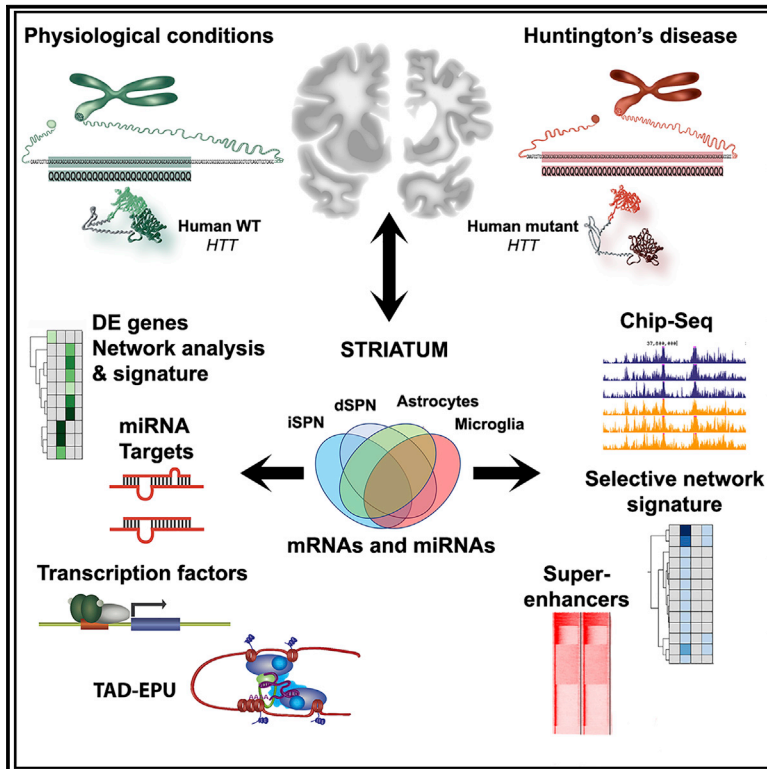


# Cell Reports

## Cell-Type-Specific Gene Expression Profiling in Adult Mouse Brain Reveals Normal and Disease-State Signatures

### Graphical Abstract



### Authors

Nicolas Merienne, Cécile Meunier, Anne Schneider, ..., Christian Neri, Karine Merienne, Nicole Déglon

### Correspondence

nicole.deglon@chuv.ch

### In Brief

The contributions of neuronal and non-neuronal populations in cerebral functions are increasingly studied, but a robust and stress-limited method for cell-type-specific isolation is lacking. Merienne et al. optimize a laser-capture microdissection approach and characterize transcriptomic and epigenetic mechanisms in normal or pathological mouse and human Huntington's disease (HD) brains.

### Highlights

- Optimized LCM method for mouse cell-type-specific dissection
- Transcriptomic and network signatures of four adult striatal subpopulations
- Role of transcription factors and DNA conformation in cell-type gene enrichment
- Cell-type-specific contribution of epigenetic mechanisms in Huntington' disease



# Cell-Type-Specific Gene Expression Profiling in Adult Mouse Brain Reveals Normal and Disease-State Signatures

Nicolas Merienne,<sup>1,2</sup> Cécile Meunier,<sup>3</sup> Anne Schneider,<sup>4</sup> Jonathan Seguin,<sup>4</sup> Satish S. Nair,<sup>5</sup> Anne B. Rocher,<sup>6</sup> Stéphanie Le Gras,<sup>7</sup> Céline Keime,<sup>7</sup> Richard Faull,<sup>8</sup> Luc Pellerin,<sup>3,9</sup> Jean-Yves Chatton,<sup>6,10</sup> Christian Neri,<sup>5</sup> Karine Merienne,<sup>4</sup> and Nicole Déglon<sup>1,2,11,\*</sup>

<sup>1</sup>Department of Clinical Neurosciences, Laboratory of Neurotherapies and Neuromodulation (LNTM), Lausanne University Hospital, 1011 Lausanne, Switzerland

<sup>2</sup>Neuroscience Research Center, LNTM, Lausanne University Hospital, 1011 Lausanne, Switzerland

<sup>3</sup>Department of Physiology, Laboratory of Neuroenergetics, University of Lausanne, 1005 Lausanne, Switzerland

<sup>4</sup>University of Strasbourg, CNRS, UMR 7364, Laboratory of Cognitive and Adaptive Neuroscience, 67000 Strasbourg, France

<sup>5</sup>Sorbonne Université, Centre National de la Recherche Scientifique, Research Unit Biology of Adaptation and Aging, Team Compensation in Neurodegenerative Diseases and Aging, 75252 Paris, France

<sup>6</sup>Department of Fundamental Neurosciences, University of Lausanne, 1005 Lausanne, Switzerland

<sup>7</sup>University of Strasbourg, CNRS, INSERM, UMR 7104, Microarray and Sequencing Platform, Institute of Genetic and Molecular and Cellular Biology, 67404 Illkirch, France

<sup>8</sup>Centre for Brain Research, Faculty of Medical and Health Science, The University of Auckland, Auckland 1023, New Zealand

<sup>9</sup>Centre de Résonance Magnétique des Systèmes Biologiques UMR 5536, CNRS-Université de Bordeaux, 146 rue Léo Saignat, Bordeaux, France

<sup>10</sup>Cellular Imaging Facility, University of Lausanne, 1005 Lausanne, Switzerland

<sup>11</sup>Lead Contact

\*Correspondence: [nicole.deglon@chuv.ch](mailto:nicole.deglon@chuv.ch)  
<https://doi.org/10.1016/j.celrep.2019.02.003>

## SUMMARY

The role of brain cell-type-specific functions and profiles in pathological and non-pathological contexts is still poorly defined. Such cell-type-specific gene expression profiles in solid, adult tissues would benefit from approaches that avoid cellular stress during isolation. Here, we developed such an approach and identified highly selective transcriptomic signatures in adult mouse striatal direct and indirect spiny projection neurons, astrocytes, and microglia. Integrating transcriptomic and epigenetic data, we obtained a comprehensive model for cell-type-specific regulation of gene expression in the mouse striatum. A cross-analysis with transcriptomic and epigenomic data generated from mouse and human Huntington's disease (HD) brains shows that opposite epigenetic mechanisms govern the transcriptional regulation of striatal neurons and glial cells and may contribute to pathogenic and compensatory mechanisms. Overall, these data validate this less stressful method for the investigation of cellular specificity in the adult mouse brain and demonstrate the potential of integrative studies using multiple databases.

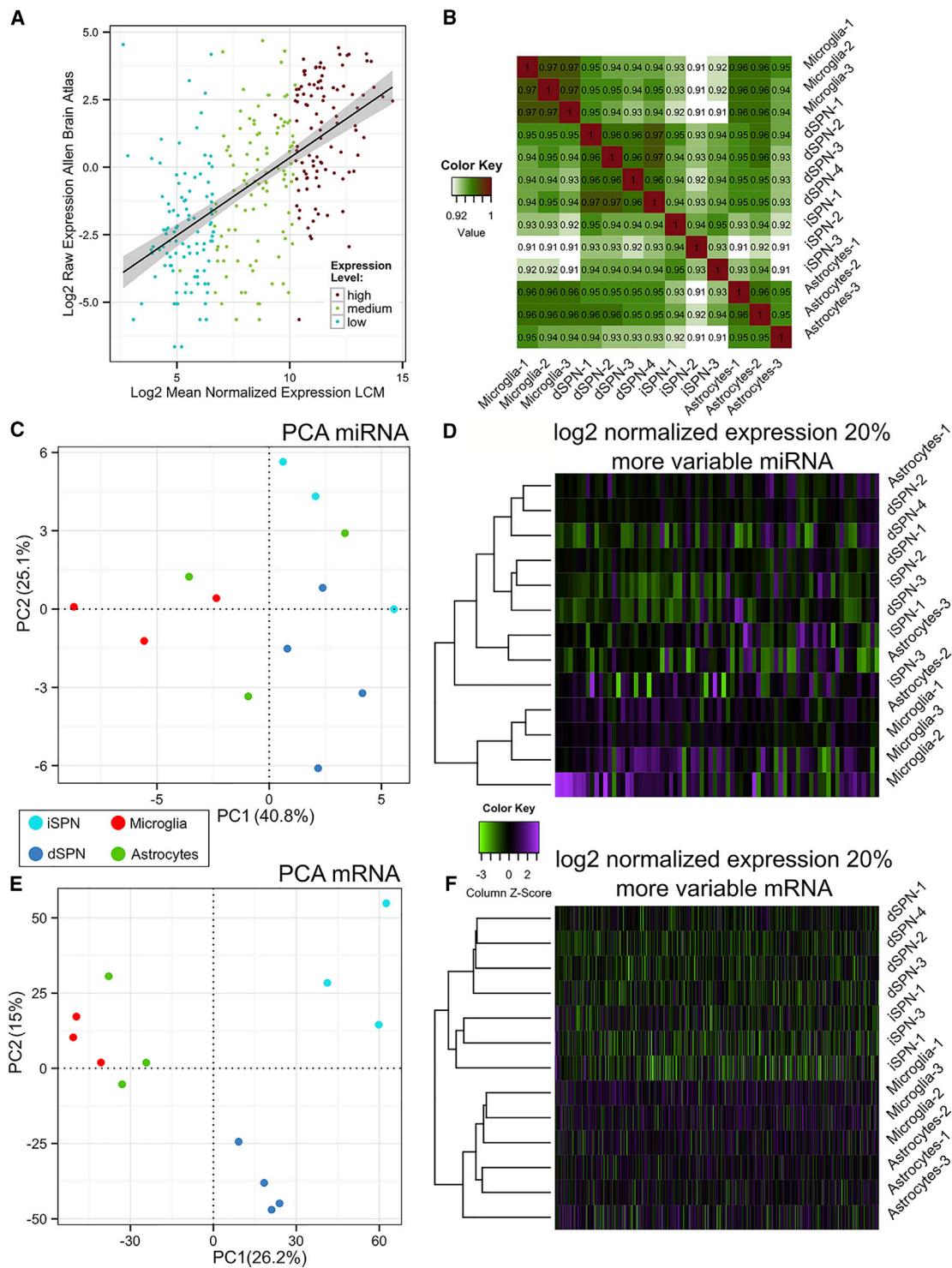
## INTRODUCTION

Neurons have long been considered to be the principal actors of thoughts and actions in the CNS. Technological developments

have highlighted the vast cellular diversity of the brain and contributions of neuronal and non-neuronal cell populations in brain functions. The striatum, in particular, is an essential relay of motor, cognitive, and limbic processes, implying highly specific roles for distinct cell subpopulations (Silberberg and Bolam, 2015). Precise dissection of cellular contributions to striatal loops is central to the comprehension of this region and its implication in normal and pathological contexts. For example, specific epigenetic and transcriptional alterations have been described in the striatum of mice and patients with Huntington's disease (HD), a genetic neurodegenerative disease affecting primarily the striatum (Achour et al., 2015; Hodges et al., 2006; Langfelder et al., 2016). However, it is still unclear whether HD mutation similarly or differently affects the epigenetic and transcriptional regulation of neurons and glial cells.

Cell-type-specific characterization faces a major challenge in the isolation of populations from both developing and adult animals with techniques compatible with high-throughput analysis. Attempts to overcome these limitations have included the purification of cell-type-specific cells from mouse striatum (Doyle et al., 2008; Heiman et al., 2008; Lobo et al., 2006) or other brain structures (Cahoy et al., 2008; Macosko et al., 2015; Molyneaux et al., 2015; Zhang et al., 2014). These methods rely either on the development of new genetically modified animals to isolate ribosome-bound RNA and determine the translational profile of CNS cell types or *ex vivo* cellular immunolabeling and sorting. However, large-scale cellular staining compatible with multiple high-throughput -omics applications (e.g., transcriptome and proteome profiling) in adult models is still associated with artificial cellular stress encountered during purification (Handley et al., 2015).





**Figure 1. Validation of the New Method for Transcriptome Profiling**

(A) The mean normalized expression levels of genes with low, medium, and high expression across samples (x axis) were compared with expression levels from the Allen Brain Atlas (y axis). Linear regression revealed a high correlation coefficient ( $r = 0.58$ ) between the two studies.

(B) Heatmap of Pearson correlations between each sample, showing stronger correlation coefficients between samples of the same group than intergroup correlations.

(C–E) Principal-component analysis (PCA) on all detected miRNAs (C) and mRNAs (E) showing a clear distinction of neuronal and glial samples for mRNA and less separation for microRNA (miRNA) based on the first principal components (PC1). Grouping of the messenger RNA dots suggests separation by cell type.

(legend continued on next page)

Here, we report a flexible method for physiological cell-type-specific profiling of adult cells, taking advantage of existing transgenic animals and widely used technologies. The method combines the advantages of fluorescent transgenic mouse models and sectioning under physiological conditions to avoid fluorescent signal loss and ensure high-quality RNA, together with laser capture microdissection (LCM) (Cheng et al., 2013). We performed transcriptome profiling of astrocytes, microglia, and spiny projection neurons (SPNs) of the striatonigral (direct pathway [dSPN]) and striatopallidal loops (indirect pathway [iSPN]). Neurons from these two pathways are morphologically identical but have distinct roles in basal ganglia networks, highlighting the need to deeply characterize their common and specific features.

We used RNA sequencing (RNA-seq) and microarray hybridization to identify differential mRNA profiles and analyze the microRNA (miRNA) pool of striatal cell populations, respectively. The analyses of differentially expressed (DE) genes highlight gene expression signatures specific to each population. The integration of these data with miRNA and epigenetic profiles emphasizes complex mechanisms of transcriptional regulation in the striatum involving the activity of selective transcription factors (TFs) and a role for genome tridimensional folding. Finally, analysis of databases on HD underlines differential transcriptomic and epigenetic changes in striatal neurons and glial cells. Specifically, we show that neuronal identity genes, and not glial genes, are highly susceptible to downregulation and histone deacetylation, thus providing new insights in HD pathogenesis.

## RESULTS

### LCM Protocol for *In Vivo* Brain Cell Population -Omics

We analyzed the transcriptomic profiles of the main cellular populations of the striatum by LCM using transgenic mice expressing EGFP in the iSPN (Drd2-eGFP) (Gong et al., 2003), astrocytes (GLT1-eGFP) (Regan et al., 2007), and microglia (Cx3cr1-eGFP) (Jung et al., 2000) and Tomato fluorescent protein in the dSPN (Drd1-Tomato) (Ade et al., 2011) (Figure S1A). As previously reported, freezing fresh, unfixed brain samples induces a strong loss of fluorescent protein due to the diffusion of the signal, whereas partial brain fixation ensures the preservation of fluorescence but is incompatible with high-quality RNA recovery (data not shown) (Rossner et al., 2006). We overcame these limitations by developing a LCM tool for *in vivo* brain cell population -omics that preserves cytoplasmic fluorescence while recovering high-quality RNA (Figures S1B and S1C). The protocol is based on vibratome slicing in cold, oxygenated artificial cerebrospinal fluid and a modified dehydration procedure.

We extracted total RNA from 10 to 12 × 10<sup>4</sup> cells (three or four animals per group) for each cell population (Table S1) and performed microarray analysis of miRNA expression and mRNA

sequencing (Figures S1D and S1E). We detected 16,058 mRNAs, including hundreds of long-intergenic non-coding RNAs (lincRNA) and 370 miRNAs. We verified the quality of our data by randomly selecting 100 genes of the Allen Brain Atlas with low, medium, or high striatal expression levels. Linear regression revealed a significant correlation ( $r = 0.58$ ,  $p < 2.2e-16$ ), showing that our data are consistent with previously published gene expression profiles (Figure 1A).

The reproducibility of the data was demonstrated by the strong correlation between samples from the different groups, as well as between all samples, which is coherent with previous studies suggesting that a small proportion of transcripts may be cell-type specific (Heinz et al., 2010; Zhang et al., 2014) (Figure 1B). In addition, principal-component analysis and hierarchical clustering demonstrated good discrimination between different cell populations across samples according to mRNA expression profiles, (Figures 1E and 1F), but less so for miRNA (Figures 1C and 1D).

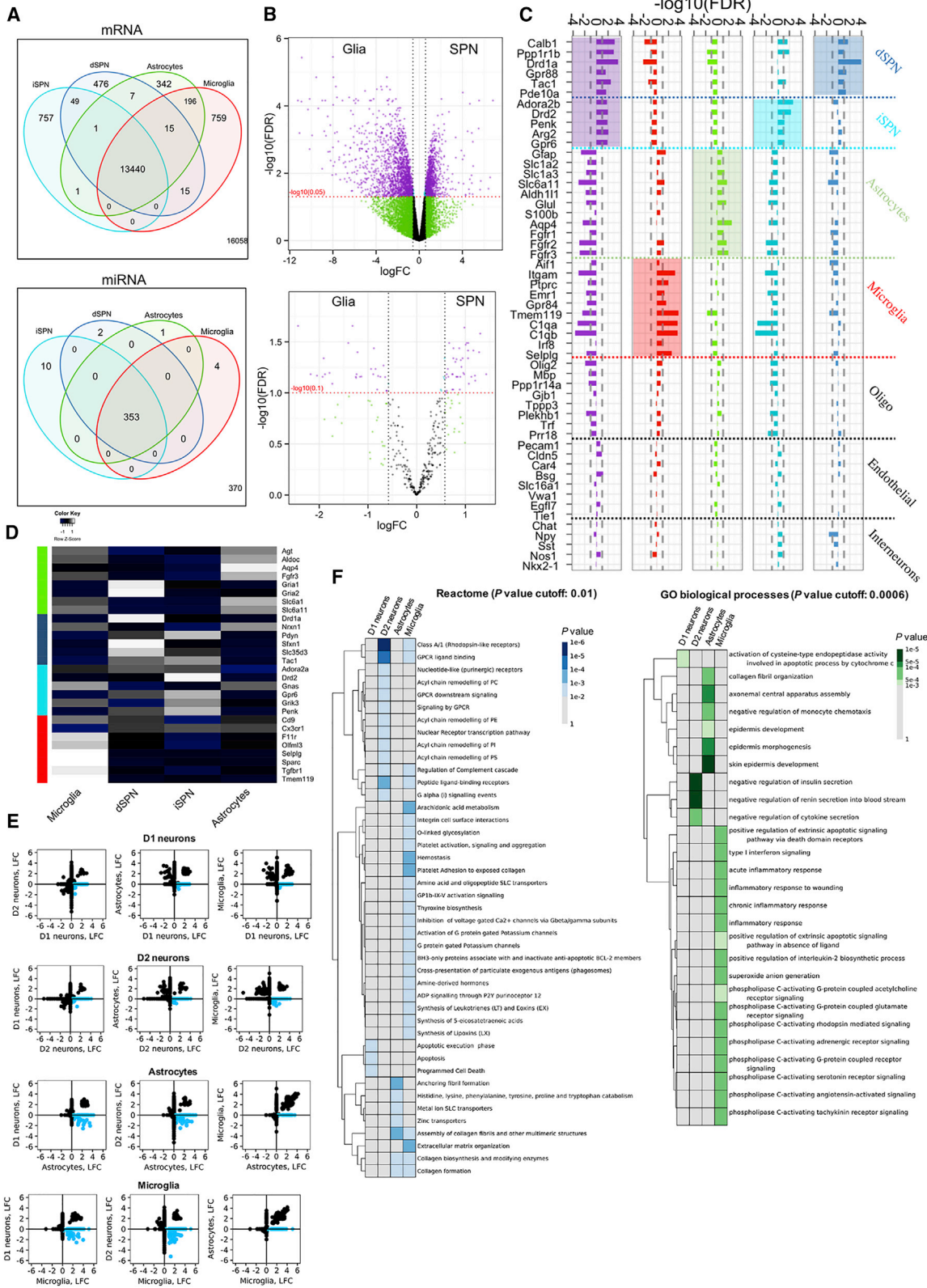
### Analysis of Cell-Type-Specific Molecular Markers and Cellular Functions

We conducted contrast analyses to identify DE mRNA and miRNA (Table S2) and obtained 757, 476, 342, and 759 mRNAs significantly enriched in the iSPN, dSPN, astrocytes, and microglial cells, respectively (Figure 2A; false discovery rate [FDR] threshold = 0.1). We also identified between 1 and 10 DE miRNAs (Figure 2A). Given that dSPN and iSPN share many functions, we compared SPNs (dSPN and iSPN samples) and glial cells (astrocytes and microglia; Figure 2B). We obtained 961 and 47 SPN-enriched mRNAs (FDR threshold = 0.05) and miRNAs (FDR threshold = 0.1) and 1,550 and 28 glia-enriched mRNAs and miRNAs, respectively. The selectivity and robustness of our data were assessed by selecting known markers of targeted and non-targeted cells (Heiman et al., 2008; Zhang et al., 2014). As expected, SPNs were enriched for *Ppp1r1b* (DARPP-32), *Adora2b*, and *Calb1* (Calbindin) (Figure 2C). More specifically, we found known markers of the striatopallidal pathway in iSPN cells (*Drd2*, *Penk*, and *Gpr6*; Figure 2C) and striatonigral markers in dSPNs (*Drd1* and *Tac1*; Figure 2C). We confirmed substantial enrichment of *Aqp4* (Aquaporin 4), *Glul* (GS), and *Slc1a3* (GLAST) in astrocytes (Figure 2C) and *Itgam* (CD11b), *Emr1* (F4/80), and *CD68* in microglia (Figure 2C). We detected marked enrichment of *Gfap*, *Slc1a2* (GLT1), and *Aldh1l1* in astrocytes. Finally, we did not observe notable enrichment of genes belonging to the non-targeted cell populations (Figure 2C). We further validated some of these DE genes by qRT-PCR and compared our DE gene (DEG) with the markers found in a previous single-cell RNA-seq study (Gokce et al., 2016; Figures 2D and S2A). A selected set of genes involved in cellular stress responses showed a lower level of expression compared to data obtained by fluorescence-activated cell sorting (FACS) (Zhang et al., 2014) (Figure S2B). Finally, our data demonstrate

(D–F) Euclidean unsupervised hierarchical clustering of the log<sub>2</sub> normalized expression levels from the 20% most variable miRNAs (D) and mRNAs (F). Color intensities are proportional to z-scaled expression levels. Strong clustering by cell population is observed for mRNA, confirming the sample distribution determined by PCA.

See also Figure S1 and Tables S1 and S2.





(legend on next page)

the absence of strong cross-contamination between cell populations due to the dissection procedure (Figure 2C).

We further validated the capacity to discriminate between striatal cell types with a network analysis. A spectral decomposition of gene expression signals was performed against the probabilistic functional network MouseNet v2, which generates unbiased and precise predictions of the biological significance of DE genes while reducing the risk of false positives and negatives (Lejeune et al., 2012; Tourette et al., 2014). Such analysis systematically retained gene nodes (Table S3, sheet 1), relative to the raw data (Figure S2C), that are highly expressed in a given cell type and show no expression or poor expression in one of the other cell types (blue genes) (Figure 2E; Table S3, sheet 2). Biological enrichment analysis of the union of the blue-gene lists for each cell type using Gene Ontology biological processes and reactome data revealed cell-type-specific patterns of biological annotations (Figure 2F). These analyses notably suggested that blue genes for dSPNs are associated with a greater susceptibility to cell death than those of iSPNs, consistent with the current view of neuronal cell vulnerability in the striatum and the higher vulnerability of D1 neurons to degeneration (Francelle et al., 2014; Han et al., 2010) (Figure 2F). More precisely, high-confidence Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database annotations of the highly enriched (log fold-change [logFC] >1.5 versus reference) genes in the blue-gene network signature of dSPNs, which includes Trim12a, Usp50 B3gnt3, Hist1h1d, Lsm7, Mrc1, Map3k7cl, and 29 other genes, representing ~35% of the 105 genes retained in this signature (including Drd1; LCF = 1.08), show these genes are implicated in mRNA splicing, extrinsic apoptosis signaling, mRNA splicing, extracellular matrix, and 1-phosphatidylinositol-3-kinase activity (Figure S7A). We detected no overlap with previously reported and larger signatures of human striatal neurons (Kelley et al., 2018) (Figure S7B). The same appears to apply to the comparison with previously reported and larger RNA-seq signatures of mouse D1 striatal neurons, as inferred from the data made available in this study (Gokce et al., 2016). These comparisons suggest that network analysis is able to extract precise functional signatures not necessarily put forth by traditional analysis. In astrocytes, blue genes were associated with cell-to-cell interactions and cell-to-extracellular matrix

processes relative to cellular differentiation and homeostasis, whereas in microglia, they were associated with inflammation and immune function, as expected (Figure 2E).

Overall, these data suggest that our method allows the purification of striatal cell types to high homogeneity, as shown by the detection of discriminative molecular signatures in each population.

### Cell-Type-Enriched miRNA Expression Levels Do Not Predict Expression Levels of Target mRNA

miRNAs are small non-coding RNAs that mediate RNA silencing through post-transcriptional regulation of gene expression (Bar-tel, 2009). Here, we recapitulated the relative expression rates of four strongly expressed miRNAs: mmu-miR-124-3p, mmu-miR-9-5p, mmu-miR-9-3p, and mmu-miR-29b-3p (Figure S3A) (He et al., 2012). We also found that SPN-enriched miRNA from chromosomes 12 and 7 were overrepresented, whereas glia-enriched miRNAs were mostly transcribed from chromosome 2 (Figures S3B and S3C). Analysis of miRNA clusters showed that the miR-379/miR-410 cluster, located on chromosome 12, was significantly enriched in SPN miRNA, as previously reported for cultured neurons, whereas no cluster reached the significance threshold in glia-enriched miRNA (Figures 3A, 3B, and S3D; Table S4) (Jovičić et al., 2013; Rago et al., 2014). Finally, we observed the expected enrichment of mmu-miR-124 in SPN relative to glia (He et al., 2012).

A single miRNA regulates multiple distinct mRNAs, whereas a single mRNA may be regulated by numerous independent miRNAs, which can be grouped by miRNA families based on their common target genes. We assessed whether the enrichment of mRNA in SPN and glia may be linked to the expression of corresponding miRNA families. According to this hypothesis, strong enrichment of one miRNA in one cell type should lead to limited expression of the corresponding target transcripts in these cells. Therefore, these transcripts should have a greater propensity to be DE in the other cell population (Figure 3C). Information in miRNA-target databases suggested that SPN-enriched miRNAs (47) could regulate 4,299 predicted target mRNAs (miRTs; Figure 3D), whereas glia-enriched miRNAs (28) could regulate 1,600 transcripts (Figure 3E; Table S4). The analysis shows that most miRNA-target genes were non-DE and that when the

### Figure 2. Differential Analysis Identifies Specific Markers of Striatal Cell Types

(A) Venn contrast diagrams for mRNA (up) and miRNA (down) of all genes for each cell type showing the number of cell-type-enriched genes (FDR threshold = 0.1) at the extremity of each area. The data show hundreds of DE mRNAs but a lower proportion of cell-type-enriched miRNAs.

(B) Volcano plots showing all detected mRNAs (up) and miRNAs (down) in the SPN-glia contrast. The horizontal red-dotted line represents the FDR threshold (0.05 for mRNA and 0.1 for miRNA) on a  $-\log_{10}$  scale, whereas the vertical black-dotted lines show the  $\log_{2}FC = 0.58$  and  $\log_{2}FC = -0.58$ . Violet dots represent transcripts with a significant FDR and an absolute  $\log_{2}FC \geq 0.58$ , turquoise dots represent transcripts with a significant FDR but a lower  $\log_{2}FC$ , and green dots represent non-significant transcripts with an absolute  $\log_{2}FC \geq 0.58$ .

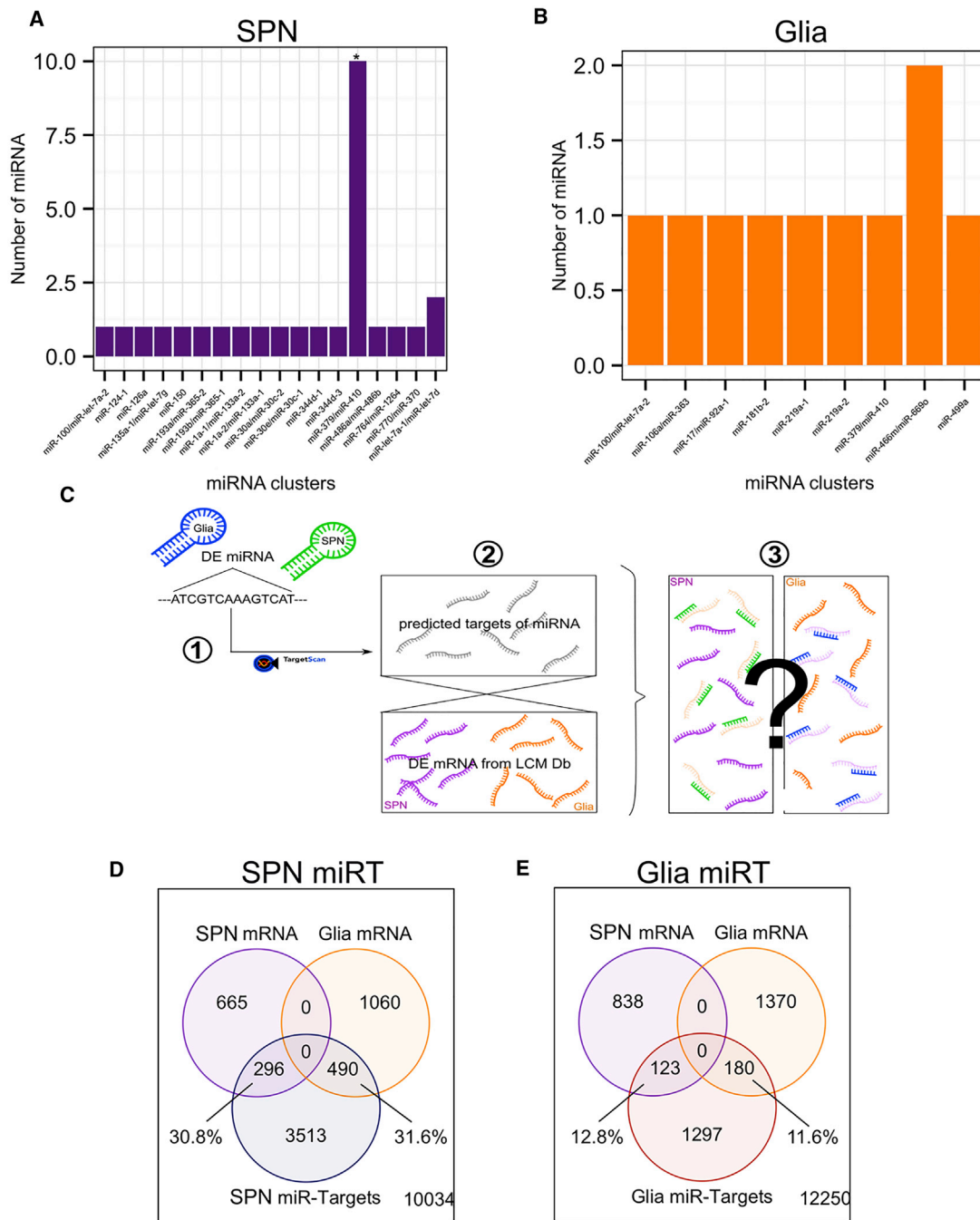
(C) Barplot showing the enrichment score for selected known cellular markers of dSPN (dark blue), iSPN (turquoise), astrocytes (green), microglia (red), and global SPN (purple). Markers for non-dissected cell types were also selected (black). The enrichment score is plotted for each gene in contrast (see STAR Methods).

(D) Comparison of our DEG with markers cell-type specific markers identified in Gokce et al. (2016). The heatmap displays the normalized average expression level in the four cell population of the key cell-type-specific markers identified by single-cell RNA-seq. Color code indicates a significant overlap of markers identified in the two studies.

(E) Cell-type specificity of gene-expression network signatures in the mouse striatum. Blue gene nodes represent those genes that are specifically enriched in the cell type considered and those that are under-expressed in the cell type to which the comparison was made.

(F) Enrichment of blue gene lists in Reactome and GO biological processes annotations. The heatmaps show biological content for the union of the blue gene groups across cell types for each cell type considered.

See also Figures S2 and S7 and Table S3.



**Figure 3. Relationship between miRNA Enrichment and Target mRNA Cellular Specificity**

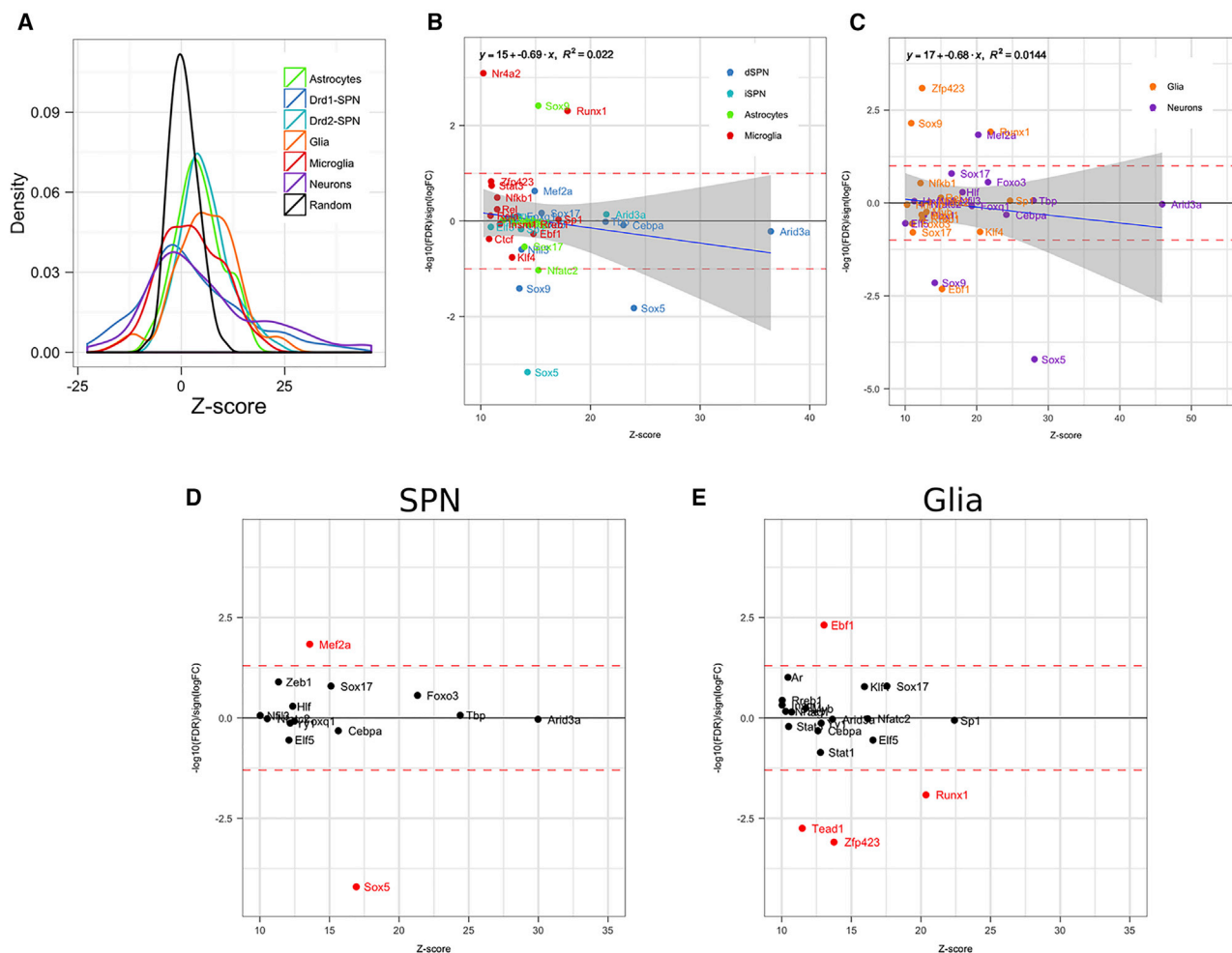
(A and B) Histogram showing the number of enriched miRNAs in each cluster of chromosome 12 (A: SPN-enriched miRNA) and chromosome 2 (B: glia-enriched miRNA). A hypergeometric analysis was performed to determine the significance of the enrichment. \*FDR < 0.1.

(C) Schematic representation of the rationale behind the analysis of mRNA regulation by cell-type-enriched miRNA. For each cell-type-enriched miRNA in SPN (green) and glia (blue), the corresponding target sequences of the transcriptome were recovered using TargetScan. Next, list of predicted mRNA targets of miRNA was crossed with lists of DEGs in SPNs (violet) or glia (orange). The hypothesis was that a miRNA highly expressed in one cell population will downregulate mRNA that, as a consequence, will be detected as enriched in the other cell population.

(D) Venn diagrams showing the overlap between miRNA targets of SPN-enriched miRNA (SPN miRT, dark blue) with SPN-enriched mRNA (violet) and glia-enriched mRNA (orange). The proportion of overlap is equal.

(E) Venn diagrams showing the overlap between miRNA targets from glia-enriched miRNA (glia miRT, dark red) with SPN-enriched mRNA (violet) and glia-enriched mRNA (orange). The proportion of overlap is equal.

See also [Figure S3](#) and [Table S4](#).



**Figure 4. Different Sets of Transcription Factors Regulate Cell-Type-Enriched Genes**

(A) Density plot of Z scores obtained after TFBS analysis using the Opposum 3.0 TFBS prediction tool. Colored lines represent the distribution of Z scores for TFs identified from a list of genes enriched in distinct cell populations, whereas the black line represents the Z score distribution obtained from randomly generated gene lists.

(B and C) Regression plots of TFs identified with an enrichment of TFBSs in cell-type-enriched genes (B) or in the neuron versus glia contrast (C). The x axis represents the Z score of the identified TF. The y axis ( $-\log_{10}(\text{FDR})\text{sgn}(\Delta)$ ) corresponds to the enrichment score for each gene in each contrast. For the cell-type-specific analysis (B), the gene is significantly enriched in the cell population if its enrichment score is  $\geq 1$ , whereas it is significantly less expressed if the score is  $\leq -1$ . Scores between  $-1$  and  $1$  are non-significant. For the neuron versus glia contrast (C), the values of the enrichment score for glia-enriched genes were inverted to facilitate the interpretation. Thus, genes that are enriched in glial cells have a positive enrichment score, whereas those that are non-enriched in glial cells have a negative enrichment score.

(D and E) Plot representing the TFs with more TFBSs close to SOX9 in genes enriched in neurons (D) or glial cells (E). The horizontal red dotted line corresponds to an FDR threshold of 0.05.

See also [Data S1](#).

miRNA-targets were DE, they were present in an equal proportion in neurons and glial cells (Figures 3D and 3E). These data suggest that miRNA regulation has only a limited role in determining global mRNA profiles and that other mechanisms might be involved in the regulation of cell-type-specific gene expression.

### The Activity of Selective TFs Is Predicted to Modulate the Cellular Specificity of Gene Expression

We next analyzed the enrichment of specific transcription-factor-binding sites (TFBSs) in the cell-type-enriched genes based

on regulatory motif over-representation across genes or chromatin immunoprecipitation sequencing (ChIP-seq) datasets (Zhou et al., 2017). We identified 23 to 35 TFs per cell population, each showing significant enrichment of TFBSs among cell-type-enriched genes (Figure 4A; Data S1). Some identified TFs are potential regulators of a large number of cell-type-enriched genes, such as RUNX1, which potentially binds to 570 genes among the 759 enriched in microglia. We then evaluated the relationship between the cell-type enrichment of TFs and the enrichment of TFBSs in DE genes in the same population. There was no



significant correlation between TF expression and enrichment of the corresponding TFBSs in the four cellular populations (Figures 4B and 4C). In contrast, we identified Sox9, Sox5, and Mef2a as potential regulators of SPN genes and Runx1, Zfp423, and Ebf1 as potential regulators of glial genes (Figure 4C).

We next evaluated whether DE genes could be regulated synchronously by different TFs sharing close proximity to transcription start sites. We focused our analysis on Sox9. In SPNs, this TF was significantly associated with Sox5 (non-enriched) and Mef2a (enriched), two TFs already shown to be major potential regulators in SPN (Figure 4D). In contrast, in glia, Sox9 is significantly associated with Runx1, Zfp423, and Ebf1, which are strong regulators of glial DE genes (Figure 4; Data S1). Overall, these results suggest that a group of distinct TFs contribute to cell-type gene enrichment, with some TFs also significantly enriched in the regulated cell type. However, for most, cell-type enrichment of the TFs cannot explain their high regulatory potential, suggesting that more integrated mechanisms contribute to TF activity.

### Conformationally Related Genomic Functional Units Could Be Involved in TF Activity and Cell-Type-Enriched Gene Expression Profiles

We finally investigated whether DNA conformation may contribute to the cell-type-specific profile of DE genes in the striatum, particularly topological adjacent domains (TADs). TADs form loops between distant parts of the genome and are known to contribute to tridimensional folding of DNA and transcriptional regulation. We retrieved the coordinates of TADs and determined the location of our DE genes (Dixon et al., 2012). Most DE genes belonged to a known TAD, and the same number of TADs was found for each cell population relative to randomly selected genes (Figure 5A; Table S5). Hypergeometric analysis did not demonstrate significant enrichment of a specific group of TADs, suggesting that DE genes for each cell type are not transcribed from distinct TADs (Figure 5B). A substantial proportion of TADs were shared between cell types. However, the distribution of DE and randomly selected genes within TADs was not significantly different (Figure 5C). Overall, our results show the absence of an association between cell-type gene enrichment and TAD membership. However, as previously shown, topological domains can be subdivided into smaller subdomains, comprising a reduced number of genes more prone to common transcriptional regulation. Enhancer-promoter units (EPUs) have been linked to boundaries between smaller domains and acting as regulators of gene transcription that belong to their domains. Therefore, we assessed whether our DE genes belonged to specific mouse cortex EPUs (Dixon et al., 2012). There was significantly less overlap between EPUs associated with SPN DE genes and those associated with glial DE genes relative to the overlap resulting from randomly selected genes (Figure 5D; Table S5). This suggests that cell-type-enriched genes in SPN and glia preferentially originate from different EPUs. We assessed whether an EPU is characterized by common regulatory transcriptional elements by performing TF analysis of DE genes within EPUs containing at least two genes and from which at least 50% were cell-type enriched. Here, we fully recapitulated the TF profile previously identified for each cell types, particularly

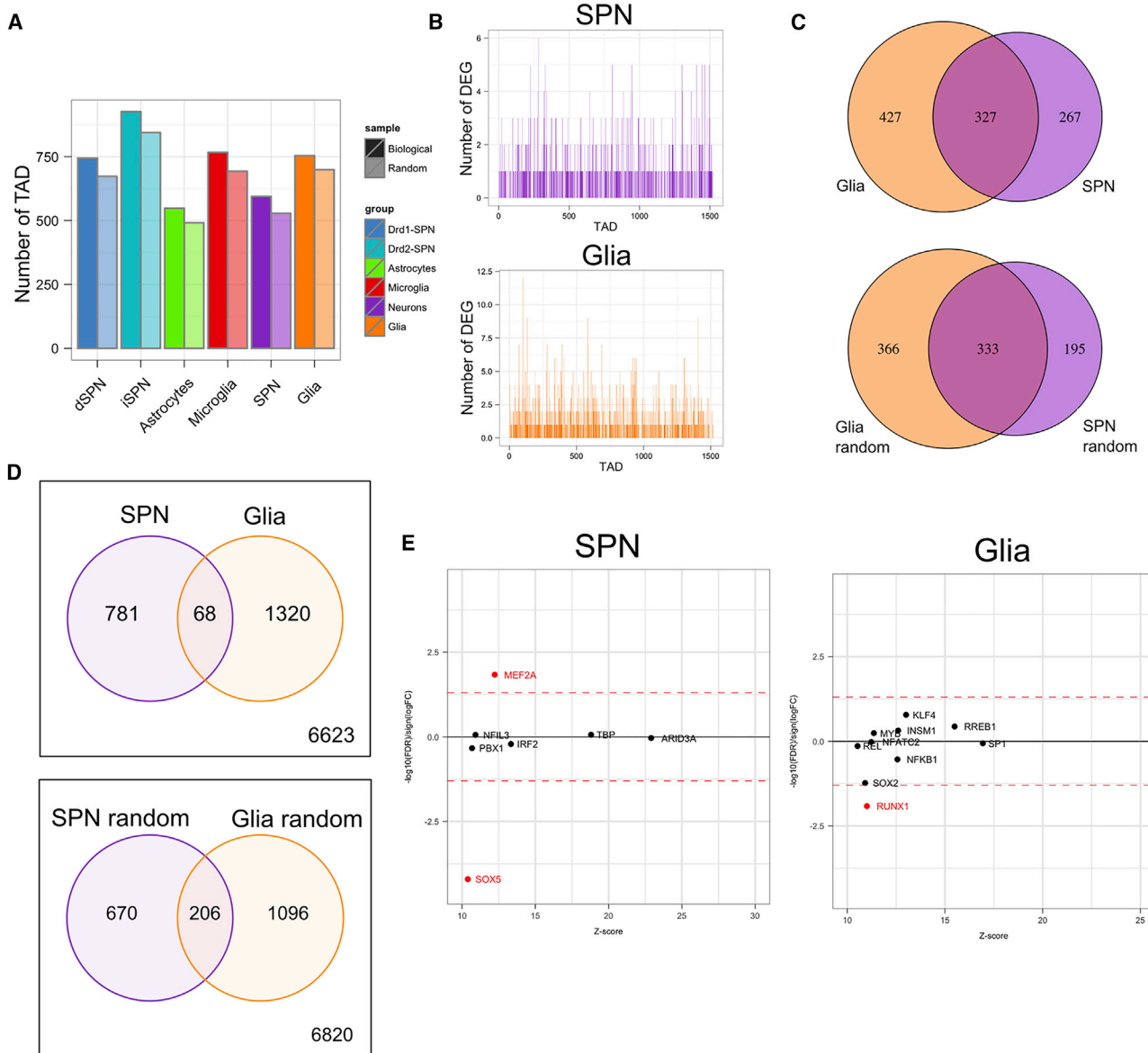
TFs with a high estimated regulatory score but non-DE in our study (Figure 5E). This suggests that the fine tridimensional regulation of DNA near genes may be involved in promoter accessibility to TFs and could serve as a regulator of TF activity on cell-type-enriched genes.

### Analysis of HD Transcriptional Signature

Striatal neurons and glial cells have been implicated in HD, highlighting the necessity to better understand the specific contribution of these cell types (Allaman et al., 2011). In particular, whether transcriptional and epigenetic signatures induced by mutant HTT differ between neurons and glial cells and within SPNs, between the dSPN and iSPN, is yet to be determined. Despite ubiquitous expression of *HTT* in striatal cells (Figure 6A), dSPNs and iSPNs display differential vulnerability to mutant huntingtin (HTT) leading to specific dysfunction in patients (Zucato et al., 2010). To gain insights into these questions, we compared DE genes with striatal transcriptomic data from HD patients and mice, including HD knockin mice and R6/1 mice (Achour et al., 2015; Hodges et al., 2006; Langfelder et al., 2016).

We extracted genes significantly up- or downregulated in the striatum compared to healthy controls ( $p \leq 0.001$ , according to Hodges et al.; Figure S4A) (Hodges et al., 2006). We retrieved 1,984 upregulated and 1,626 downregulated mouse orthologous genes. Genes significantly upregulated in HD were more highly represented in glia than SPNs (Figure 6B), whereas we observed the opposite pattern for downregulated genes (Figure 6C). We observed the same trends when considering each cell population separately (Figures 6D and 6E). In addition, we obtained similar results when intersecting our database with RNA-seq data generated in the striatum of R6/1 mice at the symptomatic stage (Achour et al., 2015) (Figure S4B; Table S6), thereby showing that neuronal and glial cells preferentially contributed to down- and upregulated genes, respectively.

We next performed Gene Ontology (GO) analyses with HD-modified genes specifically enriched in one cell population. Functions significantly downregulated in HD were mainly associated with neuronal functions (e.g., “Regulation of ion transport” or “Cell-cell signaling”; Figure S4C). In contrast, increased HD transcripts were associated with metabolism (“Negative regulation of transcription, DNA-templated”), cell responses to stimulus, and morphogenesis (Figure S4D). Downregulated transcripts specifically enriched in dSPNs and iSPNs shared common ontologies linked to neurotransmission and signaling (Figures 6F and 6G). However, a few upregulated ontologies in these neurons were significantly enriched, including cell apoptosis in dSPN and negative regulation of transcription and DNA-templated in iSPNs (Figures S4E and S4F). Upregulated HD genes in microglia displayed a strong immune-inflammatory signature, with strong implication of the JAK-STAT pathways (Figure 6H). Finally, upregulated HD genes in astrocytes were associated with developmental and metabolic processes, including cholesterol-related pathways (Figure 6I). Comparable GO results were obtained using down- and upregulated genes in HD R6/1 striatum, except that the neuroinflammatory signature was more prominent in HD patients than in HD mice (Figures S4G–S4J).



**Figure 5. Topological Adjacent Domains and Enhancer-Promoter Units**

(A) Barplot showing the total number of TADs from which cell-type-enriched genes belong. For each contrast, the position of the DEG was retrieved, intersected with the coordinates of the TADs, and the total number of unique TADs counted. For each cell type, the left bar reflects the total number of TADs for true cell-type-enriched genes, whereas the right bar shows the number of TADs obtained from the same number of randomly selected genes.

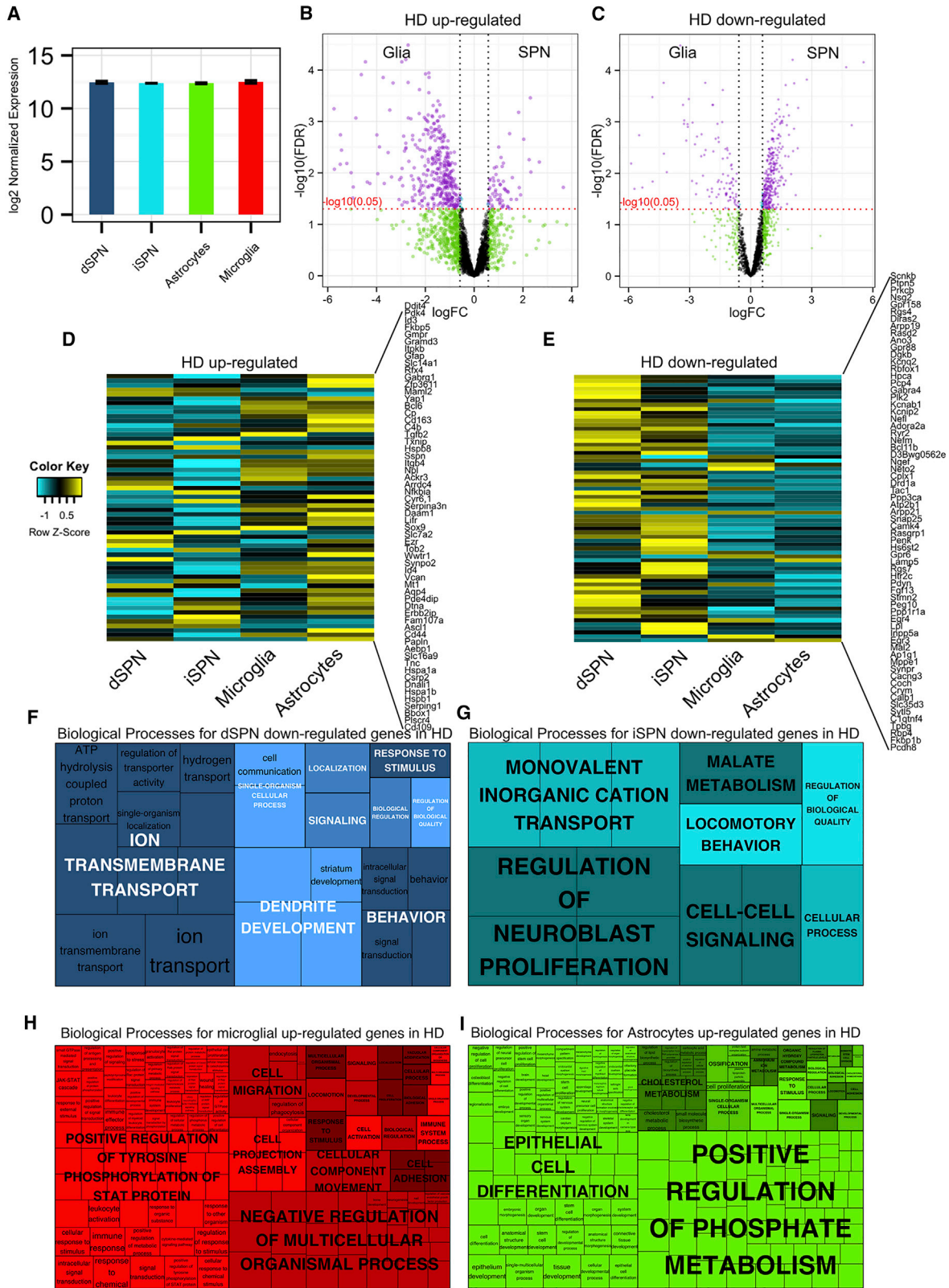
(B) Barplot representing the number of cell-type-enriched genes in each TAD detected for SPN (up) and glia (down) contrasts. Hypergeometric analysis revealed that no TAD is significantly enriched for DEGs.

(C) Venn diagram showing the sharing of TADs containing cell-type-enriched genes between SPN and glia. An equivalent analysis was performed for TADs identified from the randomly selected genes for each contrast. Statistical analyses showed an absence of a significant difference between profiles obtained with randomly selected genes and profiles obtained from true DEGs.

(D) The same analysis was performed at the EPU level. Statistical analysis shows a significantly less overlap between EPUs containing SPN- and glia-specific DEGs than EPUs identified from randomly selected genes for each contrast, revealing that at least a part of the cell-type-enriched genes may originate from EPUs specific to SPN and those specific to glia.

(E) TF-binding site prediction obtained for the DEGs belonging to EPUs with a ratio of DEG to the total number of genes >50% for the SPN and glia contrasts. The profiles are consistent with the results obtained for total DEGs, suggesting that the transcriptional control of genes belonging to EPUs containing a significant number of cell-type-enriched genes may be mediated by TFs with a high probability of regulating the DEG.

See also [Table S5](#).



(legend on next page)

Finally, we performed a comparison of the network signatures obtained in our database to the whole striatum network signatures previously obtained in the allelic series of HD knockin mice (*Hdh*) (Langfelder et al., 2016). Such a comparison showed that several cell-type-enriched genes may be strongly deregulated in a polyQ- and age-dependent manner in the striatum of *Hdh* mice (Table S3, sheet 2, last columns after selecting “True” for all columns under a specific cell type). For example, network analysis revealed that *Eda2r*, a member of the tumor necrosis factor (TNF) receptor family associated with reactive astrogliosis (Zamanian et al., 2012), and the cell-to-cell or cell-to-matrix glycoprotein *Thbs2* are two genes that are selectively and highly enriched in astrocytes and strongly upregulated in the striatum of *Hdh* mice over time. These comparisons also highlighted the glutathione peroxidase *Gpx6*, the interferon-alpha-inducible protein 27 like 2B *Ifi272b*, and the protein tyrosine phosphatase *Ptpn7* as three genes that are significantly enriched in dSPNs and strongly downregulated in the striatum of *Hdh* mice over time.

Overall, our analysis shows that transcriptional alterations caused by mutant Huntingtin (mHTT) are clearly distinct among the four main cell populations of the striatum and that mHTT may trigger a strong decrease in the expression of neuronal genes accompanied by increased expression of glial genes, suggesting that the transcriptional and/or epigenetic machinery is altered in a cell-type-specific manner.

### Application of Our Database to Study the HD Epigenetic Signature

We next studied the acetylation state of H3K27 (H3K27ac), which has been linked to gene downregulation in HD mouse striatum (Achour et al., 2015; Le Gras et al., 2017), to further understand how mHTT might alter transcriptional regulation in a cell-type-specific manner. We performed H3K27ac ChIP-seq using the striatum (e.g., caudate nucleus) of grade 2 HD patients and control individuals. H3K27ac ChIP-seq data were also generated using their cerebellum, a tissue spared by the disease (Figures S5A–S5C). As expected, H3K27ac was specifically enriched at striatal markers (e.g., PPP1R1B/DARPP32) in striatal samples and enriched at cerebellar markers (e.g., NEUROD2) in cerebellar samples (Figure S5D). We then assessed regions that were differentially enriched in H3K27ac in HD versus control striata using a method that allows computing single H3K27ac values per gene. In the striatum, 185 and 200 genes were asso-

ciated with significantly decreased and increased H3K27ac signals, respectively. No substantial changes in H3K27ac signals were observed in cerebella tissues, as expected (Table S6). GO analysis showed that genes with decreased H3K27ac levels in the striatum of HD patients were enriched in neuronal-related processes (Figure S5E). In contrast, fewer GO terms were significantly enriched when analyzing genes associated with increased H3K27ac levels, and they were associated with extracellular matrix (Figure S5E). We then integrated human H3K27ac ChIP-seq with microarray data generated in HD patients to assess H3K27ac levels at deregulated genes (Hodges et al., 2006). H3K27ac levels were lower at downregulated genes in the striatum of HD patients than in control individuals, particularly at dSPN-specific genes (Figure 7A). In contrast, H3K27ac levels were not significantly different at upregulated genes and/or glia-specific genes in HD patients, despite a trend toward their increasing (Figure 7A).

Downregulated genes in the striatum of HD R6/1 mice are preferentially controlled by super-enhancers (i.e., broad enhancers encompassing and controlling cell-type-specific genes that regulate cellular identity) (Achour et al., 2015). We investigated whether such a super-enhancer signature is present in the striatum of HD patients. Kmeans clustering was performed using a density map of H3K27ac profiles, generating six distinct clusters (A–F) (Figures 7B). Clusters A and C, which contained genes with H3K27ac signals distributed across all the genes, were enriched in super-enhancer-regulated genes (Figure 7B). Glial-specific genes were enriched in cluster A (Figure 7C), which contained genes linked to functions related to exosomes, transcription, and myelin (Figure S5F). In contrast, neuron-specific genes were not enriched in cluster A, but they were enriched in cluster C (Figure 7C), which displayed a significant neuronal signature (Figure S5G). Accordingly, clusters A and C were enriched in up- and downregulated genes in HD patients, respectively (Figure 7D). More specifically, upregulated genes originating from both astrocytes and microglia were strongly enriched in cluster A, whereas downregulated genes originating from both dSPNs and iSPNs were strongly enriched in cluster C (Figure 7E). We then performed a similar analysis, but integrating our database with the H3K27ac ChIP-seq and RNA-seq datasets previously generated using R6/1 striatum (Achour et al., 2015). R6/1 analysis essentially confirmed the human results (Figure S6). Remarkably, H3K27ac levels were lower in neuron-specific genes and higher in glia-specific genes (Figure S7). Thus, opposite epigenetic

### Figure 6. SPNs, Astrocytes, and Microglia Are Differentially Affected in Huntington's Disease

(A) Barplot showing the normalized expression levels of *Htt* in dSPNs (dark blue), iSPNs (turquoise), astrocytes (green), and microglia (red). Results are presented as the mean  $\pm$  SD.

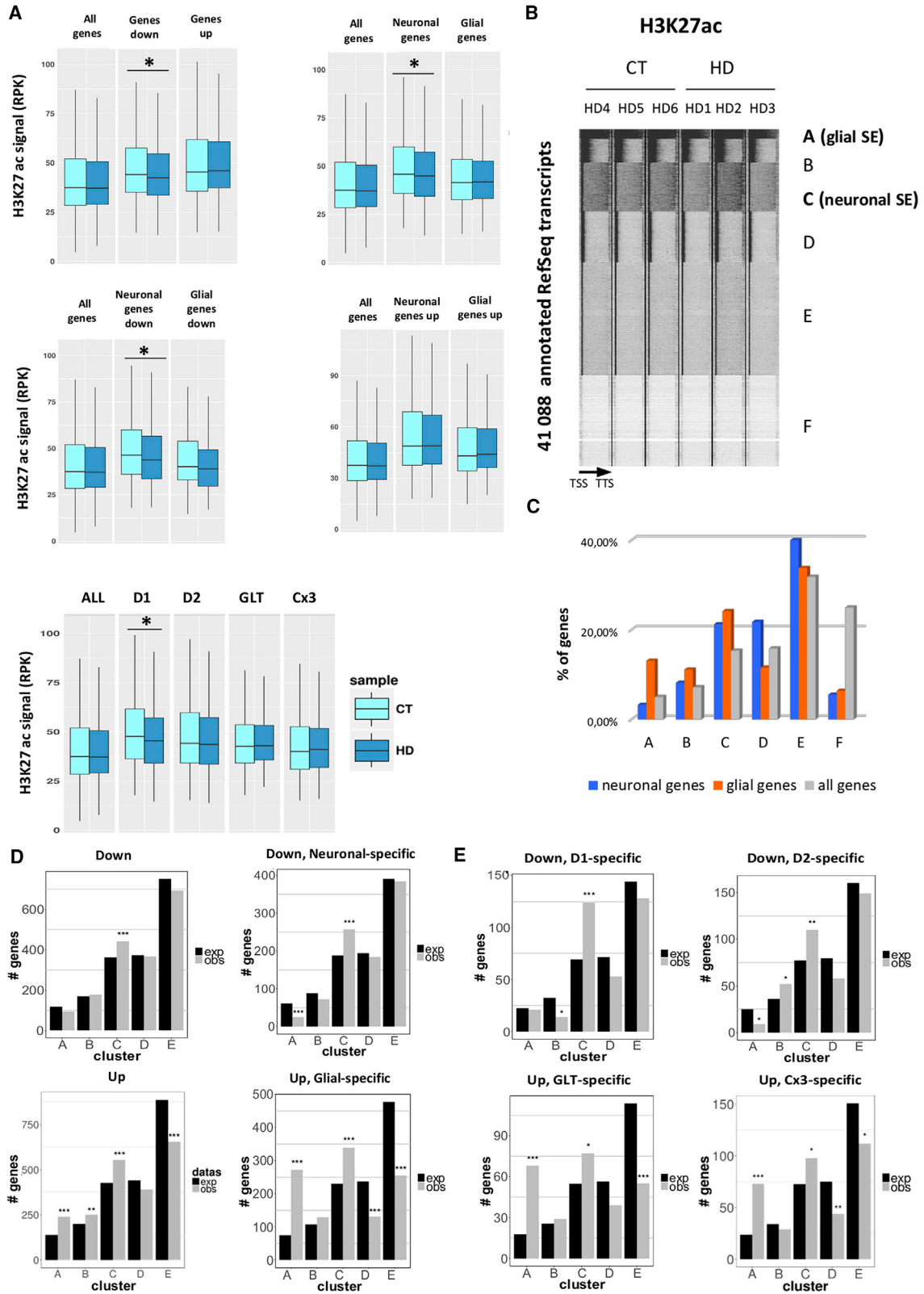
(B and C) Volcano plots representing the logFC and  $-\log_{10}(\text{FDR})$  from the SPN-glia contrasts for HD upregulated (B) or downregulated (C) genes. The horizontal red-dotted line represents the FDR threshold (0.05) on a  $-\log_{10}$  scale, whereas the vertical black-dotted lines show the  $\log_{2}\text{FC} = 0.58$  and  $\log_{2}\text{FC} = -0.58$ . Violet dots represent transcripts with a significant FDR and an absolute  $\log_{2}\text{FC} \geq 0.58$ , turquoise dots represent transcripts with a significant FDR but a lower  $\log_{2}\text{FC}$ , and green dots represent non-significant transcripts with an absolute  $\log_{2}\text{FC} \geq 0.58$ .

(D and E) Heatmaps of the top 100 significant HD upregulated (D) and HD downregulated (E) genes according to the logFC. The color scale represents the z-scaled expression levels in dSPNs, iSPNs, astrocytes, and microglia from our study. The analysis reveals a higher number of HD upregulated genes enriched in astrocytes and microglia, whereas a higher number of HD downregulated transcripts are enriched in dSPNs and iSPNs.

(F–I) Treemap representations of Gene Ontology (GO) analysis performed on HD downregulated genes significantly enriched in dSPNs (F) or iSPNs (G) and HD upregulated genes significantly enriched in astrocytes (H) or microglia (I). Each significant GO (small squares) was grouped according to its parental ontology to underline highly represented functions.

See also Figure S4 and Table S6.





(legend on next page)

mechanisms govern the principal transcriptional responses of HD neurons and glial cells (i.e., transcriptional down- and upregulation in neurons and glial cells, respectively).

## DISCUSSION

Multicellular interactions in the CNS are at the basis of complex and high-level functions. Understanding this rich cellular environment requires standardized and sensitive procedures. Breakthroughs in high-throughput profiling methods, from that of DNA to proteins, are at the basis of large international initiatives aiming to characterize specific brain regions (Hawrylycz et al., 2012; Thompson et al., 2014) (<http://portal.brain-map.org/>). However, approaches to define the roles and characterize individual CNS cell populations in adult mammals are still challenging. Here, we provide a method combining existing technologies to dissect independent cell populations in the adult CNS. Application of this method to the mouse striatum produced the first mRNA and miRNA database for the four major striatal cell types, allowing the possibility to investigate the regulation mechanisms of cell-type-specific gene expression and the differential involvement of these cells in physiological and pathological conditions.

### Power, Accessibility, and Flexibility of the Procedure

The procedure combines tissue processing, LCM, and high-throughput analysis of cellular content under physiological conditions. Previous LCM and Translating Ribosome Affinity Purification (TRAP)-based studies relied on the development of specific fluorescent strains expressing reporter genes in the targeted cell population, which is time consuming (Doyle et al., 2008; Heiman et al., 2008; Rossner et al., 2006). Our method is compatible with any strain among the hundreds developed by the Gene Expression Nervous System Atlas (GENSAT) project (Gerfen et al., 2013; Gong et al., 2003). Additionally, crossing these mice with existing models of neurological disorders makes it possible to assess cell-type-specific alterations under pathological conditions.

We used a modified tissue-slicing procedure based on electrophysiological protocols to prevent the loss of fluorescence during sample freezing (Rossner et al., 2006). We demonstrate that brain tissue slicing in oxygenated artificial cerebrospinal fluid (aCSF) at 4°C ensures preservation of both the fluorescence

and RNA integrity, offering the possibility to extract large coding and non-coding RNA (mRNA and lincRNA), as well as small non-coding RNA (miRNA) from individual adult brain cells. This method also limits the cellular stress commonly observed with trituration and homogenization procedures (Lobo et al., 2006; Molyneaux et al., 2015).

### Cell-Type-Specific Transcriptomic Signatures and Their Relevance to Transcriptional Reprogramming in HD

Previous studies have investigated cell-type-specific transcriptomic signatures in the mouse striatum (Gokce et al., 2016), providing insight into the signaling pathways that may distinguish cell types. Here, we provide a novel resource for probing the molecular diversity and studying the biology of specific cell types in the mouse striatum, covering dSPNs, iSPNs, astrocytes, and microglia. Importantly, our data include transcriptomic signatures that are retained by our network analysis (i.e., spectral decomposition), in which data are analyzed against probabilistic functional networks (i.e., MouseNet) and the risk of false positives and false negatives is reduced by virtue of the integration of orthogonal datasets. Spectral decomposition of gene expression levels against probabilistic functional networks may retain genes not retained by traditional analysis of transcriptomic data, and vice versa. These network signatures are therefore strongly associated with the signaling systems that may underlie cellular identity in the mouse striatum. We detected no overlap with previously reported and larger signatures of human striatal neurons (Kelley et al., 2018) and mouse D1 striatal neurons (Gokce et al., 2016), as inferred from the data made available in this latter study. These comparisons suggest that network analysis is able to extract precise functional signatures not necessarily put forth by traditional analysis of transcriptomic data. These comparisons also suggest that mouse and human striatal neurons could differ in terms of molecular identity and the genes that are mostly enriched in these cell types. Additionally, network analysis notably showed that *Eda2r*, a member of the TNF receptor family associated with reactive astrogliosis, and the cell-to-cell or cell-to-matrix glycoprotein *Thbs2* are two genes selectively and highly enriched in astrocytes. These genes are also strongly upregulated in the striatum of *Hdh* mice over time, highlighting their relevance to the reprogramming of transcription in the disease. Network analysis also

### Figure 7. H3K27ac Is Reduced at Neuron-Specific and Downregulated Genes in HD Patients

(A) Boxplot representation of H3K27ac signals (RPK) at down- and upregulated genes in HD patients and neuronal- and glial-specific genes in human striatum (top). Boxplot representations of H3K27ac signals are also shown for downregulated neuron-specific genes in HD striatum, downregulated glia-specific genes in HD striatum, upregulated neuron-specific genes in HD striatum, upregulated glia-specific genes in HD striatum (middle), and dSPN (D1)-, iSPN (D2)-, astrocyte (GLT)-, and microglia (Cx3)-specific genes in human striatum (bottom). H3K27ac signals for controls (CT; light blue) and patients (HD; dark blue) were compared using the Wilcoxon test (\* $p < 0.05$ ).

(B) Heatmap of the 41,088 annotated transcripts, integrating H3K27ac gene profiles and showing six distinct epigenetic profiles (clusters A–F; left). The arrow indicates the orientation of the genes. TSS, transcription start site; TTS, transcription termination site.

(C) Bar graph showing the distribution of neuron-specific, glia-specific, and all genes in clusters A–F. Human cluster A contains predominantly glial super-enhancers (glial SE), whereas human cluster C is enriched in neuronal super-enhancers (neuronal SE).

(D) Bar graphs representing the number of genes in each cluster that are neuron-specific and/or downregulated in HD striatum or glia-specific and/or upregulated in HD striatum.

(E) Bar graphs representing the number of genes in each cluster that are dSPN (D1)- or iSPN (D2)-specific and downregulated in HD striatum or astrocyte (GLT)- or microglia (Cx3)-specific and upregulated in HD striatum.

In (D) and (E), the observed numbers were compared with the expected numbers using a binomial test (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; NS, not significant). See also Figures S5 and S6 and Table S6.

showed that the glutathione peroxidase *Gpx6*, the interferon, alpha-inducible protein 27 like 2B *Iff27l2b*, and the protein tyrosine phosphatase *Ptpn7* are three genes selectively and highly enriched in dSPNs and strongly downregulated in the striatum of *Hdh* mice over time. Collectively, our data suggest that unraveling the molecular basis of striatal cell identity benefits from the use of complementary approaches for the isolation of specific cell types and unbiased assessment of molecular and cellular identity.

### Nuclear-Related Transcriptional Control: The Importance of Combining Multiple Levels of Analysis

We took advantage of our transcriptomic database to assess potential mechanisms leading to these cell-type-specific patterns. Cell-type-specific miRNA profiles did not predict mRNA enrichment profiles. This result is consistent with previous studies showing the absence of a correlation between miRNA and mRNA expression levels in cancer cells (Li et al., 2015). It has been shown that other crucial steps for miRNA activity occur before loading into the RNA-induced silencing complex (RISC), such as the control of RISC loading by the endonuclease dicer (DICER) complex or the removal of the passenger strand to generate a mature complex, as well as the regulation of miRNA activity by other noncoding RNA (for review, see Ha and Kim, 2014). However, our analysis only focused on mRNA degradation and did not consider the fact that the physiological regulation of mRNA expression by miRNA often leads to translational inhibition. Therefore, a more precise analysis with a cell-type-specific proteomic profile would be necessary to infer the role of miRNA on mRNA profiles.

We therefore performed an analysis of TFBSs to identify potential TFs responsible for the cell-type-specific profiles and found some to be coherent with the dynamics of the epigenetic control of mRNA expression. In the developing brain, SOX9 is a potent regulator of TFs implicated in the differentiation of neuronal or glial fates (Baroti et al., 2016). High SOX9 expression is associated with glial differentiation, and its inhibition blocks the cellular cycle (Kang et al., 2012). The associated activity of SOX9 with SOX5 or MEF2A was described in developmental brain or in peripheral cell populations (Gohlke et al., 2008).

RUNX1 is a major actor in microglial differentiation and proliferation, and Zfp423 was also described as a potential negative regulator of genes involved in neuronal differentiation (Alcaraz et al., 2006). In peripheral cells, Zfp423 is an inhibitor of EBF1 expression, which is consistent with our finding showing an enrichment of Zfp423 in glial cells and a significant non-enrichment of EBF1 in glial cells (Tsai and Reed, 1997). Notably, TFs with a high probability of regulating cell-type-enriched genes were often not DE. This suggests that they are good regulators of cell-type-specific profiles not because they are enriched in the cell population but rather because other, unknown factors make them potent regulators of DE genes. In particular, it has been previously demonstrated that DNA conformation is a strong regulator of TF activity and access to the promoter and that smaller EPU can be distinguished based on DNA interactions (Dixon et al., 2012; Patel et al., 2018; Zaytseva and Quinn, 2018). Indeed, we found the DNA structure in EPUs could be linked to cell-type-enriched genes in our data using previously

published databases. This is coherent data showing that samples from different organs can display different EPU profiles and suggests a smaller and more integrated level of regulation by DNA folding (Dixon et al., 2012). Our dataset on striatal cell-type-specific transcriptome in adult glial and neuronal subpopulations provides an opportunity to model and partially explain the transcriptional profiles by integrating current knowledge about TFs, their interactions with each other, and the DNA conformation that regulates TF binding at promoter levels. These analyses suggest that multiple factors, both at the epigenetic and conformational levels, may be involved in the specification of cell-type-enriched genes.

### HD: A Ubiquitously Expressed Disease Gene with a Cell-Type-Specific Response

Transcriptional dysregulation is a major feature of HD brain tissues. Although establishing causal relationships with functional outcomes has not yet been possible, transcriptional dysregulation is believed to contribute to HD pathogenesis. Many transcriptomic studies have been performed using either brain tissues of HD mouse models or postmortem tissues of HD patients (for review, see Francelle et al., 2017; Seredenina and Luthi-Carter, 2012). These studies have identified defined signatures associated with down- and upregulated genes. The dynamics of transcriptional changes has been particularly refined through a recent comprehensive transcriptomic study, based on HD knockin (KI) mice expressing murine Htt with various CAG repeat sizes (Langfelder et al., 2016). Specifically, it was shown that transcriptional changes in HD mice start at early symptomatic stages and are progressive and CAG-repeat-length dependent. Previous studies have also shown that downregulated genes in the striatum of HD patients and mice are enriched in identity genes of medium-spiny neurons (Achour et al., 2015; Hodges et al., 2006; Langfelder et al., 2016; Vashishtha et al., 2013). In addition, it was found that downregulated genes in HD mouse striatum are predominantly under the control of super-enhancers, a class of broad enhancers that regulate cellular identity genes, suggesting that altered regulation of histone acetylation might underlie the loss of striatal identity (Achour et al., 2015; Le Gras et al., 2017). However, these studies did not address the contribution of neurons and glial cells. Here, we unambiguously show that the striatal identity signature of HD downregulated genes originates from neurons, including dSPNs and iSPNs, and not from glial cells. Downregulated genes in HD striatum not only define the identity of striatal neurons but also control neuronal activity, including neuronal transmission, synaptic plasticity, and adaptive behavior. Thus, it is likely that downregulation of neuronal-identity genes in HD striatum has deleterious functional consequences.

The functional signature of upregulated genes in the striatum of HD patients and mice is less well defined and conserved across models than that of downregulated genes (Achour et al., 2015; Hodges et al., 2006; Langfelder et al., 2016). Previous studies show that upregulated genes in HD striatum include GO terms related to development, metabolism, inflammation, and more generally stress responses (Achour et al., 2015; Hodges et al., 2006; Langfelder et al., 2016). Our epigenetic and transcriptomic analyses show that glial cells, including

astrocytes and microglia, are predominantly responsible for this signature. It is possible that glial cells may adapt their metabolism through epigenetic and transcriptional regulation in response to the HD mutation, thereby activating stress resistance and survival mechanisms.

Collectively, our results demonstrate that predominant epigenetic and transcriptional responses are opposite in HD striatal neurons and glial cells, which might underlie their differential vulnerability and/or responses to the HD mutation. Indeed, it was recently shown that expression of a mutant form of a fragment of Huntingtin in either neurons or astrocytes led to distinct phenotypes in mice (Meunier et al., 2016). Differential epigenetic and transcriptional responses in these two cell types may well underlie the observed pathophysiological and behavioral differences. Future studies are required to elucidate the underlying mechanisms.

Overall, these data emphasize the importance of using cell-type-specific approaches for genome-wide analyses and provide a new level of resolution and mechanistic insight of HTT biology and cell-type-specific adaptation to the disease state. The samples generated with the new protocol are compatible with non-coding RNA profiling, alternative splicing, and genomic variant and high-throughput protein quantification (Ha and Kim, 2014; He et al., 2013; Hebbar et al., 2014). The opportunity to combine these high-throughput methods with the new procedure will contribute to our understanding of the “cellular interactome” from the gene to protein activity, which is at the core of individual cell function.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Animals
  - Human samples
- **METHOD DETAILS**
  - Animals
  - Brain processing
  - Immunostaining
  - Brain processing and laser-capture microdissection (LCM)
  - RNA purification and quality check
  - Library preparation and sequencing
  - miRNA microarray
  - RT-qPCR
- **QUANTIFICATIONS AND STATISTICAL ANALYSIS**
  - RNaseq and microarray data processing
  - DE gene analysis
  - Evaluation of cell markers, stress and strain effect
  - Sample correlation and comparison with the Allen Brain Atlas
  - Network analysis of cell-type specific mRNA-seq data
  - miRNA clustering
  - mRNA regulation by DE miRNA

- Analysis of TF binding sites on DE genes
- Evaluation of the role of DNA conformation on DE genes
- Cross-analysis with HD expression profiles
- ChipSeq data analysis
- Differential binding analysis
- **DATA AND SOFTWARE AVAILABILITY**

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found with this article online at <https://doi.org/10.1016/j.celrep.2019.02.003>.

## ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Florence D. Morgenthaler-Grand (CIF, Lausanne University) and Dr. Georg Kienzel (Arcturus, Basel) for LCM technical support. We thank Dr. Keith Harshman and the Genomic Technologies Facility team at the Center for Integrative Genomics (Lausanne University) for RNA-seq and miRNA microarray experiments. We thank Virginie Zimmer (LCMN, Lausanne University Hospital) for animal breeding and genotyping. We gratefully acknowledge Sylvain Pradervand (CIG, Lausanne University), Leonore Wiegler (CIG, Lausanne University), and the Bioconductor support team for help in bioinformatics. This work was supported by Swiss National Science Foundation grant 31003A-140945, Joint Programming Neurodegenerative Disease (JPND) grant 31ND30-166947, the European Huntington's Disease Network (EHDN; project 599), and funds from Centre National de la Recherche Scientifique (CNRS). Sequencing was performed by the genomEast Platform, a member of the France Génomique consortium (ANR-10-INBS-0009). A.A and J.S. were supported by EHDN and CHDI, respectively.

## AUTHOR CONTRIBUTIONS

N.D. and K.M. designed the project. N.M., C.M., A.B.R., and J.-Y.C. developed and optimized the new procedure. N.M. and C.M. performed the LCM experiments, immunostaining, and validation qRT-PCR. N.M. performed the RNA extractions. A.S. performed the human ChIP-seq experiments. R.F. provided post-mortem brain tissues. N.M., C.M., J.S., S.L.G., C.K., and K.M. designed and performed the bioinformatic analysis. C.N. designed and S.S.N. performed the network-based analysis. N.M., C.M., N.D., C.N., S.S.N., L.P., J.-Y.C., and K.M. contributed to the interpretation of the data. N.M., C.M., K.M., C.N., and N.D. wrote the manuscript. L.P. and J.-Y.C. performed critical reading of the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 28, 2018

Revised: December 21, 2018

Accepted: February 1, 2019

Published: February 26, 2019

## REFERENCES

- Achour, M., Le Gras, S., Keime, C., Parmentier, F., Lejeune, F.X., Bouillier, A.L., Néri, C., Davidson, I., and Merienne, K. (2015). Neuronal identity genes regulated by super-enhancers are preferentially down-regulated in the striatum of Huntington's disease mice. *Hum. Mol. Genet.* *24*, 3481–3496.
- Ade, K.K., Wan, Y., Chen, M., Gloss, B., and Calakos, N. (2011). An improved BAC transgenic fluorescent reporter line for sensitive and specific identification of striatonigral medium spiny neurons. *Front. Syst. Neurosci.* *5*, 32.
- Alcaraz, W.A., Gold, D.A., Raponi, E., Gent, P.M., Concepcion, D., and Hamilton, B.A. (2006). Zfp423 controls proliferation and differentiation of neural



- precursors in cerebellar vermis formation. *Proc. Natl. Acad. Sci. USA* *103*, 19424–19429.
- Allaman, I., Bélanger, M., and Magistretti, P.J. (2011). Astrocyte-neuron metabolic relationships: for better and for worse. *Trends Neurosci.* *34*, 76–87.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* *11*, R106.
- Baroti, T., Zimmermann, Y., Schillinger, A., Liu, L., Lommes, P., Wegner, M., and Stolt, C.C. (2016). Transcription factors Sox5 and Sox6 exert direct and indirect influences on oligodendroglial migration in spinal cord and forebrain. *Glia* *64*, 122–138.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* *136*, 215–233.
- Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., et al. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* *28*, 264–278.
- Cheng, L., Zhang, S., MacLennan, G.T., Williamson, S.R., Davidson, D.D., Wang, M., Jones, T.D., Lopez-Beltran, A., and Montironi, R. (2013). Laser-assisted microdissection in translational research: theory, technical considerations, and future applications. *Appl. Immunohistochem. Mol. Morphol.* *21*, 31–47.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* *485*, 376–380.
- Doyle, J.P., Dougherty, J.D., Heiman, M., Schmidt, E.F., Stevens, T.R., Ma, G., Bupp, S., Shrestha, P., Shah, R.D., Doughty, M.L., et al. (2008). Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell* *135*, 749–762.
- Francelle, L., Galvan, L., and Brouillet, E. (2014). Possible involvement of self-defense mechanisms in the preferential vulnerability of the striatum in Huntington's disease. *Front. Cell. Neurosci.* *8*, 295.
- Francelle, L., Lotz, C., Outeiro, T., Brouillet, E., and Merienne, K. (2017). Contribution of Neuroepigenetics to Huntington's Disease. *Front. Hum. Neurosci.* *11*, 17.
- Gerfen, C.R., Paletzki, R., and Heintz, N. (2013). GENSAT BAC cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. *Neuron* *80*, 1368–1383.
- Gohlke, J.M., Armant, O., Parham, F.M., Smith, M.V., Zimmer, C., Castro, D.S., Nguyen, L., Parker, J.S., Gradwohl, G., Portier, C.J., and Guillemot, F. (2008). Characterization of the proneural gene regulatory network during mouse telencephalon development. *BMC Biol.* *6*, 15.
- Gokce, O., Stanley, G.M., Treutlein, B., Neff, N.F., Camp, J.G., Malenka, R.C., Rothwell, P.E., Fuccillo, M.V., Südhof, T.C., and Quake, S.R. (2016). Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* *16*, 1126–1137.
- Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., and Heintz, N. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* *425*, 917–925.
- Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* *15*, 509–524.
- Han, I., You, Y., Kordower, J.H., Brady, S.T., and Morfini, G.A. (2010). Differential vulnerability of neurons in Huntington's disease: the role of cell type-specific features. *J. Neurochem.* *113*, 1073–1091.
- Handley, A., Schauer, T., Ladurner, A.G., and Margulies, C.E. (2015). Designing cell-type-specific genome-wide experiments. *Mol. Cell* *58*, 621–631.
- Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* *489*, 391–399.
- He, M., Liu, Y., Wang, X., Zhang, M.Q., Hannon, G.J., and Huang, Z.J. (2012). Cell-type-based analysis of microRNA profiles in the mouse brain. *Neuron* *73*, 35–48.
- He, J., Zhu, J., Liu, Y., Wu, J., Nie, S., Heth, J.A., Muraszko, K.M., Fan, X., and Lubman, D.M. (2013). Immunohistochemical staining, laser capture microdissection, and filter-aided sample preparation-assisted proteomic analysis of target cell populations within tissue samples. *Electrophoresis* *34*, 1627–1636.
- Hebbar, S., Schulz, W.D., Sauer, U., and Schwudke, D. (2014). Laser capture microdissection coupled with on-column extraction LC-MS(n) enables lipidomics of fluorescently labeled *Drosophila* neurons. *Anal. Chem.* *86*, 5345–5352.
- Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., et al. (2008). A translational profiling approach for the molecular characterization of CNS cell types. *Cell* *135*, 738–748.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* *38*, 576–589.
- Hodges, A., Strand, A.D., Aragaki, A.K., Kuhn, A., Sengstag, T., Hughes, G., Elliston, L.A., Hartog, C., Goldstein, D.R., Thu, D., et al. (2006). Regional and cellular gene expression changes in human Huntington's disease brain. *Hum. Mol. Genet.* *15*, 965–977.
- Jovićić, A., Roshan, R., Moiso, N., Pradervand, S., Moser, R., Pillai, B., and Luthi-Carter, R. (2013). Comprehensive expression analyses of neural cell-type-specific miRNAs identify new determinants of the specification and maintenance of neuronal phenotypes. *J. Neurosci.* *33*, 5127–5137.
- Jung, S., Aliberti, J., Graemmel, P., Sunshine, M.J., Kreutzberg, G.W., Sher, A., and Littman, D.R. (2000). Analysis of fractalkine receptor CX(3)CR1 function by targeted deletion and green fluorescent protein reporter gene insertion. *Mol. Cell. Biol.* *20*, 4106–4114.
- Kang, P., Lee, H.K., Glasgow, S.M., Finley, M., Donti, T., Gaber, Z.B., Graham, B.H., Foster, A.E., Novitch, B.G., Gronostajski, R.M., and Deneen, B. (2012). Sox9 and NFIA coordinate a transcriptional regulatory cascade during the initiation of gliogenesis. *Neuron* *74*, 79–94.
- Kelley, K.W., Nakao-Inoue, H., Molofsky, A.V., and Oldham, M.C. (2018). Variation among intact tissue samples reveals the core transcriptional features of human CNS cell classes. *Nat. Neurosci.* *21*, 1171–1184.
- Langfelder, P., Cantle, J.P., Chatzopoulou, D., Wang, N., Gao, F., Al-Ramahi, I., Lu, X.H., Ramos, E.M., El-Zein, K., Zhao, Y., et al. (2016). Integrated genomics and proteomics define huntingtin CAG length-dependent networks in mice. *Nat. Neurosci.* *19*, 623–633.
- Le Gras, S., Keime, C., Anthony, A., Lotz, C., De Longprez, L., Brouillet, E., Cassel, J.C., Boutilier, A.L., and Merienne, K. (2017). Altered enhancer transcription reveals Huntington's disease striatal transcriptional signature. *Sci. Rep.* *7*, 42875.
- Lejeune, F.X., Mesrob, L., Parmentier, F., Bicep, C., Vazquez-Manrique, R.P., Parker, J.A., Vert, J.P., Tourette, C., and Neri, C. (2012). Large-scale functional RNAi screen in *C. elegans* identifies genes that regulate the dysfunction of mutant polyglutamine neurons. *BMC Genomics* *13*, 91.
- Li, Z., Qin, T., Wang, K., Hackenberg, M., Yan, J., Gao, Y., Yu, L.R., Shi, L., Su, Z., and Chen, T. (2015). Integrated microRNA, mRNA, and protein expression profiling reveals microRNA regulatory networks in rat kidney treated with a carcinogenic dose of aristolochic acid. *BMC Genomics* *16*, 365.
- Lobo, M.K., Karsten, S.L., Gray, M., Geschwind, D.H., and Yang, X.W. (2006). FACS-array profiling of striatal projection neuron subtypes in juvenile and adult mouse brains. *Nat. Neurosci.* *9*, 443–452.
- Macosko, E.Z., Basu, A., Satija, R., Nemes, 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.
- Merienne, N., Vachey, G., de Longprez, L., Meunier, C., Zimmer, V., Perriard, G., Canales, M., Mathias, A., Herrgott, L., Beltraminelli, T., et al. (2017). The

- self-inactivating KamiCas9 system for the editing of CNS disease genes. *Cell Rep.* 20, 2980–2991.
- Meunier, C., Merienne, N., Jollé, C., Déglon, N., and Pellerin, L. (2016). Astrocytes are key but indirect contributors to the development of the symptomatology and pathophysiology of Huntington's disease. *Glia* 64, 1841–1856.
- Molyneaux, B.J., Goff, L.A., Brettler, A.C., Chen, H.H., Hrvatin, S., Rinn, J.L., and Arlotta, P. (2015). DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. *Neuron* 85, 275–288.
- Patel, A., Yang, P., Tinkham, M., Pradhan, M., Sun, M.A., Wang, Y., Hoang, D., Wolf, G., Horton, J.R., Zhang, X., et al. (2018). DNA conformation induces adaptable binding by tandem zinc finger proteins. *Cell* 173, 221–233.e212.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.
- Rago, L., Beattie, R., Taylor, V., and Winter, J. (2014). miR379-410 cluster miRNAs regulate neurogenesis and neuronal migration by fine-tuning N-cadherin. *EMBO J.* 33, 906–920.
- Regan, M.R., Huang, Y.H., Kim, Y.S., Dykes-Hoberg, M.I., Jin, L., Watkins, A.M., Bergles, D.E., and Rothstein, J.D. (2007). Variations in promoter activity reveal a differential expression and physiology of glutamate transporters by glia in the developing and mature CNS. *J. Neurosci.* 27, 6607–6619.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47.
- Rossner, M.J., Hirrlinger, J., Wichert, S.P., Boehm, C., Newrzella, D., Hiemisch, H., Eisenhardt, G., Stuenkel, C., von Ahsen, O., and Nave, K.A. (2006). Global transcriptome analysis of genetically identified neurons in the adult cortex. *J. Neurosci.* 26, 9956–9966.
- Seredenina, T., and Luthi-Carter, R. (2012). What have we learned from gene expression profiles in Huntington's disease? *Neurobiol. Dis.* 45, 83–98.
- Silberberg, G., and Bolam, J.P. (2015). Local and afferent synaptic pathways in the striatal microcircuitry. *Curr. Opin. Neurobiol.* 33, 182–187.
- Thompson, C.L., Ng, L., Menon, V., Martinez, S., Lee, C.K., Glattfelder, K., Sunkin, S.M., Henry, A., Lau, C., Dang, C., et al. (2014). A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain. *Neuron* 83, 309–323.
- Tourette, C., Li, B., Bell, R., O'Hare, S., Kaltenbach, L.S., Mooney, S.D., and Hughes, R.E. (2014). A large scale Huntingtin protein interaction network implicates Rho GTPase signaling pathways in Huntington disease. *J. Biol. Chem.* 289, 6709–6726.
- Tsai, R.Y., and Reed, R.R. (1997). Cloning and functional characterization of Roaz, a zinc finger protein that interacts with O/E-1 to regulate gene expression: implications for olfactory neuronal development. *J. Neurosci.* 17, 4159–4169.
- Vashishtha, M., Ng, C.W., Yildirim, F., Gipson, T.A., Kratter, I.H., Bodai, L., Song, W., Lau, A., Labadorf, A., Vogel-Ciernia, A., et al. (2013). Targeting H3K4 trimethylation in Huntington disease. *Proc. Natl. Acad. Sci. USA* 110, E3027–E3036.
- Zamanian, J.L., Xu, L., Foo, L.C., Nouri, N., Zhou, L., Giffard, R.G., and Barres, B.A. (2012). Genomic analysis of reactive astrogliosis. *J. Neurosci.* 32, 6391–6410.
- Zaytseva, O., and Quinn, L.M. (2018). DNA conformation regulates gene expression: the MYC promoter and beyond. *BioEssays* 40, e1700235.
- Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keefe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* 34, 11929–11947.
- Zhou, K.R., Liu, S., Sun, W.J., Zheng, L.L., Zhou, H., Yang, J.H., and Qu, L.H. (2017). ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. *Nucleic Acids Res.* 45 (D1), D43–D50.
- Zuccato, C., Valenza, M., and Cattaneo, E. (2010). Molecular mechanisms and potential therapeutic targets in Huntington's disease. *Physiol. Rev.* 90, 905–981.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
rabbit polyclonal anti-GS antibody	Sigma-Aldrich, Buchs	Cat# <b>G2781</b> , RRID: AB_259853
rabbit polyclonal anti-Iba1 antibody	Wako Chemicals GmbH	Wako Cat# 019-19741, RRID: AB_839504
rabbit polyclonal anti-pENK antibody	Neuromics, Edina, MN, USA	Neuromics Cat# RA14124, RRID:AB_2532106
goat anti-rabbit IgG anti-rabbit IgG AlexaFluor-488	Invitrogen, Life Technologies, Zug, Switzerland	Thermo Fisher Scientific Cat# A32731, RRID:AB_2633280
goat anti-rabbit IgG AlexaFluor-594	Invitrogen, Life Technologies, Zug, Switzerland	Molecular Probes Cat# A-11005, RRID:AB_141372
goat anti-mouse IgG AlexaFluor-488	Invitrogen, Life Technologies, Zug, Switzerland	Thermo Fisher Scientific Cat# A-11001, RRID:AB_2534069
goat anti-mouse IgG AlexaFluor-594	Invitrogen, Life Technologies, Zug, Switzerland	Molecular Probes Cat# A-11005, RRID:AB_141372
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
sodium pentobarbital	Esconarkon, Streuli, Uznach, Germany	N/A
artificial cerebrospinal fluid	Sigma-Aldrich, Buchs, Switzerland	N/A
phosphate-buffered saline (PBS)	GIBCO, Life Technologies, Zug, Switzerland	N/A
paraformaldehyde	PFA, Electron Microscopy Sciences, Hatfield, USA	N/A
sucrose	Sigma-Aldrich, Buchs, Switzerland	N/A
2-methylbutane	Sigma-Aldrich, Buchs, Switzerland	N/A
RNAqueous lysis buffer		N/A
normal goat serum (NGS)	Interchim, Montluçon, France	N/A
Triton X-100	Fluka, Sigma-Aldrich, Buchs, Switzerland	N/A
superfrost microscope slides	VWR International, Dietikon, Switzerland	N/A
Vectashield Fluorescence mounting medium with DAPI	Reactolab, SA, Servion, Switzerland	N/A
GeneAmp Thin-Walled Reaction tube	Applied Biosystems, Life Technologies, Buchs, Switzerland	N/A
sodium acetate	Ambion, Life Technologies, Buchs, Switzerland	N/A
RNase/DNase-free water	GIBCO, Life Technologies, Buchs, Switzerland	N/A
dNTPs	Invitrogen, Life Technologies, Zug, Switzerland	N/A
Random Hexamer primer	Invitrogen, Life Technologies, Zug, Switzerland	N/A
<b>Critical Commercial Assays</b>		
KAPA SYBRFAST kit	Sigma-Aldrich, Buchs, Switzerland	N/A
RNAqueous Micro kit for LCM	Ambion, Life Technologies, Buchs, Switzerland	N/A
NuGen Ovation kit	NuGen, San Carlos, California, USA	N/A
Illumina TruSeq Nano kit	Illumina Switzerland GmbH, Zurich, Switzerland	N/A
Agilent Mouse miRNA Microarray	Release 19.0, 8x60K; Catalog number G4872A-046065, Agilent Technologies	N/A
RNasezap	Sigma-Aldrich, Buchs, Switzerland	N/A
CapSure XS LCM caps	Arcturus, Bucher, Basel, Switzerland	N/A
DNAzap	Ambion, Life Technologies, Buchs, Switzerland	N/A
DNase I	Ambion, Life Technologies, Buchs, Switzerland	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
NCBI	<a href="https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA510761">https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA510761</a>	PRJNA510761
Experimental Models: Organisms/Strains		
BAC-GLT1-eGFP transgenic mice	kindly provided by Prof. J Rothstein, Baltimore, MD, USA	N/A
BAC Tg(Drd1a-tdTomato)6Calak	The Jackson Laboratory, ref 016204	016204
Tg(Drd2-EGFP)S118Gsat/Mmnc	Mutant Mouse Regional Resource Center, Chapel Hill, NC, USA, ref 230	230
B6.129P-Cx3cr1tm1Litt/J	The Jackson Laboratory, ref 005582	005582
Primers		
Slc1a3-F	TCTCCAGTCTCGTCACAGGAATG	GLAST1 (Solute Carrier Family 1 Member 3SLC1A3)
Slc1a3-R	TGCCAATCACCACAGCAATG	GLAST1 (Solute Carrier Family 1 Member 3SLC1A3)
Ccr111-F	TGCCCAGGAAGAGCCTGCTA	chemokine (C-C motif) receptor 1-like 1
Ccr111-R	AATCATGGCCAACAGAGGCA	chemokine (C-C motif) receptor 1-like 1
Gad2-F	ACCTATGAGATCGCCCCTGT	GAD65 (Glutamate Decarboxylase 2; GAD2)
Gad2-R	AAGATTCCATCGCCAGAGCC	GAD65 (Glutamate Decarboxylase 2; GAD2)
Gpx6-F	TATGACCAAAGCCCACAGCA	Glutathione Peroxidase 6
Gpx6-R	TAACCGGCCAGTGCTTTGAA	Glutathione Peroxidase 6
PPIA-F	ATGGCAAATGCTGGACCAAA	cyclophilin A (peptidyl propyl isomerase A; PPIA)
PPIA-R	GCCTTCTTTCACCTTCCAAA	cyclophilin A (peptidyl propyl isomerase A; PPIA)
Software and Algorithms		
LSM software	Zeiss, Carl Zeiss Microscopy GmbH, Göttingen, Germany	N/A
CASAVA software v1.8.2: Sequencing quality control and read demultiplexing	Illumina, Switzerland GmbH, Zurich, Switzerland	N/A
FastQC_0.11.5 software	Read quality: <a href="http://www.bioinformatics.babraham.ac.uk/projects/fastqc">http://www.bioinformatics.babraham.ac.uk/projects/fastqc</a>	N/A
Cutadapt	<a href="https://cutadapt.readthedocs.io/en/stable/index.html">https://cutadapt.readthedocs.io/en/stable/index.html</a>	N/A
PrinSEQ	<a href="http://prinseq.sourceforge.net/index.html">http://prinseq.sourceforge.net/index.html</a>	N/A
Tophat aligner	<a href="https://ccb.jhu.edu/software/tophat/index.shtml">https://ccb.jhu.edu/software/tophat/index.shtml</a>	N/A
SAMtools suite	<a href="http://samtools.sourceforge.net">http://samtools.sourceforge.net</a>	N/A
HTSeq software	<a href="https://htseq.readthedocs.io/en/release_0.11.1/">https://htseq.readthedocs.io/en/release_0.11.1/</a>	N/A
RTA v1.18.61		N/A
DimerRemover v0.9.2	<a href="https://sourceforge.net/projects/dimerremover">https://sourceforge.net/projects/dimerremover</a>	N/A
Bioconductor NOISeq package	<a href="https://www.bioconductor.org">https://www.bioconductor.org</a>	N/A
Agilent Feature Extraction Software v 10.7.3.1	Agilent Technologies, Basel, Switzerland	N/A
Trimmomatic-0.36 tool	<a href="http://www.usadellab.org/cms/?page=trimmomatic">http://www.usadellab.org/cms/?page=trimmomatic</a>	N/A
bowtie2_2.2.8	<a href="http://bowtie-bio.sourceforge.net/bowtie2/index.shtml">http://bowtie-bio.sourceforge.net/bowtie2/index.shtml</a>	N/A
Other		
Zeiss LSM 510 META inverted confocal microscope	Carl Zeiss Microscopy GmbH, Göttingen, Germany	N/A
Zeiss AxioVision microscope	Zeiss, Carl Zeiss Microscopy GmbH, Göttingen, Germany	N/A

(Continued on next page)



**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Semi-automated vibratome	Leica VT1000 S, Biosystems Switzerland, Nunningen, Switzerland	N/A
ArcturusXT microscope	Bucher Biotec AG, Basel, Switzerland	N/A
Illumina HiSeq 2500 system	Illumina, Switzerland GmbH, Zurich, Switzerland	N/A
sledge microtome	Leica SM2010R, Biosystems Switzerland, Nunningen, Switzerland	N/A
cryostat	Leica CM1860, Biosystems Switzerland, Nunningen, Switzerland	N/A
Fragment Analyzer system	Agilent Technologies, Basel, Switzerland	N/A
Rotor gene	QIAGEN, Basel, Switzerland	N/A

**CONTACT FOR REAGENT AND RESOURCE SHARING**

Further information and requests for reagents may be directed to, and will be fulfilled by, the corresponding author Nicole Déglon ([nicole.deglon@chuv.ch](mailto:nicole.deglon@chuv.ch)).

**EXPERIMENTAL MODEL AND SUBJECT DETAILS****Animals**

Adult male BAC Tg(Drd1a-tdTomato)6Calak (Drd1-Tomato, 3.5 months-old) ([Ade et al., 2011](#)), Tg(Drd2-EGFP)S118Gsat/Mmnc (Drd2-eGFP, 7 months old) ([Gong et al., 2003](#)), BAC GLT1-eGFP (GLT1-eGFP, 6.5 months-old, courtesy of Pr. J. Rothstein, Baltimore, MD, USA) ([Regan et al., 2007](#)), and B6.129P-Cx3cr1tm1Litt/J (Cx3cr1-eGFP, 4.5 months-old) ([Jung et al., 2000](#)) were used for the study (Jax and Mutant Mouse Regional Resource Center, Chapel Hill, NC, USA).

**Human samples**

Post-mortem brain tissue from the caudate nucleus and cerebellum were obtained from the Neurological Foundation of New Zealand Human Brain Bank, Centre for Brain Research, University of Auckland, New Zealand. Three human grade 2 HD post-mortem samples were selected for ChIPseq. Three non-neurological controls of similar age, sex, and post-mortem interval were also selected, as controls.

**METHOD DETAILS****Animals**

Mice were housed in a specific pathogen-free (SPF) facility in IVC cages GM500 (Tecniplast) or rat R.BTM.U x /R.ICV.6 cages (Innovive, Paris, France) and Innorack racks, simple face (cat# RS.5.8.40), containing corn cob bedding with five mice per cage maximum. The animals were maintained in a controlled-temperature room ( $22 \pm 1^\circ\text{C}$ ), under a 14-hour light/10-hour dark cycle. The breeding program depends on strain or productivity requests but regularly involves trio or couple breeders. Enrichments consisted of two pieces of wipes, one cardboard tunnel, and one cardboard or polysulfide house with two entrances/exits. Food (global rodent diet XP-18, vitamin-fortified, irradiated at 25 kGy (Kliba Nafag AG, Kaiseraugst, Switzerland; Cat# 3242) and water were provided *ad libitum*. All experimental procedures were performed in strict accordance with Swiss regulations concerning the care and use of laboratory animals (veterinary authorizations: 2782, 2888 and 3073). Mice were anesthetized with a lethal dose of sodium pentobarbital (NaCl, B-Braun, Sempach, Germany; Esconarkon, Streuli, Uznach, Germany) and the brains quickly extracted for slicing with a vibratome in cold ( $4^\circ\text{C}$ ) and oxygenated artificial cerebrospinal fluid (aCSF: 85 mM NaCl, 75 mM sucrose, 2.5 mM KCl, 25 mM  $\text{NaHCO}_3$ , 1.25 mM  $\text{NaH}_2\text{PO}_4$ , 3.5 mM  $\text{MgSO}_4$ , 0.5 mM  $\text{CaCl}_2$ , 10 mM glucose; Sigma-Aldrich, Buchs, Switzerland). Coronal slices were then mounted on Superfrost slides (Superfrost) and dried before progressive dehydration. Dehydrated slices were then conserved in a desiccator to avoid rewetting. LCM was performed on an ArcturusXT microscope (Arcturus). Individual cells were dissected with an infrared laser to avoid damage to the cellular content. Total RNA was directly extracted on the cap with the RNAqueous kit (Ambion) and then stored at  $-80^\circ\text{C}$  until further purification.

**Brain processing**

For immunostaining, the mice were killed by intraperitoneal injection of sodium pentobarbital (B-Braun Mediacal SA, Sempach, Switzerland) and transcardially perfused at a flow rate of 20 mL/min with phosphate-buffered saline (PBS, GIBCO, Life Technologies,

Zug, Switzerland) for 1 min and then 4% paraformaldehyde (PFA, Electron Microscopy Sciences, Hatfield, USA) for 5 min. Brains were removed and post-fixed by incubation in 4% PFA for 12 h at 4°C. They were then cryoprotected by incubation in 20% sucrose (Sigma-Aldrich, Buchs, Switzerland) in 1 x PBS for 6 h and then in 30% sucrose in 1 x PBS for 24 h. Brains were then stored at –80°C until use. We cut 20 µm-thick coronal brain sections using a sledge microtome with a freezing stage at –30°C (Leica SM2010R, Biosystems Switzerland, Nunningen, Switzerland). Slices throughout the striatum were collected and maintained in tubes as free-floating slices in anti-freeze solution (25% glycerol; Sigma-Aldrich, Buchs, Switzerland, 30% Ethylene glycol; Merck, Nottingham, UK, 25% 1 X PBS, and 20% nanopure water).

For the evaluation of fluorescence loss, fresh brain samples from GLT1-eGFP transgenic mice were directly snap frozen in cold 2-methylbutane (Sigma-Aldrich, Buchs, Switzerland) and conserved at –80°C until use. Brains were then sliced with a cryostat (Leica CM1860, Biosystems Switzerland, Nunningen, Switzerland), mounted on Superfrost microscope slides (VWR International, Dietikon, Switzerland), and directly observed under a fluorescence microscope.

For evaluation of the perfusion effect on RNA integrity, we transcidentally perfused mice with 4% PFA at a flow rate of 20 mL/min for 1 min and directly snap-froze the brain. Ten-micrometer slices were prepared with a cryostat (Leica CM1860, Biosystems Switzerland, Nunningen, Switzerland) and directly mounted on Superfrost slides. Slices were then dehydrated as previously described (Merienne et al., 2017). The feasibility of microdissection was addressed as described in the section describing the LCM. RNA stability was tested by scratching the slices with a razor blade and incubating it in RNAqueous lysis buffer (Invitrogen, Life Technologies, Zug, Switzerland) at 42°C for 30 min. The resulting RNA was purified and processed as previously described. RNA integrity was measured with the Fragment Analyzer system (Agilent Technologies, Basel, Switzerland).

### Immunostaining

Immunofluorescence staining was performed on 20-µm free-floating slices as previously described (Merienne et al., 2017). The slices were washed three times, for 10 min each, in 1 x PBS and blocked by incubation for 1 h in 1 x PBS supplemented with 10% normal goat serum (NGS)(Interchim, Montluçon, France) and 0.1% Triton X-100 (Sigma-Aldrich, Buchs, Switzerland). The following primary antibodies were used: rabbit polyclonal anti-GS antibody (1/1000, Sigma-Aldrich, Buchs, Switzerland), rabbit polyclonal anti-Iba1 antibody (1/1000, Wako Chemicals GmbH, Instrumenten Gessellschaft, Zurich, Switzerland), and rabbit polyclonal anti-pENK antibody (1/1000, Neuromics, Edina, MN, USA). The following day, slices were washed three times, for 10 min each, in 1 x PBS and incubated for 2 h at room temperature with goat anti-rabbit IgG AlexaFluor-488 (Invitrogen, Life Technologies, Zug, Switzerland), goat anti-rabbit IgG AlexaFluor-594 (Invitrogen, Life Technologies, Zug, Switzerland), goat anti-mouse IgG AlexaFluor-488 (Invitrogen, Life Technologies, Zug, Switzerland), or goat anti-mouse IgG AlexaFluor-594 (Invitrogen, Life Technologies, Zug, Switzerland), diluted in 1 x PBS-5% NGS-0.1% Triton X-100. Slices were then washed three times, for 10 minutes each, in 1 x PBS and mounted on Superfrost+ slides, in Vectashield Fluorescence mounting medium with DAPI. Images showing the specificity of transgenic mice were acquired on an inverted Zeiss LSM 510 confocal microscope with LSM software (Zeiss, Carl Zeiss Microscopy GmbH, Göttingen, Germany). Images showing the loss of eGFP following direct freezing and the preservation of eGFP fluorescence with partial 4% PFA perfusion or our new procedure were acquired with a Zeiss AxioVision microscope (Zeiss, Carl Zeiss Microscopy GmbH, Göttingen, Germany).

### Brain processing and laser-capture microdissection (LCM)

Brains were sliced at 35-µm (iSPN and dSPN) or 20-µm thickness (astrocytes and microglia), at a speed of 2 and a frequency of 10 (arbitrary units) in cold oxygenated aCSF (85 mM NaCl, 75 mM sucrose, 2.5 mM KCl, 25 mM NaHCO<sub>3</sub>, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 3.5 mM MgSO<sub>4</sub>, 0.5 mM CaCl<sub>2</sub>, 10 mM Glucose, Sigma-Aldrich, Buchs, Switzerland) with a semi-automated vibratome (Leica VT1000 S, Biosystems Switzerland, Nunningen, Switzerland). Mounted slices were maintained at room temperature (RT) in a dark environment before dehydration with the following protocol: 65% (astrocytes and microglia) or 60% (iSPN and dSPN) ethanol (Sigma-Aldrich, Buchs, Switzerland) in aCSF for 20 s with up and down shaking, 100% ethanol 20 s with up and down shaking, butanol:ethanol 25:1 for 90 s with up and down shaking (Sigma-Aldrich, Buchs, Switzerland), and xylene:butanol 25:1 for 60 s with up and down shaking (Sigma-Aldrich, Buchs, Switzerland). After dehydration, slides are stored inside a desiccator connected to a vacuum pump. The LCM area and all materials were cleaned with RNasezap before each procedure (Sigma-Aldrich, Buchs, Switzerland). For convenience, three slides were processed in parallel in series of ~30 minutes. During cell isolation, the remaining slices were maintained in a desiccator connected to a vacuum pump. Cells from three slices were isolated on three different CapSure XS LCM caps during the 30 min of serial processing (Arcturus, Bucher, Basel, Switzerland). At the end of the dissection (or 30 min), an Arcturus alignment tray was placed over each of the three caps, and 20 µL RNAqueous Micro kit lysis buffer added to the middle of the cap (Ambion, Life Technologies, Buchs, Switzerland). A 0.5 mL GeneAmp thin-walled reaction tube (Applied Biosystems, Life Technologies, Buchs, Switzerland) was plugged into the cap tray and the three caps incubated at 42°C in an Arcturus Incubation Block for 30 min. During the incubation, three other slices were processed. At the end of the extraction, the tubes linked to caps were inverted and briefly centrifuged to recover the lysis buffer containing total RNA. The caps were then discarded and extracted RNA directly frozen at –80°C before RNA. The same procedure was repeated for all slices from each animal.

### RNA purification and quality check

All samples were purified at the same time after LCM dissection of all animals. Given the high number of independent tubes (up to 80 single GeneAmp tubes per animal), samples from the different cell types were randomized and extracted in two slots to avoid batch effects. All surfaces and materials were cleaned with RNasezap (Sigma-Aldrich, Buchs, Switzerland) and DNazap (Ambion, Life Technologies, Buchs, Switzerland) to avoid RNase and exogenous DNA contamination, respectively. Each tube contained 20  $\mu$ L total RNA from one slice of each animal. Given the maximal volume allowed in RNAqueous Micro kits for LCM purification columns, 7–8 tubes were pooled together on the same column, limiting the number of columns to  $\sim$ 10 for each animal. The rest of the procedure was performed according to the manufacturer's recommendations (Ambion, Life Technologies, Buchs, Switzerland). Final RNA preparations were eluted in  $2 \times 10$   $\mu$ L elution buffer and treated with DNase I, as described in the kit instructions (Ambion, Life Technologies, Buchs, Switzerland). RNA samples from the same animal were then pooled following DNase I treatment.

RNAqueous Micro kit wash solutions contain guanidine thiocyanate, which is known to be eluted with RNA and further contaminate the RNA samples. We therefore performed an additional purification step to remove the remaining chemical contamination. After DNase I treatment of all samples (13 final tubes), four volumes of 100% cold ethanol (Sigma-Aldrich, Buchs, Switzerland) and a 1:10 volume of 3 M sodium acetate (Ambion, Life Technologies, Buchs, Switzerland) were added to each total RNA sample. No carrier was added to avoid inhibition reverse transcriptase. Tubes were vortexed and incubated overnight at  $-20^{\circ}\text{C}$ . The following day, suspensions were centrifuged at  $21,000 \times g$  at  $4^{\circ}\text{C}$  for 30 min. The supernatant was removed and the RNA pellet washed in 500  $\mu$ L 80% cold ethanol and centrifuged for 10 min at  $21,000 \times g$  at  $4^{\circ}\text{C}$ . The procedure was repeated once. After the second wash, RNA pellets were dried under the hood and resuspended in 50  $\mu$ L RNase/DNase-free water (GIBCO, Life Technologies, Buchs, Switzerland). Total RNA was kept at  $-80^{\circ}\text{C}$  until use. The RNA concentration and quality was evaluated using an Agilent Fragment Analyzer apparatus as described in the manufacturer's procedure (Agilent Technologies, Basel, Switzerland). Reverse-transcription and amplification of total RNA were performed with the NuGen Ovation kit following the supplier recommendations (NuGen, San Carlos, California, USA).

### Library preparation and sequencing

Reverse-transcription of extracted RNA was performed with the NuGene Ovation v2 kit in small quantities as described in the manufacturer's recommendations with an input of 10 ng total RNA (NuGen, San Carlos, California, USA). Reverse-transcribed cDNA was then used for sequencing library construction with the Illumina TruSeq Nano kit as indicated in the supplier's recommendations (Illumina Switzerland GmbH, Zurich, Switzerland). One hundred base-pair single-stranded reads were sequenced on an Illumina HiSeq 2500 system with multiplexing across three sequencing lanes. Sequencing was performed to obtain at least  $5 \times 10^7$  reads for each sample. Gokce et al. demonstrated that  $1 \times 10^6$  to  $5 \times 10^6$  reads are sufficient to detect most expressed genes (Gokce et al., 2016). Read quality and demultiplexing was performed with CASAVA software to generate .fastq files for each sample (Illumina Switzerland GmbH, Zurich, Switzerland).

### miRNA microarray

One hundred nanograms total RNA of each sample was labeled according to the instructions of the Agilent miRNA Complete Labeling and Hybridization Kit (Agilent Technologies, Basel, Switzerland). The labeled RNAs were hybridized to Agilent Mouse miRNA Microarrays (Release 19.0, 8x60K; Agilent Technologies, Basel, Switzerland) with 1,247 mouse miRNAs represented, for 20 h at  $55^{\circ}\text{C}$  with rotation. After hybridization and washing, the arrays were scanned with an Agilent microarray scanner using high dynamic range settings as specified by the manufacturer. Agilent Feature Extraction Software v 10.7.3.1 was used to extract the data. Samples were processed on two different microarray slides at six-month intervals, with one sample present on both batches to be used for batch-effect correction. Samples were quantile normalized based on summarized expression levels of probe sets. DE miRNAs were determined using a moderated t test implemented in the R Bioconductor package limma. P values were adjusted using the Benjamini-Hochberg method, controlling for the false discovery rate.

### RT-qPCR

NuGen reverse-transcribed cDNA was used to confirm RNaseq expression levels for some genes by RT-qPCR. One nanogram of cDNA from each sample was used with the KAPA SYBRFAST kit with primers described in the Key resource table (KAPA Biosystems, Labgene, Chatel-Saint-Denis, Switzerland). We then performed real-time quantitative PCR using a Rotor gene (QIAGEN, Basel, Switzerland) and the following cycle parameters:  $95^{\circ}\text{C}$  for 180 s, then 40 cycles of  $95^{\circ}\text{C}$  for 3 s and  $60^{\circ}\text{C}$  for 10 s. Expression levels were normalized to the expression values of cyclophilin A (PPIA), as previously described (Merienne et al., 2017). The following formula was used to determine the relative normalized expression level of target genes in each sample:

$$E = \left( \frac{2^{-Ct(GOI)}}{2^{-Ct(PPIA)}} \right) + 1$$

We then computed the linear regression between  $\log_2(E)$  and  $\log_2$  normalized and transformed RNaseq expression levels to determine the reliability between the RT-qPCR and RNaseq measurements.

## QUANTIFICATIONS AND STATISTICAL ANALYSIS

### RNaseq and microarray data processing

RNaseq raw data quality was first evaluated with FastQC\_0.11.1 software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were trimmed with Cutadapt 1.4.1 (<https://cutadapt.readthedocs.io/en/stable/index.html>) and PrinSEQ (<http://prinseq.sourceforge.net/index.html>) to remove low-quality bases and adapters, and the quality reassessed with FastQC. Alignment was performed with Tophat 2.0.12 aligner (<https://ccb.jhu.edu/software/tophat/index.shtml>) on the mouse mm10 genome release. Aligned reads from each sample were converted to BAM files with SAMtools 0.2.0 suite (<http://samtools.sourceforge.net/>) and used for gene-count measurements with HTSeq 0.6.1 software and with the *Mus musculus* GRCm38.76 release for annotation ([https://htseq.readthedocs.io/en/release\\_0.11.1/](https://htseq.readthedocs.io/en/release_0.11.1/)) The resulting count file was then imported into R software for further analysis. Post-annotation and count quality control (e.g., RNA biotype distribution) were performed with R software using custom scripts or the Bioconductor NOISeq package (<https://www.bioconductor.org/>). Data were filtered to keep only detected mRNA (16,058) with at least one correctly aligned read.

Microarray fluorescence signals were measured with the Agilent Technologies features extraction software (Agilent Technologies, Basel, Switzerland). Expression files for each animal were imported into R software and merged together. Fluorescence background and control probes were subtracted and detected miRNA (371) were filtered with custom scripts.

### DE gene analysis

RNaseq read counts were log<sub>2</sub> transformed and normalized using the TMM algorithm from the edgeR package. The resulting data were transformed with the voom function from the limma package to fit a binomial-like distribution and the mean-variance bias was determined for further correction.

MicroRNA expression levels were log<sub>2</sub> transformed and normalized with the quantile normalization from the preprocessCore package (<https://github.com/bmbolstad/preprocessCore>). In both cases, the impact of normalization and transformation was determined by comparing the data distribution before and after processing with the custom scripts. Correct data distribution was essential before performing the next steps. We used principal component analysis (PCA) and hierarchical clustering to cluster samples based on their expression levels. PCA was performed on all detected samples (16,058 mRNAs and 371 miRNAs) using the *prcomp* function with unscaled parameters. The data representation of the PCA was generated using the *ggplot2* package (<https://ggplot2.tidyverse.org/>). Euclidean distances between the cell types were determined for the top 20% mRNA and miRNA with the highest variation coefficients across cell types and plotted using the *heatmap.2* package. This analysis allowed us to detect a batch effect between the two different microarray slides for miRNA measurements. We thus used the *limma* package (*duplicateCorrelations* function) to create a linear model taking into account cell population membership and the slide batch to correct expression values. Hierarchical clustering and PCA after correction revealed a strong reduction of the batch effect and were therefore used for data representation.

For both mRNA and miRNA profiles, a linear model was created to determine group membership of the samples (dSPN, iSPN, astrocytes, microglia) with the *limma* package (Ritchie et al., 2015). We determined the transcripts enriched in one cell population compared to all the others with the following contrasts:

$$dSPN = dSPN - \frac{(iSPN + Astrocytes + Microglia)}{3}$$

$$iSPN = iSPN - \frac{(dSPN + Astrocytes + Microglia)}{3}$$

$$Astrocytes = Astrocytes - \frac{(iSPN + dSPN + Microglia)}{3}$$

$$Microglia = Microglia - \frac{(iSPN + dSPN + Astrocytes)}{3}$$

Given that dSPN and iSPN are two members of the same neuronal class, we performed an additional contrast to compare the SPN with glial cells:

$$SPN = \left( \frac{dSPN + iSPN}{2} \right) - \left( \frac{Astrocytes + Microglia}{2} \right)$$

The resulting data were transformed following the empirical Bayes (eBayes) algorithm and a fold change (in log<sub>2</sub>, logFC) and p value were computed for each gene for each contrast. We applied the Benjamini-Hochberg (BH) correction to control for false-discovery rates (FDR) across the detected transcript. Considering that the four cell-type specific contrasts are from a single biological question



and therefore performed in a single function, we also corrected for the FDR due to the repetition of these four statistical tests. We finally selected DE genes with a FDR threshold  $< 0.1$  for cell type-specific contrasts and  $0.05$  for SPN/Glia contrast. Venn diagrams showing the number of DE genes were generated with the *vennDiagram* function from *limma* package. Volcano plots for SPN/Glia contrasts were manually built with the *ggplot2* package.

### Evaluation of cell markers, stress and strain effect

We evaluated the specificity of known cell markers. Based on the literature, we manually selected validated markers for each cell type. We also selected some markers for non-dissected cells present in the striatum, i.e., oligodendrocytes, endothelial cells, and striatal interneurons. For each gene marker, we plotted the enrichment score ( $\text{abs}(-\log_{10}(\text{FDR}))$ ) adjusted for the sign of  $\log_{2}(\text{FC})$  for each contrast to confirm the presence/absence of markers in the appropriate cell population.

For the comparison with [Gokce et al. \(2016\)](#), top cell-type enriched markers were identified and used to represent the normalized average expression levels in our four cell populations.

We used mice crossed on a C57BL/6 strain background, except for the *Drd2*-eGFP transgenic mice, which are on an FvB background. We used the PCA performed during the exploratory analysis on mRNA and miRNA to evaluate the impact of mouse strain on expression levels. Variance scree plots showed most of the variance to be explained by the four first components. We therefore pairwise plotted each principal component and demonstrated that no sample clustering occurred based on the mice being of the C57BL/6 or FvB strain.

For the evaluation of the impact of our method on cellular stress, we used the list of genes supplied with the Mouse Stress and Toxicity RT2 Profiler PCR array (QIAGEN). We used RSEM tool to compute FPKM from our raw data and compared the FPKM values with expression values obtained by Zhang and collaborators ([Zhang et al., 2014](#)). An heatmap was made to compare the expression levels of the selected genes between LCM and FACS method.

### Sample correlation and comparison with the Allen Brain Atlas

Following RNaseq data preprocessing, we computed the pairwise Pearson correlations between each sample for the expression levels of all detected mRNAs to determine the reproducibility between our samples. We then used the *heatmap.2* package for graphical representation of correlations, without sample clustering.

We next evaluated the reproducibility of our method with published transcriptome studies. We separated the detected genes into low, medium, and highly expressed genes based on the quartile distribution of averaged  $\log_{2}$  transformed and normalized expression levels between all cell populations. We randomly selected 100 genes of each of three categories and compared their expression by linear regression with probe intensity data obtained by the Allen Brain Study in mouse striatum to demonstrate the reproducibility of our database with already published expression profiles (<http://mouse.brain-map.org/>).

### Network analysis of cell-type specific mRNA-seq data

The raw sequencing reads were pre-processed to discard both adaptor sequences and low-quality reads using the Trimmomatic. The filtered reads were then mapped to the mouse reference genome mm10 (UCSC 2013 release) using Bowtie2 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>). Read distributions were calculated using the featureCounts tool. We used DESeq2, available as an R package in Bioconductor ([www.bioconductor.org](http://www.bioconductor.org)), to generate the log fold-change ( $\log_{2}(\text{FC})$ ) values. The ratio used to generate the  $\log_{2}(\text{FC})$  values is the ratio between gene expression levels in individual cell types and the average gene expression levels across cell types. We generated the networks that describe each cell type by performing spectral decomposition of the RNA-seq signal ( $\log_{2}(\text{FC})$  values) against the probabilistic functional network Mousenet v2, as previously described ([Tourette et al., 2014](#)). Briefly, this approach maps entire gene expression datasets onto Mousenet v2 to identify sub-networks of closely-deregulated genes, reducing the risk of false positives and false negatives in raw gene expression data and highlighting bio-topological essentiality and centrality in these data. We then generated artificially-permuted datasets by randomly permuting the pairs of  $\log_{2}(\text{FC})$  values from the original  $\log_{2}(\text{FC})$  distribution to assess the specificity of the sub-networks generated using spectral decomposition of the signal. Using this approach, the global distribution of  $\log_{2}(\text{FC})$  values in permuted datasets remained the same compared to the original data. The number of permuted genes was set from 10% to 100% by 10% increments. For each permutation, we constructed the corresponding sub-networks and compared them to the original sub-networks by assessing the number of edges in common between the original and permuted sub-networks. Such permutation analysis showed the sub-networks generated using Mousenet v2 for spectral decomposition of the signal to be specific to the RNA-seq datasets analyzed herein (data not shown). We explored the biological content of the network signature obtained for each cell type using the Enrichr database, selecting for GO annotations (Biological processes, KEGG pathways) and Reactome annotations. Annotations supported by a minimal number of nodes set to  $N > 9$  and a P value  $< 10^{-6}$  were considered optimal. These analyses were complemented with PubMed searches.

### miRNA clustering

We selected miRNAs specifically enriched in SPN or glia from the SPN-glia contrast. We assessed the overrepresentation of miRNA clusters in these two cell populations to determine whether a common transcriptional control could explain cell-type enrichment of the miRNA. We first determined the distribution of all detected miRNAs on chromosomes, to exclude chromosomal detection bias, by plotting the miRNA density of each chromosome. We performed the same analysis with DE miRNA, between SPN and glia and

observed overrepresentation of two chromosomes between these two cell types, possibly revealing miRNA cluster enrichment. We then retrieved the composition of known miRNA clusters from the miRBase website and imported the information in R (<http://www.mirbase.org/>). For each DE miRNA, we determined their miRNA cluster membership and performed a hypergeometrical analysis (*phyper* function) to evaluate the overrepresentation of the clusters. Graphical representation of miRNA cluster enrichment was performed by plotting the number of miRNAs belonging to each cluster detected for the DE miRNA. Representations of the genomic positions of the miRNAs of chromosomes 12 and 2 for SPN and glia, respectively, were generated using the *Gviz* package. For both graphs, the positions of known miRNAs in the analyzed genomic window were recovered from *biomaRt* (upper part) and cell-type enriched miRNAs were identified based on our statistical contrasts by plotting their expression levels in each group.

### mRNA regulation by DE miRNA

We selected DE mRNAs and miRNAs from the SPN-glia contrast and hypothesized that if miRNAs strongly contribute to the cellular identity, target mRNAs of DE miRNA should be decreased in the cell population in which the miRNA is enriched. We manually recovered the known target sequence of SPN- and glia-enriched miRNA from the miRBase website. Following duplicate removal, we crossed this list of target mRNA with SPN- and glia-enriched mRNA. For each group of DE miRNAs (i.e., SPN and glia miRNA), we obtained an overlap between their potential target sequences and mRNAs enriched in SPN or glia and performed the Chi-square test to determine whether a significant overlap occurred. We represented overlaps with the *vennDiagram* function of the *limma* package.

### Analysis of TF binding sites on DE genes

For each contrast, significantly DE genes were selected as described above and enrichment for TFs binding sites (TFBS) assessed with the oPOSSUM\_3.0 tool (<http://opossum.cisreg.ca/oPOSSUM3>). Briefly, oPOSSUM 3.0 automatically recovered the genomic sequences of each gene of interest in the Ensembl database of *Mus musculus*. The sequence of nucleotides localized at  $\pm 5,000$  bp of a transcription start site was next recovered and used to predict potential binding sites for TFs, according to the position-weighted matrices of the CORE Jaspar library. For TFBS prediction, a minimum profile specificity of 8 bits was used, together with a conservation cut-off of 0.40 and a matrix score threshold of 85%. Finally, for each TF, a Z-score was calculated, which represents the enrichment of binding sites for a given TF across all genes of interest compared to background. The background corresponds to all genes detected in the LCM database (16,058 genes). The Z-score is an estimation of the likelihood that the number of TFBS nucleotides detected for the target genes is significant relative to the number of TFBS nucleotides detected for the background set. A Z-score  $\geq 10$  was considered to be significant based on the author's recommendations. In addition, for each group, an equivalent number of genes was randomly selected among the 16,058 detected genes to compare with a random distribution. The genomic localization, exact sequence, and distance to transcription start sites were extracted for each TFBS, according to the Ensembl database, and incorporated into R for further analysis.

For each group, the number of enriched TFBSs was plotted using the *ggplot2* package. The random distribution used for comparison was obtained by taking the mean of the data generated from the randomly selected genes for each group. For each contrast, the association between enrichment of the TFBSs and enrichment of the associated TFs was evaluated using a linear regression model to determine whether the TFs strongly associated with DE genes were themselves enriched in the same cell population.

We performed an anchored combination site analysis (aCSA) using oPOSSUM 3.0 to detect over-represented combinations of TFBSs in genes of interest. We selected Sox9 as the anchoring TF and the same parameters and background as in the single-site analysis. Here, oPOSSUM first evaluates all potential TFBSs for the anchoring TF on the genes of interest, and next predicts all TFBS from the CORE Jaspar library located  $\pm 100$  bp from the binding site of the anchoring TF. Finally, a Z-score is calculated that represents the probability of significant enrichment for binding sites of specific TFs close to the anchoring TF. The genomic localization, exact sequence, and distance to transcription start sites were extracted for each TFBS, according to the Ensembl database, and incorporated into R for further graphical representation, as previously described.

### Evaluation of the role of DNA conformation on DE genes

Genomic coordinates of characterized DNA topological adjacent domains (TADs) were recovered from a previous study (Dixon et al., 2012). TADs identified in the mouse cortex samples were used in this analysis. In parallel, the genomic coordinates of DE genes for each contrast were recovered and each gene was associated with an identified TAD, based on overlap of the genomic coordinates. An equivalent number of randomly selected genes were selected among the 16,058 detected genes for each contrast as a negative control for comparison.

The total number of TADs associated with cell-type-enriched genes were plotted using the *ggplot2* package and compared with number of TADs identified with the randomly selected genes. Enrichment for TADs in SPN and glia samples was assessed with the *phyper* function. Finally, Venn diagrams were constructed to evaluate the overlap of the TADs between SPN and glia, and compared with that of the randomly selected genes from these populations using Fisher's Exact test.

EPU were extracted from the same study using the same method, based on genomic coordinates. The same analysis was used to associate DE genes to EPUs for the SPN-glia contrast. The overlap of EPUs between the two cell populations was compared using Venn diagrams and Fisher's Exact test. Randomly selected genes were used as a control to show that the overlap between SPN and glia EPU was less for cell-type enriched genes. Finally, for each cell type, EPUs containing  $> 1$  DE gene were identified and the

selected genes were used for an oPOSSUM analysis to determine whether TFBSs potentially regulating these EPU displayed a typical TFBS profile, as identified in the previous analysis. Graphical representations were generated using the *ggplot2* package.

### Cross-analysis with HD expression profiles

We compared our cell-type enriched expression profiles with a published database of transcriptional alterations in HD patients (Hodges et al., 2006). Based on the publication, we selected significant genes with a p value < 0.001 and classed them as increased (logFC > 0) or decreased (logFC < 0) in HD. We next converted the RefSeq IDs of significant probes to ENSEMBL IDs in each group using *biomaRt*. For probes without a RefSeq ID, we used the HUGO gene symbol for conversion to a RefSeq ID and, finally, to an ENSEMBL ID. Probes without RefSeq or HUGO gene IDs were discarded. Human ENSEMBL IDs from each group were then converted to mouse ortholog ENSEMBL IDs using *biomaRt*. As two microarray slides were used, we removed the duplicated genes at the end of the conversion to keep only single DE genes in HD with murine IDs. Finally, we crossed these two lists with those of cell-type-enriched mRNAs to determine cellular enrichment of increased and decreased genes in HD. Volcano plots were built using the *ggplot2* package to represent the proportion of altered genes in HD patients enriched in SPN or glia in our study. Heatmaps were generated with the *heatmap.2* package, as described above. Finally, we grouped potential cell-type enriched genes in HD by GO with the *goseq* package, as previously described in the package vignette, with a FDR threshold of 0.1. Treemaps were created using the REVIGO graphical tool (<http://revigo.irb.hr/>).

### ChIPseq data analysis

For human samples, libraries were prepared using the NEXTflex ChIP-Seq Kit (V12.10 2012) and sequenced on an Illumina HiSeq 2500 sequencer as single-end 50 bp reads following the manufacturer's instructions. Image analysis and base calling were performed using RTA v1.18.61 and CASAVA v1.8.2. Adaptor dimer reads were removed using DimerRemover v0.9.2 (<https://sourceforge.net/projects/dimerremover/>). For the mouse samples, library preparation was described in (Achour et al., 2015). For human samples, reads were mapped to the human genome (hg19) using Bowtie1 v1.0.0 with the following arguments: -m 1-strata-best -y -S -l 40 -p 2. Peak calling was performed using the SICER.sh script of SICER v1.1 with the following parameters: species: hg19, threshold for redundancy allowed for chip reads: 1, threshold for redundancy allowed for control reads: 1, window size: 200 bps, effective genome size as a fraction of the reference genome of hg19: 0.77, gap size: 600 bps, E value for identification of candidate islands that exhibit clustering: 1000, false discovery rate controlling significance: 1E-2. Fragment size was set according to the value assessed by Homer v4.7.2 makeTagDirectory. Peaks were annotated using Homer v4.7.2 with annotation extracted from Ensembl v75. For mouse samples, the ChIPseq analysis was described in (Achour et al., 2015). Global clustering of the ChIP-seq data and quantitative comparisons were performed using seqMINER software. We used the core gene coordinates as reference coordinates, defined as the region between TSS -5kb and TSS +5kb; gene lists used as a reference were reduced according to the results of the cell-specificity or differential expression RNA-seq analysis in the mouse. The k-means linear method was used to group genes in six clusters called A-F. Functional analysis of gene lists was performed by searching for enrichment in KEGG and GO annotations with DAVID and GREAT.

### Differential binding analysis

We detected differential regions by generating a meta-peak list composed of the union of all H3K27ac peaks detected by SICER in all patients and control samples with bedtools merge v2.21.0 (Quinlan and Hall, 2010). Striatal meta-peak list was composed of 58,635 non-overlapping H3K27ac peaks. We then counted the number of reads falling into the genomic coordinates of the meta-peaks using bedtools intersect. The meta-peak list was annotated with Homer v4.7.2, associating one meta-peak to a single gene. We summed up read counts of all meta-peaks of a given gene using a custom R script to have one read count value per gene (number of different genes n = 25,808). Read counts were then normalized across samples with the method proposed by Anders and Huber (2010). We finally compared patients and controls using the Bioconductor package DESeq2 v1.6.3 using two factor variables: condition (i.e., patient and ctrl) and gender (i.e., male and female). Increased and decreased H3K27ac regions were selected if their adjusted p value ≤ 0.05. Similar analysis was performed using cerebellar samples.

### DATA AND SOFTWARE AVAILABILITY

The NCBI Bioproject accession number for the RNA-seq data reported in this paper is: PRJNA510761