Analyzing spatial transcriptomics and neuroimaging data in neurodegenerative diseases
Keo, D.

Citation
Keo, D. (2020, December 3). *Analyzing spatial transcriptomics and neuroimaging data in neurodegenerative diseases*. Retrieved from https://hdl.handle.net/1887/138480

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/138480 holds various files of this Leiden University dissertation.

Author: Keo, D.
Title: Analyzing spatial transcriptomics and neuroimaging data in neurodegenerative diseases
Issue Date: 2020-12-03

CHAPTER 3
CO-EXPRESSION PATTERNS BETWEEN *ATN1* AND *ATXN2* COINCIDE WITH BRAIN REGIONS AFFECTED IN HUNTINGTON'S DISEASE

Arlin Keo

N. Ahmad Aziz

Oleh Dzyubachyk

Jeroen van der Grond

Willeke M.C. Roon-Mom

Boudewijn P.F. Lelieveldt

Marcel J.T. Reinders

Ahmed Mahfouz

ABSTRACT

Cytosine-adenine-guanine (CAG) repeat expansions in the coding regions of nine polyglutamine (polyQ) genes (*HTT*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *ATN1*, *AR*, and *TBP*) are the cause of several neurodegenerative diseases including Huntington's disease (HD), six different spinocerebellar ataxias (SCAs), dentatorubral-pallidoluysian atrophy, and spinobulbar muscular atrophy. The expanded CAG repeat length in the causative gene is negatively related to the age-at-onset (AAO) of clinical symptoms. In addition to the expanded CAG repeat length in the causative gene, the normal CAG repeats in the other polyQ genes can affect the AAO, suggesting functional interactions between the polyQ genes. However, there is no detailed assessment of the relationships among polyQ genes in pathologically relevant brain regions.

We used gene co-expression analysis to study the functional relationships among polyQ genes in different brain regions using the Allen Human Brain Atlas (AHBA), a spatial map of gene expression in the healthy brain. We constructed co-expression networks for seven anatomical brain structures, as well as a region showing a specific pattern of atrophy in HD patients detected by magnetic resonance imaging (MRI) of the brain. In this HD-associated region, we found that *ATN1* and *ATXN2* were co-expressed and shared co-expression partners which were enriched for DNA repair genes. We observed a similar co-expression pattern in the frontal lobe, parietal lobe, and striatum in which this relation was most pronounced. Given that the co-expression patterns for these anatomical structures were similar to those for the HD-associated region, our results suggest that their disruption is likely involved in HD pathology. Moreover, *ATN1* and *ATXN2* also shared many co-expressed genes with *HTT*, the causative gene of HD, across the brain. Although this triangular relationship among these three polyQ genes may also be dysregulated in other polyQ diseases, stronger co-expression patterns between *ATN1* and *ATXN2* observed in the HD-associated region, especially in the striatum, may be more specific to HD.

3.1 Introduction

Polyglutamine (polyQ) diseases are a family of nine neurodegenerative disorders caused by a cytosine-adenine-guanine (CAG) trinucleotide repeat expansion in the coding region of one of the polyQ disease-associated genes. PolyQ diseases include Huntington's disease (HD), six spinocerebellar ataxias (SCAs), dentatorubral-pallidoluysian atrophy (DRPLA), and spinobulbar muscular atrophy (SBMA), each with its own causative gene: *HTT*, *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *TBP*, *ATN1*, and *AR*, respectively [1]. With the exception of SBMA, all polyQ diseases are inherited in an autosomal dominant manner [1]. The CAG repeat region is translated into a stretch of glutamine amino acids, also referred to as the polyQ tract [2,3]. In HD patients, the polyQ expansion in the huntingtin protein causes neurodegeneration that affects the striatum most severely and results in cognitive, psychiatric as well as motor disturbances and gait abnormalities [4,5]. It is thought that the expansion of the polyQ tract causes the protein to misfold and aggregate, and therefore loses its normal function. Hence, these mutant proteins also become toxic components as they trigger the misfolding of other proteins [6].

In HD and SCAs, longer CAG repeat lengths in the causative polyQ genes are associated with an earlier age-at-onset (AAO) [7,8]. In HD, up to 75% of the variability in AAO can be explained by the *HTT* CAG repeat length [7], while in SCA1, SCA2, SCA3, SCA6, and SCA7, the CAG repeat in the causative gene explains between 32 and 80% of the AAO variability [8]. There is also evidence suggesting that, in addition to the expanded CAG repeat length in the causative gene, the normal CAG repeat lengths in other non-causative polyQ genes affect the AAO in HD and SCAs [7–12]. For example, clinical observations show that the AAO in HD is not only affected by the expanded CAG repeat length in *HTT*, but also by the normal CAG repeat length in *ATN1* and *ATXN1* negatively affecting the AAO [7]. The observation that the number of CAG repeats in multiple polyQ genes can affect the AAO in HD or SCAs suggests that polyQ gene products are functionally interacting.

In addition to the suggested genetic associations based on shared effects on AAO, similar mechanisms likely contribute to polyQ disease pathogenesis. These mechanisms include misfolding of the disease protein, deleterious protein interactions, transcriptional dysregulation, mitochondrial dysfunction, aberrant neuronal signaling, cellular protein homeostasis impairment, and RNA toxicity [1]. Despite the similarities, these mechanisms seem to affect specific brain regions depending on the particular polyQ disease. To get a better understanding of how the CAG expansion in *HTT* affects the brain, multiple studies used genome-wide expression analysis of post-mortem samples collected from different brain regions. By comparing the brain region-specific expression profiles of HD patients and healthy controls, the highest number of differentially expressed genes was found in the caudate nucleus and to a lesser extent in the cerebellum [13]. This demonstrates that differential gene expression in the HD brain has a regional pattern corresponding to the known pattern of neuropathology. A more recent study used longitudinal RNA-sequencing expression data of HD knock-in mice with increasing CAG repeat lengths, and using weighted gene co-expression network analysis (WGCNA), they studied length-dependence

of molecular networks in HD-relevant brain regions [14]. They found one striatal module which showed strong patterns of downregulation that were both CAG length- and age-dependent.

In addition to patient postmortem samples and mouse models, characterizing expression patterns of wild-type genes in the healthy brain can provide useful insights into the molecular and pathological changes regarding polyQ diseases. Using gene expression data from ten different brain tissues and 101 healthy individuals from the UK Brain Expression Consortium [15], Bettencourt et al. analyzed the co-expression relationships of SCA genes in the healthy brain using WGCNA [16]. Two cerebellar modules were enriched in SCA transcripts of which one module seemed preserved across all brain regions and the other module seemed unique to the cerebellum. The study suggests that genes co-expressed with SCA genes across all brain regions give rise to more complex phenotypes (other ataxia syndromes and neurodegenerative disorders), while cerebellum-specific co-expressing genes results in pure ataxia phenotypes. While previous studies have focused on region-specific gene expression differences between patients and healthy controls [13,17], and co-expression relationships between SCA genes in the healthy brain [16], there is no detailed assessment of the relationships among polyQ genes in pathologically relevant regions of the brain.

We aim to find common mechanisms through which the nine polyQ genes could interact with each other in order to understand how these interactions could affect neuropathology in polyQ disorders. Our approach consisted of interrogating healthy brain gene expression data from the Allen Human Brain Atlas (AHBA) [18] as its high spatial resolution allows localizing interactions to specific brain regions. PolyQ genes with similar expression patterns in a specific brain area suggest their co-involvement in functions specific to that brain area. Using the AHBA to relate biological functions to disease and pinpoint pathways to specific regions of the brain has been reported previously, for example through evaluation of normal activity in the brain of genes associated to migraine [19], and autism [20–22]. Here, we follow a similar approach to detect functional relationships between the nine polyQ genes. We focus on a brain region consisting of several brain areas which were recently shown to be most severely affected in HD through an MRI-guided unbiased approach [23]. In addition, the analysis was repeated for seven anatomical brain structures to assess whether the relationships between polyQ genes within the HD-associated region are reflected in other anatomical brain structures (Figure 3.1). We analyzed gene co-expression networks within the different brain regions to identify region-specific relations among polyQ genes and their underlying functional mechanisms.
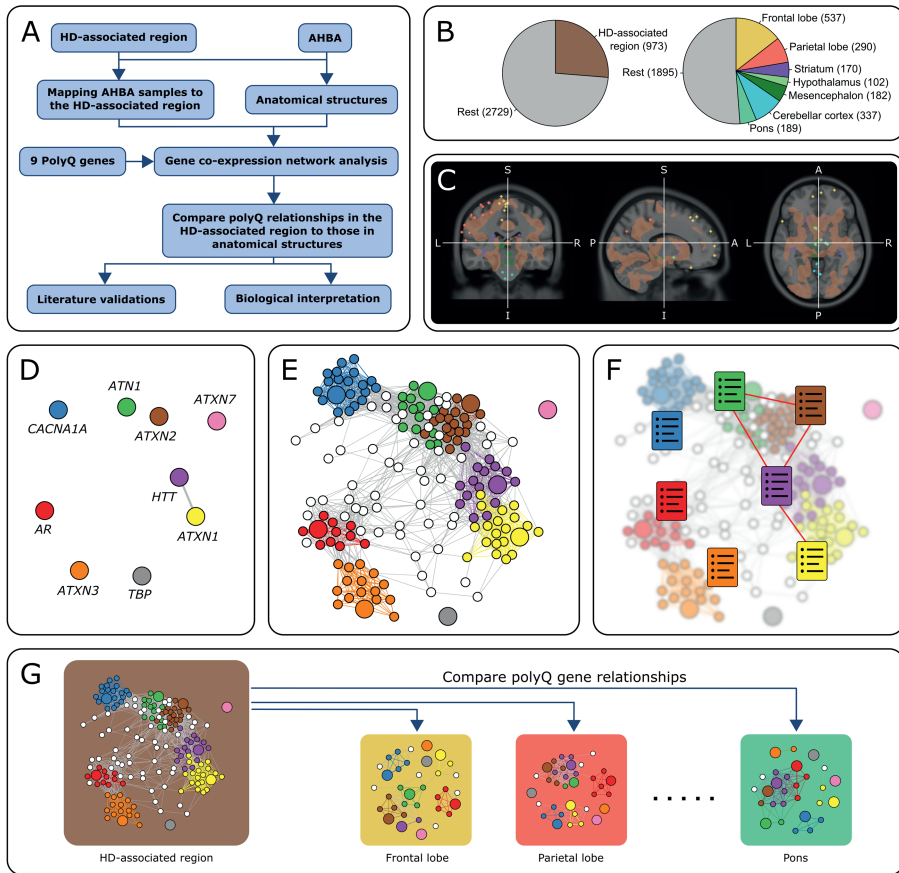
Figure 3.1 Overview of the approach. (A) Flowchart of the general approach in this study using co-expression analysis. (B) Number of samples in all six donors. The left chart shows the number of samples in the HD-associated region and the rest of the brain. The right chart shows the number of samples in seven anatomical brain structures and the rest of the brain. There are 3,702 samples in total. (C) Mapping AHBA samples to the HD-associated region. AHBA samples were spatially mapped to the HD-associated region (brown) which was based on an MRI study that looked at gray matter changes between HD patients and controls [23]. In this MRI image, we discarded all the samples that did not belong to one of the predefined regions. The colored points correspond to samples collected from anatomical structures that fall either inside or outside the HD-associated region. (D-F) Co-expression networks for the HD-associated region and each anatomical structure were constructed by retaining links between strongly correlated gene pairs (Pearson's $r > 0.5$). PolyQ genes are colored and have a larger node size. The relationships between polyQ genes were analyzed based on three methods: (D) Direct co-expression between polyQ genes. (E) Subnetworks of genes co-expressed with polyQ genes. Nodes are colored according to the co-expressed polyQ gene. (F) Functional enrichment of the set of genes co-expressed with each of the polyQ genes (colored charts). Functional links between polyQ genes were assessed based on the overlap of enriched functional terms. (G) PolyQ gene relationships in the HD-associated region were then compared to those in networks constructed for different anatomical structures.

3.2 MATERIAL AND METHODS

3.2.1 SPATIAL GENE EXPRESSION IN THE HEALTHY HUMAN BRAIN

We used the publicly available gene expression data set from AHBA [18] to exploit possibilities to study neurodegenerative diseases in the healthy brain. The spatially high-resolution data set contains genome-wide microarray expression of 3,702 samples collected from 6 healthy adult donors (5 males and 1 female, mean age 42, range 24-57 years). For each brain, 363-946 samples are available. In two out of the six brains, samples were collected from both hemispheres, while in the remaining four brains, samples were collected from the left hemisphere only.

The expression of several genes in the AHBA was assessed using multiple probes. We selected one probe per gene as follows: (1) if there was one probe for a gene, this probe was selected; (2) if there were two probes for a gene, we selected the probe with the highest variance (measured per brain and then averaged across the six brains); (3) if there were more than two probes for a gene, we chose the probe with the highest connectivity (measured as the sum of the Pearson correlations per brain and then averaged across the six brains). These steps resulted in 19,992 genes selected for further analysis.

3.2.2 MAPPING AHBA SAMPLES TO THE HD-ASSOCIATED REGION

In our analysis, we included a brain region affected in HD patients compared to healthy controls defined according to gray matter changes in MRI scans [23]. Several structural co-variance networks (SCNs), corresponding to discontinuous brain areas, were associated with HD. The SCN analysis was applied on whole brain images rather than predefined regions [24] to reveal a map of regions affected in HD. Two disjoint brain areas, referred to as the caudate nucleus and hippocampal network, showed strong significant association with pre-manifest HD and manifest HD patients. The caudate nucleus network includes the nucleus accumbens, pallidum, putamen, and precuneus. The hippocampal network includes the parahippocampal gyrus, cerebellum, pallidum, and planum polare. We combined these two regions into one, hereafter named "HD-associated region". AHBA samples were mapped to the HD-associated region using the MNI coordinate space. Based on this mapping, we identified AHBA samples located inside the HD-associated region that exhibits significant gray matter volume changes in HD compared to healthy controls.

3.2.3 DIFFERENTIAL GENE EXPRESSION IN THE HD-ASSOCIATED REGION

We examined gene expression differences between samples located inside and outside the HD-associated region. Differential expression was assessed for each brain separately using two one-tailed Mann-Whitney U-tests to identify up- and downregulated genes. Genes with a P-value lower than 0.025 in either the upper tail or lower tail in five out of six brains were considered differentially expressed.

3.2.4 REGION-SPECIFIC CO-EXPRESSION ANALYSIS

We used Pearson's correlation (r) as a measure of co-expression between two genes. Gene pairs showing a co-expression greater than 0.5 were considered to be related as they have similar expressions across tissue-specific samples. The co-expression threshold was based

on our observation of the polyQ co-expression distributions in different regions (Supplementary Figure 1). For each polyQ gene, a correlation threshold of 0.5 on average coincides with selecting the top 5% most co-expressed genes. Similar patterns of the number of co-expressed genes were observed when the co-expression threshold was set at 0.4 and 0.6 (Supplementary Figure 2). Co-expression was calculated for all possible gene pairs across samples inside the HD-associated region as well as samples representing one of the following brain structures: frontal lobe, parietal lobe, striatum, hypothalamus, mesencephalon, cerebellar cortex, and pons. Co-expression was first calculated per donor and then averaged across all donors to one consensus matrix containing all pairwise gene co-expressions.

3.2.5 Functional enrichment of co-expressed genes
We assessed enrichment in functional terms for sets of genes using DAVID Bioinformatics resources [25,26]. Enrichment analysis was done for genes differentially expressed in the HD-associated region (up- and downregulated) and sets of genes co-expressed with one of the nine polyQ genes. We selected the following annotation categories: GOTERM_BP_ALL, GOTERM_MF_ALL, GOTERM_CC_ALL. Enrichment analysis was done using Entrez IDs as gene identifiers and all 19,992 genes in the AHBA were added as a background list. Functional terms were selected when the Benjamini-Hochberg (BH) corrected *P*-value was lower than 0.05 and when at least two genes were present in the gene ontology category. To summarize the functional terms enriched in genes differentially expressed between the HD-region and the rest of the brain, we used the clustering function from DAVID and retained only terms with a cluster enrichment score > 2.

3.2.6 PolyQ interactions based on co-expression
We used three different representations of the relationships between polyQ genes, each highlighting a different way of how polyQ genes could be related. (1) Direct co-expression: co-expression between two polyQ genes indicates a strong interaction based on their spatial expression patterns. (2) Shared co-expression: a significant overlap of the two co-expressed gene sets of two polyQ genes suggests that these two polyQ genes are indirectly related to each other through other genes with which they interact. (3) Functional overlap: if two polyQ genes indirectly interact, the overlap between the functional terms enriched in their corresponding sets of co-expressed genes point out whether the two polyQ genes are also functionally related. We assessed the significance of shared co-expressed genes between two polyQ genes using one-tailed Fisher's exact test. We considered a functional overlap between two polyQ genes when their respective gene sets share at least 10 enriched functional terms.

3.2.7 Network analysis
The polyQ gene network for the HD-associated region was visualized using Gephi [27] and the rgexf R-package. In the network, nodes represent polyQ genes (each with a unique color) as well as genes co-expressed with polyQ genes (each assigned the color of the polyQ gene with which it is co-expressed). The node color is mixed when genes are co-expressed with

multiple polyQ genes. Edges are colored by the polyQ gene that they connect to. We used 'Fruchterman-Reingold' and 'label adjust' options for the layout. The circular co-expression plots were visualized using Cytoscape [28] and the RCy3 R-package [29]. Each node represents one of the nine polyQ genes and node colors indicate their mean expression across samples specific to the indicated regions per donor and then averaged across the six brains. Expression levels increase from blue to white to red. Edge thickness indicates the co-expression strength between two genes, the number of shared co-expressed genes, or number of shared functional terms between two polyQ gene sets.

3.2.8 ENRICHMENT OF UBIQUITIN LIGASES AND DNA BINDING GENES

Genes encoding ubiquitin conjugating enzymes (UBE2) were downloaded from the HUGO Nomenclature Committee [30] and includes 41 genes. DNA repair and ubiquitination gene sets, consisting of 125 and 40 genes, respectively, were obtained from the Molecular Signature Database (MSigDB) [31].

3.3 RESULTS

3.3.1 *ATXN2* IS DIFFERENTIALLY EXPRESSED IN THE HD-ASSOCIATED REGION

To assess normal gene expression changes in regions affected by HD, we examined the differential expression of genes in AHBA samples split according to being inside or outside the HD-associated region. Among the polyQ genes under study, only *ATXN2* was significantly downregulated in five out of six brains (Supplementary Figure 3). All other polyQ genes did not show consistent expression changes across the six brains. The analysis yielded 2,812 (14.1% of all 19,992 genes) significant genes (P-value < 0.025; one-sided Mann-Whitney U-test; Bonferroni-corrected). Out of the differentially expressed genes, 711 genes showed a significant higher expression in the HD-associated region (upregulated) and 2,101 genes showed a lower expression (downregulated). The set of downregulated genes was enriched in terms related to cytoplasm, mitochondrial processes, cellular component organization, DNA damage recognition, synapses, autophagy, and metabolic processes (BH-corrected P < 0.05; term cluster enrichment score > 2; Supplementary Table 1). These biological functions have been implicated in HD before [5,13,32–34]. Interestingly, there were no functional terms enriched in the set of upregulated genes. From Supplementary Figure 3 it can also be inferred that *ATXN2* had medium expression both inside and outside the HD-associated region in all six brains. Compared to all polyQ genes, *ATN1* had the highest expression across the brain, while *ATXN3* had the lowest expression.

3.3.2 *ATN1, ATXN2* AND *HTT* HAVE THE HIGHEST CONNECTIVITY AMONG POLYQ GENES

We examined the functional relations among the polyQ genes by analyzing their spatial co-expression using the high-resolution AHBA. Co-expression of the polyQ genes was analyzed in the healthy brain using the AHBA samples within regions associated with gray matter changes in HD patients (HD-associated region). In addition, we analyzed seven anatomical brain structures to assess whether polyQ gene co-expression patterns in the HD-associated

region are reflected within these structures. All anatomical structures had a sufficient number of samples to perform co-expression analysis (Table 3.1).

Table 3.1 Number of samples for the HD-associated region and seven anatomical brain structures across donors.

Donor	HD-associated region	Frontal lobe	Parietal Lobe	Striatum	Hypo-thalamus	Mesen-cephalon	Cerebellar cortex	Pons
9861	280	161	81	48	9	48	41	52
10021	183	135	57	46	22	61	76	50
12876	99	47	37	16	17	9	40	6
14380	139	70	42	24	20	22	44	26
15496	125	64	32	18	16	24	60	29
15697	147	60	41	18	18	18	76	26
Total samples	973	537	290	170	102	182	337	189

The HD-associated region overlaps with parts of different anatomical structures. None of the examined anatomical structures were overrepresented in the HD-associated region (Supplementary Table 2). We examined the cerebellar cortex instead of the cerebellum because of strong expression differences between the cerebellar cortex and cerebellar nuclei samples [35]. The cerebellar nuclei were left out of the co-expression analysis as they had too few samples. Since the HD-associated region showed little overlap of samples with anatomical structures, co-expression patterns in this region are not dominated by samples from any specific anatomical structures.

The number of co-expressed genes varied per polyQ gene and brain region (Figure 3.2). *ATN1* had the highest number of co-expressed genes in all examined regions except the mesencephalon and pons where only *HTT* had more co-expressed genes. The largest number of genes co-expressed with *ATN1* was observed in the striatum (3,428 genes), followed by the parietal lobe (3,332), and the frontal lobe (1,971). *ATXN2* had the second largest number of co-expressed genes in the parietal lobe, followed by the striatum and the frontal lobe. *TBP* had only one gene co-expressing in the cerebellar cortex. For all sets of genes co-expressed with polyQ genes we performed functional enrichment analysis to obtain sets of functional GO terms (BH-corrected $P < 0.05$). This was repeated for all inspected brain regions (Supplementary Figure 4). Similar patterns were observed for the number of functional terms and their respective gene set sizes, with *ATN1*, *ATXN2*, and *HTT* showing the highest number of GO terms.
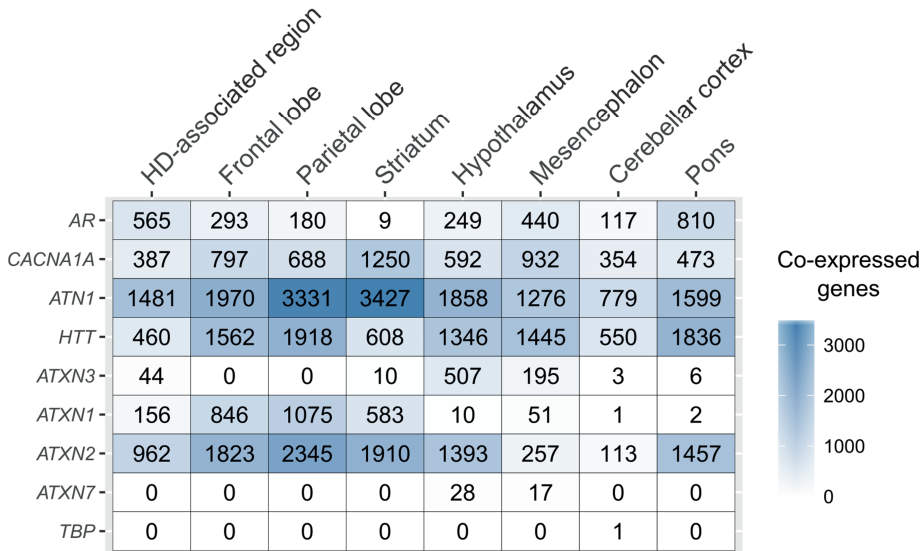
	HD-associated region	Frontal lobe	Parietal lobe	Striatum	Hypothalamus	Mesencephalon	Cerebellar cortex	Pons
AR	565	293	180	9	249	440	117	810
CACNA1A	387	797	688	1250	592	932	354	473
ATN1	1481	1970	3331	3427	1858	1276	779	1599
HTT	460	1562	1918	608	1346	1445	550	1836
ATXN3	44	0	0	10	507	195	3	6
ATXN1	156	846	1075	583	10	51	1	2
ATXN2	962	1823	2345	1910	1393	257	113	1457
ATXN7	0	0	0	0	28	17	0	0
TBP	0	0	0	0	0	0	1	0

Co-expressed genes

3000
2000
1000
0

Figure 3.2 PolyQ co-expressed gene sets across the brain. The number of genes co-expressed with each of the polyQ genes (rows) for the different, anatomical brain structures (columns).

3.3.3 STRONG CONNECTIVITY BETWEEN *ATN1* AND *ATXN2* IN THE HD-ASSOCIATED REGION

The co-expression network of polyQ genes within the HD-associated region includes all genes that are co-expressed with one of the nine polyQ genes (Figure 3.3). The numbers of shared co-expressed genes for different brain regions are shown in Figure 3.4A, and their respective co-expressed genes in the HD-associated region are listed in Supplementary Table 3. The functional overlap between two polyQ genes (Figure 3.4B) was measured as the number of shared enriched functional terms between their corresponding sets of co-expressed genes (Supplementary Figure 4).

In the HD-associated region, *ATN1-ATXN2* was the only polyQ gene pair that was directly co-expressed ($r = 0.53$, Figure 3.5). The same pair also shared most (490) co-expressed genes among all polyQ gene pairs. *TBP* and *ATXN7* were not co-expressed with other genes at all. While both genes show low mean expression levels in the examined regions, the variance is similar to other polyQ genes. Their low expression levels indicate that *TBP* and *ATXN7* are less active in these brain regions, suggesting that they are not functionally related to other polyQ genes, at least in the examined regions. The *AR* and *ATXN3* gene sets showed more indirect and distant co-expressions with other polyQ genes (Supplementary Figure 5). *HTT* shared many co-expressed genes with *ATN1*, *ATXN2*, and *ATXN1* (Figure 3.4A).

HTT-ATXN1
Cytoplasmic region
Presynapse
Dendrite
Regulation of GTPase activity
Intracellular signal transduction

HTT-ATXN2
Microtubule
Cellular component organization
Cellular protein modification process
Ubiquitin-dependent protein catabolic process
Cell cycle

ATN1-HTT
Intracellular transport
Ubiquitin protein ligase binding
Cellular protein localization
Ubiquitin-dependent protein catabolic process
Protein modification by small protein conjugation

ATN1-ATXN2
Cellular component biogenesis
Spliceosomal complex assembly
Small molecule binding
Cellular protein localization
Proteasomal protein catabolic process

Figure 3.3 PolyQ gene co-expression network in the HD-associated region. The network consists of 3,368 nodes representing genes. The poly genes are shown as larger nodes, each with a unique color. Genes were considered co-expressed when the calculated Pearson's correlations exceed 0.5. Smaller nodes are genes that are co-expressed with at least one of the nine polyQ genes. Their colors indicate the polyQ gene with which they are co-expressed. Genes co-expressed with multiple polyQ genes have a mixed color. Edges are colored according to the colors of the nodes they connect. In the HD-associated region, *ATN1-ATXN2* is the only gene pair that is directly co-expressed, while other polyQ genes are more distantly related through indirect relationships. In this network, 41 genes co-expressed with *HTT*, *ATN1*, and *ATXN2* and 488 co-expressed with *ATN1* and *ATXN2*.

3.3.4 POLYQ CO-EXPRESSION RELATIONSHIPS OVERLAP BETWEEN THE HD-ASSOCIATED REGION AND ANATOMICAL STRUCTURES INVOLVED IN HD-PATHOLOGY

Co-expression relationships of polyQ genes in the HD-associated region were also observed in other anatomical structures. The direct co-expression between *ATN1* and *ATXN2* in the HD-associated region was also observed in the frontal lobe ($r = 0.53$), parietal lobe ($r = 0.52$), as well as the striatum ($r = 0.56$) that showed the highest correlation (Figure 3.5A). In the striatum, *ATN1-ATXN2* was also the only co-expressed polyQ gene pair similar to the HD-associated region. *ATN1* and *ATXN2* had the highest correlations in the striatum and they also had the highest number of overlapping co-expressed genes and functional terms in the striatum. This indicates that the co-expression patterns between *ATN1* and *ATXN2* found in the HD-associated region are particularly pronounced in the striatum.

There is a significant overlap between sets of co-expressed genes of *HTT*, *ATN1*, and *ATXN2* across all studied regions (Figure 3.4A and Figure 3.5B). When considering the overlap between sets of enriched functional GO terms (Figure 3.4B and Figure 3.5C), the triangular relationship between these three genes is only observed in the frontal and parietal lobes of the cerebral cortex. In the striatum, we observe a very strong functional overlap between *ATN1* and *ATXN2*. Together, these observations suggest a strong regional and functional relatedness between the three genes, rather than just the measured co-expression.

In addition, *ATXN1* shows significant overlap of functional terms with *HTT* in the HD-associated region, frontal lobe, parietal lobe, and striatum. *ATXN1* was also found to share functional terms with *ATN1*, *ATXN2*, and *HTT*, forming a clique of four polyQ genes in the parietal lobe. Finally, the co-expression patterns in the HD-associated region seemed to be dominated by the co-expression patterns within the frontal lobe, parietal lobe, and striatum, which were also the regions described as part of the HD-associated region in the imaging study (Supplementary Figure 6) [23].

3.3.5 *HTT*, *ATN1*, AND *ATXN2* ARE ASSOCIATED THROUGH BINDING, LOCALIZATION, AND REGULATION OF CELLULAR COMPONENTS

To gain insight into the functional relationships between *HTT*, *ATN1*, and *ATXN2*, we examined the overlap of functional terms between all three genes (Supplementary Table 4). These three genes share 15 functional terms, in the frontal and the parietal lobe, showing that they are involved in binding of cellular components. In the striatum, they are involved in the positive regulation of catalytic activity and molecular function, while in the hypothalamus, they are involved in DNA repair and cell cycle process and regulation. In the pons, they are involved in the cytoskeleton and regulation of cellular component organization. In the HD-associated region, they are involved in ubiquitin protein ligase binding, microtubule cytoskeleton, and cellular protein and macromolecule localization. In the mesencephalon and cerebellar cortex *HTT*, *ATN1*, and *ATXN2* do not share functional terms together. Together these overlapping functions between *HTT*, *ATN1*, and *ATXN2* suggest that their connectivity in the aforementioned brain regions is involved in the binding, localization, and regulation of cellular components.

We focused on genes involved in ubiquitination or DNA repair pathways, because these pathways have been previously associated with polyQ diseases [16,36,37]. In the HD-associated region, eight *UBE2* genes that are part of the ubiquitin conjugating enzymes gene family, *UBE2G2*, *UBE2I*, *UBE2C*, *CDC34*, *UBE2W*, *UBE2Z*, *UBE2E3*, and *UBE2D4*, co-expressed with polyQ genes. Furthermore, several polyQ genes co-expressed with DNA repair genes, especially *HTT*, *ATN1*, and *ATXN2*. Most DNA repair and ubiquitination genes were co-expressed with *ATN1* in the striatum (34 and 11 genes, respectively; Supplementary Figure 7). In summary, ubiquitin-related genes and DNA repair genes co-expressed with highly connected polyQ genes *HTT*, *ATN1*, and *ATXN2* in regions associated with HD.
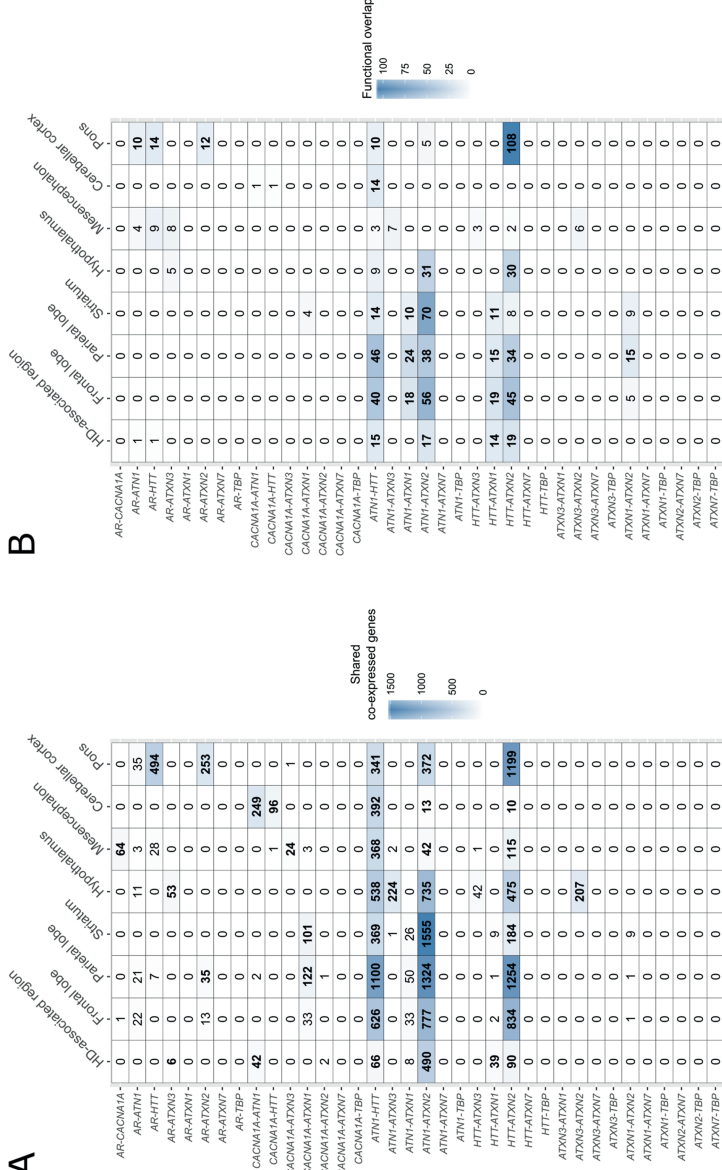
Figure 3.4 Overlap of co-expressed gene sets and functional terms between polyQ genes across the brain. (A) Number of shared co-expressed genes between each polyQ gene pair in different brain regions. Significant overlap of co-expressed genes is indicated in bold numbers (Fisher's exact test; $P < 0.05$). The respective genes co-expressed with two polyQ gene pair in the HD-associated region are given in Supplementary Table 3. (B) Functional overlap between each polyQ gene pair in different brain regions measured as the number of shared enriched functional terms between their corresponding sets of co-expressed genes. Functional overlaps of at least 10 functional terms are indicated in bold.
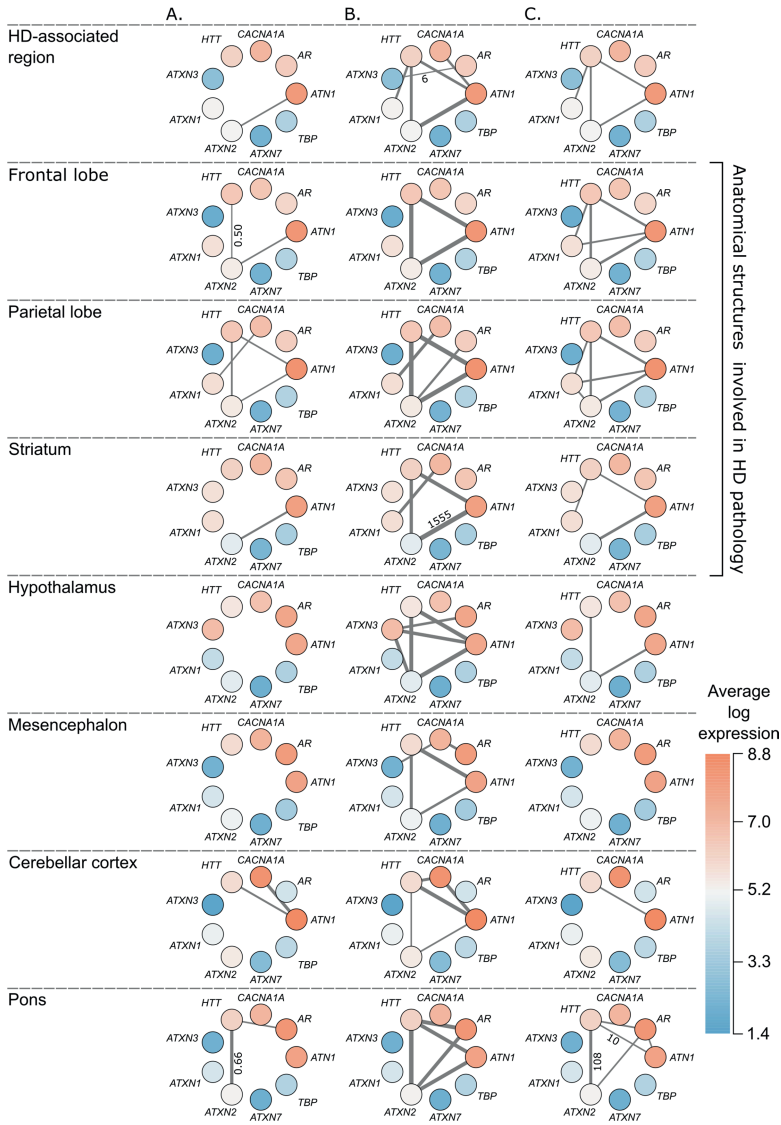
Figure 3.5 Overview of polyQ relationships in the HD-associated region and seven anatomical structures. Each node represents a polyQ gene and the color represents the expression of that polyQ gene averaged across region-specific samples and the six donors (red indicates high expression and blue low expression). Edge width indicates the strength of the detected relationship between polyQ genes. Each column represents a different type of relationship: (A) direct co-expression between polyQ genes, edge width corresponds to the co-expression value; (B) shared co-expression, edge width indicates the overlap of co-expressed genes, only for significant overlaps ($P < 0.05$); (C) functional overlap, edge width indicates the overlap size between two polyQ gene sets shown only for overlaps of at least 10 enriched terms. For each type of relationship the lowest and highest value are given along the thinnest and thickest edge.

3.4 Discussion

Polyglutamine diseases share a similar genetic basis and their phenotypes, such as AAO, are affected by CAG repeat variations in the normal range in polyQ genes other than the causative gene [7–12]. The causative polyQ genes are thought to modulate protein interactions through the role of CAG tracts in stabilizing protein interactions [38]. We analyzed the relationships between nine polyQ genes based on their co-expression within different regions of the healthy human brain.

We performed co-expression analysis in a brain region previously associated with HD specific damage in the cortical and subcortical gray matter. By combining imaging data from HD mutation carriers with spatial gene expression patterns from healthy brains, we were able to study the role of interactions between polyQ genes in the neuropathology of HD. Seven anatomical brain structures were also included in the analysis to examine regional differences compared to the HD-associated region. The anatomical structures also included the striatum, the main affected structure in HD. However, co-expression patterns in the striatum may be more specific to the structure itself rather than to HD pathology.

Among the nine polyQ genes, the relationship between *ATN1* and *ATXN2* was most pronounced based on our three representations of gene associations: it was the only gene pair that directly co-expressed, shared the most co-expressed genes, and also showed functional overlap in the HD-associated region. The same polyQ gene pair also shared co-expressed genes, and a functional overlap with *HTT*, suggesting functional connectivity with the causative gene of HD. Furthermore, out of the nine polyQ genes, only *ATXN2* had a significant lower expression in the HD-associated region compared to the rest of the brain. An expanded polyQ tract in the huntingtin protein may dysregulate pathways in which all three polyQ genes are involved. The function of *ATXN1* may also be affected in HD as this gene also shared co-expressed genes and showed functional overlap with *HTT* in the HD-associated region.

The co-expression patterns among polyQ genes within the HD-associated region overlapped with those observed within anatomical structures involved in HD neuropathology. In the striatum, the main structure affected in HD, *ATN1* and *ATXN2* were directly co-expressed and shared co-expressed genes. This suggests that this gene relationship has a more important role in HD than in other polyQ diseases. For the frontal and parietal lobes, *ATN1* and *ATXN2* were directly co-expressed and shared functional terms together with *HTT*, similar as for the patterns observed for the HD-associated region. The frontal lobe includes the primary motor cortex BA4, and is located next to the parietal lobe, which includes other motor areas. The motor cortex, involved in planning, control, and execution of voluntary movement, has been previously implicated in HD [39–42]. The motor cortex is also known to be directly connected to the caudate nucleus of the striatum [43]. In a previous study, HD gene expression profiles in the caudate nucleus and motor cortex were strikingly similar [13], suggesting that similar molecular mechanisms in different brain regions are involved in neurodegeneration. These mechanisms may include polyQ genes as we find similar co-

expression patterns in the frontal lobe and striatum between *HTT*, *ATN1*, *ATXN2*, and *ATXN1* (Figure 3.5).

There was a significant overlap in co-expressed genes between *HTT*, *ATN1*, and *ATXN2* in all examined brain regions. It would be interesting to evaluate whether this triangular relation between *HTT*, *ATN1*, and *ATXN2* is dysregulated in all polyQ diseases, especially HD, DRPLA, and SCA2. In addition, we observed strong connectivity of *ATXN2* and *AR* with *HTT* in the pons that has been described to undergo atrophy in SCA2 [44]. This relationship may possibly be disrupted in the pons of SCA2 patients. Altogether, these region-specific interactions demonstrate that co-expression analysis can reveal many interesting relations between genes based on their spatial information.

Accumulating evidence suggests that normal CAG repeat size variations in polyQ genes could act as genetic modifiers of AAO in different polyQ diseases. For example, the normal CAG repeat length in *ATXN3*, known to have a deubiquitinating function [45], has been found to have a positive effect on HD progression as it was associated with a later AAO [12]. There is also evidence for a polyQ-length dependent interference of both the mutant and normal proteins with a range of other protein binding partners [46–52]. Mutant proteins can interact either more strongly or weakly with other polyQ containing proteins and thereby inhibit there physiological function [49,52]. A case in this regard is a recent study which showed that the polyQ domain of wild-type ataxin 3 enables it to interact with beclin 1, a key initiator of autophagy, allowing the deubiquitinase activity of ataxin 3 to protect beclin 1 from proteasome-mediated degradation and thereby enabling autophagy [49]. They demonstrated that mutant huntingtin polyQ fragments competed with ataxin 3 for interaction with beclin 1 in a polyQ-length dependent manner, thereby inhibiting autophagy and contributing to neuronal dysfunction. This finding may explain the protective role of larger *ATXN3* CAG repeat sizes in HD [12]. In our analysis we did not find any relation in co-expression between *HTT* and *ATXN3* in the examined regions. This might be due to the polyQ length-dependent competing between huntingtin and ataxin 3 being relevant only with pathologically expanded polyQ tracts, which were not present in AHBA. Furthermore, atrophin 1 and the androgen receptor have been found to bind to beclin 1 in a polyQ length-dependent way. Among the polyQ genes and examined brain regions in our study, we found that *BECN1* is only co-expressed with *ATN1* in the pons. Several putative interactions with *HTT* reported in previous studies were also supported by the co-expression analysis in our study. In the HD-associated region, *ATN1* and *ATXN2* shared co-expression partners with *HTT* and both genes were also functionally related to *HTT*. For each of the genes it was shown before that variations in the normal CAG repeat length, together with the expanded CAG repeat in *HTT*, affect the AAO in HD [7]. The co-expression patterns between *HTT*, *ATN1*, and *ATXN1* suggest a functional relationship in brain regions that are involved in HD pathology.

Ubiquitin may be involved in HD pathogenesis through the relationships between *HTT*, *ATN1*, and *ATXN2*. The ubiquitin-proteasome system (UPS) has been linked extensively to

the pathogenesis of neurodegenerative diseases, including HD and other SCAs [1,6,36,53–55]. Aggregates in polyQ diseases show parts of ubiquitin and several important homeostatic proteins [56]. We found that E3 ubiquitin ligase genes *UBR4* and *CHIP/STUB1*, and ubiquitin-activating *UBA1* gene to be co-expressed with *HTT*, *ATN1*, and *ATXN2*. This suggests that the three polyQ genes contribute to the UPS in the HD-associated region of the healthy brain. Biochemical properties of *CHIP* are well-studied and this gene is known to interact with mutant polyQ proteins (including *HTT*, *ATXN1*, *ATXN3*, and *AR*) suppressing polyQ aggregation by promoting proteasomal degradation [1,2,57–59]. *CHIP* has also been implicated in several SCAs indicating it is directly involved in disease pathology [60]. The gene *UB4R* has a fundamental role in calcium signaling and neuronal survival and may contribute to neurodegenerative conditions [61]. This gene has been identified as a candidate locus in early-onset episodic ataxia [62]. It is not known whether recognition of aberrant proteins by ubiquitin ligases is beneficial or disadvantageous [63]. *UBA1* is an important regulator of cellular protein homeostasis and contributes to the pathogenesis in SBMA and HD [32,64]. Altogether, we found genes involved in the UPS that co-expressed with multiple polyQ genes which have been associated with similar pathologies in neurodegenerative disorders. The UPS is involved in protein homeostasis and may be affected early in HD through strong regional interactions. If there are changes in huntingtin protein function due to polyQ expansion, then the interaction between *ATN1* and *ATXN2* is likely to be affected. The low expression of these genes in certain brain regions could indicate their increased vulnerability to dysregulations of the UPS. We suspect co-expressed polyQ genes to have a role in AAO, as they and their shared co-expression partners seem to be involved in functions that have been associated to polyQ diseases.

There are several issues that limit a one-to-one comparison between the co-expression relationships reported in the present study and phenotype-genotype associations described previously [7–12]. First, we do not expect all phenotype-genotype associations to be explained by changes in co-expression patterns. Second, there is a wide degree of heterogeneity in both the data and methods used in the genetic association studies [7–12], limiting the possibility to combine their findings. Third, thus far, not all possible pairs of polyQ genes have been tested in the genetic association studies, highlighting the need for more comprehensive genetic association studies in larger cohorts of patients with polyQ disorders.

To validate our findings in other brain gene expression data we need a dataset with a spatial sampling resolution that allows co-expression analysis within substructures of the brain. Although the UK Brain Expression Consortium [15] sampled multiple (10) brain regions per healthy individual, the expression within a substructure, e.g., the cerebellum, is still represented by a single sample. Gene expression data of HD patients may be used to observe whether co-expression patterns are altered in disease state. However, existing datasets are particularly rich in the number of individuals they sampled and not brain regions [13]. This captures variation across individuals, while in our gene co-expression networks we capture spatial variation within a brain region of interest (e.g., striatum).

3.5 CONCLUSION

We showed that polyQ genes are co-expressed in the healthy brain and that their relationships are also specific to certain brain regions including a region associated with HD. Our aim was to find co-expression patterns between polyQ genes in different brain regions.The co-expression networks are likely altered in polyQ diseases due to interaction changes, especially in brain regions associated with the disease. The fact that these findings could not be validated on other expression datasets implies the importance of follow-up studies to understand more about the mechanisms behind polyQ diseases. We show that gene expression in the healthy brain may render specific regions vulnerable to expression changes based on gene co-expression networks.

3.6 ACKNOWLEDGEMENTS

REFERENCES

[1]   A. J. Williams and H. L. Paulson, "Polyglutamine neurodegeneration: protein misfolding revisited," *Trends Neurosci.* 31, 521–528 (2008).

[2]   N. R. Jana, P. Dikshit, A. Goswami, S. Kotliarova, S. Murata, K. Tanaka, and N. Nukina, "Co-chaperone CHIP associates with expanded polyglutamine protein and promotes their degradation by proteasomes," *J. Biol. Chem.* 280, 11635–11640 (2005).

[3]   B. P. C. van de Warrenburg, R. J. Sinke, C. C. V. Bemelmans, and H. Scheffer, "Spinocerebellar ataxias in the Netherlands," *Neurology* 58, 702–708 (2002).

[4]   J. S. Paulsen, J. D. Long, C. A. Ross, D. L. Harrington, C. J. Erwin, J. K. Williams, H. J. Westervelt, H. J. Johnson, E. H. Aylward, et al., "Prediction of manifest Huntington's disease with clinical and imaging measures: A prospective observational study," *Lancet Neurol.* 13, 1193–1201 (2014).

[5]   D. C. Rubinsztein, "Lessons from animal models of Huntington's disease," *Trends Genet.* 18, 202–209 (2002).

[6]   M. Arrasate and S. Finkbeiner, "Protein aggregates in Huntington's disease," *Exp. Neurol.* 238, 1–11 (2012).

[7]   D. Hmida-Ben Brahim, M. Chourabi, S. Ben Amor, I. Harrabi, S. Trabelsi, M. Haddaji-Mastouri, M. Gribaa, S. Sassi, F. E. Gahbiche, et al., "Modulation at age of onset in tunisian Huntington disease patients: implication of new modifier genes," *Genet. Res. Int.* 2014, 210418 (2014).

[8]   S. Tezenas, A. Durr, P. Bauer, K. P. Figueroa, C. Mariotti, B. P. C. van de Warrenburg, L. Orsi, and P. Giunti, "Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes," *Brain* 137, 2444–2455 (2014).

[9]   Z. Chen, C. Zheng, Z. Long, L. Cao, X. Li, H. Shang, X. Yin, B. Zhang, J. Liu, et al., "(CAG)n loci as genetic modifiers of age-at-onset in patients with Machado-Joseph disease from mainland China," *Brain*, 1–5 (2016).

[10]   R. M. De Castilhos, G. V. Furtado, T. C. Gheno, P. Schaeffer, A. Russo, O. Barsottini, J. L. Pedroso, D. Z. Salarini, F. R. Vargas, et al., "Spinocerebellar ataxias in Brazil - Frequencies and modulating effects of related genes," *Cerebellum* 13, 17–28 (2014).

[11]   S. M. Pulst, N. Santos, D. Wang, H. Yang, D. Huynh, L. Velazquez, and K. P. Figueroa, "Spinocerebellar ataxia type 2: PolyQ repeat variation in the CACNAIA calcium channel modifies age of onset," *Brain* 128, 2297–2303 (2005).

[12]   G. Stuitje, M. J. van Belzen, S. L. Gardiner, W. M. C. van Roon-Mom, M. W. Boogaard, S. J. Tabrizi, R. A. C. Roos, N. A. Aziz, and N. A. Aziz, "Age of onset in Huntington's disease is influenced by CAG repeat variations in other polyglutamine disease-associated genes," *Brain*, 10–12 (2017).

[13]   A. Hodges, A. D. Strand, A. K. Aragaki, A. Kuhn, T. Sengstag, G. Hughes, L. A. Elliston, C. Hartog, D. R. Goldstein, et al., "Regional and cellular gene expression changes in human Huntington's disease brain," *Hum. Mol. Genet.* 15, 965–977 (2006).

[14]   P. Langfelder, J. P. Cantle, D. Chatzopoulou, N. Wang, F. Gao, I. Al-Ramahi, X. Lu, E. M. Ramos, K. El-zein, et al., "Integrated genomics and proteomics define huntingtin CAG length – dependent networks in mice," *Nat. Neurosci.*, 1–11 (2016).

[15]   D. Trabzuni, M. Ryten, R. Walker, C. Smith, S. Imran, A. Ramasamy, M. E. Weale, and J. Hardy, "Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies," *J. Neurochem.* 119, 275–282 (2011).

[16]   C. Bettencourt, M. Ryten, P. Forabosco, S. Schorge, J. Hersheson, J. Hardy, and H. Houlden, "Insights from cerebellar transcriptomic analysis into the pathogenesis of ataxia.," *JAMA Neurol.* 71, 831–839 (2014).

[17]   A. Neueder and G. P. Bates, "A common gene expression signature in Huntington's disease patient brain regions," *BMC Med. Genom.* 7, 1–23 (2014).

[18]   M. Hawrylycz, J. A. Miller, V. Menon, D. Feng, T. Dolbeare, A. L. Guillozet-Bongaarts, A. G. Jegga, B. J. Aronow, C.-K. K. Lee, et al., "Canonical genetic signatures of the adult human brain," *Nat. Neurosci.* 18, 1832–1844 (2015).

[19]   E. Eising, S. M. H. Huisman, A. Mahfouz, L. S. Vijfhuizen, V. Anttila, B. S. Winsvold, and T. Kurth, "Gene co-expression analysis identifies brain regions and cell types involved in migraine pathophysiology: a GWAS-based study using the Allen Human Brain Atlas," *Hum. Genet.* 135, 425–439 (2016).

[20]   A. Mahfouz, M. N. Ziats, O. M. Rennert, B. P. F. Lelieveldt, and M. J. T. Reinders, "Shared Pathways Among Autism Candidate Genes Determined by Co-expression Network Analysis of the Developing Human Brain Transcriptome," *J. Mol. Neurosci.* 57, 580–594 (2015).

[21]   A. J. Willsey, S. J. Sanders, M. Li, S. Dong, A. T. Tebbenkamp, R. A. Muhle, S. K. Reilly, L. Lin, S. Fertuzinhos, et al., "Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism," *Cell* 155, 997–1007 (2013).

[22]   N. N. Parikshak, R. Luo, A. Zhang, H. Won, J. K. Lowe, V. Chandran, S. Horvath, and D. H. Geschwind, "Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism," *Cell* 155, 1008–1021 (2013).

[23]   E. M. Coppen, J. van der Grond, A. Hafkemeijer, S. A. R. B. Rombouts, and R. A. C. Roos, "Early grey matter changes in structural covariance networks in Huntington's disease," *NeuroImage Clin.* 12, 806–814 (2016).

[24]   L. Minkova, S. B. Eickhoff, A. Abdulkadir, C. P. Kaller, J. Peter, E. Scheller, J. Lahr, R. A. Roos, A. Durr, et al., "Large-scale brain network abnormalities in Huntington's disease revealed by structural covariance," *Hum. Brain Mapp.* 37, 67–80 (2016).

[25] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nat. Protoc.* 4, 44–57 (2008).

[26] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists," *Nucleic Acids Res.* 37, 1–13 (2009).

[27] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," *Third Int. AAAI Conf. Weblogs Soc. Media*, 361–362 (2009).

[28] P. Shannon, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Res.* 13, 2498–2504 (2003).

[29] P. T. Shannon, M. Grimes, B. Kutlu, J. J. Bot, and D. J. Galas, "RCytoscape: tools for exploratory network analysis," *BMC Bioinformatics* 14, 217 (2013).

[30] K. A. Gray, B. Yates, R. L. Seal, M. W. Wright, and E. A. Bruford, "Genenames.org: the HGNC resources in 2015," *Nucleic Acids Res.* 43, 1079–1085 (2015).

[31] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, and M. A. Gillette, "Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide," *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550 (2005).

[32] E. J. N. Groen and T. H. Gillingwater, "UBA1: At the crossroads of ubiquitin homeostasis and neurodegeneration," *Trends Mol. Med.* 21, 622–632 (2015).

[33] L. Jones, H. Houlden, and S. J. Tabrizi, "DNA repair in the trinucleotide repeat disorders," *Lancet Neurol.* 16, 1234 (2017).

[34] T. Seredenina and R. Luthi-Carter, "What have we learned from gene expression profiles in Huntington's disease?," *Neurobiol. Dis.* 45, 83–98 (2012).

[35] S. M. H. Huisman, B. van Lew, A. Mahfouz, N. Pezzotti, L. Michielsen, A. Vilanova, M. J. T. Reinders, H. Thomas, and P. F. Lelieveldt, "BrainScope : interactive visual exploration of the spatial and temporal human brain transcriptome," *Nucleic Acids Res.* 45, e83 (2017).

[36] C. Bettencourt, D. Hensman-Moss, M. Flower, S. Wiethoff, A. Brice, C. Goizet, G. Stevanin, and G. Koutsis, "DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases," *Ann. Neurol.* 79, 983–990 (2016).

[37] Z. Chen, C. Wang, C. Zheng, Z. Long, L. Cao, X. Li, H. Shang, X. Yin, B. Zhang, et al., "Ubiquitin-related network underlain by (CAG)n loci modulate age at onset in Machado-Joseph disease," *Brain* 140, 1–5 (2017).

[38] M. H. Schaefer, E. E. Wanker, and M. A. Andrade-Navarro, "Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks," *Nucleic Acids Res.* 40, 4273–4287 (2012).

[39] L. Fogassi and G. Luppino, "Motor functions of the parietal lobe," *Curr. Opin. Neurobiol.* 15, 626–631 (2005).

[40] G. Pfurtscheller and A. Berghold, "Patterns of cortical activation during planning of voluntary movement," *Electroencephalogr. Clin. Neurophysiol.* 72, 250–258 (1989).

[41] C. M. Stinear, J. P. Coxon, and W. D. Byblow, "Primary motor cortex and movement prevention: Where Stop meets Go," *Neurosci. Biobehav. Rev.* 33, 662–673 (2009).

[42] P. Nachev, C. Kennard, and M. Husain, "Functional role of the supplementary and pre-supplementary motor areas.," *Nat. Rev. Neurosci.* 9, 856–869 (2008).

[43] R. Muhammad, J. D. Wallis, and E. K. Miller, "A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum.," *J. Cogn. Neurosci.* 18, 974–989 (2006).

[44] S. H. Ying, S. I. Choi, S. L. Perlman, R. W. Baloh, D. S. Zee, and A. W. Toga, "Pontine and cerebellar atrophy correlate with clinical disability in SCA2," *Neurology* 66, 424–426 (2006).

[45] C. A. Matos, S. de Macedo-Ribeiro, and A. L. Carvalho, "Polyglutamine diseases: The special case of ataxin-3 and Machado-Joseph disease," *Prog. Neurobiol.* 95, 26–48 (2011).

[46] C. L. Benn, T. Sun, G. Sadri-Vakili, K. N. McFarland, D. P. DiRocco, G. J. Yohrling, T. W. Clark, B. Bouzou, and J.-H. J. Cha, "Huntingtin modulates transcription, occupies gene promoters in vivo, and binds directly to DNA in a polyglutamine-dependent manner.," *J. Neurosci.* 28, 10720–10733 (2008).

[47] L. S. Kaltenbach, E. Romero, R. R. Becklin, R. Chettier, R. Bell, A. Phansalkar, A. Strand, C. Torcassi, J. Savage, et al., "Huntingtin interacting proteins are genetic modifiers of neurodegeneration," *PLoS Genet.* 3, 689–708 (2007).

[48] X.-J. Li, M. Friedman, and S. Li, "Interacting proteins as genetic modifiers of Huntington disease.," *Trends Genet.* 23, 531–533 (2007).

[49] A. Ashkenazi, C. F. Bento, T. Ricketts, M. Vicinanza, F. Siddiqi, M. Pavel, F. Squitieri, M. C. Hardenberg, S. Imarisio, et al., "Polyglutamine tracts regulate beclin 1-dependent autophagy," *Nature* 545, 108–111 (2017).

[50] M. Yokoshi, Q. Li, M. Yamamoto, H. Okada, Y. Suzuki, and Y. Kawahara, "Direct binding of ataxin-2 to distinct elements in 3' UTRs promotes mRNA stability and protein expression," *Mol. Cell* 55, 186–198 (2014).

[51] J. Lim, J. Crespo-Barreto, P. Jafar-Nejad, A. B. Bowman, R. Richman, D. E. Hill, H. T. Orr, and H. Y. Zoghbi, "Opposing effects of polyglutamine expansion on native protein complexes contribute to SCA1," *Nature* 452, 713–718 (2008).

[52] L. A. Becker, B. Huang, G. Bieri, R. Ma, D. A. Knowles, P. Jafar-Nejad, J. Messing, H. J. Kim, A. Soriano, et al., "Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice," *Nature* 544, 367–371 (2017).

[53] N. P. Dantuma and L. C. Bott, "The ubiquitin-proteasome system in neurodegenerative diseases: precipitating factor, yet part of the solution.," *Front. Mol. Neurosci.* 7, 70 (2014).

[54] Z. Ortega and J. J. Lucas, "Ubiquitin-proteasome system involvement in Huntington's disease," *Front. Mol. Neurosci.* 7, 1–11 (2014).

[55] G. Atkin and H. Paulson, "Ubiquitin pathways in neurodegenerative disease.," *Front. Mol. Neurosci.* 7, 63 (2014).

[56] A. B. Bowman, S. Y. Yoo, N. P. Dantuma, and H. Y. Zoghbi, "Neuronal dysfunction in a polyglutamine disease model occurs in the absence of ubiquitin-proteasome system impairment and inversely correlates with the degree of nuclear inclusion formation," *Hum. Mol. Genet.* 14, 679–691 (2005).

[57] E. F. E. Kuiper, E. P. de Mattos, L. B. Jardim, H. H. Kampinga, and S. Bergink, "Chaperones in polyglutamine aggregation: Beyond the Q-stretch," *Front. Neurosci.* 11, 1–11 (2017).

[58] V. M. Miller, R. F. Nelson, C. M. Gouvion, A. Williams, E. Rodriguez-lebron, S. Q. Harper, B. L. Davidson, M. R. Rebagliati, and H. L. Paulson, "CHIP suppresses polyglutamine aggregation and toxicity in vitro and in vivo," *J. Neurosci.* 25, 9152–9161 (2005).

[59] H. Adachi, M. Waza, K. Tokui, M. Katsuno, M. Minamiyama, F. Tanaka, M. Doyu, and G. Sobue, "CHIP overexpression reduces mutant androgen receptor protein and ameliorates phenotypes of the spinal and bulbar muscular atrophy transgenic mouse model," *J. Neurosci.* 27, 5115–5126 (2007).

[60] S. M. Ronnebaum, C. Patterson, and J. C. Schisler, "Emerging evidence of coding mutations in the ubiquitin-proteasome system associated with cerebellar ataxias.," *Hum. Genome Var.* 1, 14018 (2014).

[61] K. Parsons, Y. Nakatani, and M. D. Nguyen, "P600/UBR4 in the central nervous system," *Cell. Mol. Life Sci.* 72, 1149–1160 (2015).

[62] J. Conroy, P. McGettigan, R. Murphy, D. Webb, S. M. Murphy, B. McCoy, C. Albertyn, D. McCreary, C. McDonagh, et al., "A novel locus for episodic ataxia: UBR4 the likely candidate," *Eur. J. Hum. Genet.* 22, 505–510 (2014).

[63] D. Chhangani, N. R. Jana, and A. Mishra, "Misfolded proteins recognition strategies of e3 ubiquitin ligases and neurodegenerative diseases," *Mol. Neurobiol.* 47, 302–312 (2013).

[64] H. Y. Liu and C. M. Pfleger, "Mutation in E1, the ubiquitin activating enzyme, reduces drosophila lifespan and results in motor impairment," *PLoS One* 8, 1–10 (2013).