as qRT–PCR.

Identification and characterization of novel striatal-enriched transcripts

Computational analysis of our novel candidate list of striatal-enriched SAGE tags revealed 26 known genes, 4 uncharacterized transcripts and 2 putative novel splice variants of previously annotated genes (Table 2). Trans-membrane protein 158 (*Tmem158*), also known as Ras-induced senescence 1, had the highest *P*-value score and was validated by qRT–PCR (Fig. 2B). *Tmem158* was previously implicated in cancer (23) and can elicit new mechanisms in striatal physiology. The transcript in Table 2 with the highest mean ratio was indoleamine-pyrrole 2,3 dioxygenase (*Indo*), a gene involved in tryptophan metabolism (Supplementary Material, Fig. S2). *Indo* is striatally enriched throughout striatal development and this expression pattern was validated using qRT–PCR (data for other time points not shown). Similarly, three variants of the phosphatase and actin regulator 1 (*Phactr1*) were selected as striatal-enriched transcripts. *Phactr1* has been found to inhibit protein phosphatase 1 (PP1) activity and also known to bind to actin. *Phactr1* is thought to dictate PP1 subcellular localization and in some cases serve to potentiate PP1 toward select substrate at those sites (24). A recent study by Shao *et al*. (25) demonstrated that phosphorylation of an actin-binding factor protein profilin regulates polyglutamine aggregation. Although we found no difference in the level of *Phactr1* in WT versus YAC128 animals, future investigation is required to determine the role of *Phactr1* in phosphorylation of actin cytoskeleton components and how striatal-enrichment of this gene may lead to striatal vulnerability to neurodegeneration. *Cyld* is another interesting striatal-enriched candidate gene whose

Table 3. Some genes from Desplats *et al*. and de Chaldee *et al*. were not selected through our SAGE analysis

Striatal-enriched genes not selected using SAGE	Potential explanation
Rarb, Anap5, Nbr1	Gene is striatally enriched at a developmental stage other than adult
B3gh1, Klf9, Oprd1, Drrf, Synpr, Foxp2, Hpca, Ngef, Baiap2, Nts, Zfp503, Ppp1r16b	More abundant in at least one other brain library (after normalizing)
Htr6, Chst15, Oprm1	Transcripts detected in the SAGE library as singletons
Actn2, Efna4, Osbpl8	More abundant in at least one other library in the form of singletons
Htr4, Drd3, Oprk1, Kcnip2	No valid SAGE tag for the annotated transcript

Each gene is allocated under the category which potentially explains why the gene did not get selected by SAGE.

two different variants were selected through our analysis. SAGE analysis, Allen Brain Atlas and qRT–PCR results all indicate highly enriched patterns of expression in the striatum. The deubiquitinating enzyme CYLD is a tumor suppressor protein known to repress the NF-kB activation pathway (26). The activation of this pathway is known to have anti-apoptotic role and may play a role in response to injury in the CNS. We have not observed any significant difference in *Cyld* expression in YAC128 versus wild-type animals. However, *Cyld* remains an important candidate for future studies of striatal-enriched proteins. Computational analysis also led to the selection of four unknown striatal-enriched transcripts one of which (tag no. 29 in Table 2) maps to an un-annotated region of the chromosome 19 genomic sequence. The striatal enrichment of this transcribed region has also been shown in our study using qRT–PCR (Fig. 2A). This novel striatal-enriched tag-sequence may represent an alternative splice variant of the transcriptional unit which also encompasses the SAGE tag located in a highly conserved striatal-enriched gene *Gm705* [described as a striatal-enriched EST *AI836640* in Desplats *et al*. (5)] also selected in our analysis (Table 1, tag no. 11).

Altered expression levels of candidate striatal-enriched genes in the YAC128 mouse model of HD and human post-mortem HD brain

Next, we investigated the expression levels of the candidate novel striatal-enriched transcripts selected by our analysis in both the YAC128 mouse model of HD and in human post-mortem HD caudate. Selective striatal degeneration in the YAC128 mice manifests as striatal atrophy beginning at 9 months progressing to significant neuronal loss at 12 months of age (13). Twelve-month-old YAC128 mice also have profound motor deficits and show clear patterns of cognitive dysfunction appearing as decreased pre-pulse inhibition and habituation to acoustic startle (14).

Based on the presence of striatal neuropathology, motor dysfunction and cognitive phenotypes at 12 months in YAC128 mice, striatal transcription was tested at this time point. In addition to having the advantage of a mouse model with an age-dependent selective striatal neurodegeneration, the YAC128 model provided the advantage of testing transcriptional changes induced by mutant HTT expressed from the normal genomic context of the full-length human mutant gene. Recently, it has been shown that prevention of caspase-6-mediated cleavage of the full-length mutant HTT protein *in vivo* can rescue the progressive HD phenotype of the YAC128 model (27). This suggests that specific fragments resulting from caspase-6 proteolysis of the full-length protein, and not just any N-terminal fragments, are necessary for pathogenesis in HD. Moreover, unlike N-terminal mouse models where widespread mutant HTT aggregates rapidly accumulate in the nucleus and massive transcriptional changes have been demonstrated (5,6,10), the slower progression and more modest transcriptional changes seen in YAC128 mice expressing the full-length mutant HTT protein allow investigation of these changes throughout the time-course of pathology. Comparing gene expression results between the YAC128 model and N-terminal fragment models again demonstrates that increased HTT protein length reduces the number of polyglutamine-induced gene expression changes (7).

One of the three novel striatal-enriched genes with altered mRNA expression in YAC128 mice was Cd4. The Cd4 antigen is a membrane glycoprotein of T lymphocytes that interacts with major histocompatibility class II (MHC II) antigens and is also a receptor for the human immunodeficiency virus (HIV). The gene is known to be expressed highly in thymus and spleen but also lower levels are found in microglia in the brain (28). Transgenic mice expressing the human *CD4* gene showed no evidence of neuronal damage until their brain was activated by peripheral immune challenges which led to neurodegeneration in these mice (10). In post-mortem brain tissues from AIDS patients with opportunistic infections, but without typical HIV encephalitis, human CD4 expression correlated with neurodegeneration (10). Our finding that *Cd4* is a striatal-enriched gene may explain the susceptibility of the striatum to neuroinflammation observed in HD (29) and may provide mechanistic insight into selective striatal degeneration in HD. We demonstrate that *CD4* transcript expression is significantly up-regulated in human HD caudate samples compared with controls, implying that there is altered immune activation in the striatum of HD patients. However, the observed discrepancy in Cd4 expression change between human postmortem caudate and YAC128 striatum raises interesting questions regarding the regulation of this transcript in postmortem human brain versus mouse brain and its potential link to disease progression in mouse and human. Further investigation is required to first understand which brain cell types in the striatum of YAC128 and patients express *Cd4* (i.e. glial cells versus MSNs) and to subsequently explain the role of *Cd4* expression in the context of neurodegeneration in HD.

Indo, another gene with altered mRNA expression levels in the YAC128 mouse model, is a tryptophan-catabolizing enzyme which initiates the first and rate limiting step of the kynurenine pathway. In the brain, *Indo* can be induced in microglia by interferon-γ-producing T helper (Th) 1 cells. *Indo* is thought to initiate a negative feedback loop which decreases neuroinflammation in an animal model of multiple sclerosis (MS) (30–32). On the other hand, induction of

Indo activates the, kynurenine pathway leading to production of neurotoxic metabolites such as 3 hydroxy-Kyurenine and quinolinic acid known to be involved in HD pathogenesis (33). Accordingly, Kwidzinski and Bechmann (34) have proposed that neuroinflammation and neurodegeneration are linked by IFN-γ-mediated *Indo* induction. In a recent expression study of amyloid β-peptide-stimulated human post-mortem brain microglia, the INDO protein was found to be up-regulated identifying a potential new inflammatory pathway in Alzheimer's disease (35). The observed increase in *Indo* mRNA expression in the YAC128 mice could suggest, for the first time, a similar involvement of Indo leading to neuroinflammation described in HD (29). Expression levels of *INDO* mRNA in the human post-mortem brain were also assessed by qRT–PCR; however, this analysis was not conclusive (poor amplification) due to very low levels of *INDO* mRNA in both HD and control caudate RNA samples. It is known that RNA degradation in autopsy tissue is more severe in specific subsets of transcripts or restricted to specific gene families (34), and *INDO* transcript abundance may be decreased due to postmortem changes in the brain. Further analysis of INDO protein in the context of HD is needed to draw more conclusive results.

AP1S1 is another striatal-enriched gene which is significantly down-regulated in both YAC128 striatum and human HD postmortem caudate. The protein encoded by this gene is an important subunit of the clathrin adaptor complex involved in clathrin-coated vesicle transport of endocytosis and golgi processing. A very recent study of this gene has described *AP1S1* as the gene responsible for MEDNIK syndrome (mental retardation, enteropathy, deafness, peripheral neuropathy, ichthyosis and keratodermia) and demonstrated a critical role of *AP1S1* in development of the skin and spinal cord (36). Studies of huntingtin's protein–protein interactions have provided direct links to proteins involved in cellular transport mechanisms, specifically HTT is thought to play an important role in neuronal vesicular trafficking. Impairment of this wild-type HTT function by the polyglutamine expansion in mutant HTT has been suggested to result in decreased endocytosis (37,38). These studies provide a potential link between *AP1S1* gene function and HD. The reduction of *AP1S1* mRNA expression would be predicted to cause vesicle transport defects involving clathrin-dependent endocytosis and golgi processing. This observation, however, requires much deeper investigation of the cellular function of *AP1S1* encoded protein and further validation to confirm its role in neuronal degeneration in HD.

Detailed characterization of these novel candidate striatal-enriched transcripts will be the focus of this project in the future. Detailed investigation of alterations in our three genes at the protein level is currently underway. Modulating the levels of these candidate transcripts will also provide further information regarding their function in the brain as well as their involvement in HD pathology.

Our findings demonstrate the potential value of studying novel striatal-enriched genes to identify transcripts and pathways dysregulated by the expression of mutant HTT in the brain. This study identifies new genes involved in biological pathways and processes such as *trans*-golgi network vesicle transport and trafficking (*Ap1s1*), tryptophan metabolism in the kynurenine pathway (*Indo*) and inflammatory response antigens (*Cd4*) that represent important candidates for future investigation of HD pathogenesis. Characterization of novel striatal-specific genes and their expression changes in HD will help elucidate disease pathways that have important roles in the physiological processes and integrity of striatal neurons.

Summary

Through the use of publicly available data sets and experimental techniques, we have identified candidate genes with altered expression in the striatum of a murine HD model and in human HD caudate. The products of these genes may be involved in pathways that are relevant to HD pathogenesis, be useful as biomarkers to follow disease progression, and may be potential new therapeutic targets for the treatment of HD.

MATERIALS AND METHODS

SAGE data analysis

SAGE libraries (Supplementary Material, Table S1) were generated by the Mouse Atlas of Gene Expression Project (www.mouseatlas.org). The DiscoverySpace software application (http://www.bcgsc.ca/bioinfo/software/discoveryspace/) was used to compare the striatum library to 18 other brain region libraries at dpn 84. These libraries were compared with the striatum library one at a time after tag counts in each library were normalized to that of the striatum library (total of 18 pairwise comparisons). A '*P*-value score' for each striatal tag-sequence was obtained by calculating the sum of $(1-P\text{-values})$ for that tag in all pairwise comparisons (P-values obtained using Audic Claverie statistics): P-value score for tag x $= \Sigma\ (1-P\text{-value})$.

*IF striatum count $>$ any other brain region count then

$$P = +P$$

IF striatum count $<$ any other brain region count then

$$P = -P$$

*Personal communication: This formula has been generated by Pavle Vrljicak at BC Cancer Agency, Vancouver, BC.

Candidate striatal-specific tags were selected based on the following selection criteria: (i) more abundant in striatum, by at least one tag count, compared with 18 other brain libraries at dpn 84, (ii) P-value score ≥ 13.0 at dpn 84 where $-16 \leq P\text{-value score} \leq +17$ and (iii) mean ratio ≥ 2.5 when all striatum libraries ($n = 5$) were compared with 18 other libraries at all available time points (total of 40 libraries). Tag to gene mapping was performed using the mouse Refseq, MGC and Ensembl databases accessible through DiscoverySpace. Tag position was determined as a number whose figure represented the location of the *NlaIII* site relative to the polyA tail and whose sign represented the directionality of the tag relative to the gene to which it was mapped.

In situ hybridization database

Allen Brain Atlas (AIBS) (http://www.brainatlas.org/AIBS/) *in situ* hybridization images provided a visual assessment of relative abundance of the transcript corresponding to each SAGE tag mapping to annotated gene. Clearly, only SAGE tags mapping to known genes could be filtered through this process. A demonstrated striatal-enriched/specific expression pattern in the AIBS *in situ* hybridization database was considered a necessary criterion in our analysis. SAGE tags corresponding to unknown regions were not filtered out at this stage.

Animals

Twelve-month-old female YAC128 mice and WT littermates on FVB/N (Charles River) strain were used in comparison studies. For SAGE data validation experiments, we used tissues from 3-month-old WT female mice on FVB/N strain background. Mice were group-housed with a normal light–dark cycle and given free access to food and water. Experiments were carried out in accordance with our institutional animal care guidelines. All experiments were approved by the University of British Columbia's Committee for Animal Care.

Human post-mortem caudate and putamen samples

Supplementary Material, Table S2 provides a summary of human post-mortem caudate/putamen samples used in our experiments. These samples comprised a subset of RNAs described in a previous study [Hodges *et al*. (8)] extracted from tissues obtained from the Harvard Brain Tissue Resource Center.

Quantitative real-time PCR

Mouse experiments. For SAGE validation experiments, RNA from four brain regions (striatum, cortex, cerebellum and hippocampus) as well as liver and gastrocnemius muscle was prepared from three 3-month-old mice using the RNeasy kit (Qiagen). Similarly, for transgenic and WT comparison experiments, RNA was extracted from fresh-frozen dissected striata of 11 wild-type and 8 YAC128 transgenic 12-month-old littermate mice. RNA was also extracted from six 12-month-old YAC18 animals for YAC18 versus YAC128 comparison experiments. Primers were designed using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3www.cgiT) and amplicons spanned introns where possible. Amplicons were between 110 and 150 bp for efficient amplification. Primer efficiencies were determined using a dilution series of adult striatum cDNA. Only primer pairs with an efficiency greater than 0.98 were used in subsequent analyses. The primers were also designed to encompass the SAGE tag (See Supplementary Material, Table S2 for the list of mouse primers). An ABI 7000 real-time PCR system (Applied Biosystems) and SYBR® Green Master mix (Qiagen) were used for validations of the SAGE data. An ABI 7500 real-time PCR system (standard protocol) and SYBR® Green Master mix (Qiagen) were used to compare wild-type and YAC128 mRNA. cDNA was obtained by the reverse transcription (RT) of 1 μg of total RNA from freshly dissected brain regions using the Quantitect RT kit (Qiagen). All reactions were carried out in duplicate using 2 μl of cDNA in each reaction. Duplicates with inconsistent amplification were removed from the analysis. Absolute quantity of the targets in each sample was calculated based on the standard curve method and normalized to endogenous control, β-actin (reverse primer sequence:
5′TGTGTTGGCATAGAGGTCTTTACG3′; forward primer sequence:
5′CCAGCCTTCCTTCTTGGGTAT3′).

Human post-mortem experiments. RNA was extracted from 11 brain samples listed in Supplementary Material, Table S1 using the RNeasy kit (Qiagen). Primers were designed similar to mouse primers against human genome for human *AP1S1, CD4, INDO* and *DARPP32*. Primer efficiencies were determined as in mouse experiments. An ABI 7500 real-time PCR system (Fast protocol) using ABI Fast SYBR Green Master mix (Applied Biosystems) was used. All reactions were carried out in duplicates. For the analysis, relative quantity was calculated based on the standard curve method. Supplementary Material, Table S3 contains the human specific primers designed for qRT–PCR experiments. We analyzed multiple control genes (*GAPDH,* β-*ACTIN*, phosphoglycerate kinase 1 (*PGK1*) for normalization of the human qRT–PCR data. We applied the GeNorm software analysis for calculation of the most accurate normalization factor for our data measurements (39). The normalization factor (NF) was based on geometric averaging of multiple internal control genes and calculations were based on the average expression stability and pairwise variation analysis using the GeNorm-software.

Statistical analysis

In SAGE validation experiments, a one-way ANOVA was performed for determining differences in the relative quantity of targets of interest between different tissues. Statistical comparisons were performed in a pair-wise fashion between striatum and every other tissue. In WT versus transgenic experiments, a two-tailed Student's *t*-test was performed for determining differences in the mean between the two groups and Levene's test to determine equality of variance. Differences were assumed statistically significant if they reached the 95% confidence level. All values are represented as averages with standard error of the mean.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

The authors wish to thank Ge Lu and Austin Hill for technical assistance. We also acknowledge the Harvard Brain Tissue Resource Center for human HD and control brain samples.

Conflict of Interest statement. None declared.

FUNDING

Funding for this study was provided by the NSERC (G.M.), British Columbia Cancer Foundation (E.M.S.), the National Career Institute, National Institute of Health (E.M.S.), the EPFL (R.L.C.), Canadian Institutes of Health Research (B.R.L.), CHDI Foundation (B.R.L.), B.R.L. is a CIHR Clinician-Scientist and MSFHR Scholar.

REFERENCES

1. Aylward, E.H., Sparks, B.F., Field, K.M., Yallapragada, V., Shpritz, B.D., Rosenblatt, A., Brandt, J., Gourley, L.M., Liang, K., Zhou, H. *et al.* (2004) Onset and rate of striatal atrophy in preclinical Huntington disease. *Neurology*, **63**, 66–72.
2. Rosenblatt, A., Abbott, M.H., Gourley, L.M., Troncoso, J.C., Margolis, R.L., Brandt, J. and Ross, C.A. (2003) Predictors of neuropathological severity in 100 patients with Huntington's disease. *Ann. Neurol.*, **54**, 488–493.
3. Macdonald, V. and Halliday, G. (2002) Pyramidal cell loss in motor cortices in Huntington's disease. *Neurobiol. Dis.*, **10**, 378–386.
4. Bamford, K.A., Caine, E.D., Kido, D.K., Plassche, W.M. and Shoulson, I. (1989) Clinical-pathologic correlation in Huntington's disease: a neuropsychological and computed tomography study. *Neurology*, **39**, 796–801.
5. Desplats, P.A., Kass, K.E., Gilmartin, T., Stanwood, G.D., Woodward, E.L., Head, S.R., Sutcliffe, J.G. and Thomas, E.A. (2006) Selective deficits in the expression of striatal-enriched mRNAs in Huntington's disease. *J. Neurochem.*, **96**, 743–757.
6. Thomas, E.A. (2006) Striatal specificity of gene expression dysregulation in Huntington's disease. *J. Neurosci. Res.*, **84**, 1151–1164.
7. Chan, E.Y., Luthi-Carter, R., Strand, A., Solano, S.M., Hanson, S.A., DeJohn, M.M., Kooperberg, C., Chase, K.O., DiFiglia, M., Young, A.B. *et al.* (2002) Increased huntingtin protein length reduces the number of polyglutamine-induced gene expression changes in mouse models of Huntington's disease. *Hum. Mol. Genet.*, **11**, 1939–1951.
8. Hodges, A., Strand, A.D., Aragaki, A.K., Kuhn, A., Sengstag, T., Hughes, G., Elliston, L.A., Hartog, C., Goldstein, D.R., Thu, D. *et al.* (2006) Regional and cellular gene expression changes in human Huntington's disease brain. *Hum. Mol. Genet.*, **15**, 965–977.
9. Kuhn, A., Goldstein, D.R., Hodges, A., Strand, A.D., Sengstag, T., Kooperberg, C., Becanovic, K., Pouladi, M.A., Sathasivam, K., Cha, J.H. *et al.* (2007) 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. *Hum. Mol. Genet.*, **16**, 1845–1861.
10. Luthi-Carter, R., Strand, A., Peters, N.L., Solano, S.M., Hollingsworth, Z.R., Menon, A.S., Frey, A.S., Spektor, B.S., Penney, E.B., Schilling, G. *et al.* (2000) Decreased expression of striatal signaling genes in a mouse model of Huntington's disease. *Hum. Mol. Genet.*, **9**, 1259–1271.
11. Siddiqui, A.S., Khattra, J., Delaney, A.D., Zhao, Y., Astell, C., Asano, J., Babakaiff, R., Barber, S., Beland, J., Bohacec, S. *et al.* (2005) A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. *Proc. Natl Acad. Sci. USA*, **102**, 18485–18490.
12. Pleasance, E. and Jones, S.J.M. (2005) Evaluation of SAGE Tags for Transcriptome Study. Wang, S.M. (ed.), *SAGE: Current Technologies and Applications*, Horizon Bioscience, Norfolk.
13. Slow, E.J., van Raamsdonk, J., Rogers, D., Coleman, S.H., Graham, R.K., Deng, Y., Oh, R., Bissada, N., Hossain, S.M., Yang, Y.Z. *et al.* (2003) Selective striatal neuronal loss in a YAC128 mouse model of Huntington disease. *Hum. Mol. Genet.*, **12**, 1555–1567.
14. Van Raamsdonk, J.M., Pearson, J., Slow, E.J., Hossain, S.M., Leavitt, B.R. and Hayden, M.R. (2005) Cognitive dysfunction precedes neuropathology and motor abnormalities in the YAC128 mouse model of Huntington's disease. *J. Neurosci.*, **25**, 4169–4180.
15. de Chaldee, M., Gaillard, M.C., Bizat, N., Buhler, J.M., Manzoni, O., Bockaert, J., Hantraye, P., Brouillet, E. and Elalouf, J.M. (2003) Quantitative assessment of transcriptome differences between brain territories. *Genome Res.*, **13**, 1646–1653.
16. Thierry-Mieg, D. and Thierry-Mieg, J. (2006) AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biol.*, **7**(Suppl. 1), S12–S14.
17. Van Raamsdonk, J.M., Pearson, J., Bailey, C.D., Rogers, D.A., Johnson, G.V., Hayden, M.R. and Leavitt, B.R. (2005) Cystamine treatment is neuroprotective in the YAC128 mouse model of Huntington disease. *J. Neurochem.*, **95**, 210–220.
18. Corvol, J.C., Muriel, M.P., Valjent, E., Feger, J., Hanoun, N., Girault, J.A., Hirsch, E.C. and Herve, D. (2004) Persistent increase in olfactory type G-protein alpha subunit levels may underlie D1 receptor functional hypersensitivity in Parkinson disease. *J. Neurosci.*, **24**, 7007–7014.
19. Bibb, J.A., Yan, Z., Svenningsson, P., Snyder, G.L., Pieribone, V.A., Horiuchi, A., Nairn, A.C., Messer, A. and Greengard, P. (2000) Severe deficiencies in dopamine signaling in presymptomatic Huntington's disease mice. *Proc. Natl Acad. Sci. USA*, **97**, 6809–6814.
20. Hebb, A.L., Robertson, H.A. and Denovan-Wright, E.M. (2004) Striatal phosphodiesterase mRNA and protein levels are reduced in Huntington's disease transgenic mice prior to the onset of motor symptoms. *Neuroscience*, **123**, 967–981.
21. Ghate, A., Befort, K., Becker, J.A., Filliol, D., Bole-Feysot, C., Demebele, D., Jost, B., Koch, M. and Kieffer, B.L. (2007) Identification of novel striatal genes by expression profiling in adult mouse brain. *Neuroscience*, **146**, 1182–1192.
22. Trottier, Y., Biancalana, V. and Mandel, J.L. (1994) Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. *J. Med. Genet.*, **31**, 377–382.
23. Barradas, M., Gonos, E.S., Zebedee, Z., Kolettas, E., Petropoulou, C., Delgado, M.D., Leon, J., Hara, E. and Serrano, M. (2002) Identification of a candidate tumor-suppressor gene specifically activated during Ras-induced senescence. *Exp. Cell Res.*, **273**, 127–137.
24. Allen, P.B., Greenfield, A.T., Svenningsson, P., Haspeslagh, D.C. and Greengard, P. (2004) Phactrs 1–4: a family of protein phosphatase 1 and actin regulatory proteins. *Proc. Natl Acad. Sci. USA*, **101**, 7187–7192.
25. Shao, J., Welch, W.J., Diprospero, N.A. and Diamond, M.I. (2008) Phosphorylation of profilin by ROCK1 regulates polyglutamine aggregation. *Mol. Cell. Biol.*, **28**, 5196–5208.
26. Simonson, S.J., Wu, Z.H. and Miyamoto, S. (2007) CYLD: a DUB with many talents. *Dev. Cell*, **13**, 601–603.
27. Graham, R.K., Deng, Y., Slow, E.J., Haigh, B., Bissada, N., Lu, G., Pearson, J., Shehadeh, J., Bertram, L., Murphy, Z. *et al.* (2006) Cleavage at the caspase-6 site is required for neuronal dysfunction and degeneration due to mutant huntingtin. *Cell*, **125**, 1179–1191.
28. Gillespie, F.P., Doros, L., Vitale, J., Blackwell, C., Gosselin, J., Snyder, B.W. and Wadsworth, S.C. (1993) Tissue-specific expression of human CD4 in transgenic mice. *Mol. Cell. Biol.*, **13**, 2952–2958.
29. Bjorkqvist, M., Wild, E.J., Thiele, J., Silvestroni, A., Andre, R., Lahiri, N., Raibon, E., Lee, R.V., Benn, C.L., Soulet, D. *et al.* (2008) A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington's disease. *J. Exp. Med.*, **205**, 1869–1877.
30. Kwidzinski, E., Bunse, J., Aktas, O., Richter, D., Mutlu, L., Zipp, F., Nitsch, R. and Bechmann, I. (2005) Indolamine 2,3-dioxygenase is expressed in the CNS and down-regulates autoimmune inflammation. *FASEB J.*, **19**, 1347–1349.
31. Platten, M., Ho, P.P., Youssef, S., Fontoura, P., Garren, H., Hur, E.M., Gupta, R., Lee, L.Y., Kidd, B.A., Robinson, W.H. *et al.* (2005) Treatment of autoimmune neuroinflammation with a synthetic tryptophan metabolite. *Science*, **310**, 850–855.
32. Sakurai, K., Zou, J.P., Tschetter, J.R., Ward, J.M. and Shearer, G.M. (2002) Effect of indoleamine 2,3-dioxygenase on induction of experimental autoimmune encephalomyelitis. *J. Neuroimmunol.*, **129**, 186–196.
33. Guidetti, P., Bates, G.P., Graham, R.K., Hayden, M.R., Leavitt, B.R., MacDonald, M.E., Slow, E.J., Wheeler, V.C., Woodman, B. and Schwarcz, R. (2006) Elevated brain 3-hydroxykynurenine and quinolinate levels in Huntington disease mice. *Neurobiol. Dis.*, **23**, 190–197.
34. Kwidzinski, E. and Bechmann, I. (2007) IDO expression in the brain: a double-edged sword. *J. Mol. Med.*, **85**, 1351–1359.
35. Schwarcz, R., Whetsell, W.O. Jr and Mangano, R.M. (1983) Quinolinic acid: an endogenous metabolite that produces axon-sparing lesions in rat brain. *Science*, **219**, 316–318.
36. Montpetit, A., Cote, S., Brustein, E., Drouin, C.A., Lapointe, L., Boudreau, M., Meloche, C., Drouin, R., Hudson, T.J., Drapeau, P. *et al.* (2008)

Disruption of AP1S1, causing a novel neurocutaneous syndrome, perturbs development of the skin and spinal cord. *PLoS Genetics*, **4**, e1000296.
37. Gil, J.M. and Rego, A.C. (2008) Mechanisms of neurodegeneration in Huntington's disease. *Eur. J. Neurosci.*, **27**, 2803–2820.
38. Trushina, E., Singh, R.D., Dyer, R.B., Cao, S., Shah, V.H., Parton, R.G., Pagano, R.E. and McMurray, C.T. (2006) Mutant huntingtin inhibits clathrin-independent endocytosis and causes accumulation of cholesterol in vitro and in vivo. *Hum. Mol. Genet.*, **15**, 3578–3591.
39. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, **3**, RESEARCH0034.