Supporting Information

Sathasivam et al. 10.1073/pnas.1221891110

SI Materials and Methods

Mouse Maintenance and Breeding, Genotyping, and CAG Repeat Sizing. *Hdh*$^{Q150/Q150}$ homozygous, *Hdh*$^{+/Q150}$ heterozygous mice and WT littermates on a (CBA × C57BL/6) F1 background were obtained by intercrossing *Hdh*$^{+/Q150}$ heterozygous CBA/Ca and C57BL/6J congenic lines as described previously (1). The *Hdh*Q50, *Hdh*Q100 lines were generated by selective breeding for alterations in germ-line repeat size starting with a C57BL/6 congenic of the *Hdh*Q150 lines (2). The *Hdh*$^{+/Q20}$, *Hdh*$^{+/Q80}$ (3, 4), and zQ175 (5) knock-in mice were supplied from CHDI colonies maintained at The Jackson Laboratory. The *Hdh*Q20, *Hdh*Q50, *Hdh*Q80, *Hdh*Q100, and zQ175 lines were maintained by backcrossing to C57BL/6J (Charles River), and homozygotes, heterozygotes, and WT littermates were generated by intercrossing as required. All experimental procedures were approved by the King's College London Ethical Review Committee and performed in accordance with United Kingdom Home Office regulations. All animals had unlimited access to food and water, were subject to a 12-h light/dark cycle and housing conditions; environmental enrichment were as previously described (6). Genomic DNA was isolated from an ear-punch. *Hdh*Q50, *Hdh*Q100, and *Hdh*Q150 mice were genotyped by PCR; CAG repeat length was measured as previously described (7). The *Hdh*Q20, *Hdh*Q80 mice were genotyped as described (4) using the Hotstart polymerase (Thermo Scientific). The genotyping primers for zQ175 were as in Menalled et al. (5) using the R6/2 genotyping protocol (7). Mouse CAG repeat sizes are summarized in Table S2. Dissected tissues were snap frozen in liquid nitrogen and stored at −80 °C until further analysis.

Mouse RT-PCR, Quantitative RT-PCR, and 3′RACE. RNA, RT-PCR, and quantitative RT-PCR were as described (8), except that RNA was reverse transcribed from an oligo-dT primer and quantitative RT-PCR was performed using the SsoFast Probes Supermix (Bio-Rad) with a corresponding cycler program. 3′RACE was performed as described (9). Bands were excised from gels, cloned (TA cloning kit, Invitrogen), and sequenced (Big Dye Terminator 3.1, ABI) using ABI3730xl DNA analyzer. Primer and probe sequences are detailed in Table S3.

RNA Sequencing. Frozen tissues were homogenized with VWR PowerMax AHS 200 in TRIzol Reagent (Invitrogen). RNA was extracted according to the TRIzol protocol and purified with RNeasy columns (Qiagen). Samples were prepared using a modified strand-specific version of the Illumina Tru-Seq protocol. Illumina's protocol was followed except strand-specific cDNA synthesis steps that were adapted from ref. (10), with one exception for 6-mo samples, which did not include actinomycin in first-strand cDNA synthesis. In addition, the Agencourt Ampure XP system was used to remove dNTPs between first- and second-strand synthesis. Following second-strand cDNA synthesis, samples were run on Beckman Coulter Nucleic Acid Extractor SPRI-te and digested with USER mix (New England Biolabs). Final PCR amplification was performed with KAPA HiFi polymerase and GC buffer (Kapa Biosystems). For 22-mo samples, PCR enrichment included the additive betaine to improve read coverage in the GC-rich regions of the genome. These paired-end, strand-specific cDNA libraries were multiplexed onto the Illumina HiSEq (40-bp reads). Read data were mapped to the mm9 build with the Bowtie alignment program using the "best" setting (11). Differential expression of the Htt gene was analyzed with the R DESeq package (12). Gene expression was calculated as reads per kilobase of exon per million mapped reads (RPKMs). The quality control statistics for the sequencing libraries are presented in Table S4. Splicing was analyzed using the Python/C version of MISO (13). A custom General Feature Format (GFF) file was created for the two Htt isoforms. Coordinates for the short and long isoforms, respectively: chr5(35104760-35105959) and chr5 (35251495-35255170). RPKM tracks and Psi plots were created with Sashimi-plot, part of the MISO framework. MISO is available at: http://genes.mit.edu/burgelab/miso/; Sashimi-plot is available at: http://genes.mit.edu/burgelab/miso/docs/sashimi.html.

Human 3′RACE and RT-PCR. RNA from human samples was extracted as previously described (8). A total of 2 μg total RNA was reverse transcribed (Invitrogen, Moloney murine leukemia virus) using the UAPdt18 primer. After the RT reaction, the mix was digested with 1U of RNase H (Invitrogen) for 1 h at 37 °C. The cDNA was subsequently diluted 1:10 in water and 2 μL were used as template for the 3′RACE or RT-PCRs. All PCRs were carried out using the Promega GoTaq system. Each PCR contained 5 μL of 5 × Green Flexi Buffer, 2 μL 25 mM $MgCl_2$, 0.5 μL 10 mM dNTPs, each 0.5 μL of 10 mM primers, 2 μL cDNA template, 0.125 μL GoTaq polymerase, and water to 25 μL. PCR protocols for human 3′RACE were as follows: first 3′RACE PCR: 1 cycle 94 °C for 2 min; 10 cycles 94 °C for 15 s, 60 °C for 25 s, 72 °C for 2 min; 30 cycles 94 °C for 15 s, 61 °C for 20 s, 72 °C for 1 min 45 s; 1 cycle 72 °C for 6 min followed by cooling to 10 °C. Primers were UAPnest and 7128f for Fig. 2*F*, UAPnest and 2181f for Fig. S5 and UAPnest and 6568f for Fig. 4*D*. Second 3′RACE PCR: 1 cycle 94 °C for 2 min; 35 cycles 94 °C for 15 s, 62 °C for 20 s, 72 °C for 1 min; 1 cycle 72 °C for 6 min followed by cooling to 15 °C. Primers were UAPnest and 7169f for Fig. 2*F*, UAPnest and 2357f for Fig. S5, and UAPnest and 6621f for Fig. 4*D*. Third 3′RACE PCR: 1 cycle 94 °C for 2 min; 35 cycles 94 °C for 15 s, 62 °C for 20 s, 72 °C for 20 s; 1 cycle 72 °C for 6 min followed by cooling to 15 °C. Primers were UAPnest and 7128f for Fig. 4*D*. Bands were excised from gels, cloned (TOPO-TA cloning kit, Invitrogen) and sequenced (Big Dye Terminator 3.1, ABI) using ABI3730xl DNA Analyzer. PCR protocol for human RT-PCRs was as follows: 1 cycle 94 °C for 4 min; 44 cycles 94 °C for 15 s, 59 °C for 20 s, 72 °C for 20 s; 1 cycle 72 °C for 6 min followed by cooling to 15 °C. Primer sequences are detailed in Table S5.

Polysome Gradients. The 10–40% (wt/vol) sucrose stock solutions were prepared in 50 mM Tris-Cl (pH 6.6), 140 mM NaCl, and 12 mM magnesium chloride. Immediately before use, cycloheximide (200 μg/mL) and 1 mM DTT were added. Sucrose gradients were prepared as discontinuous gradients of 2-mL layers of 40%, 32.5%, 25%, 17.5%, and 10%. Starting with 40% sucrose, each layer was frozen on dry ice before the next layer was put on top. The gradient was allowed to thaw overnight at 4 °C whereby a continuous gradient was created by diffusion. Mouse brain tissue was lysed in freshly prepared polysome buffer [10 mM Tris-Cl (pH 7.4), 140 mM NaCl, 12 mM magnesium chloride, 1% (wt/vol) Triton X-100, 1 mM DTT, 200 μg/mL cycloheximide, 0.5 U/μL RNAsin, and 10 mM ribonucleoside vanadyl complex]. Lysates were used immediately and never frozen. Samples were centrifuged twice at 13,000 × *g* at 4 °C for 5 min, and each time the supernatant was transferred to a new tube. A volume corresponding to 250 μg absorbance at 260 nm was layered on the 10–40% sucrose gradients. The gradients were centrifuged at 115,000–260,000 × *g* at 4 °C for 1 h 40 min in a SW41-Ti swing

out rotor. Fractions (18 × 570 μL) were collected and 300 μL of each fraction were extracted twice with 800 μL of a 1:1 mixture of phenol (equilibrated in 0.15 M sodium acetate pH 5.3) and chloroform/iso-amyl alcohol (49:1). For each extraction, samples were rigorously mixed, centrifuged at 13,000 × *g* at RT for 2 min and the supernatant was transferred to a new tube. RNA was precipitated overnight at −20 °C with a 1:1 mixture of ethanol and isopropanol (two times the volume of the sample) and 3 M sodium acetate pH 5.3 (one-sixth the volume of the sample). Samples were centrifuged at 13,000 × *g* at 4 °C for 1 h, washed with 0.5 mL 70% (vol/vol) ethanol, dried, and resuspended in an equal volume of water. An equal volume of each sample was reverse transcribed using random hexamers. The cDNA was diluted 1:5 with water before quantitative RT-PCR analysis. For gel visualization, RNA was mixed with 2 times the volume of loading buffer [85% (vol/vol) formamide, 10% (vol/vol) glycerol, 8.5 mM Tris-Cl pH 7.4, 0.004% (wt/vol) bromophenol blue], denatured for 5 min at 65 °C and analyzed on a 1.3% (wt/vol) agarose in 1xTAE gel (40 mM Tris-acetate, 1 mM EDTA, pH ~8.3) with 5 V/cm.

Antibodies, Immunoprecipitation, and Western Blotting. 3B5H10 is a monoclonal antibody that was raised against an N-terminal 171 amino acid fragment of HTT with 65Q and detects a polyQ tract (14) (Sigma), S830 is a sheep polyclonal antibody raised against exon 1 HTT with 53Q (15), MW8 is a monoclonal raised against the peptide AEEPLHRP at the C terminus of exon 1 HTT (16), and 1H6 is a monoclonal antibody that recognizes SLRNSPEFQKLLGI (17). Six milligrams of epoxy-activated magnetic beads (Dynabeads M-270 Epoxy; Invitrogen) were washed four times with 0.5 mL PBS. The beads were finally resuspended in 100 μL PBS and mixed with 100 μL of 3B5H10 antibody (1 mg/mL). Slowly and under constant mixing, 100 μL of 3 M ammonium sulfate (in 0.1 M sodium phosphate pH 7.4) were added. The tube was sealed and incubated at 30 °C overnight with constant motion. Beads were washed twice for 1 h with 0.5 mL of Tris-Cl pH 8.8. Following this, beads were washed two times with 0.5 mL PBS, two times with 0.5 mL PBS/0.5% Triton X-100, and finally resuspended in 400 μL PBS (supplemented with 0.2 mg/mL BSA and 0.02% sodium azide). Immunoprecipitation, Western blotting, and immunoprobing were performed as previously described (18).

SRSF6 RNA-IP. Mouse brain tissue was lysed in freshly prepared Triton buffer [50 mM Hepes/NaOH pH 7.6, 160 mM NaCl, 7 mM magnesium chloride, 3 mM calcium chloride, 5 mM potassium chloride, 1% (wt/vol) Triton X-100, 1 mM phenylmethylsulfonyl fluoride, 0.5 U/μL RNAsin and complete protease and phosphatase inhibitor mixture]. Lysates were used immediately and never frozen. Samples were centrifuged twice at 13,000 × *g* at 4 °C for 5 min, and each time the supernatant was transferred to a new tube. Supernatant corresponding to 2 mg total protein was immunoprecipitated for 5 h on a rotating wheel at 4 °C. Each reaction contained 9 μL of protein G Dynabeads (prewashed for 1 h at 4 °C in Triton buffer with 1 mg/mL BSA), 3 μg of anti-SRSF6 antibody (LS-B5712; LifeSpan BioSciences), and Triton buffer to a final volume of 400 μL. Following immunoprecipitation, the magnetic beads were washed four times with 0.5 mL Triton buffer, and finally resuspended in 300 μL of AE buffer (50 mM sodium acetate pH 5.3, 10 mM EDTA pH 8.0). RNA was extracted by adding 300 μL of phenol (equilibrated in 0.15 M sodium acetate pH 5.3) and 100 μL of chloroform/iso-amyl alcohol (49:1). Samples were rigorously mixed, centrifuged at 13,000 × *g* at room temperature for 2 min and the supernatant was transferred to a new tube. RNA was precipitated overnight at −20 °C with a 1:1 mixture of ethanol and isopropanol (2 times the volume of the sample), 3 M sodium acetate pH 5.3 (one-sixth the volume of the sample), and 40 μg of glycogen. Samples were centrifuged at 13,000 × *g* at 4 °C for 1 h, washed with 0.5 mL 70% (vol/vol) ethanol, dried, and resuspended in water. An equal volume of each sample was reverse transcribed using random hexamers. The cDNA was diluted 1:5 with water before quantitative RT-PCR analysis.

Bioinformatics. To predict splice factor binding sites, the following websites were used: RegRNA (http://regrna.mbc.nctu.edu.tw/index.php) (19, 20) and ESEfinder 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi) (21). To predict human polyadenylation sites, the intron 1 sequence of the human *HTT* gene was analyzed with the Softberry POLYAH algorithm (http://linux1.softberry.com/all.htm).

Statistics. Statistical significance was calculated by one-way ANOVA with the Bonferroni post hoc test (Fig. 2 *C*–*E*) or homoscedastic two-tailed Student *t* test (Fig. 4*C*). Differential expression of the *Htt* gene was analyzed with the R DESeq package (Fig. S3*B*) (12). DESeq employs the Benjamini-Hochberg procedure to adjust for multiple hypothesis testing. We set our false discovery rate at 10%, accepting adjusted *P* values less than 0.1 as significant.

1. Woodman B, et al. (2007) The Hdh(Q150/Q150) knock-in mouse model of HD and the R6/2 exon 1 model develop comparable and widespread molecular phenotypes. *Brain Res Bull* 72(2-3):83–97.
2. Lin CH, et al. (2001) Neurological abnormalities in a knock-in mouse model of Huntington's disease. *Hum Mol Genet* 10(2):137–144.
3. Wheeler VC, et al. (1999) Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. *Hum Mol Genet* 8(1):115–122.
4. White JK, et al. (1997) Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nat Genet* 17(4):404–410.
5. Menalled LB, Sison JD, Dragatsis I, Zeitlin S, Chesselet MF (2003) Time course of early motor and neuropathological anomalies in a knock-in mouse model of Huntington's disease with 140 CAG repeats. *J Comp Neurol* 465(1):11–26.
6. Hockly E, Woodman B, Mahal A, Lewis CM, Bates G (2003) Standardization and statistical approaches to therapeutic trials in the R6/2 mouse. *Brain Res Bull* 61(5):469–479.
7. Sathasivam K, et al. (2010) Identical oligomeric and fibrillar structures captured from the brains of R6/2 and knock-in mouse models of Huntington's disease. *Hum Mol Genet* 19(1):65–78.
8. Benn CL, Fox H, Bates GP (2008) Optimisation of region-specific reference gene selection and relative gene expression analysis methods for pre-clinical trials of Huntington's disease. *Mol Neurodegener* 3:17.
9. Scotto-Lavino E, Du G, Frohman MA (2006) 3′ end cDNA amplification using classic RACE. *Nat Protoc* 1(6):2742–2745.
10. Levin JZ, et al. (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat Methods* 7(9):709–715.
11. Langmead B (2010) Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* Chapter 11:Unit 11.7.
12. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
13. Katz Y, Wang ET, Airoldi EM, Burge CB (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* 7(12):1009–1015.
14. Peters-Libeu C, et al. (2005) Crystallization and diffraction properties of the Fab fragment of 3B5H10, an antibody specific for disease-causing polyglutamine stretches. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61(Pt 12):1065–1068.
15. Sathasivam K, et al. (2001) Centrosome disorganization in fibroblast cultures derived from R6/2 Huntington's disease (HD) transgenic mice and HD patients. *Hum Mol Genet* 10(21):2425–2435.
16. Ko J, Ou S, Patterson PH (2001) New anti-huntingtin monoclonal antibodies: Implications for huntingtin conformation and its binding proteins. *Brain Res Bull* 56(3-4):319–329.
17. Lunkes A, et al. (2002) Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions. *Mol Cell* 10(2):259–269.
18. Landles C, et al. (2010) Proteolysis of mutant huntingtin produces an exon 1 fragment that accumulates as an aggregated protein in neuronal nuclei in Huntington disease. *J Biol Chem* 285(12):8808–8823.
19. Huang HY, Chien CH, Jen KH, Huang HD (2006) RegRNA: An integrated web server for identifying regulatory RNA motifs and elements. *Nucleic Acids Res* 34(Web Server issue):W429–W434.
20. Smith PJ, et al. (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet* 15(16):2490–2508.
21. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 31(13):3568–3571.

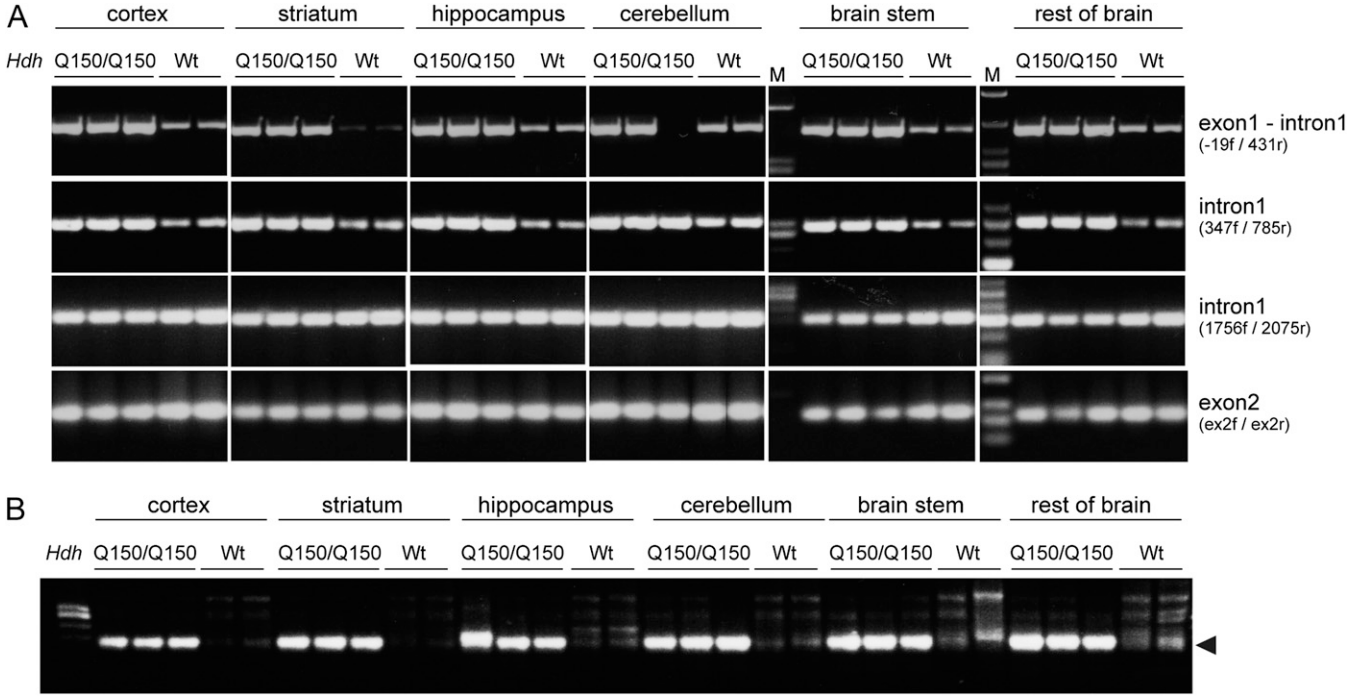


Fig. S1. Aberrant splicing of *Htt* exon 1 to exon 2 occurs in all brain regions. (*A*) RT-PCR analysis and (*B*) 3′RACE (◀) of different brain regions of *Hdh*Q150 knock-in mice at 2 mo of age. M, HaeIII-digested ΦX174.

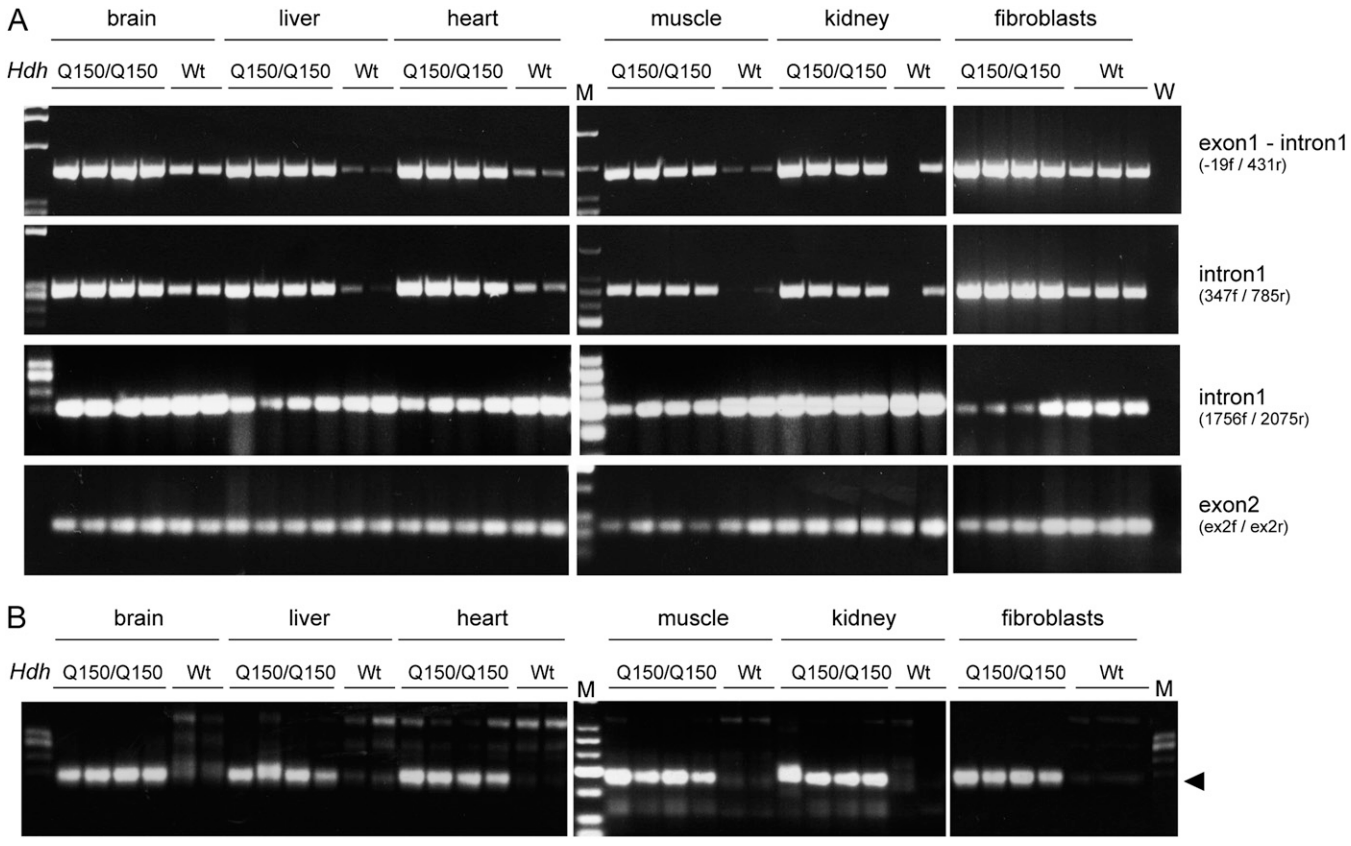


Fig. S2. Aberrant splicing of *Htt* exon 1 to exon 2 occurs in peripheral tissues. (*A*) RT-PCR analysis and (*B*) 3′RACE (◀) of brain and peripheral tissues from *Hdh*Q150 knock-in mice at 2 mo of age. W, water; M, HaeIII-digested ΦX174.

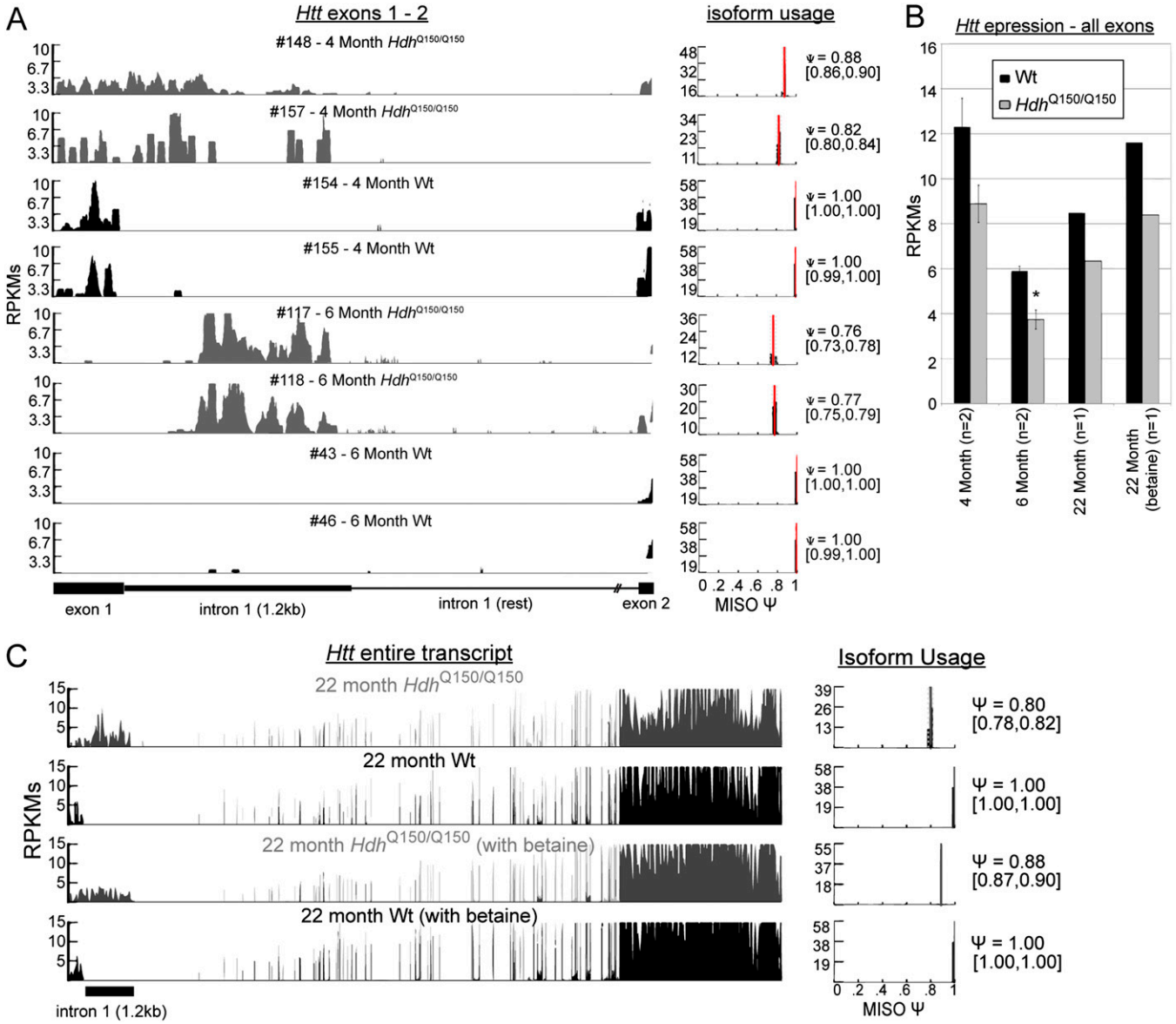
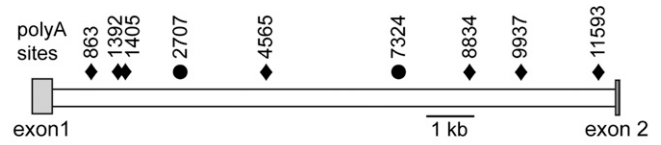


Fig. S3. RNA-Seq analysis of *Hdh*Q150 knock-in brain RNA. (*A*) Read densities in the *Htt* exon 1–exon 2 region from 4- and 6-mo-old $Hdh^{Q150/Q150}$ and WT mice and Mixture-of-Isoforms software percent spliced in (Psi) values. These libraries were prepared without betaine; coverage in exon 1 is reflective of sequencing depth of the libraries (*B*) *Htt* expression (all exons) from 4-, 6-, and 22-mo samples. $*P = 0.0027$. (*C*) RNA-Seq reads throughout the entire *Htt* transcript indicated that introns 17 and 49 were partially retained in both genotypes. Psi plots predict that normal splicing occurs in 88% and 80% of $Hdh^{Q150/Q150}$ transcripts.



> test sequence
Length of sequence - 11850
9 potential polyA sites were predicted

Pos.:	863	LDF-	3.54	sequence	GG**AATAAA**AC
Pos.:	1392	LDF-	3.09	sequence	GG**AATAAA**AT
Pos.:	1405	LDF-	2.68	sequence	GG**AATAAA**TT
Pos.:	2707	LDF-	5.30	sequence	TA**AATAAA**AT
Pos.:	4565	LDF-	1.31	sequence	GG**AATAAA**GA
Pos.:	7324	LDF-	6.73	sequence	AA**AATAAA**AT
Pos.:	8834	LDF-	1.15	sequence	GT**AATAAA**AA
Pos.:	9937	LDF-	0.76	sequence	AA**AATAAA**TA
Pos.:	11593	LDF-	1.47	sequence	AA**AATAAA**GT

Fig. S4. Prediction of cryptic polyadenylation signals in the human intron 1 of *HTT* using the SoftBerry POLYAH algorithm (http://linux1.softberry.com/all.htm). The algorithm predicts the potential position of a poly-A signal by linear discriminant functions combining characteristics describing various contextual features of these sites. The default LDF threshold ("weight" of predicted sites) in the server is equal to 0. ♦, sites that were predicted; ●, sites that were used in YAC128 and BACHD mice and in human Huntington disease tissues.

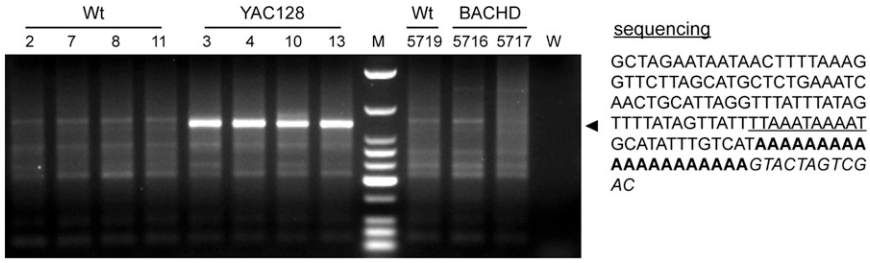


Fig. S5. 3′RACE analysis of *HTT* in YAC and BAC HD mice. This 3′RACE product was generated only from YAC128 brain RNA (◀) but not from WT controls or BACHD mice. It contained a polyA tail 2,730 bp into intron 1. The cryptic polyadenylation site in *HTT* intron 1 is underlined, the polyA tail is in bold, and the primer sequence (UAPdT18) is shown in italics. M, low-molecular-weight marker (New England Biolabs); W, water.

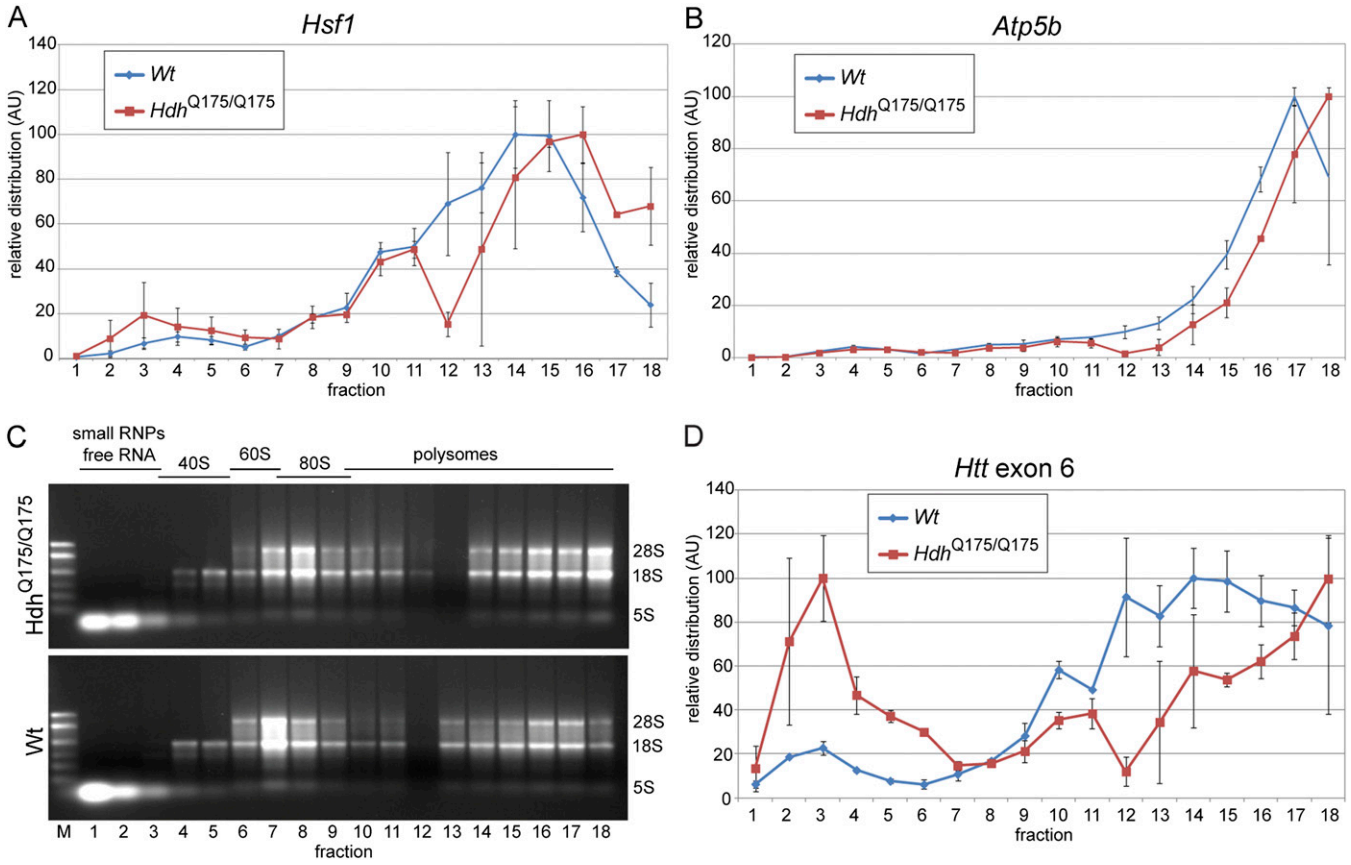


Fig. S6. Quantitative RT-PCR analysis of polysome gradients in zQ175 and WT brains at 2 mo of age showing the relative distribution of (*A*) *Hsf1* and (*B*) *Atp5b*. (*C*) RNA from gradient fractions was analyzed on a TAE/agarose gel. M, RNA ladder 6000 (Agilent). (*D*) *Htt* exon 6 transcripts. Primers and probes for *Hsf1* and *Htt* exon 6 are detailed in Table S3 and the *Atp5b* assay was from PrimerDesign. Data are shown as mean ± SEM ($n = 2$).

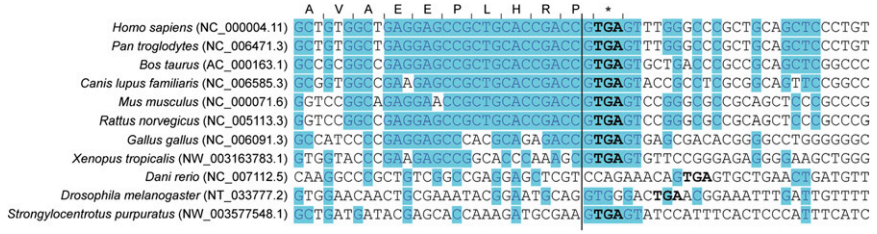


Fig. S7. Multiple sequence alignment of huntingtin exon 1–intron 1 junctions (± 29 bp). All sequences are deposited in the National Center for Biotechnology Information database; the accession number is given in brackets. Invitrogen VNTI AlignX module was used to create the multiple sequence alignment (multiple pairwise alignments with manually removed gaps). Blue background color highlights conserved nucleotides. The black vertical line is the end of exon 1. The human amino acid sequence is given as an example above the alignment. For every species analyzed exon 1 is followed by an in-frame stop codon (shown in bold).

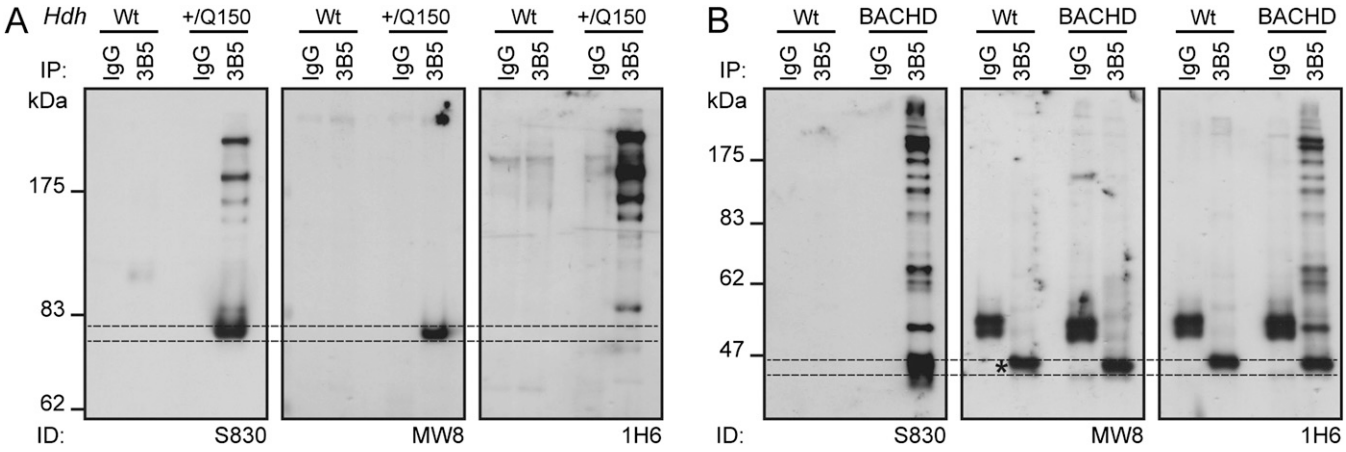


Fig. S8. Detection of HTT exon 1 in HD mouse models. (*A* and *B*) HTT proteins were immunoprecipitated with 3B5H10-coupled magnetic beads from WT and (*A*) *Hdh*Q150 and (*B*) BACHD brain lysates, and Western immunoblots were immunoprobed with S830, MW8, and 1H6 antibodies. Comparison of the S830 and MW8 blots reveals that an exon 1 protein is present in the *Hdh*Q150 brains but not in those from their WT littermates or in the IgG controls (dotted lines). On the BACHD blots, the presence of a protein immunoprecipitated with 3B5H10 and detected with the anti-mouse secondary antibody (for MW8 and 1H6) prevented detection of the predicted HTT exon 1 product (*). 3B5 = 3B5H10.

Table S1. CAG repeat size for the knock-in mouse lines

Knock-in mouse lines	CAG repeat size ± SD
*Hdh*Q20	17 ± 0.46
*Hdh*Q50	59 ± 0.62
*Hdh*Q80	81 ± 0.72
*Hdh*Q100	110 ± 1.34
*Hdh*Q150	167 ± 8.18
zQ175	189 ± 7.85

Table S2. CAG repeat lengths of human postmortem brains and human fibroblasts

Tissue	Sample identification	CAG length	Postmortem delay, h	Sex	Age, y
Brain	H132	15/19	12	F	63
	H130	Normal range	13	M	32
	HC105	15/42	9	F	67
	HC76	19/42	16	M	71
	HD1	20/72	21	F	11
	HD2	17/72	3	F	20
Fibroblasts	Da.R.	15/31			
	4845*	16/20			
	5539*	18/68			
	9197*	21/181			

F, female; M, male.
*4845, 5539, and 9197 (Coriel Cell Repository).

Table S3. Primer sequences for mouse 3′RACE and RT-PCR

Name	Sequence	Position from start of HTT intron 1; start/end in bp	Function
−19f	AGGAACCGCTGCACCGA	−19/−3	qRT-PCR/RT-PCR
135f	CTTGCGGGGTCTCTGGC	135/151	qRT-PCR/RT-PCR
200r	TCAGCGAGTCCCTGGCTG	183/200	qRT-PCR/RT-PCR
155p	CCTCAGAGGAGACAGAGCCGGGTCA	155/179	qRT-PCR/RT-PCR
347f	TCCTCATCAGGCCTAAGAGCTGG	347/369	qRT-PCR/RT-PCR
431r	GAGACCTCCTAAAAGCATTATGTCATC	405/431	qRT-PCR/RT-PCR
371p	AGTGCAGGACAGCGTGAGAGATGTG	371/395	qRT-PCR/RT-PCR
785r	TGAAAACTGAGCACCACCAATG	764/785	RT-PCR
1006f	GAAATCCATGCTGAGTGTTGAGC	1006/1028	qRT-PCR/RT-PCR
1072r	TGCCCAGAGTTGAGAGAAAGGA	1051/1072	qRT-PCR/RT-PCR
1030p	CTGTGCTCTCTTGCAGCT	1030/1047	RT-PCR
11588f	ATGATGCTGGTATGCCTCTGGG	11588/11609	RT-PCR
11779r	CATCCCACTGCCATGACACA	11760/11779	RT-PCR
571f	AACCAGGTTTTAAGCATAGCCAGA	571/594	3′RACE
622f	AGTTGGATGAGTTGTATTTGTCAAGTACAT	622/651	3′RACE
Qt	CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGC(T)18		RT
OligodT	(T)18		RT
Qo	CCAGTGAGCAGAGTGACG		3′RACE
Qi	GAGGACTCGAGCTCAAGC		3′RACE
1756f	CCTAATGCCAGGCAAGAATTTCACA	1756/1780	RT-PCR
2075r	AGTGTTCTTAGATTCCAGATAGCAG	2051/2075	RT-PCR
5′UTR 1f	CTTGGTTCCGCTTCTGCC	−323/−306	qRT-PCR/RT-PCR
5′UTR 1r	TGGAGCCTACTGGCACTACG	−241/−260	qRT-PCR/RT-PCR
5′UTR 1p	CAGAGCCCCATTCATTGCCTTGCT	−298/−275	qRT-PCR/RT-PCR
int 10kbf	CTGGGTGATCTCCTCTGCTT	11605/11624	qRT-PCR/RT-PCR
int 10kbr	ACACTGCTGCCATGTCACATG	11694/11674	qRT-PCR/RT-PCR
int 10kbp	TTTGGCATGGTGGCATGCAGACC	11631/11653	qRT-PCR/RT-PCR
ex2f	AAGAAGGAACTCTCAGCCACCA	Exon 2	qRT-PCR/RT-PCR
ex2r	CTGAGAGACTGTGCCACAATGTT	Exon 2	qRT-PCR/RT-PCR
ex2p	AGAAAGACCGTGTGAATCATTGTCTAACAATATGTGA	Exon 2	qRT-PCR/RT-PCR
ex5-6f	CTCAGAAGTGCAGGCCTTACCT	Exon 5–6	qRT-PCR/RT-PCR
ex6r	GATTCCTCCGGTCTTTTGCTT	Exon 6	qRT-PCR/RT-PCR
ex6p	TGAATCTTCTTCCATGCCTGACCCGA	Exon 6	qRT-PCR/RT-PCR
Hsf1f	ACGTCCCGGCCTTCCTAAC	*Hsf1*	qRT-PCR
Hsf1r	CCAGCAGATGAGCGCGTCT	*Hsf1*	qRT-PCR
Hsf1p	TGTGGACCCTCGTGAGCGACCC	*Hsf1*	qRT-PCR

f, forward; *P*, probe; qRT-PCR, qualitative RT-PCR; r, reverse.

Table S4. Quality control statistics for RNA-Seq libraries

Sample no.	Total	Uniquely mapping	% uniquely mapping	Ribo mapping	% ribo mapping	Exon/intron	Exon/intergenic
4-mo.							
Hom 148	116047286	96252932	83	504675	0.52	232	3645
Hom 157	76601632	64778705	85	220124	0.34	267	3721
Wt 154	96791476	81172594	84	229466	0.28	226	3347
Wt 155	91993732	78231973	85	245496	0.31	257	3505
6-mo.							
Hom 117	58088390	48599615	84	22723	0.05	70	1034
Hom 118	77528134	64398258	83	8146	0.01	61	935
Wt 43	90864660	75003852	83	14517	0.02	109	1354
Wt 46	44043910	36072275	82	6350	0.02	95	1294
22-mo.							
Hom 74	107693082	90260412	84	199891	0.22	181	3193
Hom 84	129917678	109578186	84	151925	0.14	92	2190
Wt 71	139655344	117257850	84	636391	0.54	76	1907
Wt 87	103948402	85726377	82	292999	0.34	217	3330

Table S5. Primer sequences for human 3′RACE and RT-PCR

Name	Internal name	Sequence	Position from start of HTT intron 1; start/end in bp	Function
UAPdT18	UAPdT18	GGCCACGCGTCGACTAGTAC(T)18		RT
UAPnest	UAPnest	GGCCACGCGTCGACTAGTAC		3′RACE
2181f	3124	ATCTTCCTTCCACCCCTTTCC	2181/2201	3′RACE
2357f	3124nest	AGTCTCACATTGTCACCCAGGC	2357/2378	3′RACE
6568f	7385	TCAAGACATTCTCCTGCACGG	6568/6588	3′RACE
6621f	7385nest	CACCACACCCAGCTAATTTTGTAT	6621/6644	3′RACE
7128f	7741	GAGGACTTTTGGAGATGTAAAGGC	7128/7151	3′RACE
7169f	7741nest	GTGGCTGTAGAATGTGCTGGTG	7169/7190	3′RACE
−403f	hex1for	GGGTCCAAGATGGACGGC	−403/−395	RT-PCR
−226r	HuHD 04Rv	TTGAGGGACTCGAAGGCCTT	−245/−226	RT-PCR
−17f	HuHD 05Fw	GAGCCGCTGCACCGAC	−17/−1	RT-PCR
2805f	3220for	GATTTTGGCAGTTCTGTTCACG	2805/2826	RT-PCR
2959r	3360rev	ATAAACTGAGGCCCATGCATG	2939/2959	RT-PCR
8140f	8500f	AAGATCATTTTGGCTGGGCAC	8140/8160	RT-PCR
8354r	8500r	AAGCAATCCTCCAGCCTCCT	8335/8354	RT-PCR
exon2f	hex2for	AAAGAAAGAACTTTCAGCTACCAAGAA	*HTT* exon 2	RT-PCR
exon2r	hex2rev	CTGACAGACTGTGCCACTATGTTT	*HTT* exon 2	RT-PCR