Selective Vulnerability of Layer 5a Corticostriatal Neurons in Huntington's Disease

Christina Pressl[1], Kert Mätlik [1], Laura Kus [1], Paul Darnell [1], Ji-Dung Luo [2], Matthew R. Paul[2], Alison R. Weiss[3], William Liguore[3], Thomas S. Carroll [2], David A. Davis[4], Jodi McBride[3], Nathaniel Heintz [1,*]

[1] Laboratory of Molecular Biology, The Rockefeller University, New York, United States

[2] Bioinformatics Resource Center, The Rockefeller University, New York, United States

[3] Division of Neuroscience, Oregon National Primate Research Center, Beaverton, United States

[4] Department of Neurology, University of Miami Miller School of Medicine, Miami, Florida, United States

[*] corresponding author

Email corresponding author: heintz@rockefeller.edu

ABSTRACT

The properties of the cell types that are most vulnerable in Huntington's disease (HD) cortex, the nature of somatic CAG expansions of *mHTT* in these cells, and their importance in CNS circuitry have not been delineated. Here we employed serial fluorescence activated nuclear sorting (sFANS), deep molecular profiling, and single nucleus RNA sequencing (snRNAseq) to demonstrate that layer 5a pyramidal neurons are selectively vulnerable in primary motor cortex and other cortical areas of HD donors. Extensive *mHTT*-CAG expansions occur in vulnerable layer 5a pyramidal cells, and in Betz cells, layer 6a, layer 6b neurons that are not lost in HD. Retrograde tracing experiments in macaque brains identify the vulnerable layer 5a neurons as corticostriatal pyramidal cells. Our data establish that *mHTT*-CAG expansion is not sufficient for cell loss in the cerebral cortex of HD, suggesting that cortico-striatal disconnection in HD patients may play an important role in neurodegeneration.

KEYWORDS

Huntington's Disease, human post-mortem cortex, serial fluorescence activated nuclei sorting, cell-type specific vulnerability, cortical cell-type specific CAG repeat expansion profiles, RNA-Seq, ATAC-Seq, single-nuclei RNA-Seq, RNAscope

INTRODUCTION

Neurodegeneration in Huntington's disease (HD) is evident primarily in the striatum in early stages, and it is thought that loss of medium–sized spiny neurons (MSNs) in the caudate nucleus and putamen is responsible for motor abnormalities that are cardinal features of the disease (Reiner and Deng, 2018; Vonsattel et al., 2011; Vonsattel et al., 1985). The clinical appearance of motor symptoms is used as a measure of age at onset, which is inversely correlated with the number of CAG repeats in exon 1 of the mutant *HTT* allele (*mHTT*), often defined by the CAG tract length in blood (Andrew et al., 1993; Hong et al., 2021; Monckton, 2021; Ross and Tabrizi, 2011). Neuroimaging studies of the human cerebral cortex have demonstrated that connectivity between cortical areas and the basal ganglia is disturbed very early in the HD individuals (Espinoza et al., 2018; Gargouri et al., 2016; McColgan et al., 2017) (Jernigan et al., 1991) (Hasselbalch et al., 1992; Jenkins et al., 1998; Kuhl et al., 1982; Kuwert et al., 1989, 1990; Sax et al., 1996; Unschuld et al., 2012), and that widespread cortical degeneration occurs as the disease progresses (Johnson et al., 2021). Although *HTT* is ubiquitously expressed (Landwehrmeyer et al., 1995), substantial variations in mood, motor function, and cognition in HD are associated with region specific neuropathological findings occurring in the prefrontal, motor and cingulate cortices (Hickman et al., 2022; Mehrabi et al., 2016; Nana et al., 2014; Nopoulus, 2010). Moreover, results from functional connectivity imaging studies have suggested that alterations in the connectivity within the indirect and direct basal ganglia pathways differentially affect symptoms in HD patients (Nair et al., 2022). Genome-wide association studies indicate that the effect of HD modifying genes influence the motor and cognitive domains of the disease differentially (Lee et al. 2022). Evidence of altered cortical function has also been documented in a variety of HD mouse models (MacDonald et al., 2003; Veldman and Yang, 2018) including changes in electrophysiological properties of cortical projection neurons (Blumenstock and Dudanova, 2020; Cepeda et al., 2010). Although important differences in cortical function have been noted between these models, taken together these studies are consistent with the functional disconnection between the cortex and striatum evident in HD donors.

Despite the wealth of information on cortical dysfunction and degeneration that has been obtained in studies of HD, our knowledge of histological and molecular events that occur in early stages of the disease is incomplete. It is not known, for example, which cell types are selectively vulnerable in the early HD cerebral cortex, whether the somatic expansion of the *mHTT*- exon1 CAG tract that has been demonstrated in the cortex (Shelbourne et al., 2007; Swami et al., 2009) occurs

specifically in vulnerable neurons, or how the molecular pathology of each cortical cell type differs as the disease progresses.

To address these issues, we have developed serial fluorescence activated nuclear sorting (sFANS) of cortical samples to enable deep molecular analysis of cortical neuron classes from control and HD donors. Our focused studies of the primary motor cortex (BA4) using sFANS, supplemented with single nucleus RNA seq (snRNAseq), demonstrate that Layer 5a (L5a) pyramidal cells are selectively vulnerable in the human motor cortex in early HD. Analysis of deep layer pyramidal neuron samples from a total of six control donors and thirteen HD donors establish also that loss of Layer 5a neurons occurs widely in the cerebral cortex in HD donors. Although dramatic expansion of the *mHTT* exon1 CAG tract is evident in vulnerable layer 5a cells in the HD cortex, equally robust somatic CAG expansions are also present in Betz cells, Layer 6a and Layer 6b pyramidal neurons that remain resilient during the early phases of HD degeneration. Retrograde tracing using adeno associated viral (AAV) vectors to identify striatal projecting neurons in macaque brain (Weiss et al., 2020), coupled with *in situ* hybridization probes specific for the selectively vulnerable layer 5a human cortical neurons, strongly support the conclusion that corticostriatal projection neurons are especially vulnerable in HD. These data are consistent with identification of transcriptional programs preferentially impacted in Layer 5a neurons that indicate synaptic functions may be altered in corticostriatal projection neurons in HD. Taken together, our data identify the most vulnerable cell type in HD motor cortex as striatal projecting Layer 5a neurons, they demonstrate that extensive somatic CAG expansions occur in both vulnerable and resilient deep layer cortical neurons in HD, and they suggest that synaptic dysfunction and the cortico-striatal disconnection documented in prior human studies may play an important role in HD neurodegeneration.

RESULTS

sFANSseq for profiling the major cell classes of the human cerebral cortex

To begin to understand the molecular pathology that is associated with Huntington's disease in the human cerebral cortex, we sought to develop an approach that could provide comprehensive transcriptional data, filtered for promoter accessibility as an independent measure of active genes, as well as somatic CAG tract expansion measurements of the *mHTT* allele, from isolated populations of nuclei representing the major cell classes. Given the requirement for very efficient

use of small tissue samples, especially the limited availability of samples from HD donors, we built on the fluorescence activated nuclear sorting (FANS) strategy (Xu et al., 2018). Identification of a cortex-specific multiplexed antibody panel and the discovery that a serial sorting and serial staining strategy was optimal for high resolution characterization of all major cortical cell types from a single sample in a single session allowed us to achieve these aims. Using this serial fluorescence activated nuclear sorting (sFANS) strategy (Figure 1A), we were able to isolate nuclei from 14-16 distinct cell classes depending on the cortical region being analyzed.

To identify the cognate cell class for each sFANS sorted population, RNA-sequencing was employed to generate nuclear gene expression profiles through sFANSseq. The sFANS sorting parameters and analysis of marker gene expression in the isolated nuclei identified the recovered sFANS nuclei as: excitatory neurons from Layer 2 (L2), Layer 4 (L4), Layer 5rs (region-specific; L5rs), Layer 5a (L5a), Layer 6a (L6a), Layer 6b (L6b); inhibitory interneurons expressing VIP, LAMP5, RELN, and PVALB; as well as the non-neuronal cells astrocytes, microglia, oligodendrocytes, and oligodendrocyte progenitor cells (Figure 1B). The sFANS strategy was effective for analysis of these cell types in six regions of the human cortex, including the motor cortex (BA4), prefrontal cortex (BA9), visual cortex (BA17-19), auditory cortex (BA22), insular cortex (INS, BA13-16), and cingulate cortex (BA24). Furthermore, variations in the sFANS profiles in samples from specific cortical areas allowed identification and characterization of rare and region-specific cell types, including Betz cells (Bakken et al., 2021), Von Economo neurons (Hodge et al., 2020), and CRHBP expressing interneurons (Li et al., 2016) (Supplementary Figure S1A-C). Successful enrichment for cell type specific markers in the expression profiles (Figure 1C), as well as the gene expression heatmap of these data (Figure 1D) and principal component analyses (Supplementary Figure S1D) demonstrate the reproducibility and resolution of the sFANS approach. Although further resolution of some of these cell classes into subclasses evident in snRNAseq data (Bakken et al., 2021) can be achieved using subclass specific probes for targeted sFANS isolation, the sFANS approach was used routinely in this study because it can provide deep molecular data on all cortical cell classes in a single session for each cortical sample.

Layer 5a 5-Hydroxytryptamine Receptor 2C (*HTR2C*) expressing pyramidal cells are selectively vulnerable in HD motor cortex

To study the impact of HD on each of the cell classes identified in sFANS, we used this strategy to characterize nuclei isolated from control and HD donor samples. Analysis of the sFANS data

from control samples revealed clear separation of 3 distinct subpopulations expressing high levels of the excitatory neuron markers Special AT-rich sequence-binding protein 2 (*SATB2*) (Alcamo et al., 2008) and Transducin-like enhancer protein 4 (*TLE4*) (Hevner, 2007) (Figure 2A). Analysis of transcriptional profiles collected from each of these populations relative to snRNAseq data (Bakken et al., 2021) established that a healthy *SATB2* high population contains three subpopulations of deep layer glutamatergic cell-types: *HTR2C* expressing Layer 5a neurons, as well as two separate Cellular Communication Network Factor 2 (*CCN2*) expressing deep layer 6 populations expressing Neuropeptide FF Receptor 2 **(***NPFFR2*) and Sulfatase 1 (*SULF1*), respectively. Recovery and analysis of Betz cells from the motor cortex was evident from marker expression, revealing a distinct population of large nuclei expressing POU Domain Transcription Factor (*POU3F1*) (Bakken et al., 2021) in the *SATB2* medium range of the sFANS data (Supplementary Figure S1A and Supplementary Figure S2B).

Analysis of sFANS plots from HD donor samples revealed a clear distortion in the staining pattern of nuclei expressing high levels of *SATB2* and *TLE4* (Figure2B, Supplementary Figure S2A). To quantify this observation, the relative number of nuclei stained in the initial sFANS as NeuN positive and SATB2 negative, medium, and high fluorescence was plotted for three control vs three HD donors (Figure 2C). These data revealed a decrease of SATB2 high nuclei from HD donors with a range of Vonsattel grades (VS) (Vonsattel et al., 1985), a first indication of HD-specific cell loss in the deep layer populations. To understand which cell class was most impacted in the HD samples, we used PCA of sFANSseq datasets from the four deep layer neuronal populations. Labeling the PCA plots according to cell type revealed each of the deep layer neuronal types in discrete areas of the plot (Figure 2D, left). However, labeling by donor identity and CAP score revealed that very few samples from HD brains clustered in the layer 5a domain of the PCA plot evident from control donors (Figure 2D, right). This finding was evident in 163 cell-type specific sFANSseq samples, including 85 NC and 78 HD cell-type specific samples, derived from both, male and female donors, including 6 NC and 13 HD donors (Number of male NC = 65, female NC = 20, male HD = 38, female HD = 40 donor samples. Number of cell-type specific samples across HD donor cap scores ranges were CAP100 70to90=7, CAP100 90to110=25, CAP100 110to130=15, CAP100 130to150=18, and CAP100 150to180=13. For more details see Table S2). It is apparent from these data that recovery of layer 5a neurons from the HD cortex is reduced in the sFANSseq data.

sFANS enriched snRNAseq confirms selective vulnerability of *HTR2C* expressing Layer 5a neurons

The finding that a single cell type is missing in the sFANSseq data from HD cerebral cortex, at least early in HD progression, is interesting given the cortical thinning that has been observed in neuroanatomical studies (Reiner and Deng, 2018; Vonsattel et al., 2011; Vonsattel et al., 1985). To explore further the fate of neuronal populations in the HD cortex, we employed snRNAseq (Macosko et al., 2015). Given the relatively low abundance of the Layer 5a *HTR2C* nuclei recovered from the HD donor samples, and to allow us to focus on specific populations in subsequent studies, we used sFANS to subdivide cortical nuclei into four populations prior to snRNAseq: NeuN negative, NeuN positive SATB2 negative, NeuN positive SATB2 medium, and NeuN positive SATB2 high (Figure 3A). To align our sFANSseq data with the snRNAseq data, we used the sFANS data to identify clusters from the snRNAseq. We first generated specificity indices (Dougherty et al., 2010) for all isolated cell-types (Supplementary Figure S3A). Next, we utilized the cell-type specific gene ranking from the specificity index analysis to produce correlation matrices with the snRNAseq clusters (Supplementary Figure S3B) and confirmed assignments based on known marker genes (Supplementary Figure S3C). Cell-types assigned to clusters based on Specificity Index correlations and Uniform Manifold Approximation and Projection (UMAP) were color coded based on these analyses (Figure 3B), identifying the 15 cell classes studied in the sFANS approach in both, HD and control samples. As expected, since specific cell types were targeted in our sFANS analyses, there were additional clusters evident in our snRNAseq data. Additional clusters could be assigned based on known marker gene expression, although three clusters remained unassigned. Lastly, projection of previously published human motor cortex datasets (Bakken et al., 2021) were projected onto our sFANS derived snRNAseq datasets and heatmap correlation matrices were produced to resolve the last remaining unspecific cluster identities and to confirm our cluster assignments (Supplementary Figure S3D).

To assess HD-dependent differences in abundance of specific cell-types, we analyzed sFANS-enriched snRNAseq data from a total of 18 sorted samples from motor cortex of 2 control and 3 HD donors. These samples were enriched for NeuN negative (n=4, 2 NC, 2 HD), SATB2 negative (n=4, 2 NC, 2 HD), SATB2 medium (n=4, 2 NC, 2 HD), and SATB2 plus (n=6, 3 NC, 3 HD) populations. Comparative analysis of the snRNAseq data from control and HD donors, and quantitation of the number of nuclei recovered for each cluster, revealed several interesting

features. First, and most importantly, a clear drop in the number of *HTR2C* expressing Layer 5a nuclei was evident (Figure 3C, marked with asterisk). Second, no loss of nuclei was revealed for any other snRNAseq cluster, including the unassigned clusters not evident in the sFANS data. Third, no new clusters appeared in HD samples in which loss of Layer 5a *HTR2C* expressing nuclei was clearly evident. These data confirm the finding that *HTR2C* expressing Layer 5a neurons are the most vulnerable cell type in the HD motor cortex, and they demonstrate that this is frank cell loss because we do not recover them in other snRNAseq clusters. These data also suggest that presorting for the nuclei expressing high levels of *SATB2* may be an efficient strategy for focused snRNAseq analyses of these relatively rare deep layer neuronal populations (on average 1.5% out of all NeuN positive nuclei).

Extensive somatic expansion of the *mHTT* CAG repeat in both vulnerable and resilient cortical pyramidal cell populations

To determine the degree of CAG expansion of *mHTT* in human cortical neurons and its relationship to selective vulnerability, we isolated genomic DNA from each sFANSseq cell population and measured the CAG tract from these populations using Illumina-sequencing across *HTT* exon 1 (Ciosi et al.). Data from a single donor for all analyzed cell types, with the inherited CAG size, as defined through mapping of read length-distribution in CAG-sizing data from non-expanding cell types (for example microglia and astrocytes) (red arrow) is shown in Figure 4A. These data reveal very little CAG expansions in the four glial cell types studied (astrocytes, microglia, oligodendrocytes, and Oligodendrocyte Precursor Cells (OPCs)), moderate expansions in each of the four interneuron populations (VIP, RELN, LAMP5, and PVALB), and pronounced expansions in the layer 5a vulnerable neurons, as well as Betz cells, layer 6a and layer 6b pyramidal cell populations that are more resilient in HD motor cortex. Representative plots from the motor cortex of four additional donors for the maximally expanding deep layer populations, including Betz cells, confirmed this result (Figure 4B). To extend this analysis to additional cortical regions, CAG expansions were measured in the motor, visual, prefrontal, cingulate and insular cortices of additional HD donor samples. Radar plots of these data from each cell type in each donor display the percentage of nuclei where the *mHTT* CAG tract has expanded beyond the modal inherited CAG length by 3 to 5 repeats (grey), 6 to 20 repeats (black), or by more than 20 repeat units (red). These data also document that somatic expansions occur in nearly all deep layer neurons and that, in approximately half of the deep layer neuron samples analyzed, large expansions were very frequent, so that 50% or more of *mHTT* copies reached CAG tract lengths that carry more than 20 additional CAG units (Figure 4C, see also Supplementary Figure S4).

These data demonstrate expansions of the *mHTT* CAG repeat in many cell types across the human cerebral cortex, with maximal expansion occurring in both vulnerable and resilient deep layer neurons. It is noteworthy also that the histograms of *mHTT* CAG tract size-distribution of expanded *mHTT* alleles in deep layer neurons appears less uniform in cortical neurons than MSNs. These data also generalize the conclusion that somatic CAG expansions are not restricted to vulnerable cell populations, a conclusion reached in our accompanying study of human striatal cell types in HD (Mätlik et al, see accompanying manuscript), and suggest that further analysis of the expanding vulnerable and resilient cell populations will be informative. For example, analysis of the transcription of HD modifying genes in cortical cell types does not reveal elevated expression of *MSH2* and *MSH3* (Figure 4D), as evidenced in studies of striatal MSNs.

The vulnerable layer 5a *HTR2C* expressing cells in primates are corticostriatal projection neurons

Excitatory neurons in the cerebral cortex are primarily pyramidal cells (Cajal, 1999) that send projections to a variety of brain structures and targets (Cajal, 1999; Elam et al., 2021; Usrey and Sherman, 2019). Layer 5 neurons project primarily to subcortical targets, and detailed studies in the mouse cortex have demonstrated that they have distinct physiological properties (Groh et al., 2010; Hattox and Nelson, 2007). Layer 6 neurons generally either project subcortically to target nuclei in the thalamus or have axons that ascend within the cortex (Baker et al., 2018). Since a variety of human imaging studies of HD have documented alterations in the connectivity of multiple areas of the cortex with the striatum, we investigated next the projection targets of the *HTR2C* expressing layer 5a neurons in the primate cortex. To do so, we took advantage of the very detailed mapping of striatal projecting neurons in the adult rhesus macaque previously reported by Weiss et al (2020) (Weiss et al., 2020). These studies employed a retrograde AAV2 tracer virus expressing eGFP (AAV2.retro) to transduce neurons in the caudate nucleus and putamen. Immunohistochemical detection of the eGFP gene delivered by the viral vector revealed many cortical and subcortical neurons that project to the striatum in the non-human primate (NHP) brain. The highest densities of retrogradely transported virus were detected in anterior regions of the cerebral cortex, including the prefrontal and motor areas (Weiss et al., 2020). To map *HTR2C* expressing cells in the AAV2.retro labeled macaque brains, sections from these animals were examined and sub-dissected to use for analysis of the motor and cingulate cortex. Immunofluorescence data collected from these sections revealed robust staining of eGFP in many layer 5 pyramidal neurons (Figure 5A). To simultaneously analyze expression of both eGFP and *HTR2C*, we used RNA scope *in situ* hybridization (ISH) (Wang et al., 2012). Dense signal for

eGFP (green) in layer 5 neurons using this assay was evident even in low power images (Figure 5B), as well as less robust signal from the *HTR2C* probe (red). At higher magnification, specific eGFP expression from the AAV2.retro virus as well as endogenous macaque *HTR2C* expression could be evaluated as the increased grain counts within and surrounding DAPI positive nuclei relative to background for each probe (Figure 5C). Sections stained with both RNAscope probes and DAPI were scanned at 20X magnification, tiled, and analyzed using HALO (Erben and Buonanno, 2019). To determine whether corticostriatal cells expressing eGFP from AAV2.retro also express *HTR2C*, parameters were set to reflect three classes of colocalization signal: cells strongly positive for eGFP, cells strongly positive for both eGFP and *HTR2C*, and cells strongly positive for eGFP and weakly positive for *HTR2C* (Figure 5B). Representative data for single sections from cingulate and motor cortex reveal that 41% (motor cortex) and 32% (cingulate cortex) of the eGFP expressing neurons are positive for high levels of *HTR2C*, with an additional 67% (motor cortex) and 68% (cingulate cortex) weakly positive in cingulate and motor cortex respectively (Figure 5D). These data demonstrate that macaque corticostriatal projection neurons express *HTR2C*, the marker gene specific for the vulnerable neuron type in HD.

It is noteworthy that the data we have collected in this study are consistent with both previous ISH studies of human *HTR2C* expression (Supplementary Figure S5B, brain-map.org/ish/HTR2C) and regional variation in retrogradely labeled cells in the macaque cerebral cortex (Weiss et al., 2020). Thus, the decline in the number of retrogradely labeled neurons noted as one moves from anterior to posterior regions in the macaque cortex is also reflected in the human *HTR2C* ISH data from the frontal, temporal and visual cortex in the Allen Brain Atlas (brain-map.org), and in the differences in the recovery of *HTR2C* expressing layer 5 neurons in the sFANS studies of samples from the prefrontal, motor and visual cortex (Supplementary Figure S5C). For example, although the precise location of the samples used for the sFANS from these areas is difficult to assess, very small numbers of *HTR2C* expressing nuclei were recovered from samples identified as primary visual cortex.

Transcriptional responses in the HD cerebral cortex

To gain insight into the impact of *mHTT* with long CAG tracts on cellular functions in each of these cell types, we compared their transcriptional responses to *mHTT*. Cell-type specific differential gene expression (DESeq2) analyses revealed over 1,000 differentially expressed genes in layer 5a, layer 6a, and layer 6b neurons from HD versus control donor samples. Analyses of layer 5rs neurons revealed less than 300 differentially expressed genes (Figure 6A). Principal component

analyses (PCA) of the three deep layer neuron populations, including layer 5a, layer 6a, and layer 6b neurons, revealed separation of samples based on sex (PC1) and disease status (PC2) (Supplementary Figure S6A). Samples from layer 5rs neurons however, showed no clear separation based on disease status. Analyzing Betz cell samples from motor cortex revealed a weak trend for separation in PC3 (Supplementary Figure S6B). Analyses of layer 5rs samples confirmed no clear pseudotime trajectory or trend according to disease status (Supplementary Figure S6B). Pseudotime analyses, using our sFANSeq data from the deep layer cell-types further confirmed a clear progression trajectory of molecular alterations with increasing CAP scores (Figure 6C) for layer 5a, layer 6a, and layer 6b neurons.

To reveal gene expression programs altered in the deep layer neuron cell classes that showed molecular alterations based on disease status, we performed gene ontology (GO) analyses. To this end, we first filtered the cell-type specific differential gene expression lists from our DESeq2 analyses based on the cell-type specific accessible promotor peaks revealed through ATACseq (Figure 6B). This step allowed us to focus our gene lists for the GO analyses to those genes that were truly accessible in the specific cell-types. As shown in Figure 6D, expression of genes involved in dendritic and synaptic functions are decreased in layer 5a, layer 6a, and layer 6b neurons that undergo extensive somatic expansion of the *mHTT* CAG tract in HD. These data are generally consistent with studies in mouse models of HD (Plotkin and Surmeir, 2015; Veldman and Yang 2018), and with human imaging studies indicating disconnection between the cortex and striatum (Espinoza et al., 2019; Unschuld et al., 2012), and they support the hypothesis that *mHTT* carrying long polyglutamine tracts can disrupt many aspects of dendritic and synaptic function. Analyses of differential expression of genes of interest between HD and controls revealed multiple DNA repair genes as differentially expressed between HD and controls (Figure 6E).

To provide further insight into the increased vulnerability of corticostriatal pyramidal cells in HD, we compared the transcriptional responses of layer 5a *HTR2C* neurons compared to layer 6a and layer 6b pyramidal cells that have equivalent levels of CAG expansions in HD but remain relatively resilient against cell death (Supplementary Figure S6C). In general, these data revealed similar pathways as those discussed above: depressed expression of pathways involved in neuronal membrane functions required for support of dendritic and synaptic functions. Pathways that show enrichment of upregulated gene sets in layer 5a neurons include those involved in membrane flux and delivery of cell surface proteins to their sites in the membrane. These responses are distinct

from those occurring in vulnerable striatal MSNs in HD (Mätlik et al, see accompanying manuscript). For example, the strong induction of genes involved in autophagy and lysosomal function evident in MSNs does not occur in L5a *HTR2C* expressing pyramidal cells, and the categories of genes involved in nuclear functions whose expression is suppressed in MSNs does not occur in the vulnerable cortical neurons. These data establish that the molecular responses of human cortical and striatal neurons are distinct in HD. Moreover, they indicate that altered synaptic function of L5a and L6 deep layer neurons results in their dysfunction, and they suggest that altered corticostriatal connectivity may play an important role in the selective loss of L5a corticostriatal pyramidal cells in HD.

DISCUSSION

Although degeneration of the striatum and the associated disturbances of motor control are the most visible features of Huntington's disease, cognitive decline and psychiatric symptoms occur early in HD and they are accompanied by functional and anatomical alterations in the cerebral cortex (Hickman et al., 2022; Mehrabi et al., 2016; Nana et al., 2014; Nopoulus, 2010). Cortical dysfunction in HD is thought to be the result of selective vulnerability of specific cell-types, although the molecular identity of vulnerable and resilient cell-types in the cerebral cortex and the mechanisms governing their fate in HD have not been determined. Here we have used deep molecular profiling to explore pathogenesis in six human cortical regions with a focus on the motor cortex in Huntington's disease. The transcriptional data we have collected in the cerebral cortex of control and HD donors are not sufficient to reveal directly mechanisms of HD pathogenesis. However, they are sensitive measures of molecular events that occur during HD progression, and they allow initial inferences regarding the impact of *mHTT* on specific cell types. Our studies reveal that expansions of the *mHTT*-CAG tract in the cortex, as in the striatum, is not sufficient for cell loss in HD (Mätlik et al, see accompanying manuscript), and they are consistent with the two-step model of HD pathogenesis (Hong et al., 2021). We identify layer 5 *HTR2C* expressing corticostriatal pyramidal cells as selectively vulnerable in HD and suggest that the corticostriatal connectivity deficits observed in HD result from the dysfunction and/or death of these neurons.

Serial fluorescence activated nuclear sorting (sFANS)

Studies of human tissue are challenged by the paucity of appropriate samples for analysis of early phases of disease. Given the many cell types present in the human cerebral cortex (Cajal, 1999),

the advantages of population based molecular studies evident in our studies of the human cerebellum (Xu et al., 2018) and striatum (Mätlik et al, see accompanying manuscript), and the desire to use a tissue sparing strategy for studies of HD donor samples, we developed sFANS for deep molecular analysis of all major cortical cell classes from one single tissue sample. The approach first involved sorting of nuclei into four pools based on the expression of *NEUN* and *SATB2*, followed by a second sort for each of these pools into subpopulations of nuclei based on the presence of *BCL11b*, *TLE4*, *EAAT1*, *IRF8* or *OLIG2*. Although most of these markers do not identify a single cell type, the multiplexed nature of the sort and quantitative differences in expression resulted in purification of nuclei for each of the major excitatory, inhibitory, and glial cell classes in the human cortex, as well as region specific cell types such as Betz and von Economo neurons. Importantly, the yield from the sFANS strategy was sufficient to collect *mHTT* CAG DNA expansion, epigenetic, and transcriptional data for each class of nuclei. Furthermore, although not used for these studies, probes against individual cell types can allow targeting of a single, previously identified cell type for analysis if required (Supplementary Figure S2C). Finally, as demonstrated here, distortions in the staining patterns observed on sFANS plots from HD donors can provide a preliminary readout of the relative abundance of each population in comparisons of cortical regions of donor samples. It is instructive, for example, that the differences visible on sFANS plots for deep layer neuron nuclei were the first indications of cell type specific HD pathology to emerge in this study.

Loss of layer 5a *HTR2C* expressing corticostriatal pyramidal cells in Huntington's disease

The data we have presented here strongly support the conclusion that the layer 5a *HTR2C* cells that are lost in the human cerebral cortex early in HD are pyramidal cells that project to the striatum. Although we have analyzed a relatively small number of HD donor cases (N=13), loss of *HTR2C* expressing cells is documented by both sFANSseq and snRNAseq across multiple brain regions and in many samples. Of course, we cannot exclude the loss of other cortical cell types in specific cases given the heterogeneity of cognitive symptoms in HD (Hickman et al., 2022; Mehrabi et al., 2016; Nana et al., 2014; Nopoulus, 2010), especially late in disease progression (Hedreen et al., 1991). While it remains possible that the *HTR2C* neurons in the human cortex are not homologous to those that were labeled by *HTR2C* and AAV2.retro in the macaque cortex, we view this as unlikely because of the abundance gradient of these cells along the antero-posterior axis that is common to the two primate species (Bernard et al., 2012; Weiss et al., 2020).

The selective vulnerability of corticostriatal neurons in HD that we have documented here relative to the more resilient layer 6 neurons and Betz cells is interesting given the functional imaging studies of *mHTT* carriers (Unschuld et al., 2012) (Espinoza et al., 2019). Since these cells are direct presynaptic partners of striatal MSNs, an understanding of the relative timing of their dysfunction and loss of these two cell-types in HD is essential. Longitudinal studies of the functional and connectivity changes in premanifest and symptomatic HD subjects (Georgiou-Karistianis et al., 2013; Wolf et al., 2011), as well as PET (Kuwert et al., 1990) imaging studies, have suggested that the cortical activation in pre-manifest HD may reflect a compensatory response to decreased functional connectivity between the cerebral cortex and striatum as an early sign of progression, and that the diminished activation seen in symptomatic stages may reflect the inability of the brain to continue to compensate as neuronal damage accumulates. This places an important emphasis on delineating changes in the properties of these cells in HD (see below), and in the timing of their dysfunction and loss across cortical regions. Furthermore, the robust expression of *HTR2C* in both corticostriatal pyramidal cells and striatal projection neurons suggests that further development of the *HTR2C* imaging probes (Kim et al., 2017) for use in longitudinal studies could be particularly important for understanding the relationships between corticostriatal disconnection and neurodegeneration in HD.

Widespread somatic expansion of *mHTT* CAG repeats in the HD cerebral cortex

In the striatum, somatic expansion of the *mHTT* CAG tract is evident exclusively in striatal projection neurons and cholinergic interneurons expressing choline acetyltransferase (CHAT+ INs) (Mätlik et al, see accompanying manuscript). Significant CAG expansions are rare in other striatal interneuron populations and glia. While the very limited expansion in glia observed in striatum is also evident in cerebral cortex, all classes of cortical neurons we have analyzed display somatic instability of the *mHTT* CAG tract. A significant fraction of cortical interneurons, for example, have tracts expanded by more than 5 additional repeat units. This is in contrast to the lack of somatic CAG expansions in many striatal interneurons, including those expressing *SST*, *PVALB*, and *TAC3* (Mätlik et al, see accompanying manuscript). Interestingly, striatal and cortical interneurons have a common developmental origin within the medial ganglionic eminence as progenitors of both (Knowles et al., 2021) span the cerebral cortex and striatum (Reid and Walsh, 2002). One might argue, for example, that this provides evidence that the differential CAG expansion levels in related interneuron cell types as evident in our data, may arise late in development as cortical and striatal interneurons diverge, possibly as a consequence of local cell-cell interactions in the cortex. The most dramatic CAG expansions in the HD cortex are evident

in the four (L5rs, L5a, L6a, and L6b) deep layer pyramidal cell populations we have analyzed (Figure 4B). Cortical cell types that are vulnerable (layer 5a corticostriatal neurons) and resilient to degeneration (Betz neurons and layer 6 populations) incur somatic CAG expansions. This is consistent with our finding that somatic CAG expansions are not sufficient for cell loss in the human striatum (Mätlik et al, see accompanying manuscript).

Cell-type specific patterns of somatic CAG instability

Somatic CAG instability is influenced by specific features of the locus that is expanding, its transcriptional and epigenetic state, and trans-acting factors that directly act at the repeat locus to enhance or suppress somatic instability (Nakamori et al., 2011) (Barbe and Finkbeiner, 2022). For example, in striatal medium spiny neurons (MSNs), which show the greatest CAG expansions in that structure, the expression of *MSH2* and *MSH3* is higher both at the nuclear transcript and protein level relative to other striatal cell types (Mätlik et al, see accompanying manuscript). One explanation of the expansions in MSNs and stability in striatal interneurons is that MutSβ (MSH2-MSH3), a driver of expansions, can inhibit excision of excess repeats by FAN1, a suppressor of expansions, while at lower MutSβ levels FAN1 can effectively excise the excess repeats (Mätlik et al, see accompanying manuscript). In the deep layer cortical neurons that display the most dramatic expansion (L5a, L6a, L6b, and Betz cells) elevated expression of the genes encoding *MSH2* and *MSH3* is not evident relative to neurons and glia that are relatively stable (Figure 4D). While this result may indicate that the rate limiting factors driving somatic expansion may differ between cell types, we note that the expression of *FAN1* in these neurons is low (Figure 4D). While this supports the general idea that the stoichiometric levels of DNA repair proteins can, in part explain brain region- and cell type-specific levels of somatic expansions ((Goula et al., 2012; Lopez Castel et al., 2009; Mason et al., 2014; Slean et al., 2016; Tome et al., 2013; Tome et al., 2011), Mätlik et al, see accompanying manuscript), the interactions of repair proteins and the variables affecting instability are complex. For example, it is unclear how the increased *POLD1*, a modifier of HD (Lee et al., 2022), in L5 neurons and the reduced *EXO1* in L6a and L6b neurons, being two of the most differentially expressed between HD and controls (Figure 6E) can affect CAG expansions. It is noteworthy also that the histograms of *mHTT* CAG tract size-distribution of expanded *mHTT* alleles in deep layer neurons appears less uniform in cortical neurons than MSNs (Mätlik et al, see accompanying manuscript). Thus, in all samples measured to date, the profile of CAG tract lengths expansion is more distributed and spread over a wider range of CAG repeat lengths in the cortex than in striatal MSNs. Although systematic analysis of additional donors will be required to validate this finding, and independent assays of somatic expansion will

be required to rule out technical artefacts, it seems possible that this observation reflects differences in the stoichiometry or activity of DNA repair complexes operating at the *mHTT* locus in cortical and striatal neurons. While the data do not lead us to conclude that the rate limiting steps for somatic expansion of the *mHTT* CAG tract differ between cell types, they are consistent with the recent report that variants in known HD modifying candidate genes act differentially on motor and cognitive domains of HD (Estrada-Sanchez and Rebec, 2013; Layburn et al., 2022; Lee et al., 2022; Waldvogel et al., 2015). Further in depth studies of the *mHTT* locus in deep layer cortical neurons, transcription of both *mHTT* and its antisense transcripts, and protein complexes present at the locus in these neurons will be required to understand cell type specific expansion. We believe that the ability to isolate and characterize large numbers of nuclei from HD donor samples using the sFANSseq approach is important for advancing our knowledge of this process in the human brain.

mHTT toxicity

Given the finding that extensive somatic *mHTT* CAG expansions are evident in both vulnerable and resilient deep layer pyramidal cells in the HD cerebral cortex, we believe that comparative studies of the transcriptional events occurring in these cell types may be particularly informative. The high probability that significant levels of mutant huntingtin accumulates in each of these cell types, which is supported by recent findings of regionally variable but pronounced aggregation in cortical layers 5 and 6 (Hickman et al., 2022), has important functional implications. It seems likely, for example, that these alterations cause significant dysfunction in all the expanding deep layer populations, disrupting any of the cells to which they connect. For example, many of the functional GO categories that are enriched for genes with disease-associated downregulation are involved in HD and reflect dendritic and synaptic functions (Figure 6D). Furthermore, enhanced expression of genes involved in plasma membrane function and maintenance, and assembly of important membrane complexes, suggest homeostatic responses may be engaged in the affected neurons in response to the deficits in mechanisms of synaptic support. These human data support previous studies of synaptic and connectivity defects in HD mouse models that have revealed widespread defects in cortico-striatal communication (MacDonald et al., 2003; Veldman and Yang, 2018), they are in line with the prediction that white matter changes are associated with altered transcription of synaptic genes in HD (McColgan et al., 2018), and with the suggestion that more extensive deficits in the corticostriatal thalamocortical (CSTC) circuit are likely to occur in HD (Barry et al., 2022). Although our data from the human cerebral cortex and those previously reported in mouse models (MacDonald et al., 2003; Veldman and Yang, 2018) indicate that similar

processes are impacted, the gene by gene details of neuronal responses to mutant huntingtin in human and mouse deep layer neurons differ. This is not surprising given the well documented species-specific differences that occur in many neuronal cell types (Hodge et al., 2019; Xu et al., 2018) and the ubiquitous expression of *mHTT* with long polyglutamine tract in HD mouse models, in contrast to our findings of cell-type specific somatic CAG tract expansion in human samples.

Concluding remarks

The use of sFANS to characterize cell type-specific gene expression profiles and HD-associated expression changes in the cerebral cortex illustrates the applicability of this method to the studies of this brain region in the context of other neurological diseases, including ones where tissue availability can be a limiting factor. The analyses of the molecular events that unfold in the human cerebral cortex that we have presented here are the beginning of an effort to understand in detail HD pathophysiology in the cortex. In summary, our data support the two-step model of *mHTT* toxicity (Hong et al., 2021) and provide evidence that expansion of the *mHTT* CAG tract is not the only mechanism that determines the rate of cell loss in HD cerebral cortex. We believe the depth of the data we have been able to collect by sFANS, coupled with the ability to filter out ambient RNA contamination using promoter accessibility data, is important. It will enable many additional comparative studies to gain insight into specific aspects of HD progression and into general questions surrounding selective cellular vulnerability in human disease. Our data also raise additional features for future investigation. How long are cells dysfunctional before they are lost in HD? At what stage during brain development/ageing does the predisposition to or protection from somatic CAG expansions occur? What genetic and/or epigenetic factors fix these predispositions? Are these attributes the same for the same cell types in other diseases with repeat expansions in different genes? Does the onset of corticostriatal disconnection reflect the timing of somatic CAG expansion in *mHTT* or frank loss of neurons? What is the relative timing of these processes in the cortex vs the striatum? Are the molecular mechanisms of cell loss similar in the cortex and striatum? Although additional data will be required to gain insight into these issues, sFANS allows isolation of sufficiently large numbers of cell type specific nuclei from human cortex to begin to study mechanisms suggested by these data.

Acknowledgements

This study was supported by funding from the CHDI Foundation, and we are thankful to Dr. Thomas F. Vogt and Dr. Jian Chen for helpful discussions throughout the project. We thank Nicholas Didkovsky, Cuidong Wang and Dr. Kärt Mätlik for advice on coding and data visualization. We thank Dr. Christopher E. Pearson for input and constructive criticism of the manuscript. We are also grateful to the Rockefeller University Genomics Resource Center for advice and support. C.P. Christina Pressl, MD MS is in part supported by the Stavros Niarchos Foundation as the Stavros Niarchos Foundation Scholar. The research was in part supported by grant #UL1TR001866 from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) program. We thank the NIH Neuro Bio Banks, in especially the Miami Brain Endowment Bank as well as the Harvard Brain Tissue Resource Center for providing biological samples. We thank all donors and families without whom this work would not be possible.

Author Contributions

C.P and N.H. conceptualized and designed the study. C.P. and P.D. performed the experiments. K.M. set up the CAG sizing assay and analysis strategy, optimized TrueBlack quenching, led the efforts on ATACseq-based filtering, and gave crucial guidance and support on a set of visualizations shown in this manuscript. D.D., provided human brain samples. J.MB., W.L., and A.W. provided NHP samples. C.P. analyzed the data in collaboration with J.D.L, M.R.P. and T.C.. C.P. and N.H. wrote the manuscript. All authors discussed the results and commented on the manuscript. N.H. acquired research funding.

Declaration of Interests

The authors declare no competing interests.

FIGURE TITLES AND LEGENDS

Figure 1. sFANSseq for profiling the major cell classes of the human cerebral cortex
(**A**) Schematic representation of human cortical sample preparation, nuclei extraction, and serial fluorescence activated nuclei sorting (sFANS), using multiplexed fluorescent antibody panels.
(**B**) Representative sFANS plots from one motor cortex sample, showing distribution of the major human cortical cell-type populations.
(**C**) Examples of resulting sFANSseq IGV gene expression profiles from multiple regions of cortex and donors. Populations shown include layer 2 (L2), layer 4 (L4), layer 5 region-specific neurons (L5rs), layer 5a (L5a), layer 6a (L6a), layer 6b (L6b), VIP expressing interneurons, LAMP5 expressing interneurons (LAMP5), RELN expressing interneurons (RELN), PVALB expressing interneurons (PVALB), astrocytes, microglia, oligodendrocytes (Oligos), and oligodendrocyte progenitor cells (OPCs).
(**D**) Heatmap shows log-2 transformed average DESeq-normalized expression levels for cell-type specific marker genes across all samples.

Figure 2. Layer 5a *HTR2C* expressing pyramidal cells are selectively vulnerable in HD motor cortex
(**A**) Representative sFANS panels from one control donor showing distribution of NeuNnegative, NeuNpositive/SATB2negative, NeuNpositive/SATB2medium, and NeuNpositive/SATB2high populations (left panel). Representative distribution of NeuNpositive nuclei plotted for fluorophore signal intensity of TLE4 and BCL11b, revealing three TLE4 positive deep layer excitatory neuron populations layer 5a (L5a), layer 6a (L6a), layer 6b (L6b) (right panel).
(**B**) Representative sFANS panels from one HD donor showing distribution of NeuNnegative, NeuNpositive/SATB2negative, NeuNpositive/SATB2medium, and NeuNpositive/SATB2high populations (left panel). Representative distribution of NeuNpositive nuclei plotted for fluorophore signal intensity of TLE4 and BCL11b, revealing a distortion of the staining pattern of TLE4 positive populations (right panel).
(**C**) Quantitation of the relative abundance of NeuNpositive/SATB2negative, NeuNpositive/SATB2medium, and NeuNpositive/SATB2high nuclei among all NeuNpositive nuclei isolated from three Normal Control (NC) donors, one genetically confirmed HD donor pre-diagnosis (At Risk), four HD donors with Vonsattel striatal grade 1 (VS-1), three HD donors with Vonsattel striatal grade 2 (VS-2), one HD donor with Vonsattel striatal grade 4 (VS-4).
(**D**) Principal component analyses (PCA) utilizing a total of 163 datasets from the four deep layer excitatory neuron populations layer 5 region-specific neurons (L5rs), layer 5a (L5a), layer 6a (L6a), layer 6b (L6b), color coded by cell-type (left panel) and color coded by group (right panel), with controls in green and HD samples in shades of red, grouped by the normalized CAG-age product (CAP100) (CAP100 70to90, CAP100 90to110, CAP100 110to130, CAP100 130to150, CAP100 150to180). CAP100 scores were calculated as described by Warner et al. (Warner et al., 2022):

$$CAP100score = Age\ at\ end\ of\ life * \frac{mHTT\ CAG\ length - 30}{6.49}$$

Figure 3. sFANS enriched snRNAseq confirms selective vulnerability of *HTR2C* expressing Layer 5a neurons

(**A**) Schematic representation of sFANS pre-sorting strategies and resulting Uniform Manifold Approximation and Projection (UMAP) plots of the single nuclei RNA sequencing (snRNAseq) data, color coded by sort population (NeuNnegative in gray, NeuNpositive/SATB2negative in light blue, NeuNpositive/SATB2medium in green, and NeuNpositive/SATB2high in dark red).

(**B**) Separated UMAPs for NC and HD samples, showing data from full donor datasets for which NeuNnegative, NeuNpositive/SATB2negative, NeuNpositive/SATB2medium, and NeuN positive/SATB2high nuclei could be harvested (number of donors NC=2, HD=2). Cluster assignments in bold and color coding based on sFANS-guided cell-type designations and marker gene expression. Cluster designations in normal font correspond to the correlated clusters in the publicly available human motor cortex datasets (Bakken et al., 2021).

(**C**) Quantitation of fraction nuclei in NC vs HD cell-type specific clusters as designated in (Bakken et al., 2021), reveal statistically significant loss of L5a (L5/6 NP) nuclei from the NeuNpositive/SATB2high populations in HD (adjusted P value < 0.05 in Bonferroni multiple comparison corrected T-test, marked with the asterisk).

Figure 4. Somatic expansion of the *mHTT* CAG repeat in cortical cell types

(**A**) Representative plots showing CAG length distributions in all cortical cell-types analyzed from one HD motor cortex donor. Red arrow/line indicates the modal allele length defined for each donor through genotyping and assessment of non-expanding glial data from the same donor.

(**B**) Representative plots of CAG length distributions in deep layer excitatory neuron populations (L5rs (Betz cells in motor cortex), layer 5a (L5a), layer 6a (L6a), layer 6b (L6b)) from five HD donors.

(**C**) Radar plot showing data from all cell-types of eleven HD donors, and all cortical regions analyzed. Data are color coded by length of CAG expansions greater than the modal length in striatal glia, with 3 to 5 repeats in grey, 6 to 20 repeats in black, and over 20 repeats in red.

(**D**) Heatmap shows log-2 transformed average DESeq-normalized expression levels of selected HD modifying genes across all cell-types at baseline, utilizing data from six control donors.

Figure 5. Vulnerable layer 5a *HTR2C* expressing cells in primates are corticostriatal projection neurons

(**A**) Immunohistochemistry (IHC) experiment results from macaque cortex confirm eGFP positive staining of layer 5 pyramidal neurons.

(**B**) Representative in situ hybridization (ISH) data analyses of NHP cingulate cortex, using HALO software for distribution analyses. Delineation of DAPI positive nuclei and eGFP as well as *HTR2C* RNA signal puncta quantitation, using HALO analyses tool is shown in higher magnification in top panels. Analyses results, revealing eGFP and *HTR2C* positive signal in deep layer populations at low magnification (middle and bottom panels).

(**C**) High magnification images from the cingulate cortex show positive signal for eGFP (green) and *HTR2C* (red).
(**D**) HALO quantification results from single sections from cingulate and motor cortex reveal 41% (motor cortex) and 32% (cingulate cortex) of double positive for eGFP and high *HTR2C* neurons and 67% (motor cortex) and 68% (cingulate cortex) neurons being positive for eGFP and weakly positive for *HTR2C*.

Figure 6. Transcriptional responses in the HD cerebral cortex
(**A**) Number of differentially expressed genes (adjusted P value < 0.05 by DESeq2) in the comparison of datasets from HD donors versus controls and after gene list filtering, based on the detection of accessible promoters using ATACseq.
(**B**) Representative distribution of sFANSeq and ATACseq reads from motor cortex mapped to deep layer cell-type specific marker genes.
(**C**) Bulk pseudotime analyses results for layer 5a, layer 6a, and layer 6b, depicting molecular responses along the pseudotime axis in control and in HD samples across CAP scores.
(**D**) Results from Gene Ontology Cellular Component pathway enrichment analyses of genes downregulated in HD versus controls (p adj < 0.05 by DESeq2 and filtered for genes with accessible promotors detected by ATACseq) in deep layer excitatory neuron populations, including layer 5a, layer 6a, and layer 6b nuclei.
(**E**) Differential gene expression between HD and control samples from layer 5a, layer 6a, and layer 6b are shown for a set of genes of interest.

Figure 1

sFANSseq for profiling the major cell classes of the human cerebral cortex

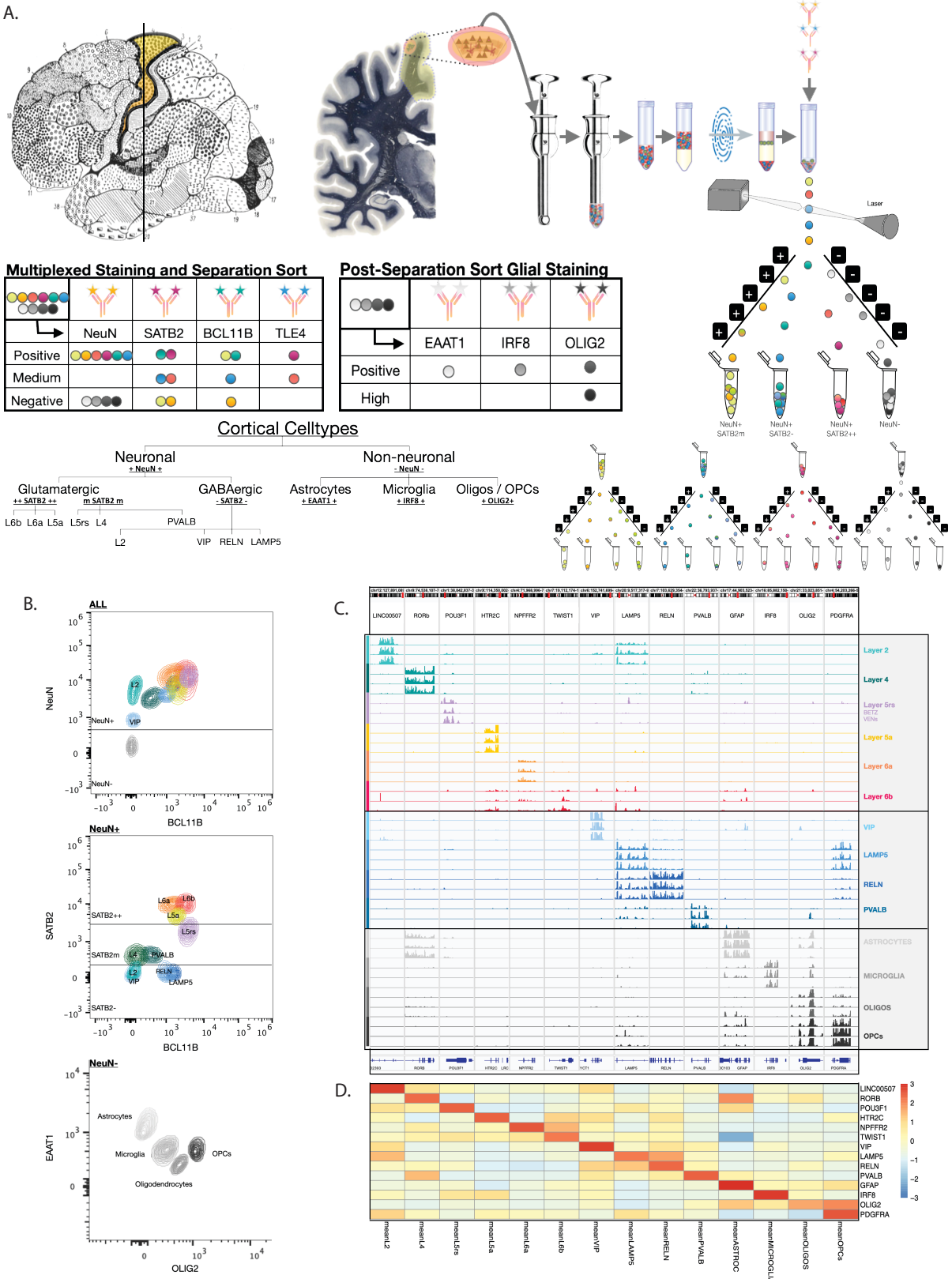
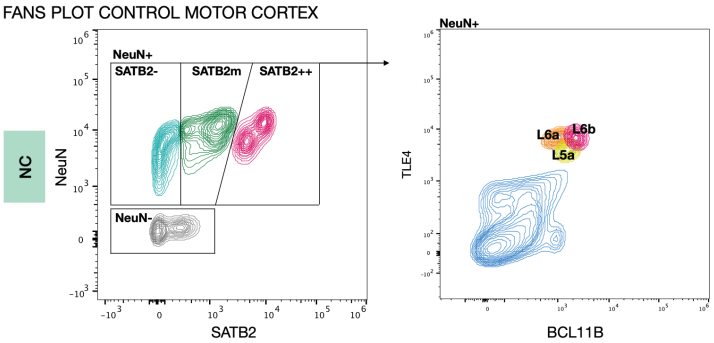
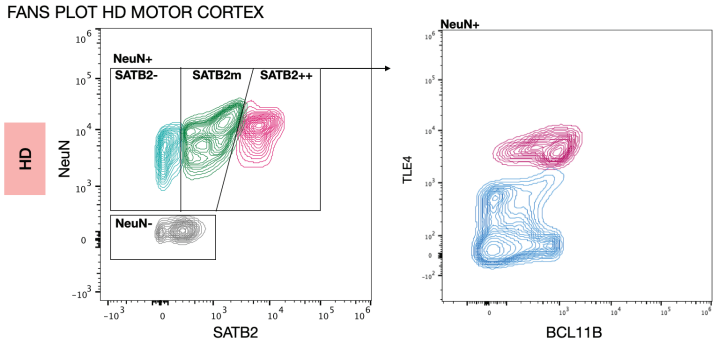


Figure 2
Layer 5a HTR2C expressing pyramidal cells are selectively vulnerable in HD motor cortex

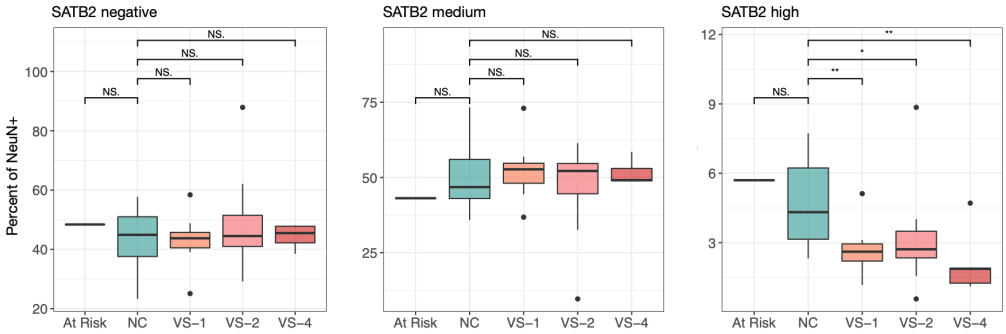
A. FANS PLOT CONTROL MOTOR CORTEX



B. FANS PLOT HD MOTOR CORTEX



C. FANS QUANTIFICATIONS ACROSS MULTIPLE CORTICAL REGIONS



D. DEEP LAYER EXCITATORY NEURONS ACROSS MULTIPLE CORTICAL REGIONS SAMPLED PCAs TOP 300 GENES

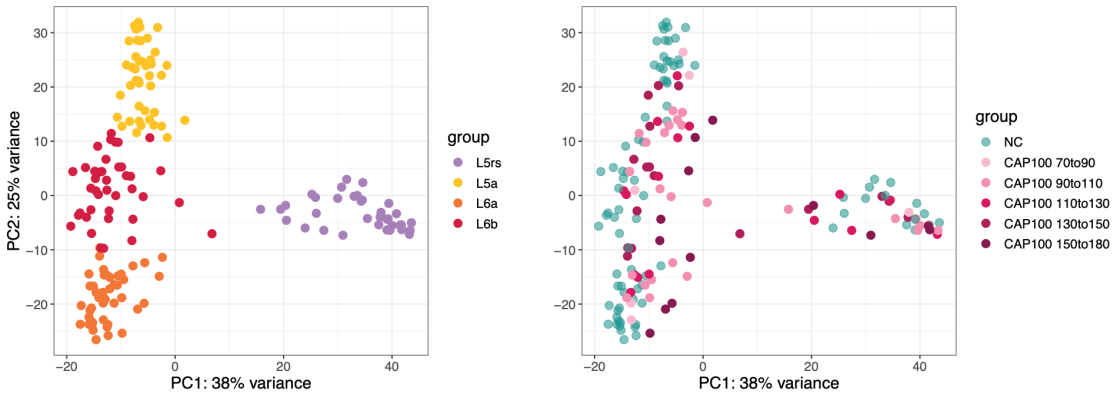


Figure 3

FANS enriched snRNAseq confirms selective vulnerability of HTR2C expressing Layer 5a neurons

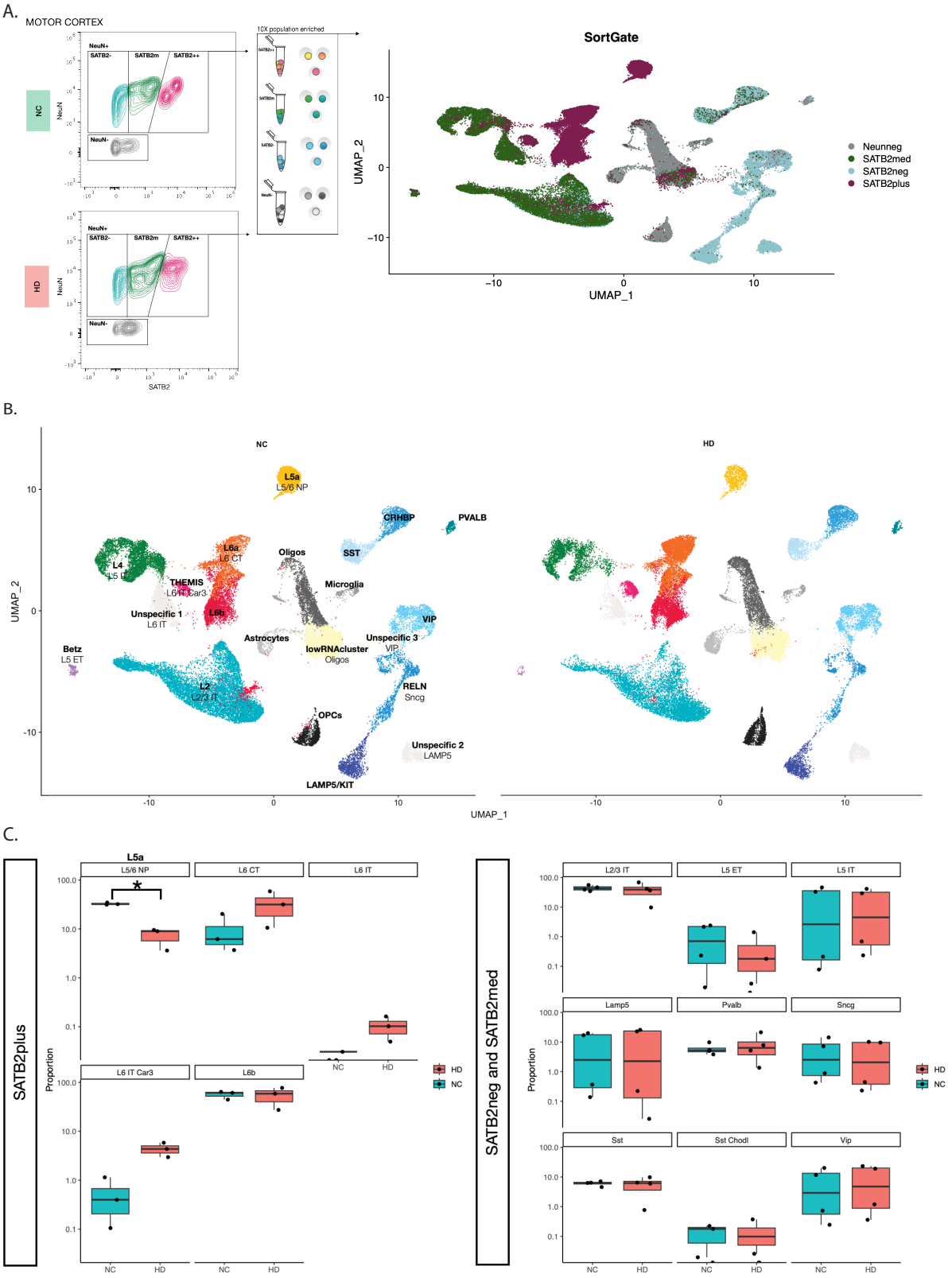


Figure 4
Somatic Expansion of the mHTT CAG Repeat in Cortical Cell Types

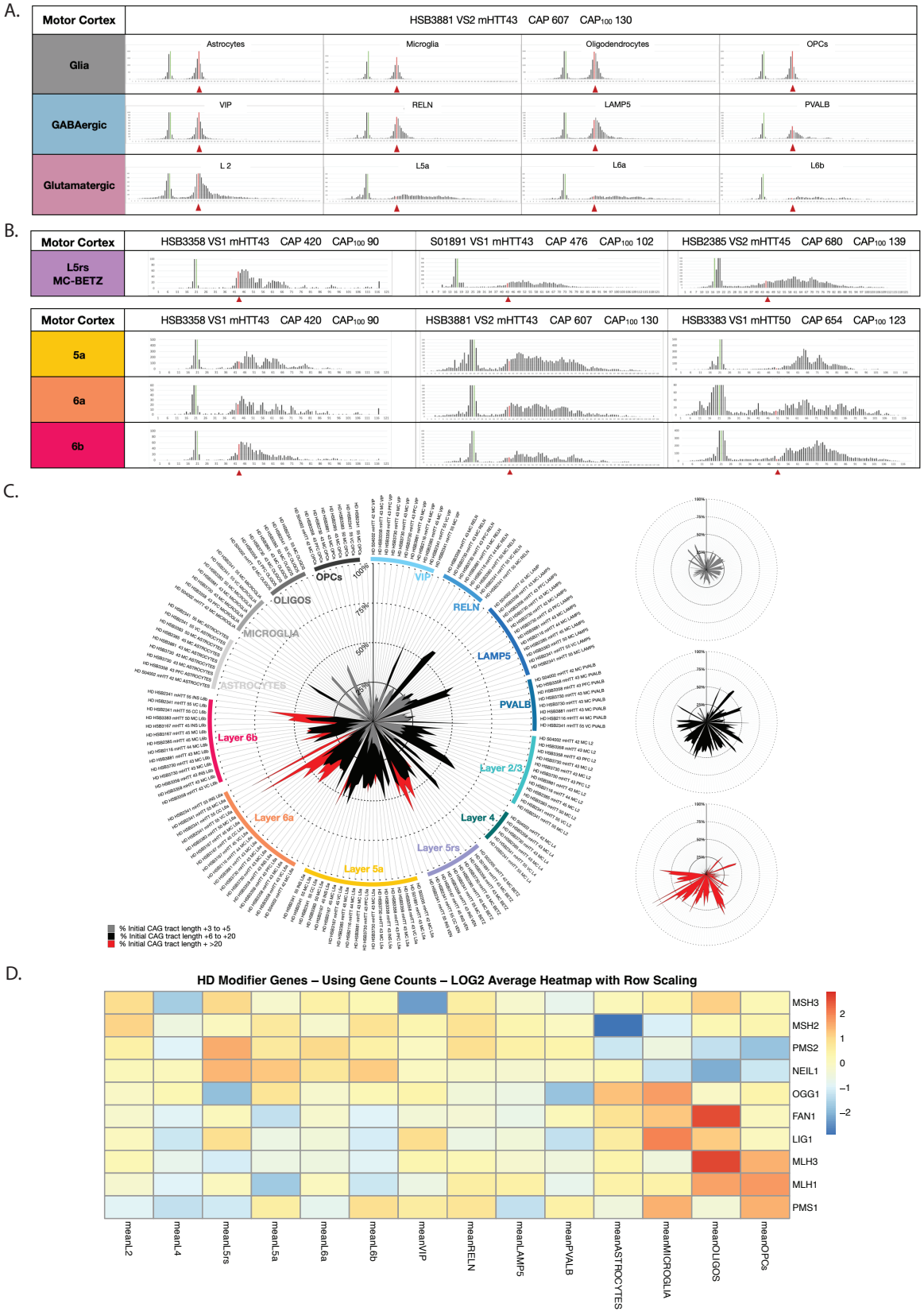


Figure 5

Vulnerable layer 5a HTR2C expressing cells in primates are corticostriatal projection neurons

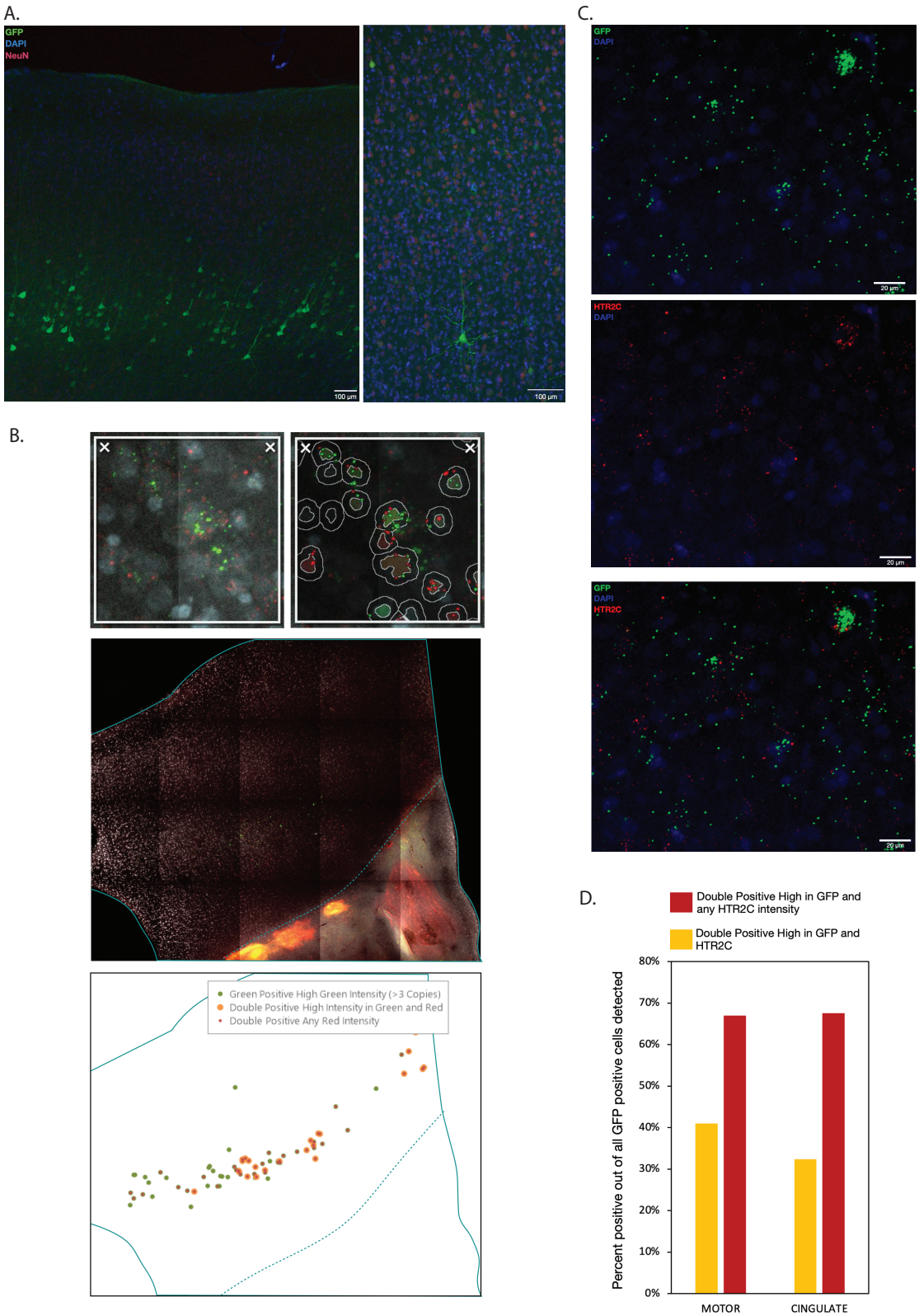
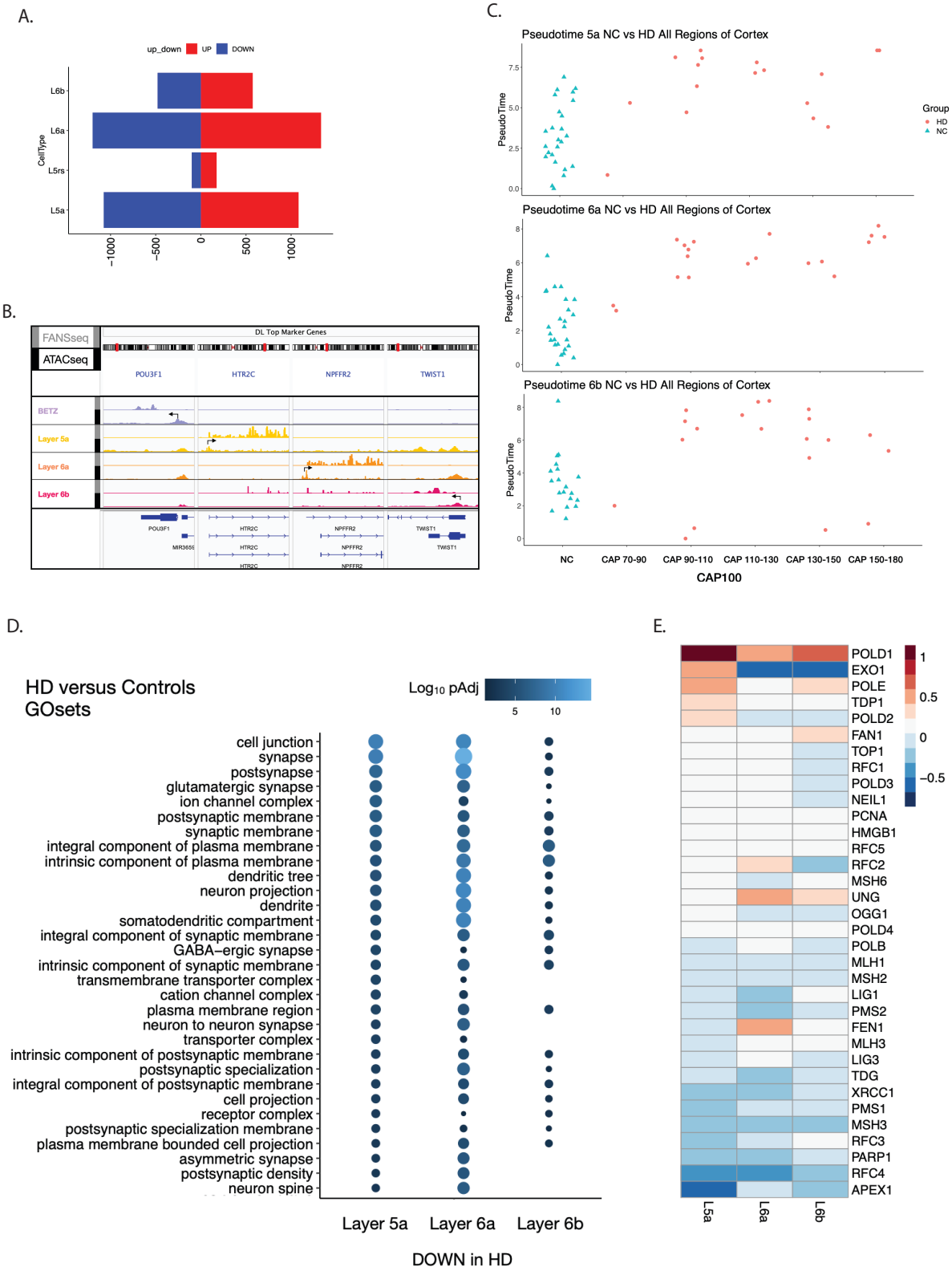


Figure 6
Transcriptional responses in the HD cerebral cortex



SUPPLEMENTAL MATERIALS

List of supplemental files:

Table S1 – Metadata for Tissue Donors.xlsx

Table S2 – Metadata for sFANSseq samples used for analyses.xlsx

SUPPLEMENTARY FIGURE TITLES AND LEGENDS

Supplementary Figure 1.
(**A**) sFANS plots show the position of a motor cortex specific population within the SATB2 medium plus range. Resulting marker gene expression profiles from this isolated population (labeled Betz Cells, purple tract) alongside example IGV tracts from layer2, layer 4, layer 5, layer 6a, and layer 6b excitatory cortical neurons, reveal the expression of *POU3F1*.
(**B**) sFANS plots from motor, prefrontal, cingulate, and insular cortex show position of isolated populations from the SATB2 medium plus range. Resulting marker gene expression profiles show enrichment for *POU3F1*, *ADRA1A*, and *GABRQ* expressing neurons.
(**C**) sFANS plots from motor and prefrontal cortex show the position of an isolated population enriched for the marker gene CRHBP as shown through visualization of resulting RNAseq profiles in IGV.
(**D**) Principal Component Analyses of control data examples for layer 2, layer 4, layer 5rs, layer 5a, layer 6a, layer 6b, VIP, RELN, LAMP5, PVALB, Astrocytes, Microglia, Oligodendrocytes, and Oligodendrocyte Precursor Cells (OPCs).

Supplementary Figure 2.
(**A**) sFANS plots showing examples of the distribution of NeuN positive nuclei along the SATB2 negative, SATB2 medium, and SATB2 high range (top row) for three controls and three HD donors. Bottom row shows the distribution of NeuN positive SATB2 high nuclei along the BCL11B and TLE4 ranges, revealing a distortion of the fluorophore signal of these populations in HD.
B) sFANS plot examples show the position of the Betz cell enriched marker gene population for three controls and three HD donors.
(**C**) Example sFANS plot and resulting IGV tracts confirm the isolation of cortical layer 2 neurons using RNA probes against the LINC00507 marker gene.

Supplementary Figure 3. Schematic representation of the steps for sFANS-guided snRNAseq cluster assignment
(**A**) Utilization of sFANS derived RNAseq data for the calculation of the cell-type specific Specificity Index for all genes. The heatmap on the right shows the top 20 genes rank ordered through the specificity index analyses for all sFANS isolated cell-types.
(**B**) Generation of correlation heatmaps between snRNAseq clusters using the cell-type specific Specificity Index gene ranking, heatmaps scaled to rows and scaled to columns are shown side by side.
(**C**) Known marker gene expression within each cluster, as shown on the heatmap, utilized for refinement and validation of cluster assignment.
(**D**) UMAP with cluster assignments and color coding based on sFANS-guided cell-type designations and marker gene expression (left panel). Scaled to row correlation heatmap shows correlation between sFANS-guided clusters and clusters as defined in the publicly available human motor cortex datasets (Bakken et al., 2021) (right panel).

Supplementary Figure 4. Histograms show the percentage of fragment lengths, detected through analyses of CAG expansion assay data. Data from multiple regions and multiple HD donors is shown for each cell-type. Each horizontal tract represents one cell-type specific sample. The percentage of fragments detected at the modal mutant allele length is shown in white, the percentage of nuclei where the *mHTT* CAG tract has expanded was calculated and is visualized as the percentage of expansion beyond the modal length by 1 to 2 repeats (light grey), 3 to 5 repeats (grey), 6 to 20 repeats (black), or by more than 20 repeat units (red).

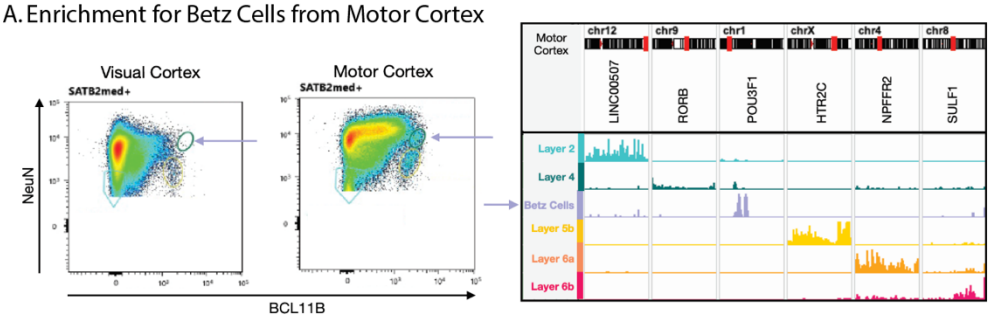
Supplementary Figure 5.
(**A**) Anatomical Mapping of experimental locations in non-human primate ISH experiments. Shown are examples for cingulate and motor cortex containing sections.
(**B**) *In situ* hybridization (ISH) data showing expression of *HTR2C* across multiple regions of human cortex, including primary and secondary visual, temporal, frontal, and cingulate cortex.
(**C**) sFANS plots showing the distribution of SATB2 high nuclei along TLE4 and BCL11B and relative abundance of layer 5a, layer 6a, and layer 6b nuclei in visual, motor, and prefrontal cortex.

Supplementary Figure 6.
(**A**) Principal Component analyses results for samples from HD and control donors are shown for layer 5rs using the top 100 most variable genes, layer 5a using the top 250 most variable genes, layer 6a using the top 100 most variable genes, and layer 6b using the top 100 most variable genes. PCA plots are color coded by gender and group status (female control (F_NC), male control (M_NC), female HD (F_HD), male HD (M_HD).
(**B**) Principal component analyses of the motor cortex specific Betz cell samples from HD and control donors showing principal components two and three using the top 100 most variable genes (left). Bulk pseudotime analyses results for layer 5rs depicting molecular responses along the pseudotime axis in control samples and in HD samples across CAP scores.
(**C**) Results from Gene Ontology Cellular Component pathway enrichment analysis of genes downregulated or upregulated in HD layer 5a versus layer 6 populations ($p\ adj < 0.05$ by DESeq2 and filtered for genes with accessible promotors detected by ATACseq).

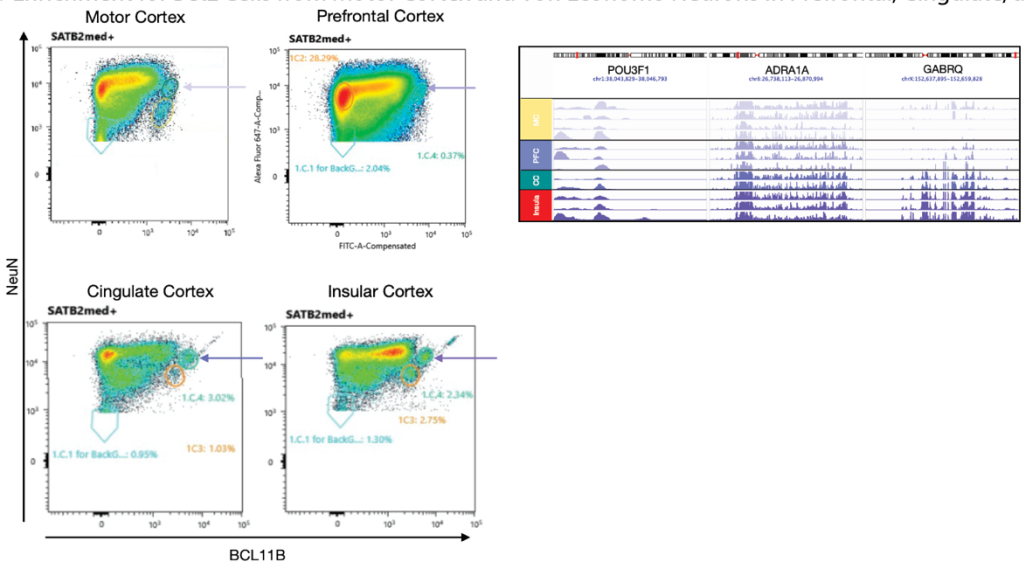
Supplementary Figure 7.
(**A**) Gating Strategies for Neuronal sFANS. Sequential steps for the isolation of neuronal cell-types are shown by the arrows.
(**B**) Gating Strategies for Glial sFANS. Sequential steps for the isolation of glial cell-types are shown by the arrows.

Supplementary Figure 1

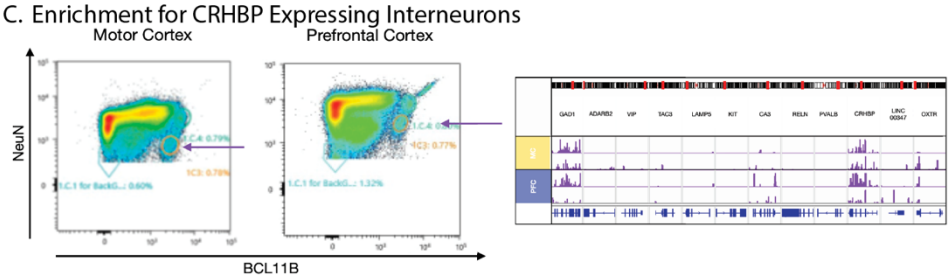
A. Enrichment for Betz Cells from Motor Cortex



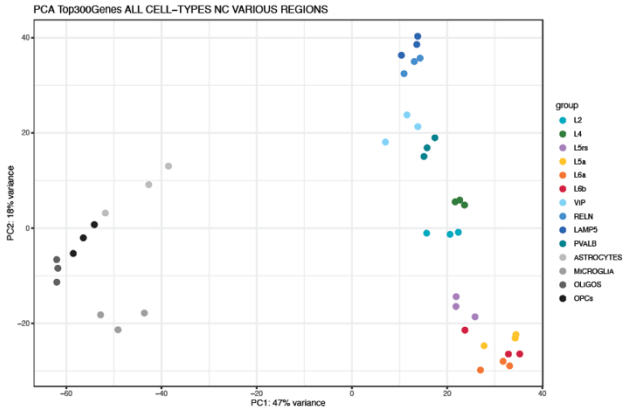
B. Enrichment for Betz Cells from Motor Cortex and Von Economo Neurons in Prefrontal, Cingulate, and Insular Cortex



C. Enrichment for CRHBP Expressing Interneurons

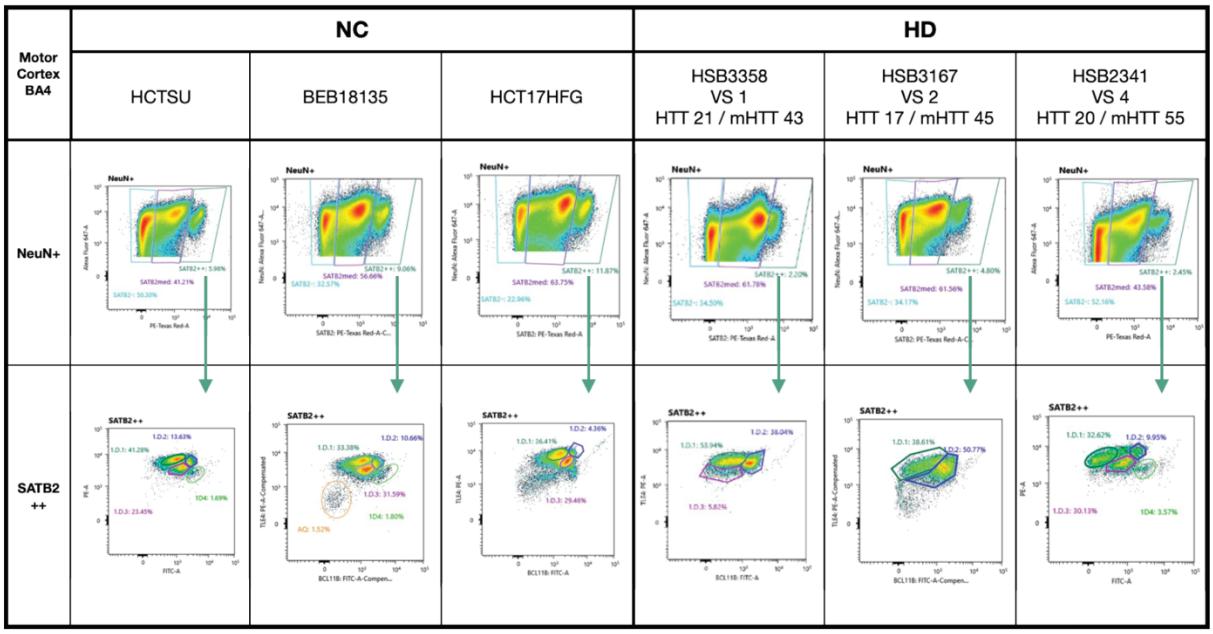


D. Principal Component Analyses of Cortical Cell-Types

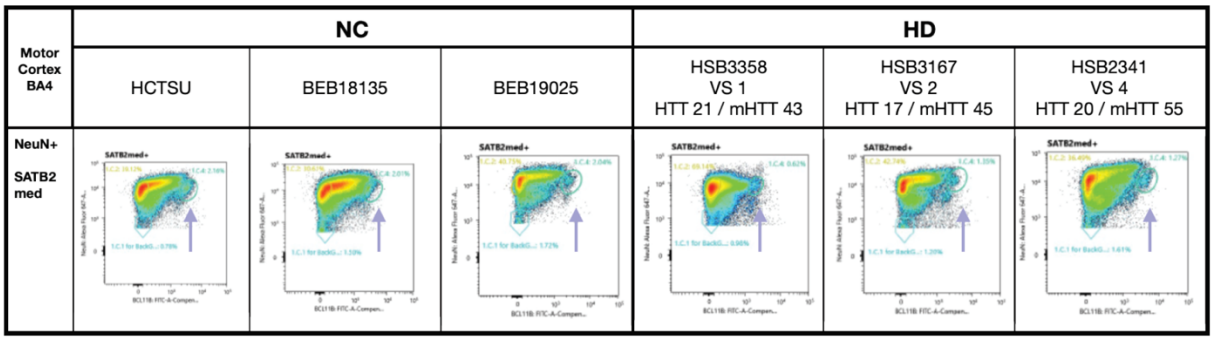


Supplementary Figure 2

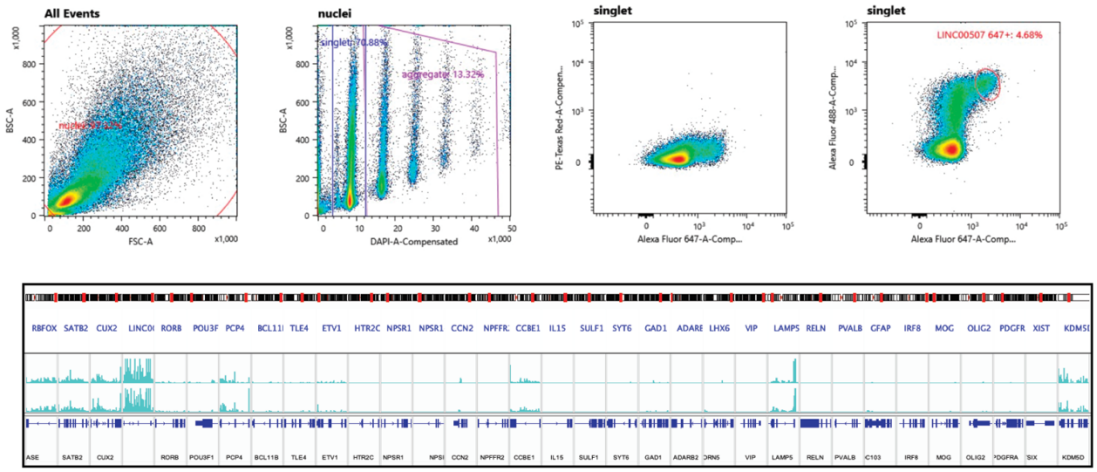
A.



B.

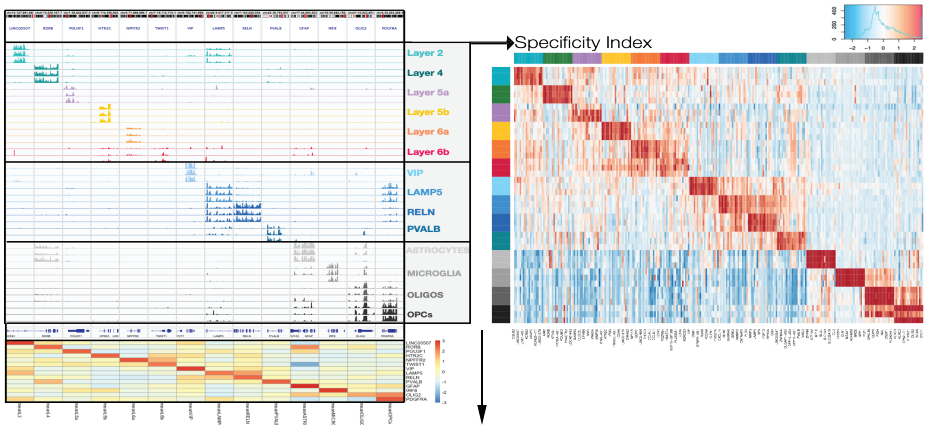


C.

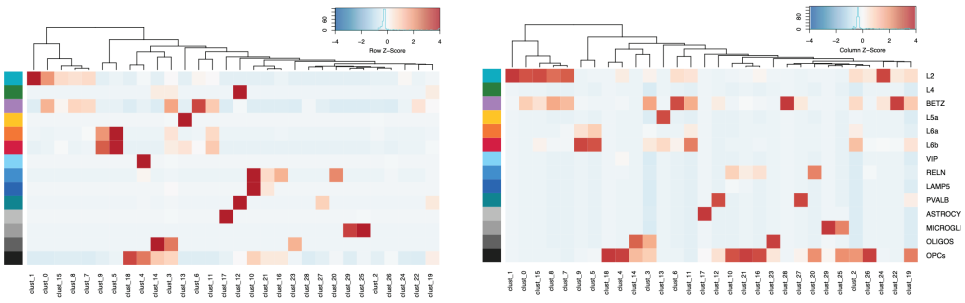


Supplementary Figure 3

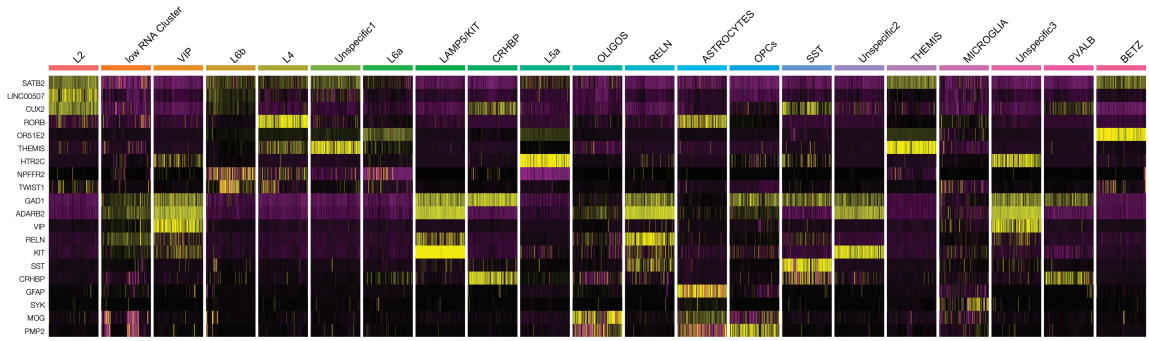
A.



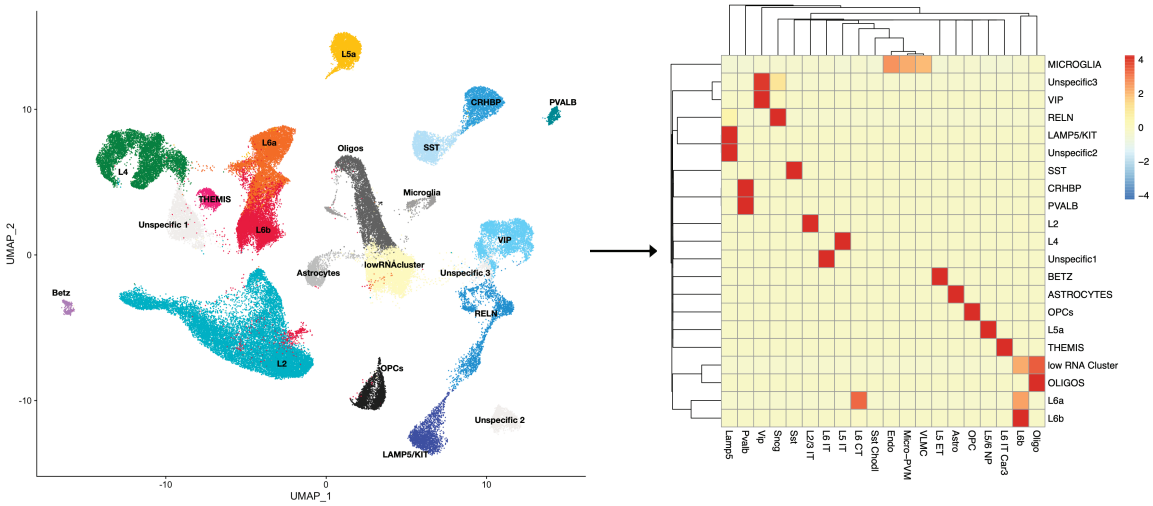
B.



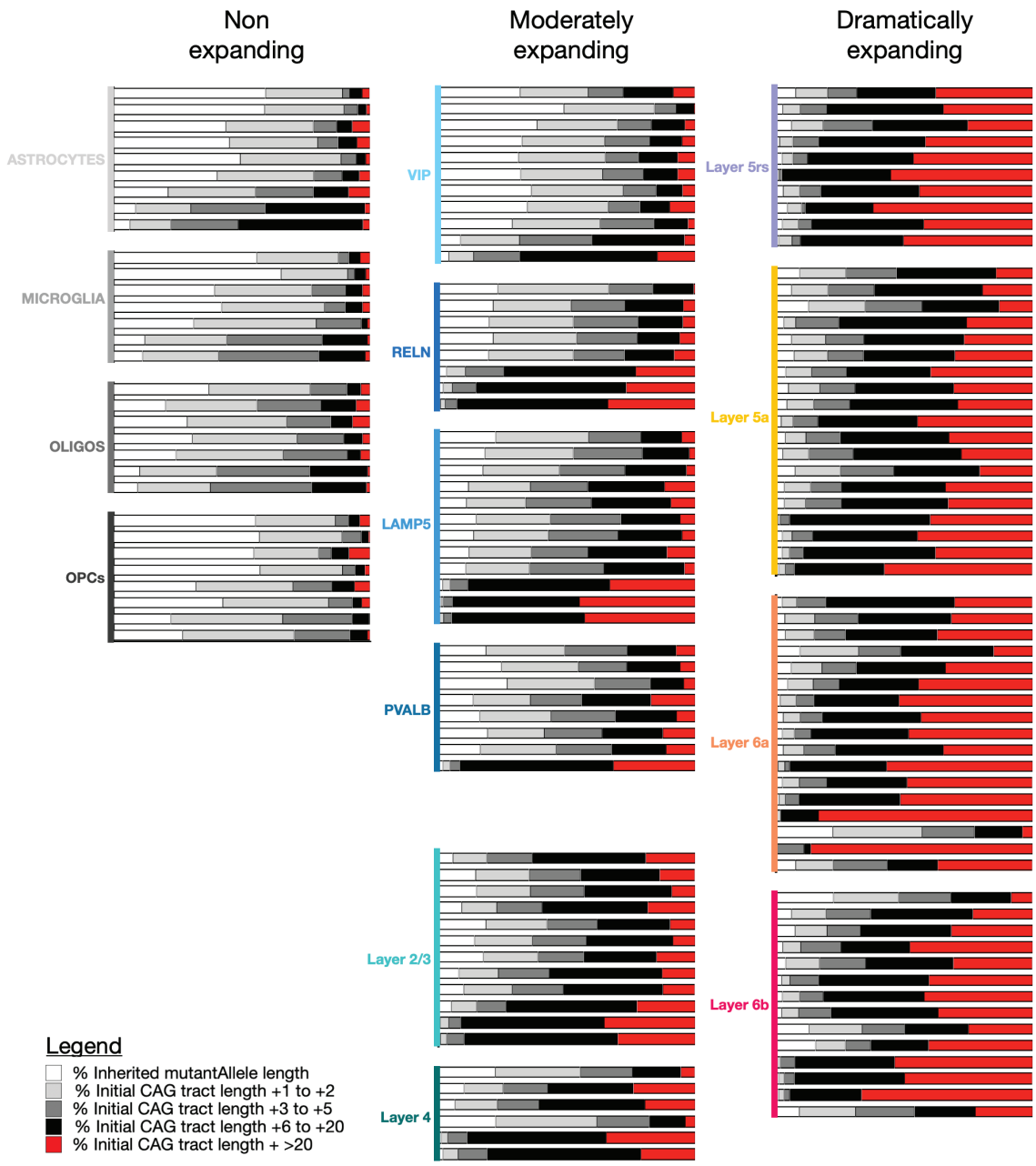
C.



D.

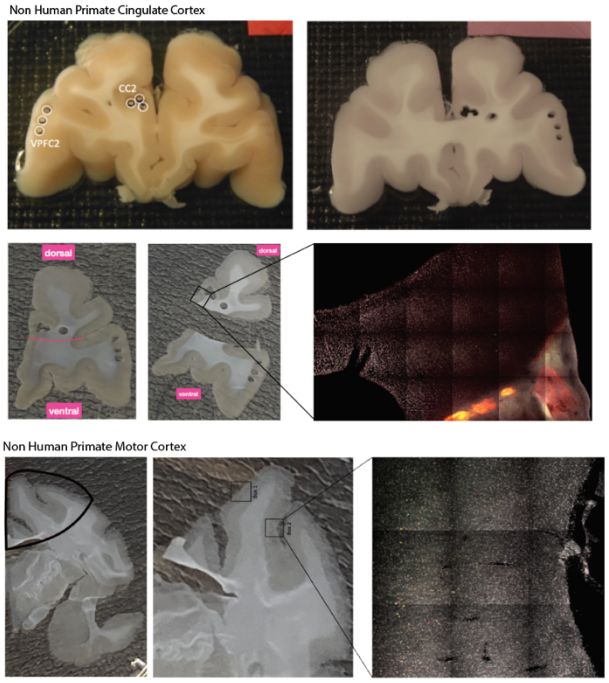


Supplementary Figure 4

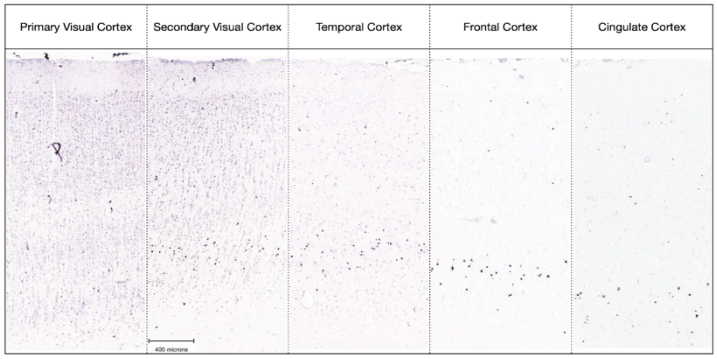


Supplementary Figure 5

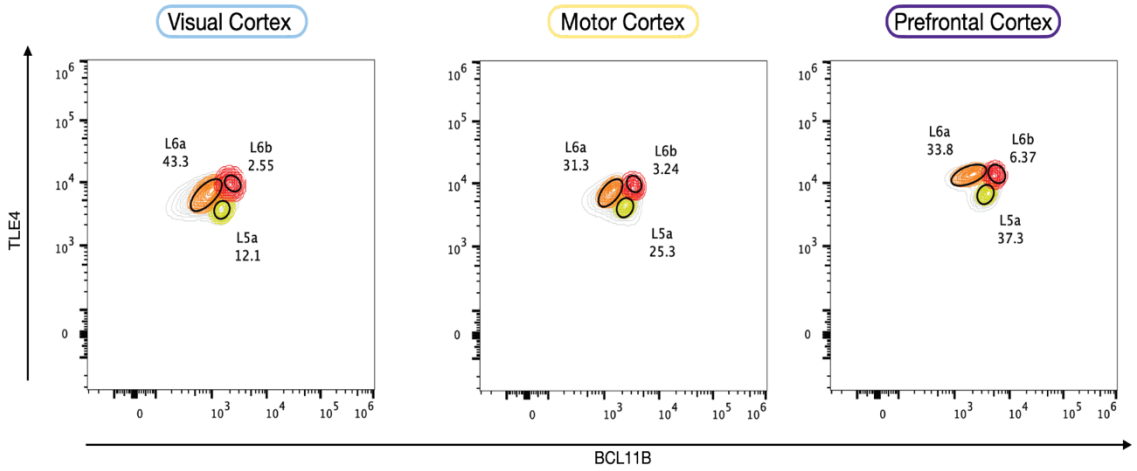
A. Non Human Primate Anatomical Mapping of Cingulate and Motor Cortex Samples



B. Human Cortex IHC in Several Anatomical Regions adapted from the Allen Brain Atlas

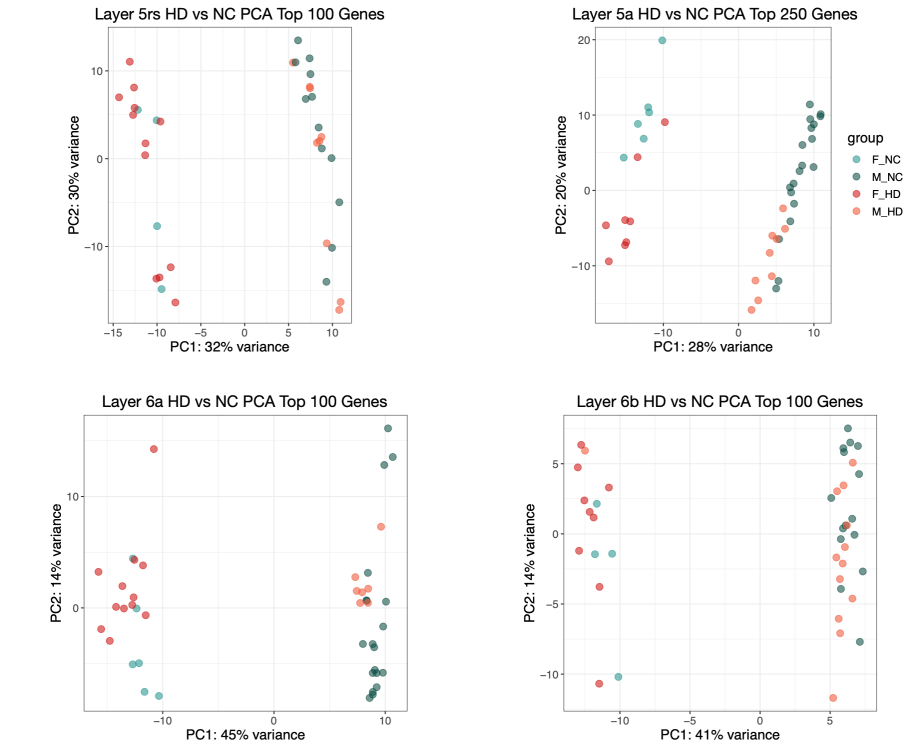


C. Human FANS of Deep Layer Neurons in Visual Motor and Prefrontal Cortex

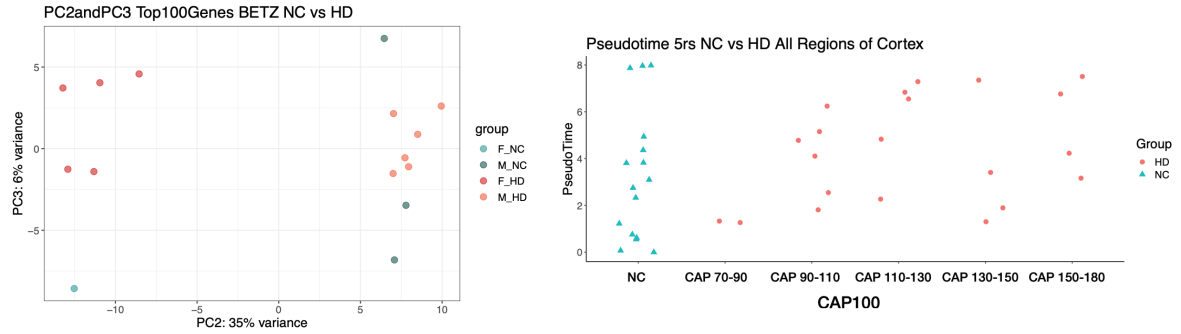


Supplementary Figure 6

A.

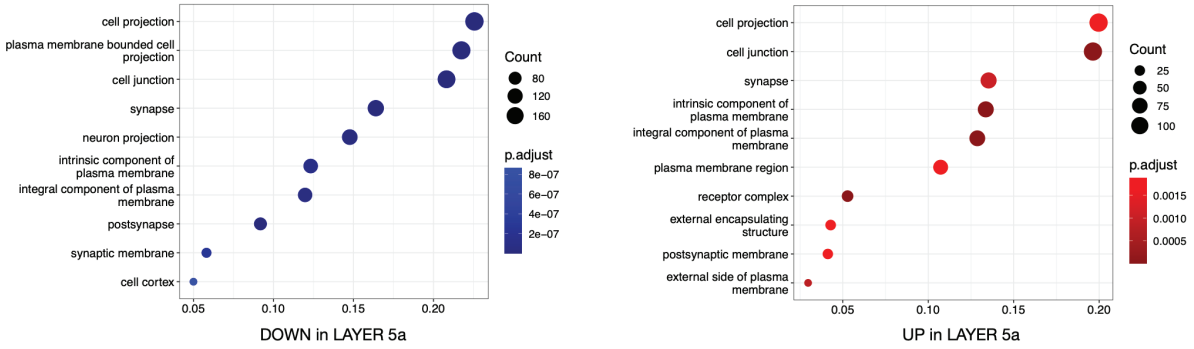


B.



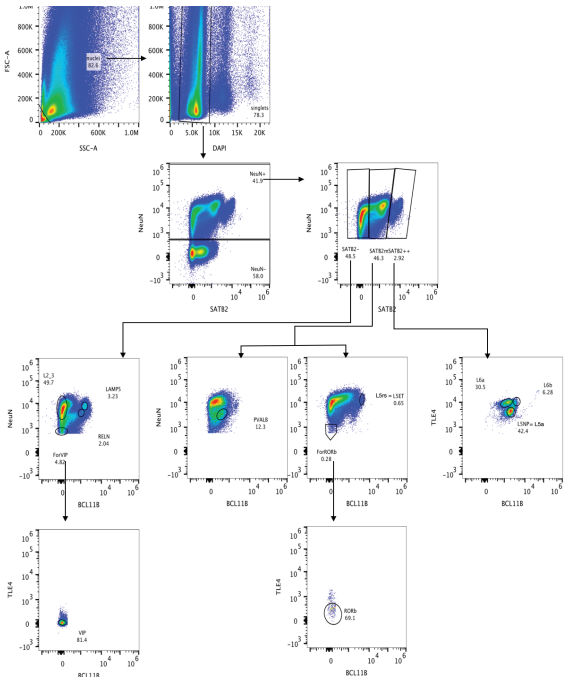
C.

Vulnerable Layer 5a Neurons versus Layer 6 Neurons

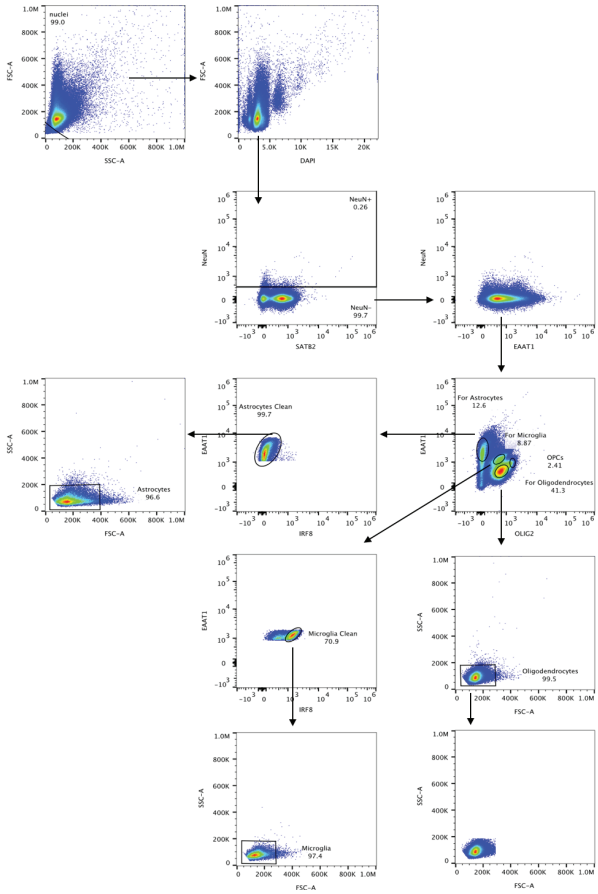


Supplementary Figure 7

A. Gating Strategies for Neuronal FANS



B. Gating Strategies for Glial FANS



STAR METHODS

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
NeuN polyclonal chicken primary antibody	Millipore	ABN91
SATB2 monoclonal mouse primary antibody	Santa Cruz	sc-81376
BCL11B monoclonal rat antibody conjugated to FITC	Abcam	ab123449
TLE4 monoclonal mouse antibody conjugated to PE	Santa Cruz	sc-365406
SLC1A3 monoclonal mouse primary antibody	Santa Cruz	sc-515839
IRF8 monoclonal mouse antibody conjugated to PE	Thermo Fisher	12-9852-82
OLIG2 polyclonal goat primary antibody	R&D Systems	AF2418
Donkey anti-Chicken A647 secondary antibody	Jackson ImmunoResearch	703-605-155
Donkey anti-Mouse A594 secondary antibody	Invitrogen	A-21203
Donkey anti-Mouse A488 secondary antibody	Invitrogen	A-21202
Donkey anti-Goat A594 secondary antibody	Invitrogen	A-11058
eGFP antibody	Thermo Fisher	A-6455
Goat anti-Rabbit A488 secondary antibody	Thermo Fisher	A-11008
Biological samples		
Healthy adult brain tissue	Miami's Brain Endowment Bank	N/A
Huntington's disease patient adult brain tissue	NIH Brain & Tissue Repository-California, Human Brain & Spinal Fluid Resource Center, VA West Los Angeles Medical Center, Los Angeles, California	N/A
	Harvard Brain Tissue Resource Center	N/A
Adult Macaquace AAV2.retro-eGFP sections	Weiss et al.	N/A
Chemicals, peptides, and recombinant proteins (RNA-seq protocol)		
Potassium Chloride	Thermo Fisher	Cat#AM9640G
Magnesium Chloride	Thermo Fisher	Cat#AM9530G
Sucrose	Thermo Fisher	Cat#BP220-1
Tricine	Sigma Aldrich	Cat#T0377
Spermidine	Sigma Aldrich	Cat#S2501
Spermine	Sigma Aldrich	Cat#S2876
EDTA-free Protease Inhibitor Cocktail	Roche	Cat#04693159001
Dithiothreitol	Thermo Fisher	Cat#165685000
SUPERase-In RNase Inhibitor	Thermo Fisher	Cat#AM2696

Recombinant RNasin Ribonuclease Inhibitor	Promega	Cat#N2515
OptiPrep Density Gradient Medium	Sigma Aldrich	Cat#D1556
TrueBlack Lipofuscin Autofluorescence Quencher	Biotium	Cat#23007
Bovine Serum Albumin	Millipore	Cat#126615
Phosphate-Buffered Saline (10x)	Invitrogen	Cat#AM9625
TritonX	Sigma Aldrich	Cat#T8787
Paraformaldehyde 16% Aqueous Solution	Electron Microscopy Sciences	Cat#15710
Glycine	Sigma Aldrich	Cat#G7126
DAPI	Sigma Aldrich	Cat#D9542
Chemicals, peptides, and recombinant proteins (ATAC-seq)		
Sodium Chloride	Sigma Aldrich	Cat#S7653
UltraPure Tris-HCl Buffer, pH 7.5	Invitrogen	Cat#15567027
IGEPAL CA-630	Millipore	Cat#I8896
Dimethyl Formamide	Sigma	Cat#D4254
EDTA	Invitrogen	Cat#AM9260G
UltraPure SDS Solution	Invitrogen	Cat#15553027
Chemicals, peptides, and recombinant proteins (CAG Repeat Expansion Assay)		
Q5 High Fidelity 2x Master Mix	New England BioLabs	Cat#M0492L
AMPure XP Reagent	Beckman Coulter	Cat#A63881
Dimethyl sulfoxide	Sigma	Cat#34869
Critical commercial assays		
AllPrep DNA/RNA FFPE Kit	Qiagen	Cat#80234
Revelo RNA-seq High Sensitivity Module	Tecan	Cat#30201359
Collibri™ Library Quantification Kit	Thermo Fisher	Cat# A38524100
Tagment DNA Enzyme and Buffer Kit	Illumina	Cat#20034198
High Sensitivity RNA ScreenTape	Agilent	Cat#5067-5579
Qubit 1X dsDNA High Sensitivity Assay Kit	Invitrogen	Cat#Q33231
MinElute Reaction Cleanup Kit	Qiagen	Cat#26206
Deposited data		
Raw and analyzed data	This paper	GEO: GSE
Human reference genome NCBI build 37, GRCh37	Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Oligonucleotides		

Primers for HTT-CAG sizing, see Table SX	M. Ciosi, S.A. Cumming, A.M. Alshammari, *et al.* Library preparation and MiSeq sequencing for the genotyping-by-sequencing of the Huntington disease *HTT* exon one trinucleotide repeat and the quantification of somatic mosaicism Protocol Exchange (2018), 10.1038/protex.2018.089	N/A
Primers for ATAC-Seq, see Table SY	Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf (2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position." Nat Methods 10(12): 1213-1218.	N/A
Software and algorithms		
Image J	Schneider et al.	https://imagej.nih.gov/ij/
HALO software	Indica Labs, Albuquerque, NM, USA	https://indicalab.com/halo/
Other		
Scripts related to the analyses in this publication	This paper	https://github.com/cpressl/Human-HD-CTX

MATERIALS AND METHODS

Sample Collection

De-identified human post-mortem brain tissue analyzed in this study was determined to be exempt from Institutional Review Board (IRB) review according to 45 CFR 46.102 (f). For this work, brain samples were obtained from Miami's Brain Endowment Bank™ or through the NIH NeuroBioBank and sourced from either the Harvard Brain Tissue Resource Center (HBTRC) or the NIH Brain & Tissue Repository-California, Human Brain & Spinal Fluid Resource Center, VA West Los Angeles Medical Center, Los Angeles, California (Rothenberg et al.), which is supported in part by National Institutes of Health and the US Department of Veterans Affairs. Inclusion criteria were; a postmortem interval (Hernandez-Garcia et al.) of less than 24 hours, a RIN number of greater than 6, male or female sex, all races and an age range between 20-70 years old. Drug addiction and schizophrenia as well as clinical evidence of brain cancers were reason for sample exclusion. While not preferred, samples from donors with a history of other non-brain cancers and diabetes were accepted and the tissue should not have been thawed and re-frozen. Fresh frozen (FF) as well as formalin-fixed and paraffin embedded (FFPE) specimen were obtained from multiple donors and several Brodmann areas. Cortical samples used in this work were derived from primary motor cortex (MC, Brodmann area BA4), prefrontal cortex (PC, Brodmann area BA9), cingulate cortex (CC, Brodmann area BA24), insular cortex, auditory cortex, specifically Wernicke's area (AC, Brodmann area BA22), visual cortex (VC, including Brodmann areas BA17, as well as Brodmann areas 18/19). Inclusion criteria for HD brain specimen were as described in the section above with the addition that HD samples from individuals who carried around 40 or more CAG repeats in the first exon of the *HTT* gene were selected and obtained from the brain banks. Vonsattel grades, when available, were defined by the brain banks, following classically assessment of striatal pathology (Vonsattel et al., 1985).

Available metadata for each donor sample used in this study is provided in our metadata tables for tissue donors (Table S1) and include information on brain bank source, sex, age, race, post mortem interval (Hernandez-Garcia et al.), year of death, clinical diagnosis, group Normal Control (NC), Huntington's Disease (HD), Vonsattel grade (when available), brain weight, and RNA Integrity Number (RIN) (when available), as well as CAG repeat tract length of the normal allele, CAG repeat tract length of the mutant allele, CAG repeat structure of the normal allele, CAG repeat structure of the mutant allele for all HD donors and a select set of NC donors that were genotyped.

Sample Preparation

Fresh frozen as well as FFPE samples from the requested Brodmann areas were provided by the brain banks and shipped to The Rockefeller University overnight. FFPE sample blocks were shipped at room temperature (RT). FF samples were shipped on dry ice. Upon arrival at The Rockefeller University all samples were photographed and catalogued. FFPE samples were stored at RT and FF specimens were stored at -80ºC. For nuclei isolation experiments, FF sample blocks were further sub-dissected. On the day of sub-dissection, sample blocks were removed from minus 80ºC and moved to minus 40ºC for two hours. After two hours, samples were removed from minus 40ºC and placed on a metal plate on dry ice. Sub-dissections were performed on frozen specimen with the goal of i) obtaining samples which contained all layers of cortex and ii) to arrive at an aliquot size of samples weighing between 150 mg and 350 mg. At no point was the tissue allowed to thaw.

Isolation of Cortical Cell-Types

Procedures for nuclei isolation and sFANS were based on protocols previously established in our lab (Kriaucionis and Heintz, 2009) (Mellen et al., 2012) (Xu et al., 2018) and were applied with modifications. Original protocols had previously been optimized in human samples for the

isolation of cerebellar cell-types and have been presented in detail elsewhere (Xu et al., 2018). General procedures for sFANS and for the application of our sFANSeq strategy will be discussed in brief in the following sections, while modifications of the original protocols for the isolation of human neocortical cell-types will be presented in detail.

Antibodies and Reagents

Seven primary antibodies were utilized to build our multiplexed sFANSeq panel. This multiplexed panel was used for the isolation of all major human cortical cell-types. The seven target proteins and antibodies for the sFANSeq panel included Neuronal Nuclei Antigen (NeuN, Millipore Cat# ABN91, RRID: AB_11205760, used at a concentration of 1:500), Special AT-Rich Sequence Binding Protein 2 (SATB2, Santa Cruz Biotechnology Cat# sc-81376, RRID:AB_1129287, used at a concentration of 1:100), BAF Chromatin Remodeling Complex Subunit (BCL11B aka CTIP2, Abcam Cat# ab123449, RRID:AB_10973033 conjugated to FITC, used at a concentration of 1:100), Transducin Like Enhancer Of Split 4 (TLE4, Santa Cruz Biotechnology Cat# sc-365406, RRID:AB_10841582 conjugated to PE, used at a concentration of 1:100), Solute Carrier Family 1 Member 3 (SLC1A3, Santa Cruz sc-515839 used at a concentration of 1:2000), Interferon Regulatory Factor 8 (IRF8, Thermo Fisher Scientific Cat# 12-9852-82, RRID:AB_2572742 conjugated to PE, used at a concentration of 1:100), Oligodendrocyte Transcription Factor 2 (OLIG2, R and D Systems Cat# AF2418, RRID:AB_2157554, used at a concentration of 1:500). Combination of these seven antibodies into a multiplexed multi-color panel, the use of hierarchical and Boolean logic and considering antibody labeling strength, as well as serial sorting and serial staining strategies allowed for the isolation of up to 16 distinct cell-types from one human cortical sample aliquot. Detailed information on the specific primary and secondary antibodies used in this work is provided in the key resources table.

Nuclei Isolation

Sample aliquots were thawed on ice for approximately 20 minutes. Samples were transferred into homogenization buffer and homogenized through dounce-homogenization and application of 30 strokes with the loose pestle A, followed by 30 strokes with the tight pestle B as described by Xu et al. (Xu et al., 2018). The homogenate was brought to 24% iodixanol concentration and layered on top of a 27% iodixanol cushion for density gradient centrifugation. Centrifugation was performed for 30 minutes in an SW41 rotor in a Coulter Beckman ultracentrifuge at 4ºC, 10,000 RPM, with maximum acceleration and slowest deceleration. Pelleted nuclei were harvested from the bottom of the ultracentrifuge tube and resuspended in homogenization buffer.

Nuclei Preparation for Fluorescence Activated Nuclei Sorting

Except for those samples intended for snRNAseq, nuclei were fixed by adding 1mL of Homogenization buffer with 1mM EDTA and 1% formaldehyde (PFA) for 8 minutes at room temperature. The reaction was quenched through the addition of 0.125 M glycine. After 5 minutes of incubation, nuclei were washed once in homogenization buffer, followed by a transfer wash into blocking buffer containing Triton 0.05% X-100. Nuclei were blocked and permeabilized in Triton X-100 containing blocking buffer for 30 minutes at RT.

True Black Treatment

After blocking, nuclei were treated with True Black (TB) to minimize autofluorescence signal. Nuclei were habituated to non-Triton X-100 containing wash buffer to avoid TB wash-out by Triton X-100 and retain suppression of autofluorescence signal through TB. Nuclei were washed twice in Wash buffer without TritonX-100 and resuspended in 100 µl of 40% ethanol containing TrueBlack Lipofuscin Autofluorescence Quencher (Biotium, #23007) for 40-50 seconds.

Antibody Staining of Nuclei

After isolation, nuclei were resuspended in resuspension buffer in 100μl total volume and incubated with primary antibodies listed in the neuronal sFANS panel for 60 minutes at RT. Antibody NeuN was added at a 1:500 concentration, anti-SATB2, anti-BCL11B conjugated to FITC, and anti-TLE4 conjugated to PE were added at a concentration of 1:100. After 60 minutes, nuclei were washed three times by pelleting (3min, 1000 x g) and resuspension in wash buffer. Secondary antibodies were added to the nuclei in 500μl total volume and at a dilution of 1:1000, for a 30-minute incubation at RT. Detailed information on the specific primary and secondary antibodies used in this work is provided in the antibody staining strategy table. Anti-NeuN antibodies were labeled with anti-chicken 2° antibodies conjugated to a A647 fluorophore. Anti-SATB2 antibodies were labeled with anti-mouse 2° antibodies conjugated to a A594 fluorophore. The concentration of secondary antibodies was kept at 1:1000 in a total volume of 500μ wash buffer and nuclei were incubated for 30 minutes. Three washes in wash buffer were followed by the last resuspension of nuclei in a DAPI (1:25,000) containing wash buffer solution. Nuclei were separated by FACS immediately or stored overnight (ON) at 4ºC.

	ANTIBODIES								
	Probe Name	**Vendor**	**Catalog Number**	**Type**	**Clonality**	**Host Species**	**Target Species**	**Gene**	**Fluorophore**
ALL FANS	NeuN Ab	Millipore	ABN91	primary antibody	Polyclonal	Chicken	/	RBFOX3	/
	Donkey anti-Chicken A647			secondary antibody		Donkey	Chicken	/	A647
Neuronal FANS	SATB2 Ab	Santa Cruz	sc-81376	primary antibody	Monoclonal	Mouse	/	SATB2	/
	Donkey anti-Mouse A594			secondary antibody		Donkey	Mouse	/	A594
	BCL11B Ab	Abcam	ab123449 - FITC	conjugated antibody	Monoclonal	Rat	/	BCL11B	FITC
	TLE4 Ab	Santa Cruz	sc-365406 PE	conjugated antibody	Monoclonal	Mouse	/	TLE4	PE
Glial FANS	SLC1A3 Ab	Santa Cruz	sc-515839	primary antibody	Monoclonal	Mouse	SLC1A3	SLC1A3	
	Donkey anti-Mouse A488			secondary antibody					
	IRF8 Ab	Thermo Fisher	12-9852-82 PE	conjugated antibody	Monoclonal	Mouse	IRF8	IRF8	PE
	OLIG2 Ab	R&D Systems	AF2418	primary antibody	Polyclonal	Goat	OLIG2	OLIG2	
	Donkey anti-Goat A594			secondary antibody					

Table. Antibody Staining Strategy

Flow Cytometry – Serial Sorting, Post-sorting, and Staining Strategies

Boolean logic and serial sorting strategies were applied to maximize the number of cell-types obtainable from each sample aliquot. Sorting strategies were developed to achieve highest efficiency within the inherent constraints of possible fluorophore combinations and technical boundaries of The SONY MA900 Fluorescence Activated Cell Sorting (FACS) instrument. The SONY MA900 was utilized for all experiments described in this work. The first sort, henceforth referred to as the separation sort, was performed to separate non-neuronal nuclei (NeuN-), from three groups of mixed neuronal nuclei populations (NeuN+SATB2-, NeuN+SATB2medium, NeuN+SATB2+). The separation sort was followed by 1) post-separation staining procedures on NeuN negative nuclei and 2) sub-sorts of the three previously sorted NeuN positive nuclei containing tubes. NeuN negative nuclei underwent post-staining procedures, including glia-specific primary and secondary antibody incubations as well as wash steps that were equivalent to the steps described above. Detailed instructions and gating steps for our sFANSeq strategy will be provided below.

sFANS Gating Strategy, Data Processing, Visualization, and Quantification

sFANS Gating and Sorting Strategies

sFANS data pre-processing, analyses, and visualization were performed using the SONY visualization tool as well as the FlowJo™ v10.8 Software (BD Life Sciences) (FlowJo™ Software for Mac. Ashland, OR: Becton, Dickinson and Company; 2023) . sFANS .fcs files were loaded into FlowJo, and gates were drawn. Fluorophore measurements were plotted in biexponential mode.

The multi-step gating strategy to isolate ten neuronal cortical cell-types included:

1. Size selection of nuclei was performed by excluding particles below approximately 100K along the forward scatter area (FSC-A) and side scatter area (SSC-A) axes and smaller than the majority of other particles measured.
2. Doublet discrimination and singlet selection was performed by gating on the lowest positive population along the liner DAPI axis.
3. Separation of NeuN positive neuronal nuclei from NeuN negative non-neuronal nuclei was performed along the NeuN axis.
4. NeuN positive nuclei were plotted along the NeuN and SATB2 axes and divided into three populations NeuN positive populations:
    A. NeuN+/SATB2- nuclei, a population located below the 10^2 mark along the SATB2 axis. This group of nuclei was plotted against NeuN and BCL11B, revealing four subpopulations.
        A.i. NeuN+low/SATB2-/BCL11B- nuclei are enriched for VIP expressing interneurons. This VIP expressing population was further purified by plotting and gating against nuclei that spread upwards along the TLE4 axis.
        A.ii. NeuN+high/SATB2-/BCL11B- nuclei are enriched for CUX2 expressing cortical layer 2 excitatory neurons.
        A.iii. NeuN+high/SATB2-/BCL11Bmedium nuclei are enriched for a RELN expressing inhibitory neuron population.
        A.iv. NeuN+high/SATB2-/BCL11B+ nuclei are enriched for a LAMP5 expressing inhibitory neuron population.
    B. NeuN+/SATB2medium nuclei, a population located above the 10^2 mark along the SATB2 axis and separated below the highest SATB2 positive population. NeuN positive SATB2 medium nuclei were further divided into

NeuN+/SATB2mediumlow and NeuN+/SATB2mediumhigh populations and plotted against NeuN and BCL11B.

B.i. NeuN+low/SATB2mediumlow/BCL11B+ nuclei are enriched for PVALB expressing inhibitory neurons.

B.ii. NeuN+low/SATB2mediumhigh/BCL11B- nuclei are enriched for layer 4 excitatory neurons. This RORb expressing population was further purified by plotting and gating against nuclei that spread upwards along the TLE4 axis.

B.iii. NeuN+high/SATB2mediumhigh/BCL11B+high is a small region-specific population, enriched for the Betz or Von Economo neuron transcription factor POU3F1. Populations isolated from this location are region-specific and are henceforth referred to as layer 5rs to allow for consistent comparisons across all regions.

B.iv. A second population within the SATB2 medium high range appeared NeuN+low/SATB2mediumhigh/BCL11B+high and was only inconsistently present across samples from the same regions. Due to its inconsistency, it has been difficult to define this population. However, since this population may represent a rare and potentially region-specific CRHBP expressing interneuron population it will continue to be of interest to further dissolve and define this population.

C. NeuN+/SATB2+, a population forming at the highest end along the SATB2 axis, contained three populations within it.

C.i. NeuN+/SATB2+/BCL11B+/TLE4+ nuclei were enriched for *HTR2C* expressing layer 5a excitatory neurons.

C.ii. NeuN+/SATB2+/BCL11B+/TLE4+high nuclei were enriched for CTGF and NPFFR2 expressing layer 6a excitatory neurons.

C.iii. NeuN+/SATB2+/BCL11B+high/TLE4+high nuclei were enriched for CTGF expressing layer 6b excitatory neurons. This cell population is sparse, does not express NPFFR2, and specific marker genes appear variable across the different regions of cortex.

Gating strategy to isolate four non-neuronal cell-types:

1. Previously separated NeuN negative nuclei underwent size selection for nuclei using the FSC and SSC scatter, as described above.
2. Single nuclei selection was performed through assessment of DAPI staining intensity as described.
3. NeuN negative nuclei staining intensity was plotted for the non-neuronal antibody markers utilized and with the goal to isolate four NeuN negative populations:
    A. NeuN negative EAAT1 (aka SLC1A3) positive nuclei, for astrocytes
    B. NeuN negative IRF8 positive nuclei, for microglia
    C. NeuN negative OLIG2 positive nuclei, for oligodendrocytes
    D. NeuN negative OLIG2 high positive nuclei, for OPCs

Supplementary Figure7 shows our gating strategy for the isolation of neuronal nuclei and non-neuronal nuclei from one control donor motor cortex sample.

sFANS Data Processing, Visualization, and Quantification Methods

sFANS data were processed using the FlowJo™ v10.8 Software (BD Life Sciences) (FlowJo™ Software for Mac. Ashland, OR: Becton, Dickinson and Company; 2023). For population quantifications, data were batch processed followed by individual sample inspection and gate correction as warranted. Gating plots were produced in batches using the layout editor function. Raw data for quantification analyses were processed in batches using the table editor function

and results were extracted in .csv format. Quantification data was read into R-Studio for statistical analyses and to produce sFANS population quantification plots.

DNA and RNA Extraction, Purification, and cDNA Library Construction from Nuclear RNA

For RNAseq experiments, RNA was extracted from fixed and sorted nuclei using the Qiagen AllPrep DNA/RNA FFPE kit (Qiagen, #80234). The original Qiagen RNeasy FFPE kit protocol had been previously optimized for our sample type. DNA from all samples was isolated and purified Genomic DNA was purified using AllPrep DNA/RNA FFPE Kit (Qiagen, #80234), purified DNA was stored at -20ºC.

RNA Sequencing

Libraries were assessed for concentration and quality using the Qubit and Tapestation assays. Samples were pooled and sequenced using the NovaSeq SP 2x150 with the goal to reach a read depth of approximately 40 million reads per sample.

Nuclei Preparation for Single Nucleus RNA Sequencing

Nuclei intended for 10X snRNAseq experiments remained unfixed but were otherwise treated as described above and identically to nuclei intended for all other down-stream applications. Further processing of nuclei isolated for snRNAseq experiments was performed by The Rockefeller University's Genomics Resource Center. In brief, the number of nuclei per sample was counted and the appropriate loading amount was calculated. Nuclei were loaded and processed on the 10X platform for snRNAseq. 1000121 Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1, 16 rxns, 1000215 Dual Index Kit TT Set A 96 rxns, and Chromium Single Cell Controller were used. Before deep sequencing, the resulting libraries were pooled in groups of 3 - 4 samples and were sequenced on MiSeq using the Nano 300 kit. The MiSeq results provided 1 million reads over all samples and were utilized to assess the ratio of empty droplets and

estimate nuclei yield. Finally, all samples were pooled and sequenced on NovaSeq using the S4 flow cell 2x100 kit to generate 8 billion reads over all samples and to allow for sufficient read depth given the total number of nuclei.

ATACseq of sFANS-Enriched Deep Layer Cortical Neurons

The ATACseq protocol utilized in this work was adapted from Corces et al. (Corces et al., 2017). Sample processing steps will be outlined in brief below. For our ATACseq experiments we focused on four deep layer cell-types of interest and performed nuclei isolations as described above from motor cortex samples of control donors. Following nuclei sorting, nuclei were spun for 5 minutes at 950 g, supernatant was removed and for samples containing up to 50,000 nuclei, 50μl lysis buffer (10 mM Tris-HCl pH 7.6, 10 mM NaCl, 3 mM MgCl2, 0.01% NP-40) was used for resuspension. For samples containing more than 50,000 nuclei, the additional amount of 10μl lysis buffer for every additional 10,000 nuclei was used. Nuclei were spun at 500 g for 10 minutes at 4ºC, supernatant was removed, and nuclei were resuspended in 200μl transposition mix (1x TD buffer containing X μl of Illumina Tagment DNA TDE1 Enzyme per every X nuclei). Samples were incubated on a heat block set to 37ºC for 30 minutes. After 30 minutes, 150μl reverse X-link solution was added, and nuclei were incubated at 55ºC overnight at 500 RPM. On day two, DNA was isolated using the Qiagen MinElute Reaction Cleanup Kit. Next, PCR amplification was performed. PCR amplification reagents were mixed in a clean strip tube before the sample was added. Amplification was performed using the following settings: (72C – 5min, 98C – 30sec, appropriate number of amplification cycles x (98C - 10sec, 63C – 30sec, 72C – 1 min) with NEBNext® High-Fidelity 2X PCR Master Mix (New England Biolabs, #M0541S) and barcoded Nextera primers (1.25 μM each) (Buenrostro et al., 2013). The number of amplification cycles was dependent on the number of nuclei contained within the sample. Less than 15,000 nuclei, 15 cycles were used, for samples containing between 15,000 and 100,000 nuclei, 13 cycles were

used. The maximum input was kept below 75,000 nuclei for 13 PCR cycles. Upon completion of the library amplification steps, double sided bead purification was performed. Qubit and Tapestation were used to confirm correct library sizing and sufficient sample concentrations before pooling. Samples were pooled for a minimum concentration of 2.6 nM total in 100μl volume and ATACseq libraries were sequenced using NovaSeq6000 SP 2x100.

Sample Sectioning for Immunohistochemistry and *In-situ* Hybridization

FFPE samples were sectioned on a Leica HistoCore AUTOCUT Rotary Microtome. Fresh frozen tissue was sectioned on a Leica CM3050 S Cryostat. Samples were mounted on SuperFrost™ Plus and FFPE slides were stored at RT, whereas fresh frozen sections were stored at -20ºC. Nonhuman primate (NHP) samples were provided by our collaborator Dr. Jodi McBride. Details on tissue processing as performed by our collaborators as well as IHC methods for GFP labeling were as described by Weiss et al. (Weiss 2020). Photo documentation, as provided by our collaborator and anatomical landmarks (Supplementary Figure S5A) were used to define the anatomical regions contained within the sections used in our experiments.

Immunohistochemistry

Standard operating procedures were based on IHC protocols sourced from the BioLegend website and applied with modifications. In brief, FFPE slides were deparaffinized with xylene, dehydrated in baths of decreasing EtOH concentrations. Endogenous peroxidase activity was blocked through incubation with 3% $_2$ in methanol and slides were rinsed. Antigen retrieval was performed in citrate buffer (pH 6.0) heated to 95-100ºC for 10 minutes. Slides were set aside for cooling for 20 minutes and rinsed two times with PBS. Next, slides were incubated in blocking buffer containing 10% normal donkey serum in PBS with 4 drops Avidin per ml buffer. Sections were coverslipped and incubated in a humidified chamber at RT for one hour. Blocking buffer was drained from the slide and primary antibodies in the appropriate dilutions (1:100) in dilution buffer

containing 1% normal donkey serum in PBS with 4 drops Avidin per ml buffer. Sections were coverslipped and incubated in a humidified chamber at 4ºC overnight. On day two, slides were washed with PBS followed by incubation of biotinylated secondary antibody (1:200) in dilution buffer. Sections were coverslipped and incubated in a humidified chamber at RT for 30 minutes. Slides were washed with PBS and Streptavidin-HRP conjugate (NEL-700 NEN Life Science Products) (1:100) in 1% normal donkey serum containing PBS was applied. Sections were coverslipped and incubated in a humidified chamber at RT for 30 minutes and protected from light. Slides were washed and incubated with TSA (Tyramide Signal Amplification Biotin System, NEL-700 NEN Life Science Products) diluted to a 1:100 concentration in amplification diluent. Sections were coverslipped and incubated in a humidified chamber at RT for 10 minutes. Slides were washed with PBS and finally, slides were cover-slipped and incubated at RT with ABC (ABC Kit Elite PK-6100, Standard Vector Labs) for 30 min at RT. Slides were washed and DAB substrate solution (0.05% DAB, 0.015% $_2$ in PBS, 2 DAB kit tablets, 1 gold & 1 silver in 15ml d) was applied to the slides, using a disposable pipet. Development time was generally kept below 5 min and sustained until the desired color intensity was reached. Slides were washed with PBS, mounted with aquamount, and dried overnight before imaging was performed.

IHC experiments for visualization of AAV2.retro treated NHP sections were performed following the methods described in Weiss et al. (Weiss et al., 2020). In brief, the 40 micron thick sections were incubated with an antibody against eGFP (Thermo Fisher, A-6455, 1:1000) over night at room temperature followed by one hour incubation with a goat anti rabbit secondary Alexa Fluor A488 (Thermo Fisher, A-11008, 1:500) conjugated secondary antibody.

In-situ Hybridization

In-situ hybridization experiments performed over the course of this work were executed using the commercially available RNAscope® assay kit. Multiplexed fluorescent ISH were performed following the vendor's protocols according to the tissue type used. Tissue types used in this work

were limited to fresh frozen NHP sections. The following ACD protocols were used: User Manual RNAscope® 2.5 Multiplex Fluorescent Reagent Kit v2 Assay (With Sample Preparation and Pretreatment, Document Number 323100-USM), Sample Preparation Technical Note for Fixed Frozen Tissue Using RNAscope® 2.5 Chromogenic Assay (Singleplex and Duplex) TN-320534/Rev B). Protocol adaptations for the 40-micron thick fixed frozen NHP sample sections were as follows: Sections were mounted onto Superfrost Plus slides and dried at RT for 1 hour. Slides were dipped in water 3 times and dried at RT for one hour followed by a backing step at 60ºC for 1 hour. Post-fixation was performed using 4% PFA at 4ºC for 1 hour. Slides were dehydrated with an increasing EtOH gradient, and slides were dried at 60ºC for 15 minutes. Next, samples were incubated with RNAscope hydrogen peroxide at RT for 10 minutes or until no active bubbling was observed. Slides were rinsed with water and dried for 15 minutes at 60ºC. Target retrieval was performed in 1X TR buffer for 10 minutes at 100ºC using a steamer apparatus. Slides were rinsed, briefly dipped in 100% EtOH and dried at 60ºC for 10 minutes or longer, until dry. The hydrophobic barrier was drawn on the slide, enclosing the tissue sample and Protease III was applied for digestion for 30 minutes at 40ºC. Slides were quickly washed with dH_2O before the samples were further processed using the appropriate RNAscope® protocol.

Imaging and Analyses of IHC and ISH Sections

All imaging of non-fluorescent samples was performed on the NanoZoomer-SQ Digital slide scanner (Hamamatsu C13140-01) and using the Hamamatsu NDPScanMonitor software program. Imaging of fluorescent samples was performed on the LSM700 Zeiss Confocal Microscope, using the Zeiss ZEN microscopy software. IHC imaging results from experiments using our cortical neuronal sFANS panel antibodies on FFPE sections and using the above described IHC DAB protocol are not shown. ISH imaging data was inspected and processed for image overlays using ImageJ (Schneider et al., 2012). ISH images were read into the HALO

software (Indica Labs, Albuquerque, NM, USA). Signal for GFP and *HTR2C* probes were automatically evaluated by the software and three categories of colocalization signal were extracted to generate anatomical maps: cells strongly positive for eGFP, cells strongly positive for both eGFP and *HTR2C*, and cells strongly positive for eGFP and weakly positive for *HTR2C*. localization maps were generated. Out of all GFP positive cells, the percentage of double positive cells high in eGFP and in *HTR2C*, as well as the percentage of double positive cells high in eGFP and with any *HTR2C* signal intensity were calculated and plotted in one bar chart showing motor and cingulate cortex sample results.

RNAseq Data Processing

Sequence and transcript coordinates for human hg38 UCSC genome and gene models were retrieved from the BSgenome.Hsapiens.UCSC.hg38 Bioconductor package (version 1.4.1) and TxDb.Haspiens.UCSC.hg38.knownGene Bioconductor libraries (version 3.4.0) respectively. RNAseq reads were aligned to the genome using Rsubread's subjunc method (version 1.30.6) (Liao et al., 2013) and exported as bigWigs normalized to reads per million using the rtracklayer package (version 1.40.6). Reads in genes were counted using the featurecounts function (Liao et al., 2014) within the Rsubread package against the full gene bodies for nuclear RNAseq (Patro et al., 2017). Normalization and rlog transformation of raw read counts in genes were performed using DESeq2 (Love et al., 2016) (Love et al., 2018) (version 1.20.0) and between sample variability was assessed with hierarchical clustering and heat maps of between sample distances implemented in the pheatmap R package (1.0.10).

ATACseq Data Processing

The ATACseq reads were aligned with the hg38 genome from the BSgenome.Hsapiens.UCSC.hg38 Bioconductor package (version 1.4.1) with Rsubread's align method in paired-end mode. Fragments between 1 and 5,000 base pairs long were considered

correctly paired. Normalized, fragment signal bigWigs were created with the rtracklayer package. Peak calls were made with MACS2 software in BAMPE mode (Feng et al., 2012) (Zhang et al., 2008). For each deep layer excitatory neuron cell-type of motor cortex the ATAC-seq consensus peaks were called from ATAC-seq datasets generated from control donor samples of BA4 motor cortex. These datasets included layer 5rs (aka Betz cells in BA4) (n=3), layer 5a (n=4), layer 6a (n=4), and layer 6b (n=4), Differential ATACseq signals were identified with the DESeq2 package (Love et al., 2016) (Love et al., 2018) (version 1.20.0). High confidence consensus peaks were derived by creating a nonredundant peak set for each cell type and disease state, and then filtering down to peaks that were present in the majority of samples. These were then annotated to TSS based on proximity using the ChIPseeker package (version 1.28.3) (Wang et al., 2022). NCBI Refseq hg38 gene annotation was used (version 109.20211119).

RNAseq and ATACseq data were inspected using the Integrative Genomics Viewer (Robinson et al., 2011). RNAseq average raw read counts were processed using the DESeq2 package (Love et al., 2016) (Love et al., 2018) (version 1.36.0). Reads were converted into normalized counts by DESeq2 before differential gene expression analysis. For deep layer cell-types, differential gene expression analysis were performed for each cell-type separately. PCA plots were generated using the transformed, variance stabilized data (using the vst() function) and plotted using the plotPCA function. Differential RNAseq gene expression analysis results were filtered using our cell-type specific ATACseq datasets to exclude genes that lacked consensus ATAC peaks over their annotated TSS positions in NCBI Refseq hg38 (version 109.20211119). The filtered lists of differentially expressed genes (DEGs) ($p\ adj < 0.05$) with accessible TSS regions were used for analysis of over-representation of Gene Ontology Cellular Compartment (GOCC) terms with clusterProfiler package (Yu et al., 2012) (version 4.4.4). The list of all genes with accessible TSS regions was used as the ‘background’ list for comparison (‘universe’). Gene lists were generated separately for layer 5a, layer 5rs, layer 6a, and layer 6b samples from differential

gene expression analyses between HD and NC samples (P adj. < 0.05) for the identification of GOCC pathways containing an over-representation of genes that showed disease-associated up- or downregulation in HD. Differential expression between HD and controls for a set of genes of interest was extracted for each cell-type separately and plotted using the pheatmap R package (1.0.10).

Bulk-pseudotime Analyses of sFANSeq Data

We generated bulk pseudotime trajectories using sFANSeq data from HD and control donor deep layer cell-types, including layer 5rs, layer 5a, layer 6a, and layer 6b. To do this, the gene expression levels of each sample were aggregated into a single matrix. This counting matrix was loaded into Seurat object with the Seurat::CreateSeuratObject() function, generating a Seurat (v 4.0, https://satijalab.org/seurat/) object for each cell-type separately. Next, data were log normalized, variable features were assessed, data was scaled, and PCAs were generated using Seurat. Next, pseudotime trajectories were modeled using slingshot (https://github.com/kstreet13/slingshot). Results were plotted to show the pseudotime trajectory of HD samples along the range of different CAP100 scores and in relation to control samples.

snRNAseq Data Processing

The snRNAseq raw data was processed using the count function in Cell Ranger (10X Genomics, version 6.0). Reads were aligned to the human genome (hg38) and including the intron regions, genes were mapped and counted in each nucleus. Counting matrices were generated in Cell Ranger and utilized in downstream analyses. Data was read into R and further processed using Seurat algorithms as well as other Bioconductor packages and custom code.

To estimate and remove ambient RNA, the counting matrices were loaded into SoupX (https://github.com/constantAmateur/SoupX). The fraction of ambient RNA contamination was

estimated and corrected by using the default settings. Doublets, defined as multiple nuclei in a single droplet, were identified by using Scrublet (https://github.com/swolock/scrublet). Details on rationale and computational steps for removal of ambient RNA can be found here: https://bioconductor.org/books/release/OSCA/droplet-processing.html.

Cell cycle phase was estimated as a quality control measure by using cyclone, a function of scran (https://bioconductor.org/packages/release/bioc/html/scran.html). For each sample, a Seurat (v 4.0, https://satijalab.org/seurat/) object was created by using the corrected matrices. The previously estimated information on doublet scores and cell cycle stages was also added to the Seurat object. After this step, the proportion of mitochondrial genes for each sample was estimated using the Seurat::PercentFeatureSet() function. Nuclei with mitochondrial contents greater than 1% were removed. Distribution of gene counts (detected), transcript count (sum), and mitochondrial content (subsets_Mito_percent) in each cell was calculated and the outliers, with extremely high/low levels in either of the three criteria, were removed. Then, gene expression was normalized by using Seurat::NormalizeData() with logNormalization and scale factor=10000. And lastly, the top 3000 variable features of each object were identified by using Seurat::FindVariableFeatures() function on the Variance Stabilization Transformed (vst) data.

The pre-processed Seurat objects were integrated into one single object by using the canonical correlation analysis (CCA) workflow. In brief, anchors, which constitute shared variable features across all samples, were identified by using the Seurat::FindIntegrationAnchors() function. Then, all datasets were integrated using Seurat::IntegrateData() function. After integration, the assay named “integrated” was set as the default assay and the integrated object was scaled by using the Seurat::ScaleData() function.

To optimize projection and clustering, we programmatically assessed the different cluster-resolution outcomes when different numbers of PCs were used. Principal component analyses (PCA) were performed using Seurat::RunPCA () with up to 40 PCs. The optimized number of PCs was determined by using the elbow plot. Then, the selected number of PCs was applied to produce projection maps using the Seurat::FindNeighbors() and Seurat::RunUMAP functions. After UMAPs were generated, we tested clustering by Seurat::FindClusters() with resolutions ranging from 0.1 to 1.0. Results were plotted using the clustree function (https://cran.r-project.org/web/packages/clustree/vignettes/clustree.html). Then, the optimal resolution with the most clusters identified and the least crosstalk between cluster lineages was applied to our data.

The merged dataset was converted into SingleCellExperiment format and cell type – clustering was performed by using the SingleR package with default settings (https://bioconductor.org/packages/release/bioc/html/SingleR.html). To annotate cell-types, specificity index results, which provide ranked marker cell-type specific gene lists as described in (Dougherty et al., 2010), for all cell-types isolated through the sFANSeq method from three representative control donor samples were calculated and applied to a reference dataset converted in SingleCellExperiment format. Correlation matrixes were calculated and plotted and snRNAseq clusters were assigned to sFANSeq cell-types based on their correlation scores. For cell-type clusters that were not isolated through our sFANSeq panels, known marker genes as well as cross-correlations with published human motor cortex datasets (Hodge et al., 2019) was performed. Lastly, previously published human motor cortex datasets (Bakken et al., 2021) were projected onto our datasets and heatmap correlation matrices were produced to resolve the last remaining unspecific cluster identities and to confirm our cluster assignments. For the quantitation of fraction nuclei in NC vs HD cell-type specific clusters we utilized full donor datasets for which NeuN negative, NeuN positive / SATB2 negative, NeuN positive / SATB2 medium, and NeuN positive / SATB2 high nuclei could be harvested. Abundance of nuclei in clusters contained within

the NeuN positive / SATB2 negative, NeuN positive / SATB2 medium gates were compared between NC and HD samples separately from those clusters contained within the NeuN positive / SATB2 high gates. Percentage of nuclei in each cluster over all nuclei sampled within the sort-gate was calculated and comparisons between HDs and NCs were performed. Comparisons of quantitation were performed using Bonferroni multiple comparison corrected T-test and results with an adjusted P value < 0.05 were marked significant.

Cell-Type Specific *HTT*-CAG Repeat Expansion Assay

The protocol for this assay was adapted from Ciosi et al. (Ciosi et al., 2019) (Ciosi et al., 2021). Nuclei were isolated following the protocol procedures described in the previous section. DNA was purified using AllPrep DNA/RNA FFPE Kit (Qiagen, #80234) and concentrated in a vacuum concentrator if required and stored at -20ºC until further processing. Before initiation of procedures to perform the cell-type specific *HTT*-CAG repeat expansion assay, DNA sample concentrations were measured with Qubit (ThermoFisher #Q32851) and 2-10 ng of genomic DNA was used as template in PCR. Up to 10 ng of gDNA was amplified in a 20 µl volume using NEBNext® High-Fidelity 2X PCR Master Mix (New England Biolabs, #M0541S) supplemented with 5% dimethyl sulfoxide and barcoded primers specific to *HTT* Exon 1 (0.5 µM each). Samples were added to the primer dilutions and a mix containing DMSO and Q5®High-Fidelity DNA Polymerase was added. Samples were PCR amplified using the following settings: (96ºC – 1min, 30 x (96ºC - 30sec, 61ºC – 45sec, 72ºC – 3 min), 72ºC – 10min). After the PCR reaction, 10µl of ddH_2O was added to each sample in 20µl and small pools were prepared with the goal for each sample to be represented at similar input concentrations within the pool. Purification was performed by adding 0.55 X AMPure beads. Samples were incubated for 10 minutes, transferred to a magnet, and incubated for 5 minutes. Supernatant was removed and beads were washed twice with 80% EtOH. Beads were dried for 5 minutes and resuspended in 32µl RNase-nuclease-free H_2O.

Samples were run on Tapestation to check for primer-dimers and the purification process was repeated if necessary. Pool concentrations were assessed by qPCR, using Collibri™ Library Quantification Kit (ThermoFisher # A38524100). Final pooling was based on the concentrations as measured by the Collibri quantification method and samples were sequenced on an Illumina MiSeq Sequencer using a MiSeq Reagent Nano Kit with 30% PhiX spike-in control and 400 nt forward read with output of both indexes.

Cell-Type Specific *HTT*-CAG Repeat Expansion Data Analyses

Data derived from the Cell-Type Specific *HTT*-CAG Repeat Expansion assay was processed following the steps as outlined by Ciosi et al. (Ciosi et al., 2021). In brief, all HD donors were genotyped for the exact structure ((CAG)n(CAACAG)n(CCGCCA)n(CCG)7(CCT)n or ((CAG)n(CAACAG)n(CCGCCA)n(CCG)10(CCT)n). Sequences derived from the Cell-Type Specific *HTT*-CAG Repeat Expansion Assay were aligned to the appropriate custom-made reference sequences according to the donor's genotype (CCG)7 or (CCG)10). The number of reads aligned to the reference sequences carrying between 1 and 120 CAG repeats was counted and R code was written to generate result outputs in the form of text files and figures.

***HTT* CAG tract sizing**

Demultiplexed sequencing reads were aligned to a set of *HTT* exon 1 reference sequences that differed by the number of CAG repeat units in the repeat tract using the Burrows-Wheeler Aligner (BWA MEM default settings except: -O 6,6 -E 4,4). The number of reads uniquely mapped to each of the reference sequences in the set was considered to reflect the distribution of CAG tract lengths in the two *HTT* alleles of the analyzed cell population. *HTT* read mapping data from each donor was inspected manually for determining the nucleotide sequence of the adjacent polyproline tract and the presence/absence of interruptions in CAG tract. If *mHTT* exon 1 structure of the donor was atypical, then sequencing reads were realigned to a set of

reference sequences matching that *mHTT* exon 1 structure. The length of CAG repeat tracts reliably mapped was limited to 113 repeat units. The ratio of somatic expansions (RoSE) (Ciosi et al., 2019) and frequency of *mHTT* copies expanded by more than 20 repeat units (RU) were calculated as follows:

$$normal\ allele\ RoSE = \frac{R_{N+1} + R_{N+2} + R_{N+3}}{R_N}$$

$$mutant\ allele\ RoSE = \frac{\sum_{i=M+1}^{113} Ri}{R_M}$$

$$\%\ of mutant\ allele\ copies\ expanded\ by > 20\ RU = \frac{\sum_{i=M+21}^{113} Ri}{\sum_{i=M}^{113} Ri} \times 100\%$$

Uninterrupted CAG tract lengths of progenitor/unexpanded *mHTT* allele (M repeat units) and normal *HTT* allele (N repeat units) were defined from the two modes of mapped read length-distribution in CAG-sizing data from non-expanding cell types (for example microglia and astrocytes). R is the number of reads mapped to a reference sequence with the specified CAG tract length.

MAIN REFERENCES

Alcamo, E. A., L. Chirivella, M. Dautzenberg, G. Dobreva, I. Farinas, R. Grosschedl, and S. K. McConnell. 2008. "Satb2 regulates callosal projection neuron identity in the developing cerebral cortex." *Neuron* 57 (3): 364-77. https://doi.org/10.1016/j.neuron.2007.12.012. https://www.ncbi.nlm.nih.gov/pubmed/18255030 https://www.cell.com/neuron/pdf/S0896-6273(07)01017-3.pdf.

Andrew, S. E., Y. P. Goldberg, B. Kremer, H. Telenius, J. Theilmann, S. Adam, E. Starr, F. Squitieri, B. Lin, M. A. Kalchman, and et al. 1993. "The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease." *Nat Genet* 4 (4): 398-403. https://doi.org/10.1038/ng0893-398. https://www.ncbi.nlm.nih.gov/pubmed/8401589 https://www.nature.com/articles/ng0893-398.

Baker, A., B. Kalmbach, M. Morishima, J. Kim, A. Juavinett, N. Li, and N. Dembrow. 2018. "Specialized Subpopulations of Deep-Layer Pyramidal Neurons in the Neocortex: Bridging Cellular Properties to Functional Consequences." *J Neurosci* 38 (24): 5441-5455. https://doi.org/10.1523/JNEUROSCI.0150-18.2018. https://www.ncbi.nlm.nih.gov/pubmed/29798890 https://www.jneurosci.org/content/jneuro/38/24/5441.full.pdf.

Bakken, T. E., N. L. Jorstad, Q. Hu, B. B. Lake, W. Tian, B. E. Kalmbach, M. Crow, R. D. Hodge, F. M. Krienen, S. A. Sorensen, J. Eggermont, Z. Yao, B. D. Aevermann, A. I. Aldridge, A. Bartlett, D. Bertagnolli, T. Casper, R. G. Castanon, K. Crichton, T. L. Daigle, R. Dalley, N. Dee, N. Dembrow, D. Diep, S. L. Ding, W. Dong, R. Fang, S. Fischer, M. Goldman, J. Goldy, L. T. Graybuck, B. R. Herb, X. Hou, J. Kancherla, M. Kroll, K. Lathia, B. van Lew, Y. E. Li, C. S. Liu, H. Liu, J. D. Lucero, A. Mahurkar, D. McMillen, J. A. Miller, M. Moussa, J. R. Nery, P. R. Nicovich, S. Y. Niu, J. Orvis, J. K. Osteen, S. Owen, C. R. Palmer, T. Pham, N. Plongthongkum, O. Poirion, N. M. Reed, C. Rimorin, A. Rivkin, W. J. Romanow, A. E. Sedeno-Cortes, K. Siletti, S. Somasundaram, J. Sulc, M. Tieu, A. Torkelson, H. Tung, X. Wang, F. Xie, A. M. Yanny, R. Zhang, S. A. Ament, M. M. Behrens, H. C. Bravo, J. Chun, A. Dobin, J. Gillis, R. Hertzano, P. R. Hof, T. Hollt, G. D. Horwitz, C. D. Keene, P. V. Kharchenko, A. L. Ko, B. P. Lelieveldt, C. Luo, E. A. Mukamel, A. Pinto-Duarte, S. Preissl, A. Regev, B. Ren, R. H. Scheuermann, K. Smith, W. J. Spain, O. R. White, C. Koch, M. Hawrylycz, B. Tasic, E. Z. Macosko, S. A. McCarroll, J. T. Ting, H. Zeng, K. Zhang, G. Feng, J. R. Ecker, S. Linnarsson, and E. S. Lein. 2021. "Comparative cellular analysis of motor cortex in human, marmoset and mouse." *Nature* 598 (7879): 111-119. https://doi.org/10.1038/s41586-021-03465-8. https://www.ncbi.nlm.nih.gov/pubmed/34616062 https://www.nature.com/articles/s41586-021-03465-8.pdf.

Barry, J., M. T. N. Bui, M. S. Levine, and C. Cepeda. 2022. "Synaptic pathology in Huntington's disease: Beyond the corticostriatal pathway." *Neurobiol Dis* 162: 105574. https://doi.org/10.1016/j.nbd.2021.105574. https://www.ncbi.nlm.nih.gov/pubmed/34848336.

Bernard, A., L. S. Lubbers, K. Q. Tanis, R. Luo, A. A. Podtelezhnikov, E. M. Finney, M. M. McWhorter, K. Serikawa, T. Lemon, R. Morgan, C. Copeland, K. Smith, V. Cullen, J. Davis-Turak, C. K. Lee, S. M. Sunkin, A. P. Loboda, D. M. Levine, D. J. Stone, M. J. Hawrylycz, C. J. Roberts, A. R. Jones, D. H. Geschwind, and E. S. Lein. 2012. "Transcriptional architecture of the primate neocortex." *Neuron* 73 (6): 1083-99. https://doi.org/10.1016/j.neuron.2012.03.002. https://www.ncbi.nlm.nih.gov/pubmed/22445337.

Blumenstock, S., and I. Dudanova. 2020. "Cortical and Striatal Circuits in Huntington's Disease." *Front Neurosci* 14: 82. https://doi.org/10.3389/fnins.2020.00082. https://www.ncbi.nlm.nih.gov/pubmed/32116525 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7025546/pdf/fnins-14-00082.pdf.

Cajal, SR. 1999. "Texture of the Nervous System of Man and the Vertebrates." *Springer Science & Business Media* Vol 1.

Cepeda, C., D. M. Cummings, V. M. Andre, S. M. Holley, and M. S. Levine. 2010. "Genetic mouse models of Huntington's disease: focus on electrophysiological mechanisms." *ASN Neuro* 2 (2): e00033. https://doi.org/10.1042/AN20090058. https://www.ncbi.nlm.nih.gov/pubmed/20396376.

Ciosi, M., A. Maxwell, S. A. Cumming, D. J. Hensman Moss, A. M. Alshammari, M. D. Flower, A. Durr, B. R. Leavitt, R. A. C. Roos, Track-Hd team, H. D. team Enroll, P. Holmans, L. Jones, D. R. Langbehn, S. Kwak, S. J. Tabrizi, and D. G. Monckton. 2019. "A genetic association study of glutamine-encoding DNA sequence structures, somatic CAG expansion, and DNA repair gene variants, with Huntington disease clinical outcomes." *EBioMedicine* 48: 568-580. https://doi.org/10.1016/j.ebiom.2019.09.020. https://www.ncbi.nlm.nih.gov/pubmed/31607598 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6838430/pdf/main.pdf.

Dougherty, J. D., E. F. Schmidt, M. Nakajima, and N. Heintz. 2010. "Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells." *Nucleic Acids Res* 38 (13): 4218-30. https://doi.org/10.1093/nar/gkq130. https://www.ncbi.nlm.nih.gov/pubmed/20308160 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910036/pdf/gkq130.pdf.

Elam, J. S., M. F. Glasser, M. P. Harms, S. N. Sotiropoulos, J. L. R. Andersson, G. C. Burgess, S. W. Curtiss, R. Oostenveld, L. J. Larson-Prior, J. M. Schoffelen, M. R. Hodge, E. A. Cler, D. M. Marcus, D. M. Barch, E. Yacoub, S. M. Smith, K. Ugurbil, and D. C. Van Essen. 2021. "The Human Connectome Project: A retrospective." *Neuroimage* 244: 118543. https://doi.org/10.1016/j.neuroimage.2021.118543. https://www.ncbi.nlm.nih.gov/pubmed/34508893.

Erben, L., and A. Buonanno. 2019. "Detection and Quantification of Multiple RNA Sequences Using Emerging Ultrasensitive Fluorescent In Situ Hybridization Techniques." *Curr Protoc Neurosci* 87 (1): e63. https://doi.org/10.1002/cpns.63. https://www.ncbi.nlm.nih.gov/pubmed/30791216.

Espinoza, F. A., J. Liu, J. Ciarochi, J. A. Turner, V. M. Vergara, A. Caprihan, M. Misiura, H. J. Johnson, J. D. Long, J. H. Bockholt, J. S. Paulsen, and V. D. Calhoun. 2019. "Dynamic functional network connectivity in Huntington's disease and its associations with motor and cognitive measures." *Hum Brain Mapp* 40 (6): 1955-1968. https://doi.org/10.1002/hbm.24504. https://www.ncbi.nlm.nih.gov/pubmed/30618191 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6865767/pdf/HBM-40-1955.pdf.

Espinoza, F. A., J. A. Turner, V. M. Vergara, R. L. Miller, E. Mennigen, J. Liu, M. B. Misiura, J. Ciarochi, H. J. Johnson, J. D. Long, H. J. Bockholt, V. A. Magnotta, J. S. Paulsen, and V. D. Calhoun. 2018. "Whole-Brain Connectivity in a Large Study of Huntington's Disease Gene Mutation Carriers and Healthy Controls." *Brain Connect* 8 (3): 166-178. https://doi.org/10.1089/brain.2017.0538. https://www.ncbi.nlm.nih.gov/pubmed/29291624.

Estrada-Sanchez, A. M., and G. V. Rebec. 2013. "Role of cerebral cortex in the neuropathology of Huntington's disease." *Front Neural Circuits* 7: 19. https://doi.org/10.3389/fncir.2013.00019. https://www.ncbi.nlm.nih.gov/pubmed/23423362 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3575072/pdf/fncir-07-00019.pdf.

Gargouri, F., A. Messe, V. Perlbarg, R. Valabregue, P. McColgan, L. Yahia-Cherif, S. Fernandez-Vidal, A. Ben Hamida, H. Benali, S. Tabrizi, A. Durr, and S. Lehericy. 2016. "Longitudinal changes in functional connectivity of cortico-basal ganglia networks in manifests and premanifest huntington's disease." *Hum Brain Mapp* 37 (11): 4112-4128. https://doi.org/10.1002/hbm.23299. https://www.ncbi.nlm.nih.gov/pubmed/27400836.

Georgiou-Karistianis, N., M. A. Gray, D. Jf Dominguez, A. R. Dymowski, I. Bohanna, L. A. Johnston, A. Churchyard, P. Chua, J. C. Stout, and G. F. Egan. 2013. "Automated differentiation of pre-diagnosis Huntington's disease from healthy control individuals based on quadratic discriminant analysis of the basal ganglia: the IMAGE-HD study." *Neurobiol Dis* 51: 82-92. https://doi.org/10.1016/j.nbd.2012.10.001. https://www.ncbi.nlm.nih.gov/pubmed/23069680.

Groh, A., H. S. Meyer, E. F. Schmidt, N. Heintz, B. Sakmann, and P. Krieger. 2010. "Cell-type specific properties of pyramidal neurons in neocortex underlying a layout that is modifiable depending on the cortical area." *Cereb Cortex* 20 (4): 826-36. https://doi.org/10.1093/cercor/bhp152. https://www.ncbi.nlm.nih.gov/pubmed/19643810.

Hasselbalch, S. G., G. Oberg, S. A. Sorensen, A. R. Andersen, G. Waldemar, J. F. Schmidt, K. Fenger, and O. B. Paulson. 1992. "Reduced regional cerebral blood flow in Huntington's disease studied by SPECT." *J Neurol Neurosurg Psychiatry* 55 (11): 1018-23. https://doi.org/10.1136/jnnp.55.11.1018. https://www.ncbi.nlm.nih.gov/pubmed/1469396.

Hattox, A. M., and S. B. Nelson. 2007. "Layer V neurons in mouse cortex projecting to different targets have distinct physiological properties." *J Neurophysiol* 98 (6): 3330-40. https://doi.org/10.1152/jn.00397.2007. https://www.ncbi.nlm.nih.gov/pubmed/17898147.

Hevner, R. F. 2007. "Layer-specific markers as probes for neuron type identity in human neocortex and malformations of cortical development." *J Neuropathol Exp Neurol* 66 (2): 101-9. https://doi.org/10.1097/nen.0b013e3180301c06. https://www.ncbi.nlm.nih.gov/pubmed/17278994.

Hickman, R. A., P. L. Faust, K. Marder, A. Yamamoto, and J. P. Vonsattel. 2022. "The distribution and density of Huntingtin inclusions across the Huntington disease neocortex: regional correlations with Huntingtin repeat expansion independent of pathologic grade." *Acta Neuropathol Commun* 10 (1): 55. https://doi.org/10.1186/s40478-022-01364-1. https://www.ncbi.nlm.nih.gov/pubmed/35440014.

Hodge, R. D., T. E. Bakken, J. A. Miller, K. A. Smith, E. R. Barkan, L. T. Graybuck, J. L. Close, B. Long, N. Johansen, O. Penn, Z. Yao, J. Eggermont, T. Hollt, B. P. Levi, S. I. Shehata, B. Aevermann, A. Beller, D. Bertagnolli, K. Brouner, T. Casper, C. Cobbs, R. Dalley, N. Dee, S. L. Ding, R. G. Ellenbogen, O. Fong, E. Garren, J. Goldy, R. P. Gwinn, D. Hirschstein, C. D. Keene, M. Keshk, A. L. Ko, K. Lathia, A. Mahfouz, Z. Maltzer, M. McGraw, T. N. Nguyen, J. Nyhus, J. G. Ojemann, A. Oldre, S. Parry, S. Reynolds, C. Rimorin, N. V. Shapovalova, S. Somasundaram, A. Szafer, E. R. Thomsen, M. Tieu, G. Quon, R. H. Scheuermann, R. Yuste, S. M. Sunkin, B. Lelieveldt, D. Feng, L. Ng, A. Bernard, M. Hawrylycz, J. W. Phillips, B. Tasic, H. Zeng, A. R. Jones, C. Koch, and E. S. Lein. 2019. "Conserved cell types with divergent features in human versus mouse cortex." *Nature* 573 (7772): 61-68. https://doi.org/10.1038/s41586-019-1506-7. https://www.ncbi.nlm.nih.gov/pubmed/31435019.

Hodge, R. D., J. A. Miller, M. Novotny, B. E. Kalmbach, J. T. Ting, T. E. Bakken, B. D. Aevermann, E. R. Barkan, M. L. Berkowitz-Cerasano, C. Cobbs, F. Diez-Fuertes, S. L. Ding, J. McCorrison, N. J. Schork, S. I. Shehata, K. A. Smith, S. M. Sunkin, D. N. Tran, P. Venepally, A. M. Yanny, F. J. Steemers, J. W. Phillips, A. Bernard, C. Koch, R. S. Lasken, R. H. Scheuermann, and E. S. Lein. 2020. "Transcriptomic evidence that von Economo neurons are regionally specialized extratelencephalic-projecting excitatory

neurons." *Nat Commun* 11 (1): 1172. https://doi.org/10.1038/s41467-020-14952-3. https://www.ncbi.nlm.nih.gov/pubmed/32127543
https://www.nature.com/articles/s41467-020-14952-3.pdf.
Hong, E. P., M. E. MacDonald, V. C. Wheeler, L. Jones, P. Holmans, M. Orth, D. G. Monckton, J. D. Long, S. Kwak, J. F. Gusella, and J. M. Lee. 2021. "Huntington's Disease Pathogenesis: Two Sequential Components." *J Huntingtons Dis* 10 (1): 35-51. https://doi.org/10.3233/JHD-200427. https://www.ncbi.nlm.nih.gov/pubmed/33579862
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7990433/pdf/jhd-10-jhd200427.pdf
https://content.iospress.com:443/download/journal-of-huntingtons-disease/jhd200427?id=journal-of-huntingtons-disease%2Fjhd200427.
Jenkins, B. G., H. D. Rosas, Y. C. Chen, T. Makabe, R. Myers, M. MacDonald, B. R. Rosen, M. F. Beal, and W. J. Koroshetz. 1998. "1H NMR spectroscopy studies of Huntington's disease: correlations with CAG repeat numbers." *Neurology* 50 (5): 1357-65. https://doi.org/10.1212/wnl.50.5.1357. https://www.ncbi.nlm.nih.gov/pubmed/9595987.
Jernigan, T. L., D. P. Salmon, N. Butters, and J. R. Hesselink. 1991. "Cerebral structure on MRI, Part II: Specific changes in Alzheimer's and Huntington's diseases." *Biol Psychiatry* 29 (1): 68-81. https://doi.org/10.1016/0006-3223(91)90211-4. https://www.ncbi.nlm.nih.gov/pubmed/1825793.
Johnson, E. B., G. Ziegler, W. Penny, G. Rees, S. J. Tabrizi, R. I. Scahill, and S. Gregory. 2021. "Dynamics of Cortical Degeneration Over a Decade in Huntington's Disease." *Biol Psychiatry* 89 (8): 807-816. https://doi.org/10.1016/j.biopsych.2020.11.009. https://www.ncbi.nlm.nih.gov/pubmed/33500176
https://www.biologicalpsychiatryjournal.com/article/S0006-3223(20)32069-2/pdf.
Kim, J., B. S. Moon, B. C. Lee, H. Y. Lee, H. J. Kim, H. Choo, A. N. Pae, Y. S. Cho, and S. J. Min. 2017. "A Potential PET Radiotracer for the 5-HT2C Receptor: Synthesis and in Vivo Evaluation of 4-(3-[(18)F]fluorophenethoxy)pyrimidine." *ACS Chem Neurosci* 8 (5): 996-1003. https://doi.org/10.1021/acschemneuro.6b00445. https://www.ncbi.nlm.nih.gov/pubmed/28194935
https://pubs.acs.org/doi/pdf/10.1021/acschemneuro.6b00445.
Knowles, R., N. Dehorter, and T. Ellender. 2021. "From Progenitors to Progeny: Shaping Striatal Circuit Development and Function." *J Neurosci* 41 (46): 9483-9502. https://doi.org/10.1523/JNEUROSCI.0620-21.2021. https://www.ncbi.nlm.nih.gov/pubmed/34789560.
Kuhl, D. E., M. E. Phelps, C. H. Markham, E. J. Metter, W. H. Riege, and J. Winter. 1982. "Cerebral metabolism and atrophy in Huntington's disease determined by 18FDG and computed tomographic scan." *Ann Neurol* 12 (5): 425-34. https://doi.org/10.1002/ana.410120504. https://www.ncbi.nlm.nih.gov/pubmed/6217782.
Kuwert, T., H. W. Lange, K. J. Langen, H. Herzog, A. Aulich, and L. E. Feinendegen. 1989. "Cerebral glucose consumption measured by PET in patients with and without psychiatric symptoms of Huntington's disease." *Psychiatry Res* 29 (3): 361-2. https://doi.org/10.1016/0165-1781(89)90090-5. https://www.ncbi.nlm.nih.gov/pubmed/2532753.
Kuwert, T., H. W. Lange, K. J. Langen, H. Herzog, A. Aulich, and L. E. Feinendegen. 1990. "Cortical and subcortical glucose consumption measured by PET in patients with Huntington's disease." *Brain* 113 (Pt 5): 1405-23. https://doi.org/10.1093/brain/113.5.1405. https://www.ncbi.nlm.nih.gov/pubmed/2147116
Landwehrmeyer, G. B., S. M. McNeil, L. S. th Dure, P. Ge, H. Aizawa, Q. Huang, C. M. Ambrose, M. P. Duyao, E. D. Bird, E. Bonilla, and et al. 1995. "Huntington's disease gene: regional and cellular expression in brain of normal and affected individuals." *Ann Neurol* 37 (2): 218-30. https://doi.org/10.1002/ana.410370213. https://www.ncbi.nlm.nih.gov/pubmed/7847863.

Layburn, F. E., A. Y. S. Tan, N. F. Mehrabi, M. A. Curtis, L. J. Tippett, C. P. Turner, N. Riguet, L. Aeschbach, H. A. Lashuel, M. Dragunow, R. L. M. Faull, and M. K. Singh-Bains. 2022. "N-terminal mutant huntingtin deposition correlates with CAG repeat length and symptom onset, but not neuronal loss in Huntington's disease." *Neurobiol Dis* 174: 105884. https://doi.org/10.1016/j.nbd.2022.105884. https://www.ncbi.nlm.nih.gov/pubmed/36220612.

Lee, J. M., Y. Huang, M. Orth, T. Gillis, J. Siciliano, E. Hong, J. S. Mysore, D. Lucente, V. C. Wheeler, I. S. Seong, Z. L. McLean, J. A. Mills, B. McAllister, S. V. Lobanov, T. H. Massey, M. Ciosi, G. B. Landwehrmeyer, J. S. Paulsen, E. R. Dorsey, I. Shoulson, C. Sampaio, D. G. Monckton, S. Kwak, P. Holmans, L. Jones, M. E. MacDonald, J. D. Long, and J. F. Gusella. 2022. "Genetic modifiers of Huntington disease differentially influence motor and cognitive domains." *Am J Hum Genet* 109 (5): 885-899. https://doi.org/10.1016/j.ajhg.2022.03.004. https://www.ncbi.nlm.nih.gov/pubmed/35325614 https://www.cell.com/ajhg/pdf/S0002-9297(22)00100-8.pdf.

Li, K., M. Nakajima, I. Ibanez-Tallon, and N. Heintz. 2016. "A Cortical Circuit for Sexually Dimorphic Oxytocin-Dependent Anxiety Behaviors." *Cell* 167 (1): 60-72 e11. https://doi.org/10.1016/j.cell.2016.08.067. https://www.ncbi.nlm.nih.gov/pubmed/27641503 https://www.cell.com/cell/pdf/S0092-8674(16)31165-5.pdf.

MacDonald, M. E., S. Gines, J. F. Gusella, and V. C. Wheeler. 2003. "Huntington's disease." *Neuromolecular Med* 4 (1-2): 7-20. https://doi.org/10.1385/NMM:4:1-2:7. https://www.ncbi.nlm.nih.gov/pubmed/14528049 https://link.springer.com/content/pdf/10.1385/NMM:4:1-2:7.pdf https://link.springer.com/article/10.1385/NMM:4:1-2:7.

Macosko, E. Z., A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, and S. A. McCarroll. 2015. "Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets." *Cell* 161 (5): 1202-1214. https://doi.org/10.1016/j.cell.2015.05.002. https://www.ncbi.nlm.nih.gov/pubmed/26000488 https://www.cell.com/cell/pdf/S0092-8674(15)00549-8.pdf.

McColgan, P., S. Gregory, A. Razi, K. K. Seunarine, F. Gargouri, A. Durr, R. A. Roos, B. R. Leavitt, R. I. Scahill, C. A. Clark, S. J. Tabrizi, G. Rees, H. D. Investigators Track On, A. Coleman, J. Decolongon, M. Fan, T. Petkau, C. Jauffret, D. Justo, S. Lehericy, K. Nigaud, R. Valabregue, A. Choonderbeek, E. P. Hart, D. J. Hensman Moss, H. Crawford, E. Johnson, M. Papoutsi, C. Berna, R. Reilmann, N. Weber, J. Stout, I. Labuschagne, B. Landwehrmeyer, M. Orth, and H. Johnson. 2017. "White matter predicts functional connectivity in premanifest Huntington's disease." *Ann Clin Transl Neurol* 4 (2): 106-118. https://doi.org/10.1002/acn3.384. https://www.ncbi.nlm.nih.gov/pubmed/28168210.

Mehrabi, N. F., H. J. Waldvogel, L. J. Tippett, V. M. Hogg, B. J. Synek, and R. L. Faull. 2016. "Symptom heterogeneity in Huntington's disease correlates with neuronal degeneration in the cerebral cortex." *Neurobiol Dis* 96: 67-74. https://doi.org/10.1016/j.nbd.2016.08.015. https://www.ncbi.nlm.nih.gov/pubmed/27569581.

Monckton, D. G. 2021. "The Contribution of Somatic Expansion of the CAG Repeat to Symptomatic Development in Huntington's Disease: A Historical Perspective." *J Huntingtons Dis* 10 (1): 7-33. https://doi.org/10.3233/JHD-200429. https://www.ncbi.nlm.nih.gov/pubmed/33579863

https://content.iospress.com:443/download/journal-of-huntingtons-disease/jhd200429?id=journal-of-huntingtons-disease%2Fjhd200429.
Nana, A. L., E. H. Kim, D. C. Thu, D. E. Oorschot, L. J. Tippett, V. M. Hogg, B. J. Synek, R. Roxburgh, H. J. Waldvogel, and R. L. Faull. 2014. "Widespread heterogeneous neuronal loss across the cerebral cortex in Huntington's disease." *J Huntingtons Dis* 3 (1): 45-64. https://doi.org/10.3233/JHD-140092. https://www.ncbi.nlm.nih.gov/pubmed/25062764
https://content.iospress.com:443/articles/journal-of-huntingtons-disease/jhd140092.
Nopoulus. 2010. "Cerebral cortex structure in prodromal huntingtons disease."
Reid, C. B., and C. A. Walsh. 2002. "Evidence of common progenitors and patterns of dispersion in rat striatum and cerebral cortex." *J Neurosci* 22 (10): 4002-14. https://doi.org/10.1523/JNEUROSCI.22-10-04002.2002. https://www.ncbi.nlm.nih.gov/pubmed/12019320.
Reiner, A., and Y. P. Deng. 2018. "Disrupted striatal neuron inputs and outputs in Huntington's disease." *CNS Neurosci Ther* 24 (4): 250-280. https://doi.org/10.1111/cns.12844. https://www.ncbi.nlm.nih.gov/pubmed/29582587
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5875736/pdf/CNS-24-250.pdf
https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/cns.12844?download=true.
Ross, C. A., and S. J. Tabrizi. 2011. "Huntington's disease: from molecular pathogenesis to clinical treatment." *Lancet Neurol* 10 (1): 83-98. https://doi.org/10.1016/S1474-4422(10)70245-3. https://www.ncbi.nlm.nih.gov/pubmed/21163446
https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(10)70245-3/fulltext.
Sax, D. S., R. Powsner, A. Kim, S. Tilak, R. Bhatia, L. A. Cupples, and R. H. Myers. 1996. "Evidence of cortical metabolic dysfunction in early Huntington's disease by single-photon-emission computed tomography." *Mov Disord* 11 (6): 671-7. https://doi.org/10.1002/mds.870110612. https://www.ncbi.nlm.nih.gov/pubmed/8914093.
Shelbourne, P. F., C. Keller-McGandy, W. L. Bi, S. R. Yoon, L. Dubeau, N. J. Veitch, J. P. Vonsattel, N. S. Wexler, U. S-Venezuela Collaborative Research Group, N. Arnheim, and S. J. Augood. 2007. "Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain." *Hum Mol Genet* 16 (10): 1133-42. https://doi.org/10.1093/hmg/ddm054. https://www.ncbi.nlm.nih.gov/pubmed/17409200.
Swami, M., A. E. Hendricks, T. Gillis, T. Massood, J. Mysore, R. H. Myers, and V. C. Wheeler. 2009. "Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset." *Hum Mol Genet* 18 (16): 3039-47. https://doi.org/10.1093/hmg/ddp242. https://www.ncbi.nlm.nih.gov/pubmed/19465745.
Unschuld, P. G., S. E. Joel, X. Liu, M. Shanahan, R. L. Margolis, K. M. Biglan, S. S. Bassett, D. J. Schretlen, G. W. Redgrave, P. C. van Zijl, J. J. Pekar, and C. A. Ross. 2012. "Impaired cortico-striatal functional connectivity in prodromal Huntington's Disease." *Neurosci Lett* 514 (2): 204-9. https://doi.org/10.1016/j.neulet.2012.02.095. https://www.ncbi.nlm.nih.gov/pubmed/22425717.
Usrey, W. M., and S. M. Sherman. 2019. "Corticofugal circuits: Communication lines from the cortex to the rest of the brain." *J Comp Neurol* 527 (3): 640-650. https://doi.org/10.1002/cne.24423. https://www.ncbi.nlm.nih.gov/pubmed/29524229.
Veldman, M. B., and X. W. Yang. 2018. "Molecular insights into cortico-striatal miscommunications in Huntington's disease." *Curr Opin Neurobiol* 48: 79-89. https://doi.org/10.1016/j.conb.2017.10.019. https://www.ncbi.nlm.nih.gov/pubmed/29125980.
Vonsattel, J. P., C. Keller, and E. P. Cortes Ramirez. 2011. "Huntington's disease - neuropathology." *Handb Clin Neurol* 100: 83-100. https://doi.org/10.1016/B978-0-444-52014-2.00004-5. https://www.ncbi.nlm.nih.gov/pubmed/21496571.
Vonsattel, J. P., R. H. Myers, T. J. Stevens, R. J. Ferrante, E. D. Bird, and E. P. Richardson, Jr. 1985. "Neuropathological classification of Huntington's disease." *J Neuropathol Exp*

Neurol 44 (6): 559-77. https://doi.org/10.1097/00005072-198511000-00003. https://www.ncbi.nlm.nih.gov/pubmed/2932539.

Waldvogel, H. J., E. H. Kim, L. J. Tippett, J. P. Vonsattel, and R. L. Faull. 2015. "The Neuropathology of Huntington's Disease." *Curr Top Behav Neurosci* 22: 33-80. https://doi.org/10.1007/7854_2014_354. https://www.ncbi.nlm.nih.gov/pubmed/25300927 https://link.springer.com/chapter/10.1007/7854_2014_354.

Wang, F., J. Flanagan, N. Su, L. C. Wang, S. Bui, A. Nielson, X. Wu, H. T. Vo, X. J. Ma, and Y. Luo. 2012. "RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues." *J Mol Diagn* 14 (1): 22-9. https://doi.org/10.1016/j.jmoldx.2011.08.002. https://www.ncbi.nlm.nih.gov/pubmed/22166544.

Warner, J. H., J. D. Long, J. A. Mills, D. R. Langbehn, J. Ware, A. Mohan, and C. Sampaio. 2022. "Standardizing the CAP Score in Huntington's Disease by Predicting Age-at-Onset." *J Huntingtons Dis* 11 (2): 153-171. https://doi.org/10.3233/JHD-210475. https://www.ncbi.nlm.nih.gov/pubmed/35466943.

Weiss, A. R., W. A. Liguore, J. S. Domire, D. Button, and J. L. McBride. 2020. "Intra-striatal AAV2.retro administration leads to extensive retrograde transport in the rhesus macaque brain: implications for disease modeling and therapeutic development." *Sci Rep* 10 (1): 6970. https://doi.org/10.1038/s41598-020-63559-7. https://www.ncbi.nlm.nih.gov/pubmed/32332773 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7181773/pdf/41598_2020_Article_63559.pdf.

Wolf, R. C., G. Gron, F. Sambataro, N. Vasic, N. D. Wolf, P. A. Thomann, C. Saft, G. B. Landwehrmeyer, and M. Orth. 2011. "Magnetic resonance perfusion imaging of resting-state cerebral blood flow in preclinical Huntington's disease." *J Cereb Blood Flow Metab* 31 (9): 1908-18. https://doi.org/10.1038/jcbfm.2011.60. https://www.ncbi.nlm.nih.gov/pubmed/21559028.

Xu, X., E. I. Stoyanova, A. E. Lemiesz, J. Xing, D. C. Mash, and N. Heintz. 2018. "Species and cell-type properties of classically defined human and rodent neurons and glia." *Elife* 7. https://doi.org/10.7554/eLife.37551. https://www.ncbi.nlm.nih.gov/pubmed/30320555 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6188473/pdf/elife-37551.pdf.

SUPPLEMENTAL REFERENCES

Alcamo, E.A., Chirivella, L., Dautzenberg, M., Dobreva, G., Farinas, I., Grosschedl, R., and McConnell, S.K. (2008). Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. Neuron *57*, 364-377.

Andrew, S.E., Goldberg, Y.P., Kremer, B., Telenius, H., Theilmann, J., Adam, S., Starr, E., Squitieri, F., Lin, B., Kalchman, M.A.*, et al.* (1993). The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nat Genet *4*, 398-403.

Baker, A., Kalmbach, B., Morishima, M., Kim, J., Juavinett, A., Li, N., and Dembrow, N. (2018). Specialized Subpopulations of Deep-Layer Pyramidal Neurons in the Neocortex: Bridging Cellular Properties to Functional Consequences. The Journal of neuroscience : the official journal of the Society for Neuroscience *38*, 5441-5455.

Bakken, T.E., Jorstad, N.L., Hu, Q., Lake, B.B., Tian, W., Kalmbach, B.E., Crow, M., Hodge, R.D., Krienen, F.M., Sorensen, S.A.*, et al.* (2021). Comparative cellular analysis of motor cortex in human, marmoset and mouse. Nature *598*, 111-119.

Barbe, L., and Finkbeiner, S. (2022). Genetic and Epigenetic Interplay Define Disease Onset and Severity in Repeat Diseases. Front Aging Neurosci *14*, 750629.

Barry, J., Bui, M.T.N., Levine, M.S., and Cepeda, C. (2022). Synaptic pathology in Huntington's disease: Beyond the corticostriatal pathway. Neurobiol Dis *162*, 105574.

Bernard, A., Lubbers, L.S., Tanis, K.Q., Luo, R., Podtelezhnikov, A.A., Finney, E.M., McWhorter, M.M., Serikawa, K., Lemon, T., Morgan, R.*, et al.* (2012). Transcriptional architecture of the primate neocortex. Neuron *73*, 1083-1099.

Blumenstock, S., and Dudanova, I. (2020). Cortical and Striatal Circuits in Huntington's Disease. Frontiers in neuroscience *14*, 82.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213-1218.

Cajal, S. (1999). Texture of the Nervous System of Man and the Vertebrates. Springer Science & Business Media *Vol 1*.

Cepeda, C., Cummings, D.M., Andre, V.M., Holley, S.M., and Levine, M.S. (2010). Genetic mouse models of Huntington's disease: focus on electrophysiological mechanisms. ASN Neuro *2*, e00033.

Ciosi, M., Cumming, S.A., Chatzi, A., Larson, E., Tottey, W., Lomeikaite, V., Hamilton, G., Wheeler, V.C., Pinto, R.M., Kwak, S.*, et al.* (2021). Approaches to Sequence the HTT CAG Repeat Expansion and Quantify Repeat Length Variation. J Huntingtons Dis *10*, 53-74.

Ciosi, M., Maxwell, A., Cumming, S.A., Hensman Moss, D.J., Alshammari, A.M., Flower, M.D., Durr, A., Leavitt, B.R., Roos, R.A.C., team, T.-H.*, et al.* (2019). A genetic association study of glutamine-encoding DNA sequence structures, somatic CAG expansion, and DNA repair gene variants, with Huntington disease clinical outcomes. EBioMedicine *48*, 568-580.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B.*, et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods *14*, 959-962.

Dougherty, J.D., Schmidt, E.F., Nakajima, M., and Heintz, N. (2010). Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. Nucleic Acids Res *38*, 4218-4230.

Elam, J.S., Glasser, M.F., Harms, M.P., Sotiropoulos, S.N., Andersson, J.L.R., Burgess, G.C., Curtiss, S.W., Oostenveld, R., Larson-Prior, L.J., Schoffelen, J.M.*, et al.* (2021). The Human Connectome Project: A retrospective. NeuroImage *244*, 118543.

Erben, L., and Buonanno, A. (2019). Detection and Quantification of Multiple RNA Sequences Using Emerging Ultrasensitive Fluorescent In Situ Hybridization Techniques. Curr Protoc Neurosci *87*, e63.
Espinoza, F.A., Liu, J., Ciarochi, J., Turner, J.A., Vergara, V.M., Caprihan, A., Misiura, M., Johnson, H.J., Long, J.D., Bockholt, J.H.*, et al.* (2019). Dynamic functional network connectivity in Huntington's disease and its associations with motor and cognitive measures. Hum Brain Mapp *40*, 1955-1968.
Espinoza, F.A., Turner, J.A., Vergara, V.M., Miller, R.L., Mennigen, E., Liu, J., Misiura, M.B., Ciarochi, J., Johnson, H.J., Long, J.D.*, et al.* (2018). Whole-Brain Connectivity in a Large Study of Huntington's Disease Gene Mutation Carriers and Healthy Controls. Brain Connect *8*, 166-178.
Estrada-Sanchez, A.M., and Rebec, G.V. (2013). Role of cerebral cortex in the neuropathology of Huntington's disease. Front Neural Circuits *7*, 19.
Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using MACS. Nat Protoc *7*, 1728-1740.
Gargouri, F., Messe, A., Perlbarg, V., Valabregue, R., McColgan, P., Yahia-Cherif, L., Fernandez-Vidal, S., Ben Hamida, A., Benali, H., Tabrizi, S.*, et al.* (2016). Longitudinal changes in functional connectivity of cortico-basal ganglia networks in manifests and premanifest huntington's disease. Hum Brain Mapp *37*, 4112-4128.
Georgiou-Karistianis, N., Gray, M.A., Dominguez, D.J., Dymowski, A.R., Bohanna, I., Johnston, L.A., Churchyard, A., Chua, P., Stout, J.C., and Egan, G.F. (2013). Automated differentiation of pre-diagnosis Huntington's disease from healthy control individuals based on quadratic discriminant analysis of the basal ganglia: the IMAGE-HD study. Neurobiol Dis *51*, 82-92.
Goula, A.V., Pearson, C.E., Della Maria, J., Trottier, Y., Tomkinson, A.E., Wilson, D.M., 3rd, and Merienne, K. (2012). The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excision repair outcome at CAG/CTG repeats. Biochemistry *51*, 3919-3932.
Groh, A., Meyer, H.S., Schmidt, E.F., Heintz, N., Sakmann, B., and Krieger, P. (2010). Cell-type specific properties of pyramidal neurons in neocortex underlying a layout that is modifiable depending on the cortical area. Cerebral cortex *20*, 826-836.
Hasselbalch, S.G., Oberg, G., Sorensen, S.A., Andersen, A.R., Waldemar, G., Schmidt, J.F., Fenger, K., and Paulson, O.B. (1992). Reduced regional cerebral blood flow in Huntington's disease studied by SPECT. Journal of neurology, neurosurgery, and psychiatry *55*, 1018-1023.
Hattox, A.M., and Nelson, S.B. (2007). Layer V neurons in mouse cortex projecting to different targets have distinct physiological properties. J Neurophysiol *98*, 3330-3340.
Hedreen, J.C., Peyser, C.E., Folstein, S.E., and Ross, C.A. (1991). Neuronal loss in layers V and VI of cerebral cortex in Huntington's disease. Neurosci Lett *133*, 257-261.
Hernandez-Garcia, C., Popmintchev, T., Murnane, M.M., Kapteyn, H.C., Plaja, L., Becker, A., and Jaron-Becker, A. (2017). Isolated broadband attosecond pulse generation with near- and mid-infrared driver pulses via time-gated phase matching. Opt Express *25*, 11855-11866.
Hevner, R.F. (2007). Layer-specific markers as probes for neuron type identity in human neocortex and malformations of cortical development. J Neuropathol Exp Neurol *66*, 101-109.
Hickman, R.A., Faust, P.L., Marder, K., Yamamoto, A., and Vonsattel, J.P. (2022). The distribution and density of Huntingtin inclusions across the Huntington disease neocortex: regional correlations with Huntingtin repeat expansion independent of pathologic grade. Acta Neuropathol Commun *10*, 55.
Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O.*, et al.* (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature *573*, 61-68.
Hodge, R.D., Miller, J.A., Novotny, M., Kalmbach, B.E., Ting, J.T., Bakken, T.E., Aevermann, B.D., Barkan, E.R., Berkowitz-Cerasano, M.L., Cobbs, C.*, et al.* (2020). Transcriptomic evidence

that von Economo neurons are regionally specialized extratelencephalic-projecting excitatory neurons. Nat Commun *11*, 1172.
Hong, E.P., MacDonald, M.E., Wheeler, V.C., Jones, L., Holmans, P., Orth, M., Monckton, D.G., Long, J.D., Kwak, S., Gusella, J.F.*, et al.* (2021). Huntington's Disease Pathogenesis: Two Sequential Components. J Huntingtons Dis *10*, 35-51.
Jenkins, B.G., Rosas, H.D., Chen, Y.C., Makabe, T., Myers, R., MacDonald, M., Rosen, B.R., Beal, M.F., and Koroshetz, W.J. (1998). 1H NMR spectroscopy studies of Huntington's disease: correlations with CAG repeat numbers. Neurology *50*, 1357-1365.
Jernigan, T.L., Salmon, D.P., Butters, N., and Hesselink, J.R. (1991). Cerebral structure on MRI, Part II: Specific changes in Alzheimer's and Huntington's diseases. Biol Psychiatry *29*, 68-81.
Johnson, E.B., Ziegler, G., Penny, W., Rees, G., Tabrizi, S.J., Scahill, R.I., and Gregory, S. (2021). Dynamics of Cortical Degeneration Over a Decade in Huntington's Disease. Biol Psychiatry *89*, 807-816.
Kim, J., Moon, B.S., Lee, B.C., Lee, H.Y., Kim, H.J., Choo, H., Pae, A.N., Cho, Y.S., and Min, S.J. (2017). A Potential PET Radiotracer for the 5-HT2C Receptor: Synthesis and in Vivo Evaluation of 4-(3-[(18)F]fluorophenethoxy)pyrimidine. ACS Chem Neurosci *8*, 996-1003.
Knowles, R., Dehorter, N., and Ellender, T. (2021). From Progenitors to Progeny: Shaping Striatal Circuit Development and Function. The Journal of neuroscience : the official journal of the Society for Neuroscience *41*, 9483-9502.
Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.
Kuhl, D.E., Phelps, M.E., Markham, C.H., Metter, E.J., Riege, W.H., and Winter, J. (1982). Cerebral metabolism and atrophy in Huntington's disease determined by 18FDG and computed tomographic scan. Annals of neurology *12*, 425-434.
Kuwert, T., Lange, H.W., Langen, K.J., Herzog, H., Aulich, A., and Feinendegen, L.E. (1989). Cerebral glucose consumption measured by PET in patients with and without psychiatric symptoms of Huntington's disease. Psychiatry Res *29*, 361-362.
Kuwert, T., Lange, H.W., Langen, K.J., Herzog, H., Aulich, A., and Feinendegen, L.E. (1990). Cortical and subcortical glucose consumption measured by PET in patients with Huntington's disease. Brain : a journal of neurology *113 (Pt 5)*, 1405-1423.
Landwehrmeyer, G.B., McNeil, S.M., Dure, L.S.t., Ge, P., Aizawa, H., Huang, Q., Ambrose, C.M., Duyao, M.P., Bird, E.D., Bonilla, E.*, et al.* (1995). Huntington's disease gene: regional and cellular expression in brain of normal and affected individuals. Annals of neurology *37*, 218-230.
Layburn, F.E., Tan, A.Y.S., Mehrabi, N.F., Curtis, M.A., Tippett, L.J., Turner, C.P., Riguet, N., Aeschbach, L., Lashuel, H.A., Dragunow, M.*, et al.* (2022). N-terminal mutant huntingtin deposition correlates with CAG repeat length and symptom onset, but not neuronal loss in Huntington's disease. Neurobiol Dis *174*, 105884.
Lee, J.M., Huang, Y., Orth, M., Gillis, T., Siciliano, J., Hong, E., Mysore, J.S., Lucente, D., Wheeler, V.C., Seong, I.S.*, et al.* (2022). Genetic modifiers of Huntington disease differentially influence motor and cognitive domains. Am J Hum Genet *109*, 885-899.
Li, K., Nakajima, M., Ibanez-Tallon, I., and Heintz, N. (2016). A Cortical Circuit for Sexually Dimorphic Oxytocin-Dependent Anxiety Behaviors. Cell *167*, 60-72 e11.
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.
Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.
Lopez Castel, A., Tomkinson, A.E., and Pearson, C.E. (2009). CTG/CAG repeat instability is modulated by the levels of human DNA ligase I and its interaction with proliferating cell nuclear antigen: a distinction between replication and slipped-DNA repair. The Journal of biological chemistry *284*, 26631-26645.

Love, M.I., Hogenesch, J.B., and Irizarry, R.A. (2016). Modeling of RNA-seq fragment sequence bias reduces systematic errors in transcript abundance estimation. Nat Biotechnol *34*, 1287-1291.
Love, M.I., Soneson, C., and Patro, R. (2018). Swimming downstream: statistical analysis of differential transcript usage following Salmon quantification. F1000Res *7*, 952.
MacDonald, M.E., Gines, S., Gusella, J.F., and Wheeler, V.C. (2003). Huntington's disease. Neuromolecular Med *4*, 7-20.
Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M.*, et al.* (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202-1214.
Mason, A.G., Tome, S., Simard, J.P., Libby, R.T., Bammler, T.K., Beyer, R.P., Morton, A.J., Pearson, C.E., and La Spada, A.R. (2014). Expression levels of DNA replication and repair genes predict regional somatic repeat instability in the brain but are not altered by polyglutamine disease protein expression or age. Human molecular genetics *23*, 1606-1618.
McColgan, P., Gregory, S., Razi, A., Seunarine, K.K., Gargouri, F., Durr, A., Roos, R.A., Leavitt, B.R., Scahill, R.I., Clark, C.A.*, et al.* (2017). White matter predicts functional connectivity in premanifest Huntington's disease. Ann Clin Transl Neurol *4*, 106-118.
McColgan, P., Gregory, S., Seunarine, K.K., Razi, A., Papoutsi, M., Johnson, E., Durr, A., Roos, R.A.C., Leavitt, B.R., Holmans, P.*, et al.* (2018). Brain Regions Showing White Matter Loss in Huntington's Disease Are Enriched for Synaptic and Metabolic Genes. Biol Psychiatry *83*, 456-465.
Mehrabi, N.F., Waldvogel, H.J., Tippett, L.J., Hogg, V.M., Synek, B.J., and Faull, R.L. (2016). Symptom heterogeneity in Huntington's disease correlates with neuronal degeneration in the cerebral cortex. Neurobiol Dis *96*, 67-74.
Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell *151*, 1417-1430.
Monckton, D.G. (2021). The Contribution of Somatic Expansion of the CAG Repeat to Symptomatic Development in Huntington's Disease: A Historical Perspective. J Huntingtons Dis *10*, 7-33.
Nair, A., Razi, A., Gregory, S., Rutledge, R.B., Rees, G., and Tabrizi, S.J. (2022). Imbalanced basal ganglia connectivity is associated with motor deficits and apathy in Huntington's disease. Brain : a journal of neurology *145*, 991-1000.
Nakamori, M., Pearson, C.E., and Thornton, C.A. (2011). Bidirectional transcription stimulates expansion and contraction of expanded (CTG)*(CAG) repeats. Human molecular genetics *20*, 580-588.
Nana, A.L., Kim, E.H., Thu, D.C., Oorschot, D.E., Tippett, L.J., Hogg, V.M., Synek, B.J., Roxburgh, R., Waldvogel, H.J., and Faull, R.L. (2014). Widespread heterogeneous neuronal loss across the cerebral cortex in Huntington's disease. J Huntingtons Dis *3*, 45-64.
Nopoulus (2010). Cerebral cortex structure in prodromal huntingtons disease.
Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods *14*, 417-419.
Reid, C.B., and Walsh, C.A. (2002). Evidence of common progenitors and patterns of dispersion in rat striatum and cerebral cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *22*, 4002-4014.
Reiner, A., and Deng, Y.P. (2018). Disrupted striatal neuron inputs and outputs in Huntington's disease. CNS Neurosci Ther *24*, 250-280.
Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.
Ross, C.A., and Tabrizi, S.J. (2011). Huntington's disease: from molecular pathogenesis to clinical treatment. Lancet Neurol *10*, 83-98.

Rothenberg, E., Sepulveda, L.A., Skinner, S.O., Zeng, L., Selvin, P.R., and Golding, I. (2011). Single-virus tracking reveals a spatial receptor-dependent search mechanism. Biophysical journal *100*, 2875-2882.
Sax, D.S., Powsner, R., Kim, A., Tilak, S., Bhatia, R., Cupples, L.A., and Myers, R.H. (1996). Evidence of cortical metabolic dysfunction in early Huntington's disease by single-photon-emission computed tomography. Mov Disord *11*, 671-677.
Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.
Shelbourne, P.F., Keller-McGandy, C., Bi, W.L., Yoon, S.R., Dubeau, L., Veitch, N.J., Vonsattel, J.P., Wexler, N.S., Group, U.S.-V.C.R., Arnheim, N.*, et al.* (2007). Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. Human molecular genetics *16*, 1133-1142.
Slean, M.M., Panigrahi, G.B., Castel, A.L., Pearson, A.B., Tomkinson, A.E., and Pearson, C.E. (2016). Absence of MutSbeta leads to the formation of slipped-DNA for CTG/CAG contractions at primate replication forks. DNA Repair (Amst) *42*, 107-118.
Swami, M., Hendricks, A.E., Gillis, T., Massood, T., Mysore, J., Myers, R.H., and Wheeler, V.C. (2009). Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Human molecular genetics *18*, 3039-3047.
Tome, S., Manley, K., Simard, J.P., Clark, G.W., Slean, M.M., Swami, M., Shelbourne, P.F., Tillier, E.R., Monckton, D.G., Messer, A.*, et al.* (2013). MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. PLoS Genet *9*, e1003280.
Tome, S., Panigrahi, G.B., Lopez Castel, A., Foiry, L., Melton, D.W., Gourdon, G., and Pearson, C.E. (2011). Maternal germline-specific effect of DNA ligase I on CTG/CAG instability. Human molecular genetics *20*, 2131-2143.
Unschuld, P.G., Joel, S.E., Liu, X., Shanahan, M., Margolis, R.L., Biglan, K.M., Bassett, S.S., Schretlen, D.J., Redgrave, G.W., van Zijl, P.C.*, et al.* (2012). Impaired cortico-striatal functional connectivity in prodromal Huntington's Disease. Neurosci Lett *514*, 204-209.
Usrey, W.M., and Sherman, S.M. (2019). Corticofugal circuits: Communication lines from the cortex to the rest of the brain. J Comp Neurol *527*, 640-650.
Veldman, M.B., and Yang, X.W. (2018). Molecular insights into cortico-striatal miscommunications in Huntington's disease. Curr Opin Neurobiol *48*, 79-89.
Vonsattel, J.P., Keller, C., and Cortes Ramirez, E.P. (2011). Huntington's disease - neuropathology. Handb Clin Neurol *100*, 83-100.
Vonsattel, J.P., Myers, R.H., Stevens, T.J., Ferrante, R.J., Bird, E.D., and Richardson, E.P., Jr. (1985). Neuropathological classification of Huntington's disease. J Neuropathol Exp Neurol *44*, 559-577.
Waldvogel, H.J., Kim, E.H., Tippett, L.J., Vonsattel, J.P., and Faull, R.L. (2015). The Neuropathology of Huntington's Disease. Curr Top Behav Neurosci *22*, 33-80.
Wang, F., Flanagan, J., Su, N., Wang, L.C., Bui, S., Nielson, A., Wu, X., Vo, H.T., Ma, X.J., and Luo, Y. (2012). RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagn *14*, 22-29.
Wang, Q., Li, M., Wu, T., Zhan, L., Li, L., Chen, M., Xie, W., Xie, Z., Hu, E., Xu, S.*, et al.* (2022). Exploring Epigenomic Datasets by ChIPseeker. Curr Protoc *2*, e585.
Warner, J.H., Long, J.D., Mills, J.A., Langbehn, D.R., Ware, J., Mohan, A., and Sampaio, C. (2022). Standardizing the CAP Score in Huntington's Disease by Predicting Age-at-Onset. J Huntingtons Dis *11*, 153-171.
Weiss, A.R., Liguore, W.A., Domire, J.S., Button, D., and McBride, J.L. (2020). Intra-striatal AAV2.retro administration leads to extensive retrograde transport in the rhesus macaque brain: implications for disease modeling and therapeutic development. Sci Rep *10*, 6970.
Wolf, R.C., Gron, G., Sambataro, F., Vasic, N., Wolf, N.D., Thomann, P.A., Saft, C., Landwehrmeyer, G.B., and Orth, M. (2011). Magnetic resonance perfusion imaging of resting-

state cerebral blood flow in preclinical Huntington's disease. J Cereb Blood Flow Metab *31*, 1908-1918.
Xu, X., Stoyanova, E.I., Lemiesz, A.E., Xing, J., Mash, D.C., and Heintz, N. (2018). Species and cell-type properties of classically defined human and rodent neurons and glia. Elife *7*.
Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284-287.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.