Published in final edited form as:

Neurobiol Dis. 2011 June ; 42(3): 459–467. doi:10.1016/j.nbd.2011.02.008.

Gene expression profiling of R6/2 transgenic mice with different CAG repeat lengths reveals genes associated with disease onset and progression in Huntington's disease

Bin Tang¹, Tamara Seredenina², Giovanni Coppola³, Alexandre Kuhn², Daniel H. Geschwind³, Ruth Luthi-Carter², and Elizabeth A. Thomas^{1,*}

¹Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA ²Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland ³Program in Neurogenetics, Department of Neurology and Semel Institute University of California at Los Angeles, Los Angeles, California

Abstract

R6/2 transgenic mice with expanded CAG repeats (>300) have a surprisingly prolonged disease progression and longer lifespan than prototypical parent R6/2 mice (carrying 150 CAGs), however, the mechanism of this phenotype amelioration is unknown. We compared gene expression profiles in the striatum of R6/2 transgenic mice carrying ~300 CAG repeats (R6/2^{Q300} transgenic mice), those carrying ~150 CAG repeats (R6/2^{Q150} transgenic mice) and littermate wt controls in order to identify genes that may play determinant roles in the time course of phenotypic expression in these mice. Of the top genes showing concordant expression changes in the striatum of both R6/2 lines, 85% were decreased in expression, while discordant expression changes were observed mostly for genes upregulated in R6/2^{Q300} transgenic mice. Upregulated genes in the R6/2^{Q300} mice were associated with the ubiquitin ligase complex, cell adhesion, protein folding and establishment of protein localization. We qPCR-validated increases in expression of genes related to the latter category, including *Lrsam1*, *Erp29*, *Nasp*, *Tap1*, *Rab9b* and *Pfdn5* in R6/2^{Q300} mice, changes that were not observed in R6/2 mice with shorter CAG repeats, even in late stages (i.e. 12 weeks of age). We further tested *Lrsam1* and *Erp29*, the two genes showing the greatest upregulation in R6/2^{Q300} transgenic mice, for potential neuroprotective effects in primary striatal cultures overexpressing a mutated human huntingtin (htt) fragment. Overexpression of *Lrsam1* prevented the loss of NeuN-positive cell bodies in htt171-82Q cultures, concomitant with a reduction of nuclear htt aggregates. *Erp29* showed no significant effects in this model. This is consistent with the distinct pattern of htt inclusion localization observed in R6/2^{Q300} transgenic mice, in which smaller cytoplasmic inclusions represent the major form of insoluble htt in the cell, as opposed to large nuclear inclusions observed in R6/2^{Q150} transgenic mice. We suggest that the prolonged onset and disease course observed in R6/2 mice with greatly expanded CAG repeats might result from differential upregulation of genes related to protein localization and clearance. Such genes may represent novel therapeutic avenues to decrease htt aggregate toxicity and cell death in HD patients, with *Lrsam1* being a promising, novel candidate disease modifier.

© 2011 Elsevier Inc. All rights reserved.

* to whom correspondence should be addressed: Elizabeth A. Thomas, Ph.D. Department of Molecular Biology, MB-10 The Scripps Research Institute 10550 N. Torrey Pines Rd. La Jolla, CA 92037 858-784-2317; FAX 858-784-2212 bthomas@scripps.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Huntington's disease; striatum; gene expression; inclusion; neuroprotection

Introduction

The presence of a single copy of an expanded CAG repeat region in the Huntington's disease (HD) gene is causative for developing this progressive, neurological disorder (Group 1993), with an inverse relationship existing between CAG length and age of disease onset (Orr and Zoghbi 2007). However, there is enormous variation in the age-at-onset, progression of disease symptoms and post mortem pathology among patients (Rosenblatt et al. 2001; Wexler et al. 2004; Li et al. 2006) and recent studies have suggested that this variation is strongly heritable (Rosenblatt et al. 2001; Djousse et al. 2003; Wexler et al. 2004; Li et al. 2006). Hence, there is widespread belief that other, currently unknown, genes can modify HD pathogenesis.

Similar phenomena are observed in mouse models of HD, whereby different lengths of the CAG repeat affect the onset and progression of the disease phenotype. For example, this has been observed for the original R6/1 vs. R6/2 transgenic mice lines, carrying 115 and 150 CAG repeats, respectively (Mangiarini et al. 1996) and YAC72 vs. YAC128 transgenic mice (Hodgson et al. 1999). However, extended breeding of R6/2 transgenic mice lines exhibiting unstable CAG repeat expansions has challenged this notion; paradoxically, R6/2 mice with highly expanded CAG repeats (i.e. >250) show a prolonged disease phenotype, with increased lifespan, as opposed to the expected decreased lifespan that might occur in humans (Thomas et al. 2008; Dragatsis et al. 2009; Morton et al. 2009). Although the mice still die prematurely of a progressive neurological disease (Thomas et al. 2008; Dragatsis et al. 2009; Morton et al. 2009), the phenotype is improved over that observed in mice with shorter repeats. The mechanism(s) for this slowed disease course remain unclear.

We reasoned that comparison between R6/2 transgenic mice, with an identical integration site for the human HD gene but with varying CAG repeat lengths, might allow us to identify genes contributing or responsible for differences in disease onset and progression in these animals. Therefore, in the present study, we have analyzed microarray data from two R6/2 lines with CAG repeat lengths of ~150 and ~300. We find that, although similar changes of gene expression were observed for downregulated genes, genes showing upregulation were distinct in each R6/2 line. These expression changes were coincident with the presence of abundant extranuclear htt inclusions in the striatum and other brain regions. Further studies on one such gene, *Lrsam1*, revealed neuroprotection against huntingtin (htt)171-82Q striatal cultures, preventing the cell death and reducing the nuclear inclusion accumulation of mutant human htt. These studies suggest that *Lrsam1*, as well as other upregulated genes identified in R6/2Q³⁰⁰ mice, and particularly those associated with protein localization and ubiquitination, may play determinant roles in the time course of the phenotypic expression, based on decreased nuclear htt levels. These studies contribute novel insights into the mitigation of htt aggregation and its ensuing toxicity by revealing genes and pathways associated with ameliorated HD onset and progression that might be targeted for therapeutic benefit.

Materials and Methods

Animals

Two R6/2 lines of the CBA × C57BL/6 strain origin (B6CBA-Tg(HDexon1)62Gpb/1J, Jackson Laboratories, Bar Harbor, ME) (Mangiarini, Sathasivam et al. 1996) have been

maintained at The Scripps Research Institute by breeding male heterozygous R6/2 mice with F1 hybrids of the same background. At the age of 4 weeks, mice were genotyped according to the Jackson Laboratories protocol to determine hemizyosity for the HD transgene. The CAG repeat lengths in these mice were verified by commercial genotyping (Laragen, Los Angeles, CA) and found to be ~ 300 (291-301) in line 1 (designated as R6/2Q³⁰⁰ transgenic mice), and ~150 (140-148) in line 2. R6/2 mice used in the microarray experiments containing ~150 repeats (139-146) were shipped directly from Jackson Laboratories to Massachusetts General Hospital, Boston, MA where they were sacrificed. (Both of these latter lines are designated as R6/2Q¹⁵⁰ transgenic mice).

Microarray Analysis

Microarray data was generated on the Illumina Mouse Mouseref-8 Expression Beadchips v1, querying the expression of ~22,000 Refseq transcripts, as described previously (Thomas et al. 2008), and analyzed using Bioconductor packages (Gentleman et al. 2004). This microarray data was part of a larger dataset (Thomas, Coppola et al. 2008). Quality assessment was performed looking at the inter-array Pearson correlation and clustering based on top variant genes was used to assess overall data coherence. Contrast analysis of differential expression was performed using the LIMMA package. After linear model fitting, a Bayesian estimate of differential expression was calculated. Raw microarray expression data previously generated from wt and R6/2 transgenic mice from the parent line containing 140-150 CAG repeats on the same Illumina platform (Luthi-Carter, Kuhn, unpublished studies) was analyzed by the identical methods. Gene Ontology (GO) terms were determined by The Database for Annotation, Visualization and Integrated Discovery (DAVID).

Real-Time PCR Analysis

Real-time PCR experiments were performed using the ABI PRISMs 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA) as described previously (Desplats et al. 2006). Specific primers for each studied sequence and for mouse and human endogenous controls were designed using Primer Express 1.5 software and their specificity for binding to the desired sequences was searched against NCBI database (Supp. Table 1). Standard curves were generated for each gene of interest using serial dilutions of mouse cDNAs. Experimental samples and no-template controls were all run in triplicate. The PCR cycling parameters were: 50°C for 2 min, 95°C for 10 min, and 40 cycles of 94°C for 15 s, 60°C for 1 min. The amount of cDNA in each sample was calculated using SDS2.1 software by the comparative threshold cycle (Ct) method and expressed as 2^{exp(Ct)} using hypoxanthine guanine phosphoribosyl transferase (Hprt) as an internal control. One-way ANOVA, followed by Student's *t* test to determine exact p-values, was used to determine differences in expression between HD transgenic mice lines and wt littermate controls.

Immunohistochemistry

Immunohistochemical experiments were performed using the anti-Htt antibody, EM48 antibody (1:500 dilution; Chemicon International). The immunoreaction was detected with Vectastain ABC kit (Vector Laboratory Inc., Burlingame, CA) according to the instructions of the manufacturer. Briefly, free floating sections were incubated with Blocking Solution (4% bovine serum albumin in 0.1% Triton X-100/PBS) for 2 hrs at room temperature followed by incubation with primary antibody in Blocking Solution for 16-20 hrs at 4°C. Sections were then washed with 0.1% Triton X-100/PBS, incubated with secondary biotinylated antibody (1:200 dilution in Blocking Solution) for 2 hrs at room temperature, washed with 0.1% Triton X-100/PBS, incubated for 1 hr with ABC reagent (1:1 in Blocking Solution), and then washed finally with 0.1% Triton X-100/PBS. Enzymatic development was performed in 0.05% diaminobenzene in PBS containing 0.003% hydrogen peroxide for 3-5 min.

Lentiviral production

Lentiviral vectors were produced in human embryonic kidney 293T (HEK293T) cells with a four-plasmid system as described previously (Rudinskiy et al., 2009). The viruses were resuspended in PBS with 1% of BSA and the particle content was assessed by p24 antigen ELISA (RETROtek; Gentaur).

Primary striatal cultures

Primary cultures of striatal neurons were prepared as in (Rudinskiy et al., 2009) and infected at DIV1 with lentiviral vectors encoding the first 171 amino acids of mutant (82Q) htt, under the control of a tetracycline-regulated element promoter (SIN-TRE-htt171) containing the Tet-response element (TRE) with seven direct repeats of the TetO operator sequence upstream of a minimal CMV promoter at a p24 concentration of 25 ng/ml. Cells were co-transduced with vectors encoding the tetracycline-controlled transactivator tTA1 under the control of the PGK promoter at 40 ng/ml p24. At DIV4 neurons were infected with a lentiviral vector encoding for Lrsam1 under the control of PGK promoter at levels of 1 ng/ml or 2.5 ng/ml p24. A vector expressing CFP under the control of PGK promoter was used as a control. After 3 weeks in culture cells were fixed in 4% paraformaldehyde for 20 min at RT, washed with PBS and blocked by incubation in PBS containing 5% BSA. Cells were incubated with primary antibodies for 1h at RT (anti-Neuronal Nuclei (NeuN) (Chemicon MAB377 clone A60; 1:500) or anti-huntingtin (Chemicon MAB5492 clone 2B4; 1:500)), washed in PBS and incubated for 1 h at RT with fluorescent secondary antibody (Alexa-Fluor 488, A21121, Invitrogen; 1:1000). Nuclei were stained with Hoechst 33342 dye (Invitrogen, Basel, Switzerland). Images were acquired using LEICA DMI 4000 microscope applying the same exposure conditions for all treatments. Image analysis was performed with ImageJ software (US NIH). Neurons with accumulation of htt in the nucleus were counted as intense 2B4 staining overlaying Hoescht-positive profiles. Bars represent mean values $\pm$ SEM. Differences in NeuN or 2B4-positive cell counts were assessed by Student's t-test, *** $p < 0.001$.

Results

Concordance of Disease-Related Changes in Gene Expression in Two Different Lines of R6/2 Transgenic Mice

We compared gene expression profiles generated from the striatum of R6/2 mice with standard (150) and highly expanded (300) CAG repeats. The integration site of the HD transgene in these two models is identical, hence the only difference is in the length of the CAG repeat region. Microarray expression data was collected on the same Illumina platform and analyzed identically, hence allowing for direct comparisons of the quality and quantity of gene expression differences. Overall, the numbers of differentially expressed genes in both lines were similar ($n=997$ genes for R6/2^{Q150} transgenic mice and $n=1334$ genes for R6/2^{Q300} transgenic mice, at $p < 0.01$), although R6/2^{Q300} transgenic mice had a higher proportion of upregulated genes, with 46.5% of genes showing upregulation in mice with 300 CAGs vs. 40.4% in mice with 150 CAGs. (Complete lists of genes with significantly altered expression in both lines are provided in Supplementary Tables 2 and 3). We first compared the concordance of the top gene expression differences in each line to each other, with respect to the significance and directionality of the expression change. The top 100 genes (based on the highest significance for the expression change) identified in the R6/2^{Q150} transgenic mice were plotted against their corresponding change in R6/2^{Q300} transgenic mice (Figure 1A). Good concordance was observed for the top 100 changes in the R6/2^{Q150} transgenic mice with a majority of the concordant expression changes being decreases. Similarly, the top 100 genes (also based on the highest significance for the expression change) altered in R6/2^{Q300} transgenic mice showed concordance to the

R6/2^{Q150} transgenic mice, although more discordant changes were observed with respect to the upregulated genes (Figure 1B). We next examined scatterplots of the overlapping genes that were differentially expressed due to mutant Htt in both R6/2 lines to look at the magnitude and direction of the expression changes (Figure 2). Excellent concordance with respect to the direction of the expression differences was observed. For example, at $p < 0.001$, none of the genes that was significantly altered in both mouse models showed an opposite direction of the expression difference, whereby at $p < 0.01$, only 4 genes showed change in the opposite direction (Figure 2). In this graph, we again see that most of discordant expression changes are those upregulated in expression in R6/2^{Q300} transgenic mice but decreased in expression in R6/2^{Q150} transgenic mice.

Functional Classification of Dysregulated Genes

We analyzed lists of genes altered in expression in both R6/2 lines for significant Gene Ontology (GO) processes, separating those genes increased in expression versus decreased in expression. Of the genes decreased in expression, there was substantial overlap of GO categories, including synaptic transmission, neurotransmitter signaling, calcium ion/calmodulin binding and metal ion transport (Figure 3). In contrast, the top GO processes, according to the number of genes in the category, associated with the genes upregulated in expression were strikingly distinct, with the exception of chromatin binding, which appeared in both mouse models (Figure 3). The complete list of significant GO categorization for upregulated genes is shown in Table 1. We became particularly interested in the GO process of Establishment of localization/protein localization, which contained 20 genes elevated in expression in R6/2^{Q300} transgenic mice, reasoning that these might have potential relevance for mutant htt altered protein aggregation and localization.

One caveat of the above comparison study was that the R6/2 mice carrying 150 CAG repeats were 6 weeks of age (of a 12 week lifespan), hence at a slightly earlier stage of disease progression than the R6/2^{Q300} transgenic mice, who were 5 months old at the time of sacrifice (lifespan ~7 months). Therefore, we compared the expression of these genes, as well as additional genes related to protein folding and ubiquitin (Tables 1 and 2), to previously published microarray data from R6/2 mice at 12 weeks of age (Kuhn et al. 2007). Although caution must be made in interpreting the magnitude of the gene expression changes as this dataset was generated on a different platform (Affymetrix) and analyzed by different statistical methods, qualitative comparisons can be formulated. Of the 36 selected genes related to these categories, 21 were also elevated in expression in the Kuhn microarray dataset (set 1 from (Kuhn et al. 2007)) whereas 15 were uniquely upregulated in R6/2^{Q300} transgenic mice (Table 2).

Real-Time PCR Validation of Up-Regulated Genes

We validated the expression differences for eight genes showing the greatest expression increases in the R6/2^{Q300} transgenic mice and in an additional colony of R6/2 with ~150 CAG repeats at late-stage (12 weeks of age). These genes include: Endoplasmic reticulum protein 29, (*Erp29*), Leucine rich repeat and sterile alpha motif containing 1, (*Lrsam1*), Nuclear autoantigenic sperm protein (histone-binding), (*Nasp*), RAB9B, member RAS oncogene family, (*Rab9b*), Peptidylprolyl isomerase (cyclophilin)-like 3, (*Ppil3*), Prefoldin 5, (*Pfdn5*), SEC63-like (*S. cerevisiae*), (*Sec63*), and Transporter 1, ATP-binding cassette, sub-family B (MDR/TAP), (*Tap1*). As predicted from the microarray data, none of these genes was elevated in expression in the striatum of R6/2^{Q150} transgenic mice, while expression increases for six of these genes, *Lrsam1*, *Erp29*, *Nasp*, *Rab9b*, *Pfdn5* and *Tap1* were validated in the R6/2^{Q300} transgenic mice (Figure 4).

Nuclear vs. Cytoplasmic Htt Inclusions

We next examined Htt inclusions in late-stage R6/2^{Q300} and R6/2^{Q150} transgenic mice (5 months and 12 weeks, respectively) by immunostaining with the EM48 antibody, which specifically recognizes aggregated forms of mutant htt. In the striata of R6/2^{Q150} transgenic mice, we find abundant, large (>3 μ m) nuclear inclusions in essentially all of the neurons (Figure 5) consistent with previous studies (Davies et al. 1997; Li et al. 2002; Meade et al. 2002). Most of the inclusion mass (>98%) was nuclear in localization. No large inclusions were detected in the cytoplasm, however scattered inclusions, roughly 1/10th the size of the nuclear inclusions, could be visualized. In contrast, the striata of R6/2^{Q300} transgenic mice exhibited much fewer nuclear inclusions (33.3 $\pm$ 4.5% of all inclusions), but abundant cytoplasmic and perinuclear inclusions (66.3 $\pm$ 5.5 % of all aggregates), generally of smaller size (~ 1 μ m), although large cytoplasmic and perinuclear inclusions (i.e. >3 μ m) were observed (Figure 5). This finding is consistent with the decreased size and change in localization reported previously for inclusions in R6/2 transgenic mice with $\geq$ 335 and $\geq$ 280 CAG repeats (Dragatsis et al. 2009; Morton et al. 2009).

Lrsam1 prevents neuronal dysfunction and cell death in vitro

We next tested the effects of Lrsam1 and Erp29, the two genes showing the greatest expression increases in R6/2^{Q300} transgenic mice, in a primary striatal culture model of HD. As previously reported (Rudinskiy et al., 2009) lentiviral-mediated over-expression of a mutated fragment of htt (htt171-82Q) resulted in a 50% reduction of NeuN-positive cells at 3 weeks compared with cultures infected with a vector expressing the wild-type htt fragment (htt171-18Q) (data not shown). Co-expression of Lrsam1 in htt171-82Q striatal cells resulted in a significant prevention of neuronal cell death elicited by mutant htt, based on the number of NeuN-positive cells (Figure 6 and 7A). Erp29 showed no significant effects on neuronal number (data not shown). In light of the altered pattern of nuclear htt inclusions observed in the R6/2^{Q300} mice, we tested whether Lrsam1 and Erp29 could alter the nuclear accumulation in htt171-82Q-infected cultures. We found that the co-expression of htt171-82Q and Lrsam1 decreased both total number of cells with observable nuclear htt accumulation, as well as the number of nuclear inclusions per cell compared to cultures infected with htt171-82Q alone (Figures 7B and C). Erp29 showed no significant effects on htt aggregation.

Discussion

The spontaneous CAG-repeat elongation observed in R6/2 mice bred via the male germ-line has led to the paradoxical discovery that highly expanded CAG repeats (i.e. >250) show a prolonged disease phenotype, with increased lifespan, as opposed to the expected decreased lifespan. This phenomenon has been observed in several different R6/2 colonies (Cowin et al. 2008; Cummings et al. 2008; Thomas et al. 2008; Dragatsis et al. 2009; Morton et al. 2009). The mechanism for this attenuated phenotype is unclear. In the current study, we compared gene expression profiles between R6/2 transgenic mice identical for the human HD transgene (a 1.9 kb genomic fragment containing promoter sequence and exon 1) but with differing CAG repeat lengths, in order to identify genes that may be related to the prolonged phenotype. We find that genes decreased in expression due to the presence of mutant Htt are similar in R6/2 lines carrying standard (~150 CAG) and highly expanded (~300 CAG) repeats, while genes upregulated in their expression are largely distinct to each line. This suggests that those genes upregulated in expression may be important modifiers of disease pathogenesis. In particular, genes found to be upregulated distinctly in R6/2^{Q300} transgenic mice were associated with protein processing, including establishment of protein localization, protein folding and ubiquitination, including *Lrsam1*, *Nasp*, *Rab9b*, *Erp29*, *Tap1* and *Pfdn5*, which we validated by real-time PCR analysis. These pathways, as they

relate to the htt protein levels, have been widely implicated as modifiers of HD pathogenesis (Finkbeiner and Mitra, 2008).

Studies on R6/2 mice with highly expanded repeats have revealed potential therapeutic relevance of the role of nuclear inclusions on HD pathology. In all three published reports (including this one) the patterns of htt aggregation in mice with >250 CAG repeats is dramatically different than that observed in mice with ~150 CAG repeats. This pattern directly correlated with disease onset and lifespan. These findings suggest that there is a length-dependency in the mechanisms underlying formation of nuclear inclusions and support the idea that nuclear inclusion formation is deleterious (Steffan and Thompson, 2003). A previous report has proposed that hindrance of nuclear entry in mice with greater than 300 CAG repeats might be due to the protein product becoming too large for passive nuclear entry (Dragatsis et al. 2009). However, the htt inclusions that are formed do not appear to correspond to the predicted CAG repeat length, as they are no larger than those observed in mice with 150 CAG repeats; rather the cytoplasmic inclusions that we detect in our mice are smaller than those observed in mice with lower CAG repeat sizes (see Figure 5). Moreover, a similar inclusion phenotype also correlates with the rescue of HD neurotoxicity in mutant htt-exposed striatal neurons by overexpression of molecular chaperones (Perrin et al. 2007). Thus, we hypothesize that the increased expression of genes related to protein localization and protein folding affects the processing and destination of mutant htt in the cell. The mechanism by which highly expanded polyglutamine repeats lead to the upregulation of these transcripts is not known; however, it is possible that this reflects a compensatory response by the cell to mobilize and redistribute the toxic protein.

In this study, we further focused on Lrsam1 (a.k.a. Tal), a multifunctional RING finger protein, with E3 ubiquitin ligase activity (Amit et al. 2004), although the exact functional role(s) for this protein are unclear. Our *in vitro* experiments in primary cultured cells demonstrated that the overexpression of Lrsam1 reduces polyQ toxicity. We show that Lrsam1 prevented the loss of primary striatal neurons expressing htt171-82Q and reduced the presence of nuclear htt inclusions. A possible mechanism for this rescue of HD neurotoxicity is the inhibition or reversal of nuclear htt aggregate load. Previous studies have shown that Lrsam1 contributes to the sorting of proteins into cytoplasm-containing vesicles that bud at the multivesicular body and at the plasma membrane (Amit et al. 2004). Hence, increased expression of Lrsam1 might act to shift the localization of mutant htt out of the nucleus and into the cytoplasm, although targeting to specific cytoplasmic organelles could not be delineated in this study. Another gene that was substantially elevated in R6/2^{Q300} transgenic mice is Erp29, which has been hypothesized to act as a chaperone that may facilitate folding and/or export of secretory proteins from the endoplasmic reticulum (Sargsyan et al. 2002). Hence, it may act as a chaperone for htt, facilitating its localization in the cytoplasm, although we did not detect any effects of Erp29 alone on htt nuclear inclusions in our *in vitro* HD model, suggesting that it may require the cooperation of other proteins.

Overall, we have identified a group of genes that are upregulated in expression in R6/2 transgenic mice with 300 CAG repeats, but not in R6/2 mice with 150 CAG repeats at both middle and late stages of illness, many of which are related to protein processing. We hypothesize that increased expression of these genes, in particular Lrsam1, results in a delayed disease onset, latent disease progression and longer lifespan. Increasing the activity of Lrsam1, and potentially other similar genes in humans with a goal of reducing nuclear inclusion load, may have therapeutic relevance in HD patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Thanks to Keith Harshman at the Lausanne DNA Array Facility for facilitating the processing of the R6/2-Q150 CAG microarray samples. These studies were funded, in part, by NIH grant NS44169 to E.A.T.

Abbreviations

HD	Huntington's Disease
HTT	human Huntington's disease gene
Htt	huntingtin
Erp29	Endoplasmic reticulum protein 29
Lrsam1	Leucine rich repeat and sterile alpha motif containing 1
Nasp	Nuclear autoantigenic sperm protein (histone-binding)
Rab9b	RAB9B, member RAS oncogene family
Ppil3	Peptidylprolyl isomerase (cyclophilin)-like 3
Pfdn5	Prefoldin 5
Sec63	SEC63-like (<i>S. cerevisiae</i>)
Tap1	Transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)

References

- Amit I, Yakir L, Katz M, Zwang Y, Marmor MD, Citri A, et al. Tal, a Tsg101-specific E3 ubiquitin ligase, regulates receptor endocytosis and retrovirus budding. *Genes Dev.* 2004; 18(14):1737–52. [PubMed: 15256501]
- Cowin R-M, Graham D, Howard C, Rohena C, Yuva-Paylor L, Monier A, et al. A relationship between polyglutamine expansion length and age of disease onset in R62 mice. *Soc Neurosci Abstr* online. 2008; 114.7
- Cummings DM, YIM MM, ALAGHBAND Y, MALVAR JS, TSURUYAMA K, JOSHI PR, et al. The R6/2 phenotype is cag repeat length-dependent. *Soc Neurosci Abstr* online. 2008 Program number 443.11/V6.
- Davies SW, Turmaine M, Cozens BA, DiFiglia M, Sharp AH, Ross CA, et al. Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell.* 1997; 90(3):537–48. [PubMed: 9267033]
- Desplats PA, Kass KE, Gilmartin T, Stanwood GD, Woodward EL, Head SR, et al. Selective deficits in the expression of striatal-enriched mRNAs in Huntington's disease. *J Neurochem.* 2006; 96(3): 743–57. [PubMed: 16405510]
- Djousse L, Knowlton B, Hayden M, Almqvist EW, Brinkman R, Ross C, et al. Interaction of normal and expanded CAG repeat sizes influences age at onset of Huntington disease. *Am J Med Genet A.* 2003; 119A(3):279–82. [PubMed: 12784292]
- Dragatsis I, Goldowitz D, Del Mar N, Deng YP, Meade CA, Liu L, et al. CAG repeat lengths > or =335 attenuate the phenotype in the R6/2 Huntington's disease transgenic mouse. *Neurobiol Dis.* 2009; 33(3):315–30. [PubMed: 19027857]
- Finkbeiner S, Mitra S. The ubiquitin-proteasome pathway in Huntington's disease. *ScientificWorldJournal.* 2008; 8:421–33. [PubMed: 18454252]
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004; 5(10)

- Huntington's Disease Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell*. 1993; 72(6):971–83. [PubMed: 8458085]
- Hodgson JG, Agopyan N, Gutekunst CA, Leavitt BR, LePiane F, Singaraja R, et al. A YAC mouse model for Huntington's disease with full-length mutant huntingtin, cytoplasmic toxicity, and selective striatal neurodegeneration. *Neuron*. 1999; 23(1):181–92. [PubMed: 10402204]
- Kuhn A, Goldstein DR, Hodges A, Strand AD, Sengstag T, Kooperberg C, et al. Mutant huntingtin's effects on striatal gene expression in mice recapitulate changes observed in human Huntington's disease brain and do not differ with mutant huntingtin length or wild-type huntingtin dosage. *Hum Mol Genet*. 2007
- Li JL, Hayden MR, Warby SC, Durr A, Morrison PJ, Nance M, et al. Genome-wide significance for a modifier of age at neurological onset in Huntington's disease at 6q23-24: the HD MAPS study. *BMC Med Genet*. 2006; 7:71. [PubMed: 16914060]
- Li SH, Cheng AL, Zhou H, Lam S, Rao M, Li H, et al. Interaction of Huntington disease protein with transcriptional activator Sp1. *Mol Cell Biol*. 2002; 22(5):1277–87. [PubMed: 11839795]
- Mangiarini L, Sathasivam K, Seller M, Cozens B, Harper A, Hetherington C, et al. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell*. 1996; 87(3):493–506. [PubMed: 8898202]
- Meade CA, Deng YP, Fusco FR, Del Mar N, Hersch S, Goldowitz D, et al. Cellular localization and development of neuronal intranuclear inclusions in striatal and cortical neurons in R6/2 transgenic mice. *J Comp Neurol*. 2002; 449(3):241–69. [PubMed: 12115678]
- Morton AJ, Glynn D, Leavens W, Zheng Z, Faull RL, Skepper JN, et al. Paradoxical delay in the onset of disease caused by super-long CAG repeat expansions in R6/2 mice. *Neurobiol Dis*. 2009; 33(3):331–41. [PubMed: 19130884]
- Orr HT, Zoghbi HY. Trinucleotide repeat disorders. *Annu Rev Neurosci*. 2007; 30:575–621. [PubMed: 17417937]
- Perrin V, Regulier E, Abbas-Terki T, Hassig R, Brouillet E, Aebischer P, et al. Neuroprotection by Hsp104 and Hsp27 in lentiviral-based rat models of Huntington's disease. *Mol Ther*. 2007; 15(5):903–11. [PubMed: 17375066]
- Rosenblatt A, Brinkman RR, Liang KY, Almqvist EW, Margolis RL, Huang CY, et al. Familial influence on age of onset among siblings with Huntington disease. *Am J Med Genet*. 2001; 105(5):399–403. [PubMed: 11449389]
- Rudinskiy N, Kaneko YA, Beesen AA, Gokce O, Régulier E, Déglon N, Luthi-Carter R. Diminished hippocalcin expression in Huntington's disease brain does not account for increased striatal neuron vulnerability as assessed in primary neurons. *J Neurochem*. 2009; 111(2):460–72. [PubMed: 19686238]
- Sargsyan E, Baryshev M, Backlund M, Sharipo A, Mkrtchian S. Genomic organization and promoter characterization of the gene encoding a putative endoplasmic reticulum chaperone, ERp29. *Gene*. 2002; 285(1-2):127–39. [PubMed: 12039039]
- Thomas EA, Coppola G, Desplats PA, Tang B, Soragni E, Burnett R, et al. The HDAC Inhibitor, 4b, Ameliorates the Disease Phenotype and Transcriptional Abnormalities in Huntington's Disease Transgenic Mice. *Proc Natl Acad Sci U S A*. 2008; 105(40):15564–9. [PubMed: 18829438]
- Wexler NS, Lorimer J, Porter J, Gomez F, Moskowitz C, Shackell E, et al. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proc Natl Acad Sci U S A*. 2004; 101(10):3498–503. [PubMed: 14993615]

A. Concordance to Q300 mice **B. Concordance to Q150 mice** **Figure 1.**

Scatterplots of concordant gene expression changes in R6/2 mice with 150 vs. 300 CAG repeats. Log₂ ratios for the top 100 genes expression changes (according to p-value) in each R6/2 line are shown, with each circle representing an individual gene. Solid circles denote non-significant gene expression changes.

Figure 2. Scatterplots of the overlap of differentially expressed genes due to mutant Htt in R6/2 mice with 150 vs. 300 CAG repeats. The number of genes overlapping at different levels of significance are shown (p<0.001 to p<0.05), with each circle representing an individual gene.

Figure 3. The top Gene Ontology (GO) categories for down-regulated and up-regulated genes in both R6/2 lines of mice. The GO categories, shown along the x-axis, in this graph were prioritized according to the number of genes in a given category (shown on y-axis).

Figure 4.

Real-time PCR validation of upregulated genes related to protein localization in R6/2 mice with ~150 and ~300 CAG repeats. The genes are indicated by their official Unigene ID. Hprt was used to normalize gene expression differences. Asterisks denote significant differences from wt or control subjects as determined by ANOVA *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; +, $p < 0.08$.

Figure 5. Huntingtin aggregates in R6/2^{Q300} and R6/2^{Q150} transgenic mice. EM48-positive aggregates in the striatum of each line of R6/2 mice are shown. Immunohistochemistry was performed on coronal mouse brain sections using EM48 antibody (1:500 dilution). Sections were counterstained with hemotoxylin. Scale bar = 10 μ m.

Figure 6. Neuroprotective effects of Lrsam1 in an *in vitro* model of Huntington's disease. The overexpression of Lrsam1 prevents the loss of NeuN-positive cell bodies in htt171-82Q-infected cultures. Lower panels: co-localization of NeuN staining with Hoechst nuclear dye is indicated by arrowheads. CFP, cyan fluorescent protein.

Figure 7.

Bar graph quantification of the neuroprotective effects of Lrsam1 in an *in vitro* model of Huntington's disease. E16 rat primary striatal cultures were co-infected with lentiviral vectors encoding wild-type (htt171 with 18Q repeats) or mutated huntingtin (htt) fragments (htt171 with 82Q repeats) and with a vector encoding Lrsam1. The overexpression of Lrsam1 prevents the loss of NeuN-positive cell bodies in htt171-82Q-expressing cultures (a), concomitant with a reduction of total cells with htt nuclear inclusions (b), and 2B4 positive-aggregates. (+) Htt nuclear inclusions/cell (c). Bars represent mean values $\pm$ SEM. Differences in NeuN or 2B4-positive cell counts were assessed by Student's t-test, *** $p < 0.001$.

Table 1

GO categories associated with upregulated genes in R6/2 mice with 150 and 300 CAG repeats. Determined using The Database for Annotation, Visualization and Integrated Discovery.

R6/2 Q150 Mice:	P:	#:	Genes:	R6/2 Q300 Mice:	P:	#:	Genes:
mitochondrial membrane	0.0002	14	NIPSNAP1, 1810027010RIK, ATP5G1, ATAD3A, 2610024G14RIK, MAGMAS, GPX4, SLC25A12, NDUFA11, MFN1, ATP5H, NDUFB3, NDUFA3, UQCRLH	ubiquitin ligase complex	0.0016	5	CBX8, ANAPC2, UBE2L3, CUL4A, ITCH
coenzyme biosynthetic process	0.0040	5	ATP5G1, GGT1L3, ATP6API, ATP5G2, ATP5H	ubiquitin-dependent protein catabolic process	0.0099	9	USP18, PSMB10, ATE1, PSMA2, ANAPC2, UBE2D2, UBE2L3, CUL4A, ITCH
nucleosome assembly	0.0043	6	HIST1H2AN, HIST1H2AO, HIST1H2AI, HIST1H2AF, HIST1H2AH, HIST2H2AC	nucleoside-triphosphatase activity	0.0104	16	ABCA7, RFC3, KIF1B, RHOJ, GNG4, GBP4, GSPT2, GNG3, ABCB4, NUBP1, PSMC4, FIGNL1, IGTTP, TAP1, EHD4, ATP6V1G2
hydrogen ion transporter activity	0.0097	5	ATP5G1, ATP6API, ATP5G2, ATP5H, UQCRLH	pyrophosphatase activity	0.0148	9	ABCA7, RFC3, ABCB4, NUBP1, PSMC4, FIGNL1, TAP1, ENP1, DOLPP1
mitochondrion	0.0172	28	NIPSNAP1, BCKDK, MRPL14, 2010100012RIK, ATP5G1, MRPS27, MRPS34, ATAD3A, ATP5G2, MRPL23, GFER, SPG7, RTN4IP1, PKM2, MAGMAS, GPX4, SLC25A12, NDUFA11, FXCL, MFN1, ATP5H, NDUFB3, MRPL2, MRPS5, NDUFA3, COQ7, UQCRLH	transcription	0.0284	56	WT1, IRX5, NKX2-3, SFMBT2, HOXC6, NKX6-1, CBX8, POLR3K, HDAC1, DLX2, ZFP30, STAT1, TAF11, HOXD8, PARP14, ZFP90, VGLL2, MYT1, NR1H3, STAT6, PSIP1, TCEAL5, SOX11, CREB3, PRIM2, POLR3A, ZFP202, SHOX2, SMARCA1, PTRF, DLX1, HES6, HIRA, TEAD2, GSC, ZBTB3, USF1, FOXI1, PHF6, KLF3, CGGBP1, ZFP27, MORF4L2, ZFP39, ZFP28, ZFP64, ZBTB8, SP4, KEAP1, IRF2, ZFP369, FOXF1A, TFB1M, GTF2A1, TCEA2, BCL11A
magnesium ion binding	0.0208	13	ATP9A, 1810034K20RIK, NME7, MAST1, GTF2F1, CDC42BPB, RPO1-4, ADPGK, PKM2, PRPSAPI, POLK, INPP1	homophilic cell adhesion	0.0223	6	PCDH12, PCDH20, PCDH10, PCDHB3, CHL1, PVRL1
lipid binding	0.0332	7	ATP5G1, CDC42BPB, ATP5G2, FABP3, BPIL2, VCP, CADPS2	protein folding	0.0380	7	FKBP14, SEC63, FKBP11, CCT6A, CCS, PPIL3, PFDN5
cell adhesion	0.0377	4	PCDH10, PCDH20, CDH15, CELSR3	chromatin binding	0.0407	6	CBX8, HDAC1, DLX2, SMARCA1, DLX1, MSH6

R6/2 Q150 Mice:									
GO Category:	P:	#:	Genes:	R6/2 Q300 Mice:	GO Category:	P:	#:	Genes:	
ribosome	0.0422	8	MRPL14, MRPS27, RPL13A, MRPS34, MRPL23, RPL39, MRPL2, MRPS5	ATPase activity	0.0412	5	ABCA7, KIF1B, ABCB4, TAP1, ATP6V1G2		
chromatin	0.0468	7	HIST1H2AN, HIST1H2AO, HIST1H2AI, HIST1H2AF, HIST1H2AH, HIST2H2AC, SATB1	establishment of protein localization	0.0414	20	APBP2, RAB9B, ARFGAP3, TRAM1, RAB39, COG8, SNX4, IPO8, NASP, LRSAM1, SEC63, TIMM10, NUP93, ERP29, NUP133, SNX11, TAP1, RAB21, EXOC8, ZFYVE20		
protein folding	0.0488	5	PFDN2, CCS, PPIL3, PFDN5, TBCA	response to DNA damage stimulus	0.0481	10	BLM, TOPORS, XPA, RAD9B, MSH6, NUDT1, RAD18, CUL4A, FEN1, 6620401K05RIK		
response to stress	0.0490	10	CIRBP, RBM3, PRKDC, PSMB5, ERCC5, GPX4, ERCC3, STIP1, POLK	GTPase activity	0.0512	7	RHOJ, GNG4, GBP4, GSPT2, GNG3, IGTP, EHD4		
DNA repair	0.0500	6	POLD1, PRKDC, ERCC5, ERCC3, POLK, 6620401K05RIK	DNA repair	0.0522	9	BLM, XPA, RAD9B, MSH6, NUDT1, RAD18, CUL4A, FEN1, 6620401K05RIK		
protein metabolic process	0.0517	6	PREP, ATE1, SPG7, RNF25, 4933406E20RIK, COQ7	translation factor activity, nucleic acid binding	0.0522	6	GSPT2, TCEAL5, TCEAL3, EIF1A, BRF2, TCEA2		

Table 2
Genes related to protein localization, folding and ubiquitination in different R6/2 lines.

Gene ID:	R6/2X-30 wks:		R6/2-6 wks:		R6/2-12 wks*:	
	Log2 ratio	p-value	Log2 ratio	p-value	Log2 ratio	p-value
ANAPC2	0.260	0.00224	0.130	0.15062	0.276	0.00253
APBP2	0.353	0.00005	0.080	0.35809	0.314	0.03512
AREGAP3	0.357	0.00010	0.050	0.64362	0.372	0.00965
ATE1	0.314	0.00014	0.330	0.00307	0.580	0.00091
CBX8	0.512	0.00001	0.050	0.59109	0.245	0.17841
CCS	0.295	0.00467	0.450	0.00002	0.299	0.01981
CCT6A	0.294	0.00410	0.170	0.06977	0.311	0.01225
COG8	0.401	0.00089	0.120	0.18293	0.319	0.01779
CUL4A	0.412	0.00793	-0.150	0.10469	-0.398	0.01912
ERP29	0.246	0.00340	0.110	0.21404	0.152	0.08300
EXOC8	0.201	0.00752	0.080	0.37959	0.196	0.08663
FKBP11	0.399	0.00311	-0.400	0.12667	0.123	0.28530
FKBP14	0.572	0.00000	0.230	0.01865	0.830	0.00117
IPO8	0.281	0.00158	0.130	0.15101	0.364	0.00510
ITCH	0.285	0.00803	0.150	0.14596	0.300	0.04116
LRSAM1	0.468	0.00213	-0.150	0.11763	0.191	0.07952
NASP	0.517	0.00189	0.160	0.09403	-0.147	0.11277
NUP133	0.258	0.00382	0.120	0.17671	0.421	0.00068
NUP93	0.426	0.00315	0.480	0.00024	0.627	0.00341
PFDN5	0.299	0.00548	0.230	0.01586	0.181	0.07470
PPIL3	0.494	0.00518	0.370	0.00101	0.421	0.08765
PSMA2	0.205	0.00032	0.070	0.47498	0.269	0.02283
PSMB10	0.595	0.00002	0.330	0.00248	0.435	0.00292
RAB21	0.334	0.00585	0.220	0.01370	0.428	0.01444
RAB39	0.443	0.00087	0.140	0.11937	0.318	0.00563
RAB9B	0.408	0.00005	0.080	0.42926	0.463	0.05677
SEC63	0.281	0.00214	0.030	0.73788	0.155	0.12045
SNX11	0.331	0.00458	0.210	0.04380	0.042	0.58907

Gene ID:	R6/2X-30 wks:		R6/2-6 wks:		R6/2-12 wks*:	
	Log2 ratio	p-value	Log2 ratio	p-value	Log2 ratio	p-value
SNX4	0.276	0.00099	NA	0.27198	0.392	0.00820
TAP1	0.333	0.00551	0.110	0.17281	0.142	0.17281
TIMM10	0.192	0.00295	0.180	0.04401	0.059	0.46005
TRAM1	0.338	0.00047	0.130	0.12538	0.541	0.01082
UBE2D2	0.229	0.00320	0.090	0.33556	0.248	0.01229
UBE2L3	0.254	0.00666	0.170	0.05702	0.300	0.00134
USP18	1.108	0.00001	0.030	0.77033	0.766	0.01461
ZFYVE20	0.230	0.00960	0.120	0.19124	0.207	0.03599

* From Kuhn et al., 2008.

The Log2 ratios are color-coded according to magnitude of change. Red indicates increased expression; green indicated decreased expression; gray indicates a non-significant change in gene expression.