Cite as: X. Wang *et al.*, *Science*
10.1126/science.aat5691 (2018).

Three-dimensional intact-tissue sequencing of single-cell transcriptional states

Xiao Wang[1]*, William E. Allen[1,2]*, Matthew A. Wright[1,3], Emily L. Sylwestrak[1], Nikolay Samusik[4], Sam Vesuna[1], Kathryn Evans[1], Cindy Liu[1], Charu Ramakrishnan[1], Jia Liu[5], Garry P. Nolan[4]†, Felice-Alessio Bava[4]†, Karl Deisseroth[1,3,6]†

[1]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [2]Neuroscience Program, Stanford University, CA 94305, USA. [3]Department of Psychiatry and Behavioral Sciences, Stanford University, CA 94305, USA. [4]Baxter Laboratory, Department of Microbiology and Immunology, Stanford University, Stanford, CA 94305, USA. [5]Department of Chemical Engineering, Stanford University, CA 94305, USA. [6]Howard Hughes Medical Institute, Stanford University, CA 94305, USA.

*These authors contributed equally to this work.

†Corresponding author. Email: deissero@stanford.edu (K.D.); alessio.bava@gmail.com (F.-A.B.); gnolan@stanford.edu (G.P.N.)

Retrieving high-content gene-expression information while retaining 3D positional anatomy at cellular resolution has been difficult, limiting integrative understanding of structure and function in complex biological tissues. Here we develop and apply a technology for 3D intact-tissue RNA sequencing, termed STARmap (Spatially-resolved Transcript Amplicon Readout Mapping), which integrates hydrogel-tissue chemistry, targeted signal amplification, and in situ sequencing. The capabilities of STARmap were tested by mapping 160 to 1,020 genes simultaneously in sections of mouse brain at single-cell resolution with high efficiency, accuracy and reproducibility. Moving to thick tissue blocks, we observed a molecularly-defined gradient distribution of excitatory-neuron subtypes across cubic millimeter-scale volumes (>30,000 cells), and discovered a short-range 3D self-clustering in many inhibitory-neuron subtypes that could be identified and described with 3D STARmap.

In biological tissues, diversity of function arises from diversity of form–in part via the complexity of cell-specific gene expression, which defines the unique three-dimensional molecular anatomy and cellular properties of each tissue. In situ transcriptomic tools for the spatial mapping of gene expression with subcellular resolution have emerged that may be applicable to probing these tissue structure-function relationships, including both multiplexed in situ RNA hybridization and in situ RNA sequencing (*1*–*10*). Current in situ sequencing approaches face the challenge of implementing enzymatic reactions in the dense, complex tissue environment and currently suffer from low efficiency (*2*), but the potential value of such intact-tissue sequencing could be enormous; in comparison to hybridization-based multiplexing/readout which utilizes multiple polynucleotide probes to encode gene identity (*3*–*5*), sequencing operates with single-nucleotide resolution, and thus inherently provides greater information. In addition, in situ sequencing methods typically utilize signal amplification, important for detection of short transcripts (e.g., neuropeptides) and for high-quality imaging in thick tissue blocks. However, current sequencing methods have not yet been successfully applicable to 3D volumes of intact tissue, due to fundamental limitations in requisite sensitivity, fidelity, and scalability for throughput in tissues such as the mammalian brain.

Hydrogels have been widely used for extracellular 3D scaffolding in applications across biology and medicine (*11*–*13*). Recently-developed hydrogel-tissue chemistry (HTC) methodologies (*14*), beginning with CLARITY (*15*), physically link in situ-synthesized polymers with selected intracellular biomolecules. This process transforms the tissue, from within its constituent cells, into a new state suitable for high-resolution volumetric imaging and analysis compatible with many kinds of molecular phenotyping for proteins, nucleic acids, and other targets (*15*). HTC-based hydrogel-embedding strategies have been extended to nucleic acid analyses in the form of in situ hybridization for RNA (*16*–*19*), but these have not yet been extended to in situ RNA sequencing–which would have the potential to reveal the full molecular complexity of the transcriptome. In non-tissue environments, however, purely synthetic hydrogels have been used to accommodate enzymatic reactions that include DNA sequencing (*20*), and if biological tissue could be converted into a hydrogel-embedded form compatible with creation, retention, and functional presentation of RNA-derived or hybridized complementary DNA (cDNA), it might be possible to perform 3D in situ sequencing within such a tissue-hydrogel formulation–leveraging the crucial attendant properties of optical transparency, reduced background, elevated diffusion rate, and greater mechanical stability. Here we achieve this goal with the development and application of a sequencing-based method (Spatially-resolved Transcript Amplicon Readout Mapping, or STARmap) for targeted 3D in situ transcriptomics in intact tissue (Fig. 1A); using STARmap we were able to identify organizational principles of a full spectrum of cell types, which would not have been otherwise accessible for identification, in the adult mammalian brain.

Results

Design and validation of STARmap principles

One component is an efficient approach for in situ-amplification of a library of cDNA probes hybridized with cellular RNAs (this approach is termed SNAIL, for Specific Amplification of Nucleic Acids via Intramolecular Ligation). Reverse transcription may be the major efficiency-limiting step for in situ sequencing (*7*, *21*), and SNAIL bypasses this step with a pair of primer and padlock probes (fig. S1A) designed such that only when both probes hybridize to the same RNA molecule, the padlock probe can be circularized and rolling-circle-amplified to generate a DNA nanoball (amplicon) containing multiple copies of the cDNA probes (Fig. 1, A to D). This mechanism ensures target-specific signal amplification and excludes noise that invariably otherwise arises from non-specific hybridization of single probes. Indeed, the outcome includes much higher absolute intensity and signal-to-noise ratio (SNR) compared to commercial single-molecule fluorescent in situ hybridization (smFISH) probes (fig. S1, B to F), and substantial improvement of detection efficiency (comparable to single-cell RNA sequencing) with simplified experimental procedures compared to previous in situ RNA sequencing methods (fig. S1, G to I).

To enable cDNA amplicon embedding in the tissue-hydrogel setting, amine-modified nucleotides were spiked into the rolling circle amplification reaction, functionalized with an acrylamide moiety using acrylic acid N-hydroxysuccinimide esters, and copolymerized with acrylamide monomers to form a distinct kind of hydrogel-DNA amplicon network (Fig. 1A and fig. S2A). The resulting tissue-hydrogel was then subjected to protein digestion and lipid removal to enhance transparency (fig. S2, B to E). This design chemistry dictates that amplicons are covalently linked with the hydrogel network, and such crosslinking is essential to maintain the position and integrity of the amplicons through many cycles of detection (fig. S2, F to H).

A 5-base barcode (library size of 1,024) was designed and built into each padlock probe as a gene-unique identifier to be sequenced, thus enabling multiplexed gene detection (Fig. 1A). Sequencing-by-synthesis paradigms were avoided as these require elevated reaction temperatures, which in turn are problematic for high-resolution imaging and sample stability (*16*) in comparison to sequencing-by-ligation methods that can be implemented at room temperature. However, none of the reported or commercially available sequencing-by-ligation methods approach exhibit the necessary SNR or accuracy for this challenging intact-tissue application: SOLiD sequencing causes strong background fluorescence in biological samples (*10*) while cPAL sequencing (*22*) lacks an error-rejection mechanism (fig. S3). For this reason, an approach we term Sequencing with Error-reduction by Dynamic Annealing and Ligation (SEDAL) was devised specifically for STARmap (fig. S3).

SEDAL employs two kinds of short, degenerate probes: reading-probes to decode bases, and fluorescence-probes to transduce decoded sequence information into fluorescence signals. The two short probes only transiently bind to the target DNA and ligate to form a stable product for imaging only when a perfect match occurs; after each cycle corresponding to a base readout, the fluorescent products are stripped by formamide, which eliminates error accumulation as sequencing proceeds (Fig. 1E and fig. S3B). In contrast to SOLiD, SEDAL exhibits minimal background (fig. S3, C to F). A 2-base encoding scheme was designed and implemented to mitigate any residual errors related to imaging high densities of spots (fig. S3, G and H). Based upon a panel of four very-highly expressed test genes in mouse brain (to mimic amplicon crowdedness as would be encountered in highly-multiplexed gene-detection), we found that the error rate of STARmap was more than an order of magnitude lower than prior cPAL methods (~1.8% versus 29.4%) (fig. S3, I to L) (*17*).

Spatial cell typing in primary visual cortex with 160-gene STARmapping

To test if STARmap could deliver on the initial goal of high-content 3D intact-tissue sequencing of single-cell transcriptional states with the necessary sensitivity and accuracy, we applied STARmap to a pressing current challenge in neuroscience: detecting and classifying cell types and corresponding tissue-organization principles in neocortex of the adult mouse brain. The anatomy and function of mouse primary visual neocortex have been extensively studied (*23*), a setting which here allows validation of our results by comparison with prior findings that span multiple papers, methodologies, and data sources (but the full diversity of deeply molecularly-defined cell types within visual cortex has not yet been spatially resolved in a single experiment, precluding identification of potentially fundamental joint statistics and organizational principles across 3D volumes). Among many examples of the experimental leverage such information could provide, joint 3D cell-typology mapping might be employed to help decode the spatiotemporal logic of neural-activity-triggered gene expression as a function of cell type and spatial location.

We therefore used 5-base barcoded SNAIL probes over six rounds of in situ SEDAL sequencing in coronal mouse brain slices (Figs. 1A and 2, A and B) to survey a large but focused and curated gene set (160 genes including 112 putative cell-type markers collated from mouse cortical single-cell RNA sequencing (*24*, *25*) and 48 activity-regulated genes (ARGs) (*26*, *27*). In one arm of the experiment, visually-evoked neural activity was provided to a cohort of mice via 1 hour light-exposure after four days of housing in the dark; other mice were kept continuously in the dark (*27*, *28*). 8 μm-thick volumes containing up to 1,000 cells covering all cortical layers were imaged. After six rounds of sequencing, fluorescent Nissl staining was used to segment cell bodies, allowing attribution of amplicons to individual cells (fig. S4, A and B). The values corresponding to amplicons-per-cell and genes-per-cell varied substantially (Fig. 2C), while the 160-gene expression pattern was consistent between biological replicates (R = 0.94-0.95) (Fig. 2D), revealing reliable detection of transcript diversity at the single-cell level. Because only 160 genes were encoded out of the 1024 possible barcodes from five bases, we were able to quantify sequencing errors that resulted in sequences being corrupted from the 160 true barcodes to the 864 invalid barcodes, which was remarkably low at 1-4%. We found that this 160-gene pilot faithfully reproduced the spatial distribution of known cortical layer markers and interneurons, illustrated here via comparison of in situ images from paired public atlases (*29*) and STARmap results (Fig. 2E).

We next performed cell classification using expression data of the 112 cell-type markers. First, >3,000 cells pooled from four biological replicates were clustered into three major cell types (excitatory neurons, inhibitory neurons and non-neuronal cells) using graph-based clustering following principal-component decomposition (*30*), and then further sub-clustered under each category (Fig. 2, F to H, and fig. S4C). Intriguingly, the richly-defined excitatory neurons segregated into four major types, here denoted (eL2/3, eL4, eL5 and eL6) (Fig. 2, I to K, and fig. S5, A and B) by spatial correspondence with anatomic cortical layers and expression profiles of known layer-specific gene markers. Although spatial organization of the four excitatory types exhibited a layered pattern, there was extensive intermixing among different cell types within each layer. Inhibitory neurons were also clustered into four major types, here denoted by the dominant interneuron marker of each subtype [VIP, SST, NPY and PV (Pvalb)] (Fig. 2, L to N, and fig. S5, C and D); the VIP and NPY type were observed to distribute more to the upper layers (L1-3) while SST and PV types were found more commonly in the lower layers (L4-6). We also detected non-neuronal cell types, including astrocytes, oligodendrocytes, endothelial cells, and smooth muscle cells (fig. S6). The number of major cell-types illustrated here (12 in total) can be further broken down (single-cell RNA sequencing can lead to classification into 40 or more subtypes, consistent with the readily-apparent heterogeneity of gene expression within each type) (figs. S5 and S6). Notably, with our targeted 112-gene set and at the size of 600-800 cells per sample, all 12 major cell types could be reliably detected without batch effects with highly similar spatial patterning among four biological replicates (defined as samples prepared from different animals) (fig. S7), and matched with published single-cell RNA sequencing results (fig. S8)

We next sought to take advantage of STARmap's quantitative capabilities at the single-cell level, to test differential gene expression analyses across experimental conditions, in molecularly-defined cell types. To this end, we assessed visual-stimulus-dependent gene expression patterns (via 48 defined ARGs with single-cell resolution in situ). Further developing the single-cell RNA sequencing procedure, mouse brains here were flash-frozen with minimal handling time after sacrifice (<5 min), for maximal preservation of native transcriptional signatures. Consistent with prior reports (*26*–*28*), we observed global induction of known immediate-early genes (*Fos*, *Egr1*, and *Egr2*) (Fig. 3A) in primary visual cortex upon 1 hour of light exposure. At single-cell resolution, the quantitative extent (fold change in expression) of ARG changes exhibited striking diversity across neuronal cell types (Fig. 3, B and C, and fig. S9) (*28*). In general, ARG expression programs in excitatory neurons across different layers were highly similar, whereas ARG expression programs in inhibitory cells exhibited much more distinct cell-type specific characteristics (fig. S9C); for example, *Egr2* exhibited light-induction across excitatory neurons (Fig. 3D) but not in inhibitory neurons, while in contrast, *Prok2* was upregulated in *Vip* inhibitory neurons (Fig. 3C) (*22*). Finally, since neural activity can trigger co-transcription of noncoding RNAs from within enhancers of ARGs (*26*, *31*), we also studied exemplars of these enhancer RNAs (here, eRNAs 1-5 of the *Fos* gene); these transcripts, not polyadenylated, would be very difficult to measure with current single-cell RNA sequencing. Intriguingly, eRNA3 was identified as the most significant and consistent ARG marker (fig. S9B).

Comparing spatial cell-type distributions in frontal and sensory cortices

We then investigated to what extent the cell types of higher cognitive cortex resemble those of sensory cortex, as exemplified by primary visual cortex. We applied the same 160-gene set to STARmapping the cell populations of medial prefrontal cortex (mPFC) (Fig. 4A) which is involved in high-level cognitive functions such as attention and memory, and is thought to be dysregulated in major psychiatric disorders (*32*). We identified 15 distinct molecular cell types including six excitatory neuron subtypes (eL2/3, eL5-1, eL5-2, eL5-3, eL6-1 and eL6-2, annotated by anatomic cortical layers), five inhibitory neuron subtypes (VIP, Reln, SST, Lhx6 and NPY, annotated by dominant gene markers) and four non-neuronal types (astrocytes, oligodendrocytes, endothelial cells and smooth-muscle cells) (Fig. 4B and fig. S10).

The spatial organization of broad cell types in mPFC resembled that of V1 with intermixed excitatory neuronal layers and sparsely distributed inhibitory neurons (Fig. 4C), however, the nature and composition of neuronal subtypes in mPFC and V1 strikingly differed (Fig. 4, D and E). For excitatory subtypes, mPFC lacks eL4 (consistent with previous reports) (*33*), and exhibits reduced eL2/3 and vast expansion of eL5 and eL6 compared to V1 (Fig. 4E). Many new types of cell were discovered, including three eL5 subtypes and two eL6 subtypes, as characterized by gene markers *Sema3e*, *Plcxd2*, *Tpbg*, *Syt6*, and *Ctgf*, respectively (Fig. 4D).

Substantially different tissue organization by cell type was also observed for inhibitory subclusters. *Sst*-, *Vip*- and *Npy*-positive subtypes in mPFC were represented similarly among all inhibitory neurons compared with those in V1, while *Pvalb*-positive cells were comparatively much sparser. In V1 *Reln*-positive neurons coexist with *Sst* and *Npy*, while in mPFC these segregate as a single cluster with ~50% co-marked by *Ndnf*; we also discovered a new inhibitory subtype labeled by *Lhx6* which in fact constitutes the most abundant inhibitory subtype in mPFC (Fig. 4E). While the 5-HT(3A) receptor (*Htr3a*) expression has been reported in cortical inhibitory neurons (*34*), *Htr3a* has not been ranked as a critical genetic marker of inhibitory subtypes in V1. In mPFC, however, we find that *Htr3a* distinguishably marks a large fraction of *Vip*$^+$ neurons and a subset of *Reln*$^+$ neurons (fig. S10D).

Superficial layers (L1-3) were found to contain *Vip*, *Reln* and *Npy* subtypes while deeper layers (L5-6) were found to contain all of the inhibitory subtypes. Indeed, all of the 15 cell types with tissue-level spatial organization could be reliably detected by STARmap across four biological replicates (fig. S11). The capability of STARmap for multidimensional cell typing in mPFC was further demonstrated in the setting of activity-dependence, supporting the possibility of defining cell types in part by communication properties including activity during behavior (*35*, *36*). 1 hour after cocaine injection (*37*), a specific subpopulation of deep-layer excitatory neurons (e.g., *Tpbg* labeled eL5-2) in mPFC was activated compared to saline-injected control mice (Fig. 4F), revealing STARmap capability for identifying functional segregation of neuronal subtypes in mPFC.

Scaling STARmap to >1,000 genes

To further test the scalability of STARmap, we extended our gene list from 160 to 1,020 genes, leveraging previously published single-cell RNA sequencing data (*24*). The 1,020-gene set was first validated in mouse hippocampal neuron culture, with successful resolution of neuronal and glial cells (fig. S12). We then probed mouse V1 neocortex with the 1,020-gene set to evaluate performance in spatial cell typing in comparison with the 160-gene set. Amplicons obtained in 1,020-gene experiment were much denser in cells compared to those in 160-gene experiments but were optically-resolvable in 3D with high-resolution imaging and post-imaging deconvolution (Fig. 5A).

We observed that a higher percentage (40%) of amplicons were filtered out in the 1020-gene experiments by our error-rejection mechanism (fig. S3H) in comparison to the 4-gene experiments (20%) (fig. S3L), indicating a more frequent initial color-misassignment potentially resulted from amplicon merging or optical resolution, and further demonstrating the importance of our designed error-rejection mechanism. Crucially though despite the read loss, we successfully clustered single cells of the imaging area into 16 annotated cell types using 1,020 genes and the same data analysis pipeline from the focused 160 gene probe set (Fig. 5, B and C, and fig. S13). Three new cell types were identified in addition to the 13 cell types detected by 160 genes (Fig. 5B): eL6 was resolved into two subtypes (eL6-1 and eL6-2); a novel hippocampal excitatory subtype (HPC) was identified; and microglial cells were cleanly identified with an expansion of non-neuronal cell type markers in the 1,020-gene set.

Beyond those advances, the 1,020-gene findings also successfully reproduced the cell types (and their spatial patterning) from the 160-gene findings, and further allowed discovery of multiple new gene markers for each cell type (e.g., *3110035E14Rik* for deep layers, *Cnot6l* for *Sst* neurons, and *Cplx1* for *Pvalb* neurons) (Fig. 5D and fig. S13). These molecularly-defined cell types were highly reproducible between biological replicates for 1,020-gene detection and were concordant with published single-cell RNA sequencing results (fig. S14). We further assessed the possibility of scaling up STARmap to accommodate higher gene numbers; while the STARmap scheme can encode and decode more than one million codes and the physical volume of mammalian cells is not limiting for amplification of more than 1,000 genes (fig. S15), the 1,020-gene experiments approached the upper limit of the optical volume of cells (fig. S15E); for those cases where more genes are needed, STARmap may cover the whole transcriptome with optical resolution enhanced by super-resolution microscopy (*38*, *39*) or the physical swelling typical of the hydrogel-tissue chemistries (*14*, *19*).

Adapting STARmap to thick tissue blocks for 3D analyses

In neuroscience, addressing the 3D complexity of both neurons and neural circuits has generally required development and use of thick tissue blocks or fully-intact brains for functional and structural readouts, including electrophysiology, imaging of activity, and analysis of morphology and connectivity. Therefore, for linking these readout measures from intact or semi-intact tissue preparations with cellular-resolution gene expression readouts from the very same preparations, methods of 3D spatial transcriptomic analysis in thick tissues have been long-sought, to achieve datastream registration as well as preserve 3D morphology and to obtain readouts from very much larger cell numbers (*2*). The initial experiments were carried out in brain slices no more than 1 cell body thick; we therefore next developed and tested STARmap to overcome limitations in diffusional access and imaging throughput for large tissue volumes, with a modified strategy for linearly reading out gene expression at cellular resolution to enable high-throughput molecular analysis in tissue volumes (Fig. 6A and fig. S16). Specificity and penetration depth of large-volume STARmap were tested initially using Thy1::YFP mouse brains, wherein STARmap successfully detected YFP mRNA across 150 μm of tissue thickness, and specifically co-localized YFP protein and mRNA at single-cell resolution (Fig. 6B) without labeling the tens of thousands of interspersed neighboring cells.

We then extended the spatial cell-typing of mouse primary visual cortex to more than 30,000 cells across volumes spanning all six layers and the corpus callosum. Using a curated gene set including 23 cell type markers and 5 ARGs read out over 7 cycles of linear SEDAL sequencing (Fig. 6, C and D, and fig. S17), we applied *K*-means clustering of marker genes (Methods) for each cell-type (recovering 11 cell types corresponding to the majority of those extracted by the 160-gene experiment). We found that 3D patterning of the 11 cell types (Fig. 6, E and F) was consistent with the 160-gene thin-section tissue findings, but provided a novel accurate and quantitative profiling of cellular distribution across space, with much larger cell numbers. As reflected by both spatial-histogram (Fig. 6E) and correlational analyses (fig. S17B), excitatory subtypes exhibited an unanticipated layered gradient distribution, with the spatial density of each subtype decaying across space into adjacent layers. In contrast, inhibitory subtypes were dispersed, albeit with layer preferences exhibited by the *Vip* subtype (largely located in layer 2/3), and the *Sst* and *Pvalb* subtypes (in layers 4 and 5). Non-neuronal cells were largely seen in layer 1 and white matter.

To discover yet-finer volumetric patterns, we further analyzed the distribution of distances from each individual cell from each sequencing-defined subtype to its nearest-neighbors, finding unexpectedly that the nearest neighbor of any inhibitory neuron tended to be its own subtype, rather than excitatory neurons or others inhibitory subtypes (Fig. 6G). If inhibitory neurons were randomly dispersed among the more abundant excitatory neurons in a purely salt-and-pepper distribution, the distance between inhibitory neurons would be larger than that from inhibitory to excitatory neurons (Fig. 6H). Remarkably instead, the actual intra-subtype distance of inhibitory neurons was much shorter (~15 μm, equivalent to the size of a single neuron, indicating direct somatic juxtaposition) (Fig. 6I), revealing a short-range self-clustering organization of inhibitory subtypes across volumes that could only be accurately measured in 3D, but not in 2D (fig. S18A). When guided by this initial STARmap observation, evidence for such patterning could be also obtained in transgenic mouse lines (fig. S18, B and C). This discovery bears considerable relevance to previous functional work; for example, electrophysiological studies have suggested that inhibitory neurons in spatial proximity tend to be connected by electric (gap) junctions important for setting up synchronized firing patterns (*40*, *41*), and in vivo imaging has suggested that inhibitory-neuron groupings in visual cortex could sharpen visual responses (*42*).

Discussion

STARmap defines a platform for 3D in situ transcriptomics, enabled by state-of-the-art DNA library preparation/sequencing and novel hydrogel-tissue chemistry. Here, STARmap was shown to be applicable to the study of molecularly-defined cell types and activity-regulated gene expression in mouse cortex, and to be scalable to larger 3D tissue blocks to visualize short- and long- range spatial organization of cortical neurons on a volumetric scale not previously accessible. In future work, STARmap may also be adapted to longer sequencing lengths or higher gene numbers; there is no intrinsic limit to the number of genes or RNA species that can be simultaneously and quantitatively accessed by STARmap (fig. S15); STARmap may also be capable of integrating cell type information with single-neuron morphology and projection anatomy (e.g., via Brainbow and MAPseq) (*43*, *44*) as well as with in vivo neural activity imaging and electrophysiology. This platform can also be generalized to study other heterogeneous cell populations in diverse tissues across the body, though the brain poses special challenges well suited to STARmap analysis. For example, the polymorphic ARG expression observed across different cell types is likely to depend on both intrinsic cell-biological properties (such as signal transduction pathway-component expression), and on extrinsic properties such as neural circuit anatomy that routes external sensory information to different cells (here in visual cortex). In general, it may not be possible to fully define brain cell typology independent of such 3D anatomy as well as activity patterns exhibited and experienced by cells during behavior; the nature of input and output communication pathways for the cells in question in fact can form the foundation for defining cell types (*35*, *36*). Toward this end, in situ transcriptomics exemplified by STARmap can effectively link this imaging-based molecular information with complementary cellular-resolution datastreams describing anatomy, natural activity, and causal significance, thus promising to fundamentally deepen our understanding of brain function and dysfunction (*2*).

Methods summary

All animal procedures followed animal care guidelines approved by Stanford University's Administrative Panel on Laboratory Animal Care (APLAC) and guidelines of the National Institutes of Health. For thin sections, animals were anesthetized and rapidly decapitated; the brain tissues were sliced using a cryostat. For thick sections, animals were anesthetized and transcardially perfused with PFA; the brain tissues were sliced using a vibratome. In STARmap experiments, tissues were hybridized with SNAIL probes, enzymatically amplified, hydrogel embedded, and sequentially imaged using SEDAL process and a confocal microscope. The resulting image datasets were registered across multiple cycles using the positions of all amplicons in each cycle and decoded. For cell-typing and single-cell gene expression analyses, the amplicons were attributed to individual cells based on segmentation images of fluorescent Nissl staining. All the detailed procedures for the experiments and data analyses are described in the supplementary materials.

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. 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
3. E. Lubeck, L. Cai, Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat. Methods* **9**, 743–748 (2012). doi:10.1038/nmeth.2069 Medline
4. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015). doi:10.1126/science.aaa6090 Medline
5. 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
6. 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
7. 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
8. A. P. Frei, F.-A. Bava, E. R. Zunder, E. W. Y. Hsieh, S.-Y. Chen, G. P. Nolan, P. F. Gherardini, Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat. Methods* **13**, 269–275 (2016). doi:10.1038/nmeth.3742 Medline
9. 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
10. M. Nagendran, D. P. Riordan, P. B. Harbury, T. J. Desai, Automated cell-type classification in intact tissues by single-cell molecular profiling. *eLife* **7**, e30510 (2018). doi:10.7554/eLife.30510 Medline
11. N. A. Peppas, J. Z. Hilt, A. Khademhosseini, R. Langer, Hydrogels in biology and medicine: From molecular principles to bionanotechnology. *Adv. Mater.* **18**, 1345–1360 (2006). doi:10.1002/adma.200501612
12. A. M. Rosales, K. S. Anseth, The design of reversible hydrogels to capture extracellular matrix dynamics. *Nat. Rev. Mater.* **1**, 1–15 (2016). Medline
13. R. Y. Tam, L. J. Smith, M. S. Shoichet, Engineering cellular microenvironments with photo- and enzymatically responsive hydrogels: Toward biomimetic 3D cell culture models. *Acc. Chem. Res.* **50**, 703–713 (2017). doi:10.1021/acs.accounts.6b00543 Medline
14. V. Gradinaru, J. Treweek, K. Overton, K. Deisseroth, Hydrogel-tissue chemistry: Principles and applications. *Annu. Rev. Biophys.* **47**, 355–376 (2018). doi:10.1146/annurev-biophys-070317-032905 Medline
15. K. Chung, J. Wallace, S.-Y. Kim, S. Kalyanasundaram, A. S. Andalman, T. J. Davidson, J. J. Mirzabekov, K. A. Zalocusky, J. Mattis, A. K. Denisin, S. Pak, H. Bernstein, C. Ramakrishnan, L. Grosenick, V. Gradinaru, K. Deisseroth, Structural and molecular interrogation of intact biological systems. *Nature* **497**, 332–337 (2013). doi:10.1038/nature12107 Medline
16. E. L. Sylwestrak, P. Rajasethupathy, M. A. Wright, A. Jaffe, K. Deisseroth, Multiplexed intact-tissue transcriptional analysis at cellular resolution. *Cell* **164**, 792–804 (2016). doi:10.1016/j.cell.2016.01.038 Medline
17. S. Shah, E. Lubeck, M. Schwarzkopf, T.-F. He, A. Greenbaum, C. H. Sohn, A. Lignell, H. M. T. Choi, V. Gradinaru, N. A. Pierce, L. Cai, Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development* **143**, 2862–2867 (2016). doi:10.1242/dev.138560 Medline
18. J. R. Moffitt, J. Hao, D. Bambah-Mukku, T. Lu, C. Dulac, X. Zhuang, High-performance multiplexed fluorescence *in situ* hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14456–14461 (2016). doi:10.1073/pnas.1617699113 Medline
19. 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
20. J. Shendure, G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra, G. M. Church, Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732 (2005). doi:10.1126/science.1117389 Medline
21. 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
22. R. Drmanac, A. B. Sparks, M. J. Callow, A. L. Halpern, N. L. Burns, B. G. Kermani, P. Carnevali, I. Nazarenko, G. B. Nilsen, G. Yeung, F. Dahl, A. Fernandez, B. Staker, K. P. Pant, J. Baccash, A. P. Borcherding, A. Brownley, R. Cedeno, L. Chen, D. Chernikoff, A. Cheung, R. Chirita, B. Curson, J. C. Ebert, C. R. Hacker, R. Hartlage, B. Hauser, S. Huang, Y. Jiang, V. Karpinchyk, M. Koenig, C. Kong, T. Landers, C. Le, J. Liu, C. E. McBride, M. Morenzoni, R. E. Morey, K. Mutch, H. Perazich, K. Perry, B. A. Peters, J. Peterson, C. L. Pethiyagoda, K. Pothuraju, C. Richter, A. M. Rosenbaum, S. Roy, J. Shafto, U. Sharanhovich, K. W. Shannon, C. G. Sheppy, M. Sun, J. V. Thakuria, A. Tran, D. Vu, A. W. Zaranek, X. Wu, S. Drmanac, A. R. Oliphant, W. C. Banyai, B. Martin, D. G. Ballinger, G. M. Church, C. A. Reid, Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science* **327**, 78–81 (2010). doi:10.1126/science.1181498 Medline
23. L. L. Glickfeld, R. C. Reid, M. L. Andermann, A mouse model of higher visual cortical function. *Curr. Opin. Neurobiol.* **24**, 28–33 (2014). doi:10.1016/j.conb.2013.08.009 Medline
24. 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
25. A. Zeisel, A. B. Muñoz-Manchado, S. Codeluppi, P. Lönnerberg, G. La Manno, A. Juréus, S. Marques, H. Munguba, L. He, C. Betsholtz, C. Rolny, G. Castelo-Branco, J. Hjerling-Leffler, S. Linnarsson, Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015). doi:10.1126/science.aaa1934 Medline
26. T. K. Kim, M. Hemberg, J. M. Gray, A. M. Costa, D. M. Bear, J. Wu, D. A. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P. F. Worley, G. Kreiman, M. E. Greenberg, Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010). doi:10.1038/nature09033 Medline
27. A. R. Mardinly, I. Spiegel, A. Patrizi, E. Centofante, J. E. Bazinet, C. P. Tzeng, C. Mandel-Brehm, D. A. Harmin, H. Adesnik, M. Fagiolini, M. E. Greenberg, Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. *Nature* **531**, 371–375 (2016). doi:10.1038/nature17187 Medline

28. S. Hrvatin, D. R. Hochbaum, M. A. Nagy, M. Cicconet, K. Robertson, L. Cheadle, R. Zilionis, A. Ratner, R. Borges-Monroy, A. M. Klein, B. L. Sabatini, M. E. Greenberg, Single-cell analysis of experience-dependent transcriptomic states in the mouse visual cortex. *Nat. Neurosci.* **21**, 120–129 (2018). doi:10.1038/s41593-017-0029-5 Medline
29. 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
30. K. Shekhar, S. W. Lapan, I. E. Whitney, N. M. Tran, E. Z. Macosko, M. Kowalczyk, X. Adiconis, J. Z. Levin, J. Nemesh, M. Goldman, S. A. McCarroll, C. L. Cepko, A. Regev, J. R. Sanes, Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323.e30 (2016). doi:10.1016/j.cell.2016.07.054 Medline
31. J. Y. Joo, K. Schaukowitch, L. Farbiak, G. Kilaru, T.-K. Kim, Stimulus-specific combinatorial functionality of neuronal *c-fos* enhancers. *Nat. Neurosci.* **19**, 75–83 (2016). doi:10.1038/nn.4170 Medline
32. M. Carlén, What constitutes the prefrontal cortex? *Science* **358**, 478–482 (2017). doi:10.1126/science.aan8868 Medline
33. H. B. Uylings, H. J. Groenewegen, B. Kolb, Do rats have a prefrontal cortex? *Behav. Brain Res.* **146**, 3–17 (2003). doi:10.1016/j.bbr.2003.09.028 Medline
34. S. Lee, J. Hjerling-Leffler, E. Zagha, G. Fishell, B. Rudy, The largest group of superficial neocortical GABAergic interneurons expresses ionotropic serotonin receptors. *J. Neurosci.* **30**, 16796–16808 (2010). doi:10.1523/JNEUROSCI.1869-10.2010 Medline
35. A. Paul, M. Crow, R. Raudales, M. He, J. Gillis, Z. J. Huang, Transcriptional architecture of synaptic communication delineates GABAergic neuron identity. *Cell* **171**, 522–539.e20 (2017). doi:10.1016/j.cell.2017.08.032 Medline
36. T. N. Lerner, L. Ye, K. Deisseroth, Communication in neural circuits: Tools, opportunities, and challenges. *Cell* **164**, 1136–1150 (2016). doi:10.1016/j.cell.2016.02.027 Medline
37. L. Ye, W. E. Allen, K. R. Thompson, Q. Tian, B. Hsueh, C. Ramakrishnan, A.-C. Wang, J. H. Jennings, A. Adhikari, C. H. Halpern, I. B. Witten, A. L. Barth, L. Luo, J. A. McNab, K. Deisseroth, Wiring and molecular features of prefrontal ensembles representing distinct experience. *Cell* **165**, 1776–1788 (2016). doi:10.1016/j.cell.2016.05.010 Medline
38. M. J. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006). doi:10.1038/nmeth929 Medline
39. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006). doi:10.1126/science.1127344 Medline
40. J. R. Gibson, M. Beierlein, B. W. Connors, Two networks of electrically coupled inhibitory neurons in neocortex. *Nature* **402**, 75–79 (1999). doi:10.1038/47035 Medline
41. Y. Amitai, J. R. Gibson, M. Beierlein, S. L. Patrick, A. M. Ho, B. W. Connors, D. Golomb, The spatial dimensions of electrically coupled networks of interneurons in the neocortex. *J. Neurosci.* **22**, 4142–4152 (2002). doi:10.1523/JNEUROSCI.22-10-04142.2002 Medline
42. T. Ebina, K. Sohya, I. Imayoshi, S.-T. Yin, R. Kimura, Y. Yanagawa, H. Kameda, H. Hioki, T. Kaneko, T. Tsumoto, 3D clustering of GABAergic neurons enhances inhibitory actions on excitatory neurons in the mouse visual cortex. *Cell Reports* **9**, 1896–1907 (2014). doi:10.1016/j.celrep.2014.10.057 Medline
43. J. Livet, T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes, J. W. Lichtman, Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56–62 (2007). doi:10.1038/nature06293 Medline
44. J. M. Kebschull, P. Garcia da Silva, A. P. Reid, I. D. Peikon, D. F. Albeanu, A. M. Zador, High-throughput mapping of single-neuron projections by sequencing of barcoded RNA. *Neuron* **91**, 975–987 (2016). doi:10.1016/j.neuron.2016.07.036 Medline
45. A. McDavid, G. Finak, P. K. Chattopadyay, M. Dominguez, L. Lamoreaux, S. S. Ma, M. Roederer, R. Gottardo, Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* **29**, 461–467 (2013). doi:10.1093/bioinformatics/bts714 Medline
46. J. L. Bentley, Multidimensional binary search trees used for associative searching. *Commun. ACM* **18**, 509–517 (1975). doi:10.1145/361002.361007
47. C. Larsson, I. Grundberg, O. Söderberg, M. Nilsson, In situ detection and genotyping of individual mRNA molecules. *Nat. Methods* **7**, 395–397 (2010). doi:10.1038/nmeth.1448 Medline
48. E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, S. A. McCarroll, Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015). doi:10.1016/j.cell.2015.05.002 Medline
49. 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

ACKNOWLEDGMENTS

We are grateful for the efforts and support of Ailey Crow, Sean Quirin, Adelaida Chibukhchyan, Connie Lee, Maisie Lo, and Cynthia Delacruz. We thank Matthew Lovett-Barron for helping with image registration, and Je H. Lee (Cold Spring Harbor Laboratory), Jing Yan (Princeton University), Noah D. Donoghue (Brown University) and Qing Dai (University of Chicago) for advice. We thank Li Ye, Aaron Andalman, Huiliang Wang and Brian Hsueh for discussion. We also thank Prof. Liqun Luo for suggestions on the manuscript. **Funding:** X.W. is supported by a Life Science Research Foundation fellowship and the Gordon and Betty Moore Foundation. W.E.A. is supported by a Fannie and John Hertz Foundation Fellowship and an NSF Graduate Research Fellowship. E.L.S. is supported by an NIMH Ruth L. Kirschstein NRSA fellowship (1F32MH110144-01). M.A.W. is supported by an NIMH Career Development Award (1K08MH113039). J.L. is supported by a Bio-X Interdisciplinary Initiatives Seed Grant. G.P.N., N.S., and F.-A. B. were supported by the Parker Institute for Cancer Immunotherapy, the FDA, and the NIH; F.-A. B was supported by the Human Frontiers Science Program. F.-A. B.'s current affiliation is Institut Curie, PSL Research University, Université Paris Sud, Université Paris-Saclay, Centre Universitaire, CNRS UMR 3348, Orsay 91405, France. K.D. is supported by NIMH (R01MH099647), NIDA (P50DA042012), the DARPA NeuroFAST program W911NF-14-2-0013, the NSF NeuroNex program, the Gatsby Foundation, the AE Foundation, the NOMIS Foundation, the Fresenius Foundation, the Wiegers Family Fund, the James Grosfeld Foundation, the Sam and Betsy Reeves Foundation, and the HL Snyder Foundation. **Author contributions:** X.W. and K.D. initiated the STARmap project to integrate hydrogel-tissue chemistry with in situ sequencing. X.W. developed the STARmap hydrogel-tissue chemistry, SEDAL sequencing, and STARmap hardware, and conducted the experiments. W.E.A. developed the STARmap software and pipeline, and analyzed the sequencing data. F.A. B., N.S. and G. P. N. designed the initial versions of the SNAIL system. X.W. and F.A.B. developed the SNAIL process for mouse brain tissue and compared SNAIL with other in situ methods. M.A.W. and E.L.S. conducted animal behavior and preparation of mouse brain tissue, and contributed valuable advice. M.A.W. compared Nissl staining with other segmentation methods. W.E.A. and E.L.S. designed the visual-stimulus and cocaine-stimulus procedures. S.V. generated the CLARITY data with PV transgenicmice. K. E. and C. R. contributed to preparation of cell cultures. C. L. assisted with experiments. J.L. assisted with experiments and graphic design. K.D. supervised all aspects of the work. X.W., W.E.A., and K.D. interpreted the STARmap data, designed and prepared the figures, and wrote the manuscript with edits from all authors. **Competing interests:** The design, steps and applications of STARmap are covered in pending patent application material from Stanford University; all methods, protocols, and sequences are freely available to nonprofit institutions and investigators. **Data and materials availability:** All data are available in the main text or the supplementary materials. Computational tools/code and other materials are available at http://clarityresourcecenter.org.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aat5691/DC1
Materials and Methods
Figs. S1 to S18
References (*45*–*49*)
Tables S1 and S2

13 March 2018; resubmitted 1 June 2018
Accepted 13 June 2018
Published online 21 June 2018
10.1126/science.aat5691

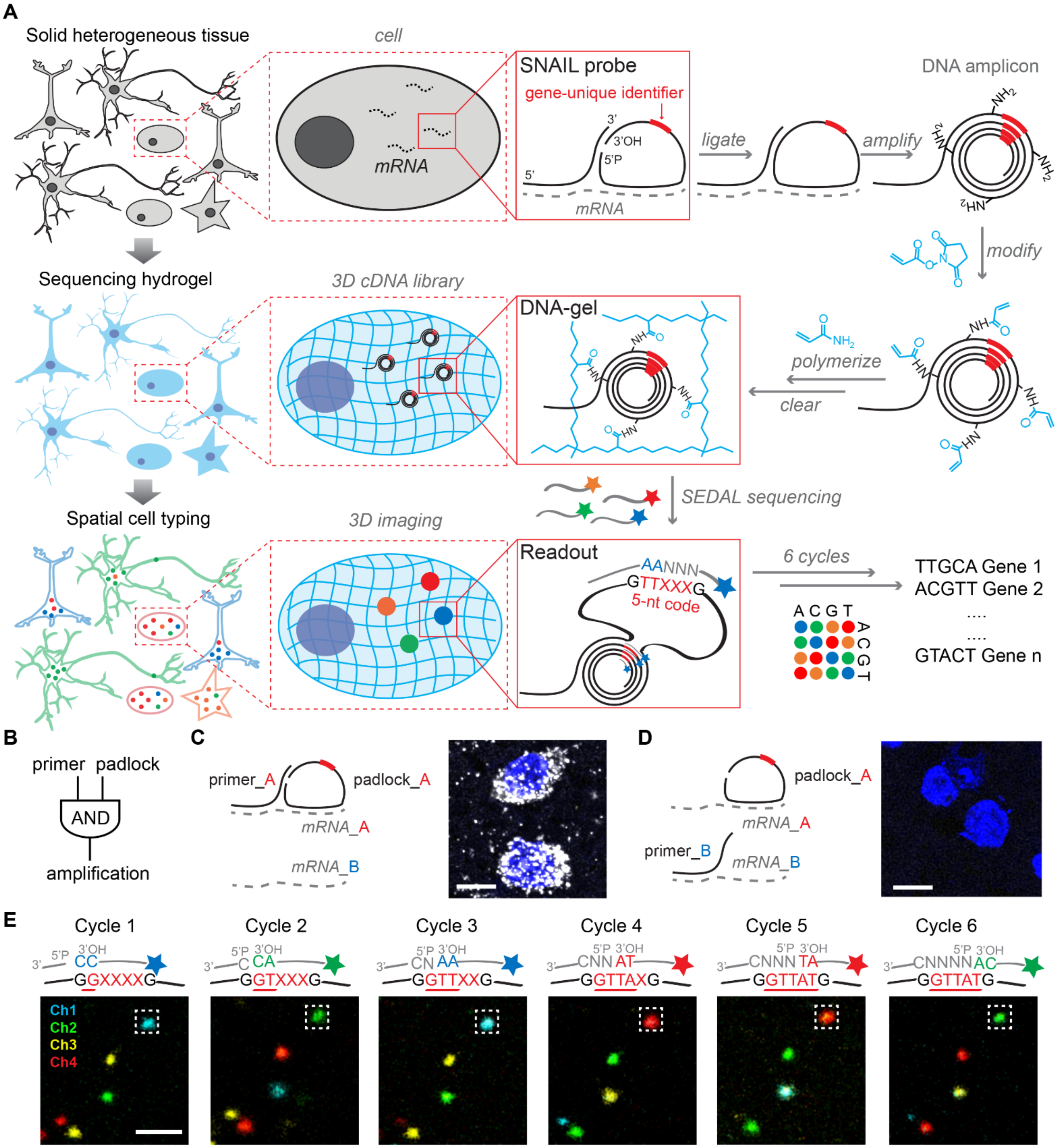
A
Solid heterogeneous tissue
cell
mRNA
SNAIL probe
gene-unique identifier
3'
3'OH
5'P
5'
mRNA
ligate
amplify
DNA amplicon
NH_2
modify
Sequencing hydrogel
3D cDNA library
DNA-gel
polymerize
clear
SEDAL sequencing
Spatial cell typing
3D imaging
Readout
AANNN
GTTXXXG
5-nt code
6 cycles
TTGCA Gene 1
ACGTT Gene 2
....
....
GTACT Gene n
A C G T
A C G T
B
primer padlock
AND
amplification
C
primer_A
padlock_A
mRNA_A
mRNA_B
D
padlock_A
mRNA_A
primer_B
mRNA_B
E
Cycle 1
Cycle 2
Cycle 3
Cycle 4
Cycle 5
Cycle 6
3' 5'P 3'OH CC
GGXXXXG
3' C 5'P 3'OH CA
GGTXXXG
3' CN 5'P 3'OH AA
GGTTXXG
3' CNN 5'P 3'OH AT
GGTTAXG
3' CNNN 5'P 3'OH TA
GGTTATG
3' CNNNN 5'P 3'OH AC
GGTTATG
Ch1
Ch2
Ch3
Ch4

Fig. 1. STARmap principles: in situ RNA sequencing for spatial transcriptomics within the 3D tissue environment. (**A**) STARmap overview schematic. After brain tissue is prepared (see Methods for mouse brain protocols), the custom SNAIL probes that encounter and hybridize to intracellular mRNAs (dashed lines) within the intact tissue are enzymatically replicated as cDNA amplicons. The amplicons are constructed in situ with an acrylic acid N-hydroxysuccinimide moiety modification (blue) and then copolymerized with acrylamide to embed within a hydrogel network (blue wavy lines), following by clearance of unbound lipids and proteins (fig. S2). Each SNAIL probe contains a gene-unique identifier segment (red) which is read-out through in situ sequencing with 2-base encoding for error correction (SEDAL) (fig. S3). Finally, highly multiplexed RNA quantification in 3D reveals gene expression and cell types in space. (**B**) SNAIL logic: a pair of primer and padlock probes amplifies target-specific signals and excludes noise known to commonly arise from non-specific hybridization of a single probe. (**C** and **D**) Only adjacent binding of primer and padlock probes leads to signal amplification. mRNA A represents *Gapdh* and mRNA B represents *Actb*. Both fluorescent images showing *Gapdh* (gray) mRNA and cell nuclei (blue) labeling in mouse brain slice; note the absence of labeling with mismatched primer and padlock (right). Scale bar, 10 μm. (**E**) In situ sequencing of DNA amplicons in the tissue-hydrogel complex via SEDAL, the novel sequencing-by-ligation method devised for STARmap: for each cycle, the reading probes (gray line without star-symbol label) contain an incrementally increasing-length run of degenerate bases (N representing an equal mixture of A, T, C and G) with phosphate at the 5′ end (5′P) to set the reading position; the decoding probes (gray line with star-symbol label) are labeled by fluorophores with color coding for the dinucleotide at the 3′ end. Only if both probes are perfectly complementary to the DNA template (black lower sequence), the two kinds of probes can then be ligated) to form a stable product with high melting-temperature, allowing later imaging after unligated probes are washed away. After each imaging cycle, probes are stripped away from the robust tissue-hydrogel using 60% formamide so that the next cycle can begin. X: unknown base to be read; red underline: decoded sequence; Ch1-4: fluorescence channels. Scale bar, 2 μm.

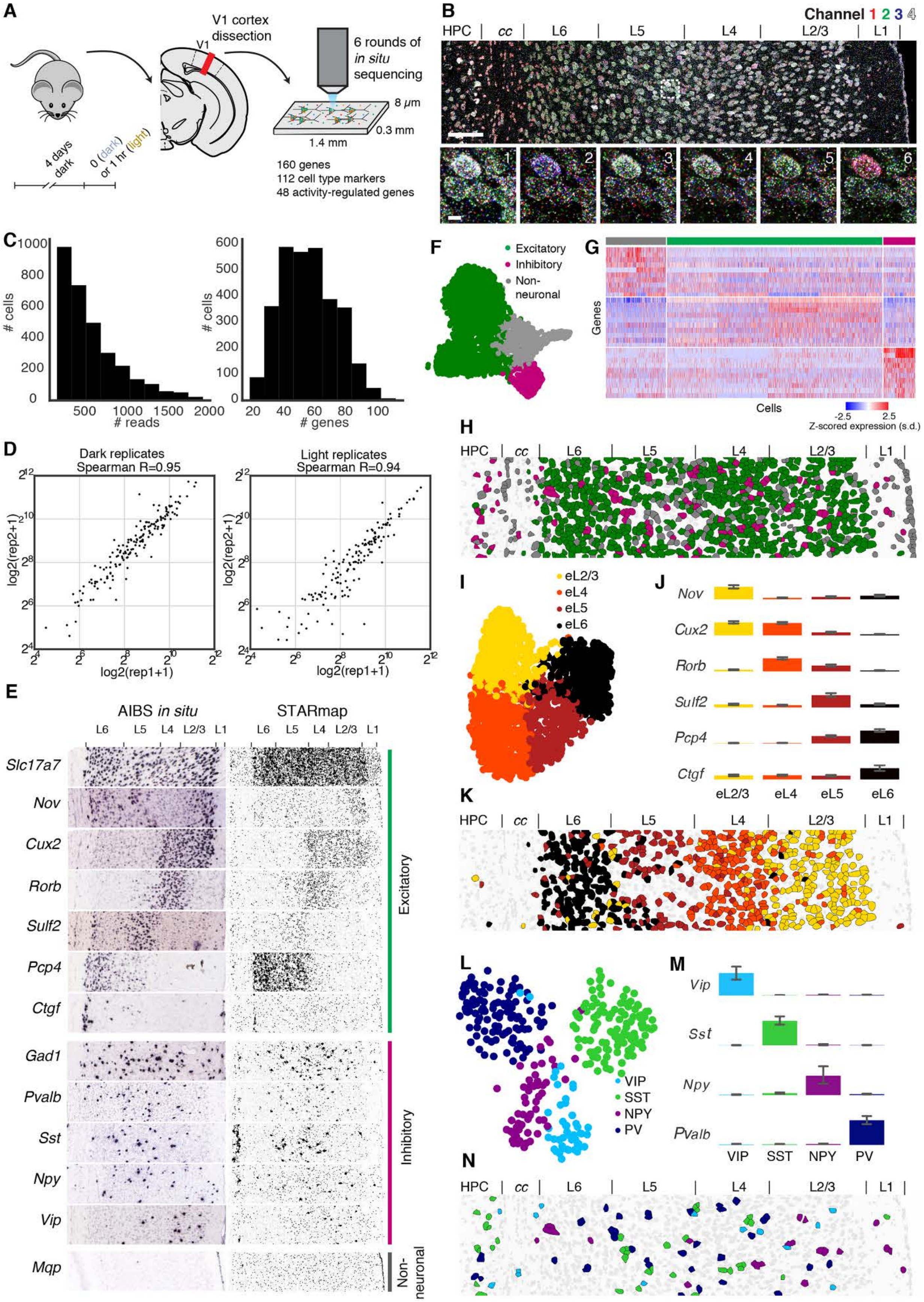
A
V1 cortex dissection
V1
6 rounds of *in situ* sequencing
8 μm
0.3 mm
1.4 mm
4 days dark
0 (dark) or 1 hr (light)
160 genes
112 cell type markers
48 activity-regulated genes
B
Channel 1 2 3 4
HPC | cc | L6 | L5 | L4 | L2/3 | L1 |
1 2 3 4 5 6
C
cells
reads
genes
D
Dark replicates Spearman R=0.95
Light replicates Spearman R=0.94
log2(rep2+1)
log2(rep1+1)
E
AIBS *in situ*
STARmap
L6 L5 L4 L2/3 L1
Slc17a7
Nov
Cux2
Rorb
Sulf2
Pcp4
Ctgf
Gad1
Pvalb
Sst
Npy
Vip
Mqp
Excitatory
Inhibitory
Non-neuronal
F
Excitatory
Inhibitory
Non-neuronal
G
Genes
Cells
-2.5 2.5
Z-scored expression (s.d.)
H
HPC | cc | L6 | L5 | L4 | L2/3 | L1 |
I
eL2/3
eL4
eL5
eL6
J
Nov
Cux2
Rorb
Sulf2
Pcp4
Ctgf
eL2/3 eL4 eL5 eL6
K
HPC | cc | L6 | L5 | L4 | L2/3 | L1 |
L
VIP
SST
NPY
PV
M
Vip
Sst
Npy
Pvalb
VIP SST NPY PV
N
HPC | cc | L6 | L5 | L4 | L2/3 | L1 |

Fig. 2. STARmapping cell types in primary visual cortex. (**A**) Experimental design; mice were dark housed, prior to sacrifice, for 4 days and then either kept in the dark or exposed to light for 1 hour. Primary visual cortex (V1) was coronally sectioned and RNAs of 112 cell type markers and 48 activity-regulated genes were quantified by STARmap. (**B**) Raw fluorescence images of in-process STARmap with the full view of Cycle 1 (top) and zoomed views across all six cycles (bottom). Full field: 1.4 mm × 0.3 mm, scale bar, 100 μm; zoomed region: 11.78 μm × 11.78 μm, scale bar, 2 μm; Channel: color code for the four fluorescence channels; L1-6: the six neocortical layers; *cc*: corpus callosum; HPC: hippocampus. (**C**) Histograms: detected reads (DNA amplicons) per cell (left), and genes per cell (right). (**D**) Quantitative reproducibility of biological replicates, whether in the light or dark condition: $\log_2$(amplicon quantity) for 160 genes across the whole imaging region plotted. Rep1: expression value in first replicate, rep2: expression value in second replicate. (**E**) Validation of STARmap: left column, in situ images from Allen Institute of Brain Science (AIBS); right column, RNA pattern of individual genes extracted from 160-gene STARmap, which reliably reproduced the spatial gene expression pattern from AIBS. (**F**) Uniform Manifold Approximation (UMAP) plot, a non-linear dimensionality reduction technique used to visualize the similarity of cell transcriptomes in two dimensions, showing consistent clustering of major cell types across 3,142 cells pooled from four biological replicates: 2,199 excitatory neurons, 324 inhibitory neurons and 619 non-neuronal cells. (**G**) Gene expression heatmap for 112 cell-type markers aligned with each cell cluster, showing clustering by inhibitory, excitatory, or non-neuronal cell types. Expression for each gene is z-scored across all genes in each cell. (**H**) Representative cell-resolved spatial map in neocortex and beyond: cell-types color-coded as in panel (F). (**I** to **N**) Clustering of excitatory and inhibitory subtypes: UMAP plots [(I) and (L)], bar plots of representative genes [(J) and (M)] (mean ± 95% confidence interval expression across all cells in that cluster, with each bar scaled to the maximum mean expression across all clusters), and in situ spatial distribution [(K) and (N)] of excitatory [(I) to (K)] and inhibitory [(L) to (N)] neurons. The number of cells in each cluster was as follows: eL2/3: 589; eL4: 649; eL5: 393; eL6: 368; PV neurons: 111; VIP neurons: 46; SST neurons: 46; NPY neurons: 56. Inclusion of cells in clusters was guided entirely by amplicon representation in each cell without using spatial information; excitatory cell clusters were then named according to the spatial layering observed for that cluster, while inhibitory cell clusters were named according to the dominant cell-type amplicon based on the strong segregation of amplicon markers.

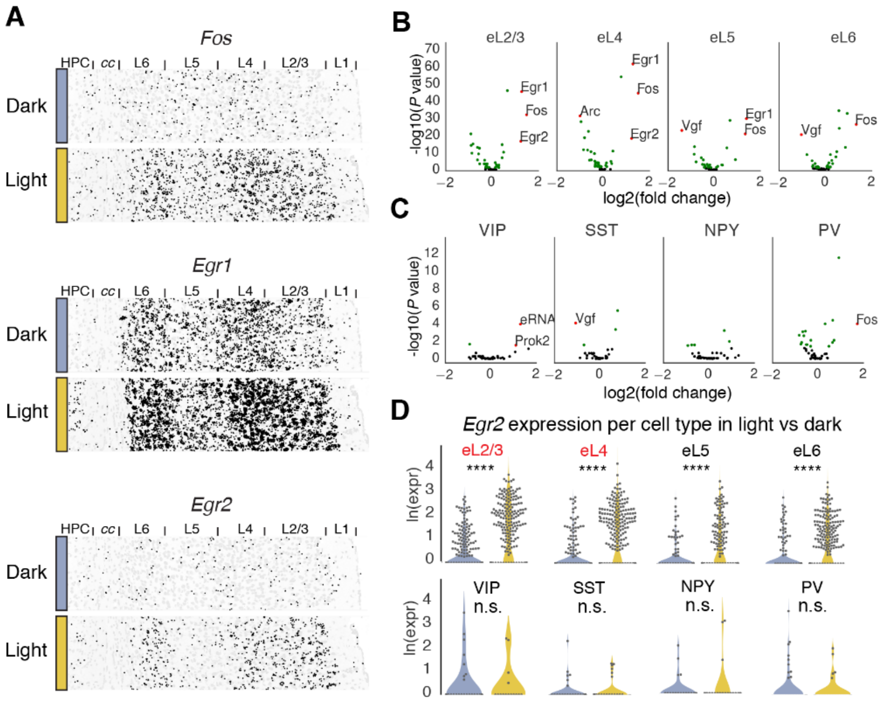

Fig. 3. STARmapping behavioral experience: detecting and quantifying cell type-specific regulation of activity-regulated genes (ARGs). (**A**) Validation: spatial expression pattern in visual cortex of prototypical ARGs known as immediate early genes (IEGs): sacrifice was in darkness or after 1 hour light exposure. (**B** and **C**) Volcano plots of log fold-change in gene expression between light and dark conditions in inhibitory and excitatory cell types. Genes with significantly increased or decreased expression (false discovery rate adjusted *P* value < 0.05, Wilcoxon rank-sum test) are labeled in green and the most significantly changed genes (P value <0.05 and fold change > 2) are labeled in red. Many ARGs showed cell-type specificity pointing to discovery of unanticipated cell type-specific logic of excitation-transcription coupling. (**D**) Violin plot of *Egr2* expression by cell type. ****$P < 0.0001$, n.s. not significant, Wilcoxon rank-sum test; red-labeled cell types, fold change >2.

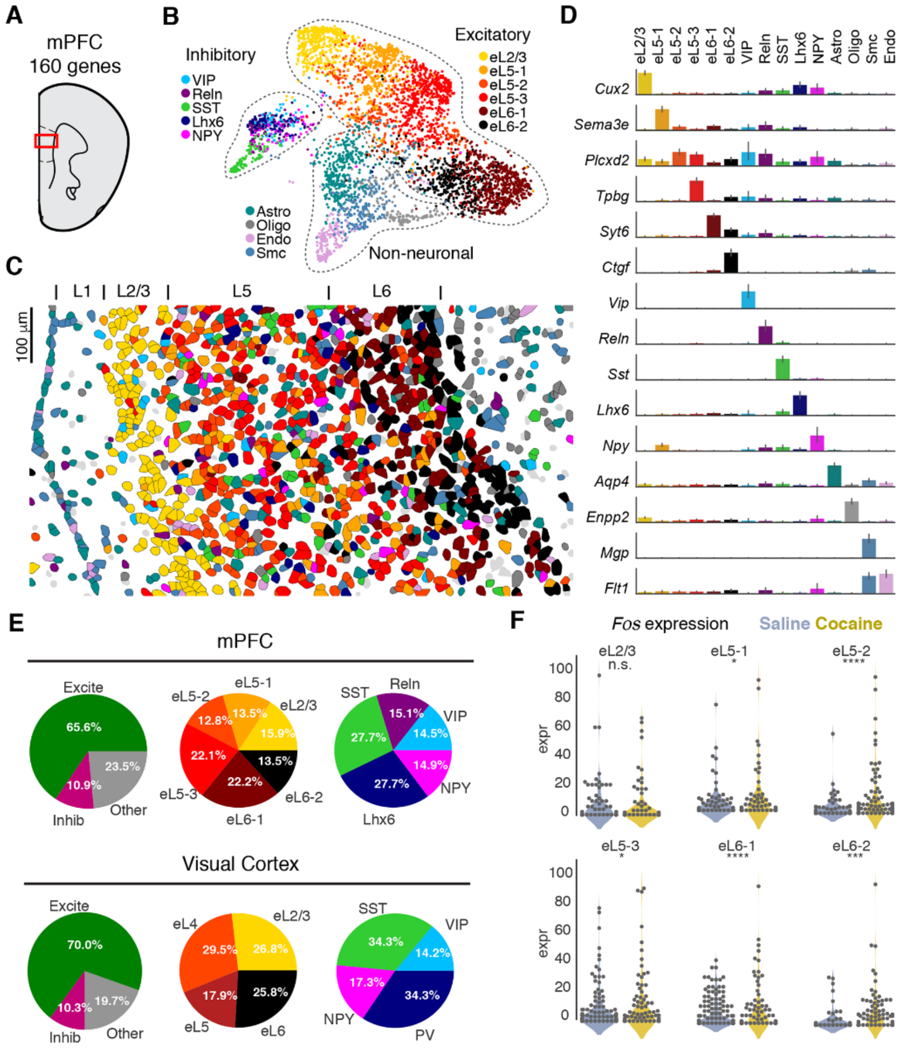

Fig. 4. STARmapping cell types and neural activity in medial prefrontal cortex. (**A**) Diagram of targeted region (red box) containing prelimbic cortex (PrL) within medial prefrontal cortex (mPFC). (**B**) UMAP visualization of all inhibitory (VIP, Reln, SST, Lhx6, NPY), excitatory (eL2/3, eL5-1, eL5-2, eL5-3, eL6-1, eL6-2), and non-neuronal (Astro, Oligo, Smc, Endo) cell types. (**C**) Spatial visualization of cell type layout in mPFC, using the same color scheme as in (B). (**D**) Barplot of representative genes per cluster (mean ± 95% confidence interval), with each bar scaled to the maximum mean expression for that gene across clusters. (**E**) Piecharts showing the relative proportion of each major and minor cell type in both mPFC and visual cortex. (**F**) Violin plots of *Fos* gene induction in different excitatory cell types in mPFC in response to cocaine. The mice were sacrificed after one hour of cocaine or saline injection. Expr = normalized expression. n.s. not significant, $^*P < 0.05$, $^{***}P < 0.001$, $^{****}P < 0.0001$, likelihood ratio test. Astro, astrocytes; Oligo, oligodendrocytes; Smc, smooth muscle cells; Endo, endothelial cells.

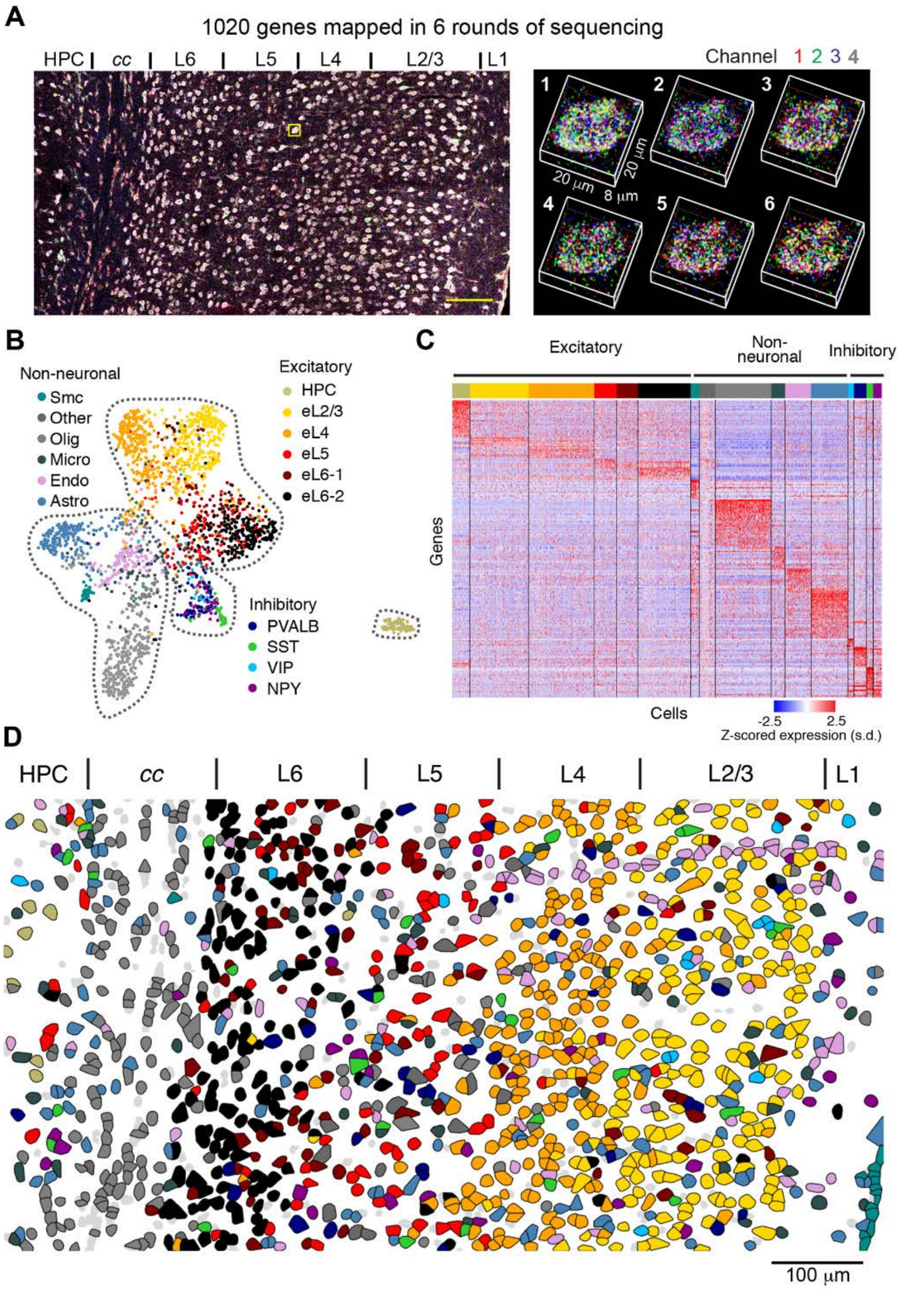

Fig. 5. Simultaneous mapping of 1,020 genes in primary visual cortex by STARmap. (**A**) Input fluorescence data. Left: maximum-intensity projection of the first sequencing round for 1020 gene experiment, showing all four channels simultaneously; yellow square, zoom region; scale bar, 100 μm. Right: zoom into a single cell showing spatial arrangement of amplicons in 3D across six sequencing rounds. (**B**) Joint UMAP plot showing all excitatory (HPC, eL2/3, eL4, eL5, eL6-1, eL6-2), non-neuronal (Smc, Other, Olig, Micro, Endo, Astro), and inhibitory (PVALB, SST, VIP, NPY) cell types. (**C**) Plot of all differentially expressed genes across every cluster, with $P < 10^{-12}$ and log fold change > 1.5. (**D**) Spatial map of all excitatory, non-neuronal, and inhibitory cell types in visual cortex using the same color code of (B). HPC, hippocampus; Smc: smooth muscle cells; Other, other unclassified cells; Oligo, oligodendrocytes; Micro, microglia; Endo, endothelia cells; Astro, astrocytes.

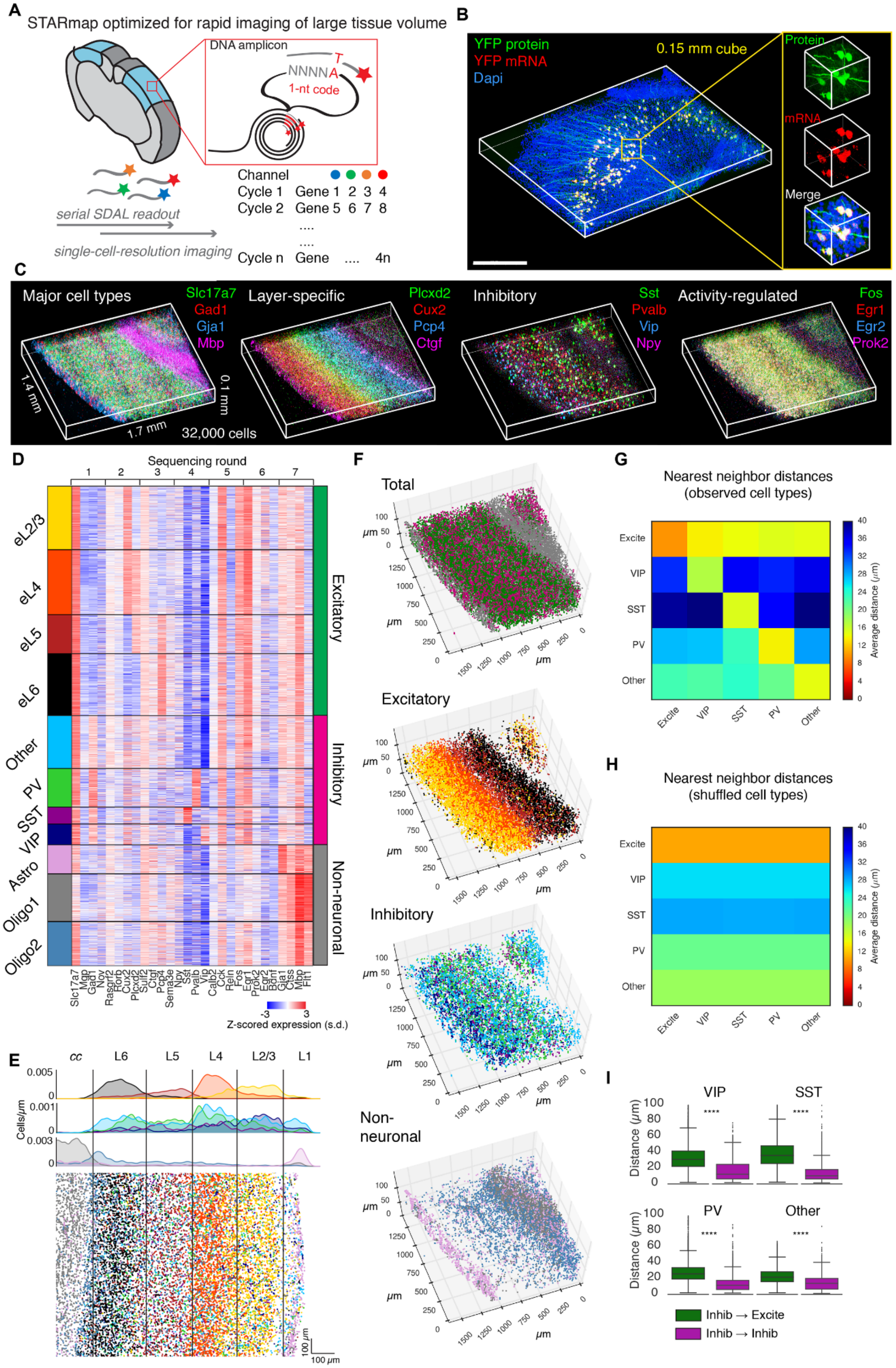
A
STARmap optimized for rapid imaging of large tissue volume
DNA amplicon
NNNNA
T
1-nt code
Channel
Cycle 1 Gene 1 2 3 4
Cycle 2 Gene 5 6 7 8
....
....
Cycle n Gene 4n
serial SDAL readout
single-cell-resolution imaging
B
YFP protein
YFP mRNA
Dapi
0.15 mm cube
Protein
mRNA
Merge
C
Major cell types
Slc17a7
Gad1
Gja1
Mbp
1.4 mm
1.7 mm
0.1 mm
32,000 cells
Layer-specific
Plcxd2
Cux2
Pcp4
Ctgf
Inhibitory
Sst
Pvalb
Vip
Npy
Activity-regulated
Fos
Egr1
Egr2
Prok2
D
Sequencing round
1 2 3 4 5 6 7
eL2/3
eL4
eL5
eL6
Other
PV
SST
VIP
Astro
Oligo1
Oligo2
Excitatory
Inhibitory
Non-neuronal
Slc17a7 Mgp Gad1 Nov Rasgrf2 Rorb Cux2 Plcxd2 Sulf2 Ctgf Pcp4 Sema3e Npy Sst Pvalb Vip Calb2 Cck Reln Fos Egr1 Prok2 Egr2 Bdnf Gja1 Ctss Mbp Flt1
-3 3
Z-scored expression (s.d.)
E
cc L6 L5 L4 L2/3 L1
Cells/μm
0.005
0
0.001
0
0.003
0
100 μm
100 μm
F
Total
Excitatory
Inhibitory
Non-neuronal
μm
100
50
0
1250
1000
750
500
250
0
1500 1250 1000 750 500 250 0
G
Nearest neighbor distances
(observed cell types)
H
Nearest neighbor distances
(shuffled cell types)
Excite
VIP
SST
PV
Other
Average distance (μm)
0 4 8 12 16 20 24 28 32 36 40
I
VIP
SST
PV
Other
Distance (μm)
0 20 40 60 80 100
****
Inhib → Excite
Inhib → Inhib

Fig. 6. Three-dimensional architecture of cell types in visual cortex volumes. (**A**) Volumetric STARmapping via sequential SEDAL gene readout. Using a modified STARmap procedure (fig. S16) and cyclic gene readout (4 genes in each cycle), large tissue volumes can be rapidly mapped at single-cell resolution without oversampling each amplicon. (**B**) Validation showing specific STARMAP labeling of YFP-expressing neurons (from transgenic Thy1::YFP mouse line) in 3D cortical volume. Scale bar, 0.5 mm. (**C**) Representative labeling of major cell types (left), layer-specific markers (left center), inhibitory markers (right center), and activity-regulated genes (right) acquired over multiple rounds in visual cortex STARmap volumes. (**D**) Per-cell expression matrix of 28 genes from 32,845 single cells from one volume clustered into multiple excitatory, inhibitory, and non-neuronal cell types, z-scored across genes for each cell in order to normalize for mean differences in total signal between cells. Columns are sorted by order of sequencing rounds as conducted, in groups of 4. (**E**) Top: Spatial histograms of excitatory, inhibitory, and non-neuronal cell types using same color labels as (D). Cells were counted in 5 μm bins in a 2D max-projection, and plotted in cell count/μm units as a function of distance from the corpus callosum (cc) to pia, averaged across the bins perpendicular to the cortical layers. Bottom: plot of max-projected cell locations color coded by cluster as in (D). (**F**) Spatial distribution of each cell type (excitatory, inhibitory, non-neuronal) and subtypes in three dimensions. Each dot represents a single cell; spatial dimensions are in μm. (**H**) Average nearest-neighbor distances computed in 3D between all excitatory cells (Excite) and each inhibitory cell type. For self-comparisons, the nearest neighbor was defined as the closest non-identical cell; note the persistent self-correlation revealing self-clustering of inhibitory subtypes. (**I**) Same distances as (H) but using shuffled (randomized) cell type labels. (**J**) Nearest-neighbor distances computed in 3D between each inhibitory cell of a certain type and any member of the same type (Inhib → Inhib, eg VIP → VIP) or any excitatory neuron (Inhib → Excite). **** $P < 0.0001$, Wilcoxon rank-sum test.

Science

Three-dimensional intact-tissue sequencing of single-cell transcriptional states

Xiao Wang, William E. Allen, Matthew A. Wright, Emily L. Sylwestrak, Nikolay Samusik, Sam Vesuna, Kathryn Evans, Cindy Liu,
Charu Ramakrishnan, Jia Liu, Garry P. Nolan, Felice-Alessio Bava and Karl Deisseroth

---

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.