Alternative processing of human *HTT* mRNA with implications for Huntington's disease therapeutics

Sandra Fienko,[1] Christian Landles,[1] Kirupa Sathasivam,[1] Sean J. McAteer,[1] Rebecca E. Milton,[1] Georgina F. Osborne,[1] Edward J. Smith,[1] Samuel T. Jones,[1] Marie K. Bondulich,[1] Emily C. E. Danby,[1] Jemima Phillips,[1] Bridget A. Taxy,[1] Holly B. Kordasiewicz[2] and Gillian P. Bates[1]

1 Department of Neurodegenerative Disease, Huntington's Disease Centre and UK Dementia Research Institute at UCL, Queen Square Institute of Neurology, UCL, London, WC1N 3BG, UK

2 Ionis Pharmaceuticals, Carlsbad, CA 92008, USA

Correspondence to: Gillian Bates

Department of Neurodegenerative Disease, Queen Square Institute of Neurology, UCL, Queen Square House, Queen Square, London WC1N 3BG, UK

E-mail: gillian.bates@ucl.ac.uk

Running title: Human HTT1a transcript in YAC128 mice

Keywords: Huntington's disease; YAC128 mice; nuclear RNA clusters; *HTT1a* transcript; exon 1 HTT and aggregation

Abbreviations: 3'UTR = 3' untranslated region; ASO = antisense oligonucleotide; BSA = bovine serum albumin; ΔF = change in fluorescence; DAPI = 4', 6-diamidino-2-phenylindole; DMSO = dimethyl sulphoxide; FBS = foetal bovine serum; HRP = horse radish peroxidase; HTRF = homogeneous time-resolved fluorescence; MEF = mouse embryonic fibroblast; NTC = non-targeting control; polyA = polyadenylated; polyQ = polyglutamine; Q = glutamine; qPCR = real-time quantitative PCR; TSA = tyramide signal amplification; YAC = yeast artificial chromosome

Abstract

Huntington disease is caused by a CAG repeat expansion in exon 1 of the huntingtin gene (*HTT*) that is translated into a polyglutamine stretch in the huntingtin protein (HTT). We previously showed that *HTT* mRNA carrying an expanded CAG repeat was incompletely spliced to generate *HTT1a*, an exon 1 only transcript, which was translated to produce the highly aggregation-prone and pathogenic exon 1 HTT protein. This occurred in all knock-in mouse models of Huntington's disease and could be detected in patient cell lines and *post-mortem* brains. To extend these findings to a model system expressing human *HTT*, we took advantage of YAC128 mice that are transgenic for a yeast artificial chromosome carrying human *HTT* with an expanded CAG repeat.

We discovered that the *HTT1a* transcript could be detected throughout the brains of YAC128 mice. We implemented RNAscope to visualise *HTT* transcripts at the single molecule level and found that full-length *HTT* and *HTT1a* were retained together in large nuclear RNA clusters, as well as being present as single transcripts in the cytoplasm. Homogeneous time-resolved fluorescence analysis demonstrated that the *HTT1a* transcript had been translated to produce the exon 1 HTT protein. The levels of exon 1 HTT in YAC128 mice, correlated with HTT aggregation, supportive of the hypothesis that exon 1 HTT initiates the aggregation process.

Huntingtin-lowering strategies are a major focus of therapeutic development for Huntington's disease. These approaches often target full-length *HTT* alone and would not be expected to reduce pathogenic exon 1 HTT levels. We have established YAC128 mouse embryonic fibroblast lines and shown that, together with our QuantiGene multiplex assay, these provide an effective screening tool for agents that target *HTT* transcripts. The effects of current targeting strategies on nuclear RNA clusters are unknown, structures that may have a pathogenic role, or alternatively could be protective by retaining *HTT1a* in the nucleus and preventing it from being translated. In light of recently halted antisense oligonucleotide trials, it is vital that agents targeting *HTT1a* are developed, and that the effects of *HTT*-lowering strategies on the subcellular levels of all *HTT* transcripts and their various HTT protein isoforms are understood.

Introduction

Huntington's disease is an hereditary neurodegenerative disorder that manifests with movement disturbances, psychiatric changes and cognitive decline.[1] It is caused by an unstable CAG repeat expansion in exon 1 of the huntingtin gene (*HTT*), that is translated to produce an unusually long polyglutamine (polyQ) stretch in the huntingtin protein (HTT).[2] Mutant HTT self-associates to form aggregates that are deposited as inclusion bodies in patient brains[3], and the neuropathology of Huntington's disease is also marked by synaptic death and neuronal cell loss in the striatum, cortex and other brain regions.[4, 5] The treatments that are currently available for Huntington's disease focus on managing symptoms as a disease-modifying therapy to delay the onset or slow the progression of the disease does not exist.

The *HTT* gene contains 67 exons and encodes a 350 kDa protein.[2] Thus far, three full-length *HTT* mRNA isoforms have been described that are produced by alternative polyadenylation in the 3' untranslated region (3'UTR).[6] We have previously shown that, in the context of an expanded CAG repeat, two small transcripts that contain exon 1 and 5' intron 1 sequences (*HTT1a*) are produced by the incomplete splicing of the *HTT* transcript.[7, 8] This occurs when one of two cryptic polyadenylation (polyA) signals in intron 1 becomes activated. These cryptic polyA signals extend 2,710 bp and 7,327 bp into intron 1 for human *HTT* [7, 8] and 680 and 1,145 bp into intron 1 for mouse *Htt.*[7] The extent to which *HTT* is subjected to this alternative mRNA processing event is dictated by CAG repeat length, in that the longer the repeat the more *HTT1a* is generated.[7] Importantly, the *HTT1a* mRNA is translated to produce the highly pathogenic and aggregation-prone exon 1 HTT protein.[9, 10]

In the recent years various therapeutic strategies have been devised to combat Huntington's disease. While therapies aimed at restoring defective molecular pathways or clearing protein aggregates have received consideration[11-14], strategies to lower *HTT* mRNA are now a focal point for Huntington's disease therapeutics. Methods that reduce *HTT* mRNA levels and / or lead to translational suppression, thereby halting pathogenic protein production, include antisense oligonucleotides (ASOs), interfering RNAs (RNAi), zinc finger proteins (ZFPs) and small molecule mRNA splicing modulators.[15] ASO-based therapies aiming to reduce either full-length wild-type and mutant *HTT* mRNAs (tominersen, Roche) or specifically mutant *HTT* mRNA (Wave Life Sciences)[16] had progressed to phase III and phase I/II clinical trials, respectively. However, in 2021 both trials were halted as the drugs failed to either show higher efficacy over placebo, as was the case for tominersen, or to

significantly lower *HTT* mRNA levels, as reported for the allele-selective ASOs.[16] Although multiple factors could have contributed to the failure of these ASOs, it is important to highlight that both therapies targeted only full-length *HTT* transcripts, leaving the *HTT1a* mRNA intact.

Mouse models that express a mutant version of the entire human *HTT* gene in the form of a yeast artificial chromosome (YAC)[17] or bacterial artificial chromosome (BAC)[18] can be used to validate therapeutic approaches targeting human *HTT* transcripts.[19] YAC128 mice are transgenic for full-length human *HTT* with an expanded CAG repeat encoding 125 glutamines.[17] To extend our understanding of *HTT* expression in this mouse line, we have performed a detailed analysis of huntingtin at both the RNA and protein levels. Our findings indicate that the human *HTT* mRNA in YAC128 mice undergoes alternative processing throughout the brain to produce *HTT1a*. This small transcript was either retained in nuclear RNA clusters, where it co-localised with full-length *HTT*, or was exported to the cytoplasm and translated to produce the exon 1 HTT protein. In young mice, aggregated HTT levels were greatest in the cerebellum, the region containing the highest level of exon 1 HTT, consistent with the aggregation-prone nature of this mutant form of HTT. Both the formation of RNA clusters and the generation of the exon 1 HTT protein have unexplored implications for development of therapies for Huntington's disease.

Materials and methods

Ethics Statement

All procedures were performed in accordance with the Animals (Scientific Procedures) Act, 1996, and approved by the University College London Ethical Review Process Committee.

Animal colony breeding and maintenance

YAC128 and zQ175 mice were maintained on a C57BL6/J (Charles River, UK) background. Mice were group-housed depending on gender and genotypes were mixed within cages. All animals were kept in individually ventilated cages containing Aspen Chips 4 Premium bedding (Datesand) with environmental enrichment in the form of chew sticks and a play tunnel (Datesand). All mice had *ad libitum* access to water and chow (Teklad global 18% protein diet, Envigo, The Netherlands). The temperature was automatically regulated at 21°C ± 1°C and animals were kept on a 12 h light / dark cycle. The animal facility was barrier-maintained and quarterly non-sacrificial FELASA (Federation of European Laboratory

Animal Science Associations) screens found no evidence of pathogens. Animals were sacrificed at 2, 3, 5, 6, 9 and 12 months of age, brains rapidly dissected, frozen in liquid nitrogen and stored at −80 °C.

DNA extractions, genotyping and repeat sizing

Genomic DNA was extracted from ear notches[20] and genotyping was performed as previously described.[21] The polyQ repeat of 125 glutamines in YAC128 mice is encoded by $(CAG)_{23}(CAA)_3CAGCAA(CAG)_{80}(CAA)_3CAGCAA(CAG)_{10}CAACAG$ which is stable on germline transmission.[22] The zQ175 mice had CAG repeat expansions of 205.

Cell culture and ASO transfection

The procedures used for the isolation of mouse embryonic fibroblasts (MEFs), their transformation with the Simian Virus-40 large tumour antigen (SV40 T-antigen) and their transfection with ASOs is outlined in detail in the Supplementary Material.

RNA extraction and cDNA synthesis

Brain tissue was homogenised in Qiazol lysis reagent for 20-30 s. The homogenising probe was washed with 100% ethanol and ddH_2O between each sample. Total RNA was extracted from brain tissue and MEFs using the RNeasy mini-kit (Qiagen) following the manufacturer's recommended protocol. DNase I treatment was carried out using the RNA-free DNase kit (Qiagen), followed by RNA elution in the suitable amount of nuclease-free H_2O (Sigma-Aldrich). The quality and quantity of RNA was determined by NanoDrop 1000 (Thermo Fisher Scientific). cDNA synthesis was performed with 1 µg RNA using M-MLV reverse transcriptase (Invitrogen) and oligo-$dT_{(18)}$ primers (Invitrogen) according to the manufacture's recommendation. A negative control was always included, to which no reverse transcriptase was added (-RT).

Relative real-time quantitative PCR (qPCR)

Primers and probes (Taqman assays) used for the quantification of *HTT* transcripts were purchased from Eurofins, whereas mouse reference gene assays were from Thermo Fisher Scientific. The amplification efficiencies of primer sets were assessed by four-fold serial dilutions of a pool of equal volumes of cDNA from 6 YAC128 animals and are listed in

Supplementary Table 1. The amplification efficiencies of the qPCR assays were assessed by four-fold serial dilution of cDNA from the cortex of 2-month-old YAC128 animals. After determination of the linear range of the qPCR reaction, a 1:10 dilution of the cDNA was chosen for all brain regions as well as MEFs. Each 15 µL reaction consisted of 3 µL diluted cDNA, TaqMan Fast Advanced Mix (Applied Biosystems) as well as the appropriate gene expression assay, distributed into 96-well thin wall Hard-Shell PCR plates (BioRad). Plates were sealed with Microseal 'B' seals (BioRad), centrifuged at 800 x *g* for 30 s and run in a BioRad CFX96 qPCR machine as follows: 40 s at 95 °C, 40 x (7 s at 95 °C, 20 s at 60 °C (or 65°C for 3'UTR)) (Supplementary Table 1). Three technical replicates of each biological sample were used for all assays. C*q* values deviating by 0.5 from the mean were excluded from further analysis. Data for genes of interest were normalised to the geometric mean of the reference genes (*Sdha*, *Ubc* and *Atp5b*) and the $2^{-\Delta Ct2}$ was calculated.[23]

Tissue lysis and QuantiGene assays

Brain samples and cell lysates were prepared as described previously.[20, 24] To assess the linearity of the QuantiGene probe set, equal volumes of brain homogenates or cell lysates of the same genotype were pooled, and a two-fold serial dilution was prepared for each pool. Subsequently, the brain homogenates or cell lysates were diluted so that the fluorescent signal for all probe sets in the QuantiGene panel fell within the linear range. The QuantiGene plex was assayed in duplicate (brain tissue) or quadruplicates (cultured cells) as per the manufacturer's protocol with the exception that the Streptavidin *R*-Phycoerythrin conjugate (SAPE) was incubated at 51ºC. Plates were read on a Magpix (Luminex). After background subtraction, the median fluorescence intensity for the genes of interest was normalised to the geometric mean of the median fluorescent intensity for the reference genes (Supplementary Table 2).

RNAscope analysis and quantification

Procedures for RNAscope probe hybridisations are outlined in detail in the Supplementary Material.

Antibodies

All antibodies are listed in Supplementary Table 3.

Immunohistochemistry

Immunohistochemistry was performed as described in detail elsewhere. [25] The S830 primary antibody was applied at a 1:3000 dilution. Images were acquired with a Zeiss AxioSkop2 plus microscope fitted with a Zeiss AxioCam HRc colour camera using Zeiss AxioVision 4.7 software.

Immunoprecipitation and western blotting

Immunoprecipitations from wild-type and YAC128 aged cortical samples was performed using the anti-polyglutamine 3B5H10 antibody (Sigma-Aldrich) exactly as described.[7, 26] For western blotting, immunoprecipitations were denatured, separated by 10% SDS-Polyacrylamide Acrylamide Gel Electrophoresis (SDS-PAGE), blotted onto nitrocellulose membrane, and detected by chemiluminescent detection, exactly as described.[25, 26] Antibody dilutions were, anti-HTT S830 (sheep): 1:2000 and anti-HTT (rabbit) CHDI-90000148: 1:1000.

Homogeneous Time-Resolved FRET (HTRF)

HTRF was performed as previously described.[21] A detailed account of the optimisation and use of assays for YAC128 tissues and MEFs is provided in the Supplementary Material.

Behavioural and Phenotype Assessment

Weight gain, grip strength, rotarod performance and locomotor activity were measured in YAC128 males (n = 11) and females (n = 10) and their wild-type littermate males (n = 9) and females (n = 10) monthly from 2 to 12 months of age. Body weight was recorded weekly to the nearest 0.1 g. Rotarod performance and grip strength measures were performed as described elsewhere.[27] Exploratory activity in the open field was measured as described previously[28] and detailed below. Mice were individually placed in a white open field arena (50 cmx 50 cm x 50 cm, Engineering & Design Plastics Ltd., Cambridge, UK) for 30 min. Behaviour was videotaped via a camera placed above the apparatus. Activity (distance travelled, cm) was tracked and analysed using EthoVision 11.5 XT software (Noldus, Netherlands). The arena was cleaned using 70% industrial methylated sprits between trials. The order in which mice were assessed for behavioural measures was mixed for genotype

and gender and operators were blind to genotype.

Statistical analysis

Data were screened for outliers using the ROUT test (Q=10%; GraphPad Prism v8) and outliers were removed from the analysis. Statistical analysis was by one-way or two-way ANOVA with either Tukeys, Dunnet's or Bonferroni *post-hoc* tests. For behavioural analysis all data were screened for statistical outliers using ROUT test (Q = 10%; GraphPad Software v8, California, USA) and outliers were removed before between-group comparisons. Statistical analysis was performed with SPSS (v26) (IMB, Portsmouth, UK) using GLM ANOVA or ANCOVA (body weight as a covariant), with Bonferroni *post-hoc* tests. Graphs were prepared using Prism v8 (GraphPad Software, California, USA). *p*-values less than 0.05 were considered statistically significant.

Data availability

The authors confirm that all the data supporting the findings of this study are available within the article and its Supplementary material. Raw data will be shared by the corresponding author on request.

Results

HTT1a is generated in the brains of YAC128 mice through alternative RNA processing

We have previously shown that mouse *Htt* transcripts carrying expanded CAG repeats are alternatively processed to generate *Htt1a* in all Huntington's disease knock-in mouse models[7, 29] YAC128 mice express human *HTT* with an expanded CAG repeat and, using 3' RACE (rapid amplification of cDNA ends), we showed that cryptic polyA sites at 2710 bp and 7327 bp within intron 1 of human *HTT* in YAC128 brains had been activated, suggesting that human *HTT1a* was generated in this mouse model.[7] To extend these findings, we set out to track the expression profile of full-length *HTT* (*FL-HTT*) and *HTT1a* in the brains of YAC128 mice by real-time quantitative PCR (qPCR). To identify the *HTT1a* transcript we applied the qPCR assay ($PolyA_2$) that binds to intron 1 before the second cryptic poly(A) site at 7327 bp.[8] We designed an assay to the 3'UTR of *HTT* to quantify *FL-HTT* and one to intron 56 to detect any contaminating pre-processed mRNA (Fig. 1A and Supplementary

Table S1). We interrogated four brain regions: cortex, striatum, hippocampus and cerebellum from YAC128 mice at 2 and 12 months of age. The *HTT1a* transcript could be readily detected in all brain regions (Fig. 1B). Quantitation of *HTT* intron 56 demonstrated that there was minimal contamination with unprocessed mRNA (Fig. 1B). The expression profile of both *HTT* transcripts, *FL-HTT* and *HTT1a* showed no consistent age-dependent changes in any of the brain regions (Fig. 1B and C).

Next, we designed a multiplex QuantiGene panel that would allow us to detect all human *HTT* transcripts simultaneously without the need for RNA extraction or error-prone cDNA synthesis. The 15-plex QuantiGene assay consisted of probes to intron 1 of human *HTT* before the first cryptic poly(A) site at 2710 bp (Intron1polyA$_1$), to intron 1 before the second cryptic poly(A) site at 7327 bp (Intron1polyA$_1$A $_2$), to sequences at the 3′ end of intron 1 (Intron1_3′), within intron 3 (Intron 3) and within intron 56 (Intron 56), as well as to exons 43-46 and the 3'UTR for the fully-spliced coding sequence (Fig. 1D and Supplementary Table S2). The QuantiGene plex was used to determine the comparative levels of each of these *HTT* transcripts between the cortex, striatum, hippocampus and cerebellum of 2-month-old YAC128 mice. In agreement with our qPCR data, the *HTT1a* mRNA was detected in all brain regions, and the pre-mRNA intron 3 and intron 56 probe-sets gave very low background levels. (Fig. 1E-I). As these assays were not quantitative, the intensities obtained for the exonic and intronic *HTT* probe-sets, within a brain region, cannot be interpreted as relative expression levels. This is exemplified by the two probe-sets that detected *FL-HTT*, which gave different median fluorescent intensity values (Fig. 1J and K), reflecting variation in probe hybridisation efficiencies.

HTT1a is retained in nuclear RNA clusters in YAC128 brains

The qPCR and QuantiGene experiments demonstrated that the *HTT1a* transcript was present in YAC128 brains, however, these methods lacked detailed spatial and subcellular context. To visualise the location of the *HTT* transcripts, we took advantage of RNAscope technology, which allows the detection of a single RNA molecule with simultaneous background suppression. We designed probes to detect either human *HTT1a* (5'sequences of intron 1), *FL-HTT* (exons 14-61) or unprocessed *HTT* pre-mRNA (intron 66) (Fig. 2A).

Cortical and hippocampal sections from YAC128 mice at 2 months of age were hybridised with the three RNAscope probes. Single *FL-HTT* transcripts were detected in the cytoplasm, consistent with the processed mRNA having been exported from the nucleus for

translation (Fig. 2B and C). Surprisingly, the *FL-HTT* probe detected large RNA clusters in both hippocampal and cortical nuclei. The intron 1 probe identified a small number of cytoplasmic transcripts, as would be expected, given that the *HTT1a* transcript is translated to generate the exon 1 HTT protein.[7] However, the vast majority of *HTT1a* was in the nucleus and colocalised with *FL-HTT* mRNA in the RNA clusters (Fig. 2 B and C, Supplementary Fig. 1A and B). It is noteworthy that the *FL-HTT* mRNA in most of the RNA clusters, as well as single mRNAs present in the nucleus, appears to be spliced, as they were rarely detected by the intron 66 probe (Fig. 2B and C).

The specificity of the human *HTT* probes was confirmed, as they gave no signal when hybridised to brain sections from wild-type mice (Supplementary Fig. 2A). In contrast to human *HTT*, mouse wild-type *Htt* was detected as single mRNA molecules both in the nucleus and cytoplasm of hippocampal and cortical sections from YAC128 and wild-type mice (Fig. 3A and B, Supplementary Fig. 2B and Supplementary Fig. 4B). Occasionally, mouse *Htt* colocalised with the large human *HTT* clusters, suggesting these structures could potentially trap other mRNA molecules (Fig. 3B and C and Supplementary Fig 1C and D).

The nuclear retention of *HTT* transcripts was emphasised by the fact that most housekeeping mRNAs, including *Ubc, Ppib and Polar2a,* were primarily localised in the cytoplasm (Supplementary Fig. 3). To confirm that the *HTT* mRNA clusters were not an artefact of the RNAscope assay, but rather reflected a biological phenomenon, we designed and validated a probe that used a smaller number of ZZ-pairs for signal amplification (Supplementary Fig. 3A). This probe also recognised the *HTT* mRNA clusters trapped in nuclear structures (Supplementary Fig. 4B).

Next, we determined whether nuclear RNA clusters could also be detected in the brains of a knock-in mouse model of Huntington's disease. RNAscope was performed on hippocampal and cortical sections from zQ175 mice, in which mouse exon 1 *Htt* has been replaced with a mutant version of human exon 1 *HTT*. The probes used were to full-length mouse *Htt* (*FL-Htt* exons 60-67), *Htt1a* (5'sequences of mouse intron 1) or unprocessed *Htt* pre-mRNA (mouse intron 2) (Supplementary Fig. 5A). In contrast to the signals obtained in YAC128 mice, RNA clusters were not observed; both the *FL-Htt* and *Htt1a* probes detected single transcripts that were predominantly cytoplasmic (Supplementary Fig. 5B).

Greatest levels of soluble exon 1 HTT and HTT aggregation were detected in the YAC128 cerebellum

We have recently established a series of homogeneous time resolved fluorescence (HTRF) assays to track changes in soluble and aggregated HTT isoforms in Huntington's disease mouse tissues.[21] Therefore, we applied these bioassays to investigate how soluble and aggregated HTT levels might change with disease progression in brain regions from YAC128 mice. HTRF relies on the detection of a fluorescent signal that is emitted when two antibodies recognising specific HTT epitopes are in close proximity.[30] The donor and receptor antibodies for the selected HTRF assays, respectively, were as follows: 4C9 and MW8 for aggregated HTT; 2B7 and MW8 for soluble mutant exon 1 HTT; MW1 and MAB5490 for soluble mutant full-length HTT; MAB5490 and MAB2166 for total full-length HTT (mutant and wild-type) (Fig. 4A).

The 4C9-MW8 assay is a well-established bioassay for detecting levels of aggregated HTT.[21, 31] Statistically significant levels of aggregation could be detected in all four brain regions at 3 months of age, which in all cases increased with disease progression (Fig. 4B). The greatest level of aggregation was detected in the cerebellum, with lower levels in the striatum, cortex and hippocampus (Fig. 5A). Interestingly, the level of aggregation in the cerebellum at 3 months of age was already at a level not reached by 12 months in the striatum and cortex (Fig. 5B).

The 2B7-MW8 and MW1-MAB5490 assays allow the levels of soluble exon 1 HTT and soluble full-length mutant HTT to be determined. The 2B7-MW8 assay is specific for the exon 1 HTT protein, as MW8 acts as a neo-epitope antibody for the C-terminus of exon 1 HTT.[21, 26] Remarkably, at 2 months of age, the level of exon 1 HTT was greatest in the cerebellum, with comparably lower levels in the cortex, striatum and hippocampus (Fig. 5C). The level of exon 1 HTT decreased from 2 to 5 months of age in all brain regions, after which it remained relatively stable (Fig. 4C and Fig. 5D). The levels of soluble mutant full-length HTT (MW1-MAB5490) were stable over the course of the disease (Fig. 4D). Finally, total full-length HTT (MAB5490-MAB2166) levels were higher in YAC128 than wild-type mice, consistent with these mice having three copies of HTT (Fig. 4E).

Nuclear huntingtin aggregates are present before three months of age in YAC128 brain regions

We next performed immunohistochemistry to determine how the temporal and spatial appearance of HTT aggregates correlated with the progressive increase in HTT aggregation detected by HTRF. Coronal sections from YAC128 and wild-type mice at 3, 6, 9 and 12 months of age were immunostained with the S830 anti-HTT antibody, that readily detects aggregated HTT. At 3 months of age, aggregated HTT, in the form of a diffuse nuclear stain, was apparent in the outer layers of the cortex, the striatum and the dentate gyrus of the hippocampus, and this increased in intensity during the course of the disease (Fig. 6 and Supplementary Fig. 6A and C and Supplementary Fig. 7B). Whilst the diffuse nuclear stain was apparent in the inner layers of the cortex in 3-month-old YAC128 mice, it remained relatively sparse in these layers up to 12 months (Fig. 6 and Supplementary Fig. 6B). Interestingly, and in keeping with the HTRF data, the cerebellum displayed a relatively high intensity of S830 staining at 3 months, which was most evident in the granular layer, with the molecular layer becoming increasingly affected over the 12-month period (Fig. 6 and Supplementary Fig. 7C). This S830 diffuse nuclear immunostain was not detected in the CA1 subfield of the hippocampus until 9 months (Fig. 6 and Supplementary Fig. 7A).

The HTT aggregation in nuclei remained diffuse, and small nuclear inclusions were only rarely apparent, for example, in the striatum at 12 months of age (Fig. 6). High-power images revealed that cytoplasmic inclusions were present in the outer and inner cortical layers and the striatum by 6 months and the CA1 subfield of the hippocampus by 9 months, and these increased in density with age (Fig. 6). They were comparatively rare in the dentate gyrus and cerebellum (Fig. 6). Importantly, wild-type sections showed no detectable S830 staining at any age under investigation (Supplementary Fig. 8).

YAC128 mice gained weight but did not develop progressive motor phenotypic deficits

A range of behavioural phenotypes have been reported in YAC128 mice but with considerable variability between laboratories in the degree of impairment detected.[17, 32-34] Therefore, we set out to assess longitudinal behavioural phenotypes in YAC128 mice using our standard battery of tests. Progressive changes in body weight, grip strength, rotarod performance and activity measures were recorded in YAC128 males (n = 11) and females (n

= 10) and their wild-type littermate males (n = 9) and females (n = 10) monthly from 2 to 12 months of age.

Consistent with previous data[32, 33] both YAC128 males and females gained weight faster than their wild-type counterparts (Supplementary Fig. 9A). Body weight is known to influence rotarod performance and locomotor activity[35] , and therefore, we introduced weight as a covariate when analysing the grip strength, rotarod and activity data. There was no difference in the change in hind- and fore-limb grip strength between YAC128 and wild-type mice up to 12 months (Supplementary Fig. 9B). To assess changes in balance and co-ordination, mice were tested for their ability to walk on an accelerating rotarod. Rotarod performance changed little with age for both YAC128 and wild-type mice, and when weight differences were taken into consideration, there was no difference in the change in rotarod performance between YAC128 and wild-type mice over the 12-month period (Supplementary Fig 9C). Exploratory activity was assessed in an open field arena for a period of 30 min. The total distance travelled decreased for both YAC128 and wild-type mice between 2- and 12-months of age. When differences in weight were taken into consideration, there were no consistent differences in the total distance travelled between wild-type and YAC128 mice (Supplementary Fig. 9D).

YAC128 MEFs: a tool for screening human *HTT* and *HTT1a* lowering agents

We have previously shown that mouse embryonic fibroblasts (MEFs), derived from the zQ175 knock-in mice constitute a reliable cellular model to study incomplete splicing events.[20] To create a system that would allow screening of agents targeting human *HTT* transcripts, we generated transformed MEF lines from YAC128 mice (n = 3). We then used qPCR to show that both the *HTT1a* and *FL-HTT* (Fig. 7A) transcripts could be detected in the three YAC128 MEFs. Our RNAscope analysis had revealed that *FL-HTT* and *HTT1a* mRNAs were retained as nuclear RNA clusters in YAC128 brains (Fig. 2B, C). These clusters may have implications for the ability of potential therapeutic agents to target *HTT* transcripts and therefore, we used RNAscope to assess the intracellular distribution of *HTT* transcripts in the YAC128 MEFs. Consistent with our *in vivo* data, housekeeping mRNAs were predominantly cytoplasmic with some single transcripts present in the nucleus (Supplementary Fig. 10A). Similarly, mouse *FL-Htt* mRNA signals were mostly detected in the cytoplasm (Fig. 7C), in accordance with previously published data for primary mouse

fibroblasts.[36] Although the large RNA clusters detected in brain were not apparent, the human *FL-HTT* and *HTT1a* mRNAs were observed to co-localise in the nucleus (Fig. 7B). However, quantification found no evidence of increased levels of human *HTT1a* in the nucleus as compared to mouse *FL-Htt* (Fig. 7C). Consistent with our *in vivo* data, the nuclear *FL-HTT* transcripts were mostly spliced, as they were rarely detected with the intron 66 RNAscope probe (Supplementary Fig. 11A). Next, we used HTRF assays to demonstrate that the wild-type and mutant soluble HTT isoforms could be detected in the three YAC lines. Levels of total soluble mutant HTT and exon 1 HTT were present in the YAC128 and not the wild-type MEFs (Fig. 7D).

To determine whether the YAC128-MEFs could be used to assess the effect of *HTT*-lowering agents, we transfected one of the YAC128 MEF lines with 20 or 200 nM *HTT*-ASO, that targets exon 42 of the *FL-HTT* transcript, or with a non-targeting control (NTC) ASO. The *HTT*-ASO degrades the *HTT* mRNA by an RNase H1-dependent mechanism.[37] The level of human *FL-HTT* and mouse *FL-Htt* knock-down was assessed by qPCR. Transfection of 20 nM *HTT*-ASO resulted in a greater decrease ($p \leq 0.05$) in human *FL-HTT* (55% ± 3%) than mouse *FL-Htt* (35% ± 4%) with a 200 nM dose resulting in comparable further reductions in human *FL-HTT* (72% ± 7%) and mouse *FL-Htt* (72% ± 3 %) as compared to the NTC-ASO (Fig. 8A).

Finally, we sought to determine whether our QuantiGene multiplex panel could be used to assess the effects of *HTT*-lowering agents, directly from cell lysates. One of the YAC128 MEF lines (MEF-7) was transfected with 2 nM, 20 nM or 200 nM of the *HTT*-ASO or of the NTC ASO. Cells were lysed and the QuantiGene 15-plex panel was used to measure mouse *FL-Htt*, human *FL-HTT*, human *HTT* intronic sequences, and the levels of five mouse housekeeping genes (Supplementary Table 2). The comparative decrease in levels of mouse *FL-Htt* (Fig. 8B) and human *FL-HTT* (Fig. 8C) in response to treatment with the *HTT*-ASO was comparable to that determined by qPCR (Fig. 8A), and the level of *HTT1a* was unchanged (Fig. 8D).

Discussion

In the presence of an expanded CAG repeat, aberrant processing of the huntingtin mRNA generates the *HTT1a* transcript which encodes the aggregation-prone and pathogenic exon 1 HTT protein.[7] This occurs in *post-mortem* brains and fibroblast cultures from HD patients [8] and has been shown to correlate with disease progression and HTT aggregate deposition in

knock-in mouse models.[29] Here we demonstrated that human *HTT1a* is produced throughout the brain in YAC128 mice, that are transgenic for the entire human *HTT* gene. We found that both full-length *HTT* and *HTT1a* were retained in the nucleus as RNA clusters as well as being present in the cytoplasm as single transcripts. *HTT1a* encodes the pathogenic exon 1 HTT protein, which was present throughout the YAC128 brain, a level that correlated with HTT aggregation in specific brain regions. These results have important implications for the design of agents aimed at lowering the levels of pathogenic forms of huntingtin.

We used both qPCR and QuantiGene analysis to show that the mutant human *HTT* mRNA in YAC128 mice was alternatively processed to generate the *HTT1a* transcript and found that the human *FL-HTT* and *HTT1a* transcripts accumulated as nuclear RNA clusters in the brains of YAC128 mice. The retention of mRNAs in nuclear 'foci' has been described for a range of tandem tri-, tetra- and penta-nucleotide repeat disorders including myotonic dystrophy types 1 and 2, fragile X tremor ataxia syndrome, spinocerebellar types 8, 10 and 31 and Huntington's disease like-2.[38] RNA foci in myotonic dystrophy type 1 caused by an expanded CUG repeat in the 3'UTR of the *DMPK* mRNA[39], colocalised with muscleblind-like splicing regulator 1 (MBNL1)[40] and were located at the periphery of nuclear speckles.[39, 41] This phenomenon has been previously reported in Huntington's disease patient-derived fibroblasts as well as lymphoblasts [42], in which 'foci' also co-localised with MBNL1 and were associated with the splicing speckle marker SC35.[42] The similarities between these diseases, would suggest that the formation of RNA foci is a consequence of mRNA structures formed by the expanded tandem tri-, tetra- and penta-nucleotide repetitive sequences. However, we did not observe nuclear RNA clusters in the brains of zQ175 knock-in mice, that contain a chimeric protein between human exon 1 *HTT* and mouse *HTT*, and harbour expansions of approximately 200 CAGs. The disease-specific differences in the huntingtin transcripts between YAC128 and zQ175 are the intron 1 sequences in *HTT1a*, which in YAC128 mice are human, extending for 7327 bp, and in zQ175 are mouse, extending for 1145 bp. The secondary and tertiary mRNA structures of the human *HTT1a* mRNA may drive cluster formation, possibly through a phase separation process involving the sequence-specific gelation of RNAs.[43]

The formation of RNA 'foci' is widely regarded as being an RNA-toxicity related phenomenon, and in the case of myotonic dystrophy, MBNL1 depletion, through sequestration into RNA foci, causes the aberrant splicing of genes linked to a wide range of clinical symptoms.[44] However, in the case of Huntington's disease, these nuclear clusters might exert a protective effect, by retaining *HTT1a* in the nucleus and preventing it from

being translated to generate the highly aggregation-prone and pathogenic exon 1 HTT protein. We applied our array of HTRF assays[21] to monitor changes in the levels of soluble HTT isoforms and aggregated HTT in YAC128 mice over the course of the disease up to 12 months of age. At 2 months, the exon 1 HTT protein could be detected in all four brain regions studied: cortex, striatum, hippocampus and cerebellum (Fig. 5). The highest exon 1 HTT levels correlated with the greatest aggregation signal, supporting the hypothesis that the mutant exon 1 HTT protein initiates the aggregation process.

We performed immunohistochemistry to visualise the temporal-spatial appearance of aggregated HTT in YAC128 brains. We have previously shown that the diffuse /granular appearance of mutant HTT in cell nuclei, as detected by the S830 antibody, represents an aggregated form of HTT.[25, 26] Aggregation of HTT in the nucleus first appears as this diffuse / granular stain that fills the entire nucleus; and subsequently, as disease progresses, this may coalesce to form nuclear inclusions, as in the case of R6/2 mice with ~200 CAGs, or alternatively remain relatively diffuse over the entire course of the disease, as for R6/2 mice with 90 CAGs.[25] Our observation that aggregated HTT in the nucleus also first appears as a diffuse staining pattern in YAC128 mice is consistent with previous studies.[34, 45, 46] We found that aggregated HTT had been deposited in the outer layers of the cortex, the striatum, the dentate gyrus of the hippocampus as well as in the granular layer of the cerebellum by 3 months of age, in keeping with our HTRF data. This appearance of HTT aggregation is earlier than previously reported for YAC128 mice on a C57BL/6J background[45, 46], a discrepancy most likely due to experimental procedures. The widespread distribution of aggregated HTT is in keeping with Bayram-Weston 2012[46] who first detected aggregation in the amygdala, thalamus, cerebellum, hippocampus, piriform cortex and ventral striatum.[46] Our data are consistent with all previous studies, in that only small and very sparse nuclear inclusions could be detected by 12 months of age, if at all [17, 46-48] and that cytoplasmic inclusions were also comparatively sparse.[46] This pattern of aggregation is remarkably similar to that in R6/2 mice with 90 CAGs.[25] Interestingly, the reduction in soluble exon 1 HTT in R6/2 $(CAG)_{90}$ mice was also similar to that in YAC128, with a large reduction occurring at a young age after which levels were relatively stable.[25] If exon 1 HTT is driving aggregation in the YAC128 mice, the level of exon 1 HTT with 125 glutamines leads to a similar pattern of aggregation as occurs with the level of exon 1 HTT with 90 glutamines in R6/2 mice, but over a more protracted timescale.

Given that widespread HTT aggregation occurs before 3 months of age, we set out to determine when we could first detect motor-phenotypes in YAC128 using our standard

battery of behavioural and phenotypic tests. As previously reported, we found that YAC128 mice on a C57BL/6J background gained weight with age[32, 33], a consequence of having an extra copy of full-length HTT.[49] As body weight can modulate the performance in behavioural tasks[35], we used weight as a co-variate in our statistical analyses. Consistent with Menalled *et al*[32], who performed a detailed analysis of YAC128 mice on both an FVB/N and C57BL/6J background, we found no evidence for a rotarod impairment, reduction in grip strength or changes in open field behaviour up to 12 months of age. However, YAC128 mice on a C57BL/6J background have been shown to display deficits in other motor[50] and cognitive tacks.[50, 51]

We had previously shown that *HTT1a* can be readily detected in MEFs derived from zQ175 mice and that a QuantiGene multiplex assay could be used to screen approaches to lower full-length mouse *Htt* and *Htt1a* levels.[20] Therefore, we established SV40 T-antigen-transformed MEF lines from YAC128 embryos and showed that *HTT1a* could be detected by qPCR, although the large nuclear RNA clusters present in YAC128 brains were not apparent. We used a well-characterised ASO targeting full-length huntingtin to test the utility of the human *HTT* QuantiGene multiplex assay as a tool for screening cell lysates. This simultaneously allowed the detection of the human *FL-HTT*, *HTT1a* and mouse *FL-Htt* transcripts as well as intronic and reference gene controls. It provides a rapid means of comparing transcript levels between control samples and those that have been genetically modified, or treated with agents designed, to modulate huntingtin transcript levels. We found that whilst full-length mouse *Htt* and human *HTT* levels were decreased, the level of *HTT1a* remained unchanged. The YAC128 MEFs allow agents to be screened against the entire human gene, before further investigation in a system in which the effects on RNA clusters can be determined.

Our work demonstrates that human *HTT* undergoes incomplete splicing in YAC128 mice to produce the *HTT1a* transcript and that this mRNA is either retained in the nucleus in RNA clusters or exported to the cytoplasm and translated to produce the exon 1 HTT protein. Moreover, we show that *HTT1a* is readily detected in YAC128-derived MEFs and that oligonucleotide-based compounds targeting *HTT* are efficient in lowering *HTT* levels in this model system. This makes YAC128-derived MEFs particularly useful as a screening tool to evaluate experimental therapies directly targeting the human *HTT* transcripts and to study the molecular underpinnings of incomplete splicing in the context of the human *HTT* gene. Our data have profound implications for the development of *HTT* lowering therapies. It is essential that agents that lower the levels of *HTT1a* are developed. These agents would have

the advantage that they specifically target the exon 1 HTT protein, a form of mutant HTT that is known to be highly pathogenic, whilst leaving levels of the wild-type HTT unchanged. This is an important consideration as the degree to which wild-type HTT can be lowered is currently unknown.[52] It remains to be established whether RNA clusters act as mediators of RNA-induced toxicity or contrarily they detain abnormal mRNA and thus curb the production of otherwise noxious mutant huntingtin proteins. Whether HTT targeting approaches will be efficient in dissolving RNA clusters, and what the biological consequences of that might be, is an essential question that needs to be elucidated.

Funding

This work was supported by grants from the Wellcome Trust (200181/Z/15/B), CHDI Foundation and UK Dementia Research Institute, which receives its funding from Dementia Research Institute Ltd, funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK.

Competing interests

HBK is an employee of Ionis Pharmaceuticals. The authors report no other competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

[1] Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR, Nance M, Ross CA, Scahill RI, Wetzel R, Wild EJ, Tabrizi SJ: Huntington disease. Nat Rev Dis Primers 2015, 1:15005.

[2] The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 1993, 72:971-83.

[3] DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP, Aronin N: Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science 1997, 277:1990-3.

[4] Waldvogel HJ, Kim EH, Tippett LJ, Vonsattel JP, Faull RL: The Neuropathology of Huntington's Disease. Curr Top Behav Neurosci 2015, 22:33-80.

[5] Vonsattel JP, Myers RH, Stevens TJ, Ferrante RJ, Bird ED, Richardson EP, Jr.: Neuropathological classification of Huntington's disease. J Neuropathol Exp Neurol 1985, 44:559-77.

[6] Romo L, Ashar-Patel A, Pfister E, Aronin N: Alterations in mRNA 3' UTR Isoform Abundance Accompany Gene Expression Changes in Human Huntington's Disease Brains. Cell Rep 2017, 20:3057-70.

[7] Sathasivam K, Neueder A, Gipson TA, Landles C, Benjamin AC, Bondulich MK, Smith DL, Faull RL, Roos RA, Howland D, Detloff PJ, Housman DE, Bates GP: Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. Proc Natl Acad Sci U S A 2013, 110:2366-70.

[8] Neueder A, Landles C, Ghosh R, Howland D, Myers RH, Faull RLM, Tabrizi SJ, Bates GP: The pathogenic exon 1 HTT protein is produced by incomplete splicing in Huntington's disease patients. Sci Rep 2017, 7:1307.

[9] Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, Bates GP, Davies SW, Lehrach H, Wanker EE: Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell 1997, 90:549-58.

[10] Mangiarini L, Sathasivam K, Seller M, Cozens B, Harper A, Hetherington C, Lawton M, Trottier Y, Lehrach H, Davies SW, Bates GP: Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. Cell 1996, 87:493-506.

[11] Sassone J, Papadimitriou E, Thomaidou D: Regenerative Approaches in Huntington's

Disease: From Mechanistic Insights to Therapeutic Protocols. Front Neurosci 2018, 12:800.
[12] Rai SN, Singh BK, Rathore AS, Zahra W, Keswani C, Birla H, Singh SS, Dilnashin H, Singh SP: Quality Control in Huntington's Disease: a Therapeutic Target. Neurotox Res 2019, 36:612-26.
[13] Gil-Mohapel JM: Screening of therapeutic strategies for Huntington's disease in YAC128 transgenic mice. CNS Neurosci Ther 2012, 18:77-86.
[14] Tabrizi SJ, Flower MD, Ross CA, Wild EJ: Huntington disease: new insights into molecular pathogenesis and therapeutic opportunities. Nat Rev Neurol 2020, 16:529-46.
[15] Tabrizi SJ, Ghosh R, Leavitt BR: Huntingtin Lowering Strategies for Disease Modification in Huntington's Disease. Neuron 2019, 101:801-19.
[16] Kingwell K: Double setback for ASO trials in Huntington disease. Nat Rev Drug Discov 2021, 20:412-3.
[17] Slow EJ, van Raamsdonk J, Rogers D, Coleman SH, Graham RK, Deng Y, Oh R, Bissada N, Hossain SM, Yang YZ, Li XJ, Simpson EM, Gutekunst CA, Leavitt BR, Hayden MR: Selective striatal neuronal loss in a YAC128 mouse model of Huntington disease. Hum Mol Genet 2003, 12:1555-67.
[18] Gray M, Shirasaki DI, Cepeda C, Andre VM, Wilburn B, Lu XH, Tao J, Yamazaki I, Li SH, Sun YE, Li XJ, Levine MS, Yang XW: Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogenesis in BACHD mice. J Neurosci 2008, 28:6182-95.
[19] Kordasiewicz HB, Stanek LM, Wancewicz EV, Mazur C, McAlonis MM, Pytel KA, Artates JW, Weiss A, Cheng SH, Shihabuddin LS, Hung G, Bennett CF, Cleveland DW: Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. Neuron 2012, 74:1031-44.
[20] Mason MA, Gomez-Paredes C, Sathasivam K, Neueder A, Papadopoulou AS, Bates GP: Silencing Srsf6 does not modulate incomplete splicing of the huntingtin gene in Huntington's disease models. Sci Rep 2020, 10:14057.
[21] Landles C, Milton RE, Jean A, McLarnon S, McAteer SJ, Taxy BA, Osborne GF, Zhang C, Duan W, Howland D, Bates GP: Development of novel bioassays to detect soluble and aggregated Huntingtin proteins on three technology platforms. Brain Commun 2021, 3:fcaa231.
[22] Pouladi MA, Stanek LM, Xie Y, Franciosi S, Southwell AL, Deng Y, Butland S, Zhang W, Cheng SH, Shihabuddin LS, Hayden MR: Marked differences in neurochemistry and aggregates despite similar behavioural and neuropathological features of Huntington disease

in the full-length BACHD and YAC128 mice. Hum Mol Genet 2012, 21:2219-32.
[23] Livak KJ, Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001, 25:402-8.
[24] Papadopoulou AS, Gomez-Paredes C, Mason MA, Taxy BA, Howland D, Bates GP: Extensive Expression Analysis of Htt Transcripts in Brain Regions from the zQ175 HD Mouse Model Using a QuantiGene Multiplex Assay. Sci Rep 2019, 9:16137.
[25] Landles C, Milton RE, Ali N, Flomen R, Flower M, Schindler F, Gomez-Paredes C, Bondulich MK, Osborne GF, Goodwin D, Salsbury G, Benn CL, Sathasivam K, Smith EJ, Tabrizi SJ, Wanker EE, Bates GP: Subcellular Localization And Formation Of Huntingtin Aggregates Correlates With Symptom Onset And Progression In A Huntington'S Disease Model. Brain Commun 2020, 2:fcaa066.
[26] Landles C, Sathasivam K, Weiss A, Woodman B, Moffitt H, Finkbeiner S, Sun B, Gafni J, Ellerby LM, Trottier Y, Richards WG, Osmand A, Paganetti P, Bates GP: Proteolysis of mutant huntingtin produces an exon 1 fragment that accumulates as an aggregated protein in neuronal nuclei in Huntington disease. J Biol Chem 2010, 285:8808-23.
[27] Bondulich MK, Fan Y, Song Y, Giorgini F, Bates GP: Ablation of kynurenine 3-monooxygenase rescues plasma inflammatory cytokine levels in the R6/2 mouse model of Huntington's disease. Sci Rep 2021, 11:5484.
[28] Pido-Lopez J, Tanudjojo B, Farag S, Bondulich MK, Andre R, Tabrizi SJ, Bates GP: Inhibition of tumour necrosis factor alpha in the R6/2 mouse model of Huntington's disease by etanercept treatment. Sci Rep 2019, 9:7202.
[29] Franich NR, Hickey MA, Zhu C, Osborne GF, Ali N, Chu T, Bove NH, Lemesre V, Lerner RP, Zeitlin SO, Howland D, Neueder A, Landles C, Bates GP, Chesselet MF: Phenotype onset in Huntington's disease knock-in mice is correlated with the incomplete splicing of the mutant huntingtin gene. J Neurosci Res 2019, 97:1590-605.
[30] Weiss A, Abramowski D, Bibel M, Bodner R, Chopra V, DiFiglia M, Fox J, Kegel K, Klein C, Grueninger S, Hersch S, Housman D, Regulier E, Rosas HD, Stefani M, Zeitlin S, Bilbe G, Paganetti P: Single-step detection of mutant huntingtin in animal and human tissues: A bioassay for Huntington's disease. Analyt Biochem 2009, 395:8-15.
[31] Reindl W, Baldo B, Schulz J, Janack I, Lindner I, Kleinschmidt M, Sedaghat Y, Thiede C, Tillack K, Schmidt C, Cardaun I, Schwagarus T, Herrmann F, Hotze M, Osborne GF, Herrmann S, Weiss A, Zerbinatti C, Bates GP, Bard J, Munoz-Sanjuan I, Macdonald D: Meso scale discovery-based assays for the detection of aggregated huntingtin. PLoS One 2019, 14:e0213521.

[32] Menalled L, El-Khodor BF, Patry M, Suarez-Farinas M, Orenstein SJ, Zahasky B, Leahy C, Wheeler V, Yang XW, MacDonald M, Morton AJ, Bates G, Leeds J, Park L, Howland D, Signer E, Tobin A, Brunner D: Systematic behavioral evaluation of Huntington's disease transgenic and knock-in mouse models. Neurobiol Dis 2009, 35:319-36.

[33] Brooks S, Higgs G, Janghra N, Jones L, Dunnett SB: Longitudinal analysis of the behavioural phenotype in YAC128 (C57BL/6J) Huntington's disease transgenic mice. Brain Res Bull 2012, 88:113-20.

[34] Van Raamsdonk JM, Pearson J, Slow EJ, Hossain SM, Leavitt BR, Hayden MR: Cognitive dysfunction precedes neuropathology and motor abnormalities in the YAC128 mouse model of Huntington's disease. J Neurosci 2005, 25:4169-80.

[35] Kudwa AE, Menalled LB, Oakeshott S, Murphy C, Mushlin R, Fitzpatrick J, Miller SF, McConnell K, Port R, Torello J, Howland D, Ramboz S, Brunner D: Increased Body Weight of the BAC HD Transgenic Mouse Model of Huntington's Disease Accounts for Some but Not All of the Observed HD-like Motor Deficits. PLoS Curr 2013, 5.

[36] Didiot MC, Ferguson CM, Ly S, Coles AH, Smith AO, Bicknell AA, Hall LM, Sapp E, Echeverria D, Pai AA, DiFiglia M, Moore MJ, Hayward LJ, Aronin N, Khvorova A: Nuclear Localization of Huntingtin mRNA Is Specific to Cells of Neuronal Origin. Cell Rep 2018, 24:2553-60 e5.

[37] Bennett CF, Kordasiewicz HB, Cleveland DW: Antisense Drugs Make Sense for Neurological Diseases. Annu Rev Pharmacol Toxicol 2021, 61:831-52.

[38] Wojciechowska M, Krzyzosiak WJ: Cellular toxicity of expanded RNA repeats: focus on RNA foci. Hum Mol Genet 2011, 20:3811-21.

[39] Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH: Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. J Cell Biol 1995, 128:995-1002.

[40] Miller JW, Urbinati CR, Teng-Umnuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS: Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. EMBO J 2000, 19:4439-48.

[41] Smith KP, Byron M, Johnson C, Xing Y, Lawrence JB: Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type I is blocked at entry into SC-35 domains. J Cell Biol 2007, 178:951-64.

[42] Urbanek MO, Jazurek M, Switonski PM, Figura G, Krzyzosiak WJ: Nuclear speckles are detention centers for transcripts containing expanded CAG repeats. Biochim Biophys Acta 2016, 1862:1513-20.

[43] Jain A, Vale RD: RNA phase transitions in repeat expansion disorders. Nature 2017, 546:243-7.
[44] Osborne RJ, Thornton CA: RNA-dominant diseases. Hum Mol Genet 2006, 15 Spec No 2:R162-9.
[45] Van Raamsdonk JM, Metzler M, Slow E, Pearson J, Schwab C, Carroll J, Graham RK, Leavitt BR, Hayden MR: Phenotypic abnormalities in the YAC128 mouse model of Huntington disease are penetrant on multiple genetic backgrounds and modulated by strain. Neurobiol Dis 2007, 26:189-200.
[46] Bayram-Weston Z, Jones L, Dunnett SB, Brooks SP: Light and electron microscopic characterization of the evolution of cellular pathology in YAC128 Huntington's disease transgenic mice. Brain Res Bull 2012, 88:137-47.
[47] Van Raamsdonk JM, Murphy Z, Slow EJ, Leavitt BR, Hayden MR: Selective degeneration and nuclear localization of mutant huntingtin in the YAC128 mouse model of Huntington disease. Hum Mol Genet 2005, 14:3823-35.
[48] Wang CE, Tydlacka S, Orr AL, Yang SH, Graham RK, Hayden MR, Li S, Chan AW, Li XJ: Accumulation of N-terminal mutant huntingtin in mouse and monkey models implicated as a pathogenic mechanism in Huntington's disease. Hum Mol Genet 2008, 17:2738-51.
[49] Van Raamsdonk JM, Gibson WT, Pearson J, Murphy Z, Lu G, Leavitt BR, Hayden MR: Body weight is modulated by levels of full-length huntingtin. Hum Mol Genet 2006, 15:1513-23.
[50] Brooks SP, Jones L, Dunnett SB: Longitudinal analyses of operant performance on the serial implicit learning task (SILT) in the YAC128 Huntington's disease mouse line. Brain Res Bull 2012, 88:130-6.
[51] Brooks SP, Janghra N, Higgs GV, Bayram-Weston Z, Heuer A, Jones L, Dunnett SB: Selective cognitive impairment in the YAC128 Huntington's disease mouse. Brain Res Bull 2012, 88:121-9.
[52] Becanovic K, Norremolle A, Neal SJ, Kay C, Collins JA, Arenillas D, Lilja T, Gaudenzi G, Manoharan S, Doty CN, Beck J, Lahiri N, Portales-Casamar E, Warby SC, Connolly C, De Souza RA, Network RIotEHsD, Tabrizi SJ, Hermanson O, Langbehn DR, Hayden MR, Wasserman WW, Leavitt BR: A SNP in the HTT promoter alters NF-kappaB binding and is a bidirectional genetic modifier of Huntington disease. Nat Neurosci 2015, 18:807-16.

Figure Legends

Figure 1 qPCR and QuantiGene analysis of full-length *HTT* and *HTT1a* transcripts in YAC128 brains. (A) Schematic of the location of the qPCR assays on the human *HTT* transcript. **(B)** *HTT1a* was detected in the cortex, striatum, hippocampus and cerebellum of 2- and 12-month-old YAC128 mice (n = 4-6/genotype) **(C)** Full-length *HTT* levels in the cortex, striatum, hippocampus and cerebellum as measured using the *HTT*_ 3'UTR assay (n = 4-6/genotype). **(D)** Schematic of the location of the QuantiGene probe sets on the human *HTT* transcript. **(E-I)** Comparison of *HTT* intronic sequences between brain regions. *HTT1a* was detected by QuantiGene in the cortex, striatum, hippocampus, and cerebellum of 2-month-old YAC128 mice (n = 4-6/genotype). **(J-K)** Full-length *HTT* detected by QuantiGene in the anterior and posterior cortex as well as in the striatum, hippocampus, and cerebellum of 2-month-old YAC128 mice (n = 4-6/genotype). Two different QuantiGene probe sets targeting exons 43-46 or the 3'UTR were designed to detect FL-*HTT*. The results are plotted on separate graphs as they render different signals due to the assay design.

Figure 2 Human *HTT* and *HTT1a* transcripts form nuclear RNA clusters in YAC128 brains. (A) Schematic showing the location of the RNAscope probes on the human *HTT* transcript: full-length *HTT* (exons 14-61), *HTT1a* (*HTTintron1*) and pre-processed mRNA (*HTTintron66*). **(B)** Full-length *HTT* (yellow) and *HTT1a* (magenta) were frequently colocalised in nuclear RNA clusters in the hippocampus and cortex of YAC128 mice at 2 months of age (white arrowheads). Full-length *HTT* was also detected in the extranuclear space as single transcripts which did not colocalise with *HTT1a* (orange arrowheads). The *HTT1a* transcripts were predominantly detected in nuclear clusters. The *HTT intron 66* (green) probe visualised non-spliced pre-mRNA which, although sparse was present in both the nucleus and cytoplasm and did not co-localise with FL-*HTT* or *HTT1a*. **(C)** Intensity profiles of FL-*HTT*, *HTT1a* and *HTTintron66* signals. Peak intensities were recorded along the white dashed line shown in the merged image in panel (B). The orange arrowheads indicate full-length human *HTT* outside of the nucleus, white arrowheads indicate human *HTT* and *HTT1a* that are colocalised in the nucleus. A statistical analysis of the transcript subcellular location is presented in Supplementary Fig. 1A and B. Nuclei were stained with DAPI (blue). The wild-type control sections at 2 months of age are shown in Supplementary Fig. 3A and B. YAC128 (n = 4). Scale bar is 20 μm in the main image and 5 μm in the

cropped magnified image.

Figure 3 Mouse *Htt* transcripts were present in the nucleus and cytoplasm, whereas human *HTT* mRNAs were retained in nuclei in YAC128 brains. (A) Schematic location of the full-length human (exons 14-61) and mouse (exons 60-67) huntingtin RNAscope probes. **(B)** Full-length human *HTT* (magenta) was present in large nuclear clusters (yellow arrowheads) and as single transcripts in the extranuclear space (orange arrowheads). Full-length mouse *Htt* (green) was predominantly detected outside of the nucleus (grey arrowheads). Human and mouse huntingtin transcripts only rarely colocalised in both the nucleus (white arrowheads) and cytoplasm (green arrowheads). **(C)** Intensity profiles of full-length mouse and human huntingtin transcript signals. Peak intensities were recorded along the white dashed line shown in the merged image in panel (B) Yellow arrowheads indicate large nuclear clusters containing human *HTT*, orange arrowheads indicate human *HTT* transcripts outside of the nucleus, grey arrowheads indicate mouse *Htt* transcripts outside of the nucleus, white arrowheads indicate colocalisation of human *HTT* and mouse *Htt* inside the nucleus and green arrowheads indicate colocalisation of human *HTT* and mouse *Htt* in the cytoplasm. A statistical analysis of the transcript subcellular location is presented in Supplementary Fig. 1C and D. Nuclei were stained with DAPI (blue). The wild-type control sections from mice at 2 months of age are shown in Supplementary Fig 3A and B. YAC128 (n = 4). Scale bar is 20 μm in the main image and 5 μm is the cropped magnified image.

Figure 4 HTRF assays detect greater levels of soluble exon 1 HTT and HTT aggregation in the cerebellum than in forebrain regions. (A) Schematic indicating the position of the HTT epitopes detected by the antibodies used in the HTRF assays (Supplementary Table 3). S830 is a polyclonal sheep antibody that was raised against exon 1 HTT with 53Q. **(B)** HTT aggregation, as detected by the 4C9-MW8 assay increased from 3 to 12 months of age in all brain regions. **(C)** Exon 1 mutant HTT (2B7-MW8) levels decreased between 2 and 5 months of age in all brain regions and then remained relatively stable. **(D)** Full-length mutant HTT (MW1-MAB5490) remained relatively stable up to 12 months of age in all brain regions. **(E)** The levels of total full-length HTT (MAB5490-MAB2166) were higher in the YAC128 mice, which contain three copies of the huntingtin gene. Samples for any given HTT assay were run on the same plate, and therefore the levels between brain regions can be compared. N = 6. Statistical analysis was one-way or two-way ANOVA with Tukey's *post-hoc* correction. Error bars = mean ± SEM *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001. M = months, WT = wild type,

aa = amino acid.

Figure 5 Comparison of aggregated and soluble HTT isoforms between brain regions and with disease progression. (A) HTT aggregation levels were greatest in the cerebellum at 3 months of age. **(B)** HTT aggregation levels were greater in the cerebellum at 3 months of age than in the cortex, striatum or hippocampus at 12 months. **(C)** Soluble exon 1 HTT levels were highest in the cerebellum. (n= 6). (**D**) Huntingtin was immunoprecipitated with 3B5H10 from YAC128 cortical lysates at 2, 5, 9 and 12 months of age and wild-type lysates at 2 months. Western blots were immunoprobed with the S830 or CHDI-90000148 antibodies. Dotted lines indicate the location of the exon 1 HTT protein. Error bars = mean ± SEM. Statistical analysis was one-way ANOVA with Tukey's *post-hoc* correction. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Figure 6 Spatiotemporal appearance of HTT aggregation in YAC128 mice. Coronal sections from 3-, 6-, 9- and 12-month-old YAC128 mice were immunostained with the S830 antibody to visualise HTT aggregates. A diffuse nuclear S830 immunostain, indicative of HTT aggregation, was readily identified in the striatum, outer cortical layers, dentate gyrus of the hippocampus and granular layer of the cerebellum at 3 months of age, but not in the CA1 region of the hippocampus until 9 months. This nuclear immunostain increased in intensity with disease progression and small nuclear inclusions were observed only rarely (see zoomed sections). By 6 months of age, extranuclear inclusions were readily apparent in the outer and inner layers of the cortex and the striatum and in the CA1 region of the hippocampus by 9 months. The location of the images from within the brain sections is illustrated in Supplementary Figs. 6 and 7. The wild-type control sections from mice at 12 months of age are shown in Supplementary Fig. 8. YAC128 (n = 3), WT (n = 1). Scale bar = 20 μm.

Figure 7 *HTT1a* is present in MEFs from YAC128 mice. qPCR analysis revealed that **(A)** *HTT1a* (*HTT* PolyA2) *FL-HTT* (FL-*HTT*3'UTR) were present in the YAC128 MEFs (n = 3 biological replicates). **(B)** Percentage of nuclear and cytoplastic localisation of different huntingtin transcripts. One-way ANOVA of percentage of transcripts in the nucleus, full-length human *HTT* compared to *HTT1a* (p = 0.002), full-length human *HTT* compared to mouse *Htt* (p = 0.02). **(C)** RNAscope on YAC128 MEFs revealed that wild-type mouse *FL-Htt* (green), human *FL-HTT* (yellow) and *HTT1a* (magenta) were present as single transcripts in the cytoplasm. Human *FL-HTT* and *HTT1a* were more likely to be found in the nucleus that mouse *FL-Htt*, where they frequently colocalised (n = 3 biological replicates). Scale bar

is 20 μm in the main image and 5 μm in the magnified image. (**D**) Total soluble mutant HTT (2B7-MW1) and exon 1 HTT (2B7-MW8) could be detected in the YAC128 MEF lines (n = 3 biological replicates) by HTRF and not in wild-type lines (n = 2 biological replicates). YAC128 line 7.1 showed the highest huntingtin expression at both the RNA and protein levels. Student's *t*-test, *p ≤ 0.05. Error bars = mean ± SEM. WT = wild type. HK= housekeeping genes.

Figure 8 YAC128 MEFs provide a tool for screening agents designed to lower *HTT* transcript levels. (A) YAC128 MEFs were transfected with ASOs targeting the human *FL-HTT* and mouse *FL-Htt* transcripts or with a non-targeting control (NTC). qPCR analysis showed that 20 nM ASO was more efficient in lowering human *FL-HTT* levels than mouse *FL-Htt*, whereas 200 nM ASO decreased *FL-HTT* and *FL-Htt* to the same extent, 48 h post transfection. N = 3 biological replicates / genotype. Error bars = mean ± SEM. Statistical analysis was two-way ANOVA with Tukey's *post-hoc* correction for the effect on huntingtin lowering, and Bonferroni *post-hoc* correction for the effect between human and mouse huntingtin, **(B-F)** YAC128 MEFs were transfected with either PBS, the *FL-HTT* / *FL-Htt* targeting ASO or non-targeting control (NTC). The 15-plex QuantiGene assays was used to measure the levels of **(B)** mouse *FL-Htt*, **(C)** human *FL-HTT*, **(D)** human *HTT1a* (human intron 1_polyA1), human intron 1 3' and **(F)** intron 3. N=3-4 technical replicates of YAC128 MEF-7.1 Error bars = mean ± SEM. Statistical analysis was one-way ANOVA with Dunnet's *post hoc* correction *p = 0.014, ***p = 0.001.

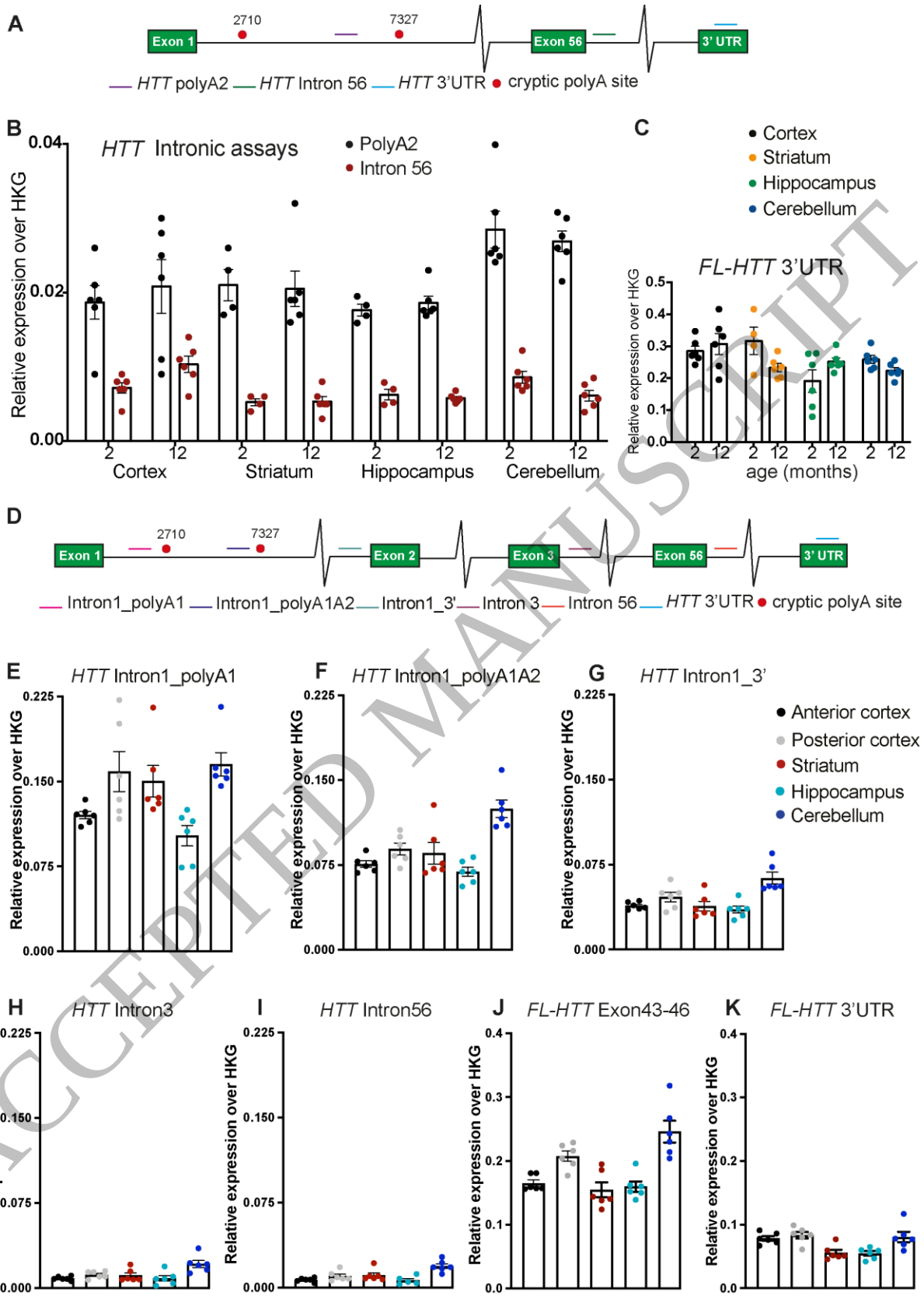

Figure 1
159x220 mm (3.4 x DPI)

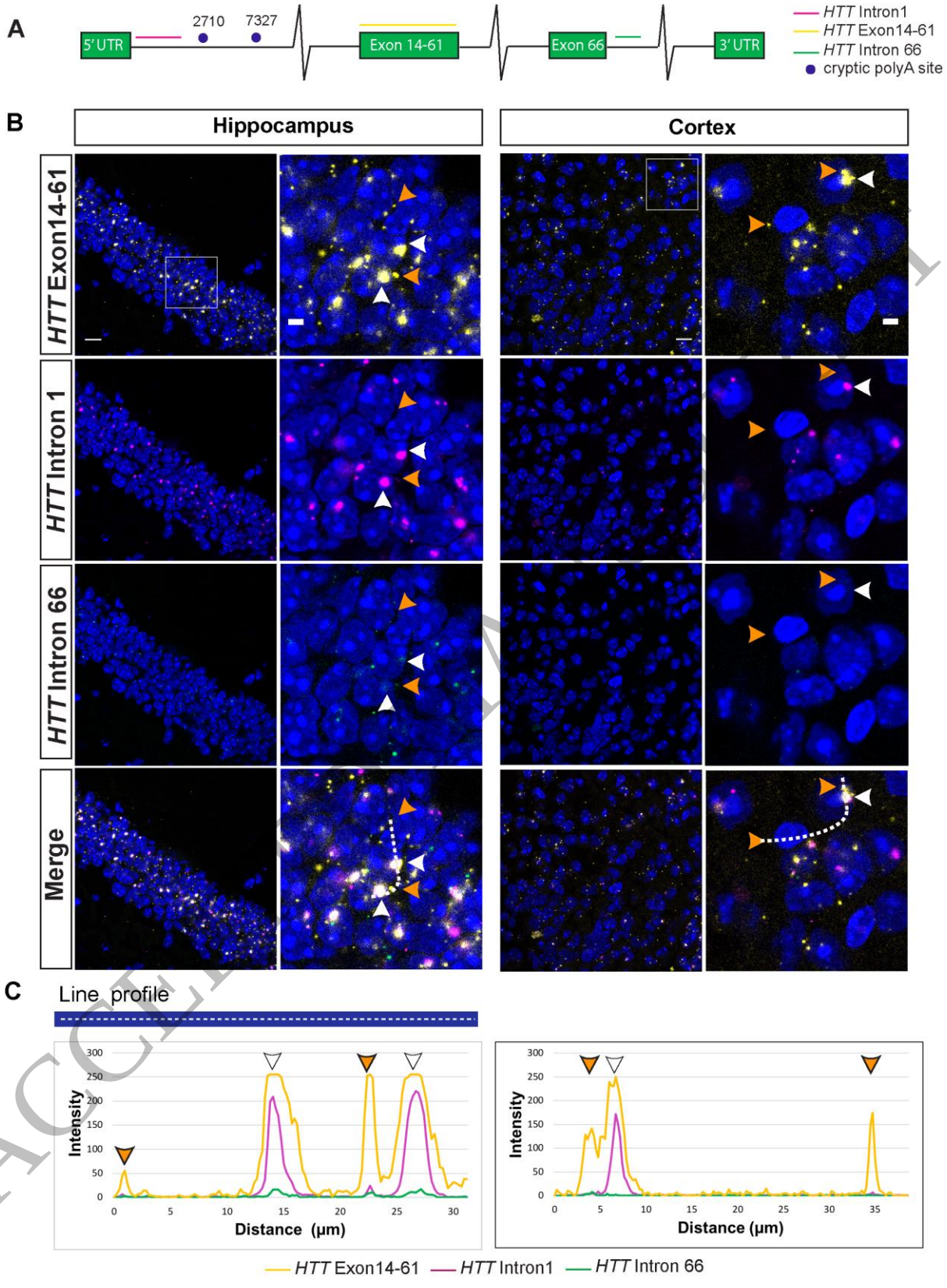

Figure 2
159x216 mm (3.4 x DPI)

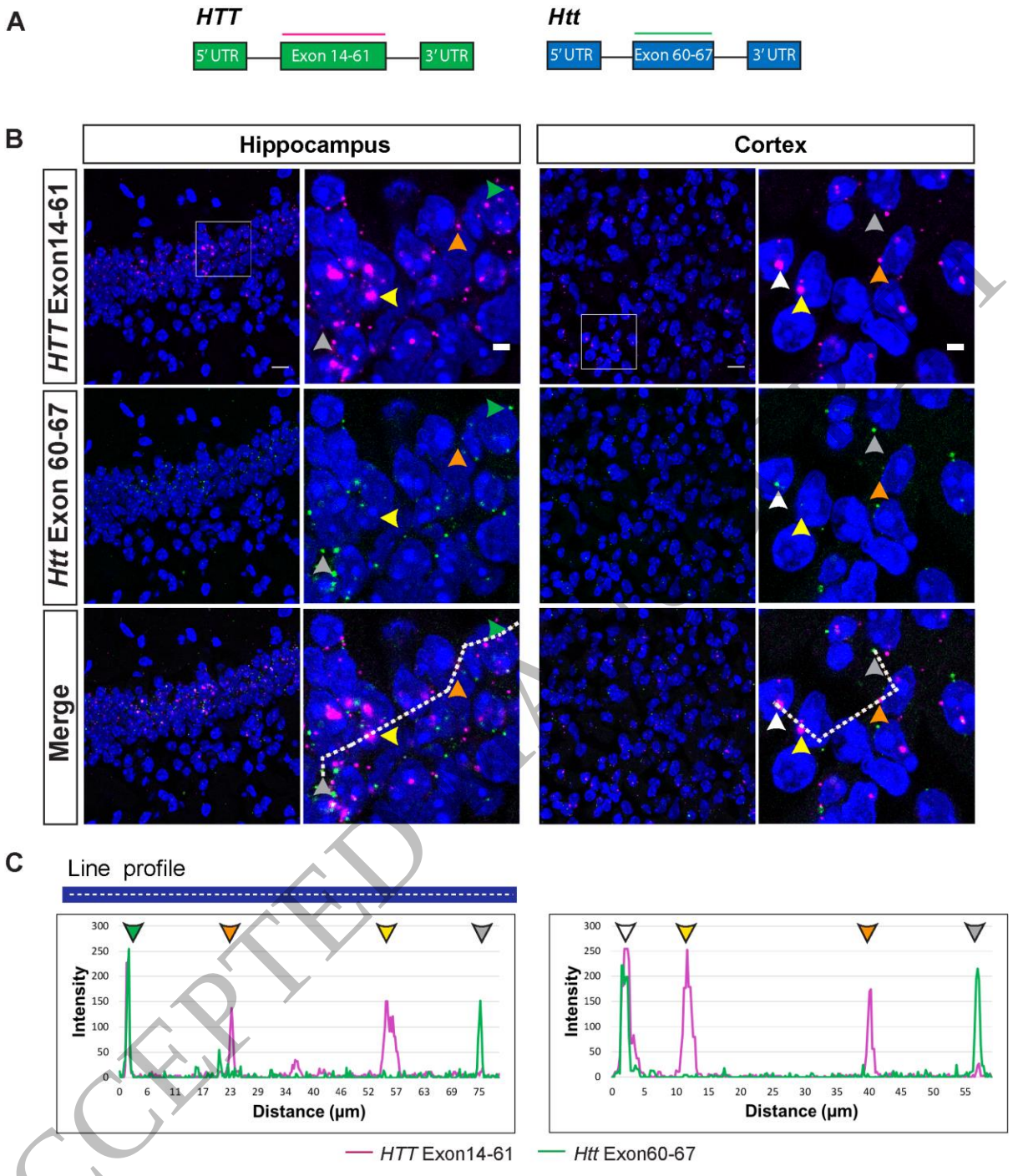

Figure 3
159x184 mm (3.4 x DPI)

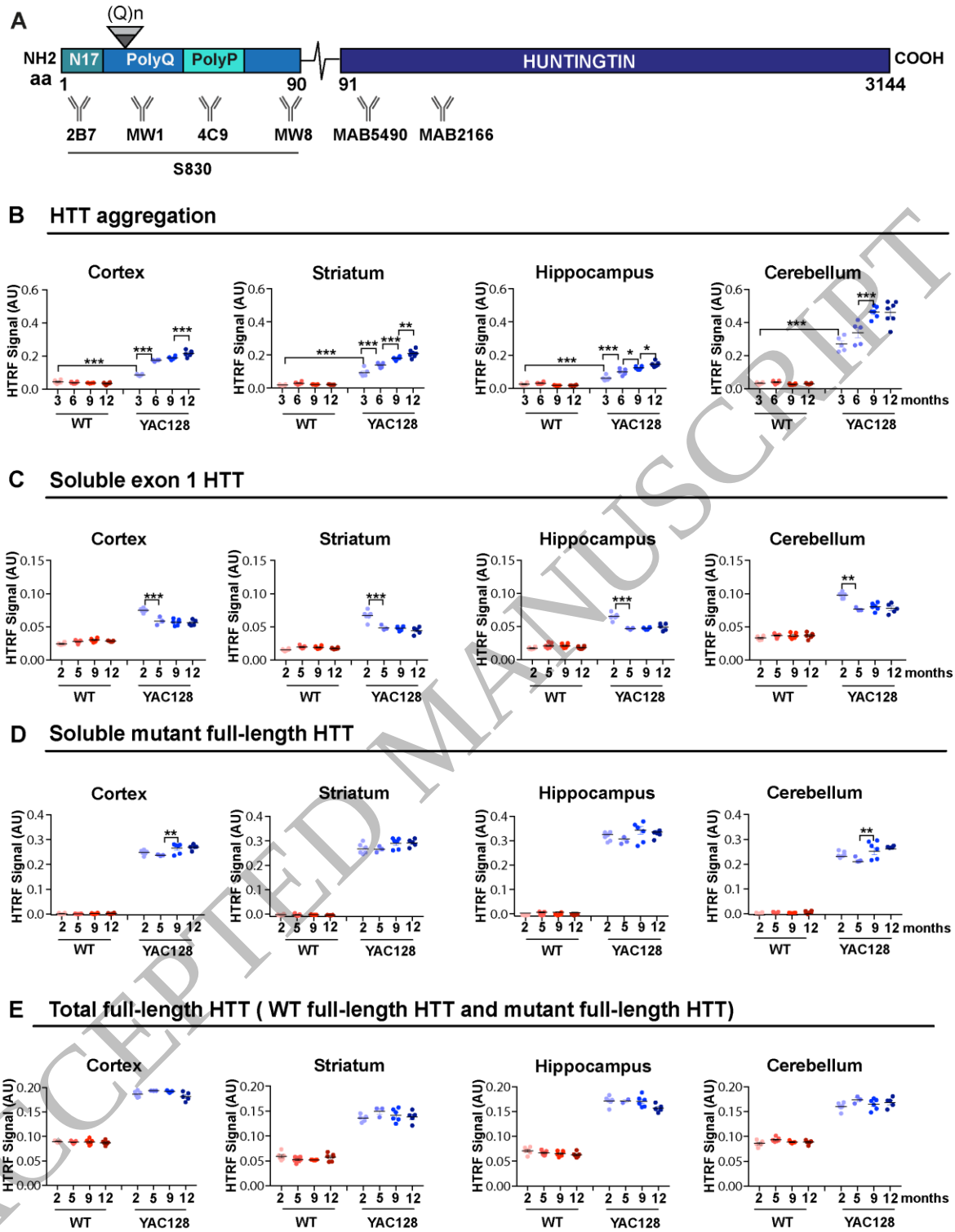

Figure 4
159x201 mm (3.4 x DPI)

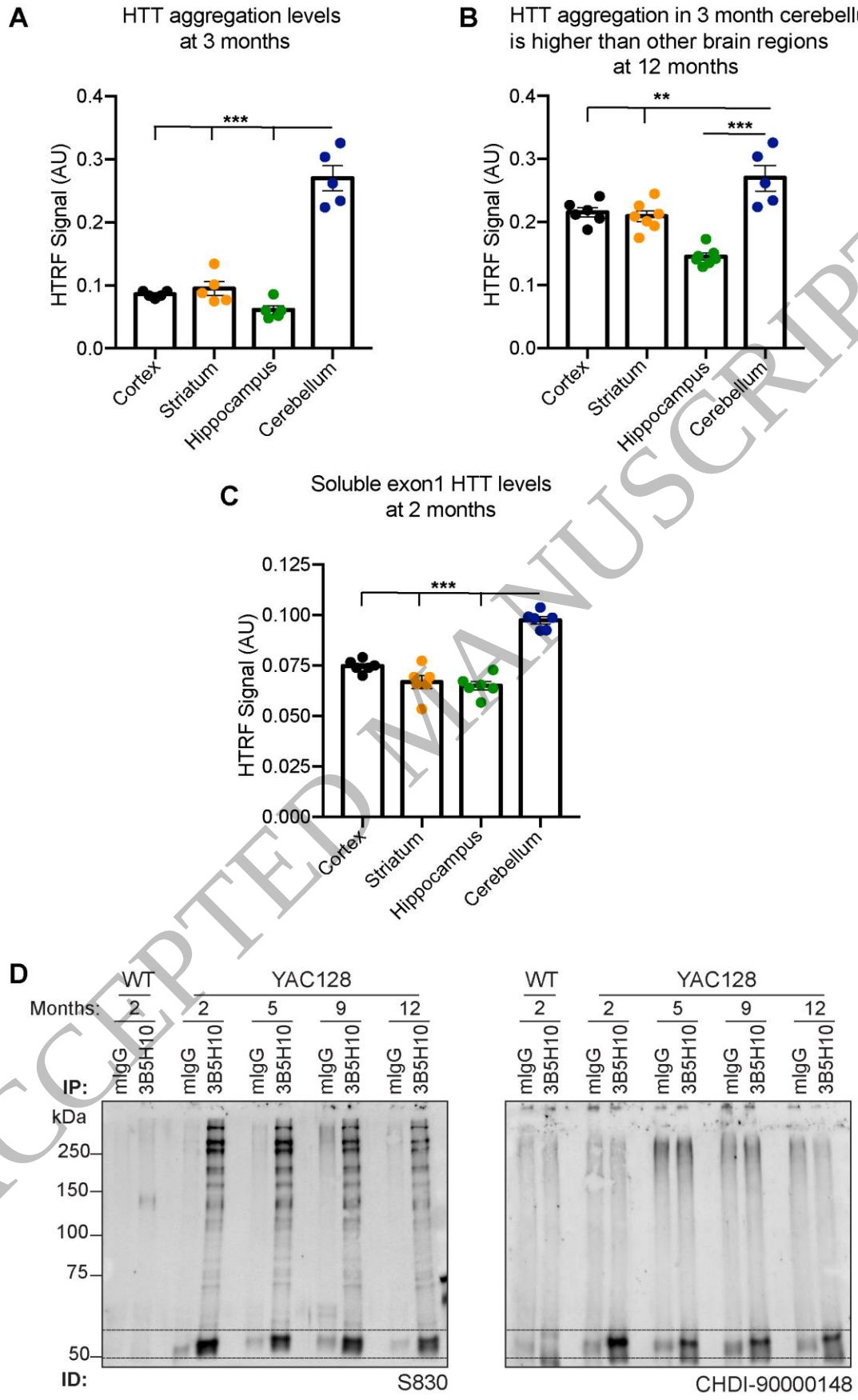

Figure 5
153x246 mm (3.4 x DPI)

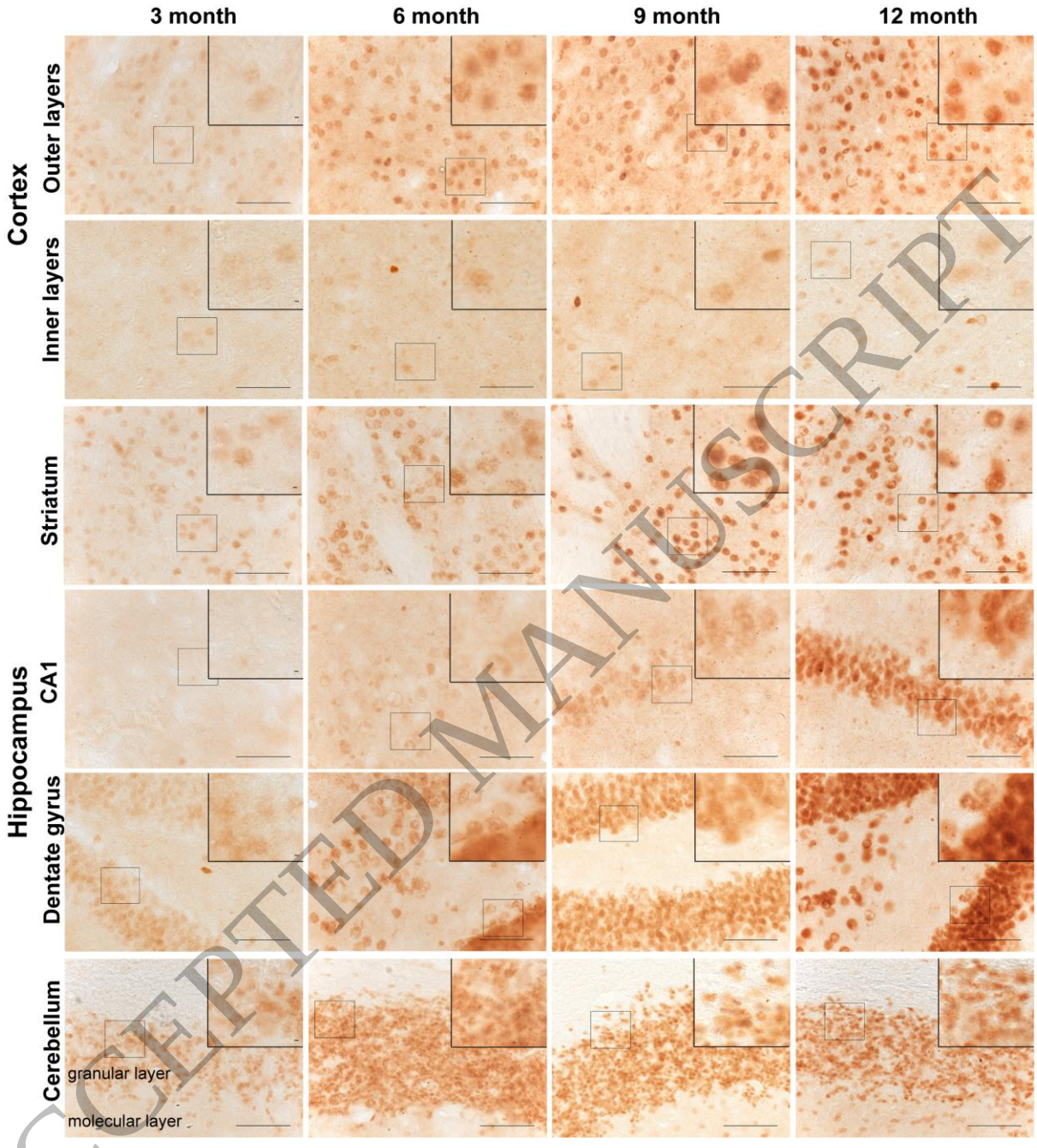

Figure 6
159x175 mm (3.4 x DPI)

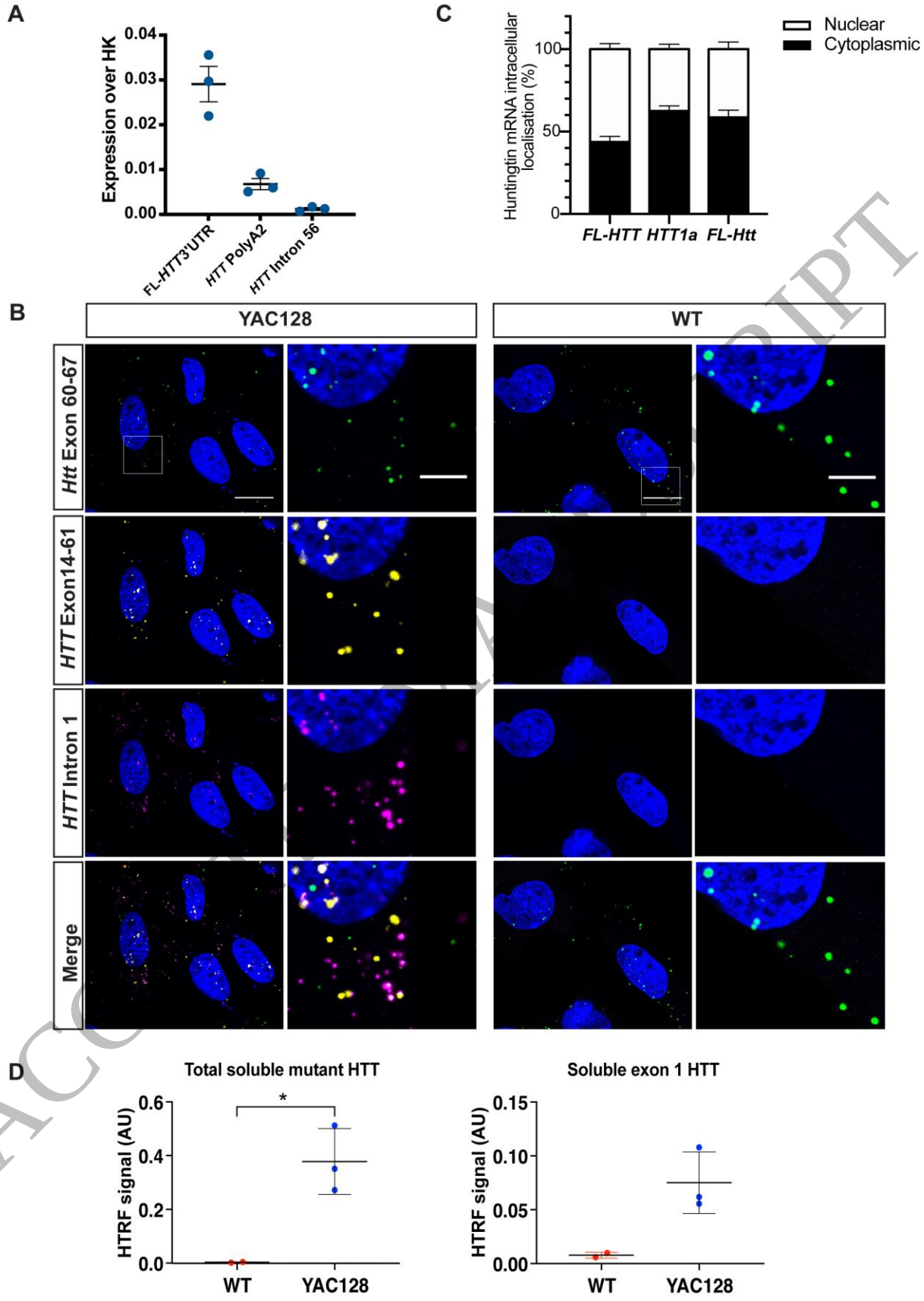

Figure 7
159x227 mm (3.4 x DPI)

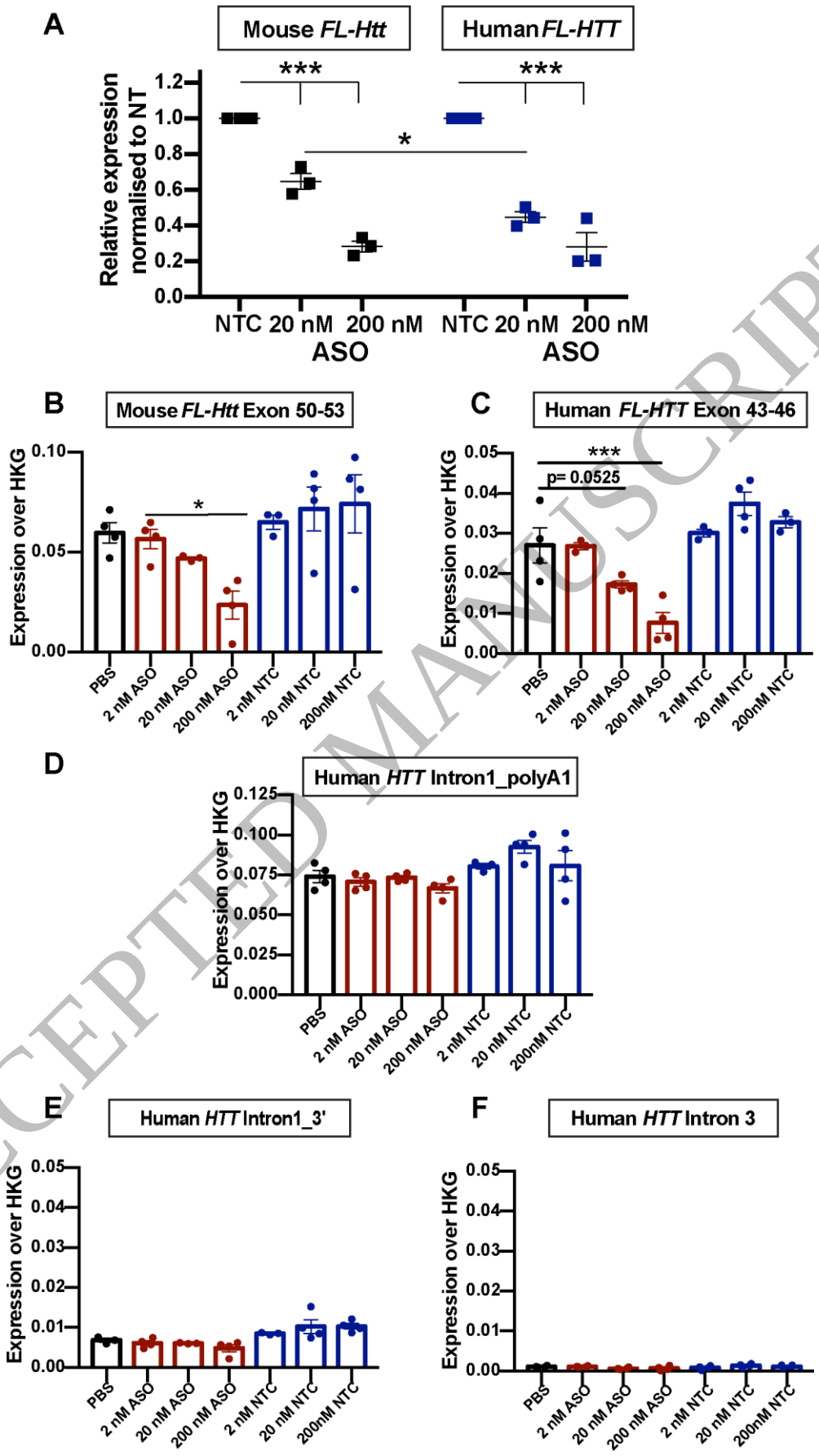

Figure 8
134x241 mm (3.4 x DPI)