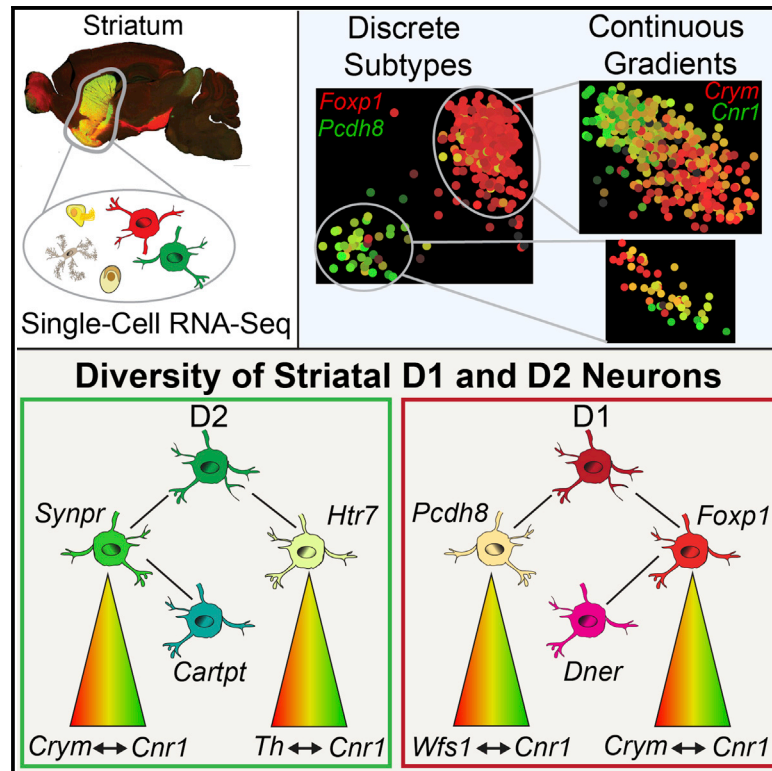# Cell Reports

Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq

Graphical Abstract

Authors

Ozgun Gokce, Geoffrey M. Stanley, Barbara Treutlein, ..., Marc V. Fuccillo, Thomas C. Südhof, Stephen R. Quake

Correspondence

tcs1@stanford.edu (T.C.S.), quake@stanford.edu (S.R.Q.)

In Brief

The striatum, the gateway to basal ganglia circuitry, is critical for motor functions. However, its cell types are incompletely characterized. Gokce et al. reveal the diversity of striatal cells using scRNA-seq. They also describe continuous expression gradients within all MSN subtypes and astrocytes that may be fundamental to transcriptional diversity.

Highlights

- Transcriptomes of 1,208 single striatal cells

- Description of previously unknown medium spiny neuron subtypes

- Discrete cell types that exist in a continuous spectrum of transcriptional states

- Neurons that have the largest transcriptome and more complex splicing patterns

Accession Numbers

GSE82187

CellPress

Cell Reports
Resource

Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq

Ozgun Gokce,[1,7,8] Geoffrey M. Stanley,[2,3,8] Barbara Treutlein,[3,6,8] Norma F. Neff,[3] J. Gray Camp,[6] Robert C. Malenka,[4] Patrick E. Rothwell,[1,4] Marc V. Fuccillo,[1,4] Thomas C. Südhof,[1,5,*] and Stephen R. Quake[3,5,*]

[1]Department of Molecular and Cellular Physiology
[2]Biophysics Program
[3]Departments of Bioengineering and Applied Physics
[4]Nancy Pritzker Laboratory, Department of Psychiatry and Behavioral Sciences
[5]Howard Hughes Medical Institute
Stanford University, Stanford, CA 94305, USA
[6]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
[7]Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-Universität LMU, 81377 Munich, Germany
[8]Co-first author
*Correspondence: tcs1@stanford.edu (T.C.S.), quake@stanford.edu (S.R.Q.)
http://dx.doi.org/10.1016/j.celrep.2016.06.059

SUMMARY

The striatum contributes to many cognitive processes and disorders, but its cell types are incompletely characterized. We show that microfluidic and FACS-based single-cell RNA sequencing of mouse striatum provides a well-resolved classification of striatal cell type diversity. Transcriptome analysis revealed ten differentiated, distinct cell types, including neurons, astrocytes, oligodendrocytes, ependymal, immune, and vascular cells, and enabled the discovery of numerous marker genes. Furthermore, we identified two discrete subtypes of medium spiny neurons (MSNs) that have specific markers and that overexpress genes linked to cognitive disorders and addiction. We also describe continuous cellular identities, which increase heterogeneity within discrete cell types. Finally, we identified cell type-specific transcription and splicing factors that shape cellular identities by regulating splicing and expression patterns. Our findings suggest that functional diversity within a complex tissue arises from a small number of discrete cell types, which can exist in a continuous spectrum of functional states.

INTRODUCTION

The striatum, the gateway to basal ganglia circuitry, is involved in translating cortical activity into adaptive motor actions. Conversely, striatal dysfunction in neuronal and non-neuronal cells contributes to many neuropsychiatric disorders, including Parkinson's and Huntington's disease, schizophrenia, obsessive-compulsive disorder, addiction, and autism (Kreitzer and Malenka, 2008; Maia and Frank, 2011; Robison and Nestler, 2011).

The principal projection neurons in the striatum are the medium spiny neurons (MSNs), which constitute 90%–95% of all neurons in the striatum. The classical model of basal ganglia circuits proposes that MSNs are composed of two subtypes with opposing circuit functions. D1 MSNs preferentially express D1-dopamine receptors and promote movement, while D2 MSNs primarily express D2-dopamine receptors and inhibit movement (DeLong and Wichmann, 2009). Anatomical and functional evidence suggests that this model, while heuristically useful, may need to be modified by incorporating a detailed characterization of the phenotypic diversity of striatal MSNs (Calabresi et al., 2014; Cui et al., 2013; Kupchik et al., 2015; Nelson and Kreitzer, 2014). Previous efforts to characterize striatal diversity have been either low-dimensional, measuring a small number of transcripts in single cells, or reliant on pooling large numbers of striatal cells for bulk RNA sequencing (RNA-seq) and obscuring heterogeneity within the pooled populations (Fuccillo et al., 2015; Heiman et al., 2008; Lobo et al., 2006).

Technological advances in single-cell mRNA sequencing (scRNA-seq) have enabled description of the cellular diversity of tissues and allowed identification of distinct cell subtypes in the developing mouse lung (Treutlein et al., 2014b), the murine spleen (Jaitin et al., 2014), the mouse and human cortex and hippocampus (Darmanis et al., 2015; Zeisel et al., 2015), other neuronal tissues (Pollen et al., 2014; Usoskin et al., 2015), and the intestine (Grün et al., 2015). Here, we use scRNA-seq of 1,208 striatal cells combined with unbiased computational analysis to reconstruct the phenotypic heterogeneity of the striatum.

RESULTS

Identification of Major Striatal Cell Types by Transcriptome Clustering

We measured the transcriptomes of 1,208 single striatal cells using two complementary approaches: microfluidic single-cell RNA sequencing (Mic-scRNA-seq) and single-cell isolation by
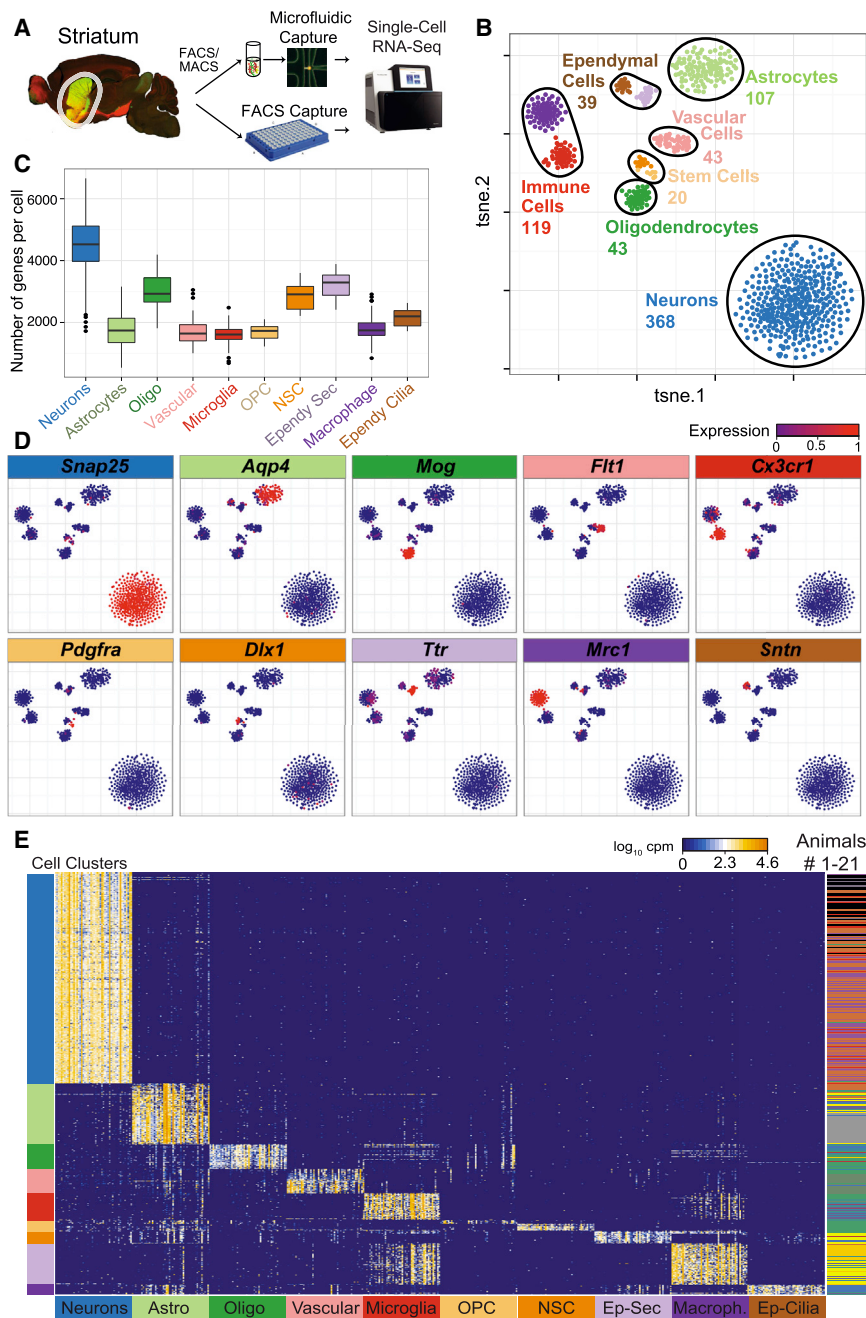
Figure 1. Reverse Engineering of Mouse Striatum by scRNA-Seq

(A) Workflow for obtaining and sequencing cDNA from single cells. Striatal slices from D1-tdTom/ D2-GFP and Aldh1l1-GFP mice were dissociated and cells collected by FACS or MACS. Cells were then captured, imaged, and cDNA amplified in microfluidic chips.

(B) Unbiased clustering of ten major classes of cells using tSNE, which distributes cells according to their whole-transcriptome correlation distance. Each cell is represented as a dot and colored by a clustering algorithm (DBSCAN).

(C) Box-and-whisker plots showing total number of genes detected per cell for major cell types.

(D) Expression of putative marker genes for each of ten major cell types. Scaled expression of marker genes is shown by the color of the cell points. Each tSNE cluster is enriched for one marker, and we were able to assign cells to one of ten major cell types.

(E) Heatmap of the top 50 genes most highly correlated to each cell type. Each row is a single cell, and each column is a single gene. The bar on the right shows the experimental origin of cells. The bar on the left shows DBSCAN clustering of cells, and the bottom bar shows the cell type assignment for each set of 50 genes. Within each 50-gene set, the genes are ordered by increasing the p value of the correlation to that cell type from left to right.

To identify distinct cell populations in Mic-scRNA-seq cells, we visualized cells using 2D t-distributed stochastic neighbor embedding (tSNE) on whole-transcriptome data weighted by a Bayesian noise model (Figure 1B) (Kharchenko et al., 2014; van der Maaten and Hinton, 2008). By examining the expression of known cell type markers, we assigned cell type names to the major clusters (Figures 1B– 1D) (Doetsch et al., 1999; Zhang et al., 2014). Our analysis revealed 368 neuronal cells, 119 immune cells (microglia and macrophages), 107 astrocytes, 43 oligodendrocytes, 43 vascular cells, 39 ependymal cells (ciliated and secretory), and 20 stem cells (Figure S3A). The stem cells were composed of neuronal stem cells (NSCs), likely captured from the rostral migratory stream (Aguirre et al., 2010), and oligodendrocyte precursor cells (OPCs).

The number of expressed genes per cell significantly differed among cell types independent of sequencing depth. Single striatal neurons expressed more genes than all other striatal cell types, suggesting that neurons possess a higher level of functional complexity (Figures 1C and S1E–S1H).

Cells captured in FACS enrichment of Aldh1l1-GFP mice were mostly confined to the astrocyte cluster, and cells from FACS enrichment of Drd1a-tdTom/Drd2-GFP cells were confined

fluorescence-activated cell sorting (FACS-scRNA-seq) (Table S1). We sampled cells either captured randomly or enriched specifically for MSNs or astrocytes using FACS from D1-tdTomato (tdTom)/D2-GFP or Aldh1l1-GFP mice, respectively (Figure 1A) (Heintz, 2004; Shuen et al., 2008). We assessed technical noise, cell quality, and dynamic range using RNA control spike-in standards (Figures S1A–S1D). Saturation analysis confirmed that our sequencing depth of 1×10^6–5×10^6 reads per cell is sufficient to detect most genes expressed (Figure S1E) and that the number of genes detected per cell is independent of the sequencing depth (Figures S1F–S1H).

largely to the neuronal cluster. However, neither tSNE nor whole-transcriptome principal-component analysis (PCA) (Figures 1B and S2), separated D1 MSNs from D2 MSNs within the neuronal cluster, suggesting that their overall gene expression pattern is similar. All identified cell types were sampled in multiple experiments involving a total of 26 mice (Figures 1E and S2; Table S1).

Unbiased Identification of Cell Type-Specific Genes

We identified marker genes for these cell types by the Spearman correlation of genes to cell cluster (Figure 1E; Table S2). NSCs expressed many astrocyte- and neuron-specific genes, further evidence for their proposed origin from astrocytes and eventual fate as neurons.

Striatal neurons show the highest correlation with *Trank1* (also called *Lba1*), whose genetic variants have been associated with bipolar disorder (Mühleisen et al., 2014). The second-highest correlating genes for striatal neurons is *Atp1a3*, which encodes the α3 subunit of Na$^+$/K+-ATPase and is associated with rapid-onset dystonia parkinsonism and alternating hemiplegia of childhood (Heinzen et al., 2014). The highest correlating gene for striatal astrocytes is *Gjb6*, which is required for intercellular trafficking of glucose and its metabolites to neurons (Pannasch et al., 2014) and is associated with non-syndromic autosomal dominant deafness (Rouach et al., 2008). For perivascular macrophages, the highest correlating gene is *Pf4* (also known as *Cxcl4*), which plays a role in coagulation and inflammation (Amiral et al., 1996). Microglia are defined by specific expression of *Tmem119*, a gene that has been proposed to clearly discriminate microglia from perivascular macrophages (Satoh et al., 2016). The highest correlating gene with OPC is *A930009A15Rik*, a gene that has not yet been functionally characterized (Table S2). Finally, ependymal cells were divided into two subtypes: ciliated ependymal cells are correlated with *Spag17*, *Armc4*, *Hydin*, and *Dnali1*, genes that are involved in cilium movement. The second subtype of secretory ependymal cells is correlated with *Npr3*, *Prlr*, and *Slc4a5*, genes that are involved in cellular secretion (Table S2).

Gene Ontology (GO) analysis based on cell type-specific correlating genes found terms associated with the given cell type function, such as synaptic transmission for neurons, vascular biology for vascular cells, cilium movement for ciliated ependymal cells, and cell division or cell-cycle exit for NSCs. GO analysis returned only two terms for OPC and none for secretary ependymal cells, highlighting the limited knowledge of their biology (Tables S3 and S4).

Discrete Subtypes of Striatal MSNs

The neuronal composition of the striatum is defined by a large population of D1- and D2 MSNs and a small population of interneurons (Kreitzer and Malenka, 2008). Using a set of known interneuron and MSN markers, we identified five interneurons, too few to explore their diversity (Figure S3B; Table S5). Using Mic-scRNA-seq, we analyzed a total of 334 MSNs either captured randomly or enriched using FACS (Figures S3C–S3E). The fluorescence signal emitted by each cell was monitored on the microfluidic chip (Figure S3E). To identify MSN subtype-specific genes with an unbiased approach, we used pairwise correlation analysis to identify genes that showed high correlation of expression within a group of genes and strong anticorrelation

to another group of correlated genes. This revealed two sets of genes that are specific to D1- or D2 MSNs (Figure 2A). *Tac1*, *Drd1a*, and *Isl1* are known to be required for D1 MSN differentiation and function (Ehrman et al., 2013; Heiman et al., 2008). The D2-specific gene cluster included the known D2 MSN markers *Drd2*, *Adora2a*, *Penk*, *Gpr6*, and *Gpr52* (Heiman et al., 2008; Lobo et al., 2006), as well as the transcription factor (TF) *Sp9* (Figure 2A). We performed robust principal-component analysis (rPCA) (Todorov and Filzmoser, 2009) using those genes and observed a clearly bimodal distribution of PC1 scores, indicating that the D1/D2 division represented discrete subtypes and was independent of cell isolation methods (Figure S3C).

We confirmed our findings using an independent approach in which we sequenced an additional 570 single MSNs from five mice using FACS-scRNA-seq. We scored MSNs based on their expression of the D1 and D2 marker genes (D1-D2 scores) (Supplemental Experimental Procedures) and observed a clearly bimodal distribution of cells based on these scores (Figures 2B and 2C), indicating the existence of two discrete subtypes: D1- and D2 MSNs (Figure 2D). Some MSNs coexpressed several D1 and D2 markers. We confirmed their existence in the tissue by in situ imaging of the striatum of D2-GFP/D1-TdTom double-reporter mice. We could identify tdTom-GFP double-positive MSNs in both nucleus accumbens and dorsal striatum using fluorescence microscopy (Figures 2E, 2F, and S3F). The rPCA of D1 MSNs revealed a set of genes that separated D1 MSNs into two subpopulations: a majority subpopulation expressing high levels of *Foxp1* and *Camk4* and a minor subpopulation defined by expression of genes such as *Pcdh8*, *Tacr1*, and *Adarb2* (*Pcdh8*-MSN) (Figures 2G–2J; Table S6). The Pcdh8-MSN subpopulation coexpressed the D2 neuropeptide *Penk*, along with *Tac1* (Figure 2H). We scored D1 MSNs based on expression of rPCA-identified genes (*Pcdh8*-Foxp1 scores), which revealed a clearly bimodal population distribution defining two discrete subpopulations (Figure 2I). Many genes enriched in *Pcdh8*-MSNs are related to neurological disorders (*Nrxn2*, *Sema3e*, *Sema4a*, *Sema5a*, *Sema5b*, *Sema6d*, *Pcdh7*, *Pcdh8*, *Ptprg*, *Ptprm*, *Ptpro*, and *Ptpru*) (Redies et al., 2012; De Rubeis et al., 2014; Südhof, 2008), tachykinin signaling (*Tac2* and *Tacr1*) (Steinhoff et al., 2014), and RNA binding (*Elavl4*, *Adarb2*, *Khdrbs3*, *Rbm20*, *Aff2*, *Lrpprc*, and *Celf4*) (Li et al., 2007; Noureddine et al., 2005). *Pcdh8*-MSNs are also depleted of important neuronal genes like *Nlgn1* and *Calb1* (Table S6).

Similar to D1 MSNs, rPCA of D2 MSNs revealed two discrete subpopulations: a small subpopulation marked by unique expression of many genes, including *Htr7* and *Agtr1a*, and the coexpression of D1 and D2 neuropeptides *Penk* and *Tac1* (*Htr7*-MSN) (Figures 2K–2N; Table S6). Genes enriched in the Htr7-MSN subpopulation included receptors *Htr7*, *Ptprt*, *Ngfr*, *Grik3*, *Cacng5*, and *Tmeff2* and TFs *Sox9*, *Sp8*, *Runx1*, *Mafb*, and *Litaf*. *Htr7*-MSNs are significantly depleted of important neuronal genes like *Cacna2d3* and *Synpr* compared to the major D2 population (Table S6) and showed coexpression of the *Gfp* and *tdTom* transgenes.

Continuous Transcriptional Gradients within MSN Subtypes

We continued our analysis of heterogeneity within subpopulations of MSNs. The rPCA on *Foxp1*-high D1 MSNs revealed a
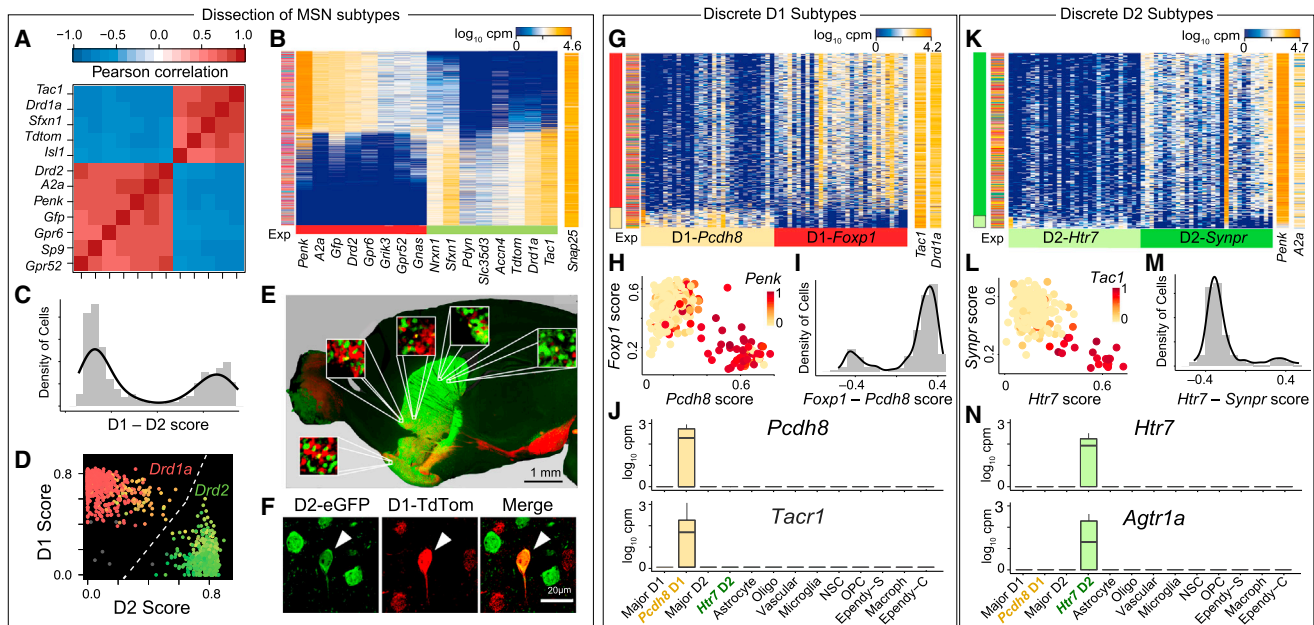
Figure 2. Characterization of Discrete MSN Subtypes

(A) Hierarchical clustering of the highest pairwise gene correlations in MSNs shows two strongly anticorrelated clusters of genes that include known MSN subtype markers for D1- and D2 MSNs.

(B) rPCA of MSNs. Cells (rows) are ordered by their projection onto PC1, and genes (columns) are ordered by their positive (left) or negative (right) contribution to PC1. This identifies three molecularly distinct populations, assigned as D1 MSN (red), D2 MSN (green), and D1/2 hybrid MSN (yellow). The bar on the left shows the experimental origin of cells.

(C) Distribution of single MSNs projected onto D1-D2 scores. The D1-D2 score is calculated by summing the scaled expression values of the genes shown in (B). D1- and D2 MSNs form distinct peaks, and the novel MSNs are distributed between the two peaks.

(D) Biplot of D1- and D2 MSN cells by their expression of D1 genes (y axis) and of D2 genes (x axis). Scaled expression of *Drd1a* is shown by the color of the cell points.

(E) Sagittal brain section of a D2-GFP/D1-TdTom double-reporter mouse showing tdTom-GFP double-positive novel MSNs in both nucleus accumbens and dorsal striatum.

(F) Confocal imaging of striatal slices of D2-GFP/D1-TdTom double-reporter mice demonstrating the existence of D1/2 hybrid MSNs in the striatum.

(G) rPCA of MSNs in the D1 part of (D) using all expressed genes. Cells (rows) are ordered by their projection onto PC1, and genes (columns) are ordered by their positive (left) or negative (right) contribution to PC1. This identifies two molecularly distinct populations assigned as major (*Foxp1*) D1 MSN (red) and *Pcdh8*-MSN (yellow).

(H) Biplot of *Foxp1*-D1 and *Pcdh8*-MSN cells by their expression of D1 genes (y axis) and Pcdh8 genes (x axis). Scaled expression of *Penk* is shown by the color of the cell points.

(I) Distribution of MSNs projected onto *Pcdh8*-*Foxp1* scores. Major-D1 and *Pcdh8*-MSNs form distinct peaks.

(J) Boxplots showing specific Pcdh8 and *Tacr1* expression in *Pcdh8*-MSNs compared to other cell types in MSN.

(K) rPCA of MSNs in the D2 part of (D). Cells (rows) are ordered by their projection onto PC1, and genes (columns) are ordered by their positive (left) or negative (right) contribution to PC1. This identifies two molecularly distinct populations assigned as D2 MSN (green) and *Htr7*-MSN (light green).

(L) Biplot of D2- and *Htr7*-MSN cells by their expression of *Htr7* genes (y axis) and *Synpr* genes (x axis). Scaled expression of *Tac1* is shown by the color of the cell points.

(M) Distribution of MSNs projected onto *Htr7*-*Synpr* scores. Major D2- and *Htr7*-MSNs form distinct peaks.

(N) Boxplots showing specific *Htr7* and *Agtr1a* expression in *Htr7*-MSNs compared to other cell types in MSN.

subpopulation expressing high levels of *Dner*, *Cxcl14*, and *Tnnt1* and lacking *Meis2* (D1-*Dner*) (Figures 3A–3C and S4A; Table S6). Similarly, rPCA on *Synpr*-high D2 MSNs revealed a small population expressing *Cartpt* and *Kcnip1* but lacking *Calb1* (Figures 3G–3H and S4C). The rPCA on the *Meis2*+/*Dner*− subpopulation of D1 MSNs revealed a gradient of transcriptome states (Figure 3D), indicated by the unimodal distribution of cells' *Crym*-*Cnr1* score (Figures 3E and S4B). Some genes defining the opposing gradients are known to be differentially expressed between striatal matrix (e.g., *Cnr1*) and striosome (e.g., *Nnat* and *Gfra1*) microzones (Figures 3D, 3E, and S4B) (Crittenden and

Graybiel, 2011). The rPCA revealed similar unimodal gene expression gradients for the other main MSN subtypes (Figures 3J–3R and S4D). The transcriptional gradients within all main MSN subtypes shared several genes (e.g., *Cnr1*, *Crym*, and *Wfs1*).

Subtypes of Non-neuronal Striatal Cell Types

We next characterized heterogeneity within vascular cells, immune cells, oligodendrocytes, and astrocytes. Within vascular cells, pairwise correlation analysis revealed two large anti-correlated gene groups (Figure 4A). The rPCA using these
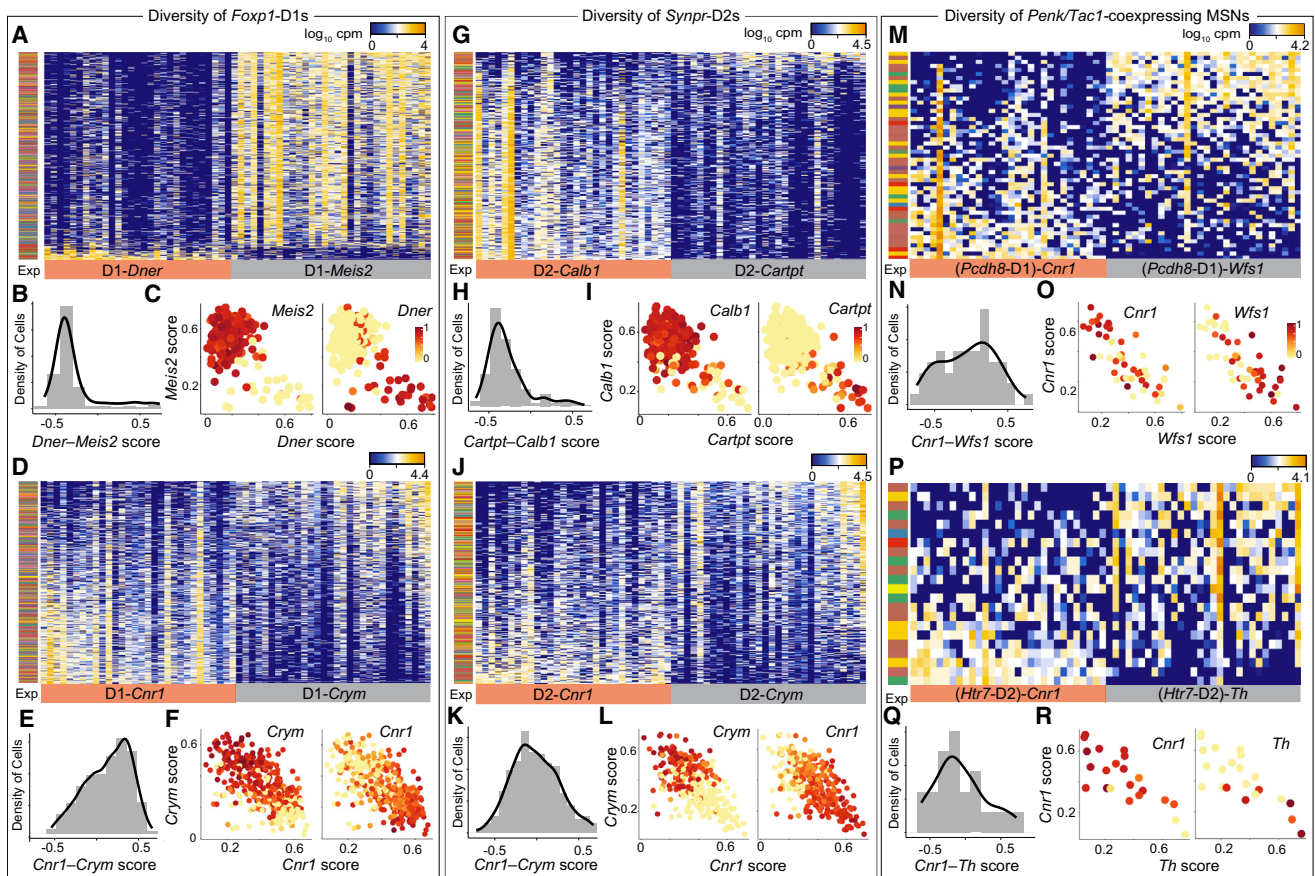
Figure 3. Identification of Heterogeneity within MSN Subtypes

(A) rPCA of major D1 MSNs using all expressed genes, revealing a molecularly distinct subpopulation of D1-*Dner* MSNs. The bar on the left shows the experimental origin of cells.

(B) Distribution of D1 MSNs projected onto *Meis2-Dner* gene group scores.

(C) Biplot of D1- and D1-*Dner* MSNs by their expression of D1-*Meis2* genes (y axis) and D1-*Tnnt1* genes (x axis). Scaled expression of *Meis2* and *Tnnt1* is shown by the color of the cell points.

(D) rPCA of D1 MSNs using all expressed genes, revealing a continuous transcriptional gradient marked by opposing expression gradients of *Cnr1* and *Crym*. The bar on the left shows the experimental origin of cells.

(E) Distribution of D1 MSNs projected onto *Cnr1-Crym* gradient scores.

(F) Biplot of D1 MSNs by their expression of *Crym* gradient genes (y axis) and *Cnr1* gradient genes (x axis). Scaled expression of *Cnr1* and *Crym* is shown by the color of the cell points.

(G) rPCA of D2 MSNs using all expressed genes, revealing a subpopulation of D2-*Cartpt* MSNs. The bar on the left shows the experimental origin of cells.

(H) Distribution of D2 MSNs projected onto *Calb1-Cartpt* gene group scores.

(I) Biplot of D2-*Calb1* and D2-*Cartpt* MSNs by their expression of *Cartpt* group genes (y axis) and *Calb1* group genes (x axis). Scaled expression of *Calb1* and *Cartpt* is shown by the color of the cell points.

(J) rPCA of D2 MSNs using all expressed genes, revealing a continuous transcriptional gradient marked by opposing expression gradients of *Cnr1* and *Crym*. The bar on the left shows the experimental origin of cells.

(K) Distribution of D1 MSNs projected onto *Cnr1-Crym* gradient scores.

(L) Biplot of D1 MSNs by their expression of *Crym* gradient genes (y axis) and *Cnr1* gradient genes (x axis). Scaled expression of *Cnr1* and *Crym* is shown by the color of the cell points.

(M) rPCA of *Pcdh8*-MSNs using all expressed genes, revealing a continuous transcriptional gradient marked by opposing expression gradients of *Cnr1* and *Wfs1* similar to 2D. The bar on the left shows the experimental origin of cells.

(N) Distribution of *Pcdh8*-MSNs projected onto *Cnr1-Wfs1* gradient scores.

(O) Biplot of Pcdh8-MSNs by their expression of *Wfs1* gradient genes (y axis) and *Cnr1* gradient genes (x axis). Scaled expression of *Cnr1* and *Wfs1* is shown by the color of the cell points.

(P) rPCA of *Htr7*-MSNs, revealing a continuous transcriptional gradient marked by opposing expression gradients of *Cnr1* and *Th*. The bar on the left shows the experimental origin of cells.

(Q) Distribution of *Htr7*-MSNs projected onto *Cnr1-Th* gradient scores.

(R) Biplot of *Htr7*-MSNs by their expression of *Th* gradient genes (y axis) and *Cnr1* gradient genes (x axis). Scaled expression of *Cnr1* and *Th* is shown by the color of the cell points.
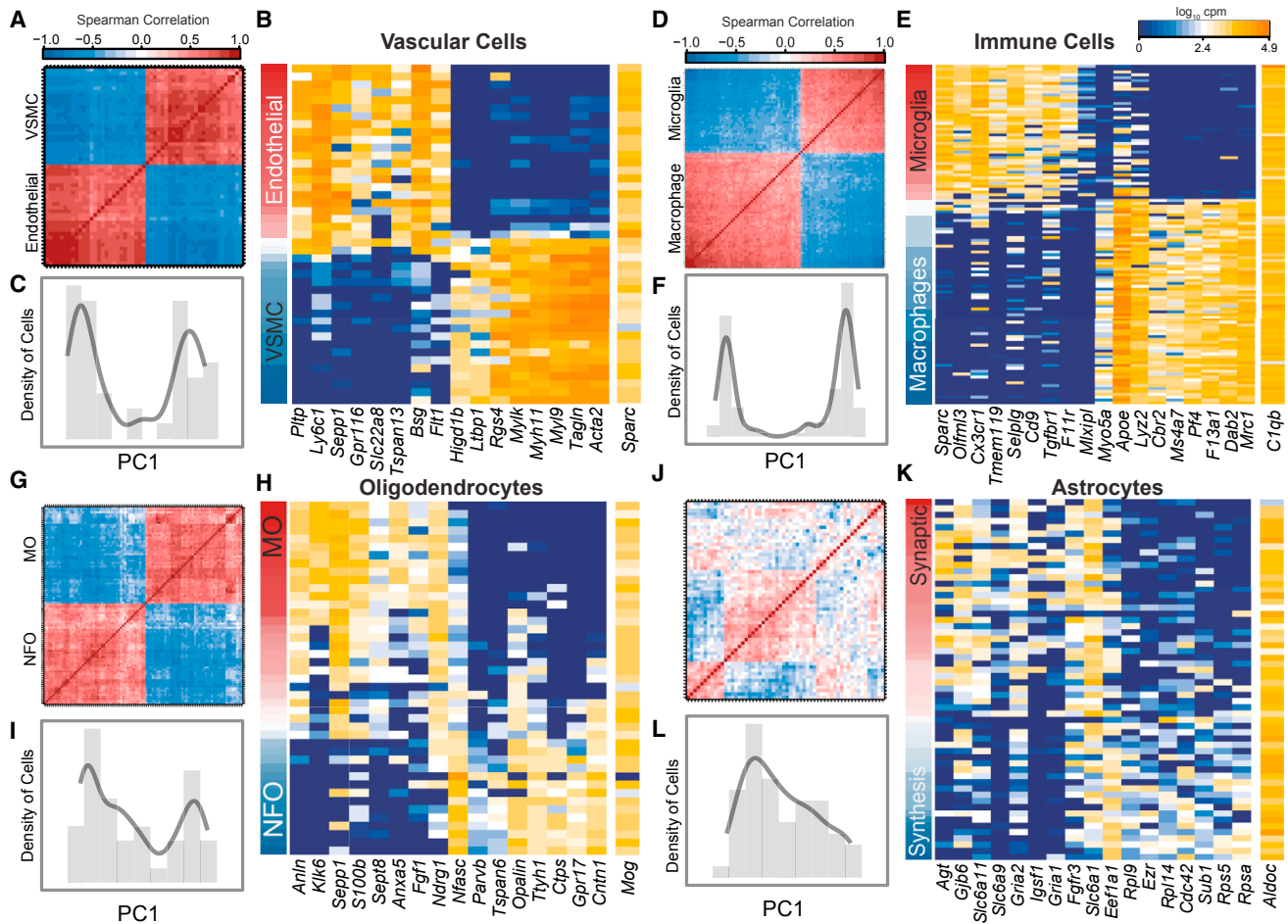
Figure 4. Characterization of Vascular Cells, Astrocytes, and Oligodendrocytes

(A) Hierarchical clustering of the highest pairwise gene correlations within vascular cells reveals two large clusters of subtype-specific transcripts.

(B) rPCA of 43 striatal vascular cells identifies two molecularly distinct populations, assigned to VSMCs that express *Myl9* and *Tagln* and endothelial cells that express *Ly6e* and *Pltp*.

(C) Histogram of vascular cells' PC1 scores shows a clearly bimodal distribution and confirms the existence of two distinct subtypes.

(D) Hierarchical clustering of the highest pairwise gene correlations within immune cells reveals two large clusters of interconnected transcripts.

(E) rPCA of 119 immune cells (microglia and macrophages) identifies two molecularly distinct populations, assigned to microglia that express *Sparc* and macrophages that express *Mrc1*.

(F) Distribution of single immune cells along the first PC1 of rPCA. Cells form two distinct peaks.

(G) Hierarchical clustering of the highest pairwise gene correlations within oligodendrocytes reveals two large clusters of interconnected transcripts.

(H) rPCA of 43 striatal oligodendrocytes identifies two distinct oligodendrocyte populations: MOs, which express *Klk6* and *Sec11c*, and NFOs, which express *Nfasc* and *Ckb*. There are also a significant number of oligodendrocytes with intermediate expression of both sets of genes, which are likely transitioning between NFOs and MOs.

(I) Distribution of single oligodendrocytes along the first PC1 of rPCA. MO and NFO cells form two distinct peaks, with transitioning oligodendrocytes bridging the two peaks.

(J) Hierarchical clustering of the highest pairwise gene correlations within astrocytes did not reveal interconnected transcripts, instead showing low correlation values between genes.

(K) rPCA of 107 single striatal astrocytes reveals transcripts that increase or decrease continuously without defining distinct subpopulations. Cells (rows) are ordered by their PC1 scores, and genes (columns) are ordered by their PC1 loading. The continuum of astrocyte transcriptional states is marked on one side by higher expression of transcripts related to synaptic communication (*Slc6a11*, *Slc6a1*, *Slc6a9*, *Gria1*, and *Gria2*) and on the other side by higher expression of transcripts related to translation (*Rpl9*, *Rpl14*, *Rps5*, and *Rpsa*) and cell polarity regulators (*Cdc42*).

(L) The distribution of single astrocytes along PC1 is unimodal, indicating striatal astrocytes exhibit continuous transcriptional variation within one discrete subtype.

genes revealed two discrete subpopulations: a population of vascular smooth muscle cells (VSMCs) (expression of *Myl9* and *Tagln*), and a population of endothelial cells (expression of *Pltp* and *Ly6c1*) (Figures 4B and 4C) (Zeisel et al., 2015). Likewise, within immune cells, pairwise correlation analysis and rPCA revealed two discrete subpopulations: microglia cells

Figure 5. Cell Type-Specific TFs Identified by Single-Cell Transcriptome Analysis
(A) Correlogram visualizing correlation of single-cell gene expression between cell type-specific TFs.
(B) Scaled expression of the TFs most specific to discrete cell types. Scaled expression of marker genes is shown by the color of the cell points.

expressing *Cx3cr1*, *Olfml3*, and *Mlxipl* and macrophages expressing *Mrc1* and *Dab2* (Figures 4D–4F).

Within oligodendrocytes, we found two large anticorrelated gene groups (Figure 4G). Projecting the cells onto PC1 revealed two discrete oligodendrocyte subtypes: a population of newly formed oligodendrocytes (NFOs) expressing *Nfasc* and a population of mature oligodendrocytes (MOs) expressing *Klk6* (Zhang et al., 2014). These two discrete populations were connected by cells that coexpressed both NFO and MO genes and may represent a transitioning population (Figures 4H and 4I).

The pairwise correlation analysis did not identify two clearly anticorrelated gene groups for astrocytes (Figure 4J); instead, whole-transcriptome rPCA revealed a continuous transcriptional gradient defined by a unimodal population distribution (Figures 4K and 4L). The genes defining this gradient were related on one side to synaptic communication, such as neurotransmitter transporters (*Slc6a11*, *Slc6a1*, and *Slc6a9*) and glutamate receptors (*Gria1* and *Gria2*), and on the other side to translation, like ri-

bosomal proteins (*Rpl9*, *Rpl14*, *Rps5*, and *Rpsa*) and cell polarity regulators (*Cdc42*) (Etienne-Manneville and Hall, 2003).

Distinct Groups of TFs Maintain Cell Type Identity
We used correlation analysis to identify TFs expressed specifically (Spearman correlation coefficient > 0.5) in single cells of a given type and visualized the pairwise correlation of the resulting 69 TFs in a correlogram (Figure 5A; Table S7). We found that striatal neurons correlated with the largest set of TFs (22) (Figure 1C), independent of the number of cells per subtype (Figure S5). The most specific neuronal TF, *Myt1l*, is able to induce neuronal cell identity in vitro when expressed in fibroblasts with two additional TFs (Vierbuchen et al., 2010). Generally, many of the cell type-specific TFs identified by our analysis have been used in TF-based direct conversion protocols. Among them are *Sox9* for astrocytes (Caiazzo et al., 2015), *Sox10* for oligodendrocytes (Yang et al., 2013), *Gata2* for endothelial cells (Elcheva et al., 2014), and *Maf* for macrophages (Hegde et al.,
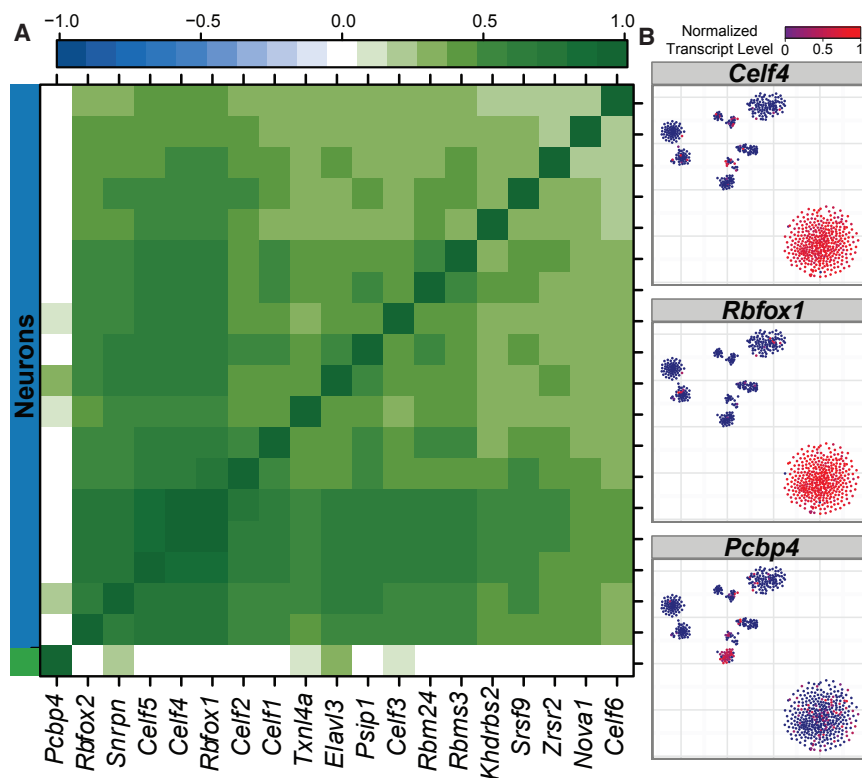
Figure 6. Single-Cell Transcriptome Analysis Reveals Splicing Factors Specific to Neurons

(A) Hierarchical clustering of pairwise gene correlations for cell type-specific splicing factors.

(B) tSNE plot of single cells colored by the scaled expression selected splicing factors: the neuronal-enriched splicing factors *Celf4* and *Rbfox1* and the oligodendrocyte-enriched splicing factor *Pcbp4*.

as Fox-3, HRNBP3, or NeuN). The RNA-binding Fox family of splicing factors has been shown to maintain mature neuronal physiology and is linked to autism (Gehman et al., 2012). *Khdrbs2* (also called SLM1), which may contribute to alternative splicing of neurexins, was also enriched in neurons (Iijima et al., 2014). Many other neuron-enriched splicing factors have not yet been well characterized for their neuronal roles (*Snrpn*, *Txnl4a*, *Psip1*, *Rbm24*, and *Rbm3*). In summary, analysis of cell type-specific splicing factors showed that the regulatory diversity of splicing is higher in neurons than in other cell types in the striatum.

To complement this analysis, we identified 45,843 sites of alternative splicing by the reads mapping to exon junctions at splice sites. Using Fisher's exact test, we defined splice junctions that are differentially spliced ($p < 10^{-5}$) in one or more cell types (Figures 7A and S7A; Table S8). These splice sites included genes such as *Csde1*, whose splicing is dysregulated in schizophrenia and for which a de novo loss-of-function mutation causes autism (Sanders et al., 2012); *Hnrnpk* (heterogeneous nuclear ribonucleoprotein K), which is associated with intellectual disability (Au et al., 2015); and *Hspa5*, which is linked to bipolar disorders and schizophrenia (Kakiuchi et al., 2003). The visualization of cell type-specific alternative splicing revealed many single cells that expressed both the major and the minor variants, a phenomenon we termed compound splicing (Figures 7A and S7A–S7D; Table S8).

It had been previously reported that single immune cells exclusively express either the major or the minor splice variant (Shalek et al., 2013). However, we had observed that individual neurons were capable of expressing both splice versions of neurexins using single-cell qPCR (Fuccillo et al., 2015). Here, we performed a global analysis to quantify and compare genes with compound splice sites by cell type (Figure 7B). We detected few compound splice sites in immune cells of the striatum but higher numbers in single neurons, as well as in NSCs, oligodendrocytes, and secretory ependymal cells. There was no relationship between sequencing depth and number of compound splice sites detected per cell, ruling out systematic differences in sequencing depth as an explanation (Figure S7B). Moreover, 3′–5′ RNA coverage bias was similar across all cell types and between bulk controls and single cells (Figures S1H and S1I).

1999). Thus, we also provide candidate TFs for stem cell transformation into cell types for which there are no existing protocols (*Lmx1a* for secretory ependymal cells, *Foxj1* for ciliated ependymal cells, *Tbx2* for VSMCs, *E2f8* for NSCs, *Mlxipl* for microglia, *Isl1* for D1 MSNs, and *Sp9* for D2 MSNs (Figure 5A; Table S7).

While *Isl1* is known to be required for normal functioning of D1 MSNs (Ehrman et al., 2013), the role of *Sp9* in MSNs had not been described previously. Therefore, we analyzed the effect of *Sp9* overexpression on expression of MSN markers in primary striatal cultures. The overexpression of *Sp9* significantly lowered the expression of *Drd1a* without significantly altering the expression of D1 MSN markers *Tac1*, *Pdyn* or D2 MSN marker *Drd2*. These results suggest that *Sp9* overexpression may disrupt the normal development of D1 MSNs (Figure S7).

Cell Type-Specific Regulation of RNA Splicing

Alternative splicing is a key mechanism for generating transcriptomic diversity. We used correlation analysis (Supplemental Experimental Procedures) to explore cell type-specific expression of RNA splicing factors. We found 18 splicing factors that were specific to neurons and only 1 splicing factor (*Pcbp4*) that correlated with oligodendrocytes (Figures 6A and 6B). No splicing factors were found with high correlation to any other cell types. The factors specific to striatal neurons include all six members of the CUG-BP, Elav-like family (*Celf*), which are associated with autism, epilepsy, and Alzheimer's disease (Ladd, 2013), as well as all three members of the RNA-binding Fox family of splicing factors: *Rbfox1* (also known as Fox-1 or A2BP1), *Rbfox2* (also known as Fox-2 or RBM9), and *Rbfox3* (also known
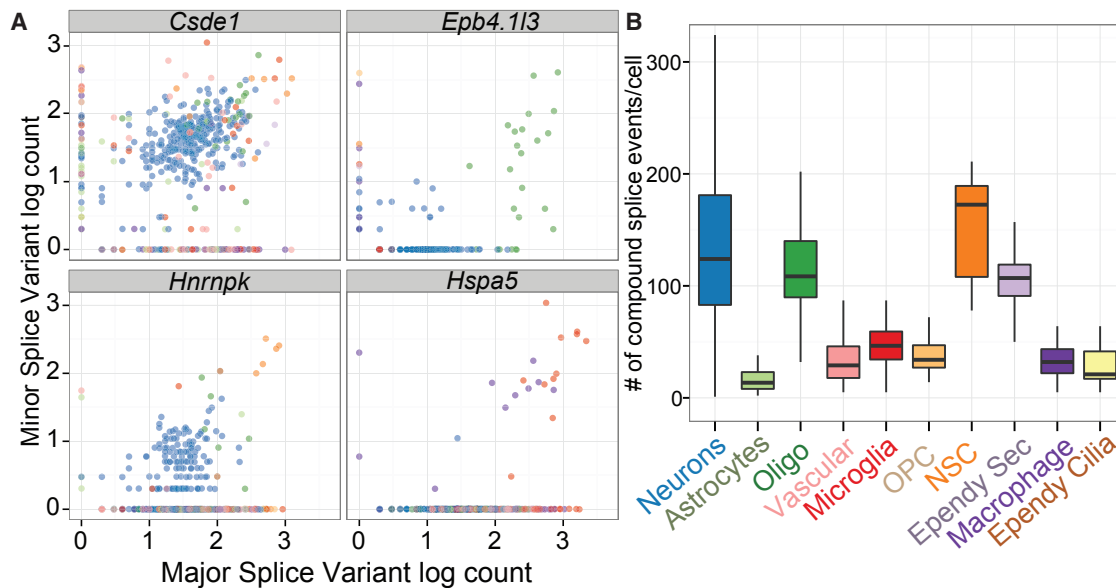
Figure 7. Differential Splicing Analysis Reveals Several Modes of Single-Cell Splicing Regulation

(A) Expression of selected splice sites that have significant cell type-specific regulation. The number of reads per cell ($\log_{10}$) aligning to the major and minor variants is plotted on the x axis and y axis, respectively. Many single cells are located on the diagonal of the plots, indicating they express both the major and the minor variants of that splice site (compound splice sites).

(B) Boxplots showing the total number of compound splice sites detected per cell across cell types.

DISCUSSION

Previous studies on the heterogeneity of striatal MSNs provided various classification schemes, including the classical distinction between D1- and D2 MSN subtypes based on expression of the D1-dopamine receptor versus the D2-dopamine receptor (Kreitzer and Malenka, 2008), the anatomical location of MSNs in dorsal striatum or nucleus accumbens (O'Doherty et al., 2004), and the position of a MSN within a striosome or matrix microzone (Crittenden and Graybiel, 2011).

We analyzed MSNs using scRNA-seq and found that D1 MSNs could be subdivided into two discrete populations: *Pchd8*-D1s and *Foxp1*-D1s. The *Foxp1*-D1 population could be further subdivided into *Foxo1*-high or *Dner*-high populations, although it was not clear how discrete the *Foxo1-Dner* division is. Similarly, we could subdivide D2 MSNs into two discrete subpopulations: *Htr7*-D2s and *Synpr*-D2s. The *Synpr*-D2 neurons could be further subdivided into *Calb1*-high or *Cartpt*-high subpopulations, although as with *Dner*-D1s, it was not clear how discrete this division is. Two marker genes for *Pchd8*-D1 and *Htr7*-D2 subtypes, *Htr7* and *Tacr1*, are linked to alcohol dependence (Blaine et al., 2013; Zlojutro et al., 2011), and repeated cocaine exposure regulates epigenetic control of *Pcdh8*-MSN marker *4932411E22Rik* (Feng et al., 2015). The *Pchd8*-MSN subtype coexpresses *Drd1a* and *Drd2* and therefore may possess atypical signaling and distinct properties. Previous studies also suggested the possible existence of a third MSN population that coexpresses *Drd1a* and *Drd2*. However, none of these studies provided specific markers for this population or resolved them into distinct subtypes (Ade et al., 2011; Bertran-Gonzalez et al., 2010; Frederick et al., 2015; Surmeier et al., 1998). Impor-

tant results of our work are that we have conclusively proved the existence of a MSN subtype that coexpress several D1- and D2-specific genes and shown that there are two distinct populations, each with specific markers. Given the involvement of MSNs in cognitive disorders, novel and discrete MSN subtypes expressing neuronal plasticity and addiction-related genes are likely to have important functions, which can be characterized in future studies using genetic tools based on the markers identified here.

Within the discrete MSN subtypes, we identified a large degree of heterogeneity that did not further separate the cells into clear subtypes. We termed this a continuous transcriptional gradient, because the neurons had a range of expression of the opposing gene sets and a roughly unimodal expression distribution. The continua within all main subtypes shared several genes (particularly *Cnr1*, *Crym*, and *Wfs1*), suggesting a common origin of these gradients. Previous scRNA-seq analyses have largely assumed that the identity of post-mitotic adult neurons is discrete: every neuron can be assigned a single subtype (Darmanis et al., 2015; Tasic et al., 2016; Usoskin et al., 2015; Zeisel et al., 2015). Here, we find that there are at least two distinct aspects to neuronal identity: discrete subtypes, in which two or more groups of neurons are clearly separated with few intermediates, and continuous gradients, in which neurons within a subtype lie on a spectrum of gene expression, with most having a range of intermediate expression values of two gene sets.

We found continuous transcriptional gradients within glial cells as well. Astrocytes have been reported to be a heterogeneous cell type (Hochstim et al., 2008; Matthias et al., 2003). We discovered a continuum (Figures 4J–4L) wherein genes for neurotransmitter transporters and receptors are anticorrelated to genes

coding for ribosomal proteins and the cell polarity regulator *Cdc42*. This may therefore represent a continuum of astrocyte states, ranging from more actively involved in synaptic communication (Matthias et al., 2003) to more inactive or more developmentally immature (Etienne-Manneville and Hall, 2003). We applied our analysis approach to other striatal cell types identified by tSNE. Pairwise correlation analysis of the vascular cells, the immune cells, and the oligodendrocytes identified subtypes and large set of subtype-specific markers (Figures 4A–4L).

We also identified cell type-specific TFs (Figures 5A and 5B; Table S7), which provides a mechanistic explanation for the maintenance of discrete cell type identities. The most cell type-specific TFs we found were often previously observed to fuel direct conversion of cells into that cell type (Caiazzo et al., 2015; Elcheva et al., 2014; Hegde et al., 1999; Pang et al., 2011; Vierbuchen et al., 2010; Yang et al., 2013). We found that the D2 MSN-specific TF identified here first, *Sp9*, disrupted D1 MSN-specific gene expression, confirming the functional relevance of these cell type-specific TFs (Figure S6).

Finally, we observed that neurons contribute to the brain's uniquely high-level alternative splicing more than any other cell type (Grosso et al., 2008; Yeo et al., 2005). We found that neurons express more specific splicing factors and that the overall complexity of alternative splicing is higher in neurons than in other cell types. For many sites of alternative splicing, we detected both variants in each neuron (compound splice sites), revealing an alternative splicing machinery that can increase the diversity of the transcriptome by altering alternative splicing ratios.

Our results show that the phenotypic diversity of striatal neurons arises from a small number of discrete subtypes, within which neurons lie on a continuum of gene expression states. Our analysis methods distinguish between discrete subtypes (with transitioning intermediates when the cells are actively differentiating) and continuous heterogeneity. These distinctions may prove fundamental to understanding the phenotypic diversity of the cells that form complex tissues.

EXPERIMENTAL PROCEDURES

Animals

All procedures conformed to the NIH Guidelines for the Care and Use of Laboratory Animals and were approved by the Stanford University Administrative Panel on Laboratory Animal Care.

Single-Cell Transcriptional Profiling

Acute brain slices were cut from the 5- to 7-week-old male mice and, after papain treatment, dissociated mechanically (Supplemental Experimental Procedures). Live cells were purified by either magnetic bead-activated cell sorting (MACS; Miltenyi) or FACS. For cell type-specific isolation, genetically labeled MSN subtypes D1 MSN and D2 MSN and astrocytes were purified by FACS. Single cells were captured on a microfluidic chip on the C1 system (Fluidigm), and whole-transcriptome-amplified cDNA was prepared on chip using the SMARTer Ultra Low RNA Kit for Illumina (Clontech Laboratories). For the smart-seq2 protocol, three MSN populations (D1 MSN tdTom+, D1 MSN GFP+, and tdTom+/GFP+ MSNs) were sorted individually into 96-well plates with lysis buffer, spun down, frozen at −80°C, and amplified using the protocol described previously (Picelli et al., 2013). Single-cell libraries were constructed as described previously (Supplemental Experimental Procedures) (Treutlein et al., 2014a).

Immunohistochemistry

Vibratome sections from perfusion-fixed mouse brains were stained with antibody for GFP (Rockland) and immunofluorescence imaging done on an Olympus fluorescent microscope or Nikon confocal microscope.

ACCESSION NUMBERS

The accession number for the single-cell sequencing raw data reported in this paper is GEO: GSE82187.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.06.059.

AUTHOR CONTRIBUTIONS

O.G., G.M.S., B.T., T.C.S., and S.R.Q. designed the study. O.G., G.M.S., B.T., and N.F.N. performed the experiments. O.G., G.M.S., and B.T. analyzed datasets. T.C.S. and S.R.Q. provided intellectual guidance in the interpretation of the data. P.E.R., R.C.M., and M.V.F. provided reagents. O.G., G.M.S., B.T., G.J.C., T.C.S., and S.R.Q. wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

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.

Aguirre, A., Rubio, M.E., and Gallo, V. (2010). Notch and EGFR pathway interaction regulates neural stem cell number and self-renewal. Nature *467*, 323–327.

Amiral, J., Wolf, M., Fischer, A., Boyer-Neumann, C., Vissac, A., and Meyer, D. (1996). Pathogenicity of IgA and/or IgM antibodies to heparin-PF4 complexes in patients with heparin-induced thrombocytopenia. Br. J. Haematol. *92*, 954–959.

Au, P.Y.B., You, J., Caluseriu, O., Schwartzentruber, J., Majewski, J., Bernier, F.P., Ferguson, M., Valle, D., Parboosingh, J.S., Sobreira, N., et al. (2015). GeneMatcher aids in the identification of a new malformation syndrome with intellectual disability, unique facial dysmorphisms, and skeletal and connective tissue abnormalities caused by de novo variants in HNRNPK. Hum. Mutat. *36*, 1009–1014.

Bertran-Gonzalez, J., Hervé, D., Girault, J.-A., and Valjent, E. (2010). What is the degree of segregation between striatonigral and striatopallidal projections? Front. Neuroanat. *4*, 1–9.

Blaine, S., Claus, E., Harlaar, N., and Hutchison, K. (2013). TACR1 genotypes predict fMRI response to alcohol cues and level of alcohol dependence. Alcohol. Clin. Exp. Res. *37 (Suppl 1)*, E125–E130.

Caiazzo, M., Giannelli, S., Valente, P., Lignani, G., Carissimo, A., Sessa, A., Colasante, G., Bartolomeo, R., Massimino, L., Ferroni, S., et al. (2015). Direct

conversion of fibroblasts into functional astrocytes by defined transcription factors. Stem Cell Reports *4*, 25–36.

Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., and Di Filippo, M. (2014). Direct and indirect pathways of basal ganglia: a critical reappraisal. Nat. Neurosci. *17*, 1022–1030.

Crittenden, J.R., and Graybiel, A.M. (2011). Basal ganglia disorders associated with imbalances in the striatal striosome and matrix compartments. Front. Neuroanat. *5*, 59.

Cui, G., Jun, S.B., Jin, X., Pham, M.D., Vogel, S.S., Lovinger, D.M., and Costa, R.M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. Nature *494*, 238–242.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proc. Natl. Acad. Sci. USA *112*, 7285–7290.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al.; Consortium (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209–215.

DeLong, M., and Wichmann, T. (2009). Update on models of basal ganglia function and dysfunction. Parkinsonism Relat. Disord. *15* (*Suppl 3*), S237–S240.

Doetsch, F., Caillé, I., Lim, D.A., García-Verdugo, J.M., and Alvarez-Buylla, A. (1999). Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. Cell *97*, 703–716.

Ehrman, L.A., Mu, X., Waclaw, R.R., Yoshida, Y., Vorhees, C.V., Klein, W.H., and Campbell, K. (2013). The LIM homeobox gene Isl1 is required for the correct development of the striatonigral pathway in the mouse. Proc. Natl. Acad. Sci. USA *110*, E4026–E4035.

Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.-H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., et al. (2014). Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nat. Commun. *5*, 4372.

Etienne-Manneville, S., and Hall, A. (2003). Cdc42 regulates GSK-3beta and adenomatous polyposis coli to control cell polarity. Nature *421*, 753–756.

Feng, J., Shao, N., Szulwach, K.E., Vialou, V., Huynh, J., Zhong, C., Le, T., Ferguson, D., Cahill, M.E., Li, Y., et al. (2015). Role of Tet1 and 5-hydroxymethylcytosine in cocaine action. Nat. Neurosci. *18*, 536–544.

Frederick, A.L., Yano, H., Trifilieff, P., Vishwasrao, H.D., Biezonski, D., Mészáros, J., Urizar, E., Sibley, D.R., Kellendonk, C., Sonntag, K.C., et al. (2015). Evidence against dopamine D1/D2 receptor heteromers. Mol. Psychiatry *20*, 1373–1385.

Fuccillo, M.V., Földy, C., Gökce, Ö., Rothwell, P.E., Sun, G.L., Malenka, R.C., and Südhof, T.C. (2015). Single-cell mRNA profiling reveals cell-type-specific expression of neurexin isoforms. Neuron *87*, 326–340.

Gehman, L.T., Meera, P., Stoilov, P., Shiue, L., O'Brien, J.E., Meisler, M.H., Ares, M., Jr., Otis, T.S., and Black, D.L. (2012). The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. Genes Dev. *26*, 445–460.

Grosso, A.R., Gomes, A.Q., Barbosa-Morais, N.L., Caldeira, S., Thorne, N.P., Grech, G., von Lindern, M., and Carmo-Fonseca, M. (2008). Tissue-specific splicing factor gene expression signatures. Nucleic Acids Res. *36*, 4823–4832.

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature *525*, 251–255.

Hegde, S.P., Zhao, J., Ashmun, R.A., and Shapiro, L.H. (1999). c-Maf induces monocytic differentiation and apoptosis in bipotent myeloid progenitors. Blood *94*, 1578–1589.

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.

Heintz, N. (2004). Gene expression nervous system atlas (GENSAT). Nat. Neurosci. *7*, 483.

Heinzen, E.L., Arzimanoglou, A., Brashear, A., Clapcote, S.J., Gurrieri, F., Goldstein, D.B., Jóhannesson, S.H., Mikati, M.A., Neville, B., Nicole, S., et al.; ATP1A3 Working Group (2014). Distinct neurological disorders with ATP1A3 mutations. Lancet Neurol. *13*, 503–514.

Hochstim, C., Deneen, B., Lukaszewicz, A., Zhou, Q., and Anderson, D.J. (2008). Identification of positionally distinct astrocyte subtypes whose identities are specified by a homeodomain code. Cell *133*, 510–522.

Iijima, T., Iijima, Y., Witte, H., and Scheiffele, P. (2014). Neuronal cell type-specific alternative splicing is regulated by the KH domain protein SLM1. J. Cell Biol. *204*, 331–342.

Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science *343*, 776–779.

Kakiuchi, C., Iwamoto, K., Ishiwata, M., Bundo, M., Kasahara, T., Kusumi, I., Tsujita, T., Okazaki, Y., Nanko, S., Kunugi, H., et al. (2003). Impaired feedback regulation of XBP1 as a genetic risk factor for bipolar disorder. Nat. Genet. *35*, 171–175.

Kharchenko, P.V., Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. Nat. Methods *11*, 740–742.

Kreitzer, A.C., and Malenka, R.C. (2008). Striatal plasticity and basal ganglia circuit function. Neuron *60*, 543–554.

Kupchik, Y.M., Brown, R.M., Heinsbroek, J.A., Lobo, M.K., Schwartz, D.J., and Kalivas, P.W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. Nat. Neurosci. *18*, 1230–1232.

Ladd, A.N. (2013). CUG-BP, Elav-like family (CELF)-mediated alternative splicing regulation in the brain during health and disease. Mol. Cell. Neurosci. *56*, 456–464.

Li, Q., Lee, J.-A., and Black, D.L. (2007). Neuronal regulation of alternative premRNA splicing. Nat. Rev. Neurosci. *8*, 819–831.

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.

van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. J. Mach. Learn. Res. *9*, 2579–2605.

Maia, T.V., and Frank, M.J. (2011). From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. *14*, 154–162.

Matthias, K., Kirchhoff, F., Seifert, G., Hüttmann, K., Matyash, M., Kettenmann, H., and Steinhäuser, C. (2003). Segregated expression of AMPA-type glutamate receptors and glutamate transporters defines distinct astrocyte populations in the mouse hippocampus. J. Neurosci. *23*, 1750–1758.

Mühleisen, T.W., Leber, M., Schulze, T.G., Strohmaier, J., Degenhardt, F., Treutlein, J., Mattheisen, M., Forstner, A.J., Schumacher, J., Breuer, R., et al. (2014). Genome-wide association study reveals two new risk loci for bipolar disorder. Nat. Commun. *5*, 3339.

Nelson, A.B., and Kreitzer, A.C. (2014). Reassessing models of basal ganglia function and dysfunction. Annu. Rev. Neurosci. *37*, 117–135.

Noureddine, M.A., Qin, X.J., Oliveira, S.A., Skelly, T.J., van der Walt, J., Hauser, M.A., Pericak-Vance, M.A., Vance, J.M., and Li, Y.J. (2005). Association between the neuron-specific RNA-binding protein ELAVL4 and Parkinson disease. Hum. Genet. *117*, 27–33.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science *304*, 452–454.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Südhof, T.C., and Wernig, M. (2011). Induction of human neuronal cells by defined transcription factors. Nature *476*, 220–223.

Pannasch, U., Freche, D., Dallérac, G., Ghézali, G., Escartin, C., Ezan, P., Cohen-Salmon, M., Benchenane, K., Abudara, V., Dufour, A., et al. (2014). Connexin 30 sets synaptic strength by controlling astroglial synapse invasion. Nat. Neurosci. *17*, 549–558.

Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods *10*, 1096–1098.

Pollen, A.A., Nowakowski, T.J., Shuga, J., Wang, X., Leyrat, A.A., Lui, J.H., Li, N., Szpankowski, L., Fowler, B., Chen, P., et al. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat. Biotechnol. *32*, 1053–1058.

Redies, C., Hertel, N., and Hübner, C.A. (2012). Cadherins and neuropsychiatric disorders. Brain Res. *1470*, 130–144.

Robison, A.J., and Nestler, E.J. (2011). Transcriptional and epigenetic mechanisms of addiction. Nat. Rev. Neurosci. *12*, 623–637.

Rouach, N., Koulakoff, A., Abudara, V., Willecke, K., and Giaume, C. (2008). Astroglial metabolic networks sustain hippocampal synaptic transmission. Science *322*, 1551–1555.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature *485*, 237–241.

Satoh, J., Kino, Y., Asahina, N., Takitani, M., Miyoshi, J., Ishida, T., and Saito, Y. (2016). TMEM119 marks a subset of microglia in the human brain. Neuropathology *36*, 39–49.

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature *498*, 236–240.

Shuen, J.A., Chen, M., Gloss, B., and Calakos, N. (2008). Drd1a-tdTomato BAC transgenic mice for simultaneous visualization of medium spiny neurons in the direct and indirect pathways of the basal ganglia. J. Neurosci. *28*, 2681–2685.

Steinhoff, M.S., von Mentzer, B., Geppetti, P., Pothoulakis, C., and Bunnett, N.W. (2014). Tachykinins and their receptors: contributions to physiological control and the mechanisms of disease. Physiol. Rev. *94*, 265–301.

Südhof, T.C. (2008). Neuroligins and neurexins link synaptic function to cognitive disease. Nature *455*, 903–911.

Surmeier, D.J., Yan, Z., and Song, W.J. (1998). Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. Adv. Pharmacol. *42*, 1020–1023.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. *19*, 335–346.

Todorov, V., and Filzmoser, P. (2009). An object-oriented framework for robust multivariate analysis. J. Stat. Softw. *32*, 1–47.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014a). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature *509*, 371–375.

Treutlein, B., Gokce, O., Quake, S.R., and Südhof, T.C. (2014b). Cartography of neurexin alternative splicing mapped by single-molecule long-read mRNA sequencing. Proc. Natl. Acad. Sci. USA *111*, E1291–E1299.

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P.V., et al. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat. Neurosci. *18*, 145–153.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035–1041.

Yang, N., Zuchero, J.B., Ahlenius, H., Marro, S., Ng, Y.H., Vierbuchen, T., Hawkins, J.S., Geissler, R., Barres, B.A., and Wernig, M. (2013). Generation of oligodendroglial cells by direct lineage conversion. Nat. Biotechnol. *31*, 434–439.

Yeo, G.W., Van Nostrand, E., Holste, D., Poggio, T., and Burge, C.B. (2005). Identification and analysis of alternative splicing events conserved in human and mouse. Proc. Natl. Acad. Sci. USA *102*, 2850–2855.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure: cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science *347*, 1138–1142.

Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keeffe, 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.

Zlojutro, M., Manz, N., Rangaswamy, M., Xuei, X., Flury-Wetherill, L., Koller, D., Bierut, L.J., Goate, A., Hesselbrock, V., Kuperman, S., et al. (2011). Genome-wide association study of theta band event-related oscillations identifies serotonin receptor gene HTR7 influencing risk of alcohol dependence. Am. J. Med. Genet. Part B Neuropsychiatr. Genet. *156B*, 44–58.