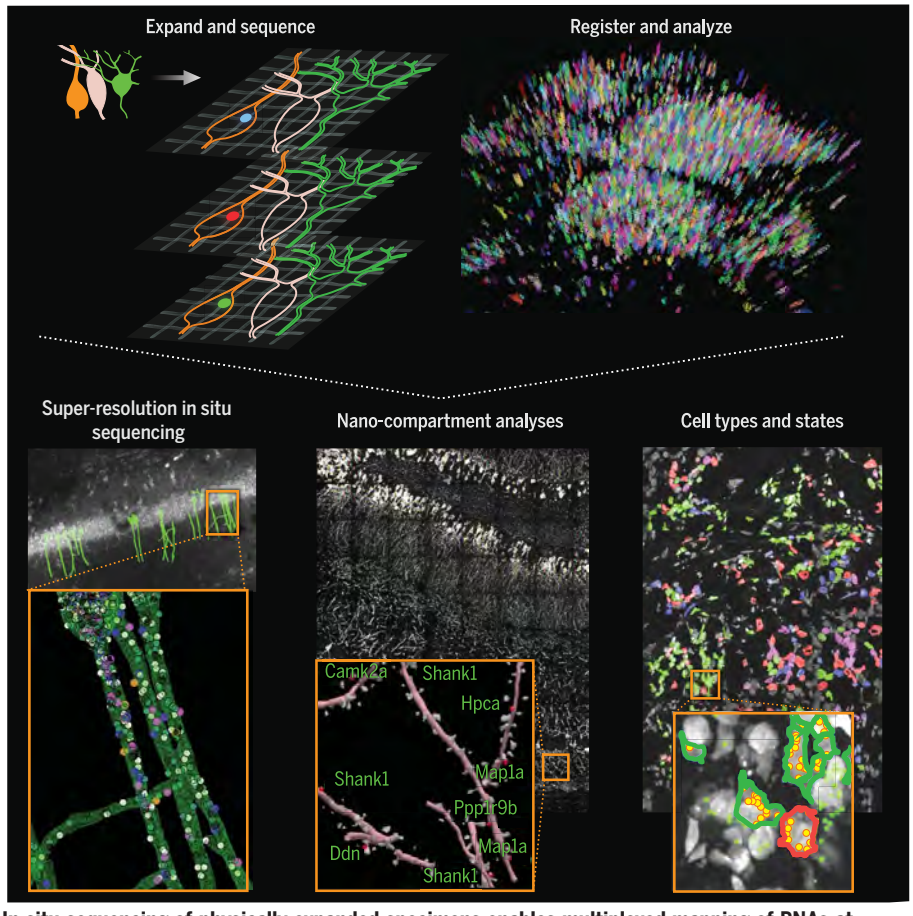
RESEARCH ARTICLE SUMMARY

IN SITU SEQUENCING

Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems

Shahar Alon*, Daniel R. Goodwin*, Anubhav Sinha*, Asmamaw T. Wassie*, Fei Chen*, Evan R. Daugharthy†, Yosuke Bando, Atsushi Kajita, Andrew G. Xue, Karl Marrett, Robert Prior, Yi Cui, Andrew C. Payne, Chun-Chen Yao, Ho-Jun Suk, Ru Wang, Chih-Chieh (Jay) Yu, Paul Tillberg, Paul Reginato, Nikita Pak, Songlei Liu, Sukanya Punthambaker, Eswar P. R. Iyer, Richie E. Kohman, Jeremy A. Miller, Ed S. Lein, Ana Lako, Nicole Cullen, Scott Rodig, Karla Helvie, Daniel L. Abravanel, Nikhil Wagle, Bruce E. Johnson, Johanna Klughammer, Michal Slyper, Julia Waldman, Judit Jané-Valbuena, Orit Rozenblatt-Rosen, Aviv Regev, IMAXT Consortium, George M. Church‡§, Adam H. Marblestone§, Edward S. Boyden‡§

INTRODUCTION: Cells and tissues are made up of diverse molecular building blocks, organized with nanoscale precision over extended length scales. Newly developed techniques that enable highly multiplexed, nanoscale, and subcellular analysis of such systems are required. Although much progress has been made on methods for multiplexed RNA imaging, these methods have been limited in their spatial precision, especially in the context of three-dimensional systems such as tissues. Because of this limitation, interrogation of tissues has been performed with either high spatial resolution or high molecular multiplexing capacity, but not both.



In situ sequencing of physically expanded specimens enables multiplexed mapping of RNAs at nanoscale, subcellular resolution throughout intact tissues. (Top) Schematics of physical expansion and in situ sequencing (left) and image analysis (right). (Bottom) Characterization of nanoscale transcriptomic compartmentalization in mouse hippocampal neuron dendrites and spines (left and middle) and maps of cell types and states in a metastatic human breast cancer biopsy (right).

RATIONALE: We reasoned that physically expanding specimens by adapting expansion microscopy could help support spatially precise in situ sequencing. The physical expansion of specimens provides two benefits: First, it enables ordinary microscopes to achieve nanoscale effective resolution. Second, by anchoring RNA molecules to a polymer network, digesting away other molecules, and then expanding the polymer in water, RNAs become more accessible. By creating a chemical process that enables enzymatic reactions to proceed in expanded specimens, we enabled in situ fluorescent sequencing of RNA with high spatial precision, which we term expansion sequencing (ExSeq). We developed both untargeted (i.e., not restricted to a predefined set of genes) and targeted versions of ExSeq.

RESULTS: Using untargeted ExSeq, we showed the presence of transcripts that retain their introns, transcription factors, and long noncoding RNAs in mouse hippocampal neuron dendrites. Using targeted ExSeq, we observed layer-specific cell types across the mouse visual cortex and RNAs in nanoscale compartments of hippocampal pyramidal neurons, such as dendritic spines and branches. We found that spines could exhibit distributions of mRNAs different from those exhibited by adjacent dendrites. Moreover, we found patterns of similarity between the dendritic profiles of RNAs in different types of hippocampal neurons. In a human metastatic breast cancer biopsy, we mapped how cell types expressed genes differently as a function of their distance from other cell types, identifying, for example, cellular states of immune cells specific to when they were close to tumor cells.

CONCLUSION: ExSeq enables highly multiplexed mapping of RNAs—from nanoscale to system scale—in intact cells and tissues. We explore how RNAs are preferentially targeted to dendrites and spines of neurons, suggesting RNA localization principles that may generalize across different cell types. We also examine gene expression differences in cell types as a function of their distance from other cell types in the context of a human cancer, which may yield insights into future therapeutic approaches that take cellular interactions into account. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†This author made key and essential contributions to the early stages of the project.
‡Corresponding author. Email: gchurch@genetics.med.harvard.edu (G.M.C.); edboyden@mit.edu (E.S.B.)
§These authors contributed equally to this work.
Cite this article as S. Alon *et al.*, *Science* **371**, eaax2656 (2021). DOI: 10.1126/science.aax2656

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aax2656

RESEARCH ARTICLE

IN SITU SEQUENCING

Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems

Shahar Alon[1,2,3]*, Daniel R. Goodwin[1,2]*, Anubhav Sinha[1,2,4]*, Asmamaw T. Wassie[1,2,5]*, Fei Chen[1,6]*, Evan R. Daugharthy[7,8]†‡, Yosuke Bando[1,9], Atsushi Kajita[10], Andrew G. Xue[1], Karl Marrett[10], Robert Prior[10], Yi Cui[1,2], Andrew C. Payne[1,6], Chun-Chen Yao[1,6], Ho-Jun Suk[1,2,4], Ru Wang[1,2], Chih-Chieh (Jay) Yu[1,2,5], Paul Tillberg[1]§, Paul Reginato[1,5,6,7,8], Nikita Pak[1,2,11], Songlei Liu[7,8], Sukanya Punthambaker[7,8], Eswar P. R. Iyer[8], Richie E. Kohman[7,8], Jeremy A. Miller[12], Ed S. Lein[12], Ana Lako[13], Nicole Cullen[13], Scott Rodig[13], Karla Helvie[14], Daniel L. Abravanel[6,15,16], Nikhil Wagle[14], Bruce E. Johnson[14], Johanna Klughammer[6], Michal Slyper[6], Julia Waldman[6], Judit Jané-Valbuena[6], Orit Rozenblatt-Rosen[6], Aviv Regev[6,17,18], IMAXT Consortium[19]¶, George M. Church[7,8]#**, Adam H. Marblestone[1]**††, Edward S. Boyden[1,2,5,17,18,20]#**

Methods for highly multiplexed RNA imaging are limited in spatial resolution and thus in their ability to localize transcripts to nanoscale and subcellular compartments. We adapt expansion microscopy, which physically expands biological specimens, for long-read untargeted and targeted in situ RNA sequencing. We applied untargeted expansion sequencing (ExSeq) to the mouse brain, which yielded the readout of thousands of genes, including splice variants. Targeted ExSeq yielded nanoscale-resolution maps of RNAs throughout dendrites and spines in the neurons of the mouse hippocampus, revealing patterns across multiple cell types, layer-specific cell types across the mouse visual cortex, and the organization and position-dependent states of tumor and immune cells in a human metastatic breast cancer biopsy. Thus, ExSeq enables highly multiplexed mapping of RNAs from nanoscale to system scale.

Tissues are made of cells of many different types and states that are regulated by and contribute to the cells' spatial organization. Multiplexed measurements of the locations and identities of RNA molecules within cells has been useful for exploring these relationships (*1*–*13*). Furthermore, mapping the subcellular locations of RNAs is important for understanding diverse biological processes (*14*, *15*), such as how RNAs in dendritic spines help regulate synaptic function (*16*–*19*).

Imaging RNAs within such compartments, and throughout detailed cellular morphologies, requires nanoscale precision. Such precision is not easily achieved in tissues with current multiplexed optical methods to image RNA. No method can currently perform multiplexed imaging of RNA within tissues in the context of nanoscale cellular morphology. Even though sequential fluorescence in situ hybridization (seqFISH+) allows high-resolution imaging of RNA molecules, it cannot resolve the detailed cellular and tissue context with nanoscale precision (*20*).

Ideally one would be able to perform the enzymatic reactions of sequencing in situ with high multiplexing capacity, while providing for fast nanoscale imaging of cellular and tissue context. Here, we present a toolbox for the untargeted (i.e., not restricted to a predefined list of gene targets) and targeted in situ sequencing of RNAs within intact tissues, in the context of nanoscale cellular morphology.

Adapting expansion microscopy to improve in situ sequencing

We created an untargeted in situ sequencing technology that enables the sequencing of arbitrary RNAs in detailed cellular and tissue contexts. Untargeted approaches have the potential to discover spatially localized sequence variants, such as splice variants and retained introns (*21*). Fluorescent in situ sequencing (FISSEQ) enables such data to be acquired from cultured cells but was not fully demonstrated in tissues (*22*). Therefore, we adapted the chemistry of expansion microscopy (ExM) (*23*, *24*) to separate RNAs from nearby molecules. We reasoned that this may facilitate the chemical access needed for in situ sequencing within tissues. We also expected that the resolution boost from ExM would enable high–spatial resolution mapping of RNAs and their cellular and tissue context on conventional microscopes.

In FISSEQ, untargeted in situ sequencing of RNA is performed to amplify RNA into nanoballs of cDNA (or amplicons), which contain many copies of an RNA sequence (*22*, *25*). These sequences are interrogated in situ with standard next-generation sequencing chemistries on a fluorescence microscope. In ExM (*23*), we isotropically separate gel-anchored biomolecules of interest by an ~4× linear expansion factor, which facilitates both nanoscale imaging with conventional optics and better chemical access to the separated biomolecules (*24*). ExM enables better resolution of normally densely packed RNA transcripts for in situ hybridization imaging (*26*, *27*).

Expanding specimens is expected to benefit FISSEQ by dividing the effective size of the FISSEQ amplicon (200 to 400 nm) (*22*) by the expansion factor. This reduces the packing density of amplicons and facilitates their tracking over many rounds of sequencing. We adapted ExM chemistry to enable FISSEQ in expanded tissues. In particular, the anchoring (Fig. 1A, i), polymerization (Fig. 1A, ii), and expansion (Fig. 1A, iii) steps, which separate RNAs for nanoscale imaging (*26*), result in charged carboxylic acid groups throughout the swellable gel. This suppresses the enzymatic reactions required for FISSEQ (fig. S1). We thus stabilized expanded specimens by re-embedding them in uncharged gels (*26*) and then chemically treated samples to result in a neutral charge environment (fig. S1). We expected that this would allow FISSEQ signal amplification (Fig. 1A, iv) and readout (Fig. 1A, v and vi, and Fig. 1B) steps to proceed.

In situ sequencing involves many rounds of adding fluorescent oligonucleotides (*22*). Accordingly, we established an automated sequencing system (*28*). Because the resultant datasets consist of a series of three-dimensional

[1]Department of Media Arts and Sciences, MIT, Cambridge, MA, USA. [2]McGovern Institute, MIT, Cambridge, MA, USA. [3]Faculty of Engineering, Gonda Brain Research Center and Institute of Nanotechnology, Bar-Ilan University, Ramat Gan, Israel. [4]Harvard-MIT Program in Health Sciences and Technology, MIT, Cambridge, MA, USA. [5]Department of Biological Engineering, MIT, Cambridge, MA, USA. [6]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Department of Genetics, Harvard Medical School, Boston, MA, USA. [8]Wyss Institute for Biologically Inspired Engineering, Boston, MA, USA. [9]Kioxia Corporation, Minato-ku, Tokyo, Japan. [10]Fixstars Solutions Inc, Irvine, CA, USA. [11]Department of Mechanical Engineering, MIT, Cambridge, MA, USA. [12]Allen Institute for Brain Science, Seattle, WA, USA. [13]Center for Immuno-Oncology (CIO), Dana-Farber Cancer Institute, Boston, MA, USA. [14]Center for Cancer Genomics, Dana-Farber Cancer Institute, Boston, MA, USA. [15]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [16]Department of Cell Biology, Harvard Medical School, Boston, MA, USA. [17]Koch Institute for Integrative Cancer Research, Department of Biology, MIT, Cambridge, MA, USA. [18]Howard Hughes Medical Institute, Chevy Chase, MD, USA. [19]CRUK IMAXT Grand Challenge Consortium, Cambridge, UK. [20]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA.
*These authors contributed equally to this work. †This author made key and essential contributions to the early stages of the project. ‡Present address: ReadCoor, part of 10x Genomics, Cambridge, MA, USA. §Present address: Janelia Research Campus, Ashburn, VA, USA. ¶The IMAXT Consortium collaborators and their affiliations are listed in the supplementary materials.
#Corresponding author. Email: gchurch@genetics.med.harvard.edu (G.M.C.); edboyden@mit.edu (E.S.B.) **These authors contributed equally to this work. ††Present address: Federation of American Scientists, Washington, DC, USA.

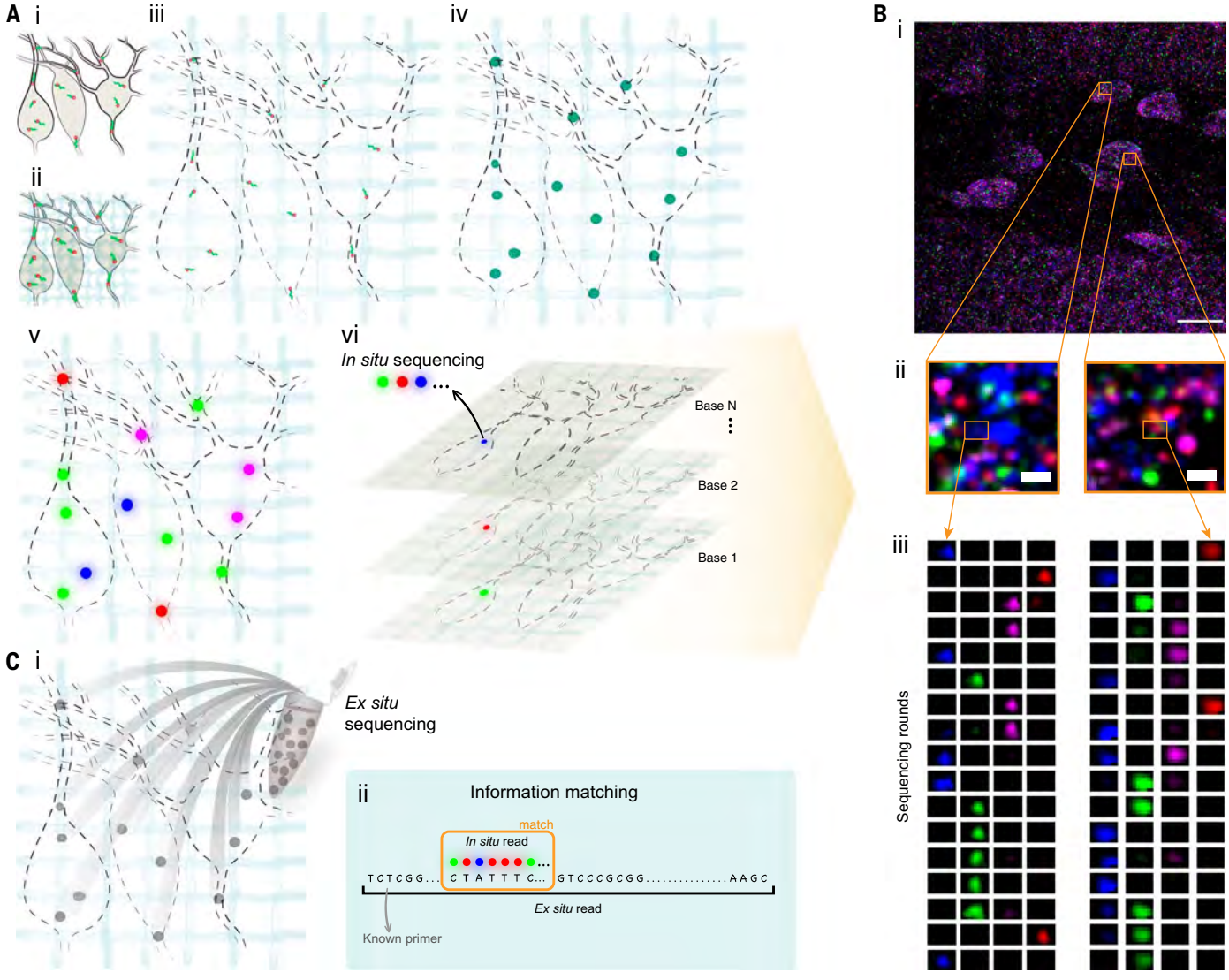


Fig. 1. Untargeted ExSeq concept and workflow. (**A**) ExSeq schematic. (i) A specimen is fixed, and RNA molecules (green) are bound by an anchor (orange). (ii) The specimen is embedded in a swellable gel material (light blue, not to scale), mechanically softened, and then expanded with water (iii). RNA molecules are anchored to the gel. (iv) RNA molecules are reverse transcribed and amplified using FISSEQ. (v) In situ sequencing. Colored dots indicate the colors used in the sequencing chemistry. (vi) In each sequencing round, colors (blue, magenta, green, and red) reveal the current base of the cDNA. (**B**) Example of ExSeq from a 50-μm-thick slice of mouse dentate gyrus. (i) One sequencing round, with two zoomed-in regions (ii) and puncta histories obtained over the course of 17 rounds of in situ sequencing (iii). (**C**) Ex situ sequencing. (i) After in situ sequencing, cDNA amplicons are eluted from the sample and resequenced ex situ with next-generation sequencing. (ii) In situ reads are matched to their longer ex situ counterparts, focusing on unique matches, augmenting the effective in situ read length. Scale bars in (B) are 17 μm in (i) (in biological, i.e., pre-expansion units used throughout, unless otherwise indicated) and 700 nm in (ii).

(3D) images, one for each successive base sequenced, we created a software pipeline (fig. S2) (*29*). This software can align—across images from many rounds—the puncta for each expressed gene to within one pixel (validated in figs. S3 and S4). Finally, puncta are segmented and bases are called (Fig. 1B, iii).

In situ sequencing has previously been limited to short reads of 5 to 30 bases (*10*, *11*, *22*). This limitation reflects laser-induced damage during imaging (*25*) and dependence of the signal for a given cycle on signals from previous cycles (known as phasing), which is caused by incomplete enzymatic reactions (*30*). Alignment of such short reads to the genome is challenging (*31*). Moreover, short reads do not easily capture mRNA complexity, such as alternative splicing.

Accordingly, we added a follow-on round of ex situ classical next-generation sequencing (Fig. 1C, i) (*28*). Notably, the random nature of untargeted sequencing (*28*) results in the creation of distinct molecular identifiers from the in situ sequenced region of the amplified cDNA (fig. S5). This allows us to use ex situ information as a dictionary to align and directly interpret the in situ reads (Fig. 1C, ii, and fig. S5A, bottom panel).

In total, 92% of all matches, and 97% of the matches aligned against nonribosomal RNA, were strictly unique. We removed the handful of in situ reads that matched to more than one ex situ library entry (*28*). Thus, one in situ read matches one ex situ library entry (fig. S5C). This allowed us to explore sequence variations in mRNA, such as alternative splicing, using the longer ex situ matched reads (fig. S6).

Biological validations of ExSeq

Expansion sequencing (ExSeq) produced data from a variety of specimens (tables S1 to S5), including mouse brain (Fig. 1B), *Caenorhabditis*

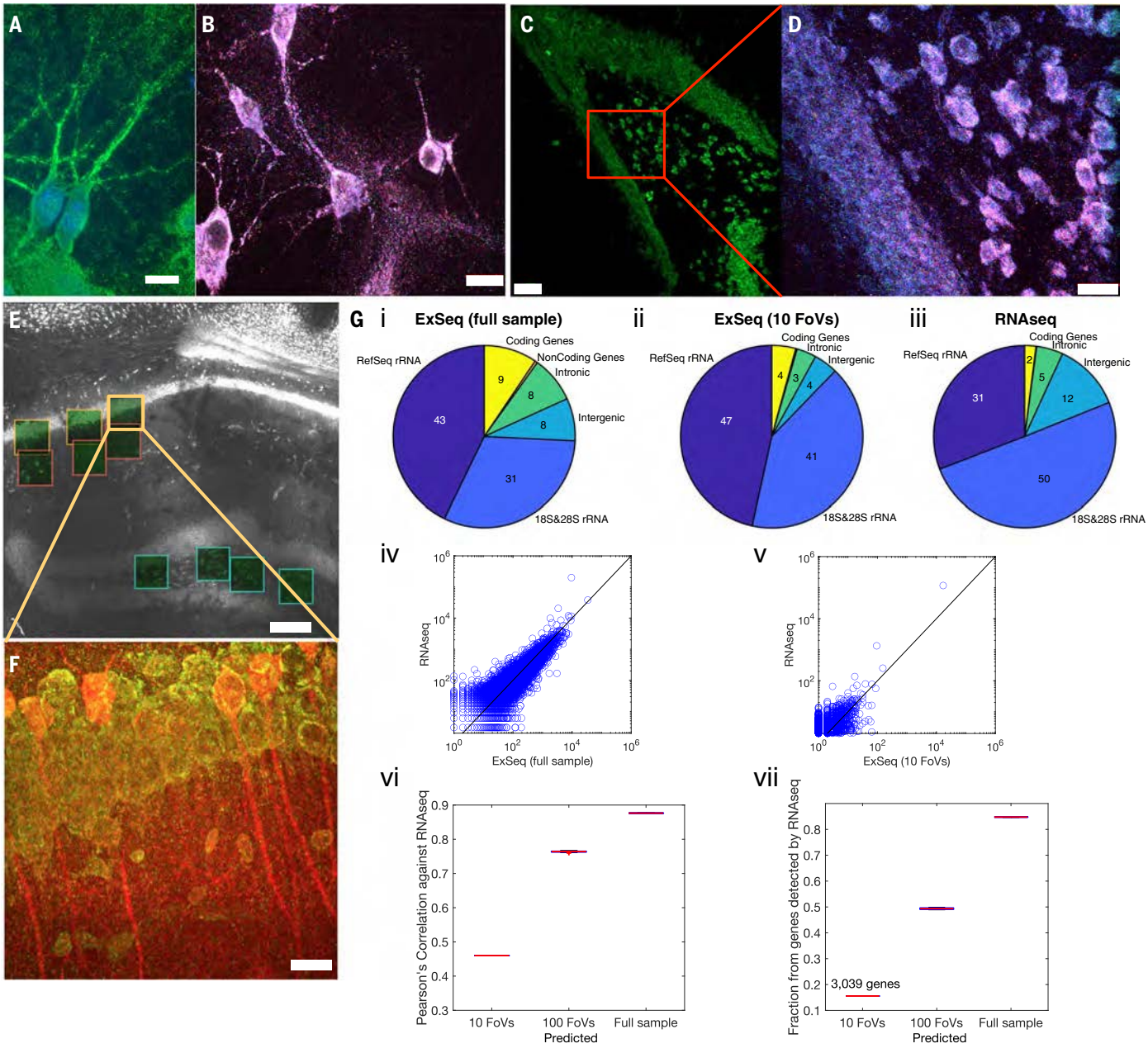


Fig. 2. In situ sequencing in cells and tissues with untargeted ExSeq. (**A**) Example of ExSeq library preparation in hippocampal culture (green, hybridization probe against amplified cDNA; blue, DAPI). (**B**) Maximum intensity projection of one sequencing round in hippocampal culture; color scheme as in Fig. 1B. (**C**) Low-magnification image of ExSeq library preparation in a 15-μm slice of mouse hippocampus (green, hybridization probe against amplified cDNA). (**D**) Maximum intensity projection of a higher-magnification image of the specimen in (C), focusing on one sequencing round; color scheme as in Fig. 1B. (**E**) Low-magnification image of ExSeq library preparation in a 50-μm slice of mouse hippocampus. Fields of view (FoVs) acquired with a higher-magnification objective are shown as green squares. White indicates hybridization probe against amplified cDNA. (**F**) Maximum intensity projection of one FoV of (E), with antibody staining after in situ sequencing (red, antibody against YFP; specimen from a Thy1-YFP mouse; green, hybridization probe against amplified cDNA). (**G**) Sequence analysis of ExSeq specimen shown in (E). (i to iii) RNA content obtained with ExSeq—either using ex situ sequencing data from the entire slice (i) or using ex situ data that correspond to in situ reads observed within the FoVs of (E) (ii)—is comparable to the RNA content of an adjacent slice obtained with standard RNA-seq (iii). Numbers inside the pie chart represent percentages of the total. (iv) Agreement between the normalized expression levels of all well-annotated genes (RefSeq genes) using RNA-seq and ExSeq with full ex situ sequencing data as in (i). (v) As in (iv), but using the 10 acquired FoVs, as in (ii). (vi) Pearson's correlation between the log-transformed expression of RefSeq genes using ExSeq and using RNA-seq, as a function of the number of acquired FoVs [estimated by sampling from the full ex situ sequencing data to simulate the number of expected reads for 100 FoVs; (*28*)]. The value for the 100 FoVs is plotted using the MATLAB boxplot function. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. (vii) Fraction of RefSeq genes detected using ExSeq versus RNA-seq, as a function of the number of acquired FoVs (estimated by sampling from the full ex situ sequencing data to simulate the number of expected reads for 100 FoVs). Scale bars in (A) to (D) and (F) are 13 μm and in (E) are 130 μm. Deconvolution was used in (D) and (F) (*28*).

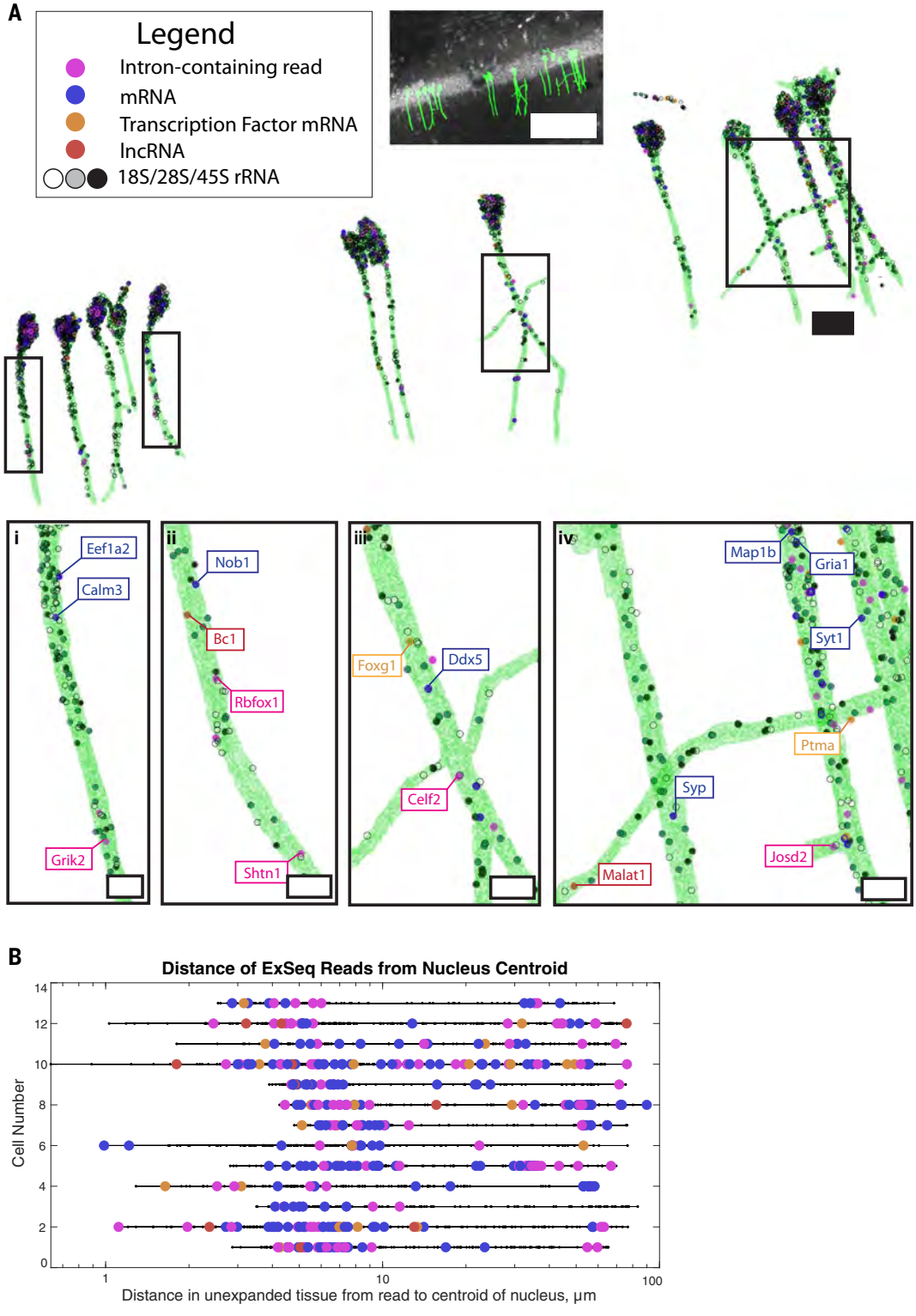


Fig. 3. Untargeted ExSeq enables mapping of RNAs and their variants in dendrites of neurons. (**A**) 3D render of Thy1-YFP CA1 neuronal morphology as determined by YFP antibody staining, containing RNA types as indicated. (i to iv) Zoomed-in dendritic regions (boxed above). Scale bars in top, middle, and bottom rows are 100, 20, and 5 µm, respectively. (**B**) Euclidean distance, relative to the center of the cell body, of sequencing reads for neurons in (A). Color code is as in (A).

elegans (fig. S7A), *Drosophila* embryos (fig. S7B), and HeLa cells (fig. S7C). To validate ExSeq, we used the following mouse specimens: cultured hippocampal neurons (Fig. 2, A and B, and fig. S8), a 15-µm-thick hippocampal slice (Fig. 2, C and D), and a 50-µm-thick hippocampal slice (Fig. 2, E and F). To improve the efficiency of cDNA circularization, we restricted the size of cDNA fragments to ~100 bases long, so ex situ Illumina reads typically contained several repeats of a given cDNA fragment that were, on average, 76 bases long (fig. S6).

Antibody staining after in situ sequencing, as with previous ExM-related protocols (*32*), enabled visualization of specific proteins. This was demonstrated by staining with antibodies against yellow fluorescent protein (YFP) in a Thy1-YFP mouse (*33*) to visualize in situ sequencing reads in neural morphology (Fig. 2F).

As a validation of ExSeq, we performed RNA sequencing (RNA-seq) with random primers on a 50-µm-thick hippocampal slice adjacent to the 50-µm-thick ExSeq specimen (Fig. 2, E and F). As expected for total RNA analysis,

most of the RNA detected in both cases was ribosomal. We observed overall agreement between the RNA types obtained with both methods (Fig. 2G, i to iii), although ExSeq exhibited a slightly higher percentage of coding RNA (4 to 9% with ExSeq versus 2% with RNA-seq). Gene ontology analysis revealed expected functional enrichments for this specimen, including categories such as synapse, neuron projection, and hippocampus (fig. S9 and table S6).

In FISSEQ, highly abundant genes were underrepresented—for example, genes involved in translation and splicing (*22*). By contrast, we did not observe this detection bias with ExSeq (*28*). The expression levels of well-annotated genes (genes from the RefSeq database) using RNA-seq and ExSeq were highly correlated (Pearson's correlation coefficient, $r = 0.89$) (Fig. 2G, iv, and fig. S8C).

The correlation between ExSeq and RNA-seq increased with the ExSeq volume imaged. For example, 10 microscope volumes (each 350 μm by 350 μm by 100 μm in size after expansion, and ~100 μm by 100 μm by 28 μm pre-expansion) resulted in a Pearson's correlation coefficient of $r = 0.47$ (Fig. 2G, v; Student's t test, $P = 9 \times 10^{-164}$) comparable to recent targeted in situ sequencing methods (*34*). Larger volumes, simulated by sampling (*28*), yielded higher correlations (Fig. 2G, vi). With 10 volumes, 3039 genes were detected, making up ~16% of all the genes detected in the sample through RNA-seq, again increasing with the volume sampled (Fig. 2G, vii). Thus, ExSeq is able to report on genome-wide expression in situ in an untargeted, highly multiplexed way.

Subcellular pinpointing of transcript locations in neurons

We next sought to utilize the improved spatial resolution of ExSeq to pinpoint RNAs relative to antibody-stained morphology. We traced 13 hippocampal CA1 pyramidal neurons (*28*). We analyzed the locations of RNAs inside identified neurons with a custom 3D viewer (Fig. 3 and fig. S10) (*28*). The number of sequencing reads per neuron was 229 ± 74 (mean ± standard deviation used throughout) including ribosomal RNA (rRNA), and 30 ± 14 for nonribosomal RNA, for cell bodies and dendrites imaged up to ~100 μm from the cell body. Not including rRNAs, 326 RefSeq genes were observed in these imaged volumes. These numbers are comparable to those obtained by the original FISSEQ protocol (*25*), applied to cultured cell lines.

Neurons contain one nucleus versus thousands of synapses. This raises the question of whether the splicing of mRNAs, such as those that contribute to synaptic function, is regulated in a spatially dependent manner along dendritic trees (*35*). We examined reads that corresponded to intronic regions and observed that, although 70% of such reads were located at the soma, introns in YFP-containing dendritic projections could be found as far down the dendrite as we looked, consistent with previous studies (*36*, *37*). For example, glutamate ionotropic receptor kainate type subunit 2 (*Grik2*), which encodes a receptor subunit involved in excitatory glutamatergic neurotransmission, appears in our data in dendrites with a retained intron (Fig. 3A, i). The *Grik1* subunit had been identified earlier as a dendritically targeted, intron-retaining sequence (*36*, *38*). Dendritic splicing of glutamate receptor subunit RNAs may contribute to the regulation of the state or plasticity of excitatory synapses. In fact, splicing in dendrites has been characterized previously in cultured neurons (*39*).

The long sequencing reads (fig. S6) and untargeted nature of ExSeq also allowed for mapping of alternative splicing in situ. We quantified the expression of known alternative splicing isoforms with ExSeq versus RNA-seq. The two methods were highly correlated (Pearson's $r = 0.944$; fig. S11A). Using only 10 confocal microscope fields of view, of sizes described above, we detected 112 sequencing reads that corresponded to known alternative splicing events. Of these sequencing reads, 67% revealed the expressed alternative splicing isoforms, including ribosomal protein S24 (*Rps24*) and microtubule-associated protein 2 (*Map2*) (fig. S11B). We also identified what are perhaps previously unidentified isoforms, for example for the gene spectrin beta (*Sptbn1*) (fig. S11B).

ExSeq provides the ability to locate these alternative splicing events in space. As an example, isoforms of *Map2*, a key dendritic protein (*40*), and the transcription factor *Cux1*, which is involved in dendrite and spine formation (*41*), could be localized to the neuronal soma outside of the nucleus (fig. S10, neurons 9 and 7, respectively, and fig. S11B).

Many genes may have unappreciated connections to neuronal signaling inside dendritic trees. mRNAs for specific transcription factors have been identified inside dendrites (*42*), for example MAX dimerization protein (*Mga*) (*43*) (fig. S10, neuron 6). However, the full complement of dendritically localized transcription factors in any neuron type is unknown.

In our hippocampus sample (table S2), 914 of the known 1675 mouse transcription factors (RIKEN transcription factor database) were detected by ExSeq. This included 32 reads localized within YFP-expressing cells and 11 reads in the dendrites of these cells. These reads include forkhead box protein G1 (*Foxg1*), which is involved with neural development (*44*), and prothymosin alpha (*Ptma*), which is involved in learning and memory and neurogenesis (*45*) (Fig. 3A, iii, and Fig. 3A, iv). We also found long noncoding RNAs (lncRNAs) and protein coding genes with unknown function in dendrites (Fig. 3A). For example, *Bc1* (Fig. 3A, ii) is a lncRNA from an RNA polymerase III transcript that complexes with proteins to form a ribonucleoprotein particle. *Bc1* is dendritically localized (*46*) and is involved with activity-dependent synaptic regulation (*47*). Additionally, *Malat1* (Fig. 3A, iv) has roles in neural growth and synaptogenesis, but its localization had not been determined in hippocampal tissue (*48*, *49*).

We localized genes that had been found in dendrites of CA1 pyramidal cells at the protein level but had not been mapped at the mRNA level, such as γ-aminobutyric acid (GABA) type A receptor gamma2 subunit (*Gabrg2*) (fig. S10, neuron 2) (*50*). Thus, ExSeq allows us to expand our knowledge of dendritically localized genes of known function, which may point to previously unknown regulatory mechanisms for their gene products. Furthermore, we identified transcripts encoding genes of unknown function in the hippocampus (e.g, *Nob1*; Fig. 3A, ii) (*51*), which may contribute to their functional analysis.

To more systematically understand how the types and identities of transcripts varied with location along a dendrite, we measured the distance from each read to the centroid of its corresponding neuron's cell body (Fig. 3B). These measurements reveal the positions of RNAs encoding for transcription factors, intron-containing reads, and lncRNAs up to 100 μm from the soma. To follow up with a more in-depth examination of specific genes, we next generated a targeted form of ExSeq.

Targeted ExSeq

Untargeted sequencing enables transcriptome-wide exploration of localized RNAs, including rare variants and those of unknown function. However, the diversity of possible reads generated by untargeted methods lead to a lower per-gene copy number of detected molecules and a larger number of biochemical and imaging cycles to distinguish among reads. Targeted methods, by contrast, detect a smaller predefined set of genes and are applicable to mapping cell types and states, mapping their spatial relationships in situ, and visualizing subcellular gene regulation.

An ideal technology for targeted multiplexed RNA mapping would satisfy the following list of criteria. First, it should have sufficient yield (probability of detecting a present molecule) to detect low copy number transcripts such as transcription factors or sparse RNA molecules. Second, the technology should have resolution below the diffraction limit both laterally and axially to resolve nanoscale morphological features, such as dendritic spines in neurons.

Third, the method should provide the ability to image both RNAs and proteins and to work with 3D tissues to localize RNAs in biological contexts. Finally, the method should work with various tissue types, including human tissues. We thus developed a targeted version of ExSeq to match these specifications (tables S7 and S8).

In targeted ExSeq, oligonucleotide padlock probes bearing barcodes hybridize to transcripts (*11*, *52*). Amplicons are then generated for readout through in situ sequencing of the barcodes (Fig. 4A and fig. S12). The inefficient (*22*) reverse transcription step required by untargeted in situ sequencing (*11*, *22*, *53*) is circumvented by the binding and ligation of padlock probes on each targeted transcript using PBCV-1 DNA ligase (also known as SplintR ligase). This enzyme can ligate DNA on an RNA template ~100× faster than T4 DNA ligase (*52*, *54*–*57*). After circularization and rolling circle amplification, the barcodes are sequenced in situ. As barcodes are sequenced across multiple rounds of imaging, the number of identifiable molecular targets scales exponentially with the number of imaging rounds.

We explored the performance of targeted ExSeq in a variety of contexts (table S5). To validate the yield, hybridization chain reaction (HCR) v3.0–amplified expansion FISH (ExFISH) and targeted ExSeq were sequentially performed for the same genes in expanded HeLa cells (*26*, *58*). Targeted ExSeq exhibited an mRNA detection yield of ~62% (Pearson's r = 0.991) relative to HCRv3.0–amplified ExFISH (Fig. 4B and tables S9 to S11), which has a detection efficiency of ~70% in tissue (*26*). For comparison, single-cell RNA sequencing (scRNA-seq) captures ~10% of mRNA (*59*, *60*).

Cell type mapping with spatial context in the visual cortex

We mapped the cell types of the mouse primary visual cortex, for which scRNA-seq data–based classification of cell types has been performed (*61*). We designed a panel of probes targeting 42 genes (tables S9 and S10) that mark key excitatory and inhibitory neuron types. We performed targeted ExSeq of these 42 genes across a coronal section of the primary visual cortex of a Thy1-YFP mouse over a volume of 0.933 mm by 1.140 mm by 0.02 mm, sequencing 265,347 reads (Fig. 4C, top, and table S12).

The spatial distribution of ExSeq reads recapitulated spatial distributions in the Allen in situ hybridization (ISH) atlas (fig. S13). Transcripts known to express in the same cell type appeared in similar positions—for example, in parvalbumin-positive (Pvalb$^+$) interneurons (PV interneurons), parvalbumin (*Pvalb*), vesicular inhibitory amino acid transporter (*Slc32a1*), and glutamate decarboxylase 2 (*Gad2*) transcripts colocalized (Fig. 4C, inset). By contrast, seizure protein 6 homolog (*Sez6*) transcripts, associated with excitatory neurons in deep cortical layers [as well as vasoactive intestinal peptide (VIP$^+$) interneurons] was not colocalized with *Pvalb*, *Slc32a1*, and *Gad2* transcripts (Fig. 4C, inset).

Segmenting cells (fig. S14) (*28*) yielded a total of 1915 cells containing a total of 220,783 reads. Out of these, 1154 cells with at least 50 reads each (177 ± 127 reads per cell) were analyzed. We k-means clustered expression profiles and embedded them into a low-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) (*62*) (Fig. 4D). Clusters were identified with known markers (*28*), such as those corresponding to excitatory neurons (labeled "Ex," and subannotated by their layer location) and inhibitory neurons (annotated with relevant cell type markers). Clusters expressed marker genes consistent with prior studies (*61*) (fig. S15).

We compared our results with a previous study of scRNA-seq of the mouse primary visual cortex (*28*, *61*) (Fig. 4E). We observed the canonical layer-by-layer stratification of excitatory neurons in the visual cortex (Fig. 4F and fig. S16). The nine ExSeq clusters of excitatory neurons corresponded, with slightly different groupings, to seven scRNA-seq clusters of excitatory neurons (Fig. 4E). We found inhibitory neuron ExSeq clusters that matched one-to-one to scRNA-seq clusters. For example, two somatostatin interneuron clusters found across the layers of the cortex—the SST cluster expressing Unc-13 homolog C (cluster SST Unc13c) and the SST cluster expressing *Chondrolectin* (cluster SST Chodl)—appeared prominently in both datasets (Fig. 4, F and G). Some ExSeq clusters of inhibitory neurons mapped onto multiple scRNA-seq clusters. For example, two ExSeq clusters, which we denoted PV and GABAergic (–PV), mapped onto multiple scRNA-seq clusters (Fig. 4E).

Such poolings of scRNA-seq clusters into ExSeq clusters (and vice versa) are likely caused by the smaller number of cells analyzed with ExSeq versus scRNA-seq, the small number of markers interrogated, and the use of a simple k-means algorithm for clustering. Some substructure is visible in the t-SNE plot for the cluster GABAergic (–PV) (Fig. 4D). This suggests that alternative clustering approaches, for instance utilizing morphological criteria or protein markers, could be devised in the future to yield more-precise delineations of cell types.

We varied the parameters used for cell segmentation of the ExSeq dataset and for clustering of the single-cell dataset and found the above conclusions to be robust (figs. S17 and S18). Nonneuronal cells (e.g., glial cells) did not highly express the interrogated markers and were likely nonspecifically clustered with other cell types.

As described (*61*, *63*), the layer-specific excitatory neuron transcription factor marker genes homeobox protein cut-like 2 (*Cux2*), RAR-related orphan receptor beta (*Rorb*), fasciculation and elongation protein zeta-2 (*Fezf2*), and forkhead box protein P2 (*Foxp2*) were expressed in cortical layers 2/3 (L2/3), L4, L5b, and L6, respectively (Fig. 4, E and F, and fig. S15). We used the clusters featuring these markers to segment the cortex into layers (fig. S14D) so that the cell types within each layer could be quantified (Fig. 4G; raw counts, fig. S19). Each cluster of inhibitory neurons was dispersed across layers (Fig. 4G and fig. S19), consistent with earlier work (*61*, *63*). Thus, targeted ExSeq enables sensitive RNA detection across circuit-relevant volumes of tissue and enables cell types to be analyzed in spatial context.

Nanoscale RNA compartmentalization in mouse hippocampal neurons

We next used targeted ExSeq to explore nanoscale RNA compartmentalization within neurons of the mouse hippocampus, where dendritic RNAs are implicated in synaptic plasticity and learning (*64*–*66*). We traced YFP in neurons to identify dendrites and spines and targeted 34 transcripts previously found in CA1 neuron dendrites for sequencing (*67*). Spines were not observed in the untargeted ExSeq hippocampus data because the antibody staining was performed after sequencing, which resulted in weaker staining, whereas here, antibody staining was performed pre-expansion (*28*).

We performed four rounds of in situ sequencing to localize these transcripts on 170 fields of view (1.7 mm by 1 mm by 0.02 mm total; table S5). This sequencing spanned a coronal section containing subfields of the hippocampus and yielded 1.2 million reads, 90,000 of which localized within YFP-expressing neurons (Fig. 5A and table S13). The distributions of expressed genes were similar to those reported in the Allen Brain Atlas in situ hybridization dataset (fig. S20).

Using the YFP signal, we segmented the CA1 pyramidal neurons and dentate gyrus granule cells (although the spines and axons of the latter exhibited low signal-to-noise ratios and were not analyzed further). We found transcripts in dendrites (CA1, DG), spines (CA1), and, to a much smaller extent, axons (CA1) (Fig. 5B). In 106,000 spines examined, we found 730 reads in dendritic spines (each spine had one RNA, except for one that had two). Through simulations (fig. S21), we concluded that it was unlikely that this sparsity of distribution was caused by chemical artifacts of the ExSeq procedure.

In CA1 neurons, as expected, genes such as the postsynaptic density protein dendrin (*Ddn*), the synaptic plasticity-associated gene *Camk2a*, and the postsynaptic scaffolding protein SH3

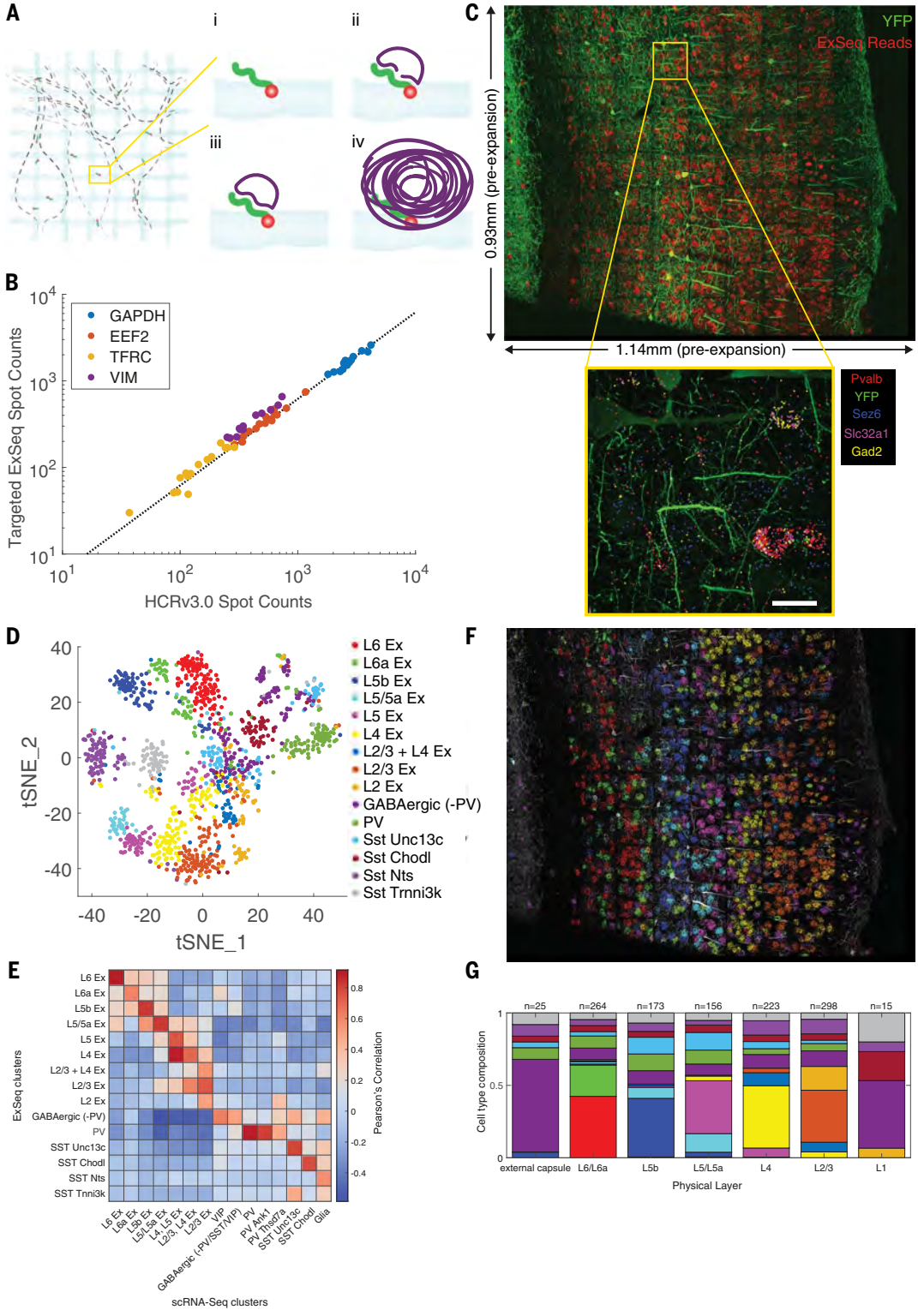


Fig. 4. Targeted ExSeq of transcripts specifying neuron types of mouse primary visual cortex. (**A**) Targeted ExSeq library preparation: (i) RNA anchoring and expansion, (ii) padlock probe hybridization, (iii) probe ligation, and (iv) rolling circle amplification. (**B**) Amplicon counts for targeted ExSeq versus HCRv3.0–amplified ExFISH for the same transcript in the same HeLa cell (60 cells) (slope, 0.62; Pearson's r = 0.991). (**C**) Targeted ExSeq of 42 cell type marker genes in Thy1-YFP mouse visual cortex. (Top) Maximum intensity projection image showing targeted ExSeq reads (red) and YFP (green). (Bottom) Localization of marker genes *Pvalb* (red), *Sez6* (cyan), *Slc32a1* (magenta), and *Gad2* (yellow) with YFP (green). (**D**) Targeted ExSeq gene expression profiles of 1154 cells clustered into 15 cell types. Cluster legend and colors apply to (D), (F), and (G). (**E**) Heatmap showing Pearson's correlation between clusters identified in targeted ExSeq versus a prior scRNA-seq study (*61*). (**F**) Spatial organization of cell types identified in (D). Cell-segmented reads are shown, colored by cluster assignment, and overlaid on YFP (white). (**G**) Layer-by-layer cell-type composition across segmented cortical layers. Scale bar in (C) (bottom) is 20 μm (pre-expansion).

and multiple ankyrin repeat domains 1 (*Shank1*) were prominent in dendrites. The neuronal calcium sensor *Hpca* and the synaptic glutamate receptor *Gria1* were amongst the most abundant in cell bodies (Fig. 5C). In spines, we found *Shank1*, Adenylyl cyclase 1 (*Adcy1*), and kinesin family member 5a (*Kif5a*) to be amongst the most abundant transcripts. We found that the distribution of reads in cell bodies, apical dendrites, basal dendrites, apical dendritic spines, and basal dendritic spines was each statistically different from the others (bootstrapped two-sample Kolmogorov-Smirnov test, $P < 0.001$), except for apical versus basal spines, which were not different from each other (Fig. 5C, iii). This suggests a common set of spine RNAs and spine RNA trafficking principles throughout these neurons.

We validated these observations through bulk RNA sequencing from hippocampal slices adjacent (±100 μm coronally) to the

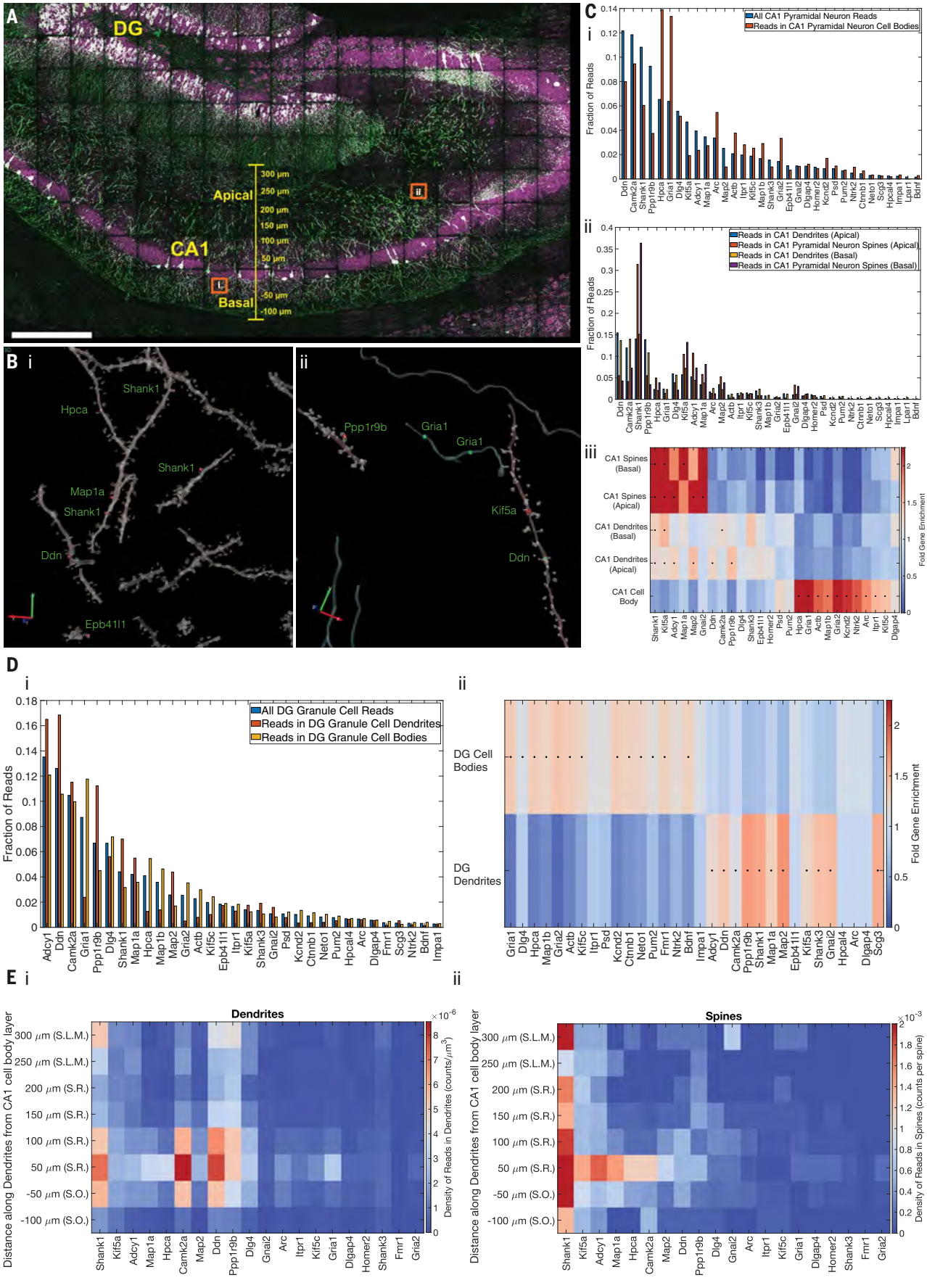


Fig. 5. Targeted ExSeq characterization of nanoscale transcriptomic compartmentalization in mouse hippocampal neuron dendrites and spines. (**A**) Confocal image showing targeted ExSeq of a 34-panel gene set across a slice of mouse hippocampus. Green indicates YFP, magenta indicates reads identified with ExSeq, and white indicates reads localized within YFP-expressing cells. DG, dentate gyrus; CA1, CA1 region of hippocampus. (**B**) 3D reconstruction of dendrites, spines, and axons showing reads localized in spines (red dots) and processes (green dots) for regions indicated by orange boxes in (A). (**C**) The abundance of transcripts in cellular compartments of CA1 pyramidal neurons: (i) abundance of transcripts in all cellular compartments versus cell bodies, (ii) abundance of transcripts in apical and basal dendrites and spines, and (iii) heatmap showing the enrichment of transcripts in apical and basal dendritic and spine compartments of CA1 pyramidal neurons, versus cell bodies. Asterisks indicate statistically significant enrichment (bootstrapped $P < 0.001$). (**D**) The abundance of transcripts in cellular compartments of dentate gyrus (DG) granule cells: (i) abundance of transcripts in the cell bodies and dendrites of DG granule cells and (ii) heatmap showing enrichment of transcripts in compartments of DG granule cells. Asterisks indicate statistically significant enrichment (bootstrapped $P < 0.001$). (**E**) Plots showing the density of transcripts in the dendrites (i) and spines (ii) of CA1 pyramidal neurons along the apical-basal axis (Euclidean distance) of CA1, including regions S.R. (stratum radiatum), S.O. (stratum oriens), and S.L.M. (stratum lacunosum moleculare). Scale bar in (A) is 300 μm, and those in (B) are 2 and 3 μm in (i) and (ii), respectively, shown as red and green arrows in (B) (pre-expansion).

section used for targeted ExSeq. We observed a high level of correlation between in situ sequencing results and bulk RNA sequencing results (Pearson's $r = 0.85$; fig. S22). For the genes studied through both untargeted and targeted versions of ExSeq, we observed high correlation between the read counts (Pearson's $r = 0.68$; fig. S22 and table S13). Using these genes, we estimated the yield of untargeted ExSeq to be 0.6% versus targeted ExSeq (table S13).

Specific genes were significantly (bootstrapped, $P < 0.001$) enriched in specific CA1 neuronal compartments (Fig. 5C). Transcripts for *Shank1*, *Kif5a*, *Adcy1*, *Map1a*, *Map2*, and *Gnai2*, were highly enriched in spines and, to a smaller extent, in apical and basal dendrites compared with cell bodies, perhaps pointing to a process through which these transcripts are enriched the closer they get to synapses. Many of these genes serve structural roles in spines and dendrites (*68*–*70*). On the other hand, a distinct set of genes, including *Hpca*, *Gria1*, *ActB*, and *Map1b* among others, were highly enriched in cell bodies compared with dendrites or spines, consistent with an earlier study (*67*). Notably, *Arc*, whose RNA is known to be dendritically targeted in plasticity contexts, was enriched in cell bodies, consistent with the highly regulated nature of its presence in dendrites (*71*, *72*). Additionally, a few genes, such as *Camk2a* and *Ddn*, were enriched in dendrites compared with both spines and cell bodies, consistent with earlier work (*67*).

In dentate gyrus dendrites, we found transcripts similar to those found in CA1 apical and basal dendrites such as *Shank1*, *Map2*, and *Ppp1r9b* (Fig. 5D). Across the entire 34-gene set, we observed similar dendritic localizations of RNAs in dentate gyrus granule cells versus CA1 pyramidal neurons (Pearson's $r = 0.91$; fig. S23). This similarity raises the possibility that there may be general rules, applicable to multiple neuron types, that govern the dendritic transport of specific RNAs.

Transcripts exhibit varied distributions along dendrites (*67*). We found that most transcripts within dendrites were close (±50 μm) to the cell body layer, and their density decayed rapidly toward distal regions of dendrites, similar to previous observations for these genes (*67*) (Fig. 5E, i, and fig. S24).

Some transcripts, such as *Shank1*, *Ddn*, and *Ppp1r9b*, were present in distal regions of dendrites. When we quantified the presence of transcripts within spines along dendrites, however, we observed a markedly different distribution (Fig. 5E, ii, and fig. S24). For most transcripts found in spines among those in our probe set, their highest density occurred close to the cell body layer. However, spine-localized *Shank1* transcripts exhibited a strong presence throughout spines in proximal and distal regions of dendrites in both apical and basal directions. *Kif5a* and *Adcy1*, to a lesser extent, were also found in the spines of distal dendrites. Thus, although spines are directly connected to dendritic branches, they can exhibit markedly different mRNA distributions.

ExSeq mapping of cell type relationships in cancer

We next explored how ExSeq might reveal spatial patterns of gene expression in the context of cancer biology and immunology. One key question is to understand how tumor microenvironments, including the state of immune cells, govern tumor growth, metastasis, and treatment resistance (*73*). Multiplexed spatial mapping of RNA performed in human tissues to date has not achieved high enough resolution for single-cell quantification, let alone subcellular resolution (*34*, *74*, *75*) (table S7).

A core biopsy was taken from a patient with metastatic breast cancer infiltration into the liver, and 297 tumor-related genes of interest (*28*) were profiled. We resolved 1.15 million reads, including 771,904 reads in 2395 4′,6-diamidino-2-phenylindole (DAPI)–segmented nuclei (Fig. 6A, counts in table S14). The high 3D spatial resolution of ExSeq allowed the detection of 516 RNA reads inside nuclear structures <1 μm in size—possibly nucleoplasmic bridges, which are challenging structures to resolve in tissue (*76*) (fig. S25).

Expression clustering of DAPI-segmented cells (*28*, *77*) revealed the expected mixture of cell types, including tumor, immune (T cell, B cell, and macrophage), and fibroblast cell clusters, characterized with known biomarkers. These biomarkers include members of the immunoglobulin family (*IGHG1*, *IGHG4*, and *IGKC*) found in B cells and genes known to be expressed in metastatic breast cancer [progesterone receptor (*PGR*) (*78*), epidermal growth factor receptor (*EGFR*) (*79*), and aldehyde dehydrogenase 1 family member A3 (*ALDH1A3*) (*80*)] (Fig. 6B).

Tumor and nontumor cells were highly intermixed (Fig. 6C). We examined spatial colocalizations (proximity within 20 μm) between cell types [Fig. 6D; results were robust to distance parameter value (fig. S26)]. Different B cell clusters tended to colocalize in space, consistent with previous observations (*81*). B cell clusters exhibited statistically significant [using bootstrapping (*28*)] colocalizations to all the other cell clusters (Fig. 6D) except for one tumor cluster expressing the gene marker *PGR* (tumor PGR). This is consistent with B cells directly interacting with tumor cells and macrophages, with such interactions contributing to humoral responses in the microenvironment (*81*, *82*). Our analysis also indicates other cell type colocalizations—for example between fibroblast clusters and macrophage, T cell, and tumor clusters (Fig. 6D). Such mappings thus may help illuminate the role of fibroblasts in supporting leukocyte aggregation at sites of cancer (*83*) or the spatial distributions of fibroblast cell types in cancers (*84*).

We finally analyzed whether one cell type could express genes differently as a function of physical proximity to another cell type. For example, one cell might change state depending on physical contact or close proximity to another cell. For each pair of cell clusters that exhibited colocalization, we searched gene expression differences between specific cells that were close (i.e., within 20 μm) versus not close using bootstrapping (*28*). Hypoxia-inducible factor (*HIF1A*) was overexpressed more than fivefold in ALDH1A3-positive tumor cells when they were in close proximity to HSPG2-positive fibroblasts (Fig. 6E, ii). Given that HIF1A serves as a proxy of hypoxic environments and is a microenvironmental cue for tumor cell maintenance (*85*), ExSeq maps may be helpful for further probing such relationships. The mRNA

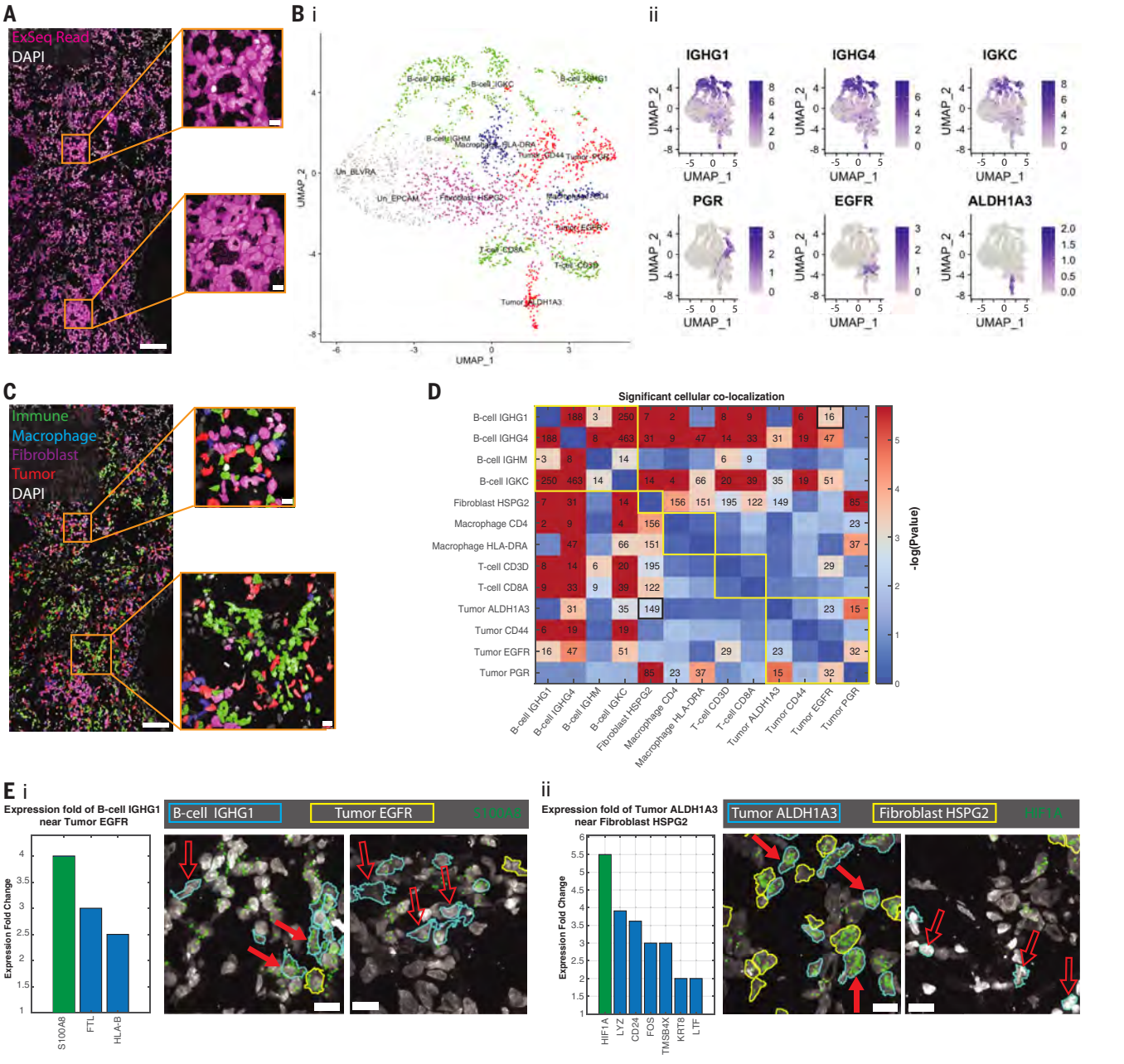


Fig. 6. Targeted ExSeq resolves maps of cell types and states in cancer. (**A**) ExSeq resolves 771,904 reads in 2395 cells (with >100 reads per cell) of 297 genes in a metastatic breast cancer biopsy. (**B**) Uniform manifold approximation and projection (UMAP) representation of principal components analysis (PCA)–based expression clustering reveals immune and tumor cell clusters, indicated by different colors: green (T cells and B cells), red (tumor cells), blue (macrophages), magenta (fibroblasts), and gray [unannotated clusters (*28*)] (i), which express known cell markers for immune cells (ii, top row) and tumor cells (ii, bottom row); expression projected onto UMAP as $\log_2(1 + \text{counts})$. (**C**) Transcriptionally defined cell clusters mapped onto tissue context [colors as in (B)(i)]. (**D**) Spatial colocalization analysis of cell clusters. Adjacency matrix text values, number of cell pairs of indicated type that are in close proximity (nucleus centroid distance of <20 μm; robustness analysis in fig. S26). Adjacency matrix heatmap, *P* value (500,000 bootstrapping iterations), relative to obtaining the same or higher number of cells in close proximity by chance. Adjacency matrix entries with text values are statistically significant (Benjamini Hochberg false-discovery rate of 1.5%). Yellow borders along the diagonal illustrate major cell type categories (B cell, fibroblast, macrophage, T cell, and tumor); the two black-bordered entries correspond to pairs shown in (E). (**E**) ExSeq analysis of cell state as a function of physical proximity, measured by calculating differential expression when cells of different kinds are spatially adjacent (<20 μm) versus far apart. The gene with the largest fold change in a specific cell type when adjacent versus nonadjacent to another specific cell type is shown in green in the histogram ($P = 0.0001$ using 100,000 bootstrapping iterations; all other genes shown in the histogram have $P < 0.05$) as well as in the image showing the gene's read locations in the original sample. (i) Fold change of gene expression in IGHG1-positive B cells when in proximity to EGFR-positive tumor cells (B cells and tumor cells shown with blue and yellow boundaries, respectively). Solid arrows indicate cells in close proximity, and hollow arrows indicate cells not in close proximity. (ii) Fold change of gene expression in ALDH1A3-positive tumor cells when in proximity to HSPG2-positive fibroblasts (tumor and fibroblast cells shown with blue and yellow boundaries, respectively). Scale bars in (A) and (C) are 100 μm, in the insets of (A) and (C) are 10 μm, and in (E) are 10 μm (pre-expansion).

level of *HIF1A* may also indicate tumor radiotherapy resistance (*86*). As a second example, the gene *S100A8*, a regulator of inflammatory processes and immune responses that may be a biomarker for relapse or progression in breast cancer patients (*87*–*89*), was overexpressed fourfold in IGHG1-positive B cells when they were close to EGFR-positive tumor cells (Fig. 6E, i).

Discussion

ExSeq adapts two techniques—ExM and in situ sequencing—to enable spatially precise, highly multiplexed imaging of RNAs in cells and tissues. ExSeq, in both untargeted and targeted forms, facilitates the investigation of scientific questions involving subcellular and even nanoscale RNA localization in intact cellular and tissue contexts (e.g., as indicated by antibody staining of proteins or DAPI staining of nuclei). It can be applied to specimens of multiple-organ systems and species, ranging from the mouse brain to human cancer biopsies, to reveal spatial relationships within and between cells. Such data may reveal principles of cellular organization and function and provide insights into potential mechanisms of how cells interact or are coordinated in complex tissues and multicellular systems. We anticipate that beyond neuroscience and cancer biology, ExSeq will find uses in other fields where many cell types are operating within a complex tissue context—ranging from developmental biology, to immunology, to aging.

Beyond spatial genomics, we expect ExSeq to be useful for in situ sequencing of lineage (*90*) and/or connectome (*91*–*93*) indexing RNA barcodes, which incorporate designed or randomized base-level variation that is not naturally addressed by a FISH approach with a fixed set of tags and targets. More generally, the approaches for re-embedding, passivation, many-round sequential probing, image analysis, and ex situ sequence matching in expanded samples that we have developed for ExSeq should be broadly applicable to other kinds of in situ enzymatic readouts—such as for the multiplexed readout of endogenous DNA or of antibody-attached tags—which may benefit from nanoscale spatial resolution in intact tissues.

Materials and methods summary

All tissues were fixed, optionally immunostained, and treated with the RNA anchoring reagent LabelX. The tissues were then gelled, digested, and expanded (*23*, *26*). Next, the tissues were re-embedded and passivated, which enabled enzymatic reactions to be performed in situ. For untargeted ExSeq, the in situ sequencing library was generated by performing reverse transcription with random primers, circularization of cDNA, and rolling circle amplification (RCA). For targeted ExSeq, padlock probes bearing barcodes were hybridized to transcripts of interest, circularized, and RCA-amplified. In situ sequencing of the cDNA amplicons was then performed through iterative rounds of sequencing chemistry and imaging. The imaging data were converted to nucleotide reads localized in 3D space by a custom image processing pipeline. Reads were ascribed to cells by using immunostaining or other morphological markers. For untargeted ExSeq, the reference for alignment of in situ reads was generated by extracting and sequencing the cDNA amplicons from the sample, which enabled augmentation of the in situ read length. Full materials and methods are available in (*28*).

REFERENCES AND NOTES

1. N. Crosetto, M. Bienko, A. van Oudenaarden, Spatially resolved transcriptomics and beyond. *Nat. Rev. Genet.* **16**, 57–66 (2015). doi: 10.1038/nrg3832; pmid: 25446315
2. T. Gregor, H. G. Garcia, S. C. Little, The embryo as a laboratory: Quantifying transcription in *Drosophila*. *Trends Genet.* **30**, 364–375 (2014). doi: 10.1016/j.tig.2014.06.002; pmid: 25005921
3. S. Alon, G. H. Huynh, E. S. Boyden, Expansion microscopy: Enabling single cell analysis in intact biological systems. *FEBS J.* **286**, 1482–1494 (2019). doi: 10.1111/febs.14597; pmid: 29938896
4. A. E. Moor *et al*., Global mRNA polarization regulates translation efficiency in the intestinal epithelium. *Science* **357**, 1299–1303 (2017). doi: 10.1126/science.aan2399; pmid: 28798045
5. K. B. Halpern *et al*., Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017). doi: 10.1038/nature21065; pmid: 28166538
6. E. Lein, L. E. Borm, S. Linnarsson, The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science* **358**, 64–69 (2017). doi: 10.1126/science.aan6827; pmid: 28983044
7. A. Regev *et al*., The Human Cell Atlas. *eLife* **6**, e27041 (2017). doi: 10.7554/eLife.27041; pmid: 29206104
8. J. R. Moffitt *et al*., Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018). doi: 10.1126/science.aau5324; pmid: 30385464
9. S. Shah, E. Lubeck, W. Zhou, L. Cai, In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* **92**, 342–357 (2016). doi: 10.1016/j.neuron.2016.10.001; pmid: 27764670
10. X. Wang *et al*., Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018). doi: 10.1126/science.aat5691; pmid: 29930089
11. R. Ke *et al*., In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013). doi: 10.1038/nmeth.2563; pmid: 23852452
12. A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998). doi: 10.1126/science.280.5363.585; pmid: 9554849
13. J. M. Levsky, S. M. Shenoy, R. C. Pezo, R. H. Singer, Single-cell gene expression profiling. *Science* **297**, 836–840 (2002). doi: 10.1126/science.1072241; pmid: 12161654
14. D.-M. Franchini *et al*., Microtubule-Driven Stress Granule Dynamics Regulate Inhibitory Immune Checkpoint Expression in T Cells. *Cell Rep.* **26**, 94–107.e7 (2019). doi: 10.1016/j.celrep.2018.12.014; pmid: 30605689
15. V. Balagopal, R. Parker, Polysomes, P bodies and stress granules: States and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.* **21**, 403–408 (2009). doi: 10.1016/j.ceb.2009.03.005; pmid: 19394210
16. C. E. Holt, K. C. Martin, E. M. Schuman, Local translation in neurons: Visualization and function. *Nat. Struct. Mol. Biol.* **26**, 557–566 (2019). doi: 10.1038/s41594-019-0263-5; pmid: 31270476
17. E. M. Schuman, mRNA trafficking and local protein synthesis at the synapse. *Neuron* **23**, 645–648 (1999). doi: 10.1016/S0896-6273(01)80023-4; pmid: 10482231
18. A. J. Rodriguez, K. Czaplinski, J. S. Condeelis, R. H. Singer, Mechanisms and cellular roles of local protein synthesis in mammalian cells. *Curr. Opin. Cell Biol.* **20**, 144–149 (2008). doi: 10.1016/j.ceb.2008.02.004; pmid: 18378131
19. A.-S. Hafner, P. G. Donlin-Asp, B. Leitch, E. Herzog, E. M. Schuman, Local protein synthesis is a ubiquitous feature of neuronal pre- and postsynaptic compartments. *Science* **364**, eaau3644 (2019). doi: 10.1126/science.aau3644; pmid: 31097639
20. C. L. Eng *et al*., Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019). doi: 10.1038/s41586-019-1049-y; pmid: 30911168
21. J. T. Morgan, G. R. Fink, D. P. Bartel, Excised linear introns regulate growth in yeast. *Nature* **565**, 606–611 (2019). doi: 10.1038/s41586-018-0828-1; pmid: 30651636
22. J. H. Lee *et al*., Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi: 10.1126/science.1250212; pmid: 24578530
23. F. Chen, P. W. Tillberg, E. S. Boyden, Expansion microscopy. *Science* **347**, 543–548 (2015). doi: 10.1126/science.1260088; pmid: 25592419
24. A. T. Wassie, Y. Zhao, E. S. Boyden, Expansion microscopy: Principles and uses in biological research. *Nat. Methods* **16**, 33–41 (2019). doi: 10.1038/s41592-018-0219-4; pmid: 30573813
25. J. H. Lee *et al*., Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015). doi: 10.1038/nprot.2014.191; pmid: 25675209
26. F. Chen *et al*., Nanoscale imaging of RNA with expansion microscopy. *Nat. Methods* **13**, 679–684 (2016). doi: 10.1038/nmeth.3899; pmid: 27376770
27. G. Wang, J. R. Moffitt, X. Zhuang, Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. *Sci. Rep.* **8**, 4847 (2018). doi: 10.1038/s41598-018-22297-7; pmid: 29555914
28. Detailed materials and methods are available as supplementary materials.
29. A. Kajita, D. Goodwin, R. Prior, dgoodwin208/ExSeqProcessing: First release of the processing software for Expansion Sequencing, version v1.0.0, Zenodo (2020); .doi: 10.5281/zenodo.4075515
30. M. Kircher, P. Heyn, J. Kelso, Addressing challenges in the production and analysis of illumina sequencing data. *BMC Genomics* **12**, 382 (2011). doi: 10.1186/1471-2164-12-382; pmid: 21801405
31. W. Li, J. Freudenberg, P. Miramontes, Diminishing return for increased Mappability with longer sequencing reads: Implications of the k-mer distributions in the human genome. *BMC Bioinformatics* **15**, 2 (2014). doi: 10.1186/1471-2105-15-2; pmid: 24386976
32. P. W. Tillberg *et al*., Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34**, 987–992 (2016). doi: 10.1038/nbt.3625; pmid: 27376584
33. G. Feng *et al*., Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000). doi: 10.1016/S0896-6273(00)00084-2; pmid: 11086982
34. J. Svedlund *et al*., Generation of in situ sequencing based OncoMaps to spatially resolve gene expression profiles of diagnostic and prognostic markers in breast cancer. *EBioMedicine* **48**, 212–223 (2019). doi: 10.1016/j.ebiom.2019.09.009; pmid: 31526717
35. K. S. Kosik, Life at Low Copy Number: How Dendrites Manage with So Few mRNAs. *Neuron* **92**, 1168–1180 (2016). doi: 10.1016/j.neuron.2016.11.002; pmid: 28009273
36. M. Khaladkar *et al*., Subcellular RNA sequencing reveals broad presence of cytoplasmic intron-sequence retaining transcripts in mouse and rat neurons. *PLOS ONE* **8**, e76194 (2013). doi: 10.1371/journal.pone.0076194; pmid: 24098440
37. H. Saini, A. A. Bicknell, S. R. Eddy, M. J. Moore, Free circular introns with an unusual branchpoint in neuronal projections. *eLife* **8**, e47809 (2019). doi: 10.7554/eLife.47809; pmid: 31697236
38. P. T. Buckley *et al*., Cytoplasmic intron sequence-retaining transcripts can be dendritically targeted via ID element retrotransposons. *Neuron* **69**, 877–884 (2011). doi: 10.1016/j.neuron.2011.02.028; pmid: 21382548
39. J. Glanzer *et al*., RNA splicing capability of live neuronal dendrites. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16859–16864 (2005). doi: 10.1073/pnas.0503783102; pmid: 16275927
40. A. Harada, J. Teng, Y. Takei, K. Oguchi, N. Hirokawa, MAP2 is required for dendrite elongation, PKA anchoring in dendrites, and proper PKA signal transduction. *J. Cell Biol.* **158**, 541–549 (2002). doi: 10.1083/jcb.200110134; pmid: 12163474

41. B. Cubelos *et al*., Cux1 and Cux2 regulate dendritic branching, spine morphology, and synapses of the upper layer neurons of the cortex. *Neuron* **66**, 523–535 (2010). doi: 10.1016/j.neuron.2010.04.038; pmid: 20510857
42. P. Crino *et al*., Presence and phosphorylation of transcription factors in developing dendrites. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2313–2318 (1998). doi: 10.1073/pnas.95.5.2313; pmid: 9482882
43. S. A. Middleton, J. Eberwine, J. Kim, Comprehensive catalog of dendritically localized mRNA isoforms from sub-cellular sequencing of single mouse neurons. *BMC Biol.* **17**, 5 (2019). doi: 10.1186/s12915-019-0630-z; pmid: 30678683
44. L. Pancrazi *et al*., Foxg1 localizes to mitochondria and coordinates cell differentiation and bioenergetics. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13910–13915 (2015). doi: 10.1073/pnas.1515190112; pmid: 26508630
45. H. Ueda *et al*., Prothymosin alpha-deficiency enhances anxiety-like behaviors and impairs learning/memory functions and neurogenesis. *J. Neurochem.* **141**, 124–136 (2017). doi: 10.1111/jnc.13963; pmid: 28122138
46. H. Tiedge, R. T. Fremeau Jr.., P. H. Weinstock, O. Arancio, J. Brosius, Dendritic location of neural BC1 RNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2093–2097 (1991). doi: 10.1073/pnas.88.6.2093; pmid: 1706516
47. I. A. Muslimov, G. Banker, J. Brosius, H. Tiedge, Activity-dependent regulation of dendritic BC1 RNA in hippocampal neurons in culture. *J. Cell Biol.* **141**, 1601–1611 (1998). doi: 10.1083/jcb.141.7.1601; pmid: 9647652
48. X. Zhang, M. H. Hamblin, K.-J. Yin, The long noncoding RNA Malat1: Its physiological and pathophysiological functions. *RNA Biol.* **14**, 1705–1714 (2017). doi: 10.1080/15476286.2017.1358347; pmid: 28837398
49. J. A. Briggs, E. J. Wolvetang, J. S. Mattick, J. L. Rinn, G. Barry, Mechanisms of Long Non-coding RNAs in Mammalian Nervous System Development, Plasticity, Disease, and Evolution. *Neuron* **88**, 861–877 (2015). doi: 10.1016/j.neuron.2015.09.045; pmid: 26637795
50. H. Hörtnagl *et al*., Patterns of mRNA and protein expression for 12 GABAA receptor subunits in the mouse brain. *Neuroscience* **236**, 345–372 (2013). doi: 10.1016/j.neuroscience.2013.01.008; pmid: 23337532
51. Y. Han *et al*., Up-regulation of Nob1 in the rat auditory system with noise-induced hearing loss. *Neurosci. Lett.* **491**, 79–82 (2011). doi: 10.1016/j.neulet.2011.01.010; pmid: 21219967
52. N. Schneider, M. Meier, Efficient in situ detection of mRNAs using the Chlorella virus DNA ligase for padlock probe ligation. *RNA* **23**, 250–256 (2017). doi: 10.1261/rna.057836.116vpmid: 27879431
53. D. Fürth, V. Hatini, J. H. Lee, In Situ Transcriptome Accessibility Sequencing (INSTA-seq). bioRxiv 722819 [Preprint]. 6 August 2019. .doi: 10.1101/722819
54. A. E. Cartier *et al*., Regulation of synaptic structure by ubiquitin C-terminal hydrolase L1. *J. Neurosci.* **29**, 7857–7868 (2009). doi: 10.1523/JNEUROSCI.1817-09.2009; pmid: 19535597
55. G. J. S. Lohman, Y. Zhang, A. M. Zhelkovsky, E. J. Cantor, T. C. Evans Jr.., Efficient DNA ligation in DNA-RNA hybrid helices by Chlorella virus DNA ligase. *Nucleic Acids Res.* **42**, 1831–1844 (2014). doi: 10.1093/nar/gkt1032; pmid: 24203707
56. R. Deng, K. Zhang, Y. Sun, X. Ren, J. Li, Highly specific imaging of mRNA in single cells by target RNA-initiated rolling circle amplification. *Chem. Sci.* **8**, 3668–3675 (2017). doi: 10.1039/C7SC00292K; pmid: 28580104
57. E. P. R. Iyer *et al*., Barcoded oligonucleotides ligated on RNA amplified for multiplex and parallel in-situ analyses. bioRxiv 281121 [Preprint]. 20 March 2018. .doi: 10.1101/281121
58. H. M. T. Choi *et al*., Third-generation *in situ* hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753 (2018). doi: 10.1242/dev.165753; pmid: 29945988
59. G. K. Marinov *et al*., From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. *Genome Res.* **24**, 496–510 (2014). doi: 10.1101/gr.161034.113; pmid: 24299736
60. E. Torre *et al*., Rare Cell Detection by Single-Cell RNA Sequencing as Guided by Single-Molecule RNA FISH. *Cell Syst.* **6**, 171–179.e5 (2018). doi: 10.1016/j.cels.2018.01.014; pmid: 29454938
61. B. Tasic *et al*., Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016). doi: 10.1038/nn.4216; pmid: 26727548
62. L. van der Maaten, G. Hinton, Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008); http://jmlr.org/papers/v9/vandermaaten08a.html.
63. E. S. Lein *et al*., Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007). doi: 10.1038/nature05453; pmid: 17151600
64. L. E. Ostroff, J. C. Fiala, B. Allwardt, K. M. Harris, Polyribosomes redistribute from dendritic shafts into spines with enlarged synapses during LTP in developing rat hippocampal slices. *Neuron* **35**, 535–545 (2002). doi: 10.1016/S0896-6273(02)00785-7; pmid: 12165474
65. O. Steward, P. Worley, Local synthesis of proteins at synaptic sites on dendrites: Role in synaptic plasticity and memory consolidation? *Neurobiol. Learn. Mem.* **78**, 508–527 (2002). doi: 10.1006/nlme.2002.4102; pmid: 12559831
66. Y. J. Yoon *et al*., Glutamate-induced RNA localization and translation in neurons. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6877–E6886 (2016). doi: 10.1073/pnas.1614267113; pmid: 27791158
67. I. J. Cajigas *et al*., The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453–466 (2012). doi: 10.1016/j.neuron.2012.02.036; pmid: 22578497
68. A. Y. Hung *et al*., Smaller dendritic spines, weaker synaptic transmission, but enhanced spatial learning in mice lacking Shank1. *J. Neurosci.* **28**, 1697–1708 (2008). doi: 10.1523/JNEUROSCI.3032-07.2008; pmid: 18272690
69. N. Hirokawa, S. Niwa, Y. Tanaka, Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* **68**, 610–638 (2010). doi: 10.1016/j.neuron.2010.09.039; pmid: 21092854
70. A. J. Koleske, Molecular mechanisms of dendrite stability. *Nat. Rev. Neurosci.* **14**, 536–550 (2013). doi: 10.1038/nrn3486; pmid: 23839597
71. O. Steward, C. S. Wallace, G. L. Lyford, P. F. Worley, Synaptic activation causes the mRNA for the IEG Arc to localize selectively near activated postsynaptic sites on dendrites. *Neuron* **21**, 741–751 (1998). doi: 10.1016/S0896-6273(00)80591-7; pmid: 9808461
72. O. Steward, P. F. Worley, Selective targeting of newly synthesized Arc mRNA to active synapses requires NMDA receptor activation. *Neuron* **30**, 227–240 (2001). doi: 10.1016/S0896-6273(01)00275-6; pmid: 11343657
73. D. S. Chen, I. Mellman, Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017). doi: 10.1038/nature21349; pmid: 28102259
74. P. L. Ståhl *et al*., Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016). doi: 10.1126/science.aaf2403; pmid: 27365449
75. S. G. Rodriques *et al*., Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019). doi: 10.1126/science.aaw1219; pmid: 30923225
76. M. Fenech *et al*., Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis* **26**, 125–132 (2011). doi: 10.1093/mutage/geq052; pmid: 21164193
77. T. Stuart *et al*., Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118
78. M. Kitahara *et al*., HER2-Positive Conversion in a Metastatic Liver Focus in Late Recurrent Breast Cancer. *Case Rep. Oncol.* **12**, 473–479 (2019). doi: 10.1159/000501306; pmid: 31320870
79. H. Masuda *et al*., Role of epidermal growth factor receptor in breast cancer. *Breast Cancer Res. Treat.* **136**, 331–345 (2012). doi: 10.1007/s10549-012-2289-9; pmid: 23073759
80. P. Marcato *et al*., Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells* **29**, 32–45 (2011). doi: 10.1002/stem.563; pmid: 21280157
81. S. Garaud *et al*., Tumor-infiltrating B cells signal functional humoral immune responses in breast cancer. *JCI Insight* **4**, e129641 (2019). doi: 10.1172/jci.insight.129641; pmid: 31408436
82. E. Pipi *et al*., Tertiary Lymphoid Structures: Autoimmunity Goes Local. *Front. Immunol.* **9**, 1952 (2018). doi: 10.3389/fimmu.2018.01952; pmid: 30258435
83. S. Nayar *et al*., Immunofibroblasts are pivotal drivers of tertiary lymphoid structure formation and local pathology. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13490–13497 (2019). doi: 10.1073/pnas.1905301116; pmid: 31213547
84. M. Bartoschek *et al*., Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat. Commun.* **9**, 5150 (2018). doi: 10.1038/s41467-018-07582-3; pmid: 30514914
85. N. Colwell *et al*., Hypoxia in the glioblastoma microenvironment: Shaping the phenotype of cancer stem-like cells. *Neuro-Oncol.* **19**, 887–896 (2017). doi: 10.1093/neuonc/now258; pmid: 28339582
86. F. Cimmino *et al*., HIF-1 transcription activity: HIF1A driven response in normoxia and in hypoxia. *BMC Med. Genet.* **20**, 37 (2019). doi: 10.1186/s12881-019-0767-1; pmid: 30808328
87. C. Gebhardt, J. Németh, P. Angel, J. Hess, S100A8 and S100A9 in inflammation and cancer. *Biochem. Pharmacol.* **72**, 1622–1631 (2006). doi: 10.1016/j.bcp.2006.05.017; pmid: 16846592
88. J.-M. Zhong *et al*., Protein S100-A8: A potential metastasis-associated protein for breast cancer determined via iTRAQ quantitative proteomic and clinicopathological analysis. *Oncol. Lett.* **15**, 5285–5293 (2018). doi: 10.3892/ol.2018.7958; pmid: 29552168
89. D. Wang *et al*., Clinical Significance of Elevated S100A8 Expression in Breast Cancer Patients. *Front. Oncol.* **8**, 496 (2018). doi: 10.3389/fonc.2018.00496; pmid: 30456203
90. R. Kalhor *et al*., Developmental barcoding of whole mouse via homing CRISPR. *Science* **361**, eaat9804 (2018). doi: 10.1126/science.aat9804; pmid: 30093604
91. X. Chen, Y.-C. Sun, G. M. Church, J. H. Lee, A. M. Zador, Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. *Nucleic Acids Res.* **46**, e22 (2018). doi: 10.1093/nar/gkx1206; pmid: 29190363
92. I. D. Peikon *et al*., Using high-throughput barcode sequencing to efficiently map connectomes. *Nucleic Acids Res.* **45**, e115 (2017). doi: 10.1093/nar/gkx292; pmid: 28449067
93. A. M. Zador *et al*., Sequencing the connectome. *PLOS Biol.* **10**, e1001411 (2012). doi: 10.1371/journal.pbio.1001411; pmid: 23109909
94. D. Goodwin, dgoodwin208/ExSeqOF3DVisualizer: First release of 3D visualization software for Expansion Sequencing, version v1.0.0, Zenodo (2020); .doi: 10.5281/zenodo.4075649

ACKNOWLEDGMENTS

We acknowledge A. Lin for the help with Illumina sequencing, D. Park for providing cultured neurons, E. Murray for providing cultured HeLa cells, K. Piatkevich for performing transcardial perfusions, R. Kalhor for helpful discussions, and E. A. Pnevmatikakis for helpful discussions on image processing. We also acknowledge the SpaceTx analysis working group for help in clustering (T. Bakken, Z. Yao, and P. Kharchenko) and for help in gene selection (E. D. Vaishnav, B. Aevermann, R. Scheuermann, and K. Harris). **Funding:** R.E.K. and A.H.M. acknowledge IARPA MICrONS (D16PC0008); E.R.D. acknowledges government support under HG005550 and HG008525 awarded by the National Institutes of Health and under DGE1144152 awarded by the National Science Foundation; S.A. was supported by the Howard Hughes Medical Institute (HHMI) fellowship of the LSRF; D.R.G. is an NSF GRFP fellow; F.C. is supported by the Schmidt Fellows Program at the Broad Institute; A.T.W. was supported by the Hertz Foundation Fellowship and the Siebel Scholarship; A.S. was supported by the NIH Neuroimaging Training Program T32 grant 5T32EB001680; and J.A.M. and E.S.L. thank the Chan Zuckerberg Initiative, an advised fund of Silicon Valley Community Foundation, for their support (award no. 2017-0525). E.S.B. was supported by L. Yang, J. Doerr, the Open Philanthropy Project, Cancer Research UK Grand Challenge grant C31893/A25050, U.S. Army Research Laboratory and the U.S. Army Research Office under contract no. W911NF1510548, NSF grant 1734870, NIH UF1NS107697, NIH IDIQ17X149, NIH 1RM1HG008525, NIH 1R01MH103910, NIH 1R01MH114031, NIH 1R01MH110932, NIH 1R01EB024261, NIH 1R01DA045549, IARPA D16PC00008, NIH 1U19MH114821, NIH 1R01NS102727, the Chan Zuckerberg Initiative Human Cell Atlas pilot program, the Ludwig Foundation, the HHMI-Simons Faculty Scholars Program, and the HHMI investigator program. This project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health, task order no. HHSN261100039 under contract no. HHSN261201500003I. J.K. was supported by an HFSP long-term fellowship (LT000452/2019-L). **Author contributions:** S.A., D.R.G., A.S., A.T.W., F.C., E.R.D., G.M.C., A.H.M., and E.S.B. contributed key ideas and designed experiments; S.A., A.T.W., A.S., and F.C. performed experiments; S.A., D.R.G., A.T.W., A.S., F.C., A.H.M., and E.S.B. analyzed data and wrote the paper; D.R.G., F.C., Y.B., A.K., and A.G.X. conceived and implemented the image analysis pipeline with assistance from S.A. and initial discussions with E.R.D; D.R.G., Y.B., A.K., A.G.X., K.M., and R.P. analyzed the imaging data; A.T.W., A.S., F.C., and A.H.M. conceived and implemented targeted ExSeq;

D.R.G. designed and implemented the 3D visualization tool; D.R.G. analyzed cDNA movement experiments; S.A. and D.R.G. performed 3D tracing; Y.C., A.C.P., C.-C.Ya., P.T., P.R., and R.E.K. contributed to protocol optimization; S.A., F.C., H.-J.S., and R.W. automated the in situ sequencing; F.C., E.R.D., and A.H.M. conceived passivation; S.A. and F.C. implemented passivation; S.A. and F.C. conceived and implemented ex situ sequencing with initial discussions with E.R.D and A.C.P.; S.A. and F.C. implemented ex situ and in situ sequencing matching with assistance from D.R.G.; S.A., A.T.W., A.S., and F.C. conceived and implemented antibody staining; E.R.D. conceived and implemented cDNA fragmentation protocol and performed experiments on earlier preliminary protocols; S.A., F.C., and E.R.D. optimized FISSEQ enzymatics for the expanded gel; S.A. and F.C. conceived and implemented cDNA anchoring with initial discussions with E.R.D.; C.-C.Yu performed *C. elegans* fixation, cuticle reduction, and permeabilization; N.P. carried out mouse surgeries; J.A.M. and E.S.L. designed the gene list for the visual cortex experiment; A.L., N.C., S.R., K.H., D.L.A., N.W., B.E.J., J.K., M.S., J.W., J.J.-V., O.R.-R., and A.R. designed the gene list and provided the human sample for the cancer experiment; S.L., S.P., and E.P.R.I. contributed to targeted ExSeq; A.H.M., E.S.B., and G.M.C. initiated the project; and E.S.B. supervised the project. **Competing interests:** E.S.B. is a cofounder of Expansion Technologies, which has commercial interests in the space of ExM. F.C. is a paid consultant of Celsius Therapeutics and is a member of Celsius Thinklab. E.R.D. is a cofounder of ReadCoor, part of 10x Genomics, which has commercial interests in the space of ExM and FISSEQ. From October 2019, A.L. is an employee of Bristol Myers Squibb. From June 2018, E.P.R.I. is an employee of 10x Genomics. N.W. is a stockholder of Relay Therapeutics, is in the scientific advisory board of Relay Therapeutics, is a consultant to Eli Lilly, and received a grant from Puma Biotechnologies. A.R. is a founder and equity holder of Celsius Therapeutics; an equity holder in Immunitas Therapeutics; and, until 31 August 2020, was an SAB member of Syros Pharmaceuticals, Neogene Therapeutics, Asimov, and ThermoFisher Scientific. From 1 August 2020, A.R. is an employee of Genentech, a member of the Roche Group. G.M.C. is a cofounder and SAB member of ReadCoor and is an adviser to 10x Genomics after their acquisition of ReadCoor. Conflict of interest link for G.M.C.: http://arep.med.harvard.edu/gmc/tech.html. S.A., D.R.G., A.T.W., A.S., F.C., E.R.D., A.C.P., P.T., G.M.C., A.H.M., and E.S.B. are inventors on U.S. patents 10526649, 10364457, and 10059990 and U.S. patent application 15/876,347, submitted by MIT and Harvard, that cover ExSeq. **Data and materials availability:** All of the raw Illumina sequencing data were deposited to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA), BioProject PRJNA663046. The entire MATLAB pipeline to process ExSeq datasets from the microscope to spatial analysis of gene expression is publicly accessible at Zenodo (*29*). From there, a tutorial with step-by-step instructions on how to run the pipeline and a tutorial set of targeted ExSeq data from the mouse visual cortex can be accessed. The 3D visualization tool is hosted at Zenodo (*94*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/371/6528/eaax2656/suppl/DC1
IMAXT Consortium Collaborator List
Materials and Methods
Figs. S1 to S26
Tables S1 to S14
References (*95*–*137*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

8 March 2019; resubmitted 13 May 2020
Accepted 20 November 2020
10.1126/science.aax2656

Science

Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems

Shahar Alon, Daniel R. Goodwin, Anubhav Sinha, Asmamaw T. Wassie, Fei Chen, Evan R. Daugharthy, Yosuke Bando, Atsushi Kajita, Andrew G. Xue, Karl Marrett, Robert Prior, Yi Cui, Andrew C. Payne, Chun-Chen Yao, Ho-Jun Suk, Ru Wang, Chih-Chieh (Jay) Yu, Paul Tillberg, Paul Reginato, Nikita Pak, Songlei Liu, Sukanya Punthambaker, Eswar P. R. Iyer, Richie E. Kohman, Jeremy A. Miller, Ed S. Lein, Ana Lako, Nicole Cullen, Scott Rodig, Karla Helvie, Daniel L. Abravanel, Nikhil Wagle, Bruce E. Johnson, Johanna Klughammer, Michal Slyper, Julia Waldman, Judit Jané-Valbuena, Orit Rozenblatt-Rosen, Aviv Regev, IMAXT Consortium, George M. Church, Adam H. Marblestone and Edward S. Boyden

Science **371** (6528), eaax2656.
DOI: 10.1126/science.aax2656

Identifying transcript location in cells

Identifying where specific RNAs occur within a cell or tissue has been limited by technology and imaging capabilities. Expansion microscopy has allowed for better visualization of small structures by expanding the tissues with a polymer- and hydrogel-based system. Alon *et al.* combined expansion microscopy with long-read in situ RNA sequencing, resulting in a more precise visualization of the location of specific transcripts. This method, termed "ExSeq" for expansion sequencing, was used to detect RNAs, both new transcripts and those previously demonstrated to localize to neuronal dendrites. Unlike other in situ sequencing methods, ExSeq does not target sets of genes. This technology thus unites spatial resolution, multiplexing, and an unbiased approach to reveal insights into RNA localization and its physiological roles in developing and active tissue.

Science, this issue p. eaax2656

science.sciencemag.org/content/371/6528/eaax2656/suppl/DC1

Supplementary Materials for

Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems

Shahar Alon*, Daniel R. Goodwin*, Anubhav Sinha*, Asmamaw T. Wassie*, Fei Chen*, Evan R. Daugharthy†, Yosuke Bando, Atsushi Kajita, Andrew G. Xue, Karl Marrett, Robert Prior, Yi Cui, Andrew C. Payne, Chun-Chen Yao, Ho-Jun Suk, Ru Wang, Chih-Chieh (Jay) Yu, Paul Tillberg, Paul Reginato, Nikita Pak, Songlei Liu, Sukanya Punthambaker, Eswar P. R. Iyer, Richie E. Kohman, Jeremy A. Miller, Ed S. Lein, Ana Lako, Nicole Cullen, Scott Rodig, Karla Helvie, Daniel L. Abravanel, Nikhil Wagle, Bruce E. Johnson, Johanna Klughammer, Michal Slyper, Julia Waldman, Judit Jané-Valbuena, Orit Rozenblatt-Rosen, Aviv Regev, IMAXT Consortium, George M. Church‡§, Adam H. Marblestone§, Edward S. Boyden‡§

*These authors contributed equally to this work.
†This author made key and essential contributions to the early stages of the project.
‡Corresponding author. Email: gchurch@genetics.med.harvard.edu (G.M.C.); edboyden@mit.edu (E.S.B.)
§These authors contributed equally to this work.

Published 29 January 2021, *Science* **371**, eaax2656 (2021)
DOI: 10.1126/science.aax2656

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/content/371/6528/eaax2656/suppl/DC1)

Tables S1 to S6 and S9 to S14 (.xlsx)
MDAR Reproducibility Checklist (.pdf)

Author: IMAXT Consortium: Members and Affiliations

Ali HR[1,2], Al Sa'd M[3], Alon S[4], Aparicio S[5,6], Battistoni G[1], Balasubramanian S[1,7], Becker R[8], Bodenmiller B[2], Boyden ES[4], Bressan D[1], Bruna A[9], Burger Marcel[2], Caldas C[9], Callari M[1], Cannell IG[1], Casbolt H[1], Chornay N[3], Cui Y[4], Dariush A[3], Dinh K[10], Emenari A[4], Eyal-Lubling Y[9], Fan J[11], Fatemi A[1], Fisher E[1], González-Solares EA[3], González-Fernández C[3], Goodwin D[4], Greenwood W[1], Grimaldi F[8], Hannon GJ[1], Harris O[8], Harris S[8], Jauset C[1], Joyce JA[12], Karagiannis ED[4], Kovačević T[1], Kuett L[2], Kunes R[10], Küpcü Yoldaş A[3], Lai D[5,6], Laks E[5,6], Lee H[11], Lee M[1,7], Lerda G[1], Li Y[5], McPherson A[5,6,13], Millar N[3], Mulvey CM[1], Nugent F[1], O'Flanagan CH[5], Paez-Ribes M[1], Pearsall I[1], Qosaj F[1], Roth AJ[5,6,14], Rueda OM[9], Ruiz T[5], Sawicka K[1], Sepúlveda LA[11], Shah SP[5,6,13], Shea A[9], Sinha A[4], Smith A[5], Tavaré S[1,10,15], Tietscher S[2], Vázquez-García I[10,13], Vogl SL[8], Walton NA[3], Wassie AT[4], Watson SS[12], Weselak J[8], Wild SA[1], Williams E[1], Windhager J[2], Whitmarsh T[3], Xia C[11], Zheng P[11], Zhuang X[11].

[1]Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge CB2 0RE, UK.

[2]Department of Quantitative Biomedicine, University of Zurich, Zurich 8054, Switzerland.

[3]Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK.

[4]McGovern Institute, Departments of Biological Engineering, Media Arts and Sciences, and Brain and Cognitive Sciences, Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

[5]Department of Molecular Oncology, BC Cancer, part of the Provincial Health Services Authority, Vancouver, BC, Canada.

[6]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

[7]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.

[8]Súil Interactive Ltd, Dame Lane, Dublin, Ireland.

[9]Department of Oncology and Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, CB2 0RE, UK.

[10]Herbert and Florence Irving Institute for Cancer Dynamics, Columbia University, New York, NY, USA.

[11]Howard Hughes Medical Institute, Department of Physics and of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.

[12]Department of Oncology and Ludwig Institute for Cancer Research, University of Lausanne, Lausanne, Switzerland.

[13]Computational Oncology, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, USA

Materials and Methods

Untargeted ExSeq (Fig. 1-3)

Overview of fixation and use of mice

The first step of ExSeq is tissue fixation. For mouse tissues, we have demonstrated ExSeq using three different fixation conditions: hippocampal culture fixation (**Fig. 2A**), fixation of fresh frozen brain slices (**Fig. 2C**), and transcardially paraformaldehyde (PFA) perfused mouse brain (**Fig. 2E**). All methods for animal care and use were approved by the Massachusetts Institute of Technology Committee on Animal Care (CAC protocol 1208-100-21) and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. All solutions below were prepared from nuclease-free reagents. The mice used for **Fig. 2** and **Fig. 3** (excluding **Fig. 2A-B)** were Thy1–YFP (Tg(Thy1–YFP)16Jrs) (Jackson Labs stock #003709) male mice in the age range 6–8 weeks. Hippocampal neurons (**Fig. 2A-B**) were prepared from postnatal day 0 or day 1 Swiss Webster (Taconic SW-F, SW-M) mice without regard to sex. No sample-size estimate was performed, since the goal was to demonstrate a technology. As noted in (*95*), "in experiments based on the success or failure of a desired goal, the number of animals required is difficult to estimate." As was also noted in this paper, "the number of animals required is usually estimated by experience instead of by any formal statistical calculation, although the procedures will be terminated [when the goal is achieved]." No exclusion, randomization, or blinding of samples was performed.

Fixation of hippocampal culture (**Fig. 2A**)

Hippocampal neurons were prepared as previously described (*96*, *97*). ~1000 hippocampal neurons were cultured per coverslip. Two weeks after the culture was prepared, the neurons were fixed using 10% formalin in 1X PBS for 15 min at 25°C, then washed with 1X PBS three times and finally stored in 70% ethanol at 4°C until use.

Fixation of fresh frozen brain slice (**Fig. 2C**)

Mice were terminally anesthetized with isoflurane, then decapitated, and the brain dissected out into a cryomold with OCT. The cryomold was then placed in a dry ice/isopentane bath. Overall, freezing of the brain was completed within 5 min after euthanasia. 15 μm slices were sliced on a Cryotome (Leica) and then immediately fixed with ice cold 10% formalin in 1X PBS for 12 min. Slices were washed 3 times for 5 minutes each with ice cold 1X PBS, treated with ice cold 4% SDS in 1X PBS for 2 minutes, washed again with ice cold 1X PBS three times, and finally stored at 4°C in 70% ethanol until use.

Fixation via transcardial perfusion (**Fig. 2E**)

Mice were terminally anesthetized with isoflurane and perfused transcardially with ice-cold 4% paraformaldehyde, and the brain dissected out, and left to post-fix in 4% paraformaldehyde at 4 °C for one day, before moving it into 1X PBS containing 100 mM glycine. 50 μm slices were sliced on a vibratome (Leica VT1000S) and stored at 4°C in 70% ethanol until use.

RNA anchoring

RNAs were anchored to the hydrogel network using the chemical LabelX as previously described (*26*). Briefly, LabelX is the commercial chemical reagent Label IT® (Mirus Bio LLC) modified, by simple mixing, with the hydrogel-anchorable group acryloyl-X. In more detail, specimens were washed with 1x PBS three times to remove residual ethanol, then once with 1X MOPS buffer (20 mM MOPS pH 7.7) for 45 min (15 min for neuronal culture samples), and finally reacted overnight at 37°C with 0.018 g/L LabelX in 1X MOPS buffer.

Gelling, digestion and expansion

To allow expansion, the specimens were gelled as previously described (*23*). Briefly, tissue specimens were incubated in gelling solution (1X PBS, 2 M NaCl, 8.625% (w/v) sodium acrylate, 2.5% (w/v) acrylamide, 0.15% (w/v) N,N′-methylenebisacrylamide, 0.01% 4-hydroxy-TEMPO, 0.2% (w/v) ammonium persulfate (APS) and 0.2% (w/v) tetramethylethylenediamine (TEMED)) for 10 min at 4°C (for neuronal culture samples this incubation step is not needed). The specimens were then placed in a gel chamber, constructed by sandwiching the sample between a slide and a coverglass, with #0 coverglass spacers on either side of the sample to prevent contact between the biological specimen and the coverglass. Specimens were incubated with gelling solution at 37°C for 1.5-2 hours until gelling was completed. Next, the specimens were incubated overnight at 37°C in digestion buffer (8 units/mL Proteinase K (New England Biolabs, cat. no. P8107S), 50 mM Tris pH 8.0, 1 mM EDTA, 0.5% Triton X-100, and 5 mM 2-Amino-5-methoxybenzoic acid (for formaldehyde adduct removal (*98*)). Specimens were then expanded by incubation in excess volumes of ddH2O 3 times, for 45 mins each time.

Re-embedding

To prevent gel conformational changes during sequencing, expanded gels were re-embedded into non-expanding polyacrylamide gels, as previously described (*26*), resulting in a final expansion factor of ~3.3x. Briefly, the specimens were incubated while rocking for 30 min at 25°C with re-embedding monomer solution (acrylamide and N,N-Methylenebisacrylamide (3% and 0.15% (w/v), respectively), 5 mM Tris base, 0.075% (w/v) TEMED, 0.075% (w/v) APS). Specimens were then placed in gel chambers, constructed by sandwiching the sample between slides and coverglasses, with two #1.5 coverglass spacers (placed one on top of the other) on either side of the sample to prevent contact. Gel chambers were placed inside a closed Tupperware-like container which was filled with nitrogen for about 5 min. Gel chambers were then incubated for 1.5 hr at 37°C.

Passivation

We discovered that the carboxylic acid groups of sodium acrylate, used in the polymer of the expanding gel, could inhibit downstream enzymatic reactions important for *in situ* sequencing (**Fig. S1**). This might have been due either to a chelation effect or to interaction between the negative charges of the carboxylic group and the enzymes tested. A passivation procedure was designed to block the negative charges of carboxylic groups after expansion. The samples were treated with 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), N-Hydroxysuccinimide (NHS) to covalently react ethanolamine to the carboxylic groups, converting them to amides with no charge. This reaction was performed in two steps (*99*): the first was incubation at 25°C for 2 hours with 2 M ethanolamine hydrochloride, 150 mM EDC, and 150 mM NHS in 100 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer at pH 6.5; the second step was incubation at 25°C for 40 min with 2 M ethanolamine hydrochloride and 60 mM

sodium borate (SB) buffer at pH 8.5. The samples were then washed three times with 1X PBS for at least an hour overall. We note that as EDC can react with guanine in RNA (*100*), therefore higher concentrations of EDC, compared to the concentration used here, might have undesired side effects.

The *in vitro* experiments in **Fig. S1** were performed with 0.1-0.5% (w/v) polyacrylic acid, to mimic the concentration inside the expanded gel which is ~0.2% (8.625% (w/v) sodium acrylate before expansion corresponds to 8.625/3.5^3 or ~0.2% with expansion factor ~3.5, and between ~0.1% to ~0.5% with expansion factors between 2.5 to 4). We validated the passivation with a range of EDC concentrations, 50-150mM, that are higher than the polyacrylic acid concentration of ~30mM (~0.2% polyacrylic acid); however, we didn't want the EDC concentration to be more than a few fold higher than the polyacrylic acid concentration as EDC can react with guanine in RNA (*100*). The reverse transcription reaction was performed with M-MuLV (Enzymatics, cat. no. P7040L) according to the manufacturer's protocol, with 1.2kb Kanamycin Positive Control RNA (Promega, cat. no. C1381) as a template, and random primers. The rolling circle amplification reaction was performed with Phi29 (Enzymatics, cat. no. P7020-HC-L) according to the manufacturer's protocol. The template used was the 55-base long CircLigase II ssDNA Control Oligo, after circularization with CircLigase II (epicentre, cat. no. CL9021K) according to the manufacturer's protocol.

Library preparation for *in situ* sequencing

Following expansion, re-embedding and passivation, several enzymatic steps were carried out to prepare the samples for *in situ* sequencing. Some of these steps were adapted from the FISSEQ protocol (*25*), but modified to better work with expanded gels. Briefly, anchored RNA underwent reverse transcription, cDNA circularization, and rolling circle amplification with phi29 polymerase. The details of the reactions are below.

First, endogenous DNA was removed to allow for later *ex situ* sequencing of cDNA without DNA contamination after library preparation was complete. The specimens were incubated with DNase I (Roche; MilliporeSigma cat. no. 4716728001) at 0.5 U/µL and RNase inhibitor (Enzymatics, cat. no. Y9240L) at 0.4 U/µL in 1X DNase I buffer for 2 hours at 25°C. The reaction underwent heat inactivation at 75°C for 5 min, and finally washed for 1 hour with 1X PBS.

Next, anchored RNAs were further biochemically modified into *in situ* sequence-able form via reverse transcription. Specimens were incubated with 10 U/µl SuperScript IV (SSIV) reverse transcriptase (ThermoFisher, cat. no. 18090050), 0.4 U/µL RNase Inhibitor (Enzymatics), 5 mM DTT, 250 µM dNTP, 10 µM inosine, 1X SSIV RT Buffer and 2.5 µM random octamer reverse transcription primer. The primer sequence was /5Phos/ACTTCAGCTGCCCCGGGTGAAGANNNNNNNN for the 15 micron thick hippocampal slice (**Fig. 2C**), and /5Phos/TCTCGGGAACGCTGAAGANNNNNNNN for the 50 micron thick hippocampal slice and the hippocampal culture (**Fig. 2E** and **2A**, respectively). (The latter primer is shorter, so less self-circularization is anticipated.) The inclusion of inosine allowed generation of ~100 base long cDNA fragments in a subsequent step using Endonuclease V cutting. We performed this step because CircLigase has lower efficiency for longer cDNA strands, and therefore the overall yield was improved by including inosine. The reverse transcription reaction was done overnight at 37°C after 15 min of 4°C incubation. For the 15 micron thick hippocampal slice, to save time, the reverse transcription reaction was done with

thermo-cycling using the following program: (a) 8°C for 12 min, (b) 8°C for 1 min, (c) 37°C for 4 min, (d) back to (b) 70 times.

For the hippocampal culture and the 50 micron thick hippocampal slice (**Fig. 2A** and **2E**), 40 μM aminoallyl-dUTP was also included in the reverse transcription mixture. These samples were formalin fixed after the reverse transcription to anchor aminoallyl-dUTP-modified cDNAs to the acrylamide moieties in the gel, to ensure that cDNAs would not move during subsequent steps (**Fig. S4**). The samples were washed with 1X PBS twice for 15 min, then incubated with 4% formaldehyde in 1X PBS for 1 hour at 25°C, and finally washed three times with 1X PBS for 15 min. For the 15 micron slice, **Fig. 2C**, no cDNA anchoring was performed; some small amount of motion of resultant cDNAs may have occurred (**Fig. S4**). (We note that since cDNA is not generated in targeted ExSeq, and the padlock probes are expected to remain stationary around the anchored RNA transcript (*101*), no post-fixation was performed in these samples; see below).

As single-stranded cDNA is required for the circularization step, the samples were treated with RNase. The specimens were incubated for 2 hours at 37°C with 0.01 U/μL Riboshredder RNase blend (Epicentre, cat. no. 12500), 0.25 U/μL RNase H (Enzymatics, cat. no. Y9220L), 0.05 U/μL Endonuclease V (New England Biolabs, cat. no. M0305S) and 1X NEB 4 buffer. For the 15 micron thick hippocampal slice (**Fig. 2C**), 1 mg/mL RNase DNase-free (Roche; MilliporeSigma cat. no. 11579681001) was used instead of Riboshredder RNase blend, as the latter was temporarily out of stock.

Next, the specimens were rinsed with nuclease-free water twice to remove traces of phosphate and the cDNA underwent circularization. The circularization reaction ran for 3 hours at 60°C with 3 U/μl CircLigase II (Epicentre, cat. no. CL9025K), 1 M betaine, 2.5 mM $MnCl_2$ and 1X CircLigase buffer.

The specimens were then hybridized with the rolling circle amplification primer, which is reverse complement to the reverse transcription primers mentioned above: TCTTCAGCGTTCCCGA*G*A (* is phosphorothioate to block the exonuclease activity of phi29) for the hippocampal culture and the 50 micron thick hippocampal slice (**Fig. 2A** and **2E**), and TCTTCACCCGGGGCAGCTGAA*G*T (* is phosphorothioate) for the 15 micron thick hippocampal slice (**Fig. 2C**). The hybridization was done for 2 hours at 37°C with 0.5 μM of the rolling circle amplification primer, 30% formamide and 2X SSC buffer. The specimens were then washed for 30 min at 37°C with 30% formamide and 2X SSC buffer, and then with 2X SSC, 1X SSC and 1X PBS, each for 5 min at 25°C.

Next, rolling circle amplification was performed overnight at 30°C with 1 U/μl phi29 DNA polymerase (Enzymatics, cat. no. P7020-HC-L), 250 μM dNTP, 40 μM aminoallyl dUTP and 1X Phi29 buffer.

To cross-link the amplified cDNA molecules containing aminoallyl dUTP, the specimens were washed with 1X PBS, and then incubated for 2 hours at 25°C with 5 mM BS(PEG)9 in 1X PBS. Then the samples were washed with 1X PBS and the reaction was quenched with 1 M Tris pH 8.0 for 45 min at 25°C.

Finally, to generate **Fig. 2A**, **C** and **E**, the specimens were incubated with the ATTO 565-labeled hybridization probe, matching the reverse transcription primers mentioned above: /5ATTO565N/TCTCGGGAACGCTGAAGA for the hippocampal culture and the 50 micron thick hippocampal slice (**Fig. 2A** and **2E**), and /5ATTO565N/ACTTCAGCTGCCCCGGGTGAAGA for the 15 micron thick hippocampal slice (**Fig. 2C**). The hybridization was done for 45 min at 25°C with 0.1 μM hybridization probe, 10%

formamide, 4X SSC and then the samples were washed with 2X SSC, 1X SSC and 1X PBS at 25°C for 5 min each.

BrdU staining protocol for cDNA visualization (**Fig. S4**)

The BrdU staining protocol was performed as follows: 5-bromo-2′-deoxyuridine triphosphate was mixed with G, A, C to a concentration of 5mM. This mix was used instead of dNTPs in the reverse transcription reaction with final concentration of 0.25mM. After the reverse transcription, the RNase treatment, and the 12 hour wash in PBS, 5μg/ml anti-BrdU mouse antibody (Anti-Bromodeoxyuridine from mouse IgG1; Roche; MilliporeSigma, 11170376001) in 1x PBS was incubated for 1 hour at room temperature. After washing with 1x PBS, 10μg/ml of Goat anti-Mouse Cy5 antibody (Abcam ab6563) in 1x PBS was incubated for 1 hour at room temperature. Finally, the sample was washed again with 1x PBS. The cDNA anchoring was done as described in the **Methods** section 'Library preparation for in situ sequencing'.

Automated *in situ* sequencing

SOLiD chemistry was utilized to sequence the cDNA amplicons (*25*). We automated the sequencing procedure by combining a dedicated spinning disk confocal microscope with computer-controlled fluidics. The spinning disk used was a Yokogawa CSU-W1 coupled with a Nikon Ti-E inverted microscope with Borealis modification. The fluidics system was composed of a FCS2 flow cell (Bioptechs), modular valve positioner with HVXM 8-5 valve (Hamilton) and PTFE laboratory tubing (S1810-12, Finemech). The flow was controlled with a syringe pump for the 15 micron thick hippocampal slice (**Fig. 2C**), and peristaltic pump for the hippocampal culture and the 50 micron thick hippocampal slice (**Fig. 2A** and **2E**). (The peristaltic pump is accurate enough for high flow rates, and less prone to stop pumping in the middle of an experiment, than the syringe pump.) In both cases, a National Instruments Data Acquisition (NI-DAQ) card was used to connect the pump and the valve positioner to the scope computer. To mitigate movement during sequencing, the specimens were re-embedded again onto a Bind-silane (GE17-1330-01, GE Healthcare) treated coverslip, and the coverslip with the specimen then placed inside the flowcell. The treatment of coverslips with Bind-silane is described in (*26*), and the re-embedding was performed using the same protocol as in the 'Re-embedding' section above with the addition of fluorescent beads (Tetraspeck 0.2 μm, Life Technologies (T7280), diluted 1:100 into the re-embedding solution), to allow color correction as described below. The hippocampal culture and the 50 micron thick hippocampal slice (**Fig. 2A** and **2E**) were re-embedded on the same coverslip and sequenced together, whereas the 15 micron thick hippocampal slice (**Fig. 2C**) was re-embedded on a different coverslip and sequenced separately. The lines of the valve positioner were populated with the chemicals required to perform the SOLiD sequencing by ligation *in situ*. The chemicals that required temperature other than 25°C were temperature controlled with Mini Dry Baths (Fisher Scientific).

For SOLiD sequencing by ligation chemistry, the fluidics setup was utilized to perform the following sequence of reactions; for each one of the 5 sequencing primers (**Table S1**), the specimen was first stripped for 30 min to remove the hybridization probe or the previous sequencing primer with strip solution (80% formamide and 0.01% Triton-X in water). Next, the specimen was washed with 1X instrument buffer (SOLiD Buffer F, 1:10 diluted) for 10 min, and incubated for 20 min with 2.5 μM of sequencing primer in 5X SASC (0.75 M sodium acetate, 75 mM tri-sodium citrate, pH 7.5). After a 5 min wash with 1X instrument buffer, the specimen was reacted for 1 hour with T4 DNA ligation mixture (6 U/μl T4 DNA ligase (Enzymatics, cat. no.

L6030-HC-L) and 1:40 diluted SOLiD sequencing oligos in 1X T4 DNA ligase buffer). The specimen was then washed for 1 hour with 1X instrument buffer and imaged with SOLiD imaging buffer. To acquire the next base a two-step cleave reaction was performed, first with SOLiD buffer C for 30 min (part #4458932) and then SOLiD buffer B for 15 min (part #4463021). The cleave reaction was followed by a 10 min wash with 1X instrument buffer, 1 hour with T4 DNA ligation mixture, 1 hour wash with 1X instrument buffer and finally imaging with SOLiD imaging buffer (SOLiD buffer A, part #4463024). The cleave-ligation-wash-imaging cycle was repeated 3 times for each one of the 5 sequencing primers. For the hippocampal culture and the 50 micron thick hippocampal slice, a dephosphorylation reaction was performed before each cleave reaction to reduce phasing. (The 15 micron slice experiment was done earlier, before we realized that phasing could be an issue; we recommend the phosphatase for routine work.) The reaction was for 30 min with 1:20 dilution of Quick CIP (NEB, cat. no. M0508L) in 1X CutSmart buffer. All reactions were done at room temperature; the strip solution and the sequencing primers were kept at 80°C; the ligation mixture, the imaging buffer, and SOLiD buffer B were kept at 4°C.

The following optical configuration was used for *in situ* sequencing: (a) laser lines - 150 mW solid state OPSL 488 laser, 100 mW solid state OPSL 560 laser, 100 mW solid state DPSS 594 laser, 110 mW solid state OPSL 642 laser; (b) emission filters - 525/50, 582/15, 624/40, 685/40; (c) camera: Zyla sCMOS plus 4.2 megapixel with 100 msec exposure time; (d) objective: Nikon 40X CFI Apo, water immersion with long working distance, NA 1.15; (e) laser power and resolution: for the 15 micron thick hippocampal slice 50% laser power was used for all laser lines and the resolution was set to 0.5 microns in the Z axis and 0.17 microns in X & Y axis. For the hippocampal culture and the 50 micron thick hippocampal slice 100% laser power was used for all laser lines and the resolution was set to 0.4 microns in the Z axis and 0.17 microns in X & Y axis. The imaging time for one 40X field of view with the above configurations and 150 z-sections was ~7 min using a piezo stage. As SOLiD chemistry takes ~3.5 hours per sequenced base (see above), the chemistry+imaging time is ~5hours for 10 fields of view per base and ~100hours for a 20 base long sequencing experiment. We note that total experiment time scales with the number of bases sequenced, and therefore sequencing of short barcodes is faster, see **Methods** section 'Targeted ExSeq of Visual Cortex and Hippocampus' below.

Morphology (**Fig. 2F** and **Fig. 3**)

Following *in situ* sequencing, the 50 micron thick hippocampal slice from the Thy1-YFP mouse was first stripped for 30 min to remove the sequencing primer with strip solution. Next, the specimen was washed with 1X instrument buffer for 10 min, and incubated for 50 min with 2.5 μM of hybridization probe in 4X SASC and 10% formamide. After a 10 min wash with 1X instrument buffer, the specimen was treated with 10 μg/mL primary antibody against GFP (Rabbit Anti-GFP, Invitrogen A-11122) in 5X SASC and 0.1% Triton-X for 24 hours. After a 3 hour wash with 1X instrument buffer, the specimen was reacted with secondary antibody, 4 μg/mL donkey anti-rabbit IgG CF®633 dye (Biotium, 20125-1) with 5X SASC and 0.1% Triton-X, for 24 hours. Finally, after an additional 3 hour wash with 1X instrument buffer, the sample was imaged in an imaging buffer. All reactions were done at room temperature, the strip solution was kept at 80°C, and the imaging buffer was kept at 4°C.

Software overview and open source accessibility

The entire MATLAB library to process ExSeq datasets from the microscope to spatial analysis of gene expression is accessible at Zenodo (*29*). Specifically, this includes a linked tutorial wiki which includes instructions on how to run the pipeline and a tutorial set of targeted ExSeq data from the visual cortex. **Fig. S2** is an illustrated overview of the pipeline.

Image processing - Deconvolution

Deconvolution was used in all imaging channels for the two untargeted ExSeq experiments of the hippocampus (**Fig. 2D,F**) because the morphology signal was low, but was not used for other experiments. Deconvolution was done with Huygens software from Scientific Volume Imaging, using the CMLE mode with 10 iterations and a signal to noise ratio of 10.

Image processing - Color Correction

For color correction, each 3D image volume from a round of SOLiD sequencing was acquired in four fluorescence channels. In the case of rigid offsets for the images between fluorescence channels, which can occur depending on the acquisition order on the microscope, color correction was done using the fluorescent beads as a reference (see **Methods** section 'Automated *in situ* sequencing'). This function is included in the software pipeline repository published with this paper (see **Fig. S2**).

Image processing - Feature-based 3D Registration

Feature-based registration pipelines have three fundamental parts: keypoint detection, feature construction at a keypoint, and feature matching. Matched features then create corresponding points between different image volumes, which can then be used to calculate a warp for one image into the coordinate space of another image. We implemented this pipeline, as described below. We selected the round with the best signal to noise ratio (typically one of the early SOLiD sequencing rounds) as the reference round for registration, and warped the other rounds of imaging to match the reference. The image data was downsampled for the first four steps, providing matched features that were re-mapped into the original space for the warp calculation. All code is included in the open source repository given with this paper.

(a) Normalization: The four fluorescence channels from the SOLiD chemistry were combined into one 3D image before registration. Because each channel has a different fluorescence distribution, quantile normalization was used to normalize each fluorescence image and the adjusted color channels are then summed into one grayscale image.

(b) Calculation of 3D keypoints: The Harris Corner Detector (*102*), adapted to 3D as previously applied to 3D hydrogel registration (*103*), identified candidate points in the image that would be identifiable across all sequencing rounds.

(c) Calculation of 3D SIFT features: a feature vector was constructed at each keypoint to describe the point and its local spatial context (*104*). For a complete description, please refer to (*105*); briefly, the image gradients were calculated in the immediate neighborhood of each keypoint, at multiple scales and in multiple directions, and a vector was created from the polar histograms of the various gradients.

(d) Matching of 3D keypoints: (*104*) presented a matching algorithm for SIFT features in 2D which also is effective in the 3D SIFT features used here. To additionally filter the matches, we applied a RANSAC operation to discard correspondences that are outliers to the resulting affine warp estimated by the majority of keypoint correspondences.

(e) Warping to the reference round: once the quality correspondences were identified, a warp was calculated and applied to all SOLiD sequencing channels of the original data. We implemented both an affine transform (a global, linear operation) and a Thin Plate Spline (*106*) (which can produce much finer, localized adjustments at the cost of computational complexity), and ultimately an affine transform was sufficient when used in combination with the second re-embedding of the sample to the coverslip (see **Methods** section 'Automated *in situ* sequencing' above). For less than 1% of the fields of view, the affine transform was insufficient as determined by visual inspection and the thin plate spline was then used to improve the registration.

Image processing - Background Subtraction

When manual inspection using FIJI (*107*) determined the presence of a non-uniform background signal, a background subtraction was performed by applying a morphological opening operation (of size 5 pixels, roughly the radius of an amplicon), then subtracting the opened image from the original image and clamping any negative values at zero.

Image processing - Puncta segmentation

To identify and extract the pixels associated with each amplified cDNA, we utilized the long-established methodology of a watershed transform. To increase the signal, we first summed the registered image volumes (the grayscale 3D images from the combined normalized channel volumes then applied a background subtraction method) of the *in situ* sequencing rounds into a single volume. This composite image was then interpolated via a shape-preserving piecewise cubic interpolation in Z to achieve isotropic voxel size and the punctate signal was amplified using a Difference of Gaussians filter (similar to the commonly used Laplacian of Gaussians, used in (*10*, *108*)). The filtered image was thresholded using the Otsu method, and the binarized image was segmented using the watershed transform. After segmentation, the data was un-interpolated in Z back into the original image size. We note that by design, this segmentation strategy was less likely to generate false negatives (i.e. puncta without segmentation) and more likely to generate false positives (i.e. not existing puncta); the latter are then removed by the following quality control steps. The resulting sections were identified using connected components, and candidate puncta that were too small or too large were removed, based on a manually set threshold (less than 30 pixels, greater than 2000 pixels). Additionally, we enforced the constraint that a puncta was present in all sequencing rounds; because the background subtraction clamps pixel values at 0, the presence of fluorescent signal in any channel for a particular *in situ* sequencing round was readily detected. If a puncta was missing a fluorescent signal in any channel for more than one sequencing round, it was discarded from consideration.

Data Analysis - Basecalling

After cDNA amplicons were segmented using a 3D watershed approach, the pixels comprising the amplicons were processed for base calling. For each sequencing round, the pixels for the four color channels were quantile normalized to account for differences in fluorophore characteristics. For each channel in each amplicon, the normalized intensities of the channels were sorted and the average of the top 30 pixels were compared between channels, creating a sorted vector V. The highest channel was selected, and the confidence calculation was calculated as follows:

Chastity = $(V_1 - V_4)/ [(V_1-V_4) + (V_2-V_4)]$
Confidence = -10*log10(1-Chastity)

That is, the color intensities for each puncta were shifted to zero according to the dimmest color channel, V_4, then the chastity was calculated as the ratio of the brightest color to the sum of the top two brightest colors. This is similar to Illumina's chastity calculation, and the confidence, in rough analogy to a Phred score, is scored logarithmically. We then applied a filter to the candidate *in situ* reads, requiring that no more than 6 bases be called with a confidence less than 5.

Ex situ sequencing

The specimens were first stripped for 30 min to remove the sequencing primer with strip solution, and washed with 1X instrument buffer for 10 min. The hydrogel was then digested with 20 mM sodium meta-periodate in 1X PBS pH 6 at 37°C for 12 hours, followed by 2 min of vortexing. The DNA was extracted and purified from the digested hydrogel with Genomic DNA Clean & Concentrator Kit (Zymo Research), and then the single stranded DNA was transformed into dsDNA with NEBNext (New England Biolabs). After additional purification with the Genomic DNA Clean & Concentrator Kit, the Nextera XT DNA Library Preparation Kit (Illumina) was used to simultaneously fragment and tag the dsDNA with the adapter sequences required for the Illumina MiSeq/NextSeq *in vitro* sequencing. The Nextera XT Index Kit (Illumina) was used to add Illumina sequencing barcodes during the PCR amplification of the extracted DNA. For the 15 micron thick hippocampal slice a MiSeq instrument was used to generate paired-end sequencing reads from the extracted DNA with MiSeq® Reagent Kit v3 (2X300 cycles; Illumina). For the hippocampal culture and the 50 micron thick hippocampal slice a NextSeq instrument was used to generate paired-end sequencing reads from the extracted DNA with NSQ® 500 Mid Output Kit v2 (2X150 cycles; Illumina).

Data Analysis - Annotation of the *ex situ* sequencing data

The paired-end reads from the *ex situ* sequencing were aligned against both the mouse mRNA database (RefSeq genes, downloaded from the UCSC genome browser on 2018/08/14) and the mouse genome (version GRCm38/mm10). In both cases local alignment was used with Bowtie2 software (*109*) using default settings, and the output format was converted into browser extensible data (bed) format using the samtools package (*110*). For alignment against the genome, to avoid alignments in repetitive regions reads that were given a Bowtie mapping quality score of 10 and below were filtered out. Minimum mapping quality was not set for the alignment against mRNA to allow alignments against genes with several alternative splicing isoforms. To unify the alignments against the mRNA and the genome, the alignment against the mRNA was converted into genomic coordinates. This was done with a custom Perl script, using the alignment of the RefSeq genes against the mouse genome in bed format (downloaded from the UCSC genome browser on 2018/08/14). After converting all alignments to genomic coordinates, reads aligned against genomic regions with low complexity, detected with the WindowMasker software with SDust module (*111*) (coordinates downloaded from the UCSC genome browser on 2018/08/15), were filtered out using the intersect function of BEDTools (*112*); default settings were used with '-f 0.2', i.e. no more than 20% of the read was allowed to overlap with the low complexity regions. In addition, reads aligned against mRNA regions that couldn't be mapped in the genome were discarded. The aligned reads that passed the above

mentioned filters were annotated using UCSC gene information to rRNA (28S and 18S), tRNA, introns, exons of non-coding genes, and exons of coding genes. In addition, we detected all reads that mapped to known alternative splicing events (UCSC alternative splicing events dataset). All the gene information and the known alternative splicing events were downloaded in bed format from the UCSC table browser (on 2018/08/22), and the intersection was determined with BEDTools as described above. RefSeq annotations, i.e. transcripts names (NM and NR accessions) and associated gene symbols, were downloaded from the NCBI website (on 2018/08/20). Cases in which the two mates of a paired-end read were aligned against two different transcript names were allowed if they represented the same gene symbol (as in the case of two splice variants of the same gene); however, paired-end reads were filtered if the two mates were aligned against two different gene symbols.

Data Analysis - *Ex situ* and *in situ* sequence matching

The *in situ*-generated 20 base long sequences (see the **Methods** section 'Data Analysis – Basecalling'), in SOLiD color space (henceforth referred to as *in situ* color vectors), and the annotated *ex situ* sequences were matched as follows; first, as the SOLiD sequencing started from the 5' of the sequencing primer, this primer sequence was identified in the paired-end *ex situ* reads. The software BLASTn using default settings and e-value cutoff of 0.05 was utilized for primer identification. For each paired-end read, we required that at least one of the paired-end mates matched the sequencing primer and that the match included the primer 5' end. To generate 20 base long fragments equivalent to the *in situ* sequences, the 17 bases at the 5' end of the sequencing primer were then collected from the *ex situ* read, and combined with the three 5' bases of the sequencing primer, which are the first bases read in the SOLiD sequencing chemistry. To minimize sequencing errors, these 17 bases were collected only if each base had a Phred score of 30 or above. Next, the resulting 20 base long sequences from the annotated *ex situ* reads were converted into SOLiD color space (henceforth referred to as the *ex situ* library). We note that converting sequences in base space to SOLiD color space is straightforward and direct, especially as only bases with low probability of sequencing errors were used, and therefore the ambiguity of converting SOLiD color space to base space was avoided (*22*, *113*); for example, in converting from SOLiD color space to base space, all bases after a sequencing error would be converted incorrectly (*113*). To avoid non-informative *ex situ* and *in situ* matching, a low-complexity filter was applied both to the *in situ* color vectors and the *ex situ* library, requiring a minimum Shannon entropy of 1. Overall, the resulting annotated *ex situ* library had 710,659 entries.

To match the *in situ* color vectors to the annotated *ex situ* library, we employed a continuous distance function. Briefly, we reshaped each *in situ* color vector and *ex situ* library entry into 68-length vectors (17 *in situ* sequencing rounds x 4 colors) and normalized each sequencing round to 1 by dividing by the brightest channel's value (setting the max value to one every four entries in the 68-length vector). The *ex situ* library was transformed into a matrix of size 710,659x68, termed G, and the *in situ* color vectors were transformed into a similar matrix of size 394,957x68, termed C. All alignments were done independently for each one of the fields of view. The alignment distance, A, is calculated by $A = 17\text{-}CG^T$. Each row of A is the alignment distance to each entry of G, so the column index for the minimum distance in each row of A is the index for the matching *ex situ* library entry. This distance calculation requires a threshold under which a putative match could be considered an actual match, and that threshold was determined by setting a target false positive rate. To quantify the false positive rate of the *ex situ*

and *in situ* sequence matching, we shuffled the order of all the *in situ* color vectors and re-aligned using the same method as described above. Because there is a monotonic relationship between threshold distance and the false positive rate, we can set the threshold distance based on a desired false positive rate, which was set at 15%. We demanded that the *in situ* color vectors will uniquely match the *ex situ* library. This was achieved by removing all the entries in G (the *ex situ* library) that were aligned to C (the *in situ* color vectors), producing G'. We then realigned the *in situ* color vectors against G', and counted any entries of C that align to G' as a non-unique alignment that will be removed. We determined that 92% of all alignments (including rRNA) were unique, and importantly, 97% of the non-rRNA alignments were unique. As expected, this uniqueness is dependent on the number of bases sequenced *in situ* (**Fig. S5C**). All non-unique matches were discarded from further analysis.

Overall, we matched 115,075, 15,709, and 35,835 *in situ* color vectors with a false positive rate of 15%, 5%, and 15%, for the 50 micron thick hippocampal slice (**Fig. 2E**), 15 micron thick hippocampal slice (**Fig. 2C**), and the hippocampal culture (**Fig. 2A**), respectively. We note that reducing the false positive rate (FDR) reduces the number of aligned reads; for example, reducing the false positive rate to 10% and 5% for the 50 micron thick hippocampal slice results in 96,857 and 63,306 aligned reads, respectively. However, reducing the FDR to 10% or even 5% doesn't change the results presented in this paper. Qualitatively, expected synaptic and cytoskeletal RNAs, such as Kcnq2 and Map1b, are still obtained, as well as almost all the genes discussed in the main text: Map2, Sptbn1, Mga, Ptma, BC1, Malat1, Grik2, Gabrg2, Nob1, Map1b, Gria1, Eef1a2, Calm3, Nob1, Rbfox1, Shtn1, Ddx5, Celf2, Syt1, Syp, Josd2, Rps24 (with FDR of 10%), excluding only Cux1 and Foxg1. Quantitatively, the functional enrichment groups in the lists of genes obtained with the different FDR are very similar to one another, and to the functional enrichment groups obtained with the RNAseq dataset (**Table S6**). Moreover, the correlation between RNAseq and untargeted ExSeq with 10 volumes (**Fig. 2Gv**) still holds with the different FDR: Pearson's r of 0.41 (p-value 6.2×10^{-93}) and 0.39 (p-value 5×10^{-50}) with 10% and 5% FDR, respectively, compared to Pearson's r of 0.47 (p-value 9×10^{-164}) with FDR of 15%. Finally, the agreement between untargeted ExSeq and targeted ExSeq (**Fig. S22D**) still holds with the different FDR: Pearson's r of 0.57 (p-value 4.1×10^{-4}) and 0.53 (p-value 1.4×10^{-3}) with 10% and 5% FDR, respectively, compared to Pearson's r of 0.68 (p-value 9.75×10^{-6}) with FDR of 15%.

The expression levels of all RefSeq genes detected using ExSeq in all these conditions are given in **Table S2-S4**. For the 50 micron thick hippocampal slice, the RNAseq is also included in **Table S2**. All the raw Illumina sequencing data was deposited to NCBI Sequence Read Archive (SRA), BioProject PRJNA663046.

Data Analysis - *Ex situ* sequencing data sampling

Sampling from the full *ex situ* sequencing data, for **Fig. 2G**, was performed as follows:

(a) Paired-end reads were randomly selected from the full *ex situ* sequencing fastq file. The number of lines selected was such that the total number of counts for all RefSeq genes in the random set was 10 times that of the number in 10 FoVs, to simulate the datasets for 100 FoVs.

(b) The random set of paired-end reads was then analyzed similarly to the procedure described in the **Methods** section 'Annotation of the *ex situ* sequencing data'; briefly, the reads were aligned against the mouse RefSeq genes using local alignment with Bowtie2 software (*109*). The alignments against the mRNA were converted into genomic coordinates, and reads aligned against genomic regions with low complexity were filtered out. Cases in which the two

mates of a paired-end read were aligned against two different transcript names were allowed if they represented the same gene symbol (as in the case of two splice variants of the same gene); however, paired-end reads were filtered if the two mates were aligned against two different gene symbols.

(c) For each RefSeq gene, the number of aligned paired-end reads was counted, resulting in an expression vector that was compared to the expression vector generated using the RNAseq data; the Pearson's correlation between the log-transformed expression vectors was calculated. In addition, the fraction of RefSeq genes detected using the random set of paired-end reads, compared to RNAseq, was calculated. (d) Steps (a-c) were repeated 10 times. Box plots are presented in **Fig. 2Gvi-vii**.

Comparison of untargeted ExSeq to bulk RNAseq (**Fig. 2G**)

To perform a comparison of untargeted ExSeq to bulk RNAseq of a matching tissue sample, a coronal hippocampal slice was selected that was within 100 μm of the original 50 micron thick hippocampal slice (**Fig. 2-3**) along the antero-posterior axis. The coronal hippocampal slice was fixed with 4% PFA in exactly the same way as the original slice, washed with 1X PBS and was kept in 70% ethanol at 4℃ until RNA extraction; throughout the experimental procedure of sequencing the RNA from a matching sample, we followed the steps in the **Methods** section 'Library preparation for *in situ* sequencing' when possible. Accordingly, the DNA was removed using DNase I (Roche). RNA was extracted from these slices using an RNeasy FFPE Kit (Qiagen) following the "Deparaffinization using Melting" protocol indicated in the kit. The extracted RNA was then reverse transcribed using SSIV reverse transcriptase with random primers, following the manufacturer's protocol. The RNA was removed using RNase DNase-free (Roche) and RNase H. The single stranded cDNA was transformed into ds-cDNA with NEBNext second strand synthesis (New England Biolabs). After purification with the Genomic DNA Clean & Concentrator Kit (Zymo), the Nextera XT DNA Library Preparation Kit (Illumina) was used to simultaneously fragment and tag the dsDNA with the adapter sequences required for the Illumina *in vitro* sequencing. The Nextera XT Index Kit (Illumina) was used to add Illumina sequencing barcodes during the PCR amplification of the extracted DNA. The library was sequenced using a MiSeq instrument to generate paired-end sequencing reads with the MiSeq® Reagent Kit v3. The resulting sequencing reads were annotated as described in the **Methods** section 'Data Analysis - Annotation of the *ex situ* sequencing data' above.

To perform the comparison of untargeted ExSeq to bulk RNAseq of a matching hippocampal culture (**Fig. S8**), we followed the same steps as for the hippocampal slice above, with the following differences: (1) whereas one coverslip with hippocampal neurons was used for ExSeq, five such coverslips (cultured and fixed at the same time as the one used for ExSeq, see **Methods** section 'Fixation of hippocampal culture') were used for the matching RNAseq experiment, to compensate for the fact that the RNA is not RCA amplified in the RNAseq experiment; (2) the neurons in the five matching coverslips were first treated with ProK to allow detachment from the coverslips and collection in an eppendorf tube.

In FISSEQ, highly abundant genes, such as genes involved in translation and splicing, were underrepresented (*22*). We examined if such a detection bias exists in the untargeted ExSeq dataset by comparing it with the matching RNAseq dataset (**Table S2**). First, we detected genes that were expressed in the RNAseq dataset but were not expressed in the untargeted ExSeq with full *ex situ* sequencing data in a 50 micron thick hippocampus slice; three expression cutoffs were used: 10, 20, or 50 counts, and for each cutoff genes with counts higher or equal to the

cutoff in the RNAseq, but not in the untargeted ExSeq dataset (normalized to the RNAseq dataset by the total number of reads), were detected. Gene ontology analysis using the software package DAVID (*114*) was then performed on the detected genes with each cutoff, revealing only 2, 0, and 0 enriched ontologies (Benjamini p-value<0.01) with the expression cutoffs of 10, 20, or 50 counts, respectively. Therefore the underrepresentation of highly expressed genes from specific functional groups, reported in FISSEQ, is not observed with untargeted ExSeq. The two functional enrichment groups (Benjamini p-value<0.01) as obtained by DAVID for the expression cutoff of 10 counts are given below:

Category	Term	Benjamini p-value
GOTERM_CC_DIRECT	extracellular exosome	5.7E-3
UP_KEYWORDS	Glycoprotein	6.1E-3

Calculation of cDNA amplicon lengths (**Fig. S6**)

The Illumina reads generated from *ex situ* sequencing typically contain several repeats of the cDNA fragment, as the size of the cDNA fragments was restricted to be ~100 bases long (see **Methods** section 'Library preparation for *in situ* sequencing') and the Illumina reads were 300-600 bases long. To calculate the length of the cDNA fragments, the 300x2 paired-end reads from the 15 micron thick mouse hippocampal slice were merged using the software Pear (*115*) with default settings, resulting in the successful merge of 92.7% of the paired-end reads. The cDNA fragment repeats are flanked by the sequencing primer (introduced during reverse transcription), and therefore nucleotide-nucleotide BLAST version 2.8.1+ was used to detect all the sequencing primer locations in the merged reads. Default settings were used, with no limit on the number of possible matches. We note that setting a more stringent e-value compared to the default value of 10, for example an e-value of 0.05 which corresponds to a perfect match against the primer, had a minor effect on the resulting cDNA amplicon length distribution. The cDNA amplicon length was defined as the distance between two sequencing primer matches, using all the merged reads that contained more than one sequencing primer match; if more than two sequencing primer matches were present in a merged read, the average distance was calculated. Overall, 23.7% of the merged reads contained more than one sequencing primer match, and therefore were used in the calculation of the cDNA amplicon lengths.

The calculation of the cDNA fragment lengths was also performed on the 150x2 paired-end reads from the 50 micron thick mouse hippocampal slice; however, because the Illumina reads were shorter, the merging of the paired-end reads was less successful (only 76.3% of the reads could be merged), and the number of merged reads with more than one sequencing primer match was only 2.2% of the merged reads. Therefore, the cDNA fragment length distribution from the *ex situ* sequencing of the 50 micron thick mouse hippocampal slice is not as representative as the data presented for the 15 micron thick mouse hippocampal slice.

Demonstration of untargeted ExSeq with C. elegans, Drosophila and HeLa cell line (**Fig. S7**)

Worm fixation and cuticle reduction was adopted from the published Bouin's tube fixation protocol (*116*). The strain used in the figure was CZ1632 (WormBase; genotype: juIs76 [unc-25p::GFP + lin-15(+)] II). The strain was maintained at 20°C under standard conditions (*117*). The worms were collected from agar plates with M9 buffer (3g KH_2PO_4, 6g Na_2HPO_4, 5g NaCl, 1ml 1M $MgSO_4$, water to 1 liter, sterilized by autoclaving) into a 15 mL tube. The tube was spun down at 1000g for 2 min, and the supernatant was replaced with 10 mL of fresh M9. The M9 wash step was repeated 2 more times. The worms were then transferred to a 1.5 mL tube and

spun down to remove as much supernatant as possible without disturbing the worm pellet. The worms were placed on ice for 5 min. 1 mL of Bouin's Fixative (0.46% picric acid, 4.4% paraformaldehyde, 2.4% acetic acid, 50% methanol, 1.2% 2-mercaptoethanol; as prepared in the published protocol), prepared fresh and pre-chilled to 4°C, was then added. The pellet was resuspended and mixed well. The sample was then placed on a tube rotator and mixed vigorously for 30 min at 25°C, followed by 4 hours of incubation at 4°C. The sample was then washed 3 times with 1mL Borate Triton β-mercaptoethanol solution (BTB; see recipe below); each time the sample was spun down, the supernatant was removed, the buffer was added and the sample was mixed thoroughly. BTB was prepared fresh using 1 mL 40x Borate Buffer Stock (3.1g boric acid, 1g NaOH, water to 50 mL), 1 mL 20% Triton X-100, 0.8 mL 2-mercaptoethanol, and 37.2 mL water. The sample was then further incubated three times, for 1 hour each, in 1 mL fresh BTB on a tube rotator at 25°C. Finally, the sample was washed six times: twice with 1 mL BT (1 mL 40x Borate Buffer Stock, 1 mL 20% Triton X-100, 38 mL water), twice with 1 mL 1x PBST (1x PBS, 0.5% Triton X-100), and twice with 1x PBS. The worms were permeabilized for 1 hour with 0.25% Triton X-100 in 1X PBS at 25°C.

Drosophila larvae w1118 (Bloomington BL#5905) were kindly provided by the lab of Aravinthan DT Samuel (Harvard University). Drosophila were raised in vials or bottles with standard yeast-containing medium at 22°C with alternating 12-h cycles of dark and light.

HeLa (ATCC CCL-2) cells were cultured on CultureWell Chambered 16 wells Coverglass (Invitrogen) in D10 medium (Cellgro) supplemented with 10% fetal bovine serum (FBS) (Invitrogen), 1% penicillin–streptomycin (Cellgro), and 1% sodium pyruvate (BioWhittaker). Cultured cells were washed once with DPBS (Cellgro), fixed with 10% formalin in PBS for 15 min at 25°C, and washed three times with 1× PBS. Fixed cells were then stored in 70% ethanol at 4°C until use.

ExSeq experimental procedures for the worms, Drosophila and HeLa cells were performed according to the following **Methods** sections: 'RNA anchoring', 'Gelling, digestion and expansion', 'Re-embedding', 'Passivation', and 'Library preparation for *in situ* sequencing'. For the reverse transcription, instead of SSIV, M-MuLV (10U/µl; Enzymatics, cat. no. P7040L) was used for the worms and HeLa cells, whereas Maxima (10U/µl; Thermo Scientific, cat. no. EP0741) was used for Drosophila. Aminoallyl-dUTP was not included in the reverse transcription mix (and therefore the cDNA were not formalin-fixed), and the following primer sequence was used: /5Phos/ACTTCAGCTGCCCCGGGTGAAGANNNNNNNN. For the worms and HeLa cells, 2U/µl CircLigase II were used for circularization. To account for possible self-circularization of the primers, control samples with no reverse transcription enzyme were also processed for the worms, Drosophila, and HeLa cells, and as expected produced only a weak signal with hybridization probe after the library preparation. *In situ* sequencing was performed manually using the reagents and the enzymatic reactions outlined in the **Methods** section 'Automated *in situ* sequencing'. Imaging was performed on a Zeiss Laser Scanning Confocal (LSM710) with Nikon 40X CFI Apo, water immersion with long working distance, NA 1.15 objective, and excitation light sources and emission filters as in (*25*).

Gene ontology analysis of ExSeq data from hippocampal culture (**Fig. S8**)

Gene ontology analysis was performed using the software DAVID (*114*) on all 127 expressed genes detected with ExSeq using 10 acquired FoVs (each of the size shown in **Fig.**

S8A), in a hippocampal culture. The first few functional enrichment groups (i.e. the functions with the lowest p-values) as obtained by DAVID are given below.

Category	Term	Benjamini p-value
UP_TISSUE	Brain	3.8E-10
UP_TISSUE	Brain cortex	9.2E-7
SP_PIR_KEYWORDS	acetylation	5.5E-6
GOTERM_CC_FAT	neuron projection	1.1E-5
SP_PIR_KEYWORDS	phosphoprotein	6.4E-6
GOTERM_CC_FAT	presynaptic membrane	9.5E-5
GOTERM_CC_FAT	cell projection	2.5E-4
GOTERM_CC_FAT	synapse	3.3E-4
GOTERM_CC_FAT	dendrite	2.9E-4
GOTERM_CC_FAT	cell projection part	4.9E-4
GOTERM_CC_FAT	dendritic spine	7.9E-4
UP_TISSUE	Hippocampus	1.6E-3

Image Processing - 3D Tracing (**Fig. 3**)

To be considered a read inside a cell, the pixels of an amplicon must overlap with the annotated neuron morphology. The annotations were created by manually tracing Thy1-YFP antibody signals using Vast Lite (*118*). A cell was annotated as three parts: the soma, the dendrites and the nucleus (visually determined by the high perinuclear density of amplicons, and the reduction of puncta density inside the nucleus).

3D Viewer (**Fig. 3**)

In order to visualize the richness of the ExSeq information, we developed a 3D visualization tool using OpenFrameworks (an open-source platform for graphics and interactivity, v.0.9.8, (*119*)). Code is hosted at Zenodo (*94*). We used the cytoplasmic YFP antibody data to produce a 3D mesh of the exterior of the neuron. We then loaded each *in situ* read in space, along with its readtype (exon, intron, etc.) and gene symbol (e.g, '*Atp2c1*','*Camk2a*'). Using the exportTo3DVisualizer.m function included in the Zenodo repository, the *in situ* reads were converted from MATLAB objects into a format readable by ExSeq viewer for exploration.

Targeted ExSeq of Visual Cortex and Hippocampus (Fig. 4-5)

Gene selection

For the mouse primary visual cortex, gene panels were selected using a combination of manual and algorithm-based strategies, as described previously (*120*) and below, and required a reference single cell and single nucleus RNAseq data set from the same kind of tissue (in this case, ~12,000 single cells in mouse primary visual cortex) (*121*). First, cells were re-assigned to a more refined set of 192 types, using the published types as a starting point and a consensus of several computational methods to "over-split" the data. Second, an initial set of high-confidence marker genes were selected through a combination of literature search and analysis of the reference data. These genes were used as input for a greedy algorithm (detailed below). Third, the reference RNAseq data set was filtered to only include genes compatible with single-molecule FISH (smFISH). Retained genes had to be: 1) long enough to allow probe design (> 960 base pairs); 2) expressed highly enough to be detected (fragments per kilobase of exon model per million reads mapped (FPKM) >= 10), but not so high as to overcrowd the signal of other genes in a cell (FPKM < 500); 3) expressed with low expression in off-target cells (FPKM < 50 in non-neuronal cells); and 4) differentially expressed between cell types (top 1000 remaining genes by marker score). To more evenly sample each cell type, the reference data set was also filtered to include a maximum of 50 cells per cluster. We note that these genes were selected to be compatible with both smFISH (*120*) and ExSeq and therefore are more stringent than is required for ExSeq alone.

The main step of gene selection used a greedy algorithm to iteratively add genes to the initial set. To do this, each cell in the filtered reference data set was mapped to a cell type by taking the Pearson's correlation of its expression levels with each cluster median using the initial gene set of size n, and the cluster corresponding to the maximum value was defined as the "mapped cluster". The "mapping distance" was then defined as the average cluster distance between the mapped cluster and the originally assigned cluster for each cell. In this case a weighted cluster distance, defined as one minus the Pearson's correlation between cluster medians calculated across all filtered genes, was used to penalize cases where cells are mapped to very different types, but an unweighted distance, defined as the fraction of cells that are not mapped to their assigned cluster, could also be used. This mapping step was repeated for every possible n+1 gene set in the filtered reference data set, and the set with minimum cluster distance was retained as the new gene set. These steps were repeated using the new gene set (of size n+1) until a gene panel of 42 genes was attained. Code for reproducing this gene selection strategy is available as part of the mfishtools R library (https://github.com/AllenInstitute/mfishtools).

For experiments on the hippocampus, 35 genes of interest that were highly expressed in the synaptic neuropil in rats were selected from a prior study (*67*), and converted to their mouse homologs.

Probe Design - Overview

For each gene selected to be interrogated *in situ*, a set of DNA oligonucleotide padlock probes directly targeting the RNA transcript was designed. Each probe had four key parts: (1) a 32 nucleotide (nt) homology region (split into two 16 nt portions spanning the ligation junction at the 5' and 3' ends of the padlock probe); (2) a constant backbone region for RCA initiation and *in situ* sequencing primer binding; (3) a barcode 5' to the constant region for SOLiD sequencing readout; and (4) a barcode 3' to the constant region for Illumina sequencing readout.

The homology regions were linked to the SOLiD and Illumina barcodes with a short linker sequence 'AAA'. Transcripts were assigned barcodes in a logical barcode space (described below), which was represented in sequence space by two nucleotide sequences, one for readout using the SOLiD chemistry and the other for readout using the Illumina chemistry. This provided flexibility, enabling either chemistry to be used to read out the sequences. A schematic of the probe and amplicon product is shown in **Fig. S12.**

Probe Design - Homology and constant region

Sequences of interest were downloaded from RefSeq in GenBank format. A sliding-window sequence walk was performed down the length of a transcript, in which 32-mer regions were serially selected, and tested to identify regions passing our QC criterion. If the region passed, the window was advanced to start 5 nt past the end of the current window. If the region did not pass, the window was advanced 1 nt. As the 32-mer sequence was split between two 16-mer regions on the 5' and 3' ends of the probe, part of the QC screened each half individually. Our QC criterion for homology regions excluded any homology region with: (1) repetitive sequences (>4 consecutive identical bases); (2) GC content for either 16-mer half < 40% or > 65%; (3) melting temperature for both 16-mer halves > threshold (for visual cortex probes, the threshold was originally 55^{o}C, then subsequently lowered (to a minimum Tm of 52^{o}C) if fewer than six probes per gene were generated; for hippocampus probes, the threshold was fixed at 50^{o}C); (4) significant hairpin or dimer secondary structure; (5) a BLAST hit of the homology region against the mouse transcriptome (excluding the gene of interest) of >12 nt in length, that spanned the ligation junction by at least three nt on either side of the ligation junction. 32-mers passing the QC criterion were saved.

The constant region on the probe backbone was TCT CGG GAA CGC TGA AGA CGG C, a modified version of the universal primer sequence from FISSEQ (*22*), that was extended to increase its melting temperature for compatibility with Illumina sequencing chemistry.

Probe Design - Barcode design

Transcript barcodes uniquely identifying transcripts were designed in a logical space, consisting of four rounds of imaging, in which a logical 0, 1, 2, or 3 was decoded in each round of imaging, i.e. 3021 corresponded to reading out a 3 on the first round of imaging, a 0 in the second round of imaging, and so on. Each logical color corresponded to a physical color when read out using SOLiD or Illumina chemistries. Barcodes for transcripts were designed with the first three rounds R1, R2, R3 being independent, and with the last round being a check-sum round equal to R4 = (R1 + R2 + R3) (mod 4), i.e. 3021 is a valid barcode since 3 + 0 + 2 (mod 4) = 1, while 3022 is not a valid transcript barcode. For the hippocampus probeset, bulk expression data (from the original study (*67*)) was used to assign barcodes such that the barcodes were more evenly distributed in colorspace across all rounds of imaging. For the visual cortex dataset, the barcode sequences were evenly distributed in colorspace without weighting by gene expression.

Probe Design - Barcode implementation in SOLiD and Illumina chemistry

SOLiD barcodes were designed to be seven nt long, with two ligations on the first primer (SeqN), and two ligations on the second primer (SeqN-1). Dibase encodings correspond to logical bases such that longer wavelengths correspond to higher logical bases, i.e. dibase encodings for the lowest wavelength fluorophore corresponds to logical 0, while dibase encodings for the longest wavelength fluorophore corresponds to logical 3, except in the first

ligation on the SeqN-1 primer, in which case the color-mapping is reversed (due to a technical error in one base of the barcode design; we note that since the Illumina chemistry was used with these probes, this error has no effect on the result). All possible 7 nt barcode sequences were interpreted, and formed a library of 64 sequences per 4-round barcode.

Illumina barcodes were designed to be seven nt long, with the first four bases synthesized directly corresponding to the first four readout rounds, with the mapping between logical bases and sequence bases being ({0, 1, 2, 3} -> {A, C, G, T}) (and with the shortest to longest wavelength order being G, T, A, C). All possible 7 nt barcode sequences were interpreted, and formed a library of 64 sequences per barcode.

Probe Design - Probe assembly, negative control probes, ordering and pooling

For the visual cortex probeset, a maximum of 8 probes per transcript were selected, and for the hippocampus probeset, a maximum of 16 probes per transcript were selected (we expect that using 16 probes per gene will result in higher yield and is recommended to the user). To assemble probe sequences, the correct number of homology regions were randomly selected from the set of acceptable homology regions for a particular transcript. Each homology sequence was associated with a SOLiD and an Illumina barcode sequence that corresponds to the logical barcode associated with the transcript. The construction of the probe is described above.

In parallel, three negative control probe sets targeting the mCherry-expressing plasmid pMExt589 (*122*), the *D. melanogaster* gene Vg, and random barcodes from a transcriptome-orthogonal barcode set (*123*) were also designed as described above, with BLAST screening against the both the human and mouse transcriptome. These probes were given distinct barcodes from the transcript barcodes, enabling detection of these negative control probes.

Probes were ordered with 5' phosphate modifications in 96-well plate format from either Integrated DNA Technologies (IDT) or Eurofins Genomics. Probes for each gene were pooled together to form 100 μM or 200 μM subpool solutions. Subpools were pooled together to form stock solutions containing all probes of a particular probeset. Visual cortex and hippocampus barcodes are in **Table S9**, with probe sequences in **Table S10**. For the hippocampus probeset, one gene, Rgs5, was assigned to a highly repetitive barcode (2222), and therefore was excluded *in silico* (as it was difficult to distinguish this simple repeat from imaging artifacts) after the sequencing data was collected by removing the entry from the barcode library used to match *in situ* reads.

Tissue preparation

For experiments involving study of the visual cortex (**Fig. 4**), one seven week old C57BL/6 Thy1-YFP female mouse was terminally anesthetized with isoflurane and perfused transcardially with ice-cold 4% paraformaldehyde. The brain was dissected out and left in 4% paraformaldehyde at 4°C for 12-16 hrs. After briefly washing the brain with 1X PBS, 50 μm slices were then prepared on a vibratome (Leica VT1000s) and stored in 70% ethanol at 4°C until use. For experiments involving study of the hippocampus (**Fig. 5**), one 14 week old C57BL/6 Thy1-YFP male mouse was similarly utilized.

Targeted ExSeq library preparation

Brain sections stored in 70% ethanol were rehydrated with two 15 minute washes with PBST (1X PBS, 0.1% Triton-X) at room temperature. A coronal slice was selected that contained the primary visual cortex, and another coronal slice from a different mouse (see above)

was selected that contained the hippocampus. These were labeled with a primary antibody against GFP (which labels the YFP protein) at a concentration of 10 μg/mL Rabbit Anti-GFP (Thermo Fisher, A-11122), in PBST overnight at 4°C followed by staining with a biotinylated secondary antibody at 10 μg/mL (Thermofisher, B-2770), in PBST overnight at 4°C. We note that antibody staining which is performed pre-expansion (as described in this section) results in a better staining compared to antibody staining which is performed post-sequencing (as described in the **Methods** section 'Morphology' above); however, in our hands, *in situ* sequencing with SOLiD chemistry was not successful for slices that were antibody stained pre-expansion. Therefore, for the samples described in this section, the Illumina chemistry was utilized for *in situ* sequencing instead of the SOLiD chemistry (see **Methods** section 'Targeted ExSeq *in situ* sequencing by synthesis' below).

To allow the retention of RNA, the sections were treated with 0.1 mg/mL LabelX overnight in 1X MOPS buffer (20 mM MOPS pH 7.7), as previously described. Gelation and digestion were performed as described above for untargeted ExSeq. Re-embedding was also performed as described above except with the following re-embedding solution composition: Acrylamide and N,N-Methylenebisacrylamide, (4% and 0.2% (w/v), respectively), 5 mM Tris base, 0.05% (w/v) TEMED, 0.05% (w/v) APS. Passivation was carried out as described above.

Probes were pooled together to form 200 μM stock libraries for the visual cortex probeset and 81 μM for the hippocampus probeset (these two probeset were purchased from different vendors with different probe stock concentrations). The gelled and passivated samples were then pre-incubated for 30 minutes at room temperature with wash buffer (20% formamide, 2X SSC buffer), then incubated overnight at 37°C with the respective hybridization mix. For the visual cortex experiments, the hybridization mix consisted of 54 μM pooled visual cortex library probeset (total of 334 probes; ~162 nM each probe) in 20% formamide, 2X SSC buffer. For the hippocampus experiments, the hybridization mix consisted of 54 μM pooled hippocampus library probeset (total of 540 probes; 100 nM each probe), 4.4 μM negative control probes (total of 44 probes; 100 nM each probe) in 20% formamide, 2X SSC buffer.

The samples were then washed twice with wash buffer at 37°C for 30 minutes each, followed by a wash with 1X PBS at 37°C for 30 minutes, and a pre-incubation with 1X SplintR Ligase Buffer for 30 minutes at room temperature. The samples were then incubated with 1250 U/mL SplintR ligase (NEB, cat. no. M0375L) in 1X SplintR ligase buffer at 4°C for 6 hours. The samples were then incubated with a freshly prepared solution of 1250 U/mL SplintR ligase in 1X SplintR ligase buffer overnight at 37°C. Following ligation, the samples were washed with 2X SSC for 30 minutes at room temperature followed by a pre-hybridization with wash buffer for 15 minutes at room temperature.

The samples were then incubated with 500 nM rolling circle amplification primer (TCT TCA GCG TTC CCG A*G*A, where * denotes phosphorothioate backbone modification) in wash buffer for 2 hours at 37°C, followed by a 30 minute wash with wash buffer at 37°C, and another wash with 1X PBS for 15 minutes at 37°C.

After a pre-incubation with 1X Phi29 buffer for 15 minutes at room temperature, the samples were incubated with 1000 U/mL Phi29 polymerase (Enzymatics) in 1X Phi29 buffer for 6 hours at 4°C. Following this step, the samples were incubated with 1000 U/mL Phi29 polymerase, 250 μM dNTP, 40 μM aminoallyl dUTP in 1X Phi29 buffer at 30°C overnight. The next day, the samples were washed once with 1X PBS at room temperature for 30 minutes, and then treated with 5 mM BS(PEG)9 (ThermoFisher, cat. no. 21582) in 1X PBS for 2 hours at room temperature. The samples were then again washed with 1X PBS for 15 minutes at room

temperature followed by another wash with 1 M Tris, pH 8 for 15 minutes at room temperature. After this step, the samples were then washed once with 1X PBS.

Targeted ExSeq *in situ* sequencing by synthesis

For the stable, multi-round imaging required for *in situ* sequencing, samples were immobilized to the bottom of a 24-well glass-bottom plate. To prepare the plate, wells were treated with Bind-silane (GE17-1330-01, GE Healthcare) as described previously (*26*). Subsequently, the gelled samples were re-embedded in individual wells within a re-embedding gel (Acrylamide and N,N-Methylenebisacrylamide (4% and 0.2% (w/v), respectively), 5 mM Tris base, 0.05% (w/v) TEMED, 0.05% (w/v) APS) along with 0.2 µm TetraSpeck beads (Life Technologies, cat. no. T7280), which were diluted 1:100 in the re-embedding solution. A droplet of the re-embedding solution was transferred to a well of a bind-silane treated plate. The gel containing the sample was placed on top of the drop, and oriented so that the tissue was on the top of the sample (farthest from the glass). Another droplet of the re-embedding solution was added on top of the sample, and a 10 mm circular coverglass was placed on top. The remaining volume underneath the coverglass and around the gel was back-filled. The plate was placed into a Tupperware container which was purged with nitrogen gas for 5 min, then gelled at 37ºC for 1.5 hours. Following gelation, the sample was washed with 1X PBS for 1 hour at room temperature.

To prevent background from base addition to exposed 3' DNA ends of cellular DNA during sequencing, the surface-attached samples were blocked with dideoxynucleotides using Terminal deoxynucleotidyl transferase (TdT, New England Biolabs, cat. no. M0315L). First the samples were pre-incubated with 50 µM ddNTP, 250 µM $CoCl_2$, in 1X TdT buffer for 20 minute at room temperature. Then, the samples were incubated with 400 U/mL, 50 µM ddNTP, 250 µM $CoCl_2$ in 1X TdT buffer for 90 minutes at 37ºC. After TdT treatment, the samples were washed with 1X PBS for 30 minutes at room temperature, followed by a wash with 4X SSC for 20 minutes at room temperature. The Illumina sequencing primer (TCT CGG GAA CGC TGA AGA CGG C) was then hybridized to the amplicons at 2.5 µM sequencing primer in 4X SSC at 37ºC for one hour. The samples were then washed four times with 4X SSC for 10 minutes each wash at 37ºC.

At this step, the samples were ready for *in situ* sequencing via Illumina Sequencing-by-Synthesis. For this purpose, we collected the incorporation mix buffer (IMT), imaging buffer (SRE), and cleavage buffer (EMS) solutions from a MiSeq V2 kit, and aliquoted and stored them at -20ºC. Each round of sequencing involves a step of base addition, imaging, and cleavage. For adding a base, samples were first washed three times for 10 minutes at room temperature with PR2 buffer (part of the MiSeq V2 kit), and then treated with MiSeq V2 IMT buffer for 20 minutes at room temperature, followed by another wash with IMT for 10 minutes at room temperature. Then, the plates with the re-embedded samples incubating in IMT were transferred to an incubator heated to 60ºC for 10 minutes. Following this step, the samples were washed three times, 15 minutes per wash, with PR2 buffer at 60ºC.

After base addition, samples were ready for imaging. The samples were stained with DAPI (1 mg/L) in PR2 buffer, and washed twice for 15 minutes at room temperature with the MiSeq V2 SRE imaging buffer. The samples were washed again with SRE imaging buffer, and were then imaged on an Andor Dragonfly Spinning Disk confocal with laser lines 100 mW solid state 405 laser, 150 mW solid state 488 laser 150 mW solid state 561 laser, 160 mW solid state 633 laser; emission filters, 450/50, 540/30, 631/36, 676/37, 775/140; Andor Zyla sCMOS 4.2 plus with 200 msec exposure time; and a Nikon 40X CFI Apo, water immersion with long working

distance, NA 1.15 objective. Tiled images of the entire primary visual cortex and the hippocampus were acquired using the Andor Fusion software. For each tile, 151 Z-sections were imaged at a spacing of 0.4 μm.

After imaging, samples were then incubated with the Illumina cleavage buffer, first with two washes each for 10 minutes at room temperature, and then finally with an incubation with the cleavage buffer at 60°C for 20 minutes. Following this step, the samples were washed three times for 10 minutes with PR2 buffer. At this point, the cleavage step is complete and the process was repeated starting from the base addition step. A total of four rounds of base addition, imaging, and cleavage were performed for the visual cortex and hippocampus samples, though cleavage was not performed after the final round of sequencing in order to retain the signal from the final base addition. After the final round of sequencing was imaged, the samples were washed with PR2 buffer for 30 minutes at room temperature, followed by an overnight incubation at 4°C with 10 μg/mL Alexa 488-labeled streptavidin (ThermoFisher, cat. no. S11223) in PR2 to visualize YFP immunostaining. The samples were washed three times for 30 minutes with PR2 at room temperature, and stained with DAPI and imaging buffer (as described above for sequencing). Afterwards, all the fields of view in the samples were imaged including the streptavidin staining marking YFP as well as the signal from the final round of sequencing. By imaging the YFP morphology along with the final round of sequencing, the YFP protein signal was co-registered to the sequencing data.

Image Processing - Color correction, registration, segmentation

Color correction, registration and segmentation were all done as described in the **Methods** sections 'Image processing' above, with the sole exception that background subtraction was not used because it was not deemed necessary by manual inspection.

Image Processing - Basecalling and alignment to barcodes

Basecalling was done as described for the untargeted ExSeq (see **Methods** section 'Data Analysis – Basecalling') with one exception that candidate puncta were not filtered by confidence. Instead of filtering on confidence as was done in longer reads of untargeted ExSeq, puncta for targeted ExSeq were required to have at least one color channel "present" per sequencing round, where presence for a channel is defined as a signal above a channel's median value across all puncta in a sequencing round. Puncta that did not have a present signal for each round were discarded. Alignment was done by determining the base for each amplicon per round, then calculating the Hamming distance of that sequence to the list of barcodes in the probeset. Only perfect matches (agreements across all four bases) were kept.

Image Processing - Visual cortex cell segmentation

Using the DAPI nuclear stain, 3D centroids of each neuron were determined via a thresholding of the nuclear stain. Briefly, each volumetric image had outlier pixel intensities removed by setting a maximum value at the 99% percentile of nuclear stain values. The image was then blurred to account for nonuniformity of nuclear structure, and then a threshold value was determined using Otsu's method of all non-zero voxels in the image. This threshold value was then used to binarize the image, and connected components was used to filter out putative nuclear images of insufficient size (set as 50000 voxels, corresponding to a volume of 15.5 μm^3). As some nuclei were erroneously split, nuclei with centroids within a distance of 5 microns of each other were merged into a single object (shown in **Fig. S14A**). The final set of nuclei then

had their centroids calculated, which then were used to estimate each cell as a sphere of radius 12.5 μm. ExSeq reads within a cell's volume were assigned to that cell. If a read was within the volume of multiple cells, it was assigned to the cell with the closest nuclear centroid (shown in **Fig. S14B**). This segmentation was repeated with a radius of 7.5 μm for the robustness analysis (**Fig. S18**).

Image Processing - Localization of reads within spines, dendrites, and cell bodies of the hippocampus (**Fig. 5**)

The locations of spine heads were identified using the commercially available software, MBF Neurolucida 360. Fields of view of interest were first loaded into the software one at a time. Dendrites of YFP expressing neurons were traced using the software's semi-automated tracing modality. Following this step, spines on traced dendrites were automatically identified. The output of the software consists of the coordinate of spine heads as well as other physical characteristics which were used for subsequent analysis. Axons were traced using the software's semi-automated tracing modality.

To localize reads within spine heads, a threshold was applied to the deconvolved images of YFP expressing cells to generate a binary mask. Reads were attributed to spine heads if the centroid of the transcript resided within a positive region of the mask and within 250 nm from the centroid of a spine head.

Similarly, to localize reads within dendrites, a threshold was applied to the deconvolved images to generate a binary mask to be applied to the sequencing data. To identify dendrites in different regions of the hippocampus, the software VAST was used to annotate regions of the hippocampus on Z-projected stitched images of the entire dataset. Reads were then assigned to dendrites within annotated regions of the hippocampus depending on whether their centroid resided within dendrites. To identify reads within YFP cell bodies, VAST was used to manually create masks for YFP neuronal cell bodies within CA1 and the dentate gyrus. VAST was also used to segment the neuropil region of CA1 into 50 micron segments (**Fig. 5E**).

Data Analysis - ExFISH and ExSeq comparison in cultured cells (**Fig. 4B**)

HeLa cells (ATCC CCL-2, RRID CVCL_0030) were cultured on CultureWell Chambered Coverglasses (Thermo Scientific, C37000) in D10 medium (Cellgro) supplemented with 10% fetal bovine serum (FBS) (Invitrogen), 1% penicillin–streptomycin (Cellgro), and 1% sodium pyruvate (BioWhittaker). Cultured cells were washed once with DPBS (Cellgro), fixed with 10% formalin for 10 min, and washed twice with 1X PBS. Fixed cells were then stored in 70% ethanol at 4°C until use. Cultured cells were treated with LabelX (0.01 mg/mL), and expanded, as described previously following the tissue gelation/expansion protocol (*26*). Samples were subsequently re-embedded and passivated as described above in the targeted ExSeq library preparation protocol.

A direct comparison of ExFISH to targeted ExSeq was performed by carrying out these processes sequentially within the same expanded cells. HCRv3.0-amplified ExFISH (HCRv3.0-ExFISH) was performed using a modified ExFISH protocol (as described below), with HCRv3.0 probes and reagents (Molecular Instruments) (*26*, *58*). Modifications to the original protocol were designed to avoid using dextran sulfate in hybridization/amplification steps, as it was found to partly inhibit downstream enzymatic reactions. All steps were performed with gels in PCR strips, with 200 μL liquid volumes unless otherwise noted. To perform HCRv3.0 labeling, samples were pre-incubated with HCRv3.0 Wash Buffer for 30 minutes at room temperature.

Samples were then hybridized with HCRv3.0 probes targeting a single gene that were diluted in HCRv3.0 Wash Buffer (8 nM total final probe concentration; 1.6 pmol probes into 200 μL total volume) at 37ºC overnight. Four samples were prepared, using HCRv3.0 probes against one of GAPDH, EEF2, TFRC, VIM (**Tables S9, S10**). Samples were washed with HCRv3.0 Wash Buffer four times for 30 minutes at 37ºC, followed by four 15 minute washes with 5X SSCT (5X SSC buffer with 0.1% Tween-20) at RT. HCR amplification was carried out with the corresponding HCR amplifiers (hairpins) labeled with Alexa 546. Briefly, a pair of HCR amplifiers (i.e. BxH1 and BxH2, where x denotes the initiator type; stock solution 3 uM) were individually snap cooled by heating to 95ºC for 90 seconds, then annealed at room temperature for 30 minutes in a dark drawer. The amplification buffer was prepared by mixing 2 μL of each snap cooled HCR amplifier with 96 μL of 5X SSCT, forming a total volume of 100 μL with final concentration 60 nM each hairpin. Samples were incubated in the amplification buffer for four hours at room temperature. After washing four times with 5X SSCT for 30 minutes each, gels were transferred to individual wells of a 24-well glass-bottomed plate. Gels were stained with DAPI (1 mg/L) in 5X SSCT for 10 minutes at room temperature. After staining, samples were imaged using a spinning-disc confocal microscope (see 'Automated *in situ* sequencing' above) above with a Z-step of 0.5 μm.

Following imaging, HCRv3.0 probes and amplifiers were stripped, and a targeted ExSeq library was prepared. To strip HCR reagents, samples were pre-incubated with 500 μL strip buffer (80% formamide, 20% water (v/v)) for 15 minutes at RT in the 24-well plate. Samples were then washed with 500 μL of strip buffer, and incubated at 37ºC for 2 hours, followed by six 20 minute washes with 500 μL strip buffer at 37ºC. After stripping, samples were transferred back to PCR tubes. Targeted ExSeq library preparation was then performed for the same gene that was interrogated with HCRv3.0-ExFISH. The library preparation was carried out as described above, using 8 padlock probes per transcript, at a concentration of 100 nM for each probe. After library preparation (ending with washing with PBS after BS-PEG9 cross-linking), amplicons were visualized by hybridizing an Alexa 546-labeled detection oligo (/5Alex546N/TCTCGGGAACGCTGAAGA, where /5Alex546N/ is the Alexa 546 modification) at 100 nM in 2X SSC with 10% formamide at 37ºC for 1 hour. Samples were then washed twice with 2X SSC, 10% formamide for 15 minutes each at 37ºC. Finally, the samples were washed with 1X PBS for 15 minutes at RT, transferred to individual wells of a glass-bottom 24-well plate, and stained with DAPI (1 mg/L). The same regions that were previously imaged were identified and imaged again on the same spinning disk confocal microscope, with a Z-step of 0.5 um.

Analysis for each cell was spot-counting quantification for ExFISH amplicons and for targeted ExSeq amplicons for the same gene in the same cell. Using a custom spot-counting MATLAB code developed by the Raj lab (complete source code and instructions can be found at https://bitbucket.org/arjunrajlaboratory/rajlabimagetools/wiki/Home), spots were identified using manual thresholding, counted, and extracted from pairs of image stacks corresponding to ExFISH and targeted ExSeq for the same gene in the same cell. Pearson's correlation was computed for the log10-transformed data.

Data Analysis - Clustering and t-SNE embedding (**Fig. 4D**)

Analysis was performed using a custom MATLAB script. Segmented cells (as described previously) with greater than 50 ascribed reads (thresholded after inspecting the data) were retained for subsequent analysis. The gene expression profiles of the cells were standardized

using the Z-transform. Z-score values were used for subsequent analysis unless explicitly noted. *K*-means clustering was performed on the expression profiles using *K*=15 clusters. Briefly, *K*=15 was selected based on a prior study (*61*); the 49 clusters of the original paper were grouped into 15 groups: (1, 2) two groups of non-SST, VIP, PVALB GABAergic neurons; (3, 4) two groups of VIP neurons; (5, 6) two groups of SST neurons; (7, 8) two groups of PV neurons; (9) L2/3 neurons; (10) L4 neurons; (11, 12) two groups of L5 neurons; (13, 14) two groups of L6 neurons; (15) glia. The cells were embedded into a two-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) directly on their gene expression profiles, with coloring of points in the t-SNE plot corresponding to their *K*-means cluster.

Clusters were annotated based on the expression of known and novel marker genes (using the dataset from a prior paper (*61*) or the Allen Mouse Brain ISH Atlas (*63*)), and physical location of cells in the ExSeq dataset. Excitatory clusters (those with an annotation ending in "Ex") were identified as those highly expressing at least one of the following marker genes (for corresponding cortical layer(s) indicated in parenthesis): Foxp2 (L6), Sez6 (L5, L6), Galnt14 (L5, L6), Fezf2 (L5), Kcnk2 (L5), Rorb (L4, L5), Lingo2 (L2/3), Cux2 (L2/3), or Pcdh8 (L2). Excitatory clusters were ordered and named based on the physical ordering of the layers, and the known expression patterns of the marker genes. This resulted in the nine ExSeq clusters L6 Ex, L6a Ex, L5b Ex, L5/L5a Ex, L5 Ex, L4 Ex, L2/3 + L4 Ex, L2/3 Ex, and L2 Ex. Inhibitory clusters were identified by their low expression of the marker genes for excitatory neurons listed above. Of these six clusters, five had clear and unique marker genes. The cluster strongly expressing Pvalb was annotated as PV. The four clusters with marker genes Unc13c, Chodl, Nts, and Tnni3k were annotated as SST Unc13c, SST Chodl, SST Nts, and SST Tnni3k respectively based on the the specific expression of the marker genes in minimally-overlapping subsets of SST neurons in the prior study. The remaining cluster, ultimately annotated as GABAergic (-PV), indicating non-PV GABAergic neurons, expressed a number of marker genes including Gad2, Prox1, Npas1, and Nr2f2. These genes were studied in prior work (*61*). Gad2 was strongly expressed in nearly all interneuron types, Prox1 in VIP interneurons, Npas1 in a subset of VIP interneurons and a subset of non-PV/SST/VIP GABAergic neurons, and Nr2f2 in subsets of VIP interneurons and non-PV/SST/VIP GABAergic neurons (that did not overlap with those expressing Npas1). As Pvalb (and strongly correlated markers Ank1, Slc32a1, Lhx6, Kcnmb2) were not strongly expressed in this ExSeq cluster, the cluster was annotated as GABAergic (-PV) to denote the mixed GABAergic neuron composition, excluding $Pvalb^{+}$ interneurons.

Annotation of the clusters for the robustness analysis was performed similarly. The same set of marker genes as above was used to identify excitatory clusters, which were named and ordered based on the physical ordering of the layers in space and the known expression patterns of the marker genes, resulting in the eight ExSeq clusters L6 Ex, L6a Ex, L5b Ex, L5/5a Ex, L4 Ex, L2/3 Ex, L2/3 + L4 Ex, L2/3 + L5b Ex. The latter two clusters included cells from multiple cortical layers. Annotation of the seven remaining clusters was performed similarly as above. Six of the seven clusters had strong marker genes. The cluster expressing Pvalb was annotated as PV, and clusters expressing Grin3a, Tnni3k, Unc13c, Chodl, and Nts were annotated as SST Grin3a, SST Tnni3k, SST Unc13c, SST Chodl, and SST Nts, respectively, based on the specific expression of the marker genes in minimally overlapping subsets of SST neurons in the prior study. The remaining cluster expressed a number of marker genes including Gad2, Prox1, and Nr2f2. As described above, these encompass multiple interneuron types. As markers of PV neurons (Pvalb, and strongly correlated markers Ank1, Slc32a1, Lhx6, Kcnmb2) and SST

neurons (correlated markers Chodl, Cdh9, Thsd7a, Tnni3k, Nts) were not highly expressed in this cluster, the cluster was annotated as GABAergic (-PV, SST).

Image Processing - Making multicolor images of reads in space (**Fig. 4C, 4F**)

Raw images were generated using a custom MATLAB script and merged using FIJI (*107*). For images showing reads, the relevant set of reads was loaded into a custom MATLAB script that Z-projected the reads into a 2D image, with centroids rounded to the nearest pixel (corresponding to 50 nm pre-expansion). For each read, a 2D circle with radius corresponding to 0.5 microns (pre-expansion) was drawn at the centroid location in the relevant color. The script produced three images representing RGB channels. Images were downsampled by a factor of 2 before saving. For **Fig. 4C**, all reads were used and colored identically; for **Figs. 4F, S14C-D, S18B**, reads were colored depending on the cluster assignment of the cell or randomly by cell as indicated. Images of the YFP morphology across the entire dataset were maximum intensity Z-projected and stitched together. The RGB images and YFP morphology images were merged in FIJI, with the YFP morphology shown in grey.

Data Analysis - Correlation to previously known cortical cell types (**Fig. 4E**)

Data from a previous study was downloaded (http://casestudies.brain-map.org/celltax) and RPKM data was imported into a custom MATLAB script (*61*). The 42 genes in the visual cortex dataset were identified and used for subsequent clustering analysis. Gene expression profiles for each cell were standardized using the Z-transform, and clustered using *K*-means with $K=15$ clusters. Clusters identified by performing *k*-means clustering on the previous dataset were annotated by examining the correlation of cells within the cluster to the 49 clusters previously identified. Clusters corresponding to excitatory neurons were annotated by their layer (L6 Ex; L6a Ex; L5b Ex; L5/L5a Ex; L4, L5 Ex; L2/3, L4 Ex; and L2/3 Ex). A glia cluster (annotated Glia) was also identified. The remaining clusters were identified as interneuron clusters. Four clusters had strong marker genes, and were clearly attributable to classical interneuron types, and were annotated by their classical type and marker gene: PV Ank1, PV Thsd7a, SST Unc13c, SST Chodl. Of the three remaining clusters, two corresponded to classical types: VIP (annotated without a marker gene as it was the only strongly VIP-correlated cluster), and PV (annotated without a marker gene as a strong subtype marker gene was not identified. The remaining cluster strongly correlated to non-PV/SST/VIP GABAergic neurons and was annotated as GABAergic (-PV/SST/VIP).

The mean expression profile for each cluster was computed. Mean expression profiles were also computed for clusters in the targeted ExSeq dataset, and the Pearson's correlation matrix between mean cluster expression profiles was computed. This was also performed for the segmentation robustness analysis in **Fig. S18**.

The data from the prior study was also clustered using the variable genes within the 24,057 genes in the original dataset (*61*) (**Fig. S17**), instead of clustering using the 42 genes interrogated by ExSeq. First, genes that exceeded a minimal level of expression were selected by summing the genes' RPKM values across all cells, and selecting genes with a sum greater than 1000, resulting in a subset of 13,392 genes. Coefficients of variance (CV) were computed for each one of these genes, and genes with a CV greater than 0.75 were retained, resulting in a final subset of 12,604 genes. Downstream analysis was performed as described above, with a similar annotation approach. Six clusters corresponding to excitatory neurons were annotated by their layer(s) (L6a Ex, L5b Ex, L5a + L6b Ex, L4 Ex(1), L4 Ex(2), L2/3 Ex). The L5a + L6b Ex cluster spanned

multiple cortical layers, and two similar L4 clusters were identified (L4 Ex(1) and L4 Ex(2)). Seven interneuron clusters were identified and annotated by the expression of marker genes (VIP Nr2f2, VIP Dlx4, VIP Kcnip4, PV, SST Npas1, SST Thsd7a, SST Chodl). The remaining interneuron cluster corresponded to a mix of non-PV/VIP/SST interneuron types, and was annotated GABAergic (-PV/VIP/SST). Finally, a cluster corresponding to glial cells was annotated as Glia.

Data Analysis - Distribution of cell types across cortical layers (**Fig. 4G**)

Images of the transcriptomically-defined clusters corresponding to excitatory neurons (from the clustering analysis) were exported into RGB images using custom MATLAB scripts and FIJI, as described above. The images were partitioned into layers with VAST (*118*), using the DAPI nuclear staining and coronal section images from the Allen Reference Atlas (https://mouse.brain-map.org/static/atlas)(*63*) to guide the segmentation boundaries. Seven compartments were identified: (1) external capsule (corresponding to white matter below the visual cortex); (2) L6/L6a; (3) L5b; (4) L5/L5a; (5) L4; (6); L2/3; and (7) L1. The layer segmentations were exported as images and loaded into a custom MATLAB script, which first computed the centroids of cells as the mean of the reads ascribed to a cell, and then assigned cells to layers according to the location of the centroid. Bar plots of the cluster distribution by physical layer and the physical layer distribution by cluster were subsequently generated.

Data Analysis - Comparison of targeted ExSeq to *ex situ* sequencing and bulk RNAseq

To perform comparison of targeted ExSeq to *ex situ* sequencing of libraries, the expanded hippocampal slice processed with targeted ExSeq was processed for *ex situ* sequencing via digestion, DNA fragmentation, amplification, and MiSeq Illumina sequencing as described above in the Methods section, "*Ex situ* sequencing". The resulting sequencing reads were aligned against the probe sequences using Bowtie2 (*109*) with local alignment and default settings. The output of Bowtie2 was parsed to calculate the number of sequencing reads aligned against each probe sequence, and this information was then merged to give the number of sequencing reads aligned against each targeted gene (as several probes were used for each targeted gene).

To perform a comparison of targeted ExSeq to bulk RNAseq of a matching tissue sample, two coronal hippocampal slices, from the same mouse, were selected that were within 100 μm of the hippocampal slice along the coronal axis. RNA was extracted from these slices using an RNeasy FFPE Kit (Qiagen) following the "Deparaffinization using Melting" protocol indicated in the kit. The extracted RNA was then reverse transcribed and PCR amplified using the NEBNext Ultra II RNA Library Prep Kit using the universal PCR primers of the kit. The amplified cDNA was then sequenced on a MiSeq system following the instructions for a "4 nM library". The resulting sequencing reads were aligned against mouse mRNA (RefSeq genes, downloaded from the UCSC genome browser on 2018/08/14), using Bowtie2 (*109*) with local alignment and default settings. The output of Bowtie2 was parsed to calculate the number of sequencing reads aligned against each mouse mRNA.

Targeted ExSeq of Metastatic Breast Tumor Sample (Fig. 6)

Gene selection and probe design

To select a set of 300 genes for spatial profiling in metastatic breast cancer (MBC) samples, a preliminary list of ~ 600 potentially relevant genes was first assembled based on prior knowledge and literature, as well as MBC single cell and single nucleus datasets. Genes were chosen to represent various aspects of breast cancer biology, metastasis, and the tumor-immune-microenvironment, as well as cell types and programs discovered from the single cell and single nucleus RNAseq data. The preliminary list was then filtered down to 300 genes based on expression statistics as measured in the MBC single cell RNAseq data set. During probe design, three of the selected 300 genes were excluded as they did not meet technical criteria (specifically all three transcripts were too short), reducing the final gene set to 297 genes. For each selected gene we designed up to 16 probes and no less than 9 probes. The design of the probes was as described in the **Methods** section 'Targeted ExSeq of Visual Cortex and Hippocampus' above, with the following two exceptions: (a) the barcodes were generated using the R package DNABarcodes (*124*); we used barcodes of length 7 that had a hamming distance of 3 between them (with the following command: create.dnabarcodes(7, dist=3, heuristic="ashlock",cores=12, filter.gc=FALSE,population=500, iterations=500), therefore allowing correction of one substitution error in the *in situ* sequencing step; (b) we followed the design in **Fig. S12**, with the SOLiD and Illumina barcodes having the same sequence; the SOLiD one was encoded in color space and the Illumina one is in base space. Accession numbers and probe sequences are in **Tables S9** and **S10**, respectively.

Tissue preparation

As part of an ongoing research study on metastatic breast cancer, biopsies are collected from patients at Dana Farber Cancer Institute. Prior to any study procedures, the patients provide written informed consent for a research biopsy and subsequent analysis of tumor and normal samples, as approved by the Dana-Farber/Harvard Cancer Center Institutional Review Board (DF/HCC Protocol 05-246). For this study, we used an 18-gauge core needle biopsy (~6x0.8 mm) of a liver metastasis obtained from a 66-year-old woman with a known diagnosis of hormone receptor positive metastatic breast cancer. Surgically dissected metastatic tumor samples were quickly washed with 1X PBS and placed into a cryomold with OCT. The cryomold was then placed in a dry ice/isopentane bath. 8 μm slices were prepared on a Cryotome (Leica) and adhered to Superfrost Plus glass sides, which were then immediately fixed with ice cold 10% formalin in 1X PBS for 12 min. Slices were washed 3 times for 5 minutes each with ice cold 1X PBS and finally stored at 4°C in 70% ethanol until use.

Targeted ExSeq library preparation

To allow the retention of RNA, tumor sections were treated with 0.1 mg/mL LabelX overnight in 1X MOPS buffer (20 mM MOPS pH 7.7). Gelation was performed as described above for untargeted ExSeq of brain slices. A modified Proteinase-K digestion protocol was used, consisting of 1 hour of digestion at 60°C, followed by 24 hours of digestion at 37°C, until the gel came off the glass slide. The digestion buffer was modified to include guanidine hydrochloride; final composition 50 mM Tris pH 8.0, 1 mM EDTA, 0.5% Triton X-100, 0.8 M guanidine hydrochloride, 8 U/mL Proteinase-K (NEB). Re-embedding and passivation were performed as described above for targeted ExSeq of the mouse visual cortex and hippocampus.

Probes were pooled together to form 200 μM libraries for the MBC probeset. The gelled and passivated samples were pre-incubated for 30 minutes at room temperature with wash buffer (20% formamide, 2X SSC), then incubated overnight at 37°C with the hybridization mix, consisting of 140 μM pooled library probeset (total of 4353 probes; ~32 nM each probe) in 20% formamide, 2X SSC buffer. The samples were then washed three times with wash buffer at 37°C for 30 minutes each wash, followed by one wash with 1X PBS for 30 minutes at 37°C. The samples were then pre-incubated with 1X SplintR ligase buffer for 30 minutes at room temperature. The samples were then incubated with 1250 U/mL SplintR ligase in 1X SplintR ligase buffer at 4°C for 3 hours. The samples were then incubated with a freshly prepared solution of 1250 U/mL SplintR ligase (NEB) in 1X SplintR ligase buffer overnight at 37°C. Following ligation, the samples were washed with 2X SSC for 30 minutes at room temperature followed by a pre-hybridization with wash buffer for 15 minutes at room temperature. The samples were then incubated with 500 nM rolling circle amplification primer (TCT TCA GCG TTC CCG A*G*A, where * denotes phosphorothioate backbone modification) in wash buffer for 2 hours at 37°C, followed by a 30 minute wash with wash buffer at 37°C, and another wash with 1X PBS for 15 minutes at 37°C. After a pre-incubation with 1X Phi29 buffer for 15 minutes at room temperature, the samples were incubated with 1000U/mL Phi29 polymerase (Enzymatics) in 1X Phi29 buffer for 3 hours at 4°C. Following this step, the samples were incubated with 1000 U/mL Phi29 polymerase, 250 μM dNTP, 40 μM aminoallyl dUTP in 1X Phi29 buffer at 30 °C overnight. The next day, the samples were washed once with 1X PBS at room temperature for 30 minutes, and then treated with 5 mM BS(PEG)9 in 1X PBS for 2 hours at room temperature. The samples were then again washed with 1X PBS for 15 minutes at room temperature followed by another wash with 1 M Tris, pH 8 for 15 minutes at room temperature. After this step, the samples were then washed once with 1X PBS.

Targeted ExSeq *in situ* SOLiD sequencing

To enable stable, multi-round imaging, gelled samples were re-embedded in individual wells of a Bind-silane treated 6-well glass-bottom plate as described above for targeted ExSeq of mouse visual cortex and hippocampus. Because of the larger sample size, bind-silane was diluted 1:250 for the treatment to ensure sample adherence.

For SOLiD sequencing by ligation chemistry, for each one of the 5 sequencing primers, the specimen was first stripped for 30 min to remove the hybridization probe or the previous sequencing primer with strip solution (80% formamide and 0.01% Triton-X in water). Next, the specimen was washed with 1X instrument buffer (SOLiD Buffer F, 1:10 diluted) for 10 min, and incubated for 20 min with 2.5 μM of sequencing primer in 5X SASC (0.75 M sodium acetate, 75 mM tri-sodium citrate, pH 7.5). After two washes for 5 min total with 1X instrument buffer, the specimen was reacted for 1 hour with T4 DNA ligation mixture (6 U/μl T4 DNA ligase (Enzymatics) and 1:40 diluted SOLiD sequencing oligos in 1X T4 DNA ligase buffer). The specimen was then washed 4 times for 3 hours overall with 1X instrument buffer and imaged with SOLiD imaging buffer (the extended wash was needed since we didn't use the flowcell but rather did manual sequencing).

After each ligation, samples were ready for imaging. The samples were stained with DAPI (1 mg/L) in 1X instrument buffer, and washed twice with an imaging buffer. The samples were then imaged on an Andor Dragonfly Spinning Disk confocal with laser lines 100 mW solid state 405 laser, 150 mW solid state 488 laser 150 mW solid state 561 laser, 160 mW solid state 633 laser; emission filters, 525/50, 582/15, 624/40, 685/40; Andor Zyla sCMOS 4.2 plus with 200

msec exposure time; and a Nikon 40X CFI Apo, water immersion with long working distance, NA 1.15 objective. Tiled images of the sample were acquired using the Andor Fusion software. For each tile, 151 Z-sections were imaged at a spacing of 0.4 μm.

Following each imaging round, a dephosphorylation reaction was performed before each cleave reaction to reduce phasing. The reaction was for 30 min with 1:20 dilution of Quick CIP (NEB, M0508L) in 1X CutSmart buffer. To acquire the next base, a two-step cleave reaction was performed, first with SOLiD buffer C twice for 30 min total (part #4458932) and then SOLiD buffer B twice for 15 min total (part #4463021). The cleave reaction was followed by three washes for 20 min total with 1X instrument buffer, 1 hour with T4 DNA ligation mixture, 4 washes for 3 hours total with 1X instrument buffer, and finally washing with SOLiD imaging buffer (SOLiD buffer A, part #4463024) before imaging. The cleave-ligation-wash-imaging cycle was done once for each one of the 5 sequencing primers. All reactions were done at room temperature. During the *in situ* sequencing reactions the strip solution and the sequencing primers were kept at 80°C while the ligation mixture, the imaging buffer, and SOLiD buffer B were kept at 4°C.

Image Processing - Color correction, registration, segmentation, basecalling and alignment

Color correction, registration, segmentation basecalling were done as described for the targeted hippocampus and visual cortex experiments with two modifications. Background subtraction, as described in the untargeted hippocampus **Methods** sections, was performed. The second modification is that alignment allowed up to one mismatch in the 7-base barcodes, due to the barcode design (described in the **Methods** section 'Gene selection and probe design' above) which allows correction of one substitution error.

Image Processing - Cell segmentation (**Fig. 6**)

Because of the high density of cells of the tumor biopsy, the automated cell segmentation was unreliable. We utilized VAST (*118*) to implement a 2D manual segmentation on a stitched image of maximum intensity projections of the DAPI stain for the 78 fields of view. A single person was able to annotate the cells in one day by utilizing the "conditional" feature of VAST, allowing only pixels of a trained intensity regime to be annotated, which accelerates the speed of annotation. Additionally, because MATLAB's 'bwlabel' function can ascribe unique values to all non-touching binary objects, the speed of manual annotation was increased by only using a single annotation color in VAST, and therefore the manual task was mainly to remove touching annotations of neighboring cells. With this pipeline, we segmented 5,862 cells containing at least one ExSeq read, and retained 2,395 cells with >100 ExSeq reads (thresholded after inspecting the data), containing a total of 771,904 reads. For an ExSeq read to be counted as within a cell, given the sample is 8 microns thick (comparable to the cell size), we demanded that the XY portion of the 3D centroid of the ExSeq read would be within the segmented nucleus of a cell.

Nuclear structures (possibly nucleoplasmic bridges) (**Fig. S25**) were manually identified in 2D using VAST and then extended into 3D using the DAPI stain, creating a 3D mask of the nuclear structures. For an ExSeq read to be considered inside the nuclear structure, we enforced that at least 10% of the puncta voxels overlapped with the 3D mask. We note that the relatively small number of reads (516) obtained in these nanoscale regions limited our ability to systematically classify these reads into specific cell types.

Data Analysis - Expression clustering (**Fig. 6B** and **6C**)

In order to identify and cluster the 2,395 cells according to their expression pattern, we utilized the R toolkit Seurat (*77*, *125*). Following (*126*), we used a supervised approach of using selected genes for dimension reduction (instead of choosing the genes with the highest variability). These genes were: non-tumor marker genes (*126*)- CD3G, CD68, FOXP3, CD4 ,CD8A, CD3D, CD3E, HLA-DRA; tumor marker genes (*126*)- EGFR, GRB7, ERBB2, PGR, CD44, CD24, ALDH1A3, EPCAM, KRT19, KRT18, CDH1; B-cells- IGHG1, IGHG4, IGKC, IGHM; Fibroblast- HSPG2. The rest of the analysis was according to the Seurat analysis pipeline, i.e. creating a K-nearest neighbor graph based on the euclidean distance in PCA space, and then finding clusters in the PCA space using the Louvain algorithm. Genes that can 'mark' each cluster (i.e. with an expression level which is higher in a given cluster compared to the other clusters) were discovered using the Seurat 'FindAllMarkers' function. The 'FindAllMarkers' function reports a p-value for each putative gene marker in each cluster; for putative gene markers with p-values less than 1E-10, we assigned the cluster with the known annotation of the marker gene, otherwise we marked the cluster as "unknown".

Data Analysis - Determining adjacent cell clusters (**Fig. 6D**)

For all 2,395 cells classified by Seurat, cell centroids were calculated as the average position of the reads inside that cell. An adjacency graph was then calculated by counting all instances of two cell centroids, from two different cell clusters, being within 20 microns of each other. Cluster labels were shuffled and the adjacency graph of cell clusters was recalculated 500,000 times. From this bootstrapped data, p-values were calculated for each pair of cell clusters as (1+x)/500,000 where x=number of bootstrapped iterations in which the randomized adjacency graph for a given pair exceeded that of the original adjacency graph. The distance of 20 microns was empirically chosen, and 10 microns and 40 microns (39 microns was used in (*126*)) were also computed at 10,000 iterations each, producing a similar adjacency graph structure (**Fig. S26**).

Data Analysis - Detecting upregulated genes (**Fig. 6E**)

Using the cell clusters and adjacency matrix as described above, for any pair of cell clusters A and B, cluster A was partitioned into two subsets: a subset of A cells that are adjacent to B cells, and a subset of A cells that are not adjacent to B cells. Gene expression change (fold change) per gene was then calculated as the ratio of the median expression in the subset of A near B and the median expression in the subset of A not near B, ignoring any genes with median value of zero. To calculate statistical significance, the A cells were randomly partitioned into two subsets, of the same size as the two original subsets, and the fold changes of all genes between the two subsets were recalculated 100,000 times as described above. From this bootstrapped data, p-values were calculated for each gene as (1+x)/100,000 where x=number of bootstrapped iterations in which the fold-change for any gene exceeded that of the original detected gene.

We note that a discontinuity is clear in the right image in panel **Fig. 6Ei** -- an horizontal line is evident in the top part of this image; this is a result of imperfect stitching between two fields of view. This discontinuity should not have any effect on the data presented or the analysis of the data.

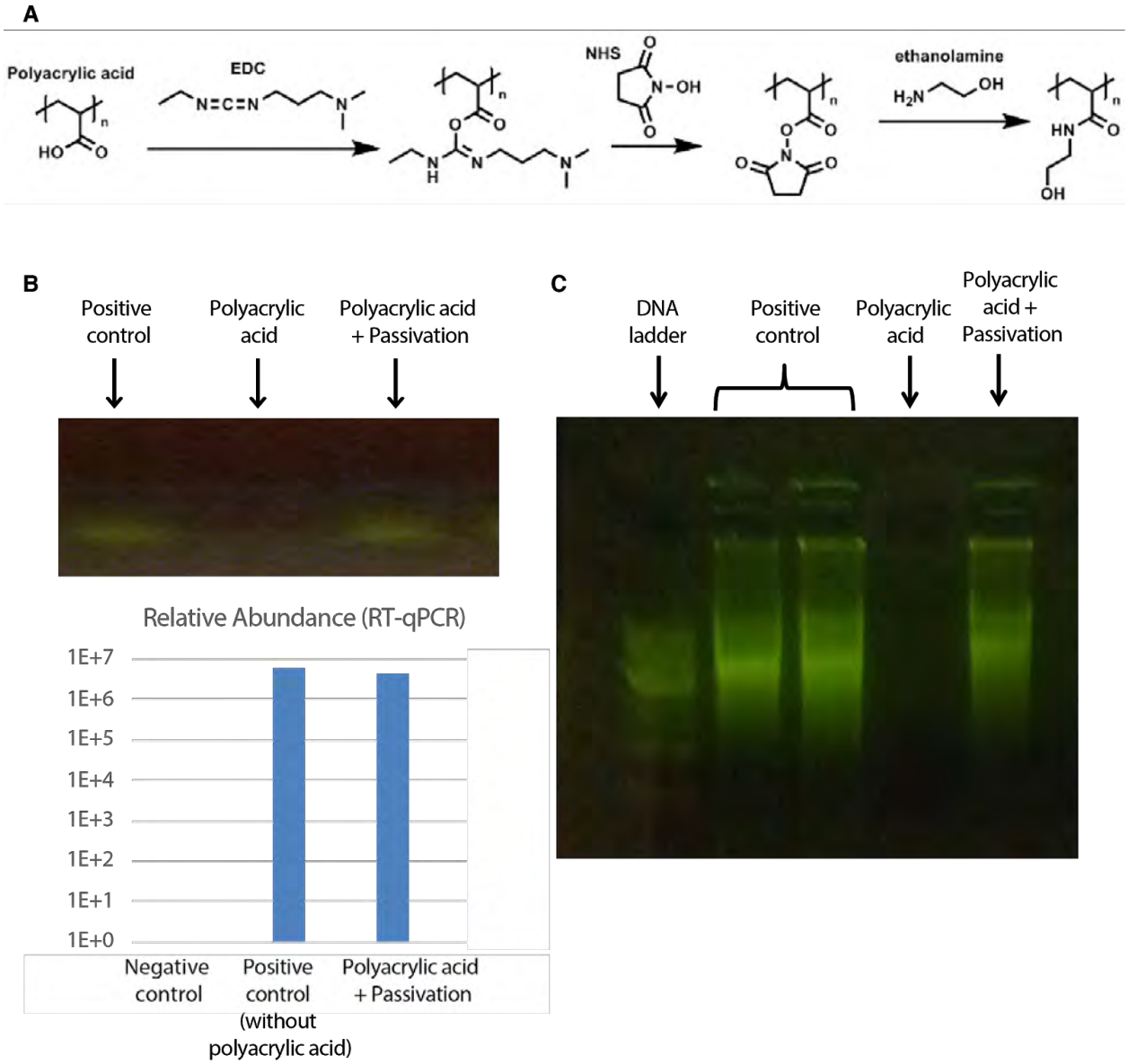


Fig. S1.

In vitro experiments demonstrating that polyacrylic acid, very similar to the polymer used in ExM, inhibits both the reverse transcription step and the rolling circle amplification step, two key enzymatic reactions of FISSEQ, but that the enzymatic activity can be restored with a passivation reaction that cancels out the charge of moieties on the polymer backbone. The passivation reaction (*99*), in which ethanolamine is reacted with carboxylic groups and converts them to amides with no charge (A), restores reverse transcription activity (B) and rolling circle amplification activity (C).

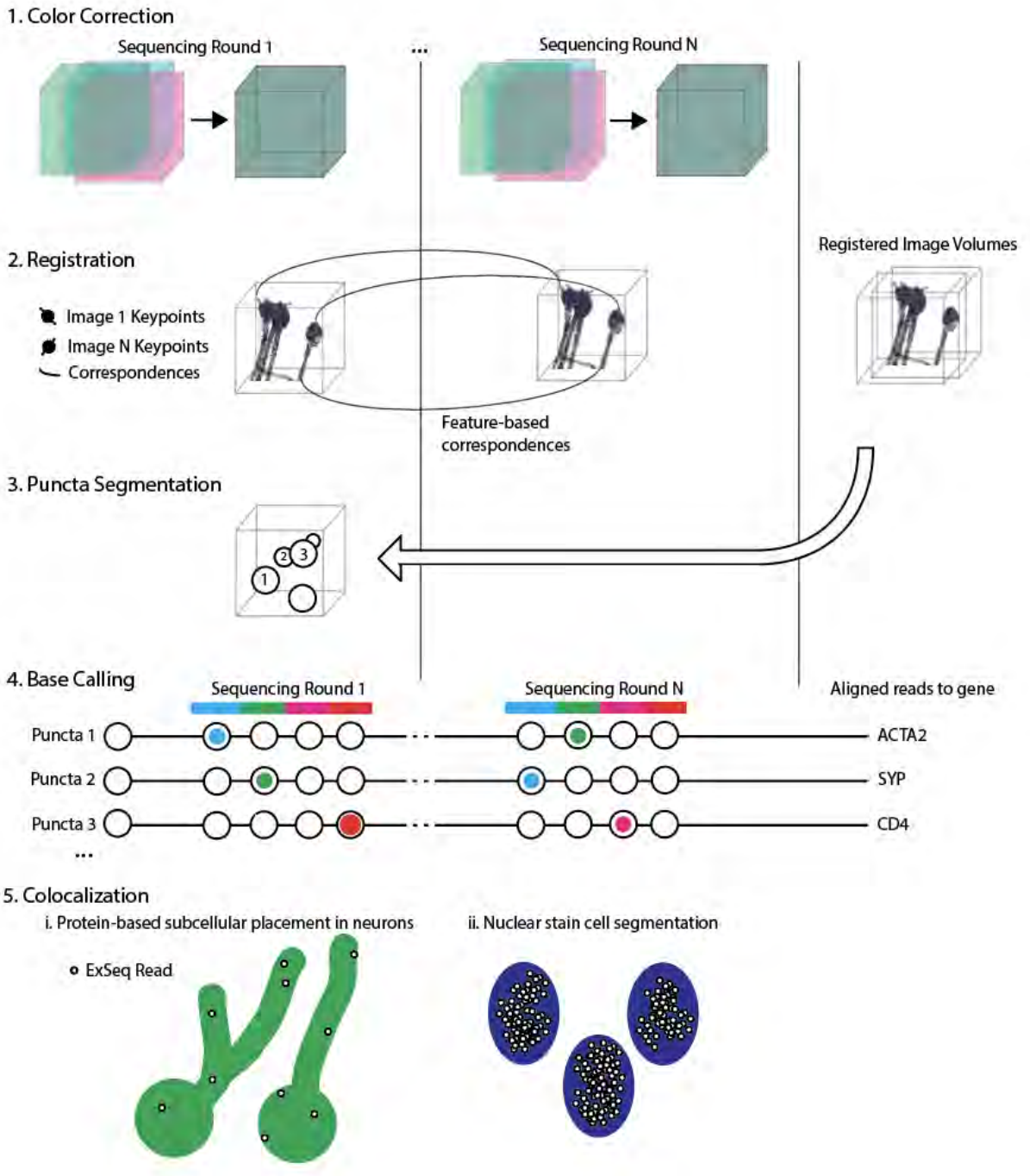


Fig. S2.
Schematic, for a single field of view, of the multi-step image processing pipeline that automatically processed 362 fields of view, across 6 ExSeq experiments (three untargeted, including 50 and 15 micron thick hippocampal slices, and one hippocampal culture, and three targeted, listed in **Table S5**), from volumetric images to spatially localized RNA reads. Each volumetric image for a given sequencing round is first rigidly color corrected to account for piezo stage drift and optical differences between color channels (1). Then (2), a precise

registration (**Fig. S3**) is performed by detecting salient keypoints, describing each keypoint using a 3D SIFT descriptor, and discovering correspondences between keypoints in different sequencing rounds using the SIFT matching algorithm followed by RANSAC. The correspondences are then used to calculate an affine transformation. All the registered images are then combined to detect the amplicons; this puncta segmentation task is performed using watersheding (3). Next, each amplicon is described as a sequence of colors across rounds, by identifying the dominant color channel per round (4). These sequences are then aligned to either an *ex situ* library (untargeted ExSeq) or known barcode library (targeted ExSeq) to convert the color sequence into a gene identity. (5) The reads are then studied in space with the registered morphology or nuclear stain. In (5), all analyses were done in 3D, except for the 8 micron thick tumor tissue (**Fig. 6**), in which the reads were further studied in 2D. The entire MATLAB library to process ExSeq datasets from the microscope to spatial analysis of gene expression is accessible at Zenodo (*29*). Specifically, the Zenodo repository includes a tutorial wiki with step-by-step instructions on how to run the pipeline and a tutorial set of targeted ExSeq data from the mouse visual cortex.

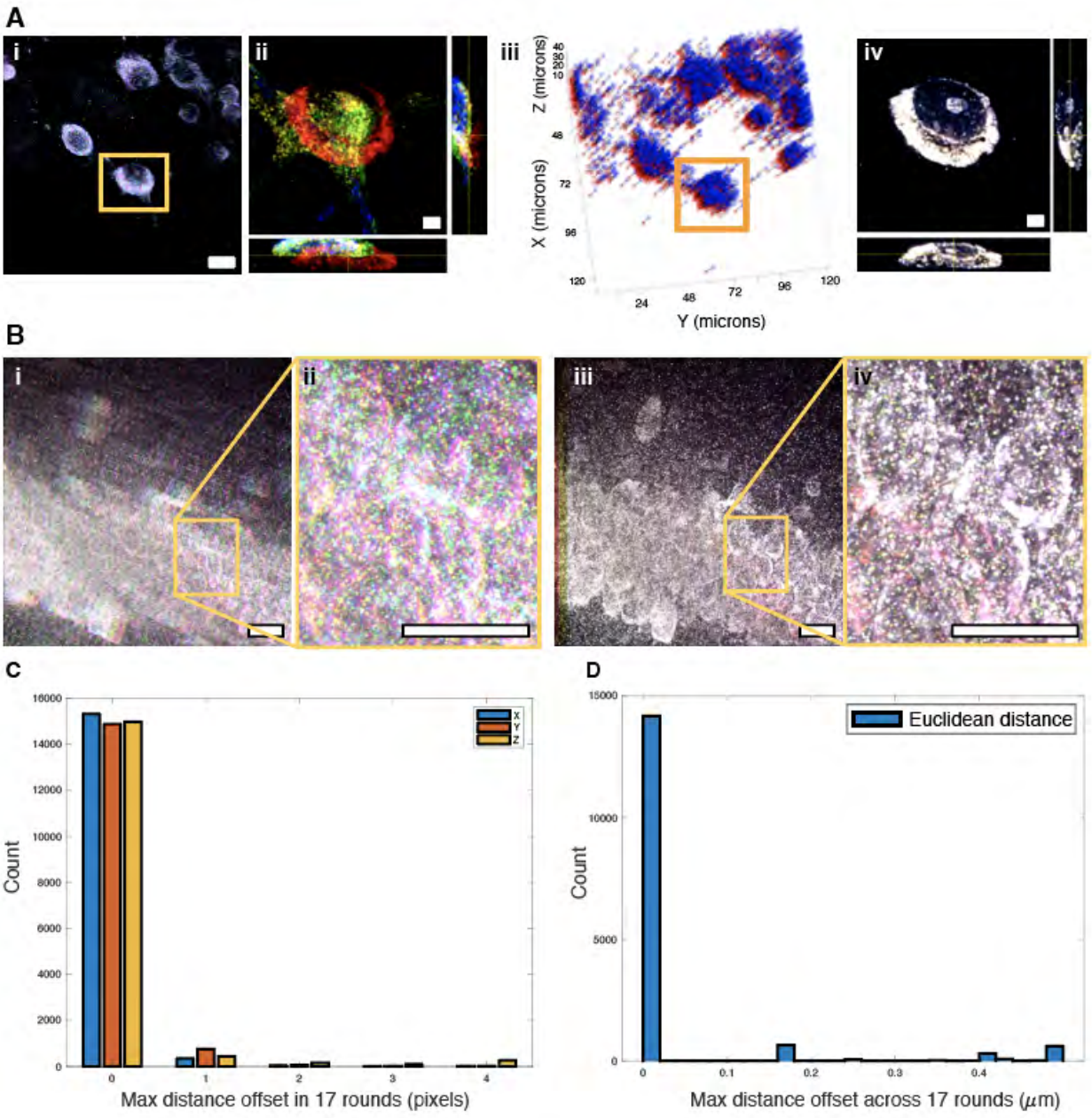


Fig. S3.
Image registration aligns puncta across 20 *in situ* sequencing rounds for untargeted ExSeq to within 1 pixel on average. Using a sparse neuron culture for demonstration purposes, a single round of sequencing is shown in (Ai), with the four SOLiD sequencing fluorescence channels shown (blue, green, red, magenta). Yellow box, region explored in depth in Aii. (Aii) Combined fluorescence channels of four sequencing rounds, for a set of cultured neurons, are shown as orthogonal views in four different colors: red, green, blue and yellow. Color indicates sequencing round, by summing all fluorescence channels in a given round, and white indicates overlap between all rounds. The image data is not registered. Features (3D SIFT descriptors, see **Methods**) are then calculated in each round and the correspondences to a reference round (the first sequencing round beyond the primer, chosen for its high image quality) are discovered via SIFT matching. (Aiii) 3,311 correspondence points between the reference round (shown in blue)

and one of the 19 non-reference rounds (red) were calculated, with the specific neuron shown in (Ai) and (Aii) highlighted in the orange square. Axis labels are pre-expansion distances. (Aiv) After the affine warp is applied, the four sequencing rounds are in high agreement: colors are as in (Aii). (Bi) An example field of view from the intact mouse hippocampus showing six rounds (red, blue, green, cyan, magenta, yellow; all bases are given the same color so that the reader can focus on alignment) without registration, with zoomed in region showing approximately three cell bodies (Bii). White indicates overlap between rounds. (Biii-iv) The same regions as in Bi-ii, after registration. (C-D) To quantify the quality of registration, we calculated a normalized cross correlation of 15,943 subvolumes (each of size 41x41x19 pixels), randomly chosen across the imaged field of view (350x350x100 microns in size, post-expansion). (C) shows the maximum offsets, after registration, across the sequencing rounds for each subvolume in each dimension. The average of these maximum offsets per subvolume was 0.19 ± 1.52 pixels in X, 0.24 ± 1.62 pixels in Y and 0.16 ± 0.81 pixels in Z (mean ± standard deviation)). (D) Histograms of the maximum offsets shown in euclidean distance, with an average distance of 0.10 ± 0.5 microns (SD). Scale bars: (Ai,Bi-Biv) 13μm pre-expansion, (Aii, Aiv) 3μm pre-expansion.

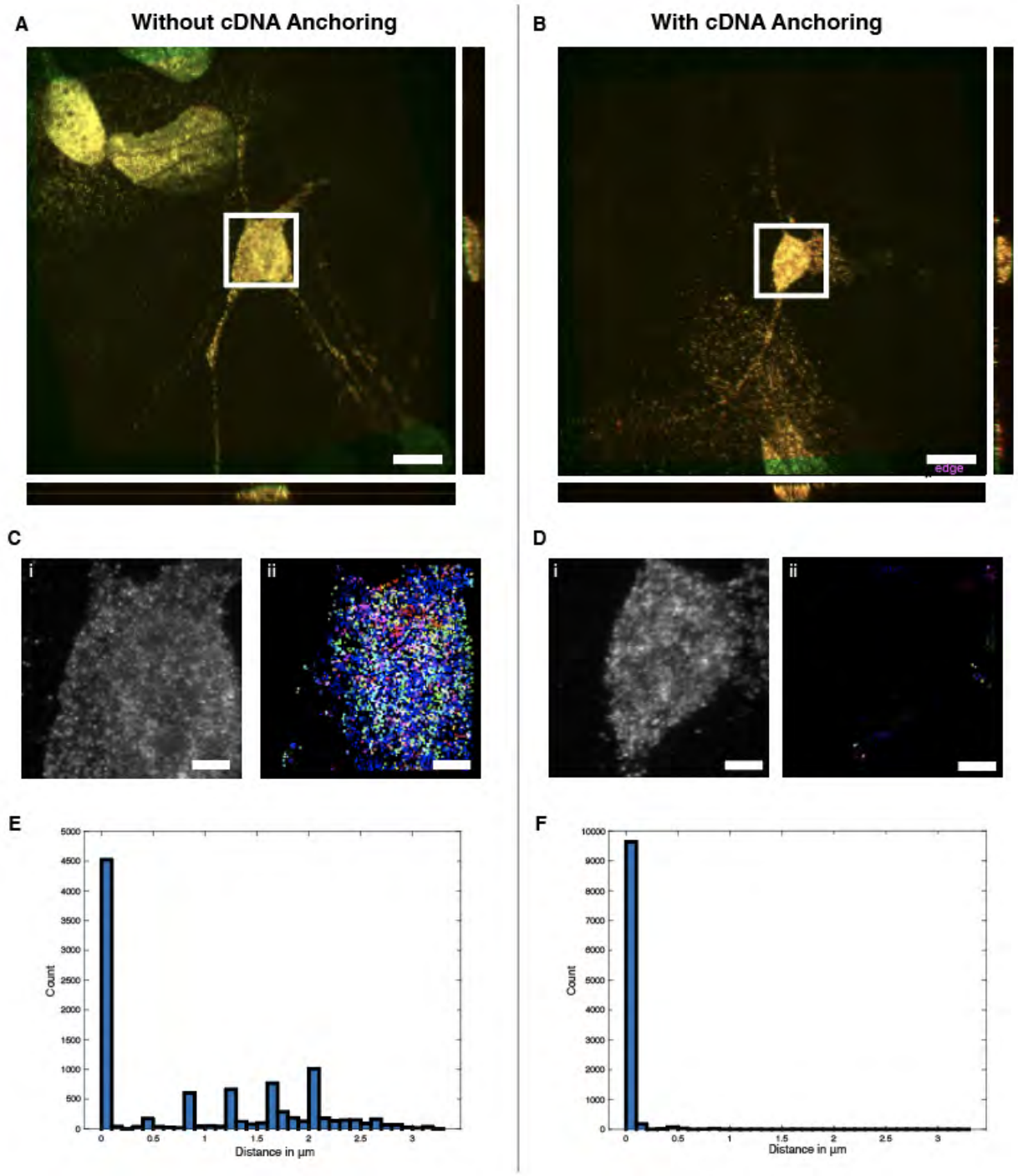


Fig. S4.

Anchoring of cDNA ensures precise spatial capture of cDNA location. In principle, without cDNA anchoring, the cDNA can move after the RNase digestion, as they are no longer linked to the gel in any way. To quantify this effect, we measured the location of the cDNA before the RNase digestion, and then 12 hours after RNase digestion, to create an extreme scenario in which the cDNA could, in principle, move quite a bit. The experiment was done both with and

without cDNA anchoring, and performed in cultured hippocampal neurons. (A) BrdU antibody (see **Methods** section 'BrdU staining protocol for cDNA visualization (**Fig. S4**)') against cDNA before the RNAse step (green) and 12 hours after the RNase step (red), registered together (yellow overlap). (B) Same as (A), but the cDNA are anchored to the gel before the RNase step. Focusing just on a single soma, (C) and (D), a normalized cross-correlation method using 21x21x13pixel subvolumes from random locations within the volume shown was used to calculate cDNA drifts between the two imaging times. The locations of the 10,000 21x21x13pixel subvolumes were randomly chosen and required to have a minimum signal intensity to ensure sufficient signal-to-noise ratio for the autocorrelation calculation to calculate a correct offset. As shown in Cii and Dii, the direction and magnitude of the offsets are encoded in an RGB image. For each subvolume, the offset of the peak of the normalized cross-correlation is displayed in red for the X-direction offset (black=zero offset, bright red = offset of 10 pixels or 1.7μm), green for Y-direction offset (black=zero offset, bright green = offset of 10 pixels or 1.7μm) and blue for Z-direction offset (black=zero offset, bright blue = offset of 5 pixels or 2.0μm). (E) Only 52% of the subvolumes for the without-cDNA-anchoring condition are within 0.5μm, whereas 99% of the subvolumes match that criteria when cDNA anchoring was used (F). Scale bar: (A, B) 13μm pre-expansion, (C, D) 3μm pre-expansion.

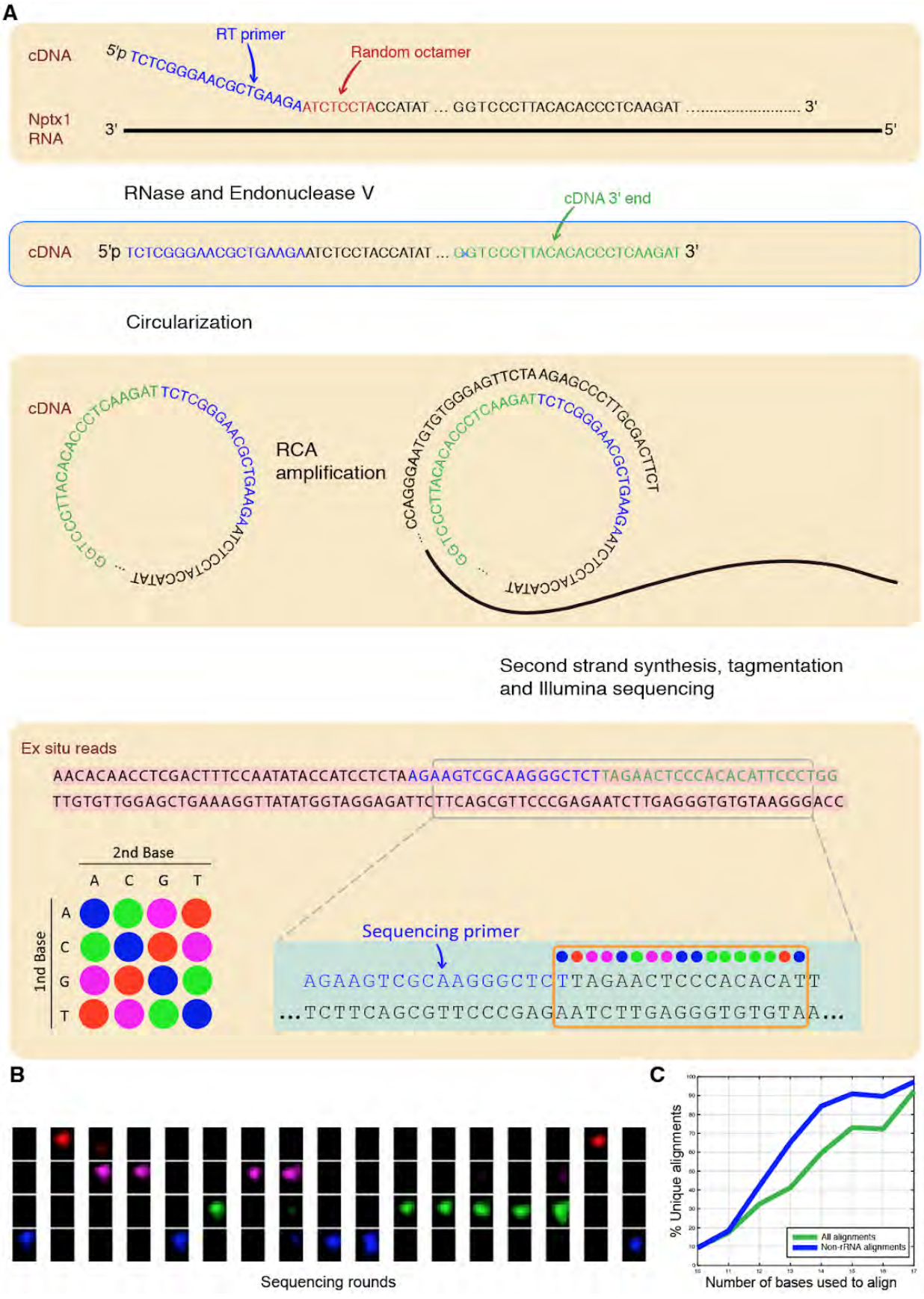

A
cDNA
5'p TCTCGGGAACGCTGAAGAATCTCCTACCATAT … GGTCCCTTACACACCCTCAAGAT ..... 3'
RT primer
Random octamer
Nptx1 RNA
3'
5'
RNase and Endonuclease V
cDNA 3' end
cDNA 5'p TCTCGGGAACGCTGAAGAATCTCCTACCATAT … GGTCCCTTACACACCCTCAAGAT 3'
Circularization
cDNA
RCA amplification
Second strand synthesis, tagmentation and Illumina sequencing
Ex situ reads
AACACAACCTCGACTTTCCAATATACCATCCTCTAAGAAGTCGCAAGGGCTCTTAGAACTCCCACACATTCCCTGG
TTGTGTTGGAGCTGAAAGGTTATATGGTAGGAGATTCTTCAGCGTTCCCGAGAATCTTGAGGGTGTGTAAGGGACC
2nd Base
A C G T
1nd Base
Sequencing primer
AGAAGTCGCAAGGGCTCTTAGAACTCCCACACATT
…TCTTCAGCGTTCCCGAGAATCTTGAGGGTGTGTAA…
B
Sequencing rounds
C
% Unique alignments
All alignments
Non-rRNA alignments
Number of bases used to align

Fig. S5.

Ex situ library construction and *in situ* matching. (A) The random nature of untargeted reverse transcription priming, the cDNA size restriction, and the circularization of the cDNA, create unique molecular identifiers in the *in situ* sequenced region of the amplified cDNA, thus enabling later matching to *ex situ* reads. One sequenced region of the gene Neuronal Pentraxin 1 (*Nptx1*) is presented as an example. First, the cDNA is generated using random priming; the exact location of priming was reconstructed from the actual *ex situ* reads presented below. Inosine is included in the reverse transcription process, allowing generation of ~100 base long cDNA fragments following Endonuclease V treatment. Next, RNase treatment and Endonuclease V cutting (at inosine) create cDNA fragments which are 95 bases long in this specific example, not including the constant region of the reverse transcription (RT) primer (the exact location of the cDNA 3' end, reconstructed from the actual *ex situ* reads, is shown). cDNA circularization allows rolling circle amplification (RCA) with a RCA primer which is complementary to the constant region of the RT primer. Note that the cDNA anchoring to the hydrogel is not illustrated, for simplicity. After *in situ* sequencing, the amplified cDNA is extracted from the gel, the ss-cDNA is converted to ds-cDNA, and the ds-cDNA is tagmented and sequenced *ex situ*. The actual resulting *ex situ* reads are presented — note that in this case the paired-end reads give identical reverse complement sequences. To allow *in situ* and *ex situ* sequencing reads to be matched, we reconstruct the *in situ* read from the *ex situ* sequence; first we identify the sequencing primer sequence in the *ex situ* read, and then we convert the bases at the 5' end of the sequencing primer to SOLiD color space according to the SOLiD 2 base encoding diagram (lower left). The sequences of colors resulting from the *ex situ* information ('*ex situ* library') are then used as a dictionary to align and directly match the *in situ* sequencing reads. (B) The sequences of colors resulting from the *ex situ* information (panel A) has one perfect match to the *in situ* sequencing of the gene *Nptx1*, sequenced in a dentate gyrus neuron in a 50 micron mouse hippocampus slice and shown on the left side of **Fig. 1Biii**. (C) The matching of the *in situ* reads to the *ex situ* library is detailed in the **Methods** section 'Data Analysis - *Ex situ* and *in situ* sequence matching'. Overall, 92% of the matches, and 97% of the non-rRNA matches, are strictly unique in the sense that if a matching *ex situ* read is removed from *ex situ* library, the *in situ* read does not match to another *ex situ* read. Importantly, all *in situ* reads that are not strictly unique are removed (see **Methods** section 'Data Analysis - *Ex situ* and *in situ* sequence matching'). This allows to explore sequence variations in mRNA, such as alternative splicing, using the longer *ex situ* matched reads.

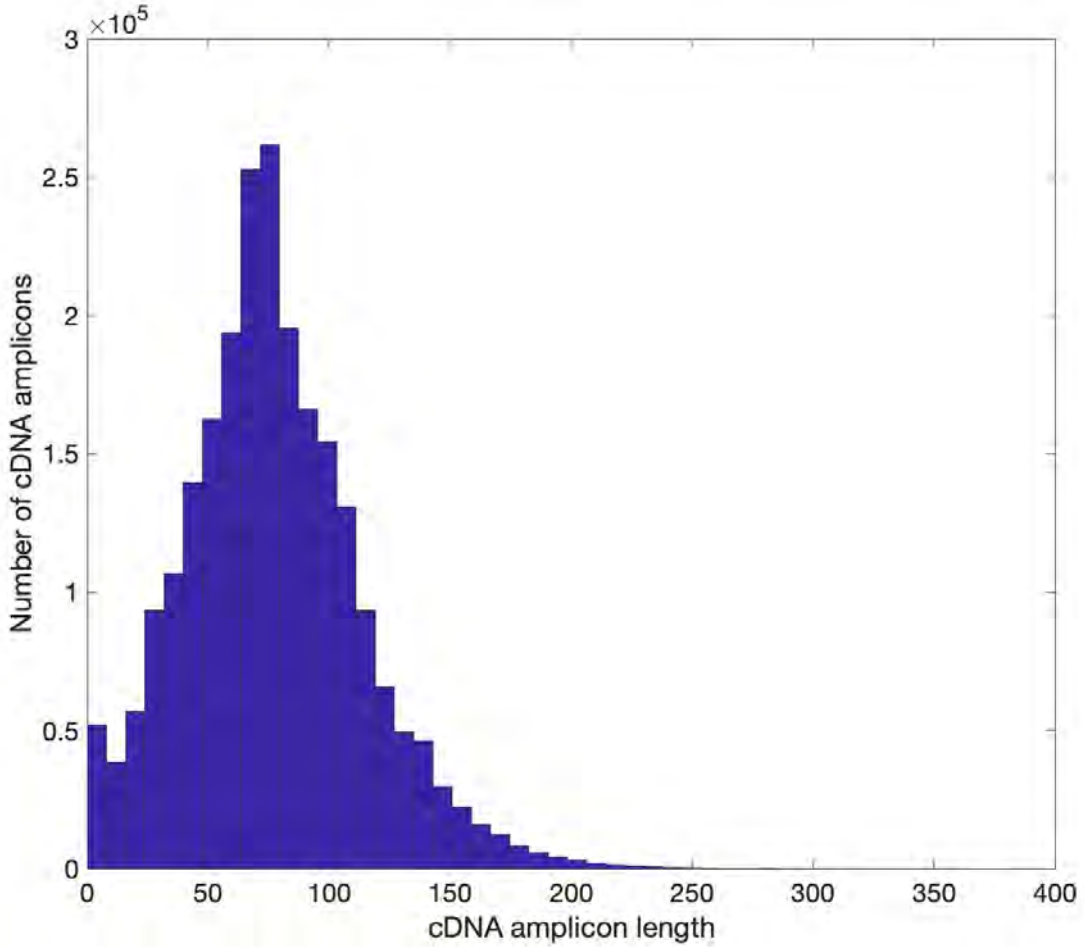


Fig. S6.
Histogram of cDNA amplicon lengths. The average cDNA amplicon length was 76.6 bases, with standard deviation of 36.6 bases, and maximal length of 397 bases.

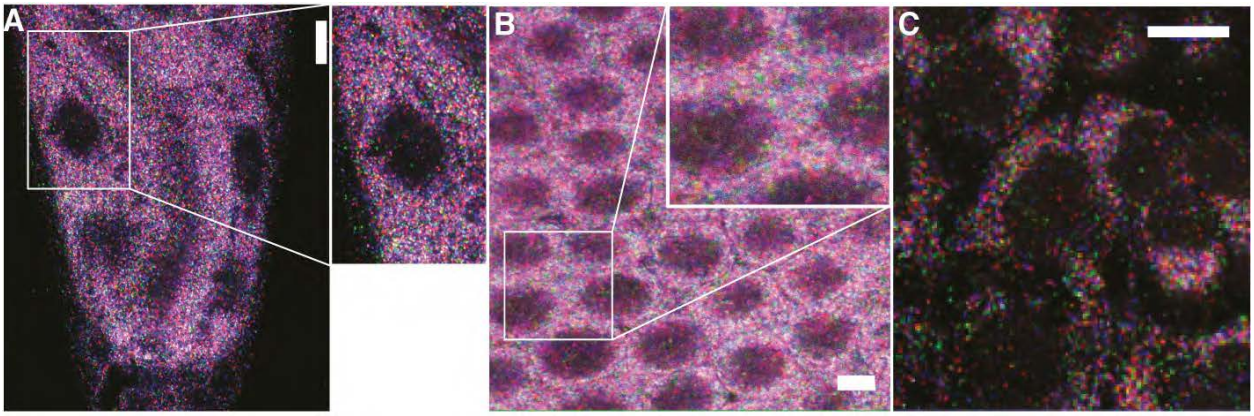


Fig. S7.

Demonstration of untargeted ExSeq with C. elegans (A), a Drosophila embryo (B) and the HeLa human cell line (C). The first *in situ* sequencing round is shown and the different colors (blue, magenta, green, and red) reveal the current base of the amplified cDNA (SOLiD sequencing was used). Scale bars: 20, 10 and 30 microns for panels A, B and C, respectively, in post-expansion (e.g., actual size) units.

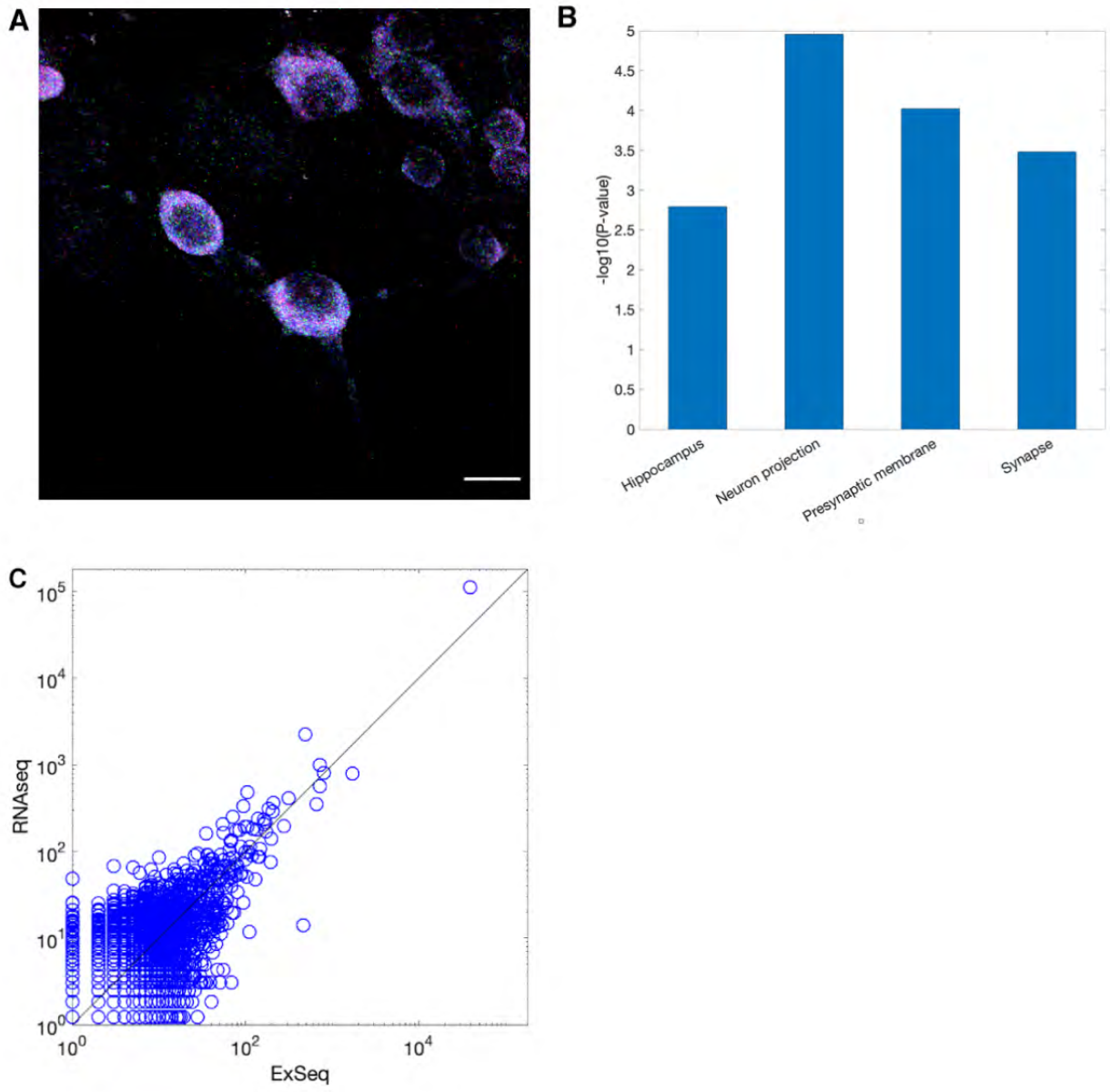


Fig. S8.

Untargeted ExSeq of hippocampal neurons in culture. (A) Maximum intensity projection of the first in situ sequencing round in a hippocampal culture. The different colors (blue, magenta, green, and red) reveal the current base of the amplified cDNA, in a SOLiD sequencing step. (B) Gene ontology analysis using the software DAVID (*114*) on all 127 expressed genes detected with ExSeq using 10 acquired FoVs (each of the size shown in A), in a hippocampal culture. Ontologies with low p-values are presented, revealing the expected functional enrichments for hippocampal neurons. See **Methods** section 'Gene ontology analysis of ExSeq data from hippocampal culture (**Fig. S8**)' for the list of enriched GO terms as obtained by DAVID. (C)

Agreement between the normalized expression levels of all well-annotated genes (RefSeq genes) using RNAseq, and ExSeq with full *ex situ* sequencing data, from the same hippocampal culture. The Pearson's correlation between the log-transformed expression of RefSeq genes using ExSeq and using RNAseq is 0.621. Scale bar: (A) 13μm, pre-expansion.

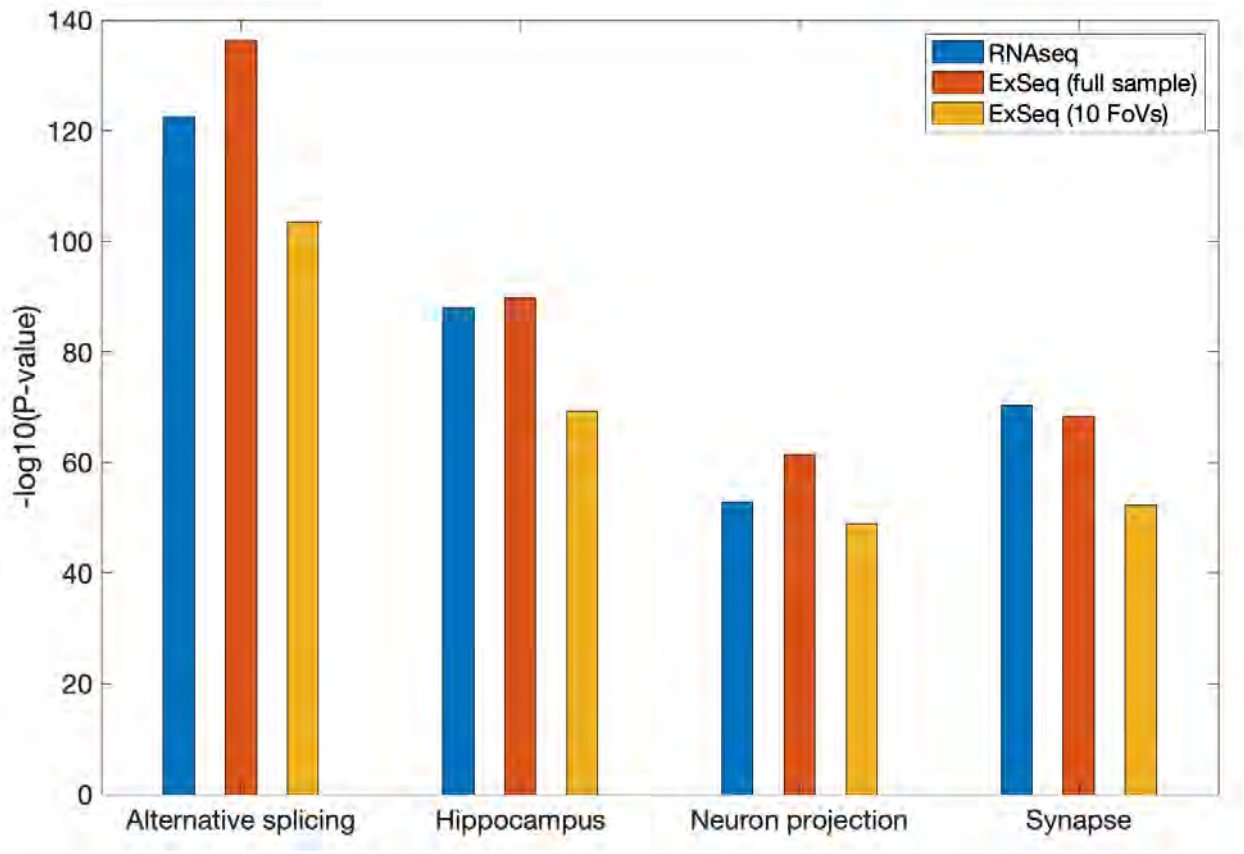


Fig. S9.

Gene ontology analysis using the software package DAVID (*114*) on the highly expressed genes detected with RNAseq, untargeted ExSeq with full *ex situ* sequencing data, and untargeted ExSeq using the *ex situ* data corresponding to the 10 acquired FoVs, in a 50 micron thick hippocampus slice. 3,039 genes were detected via ExSeq using the 10 acquired FoVs, and therefore the top 3,039 genes (sorted according to expression level) were analyzed in each one of the three analysis styles shown. Ontologies with low p-values are presented, revealing the expected functional enrichments for the hippocampus region. Importantly, the functionally enriched groups are common to the three tested conditions. The full list of the first 25 functional enrichment groups (i.e. the functions with the lowest p-values) as obtained by DAVID for each dataset is given in **Table S6**.

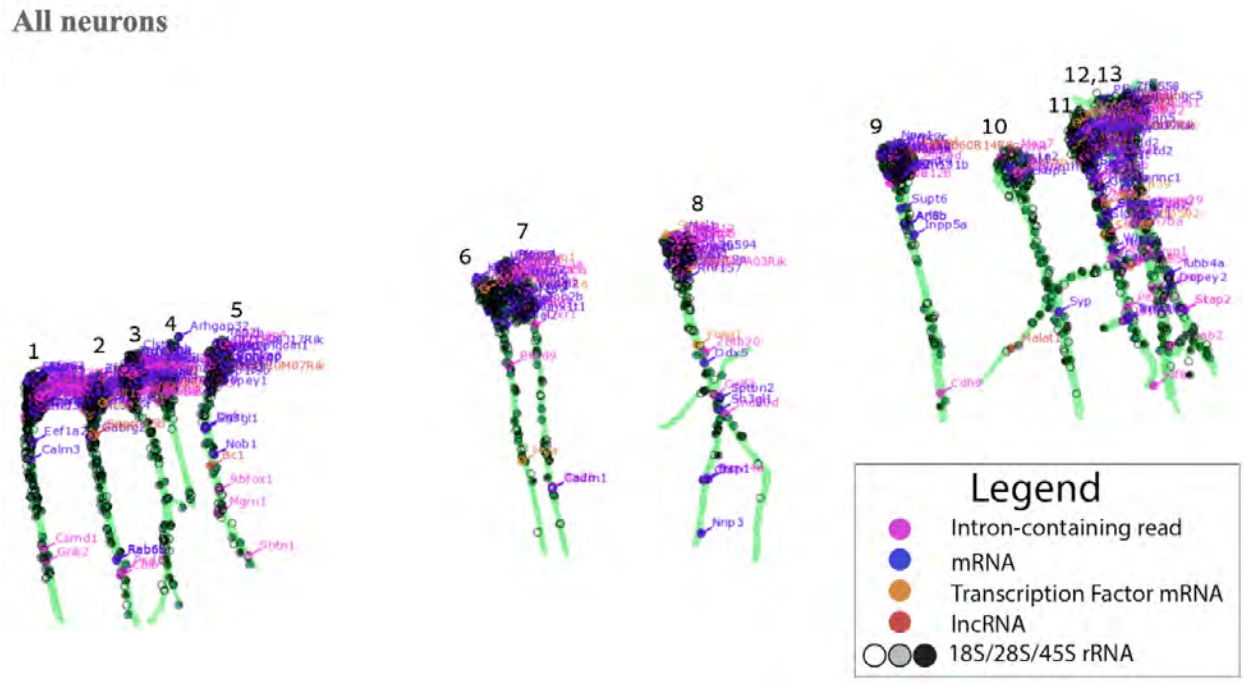


Fig. S10.

Complete list of the transcripts detected in each hand-traced neuron, not including rRNA. Each neuron is presented separately, numbered from left to right. For each neuron, the transcripts are listed according to their position from top (soma) to bottom (end of dendrites). Each ExSeq read is assigned ('Cell Loc' below) into the dendrite, soma, or nucleus (see **Methods** section 'Image Processing - 3D Tracing'), and an Euclidean distance ('Cell Dist' below) is measured from the centroid of the nucleus (in microns, pre-expansion).

Neuron 1: (392 RNA reads, 59 mRNA reads, 55 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
exon	Kcnq2	potassium voltage-gated channel subfamily Q member 2 (Kcnq2)	nucleus	0.9
mrna-tf	Mapk1	mitogen-activated protein kinase 1 (Mapk1)	nucleus	1.4
intron	Camta1	calmodulin binding transcription activator 1 (Camta1)	nucleus	1.8
exon	Scamp5	secretory carrier membrane protein 5 (Scamp5)	nucleus	2
intron	Grin2b	glutamate receptor ionotropic NMDA2B (epsilon 2) (Grin2b)	nucleus	2
mrna-tf	Hivep2	human immunodeficiency virus type I enhancer binding protein 2 (Hivep2)	nucleus	2.2
exon	Spin1	spindlin 1 (Spin1)	nucleus	2.4
intron	Cenpp	centromere protein P (Cenpp)	nucleus	2.4
intron	Wdr55	WD repeat domain 55 (Wdr55)	nucleus	2.7
exon	Eml6	echinoderm microtubule associated protein like 6 (Eml6)	nucleus	3.1
exon	Kcnq2	potassium voltage-gated channel subfamily Q member 2 (Kcnq2)	nucleus	3.1
exon	Acox1	acyl-Coenzyme A oxidase 1 palmitoyl (Acox1)	nucleus	3.3
exon	Taok1	TAO kinase 1 (Taok1)	nucleus	3.5
exon	Plxnd1	plexin D1 (Plxnd1)	nucleus	3.5
exon	Glg1	golgi apparatus protein 1 (Glg1)	nucleus	3.7
exon	Sh3gl2	SH3-domain GRB2-like 2 (Sh3gl2)	nucleus	3.9
exon	Hnrnpul2	heterogeneous nuclear ribonucleoprotein U-like 2 (Hnrnpul2)	nucleus	4
intron	Josd2	Josephin domain containing 2 (Josd2)	soma	4.2
mrna-tf	Carm1	coactivator-associated arginine methyltransferase 1 (Carm1)	soma	4.3
exon	Tro	trophinin (Tro)	soma	4.5
exon	Ncald	neurocalcin delta (Ncald)	nucleus	4.6
intron	Dazap1	DAZ associated protein 1 (Dazap1)	soma	4.6
intron	Rbfox1	RNA binding protein fox-1 homolog (C. elegans) 1 (Rbfox1)	soma	4.6
intron	Rbfox1	RNA binding protein fox-1 homolog (C. elegans) 1 (Rbfox1)	soma	4.6
exon	Rps16	ribosomal protein S16 (Rps16)	soma	5
ncRNA	Snrnp70	small nuclear ribonucleoprotein 70 (U1) (Snrnp70)	soma	5
intron	Glis3	GLIS family zinc finger 3 (Glis3)	soma	5.1
intron	Dock2	dedicator of cyto-kinesis 2 (Dock2)	soma	5.3
intron	Kcnh7	potassium voltage-gated channel subfamily H (eag-related) member 7 (Kcnh7)	soma	5.4
exon	Map1b	microtubule-associated protein 1B (Map1b)	soma	5.4
intron	Arl3	ADP-ribosylation factor-like 3 (Arl3)	soma	5.8
mrna-tf	Hivep3	human immunodeficiency virus type I enhancer binding protein 3 (Hivep3)	soma	5.9
exon	Msi2	musashi RNA-binding protein 2 (Msi2)	soma	5.9
exon	Sparcl1	SPARC-like 1 (Sparcl1)	soma	5.9
exon	Socs7	suppressor of cytokine signaling 7 (Socs7)	soma	6
exon	Enc1	ectodermal-neural cortex 1 (Enc1)	soma	6.1
intron	Nell1	NEL-like 1 (Nell1)	soma	6.1
intron	Kcnip4	Kv channel interacting protein 4 (Kcnip4)	soma	6.3
exon	Tmx3	thioredoxin-related transmembrane protein 3 (Tmx3)	soma	6.3
intron	Lrpprc	leucine-rich PPR-motif containing (Lrpprc)	soma	6.3
exon	Chmp1b	charged multivesicular body protein 1B (Chmp1b)	nucleus	6.4
exon	Abr	active BCR-related gene (Abr)	soma	6.4
exon	Gpd1	glycerol-3-phosphate dehydrogenase 1 (soluble) (Gpd1)	soma	6.4
exon	Fbxo22	F-box protein 22 (Fbxo22)	soma	6.7
exon	Eml6	echinoderm microtubule associated protein like 6 (Eml6)	soma	6.8
intron	Grik4	glutamate receptor ionotropic kainate 4 (Grik4)	soma	6.9
intron	Camk2a	calcium/calmodulin-dependent protein kinase II alpha (Camk2a)	soma	6.9
exon	Sparcl1	SPARC-like 1 (Sparcl1)	soma	7.2
intron	Mast4	microtubule associated serine/threonine kinase family member 4 (Mast4)	soma	7.3
exon	Heatr3	HEAT repeat containing 3 (Heatr3)	soma	7.4
intron	Dlgap2	DLG associated protein 2 (Dlgap2)	soma	7.4
exon	Slc1a2	solute carrier family 1 (glial high affinity glutamate transporter) member 2 (Slc1a2)	soma	7.4
exon	Lmo3	LIM domain only 3 (Lmo3)	soma	7.5
exon	Zfp62	zinc finger protein 62 (Zfp62)	soma	8.4
intron	Rock1	Rho-associated coiled-coil containing protein kinase 1 (Rock1)	soma	9.1
exon	Eef1a2	eukaryotic translation elongation factor 1 alpha 2 (Eef1a2)	dendrite	16.9
exon	Calm3	calmodulin 3 (Calm3)	dendrite	23.4
intron	Csmd1	CUB and Sushi multiple domains 1 (Csmd1)	dendrite	54.9
intron	Grik2	glutamate receptor ionotropic kainate 2 (beta 2) (Grik2)	dendrite	59.7

Fig. S10.
Continued.

Neuron 2:(317 RNA reads, 43 mRNA reads, 42 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
intron	Cntnap5b	contactin associated protein-like 5B (Cntnap5b)	nucleus	1.1
intron	Ctnnd2	catenin (cadherin associated protein) delta 2 (Ctnnd2)	nucleus	2
ncRNA	1600010M07Rik	RIKEN cDNA 1600010M07 gene (1600010M07Rik)	nucleus	2.4
exon	Lgalsl	lectin galactoside binding-like (Lgalsl)	nucleus	2.7
intron	Tbc1d1	TBC1 domain family member 1 (Tbc1d1)	nucleus	2.8
exon	Atp2b2	ATPase Ca++ transporting plasma membrane 2 (Atp2b2)	nucleus	3
exon	Prickle2	prickle planar cell polarity protein 2 (Prickle2)	nucleus	3.9
exon	Nisch	nischarin (Nisch)	soma	4
exon	Syt7	synaptotagmin VII (Syt7)	soma	4.1
intron	B230303A05Rik	RIKEN cDNA B230303A05 gene (B230303A05Rik)	soma	4.2
exon	Ccp110	centriolar coiled coil protein 110 (Ccp110)	soma	4.2
exon	Add2	adducin 2 (beta) (Add2)	soma	4.5
intron	Epha7	Eph receptor A7 (Epha7)	soma	4.9
exon	Sorbs1	sorbin and SH3 domain containing 1 (Sorbs1)	soma	5
exon	Pten	phosphatase and tensin homolog (Pten)	soma	5
intron	Ank3	ankyrin 3 epithelial (Ank3)	soma	5.2
exon	Pnn	pinin (Pnn)	soma	5.2
exon	Ntrk2	neurotrophic tyrosine kinase receptor type 2 (Ntrk2)	soma	5.3
intron	Klhl32	kelch-like 32 (Klhl32)	soma	5.5
intron	Spock1	sparc/osteonectin cwcv and kazal-like domains proteoglycan 1 (Spock1)	soma	5.7
intron	Prkcb	protein kinase C beta (Prkcb)	soma	5.9
exon	Calm3	calmodulin 3 (Calm3)	soma	6.1
exon	Cnp	2'3'-cyclic nucleotide 3' phosphodiesterase (Cnp)	soma	6.1
exon	Tmem254c	transmembrane protein 254c (Tmem254c)	soma	6.4
exon	Slc35b4	solute carrier family 35 member B4 (Slc35b4)	soma	6.6
exon	Dst	dystonin (Dst)	soma	6.7
intron	Celf2	CUGBP Elav-like family member 2 (Celf2)	soma	6.9
intron	Clvs2	clavesin 2 (Clvs2)	soma	7
mrna-tf	Elf1	E74-like factor 1 (Elf1)	soma	7
mrna-tf	Hivep3	human immunodeficiency virus type I enhancer binding protein 3 (Hivep3)	soma	7.2
exon	Slc1a2	solute carrier family 1 (glial high affinity glutamate transporter) member 2 (Slc1a2)	soma	7.3
exon	Numbl	numb-like (Numbl)	soma	7.5
mrna-tf	Nfat5	nuclear factor of activated T cells 5 (Nfat5)	soma	8.1
exon	Prpf6	pre-mRNA splicing factor 6 (Prpf6)	soma	9.2
exon	Rcbtb2	regulator of chromosome condensation (RCC1) and BTB (POZ) domain containing	soma	9.4
exon	Zfp646	zinc finger protein 646 (Zfp646)	soma	10.1
ncRNA	Snord15b	small nucleolar RNA C/D box 14B (Snord15b)	dendrite	13
mrna-tf	Prox1	prospero homeobox 1 (Prox1)	dendrite	13.3
exon	Gabrg2	gamma-aminobutyric acid (GABA) A receptor subunit gamma 2 (Gabrg2)	dendrite	14.1
exon	Rab6b	RAB6B member RAS oncogene family (Rab6b)	dendrite	57.3
exon	Rab6b	RAB6B member RAS oncogene family (Rab6b)	dendrite	57.5
intron	Psd3	pleckstrin and Sec7 domain containing 3 (Psd3)	dendrite	61.4
intron	Cblb	Casitas B-lineage lymphoma b (Cblb)	dendrite	63

Fig. S10.
Continued.

Neuron 3: (239 RNA reads, 30 mRNA reads, 29 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
intron	Tmem229b	Mus musculus transmembrane protein 229B (Tmem229b)	nucleus	1.3
intron	Atp11c	ATPase class VI type 11C (Atp11c)	nucleus	1.7
intron	Tdrkh	tudor and KH domain containing protein (Tdrkh)	nucleus	2
intron	Tanc2	tetratricopeptide repeat ankyrin repeat and coiled-coil containing 2 (Tanc2)	nucleus	2.1
intron	Josd2	Josephin domain containing 2 (Josd2)	nucleus	2.4
exon	Gnao1	guanine nucleotide binding protein alpha O (Gnao1)	nucleus	2.4
exon	Tpi1	triosephosphate isomerase 1 (Tpi1)	nucleus	2.5
intron	Cntn4	contactin 4 (Cntn4)	nucleus	2.5
intron	Pcdh9	protocadherin 9 (Pcdh9)	nucleus	2.7
intron	Cacna1c	calcium channel voltage-dependent L type alpha 1C subunit (Cacna1c)	nucleus	2.8
intron	Mttp	microsomal triglyceride transfer protein (Mttp)	nucleus	2.9
exon	Sptbn4	spectrin beta non-erythrocytic 4 (Sptbn4)	nucleus	3
intron	Plxna4	plexin A4 (Plxna4)	nucleus	3.3
intron	Spock1	sparc/osteonectin cwcv and kazal-like domains proteoglycan 1 (Spock1)	nucleus	3.3
exon	Ppp2r5c	protein phosphatase 2 regulatory subunit B' gamma (Ppp2r5c)	nucleus	3.8
exon	Zfp865	zinc finger protein 865 (Zfp865)	nucleus	3.8
exon	Trmt2b	TRM2 tRNA methyltransferase 2B (Trmt2b)	soma	4.1
exon	Ralgapa2	Ral GTPase activating protein alpha subunit 2 (catalytic) (Ralgapa2)	nucleus	4.1
intron	Ngef	neuronal guanine nucleotide exchange factor (Ngef)	nucleus	4.1
exon	Trmt10a	tRNA methyltransferase 10A (Trmt10a)	nucleus	4.2
exon	Trmt10a	tRNA methyltransferase 10A (Trmt10a)	nucleus	4.2
intron	Pik3c2b	phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 beta (Pik3c2b)	soma	4.4
exon	Oaz1	ornithine decarboxylase antizyme 1 (Oaz1)	soma	4.4
exon	Camta1	calmodulin binding transcription activator 1 (Camta1)	soma	4.8
exon	Mrrf	mitochondrial ribosome recycling factor (Mrrf)	soma	5
exon	Ctdsp2	CTD (carboxy-terminal domain RNA polymerase II polypeptide A) small phosphatase 2 (Ctdsp2)	soma	5.2
exon	Clstn1	calsyntenin 1 (Clstn1)	soma	6.2
exon	Dynll2	dynein light chain LC8-type 2 (Dynll2)	soma	7.4
intron	Ctnna3	catenin (cadherin associated protein) alpha 3 (Ctnna3)	soma	9.2
intron	Rasgef1a	RasGEF domain family member 1A (Rasgef1a)	dendrite	11.5

Fig. S10.
Continued.

Neuron 4: (128 RNA reads, 10 mRNA reads, 9 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
mrna-tf	Rora	RAR-related orphan receptor alpha (Rora)	nucleus	1.6
intron	Fam168a	family with sequence similarity 168 member A (Fam168a)	nucleus	2.5
mrna-tf	Ep300	E1A binding protein p300 (Ep300)	nucleus	3.1
exon	Fam168a	family with sequence similarity 168 member A (Fam168a)	nucleus	4.2
intron	Xkr4	X-linked Kx blood group related 4 (Xkr4)	soma	5.5
exon	Gm20594	predicted gene 20594 (Gm20594)	soma	5.7
exon	Scd2	stearoyl-Coenzyme A desaturase 2 (Scd2)	soma	5.7
intron	Grik3	glutamate receptor ionotropic kainate 3 (Grik3)	soma	6.3
exon	Numb	NUMB endocytic adaptor protein (Numb)	dendrite	13.2
exon	Arhgap32	Rho GTPase activating protein 32 (Arhgap32)	dendrite	17.6

Fig. S10.
Continued.

Neuron 5: (186 RNA reads, 28 mRNA reads, 26 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
exon	Atp9a	ATPase class II type 9A (Atp9a)	soma	3.9
exon	Tgoln2	trans-golgi network protein 2 (Tgoln2)	soma	4.5
exon	Syt1	synaptotagmin I (Syt1)	soma	5
exon	Hpca	hippocalcin (Hpca)	soma	5.3
exon	Rab6b	RAB6B member RAS oncogene family (Rab6b)	soma	5.6
intron	Camk2b	calcium/calmodulin-dependent protein kinase II beta	soma	6.1
exon	Ppp1r9b	protein phosphatase 1 regulatory subunit 9B (Ppp1r9b)	soma	6.4
exon	Camk2b	calcium/calmodulin-dependent protein kinase II beta	soma	6.8
ncRNA	1600010M07Rik	RIKEN cDNA 1600010M07 gene (1600010M07Rik)	soma	6.9
exon	Vxn	vexin (Vxn)	soma	7.1
exon	4921524J17Rik	RIKEN cDNA 4921524J17 gene (4921524J17Rik)	soma	7.6
exon	Calb1	calbindin 1 (Calb1)	soma	7.6
intron	Inpp4a	inositol polyphosphate-4-phosphatase type I (Inpp4a)	soma	7.9
exon	Ftsj3	FtsJ RNA methyltransferase homolog 3 (E. coli) (Ftsj3)	soma	8.1
intron	Lsamp	limbic system-associated membrane protein (Lsamp)	soma	8.7
exon	Aak1	AP2 associated kinase 1 (Aak1)	soma	9.3
exon	Dopey1	dopey family member 1 (Dopey1)	dendrite	9.6
exon	Sphkap	SPHK1 interactor AKAP domain containing (Sphkap)	soma	9.6
exon	Sphkap	SPHK1 interactor AKAP domain containing (Sphkap)	soma	9.7
exon	Top2b	topoisomerase (DNA) II beta (Top2b)	soma	11.2
intron	Cntnap4	contactin associated protein-like 4 (Cntnap4)	soma	11.5
exon	Dgki	diacylglycerol kinase iota (Dgki)	dendrite	21.8
exon	Sh3gl1	SH3-domain GRB2-like 1 (Sh3gl1)	dendrite	22.5
exon	Nob1	NIN1/RPN12 binding protein 1 homolog (Nob1)	dendrite	31.2
ncRNA	Bc1	brain cytoplasmic RNA 1 (Bc1)	dendrite	35
intron	Rbfox1	RNA binding protein fox-1 homolog (C. elegans) 1 (Rbfox1)	dendrite	43.3
intron	Mgrn1	mahogunin ring finger 1 (Mgrn1)	dendrite	50.8
intron	Shtn1	shootin 1 (Shtn1)	dendrite	66.8

Fig. S10.
Continued.

Neuron 6: (151 RNA reads, 13 mRNA reads, 13 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
exon	Lsm7	LSM7 homolog U6 small nuclear RNA and mRNA degradation associated (Lsm7)	nucleus	1
exon	Snap25	synaptosomal-associated protein 25 (Snap25)	nucleus	1.2
exon	Tomm20	translocase of outer mitochondrial membrane 20 (Tomm20)	nucleus	4.3
intron	Kalrn	kalirin RhoGEF kinase (Kalrn)	soma	5.9
exon	Diexf	digestive organ expansion factor homolog (zebrafish) (Diexf)	soma	7.7
mrna-tf	Prkar1a	protein kinase cAMP dependent regulatory type I alpha (Prkar1a)	soma	7.7
intron	Ankrd45	ankyrin repeat domain 45 (Ankrd45)	soma	7.8
exon	Fry	FRY microtubule binding protein (Fry)	soma	8.3
exon	Tufm	Tu translation elongation factor mitochondrial (Tufm)	dendrite	8.4
exon	Kmt2e	lysine (K)-specific methyltransferase 2E (Kmt2e)	soma	9.2
exon	Rasal2	RAS protein activator like 2 (Rasal2)	dendrite	9.8
intron	Btbd9	BTB (POZ) domain containing 9 (Btbd9)	dendrite	22.4
mrna-tf	Mga	MAX gene associated (Mga)	dendrite	53.1

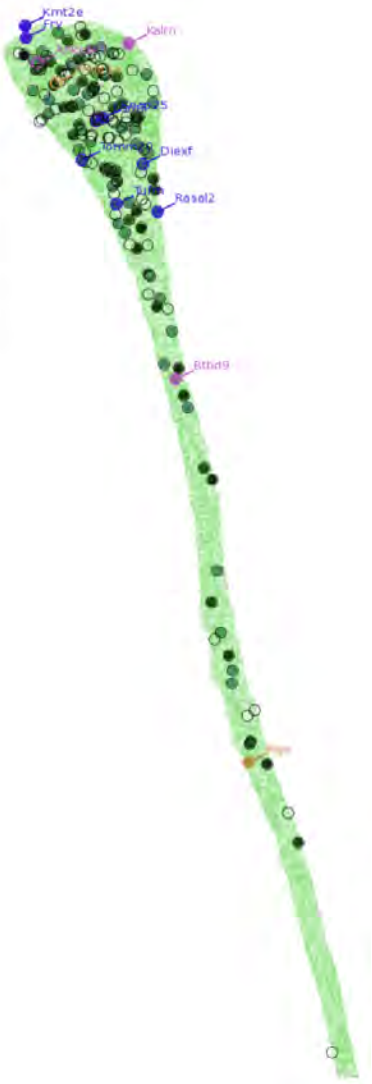


Fig. S10.
Continued.

Neuron 7: (209 RNA reads, 32 mRNA reads, 32 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
exon	Ywhaq	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta (Ywhaq)	nucleus	1.6
exon	Mdm2	transformed mouse 3T3 cell double minute 2 (Mdm2)	nucleus	2.7
exon	Eef1a2	eukaryotic translation elongation factor 1 alpha 2 (Eef1a2)	nucleus	3.1
intron	Ppp3ca	protein phosphatase 3 catalytic subunit alpha isoform (Ppp3ca)	nucleus	3.6
intron	Fam155a	family with sequence similarity 155 member A (Fam155a)	nucleus	3.8
intron	Tnik	TRAF2 and NCK interacting kinase (Tnik)	nucleus	4.1
intron	Atp2b1	ATPase Ca++ transporting plasma membrane 1 (Atp2b1)	nucleus	4.7
mrna-tf	Prkar1a	protein kinase cAMP dependent regulatory type I alpha (Prkar1a)	soma	5.1
intron	Tox	thymocyte selection-associated high mobility group box (Tox)	nucleus	5.6
exon	Cux1	cut-like homeobox 1 (Cux1)	soma	5.9
exon	Reep2	receptor accessory protein 2 (Reep2)	soma	6
exon	Vangl2	VANGL planar cell polarity 2 (Vangl2)	soma	6.2
intron	Cabp1	calcium binding protein 1 (Cabp1)	soma	6.5
intron	Mir3473a	microRNA 3473a (Mir3473a)	soma	6.7
intron	Igf1r	insulin-like growth factor I receptor (Igf1r)	soma	6.8
mrna-tf	Rfx3	regulatory factor X 3 (influences HLA class II expression) (Rfx3)	soma	6.9
exon	H3f3b	H3 histone family 3B (H3f3b)	soma	7.2
exon	Islr2	immunoglobulin superfamily containing leucine-rich repeat 2 (Islr2)	soma	7.2
intron	Crebrf	CREB3 regulatory factor (Crebrf)	soma	8.1
intron	Grm1	glutamate receptor metabotropic 1 (Grm1)	soma	8.1
mrna-tf	Trp53bp1	transformation related protein 53 binding protein 1 (Trp53bp1)	soma	8.3
intron	Kcnd2	potassium voltage-gated channel Shal-related family member 2 (Kcnd2)	soma	8.4
exon	Top2b	topoisomerase (DNA) II beta (Top2b)	soma	8.4
exon	Dopey2	dopey family member 2 (Dopey2)	soma	8.4
exon	Mpc2	mitochondrial pyruvate carrier 2 (Mpc2)	soma	8.9
exon	Runx1t1	runt-related transcription factor 1; translocated to 1 (cyclin D-related) (Runx1t1)	soma	9.4
exon	Pippr4	phospholipid phosphatase related 4 (Pippr4)	soma	9.8
intron	Celf2	CUGBP Elav-like family member 2 (Celf2)	soma	10.1
exon	Ufl1	UFM1 specific ligase 1 (Ufl1)	soma	10.2
intron	Oxr1	oxidation resistance 1 (Oxr1)	dendrite	12.4
intron	Tra2b	transformer 2 beta (Tra2b)	dendrite	64.5
exon	Cadm1	cell adhesion molecule 1 (Cadm1)	dendrite	64.8

Fig. S10.
Continued.

Neuron 8: (212 RNA reads, 37 mRNA reads, 37 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
exon	Grin2b	glutamate receptor ionotropic NMDA2B (epsilon 2) (Grin2b)	nucleus	2.1
intron	Sorl1	sortilin-related receptor LDLR class A repeats-containing (Sorl1)	nucleus	2.6
exon	Gm20594	predicted gene 20594 (Gm20594)	nucleus	2.6
intron	Cnksr2	connector enhancer of kinase suppressor of Ras 2 (Cnksr2)	nucleus	3.3
intron	Adcy5	adenylate cyclase 5 (Adcy5)	nucleus	3.7
exon	Nf1	neurofibromin 1 (Nf1)	nucleus	4
intron	Phf21a	PHD finger protein 21A (Phf21a)	soma	4.4
exon	Guk1	guanylate kinase 1 (Guk1)	nucleus	4.8
exon	Hspa8	heat shock protein 8 (Hspa8)	soma	4.9
intron	Zhx2	zinc fingers and homeoboxes 2 (Zhx2)	nucleus	5
exon	Kbtbd11	kelch repeat and BTB (POZ) domain containing 11 (Kbtbd11)	soma	5.1
mrna-tf	Mapk1	mitogen-activated protein kinase 1 (Mapk1)	soma	5.5
intron	2700049A03Rik	RIKEN cDNA 2700049A03 gene (2700049A03Rik)	soma	5.6
exon	Kcnc3	potassium voltage gated channel Shaw-related subfamily member 3 (Kcnc3)	soma	5.7
intron	Ppib	peptidylprolyl isomerase B (Ppib)	soma	5.9
exon	Dusp3	dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related) (Dusp3)	soma	6.3
intron	Kcnip1	Kv channel-interacting protein 1 (Kcnip1)	soma	6.6
exon	Ppp1r9a	protein phosphatase 1 regulatory subunit 9A (Ppp1r9a)	nucleus	6.8
intron	Grin2a	glutamate receptor ionotropic NMDA2A (epsilon 1) (Grin2a)	dendrite	7
intron	Ica1	islet cell autoantigen 1 (Ica1)	soma	7
exon	Msl1	male specific lethal 1 (Msl1)	soma	7.4
intron	Stap2	signal transducing adaptor family member 2 (Stap2)	soma	7.4
exon	Rnf157	ring finger protein 157 (Rnf157)	dendrite	7.9
mrna-tf	St18	suppression of tumorigenicity 18 (St18)	soma	7.9
exon	Dsel	dermatan sulfate epimerase-like (Dsel)	soma	8.8
intron	Zfp512	zinc finger protein 512 (Zfp512)	soma	9
mrna-tf	Foxg1	forkhead box G1 (Foxg1)	dendrite	29.3
intron	Zbtb20	zinc finger and BTB domain containing 20 (Zbtb20)	dendrite	32.2
exon	Ddx5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5 (Ddx5)	dendrite	35.5
intron	Celf2	CUGBP Elav-like family member 2 (Celf2)	dendrite	45.8
exon	Sptbn2	spectrin beta non-erythrocytic 2 (Sptbn2)	dendrite	47.2
exon	Sh3gl1	SH3-domain GRB2-like 1 (Sh3gl1)	dendrite	50.7
intron	Ino80d	INO80 complex subunit D (Ino80d)	dendrite	52.2
intron	Frmd4a	FERM domain containing 4A (Frmd4a)	dendrite	72.3
exon	Itsn1	intersectin 1 (SH3 domain protein 1A) (Itsn1)	dendrite	72.6
exon	Chl1	cell adhesion molecule L1-like (Chl1)	dendrite	73.1
exon	Nrip3	nuclear receptor interacting protein 3 (Nrip3)	dendrite	89.9

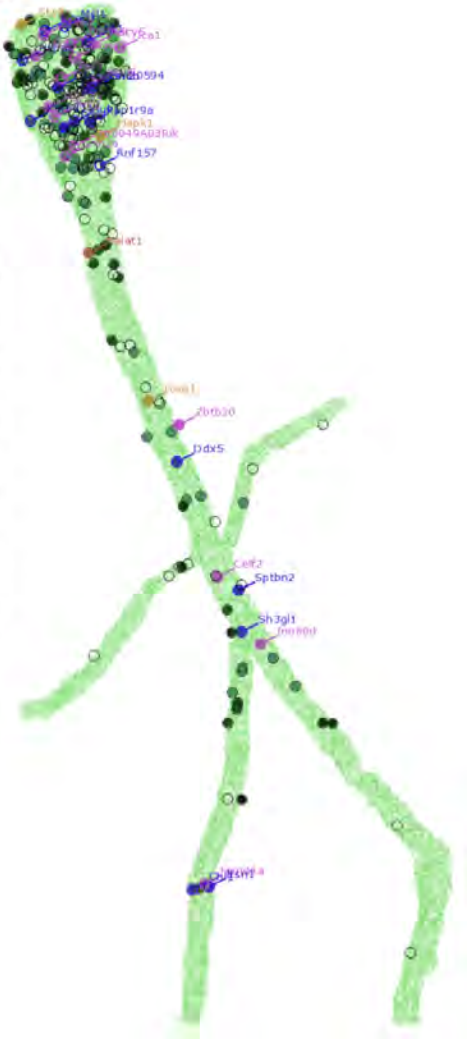


Fig. S10.
Continued.

Neuron 9: (208 RNA reads, 27 mRNA reads, 25 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
intron	Plxna2	plexin A2 (Plxna2)	nucleus	2.4
intron	Pitpnc1	phosphatidylinositol transfer protein cytoplasmic 1 (Pitpnc1)	nucleus	3
exon	Map1b	microtubule-associated protein 1B (Map1b)	nucleus	3.3
intron	Pde4d	phosphodiesterase 4D cAMP specific (Pde4d)	soma	4.7
exon	Fam131b	family with sequence similarity 131 member B (Fam131b)	soma	4.7
exon	Kcnk9	potassium channel subfamily K member 9 (Kcnk9)	nucleus	4.7
exon	Cul3	cullin 3 (Cul3)	soma	4.8
exon	Map2	microtubule-associated protein 2 (Map2)	soma	4.8
ncRNA	2900060B14Rik	RIKEN cDNA 2900060B14 gene (2900060B14Rik)	soma	4.8
exon	Map2	microtubule-associated protein 2 (Map2)	soma	4.9
exon	Prkaa2	protein kinase AMP-activated alpha 2 catalytic subunit (Prkaa2)	soma	5.3
exon	Dync1h1	dynein cytoplasmic 1 heavy chain 1 (Dync1h1)	soma	5.3
mrna-tf	Tcerg1	transcription elongation regulator 1 (CA150) (Tcerg1)	soma	5.3
exon	Prrc2c	proline-rich coiled-coil 2C (Prrc2c)	soma	5.3
intron	Dnah7b	dynein axonemal heavy chain 7B (Dnah7b)	soma	5.6
intron	Copg2	coatomer protein complex subunit gamma 2 (Copg2)	soma	6
intron	Mvb12b	multivesicular body subunit 12B (Mvb12b)	soma	6.4
exon	Adar	adenosine deaminase RNA-specific (Adar)	soma	6.5
intron	Erc2	ELKS/RAB6-interacting/CAST family member 2 (Erc2)	soma	6.7
exon	Neo1	neogenin (Neo1)	soma	6.7
exon	Lsm7	LSM7 homolog U6 small nuclear RNA and mRNA degradation associated (Lsm7)	soma	6.9
exon	Uchl1	ubiquitin carboxy-terminal hydrolase L1 (Uchl1)	soma	7.2
exon	Supt6	suppressor of Ty 6 (Supt6)	dendrite	15.7
exon	Arl8b	ADP-ribosylation factor-like 8B (Arl8b)	dendrite	21.6
exon	Arl8b	ADP-ribosylation factor-like 8B (Arl8b)	dendrite	21.9
exon	Inpp5a	inositol polyphosphate-5-phosphatase A (Inpp5a)	dendrite	24.5
intron	Cdh9	cadherin 9 (Cdh9)	dendrite	71.5

Fig. S10.
Continued.

Neuron 10: (358 RNA reads, 58 mRNA reads, 56 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
ncRNA	1600010M07Rik	RIKEN cDNA 1600010M07 gene (1600010M07Rik)	nucleus	1.8
intron	Ampd3	adenosine monophosphate deaminase 3 (Ampd3)	nucleus	2.7
exon	Matr3	matrin 3 (Matr3)	nucleus	3
intron	Stx7	syntaxin 7 (Stx7)	nucleus	3.1
intron	Dcc	deleted in colorectal carcinoma (Dcc)	nucleus	3.2
exon	Sphkap	SPHK1 interactor AKAP domain containing (Sphkap)	nucleus	3.3
exon	Camk2a	calcium/calmodulin-dependent protein kinase II alpha (Camk2a)	nucleus	3.3
exon	Unc13b	unc-13 homolog B (Unc13b)	nucleus	3.4
mrna-tf	Foxo3	forkhead box O3 (Foxo3)	soma	3.6
exon	Ryr2	ryanodine receptor 2 cardiac (Ryr2)	soma	4
exon	Hsp90ab1	heat shock protein 90 alpha (cytosolic) class B member 1 (Hsp90ab1)	soma	4.3
intron	Ankrd33b	ankyrin repeat domain 33B (Ankrd33b)	nucleus	4.3
exon	Prdm2	PR domain containing 2 with ZNF domain (Prdm2)	soma	4.5
ncRNA	Malat1	metastasis associated lung adenocarcinoma transcript 1 (non-coding RNA) (Malat1)	nucleus	4.7
exon	Arfip2	ADP-ribosylation factor interacting protein 2 (Arfip2)	soma	4.8
intron	Ppp6r3	protein phosphatase 6 regulatory subunit 3 (Ppp6r3)	soma	5.5
intron	Nin	ninein (Nin)	soma	5.6
intron	Klhl12	kelch-like 12 (Klhl12)	soma	5.8
intron	Rock1	Rho-associated coiled-coil containing protein kinase 1 (Rock1)	soma	5.9
exon	Fyn	Fyn proto-oncogene (Fyn)	soma	5.9
mrna-tf	Snapc4	small nuclear RNA activating complex polypeptide 4 (Snapc4)	soma	6
intron	H2afy	H2A histone family member Y (H2afy)	soma	6.1
exon	Zmynd8	zinc finger MYND-type containing 8 (Zmynd8)	soma	6.1
exon	D430042O09Rik	RIKEN cDNA D430042O09 gene (D430042O09Rik)	soma	6.7
exon	Rev3l	REV3 like DNA directed polymerase zeta catalytic subunit (Rev3l)	soma	7.1
exon	Celf2	CUGBP Elav-like family member 2 (Celf2)	soma	7.2
exon	Scamp4	secretory carrier membrane protein 4 (Scamp4)	soma	7.7
mrna-tf	Ncor2	nuclear receptor co-repressor 2 (Ncor2)	soma	7.9
exon	Sephs1	selenophosphate synthetase 1 (Sephs1)	soma	9.9
exon	Pclo	piccolo (presynaptic cytomatrix protein) (Pclo)	soma	10.8
intron	Shisa6	shisa family member 6 (Shisa6)	dendrite	11.8
intron	Dgkh	diacylglycerol kinase eta (Dgkh)	dendrite	12.3
exon	Slc2a3	solute carrier family 2 (facilitated glucose transporter) member 3 (Slc2a3)	dendrite	12.5
intron	Prkce	protein kinase C epsilon (Prkce)	dendrite	13.5
exon	Map1b	microtubule-associated protein 1B (Map1b)	dendrite	14.8
exon	Gria1	glutamate receptor ionotropic AMPA1 (alpha 1) (Gria1)	dendrite	16.1
intron	Ksr1	kinase suppressor of ras 1 (Ksr1)	dendrite	18.4
intron	Mdn1	midasin AAA ATPase 1 (Mdn1)	dendrite	19.7
mrna-tf	Ank2	ankyrin 2 brain (Ank2)	dendrite	20.5
exon	Serinc1	serine incorporator 1 (Serinc1)	dendrite	22.7
intron	Epc2	enhancer of polycomb homolog 2 (Epc2)	dendrite	22.9
exon	Serinc1	serine incorporator 1 (Serinc1)	dendrite	22.9
exon	Slc1a2	solute carrier family 1 (glial high affinity glutamate transporter) member 2 (Slc1a2)	dendrite	22.9
exon	Slc25a4	solute carrier family 25 (mitochondrial carrier adenine nucleotide translocator) member 4 (Slc25a4)	dendrite	25.6
intron	Tomm70a	translocase of outer mitochondrial membrane 70A (Tomm70a)	dendrite	28.5
mrna-tf	Ssbp2	single-stranded DNA binding protein 2 (Ssbp2)	dendrite	29
exon	Wipi1	WD repeat domain phosphoinositide interacting 1 (Wipi1)	dendrite	33.3
exon	Nuak1	NUAK family SNF1-like kinase 1 (Nuak1)	dendrite	34.8
intron	Kalrn	kalirin RhoGEF kinase (Kalrn)	dendrite	36.3
exon	Fancm	Fanconi anemia complementation group M (Fancm)	dendrite	38.7
mrna-tf	Lmo4	LIM domain only 4 (Lmo4)	nucleus	46.4
mrna-tf	Ep300	E1A binding protein p300 (Ep300)	nucleus	49.3
intron	Pitpnc1	phosphatidylinositol transfer protein cytoplasmic 1 (Pitpnc1)	dendrite	52
intron	Dnah7b	dynein axonemal heavy chain 7B (Dnah7b)	nucleus	54.1
exon	Trmt10a	tRNA methyltransferase 10A (Trmt10a)	dendrite	54.9
intron	Josd2	Josephin domain containing 2 (Josd2)	dendrite	55.5
exon	Trmt10a	tRNA methyltransferase 10A (Trmt10a)	dendrite	55.6
intron	Klf8	Kruppel-like factor 8 (Klf8)	dendrite	76.3

Fig. S10.
Continued.

Neuron 11: (121 RNA reads, 18 mRNA reads, 18 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
mrna-tf	Sin3b	transcriptional regulator SIN3B (yeast) (Sin3b)	soma	3.8
exon	Gdap1	ganglioside-induced differentiation-associated-protein 1 (Gdap1)	soma	4.1
exon	Pfkp	phosphofructokinase platelet (Pfkp)	soma	5
exon	Tagap	T cell activation Rho GTPase activating protein (Tagap)	soma	5.5
intron	Naa30	N(alpha)-acetyltransferase 30 NatC catalytic subunit (Naa30)	soma	5.8
exon	Zfp558	zinc finger protein 558 (Zfp558)	soma	7
exon	Lrp4	low density lipoprotein receptor-related protein 4 (Lrp4)	soma	8.3
exon	Add2	adducin 2 (beta) (Add2)	dendrite	11.5
intron	Ryr2	ryanodine receptor 2 cardiac (Ryr2)	dendrite	14.1
exon	Lyst	lysosomal trafficking regulator (Lyst)	dendrite	14.5
exon	Serinc1	serine incorporator 1 (Serinc1)	dendrite	22.2
mrna-tf	Zfp39	zinc finger protein 39 (Zfp39)	dendrite	23.4
intron	Arhgap29	Rho GTPase activating protein 29 (Arhgap29)	dendrite	28.7
exon	Sf3b2	splicing factor 3b subunit 2 (Sf3b2)	dendrite	30.5
intron	Adarb2	adenosine deaminase RNA-specific B2 (Adarb2)	dendrite	30.7
exon	Syt1	synaptotagmin I (Syt1)	dendrite	32.9
intron	Nup93	nucleoporin 93 (Nup93)	dendrite	52.6
intron	Prkab2	protein kinase AMP-activated beta 2 non-catalytic subunit (Prkab2)	dendrite	69.4

Fig. S10.
Continued.

Neuron 12: (314 RNA reads, 24 mRNA reads, 21 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
intron	Dmtn	dematin actin binding protein (Dmtn)	nucleus	2.4
ncRNA	Meg3	maternally expressed 3 (Meg3)	nucleus	3.2
intron	Clvs2	clavesin 2 (Clvs2)	nucleus	4.1
ncRNA	Malat1	metastasis associated lung adenocarcinoma transcript 1 (non-coding RNA) (Malat1)	soma	4.3
intron	Stap2	signal transducing adaptor family member 2 (Stap2)	soma	4.5
intron	Mob2	MOB kinase activator 2 (Mob2)	soma	4.7
intron	Klhl32	kelch-like 32 (Klhl32)	soma	4.7
intron	Adora1	adenosine A1 receptor (Adora1)	soma	5.1
exon	Pcdhgc5	protocadherin gamma subfamily C 5 (Pcdhgc5)	soma	5.1
exon	Tspan5	tetraspanin 5 (Tspan5)	soma	5.1
exon	Caln1	calneuron 1 (Caln1)	soma	5.2
exon	Eif1	eukaryotic translation initiation factor 1 (Eif1)	soma	5.3
intron	Opcml	opioid binding protein/cell adhesion molecule-like (Opcml)	soma	5.6
exon	Setd2	SET domain containing 2 (Setd2)	dendrite	12.8
intron	Cecr2	CECR2 histone acetyl-lysine reader (Cecr2)	dendrite	28.3
mrna-tf	Zfp592	zinc finger protein 592 (Zfp592)	dendrite	31.8
intron	Kcnip1	Kv channel-interacting protein 1 (Kcnip1)	dendrite	42.7
mrna-tf	Ptma	prothymosin alpha (Ptma)	dendrite	43.9
intron	Fam155a	family with sequence similarity 155 member A (Fam155a)	dendrite	44.7
exon	Tubb4a	tubulin beta 4A class IVA (Tubb4a)	dendrite	47.5
exon	Dopey2	dopey family member 2 (Dopey2)	dendrite	51.4
intron	Stap2	signal transducing adaptor family member 2 (Stap2)	dendrite	58.6
intron	Stap2	signal transducing adaptor family member 2 (Stap2)	dendrite	58.8
ncRNA	Malat1	metastasis associated lung adenocarcinoma transcript 1 (non-coding RNA) (Malat1)	dendrite	76.1

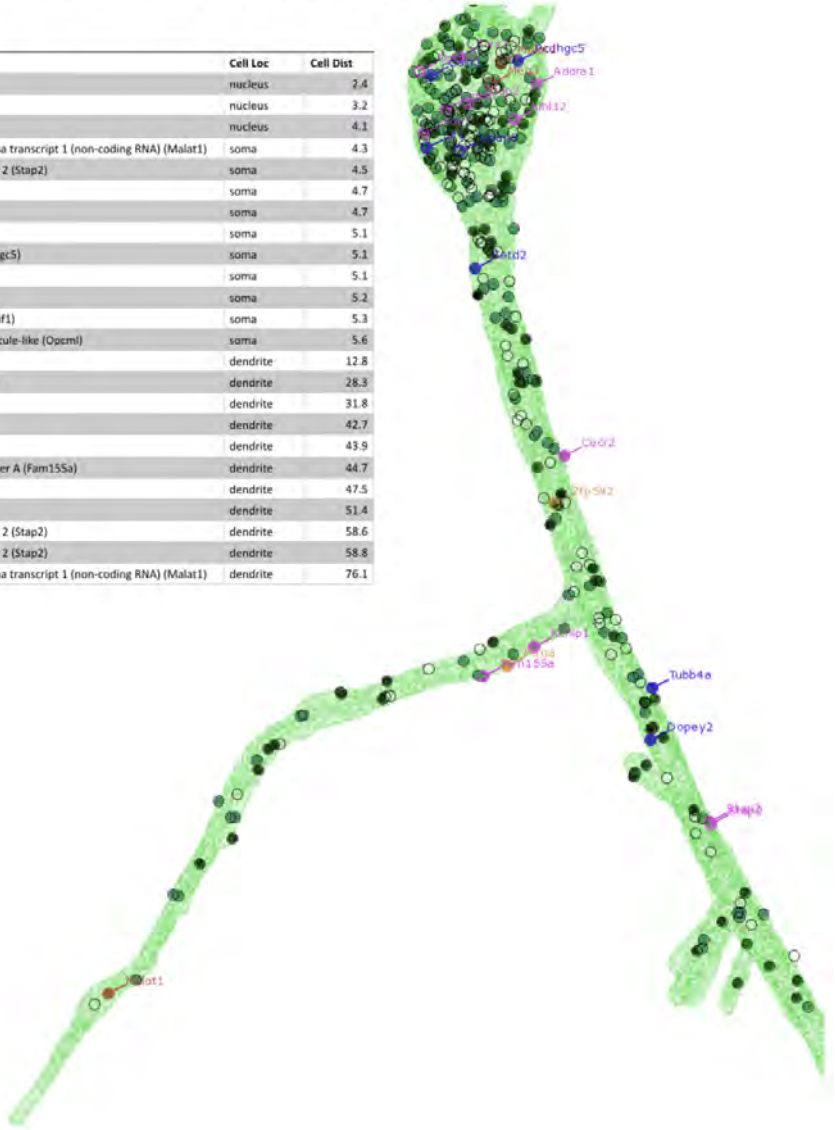


Fig. S10.
Continued.

Neuron 13: (160 RNA reads, 10 mRNA reads, 10 different genes)

ReadType	Gene Name	Gene Desc	Cell Loc	Cell Dist
exon	Nckap1	NCK-associated protein 1 (Nckap1)	nucleus	2.9
mrna-tf	Dazap2	DAZ associated protein 2 (Dazap2)	nucleus	3.2
exon	Dync1h1	dynein cytoplasmic 1 heavy chain 1 (Dync1h1)	nucleus	3.3
exon	Slc1a2	solute carrier family 1 (glial high affinity glutamate transporter) member 2 (Slc1a2)	nucleus	3.9
intron	Dpp6	dipeptidylpeptidase 6 (Dpp6)	nucleus	4.1
exon	Got1	glutamic-oxaloacetic transaminase 1 soluble (Got1)	soma	4.5
intron	Kiz	kizuna centrosomal protein (Kiz)	soma	4.8
intron	Map7	microtubule-associated protein 7 (Map7)	soma	5.6
intron	Cntn4	contactin 4 (Cntn4)	soma	6
exon	Syp	synaptophysin (Syp)	dendrite	43.7

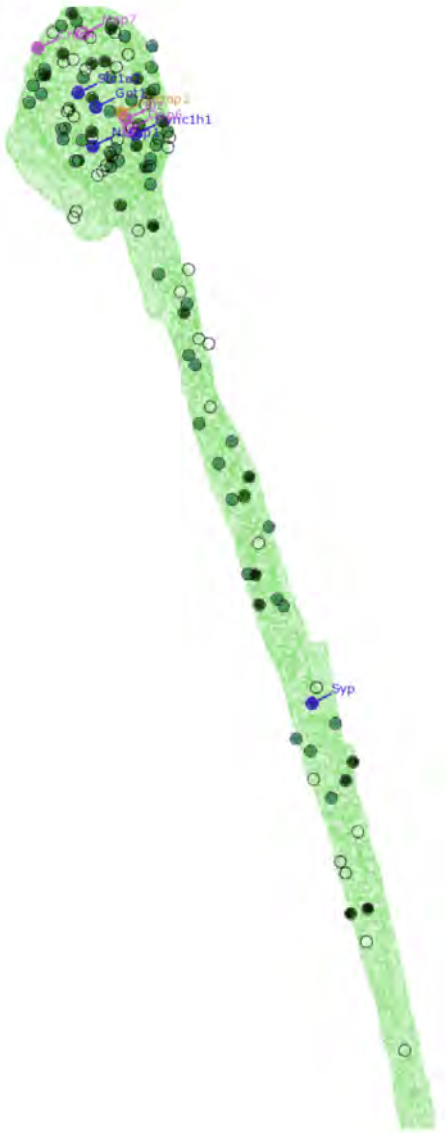


Fig. S10.
Continued.

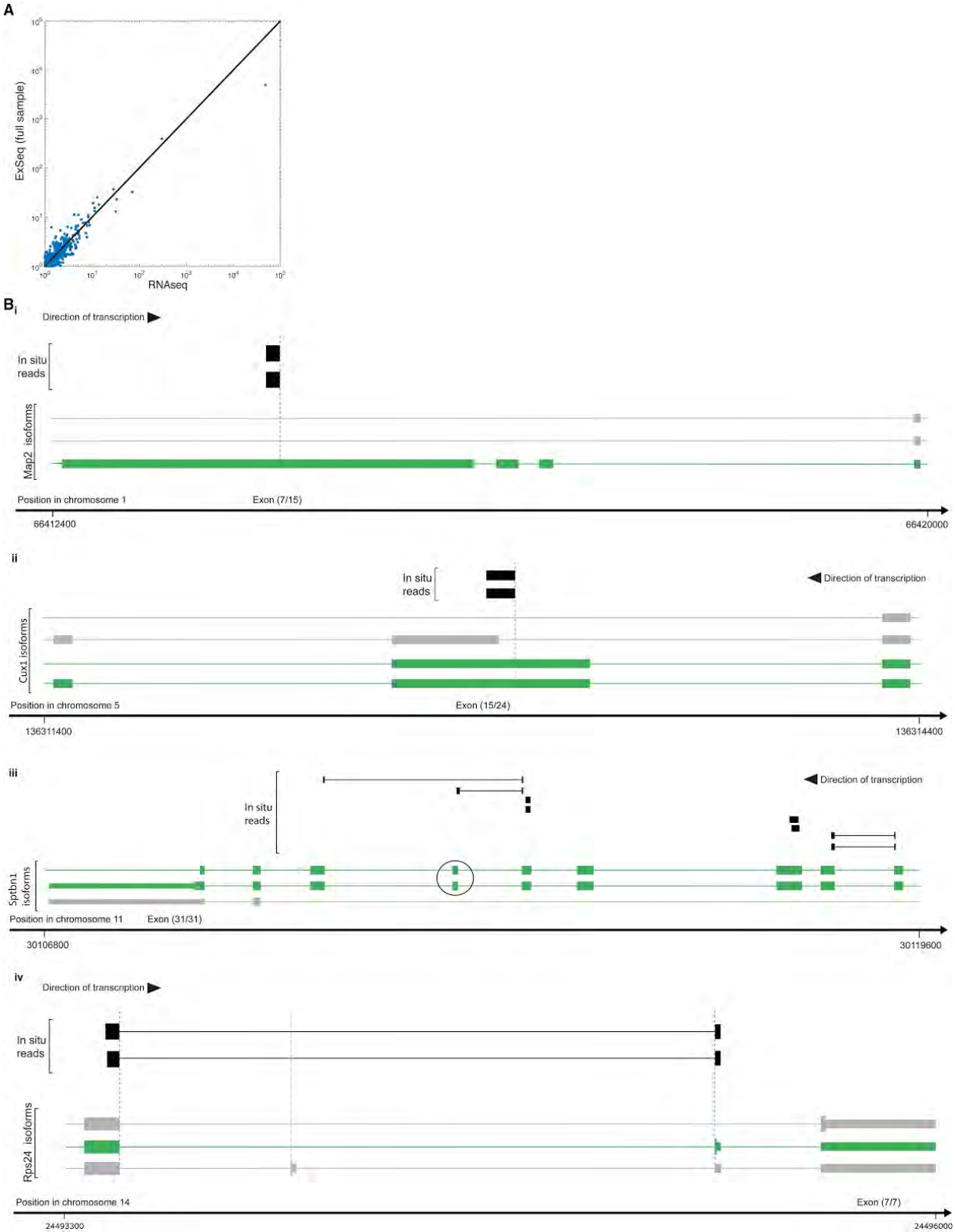

A
ExSeq (full sample)
RNAseq
Bi
Direction of transcription
In situ reads
Map2 isoforms
Position in chromosome 1
Exon (7/15)
66412400
66420000
ii
In situ reads
Direction of transcription
Cux1 isoforms
Position in chromosome 5
Exon (15/24)
136311400
136314400
iii
In situ reads
Direction of transcription
Sptbn1 isoforms
Position in chromosome 11
Exon (31/31)
30106800
30119600
iv
Direction of transcription
In situ reads
Rps24 isoforms
Position in chromosome 14
Exon (7/7)
24493300
24496000

Fig. S11.

Untargeted ExSeq mapping of alternative splicing events in neurons. (A) ExSeq resolves alternative splicing isoforms with high correlation to those seen with standard RNAseq. For each RefSeq gene identified in untargeted ExSeq, the counts of the highest frequency isoform divided by the counts of the second highest frequency isoform are calculated for ExSeq (y-axis) and for RNAseq of the adjacent slice (x-axis). (B) Examples of alternative splice events detected *in situ*, shown for the genes Microtubule Associated Protein 2 (Map2) (i), Cut Like Homeobox 1 (Cux1) (ii), Spectrin Beta, Non-Erythrocytic 1 (*Sptbn1*) (iii) and Ribosomal Protein S24 (Rps24) (iv). For each sub-figure the genomic coordinates are displayed at the bottom, the known gene isoforms (RefSeq annotations) that align to that genomic region are in the middle (one row for each isoform), and the aligned *in situ* reads are at the top (each row corresponds to one sequenced puncta). For each sub-figure, the relative position of one of the exons is displayed in the following format: x/y, i.e. exon number (starting from the 5' of the mRNA) / total number of exons. Thick lines for the known gene isoforms and *in situ* reads represent the sequences matching to the genomic coordinates at the bottom, whereas thin lines represent gaps in the alignment due to intronic regions. For the mRNA, thin lines represent introns, medium-thick lines represent untranslated regions, and fully thick lines represent coding regions. For all four genes, the *in situ* reads are matching the structure of specific mRNA isoforms (shown in green) and not other isoforms (grey), therefore revealing the expressed alternative splicing isoforms. In addition, for *Sptbn1*, the top *in situ* read possibly reveals a previously uncharacterized exon skipping splicing event, as the exon marked in a circle is not present in the read whereas the two flanking exons are present.

Fig. S12.

Architecture of padlock probes for targeted ExSeq. Targeted ExSeq padlock probes consist of four key elements: (1) a 32 nt homology region, split into two 16 nt halves (H1, H2) on the 5' and 3' ends of the padlock probe; (2) a constant RCA/Sequencing primer site on the backbone; (3) a barcode ('BC') sequence for readout with SOLiD sequencing chemistry; and (4) a barcode sequence for readout with Illumina chemistry. After ligation and rolling circle amplification, the amplicon concatamer is formed.

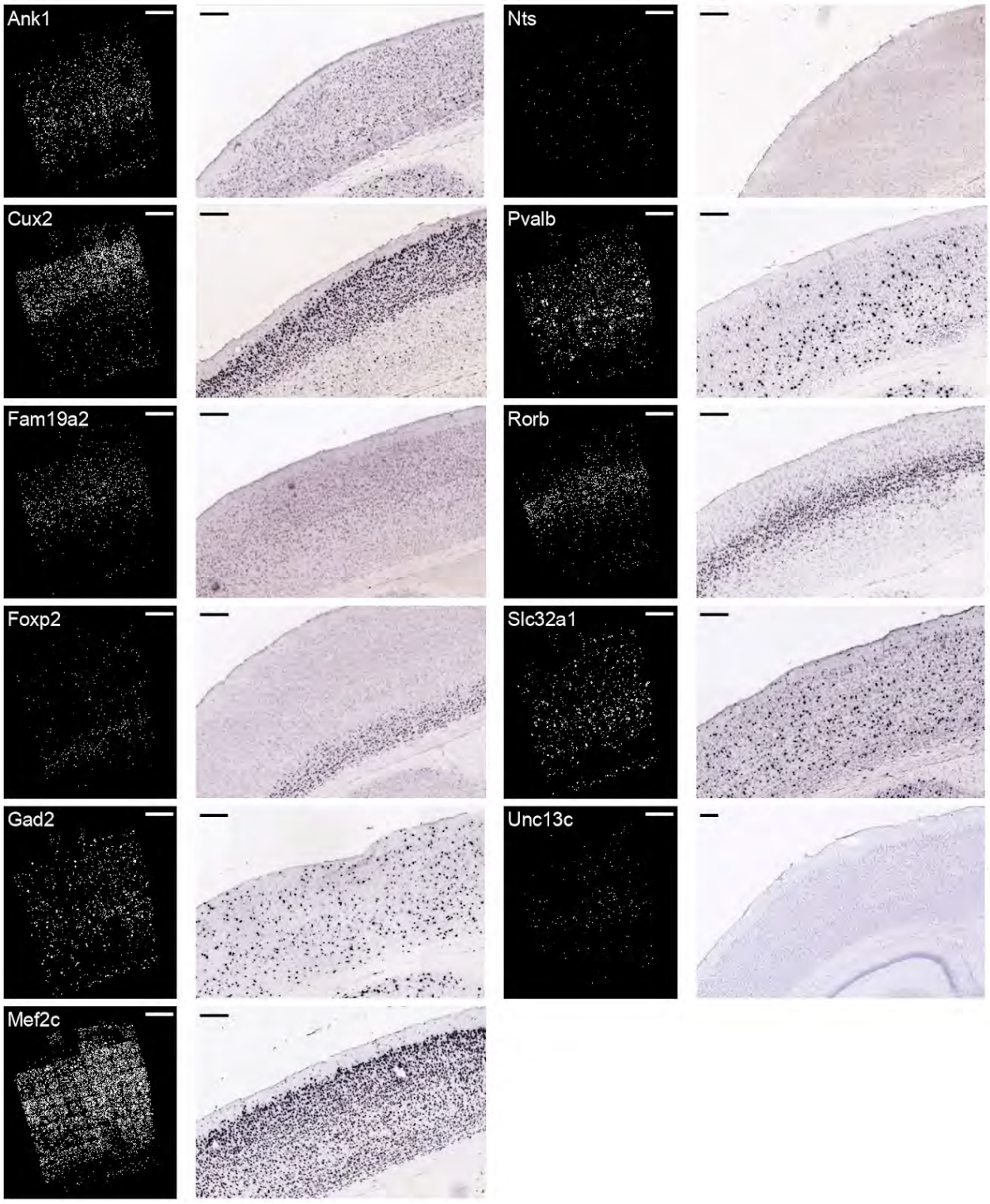

Ank1
Nts
Cux2
Pvalb
Fam19a2
Rorb
Foxp2
Slc32a1
Gad2
Unc13c
Mef2c

Fig. S13.

Comparison between targeted ExSeq and Allen Institute *in situ* hybridization (ISH) atlas (https://mouse.brain-map.org/)(*63*) for selected marker genes in the primary visual cortex. Scale bars: targeted ExSeq, 200 microns pre-expansion; Allen Institute ISH, 200 microns.

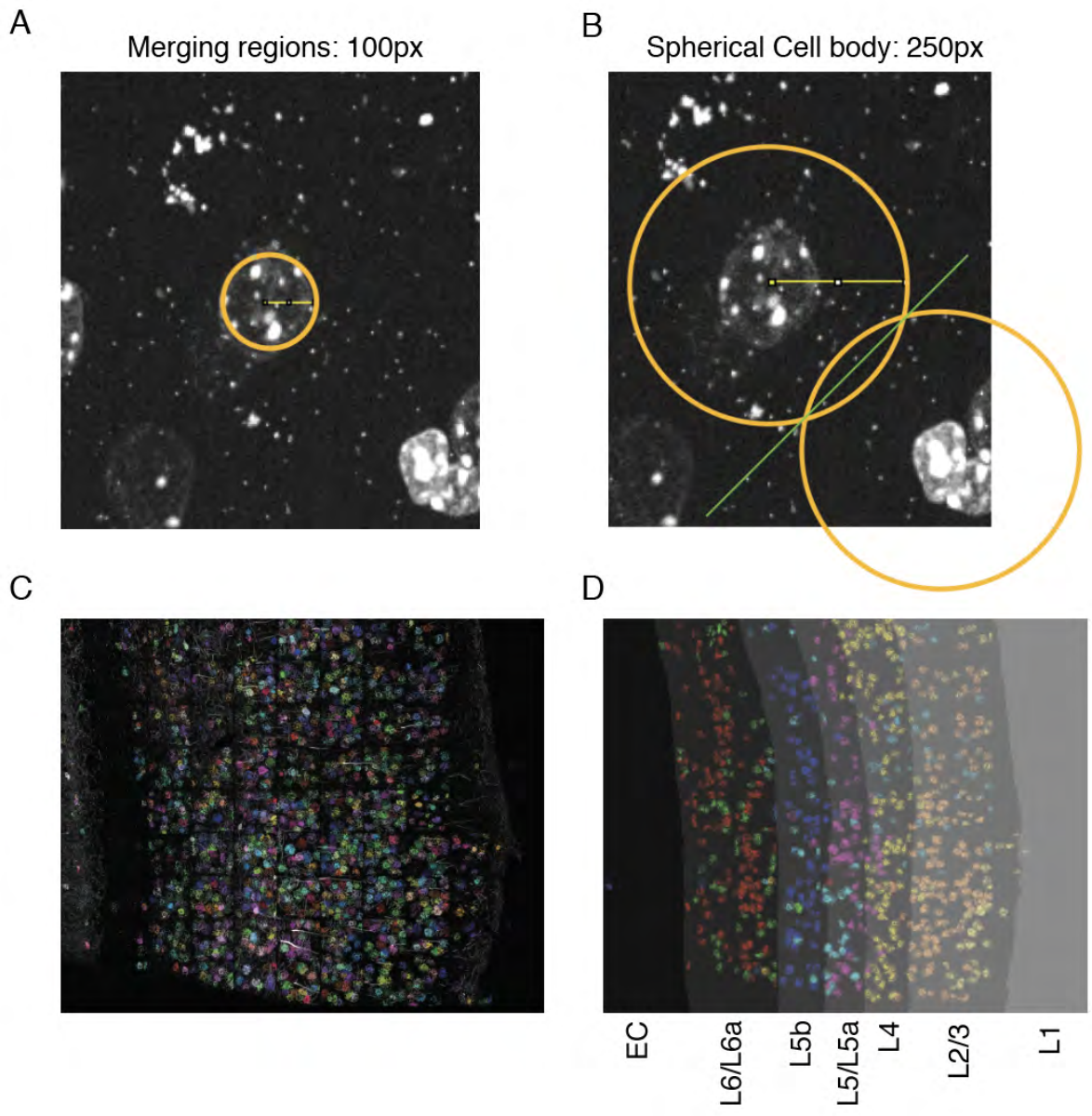


Fig. S14.

Cell and layer segmentation in the visual cortex. (A) DAPI staining marks putative nuclei objects, which were merged together if the centroids of the objects were within a radius of 5 microns (pre-expansion, 100 px) (shown). (B) Reads within 12.5 microns (pre-expansion) of a centroid were ascribed to that cell; if two centroids were within 12.5 microns of a read (as shown), the reads were ascribed to their nearest neighbor. (C) Demonstration of the cell segmentation pipeline by coloring all reads assigned to an individual cell with a random color. (D) The segmentation of the visual cortex into layers. The excitatory neuron clusters from **Fig. 4D-E** are shown, alongside the layer segmentation in gradations of grey.

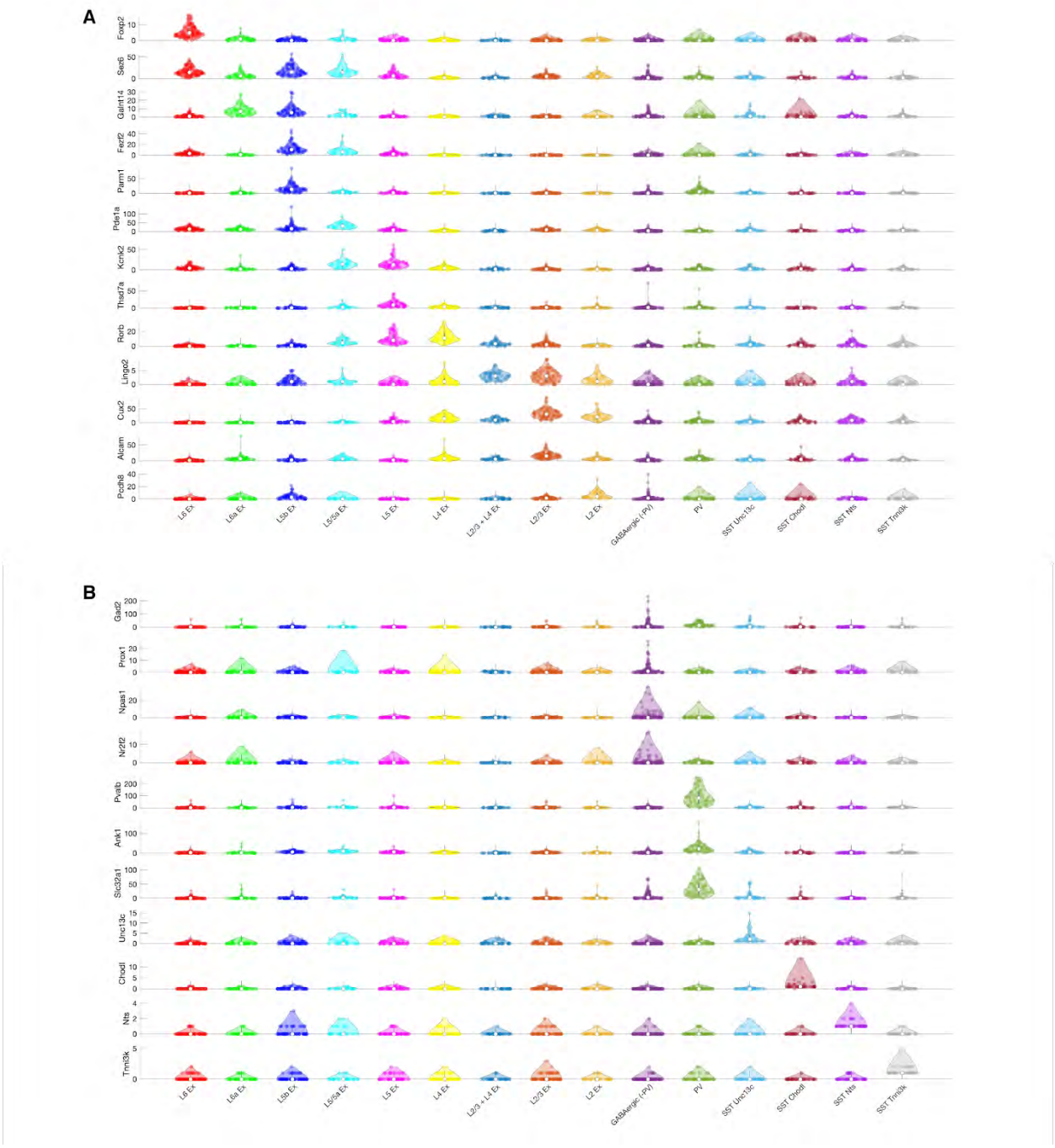


Fig. S15.

Expression of marker genes in the visual cortex clusters. (A) Violin plots of markers for excitatory neurons in **Fig. 4D**, showing the number of reads for each marker gene in each cell within a cluster. (B) Violin plots of markers for inhibitory neurons identified in **Fig. 4D**, showing the number of reads for each marker gene in each cell within a cluster. These results are consistent with previous studies (including (*61*)). For example: Foxp2 in layer 6 excitatory neurons (L6 Ex) (*132*), Sez6 in layer 5 and 6 excitatory neurons (L6 Ex, L6a Ex, L5b Ex, L5/L5a

Ex, L5 Ex) (*133*-*134*); Kcnk2 in layer 5 excitatory neurons (L5/L5a Ex, L5 Ex) (*135*, *136*); Lingo2 in layer 2/3 excitatory neurons (L2 Ex, L2/3 Ex, L2/3 + L4 Ex) (*61*); Cux2 in layer 2/3 excitatory neurons (L2/3 Ex, L2 Ex) (*41*); Pvalb in PV inhibitory neurons (PV) (*137*); Prox1 in VIP inhibitory neurons (GABAergic (-PV)) (*61*); and Chodl in a subset of SST neurons (SST Chodl) (*61*).

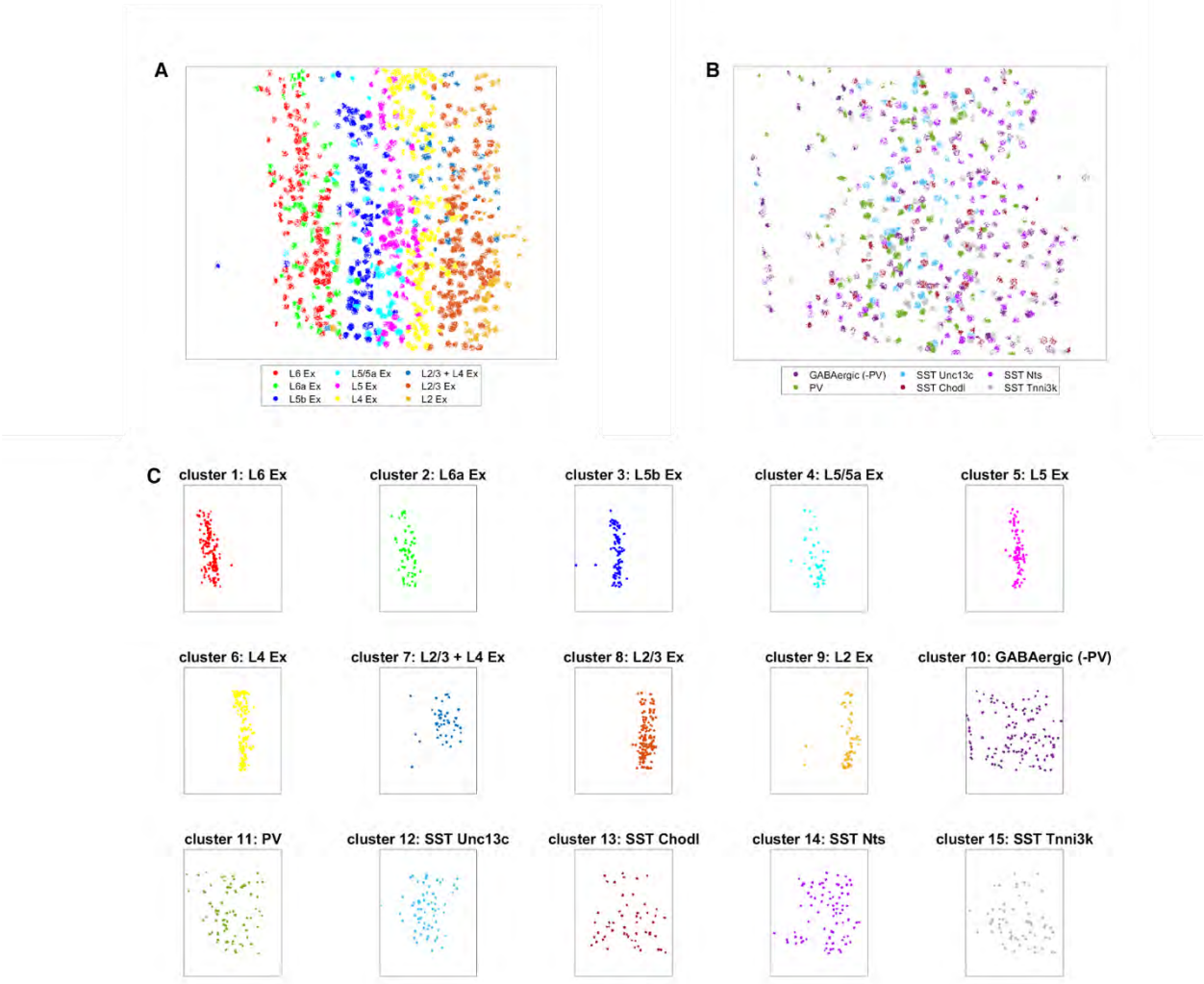


Fig. S16.

Spatial organization of clusters in the visual cortex. (A) The reads corresponding to the excitatory neuron clusters are colored by their cluster identity and plotted in their spatial locations in the visual cortex slice. (B) Reads corresponding to inhibitory neurons are colored by their cluster identity and plotted in their spatial locations in the visual cortex slice. (C) Reads within cells of each cluster are shown independently, plotted in their spatial locations in the visual cortex slice.

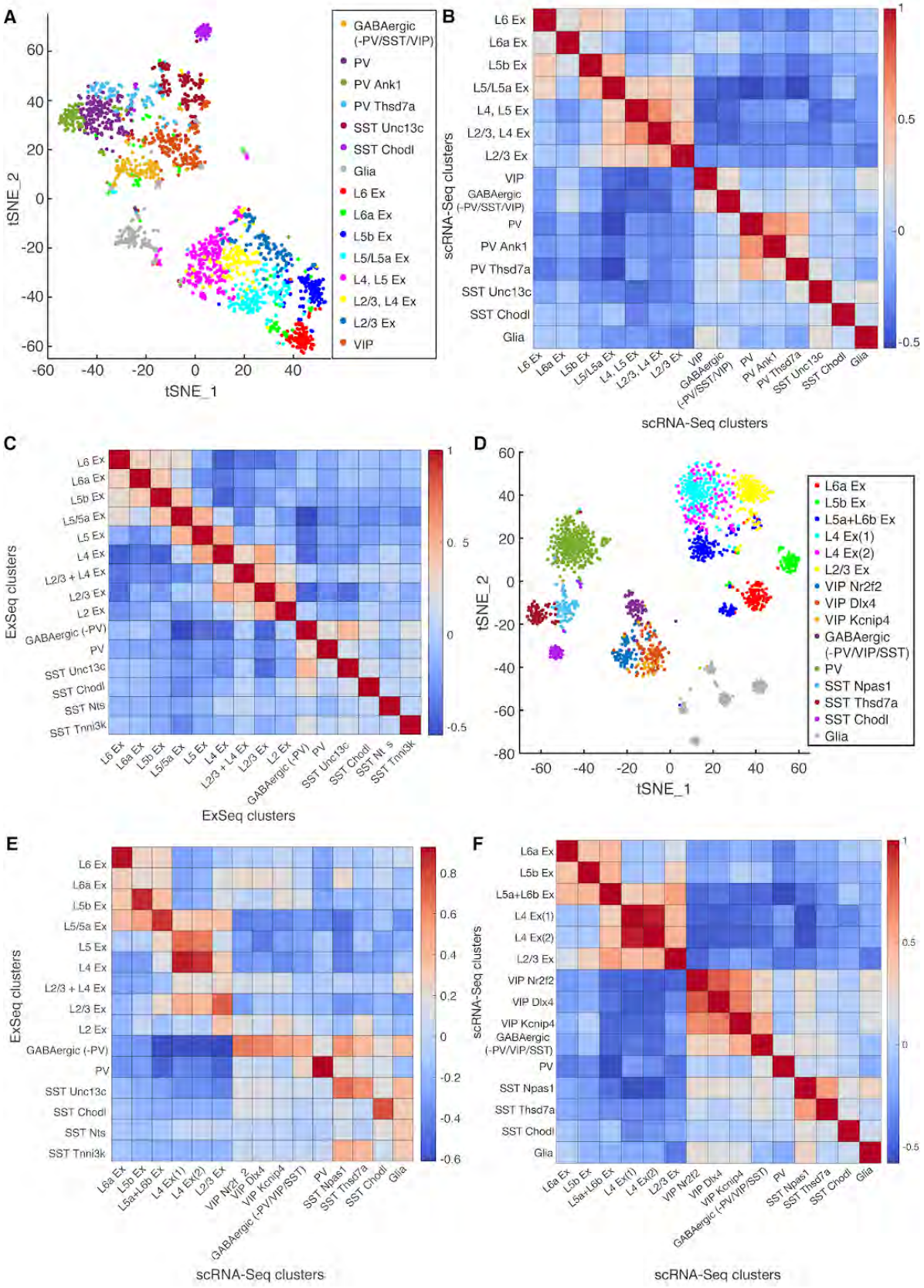

A
tSNE_2
tSNE_1
GABAergic (-PV/SST/VIP)
PV
PV Ank1
PV Thsd7a
SST Unc13c
SST Chodl
Glia
L6 Ex
L6a Ex
L5b Ex
L5/L5a Ex
L4, L5 Ex
L2/3, L4 Ex
L2/3 Ex
VIP
B
scRNA-Seq clusters
C
ExSeq clusters
L5 Ex
L4 Ex
L2/3 + L4 Ex
L2 Ex
SST Nts
SST Tnni3k
D
L6a Ex
L5b Ex
L5a+L6b Ex
L4 Ex(1)
L4 Ex(2)
VIP Nr2f2
VIP Dlx4
VIP Kcnip4
GABAergic (-PV/VIP/SST)
SST Npas1
SST Thsd7a
E
F

Fig. S17.

Additional validation of single-cell clustering approach. (A) t-SNE plot of the scRNA-Seq single-cell dataset used in the clustering analysis (*61*), restricted to the 42 genes utilized for the targeted ExSeq of visual cortex. (B) Heatmap showing Pearson's correlation between pairs of single-cell clusters from A. (C) Heatmap showing Pearson's correlation between pairs of ExSeq clusters. (D) t-SNE plot of the scRNA-Seq single-cell dataset used in the clustering analysis, clustered using variable genes (see **Methods**). (E) Heatmap showing Pearson's correlation between ExSeq clusters and single-cell clusters generated using variable genes. (F) Heatmap showing Pearson's correlation between pairs of scRNA-Seq single-cell clusters, when they are clustered using highly variable genes, in D.

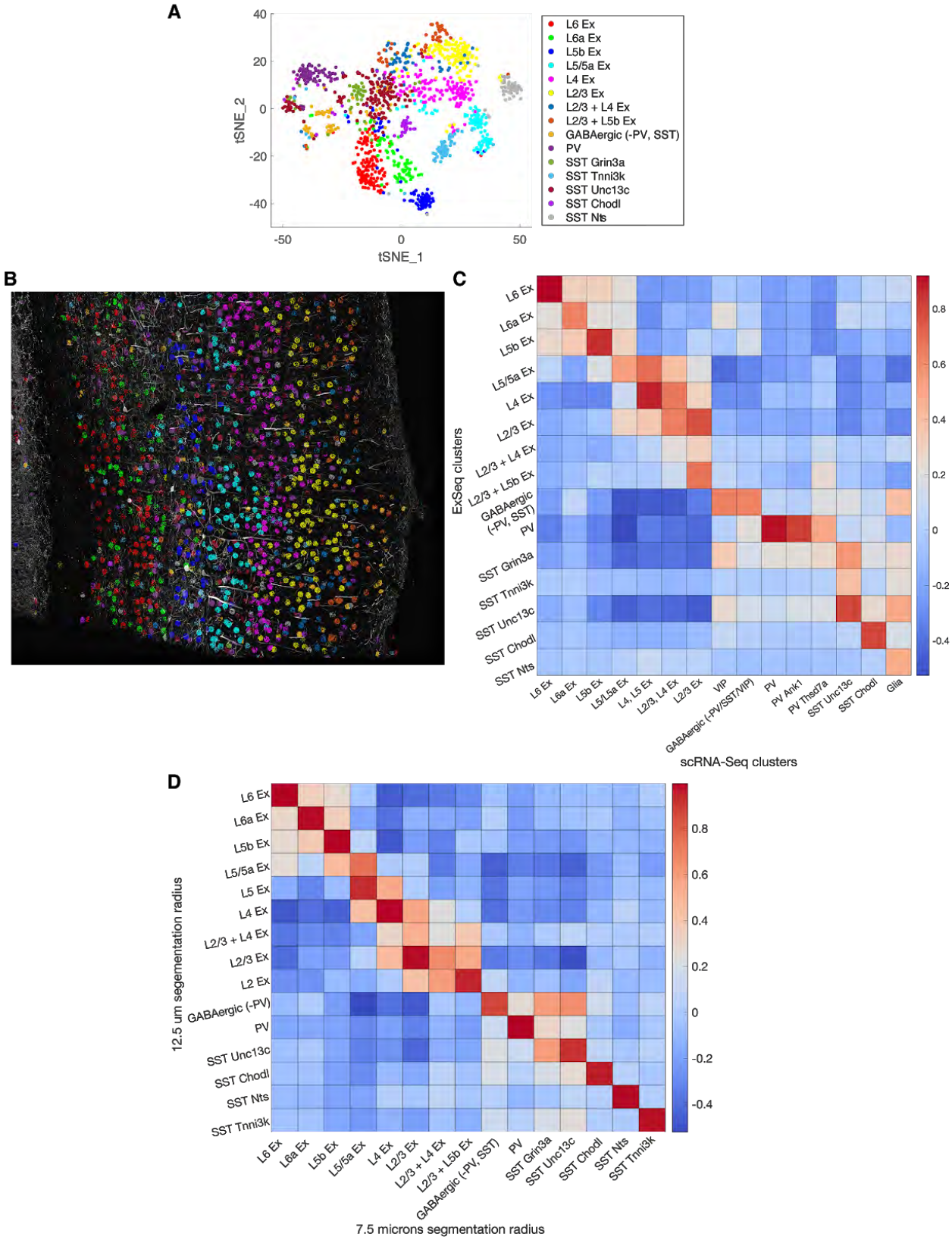

A
tSNE_2
tSNE_1
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L4 Ex
L2/3 Ex
L2/3 + L4 Ex
L2/3 + L5b Ex
GABAergic (-PV, SST)
PV
SST Grin3a
SST Tnni3k
SST Unc13c
SST Chodl
SST Nts
B
C
ExSeq clusters
scRNA-Seq clusters
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L4 Ex
L2/3 Ex
L2/3 + L4 Ex
L2/3 + L5b Ex
GABAergic (-PV, SST)
PV
SST Grin3a
SST Tnni3k
SST Unc13c
SST Chodl
SST Nts
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L4, L5 Ex
L2/3, L4 Ex
L2/3 Ex
VIP
GABAergic (-PV/SST/VIP)
PV
PV Ank1
PV Thsd7a
SST Unc13c
SST Chodl
Glia
D
12.5 um segmentation radius
7.5 microns segmentation radius
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L5 Ex
L4 Ex
L2/3 + L4 Ex
L2/3 Ex
L2 Ex
GABAergic (-PV)
PV
SST Unc13c
SST Chodl
SST Nts
SST Tnni3k
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L4 Ex
L2/3 Ex
L2/3 + L4 Ex
L2/3 + L5b Ex
GABAergic (-PV, SST)
PV
SST Grin3a
SST Unc13c
SST Chodl
SST Nts
SST Tnni3k

Fig. S18.

Robustness of cell segmentation. (A) t-SNE plot of targeted ExSeq gene expression profiles when cells were processed with a segmentation radius of 7.5 microns (compared to 12.5 microns in **Fig. 4D**). (B) Spatial organization of cell types identified in (A). Cell-segmented reads are shown, colored by cluster assignment, and overlaid on the YFP morphology (white). (C) Heatmap of Pearson's correlation between targeted ExSeq clusters with physical radius 7.5 microns, as in A, and clusters identified in the single-cell RNA-Seq dataset (*61*). (D) Heatmap of Pearson's correlation between targeted ExSeq clusters using a physical radius of 12.5 microns for cell segmentation (as in **Fig. 4D-G**) and targeted ExSeq clusters using a radius of 7.5 microns for cell segmentation, as in A.

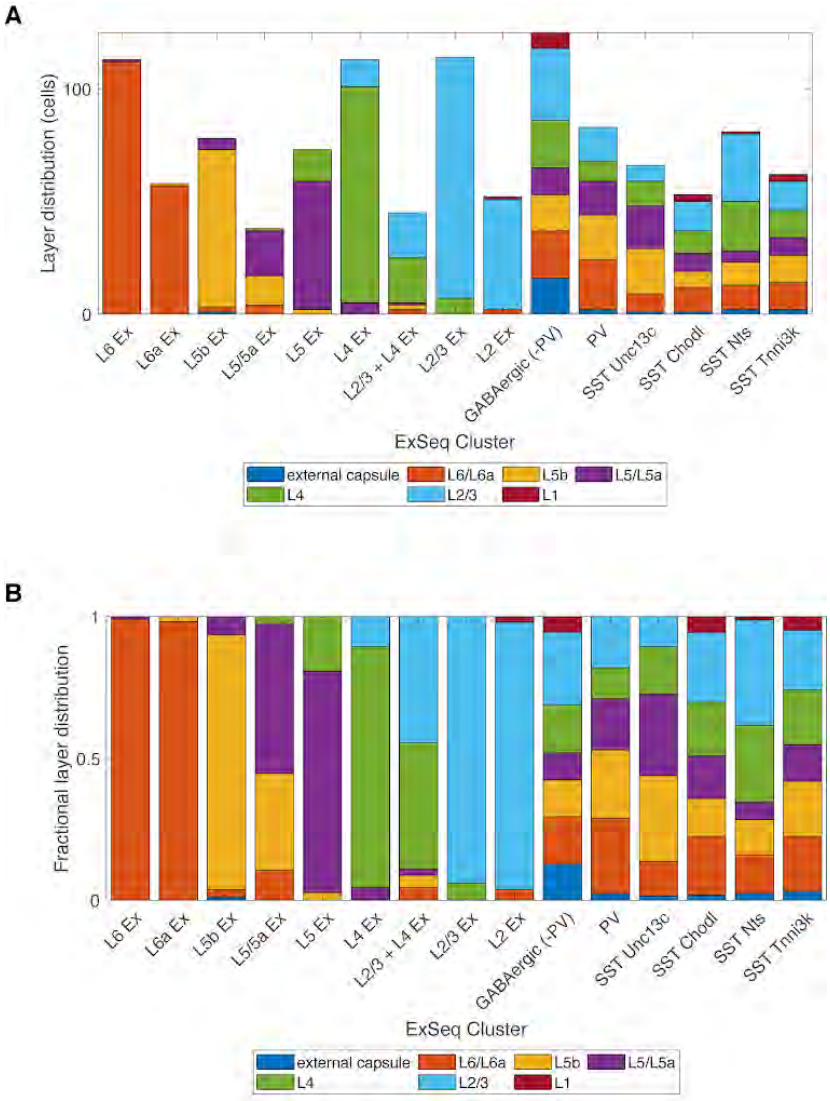

A
Layer distribution (cells)
100
0
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L5 Ex
L4 Ex
L2/3 + L4 Ex
L2/3 Ex
L2 Ex
GABAergic (-PV)
PV
SST Unc13c
SST Chodl
SST Nts
SST Tnni3k
ExSeq Cluster
external capsule
L6/L6a
L5b
L5/L5a
L4
L2/3
L1
B
Fractional layer distribution
1
0.5
0
ExSeq Cluster
C
Cell type counts
300
200
100
0
external capsule
L6/L6a
L5b
L5/L5a
L4
L2/3
L1
Physical layer
L6 Ex
L6a Ex
L5b Ex
L5/5a Ex
L5 Ex
L4 Ex
L2/3 + L4 Ex
L2/3 Ex
L2 Ex
GABAergic (-PV)
PV
SST Unc13c
SST Chodl
SST Nts
SST Tnni3k

Fig. S19.

Organization of cell type clusters into layers. (A) Physical layer distribution of clusters by absolute number of cells. (B) Fractional layer distribution of clusters. (C) Cell type counts by physical layer. Whereas the excitatory neuron clusters are each preferentially found in one layer (not surprisingly, because these clusters define the layers), the inhibitory neuron clusters (‘GABAergic (-PV)’, ’PV’, ’SST Unc13c’, ’SST Chodl’, ‘SST Nts’, ‘SST Tnni3k’) are more evenly distributed. For example, cells in the “GABAergic (-PV)” cluster were relatively evenly distributed across layers L2-L6.

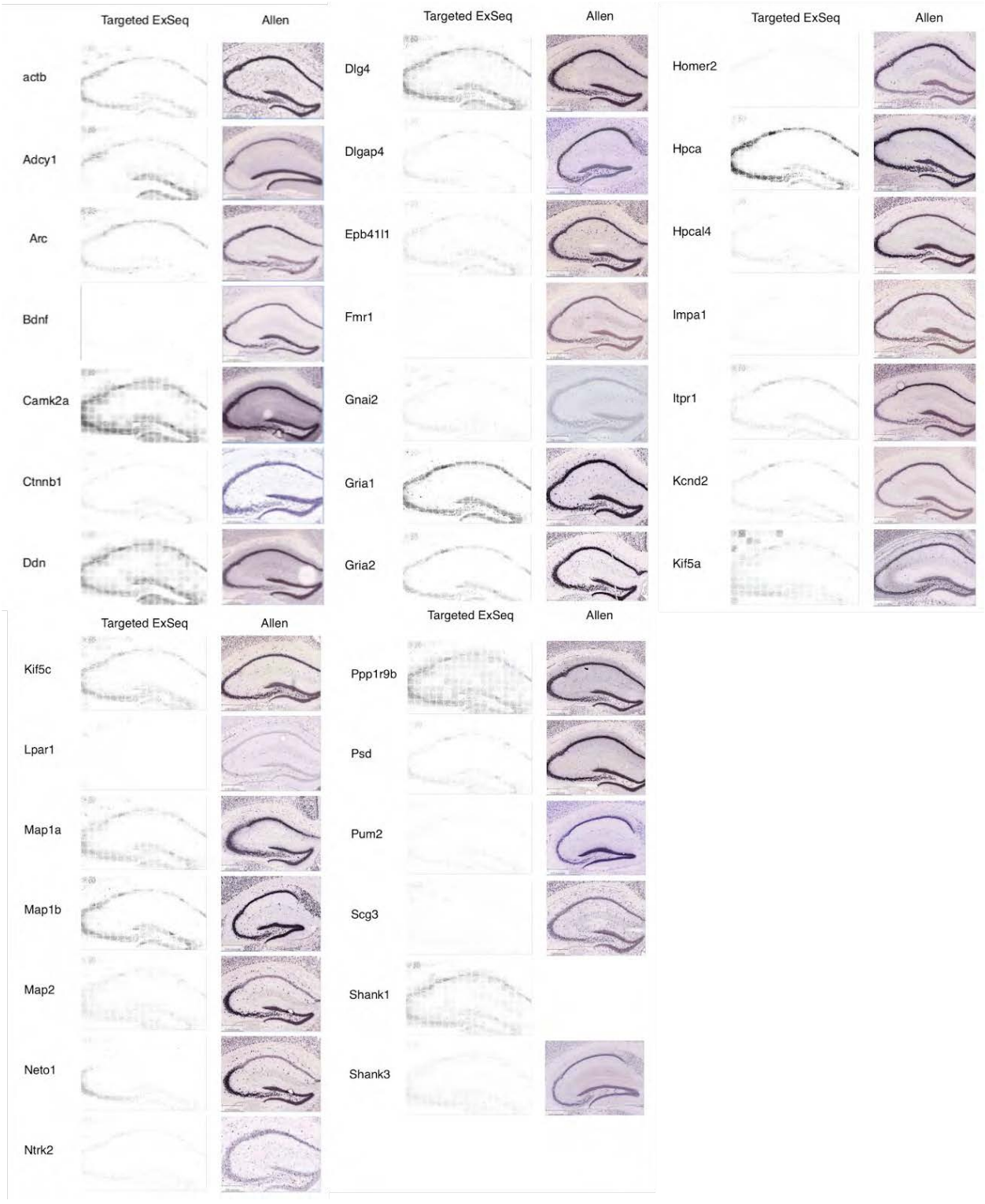

Targeted ExSeq
Allen
actb
Adcy1
Arc
Bdnf
Camk2a
Ctnnb1
Ddn
Dlg4
Dlgap4
Epb41l1
Fmr1
Gnai2
Gria1
Gria2
Homer2
Hpca
Hpcal4
Impa1
Itpr1
Kcnd2
Kif5a
Kif5c
Lpar1
Map1a
Map1b
Map2
Neto1
Ntrk2
Ppp1r9b
Psd
Pum2
Scg3
Shank1
Shank3

Fig. S20.

Gene expression patterns in the mouse hippocampus identified via targeted ExSeq are similar to those observed in the Allen Institute Brain Atlas (https://mouse.brain-map.org/)(*63*). These series of images show transcript localization in the hippocampus via targeted ExSeq (left), comparing them to their respective coronal slice images in the Allen Institute Brain Atlas *in situ* hybridization dataset (right).

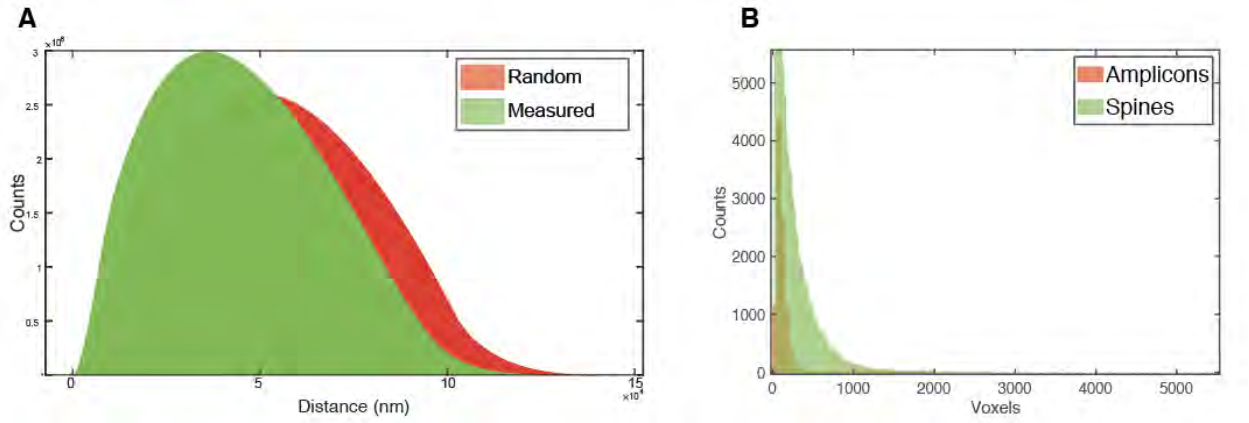


Fig. S21.

Analysis of amplicon formation density and volume, in the hippocampus targeted ExSeq dataset. (A) Histogram showing measured pairwise distances between targeted ExSeq amplicons in the hippocampal dataset (green) (associated with **Fig. 5**), and histogram showing the distribution of pairwise distances when the same amplicons were randomly placed within the same volume (red). The inability of amplicons to form close to one another (e.g., if forming one amplicon "laterally inhibits" the formation of a nearby amplicon) would have resulted in the measured pairwise distances tapering off quickly closer to 0, which we do not see here. (B) Histograms showing the distribution of volumes for spine heads (green) in the hippocampal dataset (**Fig. 5**) and amplicons (red). Based on these volumes, it is possible for multiple amplicons to occupy the same spine head, indicating that our observation of one amplicon per spine reflects the infrequent presence of the transcripts studied here within spines.

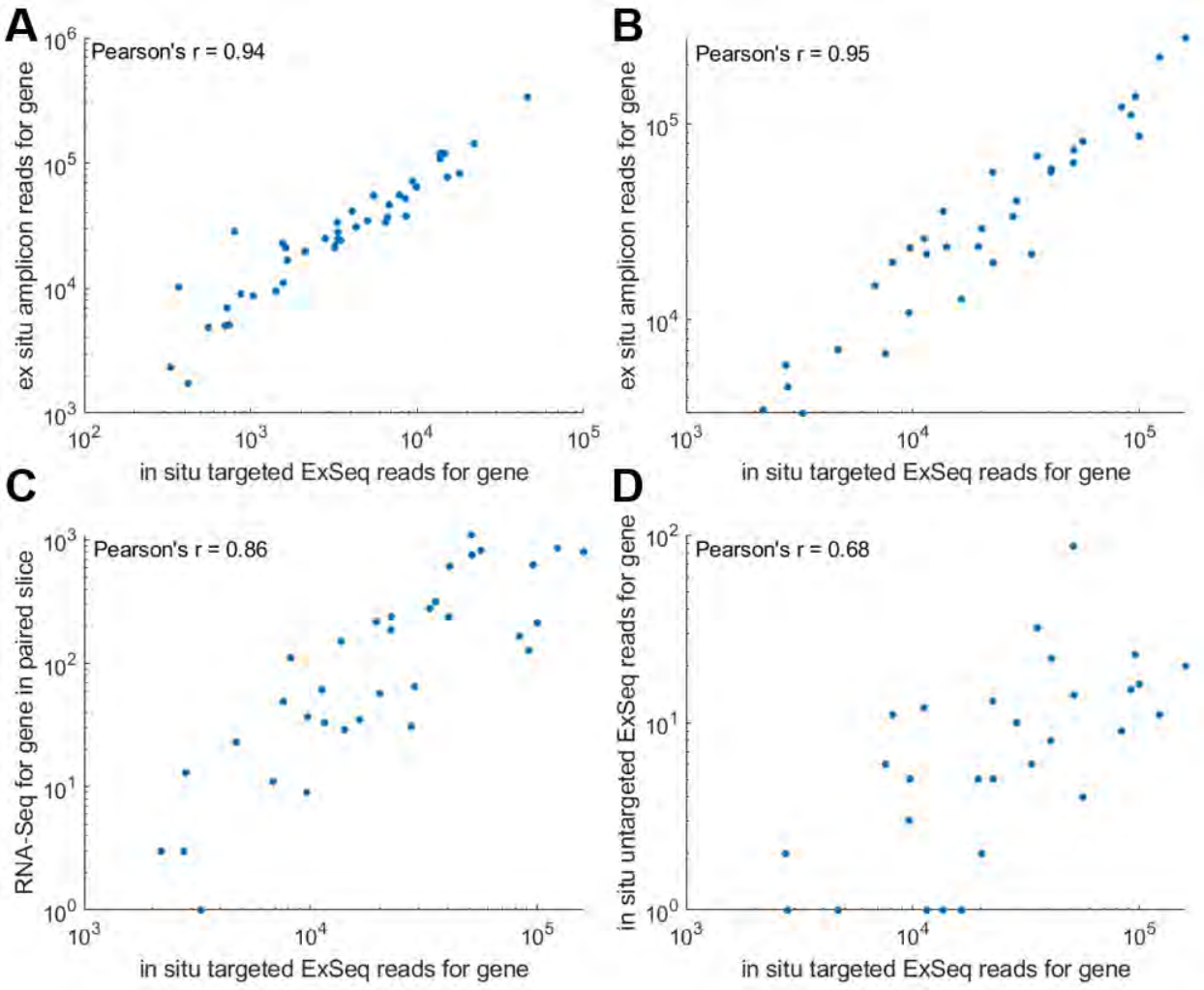


Fig. S22.

Targeted ExSeq reads are highly correlated with *ex situ* amplicon sequencing, RNAseq, and untargeted ExSeq datasets. (A) Scatterplot showing targeted ExSeq counts for transcripts studied in **Fig. 4** (mouse visual cortex tissue) along with their abundance upon *ex situ* extraction and barcode sequencing of amplicons from the same sample (see **Methods** section 'Comparison of targeted ExSeq to *ex situ* sequencing and bulk RNAseq'). Each point is a gene among the 42 genes examined in the visual cortex dataset (Pearson's r = 0.94, p-value = 8.68×10^{-21}; counts in **Table S12**). (B) Scatterplot showing targeted ExSeq counts for transcripts studied in **Fig. 5** (mouse hippocampus tissue) along with their abundance upon *ex situ* extraction and barcode sequencing of amplicons from the same sample. Each point is a gene among the 34 genes examined in the hippocampal dataset (Pearson's r = 0.95, p-value = 2.85×10^{-18}; counts in **Table S13**). (C) Scatter plot showing targeted ExSeq counts for transcripts studied in **Fig. 5** against their expression with RNAseq using a paired (i.e., adjacent) hippocampal slice (Pearson's r = 0.86, p-value = 9.38×10^{-11}; counts in **Table S13**) . (D) Scatter plot showing targeted vs untargeted ExSeq counts for the genes studied in **Fig. 5** in the mouse hippocampus, taken from two different mice but from identical coordinates (Pearson's r = 0.68, p-value = 9.75×10^{-6}; counts in **Table S13**).

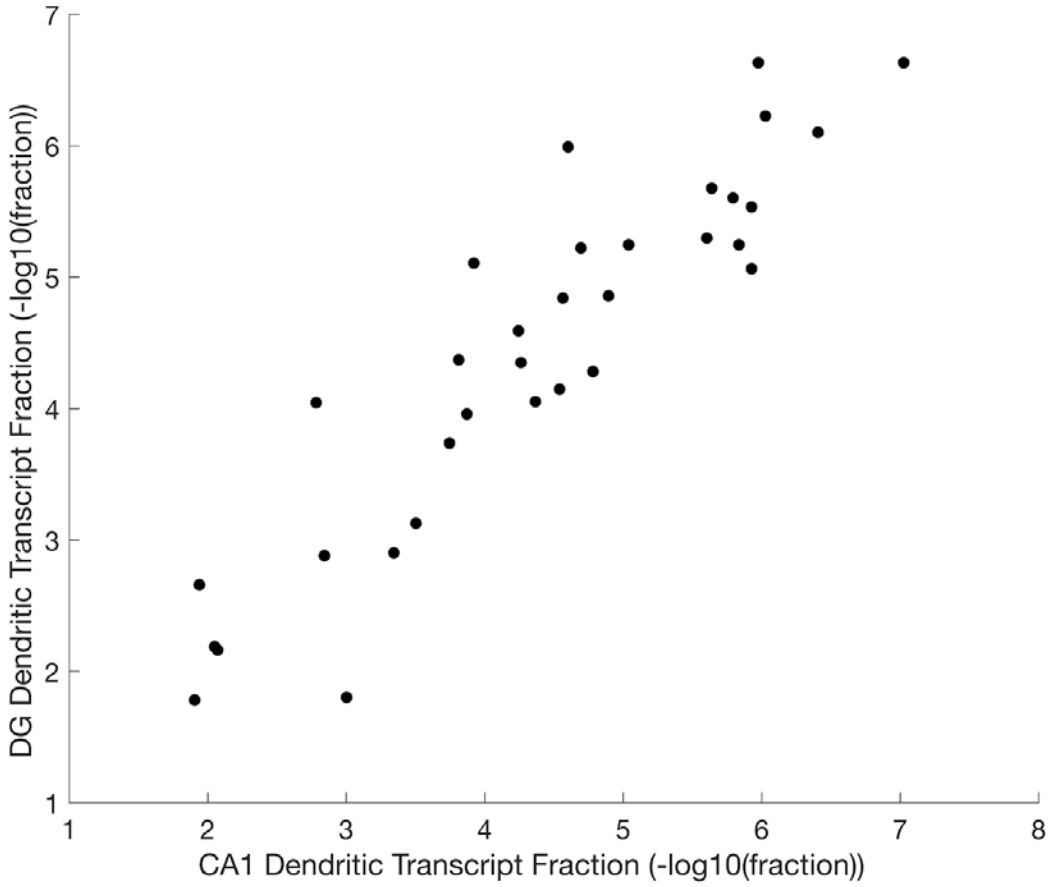


Fig. S23.
High correlation between the dendritic expression of genes for granule cells of the dentate gyrus (DG) vs. pyramidal neurons of CA1 (Pearson's r = 0.91; p-value 1.2×10^{-13}). Scatter plot showing the 'dendritic transcript fraction', defined as the number of reads of a given gene that are located in the dendrites divided by the total number of reads found in dendrites (for all examined genes combined), in the dentate gyrus vs CA1 for the genes studied in **Fig. 5** using targeted ExSeq.

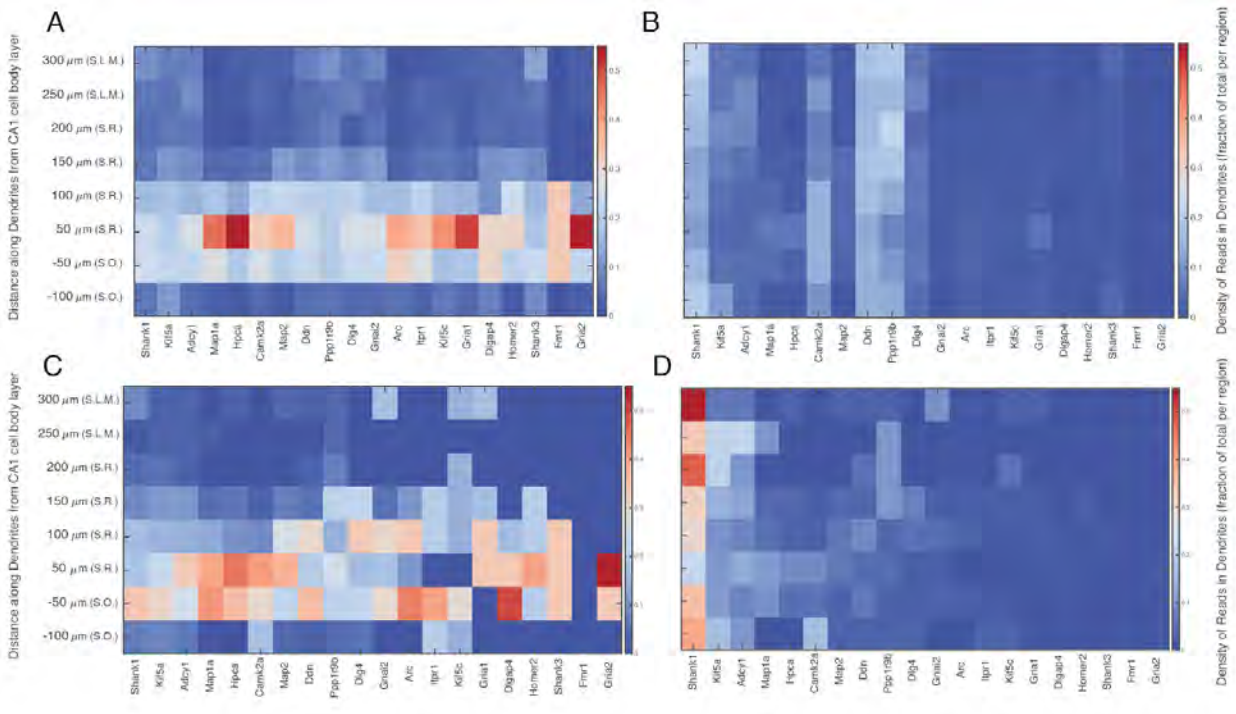


Fig. S24.

Normalized dendrite and spine reads densities in mouse hippocampus. Heatmaps showing the density of transcripts in the dendrites (A,B) and spines (C,D) of CA1 pyramidal neurons corresponding to **Fig. 5E** along the apical-basal axis (Euclidean distance) of hippocampal area CA1, normalized by either the total transcript counts per gene (for A and C), or by the total transcript counts at a given distance from the cell body layer (for B and D).

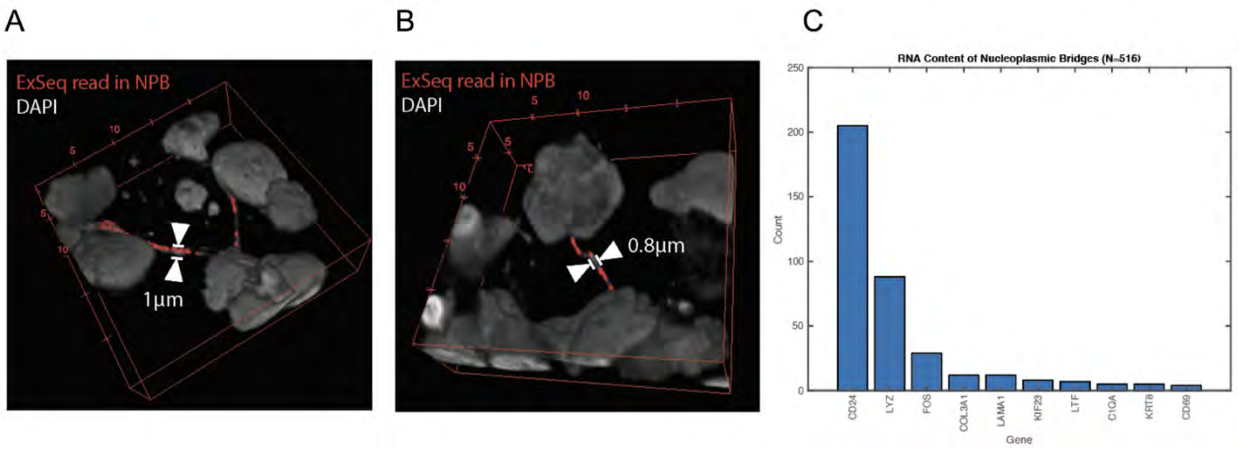
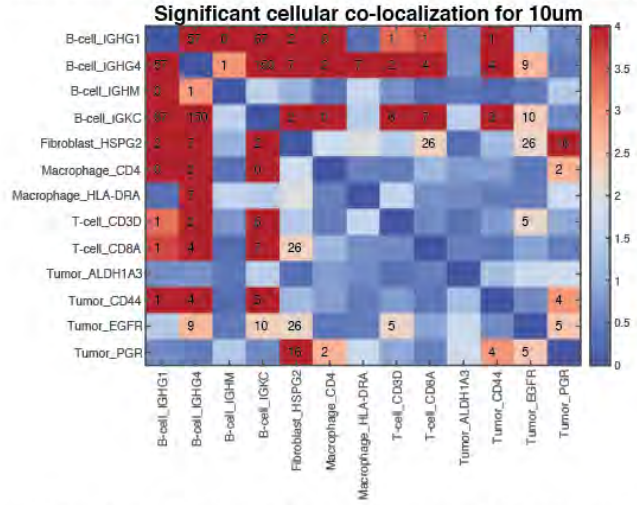


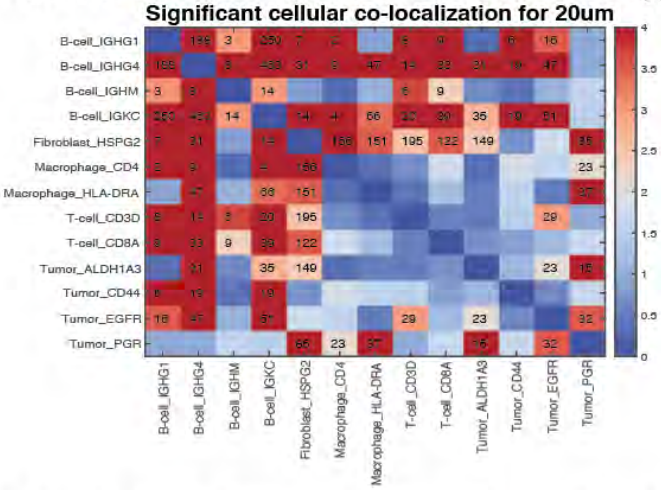
Fig. S25.

Targeted ExSeq of a human metastatic breast tumor biopsy reveals transcripts within nuclear structures (possibly nucleoplasmic bridges) that are under one micron in size. (A,B) 3D renders showing ExSeq reads (red) localized within the volumes of the nuclear structures overlaid on a confocal image of DAPI nuclear staining. Arrows show the nuclear structures and their width. (C) Bar plot showing the abundance of genes localized within the nuclear structures. A total of 516 reads were identified in 83 nuclear structures. We note that the relatively small number of reads obtained in these nanoscale regions limited our ability to systematically classify these reads into specific cell types.

A 854 cell-cell interactions between different subtypes (200pix threshold)



B 4970 cell-cell interactions between different subtypes (400pix threshold)



C 17,988 cell-cell interactions between different subtypes (800pix threshold)

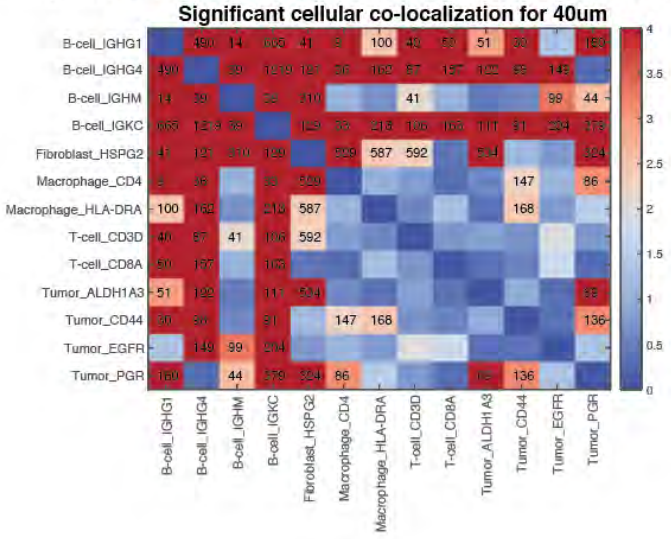


Fig. S26.

The cell clusters in metastatic breast cancer tissue exhibit robust, non-random spatial colocalizations. The adjacency matrix text values indicate the number of cell pairs across different clusters that are in close proximity as determined by a Euclidean distance threshold between cell centroids of 10 (A), 20 (B) and 40 (C) microns. The adjacency matrix heatmap shows the p-value (10,000 bootstrapping iterations) relative to obtaining the same or higher number of cells in close proximity by chance. All adjacency matrix entries with text values are statistically significant (Benjamini Hochberg false-discovery rate of 1.5%). Note that the structure of the matrix is preserved across the different thresholds.

Tables S1-S6 are in an online excel spreadsheet.

Table S1.
Sequencing primers for untargeted ExSeq.

Table S2.
Sequencing read counts in untargeted ExSeq of hippocampus (50 microns).

Table S3.
Sequencing read counts in untargeted ExSeq of hippocampus (15 microns).

Table S4.
Sequencing read counts in untargeted ExSeq of neuronal culture.

Table S5.
Summary of ExSeq experimental details for Figs. 3-6.

Table S6.
DAVID Gene Ontology term enrichment for untargeted ExSeq and RNAseq.

Technology	Maximum tissue thickness demonstrated[1]	Resolution demonstrated in tissue	Axial optical sectioning[2]	Multiplexing with antibody stains shown?[3]	Demonstrated in human tissue?[1]	Ref.
Targeted ExSeq	50 microns[4]	100 nm x 100 nm x 300 nm[5]	0.12 microns/ section	yes	yes	This paper
STARMAP	8 microns	diffraction limited	0.3 microns/ section	no	no	(*10*)
Nilsson ISS	10 microns	diffraction limited	1.5 microns /section	no	yes	(11, *34*, *127*)
BOLORAMIS	N/A[6]	diffraction limited	0.3 microns /section	no	no	(*57*)
osmFISH	10 microns	diffraction limited	0.3 microns/ section	yes	no	(*128*)
MERFISH[7]	10 microns	diffraction limited	1.5 microns/ section	yes	no	(*8*, *27*, *129*, *130*)
SeqFISH	15 microns	diffraction limited	unreported	no	no	(*9*)
SeqFISH+	5 microns	diffraction limited[8]	5 microns /section	no	no	(*20*)
Spatial transcriptomics (ST)	16 microns	100 micron spacing of RNA capture spots	N/A	no	yes	(*74*)
High Definition ST (HDST)	16 microns	"single-cell like"[9]	N/A	no	yes	(*131*)
Slide-Seq	10 microns	10 micron beads	N/A	no	no	(*75*)

Table S7.

Evaluation of highly multiplexed targeted transcriptomic technologies for various criteria. Highlighted in green are the technologies with the best result for each examined criterion. Not shown in this table are untargeted ExSeq, FISSEQ (*25*) and INSTA-seq (*53*) which are

untargeted technologies (as defined by their use of random primers to initialize their library preparation for *in situ* sequencing).

[1]Demonstrated for multiplexed, single-molecule-identified experiments.
[2]Z-step length between adjacent optical sections (in pre-expansion coordinates, if applicable). To fully reconstruct volumetric information, axial sectioning must be less than half of Z-resolution (Nyquist criterion).
[3]Demonstrated with at least one antibody and multiple RNA targets in the same experiment.
[4]Shown in Fig. 3 using untargeted ExSeq.
[5]Pre-expansion units (diffraction limit divided by 3.3x expansion factor).
[6]The authors demonstrate library preparation of an intact tissue specimen but do not perform a multiplexed experiment on the intact tissue.
[7]Expansion MERFISH allows 2X expansion in cultured cells, however, it was not demonstrated in tissues and therefore is not listed in this table.
[8]Although seqFISH+ allows high resolution localization of RNA molecules using Gaussian centroid fitting and rounds of barcode stratification, the detailed tissue context, i.e. the protein and morphological information, is not nanoscale-resolved.
[9]The stated resolution in High Definition ST is 2 microns, and it refers to the size of the wells containing oligonucleotides; however, the data from neighboring wells is then binned to regions in space which according to the authors are “single-cell like” in size.

	Yield (percent of molecules detected compared to smFISH)
Targeted ExSeq	62% in cultured HeLa cells[1]
STARMAP	not reported[2]
Nilsson ISS	estimated at ~5%[3]
BOLORAMIS	not reported
osmFISH	Estimated to be similar to smFISH in mouse brain tissue[4]
MERFISH	Measured ~21-94% in U-2 OS cells
Expansion MERFISH	This technology was not demonstrated in tissues yet. Measured ~105% in expanded U-2 OS cells[5]
SeqFISH	Measured 71-84% in mouse brain tissue[6]
SeqFISH+	Measured 49% in cultured NIH/3T3 cells[7]
Spatial transcriptomics	6.9% in mouse brain tissue[8]
High Definition Spatial Transcriptomics (HDST)	1.3% in mouse brain tissue[9]
SlideSeq	1% in mouse brain tissue[10]

Table S8.

Comparison of yield, i.e. percent of molecules detected compared to single-molecule FISH (smFISH), and methods of determining yield for highly multiplexed, targeted, RNA localization technologies.

[1]Yield determined by performing HCR-ExFISH, followed by targeted ExSeq, for the same gene in the same expanded HeLa cells, and comparing HCR-ExFISH spot counts with targeted ExSeq spot counts.
[2]Estimated to be comparable to scRNA-Seq by Rank-Sum test average expression of 151 genes across cells in mouse brain tissue; this estimation was done by computing mean counts per cell for genes targeted with STARMAP, and with scRNA-Seq of similar brain region, and showing similarity of overall distribution. No calculation of the yield (as defined here) was performed.
[3]Not directly measured against smFISH; yield estimated to be 5% in (*6*).
[4]No direct measurement was performed.
[5]Yield determined by comparing mean expansion MERFISH spot counts to mean smFISH spot counts in non-expanded U-2 OS cells. The value is greater than 100% due to decrowding of high-expression genes. Expansion MERFISH has not been demonstrated in tissue.
[6]Yield determined by sequentially performing SeqFISH, followed by HCR-amplified smFISH for a subset of genes in the same cells in the same section, and comparing spot counts (71% - 84%).
[7]Yield determined by comparing mean SeqFISH+ spot counts to mean smFISH spot counts in NIH/3T3 cells.

[8]Yield determined by performing smFISH in a paired mouse brain slice, and comparing spatially binned spot counts to spatial transcriptomics detection events.
[9]Yield determined by using the smFISH dataset from (*74*) on a similar mouse brain slice, and comparing spatially binned spot counts to HDST detection events.
[10]Yield determined by performing HCR-amplified smFISH in a paired mouse brain section, and comparing spot counts to Slide-Seq detection events.

Tables S9-S14 are in an online excel spreadsheet.

Table S9.
Genes, accession numbers, and barcodes used with targeted ExSeq.

Table S10.
Targeted ExSeq probe sequences.

Table S11.
ExFISH and targeted ExSeq comparison data.

Table S12.
Targeted ExSeq reads in the visual cortex.

Table S13.
Targeted ExSeq reads in the hippocampus.

Table S14.
Targeted ExSeq reads in human cancer.

References and Notes

1. N. Crosetto, M. Bienko, A. van Oudenaarden, Spatially resolved transcriptomics and beyond. *Nat. Rev. Genet.* **16**, 57–66 (2015). doi:10.1038/nrg3832 Medline

2. T. Gregor, H. G. Garcia, S. C. Little, The embryo as a laboratory: Quantifying transcription in *Drosophila*. *Trends Genet.* **30**, 364–375 (2014). doi:10.1016/j.tig.2014.06.002 Medline

3. S. Alon, G. H. Huynh, E. S. Boyden, Expansion microscopy: Enabling single cell analysis in intact biological systems. *FEBS J.* **286**, 1482–1494 (2019). doi:10.1111/febs.14597 Medline

4. A. E. Moor, M. Golan, E. E. Massasa, D. Lemze, T. Weizman, R. Shenhav, S. Baydatch, O. Mizrahi, R. Winkler, O. Golani, N. Stern-Ginossar, S. Itzkovitz, Global mRNA polarization regulates translation efficiency in the intestinal epithelium. *Science* **357**, 1299–1303 (2017). doi:10.1126/science.aan2399 Medline

5. K. B. Halpern, R. Shenhav, O. Matcovitch-Natan, B. Tóth, D. Lemze, M. Golan, E. E. Massasa, S. Baydatch, S. Landen, A. E. Moor, A. Brandis, A. Giladi, A. S. Avihail, E. David, I. Amit, S. Itzkovitz, Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017). doi:10.1038/nature21065 Medline

6. E. Lein, L. E. Borm, S. Linnarsson, The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science* **358**, 64–69 (2017). doi:10.1126/science.aan6827 Medline

7. A. Regev, S. A. Teichmann, E. S. Lander, I. Amit, C. Benoist, E. Birney, B. Bodenmiller, P. Campbell, P. Carninci, M. Clatworthy, H. Clevers, B. Deplancke, I. Dunham, J. Eberwine, R. Eils, W. Enard, A. Farmer, L. Fugger, B. Göttgens, N. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg, J. Lundeberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea, G. Nolan, D. Pe'er, A. Phillipakis, C. P. Ponting, S. Quake, W. Reik, O. Rozenblatt-Rosen, J. Sanes, R. Satija, T. N. Schumacher, A. Shalek, E. Shapiro, P. Sharma, J. W. Shin, O. Stegle, M. Stratton, M. J. T. Stubbington, F. J. Theis, M. Uhlen, A. van Oudenaarden, A. Wagner, F. Watt, J. Weissman, B. Wold, R. Xavier, N. Yosef, Human Cell Atlas Meeting Participants, The Human Cell Atlas. *eLife* **6**, e27041 (2017). doi:10.7554/eLife.27041 Medline

8. J. R. Moffitt, D. Bambah-Mukku, S. W. Eichhorn, E. Vaughn, K. Shekhar, J. D. Perez, N. D. Rubinstein, J. Hao, A. Regev, C. Dulac, X. Zhuang, Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* **362**, eaau5324 (2018). doi:10.1126/science.aau5324 Medline

9. S. Shah, E. Lubeck, W. Zhou, L. Cai, In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* **92**, 342–357 (2016). doi:10.1016/j.neuron.2016.10.001 Medline

10. X. Wang, W. E. Allen, M. A. Wright, E. L. Sylwestrak, N. Samusik, S. Vesuna, K. Evans, C. Liu, C. Ramakrishnan, J. Liu, G. P. Nolan, F.-A. Bava, K. Deisseroth, Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018). doi:10.1126/science.aat5691 Medline

11. R. Ke, M. Mignardi, A. Pacureanu, J. Svedlund, J. Botling, C. Wählby, M. Nilsson, In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013). doi:10.1038/nmeth.2563 Medline

12. A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998). doi:10.1126/science.280.5363.585 Medline

13. J. M. Levsky, S. M. Shenoy, R. C. Pezo, R. H. Singer, Single-cell gene expression profiling. *Science* **297**, 836–840 (2002). doi:10.1126/science.1072241 Medline

14. D.-M. Franchini, O. Lanvin, M. Tosolini, E. Patras de Campaigno, A. Cammas, S. Péricart, C.-M. Scarlata, M. Lebras, C. Rossi, L. Ligat, F. Pont, P. B. Arimondo, C. Laurent, M. Ayyoub, F. Despas, M. Lapeyre-Mestre, S. Millevoi, J.-J. Fournié, Microtubule-Driven Stress Granule Dynamics Regulate Inhibitory Immune Checkpoint Expression in T Cells. *Cell Rep.* **26**, 94–107.e7 (2019). doi:10.1016/j.celrep.2018.12.014 Medline

15. V. Balagopal, R. Parker, Polysomes, P bodies and stress granules: States and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.* **21**, 403–408 (2009). doi:10.1016/j.ceb.2009.03.005 Medline

16. C. E. Holt, K. C. Martin, E. M. Schuman, Local translation in neurons: Visualization and function. *Nat. Struct. Mol. Biol.* **26**, 557–566 (2019). doi:10.1038/s41594-019-0263-5 Medline

17. E. M. Schuman, mRNA trafficking and local protein synthesis at the synapse. *Neuron* **23**, 645–648 (1999). doi:10.1016/S0896-6273(01)80023-4 Medline

18. A. J. Rodriguez, K. Czaplinski, J. S. Condeelis, R. H. Singer, Mechanisms and cellular roles of local protein synthesis in mammalian cells. *Curr. Opin. Cell Biol.* **20**, 144–149 (2008). doi:10.1016/j.ceb.2008.02.004 Medline

19. A.-S. Hafner, P. G. Donlin-Asp, B. Leitch, E. Herzog, E. M. Schuman, Local protein synthesis is a ubiquitous feature of neuronal pre- and postsynaptic compartments. *Science* **364**, eaau3644 (2019). doi:10.1126/science.aau3644 Medline

20. C. L. Eng, M. Lawson, Q. Zhu, R. Dries, N. Koulena, Y. Takei, J. Yun, C. Cronin, C. Karp, G.-C. Yuan, L. Cai, Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019). doi:10.1038/s41586-019-1049-y Medline

21. J. T. Morgan, G. R. Fink, D. P. Bartel, Excised linear introns regulate growth in yeast. *Nature* **565**, 606–611 (2019). doi:10.1038/s41586-018-0828-1 Medline

22. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, J. L. Yang, T. C. Ferrante, R. Terry, S. S. F. Jeanty, C. Li, R. Amamoto, D. T. Peters, B. M. Turczyk, A. H. Marblestone, S. A. Inverso, A. Bernard, P. Mali, X. Rios, J. Aach, G. M. Church, Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi:10.1126/science.1250212 Medline

23. F. Chen, P. W. Tillberg, E. S. Boyden, Expansion microscopy. *Science* **347**, 543–548 (2015). doi:10.1126/science.1260088 Medline

24. A. T. Wassie, Y. Zhao, E. S. Boyden, Expansion microscopy: Principles and uses in biological research. *Nat. Methods* **16**, 33–41 (2019). doi:10.1038/s41592-018-0219-4 Medline

25. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, G. M. Church, Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015). doi:10.1038/nprot.2014.191 Medline

26. F. Chen, A. T. Wassie, A. J. Cote, A. Sinha, S. Alon, S. Asano, E. R. Daugharthy, J.-B. Chang, A. Marblestone, G. M. Church, A. Raj, E. S. Boyden, Nanoscale imaging of RNA with expansion microscopy. *Nat. Methods* **13**, 679–684 (2016). doi:10.1038/nmeth.3899 Medline

27. G. Wang, J. R. Moffitt, X. Zhuang, Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. *Sci. Rep.* **8**, 4847 (2018). doi:10.1038/s41598-018-22297-7 Medline

28. Detailed materials and methods are available as supplementary materials.

29. A. Kajita, D. Goodwin, R. Prior, dgoodwin208/ExSeqProcessing: First release of the processing software for Expansion Sequencing, version v1.0.0, Zenodo (2020); http://doi.org/10.5281/zenodo.4075515.

30. M. Kircher, P. Heyn, J. Kelso, Addressing challenges in the production and analysis of illumina sequencing data. *BMC Genomics* **12**, 382 (2011). doi:10.1186/1471-2164-12-382 Medline

31. W. Li, J. Freudenberg, P. Miramontes, Diminishing return for increased Mappability with longer sequencing reads: Implications of the k-mer distributions in the human genome. *BMC Bioinformatics* **15**, 2 (2014). doi:10.1186/1471-2105-15-2 Medline

32. P. W. Tillberg, F. Chen, K. D. Piatkevich, Y. Zhao, C.-C. Yu, B. P. English, L. Gao, A. Martorell, H.-J. Suk, F. Yoshida, E. M. DeGennaro, D. H. Roossien, G. Gong, U. Seneviratne, S. R. Tannenbaum, R. Desimone, D. Cai, E. S. Boyden, Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34**, 987–992 (2016). doi:10.1038/nbt.3625 Medline

33. G. Feng, R. H. Mellor, M. Bernstein, C. Keller-Peck, Q. T. Nguyen, M. Wallace, J. M. Nerbonne, J. W. Lichtman, J. R. Sanes, Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000). doi:10.1016/S0896-6273(00)00084-2 Medline

34. J. Svedlund, C. Strell, X. Qian, K. J. C. Zilkens, N. P. Tobin, J. Bergh, A. M. Sieuwerts, M. Nilsson, Generation of in situ sequencing based OncoMaps to spatially resolve gene expression profiles of diagnostic and prognostic markers in breast cancer. *EBioMedicine* **48**, 212–223 (2019). doi:10.1016/j.ebiom.2019.09.009 Medline

35. K. S. Kosik, Life at Low Copy Number: How Dendrites Manage with So Few mRNAs. *Neuron* **92**, 1168–1180 (2016). doi:10.1016/j.neuron.2016.11.002 Medline

36. M. Khaladkar, P. T. Buckley, M. T. Lee, C. Francis, M. M. Eghbal, T. Chuong, S. Suresh, B. Kuhn, J. Eberwine, J. Kim, Subcellular RNA sequencing reveals broad presence of cytoplasmic intron-sequence retaining transcripts in mouse and rat neurons. *PLOS ONE* **8**, e76194 (2013). doi:10.1371/journal.pone.0076194 Medline

37. H. Saini, A. A. Bicknell, S. R. Eddy, M. J. Moore, Free circular introns with an unusual branchpoint in neuronal projections. *eLife* **8**, e47809 (2019). doi:10.7554/eLife.47809 Medline

38. P. T. Buckley, M. T. Lee, J.-Y. Sul, K. Y. Miyashiro, T. J. Bell, S. A. Fisher, J. Kim, J. Eberwine, Cytoplasmic intron sequence-retaining transcripts can be dendritically targeted via ID element retrotransposons. *Neuron* **69**, 877–884 (2011). doi:10.1016/j.neuron.2011.02.028 Medline

39. J. Glanzer, K. Y. Miyashiro, J.-Y. Sul, L. Barrett, B. Belt, P. Haydon, J. Eberwine, RNA splicing capability of live neuronal dendrites. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16859–16864 (2005). doi:10.1073/pnas.0503783102 Medline

40. A. Harada, J. Teng, Y. Takei, K. Oguchi, N. Hirokawa, MAP2 is required for dendrite elongation, PKA anchoring in dendrites, and proper PKA signal transduction. *J. Cell Biol.* **158**, 541–549 (2002). doi:10.1083/jcb.200110134 Medline

41. B. Cubelos, A. Sebastián-Serrano, L. Beccari, M. E. Calcagnotto, E. Cisneros, S. Kim, A. Dopazo, M. Alvarez-Dolado, J. M. Redondo, P. Bovolenta, C. A. Walsh, M. Nieto, Cux1 and Cux2 regulate dendritic branching, spine morphology, and synapses of the upper layer neurons of the cortex. *Neuron* **66**, 523–535 (2010). doi:10.1016/j.neuron.2010.04.038 Medline

42. P. Crino, K. Khodakhah, K. Becker, S. Ginsberg, S. Hemby, J. Eberwine, Presence and phosphorylation of transcription factors in developing dendrites. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2313–2318 (1998). doi:10.1073/pnas.95.5.2313 Medline

43. S. A. Middleton, J. Eberwine, J. Kim, Comprehensive catalog of dendritically localized mRNA isoforms from sub-cellular sequencing of single mouse neurons. *BMC Biol.* **17**, 5 (2019). doi:10.1186/s12915-019-0630-z Medline

44. L. Pancrazi, G. Di Benedetto, L. Colombaioni, G. Della Sala, G. Testa, F. Olimpico, A. Reyes, M. Zeviani, T. Pozzan, M. Costa, Foxg1 localizes to mitochondria and coordinates cell differentiation and bioenergetics. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13910–13915 (2015). doi:10.1073/pnas.1515190112 Medline

45. H. Ueda, K. Sasaki, S. K. Halder, Y. Deguchi, K. Takao, T. Miyakawa, A. Tajima, Prothymosin alpha-deficiency enhances anxiety-like behaviors and impairs learning/memory functions and neurogenesis. *J. Neurochem.* **141**, 124–136 (2017). doi:10.1111/jnc.13963 Medline

46. H. Tiedge, R. T. Fremeau Jr., P. H. Weinstock, O. Arancio, J. Brosius, Dendritic location of neural BC1 RNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2093–2097 (1991). doi:10.1073/pnas.88.6.2093 Medline

47. I. A. Muslimov, G. Banker, J. Brosius, H. Tiedge, Activity-dependent regulation of dendritic BC1 RNA in hippocampal neurons in culture. *J. Cell Biol.* **141**, 1601–1611 (1998). doi:10.1083/jcb.141.7.1601 Medline

48. X. Zhang, M. H. Hamblin, K.-J. Yin, The long noncoding RNA Malat1: Its physiological and pathophysiological functions. *RNA Biol.* **14**, 1705–1714 (2017). doi:10.1080/15476286.2017.1358347 Medline

49. J. A. Briggs, E. J. Wolvetang, J. S. Mattick, J. L. Rinn, G. Barry, Mechanisms of Long Non-coding RNAs in Mammalian Nervous System Development, Plasticity, Disease, and Evolution. *Neuron* **88**, 861–877 (2015). doi:10.1016/j.neuron.2015.09.045 Medline

50. H. Hörtnagl, R. O. Tasan, A. Wieselthaler, E. Kirchmair, W. Sieghart, G. Sperk, Patterns of mRNA and protein expression for 12 GABAA receptor subunits in the mouse brain. *Neuroscience* **236**, 345–372 (2013). doi:10.1016/j.neuroscience.2013.01.008 Medline

51. Y. Han, L. Hong, Y. Chen, C. Zhong, Y. Wang, D. Zhao, T. Xue, L. Qiao, J. Qiu, Up-regulation of Nob1 in the rat auditory system with noise-induced hearing loss. *Neurosci. Lett.* **491**, 79–82 (2011). doi:10.1016/j.neulet.2011.01.010 Medline

52. N. Schneider, M. Meier, Efficient in situ detection of mRNAs using the Chlorella virus DNA ligase for padlock probe ligation. *RNA* **23**, 250–256 (2017). doi:10.1261/rna.057836.116 Medline

53. D. Fürth, V. Hatini, J. H. Lee, In Situ Transcriptome Accessibility Sequencing (INSTA-seq). bioRxiv 722819 [Preprint]. 6 August 2019. https://doi.org/10.1101/722819.

54. A. E. Cartier, S. N. Djakovic, A. Salehi, S. M. Wilson, E. Masliah, G. N. Patrick, Regulation of synaptic structure by ubiquitin C-terminal hydrolase L1. *J. Neurosci.* **29**, 7857–7868 (2009). doi:10.1523/JNEUROSCI.1817-09.2009 Medline

55. G. J. S. Lohman, Y. Zhang, A. M. Zhelkovsky, E. J. Cantor, T. C. Evans Jr., Efficient DNA ligation in DNA-RNA hybrid helices by Chlorella virus DNA ligase. *Nucleic Acids Res.* **42**, 1831–1844 (2014). doi:10.1093/nar/gkt1032 Medline

56. R. Deng, K. Zhang, Y. Sun, X. Ren, J. Li, Highly specific imaging of mRNA in single cells by target RNA-initiated rolling circle amplification. *Chem. Sci.* **8**, 3668–3675 (2017). doi:10.1039/C7SC00292K Medline

57. E. P. R. Iyer, S. Punthambaker, S. Liu, K. Jindal, M. Farrell, J. Murn, T. Ferrante, S. Rudnicki, R. E. Kohman, A. T. Wassie, D. Goodwin, F. Chen, S. Alon, A. Sinha, D. Milanova, L. Aron, C. Camplisson, A. Skrynnyk, P. L. Reginato, N. Conway, J. Aach, B. Yankner, E. S. Boyden, G. M. Church, Barcoded oligonucleotides ligated on RNA amplified for multiplex and parallel in-situ analyses. bioRxiv 281121 [Preprint]. 20 March 2018. https://doi.org/10.1101/281121.

58. H. M. T. Choi, M. Schwarzkopf, M. E. Fornace, A. Acharya, G. Artavanis, J. Stegmaier, A. Cunha, N. A. Pierce, Third-generation *in situ* hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753 (2018). doi:10.1242/dev.165753 Medline

59. G. K. Marinov, B. A. Williams, K. McCue, G. P. Schroth, J. Gertz, R. M. Myers, B. J. Wold, From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. *Genome Res.* **24**, 496–510 (2014). doi:10.1101/gr.161034.113 Medline

60. E. Torre, H. Dueck, S. Shaffer, J. Gospocic, R. Gupte, R. Bonasio, J. Kim, J. Murray, A. Raj, Rare Cell Detection by Single-Cell RNA Sequencing as Guided by Single-Molecule RNA FISH. *Cell Syst.* **6**, 171–179.e5 (2018). doi:10.1016/j.cels.2018.01.014 Medline

61. B. Tasic, V. Menon, T. N. Nguyen, T. K. Kim, T. Jarsky, Z. Yao, B. Levi, L. T. Gray, S. A. Sorensen, T. Dolbeare, D. Bertagnolli, J. Goldy, N. Shapovalova, S. Parry, C. Lee, K.

Smith, A. Bernard, L. Madisen, S. M. Sunkin, M. Hawrylycz, C. Koch, H. Zeng, Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016). doi:10.1038/nn.4216 Medline

62. L. van der Maaten, G. Hinton, Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008); http://jmlr.org/papers/v9/vandermaaten08a.html.

63. E. S. Lein, M. J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A. F. Boe, M. S. Boguski, K. S. Brockway, E. J. Byrnes, L. Chen, L. Chen, T.-M. Chen, M. C. Chin, J. Chong, B. E. Crook, A. Czaplinska, C. N. Dang, S. Datta, N. R. Dee, A. L. Desaki, T. Desta, E. Diep, T. A. Dolbeare, M. J. Donelan, H.-W. Dong, J. G. Dougherty, B. J. Duncan, A. J. Ebbert, G. Eichele, L. K. Estin, C. Faber, B. A. Facer, R. Fields, S. R. Fischer, T. P. Fliss, C. Frensley, S. N. Gates, K. J. Glattfelder, K. R. Halverson, M. R. Hart, J. G. Hohmann, M. P. Howell, D. P. Jeung, R. A. Johnson, P. T. Karr, R. Kawal, J. M. Kidney, R. H. Knapik, C. L. Kuan, J. H. Lake, A. R. Laramee, K. D. Larsen, C. Lau, T. A. Lemon, A. J. Liang, Y. Liu, L. T. Luong, J. Michaels, J. J. Morgan, R. J. Morgan, M. T. Mortrud, N. F. Mosqueda, L. L. Ng, R. Ng, G. J. Orta, C. C. Overly, T. H. Pak, S. E. Parry, S. D. Pathak, O. C. Pearson, R. B. Puchalski, Z. L. Riley, H. R. Rockett, S. A. Rowland, J. J. Royall, M. J. Ruiz, N. R. Sarno, K. Schaffnit, N. V. Shapovalova, T. Sivisay, C. R. Slaughterbeck, S. C. Smith, K. A. Smith, B. I. Smith, A. J. Sodt, N. N. Stewart, K.-R. Stumpf, S. M. Sunkin, M. Sutram, A. Tam, C. D. Teemer, C. Thaller, C. L. Thompson, L. R. Varnam, A. Visel, R. M. Whitlock, P. E. Wohnoutka, C. K. Wolkey, V. Y. Wong, M. Wood, M. B. Yaylaoglu, R. C. Young, B. L. Youngstrom, X. F. Yuan, B. Zhang, T. A. Zwingman, A. R. Jones, Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007). doi:10.1038/nature05453 Medline

64. L. E. Ostroff, J. C. Fiala, B. Allwardt, K. M. Harris, Polyribosomes redistribute from dendritic shafts into spines with enlarged synapses during LTP in developing rat hippocampal slices. *Neuron* **35**, 535–545 (2002). doi:10.1016/S0896-6273(02)00785-7 Medline

65. O. Steward, P. Worley, Local synthesis of proteins at synaptic sites on dendrites: Role in synaptic plasticity and memory consolidation? *Neurobiol. Learn. Mem.* **78**, 508–527 (2002). doi:10.1006/nlme.2002.4102 Medline

66. Y. J. Yoon, B. Wu, A. R. Buxbaum, S. Das, A. Tsai, B. P. English, J. B. Grimm, L. D. Lavis, R. H. Singer, Glutamate-induced RNA localization and translation in neurons. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6877–E6886 (2016). doi:10.1073/pnas.1614267113 Medline

67. I. J. Cajigas, G. Tushev, T. J. Will, S. tom Dieck, N. Fuerst, E. M. Schuman, The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* **74**, 453–466 (2012). doi:10.1016/j.neuron.2012.02.036 Medline

68. A. Y. Hung, K. Futai, C. Sala, J. G. Valtschanoff, J. Ryu, M. A. Woodworth, F. L. Kidd, C. C. Sung, T. Miyakawa, M. F. Bear, R. J. Weinberg, M. Sheng, Smaller dendritic spines, weaker synaptic transmission, but enhanced spatial learning in mice lacking Shank1. *J. Neurosci.* **28**, 1697–1708 (2008). doi:10.1523/JNEUROSCI.3032-07.2008 Medline

69. N. Hirokawa, S. Niwa, Y. Tanaka, Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* **68**, 610–638 (2010). doi:10.1016/j.neuron.2010.09.039 Medline

70. A. J. Koleske, Molecular mechanisms of dendrite stability. *Nat. Rev. Neurosci.* **14**, 536–550 (2013). doi:10.1038/nrn3486 Medline

71. O. Steward, C. S. Wallace, G. L. Lyford, P. F. Worley, Synaptic activation causes the mRNA for the IEG Arc to localize selectively near activated postsynaptic sites on dendrites. *Neuron* **21**, 741–751 (1998). doi:10.1016/S0896-6273(00)80591-7 Medline

72. O. Steward, P. F. Worley, Selective targeting of newly synthesized Arc mRNA to active synapses requires NMDA receptor activation. *Neuron* **30**, 227–240 (2001). doi:10.1016/S0896-6273(01)00275-6 Medline

73. D. S. Chen, I. Mellman, Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017). doi:10.1038/nature21349 Medline

74. P. L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J. F. Navarro, J. Magnusson, S. Giacomello, M. Asp, J. O. Westholm, M. Huss, A. Mollbrink, S. Linnarsson, S. Codeluppi, Å. Borg, F. Pontén, P. I. Costea, P. Sahlén, J. Mulder, O. Bergmann, J. Lundeberg, J. Frisén, Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016). doi:10.1126/science.aaf2403 Medline

75. S. G. Rodriques, R. R. Stickels, A. Goeva, C. A. Martin, E. Murray, C. R. Vanderburg, J. Welch, L. M. Chen, F. Chen, E. Z. Macosko, Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019). doi:10.1126/science.aaw1219 Medline

76. M. Fenech, M. Kirsch-Volders, A. T. Natarajan, J. Surralles, J. W. Crott, J. Parry, H. Norppa, D. A. Eastmond, J. D. Tucker, P. Thomas, Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis* **26**, 125–132 (2011). doi:10.1093/mutage/geq052 Medline

77. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck 3rd, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21 (2019). doi:10.1016/j.cell.2019.05.031 Medline

78. M. Kitahara, Y. Hozumi, A. Nakamura, K. Tachi, H. Saitoh, T. Iijima, HER2-Positive Conversion in a Metastatic Liver Focus in Late Recurrent Breast Cancer. *Case Rep. Oncol.* **12**, 473–479 (2019). doi:10.1159/000501306 Medline

79. H. Masuda, D. Zhang, C. Bartholomeusz, H. Doihara, G. N. Hortobagyi, N. T. Ueno, Role of epidermal growth factor receptor in breast cancer. *Breast Cancer Res. Treat.* **136**, 331–345 (2012). doi:10.1007/s10549-012-2289-9 Medline

80. P. Marcato, C. A. Dean, D. Pan, R. Araslanova, M. Gillis, M. Joshi, L. Helyer, L. Pan, A. Leidal, S. Gujar, C. A. Giacomantonio, P. W. K. Lee, Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells* **29**, 32–45 (2011). doi:10.1002/stem.563 Medline

81. S. Garaud, L. Buisseret, C. Solinas, C. Gu-Trantien, A. de Wind, G. Van den Eynden, C. Naveaux, J.-N. Lodewyckx, A. Boisson, H. Duvillier, L. Craciun, L. Ameye, I. Veys, M. Paesmans, D. Larsimont, M. Piccart-Gebhart, K. Willard-Gallo, Tumor-infiltrating B cells signal functional humoral immune responses in breast cancer. *JCI Insight* **4**, e129641 (2019). doi:10.1172/jci.insight.129641 Medline

82. E. Pipi, S. Nayar, D. H. Gardner, S. Colafrancesco, C. Smith, F. Barone, Tertiary Lymphoid Structures: Autoimmunity Goes Local. *Front. Immunol.* **9**, 1952 (2018). doi:10.3389/fimmu.2018.01952 Medline

83. S. Nayar, J. Campos, C. G. Smith, V. Iannizzotto, D. H. Gardner, F. Mourcin, D. Roulois, J. Turner, M. Sylvestre, S. Asam, B. Glaysher, S. J. Bowman, D. T. Fearon, A. Filer, K. Tarte, S. A. Luther, B. A. Fisher, C. D. Buckley, M. C. Coles, F. Barone, Immunofibroblasts are pivotal drivers of tertiary lymphoid structure formation and local pathology. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 13490–13497 (2019). doi:10.1073/pnas.1905301116 Medline

84. M. Bartoschek, N. Oskolkov, M. Bocci, J. Lövrot, C. Larsson, M. Sommarin, C. D. Madsen, D. Lindgren, G. Pekar, G. Karlsson, M. Ringnér, J. Bergh, Å. Björklund, K. Pietras, Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat. Commun.* **9**, 5150 (2018). doi:10.1038/s41467-018-07582-3 Medline

85. N. Colwell, M. Larion, A. J. Giles, A. N. Seldomridge, S. Sizdahkhani, M. R. Gilbert, D. M. Park, Hypoxia in the glioblastoma microenvironment: Shaping the phenotype of cancer stem-like cells. *Neuro-Oncol.* **19**, 887–896 (2017). doi:10.1093/neuonc/now258 Medline

86. F. Cimmino, M. Avitabile, V. A. Lasorsa, A. Montella, L. Pezone, S. Cantalupo, F. Visconte, M. V. Corrias, A. Iolascon, M. Capasso, HIF-1 transcription activity: HIF1A driven response in normoxia and in hypoxia. *BMC Med. Genet.* **20**, 37 (2019). doi:10.1186/s12881-019-0767-1 Medline

87. C. Gebhardt, J. Németh, P. Angel, J. Hess, S100A8 and S100A9 in inflammation and cancer. *Biochem. Pharmacol.* **72**, 1622–1631 (2006). doi:10.1016/j.bcp.2006.05.017 Medline

88. J.-M. Zhong, J. Li, A.-D. Kang, S.-Q. Huang, W.-B. Liu, Y. Zhang, Z.-H. Liu, L. Zeng, Protein S100-A8: A potential metastasis-associated protein for breast cancer determined via iTRAQ quantitative proteomic and clinicopathological analysis. *Oncol. Lett.* **15**, 5285–5293 (2018). doi:10.3892/ol.2018.7958 Medline

89. D. Wang, G. Liu, B. Wu, L. Chen, L. Zeng, Y. Pan, Clinical Significance of Elevated S100A8 Expression in Breast Cancer Patients. *Front. Oncol.* **8**, 496 (2018). doi:10.3389/fonc.2018.00496 Medline

90. R. Kalhor, K. Kalhor, L. Mejia, K. Leeper, A. Graveline, P. Mali, G. M. Church, Developmental barcoding of whole mouse via homing CRISPR. *Science* **361**, eaat9804 (2018). doi:10.1126/science.aat9804 Medline

91. X. Chen, Y.-C. Sun, G. M. Church, J. H. Lee, A. M. Zador, Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. *Nucleic Acids Res.* **46**, e22 (2018). doi:10.1093/nar/gkx1206 Medline

92. I. D. Peikon, J. M. Kebschull, V. V. Vagin, D. I. Ravens, Y.-C. Sun, E. Brouzes, I. R. Corrêa Jr., D. Bressan, A. M. Zador, Using high-throughput barcode sequencing to efficiently map connectomes. *Nucleic Acids Res.* **45**, e115 (2017). doi:10.1093/nar/gkx292 Medline

93. A. M. Zador, J. Dubnau, H. K. Oyibo, H. Zhan, G. Cao, I. D. Peikon, Sequencing the connectome. *PLOS Biol.* **10**, e1001411 (2012). doi:10.1371/journal.pbio.1001411 Medline

94. D. Goodwin, dgoodwin208/ExSeqOF3DVisualizer: First release of 3D visualization software for Expansion Sequencing, version v1.0.0, Zenodo (2020); http://doi.org/10.5281/zenodo.4075649.

95. R. B. Dell, S. Holleran, R. Ramakrishnan, Sample size determination. *ILAR J.* **43**, 207–213 (2002). doi:10.1093/ilar.43.4.207 Medline

96. N. C. Klapoetke, Y. Murata, S. S. Kim, S. R. Pulver, A. Birdsey-Benson, Y. K. Cho, T. K. Morimoto, A. S. Chuong, E. J. Carpenter, Z. Tian, J. Wang, Y. Xie, Z. Yan, Y. Zhang, B. Y. Chow, B. Surek, M. Melkonian, V. Jayaraman, M. Constantine-Paton, G. K.-S. Wong, E. S. Boyden, Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014). doi:10.1038/nmeth.2836 Medline

97. B. Y. Chow, X. Han, A. S. Dobry, X. Qian, A. S. Chuong, M. Li, M. A. Henninger, G. M. Belfort, Y. Lin, P. E. Monahan, E. S. Boyden, High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* **463**, 98–102 (2010). doi:10.1038/nature08652 Medline

98. S. Karmakar, E. M. Harcourt, D. S. Hewings, F. Scherer, A. F. Lovejoy, D. M. Kurtz, T. Ehrenschwender, L. J. Barandun, C. Roost, A. A. Alizadeh, E. T. Kool, Organocatalytic removal of formaldehyde adducts from RNA and DNA bases. *Nat. Chem.* **7**, 752–758 (2015). doi:10.1038/nchem.2307 Medline

99. G. T. Hermanson, *Bioconjugate Techniques* (Academic Press, 2013).

100. D. Mitchell 3rd, A. J. Renda, C. A. Douds, P. Babitzke, S. M. Assmann, P. C. Bevilacqua, In vivo RNA structural probing of uracil and guanine base-pairing by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC). *RNA* **25**, 147–157 (2019). doi:10.1261/rna.067868.118 Medline

101. S. H. Rouhanifard, I. A. Mellis, M. Dunagin, S. Bayatpour, C. L. Jiang, I. Dardani, O. Symmons, B. Emert, E. Torre, A. Cote, A. Sullivan, J. A. Stamatoyannopoulos, A. Raj, ClampFISH detects individual nucleic acid molecules using click chemistry–based amplification. *Nat. Biotechnol.* **37**, 84–89 (2019). doi:10.1038/nbt.4286 Medline

102. C. Harris, M. Stephens, "A combined corner and edge detector" in *Proceedings of the Alvey Vision Conference*, C. J. Taylor, Ed. (Alvety Vision Club, 1988), pp. 23.1–23.6; https://doi.org/10.5244/C.2.23.

103. E. Murray, J. H. Cho, D. Goodwin, T. Ku, J. Swaney, S.-Y. Kim, H. Choi, Y.-G. Park, J.-Y. Park, A. Hubbert, M. McCue, S. Vassallo, N. Bakh, M. P. Frosch, V. J. Wedeen, H. S. Seung, K. Chung, Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell* **163**, 1500–1514 (2015). doi:10.1016/j.cell.2015.11.025 Medline

104. D. G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **60**, 91–110 (2004). doi:10.1023/B:VISI.0000029664.99615.94

105. P. Scovanner, S. Ali, M. Shah, in *Proceedings of the 15th ACM International Conference on Multimedia* (Association for Computing Machinery, 2007), pp. 357–360.

106. F. L. Bookstein, Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567–585 (1989). doi:10.1109/34.24792

107. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi:10.1038/nmeth.2019 Medline

108. A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008). doi:10.1038/nmeth.1253 Medline

109. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi:10.1038/nmeth.1923 Medline

110. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi:10.1093/bioinformatics/btp352 Medline

111. A. Morgulis, E. M. Gertz, A. A. Schäffer, R. Agarwala, WindowMasker: Window-based masker for sequenced genomes. *Bioinformatics* **22**, 134–141 (2006). doi:10.1093/bioinformatics/bti774 Medline

112. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). doi:10.1093/bioinformatics/btq033 Medline

113. B. D. Ondov, A. Varadarajan, K. D. Passalacqua, N. H. Bergman, Efficient mapping of Applied Biosystems SOLiD sequence data to a reference genome for functional genomic applications. *Bioinformatics* **24**, 2776–2777 (2008). doi:10.1093/bioinformatics/btn512 Medline

114. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009). doi:10.1038/nprot.2008.211 Medline

115. J. Zhang, K. Kobert, T. Flouri, A. Stamatakis, PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014). doi:10.1093/bioinformatics/btt593 Medline

116. J. S. Duerr, Immunohistochemistry. *WormBook* **2006**, 1–61 (2006). doi:10.1895/wormbook.1.105.1 Medline

117. T. Stiernagle, Maintenance of *C. elegans*. *WormBook* **2006**, 1–11 (2006). doi:10.1895/wormbook.1.101.1 Medline

118. D. R. Berger, H. S. Seung, J. W. Lichtman, VAST (Volume Annotation and Segmentation Tool): Efficient Manual and Semi-Automatic Labeling of Large 3D Image Stacks. *Front. Neural Circuits* **12**, 88 (2018). doi:10.3389/fncir.2018.00088 Medline

119. OpenFrameworks Community, "OpenFrameworks"; https://openframeworks.cc.

120. P. R. Nicovich, M. J. Taormina, C. A. Baker, T. N. Nguyen, E. R. Thomsen, E. Garren, B. Long, M. Gorham, J. A. Miller, T. Hage, A. Bosma-Moody, G. J. Murphy, B. P. Levi, J. L. Close, B. Tasic, E. S. Lein, H. Zeng, Multimodal cell type correspondence by intersectional mFISH in intact tissues. bioRxiv 525451 [Preprint]. 30 January 2019. https://doi.org/10.1101/525451.

121. B. Tasic, Z. Yao, L. T. Graybuck, K. A. Smith, T. N. Nguyen, D. Bertagnolli, J. Goldy, E. Garren, M. N. Economo, S. Viswanathan, O. Penn, T. Bakken, V. Menon, J. Miller, O. Fong, K. E. Hirokawa, K. Lathia, C. Rimorin, M. Tieu, R. Larsen, T. Casper, E. Barkan, M. Kroll, S. Parry, N. V. Shapovalova, D. Hirschstein, J. Pendergraft, H. A. Sullivan, T. K. Kim, A. Szafer, N. Dee, P. Groblewski, I. Wickersham, A. Cetin, J. A. Harris, B. P. Levi, S. M. Sunkin, L. Madisen, T. L. Daigle, L. Looger, A. Bernard, J. Phillips, E. Lein, M. Hawrylycz, K. Svoboda, A. R. Jones, C. Koch, H. Zeng, Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018). doi:10.1038/s41586-018-0654-5 Medline

122. F. C. Fang, E. R. Frawley, T. Tapscott, A. Vázquez-Torres, Bacterial Stress Responses during Host Infection. *Cell Host Microbe* **20**, 133–143 (2016). doi:10.1016/j.chom.2016.07.009 Medline

123. Q. Xu, M. R. Schlabach, G. J. Hannon, S. J. Elledge, Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2289–2294 (2009). doi:10.1073/pnas.0812506106 Medline

124. T. Buschmann, DNABarcodes: An R package for the systematic construction of DNA sample tags. *Bioinformatics* **33**, 920–922 (2017). doi:10.1093/bioinformatics/btw759 Medline

125. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018). doi:10.1038/nbt.4096 Medline

126. L. Keren, M. Bosse, D. Marquez, R. Angoshtari, S. Jain, S. Varma, S.-R. Yang, A. Kurian, D. Van Valen, R. West, S. C. Bendall, M. Angelo, A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373–1387.e19 (2018). doi:10.1016/j.cell.2018.08.039 Medline

127. X. Qian, K. D. Harris, T. Hauling, D. Nicoloutsopoulos, A. B. Muñoz-Manchado, N. Skene, J. Hjerling-Leffler, M. Nilsson, Probabilistic cell typing enables fine mapping of closely related cell types in situ. *Nat. Methods* **17**, 101–106 (2020). doi:10.1038/s41592-019-0631-4 Medline

128. S. Codeluppi, L. E. Borm, A. Zeisel, G. La Manno, J. A. van Lunteren, C. I. Svensson, S. Linnarsson, Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods* **15**, 932–935 (2018). doi:10.1038/s41592-018-0175-z Medline

129. C. Xia, H. P. Babcock, J. R. Moffitt, X. Zhuang, Multiplexed detection of RNA using MERFISH and branched DNA amplification. *Sci. Rep.* **9**, 7721 (2019). doi:10.1038/s41598-019-43943-8 Medline

130. J. R. Moffitt, J. Hao, G. Wang, K. H. Chen, H. P. Babcock, X. Zhuang, High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11046–11051 (2016). doi:10.1073/pnas.1612826113 Medline

131. S. Vickovic, G. Eraslan, F. Salmén, J. Klughammer, L. Stenbeck, D. Schapiro, T. Äijö, R. Bonneau, L. Bergenstråhle, J. F. Navarro, J. Gould, G. K. Griffin, Å. Borg, M. Ronaghi, J. Frisén, J. Lundeberg, A. Regev, P. L. Ståhl, High-definition spatial transcriptomics for

in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019). doi:10.1038/s41592-019-0548-y Medline

132. R. J. Kast, A. L. Lanjewar, C. D. Smith, P. Levitt, FOXP2 exhibits projection neuron class specific expression, but is not required for multiple aspects of cortical histogenesis. *eLife* **8**, e42012 (2019). doi:10.7554/eLife.42012 Medline

133. J. M. Gunnersen, M. H. Kim, S. J. Fuller, M. De Silva, J. M. Britto, V. E. Hammond, P. J. Davies, S. Petrou, E. S. L. Faber, P. Sah, S.-S. Tan, Sez-6 proteins affect dendritic arborization patterns and excitability of cortical pyramidal neurons. *Neuron* **56**, 621–639 (2007). doi:10.1016/j.neuron.2007.09.018 Medline

134. M. L. S. Tantirigama, M. J. Oswald, A. J. Clare, H. E. Wicky, R. C. Day, S. M. Hughes, R. M. Empson, *Fezf2* expression in layer 5 projection neurons of mature mouse motor cortex. *J. Comp. Neurol.* **524**, 829–845 (2016). doi:10.1002/cne.23875 Medline

135. H. Zeng, E. H. Shen, J. G. Hohmann, S. W. Oh, A. Bernard, J. J. Royall, K. J. Glattfelder, S. M. Sunkin, J. A. Morris, A. L. Guillozet-Bongaarts, K. A. Smith, A. J. Ebbert, B. Swanson, L. Kuan, D. T. Page, C. C. Overly, E. S. Lein, M. J. Hawrylycz, P. R. Hof, T. M. Hyde, J. E. Kleinman, A. R. Jones, Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* **149**, 483–496 (2012). doi:10.1016/j.cell.2012.02.052 Medline

136. T. L. Daigle, L. Madisen, T. A. Hage, M. T. Valley, U. Knoblich, R. S. Larsen, M. M. Takeno, L. Huang, H. Gu, R. Larsen, M. Mills, A. Bosma-Moody, L. A. Siverts, M. Walker, L. T. Graybuck, Z. Yao, O. Fong, T. N. Nguyen, E. Garren, G. H. Lenz, M. Chavarha, J. Pendergraft, J. Harrington, K. E. Hirokawa, J. A. Harris, P. R. Nicovich, M. J. McGraw, D. R. Ollerenshaw, K. A. Smith, C. A. Baker, J. T. Ting, S. M. Sunkin, J. Lecoq, M. Z. Lin, E. S. Boyden, G. J. Murphy, N. M. da Costa, J. Waters, L. Li, B. Tasic, H. Zeng, A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. *Cell* **174**, 465–480.e22 (2018). doi:10.1016/j.cell.2018.06.035 Medline

137. B. Rudy, G. Fishell, S. Lee, J. Hjerling-Leffler, Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev. Neurobiol.* **71**, 45–61 (2011). doi:10.1002/dneu.20853 Medline