SINGLE-CELL GENOMICS

Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding

Alexander B. Rosenberg,[1]*† Charles M. Roco,[2]* Richard A. Muscat,[1] Anna Kuchina,[1] Paul Sample,[1] Zizhen Yao,[3] Lucas T. Graybuck,[3] David J. Peeler,[2] Sumit Mukherjee,[1] Wei Chen,[4] Suzie H. Pun,[2] Drew L. Sellers,[2,5] Bosiljka Tasic,[3] Georg Seelig[1,4,6]†

To facilitate scalable profiling of single cells, we developed split-pool ligation-based transcriptome sequencing (SPLiT-seq), a single-cell RNA-seq (scRNA-seq) method that labels the cellular origin of RNA through combinatorial barcoding. SPLiT-seq is compatible with fixed cells or nuclei, allows efficient sample multiplexing, and requires no customized equipment. We used SPLiT-seq to analyze 156,049 single-nucleus transcriptomes from postnatal day 2 and 11 mouse brains and spinal cords. More than 100 cell types were identified, with gene expression patterns corresponding to cellular function, regional specificity, and stage of differentiation. Pseudotime analysis revealed transcriptional programs driving four developmental lineages, providing a snapshot of early postnatal development in the murine central nervous system. SPLiT-seq provides a path toward comprehensive single-cell transcriptomic analysis of other similarly complex multicellular systems.

More than 300 years have passed since van Leeuwenhoek first described living cells, yet we still do not have a complete catalog of cell types or their functions. Recently, transcriptomic profiling of individual cells has emerged as an essential tool for characterizing cellular diversity (*1*–*3*). Single-cell RNA-sequencing (scRNA-seq) methods have profiled tens of thousands of individual cells (*4*–*6*), revealing new insights about cell types within both healthy (*7*–*14*) and diseased tissues (*15*–*18*). Unfortunately, since these methods require cell sorters, custom microfluidics, or microwells, throughput is still limited and experiments are costly. We introduce split-pool ligation-based transcriptome sequencing (SPLiT-seq), a low-cost, scRNA-seq method that enables transcriptional profiling of hundreds of thousands of fixed cells or nuclei in a single experiment. SPLiT-seq does not require partitioning single cells into individual compartments (droplets, microwells, or wells) but relies on the cells themselves as compartments. The entire workflow before sequencing consists just of pipetting steps, and no complex instruments are needed.

In SPLiT-seq, individual transcriptomes are uniquely labeled by passing a suspension of formaldehyde-fixed cells or nuclei through four rounds of combinatorial barcoding. In the first round of barcoding, cells are distributed into a 96-well plate, and cDNA is generated with an in-cell reverse transcription (RT) reaction using well-specific barcoded primers. Each well can contain a different biological sample, thereby enabling multiplexing of up to 96 samples in a single experiment. After this step, cells from all wells are pooled and redistributed into a new 96-well plate, where an in-cell ligation reaction appends a second well-specific barcode to the cDNA. The third-round barcode, which also contains a unique molecular identifier (UMI), is then appended with another round of pooling, splitting, and ligation. After three rounds of barcoding, the cells are pooled and split into sublibraries, and sequencing barcodes are introduced by polymerase chain reaction (PCR). This final step provides a fourth barcode, while also making it possible to sequence different numbers of cells in each sublibrary. After sequencing, each transcriptome is assembled by combining reads containing the same four-barcode combination (Fig. 1A and fig. S1A).

Four rounds of combinatorial barcoding can yield 21,233,664 barcode combinations (three rounds of barcoding in 96-well plates followed by a fourth round with 24 PCR reactions), enough to uniquely label over 1 million cells. Even larger numbers of barcode combinations can be achieved by performing experiments in 384-well plates or through additional rounds of barcoding (fig. S1B). In addition, by performing the first step in a 384-well plate, up to 384 different biological samples could be combined in a single experiment.

SPLiT-seq validation

To test SPLiT-seq's ability to generate uniquely barcoded cells (UBCs), we performed a species-mixing experiment. We mixed cells from one mouse and two human cell lines (NIH/3T3, HEK293, and Hela-S3), fixed them, and used SPLiT-seq to generate a scRNA-seq library with 1758 UBCs. The library was sequenced, and reads were aligned to a combined mouse-human genome. Nearly all (99.9%) of the UBCs were unambiguously assigned to a single species (>90% of reads aligned to a single genome), with the remaining 0.1% of UBCs representing barcode collisions between mouse and human cells (Fig. 1B). At saturating read coverage (>500,000 reads per cell), we identified a median of 15,365 UMIs and 5498 genes per human cell and 12,243 UMIs and 4497 genes per mouse cell. The species purity in both human and mouse UBCs was high: 99.6% of reads in human UBCs and 99.0% of reads in mouse UBCs aligned to their respective genomes. We also performed single-nucleus RNA-seq (snRNA-seq) experiments using SPLiT-seq with freshly prepared nuclei, as well as nuclei and cells that had been preserved at −80°C for 2 weeks. In all samples, we detected similar numbers of transcripts and genes per cell (Fig. 1C, fig. S2, and table S1). Gene expression was highly correlated between preserved and freshly prepared cells (Fig. 1D and fig. S2) (Pearson *r*, 0.987), as well as between cells and nuclei (fig. S2) (Pearson *r*, 0.952). We also examined gene and UMI detection at different sequencing depths and found that the sensitivity of SPLiT-seq is comparable to droplet-based scRNA-seq methods (fig. S3).

Single-nuclei RNA-seq of developing mouse brain and spinal cord

We used SPLiT-seq to profile nuclei from the developing brain and spinal cord of postnatal day 2 and 11 (P2 and P11) mice. The first round of barcoding assigned identifiers for the P2 brain, P2 spinal cord, P11 brain, and P11 spinal cord samples (Fig. 2A and fig. S4). In total, four rounds of barcoding (48 × 96 × 96 × 14) generated more than 6 million distinct barcode combinations, making it possible to process hundreds of thousands of nuclei in a single experiment with minimal barcode collisions (2.5% expected collisions for 150,000 nuclei).

To determine how many transcripts SPLiT-seq detects within nuclei from the central nervous system, we performed deep sequencing on a sublibrary containing only 131 nuclei. We detected 4943 UMIs and 2055 genes per nucleus (UMI duplication, 95%). We then sequenced the rest of the library at lower depth, resulting in a median of 677 genes and 1022 UMIs per nucleus (UMI duplication, 58%) (table S2). Low-quality transcriptomes were removed from analysis (*19*), yielding 156,049 single-nucleus transcriptomes (74,862 P2 brain; 7028 P2 spinal cord; 58,573 P11 brain; 15,586 P11 spinal cord).

Unsupervised clustering grouped transcriptomes into 73 distinct clusters (*19*) (tables S3 to S5), which were visualized by t-distributed stochastic neighbor embedding (t-SNE) (Fig. 2A). Each of these 73 clusters was assigned to a cell class on the basis of expression of established

[1]Department of Electrical Engineering, University of Washington, Seattle, WA, USA. [2]Department of Bioengineering, University of Washington, Seattle, WA, USA. [3]Allen Institute for Brain Science, Seattle, WA, USA. [4]Molecular Engineering and Sciences Institute, University of Washington, Seattle, WA, USA. [5]Institute for Stem Cell and Regenerative Medicine, Seattle, WA, USA. [6]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: alex.b.rosenberg@gmail.com (A.B.R.); gseelig@uw.edu (G.S.)

Downloaded from http://science.sciencemag.org/ on July 15, 2019

marker genes (Fig. 2B). Neurons accounted for 83% of the profiled transcriptomes (54 clusters), with most clusters expressing *Meg3*.

The 27,096 non-neuronal transcriptomes spanned 19 different clusters, each assigned to a specific cell type. Four astrocyte types (Fig. 2C) accounted for 50% of all non-neuronal nuclei (n = 13,481). Oligodendrocytes (six types, n = 4294) and oligodendrocyte precursor cells (OPC) (one type, n = 5793) formed the second most abundant population. We further identified two vascular and leptomeningeal cell (VLMC) types (fig. S5A), endothelial cells, smooth muscle cells (fig. S5B), microglia, macrophages (fig. S5C) (*20*, *21*), ependymal cells, and olfactory ensheathing cells (OEC).

Previous work has observed that t-SNE can order cells in two-dimensional space according to stages of differentiation (*9*). Moving through t-SNE space along the path of differentiation can then be viewed as moving through "pseudotime" (*22*). As oligogenesis spans the first two postnatal weeks of murine development (*23*), we asked whether the oligodendrocyte and OPC clusters might reflect a continuous developmental trajectory. When we examined the oligodendrocyte clusters, we found that they formed an overlapping elongated shape in the t-SNE visualization. OPCs and oligodendrocytes from the P2 mouse were enriched at one end of the structure, whereas oligodendrocytes from the P11 mouse were enriched at the opposite end (fig. S6), indicative of a lineage (*19*, *22*).

We then performed a more thorough analysis of this putative lineage. To ensure that our ordering of oligodendrocytes was determined exclusively by their relationship to other oligodendrocytes, rather than all cells, we re-embedded only transcriptomes within these seven clusters with t-SNE (Fig. 2D and fig. S7A). We calculated the moving average of gene expression in the resulting pseudotime ordering (Fig. 2E and fig. S8). Analysis of these expression patterns confirmed that proliferating OPCs segregated to one end of the t-SNE, whereas mature oligodendrocytes segregated to the opposite end (fig. S7B). We also detected previously reported intermediate stages of oligodendrocyte development, with the order of gene expression across pseudotime nearly identical to the one defined previously (*9*) (fig. S7C) (Spearman r, 0.94). When analyzing spinal cord– and brain-derived cells separately, we found more mature oligodendrocytes in the spinal cord than in the brain (fig. S7D), indicating that oligodendrocyte maturation occurs earlier in the spinal cord.

Neuronal cell types

Using known gene markers, we were able to assign most neuronal clusters to specific cell types (*19*). Although some clusters corresponded to abundant cell types, such as cerebellar granule cells (CGCs), others mapped to rare and often less-characterized cell types, such as mitral/tufted cells. Previously characterized regional markers were used to assign the majority of clusters to a specific region of the brain (*24*) (Fig. 3A). Regional assignments were validated with RNA in situ hybridization (ISH) from the Allen Institute's Developing Mouse Brain Atlas (Allen DMBA) (*25*). Specifically, we generated composite ISH maps by averaging across the five most highly enriched genes from each of our clusters (tables S6 and S7). For clusters primarily containing P2 or P11 nuclei, we used the P4 or P14 atlases, respectively. The resulting composite maps confirmed the high regional specificity of most types (Fig. 3B and figs. S9 and S10). Cortical pyramidal neuronal types could be further assigned to specific layers using marker genes (Fig. 3C) (*7*, *8*).

Granule cell fate in the hippocampus

In the hippocampus, immature granule cells originating in the dentate gyrus give rise not only to mature granule cells but also to pyramidal neurons (*26*). This process is one of two instances of neurogenesis that continues into adulthood (*27*), but little is known about the underlying transcriptional program. We determined that three neuronal cell types from the hippocampus likely constituted a developmental trajectory (*19*). Analysis of only these transcriptomes with t-SNE revealed a clear branching structure (Fig. 3D and fig. S11A). The transcription factor *Prox1*, suspected to be necessary for granule cell identity (*28*), was exclusively expressed in one branch,

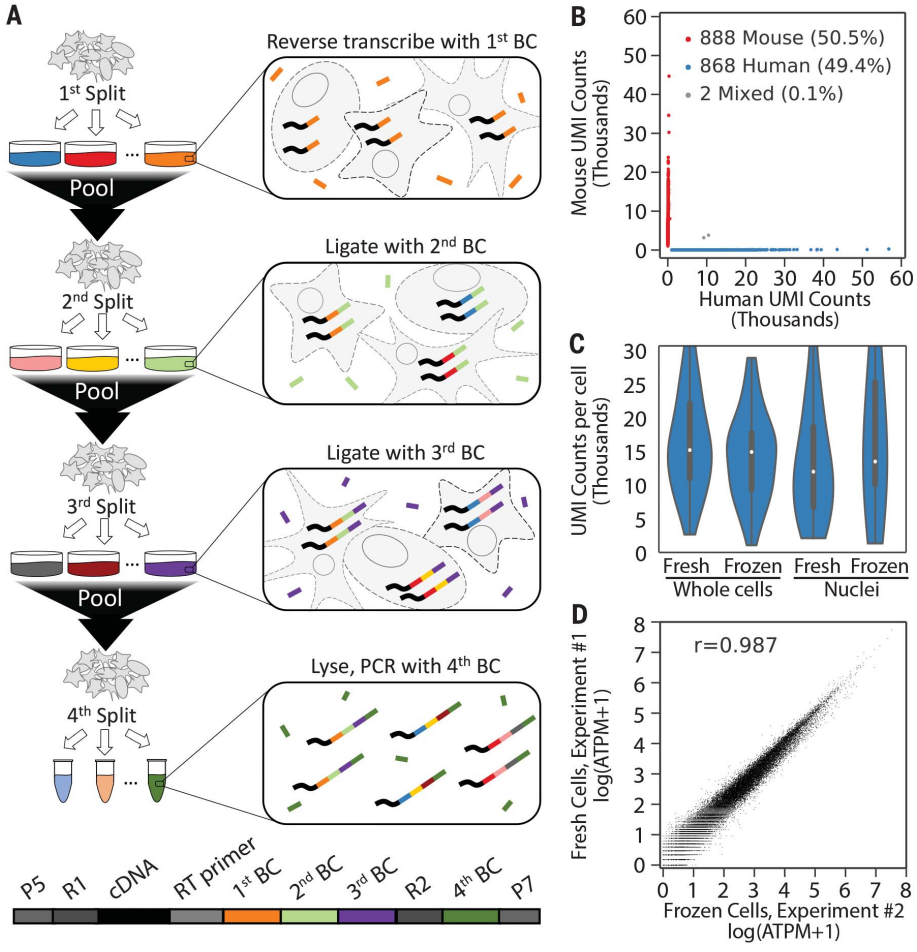

Fig. 1. Overview of SPLiT-seq. (**A**) Labeling transcriptomes with split-pool barcoding. In each split-pool round, fixed cells or nuclei are randomly distributed into wells, and transcripts are labeled with well-specific barcodes. Barcoded RT primers are used in the first round. Second- and third-round barcodes are appended to cDNA through ligation. A fourth barcode is added to cDNA molecules by PCR during sequencing library preparation. The bottom schematic shows the final barcoded cDNA molecule. (**B**) Species-mixing experiment with a library prepared from 1758 whole cells. Human UBCs are blue, mouse UBCs are red, and mixed-species UBCs are gray. The estimated barcode collision rate is 0.2%, whereas species purity is >99%. (**C**) UMI counts from mixing experiments performed with fresh and frozen (stored at −80°C for 2 weeks) cells and nuclei. Median human UMI counts for fresh cells: 15,365; frozen cells: 15,078; nuclei: 12,113; frozen nuclei: 13,636. (**D**) Measured gene expression by SPLiT-seq is highly correlated between frozen cells and cells processed immediately (Pearson r, 0.987). Frozen and fresh cells were processed in two different SPLiT-seq experiments.

Downloaded from http://science.sciencemag.org/ on July 15, 2019

whereas genes known to be specific to CA3 pyramidal neurons such as *Spock1* (*29*) were expressed exclusively in the other branch. Markers of dividing neuronal progenitors were expressed before the branching point, and genes in the Slit-Robo signaling pathway were differentially expressed between the two lineages (fig. S11B). We used these data to identify specific temporal dynamics of transcription factors across the two lineages, with *Meis2* as a candidate marker of early pyramidal cell differentiation (Fig. 3E and fig. S12).

Profiling cells in the developing cerebellum

The cerebellum accounts for only 9% of the brain mass in adult mice but contains nearly 85% of all neurons (*30*). Despite the wide range of functions performed by the cerebellum, many of the gene expression programs driving development of cerebellar cell types remain unknown. We identified the four main cerebellar neuronal types (Fig. 4A): Purkinje cells, Golgi cells, stellate/basket cells, and CGCs. Two types of Purkinje cells (Fig. 4B) were segregated primarily by age (P2 versus P11) and did not form a continuous trajectory in t-SNE but rather two clearly segregated clusters. The absence of cells at intermediate stages of maturation suggests that Purkinje cell development may be more synchronous than other processes of neurogenesis captured by our data set.

CGCs, the most numerous type of neuron in the brain (*31*), drive the postnatal foliation of the cerebellar cortex by migrating from the external granule layer (EGL) through the molecular layer (ML) and the Purkinje cell layer (PcL) to the internal granule layer (IGL) (*32*, *33*). We created a pseudotime ordering of 15,360 CGCs (Fig. 4C and fig. S13) and measured gene expression across this lineage. We defined genes with specific expression at different points in pseudotime (fig. S14) and then used RNA ISH to map these genes to layers of the developing cerebellar cortex. Genes ordered from early to late in pseudotime were progressively expressed from outer to inner layers, consistent with the known direction of CGC migration (Fig. 4D). Our analysis revealed previously unknown pseudotime and layer-specific gene expression patterns within pathways related to axonal development and neuronal migration (fig. S15).

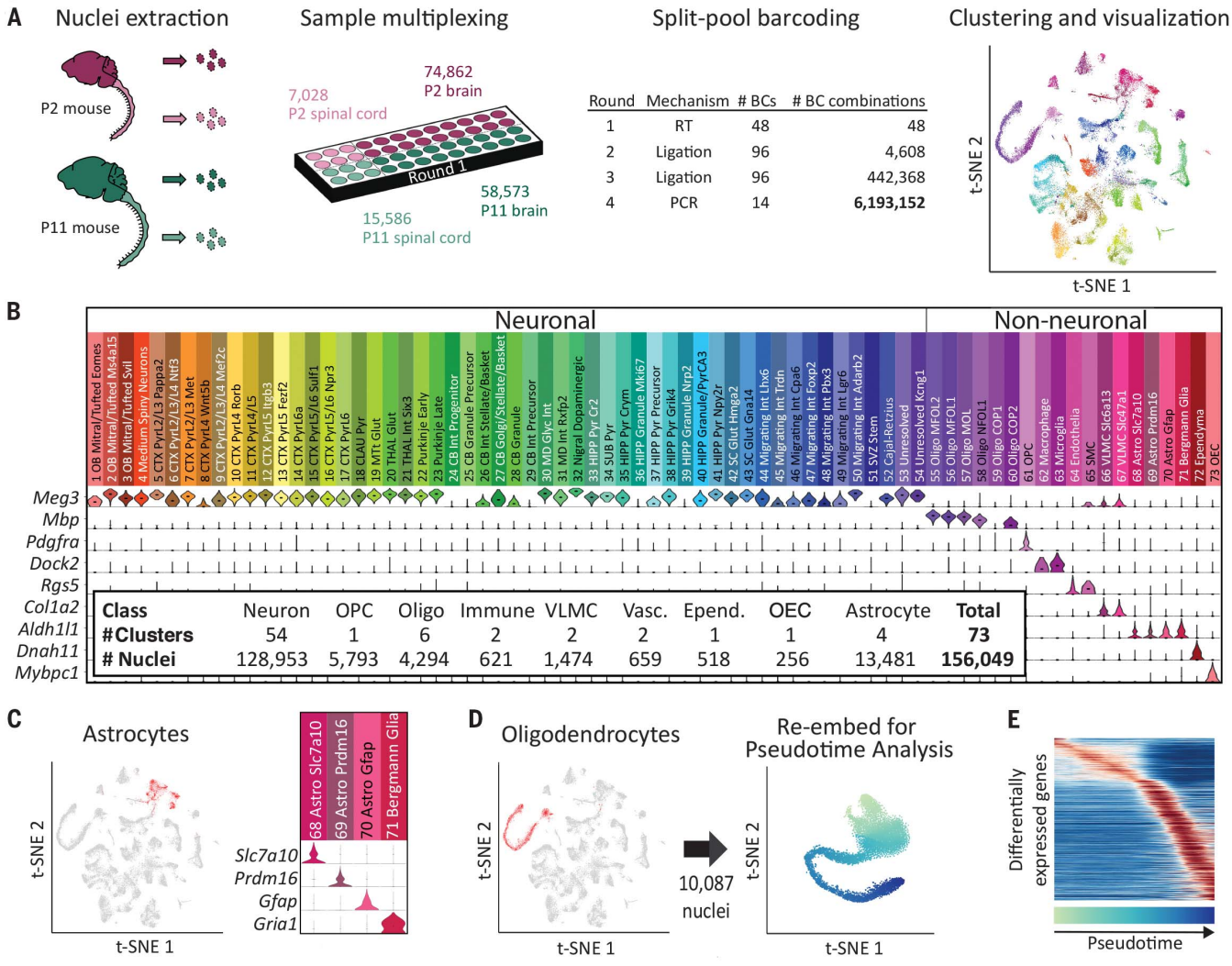

Fig. 2. Single-cell transcriptome landscape of postnatal brain and spinal cord development by SPLiT-seq. (**A**) More than 150,000 nuclei from P2 and P11 mouse brains and spinal cords were profiled in a single experiment employing more than 6 million barcode combinations. Transcriptomes were clustered and then visualized using t-SNE. Cells are colored according to cell type. Each cluster was downsampled to 1000 cells for visualization. (**B**) A total of 73 distinct clusters were assigned to nine cell classes based on expression of established markers. The violin plots show marker gene expression in each cluster. (**C**) Astrocyte clusters are highlighted in red in the t-SNE. The violin plots show markers that are differentially expressed between astrocyte subtypes. (**D**) Seven OPC and oligodendrocyte clusters (containing 10,087 nuclei) colocalized in the original t-SNE (highlighted in red), forming a lineage. Cells from these clusters were re-embedded with t-SNE. (**E**) The heat map shows genes expressed differentially across pseudotime in the oligodendrocyte lineage.

Downloaded from http://science.sciencemag.org/ on July 15, 2019

Origins of cerebellar inhibitory interneurons

The question of whether all cerebellar inhibitory interneurons arise from the same progenitor population has been a point of contention (*34*). Early hypotheses proposed that stellate/basket cells originated from precursors in the EGL, whereas Golgi cell precursors resided in the ventricular epithelium (*35*). Later evidence indicated that these two interneurons shared a common precursor in the cerebellar white matter (*36*, *37*). However, the molecular profile of the inhibitory neuron lineage in the cerebellum remains largely unknown.

We found a cerebellar inhibitory interneuron lineage (1517 cells) (Fig. 4E and fig. S16A) with a

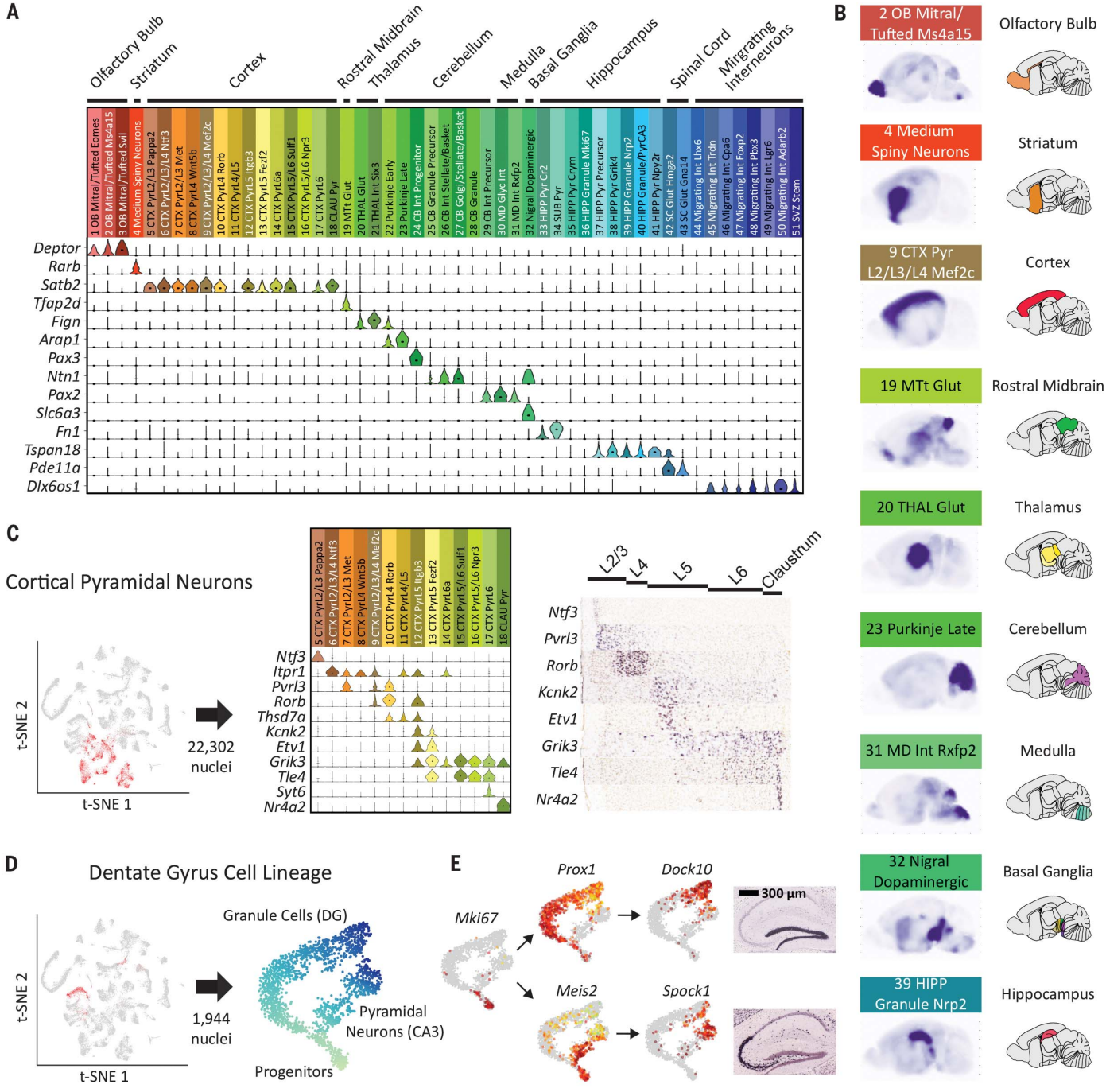

Fig. 3. Neuronal clusters exhibit regional specificity. (**A**) Marker gene expression was used to map neuronal clusters to specific brain regions. (**B**) Sagittal composite RNA ISH maps for nine representative clusters from distinct areas. For each cell type, we averaged ISH intensities from the Allen DMBA across the top five differentially expressed genes. (**C**) Types of pyramidal neurons in the cortex display layer-specific enrichments according to marker genes; cortical pyramidal neurons are highlighted in red in the t-SNE. Expression of example marker genes in pyramidal clusters is shown in the middle, and corresponding available RNA ISH results are on the right. (**D**) Three clusters constitute a developmental trajectory in the hippocampus. Re-embedding these clusters highlights the branching of the two differentiation trajectories in pseudotime. (**E**) Expression of differentiation marker genes is overlaid on the t-SNE. RNA ISH maps (Allen DMBA) show the regional specificity of granule cell and pyramidal neuron markers.

Downloaded from http://science.sciencemag.org/ on July 15, 2019

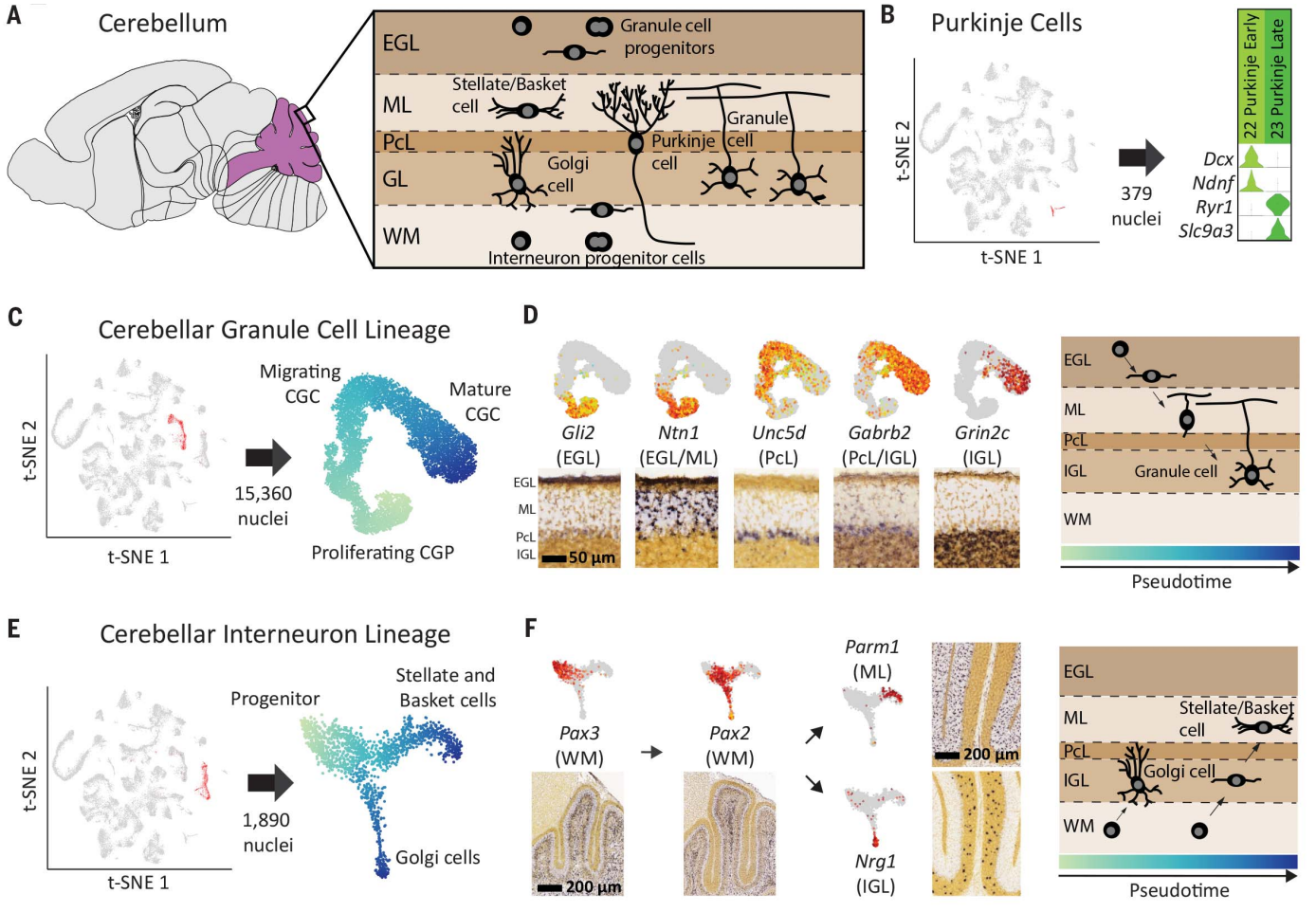

Fig. 4. Neuronal differentiation trajectories in the cerebellum revealed by SPLiT-seq. (**A**) Major cell types and their locations in the cerebellum. (**B**) Two types of Purkinje cells with distinct gene expression programs were identified. Early Purkinje cells are primarily found in the P2 brain and late Purkinje cells in the P11 brain. (**C**) t-SNE re-embedding of 15,360 nuclei suggests a pseudotime ordering from proliferating, to migrating, to mature CGCs. (**D**) Expression of marker genes is overlaid on the t-SNE, and the corresponding RNA ISH from Allen DMBA is shown below. Marker genes associated with different layers of the cerebellum are expressed at different points in pseudotime. Gene expression order is consistent with ordering of the physical layers. RNA ISH maps confirm regional specificity of marker genes. (**E**) t-SNE re-embedding of 1890 nuclei reveals a branching differentiation trajectory. Progenitors can either become Golgi cells or stellate/basket cells. (**F**) Markers for progenitors and mature cell types are expressed at different points in pseudotime and have layer specificity.

shared progenitor branching into either Golgi or stellate/basket cells (fig. S17). This lineage includes a known precursor cell type expressing *Pax2* (*36*) but also a previously unknown, earlier precursor expressing *Pax3* (Fig. 4F). RNA ISH analysis suggests that this Pax3^{+} precursor is located deep within the cerebellar white matter. Moreover, we found that stellate/basket cells expressed genes specific to the molecular layer, whereas Golgi cells expressed genes specific to the granule cell layer (Fig. 4F and fig. S18). The distribution of P2 and P11 nuclei within the lineage clearly demonstrated that the maturation of Golgi cells was well under way by P2 and complete by P11 (fig. S16B). In contrast, stellate/basket cells had not begun to differentiate at P2 and were still not fully mature by P11. These results indicate that the same molecularly defined precursor gives rise to two distinct interneurons at different stages of development.

Cell types in the developing spinal cord

The original clustering was dominated by cells in the brain, and many spinal cord cells did not segregate into well-defined clusters (fig. S19). To resolve more cell types in the spinal cord, we selected all the nuclei originating from the spinal cord and reclustered them (*19*), resulting in 44 clusters: 14 non-neuronal types (12 of which were also found in the brain) and 30 neuronal types (Fig. 5A and tables S8 to S10). We identified 11 different types of γ-aminobutyric acid–releasing (GABAergic) neurons, of which several were also glycinergic (Fig. 5B). One GABAergic type was identified as cerebrospinal fluid–contacting neurons (CSF-cNs) (*38*), with the other 10 types corresponding to inhibitory interneurons. Glutamatergic interneurons accounted for 15 additional types. We also identified two clusters of cholinergic motor neuron types (alpha and gamma) (*39*). To date, known markers exist only for gamma motor neurons (e.g., *Esrrg*) (*40*); however, we identified specific markers for both alpha and gamma neurons (Fig. 5C).

To infer the spatial origin of neuronal types in the spinal cord, we identified the 10 most enriched genes in each type according to our snRNA-seq data and created composite ISH maps based on the Allen Mouse Spinal Cord Atlas (*41*) (Fig. 5D and fig. S20). Some interneuron subtypes appeared to originate primarily from laminae 1 to 3, with others originating from laminae 4 to 6. We found both inhibitory and excitatory neurons in each region. Motor neurons expressed genes found in laminae 9, whereas CSF-cNs were the only neuronal type expressing genes found in the central canal. These data allowed us to create an atlas of gene expression in the early spinal cord, providing a rich resource for further understanding development of the central nervous system.

Downloaded from http://science.sciencemag.org/ on July 15, 2019

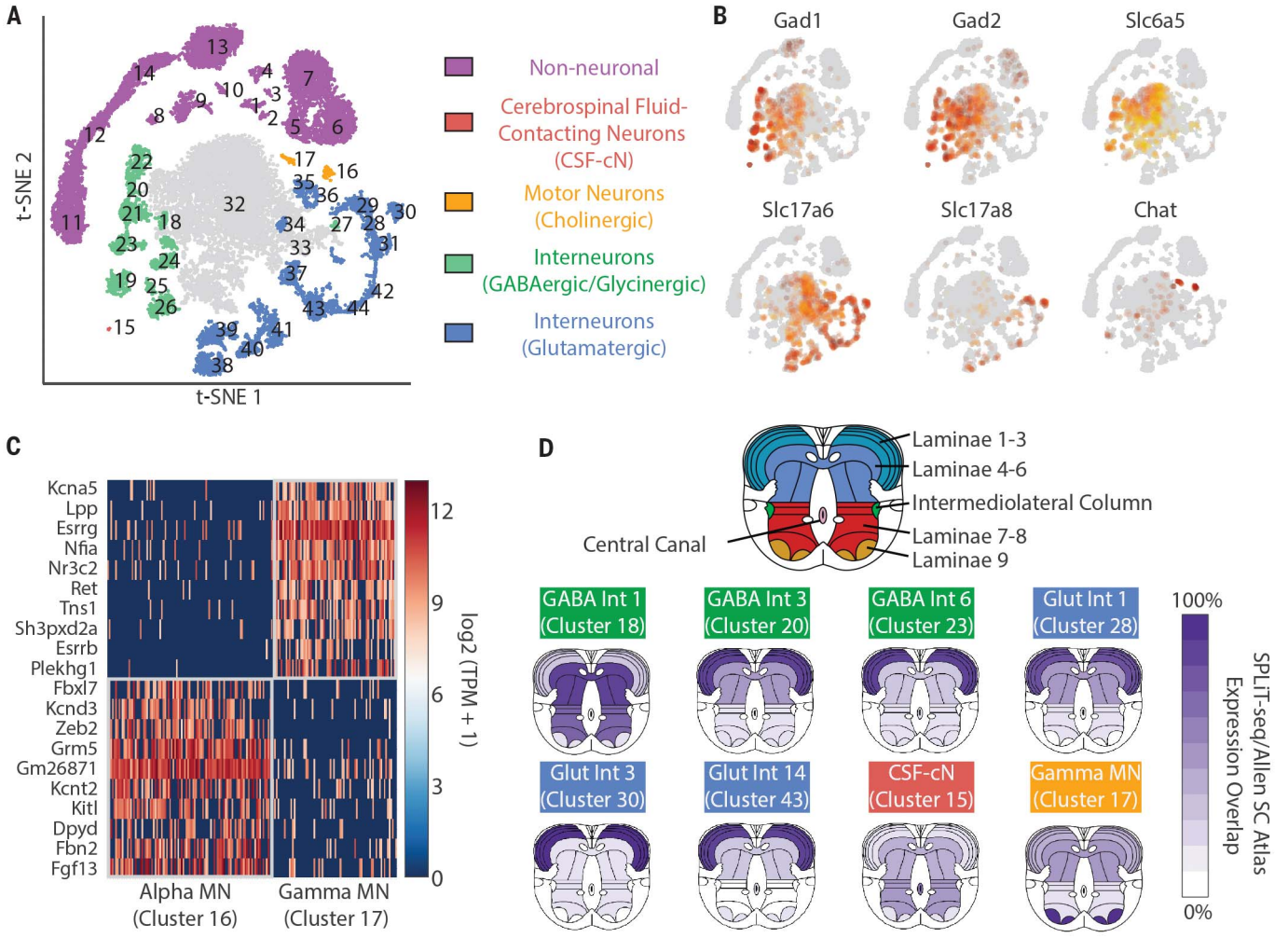

Fig. 5. Gene expression patterns and spatial origin of cell types in the spinal cord. (**A**) Reclustering spinal cord nuclei resulted in 30 neuronal and 14 non-neuronal clusters. (**B**) GABAergic neurons were defined by expression of *Gad1* and *Gad2*. A subset of GABAergic neurons are also glycinergic, based on expression of *Slc6a5*. Glutamatergic neurons were defined by expression of VGLUT2 (*Slc17a6*), whereas cholinergic motor neurons express *Chat*. (**C**) Novel gene markers distinguish gamma motor neurons from alpha motor neurons. (**D**) Inferred spatial origin of neuronal clusters within the spinal cord. We analyzed the Allen Spinal Cord Atlas expression patterns of the top 10 enriched genes in each cluster. Dark purple indicates expression of all 10 genes in the given region, whereas white indicates that none of the 10 genes were expressed in the given region.

Discussion

In this work, we profiled hundreds of thousands of cells using only basic laboratory equipment with a library preparation cost of ~$0.01 per cell (fig. S21 and table S11). In our analysis of more than 150,000 single-nucleus transcriptomes from two early postnatal stages, we identified 69 types of cells in the brain and 44 types in the spinal cord. We defined many new molecular markers for specific cell types and explored gene expression in four different developmental lineages.

SPLiT-seq's compatibility with fixed cells and fixed nuclei overcomes challenges faced by other scRNA-seq methods. Fixation can reduce perturbations to endogenous gene expression during cell handling (*42*) and makes it possible to store cells for future experiments. Moreover, the use of nuclei bypasses the need to obtain intact single cells, which can be challenging for many complex tissues. SPLiT-seq's compatibility with formaldehyde-fixed nuclei suggests that it may be used to profile single nuclei from formalin-fixed, paraffin-embedded tissue (*43*).

SPLiT-seq enables flexible and scalable cell and sample multiplexing. The use of the first-round barcode as a sample identifier makes it possible to profile a large number and variety of samples in parallel, thus minimizing batch effects. As the number of unique barcodes grows exponentially with the number of barcoding rounds, larger numbers of cells than presented here could be processed by adding a fifth barcoding round or by switching to a 384-well plate format. Although for such large cell numbers, sequencing cost may currently be forbidding, it is easy to imagine applications, such as targeted sequencing of gene panels, which would even now benefit from very large cell numbers and only require shallow sequencing depth.

Our hope is that the increased scale and accessibility provided by the low cost and minimal equipment requirements of SPLiT-seq will further accelerate the widespread adoption of scRNA-seq.

REFERENCES AND NOTES

1. S. Picelli *et al.*, *Nat. Methods* **10**, 1096–1098 (2013).
2. T. Hashimshony, F. Wagner, N. Sher, I. Yanai, *Cell Reports* **2**, 666–673 (2012).
3. D. A. Jaitin *et al.*, *Science* **343**, 776–779 (2014).
4. E. Z. Macosko *et al.*, *Cell* **161**, 1202–1214 (2015).
5. A. M. Klein *et al.*, *Cell* **161**, 1187–1201 (2015).
6. G. X. Y. Zheng *et al.*, *Nat. Commun.* **8**, 14049 (2017).
7. B. Tasic *et al.*, *Nat. Neurosci.* **19**, 335–346 (2016).
8. A. Zeisel *et al.*, *Science* **347**, 1138–1142 (2015).
9. S. Marques *et al.*, *Science* **352**, 1326–1329 (2016).
10. S. Darmanis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7285–7290 (2015).
11. B. B. Lake *et al.*, *Science* **352**, 1586–1590 (2016).
12. A. K. Shalek *et al.*, *Nature* **498**, 236–240 (2013).
13. D. Grün *et al.*, *Nature* **525**, 251–255 (2015).
14. V. Moignard *et al.*, *Nat. Biotechnol.* **33**, 269–276 (2015).

Downloaded from http://science.sciencemag.org/ on July 15, 2019

15. A. S. Venteicher *et al.*, *Science* **355**, eaai8478 (2017).
16. I. Tirosh *et al.*, *Science* **352**, 189–196 (2016).
17. L. Sang, H. A. Coller, J. M. Roberts, *Science* **321**, 1095–1100 (2008).
18. C. Zheng *et al.*, *Cell* **169**, 1342–1356.e16 (2017).
19. Materials and methods are provided as supplementary materials.
20. S. E. Hickman *et al.*, *Nat. Neurosci.* **16**, 1896–1905 (2013).
21. O. Matcovitch-Natan *et al.*, *Science* **353**, aad8670 (2016).
22. C. Trapnell *et al.*, *Nat. Biotechnol.* **32**, 381–386 (2014).
23. S. W. Levison, J. E. Goldman, *Neuron* **10**, 201–212 (1993).
24. E. S. Lein *et al.*, *Nature* **445**, 168–176 (2007).
25. Allen Institute for Brain Science, Developing Mouse Brain Atlas (2008); available at http://developingmouse.brain-map.org/.
26. T. Iwano, A. Masuda, H. Kiyonari, H. Enomoto, F. Matsuzaki, *Development* **139**, 3051–3062 (2012).
27. C. Zhao, W. Deng, F. H. Gage, *Cell* **132**, 645–660 (2008).
28. A. Lavado, O. V. Lagutin, L. M. L. Chow, S. J. Baker, G. Oliver, *PLOS Biol.* **8**, e1000460 (2010).
29. F. Bonnet *et al.*, *J. Biol. Chem.* **271**, 4373–4380 (1996).
30. S. Herculano-Houzel, *Front. Hum. Neurosci.* **3**, 31 (2009).
31. K. Nakashima, H. Umeshima, M. Kengaku, *Dev. Dyn.* **244**, 748–758 (2015).
32. A. Sudarov, A. L. Joyner, *Neural Dev.* **2**, 26 (2007).
33. J. C. Chang *et al.*, *J. Neuropathol. Exp. Neurol.* **74**, 261–272 (2015).
34. K. Schilling, J. Oberdick, F. Rossi, S. L. Baader, *Histochem. Cell Biol.* **130**, 601–615 (2008).
35. J. Altman, S. A. Bayer, *J. Comp. Neurol.* **257**, 477–489 (1987).
36. S. M. Maricich, K. Herrup, *J. Neurobiol.* **41**, 281–294 (1999).
37. G. Weisheit *et al.*, *Eur. J. Neurosci.* **24**, 466–478 (2006).
38. Y. L. Petracca *et al.*, *Development* **143**, 880–891 (2016).
39. A. Enjin *et al.*, *J. Comp. Neurol.* **518**, 2284–2304 (2010).
40. M. Lalancette-Hebert, A. Sharma, A. K. Lyashchenko, N. A. Shneider, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E8316–E8325 (2016).
41. Allen Institute for Brain Science, Allen Mouse Spinal Cord Atlas (2008); available at http://mousespinal.brain-map.org/imageseries/showref.html.
42. B. Lacar *et al.*, *Nat. Commun.* **7**, 11022 (2016).
43. E. R. Thomsen *et al.*, *Nat. Methods* **13**, 87–93 (2016).

ACKNOWLEDGMENTS

We thank T. N. Nguyen for help with cluster identity assignment. **Funding:** This work was supported by NIH R01CA207029 and NSF CCF-1317653 to G.S. and NIH R01NS064404 and R21NS086500 to S.H.P. Z.Y., L.T.G., and B.T. were supported by the Allen Institute for Brain Science. C.M.R. was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award number TL1 TR002318. **Author contributions:** A.B.R., C.M.R., R.A.M., and G.S. were mainly responsible for developing the method. A.B.R., C.M.R., R.A.M., A.K., P.S., S.M., W.C., and G.S. designed experiments. A.B.R., C.M.R., R.A.M., A.K., P.S., S.M., and W.C. performed experiments. Brain and spinal cord extraction was performed and supported by D.J.P., S.H.P., and D.L.S. A.B.R., C.M.R., Z.Y., and L.G. performed data analysis. Cluster annotation was performed by A.B.R., C.M.R., Z.Y., L.G., and B.T. A.B.R., C.M.R., B.T., and G.S. wrote the manuscript. **Competing interests:** A.B.R., R.M., and G.S. are inventors on a patent application (14/941,433) submitted by the University of Washington that covers the SPLiT-seq method. **Data availability:** All relevant sequencing files were deposited to the Gene Expression Omnibus under accession number GSE110823.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/360/6385/176/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S21
Tables S1 to S12
References (*44*–*92*)

1 February 2017; resubmitted 30 September 2017
Accepted 26 February 2018
Published online 15 March 2018
10.1126/science.aam8999

Downloaded from http://science.sciencemag.org/ on July 15, 2019

Science

Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding

Alexander B. Rosenberg, Charles M. Roco, Richard A. Muscat, Anna Kuchina, Paul Sample, Zizhen Yao, Lucas T. Graybuck, David J. Peeler, Sumit Mukherjee, Wei Chen, Suzie H. Pun, Drew L. Sellers, Bosiljka Tasic and Georg Seelig

Science **360** (6385), 176-182.
DOI: 10.1126/science.aam8999originally published online March 15, 2018

Identifying single-cell types in the mouse brain

The recent development of single-cell genomic techniques allows us to profile gene expression at the single-cell level easily, although many of these methods have limited throughput. Rosenberg *et al.* describe a strategy called split-pool ligation-based transcriptome sequencing, or SPLiT-seq, which uses combinatorial barcoding to profile single-cell transcriptomes without requiring the physical isolation of each cell. The authors used their method to profile >100,000 single-cell transcriptomes from mouse brains and spinal cords at 2 and 11 days after birth. Comparisons with in situ hybridization data on RNA expression from Allen Institute atlases linked these transcriptomes with spatial mapping, from which developmental lineages could be identified.

Science, this issue p. 176

ARTICLE TOOLS http://science.sciencemag.org/content/360/6385/176

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2018/03/14/science.aam8999.DC1

REFERENCES This article cites 88 articles, 23 of which you can access for free
http://science.sciencemag.org/content/360/6385/176#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.

Downloaded from http://science.sciencemag.org/ on July 15, 2019