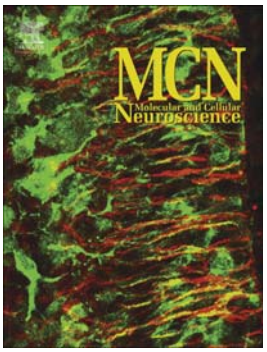
Accepted Manuscript



Transcriptional profiles for distinct aggregation states of mutant Huntingtin exon 1 protein unmask new Huntington's disease pathways

Nagaraj S. Moily, Angelique R. Ormsby, Aleksandar Stojilovic, Yasmin M. Ramdzan, Jeannine Diesch, Ross D. Hannan, Michelle S. Zajac, Anthony J. Hannan, Alicia Oshlack, Danny M. Hatters

PII:	S1044-7431(17)30130-6
DOI:	doi: 10.1016/j.mcn.2017.07.004
Reference:	YMCNE 3214
To appear in:	*Molecular and Cellular Neuroscience*
Received date:	12 April 2017
Revised date:	23 June 2017
Accepted date:	21 July 2017

Please cite this article as: Nagaraj S. Moily, Angelique R. Ormsby, Aleksandar Stojilovic, Yasmin M. Ramdzan, Jeannine Diesch, Ross D. Hannan, Michelle S. Zajac, Anthony J. Hannan, Alicia Oshlack, Danny M. Hatters , Transcriptional profiles for distinct aggregation states of mutant Huntingtin exon 1 protein unmask new Huntington's disease pathways, *Molecular and Cellular Neuroscience* (2017), doi: 10.1016/j.mcn.2017.07.004

Transcriptional profiles for distinct aggregation states of mutant Huntingtin exon 1 protein unmask new Huntington's disease pathways

Nagaraj S. Moily[1], Angelique R. Ormsby[1], Aleksandar Stojilovic[1], Yasmin M. Ramdzan[1], Jeannine Diesch[2,3], Ross D. Hannan[2,4], Michelle S. Zajac[5], Anthony J. Hannan[5], Alicia Oshlack[6,*] and Danny M. Hatters[1,*]

[1]Department of Biochemistry and Molecular Biology; and Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, VIC 3010. Australia

[2]Research Division, Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne VIC 3000. Australia.

[3]Josep Carreras Leukaemia Research Institute, ICO-Hospital Germans Trias i Pujol, Badalona, Spain

[4]The John Curtin School of Medical Research, Australian National University, Acton, ACT, Australia

[5]Florey Institute of Neuroscience and Mental Health, Melbourne Brain Centre, University of Melbourne, Parkville, VIC, Australia; and Department of Anatomy and Neuroscience, University of Melbourne, Parkville, VIC, Australia.

[6]Murdoch Childrens Research Institute, Royal Children's Hospital, Flemington Road, Parkville, Victoria 3052. Australia

*To whom correspondence should be addressed. Email:dhatters@unimelb.edu and alicia.oshlack@mcri.edu.au

Running title: Transcriptional signatures of Httex1 aggregation

ABBREVIATIONS

Polyglutamine (PolyQ), Huntingtin (Htt), Htt Exon 1 (Httex1), Huntington's Disease (HD), Pulse Shape Analysis (PulSA), Cells With Inclusions (i), Cells With Diffuse Httex1 (ni), Relative Fluorescence Units (RFU), Ingenuity Pathway Analysis (IPA), False Discovery Rate (FDR), Wild-Type (WT), Green Fluorescent Protein (GFP), Region Of Interest (ROI).

ABSTRACT

Huntington's disease is caused by polyglutamine (polyQ)-expansion mutations in the CAG tandem repeat of the *Huntingtin* gene. The central feature of Huntington's disease pathology is the aggregation of mutant Huntingtin (Htt) protein into micrometer-sized inclusion bodies. Soluble mutant Htt states are most proteotoxic and trigger an enhanced risk of death whereas inclusions confer different changes to cellular health, and may even provide adaptive responses to stress. Yet the molecular mechanisms underpinning these changes remain unclear. Using the flow cytometry method of pulse-shape analysis (PulSA) to sort neuroblastoma (Neuro2a) cells enriched with mutant or wild-type Htt into different aggregation states, we clarified which transcriptional signatures were specifically attributable to cells before versus after inclusion assembly. Dampened CREB signalling was the most striking change overall and invoked specifically by soluble mutant Httex1 states. Toxicity could be rescued by stimulation of CREB signalling. Other biological processes mapped to different changes before and after aggregation included NF-kB signalling**,** autophagy, SUMOylation, transcription regulation by histone deacetylases and BRD4, NAD+ biosynthesis, ribosome biogenesis and altered HIF-1 signalling. These findings open the path for therapeutic strategies targeting key molecular changes invoked prior to, and subsequently to, Httex1 aggregation.

Keywords: Protein misfolding, amyloid, neurodegenerative disease, Huntington's Disease

INTRODUCTION

Protein misfolding and aggregation into amyloids underlies many of the major human neurodegenerative diseases, including Alzheimer's, Parkinson's, Huntington's and motor neuron disease (1). In these diseases, the appearance of macroscopic intracellular protein aggregates known as inclusions form as a hallmark of disease development. Huntington's disease (HD) is one of the classic examples of the amyloid diseases involving intracellular protein aggregating into inclusions. HD involves an expansion of a CAG trinucleotide repeat sequence, encoding polyglutamine (polyQ), within exon 1 of huntingtin (Htt) beyond the disease threshold of 35 repeats (2). The polyQ expansion leads to a greatly enhanced capacity of the Htt exon 1 (Httex1) domain to aggregate into amyloid fibrils in vitro (3) and for N-terminal fragments similar in length to Httex1 to form microscopic inclusion bodies in pathology (4, 5). The link to disease from polyQ-expanded Httex1 has been well documented. Namely, the transgenic expression of polyQ-expanded Httex1 is sufficient to produce a HD-like pathology in rodent and primate models (6-8) and confers toxicity in cell culture models as well (9). As such Httex1 has been used extensively to model HD biology.

An outstanding question remains as to the role of aggregation in pathogenesis. Evidence increasingly suggests that cellular malfunctioning originates prior to aggregation of Htt into the inclusion body – indeed formation of inclusions seems to aid survival (reviewed in (10, 11)). Therefore, soluble polyQ-expanded Httex1 states (including monomers and oligomers) may directly confer cellular toxicity and that toxicity could become nullified as the soluble states are sequestered into the visible aggregates. The mechanisms mediating the origins of cellular toxicity and inclusion assembly remain to be properly elucidated.

Until recently, it has not been straightforward to decipher the cellular changes arising from polyQ-expanded Httex1 as it shifts from soluble states into the inclusion body because cell populations contain mixtures of cells enriched with Httex1 in different aggregation states. This heterogeneity prevents a clean delineation of effects attributable to mutation from aggregation state.

We previously developed a flow cytometry based method (Pulse Shape Analysis: PulSA) to separate cells expressing polyQ-expanded Httex1 into two major groups: those enriched with Httex1 in inclusions (i) and those enriched with soluble Httex1 states (predominantly monomer) (ni) (12) (Fig 1a). This methodology provides a unique opportunity to explicitly understand how aggregation state

of Httex1 (or indeed any mutant protein that forms an inclusion) relates to molecular dysfunction. Therefore, in this study, we sought to comprehensively determine how the transcriptome is altered by Httex1 aggregation in a cell culture model that is both sensitive to mutant Httex1 toxicity and which can be robustly controlled for co-variates such as time and expression level. We show that Httex1 provides the most profound changes to the transcriptome prior to aggregation and that the toxicity can be accounted for by a sustained shutdown of baseline CREB signalling from soluble Httex1.

MATERIALS AND METHODS

DNA vectors and constructs. pT-REx vectors (Life Technologies) encoding Httex1 with 25Q or 72Q polyQ sequence lengths were prepared as described previously (12, 13). pGW1-based Httex1-mCherry constructs were prepared as described previously (14).

Cell culture. Neuro2a cells, obtained originally from the American Type Culture Collection, were maintained in Optimem medium (Life Technologies) supplemented with 10% v/v FCS, 1 mM glutamine, 200 U/ml penicillin and 200 µg/ml streptomycin in a humidified incubator with 5% v/v atmospheric CO_2. Neuro2a cells were transiently transfected with the vectors using Lipofectamine 2000 reagent (Life Technologies) using the manufacturer's recommendations.

Collection of cells by Pulse Shape Analysis, RNA extraction and sequencing. 48 h after transfection, cells were fluorescence-activated cell sorted using a BD FACS ARIA III cell sorter. In essence, the cells were gated for live cell population by SSC and FSC gates and then the 'ni' and 'i' gates by PuLSA as described (12). To match for expression, cells were further gated to the same median GFP intensity of 2,200 RFU by varying the window of expression (Fig S1). The 25Q and 72Q Httex1 constructs were transfected and sorted in parallel across four days and performed as four matched replicates. 50,000 cells were collected for each sample (25Q-ni, 72Q-ni and 72Q-i) and sorted directly into TRI-reagent (Zymo Research, Irvine, CA). For imaging, a parallel sample set was collected and imaged as described previously by confocal microscopy (12). Total RNA was purified with the Direct Zol RNA extraction kit (Zymo Research, Irvine, CA) and verified for quality and yield (greater than 1 µg and RNA Integrity number of > 9) using the Bioanalyzer 2100 and RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, CA). rRNA was depleted and the library prepared per the manufacturer's protocols (TruSeq, Illumina) for whole transcriptome sequencing using Illumina Hi-Seq

2000 (Illumina, San Diego, CA, USA). FastQC report generated from the RNA Sequencing analysis showed good per base quality scores.

RNA sequence processing. 50 base pair paired-end reads with an average library size of 15 million tags with at least one reported alignment (~90% mapping rate) was obtained. The reads were mapped to the reference mouse genome (GRCm38/mm10) using Tophat2 (2.0.13) (15) aligner. The gene level read counts were obtained (Ensembl transcripts release 78) using Rsubread - featureCounts (v1.4.6-p1) (16) using RStudio (Version 0.99.1246) /R (Version 3.3.1).

Differential expression analysis of RNA sequencing data. Differential expression analysis of RNA sequencing data was conducted using voom transformation and limma differential expression analysis pipeline in R-bioconductor (Version 3.3). Genes were first filtered if they did not have at least 1 counts per million mapped fragments, in more than three libraries. Counts were converted to $\log_2$ counts per million, normalized using the trimmed mean of M-values normalization method (17) and precision weighted with the 'voom' function of the limma package (v3.24.15) (18). A linear model was fitted to each gene with effects for the sample types (25Q-ni, 72Q-ni and 72Q-i) as well a further fixed effect for the four replicate batches that the samples were PuLSA sorted across four days. Empirical Bayes moderated t-statistics were used to assess differences in gene expression. Statistical tests for differences in expression levels were evaluated by comparing between the sample groups: i.e. the 72Q-ni vs. 25Q-ni, 72Q-ni vs 72Q-i, 72Q-i vs 25Q-ni and total 72Q (72Q-i + 72Q-ni) vs.25Q. Genes were adjusted for multiple testing and significant genes were designated as having a false discovery rate (FDR) <0.05 using decideTest function in limma. The upregulated genes were labelled 1; downregulated genes -1; and not differentially expressed genes 0 (Table S1c). KEGG pathway analysis was performed with DAVID v6.8 using a minimum gene count of 10 and a modified Fisher Exact P-Value of <0.01 with FDR correction (19 , 20).

STRING analysis. The differentially expressed genes between the 72Q-ni and 25Q-ni with FDR corrected *p*-values <0.05 and fold change of >2 were used to create a protein: protein interaction network with the STRING v.10 algorithm (21). Data were analysed with parameters including only experimentally obtained data and a medium confidence interaction score of 0.40. 1st shell interactors of no more than 5 proteins was used.

Upstream regulator analysis. Upstream regulator analysis was performed on the genes differentially expressed between the 72Q-ni and 25Q-ni with Ingenuity Pathway Analysis (IPA) (Qiagen). IPA compares gene expression in a dataset with lists of genes that are known to be regulated by specific upstream transcriptional regulators, to infer whether a regulator is likely to be activated or inhibited. Since there is a considerable overlap in the lists of gene regulated by different transcriptional regulators, multiple points along a pathway can be suggested as key regulators, which can lead to a significant redundancy in the analysis. In IPA, this effect is quantified by the activation z-score statistics, which increases or decreases depending on the inferred activation/inhibition state of an upstream regulator. Significance in IPA was attributed to upstream regulators that have an overlap p-value of < 0.01 and an activation z-score of greater than 2 or less than −2.

Heuristic post hoc analysis of gene expression data. Gene expression data were classified into different transcriptional signatures using significant results (FDR<0.05) obtained by decideTest function in limma on different comparisons. The transcription signature groups included genes with the criteria as described below (in both positive and negative directions (i.e –1 or 1) for each signature. Progressive fold change to determine the direction of expression change across aggregation states was calculated using the formula (fold change 72Q-ni vs 25Q-ni) * [Absolute(fold change of 72Q-i vs 25Q-ni – fold change of 72Q-ni vs 25Q-ni)].

Gene expression signature	**Decide test 72Q-ni vs. 25Q-ni**	**Decide test 72Q-i vs. 25Q-ni**	**Decide test 72Q-ni vs. 72Q-i**	**Progressive fold change**
Transcription signature group 1	1	1	0	–
Transcription signature group 2	0	1	1	–
Transcription signature group 3	1	1	1	>0
Transcription signature group 4	1	0	–	<0

For each transcription signature group 1 the genes were subdivided into upregulated and downregulated gene lists with regard to direction of expression in 72Q-ni vs 25Q-ni comparison. The genes which did not fall into the categories mentioned above were excluded from the analysis.

Network Analyst – Minimal interaction networks. The gene expression signatures were mapped to known protein: protein interaction networks using network analyst toolkit (22, 23). In essence, genes in a transcriptome signature group were mapped to a validated network of all known binary protein: protein interactions. The network was trimmed to contain only proteins that connect the input genes. The most connected genes remaining in the network that had a degree centrality of >30 formed the hubs. The most locally well-connected subnetworks with a statistical threshold of Wilcoxson rank-sum P-values <0.05) were extracted as modules.

R6/1 mouse tissue processing, RNA extraction and microarray analysis. R6/1 hemizygote males (24) were obtained from the Jackson Laboratory (Bar Harbor, ME, USA) and bred with CBB6 (CBAxC57/B6) F1 females to establish an R6/1 colony. Genotypes were determined by PCR (24) with genomic DNA obtained from toe clips. Wild-type (WT) and R6/1 HD male mice were collected after weaning at 3.5 weeks of age and transferred into standard mouse cage boxes, 10 cm × 16 cm × 31 cm in size, in groups of 3–6.

WT and HD mice were killed at 12 weeks of age by cervical dislocation between 9–11:30 am. Different experimental groups were randomly allocated into balanced batches to eliminate batch effects. Brains were removed, rinsed in PBS and placed on ice for no more than 5–10 minutes prior to dissection. Brains were bisected and striatum was dissected out. Samples were immediately frozen on dry ice and transferred to a –80 °C freezer for storage until required.

Total RNA from the striatum was extracted using RNeasy Mini kits (Qiagen). On-column DNase clean-up was performed on all samples using RNase-Free DNase set (Qiagen). The concentration and integrity of the extracted total RNA present in the final eluates were determined using an Agilent Bioanalyser 2100. All the samples had a RIN of greater than 9.3. GeneChip Mouse Exon 1.0ST arrays (Affymetrix, Santa Clara, CA, USA) were used, with hybridization and scanning conducted by the Australian Genomics Research Facility (Parkville, Australia) per standard Affymetrix protocols, with appropriate quality control (QC).

Microarray data was analysed using the limma package (v3.24.15) in R-bioconductor (Version 3.3). The arrray data was preprocessed and quality control was done using the oligo package . The normalization was done at the probe level using Robust Multi-Array Average method in oligo (25). The probeset annotation was done using the Affymetrix mogene10 annotation data package (version

8.4.0) (26). The statistical testing of changes in gene expression of transgenic R6/1 HD mice was compared to wild-type littermates using a linear model in limma (Supplementary Table S2a). The pathways were obtained using IPA core analysis with gene changes attributed specifically to neuronal tissue in IPA knowledge base.

Human caudate gene analysis. Previously published raw microarray data (27) was retrieved from the EBI Array Express database (28) (http://www.ebi.ac.uk/arrayexpress) under accession number E-AFMX-6. The raw data was processed and normalized using the bioconductor affy package (v1.52.0) (29) with the Robust Multi-Array Average method (30) and differential expression was tested in limma (v3.24.15) in R-bioconductor (Version 3.3). Gene-set enrichment analyses were conducted using the caudate genes as the reference set and our cell-based RNA-seq data as the gene set with the Roast function in limma (31).

Longitudinal live cell imaging. The cells were plated at a density of 1×10^5 cells/ well in a 24 well plate. After 24 h, cells were co-transfected with 72Q-Httex1-mCherry, 25Q-Httex1-mCherry or mCherry alone and GFP using Lipofectamine 3000 reagent (Life Technologies). Media was refreshed 6 h after transfection. The GFP was used as a cellular tracer of death and proxy for Httex1 expression level. After 24 h cells were imaged using a JuLI stage live cell imaging system with fluorescent images acquired at 15 min intervals for 96 h (Nanoentek, Seoul, South Korea). Cells were tracked for visible aggregate formation and the time point of death (by loss of GFP fluorescence) with image processing tool Fiji (Version 2.0.0-rc-44) (32). Survival curves were analyzed using Prism software. Cells that drifted from focus were censored in the analysis. Expression level was calculated by measuring mean GFP levels in an ROI within the cytosol. The cells included for analysis had a mean fluorescence ± 2 standard deviation of the sample with the lowest variance. *P*-values were determined using log-rank (Mantel Cox) test. For drug treatments, forskolin (Cayman Chemical) was added after transfection to a concentration of 0.5 μM, which corresponds to the EC_{50} dosage. CREB inhibitors SGC-CBP30 (20 nM) and I-CBP112 (150 nM) were added after transfection (Sigma-Aldrich). Both drugs were added at 6 h post transfection during media refreshment.

RESULTS & DISCUSSION

Cell populations expressing matched Httex1 levels divided into those with inclusions versus those without

Our model system was a mouse neuroblastoma cell line (Neuro2a), which is a widely used model for studying Htt-induced pathogenic mechanisms. These cells were selectively vulnerable to toxicity of polyQ-expanded Httex1 (72Q) over the wild-type polyQ length (25Q) (both as fusions to mEmerald fluorescent protein (33)) (Fig 1b). In accord with prior observations, the formation of inclusions repressed the rate of death relative to cells retaining exclusively diffuse mutant Httex1 (9).

In transfected cell populations the levels of Httex1 and aggregation rates are heterogeneous (12). To remove this heterogeneity and enrich for cells with aggregates versus those without we used PulSA (12). PulSA enabled the separation of cells with inclusions (i) from those without (ni) in non-expanded, wild-type (25Q) and expanded, mutant (72Q) forms of Httex1, using a gating strategy that matched expression levels across our three groups (25Q-ni, 72Q-ni and 72Q-i) (Fig 1c; Fig S1). This strategy captured a snapshot of cells representing the extreme ends of the spectrum of aggregation states, and which enriches cell groups with greater than 90-95% fidelity (34, 35).

The largest changes in the transcriptome are stimulated by soluble mutant Httex1

Next-generation RNA sequencing (RNAseq) yielded 12,876 high abundance genes on four biological matched replicates. To first test whether our dataset yielded results in accord with previous HD datasets, we tested for differential expression between the disease like 72Q samples (72Q-ni+ 72Q-i) and 25Q-ni as control. Comparing these two groups with the previously described human caudate microarray data from HD brain (versus non-HD controls) (27) we observed 1,109 differentially expressed genes in common to both. The upregulated genes detected in mouse were found to be significantly enriched for upregulated expression in the human samples. That is, we took the significantly upregulated genes detected in our mouse study (n=428 genes) and used this as a gene set for testing with Roast (31) in the human study (FDR =0.01). These results are indicative of a general consistency in upregulated genes between our cell model and human caudate samples.

As an additional test, we found that our 72Q samples treated as one group (72Q-ni+ 72Q-i) versus the 25Q-ni group yielded several HD-relevant KEGG pathways (Supplementary Table S1a). This included DNA repair (36), p53 signalling (37) and RNA transport (38).

To understand how aggregation state affected transcription, we measured the differential expression patterns for all pair-wise comparisons between the three sample Httex1 groups: 25Q-ni, 72Q-ni and 72Q-i. The Httex1 72Q-ni versus the 25Q-ni comparison yielded changes to how soluble mutant Httex1 affected expression while the 72Q-i versus the 72Q-ni comparison assessed more specifically how aggregation of mutant Httex1 into inclusions affected expression (Fig 2). These analyses indicated that the largest and most significant changes occurred from mutant Httex1 when it was soluble prior to inclusion assembly. Furthermore, most of the changes were sustained in cells with inclusions (Fig 2 and Table S1c).

Inactivated CREB-CREM signalling is the most profound impact arising from soluble Httex1

Next, we probed the dataset for clues as to how the transcriptional changes triggered by the soluble mutant Httex1 relate to mechanism of toxicity. First, we performed a STRING network analysis (21) (on human protein counterparts) of the genes most altered in expression to see if they correlated with validated protein-protein interactions. Of the 9 genes upregulated more than 2-fold in the 72Q-ni condition (versus 25Q-ni), 3 formed part of the CREB signalling network (NCOA3, MDM4 and RUNX2) (Fig 3a). CREB is a key transcription factor that has been previously implicated as defective in HD (39, 40, 41 , 42-45). Furthermore, CREB signalling has been suggested to be mediated by deactivation of CREB binding protein (CREBBP) activity through abnormal interactions with soluble Httex1 (44, 46-48). There was no significant change in CREBBP or CREB transcript levels.

We next sought to measure the extent to which the changes in transcriptome can be explained by alteration of CREB activity. To do this we performed an unbiased Upstream Regulator Analysis using Ingenuity Pathway Analysis (IPA). The principle behind this analysis is to search for changes in transcription factor activity that accounts for the changes in downstream gene expression as represented by our dataset. 16 transcription factors were identified that are known to regulate the genes in our dataset. Critically, only 2 of these transcription factors, CREB and the closely functionally related CREM, could explain the altered expression patterns of our data with high confidence by their inactivation (Fig 3b, Table S1b). These data collectively suggest that CREB signalling defects are among the most pervasive of the changes in the transcriptome (including before and after aggregation) and that these effects are attributable to mutant Httex1 before aggregation.

To test for the in vivo relevance of these findings, we also examined the mRNA of R6/1 Httex1 transgenic mouse (24) striatum tissue at a time point prior to motor onset (12 weeks of age) by microarray analysis. Analysis of the gene expression differences in the R6/1 genotype (*n*=3) versus non-transgenic littermates (*n*=4) indicated CREB signalling (and related cAMP signalling) were in the top 5 pathways negatively impacted in an IPA canonical pathway analysis (Table S2a,b).

Dampened CREB signalling explains the molecular basis of toxicity

To further examine the importance of CREB inactivation in toxicity we performed survival curve analyses focusing on manipulating CREB signalling (Fig 3c). Prior studies have found that deletion of CREB-CREM expression in postnatal forebrain results in extensive neurodegeneration reminiscent of HD pathology (42). Inhibition of CREB signalling with inhibitors SGC-CBP30 and I-CBP112, which target CREBBP and EP300 bromodomains with high specificity (49, 50), in our Neuro2a culture system led to a high death rate indicating that baseline CREB signalling is critical for cell survival (Fig 3d). Cells expressing mutant Httex1 were negatively impacted by inhibition of CREB signalling, as anticipated (Fig 3c). However, activation of CREB signalling with forskolin, which triggers downstream phosphorylation and activation of CREB (51), significantly rescued the toxicity caused by mutant Httex1 and had otherwise no effect on baseline survival rates in cells with the wild-type Httex1. Collectively, these results suggested that toxicity of soluble mutant Httex1 manifests predominately through impairment of baseline CREB signalling and that this be remedied by triggering the CREB signalling.

Additional novel pathways unearthed as transcriptional signatures for Httex1 aggregation

We next focused on extracting other patterns in the dataset using a transcriptional signature approach as shown in Fig 4a using a post hoc heuristic analysis. These signatures can be used to make predictions about how cells respond to Httex1 aggregation state in both pathogenic and adaptive manners. This analysis yielded hundreds of genes in each transcriptional signature group. Features were extracted from this data using the Network analyst software (22), which links together genes based on their known protein-protein interactions. The most connected genes in the networks (set with a minimum threshold of 30 interactions) were identified as hub genes of interest (Table S3). Functional subnetworks (modules) that relate more discretely to biological functions were also identified using a Walktrap algorithm in Network analyst (Fig 4b). The final list (as discussed below)

of modules and hub genes has an extensive overlap with established HD mechanisms. There was also a strong overlay with CREB signalling, which suggests these modules are affected directly by the dampened CREB signalling response or are adaptive responses to counteract CREB signalling defects.

We explain the key details of these modules below and provide a discussion of how they may connect to CREB signalling and HD biology (and expanded on in Supplementary Note 1). This discussion outlines what we see are the key features of the data set and point to connections with established links to Huntington's disease biology and other features of potential interest.

Transcriptional signature 1: This signature describes the changes invoked by soluble mutant Httex1 and which have no further change when inclusions form. One explanation for this signature is to describe changes occurring as a result of toxic responses. This may apply for the HDAC5 and BRD4 modules, for which HDAC5 and BRD4 genes are chromatin modifiers that have been shown to influence CREB signalling, which are consistent with the broader changes seen in CREB signalling (52-54). HDAC proteins have also been previously linked to underlying HD pathogenesis (see Supplementary Note 1 for more discussion).

Other modules are more likely to reflect responses to counter the toxic effects of Httex1. Such upregulated modules (named after the key genes in them) included HDAC5, BRD4, NAMPT, ribosome-associated proteins and GABARAP. GABARAP, ELAVL1 and SQSTM1 genes have roles in regulating macroautophagy and proteasomal-mediated degradation and have been shown to be involved in HD pathology (see Supplementary Note 1 for more discussion). Their upregulation thus may indicate attempted increased clearance mechanisms to cope with the toxic Httex1 response. However, for GABARAP the changes may also reflect a block in autophagy (discussed in more detail in Supplementary Note 1). NAMPT gene encodes the rate-limiting enzyme in NAD production, which is an essential cellular metabolite involved in energy production, redox homeostasis, Ca^{2+} signalling and post-translational modifications. Prior work has shown that misfolded prion protein can reduce oxidised NAD (NAD+) levels (55), hence the upregulation of NAD+ may reflect a compensation response. Increased SUMOylation of Htt has been suggested to facilitate disaggregation of Htt, thereby releasing the soluble mutant Htt monomers and increasing neurotoxicity (56). Hence reduction of SUMO pathways may indicate an adaptive response to aid sequestration of the soluble Httex1 states.

Transcriptional signature 2: This signature describes the transcriptional changes that are only altered when inclusions form. One possibility to explain this signature are the genes that change expression upon adaptation responses as inclusions form on the basis of prior work suggesting inclusions are actively formed to sequester toxic soluble protein states (9, 57, 58). Up-regulated modules include RELA, which is a subnetwork within the NF-κB signalling pathway that has been implicated in HD (59) and ribosomal associated protein module, which has functions in regulating ribosome biogenesis (60). As such their upregulation may be part of mechanisms to globally regulate protein translation. Of note to the broader connection to CREB signalling alterations, CREB signalling regulates ribosomal protein production and aberrant production has been linked to ribosomal stress and apoptosis (61). Five modules were downregulated in this group. Within these modules, the PAXIP1 gene is involved in DNA repair and the CDK1 gene is involved in cell cycle re-entry. Both have been previously linked to HD pathology and in toxicity (explained further in Supplementary Note 1). A notable downregulated module was EP300. The EP300 gene encodes a co-activator of CREB, which suggests this response is influenced by the broader deregulation of CREB signalling. Also downregulated is the UBC module. UBC encodes ubiquitin, which is involved in ubiquitin-proteasome mediated protein degradation. This response is likely part of broader adaptation changes to protein folding stress (more details are explained in Supplementary Note 1).

Transcriptional signature 3: This signature describes a change of transcript levels triggered by soluble Httex1 and extended further when inclusions form. This signature may reflect a sustained or delayed toxicity response; or alternatively a delayed or amplifying adaptive response as aggregation proceeds. Upregulated modules included NF-kB pathway (more generally than the RELA subset). This may reflect a compensation response since increased NF-kB signalling aids cell survival (62). Down-regulated modules included KIAA0101, HNF4A and MYC. KIAA0101 gene is involved in oxidative stress mediated DNA repair (63), HNF4A gene is involved in the metabolic regulation of glucose and lipids has been shown to be impaired in HD (64, 65) and MYC gene is a master regulator of protein transcription (and indirectly translation) and is influenced by CREB signalling (66). Thus, the changes in MYC likely reflect an amplifying adaptive response to bypass defects in CREB signalling as well as aiding cell survival by reducing translation under stressed conditions.

Transcriptional signature 4: This signature describes transcript changes invoked by soluble 72Q Httex1 that then return to control levels when inclusions form. This signature could reflect genes that

are specifically rebalanced by aggregation in adaptive responses from inclusion formation. The upregulated modules included STAT3, of which the STAT3 gene works in concert with NF-κB to regulate cell survival (67) and VHL, which directs Hypoxia-inducible factor-1 (HIF-1) for degradation via the ubiquitin-proteasome system and suppresses apoptosis. The sole down-regulated module included HSP90AA. HSP90 and related chaperones link closely to systems regulating protein folding, degradation and potentially disaggregation (68).

Conclusions

In conclusion, our study has enabled the first comprehensive view of how aggregation state of Httex1 influences the transcriptome in HD. Most noteworthy was the finding that soluble Httex1 propagates the most substantial changes in the transcriptome by blocking CREB signalling, which supports an extensive body of work indicating CREB signalling defects as central to HD pathomechanisms. Importantly, our study is the first to our knowledge that has been able to identify this defect as specifically attributed to the soluble mutant Httex1 and in context of the system-wide transcriptional changes. Indeed, the blockage of CREB signalling seemingly accounted for the majority of the toxicity caused by Httex1. As such, our data reaffirms CREB signalling as a highly attractive therapeutic target in Huntington's disease. This conclusion is supported findings that an enhancement of CREB levels via Sirt1 overexpression in the R6/2 HD mouse model can improve survival time (69). In contrast, the formation of inclusions offers (by comparison) trivial changes in the magnitude of transcriptional changes. Moreover, of the changes observed, the signatures point to those involved in adapting to defects in CREB signalling such as the rewiring of CREB-related signally networks to compensate for CREB deficiency. In addition, other changes associated with protein homeostasis responses were also documented in these signatures.

Another feature of our dataset was the pronounced enrichment of genes encoding polyQ. Five of the 9 genes upregulated more than 2-fold by the soluble polyQ-expanded Httex1 contained endogenous polyQ sequences suggesting that polyQ-expanded Htt may more broadly deactivate by soluble co-aggregation these proteins leading to their compensatory upregulation (Fig 2). Three of the upregulated polyQ-encoding genes are involved in CREB signalling, which may explain why CREB signalling appears to be the most profoundly affected pathway (Fig 3a).

Our new roadmaps of the transcriptional changes associated with Htt aggregation states offer a powerful resource to understanding how Httex1, notably the soluble forms, propagates toxicity as well as in yielding CREB signalling as a therapeutic target for further investigation.

AUTHOR CONTRIBUTIONS

NSM, JD, RDH, AJH, AO, and DMH designed research; NSM, ARO, AS, YMR, and MSZ performed research; NSM, ARO, AS, AO and DMH analyzed data; and NSM, A O, and DMH co-wrote the paper.

ACKNOWLEDGEMENTS

This work was supported by grants to DMH from the Australian Research Council (grant number FT120100039); grants/fellowships from the National Health and Medical Research Council Project to DMH (grant numbers APP1049458, APP1049459 and APP1102059), and a grant from the Hereditary Disease Foundation (USA). AJH is an NHMRC Principal Research Fellow.

The authors declare that they have no competing interests.

REFERENCES

1. Sipe, J. D., Benson, M. D., Buxbaum, J. N., Ikeda, S. I., Merlini, G., Saraiva, M. J., and Westermark, P. (2016) Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines. *Amyloid* **23**, 209-213
2. MacDonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., Barnes, G., Taylor, S. A., James, M., Groot, N., MacFarlane, H., Jenkins, B., Anderson, M. A., Wexler, N. S., Gusella, J. F., Bates, G. P., Baxendale, S., Hummerich, H., Kirby, S., North, M., Youngman, S., Mott, R., Zehetner, G., Sedlacek, Z., Poustka, A., Frischauf, A.-M., Lehrach, H., Buckler, A. J., Church, D., Doucette-Stamm, L., O'Donovan, M. C., Riba-Ramirez, L., Shah, M., Stanton, V. P., Strobel, S. A., Draths, K. M., Wales, J. L., Dervan, P., Housman, D. E., Altherr, M., Shiang, R., Thompson, L., Fielder, T., Wasmuth, J. J., Tagle, D., Valdes, J., Elmer, L., Allard, M., Castilla, L., Swaroop, M., Blanchard, K., Collins, F. S., Snell, R., Holloway, T., Gillespie, K., Datson, N., Shaw, D., and Harper, P. S. (1993) A Novel Gene Containing a Trinucleotide Repeat that is Expanded and Unstable on Huntington's Disease Chromosomes. *Cell* **72**, 971-983
3. Scherzinger, E., Sittler, A., Schweiger, K., Heiser, V., Lurz, R., Hasenbank, R., Bates, G. P., Lehrach, H., and Wanker, E. E. (1999) Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4604-4609
4. DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N. (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* **277**, 1990-1993
5. Kazantsev, A., Preisinger, E., Dranovsky, A., Goldgaber, D., and Housman, D. (1999) Insoluble Detergent-resistant Aggregates Form Between Pathological and Nonpathological Lengths of Polyglutamine in Mammalian Cells. *Proc. Natl. Acad. Sc. USA* **96**, 11404-11409
6. Davies, S. W., Turmaine, M., Cozens, B. A., DiFiglia, M., Sharp, A. H., Ross, C. A., Scherzinger, E., Wanker, E. E., Mangiarini, L., and Bates, G. P. (1997) Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* **90**, 537-548

7. von Horsten, S., Schmitt, I., Nguyen, H. P., Holzmann, C., Schmidt, T., Walther, T., Bader, M., Pabst, R., Kobbe, P., Krotova, J., Stiller, D., Kask, A., Vaarmann, A., Rathke-Hartlieb, S., Schulz, J. B., Grasshoff, U., Bauer, I., Vieira-Saecker, A. M., Paul, M., Jones, L., Lindenberg, K. S., Landwehrmeyer, B., Bauer, A., Li, X. J., and Riess, O. (2003) Transgenic rat model of Huntington's disease. *Hum. Mol. Genet.* **12**, 617-624

8. Yang, S.-H., Cheng, P.-H., Banta, H., Piotrowska-Nitsche, K., Yang, J.-J., Cheng, E. C. H., Snyder, B., Larkin, K., Liu, J., Orkin, J., Fang, Z.-H., Smith, Y., Bachevalier, J., Zola, S. M., Li, S.-H., Li, X.-J., and Chan, A. W. S. (2008) Towards a transgenic model of Huntington/'s disease in a non-human primate. *Nature* **453**, 921-924

9. Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* **431**, 805-810

10. Arrasate, M., and Finkbeiner, S. (2012) Protein aggregates in Huntington's disease. *Exp. Neurol.* **238**, 1-11

11. Hatters, D. M. (2012) Putting huntingtin "aggregation" in view with windows into the cellular milieu. *Curr Top Med Chem* **12**, 2611-2622

12. Ramdzan, Y. M., Polling, S., Chia, C. P., Ng, I. H., Ormsby, A. R., Croft, N. P., Purcell, A. W., Bogoyevitch, M. A., Ng, D. C., Gleeson, P. A., and Hatters, D. M. (2012) Tracking protein aggregation and mislocalization in cells with flow cytometry. *Nat Methods* **9**, 467-470

13. Ramdzan, Y. M., Nisbet, R. M., Miller, J., Finkbeiner, S., Hill, A. F., and Hatters, D. M. (2010) Conformation sensors that distinguish monomeric proteins from oligomers in live cells. *Chem. Biol.* **17**, 371-379

14. Tsvetkov, A. S., Miller, J., Arrasate, M., Wong, J. S., Pleiss, M. A., and Finkbeiner, S. (2010) A small-molecule scaffold induces autophagy in primary neurons and protects against toxicity in a Huntington disease model. *Proc. Natl. Acad. Sc. USA* **107**, 16982-16987

15. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36

16. Liao, Y., Smyth, G. K., and Shi, W. (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* **41**, e108

17. Robinson, M. D., and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25

18. Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29

19. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13

20. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57

21. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447-452

22. Xia, J., Benner, M. J., and Hancock, R. E. (2014) NetworkAnalyst--integrative approaches for protein-protein interaction network analysis and visual exploration. *Nucleic Acids Res* **42**, W167-174

23. Xia, J., Gill, E. E., and Hancock, R. E. (2015) NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat Protoc* **10**, 823-844

24. Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trottier, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996) Exon 1 of the HD Gene with an Expanded CAG Repeat Is Sufficient to Cause a Progressive Neurological Phenotype in Transgenic Mice. *Cell* **87**, 493-506

25. Carvalho, B. S., and Irizarry, R. A. (2010) A framework for oligonucleotide microarray preprocessing. *Bioinformatics* **26**, 2363-2367

26. MacDonald, J. W. (2016) mogene10stprobeset.db.

27. Hodges, A., Strand, A. D., Aragaki, A. K., Kuhn, A., Sengstag, T., Hughes, G., Elliston, L. A., Hartog, C., Goldstein, D. R., Thu, D., Hollingsworth, Z. R., Collin, F., Synek, B., Holmans, P. A., Young, A. B., Wexler, N. S., Delorenzi, M., Kooperberg, C., Augood, S. J., Faull, R. L., Olson, J. M., Jones, L., and Luthi-Carter, R. (2006) Regional and cellular gene expression changes in human Huntington's disease brain. *Hum. Mol. Genet.* **15**, 965-977

28. Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y. A., Williams, E., Dylag, M., Kurbatova, N., Brandizi, M., Burdett, T., Megy, K., Pilicheva, E., Rustici, G., Tikhonov, A., Parkinson, H., Petryszak, R., Sarkans, U., and Brazma, A. (2015) ArrayExpress update--simplifying data submissions. *Nucleic Acids Res* **43**, D1113-1116

29. Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004) affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-315

30. Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193

31. Wu, D., Lim, E., Vaillant, F., Asselin-Labat, M. L., Visvader, J. E., and Smyth, G. K. (2010) ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176-2182

32. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682

33. Olshina, M. A., Angley, L. M., Ramdzan, Y. M., Tang, J., Bailey, M. F., Hill, A. F., and Hatters, D. M. (2010) Tracking mutant huntingtin aggregation kinetics in cells reveals three major populations that include an invariant oligomer pool. *J. Biol. Chem.* **285**, 21807-21816

34. Polling, S., Mok, Y. F., Ramdzan, Y. M., Turner, B. J., Yerbury, J. J., Hill, A. F., and Hatters, D. M. (2014) Misfolded polyglutamine, polyalanine, and superoxide dismutase 1 aggregate via distinct pathways in the cell. *J. Biol. Chem.* **289**, 6669-6680

35. Polling, S., Ormsby, A. R., Wood, R. J., Lee, K., Shoubridge, C., Hughes, J. N., Thomas, P. Q., Griffin, M. D., Hill, A. F., Bowden, Q., Bocking, T., and Hatters, D. M. (2015) Polyalanine expansions drive a shift into alpha-helical clusters without amyloid-fibril formation. *Nat Struct Mol Biol* **22**, 1008-1015

36. Genetic Modifiers of Huntington's Disease, C. (2015) Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell* **162**, 516-526

37. Bae, B. I., Xu, H., Igarashi, S., Fujimuro, M., Agrawal, N., Taya, Y., Hayward, S. D., Moran, T. H., Montell, C., Ross, C. A., Snyder, S. H., and Sawa, A. (2005) p53 mediates cellular dysfunction and behavioral abnormalities in Huntington's disease. *Neuron* **47**, 29-41

38. Gasset-Rosa, F., Chillon-Marinas, C., Goginashvili, A., Atwal, R. S., Artates, J. W., Tabet, R., Wheeler, V. C., Bang, A. G., Cleveland, D. W., and Lagier-Tourenne, C. (2017) Polyglutamine-Expanded Huntingtin Exacerbates Age-Related Disruption of Nuclear Integrity and Nucleocytoplasmic Transport. *Neuron* **94**, 48-57 e44

39. Luthi-Carter, R., Strand, A., Peters, N. L., Solano, S. M., Hollingsworth, Z. R., Menon, A. S., Frey, A. S., Spektor, B. S., Penney, E. B., Schilling, G., Ross, C. A., Borchelt, D. R., Tapscott, S. J., Young, A. B., Cha, J. H., and Olson, J. M. (2000) Decreased expression of striatal signaling genes in a mouse model of Huntington's disease. *Hum. Mol. Genet.* **9**, 1259-1271

40. Shimohata, T., Nakajima, T., Yamada, M., Uchida, C., Onodera, O., Naruse, S., Kimura, T., Koide, R., Nozaki, K., Sano, Y., Ishiguro, H., Sakoe, K., Ooshima, T., Sato, A., Ikeuchi, T., Oyake, M., Sato, T., Aoyagi, Y., Hozumi, I., Nagatsu, T., Takiyama, Y., Nishizawa, M., Goto, J., Kanazawa, I., Davidson, I., Tanese, N., Takahashi, H., and Tsuji, S. (2000) Expanded polyglutamine stretches interact with TAFII130, interfering with CREB-dependent transcription. *Nat. Genet.* **26**, 29-36

41. Wyttenbach, A., Swartz, J., Kita, H., Thykjaer, T., Carmichael, J., Bradley, J., Brown, R., Maxwell, M., Schapira, A., Orntoft, T. F., Kato, K., and Rubinsztein, D. C. (2001) Polyglutamine expansions cause decreased CRE-mediated transcription and early gene expression changes prior to cell death in an inducible cell model of Huntington's disease. *Hum. Mol. Genet.* **10**, 1829-1845

42. Mantamadiotis, T., Lemberger, T., Bleckmann, S. C., Kern, H., Kretz, O., Martin Villalba, A., Tronche, F., Kellendonk, C., Gau, D., Kapfhammer, J., Otto, C., Schmid, W., and Schutz, G. (2002) Disruption of CREB function in brain leads to neurodegeneration. *Nat. Genet.* **31**, 47-54

43. Steffan, J. S., Kazantsev, A., Spasic-Boskovic, O., Greenwald, M., Zhu, Y. Z., Gohler, H., Wanker, E. E., Bates, G. P., Housman, D. E., and Thompson, L. M. (2000) The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6763-6768

44. Steffan, J. S., Bodai, L., Pallos, J., Poelman, M., McCampbell, A., Apostol, B. L., Kazantsev, A., Schmidt, E., Zhu, Y. Z., Greenwald, M., Kurokawa, R., Housman, D. E., Jackson, G. R., Marsh, J. L., and Thompson, L. M. (2001) Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. *Nature* **413**, 739-743

45. Dunah, A. W., Jeong, H., Griffin, A., Kim, Y. M., Standaert, D. G., Hersch, S. M., Mouradian, M. M., Young, A. B., Tanese, N., and Krainc, D. (2002) Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. *Science* **296**, 2238-2243

46. Jiang, H., Poirier, M. A., Liang, Y., Pei, Z., Weiskittel, C. E., Smith, W. W., DeFranco, D. B., and Ross, C. A. (2006) Depletion of CBP is directly linked with cellular toxicity caused by mutant huntingtin. *Neurobiol. Dis.* **23**, 543-551

47. Schaffar, G., Breuer, P., Boteva, R., Behrends, C., Tzvetkov, N., Strippel, N., Sakahira, H., Siegers, K., Hayer-Hartl, M., and Hartl, F. U. (2004) Cellular Toxicity of Polyglutamine Expansion Proteins: Mechanism of Transcription Factor Deactivation. *Mol. Cell* **15**, 95-105

48. Cong, S. Y., Pepers, B. A., Evert, B. O., Rubinsztein, D. C., Roos, R. A., van Ommen, G. J., and Dorsman, J. C. (2005) Mutant huntingtin represses CBP, but not p300, by binding and protein degradation. *Mol. Cell. Neurosci.* **30**, 560-571

49. Hammitzsch, A., Tallant, C., Fedorov, O., O'Mahony, A., Brennan, P. E., Hay, D. A., Martinez, F. O., Al-Mossawi, M. H., de Wit, J., Vecellio, M., Wells, C., Wordsworth, P., Muller, S., Knapp, S., and Bowness, P. (2015) CBP30, a selective CBP/p300 bromodomain inhibitor, suppresses human Th17 responses. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 10768-10773

50. Picaud, S., Fedorov, O., Thanasopoulou, A., Leonards, K., Jones, K., Meier, J., Olzscha, H., Monteiro, O., Martin, S., Philpott, M., Tumber, A., Filippakopoulos, P., Yapp, C., Wells, C., Che, K. H., Bannister, A., Robson, S., Kumar, U., Parr, N., Lee, K., Lugo, D., Jeffrey, P., Taylor, S., Vecellio, M. L., Bountra, C., Brennan, P. E., O'Mahony, A., Velichko, S., Muller, S., Hay, D., Daniels, D. L., Urh, M., La Thangue, N. B., Kouzarides, T., Prinjha, R., Schwaller, J., and Knapp, S. (2015) Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy. *Cancer Res.* **75**, 5106-5119

51. Insel, P. A., and Ostrom, R. S. (2003) Forskolin as a Tool for Examining Adenylyl Cyclase Expression, Regulation, and G Protein Signaling. *Cell. Mol. Neurobiol.* **23**, 305-314

52. Ha, C. H., Kim, J. Y., Zhao, J., Wang, W., Jhun, B. S., Wong, C., and Jin, Z. G. (2010) PKA phosphorylates histone deacetylase 5 and prevents its nuclear export, leading to the inhibition of gene transcription and cardiomyocyte hypertrophy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15467-15472

53. Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E. T., Buhler, M., Margueron, R., and Schneider, R. (2013) Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* **152**, 859-872

54. Devaiah, B. N., Case-Borden, C., Gegonne, A., Hsu, C. H., Chen, Q., Meerzaman, D., Dey, A., Ozato, K., and Singer, D. S. (2016) BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin. *Nat Struct Mol Biol* **23**, 540-548

55. Zhou, M., Ottenberg, G., Sferrazza, G. F., Hubbs, C., Fallahi, M., Rumbaugh, G., Brantley, A. F., and Lasmezas, C. I. (2015) Neuronal death induced by misfolded prion protein is due to NAD+ depletion and can be relieved in vitro and in vivo by NAD+ replenishment. *Brain* **138**, 992-1008

56. Steffan, J. S., Agrawal, N., Pallos, J., Rockabrand, E., Trotman, L. C., Slepko, N., Illes, K., Lukacsovich, T., Zhu, Y. Z., Cattaneo, E., Pandolfi, P. P., Thompson, L. M., and Marsh, J. L. (2004) SUMO modification of Huntingtin and Huntington's disease pathology. *Science* **304**, 100-104

57. Kopito, R. R. (2000) Aggresomes, Inclusion Bodies and Protein Aggregation. *Trends Cell Biol.* **10**, 524

58. Kaganovich, D., Kopito, R., and Frydman, J. (2008) Misfolded Proteins Partition Between Two Distinct Quality Control Compartments. *Nature* **454**, 1088-1095

59. Trager, U., Andre, R., Lahiri, N., Magnusson-Lind, A., Weiss, A., Grueninger, S., McKinnon, C., Sirinathsinghji, E., Kahlon, S., Pfister, E. L., Moser, R., Hummerich, H., Antoniou, M., Bates, G. P., Luthi-Carter, R., Lowdell, M. W., Bjorkqvist, M., Ostroff, G. R., Aronin, N., and Tabrizi, S. J. (2014) HTT-lowering reverses Huntington's disease immune dysfunction caused by NFkappaB pathway dysregulation. *Brain* **137**, 819-833

60. Warner, J. R., and McIntosh, K. B. (2009) How Common are Extra-ribosomal Functions of Ribosomal Proteins? *Mol. Cell* **34**, 3-11

61. Deisenroth, C., and Zhang, Y. (2011) The Ribosomal Protein-Mdm2-p53 Pathway and Energy Metabolism: Bridging the Gap between Feast and Famine. *Genes Cancer* **2**, 392-403

62. Yu, Z., Zhou, D., Cheng, G., and Mattson, M. P. (2000) Neuroprotective role for the p50 subunit of NF-kappaB in an experimental model of Huntington's disease. *J. Mol. Neurosci.* **15**, 31-44

63. Klein, J. A., and Ackerman, S. L. (2003) Oxidative stress, cell cycle, and neurodegeneration. *J. Clin. Invest.* **111**, 785-793

64. Olah, J., Klivenyi, P., Gardian, G., Vecsei, L., Orosz, F., Kovacs, G. G., Westerhoff, H. V., and Ovadi, J. (2008) Increased glucose metabolism and ATP level in brain tissue of Huntington's disease transgenic mice. *FEBS J* **275**, 4740-4755

65. Block, R. C., Dorsey, E. R., Beck, C. A., Brenna, J. T., and Shoulson, I. (2010) Altered Cholesterol and Fatty Acid Metabolism in Huntington Disease. *J Clin Lipidol* **4**, 17-23

66. Riggelen, J. v., Yetil, A., and Felsher, D. W. (2010) MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer* **10**, 301-309

67. Grivennikov, S. I., and Karin, M. (2010) Dangerous liaisons: STAT3 and NF-kappaB collaboration and crosstalk in cancer. *Cytokine Growth Factor Rev.* **21**, 11-19

68. Luo, W., Sun, W., Taldone, T., Rodina, A., and Chiosis, G. (2010) Heat shock protein 90 in neurodegenerative diseases. *Mol Neurodegener* **5**, 24

69. Cohen, D. E., Cui, L., Supinski, A., Savas, J. N., Mazzulli, J. R., Yates, J. R., Bordone, L., Guarente, L. P., and Krainc, D. (2012) Sirt1 mediates neuroprotection from mutant huntingtin by activation of TORC1 and CREB transcriptional pathway. *Nat. Med.* **18**, 159-165

LEGENDS TO FIGURES

Fig 1: Strategy to separate cells into populations enriched with polyQ-expanded Httex1 in distinct aggregation states. **a**) Pulse Shape Analysis (PulSA) by flow cytometry can separate cell populations heterogeneous in aggregation state. **b)** Neuro2a cells that form inclusions live longer than those that do not. Shown are Kaplan-Meier survival curves tracked by longitudinal imaging of Httex1-Emerald fusions from 20 h after transfection. *P*-values refer to log-rank test (Mantel Cox) on survival curves. **c)** Flow cytograms showing the PulSA gating strategy to sort the mixed cell populations for analysis by whole transcriptome analysis. Shown are gates for i and ni and for matched expression (red; Further details in Fig S1). The images show representative cells collected by sorting flow cytometry and imaged for GFP fluorescence by confocal microscopy. Scale bar, 20 µm.

Fig 2: Volcano plots of gene expression changes attributable to aggregation states of Httex1. Shown are comparisons between the three treatment groups: cells lacking visible aggregates but differing in polyQ length (25Q-ni and 72Q-ni cells) and cells with expanded polyQ length that differ in aggregation state (72Q-i). Genes with a FDR less than *P*=0.05 are shown in red. Genes with greater than 2-fold change are shown in blue. Genes encoding proteins with an intrinsic native polyQ sequence have orange labels.

Fig 3: CREB signaling is dampened by soluble mutant Httex1 and explains the major source of toxicity to Neuro2a cells. a) A protein-protein interacting network of the genes enriched more than 2-fold in cells enriched with soluble mutant Httex1. Shown are STRING (v10) validated protein-protein interactions of genes corresponding to those enriched in the dataset. The network was built to include first shell interactors (i.e. interacting proteins corresponding to genes not enriched in the dataset). **b)** Upregulator analysis (Ingenuity) to explain how changes in differential expression patterns of our dataset relate to changed protein activity of transcription factors. The inner circle shows transcription factors that control genes seen differentially expressed in the dataset. The outer ring shows the key genes in our dataset that are targets of the transcription factors and their patterns of expression. **c)** Schematic of CREB signaling activity and role of activator forskolin. **d)** Effect of upregulating (with forskolin) or suppressing (with inhibitors CREB signaling SGC-CBP30 and I-CBP112) on survival curves. *P*-values corresponds to log-rank (Mantel Cox) tests.

Fig.4: Transcriptional signatures for Httex1 aggregation state. a) Shown is our model for how transcriptome data can be mapped to the changes attributable to mutation, and then aggregation into inclusions. Data are filtered into transcriptional signatures as represented by the bar graphs, based on changes with FDR <0.05 (*n*=4). **b)** Shown are functional modules (with Wilcoxson rank-sum *P*-values <0.05), and their designated names, extracted from protein: protein interaction networks corresponding to all the genes in each transcriptional signature using Network Analyst Walktrap software. These modules can be upregulated or downregulated in the transcriptional signature (and the direction is shown). Each module shows the genes (points) and their connections. The most connected genes in the modules are labelled.

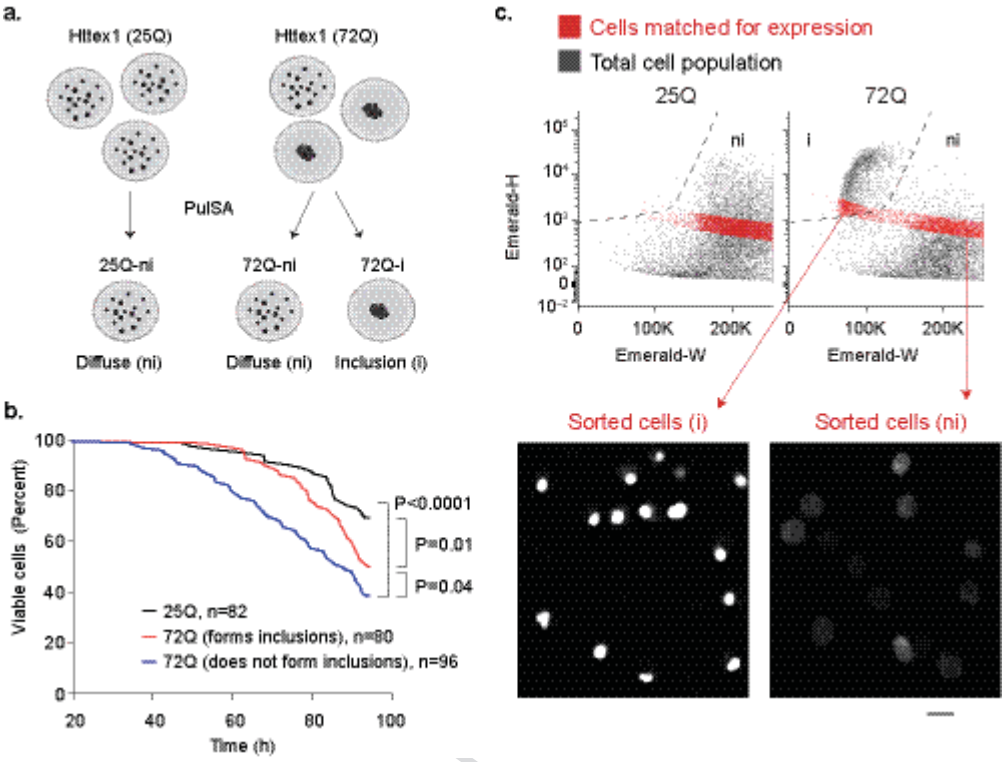


Fig. 1

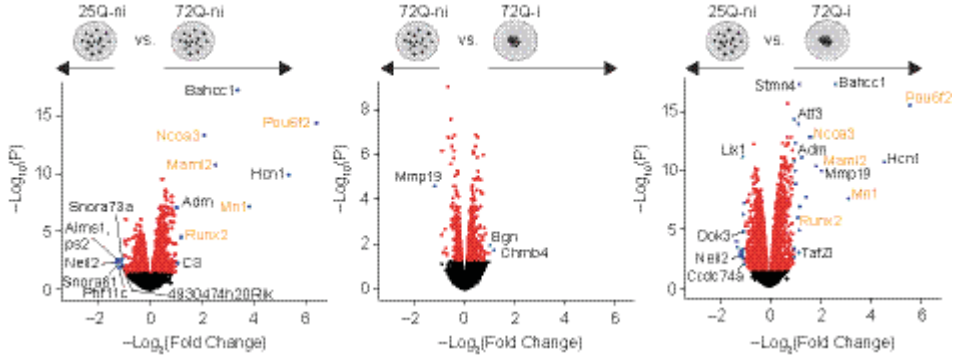


Fig. 2

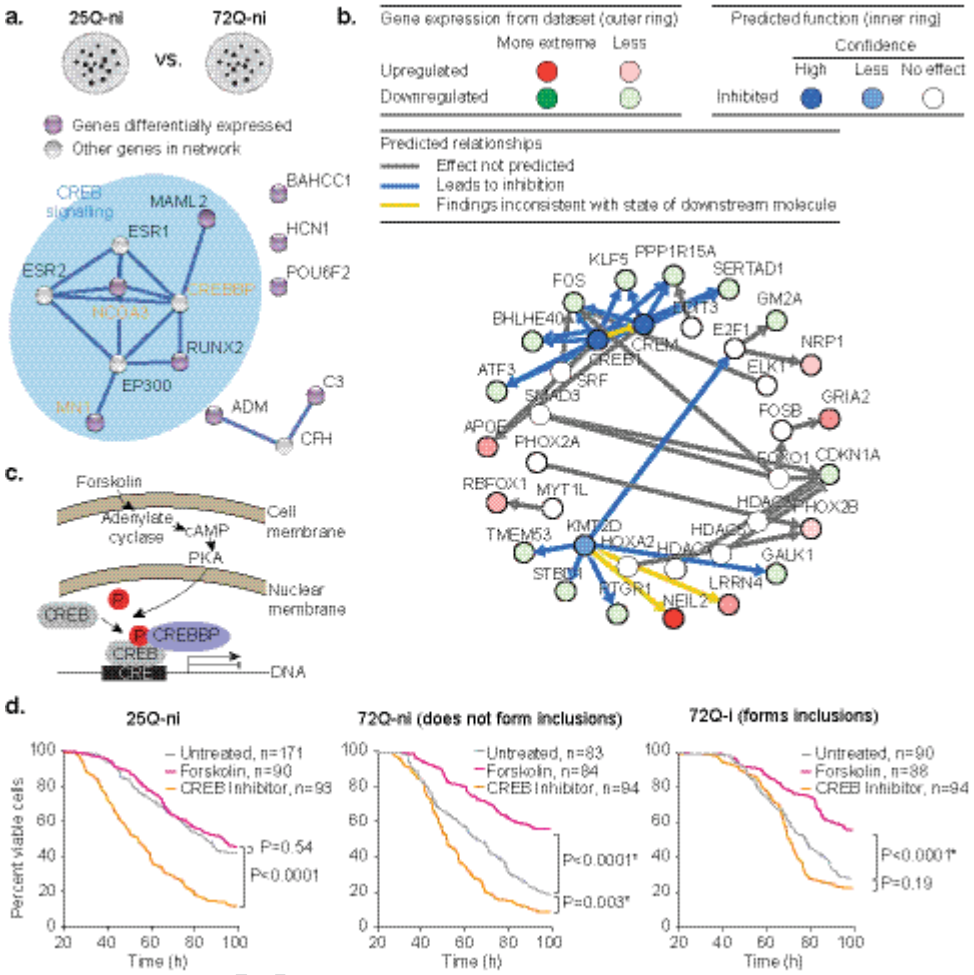


Fig. 3

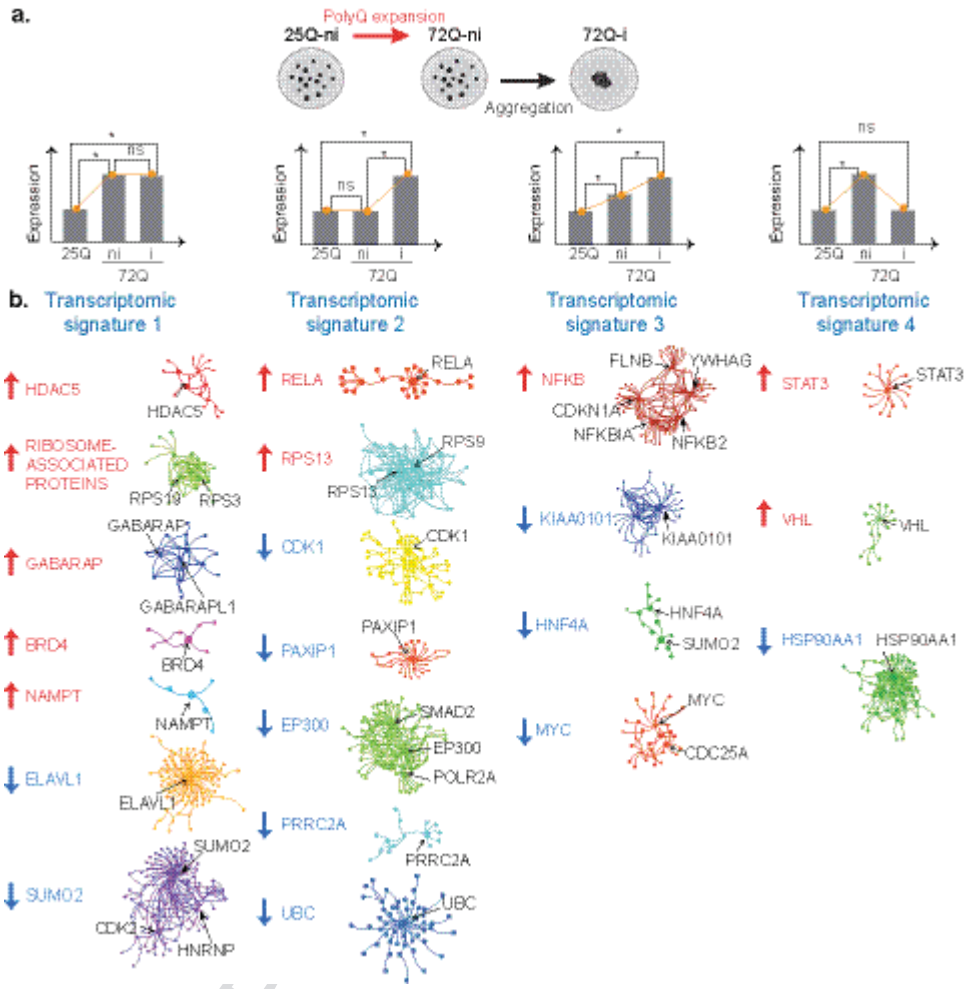


Fig. 4

Highlights

- Transcriptional changes due to aggregation of mutant Httex1 are described
- The largest changes in the transcriptome are stimulated by soluble mutant Httex1
- Inactivated CREB signalling is the most profound impact arising from soluble Httex1
- Dampened CREB signalling explains the molecular basis of toxicity
- Additional pathways unearthed as transcriptional signatures for Httex1 aggregation