Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution

Samuel G. Rodrigues[1,2,3]*, Robert R. Stickels[3,4,5]*, Aleksandrina Goeva[3],
Carly A. Martin[3], Evan Murray[3], Charles R. Vanderburg[3], Joshua Welch[3],
Linlin M. Chen[3], Fei Chen[3]†‡, and Evan Z. Macosko[3,6]†‡

Spatial positions of cells in tissues strongly influence function, yet a high-throughput, genome-wide readout of gene expression with cellular resolution is lacking. We developed Slide-seq, a method for transferring RNA from tissue sections onto a surface covered in DNA-barcoded beads with known positions, allowing the locations of the RNA to be inferred by sequencing. Using Slide-seq, we localized cell types identified by single-cell RNA sequencing datasets within the cerebellum and hippocampus, characterized spatial gene expression patterns in the Purkinje layer of mouse cerebellum, and defined the temporal evolution of cell type–specific responses in a mouse model of traumatic brain injury. These studies highlight how Slide-seq provides a scalable method for obtaining spatially resolved gene expression data at resolutions comparable to the sizes of individual cells.

T he functions of complex tissues are fundamentally tied to the organization of their resident cell types. Multiplexed in situ hybridization and sequencing-based approaches can measure gene expression with subcellular spatial resolution (1–3) but require specialized knowledge and equipment, as well as the upfront selection of gene sets for measurement. By contrast, technologies for spatially encoded RNA sequencing with barcoded oligonucleotide capture arrays are limited to resolutions in hundreds of micrometers (4), which are insufficient to detect important tissue features.

To develop our Slide-seq technology for high-resolution genome-wide expression analysis, we first packed uniquely DNA-barcoded 10-μm microparticles ("beads")—similar to those used in the Drop-seq approach for single-cell RNA sequencing (scRNA-seq) (5)—onto a rubber-coated glass coverslip to form a monolayer we termed a "puck" (fig. S1). We found that each bead's distinct barcode sequence could be determined via SOLiD (sequencing by oligonucleotide ligation and detection) chemistry (Fig. 1A and fig. S1) (6–8). We next developed a protocol wherein 10-μm fresh-frozen tissue sections were transferred onto the dried bead surface via cryosectioning (7). mRNA released from the tissue was captured onto the

beads for preparation of 3′-end, barcoded RNA-seq libraries (5) (Fig. 1B). Clustering of individual bead profiles from a coronal section of mouse hippocampus (7) yielded assignments reflecting known positions of cell types in the tissue (Fig. 1C). Very fine spatial features were resolved, including the single-cell ependymal cell layer between the central ventricle and the habenula in the mouse brain (Fig. 1C, inset). Moreover, Slide-seq could be applied to a range of tissues, including the cerebellum and olfactory bulb, where layered tissue architectures were immediately detectable (Fig. 1D and fig. S2), as well as the liver and kidney, where the identified clusters revealed hepatocyte zonation patterns (9) and the cellular constituents of the nephron, respectively. Slide-seq on postmortem human cerebellum was also successful in capturing the same architectural features observed in the cognate mouse tissue (fig. S3). Expression measurements by Slide-seq agreed with those from bulk mRNA-seq and scRNA-seq, and average mRNA transcript capture per cell was consistent across tissues and experiments (fig. S4). Finally, we found no detectable difference in the dimensions of brain structures analyzed by Slide-seq and fluorescence in situ hybridization (fig. S5), implying that mRNA is transferred from the tissue to the beads with minimal lateral diffusion.

To map scRNA-seq cell types onto Slide-seq data, we developed a computational approach called non-negative matrix factorization regression (NMFreg) that reconstructs expression of each Slide-seq bead as a weighted combination of cell type signatures defined by scRNA-seq (Fig. 2A). Application of NMFreg to a coronal mouse cerebellar puck recapitulated the spatial distributions of classical neuronal and non-neuronal cell types (10), such as granule cells, Golgi interneurons, unipolar brush cells, Purkinje

cells, and oligodendrocytes (Fig. 2B and fig. S6A). The mapping by NMFreg was found to be reliable across a range of factor numbers and random restarts (fig. S6, B and C). We found that 65.8 ± 1.4% of beads could be matched with a single cell type (7), whereas 32.6 ± 1.2% showed mRNA from two cell types (mean ± SD, $N = 7$ cerebellar pucks) (Fig. 2C and fig. S7). The high spatial resolution of Slide-seq was key to mapping cell types: When data were aggregated into larger feature sizes, cell types in heterogeneous regions of tissue could not be resolved (fig. S8). Slide-seq collects a two-dimensional (2D) spatial sample of 3D tissue volumes, and thus caution should be taken when making absolute counting measurements throughout the 3D volume in the absence of proper stereological controls and sampling methods (11).

We first sequenced pucks capturing 66 sagittal tissue sections from a single dorsal mouse hippocampus, covering a volume of 9 mm³, with ~10-μm resolution in the dorsal-ventral and anterior-posterior axes and ~20-μm resolution (alternate 10-μm sections) in the medial-lateral axis (fig. S9, A to D). 1.5 million beads, of which 770,000 could be associated with a single scRNA-seq–defined cell type using NMFreg, were mapped to short-read sequencing data in the volume. We computationally co-registered pucks along the medial-lateral axis, allowing for visualization of the cell types and gene expression in the hippocampus in three dimensions (Fig. 2D and fig. S9, E and F, and movie S1). We plotted metagenes composed of previously defined markers (10) for the dentate gyrus, CA2, CA3, a subiculum subpopulation, an anteriorly localized CA1 subset (exemplified by the marker Tenm3), and cells undergoing mitosis and neurogenesis. The metagenes were highly expressed and specific for the expected regions (Fig. 2E), confirming the ability of Slide-seq to localize both common cell types as well as finer cellular subpopulations. The entire experimental processing of these 66 pucks (excluding puck generation) required ~40 person-hours (7) and only standard experimental apparatus.

We then developed a nonparametric, kernel-free algorithm to identify genes with spatially nonrandom distribution across the puck (fig. S10) (7). Application of this algorithm to coronally sliced cerebellum identified Ogfrl1, Prkcd, and Atp2b1 as highly localized to a region just inferior to the cerebellum (fig. S11A). We found Ogfrl1, in particular, to be a specific and previously unrecognized marker for parvalbumin interneurons in the molecular and fusiform layers of the dorsal cochlear nucleus (fig. S11B), likely the cartwheel cells of the dorsal cochlear nucleus (12, 13). Our algorithm also identified Rasgrf1 as expressed only in granule cells anterior to the primary fissure [fig. S11, C (cyan) and D (left)] (14), and further analysis revealed four previously uncharacterized genes expressed only posterior to the primary fissure (7) [table S2 and fig. S11, C (yellow) and D (right)].

The cerebellum is marked by parasagittal bands of gene expression in the Purkinje layer that correlate with heterogeneity in Purkinje cell physiology and projection targets (15–18). Several

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [3]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. [4]Graduate School of Arts and Sciences, Harvard University, Cambridge, MA 02138, USA. [5]Division of Medical Science, Harvard Medical School, Boston, MA 02115, USA. [6]Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA.
*These authors contributed equally to this work.
†These authors contributed equally to this work.
‡Corresponding author. Email: chenf@broadinstitute.org (F.C.); emacosko@broadinstitute.org (E.Z.M.)

genes, including *Aldoc* (also known as the antigen of the zebrin II antibody), show similar or complementary parasagittal expression (*17, 19, 20*), but the extent of this form of expression variation is unknown, and these patterns have not previously been identified in single-cell sequencing studies. Using the spatial information collected by Slide-seq, we identified 669 spatially non-random genes in the Purkinje layer (table S2), of which 126 appeared either correlated or anti-correlated with the zebrin pattern, using *Aldoc* and *Plcb4* as markers of zebrin II–positive or zebrin II–negative bands, respectively (Fig. 3A). Among the anticorrelated genes were four adenosine triphosphatases and four potassium channels, including some which may explain differences in electrophysiology between zebrin II–positive and zebrin II–negative Purkinje neurons (table S2). Moreover, we identified several other patterns, besides the zebrin pattern, of spatial gene expression. Whereas most of the identified genes displayed a pattern consistent with zebrin II staining (Fig. 3, B and C), several were exclusively expressed in or excluded from the

vestibulocerebellar region (lobules IX and X) (*21, 22*) (Fig. 3D and table S2), confirming that lobules IX and X have a distinct program of gene expression. Still other genes showed either exclusive expression in [e.g., *B3galt5* (*7, 23*)] or exclusion from (e.g., *Gnai1*) lobules IX and X and lobules VI and VII (fig. S11, E and F), suggesting that these regions might share a pattern of gene expression, despite the disparate cognitive roles associated with them (table S2) (*24*). Finally, although only Purkinje cells have previously been associated with the *Aldoc* pattern, we found that *Mybpc1*, a Bergmann cell marker previously studied primarily in muscle, appears to have a zebrin pattern of expression in both Slide-seq (fig. S11G) and in situ data (fig. S11H). We thus conclude that the banded gene expression patterns divide many cerebellar cell types—including Purkinje cells, Bergmann glia, and granule cells—into spatially defined subpopulations, which was not indicated in previous single-cell sequencing studies (*10, 25*).

Finally, we applied Slide-seq to quantify the brain's response to traumatic brain injury over

time. Cortical injuries were visualized in Slide-seq data by the presence of hemoglobin transcripts 2 hours after injury (Fig. 4A) or by transcripts of *Vim*, *Gfap*, and *Ctsd* 3 days and 2 weeks after injury (Fig. 4, B and C) (*7*). *Vim*, *Gfap*, and *Ctsd* were chosen because they are known markers of the astrocytic (*Vim* and *Gfap*) or microglial (*Vim* and *Ctsd*) responses that were found to be highly up-regulated at the injury in the Slide-seq data (fig. S13). We applied an algorithm to identify all genes that correlate spatially with those transcripts (*7*). At the 2-hour time point, only *Fos* and ribosomal RNA (rRNA) (*26*) were found to correlate spatially with the injury (Fig. 4A and fig. S14). By contrast, at the 3-day time point, we found microglia- and macrophage-assigned beads localized to the injury, bordered by a distinct layer of cells (thickness: $92.4 \pm 11.3\ \mu m$, mean $\pm$ SE, $N = 3$ measurements) expressing mitosis-associated factors, followed by a layer of astrocyte-assigned beads (Fig. 4D). Finally, at the 2-week time point, we observed microglia- and macrophage-assigned beads filling the injury, surrounded by an astrocytic scar (thickness: $36.6 \pm$
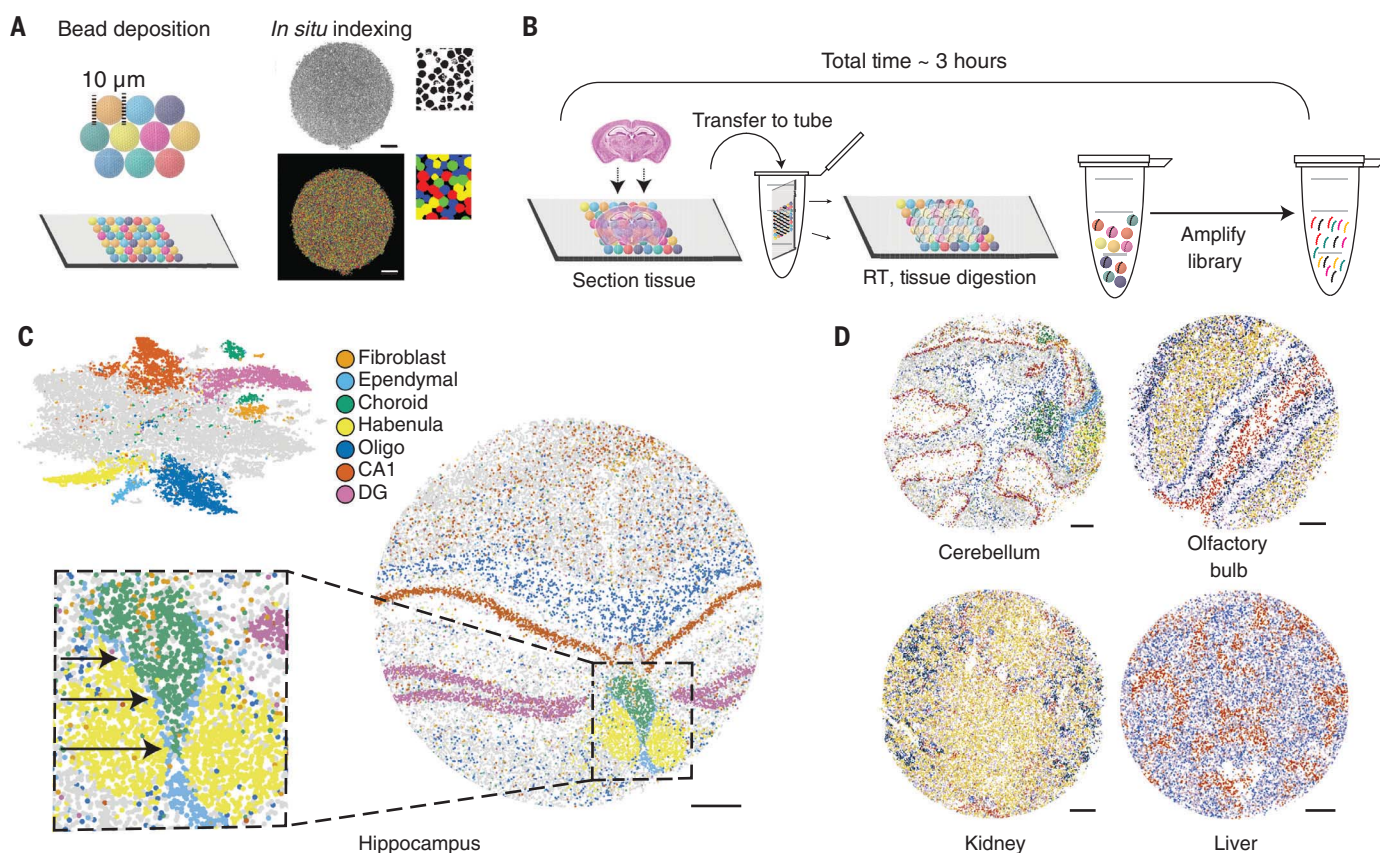
Fig. 1. High-resolution RNA capture from tissue by Slide-seq.
(**A**) (Left) Schematic of array generation. A monolayer of randomly deposited, DNA barcoded beads (a "puck") is spatially indexed by SOLiD sequencing. (Top right) Representative puck with sequenced barcodes shown in black. (Bottom right) Composite image of the same puck colored by the base calls for a single base of SOLiD sequencing. (**B**) Schematic of the sample preparation procedure. RT, reverse transcription. (**C**) (Top left) t-distributed stochastic neighbor embedding (tSNE) representation of Slide-seq beads from a coronal mouse hippocampus slice with colors indicating clusters. GD, dentate gyrus. (Right) The spatial position of each bead is shown, colored by the cluster assignments shown in the tSNE. (Bottom left) Inset indicating the position of a single-cell-thick ependymal cell layer (arrows). (**D**) As in (C), but for the indicated tissue type (see fig. S2 for clustering and cluster identities). All scale bars: 500 μm.

13.4 μm, mean ± SE, $N = 6$ measurements), with evidence of microglia (but not macrophages) penetrating 39 ± 17.8 μm (mean ± SE, $N = 6$ measurements) through the astrocytic scar and into neuron-rich regions (Fig. 4, E and F). Macrophages were visualized using *Lyz2*, a specific marker for macrophages and granulocytes; however, we interpret this as a marker of macrophages, because other granulocyte-specific markers were not found to colocalize with *Gfap*, *Ctsd*, and *Vim*.

To investigate other changes in gene expression between the 3-day and 2-week time points, we identified genes that correlated spatially with *Vim*, *Gfap*, and *Ctsd* at the 3-day time point or the 2-week time point (7). Applying gene ontology
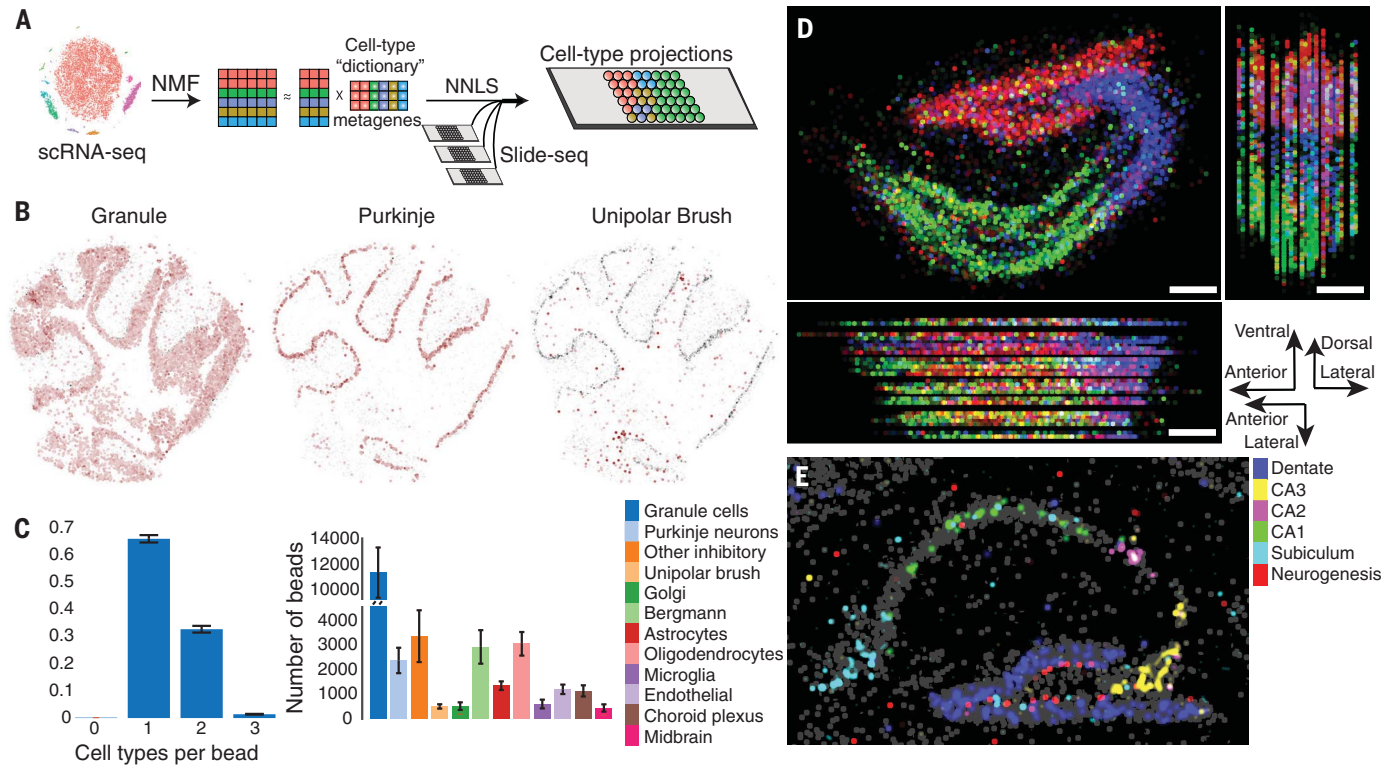
Fig. 2. Localization of cell types in the cerebellum and hippocampus using Slide-seq. (**A**) Schematic for assigning cell types from scRNA-seq datasets to Slide-seq beads using NMF and NNLS (non-negative least squares) regression (NMFreg). (**B**) Loadings of individual cell types, defined by scRNA-seq of the cerebellum (10), on each bead of one 3-mm-diameter coronal cerebellar puck (red, cell type location; gray, Purkinje loadings plotted as a counterstain). Other cell types are in fig. S6. (**C**) (Left) Number of cell types assigned per bead (fig. S3). (Right) Number of beads called as each scRNA-seq–defined cell type for cerebellar pucks (mean ± SD, $N = 7$ pucks). (**D**) Projections of hippocampal volume with NMFreg cell type calls for CA1 (green), CA2 and CA3 (blue), and dentate gyrus (red). Top left: Sagittal projection. Top right: Coronal projection. Bottom left: Horizontal projection. Bottom right: Axis orientations for each of the projections. (**E**) Composite image of metagenes for six different cell types. All scale bars: 250 μm. All metagenes are listed in table S2.
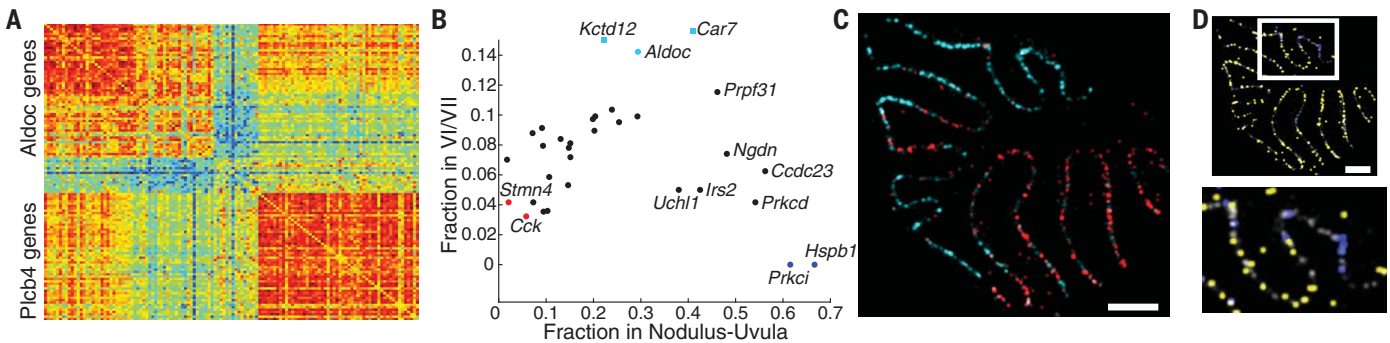

Fig. 3. Identification of variation in cerebellar gene expression by Slide-seq. (**A**) Heatmap illustrating the separation of Purkinje-expressed genes into two clusters by spatial gene correlation. The i,jth entry is the number of genes found to overlap with both genes i and j in the Purkinje cluster (7). (**B**) For genes with significant expression ($P < 0.001$, Fisher exact test) in the nodulus-uvula region (7), the fraction of reads localized to the nodulus and uvula as well as to the VI/VII boundary is shown. *Pcp4*, a ubiquitous marker for Purkinje cells, is in gray. (**C**) An *Aldoc* metagene (cyan); a *Cck* metagene (red). (**D**) A *H2-D1* metagene (yellow); a *Hspb1* metagene (blue). The bottom image is a close-up view of the boxed region. All scale bars: 250 μm. All metagenes are listed in table S2.
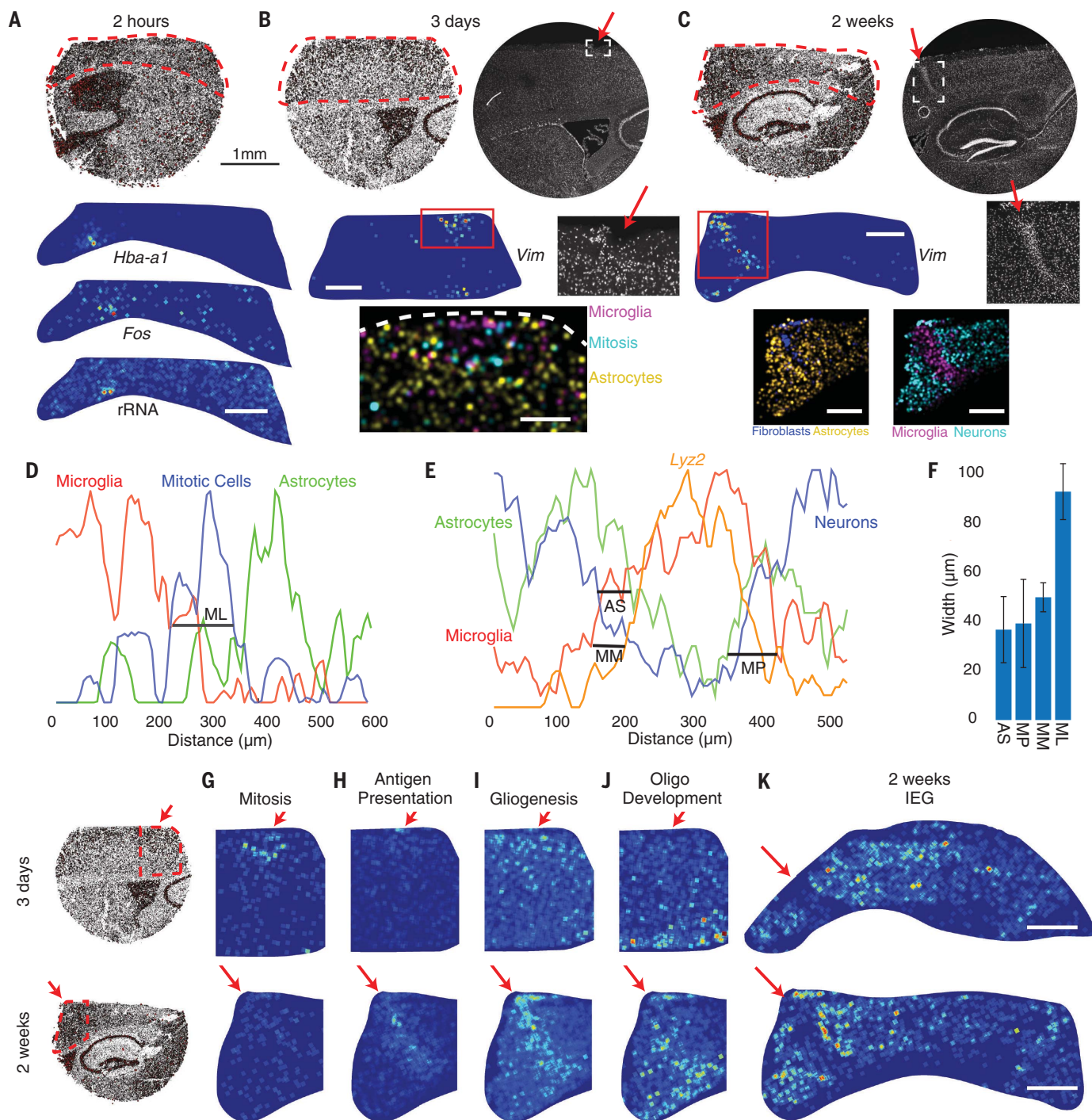
Fig. 4. Slide-seq identifies local transcriptional responses to injury.
(**A**) (Top) All mapped beads for a coronal hippocampal slice from a mouse euthanized 2 hours after injury, with circle radius proportional to transcripts. (Bottom) Genes marking the injury. (**B**) As in (A) for a mouse euthanized 3 days after injury. (Top and middle right) DAPI (4′,6-diamidino-2-phenylindole) stained image of an adjacent slice. Panels with black backgrounds show NMFreg cell types as density plots. Scale bar: 250 μm (7). (**C**) As in (B) for a mouse euthanized 2 weeks after injury. Bottom scale bars: 500 μm. (**D**) Spatial density profiles for the puck in (B) (7). (**E**) Spatial density profiles for the puck in (C). *Lyz2* is plotted as a marker

of macrophages. The vertical axis in (D) and (E) represents cell type density in arbitrary units (7). ML, mitotic layer; AS, astrocytic scar; MM, microglia-macrophage distance; MP, microglia penetration. (**F**) Thickness of features in (D) and (E) (mean ± SD, N = 6 measurements for scar, N = 6 for penetration, N = 3 for mitosis layer). (**G** to **J**) Gene ontology–derived metagenes for the puck in (B) (top) or (C) (bottom). (**K**) The IEG metagene (table S2) for two 2-week pucks. In (G) to (K), warmer colors correspond to greater metagene counts. The 1-mm scale bar in (A) pertains to the circular images in (A) to (C). All scale bars for images with blue backgrounds: 500 μm. Red arrows indicate the injury.

analysis to these gene sets revealed enrichment of annotations relating to chromatid segregation, mitosis, and cell division at the 3-day time point (Fig. 4G), as well as enrichment relating to the immune response (Fig. 4H), gliogenesis (Fig. 4I), and oligodendrocyte development (Fig. 4J) at the 2-week time point. This finding suggests that cell proliferation occurs in the first few days after injury and transitions to differentiation on the order of weeks. For example, although the degree to which oligodendrocyte progenitor cells differentiate into oligodendrocytes after a focal gray matter injury is controversial (27), we confirmed that both *Sox4* and *Sox10* localize to the region surrounding the injury at the 2-week time point, indicating the presence of immature oligodendrocytes (fig. S15). We also discovered evidence that several immediate early genes, including highly neuron-specific genes such as *Npas4* (table S2), are up-regulated in a region of width measuring 0.72 ± 0.19 mm (mean $\pm$ SE, $N = 4$ measurements) around the injury at both the 3-day and the 2-week time points (28–30) (Fig. 4K and table S2), suggesting persistent effects of the injury on neural activity in a large area around the injury.

Our study demonstrates that Slide-seq enables the spatial analysis of gene expression in frozen tissue with high spatial resolution and scalability to large tissue volumes. Slide-seq is easily integrated with large-scale scRNA-seq datasets and facilitates discovery of spatially defined gene expression patterns in normal and diseased tissues. The primary cost of Slide-seq is the cost of short-read sequencing, which is ~$200 to $500 for the pucks presented here. As the cost of sequencing drops further, we expect to be able to scale Slide-seq to entire organs or even entire organisms. We anticipate that Slide-seq will be instrumental in positioning molecularly defined cell types in complex tissues and defining molecular pathways involved in neuropathological states.

REFERENCES AND NOTES

1. S. Shah, E. Lubeck, W. Zhou, L. Cai, *Neuron* **92**, 342–357 (2016).
2. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, *Science* **348**, aaa6090 (2015).
3. J. H. Lee et al., *Science* **343**, 1360–1363 (2014).
4. P. L. Ståhl et al., *Science* **353**, 78–82 (2016).
5. E. Z. Macosko et al., *Cell* **161**, 1202–1214 (2015).
6. J. H. Lee et al., *Nat. Protoc.* **10**, 442–458 (2015).
7. Materials and methods are available as supplementary materials.
8. K. L. Gunderson et al., *Genome Res.* **14**, 870–877 (2004).
9. K. B. Halpern et al., *Nature* **542**, 352–356 (2017).
10. A. Saunders et al., *Cell* **174**, 1015–1030.e16 (2018).
11. M. J. West, *Neurobiol. Aging* **14**, 275–285 (1993).
12. L. O. Trussell, D. Oertel, in *The Mammalian Auditory Pathways: Synaptic Organization and Microcircuits*, D. N. Oliver, N. B. Cant, R. R. Fay, A. N. Popper, Eds. (vol. 65 of Springer Handbook of Auditory Research Series, Springer, 2018), pp. 73–99.
13. M. T. Roberts, L. O. Trussell, *J. Neurophysiol.* **104**, 2462–2473 (2010).
14. E. S. Lein et al., *Nature* **445**, 168–176 (2007).
15. C. Gravel, R. Hawkes, *J. Comp. Neurol.* **291**, 79–102 (1990).
16. J. Xiao et al., *PLOS ONE* **9**, e105633 (2014).
17. N. H. Barmack, Z. Qian, J. Yoshimura, *J. Comp. Neurol.* **427**, 235–254 (2000).
18. N. L. Cerminara, E. J. Lang, R. V. Sillitoe, R. Apps, *Nat. Rev. Neurosci.* **16**, 79–93 (2015).
19. A. Demilly, S. L. Reeber, S. A. Gebre, R. V. Sillitoe, *Cerebellum* **10**, 409–421 (2011).
20. G. Brochu, L. Maler, R. Hawkes, *J. Comp. Neurol.* **291**, 538–552 (1990).
21. K. Maekawa, J. I. Simpson, *J. Neurophysiol.* **36**, 649–666 (1973).
22. D. R. W. Wylie, M. R. Brown, I. R. Winship, N. A. Crowder, K. G. Todd, *J. Comp. Neurol.* **465**, 179–194 (2003).
23. C. Chung et al., *PLOS Genet.* **12**, e1006042 (2016).
24. C. J. Stoodley, J. D. Schmahmann, *Cortex* **46**, 831–844 (2010).
25. R. A. Carter et al., *Curr. Biol.* **28**, 2910–2920.e2 (2018).
26. P. D. Storer, K. J. Jones, *J. Comp. Neurol.* **458**, 326–333 (2003).
27. K. L. Adams, V. Gallo, *Nat. Neurosci.* **21**, 9–15 (2018).
28. A. M. Kenney, J. D. Kocsis, *J. Neurosci.* **18**, 1318–1328 (1998).
29. G. A. Robinson, *Brain Res. Mol. Brain Res.* **24**, 43–54 (1994).
30. J. Honkaniemi, S. M. Sagar, I. Pyykönen, K. J. Hicks, F. R. Sharp, *Brain Res. Mol. Brain Res.* **28**, 157–163 (1995).
31. S. G. Rodriques, L. M. Chen, broadchenf/Slideseq: Slide-seq 2019 Archive, Version 1.0, Zenodo (2019); https://doi.org/10.5281/zenodo.2571615.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/363/6434/1463/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S3
References (32–39)
Movie S1