



OPEN

## A single-cell trajectory atlas of striatal development

Ashley G. Anderson<sup>1,4</sup>, Ashwinikumar Kulkarni<sup>2,3,4</sup> & Genevieve Konopka<sup>1,2,3</sup>✉

The striatum integrates dense neuromodulatory inputs from many brain regions to coordinate complex behaviors. This integration relies on the coordinated responses from distinct striatal cell types. While previous studies have characterized the cellular and molecular composition of the striatum using single-cell RNA-sequencing at distinct developmental timepoints, the molecular changes spanning embryonic through postnatal development at the single-cell level have not been examined. Here, we combine published mouse striatal single-cell datasets from both embryonic and postnatal timepoints to analyze the developmental trajectory patterns and transcription factor regulatory networks within striatal cell types. Using this integrated dataset, we found that dopamine receptor-1 expressing spiny projection neurons have an extended period of transcriptional dynamics and greater transcriptional complexity over postnatal development compared to dopamine receptor-2 expressing neurons. Moreover, we found the transcription factor, *FOXP1*, exerts indirect changes to oligodendrocytes. These data can be accessed and further analyzed through an interactive website (<https://mouse-striatal-dev.cells.ucsc.edu>).

The striatum is a highly conserved forebrain structure important for regulating a wide range of motor and cognitive behaviors<sup>1</sup>. This region receives dense glutamatergic and neuromodulatory inputs from several brain regions, including the cortex, thalamus, and substantia nigra<sup>1</sup>. These diverse inputs are integrated and propagated to downstream basal ganglia nuclei via distinct classes of striatal GABAergic spiny projection neurons (SPNs), interneurons, and glial cell types. Disruption of striatal cell types has been observed across several neurodevelopmental and neurodegenerative disorders, including autism spectrum disorders (ASD) and Huntington's disease (HD)<sup>2,3</sup>. Uncovering the molecular mechanisms regulating striatal cell type development in the brain is therefore an important step towards identifying mechanisms altered in disease states to ultimately improve therapeutics.

High throughput single-cell RNA-sequencing (scRNA-seq) technology has advanced our understanding of the cellular composition and molecular characterization of the brain<sup>4</sup>. In the striatum specifically, recent scRNA-seq studies have led to important insights into striatal cellular composition and further unraveled the molecular differences between the principal striatal spiny projection neurons and aspiny interneuron subtypes<sup>5–11</sup>. SPNs and interneurons are derived from separate progenitor pools from either the lateral ganglionic eminence (LGE) or medial and caudal ganglionic eminence (MGE, CGE), respectively. SPNs are classically divided into neurons that express dopamine receptor-1 receptor (D1) and project along the *direct* pathway (dSPNs) or SPNs that express dopamine receptor-2 (D2) and project along the *indirect* pathway (iSPNs)<sup>1</sup>. scRNA-seq studies have found more diversity among SPN subtypes than previously appreciated, including a small population of neurons that express either D1 or D2 receptors but have distinct molecular profiles from canonical SPNs (“eccentric” SPNs, or eSPNs)<sup>9</sup>. This population was masked by previous studies relying on fluorescent-reporter-driven techniques<sup>12</sup> to identify and separate dSPNs versus iSPNs for molecular characterization followed by bulk RNA-sequencing approaches, showing the importance of single-cell methodologies<sup>13–15</sup>. Striatal scRNA-seq studies have also shed light on the cell type specific molecular changes that occur upon disrupting genes important for striatal development, such as *FOXP1*, a gene strongly associated with autism and intellectual disability in humans<sup>5</sup>. Though highly expressed in both dSPNs and iSPNs, deletion of *Foxp1* severely affected iSPNs and significantly reduced that cellular population specifically<sup>5</sup>. Striatal interneurons make up ~5% of the striatal neuron population and are largely divided into subtypes that include *Pvalb*-expressing, *Sst/Npy/Nnos*-expressing, *Calb2*-expressing, and *Th*-expressing groups<sup>8,16</sup>. Single-cell studies of striatal interneurons have found interneurons subgroups that are molecularly discrete (i.e. *Npy*<sup>+</sup>/*Sst*<sup>+</sup>, *Chat*<sup>+</sup>, *Th*<sup>+</sup>, *Npy*<sup>+</sup>/*Sst*<sup>-</sup>, *Cck*<sup>+</sup>) or display continuous gradients of gene expression (*Pvalb*)<sup>8</sup>. While these studies have furthered our understanding of striatal cellular and molecular development, each study was performed at a single time point and the developmental trajectory of striatal cell

<sup>1</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. <sup>2</sup>Department of Neuroscience, UT Southwestern Medical Center, Dallas, TX 75390-9111, USA. <sup>3</sup>Peter O'Donnell Jr. Brain Institute, UT Southwestern Medical Center, Dallas, TX 75390, USA. <sup>4</sup>These authors contributed equally: Ashley G. Anderson and Ashwinikumar Kulkarni. ✉email: Genevieve.Konopka@utsouthwestern.edu

types spanning development remains incomplete. A recent single-cell study has made some progress along these lines by profiling cells from the human fetal lateral ganglionic eminence from 7 to 11 post conceptual weeks and identified key transcription factors important for governing D1 and D2 lineage specification, highlighting the benefit of trajectory-level analyses<sup>17</sup>.

To identify the key molecular mechanisms spanning striatal development, we combine previously published single-cell or single-nuclei RNA-sequencing datasets at distinct embryonic and postnatal timepoints in the mouse brain to build a striatal cell type-specific developmental trajectory map<sup>5–11</sup>. From this integrated dataset, we investigate the trajectory pattern and gene regulatory networks within both neuronal and glial populations. We find that dSPNs and iSPNs diverge in their postnatal pseudotime trajectory, with dSPNs exhibiting greater transcriptional complexity compared to iSPNs. We further show how interneuron subtypes and oligodendrocytes change their molecular composition over development. Moreover, we show that FOXP1 may indirectly alter oligodendrocyte maturation via SPN-specific disruption. We created an interactive website to easily access and further analyze these datasets. This resource is an important step towards compiling single-cell data from across labs and methodologies to further our understanding of neural development.

## Results

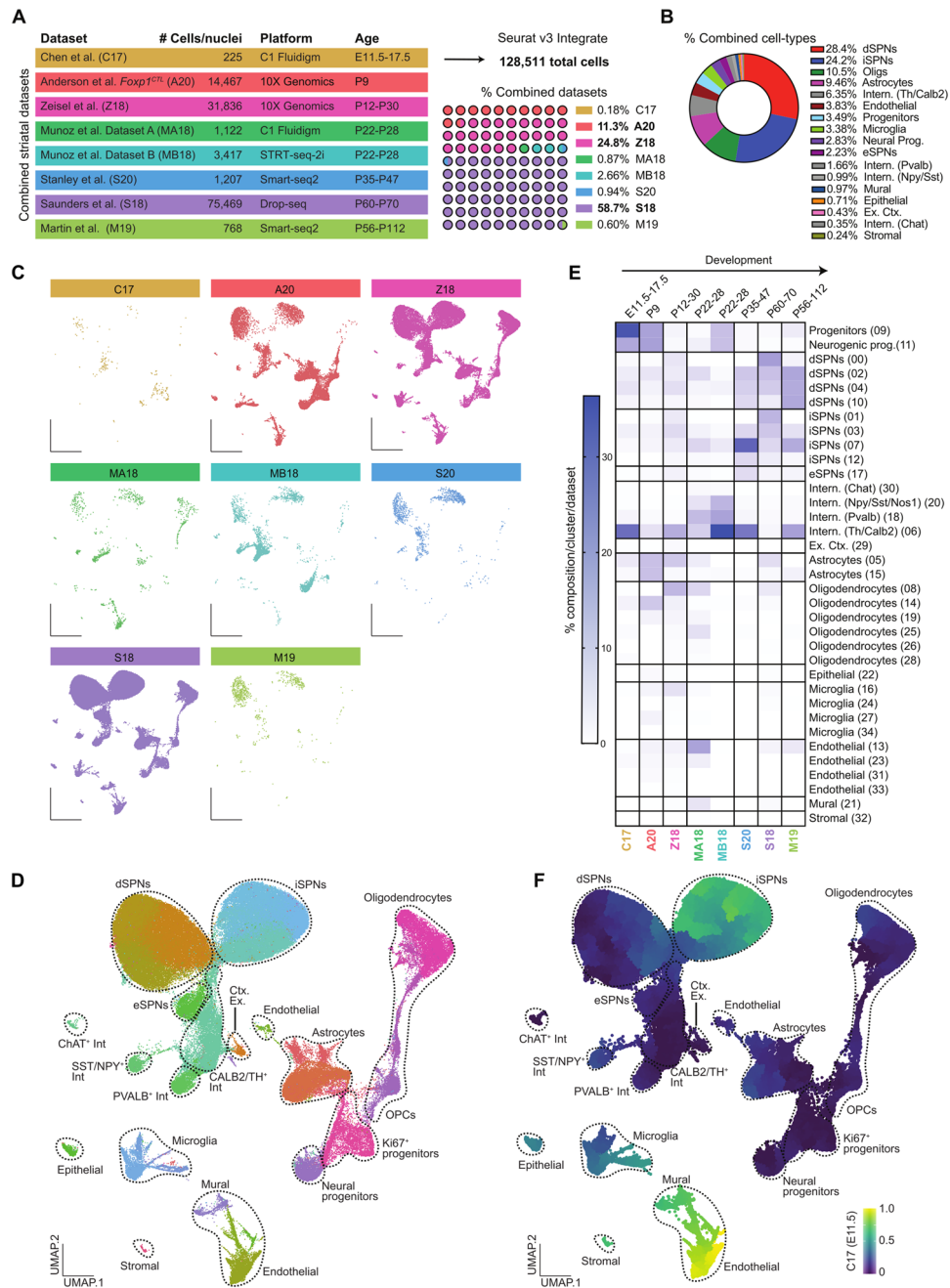
**Combined striatal single-cell datasets across development.** To build a single-cell developmental trajectory map of striatal cell types, we integrated previously published striatal single-cell datasets from seven studies that collected data from mouse brain at different timepoints during striatal development (Fig. 1A). These data include single cells from medial and lateral ganglionic eminences and mature striatal tissue between the ages of embryonic day (E) E11.5–E17.5<sup>6</sup> (C17, 225 cells), postnatal day (P) P9<sup>5</sup> (A20, 14,467 cells), P12–P30<sup>11</sup> (Z18, 31,836 cells), P22–P28<sup>8</sup> (Dataset A, MA18: 1122 cells and Dataset B, MB18: 3417 cells), P35–47<sup>10</sup> (S20, 1207 cells), P60–P70<sup>9</sup> (S18, 75,469 cells), and P56–112<sup>7</sup> (M19, 768 cells). The number of genes detected per dataset was related to the number of cells sequenced, with more genes present in datasets with fewer cells that were more deeply sequenced using Smart-seq2 (C17, S20, and M19, Fig. S1A).

The combined dataset resulted in 128,511 total cells with 35 unique clusters (Fig. S1C). Three datasets across postnatal development contributed greater than 95% of cells to the combined analysis (Fig. 1A). dSPNs (28.4%) and iSPNs (24.2%) comprised ~52.6% of the total dataset, followed by oligodendrocytes (10.5%), astrocytes (9.5%), and interneurons (6.35%) (Fig. 1B). No clusters were unique to a given dataset (Fig. 1C and Fig. S1C), with cells clustering primarily by cell type identity (Fig. 1D and Fig. S1D). The number of genes were greater in neuronal cell types (Fig. S1B) or across neuronal clusters (Fig. S1C) as seen in previous single-cell studies of brain tissue<sup>18</sup>. We found that the percentage of cells from the embryonic and early postnatal timepoints were more abundant in the progenitor and neural progenitor clusters compared to P12–P112 timepoints (Fig. 1E). Four datasets used cellular isolation methods to enrich for distinct cell types, interneurons for MA18 and MB18<sup>8</sup> and SPNs for M19 and S20<sup>7,10</sup>, which is observed in the percent composition of cell types within these studies (Fig. 1E).

To examine the differentiation trajectory pattern of the combined dataset, we used Monocle3<sup>19</sup> to organize cells along a pseudotime scale by setting the root as the biologically earliest cells (E11.5). We then projected their pseudotime values onto UMAP coordinates (Fig. 1F). Using this method, we observed the most dynamic changes in cellular trajectory patterns within three cell types: SPNs, microglia, and vascular cells (Fig. 1F). Within the SPN population, a distinct change in trajectory pattern was observed between dSPNs and iSPNs, with iSPNs progressing faster along the differentiation trajectory compared to dSPNs. These findings suggest dSPNs and iSPNs have distinct developmental trajectory patterns.

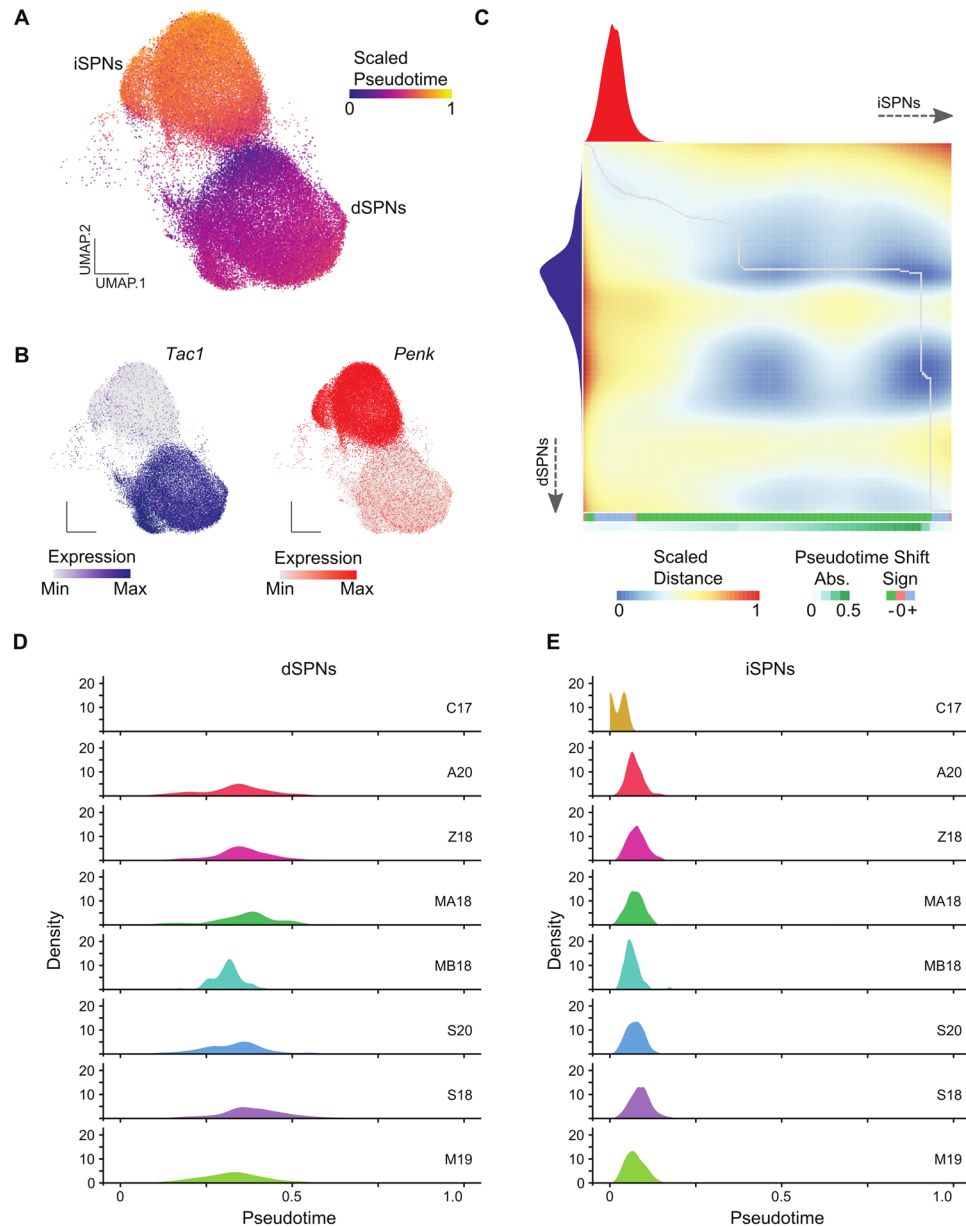
**Extended period of gene expression dynamics in dSPNs relative to iSPNs.** To further study the pseudotime trajectories of dSPNs versus iSPNs, we isolated SPNs from the combined dataset, using clusters 0, 2, 4, 10 for dSPNs and 1, 3, 7, 12 for iSPNs (Fig. 1E). We used PHATE<sup>20</sup> to perform a pseudotime analysis only on SPNs (Fig. 2A). Similar to the results found using Monocle3<sup>19</sup> on all cells, iSPNs were farther along in pseudotime compared to dSPNs (Fig. 2A,B). To quantitatively compare the trajectory dynamics between dSPNs and iSPNs we used cellAlign<sup>21</sup> to compare single-cell pseudotime trajectories. The outputs of this analysis are a global alignment-based dissimilarity matrix and a pseudotime shift score indicating differences between pseudotime values (Fig. 2C). Using this method, we observed a distinct pseudotime shift between iSPNs and dSPNs, indicating that faster gene expression dynamics occur within iSPNs relative to dSPNs (Fig. 2C). This pseudotime shift hints at an extended period of gene expression dynamics occurring in dSPNs compared to iSPNs over postnatal timepoints (Fig. 2C). This change in pseudotime dynamics is observed across each dataset when plotting the expression of dSPN markers (*Drd1*, *Tac1*, Fig. 2D, Fig. S2A top panel) or iSPN markers (*Penk*, *Drd2*, Fig. 2E, Fig. S2A bottom panel) across pseudotime. We note that the embryonic C17 dataset has low signal for dSPNs. Therefore, we used the early postnatal (P9) cells to set the pseudotime trajectory root and observed the same pseudotime shifts, suggesting that this difference in relative gene expression dynamics occurs during postnatal development (Fig. S2B,C).

**dSPNs have more discrete transcriptional networks.** We next used a gene regulatory network (GRN) analysis to identify key transcription factors (TFs) involved in dSPN and iSPN development (Fig. 3). dSPNs and iSPNs have several shared hub TFs including *Foxp1*, *Myt1l*, *Meis2*, and *Csde1*. We also identified hub TFs unique to each subpopulation. dSPNs unique hub TFs included *Sox11*, *Bcl11b*, *Ybx1*, and *Ebf1* (Fig. 3A). iSPNs unique hub TFs included *Rarb*, *Nr1d1*, and *Tef* (Fig. 3B). We observed that dSPNs had more discrete transcriptional networks, compared to iSPNs whose hub genes were more interconnected. Moreover, the dSPNs TF hub genes were enriched with markers of early-born neurons, including *Sox4* and *Sox11* (Fig. 3A). These results suggest that dSPNs have more transcriptional complexity compared to mature iSPNs.



**Figure 1.** Integrated striatal single-cell datasets across timepoints and single-cell methodologies. **(A)** Table describing the published datasets integrated in our analysis and the percent composition of each study to the combined dataset. **(B)** The percent composition of cell types within the combined dataset. **(C)** UMAP plots showing where cells from each dataset clustered in the combined analysis. **(D)** UMAP of combined dataset colored by cluster affiliation and annotated with cell type identification. **(E)** The percent contribution of cells per cluster (35 total) for each dataset ordered by developmental time. **(F)** Pseudotime analysis using Monocle3 of the combined dataset plotted with UMAP coordinates.

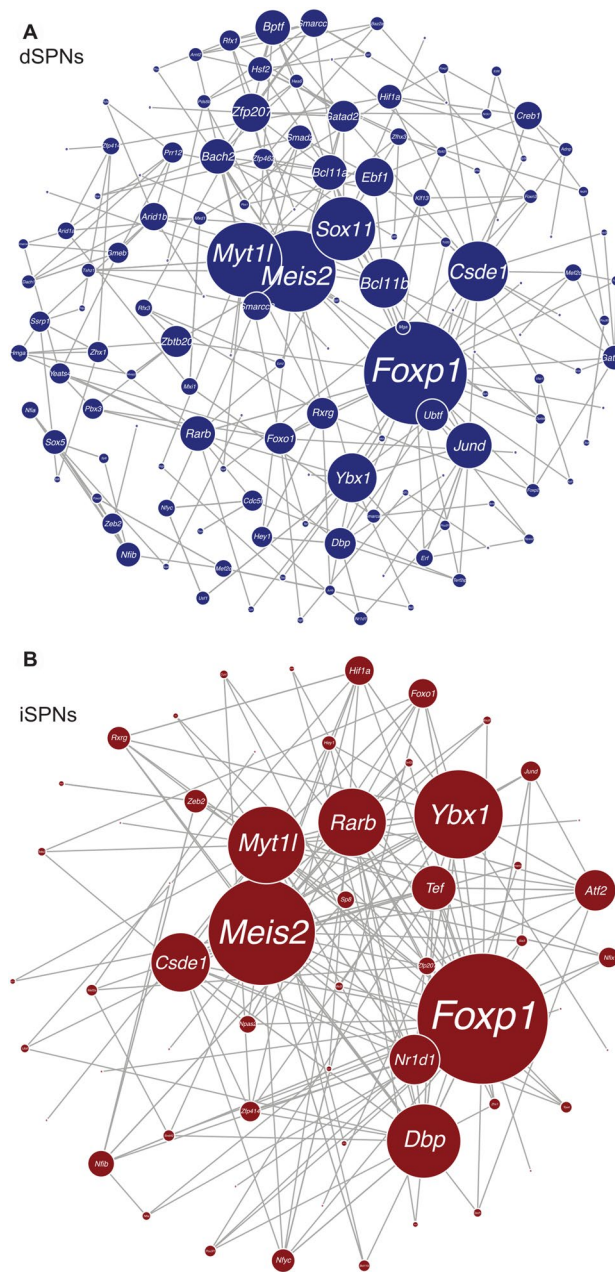
**Interneuron pseudotime trajectories over striatal development.** Several interneuron subtypes populate the striatum and are critical for striatal function. We isolated both interneuron and neural progenitor clusters from the integrated dataset to analyze the pseudotime trajectory pattern of these interneuron subtypes (Fig. 4A). We identified interneurons by key molecular markers including *Chat* (cholinergic interneurons), *Npy* (Neuropeptide Y), *Nos1* (Nitric oxide synthase 1), *Sst* (somatostatin), *Pvalb* (Parvalbumin), *Th* (Tyrosine hydroxylase), and *Calb1* (Calbindin 1) (Fig. 4B). Several of these markers colocalize in the same cells (*Nos1*, *Npy*, *Sst*), whereas *Th*, *Pvalb*, *Calb1*, and *Chat* interneuron clusters were distinct. Using PHATE<sup>20</sup>, we found distinct differences in the pseudotime differentiation trajectory of interneuron subtypes (Fig. 4C). *Chat* and *Npy/Sst* interneurons were further along in pseudotime compared to the other subtypes, followed by *Pvalb*, *Th*,



**Figure 2.** Comparing pseudotime trajectory dynamics between dSPNs and iSPNs. (A) UMAP plot colored by PHATE pseudotime scale with (B) feature plots showing the expression of dSPN marker (*Tac1*) and iSPN marker (*Penk*). (C) Dissimilarity matrix and global alignment of pseudotime trajectories between dSPNs (x-axis) and iSPNs (y-axis) with pseudotime shifts labelled below. (D) Plots of *Tac1* (dSPNs) or (E) *Penk* (iSPN) expression across pseudotime separated by dataset.

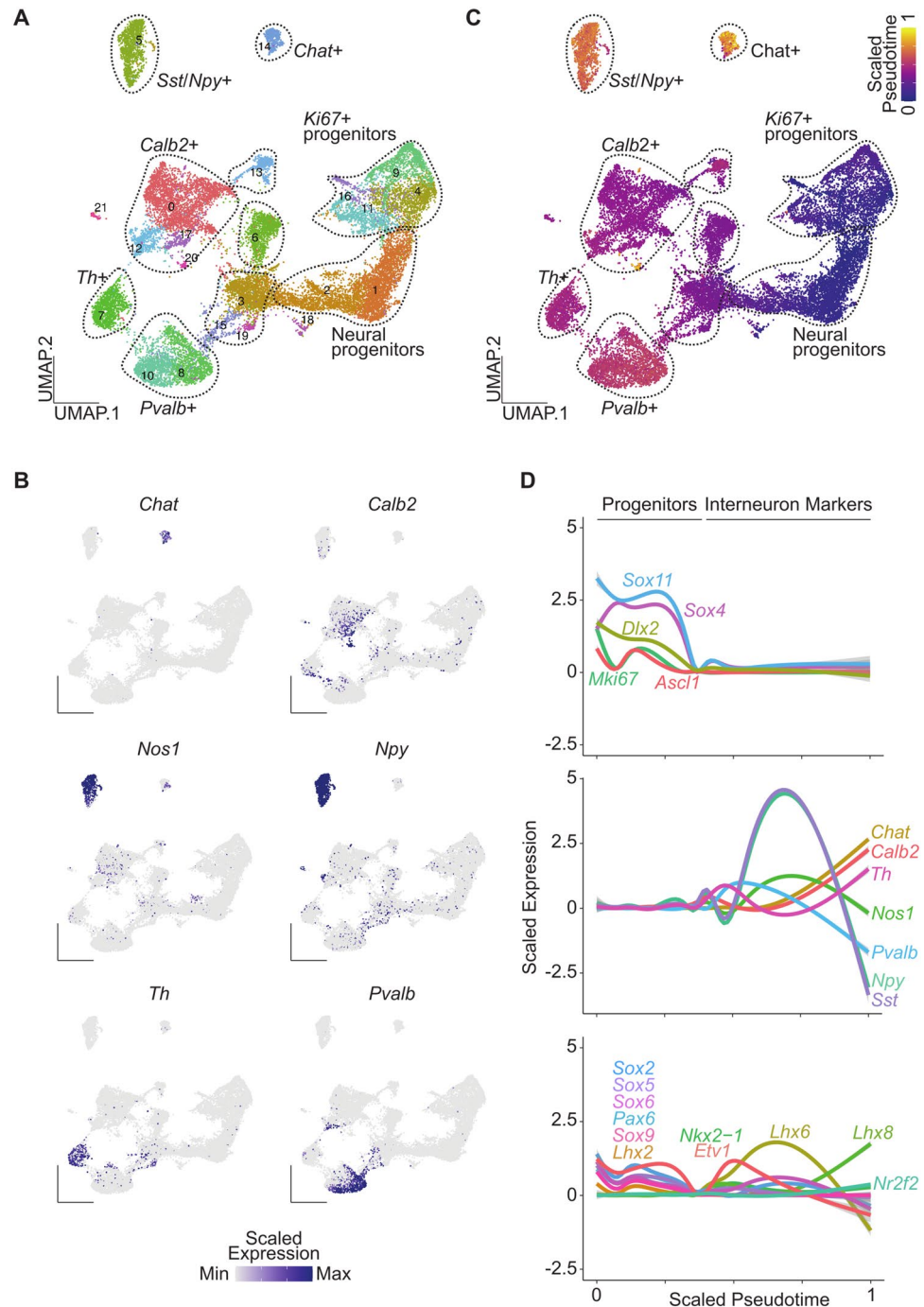
and *Calb2* expressing interneurons (Fig. 4C). We next plotted the expression of key markers of the progenitor state (*Sox4*, *Sox11*, *Dlx2*, *Mki67*, *Ascl1*), interneuron markers, and the TFs highly associated with interneuron development (*Sox2/5/6/9*, *Pax6*, *Lhx2*, *Nkx2.1*, *Etv1*, *Lhx6/8*, and *Nr2f2*) (Fig. 4D)<sup>22</sup>. As expected, peak expression of progenitor markers occurred early in pseudotime, whereas interneuron markers peaked later in pseudotime with little to no overlap. We also saw that *Lhx6* and *Lhx8* expression peaked along the scaled pseudotime after *Nkx2.1* expression, since both are downstream of *Nkx2.1*. Moreover, *Lhx6* is critical for *Pvalb* and *Sst/Npy/Nos1* interneuron specification. *Lhx8* increased over pseudotime following the same trend as *Chat*, an expected finding given that *Lhx8* is important for *Chat* interneuron development and function. Interestingly, we found a bimodal pseudotime pattern of many TFs associated with interneuron development, suggesting successive or distinct waves of interneuron development. This patterning could potentially represent regional differences from interneurons derived from different subregions within the medial or caudal GE, since cell types from both regions have unique cellular trajectories<sup>23</sup>. These results indicate that interneuron subtypes develop along distinct trajectory patterns and provide a rich resource for researchers to further investigate molecular development of striatal interneurons.





**Figure 3.** Gene regulatory analysis of dSPN and iSPNs. Visualization of network transcription factors in dSPNs (A) or iSPNs (B). Gray lines between hubs indicate degree of interconnectivity.

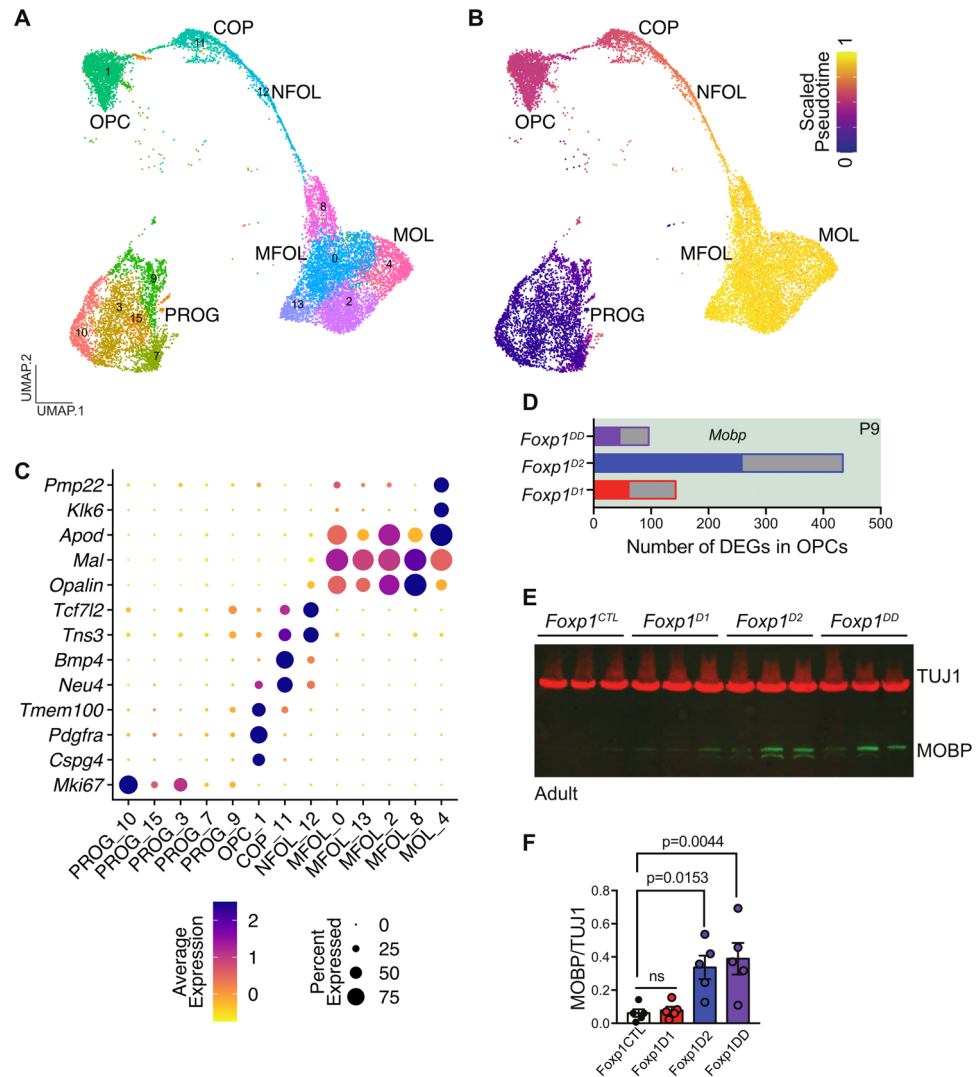
**Oligodendrocyte pseudotime trajectories over striatal development.** We next wanted to examine the developmental trajectory of the second most abundant cell type within this integrated dataset, oligodendrocytes. We isolated both the oligodendrocyte and progenitor clusters from the integrated dataset and identified clusters for oligodendrocyte precursors (OPC, cluster 1), committed oligodendrocyte precursors (COP, cluster 11), newly-formed oligodendrocytes (NFOL, cluster 12), myelin-forming oligodendrocytes (MFOL, clusters 0, 2, 8 and 13), mature oligodendrocytes (MOL, cluster 4) along with progenitors (PROG, clusters 3, 7, 9, 10 and 15) (Fig. 5A). Using PHATE to obtain pseudotime trajectory values for each cell, we found distinct trajectory originating from progenitors (*Mki67*<sup>+</sup>) to clusters enriched for markers of mature oligodendrocytes (*Klk6*<sup>+</sup>, *Apod*<sup>+</sup>) progressing through OPCs, COPs, NFOL and MFOL (Fig. 5B). OPCs were marked by the expression of gene markers such as *Pdgfra* and *Cspg4* (Fig. 5C). Genes previously associated with astrocytes or radial glia (*Tmem100*) also appeared to be enriched in OPCs consistent with the origin of OPCs from radial glia-like cells and their ability to generate astrocytes in an event of injury<sup>24,25</sup> (Fig. 5C). COPs were distinct from OPCs and expressed *Neu4* and genes involved in keeping oligodendrocytes undifferentiated such as *Bmp4*<sup>25,26</sup> (Fig. 5C). NFOLs expressed genes involved in oligodendrocyte differentiation such as *Tcf7l2*<sup>25,27</sup> (Fig. 5C). Both COPs and NFOLs also showed expression of genes involved in migration such as *Ths3*<sup>25</sup> (Fig. 5C). MFOLs expressed



**Figure 4.** Pseudotime analysis of striatal interneuron subtypes. **(A)** Annotated UMAP clusters of interneuron and neural progenitor clusters isolated from the integrated dataset. **(B)** Scaled expression of genes enriched in distinct interneuron populations. **(C)** UMAP of cells colored by PHATE pseudotime values. **(D)** Expression of progenitor markers (top panel), interneuron subtype markers (middle panel), and key transcription factor important for interneuron development (bottom panel) across pseudotime.

genes such as *Mal* and *Opalin* known to be responsible for myelin formation whereas MOLs expressed late oligodendrocyte differentiation genes (*Klk6*, *Apod*) including genes enriched in myelinating cells (*Pmp22*)<sup>25,28</sup> (Fig. 5C). These findings show that striatal oligodendrocytes have distinct subtypes with unique gene expression and trajectory profiles.

**Deletion of *Foxp1* upregulates oligodendrocyte marker *MOBP* in striatum.** To better understand the non-cell autonomous effects of disrupting a key transcription factor in striatal SPN development (Fig. 3), we performed a pseudobulk differential gene expression analysis within oligodendrocytes in the P9 striatal single-



**Figure 5.** Pseudotime analysis of oligodendrocytes across striatal development and with deletion of *Foxp1*. **(A)** Annotated UMAP of oligodendrocyte progenitor cells (OPCs), committed oligodendrocyte precursors (COPs), newly-formed oligodendrocytes (NFOL), myelin-forming oligodendrocytes (MFOL), mature oligodendrocytes (MOL) and progenitor clusters from the combined dataset. **(B)** Cells colored by PHATE pseudotime values. **(C)** Scaled expression of genes specific to each oligodendrocyte subtype. **(D)** Number of differentially expressed genes (DEGs) in OPCs from P9 dataset with deletion of *Foxp1* from dSPNs, iSPNs, or both. Filled bars are upregulated DEGs (such as *Mobp*) and grey bars are downregulated DEGs. **(E)** Western blot of TUJ1 (housekeeping gene) and MOBP (oligodendrocyte marker) in striatal tissue in *Foxp1<sup>CTL</sup>*, *Foxp1<sup>D1</sup>*, *Foxp1<sup>D2</sup>*, and *Foxp1<sup>DD</sup>* mice and **(F)** quantification of MOBP levels relative to TUJ1 across genotypes (N = 5/genotype). Data is represented as mean  $\pm$  SEM. P-values determined using a one-way ANOVA with Dunnett's multiple comparison.

cell dataset with *Foxp1* deleted from either dSPNs (*Foxp1<sup>D1</sup>*), iSPNs (*Foxp1<sup>D2</sup>*), or both (*Foxp1<sup>DD</sup>*). We found both upregulated and downregulated DEGs in oligodendrocytes across all genotypes, but more DEGs were observed when *Foxp1* was deleted specifically from iSPNs (Fig. 5D). The mature oligodendrocyte marker, *Mobp*, was upregulated in *Foxp1<sup>D2</sup>* samples and we confirmed this finding at the protein level (Fig. 5E,F, Fig. S3). While deletion of *Foxp1* in iSPNs was shown to have non-cell-autonomous effects on dSPNs<sup>5</sup>, we now show that loss of *Foxp1* in iSPNs also exerts non-cell-autonomous effects on oligodendrocytes in the striatum.

## Discussion

The striatum is a hub for propagating signals from multiple brain regions to modulate complex learning and motor behaviors. Here, we have developed a single-cell transcriptome resource with the goal of increasing understanding of striatal molecular development at cellular resolution. We have developed an interactive website that integrates previously published striatal single-cell datasets across timepoints and technological modalities. This resource can also be expanded to include additional datasets and can be easily navigated by bench scientists.

Using this integrated striatal single-cell dataset, we analyzed trajectory information for the main neuronal cell types (dSPNs, iSPNs, and interneurons) and one major glial cell type (oligodendrocytes) of the striatum.

Our findings suggest that dSPNs have greater transcriptional complexity compared to iSPNs during postnatal development. In line with our findings, a study of human embryonic striatal scRNA-seq found that dSPNs had slower differentiation kinetics compared to iSPNs and dSPNs had a greater number of transcriptionally distinct clusters<sup>17</sup>. This is interesting given the enrichment of dSPNs in distinct neurochemical compartments of the striatum, known as the striosome (or “patch”), compared to iSPNs. Striosomes receive dense dopaminergic innervation from the VTA and substantia nigra. This innervation becomes more dense over postnatal development, which might play a role in the different trajectory patterns and transcriptional states observed between dSPNs and iSPNs in our analysis.

The striatum contains a substantial population of oligodendrocytes and these cells likely constitute the increased amount of myelination that occurs postnatally on axon tracts targeting and passing through the striatum. Oligodendrocytes are responsible for generating myelin sheaths for the optimization of signal conduction, maturation, survival, and regenerative properties of axons. They are also vulnerable to dysfunction in numerous disorders, including ASD and HD. For example, oligodendrocyte density is increased within HD post-mortem striatum compared to healthy controls. A mouse model of Timothy syndrome, a severe congenital syndrome associated with autism and caused by mutations in an L-type voltage-gated Ca<sup>+</sup> channel (Cav1.2), exhibits accelerated oligodendrocyte development and myelination in the striatum<sup>29</sup>. How oligodendrocytes mature in the striatum over development at the single-cell level is unknown. We found that striatal oligodendrocytes have a distinct lineage with different developmental stages.

Non-neurons, including oligodendrocytes, can send and receive signals to neurons. Such interactions are ultimately important for normal development and function of neurons. Single cell genomics can be harnessed to uncover non-cell autonomous effects on gene expression with the alteration of individual genes in specific cell types. Thus, we asked whether manipulation of striatal SPNs might alter non-neuronal populations in the striatum. We examined how deletion of the transcription factor *Foxp1*, a hub transcription factor in our GRN analysis of SPNs, alters the trajectory pattern of striatal oligodendrocytes. We identified non-cell autonomous gene expression changes in oligodendrocytes with deletion of *Foxp1* in dSPNs, iSPNs, or both cell types. Similar to the Timothy syndrome mouse model, we found that loss of FOXP1 specifically in iSPNs enhanced the maturation of oligodendrocytes and significantly increased the mature oligodendrocyte marker MOBP in adult striatum. These findings are just one example of how this resource can be queried to understand the role of individual genes on cell type specific patterns of expression over striatal development in both a cell autonomous and non-cell autonomous manner. Ultimately, this resource should further our understanding of striatal neurobiology at the single-cell level and aid in addressing therapeutic challenges facing neurodevelopmental and degenerative disorders that alter striatal function.

## Materials and methods

**Integration analysis.** First, raw counts, matching cell type and meta information for each of the datasets was downloaded from respective sources. After checking the integrity of the datasets, raw counts for only common protein-coding genes across all the datasets were retained. Data processing and analysis was performed using R. Individual datasets were first filtered following cutoffs mentioned in each published paper (see table below). For dataset(s) with ‘NA’ cutoffs, either the datasets were already filtered and/or mitochondrial genes were already filtered out. Also, genes with no expression in any of the cells and genes from chromosomes X, Y and M were removed. Following the filtering, each dataset was processed through the standard Seurat (v3) pipeline (*NormalizeData*, *FindVariableFeatures*, *ScaleData*, *FindNeighbors*, *RunUMAP*, *FindClusters*) regressing for the number of UMIs and percent mitochondrial content ([https://satijalab.org/seurat/archive/v3.0/pbmc3k\\_tutorial.html](https://satijalab.org/seurat/archive/v3.0/pbmc3k_tutorial.html))<sup>30</sup>. Seurat objects for each of the datasets were then combined using Seurat’s integration (<https://satijalab.org/seurat/archive/v3.0/integration.html>)<sup>30</sup> approach (*FindIntegrationAnchors*, *IntegrateData*) with 30 principal components. Data were clustered (*FindNeighbors*, *FindClusters*) using the original Louvain algorithm with a resolution of 0.8 and the clusters were visualized with Uniform Manifold Approximation and Projection (UMAP)<sup>4,31</sup> in two dimensions (*RunUMAP*) for a total of 128,511 single cells or nuclei from the mouse striatum. Gene markers enriched for each cluster were identified using ‘*FindAllMarkers*’. Clusters were then annotated using the ‘*LabelTransfer*’ approach from Seurat (<https://satijalab.org/seurat/archive/v3.0/integration.html>)<sup>30</sup> using cell types defined in S18 as reference.

**Pseudotime trajectory analysis for all cell types.** The integrated Seurat object with all cell types for all datasets was converted into a Monocle (v3) compatible object using the ‘*as.cell\_data\_set*’ command. The Monocle object was then pre-processed (*cluster\_cells*, *learn\_graph*) using the standard Monocle pipeline<sup>19,32</sup> (<https://cole-trapnell-lab.github.io/monocle3/docs/trajectories/>). Further, E11.5 cells from C17 were selected as the root population for performing pseudotime trajectory analysis (*order\_cells*). UMAP plots colored by scaled pseudotime values were then generated.



Dataset	# UMI cutoff	% Mito cutoff	# Cells
C17	3,000,000	NA	225
A20	≤ 50,000	≤ 10	14,467
Z18	NA	≤ 10	31,836
MA18	≤ 20,000	NA	1122
MB18	≤ 20,000	NA	3417
S20	NA	NA	1207
S18	≤ 40,000	≤ 10	75,469
M19	NA	NA	768
Total			128,511

**SPN sub-clustering and pseudotime trajectory analysis.** Cell barcodes corresponding to clusters annotated as SPNs (dSPNs, iSPNs and eSPNs) were then used to subset the SPN population from all the cells. Using raw counts corresponding to the identified SPN population for all datasets were then used to run the Seurat integration approach (*FindIntegrationAnchors*, *IntegrateData*) to identify sub-populations among major SPN categories (<https://satijalab.org/seurat/archive/v3.0/integration.html>)<sup>30</sup>. Using 30 principal components, SPNs were clustered (*FindNeighbors*, *FindClusters*) using the original Louvain algorithm with a resolution of 0.8 and clusters were visualized using UMAP<sup>4,31</sup>. Using the clustering information for SPNs, subsets for dSPNs and iSPNs were further created. The dSPN and iSPN subsets were then subjected to pseudotime trajectory analysis using PHATE<sup>20</sup>. First, loom objects corresponding to dSPN and iSPN subsets were exported. Using loom objects as input and *scanpy*'s python implementation of PAGA/PHATE<sup>20</sup> (<https://scanpy-tutorials.readthedocs.io/en/latest/paga-paul15.html>, <https://scanpy.readthedocs.io/en/stable/generated/scanpy.external.tl.phate.html>), pseudotime trajectories were computed using C17 cells as root populations. A UMAP visualization of cells colored by scaled diffusion pseudotime (DPT)<sup>33</sup> was also generated. Pseudotime information for dSPNs and iSPNs was then used to align the trajectories using the *'cellAlign'* approach<sup>21</sup> (<https://github.com/shenorrLab/cellAlign>) and data were visualized using a heatmap accompanied with pseudotime densities. A similar approach was also used to perform sub-clustering and trajectory analysis of SPNs populations using cells from the P9 dataset as root population (Fig. S2B,C).

**Gene regulatory network analysis for SPNs.** A list of mouse transcription factors (TFs) was obtained from a mouse tissue transcription factor atlas<sup>34</sup>. A unique list of 471 TFs falling into fetal brain and adult brain tissue categories were retained for gene regulatory network analysis using an *Arboreto* and *grnboost2* based approach<sup>35</sup>. First, raw counts corresponding to expressed (446/471) TFs was fetched separately for both dSPNs and iSPNs. A gene regulatory network (GRN) was built with raw expression data for dSPNs and iSPNs separately using python implementation of *Arboreto* and *grnboost2* (<https://arboreto.readthedocs.io/en/latest/examples.html>). The GRN output was then filtered following a previously published approach<sup>35</sup>, ([https://github.com/bradleycolquitt/songbird\\_cells/tree/master/grn](https://github.com/bradleycolquitt/songbird_cells/tree/master/grn)) to retain the top one percent of the TF-gene interactions, which were then visualized using the *igraph* R package (<https://igraph.org/r/>).

**Interneurons and oligodendrocytes sub-clustering and pseudotime trajectory analysis.** Cell barcodes corresponding to clusters annotated as interneurons (*Pvalb*<sup>+</sup>, *Sst*<sup>+</sup>/*Npy*<sup>+</sup>, *Chat*<sup>+</sup>, *Calb2*<sup>+</sup>/*Th*<sup>+</sup>) were then used to subset from all the cells along with *Mki67*<sup>+</sup> progenitors and *Sox4*<sup>+</sup>/*Sox11*<sup>+</sup> neurogenic progenitors. Raw counts corresponding to the identified subset population for all datasets were then used to run Seurat integration approach (*FindIntegrationAnchors*, *IntegrateData*) to identify sub-populations among major interneuron categories (<https://satijalab.org/seurat/archive/v3.0/integration.html>)<sup>30</sup>. Using 30 principal components, interneuron cells were clustered (*FindNeighbors*, *FindClusters*) using the original Louvain algorithm with a resolution of 0.8 and clusters were visualized using UMAP<sup>4,31</sup>. The interneuron sub-clustering data were then subjected to pseudotime trajectory analysis using PHATE<sup>20</sup>. First, loom objects corresponding to interneuron clusters were exported. Using loom objects as input and *scanpy*'s python implementation of PAGA/PHATE<sup>20</sup>, (<https://scanpy-tutorials.readthedocs.io/en/latest/paga-paul15.html>, <https://scanpy.readthedocs.io/en/stable/generated/scanpy.external.tl.phate.html>), pseudotime trajectories were computed using cells expressing *Mki67* as root populations. UMAP visualization of cells colored by scaled diffusion pseudotime (DPT)<sup>33</sup> was also generated. Gene expression patterns for specific sets of progenitor markers, interneuron markers and transcription factors were generated across scaled pseudotime. Similar to interneurons, sub-clustering and pseudotime trajectory analysis was also performed for the oligodendrocyte population including *Mki67*<sup>+</sup> progenitors and a UMAP colored by scaled pseudotime was also generated.

**Mice.** All experiments were approved by UT Southwestern IACUC # 2016-101-825. All methods were performed in accordance with all relevant guidelines and regulations as specified by UTSW IACUC and the American Veterinary Medical Association guidelines. This study is reported in accordance with ARRIVE guidelines. *Foxp1*<sup>flox/flox</sup> mice were provided by Dr. Haley Tucker and backcrossed to C57BL/6J for at least 10 generations to obtain congenic animals as previously described<sup>5</sup>. *Drd1a-Cre* (262Gsat, 030989-UCD) and *Drd2-Cre* (ER44Gsat, 032108-UCD) mice were obtained from MMRC.

**Pseudobulk differential gene expression analysis for oligodendrocytes in P9 striatal scRNA-seq data.** Oligodendrocyte clusters were identified based on known marker genes (*Olig1+*) and cells from each cluster were pooled by genotype (*Foxp1<sup>D1</sup>*, *Foxp1<sup>D2</sup>*, *Foxp2<sup>DD</sup>*, or *Foxp1<sup>CTL</sup>*). Differential expression was performed using the Poisson likelihood ratio test from Seurat R analysis pipeline between *Foxp1<sup>CTL</sup>* and *Foxp1<sup>D1</sup>*, *Foxp1<sup>D2</sup>*, or *Foxp2<sup>DD</sup>* oligodendrocytes. Significant expression cutoffs were adj. p-value  $\leq 0.05$  and  $\text{abs}(\log_2\text{FC}) \geq 0.3$ .

**Protein isolation and immunoblotting.** Striatal tissue from adult mice (P56) was harvested as previously described<sup>5</sup>. Briefly, tissue was flash frozen, and protein was extracted using 1X RIPA buffer (750 mM NaCl, 250 mM Tris-HCl pH7.4, 0.5% SDS, 5% Igepal, 2.5% Sodium deoxycholate, 5 mM EDTA, 5 mM NaVO<sub>4</sub>) with fresh protease inhibitor cocktail (10  $\mu\text{l}/\text{ml}$ ), 100 mM PMSF (10  $\mu\text{l}/\text{ml}$ ), and 200 mM sodium orthovanadate (25  $\mu\text{l}/\text{ml}$ ). Tissue was homogenized using a QIAGEN TissueLyser LT, rotated for 1 h at 4 °C, and spun down at max speed for 15 min. Protein was quantified using a standard Bradford assay and 20  $\mu\text{g}$  of protein was loaded into a 10% SDS-Page gel. Protein samples were transferred to a PVDF membrane and then membrane was blocked in a 5% milk TBST solution. The following antibodies were used for immunoblots (IB) experiments: rabbit anti-MOBP (1:2000; Sigma HPA035152) or mouse anti-TUJ1 (1:10,000; Covance MMS-435P). Using an Odyssey infrared imaging system, rectangles were drawn around individual samples in either 800 or 700 IR channels to quantify the intensity signal after setting a background reference rectangle. MOBP signal was normalized to TUJ1 within each sample.

### Data availability

These data can be accessed and further analyzed through an interactive website (<https://mouse-striatal-dev.cells.ucsc.edu>).

Received: 13 December 2022; Accepted: 31 May 2023

Published online: 03 June 2023

### References

- Gerfen, C. R. & Surmeier, D. J. Modulation of striatal projection systems by dopamine. *Annu. Rev. Neurosci.* **34**, 441–466. <https://doi.org/10.1146/annurev-neuro-061010-113641> (2011).
- Crittenden, J. R. & Graybiel, A. M. Basal ganglia disorders associated with imbalances in the striatal striosome and matrix compartments. *Front. Neuroanat.* **5**, 59. <https://doi.org/10.3389/fnana.2011.00059> (2011).
- Fuccillo, M. V. Striatal circuits as a common node for autism pathophysiology. *Front. Neurosci.* **10**, 27. <https://doi.org/10.3389/fnins.2016.00027> (2016).
- Kulkarni, A., Anderson, A. G., Merullo, D. P. & Konopka, G. Beyond bulk: A review of single cell transcriptomics methodologies and applications. *Curr. Opin. Biotechnol.* **58**, 129–136. <https://doi.org/10.1016/j.copbio.2019.03.001> (2019).
- Anderson, A. G., Kulkarni, A., Harper, M. & Konopka, G. Single-cell analysis of Foxp1-driven mechanisms essential for striatal development. *Cell Rep.* **30**(9), 3051–3066 e3057. <https://doi.org/10.1016/j.celrep.2020.02.030> (2020).
- Chen, Y. J. et al. Single-cell RNA sequencing identifies distinct mouse medial ganglionic eminence cell types. *Sci. Rep.* **7**, 45656. <https://doi.org/10.1038/srep45656> (2017).
- Martin, A. et al. A spatiomolecular map of the striatum. *Cell Rep.* **29**(13), 4320–4333 e4325. <https://doi.org/10.1016/j.celrep.2019.11.096> (2019).
- Munoz-Manchado, A. B. et al. Diversity of interneurons in the dorsal striatum revealed by single-cell RNA sequencing and Patch-Seq. *Cell Rep.* **24**(8), 2179–2190 e2177. <https://doi.org/10.1016/j.celrep.2018.07.053> (2018).
- Saunders, A. et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**(4), 1015–1030 e1016. <https://doi.org/10.1016/j.cell.2018.07.028> (2018).
- Stanley, G., Gokce, O., Malenka, R. C., Sudhof, T. C. & Quake, S. R. Continuous and discrete neuron types of the adult murine striatum. *Neuron* **105**(4), 688–699 e688. <https://doi.org/10.1016/j.neuron.2019.11.004> (2020).
- Zeisel, A. et al. Molecular architecture of the mouse nervous system. *Cell* **174**(4), 999–1014 e1022. <https://doi.org/10.1016/j.cell.2018.06.021> (2018).
- Valjent, E., Bertran-Gonzalez, J., Herve, D., Fisone, G. & Girault, J. A. Looking BAC at striatal signaling: Cell-specific analysis in new transgenic mice. *Trends Neurosci.* **32**(10), 538–547. <https://doi.org/10.1016/j.tins.2009.06.005> (2009).
- Heiman, M. et al. A translational profiling approach for the molecular characterization of CNS cell types. *Cell* **135**(4), 738–748. <https://doi.org/10.1016/j.cell.2008.10.028> (2008).
- Lobo, M. K. et al. Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* **330**(6002), 385–390. <https://doi.org/10.1126/science.1188472> (2010).
- Maze, I. et al. G9a influences neuronal subtype specification in striatum. *Nat. Neurosci.* **17**(4), 533–539. <https://doi.org/10.1038/nn.3670> (2014).
- Tepper, J. M. et al. Heterogeneity and diversity of striatal GABAergic interneurons: Update 2018. *Front. Neuroanat.* **12**, 91. <https://doi.org/10.3389/fnana.2018.00091> (2018).
- Bocchi, V. D. et al. The coding and long noncoding single-cell atlas of the developing human fetal striatum. *Science* **372**(6542), eabf5759. <https://doi.org/10.1126/science.abf5759> (2021).
- Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**(7772), 61–68. <https://doi.org/10.1038/s41586-019-1506-7> (2019).
- Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**(7745), 496–502. <https://doi.org/10.1038/s41586-019-0969-x> (2019).
- Moon, K. R. et al. Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* **37**(12), 1482–1492. <https://doi.org/10.1038/s41587-019-0336-3> (2019).
- Alpert, A., Moore, L. S., Dubovik, T. & Shen-Orr, S. S. Alignment of single-cell trajectories to compare cellular expression dynamics. *Nat. Methods* **15**(4), 267–270. <https://doi.org/10.1038/nmeth.4628> (2018).
- Lim, L., Mi, D., Llorca, A. & Marin, O. Development and functional diversification of cortical interneurons. *Neuron* **100**(2), 294–313. <https://doi.org/10.1016/j.neuron.2018.10.009> (2018).
- Lee, D. R. et al. Transcriptional heterogeneity of ventricular zone cells in the ganglionic eminences of the mouse forebrain. *Elife* **11**, e71864. <https://doi.org/10.7554/eLife.71864> (2022).
- Dimou, L. & Gotz, M. Glial cells as progenitors and stem cells: New roles in the healthy and diseased brain. *Physiol. Rev.* **94**(3), 709–737. <https://doi.org/10.1152/physrev.00036.2013> (2014).

25. Marques, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**(6291), 1326–1329. <https://doi.org/10.1126/science.aaf6463> (2016).
26. Samanta, J. & Kessler, J. A. Interactions between ID and OLIG proteins mediate the inhibitory effects of BMP4 on oligodendroglial differentiation. *Development* **131**(17), 4131–4142. <https://doi.org/10.1242/dev.01273> (2004).
27. Ye, F. *et al.* HDAC1 and HDAC2 regulate oligodendrocyte differentiation by disrupting the beta-catenin-TCF interaction. *Nat. Neurosci.* **12**(7), 829–838. <https://doi.org/10.1038/nn.2333> (2009).
28. Cahoy, J. D. *et al.* A transcriptome database for astrocytes, neurons, and oligodendrocytes: A new resource for understanding brain development and function. *J. Neurosci.* **28**(1), 264–278. <https://doi.org/10.1523/JNEUROSCI.4178-07.2008> (2008).
29. Cheli, V. T. *et al.* Enhanced oligodendrocyte maturation and myelination in a mouse model of Timothy syndrome. *Glia* **66**(11), 2324–2339. <https://doi.org/10.1002/glia.23468> (2018).
30. Stuart, T. *et al.* Comprehensive integration of single-cell data. *Cell* **177**(7), 1888–1902 e1821. <https://doi.org/10.1016/j.cell.2019.05.031> (2019).
31. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* <https://doi.org/10.1038/nbt.4314> (2018).
32. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**(4), 381–386. <https://doi.org/10.1038/nbt.2859> (2014).
33. Haghverdi, L., Buttner, M., Wolf, F. A., Büttner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**(10), 845–848. <https://doi.org/10.1038/nmeth.3971> (2016).
34. Zhou, Q. *et al.* A mouse tissue transcription factor atlas. *Nat. Commun.* **8**, 15089. <https://doi.org/10.1038/ncomms15089> (2017).
35. Colquitt, B. M., Merullo, D. P., Konopka, G., Roberts, T. F. & Brainard, M. S. Cellular transcriptomics reveals evolutionary identities of songbird vocal circuits. *Science* **371**(6530), eabd9704. <https://doi.org/10.1126/science.abd9704> (2021).

## Acknowledgements

This work was supported by the NIH (NS122920) to A.G.A.; and the Simons Foundation for Autism Research (Award #573689), the James S. McDonnell Foundation 21<sup>st</sup> Century Science Initiative in Understanding Human Cognition (Scholar Award 220020467) and NIH (MH126481, HG011641, NS115821, MH102603) to G.K.

## Author contributions

A.G.A. conceived the project, conducted experiments, and drafted the manuscript; A.K. analyzed data, generated figures, and edited the manuscript; G.K. supervised the study, edited the manuscript, and provided financial support.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-36255-5>.

**Correspondence** and requests for materials should be addressed to G.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023