Alterations in mRNA 3′UTR isoform abundance accompany gene expression changes in human Huntington's disease brains

Lindsay Romo1,4,* , Ami Ashar-Patel2, Edith Pfister1, and Neil Aronin1,2,*

1 Dept. of Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA
2 RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, MA 01605, USA

Summary

The huntingtin gene has two mRNA isoforms that differ in their 3′UTR length. The relationship of these isoforms with Huntington's disease is not established. We provide evidence that the abundance of huntingtin 3′UTR isoforms differs between patient and control neural stem cells, fibroblasts, motor cortex, and cerebellum. Huntingtin 3′UTR isoforms, including a mid-3′UTR isoform, have different localizations, half-lives, polyA tail lengths, microRNA sites, and RNA binding protein sites. Isoform shifts in Huntington's disease motor cortex are not limited to huntingtin. Eleven percent of alternatively polyadenylated genes change the abundance of their 3′UTR isoforms. Altered expression of RNA binding proteins may be associated with aberrant isoform abundance; knockdown of the RNA binding protein CNOT6 in control fibroblasts leads to huntingtin isoform differences similar to those in disease fibroblasts. These findings demonstrate mRNA 3′UTR isoform changes are a feature of molecular pathology in the Huntington's disease brain.

Graphical abstract

Romo et al. show the abundance of mRNA 3′UTR isoforms of huntingtin and a subset of other genes changes in Huntington's disease motor cortex. Genes with isoform alterations occur in pathways implicated in disease pathogenesis. Differential expression of RNA binding proteins may cause these isoform changes in Huntington's disease.

*Correspondence: lindsay.romo@umassmed.edu (L.R.), neil.aronin@umassmed.edu (N.A.).
4Lead Contact

Author Contributions. L.R. and N.A. conceived the project, L.R. and E.P. designed experiments, L.R. performed experiments and interpreted results, A.A. designed the PAS-seq method, and L.R. wrote the manuscript with comments from all authors.

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.
Introduction

Huntington’s disease (HD) is a neurodegenerative disorder caused by expansion of the CAG repeat in huntingtin (HTT) exon 1 (Macdonald et al., 1993). The expansion is dominantly inherited and fully penetrant. Individuals usually develop symptoms in their mid-thirties, and death occurs within two decades (reviewed in Suchowersky, 2014). Neurodegeneration starts in the striatum and motor cortex, but spreads to other brain regions by the time of death (Vonsattel et al., 1985). Neuronal toxicity is largely due to the effect of mutant HTT protein, although mutant messenger RNA (mRNA) may also be toxic (Bañez-Coronel et al., 2015; DiFiglia et al., 1997; Li et al., 2008; Zuccato et al., 2010). Emerging therapies aim to decrease mutant HTT mRNA (Hu et al., 2010; Pfister et al., 2009).

Three isoforms of HTT mRNA are produced via alternative polyadenylation. A truncated (7.9 kb) isoform is terminated at a cryptic polyadenylation (polyA) signal in the first intron (Sathasivam et al., 2013). This isoform is of low abundance. Its presence in HD patients and mouse models, but not controls, suggests it may be associated with disease (Sathasivam et al., 2013). The two predominant HTT mRNA isoforms are generated by alternative polyadenylation in the 3′ untranslated region (3′ UTR) (Lin et al., 1993). The short (10.3kb) 3′ UTR isoform predominates in dividing cells, whereas the long (13.7kb) isoform predominates in nondoning cells (Lin et al., 1993). It has not been established whether the relative abundance of HTT mRNA 3′UTR isoforms changes in HD. Although many genes exhibit expression and splicing differences in HD, 3′ UTR isoform expression alterations have not been investigated (Hodges et al., 2006; Labadow et al., 2015; Lin et al., 2016; Mykowska et al., 2011).

3′ UTR length is important for mRNA localization, stability, and translation (Di Giammartino et al., 2011). Many mRNAs use 3′ UTR alternative polyadenylation to achieve tissue-specific expression and function (Lianoglou et al., 2013; Smibert et al., 2012; Zhang
et al., 2005). Long isoforms contain binding sites for transacting factors such as microRNAs that exert dynamic changes in steady-state mRNA levels (Sandberg et al., 2008). Ubiquitously expressed genes, like HTT, are the most likely to use alternative polyadenylation to alter mRNA expression in the brain (Lianoglou et al., 2013; Lin et al., 1993). Changes in mRNA 3′ UTR length occur in Parkinson’s disease and cancer (Mayr and Bartel, 2009; Rhinn et al., 2012). We sought to determine if alterations in 3′ UTR length are a feature of HD.

We provide evidence that the abundance of HTT mRNA 3′ UTR isoforms differs between HD patients and non-HD controls, including a mid-length 3′ UTR isoform. We show that HTT 3′ UTR isoforms have different localizations, half-lives, polyA tail lengths, microRNA sites, and RNA binding protein sites. Isoform alterations in HD are not unique to HTT; there are widespread changes in mRNA 3′ UTR isoform expression in HD motor cortex. Genes with isoform changes are associated with pathways dysfunctional in HD. Knockdown of the CNOT6 RNA binding protein leads to isoform shifts similar to those in HD motor cortex. Our findings indicate that altered 3′ UTR isoform expression of HTT and a subset of other genes is a feature of molecular pathology in HD motor cortex.

Results

The abundance of HTT mRNA 3′ UTR isoforms changes in HD

We identified two abundant HTT mRNA 3′ UTR isoforms in public 3′ sequencing data (Figure 1A) (Lianoglou et al., 2013). As in previous studies, we found HTT mRNA isoform abundance varies across normal tissues: the longer isoform predominates in the brain, neuronal precursor (NTERA2) cells, breast, and ovary, whereas the shorter isoform predominates in testes, B-cells, muscle, and human embryonic kidney (HEK293) cells (Figure 1B) (Lin et al., 1993).

To determine whether HTT mRNA isoform abundance changes in HD, we performed qRT-PCR of the HTT long isoform normalized to total HTT expression on patient cerebellum, motor cortex, fibroblasts, and neural stem cells. HD cerebellum samples were from grade 2, 3, or 4 brains, whereas motor cortex samples were from grade 1 and 2 brains (Table S1). HD brain grades are based on the degree of striatal degeneration. Grades range from 1 (50% neuronal loss) to 4 (95% neuronal loss) (Vonsattel et al., 1985). There is limited involvement of the cerebellum in all grades (Rüb et al., 2013; Vonsattel, 2008; Vonsattel et al., 1985). Cortical pathology is variable, but typically 10% of neurons are lost in grade 1, and up to 40% are lost in grade 2 (Thu et al., 2010; Vonsattel et al., 1985). Neuronal loss can confound gene expression studies. We chose to analyze HTT isoform expression in cortex and cerebellum rather than in the striatum because neuronal loss is limited in these regions.

We found a 1.6-fold increase in the long 3′ UTR isoform relative to total HTT mRNA in HD patient cerebellum. Conversely, there is a 2.5-fold decrease in the long isoform in HD motor cortex (Figure 1C). Patient fibroblasts and neural stem cells also exhibit a decrease in the long isoform (Figure 1D, 2-fold and 3.3-fold, respectively). These results demonstrate the abundance of HTT mRNA 3′ UTR isoforms changes in HD in a tissue and cell-specific manner.
**HTT mRNA isoform changes extend to liver and muscle and arise from both alleles**

We used HD model mice to determine if disease-associated isoform changes extend beyond brain. We bred mice that lack murine Htt, but harbor full length human HTT with either 18 or 128 CAG repeats on a yeast artificial chromosome (Yac18, Yac128) (Slow et al., 2003; Zeitlin et al., 1995). Like patients, Yac128 cerebellum exhibits an increase (1.7-fold) in the long HTT 3′ UTR isoform compared to yac18 cerebellum. Isoform differences extend to muscle (quadriceps) and liver, with a 2.2-fold increase and a 1.6-fold decrease in the HTT long mRNA isoform, respectively (Figure 2A). These findings indicate that at least two non-brain tissues exhibit tissue-specific HTT mRNA isoform changes.

Altered HTT mRNA isoform abundance in HD brains could arise from the wild-type allele, the mutant allele, or both. To determine which allele is responsible for the changes, we performed allele-specific qPCR on human cerebellum and motor cortex samples heterozygous for single nucleotide polymorphism (SNP) rs362267 in the HTT long, but not short, 3′UTR. This SNP heterozygosity is not linked to the CAG repeat; in some patients, the C allele is expanded, whereas in others the T allele is expanded. The long isoform qPCR probe matched either the C or the T SNP. The amounts of both the mutant and wild-type HTT long 3′UTR isoform are changed in HD cerebellum and motor cortex compared to controls (Figure 2B, C). The direction of isoform changes is consistent between alleles in each patient (Figure 2D, correlation coefficient 0.83). Thus the change in HTT mRNA isoforms arises from both alleles in HD.

**PolyA site sequencing identified a conserved mid-3′UTR isoform of HTT mRNA**

To identify other 3′ UTR isoforms of HTT, we used a high quality polyA site sequencing method, PAS-seq (Figure 3A-B, S1) (Ashar-Patel et al., 2017). PAS-seq confirmed there is a decrease in the long isoform (1.4-fold) in HD motor cortex, whereas there is a decrease in the short isoform (1.8-fold) in HD cerebellum (Figure 3G-H, J). PAS-seq identified a third 12.5kb isoform of HTT expressed in both control and HD human cerebellum and motor cortex (Figure 3C, top). We performed PAS-seq on mice to test if this mid-3′ UTR isoform is conserved. We used wild-type and Q140 HD mice, which harbor 140 CAG repeats knocked into mouse HTT (Menalled et al., 2003). All three HTT 3′ UTR isoforms are present in wild-type and Q140 mice (Figure 3C, bottom). There is a putative polyA signal (AAUGAA) located twenty nucleotides upstream of the mid-3′ UTR isoform polyA site (Sheets et al., 1990; Tian et al., 2005). This polyA signal and site are conserved between mice and humans (Figure 3D). The regions up and downstream of the isoform polyA sites are also conserved (Figure 3E). The abundance of the mid-3′ UTR isoform varies across tissues and increases 2.8-fold in HD motor cortex compared to control (Figure 3F-H). Isoform changes cannot be explained by gene expression alterations: total HTT expression is decreased by 1.4-fold in HD motor cortex but unchanged in cerebellum compared to controls (Figure 3I). In addition, isoform changes are still significant after normalizing to total HTT expression (Figure 3J).

**HTT mRNA 3′UTR isoforms are metabolized differently**

3′ UTR length affects mRNA localization, stability, and translation (Di Giammartino et al., 2011). Isoform shifts in HD may affect HTT mRNA metabolism. To assay the localization of HTT 3′ UTR isoforms, we performed isoform specific qRT-PCR (Figure S2A-B) on
nuclear or cytoplasmic mRNA from SH-SY5Y (human neuroblastoma) cells and HD or control fibroblasts. SH-SY5Y cells have neuronal characteristics, making them a model of mRNA metabolism in brain (Datta et al., 2013). In SH-SY5Y cells, all HTT isoforms were three times more abundant in the cytoplasm than in the nucleus, whereas the cytoplasmic to nuclear expression ratio was lower in fibroblasts (Figure 4A). In fibroblasts, the HTT mid isoform was more abundant in the cytoplasm than the other isoforms. The ratio of cytoplasmic to nuclear transcripts is lower in HD fibroblasts than in wild-type fibroblasts for all three isoforms, indicating HTT mRNA shifts to the nucleus in HD, as has been reported (de Mezer et al., 2011). To test HTT 3′UTR isoform stability, we incubated SH-SY5Y cells with labeled uridine and collected labeled mRNA at various times after washing away the modified nucleotide. Isoform specific qRT-PCR revealed all HTT isoforms have significantly different half-lives; the short isoform is most stable, whereas the mid isoform is least stable (Figure 4B). These findings suggest that altered HTT 3′UTR isoform abundance in HD results in changes in total HTT mRNA localization and stability.

Differences in HTT 3′UTR isoform metabolisms could be due to different polyA tail lengths, microRNA sites, or RNA binding protein sites (Di Giammartino et al., 2011; Norbury, 2013). To determine the polyA tail length of the isoforms, we added an adaptor to the 3′ end of the polyA tail that was used as a priming site during isoform-specific PCR. Products were resolved on a gel, enabling quantification of the polyA tail (Figure 4C). We found the HTT short isoform has a polyA tail length of fifty to eighty in both SH-SY5Y cells and in HD and control fibroblasts, whereas the HTT mid and long isoforms have polyA tail lengths of about five and ten (Figure 4D, S2C). Studies show mRNA stability and translation rate drop as the polyA tail decreases below twenty nucleotides (Chang et al., 2014; Park et al., 2016). The short polyA tail length of the mid and long HTT isoforms may explain their low half-lives compared to the short isoform.

We used program TargetScan to identify putative microRNA sites in the HTT 3′UTR (Lewis et al., 2003). We found several microRNAs expressed in human brain with sites exclusive to the mid and long HTT 3′UTR (Figure 4E). To test if these microRNAs regulate HTT isoform abundance, we transfected SH-SY5Y cells with mimics of microRNAs 221, 137, or both. MicroRNA 137 exclusively binds the long isoform, while microRNA 221 has an 8-mer binding site in the long isoform and a 6-mer binding site common to the mid and long isoform. The microRNA 137 mimic had no effect on isoform abundance, but transfection of the microRNA 221 or both mimics reduced the abundance of the long and mid but not short isoform relative to total HTT or GAPDH (Figure 4F-G). These results show at least one microRNA exclusively degrades the long and mid HTT transcripts, potentially contributing to the lower stability of these isoforms compared to the short isoform.

To identify RNA binding protein sites exclusive to the mid and long isoform, we used CLIP-Seq databases (Li et al., 2014; Yang et al., 2015). We found several brain-expressed RNA binding proteins adhere to the HTT 3′UTR (Figure S2D). These RNA binding proteins may differentially affect the stability and localization of HTT isoforms. Based on these results, we expect the short isoform is translated more efficiently than the mid and long isoform due to its longer polyA tail and increased stability. Indeed, a study published during preparation...
of this manuscript found the short *HTT* 3′UTR isoform may be translated more efficiently than the long isoform (Xu et al., 2017).

**Many other genes exhibit changes in isoform abundance**

We studied whether alterations in mRNA isoform expression are unique to *HTT*, or features of many genes in HD. To assay transcriptome-wide isoform expression, we performed PAS-seq on motor cortex from grade 1 HD brains (n=6) versus controls (n=5), and cerebellum from grade 2-4 patient brains (n=9) versus controls (n=7). We identified genes with multiple 3′UTR isoforms, and normalized isoform expression to gene expression. We compared the normalized isoform expression between HD and control samples.

PAS-seq uses oligo(dT) to capture polyA tails during reverse transcription. Oligo(dT) also anneals to genomic polyA stretches. These genomic sites are artefacts that can lead to overestimation of the number of alternatively polyadenylated genes, and we discarded them from our analysis (Sheppard et al., 2013). We found the number of alternatively polyadenylated genes in our data (about 30%, Figure S3) is similar to that obtained via the 3Pseq 3′ sequencing method, which avoids oligo(dT) priming (Jan et al., 2011). Consistency with 3Pseq data indicates PAS-seq accurately captured reads primed from mRNA polyA tails while excluding reads primed from genomic regions.

In the motor cortex, 11% of alternatively polyadenylated genes showed a significant (false discovery rate<10%) change in abundance of at least one 3′UTR isoform in HD (Figure 5A). Many isoform relative abundances changed over 25% in HD motor cortex versus controls, indicating a major shift in isoform expression for the corresponding gene (Figure 5B). In the cerebellum, none of the alternatively polyadenylated genes showed a significant change in abundance of at least one isoform (Figure 5C). The false discovery rate threshold corrects for false-positives due to multiple testing, but may reject true positives. To allow comparison of isoform changes between motor cortex and cerebellum, we also included a more sensitive p-value threshold (p<0.01). Using this threshold, 13% of alternatively polyadenylated genes exhibit isoform changes in HD motor cortex, whereas 5% of genes exhibit changes in cerebellum. Isoform abundance changes in the cerebellum were not as large as in the motor cortex (Figure 5D). These results indicate many genes alter their 3′UTR isoform abundances in HD motor cortex, whereas changes in the cerebellum are negligible.

Neurons die and glia proliferate in the HD brain (Vonsattel and DiFiglia, 1998). Neuronal loss and gliosis are limited in motor cortex from grade 1 brains (Thu et al., 2010). However, it is possible some or all isoform changes we see are due to changes in cell populations. We found isoform abundance is highly correlated (r=0.87) between the motor cortex and cerebellum despite the different cell milieu and trace neuronal loss in the cerebellum (Figure S4). Previous studies that used HD caudate, which suffers extensive neuronal loss (50-95%), have demonstrated mRNA expression is similar in HD caudate tissue and HD caudate neurons (Hodges et al., 2006). These results suggest the isoform changes we observe in HD motor cortex are not solely due to neuronal loss.
To determine whether genes with significant (p<0.01) isoform changes shifted towards longer or shorter mRNA 3′ UTR isoforms, we calculated a weighted change (see Experimental Procedures). If the weighted change is positive, the gene shifts towards longer mRNA isoforms in HD patients, whereas if it is negative, the gene shifts towards shorter isoforms. We found 129 genes shift towards longer isoforms in HD motor cortex, whereas 110 genes shift towards shorter isoforms (Figure 5E). In the cerebellum, 31 genes shifted to the longer isoform, whereas 83 shifted to the shorter isoform (Figure 5F). We compared significant (p<0.01) isoform changes in the motor cortex and cerebellum. Only 11 genes exhibit mRNA 3′ UTR isoform shifts in both brain areas (Figure 5G, purple). These findings indicate that there is no generalized trend towards longer or shorter 3′ UTR isoforms in HD, and isoform alterations are region-specific.

To identify affected pathways, we performed gene ontology (GO) analysis on genes with significant (p<0.01) isoform changes (Figure 5H). We found several pathways that have been reported to be enriched among genes differentially expressed in HD, including cytokine signaling and production, RNA pol2 transcription, translation, calcium signaling, vesicle-mediated transport, microtubule organization, and DNA binding (Hodges et al., 2006; Labadorf et al., 2015). Some enriched pathways are thought to be involved in HD pathogenesis, including calcium signaling, cytokine signaling, histone acetylation, transcription factor binding, axonal transport, synaptic vesicle localization, dendritic spine morphogenesis, microtubule organizing, and phagocytic vesicles (Ellrichmann et al., 2013; Martinez-Vicente et al., 2010; Mounné et al., 2013; Trushina et al., 2003; Zuccato et al., 2010). Aberrant 3′ UTR isoform expression may influence pathogenesis in these disease pathways.

**Most genes with isoform changes in HD are not differentially expressed**

Widespread gene expression changes are reported in the cortex of late-stage HD patients (Hodges et al., 2006; Labadorf et al., 2015). We sought to determine if the extensive 3′ UTR isoform changes in HD motor cortex co-occur with gene expression changes. To measure steady-state gene expression, we combined all PAS-seq reads from each gene (see Experimental Procedures). Consistent with previous studies, many genes were differentially expressed in HD patient motor cortex relative to controls (false discovery rate<10%) (Figure 6A) (Hodges et al., 2006; Labadorf et al., 2015). Quantitative PCR confirmed differential expression of three genes (Figure S5). Of differentially expressed genes, 700 exhibited increased expression in HD motor cortex, whereas 1008 exhibited decreased expression (Figure 6C). The percent of detected genes exhibiting expression changes is identical to the percent of alternatively polyadenylated genes exhibiting isoform changes (11%). In contrast, only 1% of genes are differentially expressed (false discovery rate<10%) in patient cerebellum (Figure 6B). Of differentially expressed genes, 89 increased expression in HD patient cerebellum whereas 64 decreased (Figure 6D). Thus, isoform changes are proportional to gene expression changes in the HD motor cortex, whereas isoform and gene expression changes are minimal in the cerebellum.

To determine if genes with isoform changes also exhibit expression changes, we compared gene expression to isoform expression for all alternatively polyadenylated using a
significance cutoff of p<0.01 (Figure 6E-F). In the motor cortex, only 17% of genes with isof orm shifts are differentially expressed (Figure 6E, red). Of genes exhibiting isof orm and gene expression changes, we found no association between isof orm shift and gene expression direction. This dissociation of isof orm and gene expression is consistent with studies showing the relationship between 3′UTR length and isof orm stability is genespecific (Spies et al., 2013). In the cerebellum, only 4% of genes with significant isof orm changes are differentially expressed (Figure 6F, red). These results indicate that most genes with isof orm changes are not differentially expressed in HD motor cortex and cerebellum.

We sought to determine if isof orm and gene expression changes are associated with common cell pathways. We found several ontology categories enriched among genes with isof orm changes and genes with expression changes including mRNA 3′UTR binding, cell junction, ubiquitination, mRNA binding, and protein transport (Figure 5H, Figure 6H). Within each category, genes with isof orm differences and genes with expression changes do not overlap; i.e., the same gene does not exhibit aberrant isof orm and gene expression. Isoform alterations and gene expression changes are separate, but they may affect common processes in HD brains.

**Decreasing expression of the RNA binding protein CNOT6 leads some genes to change isof orm abundance**

A change in the expression of an RNA binding protein may lead to changes in isof orm abundance by altering isof orm production or decay. To identify candidate RNA binding proteins, we searched for genes differentially expressed in HD motor cortex with the RNA binding protein GO-term (Figure 7A). Of those, twenty were involved in mRNA alternative polyadenylation or stability. We reasoned an RNA binding protein responsible for HTT isof orm shifts would show opposite expression changes in HD motor cortex and cerebellum. Of nine with opposite expression changes, two (CNOT6 and CNOT7) are subunits of the Ccr4-not complex that plays a widespread role in mRNA metabolism (Mittal et al., 2011; Shirai et al., 2014). CNOT6, the subunit that catalyzes deadenylation of mRNAs, is most differentially expressed in HD motor cortex and cerebellum (Yamashita et al., 2005).

To determine if changes in CNOT6 expression exert changes in isof orm expression, we transfected control human fibroblasts with siRNAs targeting CNOT6 or an off-target control (MAP4K4). The CNOT6 siRNAs reduced CNOT6 mRNA expression over 80% but did not affect total HTT mRNA levels (Figure 7B-C). We chose HTT and SECRISBP2L (SECIS Binding Protein 2 Like) as candidate CNOT6 targets. HTT and SECRISBP2L shift to longer 3′UTR isoforms in HD cerebellum and shorter isoforms in HD motor cortex. Knock down of CNOT6 resulted in a relative decrease in the HTT and SECRISBP2L long isoforms similar to that seen in HD patient motor cortex and in HD fibroblasts (Figure 7D). HTT isof orm-specific qRT-PCR revealed that the abundance of the short isof orm didn't change whereas abundance of the long isof orm decreased by almost two-fold, as in HD patient motor cortex (Figure 7E). These results suggest changes in the expression of CNOT6 may influence the abundance of some 3′UTR isoforms in HD patient motor cortex and cerebellum.
Widespread changes in isoform abundance may lead to aberrant mRNA metabolism in HD. Some 3′ UTR isoforms exhibit unique localizations and functions. The long but not short isoform of brain-derived neurotrophic factor mRNA is translated in dendrites, where the protein is vital for pruning and spine enlargement (An et al., 2008). The long isoform of CD47 mRNA acts as a scaffold for proteins that translocate the CD47 protein to the cell membrane, whereas the short isoform localizes the protein to the endoplasmic reticulum (Berkovits and Mayr, 2015). Loss of normal 3′ UTR isoform abundance can contribute to disease. The alpha synuclein long isoform is localized to mitochondria; increased abundance of the extended alpha synuclein mRNA may contribute to mitochondrial dysfunction in Parkinson's disease (Rhinin et al., 2012). Many genes in cancer cells shift towards more stable isoforms, increasing translation of tumorigenic proteins (Mayr and Bartel, 2009). In myotonic dystrophy, triplet repeat RNAs sequester muscleblind, a regulator of 3′ UTR length, resulting in an altered polyadenylation profile (Batra et al., 2014). We demonstrate that genes with isoform changes in HD motor cortex are associated with pathways disrupted in HD. Altered metabolism of mRNAs in disease pathways could result in aberrant localization or function of proteins and contribute to HD pathogenesis.

Altered expression of RNA binding proteins in the HD brain may lead to changes in the expression of mRNA 3′ UTR isoforms. RNA binding proteins may interact with nascent mRNA during transcription and alternative polyadenylation, resulting in changes in polyA site selection and isoform production. For instance, aberrant expression of 3′ end modulatory factors may lead to isoform abundance changes (Gruber et al., 2014). We found three 3′ end processing proteins are differentially expressed in HD grade 1 motor cortex: CPSF2, PCBP2, and THOC5. RNA binding proteins may also adhere to isoforms post-transcriptionally, affecting the stability and decay of isoforms and leading to changes in isoform abundance. We show a decrease in the CNOT6 deadenylase is associated with a decrease in the HTT long isoform but no change in the short or mid isoforms. Further studies are necessary to determine the mechanism of 3′ UTR isoform changes in HD. We found HTT and SECISBP2L isoform amounts are responsive to changes in CNOT6 mRNA levels. However, altered CNOT6 expression in HD may not explain all isoform changes we identified. Other RNA binding proteins, such as 3′ end processing factors, likely affect isoform abundance in HD. Exploration of the causes of isoform changes in HD motor cortex may result in novel therapeutic targets.

Our findings have important implications for HD and may apply to other diseases. The increased relative abundance of the HTT short isoform in HD motor cortex might contribute to pathology; a study published during preparation of this manuscript found transfected HTT exon 1-short 3′ UTR constructs form more aggregates than exon 1-long 3′ UTR constructs (Xu et al., 2017). New therapies aim to disrupt the mutant but not wild-type HTT mRNA by targeting SNPs heterozygous in patients (Østergaard et al., 2013; Pfister et al., 2009). Our results suggest targeting SNPs exclusive to the long isoform may be less effective than targeting SNPs common to all isoforms, as HTT long isoform abundance is reduced in HD patient motor cortex. Further research is necessary to determine if isoform alterations are unique to HD, or if they frequently accompany gene expression changes in human disease.
Experimental Procedures

Samples

Human samples were supplied by the New York and Neurological Foundation of New Zealand brain banks. Fibroblast lines were obtained from the Coriell Repository. Mouse tissues were dissected from 7-9 month male and female mice. All protocols were reviewed and approved by the institutional biosafety committee (human samples) or institutional animal care and use committee (mouse samples).

Public 3′Seq data analysis

Raw 3′ sequencing data was downloaded from the NCBI short sequence read archive, accession SRP029953 (Lianoglou et al., 2013), and analyzed as below.

RNA extraction

Total RNA was extracted with Trizol (Ambion), cleaned on Clean and Concentrator columns (Zymo), analyzed via Bioanalyzer (Agilent), and treated with TurboDNase (Ambion).

qRT-PCR

RNA was reverse transcribed by SuperScript IV (Invitrogen). Quantitative PCR was performed by Qiagen QuantiFast SYBR green master mix (non allele-specific), or Qiagen 1× Type-it Fast SNP PCR master mix (allele-specific). Product amounts were calculated using the ΔΔCt method. Primer and probe sequences are listed in the Supplemental Experimental Procedures.

Nuclear and cytoplasmic isoform abundance

SH-SY5Y cell (ATCC) nuclear and cytoplasmic RNA fractions were collected as described in Supplemental Experimental Procedures, and RNA was submitted to isoform-specific qRT-PCR.

EU pulse chase and isoform-specific qRT-PCR

SH-SY5Y cells (ATCC) were pulsed in 200μM EU for 14 hours, washed in PBS, and chased with unmodified media. We isolated labeled RNA using the Click-iT Nascent RNA Capture Kit (Life Technologies) as detailed in the Supplemental Experimental Procedures. qRT-PCR was performed, and isoform amounts were normalized to those of a firefly luciferase spike-in.

PolyA tail length assay

The HTT short and long isoform polyA tails were determined using the Affymetrix PolyA Tail Length Assay as described in the Supplemental Experimental Procedures.

MicroRNA and RNA binding protein target site prediction

We used the CLIPdb and starBase v2.0 browsers to identify HTT 3′ UTR RNA binding protein sites supported by more than five reads (Li et al., 2014; Yang et al., 2015). Brain expression was confirmed with the human protein atlas, www.proteinatlas.org (Uhlen et al.,
To identify predicted microRNA 7mer-1A, 7mer-m8, 8mer binding sites binding sites, we used TargetScanHuman version 7.0 (Lewis et al., 2003). Brain expression was confirmed with miRmine Human miRNA Expression Database (Panwar et al., 2017).

**MicroRNA transfection**

SH-SY5Y cells (ATCC) were transfected with mimics (Dharmacon) to microRNAs 137, 221, a C. Elegans microRNA, or an HTT ORF-targeting siRNA via Lipofectamine RNAimax (Invitrogen). After 72 hours, RNA was extracted and submitted to qRT-PCR.

**PAS-seq library preparation**

For details, see Supplemental Experimental Procedures. Briefly, RNA was fragmented and reverse-transcribed by Superscript III (Invitrogen) from the PAS-seq oligo d(T) primer. cDNA fragments 160-200 nucleotides long were circularized by CircLigase II (Epicentre), and Phusion polymerase (NEB) amplified the cDNA and adaptors with PE 1.0 and PE 2.0 primers (Illumina). The 215-254 nucleotide library was size-selected via gel electrophoresis and subjected to single-end 100 base-pair sequencing.

**PAS-seq analysis**

For details, see Supplemental Experimental Procedures. Briefly, sequences were mapped to the genome, and reads corresponding to the same 3′UTR isoform were combined. We removed genomically-primed reads and normalized isoform expression to gene expression. For isoforms represented in at least five samples, the normalized abundance was compared between HD and control samples. To calculate the isoform weighted change for each gene, we weighed isoform fractions by the isoform ordinal number multiplied by ten and determined the difference in the average weighted isoform fractions between HD and control samples. For differential gene expression analysis, the total reads mapping to each gene were compared between patients and controls using DESeq2 (Love et al., 2014). We considered genes with a change greater than 50% and false discovery rate <10% significantly differently expressed.

**Gene ontology analysis**

The list of genes with significant changes in total or 3′UTR isoform expression in HD brains was compared to the background total or alternatively polyadenylated expressed genes using Panther (Mi et al., 2013). We considered categories enriched if they had five or more genes, enrichment greater than 1.5 fold, and p<0.01.

**CNOT6 siRNA transfection**

Wild-type fibroblasts were transfected with three siRNAs targeting CNOT6 or MAP4K4 (Qiagen FlexiTube) via Lipofectamine RNAimax (Invitrogen). After 72 hours, RNA was extracted and submitted to qRT-PCR.

**Statistical methods**

Two-tailed t-tests were used to determine statistical significance for all figures except Figure 4, where an ANOVA test was used, and Figure 6, where DESeq2 was used for differential
gene expression analysis (Love et al., 2014). p<0.05 was considered significant unless otherwise stated.

**Accession Numbers**

PAS-seq data can be accessed at GEO: GSE96099. The PAS-seq data analysis pipeline can be accessed at [http://dx.doi.org/10.17632/99gfdrgbxc.2](http://dx.doi.org/10.17632/99gfdrgbxc.2).

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

We thank the M. DiFiglia lab (Massachusetts General Hospital) for providing neural stem cell pellets, X.W. Yang (University of California Los Angeles) for supplying null Htt mice, M. Hayden (University of British Columbia and Center for Molecular Medicine and Therapeutics) for supplying Yac18 mice, and the Cure Huntington’s Disease Initiative for supplying Q140 mice. We thank the New York and Neurological Foundation of New Zealand brain banks for supplying human samples. A. Bicknell provided the EU pulse chase protocol, and J. Alterman provided the HTT-targeting siRNA. S. Yang and C. Smith assisted with allele-specific qPCR. K. Chase, F. Conroy, and R. Miller performed mouse husbandry and genotyping. We thank S. Ryder, S. Corvera, V. Budnik, and A. Jacobson for their comments on experimental design. This study was supported by NIH grant 5 RO1 NS038194-15 (NA), NIH NINDS grant 1 F30 NS095450-01 (LR), and private donations. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

**References**


Ashar-Patel, A., Kaymaz, Y., Rajakumar, A., Bailey, JA., Karumanchi, SA., Moore, MJ. FLT1 and transcriptome-wide polyadenylation site (PAS) analysis in preeclampsia. Sci Rep. 2017. [http://dx.doi.org/10.1038/s41598-017-11639-6](http://dx.doi.org/10.1038/s41598-017-11639-6)


Cell Rep. Author manuscript; available in PMC 2017 October 03.


**Highlights**

- The relative abundance of huntingtin 3′UTR isoforms changes in Huntington's disease
- Huntingtin 3′UTR isoforms are metabolized differently
- Many other genes associated with Huntington's pathogenesis exhibit isoform changes
- Knockdown of the CNOT6 RNA binding protein recapitulates some isoform alterations
Figure 1. **HTT 3′ UTR isoforms change their relative amounts in HD**

(A) NTERA2 cell 3′ sequencing reads mapping to the HTT 3′ UTR. (B) Quantification of HTT 3′ UTR isoform 3′ sequencing reads relative to total HTT 3′ UTR reads across human tissues and cells. (C) Fold change in expression of the HTT mRNA long 3′ UTR isoform (long) relative to total HTT expression (total) between HD patient and control cerebellum (CB) and motor cortex (MCx), measured by qPCR. *, **, and *** signify p<0.05, 0.005, and 0.0005. (D) qPCR as in C, using patient or control fibroblasts or neural stem cells (NSCs).
Figure 2. *HTT* isoform changes extend to muscle and liver, and both *HTT* alleles change 3′UTR isoform amounts in HD

(A) Fold change in expression of the *HTT* long 3′UTR isoform (long) relative to total *HTT* expression (total) between Yac128 and Yac18 mouse tissues, measured by qPCR. CB=cerebellum. *, **, and *** signify p<0.05, 0.005, and 0.0005. (B) Allele-specific qPCR using SNP rs362267; shown is fold change in the *HTT* long 3′UTR isoform (long) relative to total *HTT* expression between patient and control cerebellum. (C) Allele-specific qPCR in patient and control motor cortex, as in B. (D) Correlation of *HTT* allele expression within each HD or control (con) cerebellum (CB) or motor cortex (MCx); r=0.83.
Figure 3. PAS-seq revealed a third HTT 3′ UTR isoform whose relative abundance also changes in HD

(A) PAS-seq library preparation. (B) PAS-seq data analysis on Tbr1, a gene with two 3′ UTR isoforms. See also Figure S1. (C) HTT 3′ UTR PAS-seq peak coordinates in HD patient and control cerebellum (CB) and motor cortex (MCx), and Q140 and wild-type (WT) mouse striatum (Str). (D) Alignment of the human (top) and murine (bottom) HTT 3′ UTR around the 3′ UTR isoform polyA signals (blue) and sites (yellow). (E) NCBI BLAST alignment score of the human and mouse HTT 3′ UTR; arrows indicate polyA sites. (F) HTT mid 3′ UTR isoform 3′ sequencing reads relative to total HTT 3′ UTR reads across human tissues and cells. (G-H) HTT 3′ UTR isoform read counts in HD motor cortex (G) or cerebellum (H) compared to controls, normalized to sequencing depth. * signifies p<0.05. (I) Total HTT read counts in HD motor cortex or cerebellum compared to controls. (J) HTT isoform read counts in HD motor cortex (left) or cerebellum (right) compared to controls, normalized to sequencing depth (top) or total HTT expression (bottom).
Figure 4. HTT 3′UTR isoforms are metabolized differently
See also Figure S2. (A) Ratio of cytoplasmic to nuclear HTT 3′UTR isoform expression, measured by qPCR. Data points are the average of triplicates with standard deviation. *, **, and *** signify p<0.05, 0.005, and 0.0005. (B) Relative abundance of labeled HTT 3′UTR isoforms in SH-SY5Y cells after a labeled uridine pulse, measured by qPCR. Data points are the average of triplicates with standard deviation. (C) PolyA tail length assay. (D) Agarose gel of PCR products from the polyA tail length assay. (E) Brain-expressed microRNA sites in the HTT 3′UTR. (F-G) Fold change in expression of HTT 3′UTR isoforms relative to total HTT (F) or GAPDH (G) expression between SH-SY5Y cells transfected with microRNA mimics (miRs) targeting HTT or a non-targeting (NT) control, measured by qPCR. Data points are the average of triplicates with standard deviation. siRNA is a small interfering RNA targeting the HTT open reading frame.
Figure 5. Transcriptome-wide PAS-seq analysis identifies a large subset of genes with 3′ UTR isoform changes in HD motor cortex

See also Figures S3 and S4. (A) Scatter plot of PAS-seq reads from control (n=5) and HD (n=6) motor cortex. Each dot is a 3′UTR isoform. Shown is the number of isoform reads divided by the gene’s total 3′UTR reads (isoform fraction). Orange and red dots indicate isoforms with p<0.01 or false discovery rate<10%, respectively. (B) Volcano plot of reads from control and HD motor cortex, colors as in A. The percent change is the difference in the isoform fraction between HD and control samples multiplied by 100. (C) Scatter plot of PAS-seq reads from control (n=7) and HD (n=9) cerebellum, as in A. Green dots indicate isoforms with p<0.01. (D) Volcano plot of reads from control and HD cerebellum, colors as in C. (E-F) Heat map of the weighted change of genes with one or more isoform differentially expressed (p<0.01) in patient motor cortex (E) or cerebellum (F). A positive number indicates the gene shifts towards longer isoforms in HD, whereas a negative number indicates the gene shifts towards shorter isoforms. (G) Comparison of gene weighted changes between patient versus control motor cortex (MCx) and cerebellum (CB). Green and orange dots have significant changes in HD cerebellum and motor cortex, respectively, and purple dots have both (p<0.01). (H) Gene ontology analysis of genes exhibiting significant (p<0.01) shifts to longer (long) or shorter (short) isoforms in HD motor cortex (MCx) and cerebellum (CB).
Figure 6. Gene expression analysis shows most genes with 3′UTR isoform changes do not exhibit expression changes in HD

See also Figures S4 and S5. (A) Scatter plot of gene expression normalized to sequencing depth in control (n=5) and HD (n=6) motor cortex, measured by PAS-seq. Orange and red dots indicate genes with p<0.01 or false discovery rate<10%, respectively. (B) Scatter plot of gene expression from control (n=7) and HD (n=9) cerebellum, as in A. Green and blue dots indicate isoforms with p<0.01 or false discovery rate<10%, respectively. (C-D) Heat map of genes significantly differentially expressed (false discovery rate<10%) in patient motor cortex (C) or cerebellum (D). Shown are fold changes in normalized gene expression. (E) Comparison of gene isoform weighted changes and differential expression (DE) log 2 fold changes in HD versus control motor cortex. Green and orange dots have significant isoform or expression changes in HD, respectively, and red dots have both (p<0.01). (F) Comparison of gene isoform weighted changes and expression log 2 fold changes in HD versus control cerebellum, as in E. (G) Comparison of gene expression log 2 fold changes in motor cortex (MCx) and cerebellum (CB). Red and blue dots have significant expression changes in HD motor cortex or cerebellum, respectively, and light blue dots have both (false discovery rate<10%). (H) Gene ontology analysis of genes with significant (false discovery rate<10%) expression changes in HD motor cortex (MCx) and cerebellum (CB).
Figure 7. Changes in the expression of RNA binding proteins may influence alterations in isoform abundances

(A) Algorithm to identify proteins that may cause 3' UTR isoform changes in HD. Numbers are fold changes (FC) in gene expression between HD and control motor cortex (MCx) or cerebellum (CB).

(B) Fold change in CNOT6 expression relative to GAPDH expression after transfection of CNOT6 siRNA (gray), no siRNA (white), or non-targeting (NT) MAP4K4 control siRNA (black) into wild-type fibroblasts, measured by qPCR. *, **, and *** signify p<0.05, 0.005, and 0.0005. Shown is the average of at least three transfections with standard deviation.

(C) Fold change in total HTT expression relative to GAPDH expression after transfection as in (B).

(D) Fold change in expression of the HTT long isoform (long) relative to total HTT expression (total) or of the SECISBP2L long isoform (long) relative to total SECISBP2L expression (total) after transfection as in (B).

Red and blue bars are untreated fibroblasts. (D) Fold change in each HTT isoform isoform (long) relative to total HTT expression after transfection of CNOT6 siRNA (colored), no siRNA (white), or non-targeting (NT) MAP4K4 control siRNA (black) into wild-type fibroblasts.