

# Neurobiology: gene expression captured on-site

Vivien Marx

A wealth of spatial techniques capture the ‘where’ of gene expression.

Brain tissue in vertebrates is mighty complex: many different types of cells connect and relay signals in a hum of activity that lets birds fly, dogs play and humans dream. To characterize this activity, researchers can apply electrophysiology or imaging. By also leveraging molecular approaches such as gene expression analysis or targeted gene activation with optogenetics techniques, they hope to get ‘what happens where’ data to discern which cells express which genes and how a certain positioning of cells or cell types matters. Recent methods to profile gene expression at its point of origin add to these data. Researchers anticipate that such spatial transcriptomic information will lead to a better, *in situ* understanding of cells and cell types in brain tissue and shed light on the way cells team up in neural circuits. Among the challenges are cellular heterogeneity and the multitude of cell types in the brain. The gain from spatial information is the potential for targeting neuronal populations.

Spatial transcriptomic techniques are still maturing, and viewpoint exchanges about tool selection and data interpretation are openly discussed. Despite their differing views, scientists speaking with *Nature Methods* share an optimistic view of the role these methods can play in neurobiology.

## Many flavors

Spatial transcriptomics can let researchers assay many cells, perhaps tens of thousands of them at a time, in their original location within the tissue and thereby capture cells “possibly closer to their native states,” says Hongkui Zeng, who directs large-scale projects at the Allen Institute for Brain Science that include assay pipelines for building atlases and other community resources. This native state can change as a result of, for example, a learning experi-

ment. Scientists want to explore the changes in these very cells *in situ* in terms of both function and gene expression.

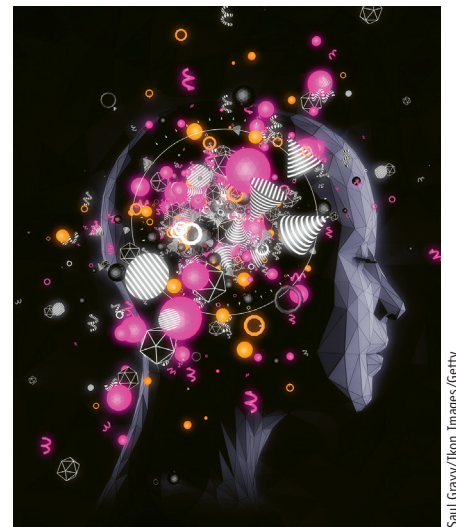
Zeng and colleagues have established a single-cell RNA sequencing (scRNA-seq) pipeline to profile isolated single cells or single nuclei from mouse and human brains. The researchers now want to scale up spatial transcriptomic approaches—*in situ* hybridization or sequencing—and to do so, she says, “we are eagerly waiting for the technologies to become more mature.” There are “many flavors” of techniques to choose from, and selection needs to be tailored to a specific research goal. “Even at Allen we are considering several different approaches for different purposes,” she says.

In some cases, Zeng and her team need a method for simultaneously profiling many genes, perhaps hundreds of them, to then decide which cell types make up the tissue. For other work, the scientists seek to simultaneously profile a dozen or so genes in a sample with few labeled cells and want to then run the method repeatedly on hundreds, even thousands, of samples. She does not wish to “make generic comments about one technique being better than another,” or indicate which techniques she might have chosen thus far, but she acknowledges that the research community closely watches the Allen’s choices. That makes her “extra careful” also because she and her team want to generate data applicable in many labs that get “as close to the basic ground truth as possible.”



Allen Institute for Brain Science

Spatial transcriptomics can let researchers assay many cells in their original location, says Hongkui Zeng.



Saul Gravy/ikon Images/Getty

‘What-happens-where’ data can be used to see how cell positioning and gene expression matter.

Among the spatial techniques that can be used to profile brain tissue and cells *in situ* are fluorescence *in situ* hybridization (FISH) and *in situ* RNA-sequencing methods (see **Box 1**, “RNA *in situ*”).

## Circuits and cells

As he teases apart neural circuits involved in rodent behavior, such as reactions to sounds, neuroscientist Anthony Zador at Cold Spring Harbor Laboratory wants multidimensional insight about circuits and cells. Neighboring neurons might fulfill similar functions, which is an aspect that has long been addressed with Cre lines, transgenic mice engineered to express reporter genes. Cre is powerful when the markers inform well on brain function, says Zador. But that is not yet the case for the excitatory neurons in the cortex he studies. He decided to recast his research strategy. The better way to explore neurons, their circuits and the vast heterogeneity of brain

## BOX 1 RNA *IN SITU*: SOME RNA LOCALIZATION METHODS

**BaristaSeq:** An approach using padlock probes for *in situ* barcode-based RNA sequencing.

**ExFISH:** Expansion-microscopy-based technique in which a small-molecule linker attaches to RNA *in situ*.

**FISSEQ:** Fluorescence *in situ* RNA sequencing. Cross-linked cDNA amplicons are sequenced and imaged.

***In situ* sequencing for RNA analysis:** *In situ* RNA sequencing with barcodes, rolling-circle amplification and sequencing by ligation chemistry.

**MAPseq:** A neural tracing method, using barcoded mRNAs.

**MERFISH:** Imaging-based single-cell transcriptomic approach involving *in situ* hybridization and combinatorial labeling.

**SeqFISH:** Multiplexed mRNA analysis using barcodes, transcript fixation and sequential rounds of hybridization and imaging.

**Spatial transcriptomics:** Histological sections are positioned on a barcoded array, and mRNAs are captured and then sequenced.

**TIVA:** Transcriptome *in vivo* analysis. With laser-capture microdissection and a photoactivatable tag, mRNA transcripts are captured from a single cell in live tissue and then sequenced.

**Tomo-seq:** RNA-tomography approach to obtain gene expression patterns in cryosectioned tissue.

tissue is, he thinks, to apply molecular characterization tools such as gene expression analysis and optogenetics techniques and to convert “the problem of neural wiring into a problem of DNA sequencing.” Sequencing is a mature technology and “if you can hop onto a widely used technology, that’s a win,” he says.

One method from his lab in this vein is MAPseq, a sequencing-based neuroanatomic tracing method in which neurons are labeled with RNA barcodes. It’s an alternative to the laborious approach used to capture neuronal message transmission with tracing techniques and dyes. With MAPseq, the readout is barcoded mRNA from the viral library injection site and mRNA from the axon that the barcoded mRNA traveled to.

BaristaSeq, in which DNA and RNA barcodes are used for *in situ* barcode sequencing at cellular resolution, is a recent approach Zador developed with George Church and others at Harvard Medical School. The team believes BaristaSeq addresses some efficiency issues users face with padlock-probe-based targeted *in situ* RNA amplification. Padlock probes are oligos with target DNA sequence on each end. The BaristaSeq tweaks include a way to sequence longer barcodes, and they involve cross-linking to stabilize the colonies, the DNA ‘balls’ generated with rolling-circle amplification.

Zador sees great value for neurobiology in combining scRNA-seq data with spatial information. BaristaSeq, a modified form of FISSEQ, lets researchers capture mRNA and, he says, “nail it down where it is.” In both BaristaSeq and FISSEQ, RNA is reverse-transcribed in cells and sequenced, and the sequence is read out by imaging, says Zador. BaristaSeq shows high sensitivity, says Zador, but it is “still considerably less sensitive than *in situ* hybridization approaches.”

When weighing *in situ* RNA sequencing and *in situ* hybridization (ISH) technologies, there is much to consider. One issue with *in situ* RNA sequencing is that the conversion of RNA to cDNA is “pretty inefficient,” says Zador. Ideally *in situ* sequencing “is the coolest” approach for spatially resolved gene expression data. It delivers RNA data sequence where it originates. “In practice, it is limited first and foremost by sensitivity,” he says.

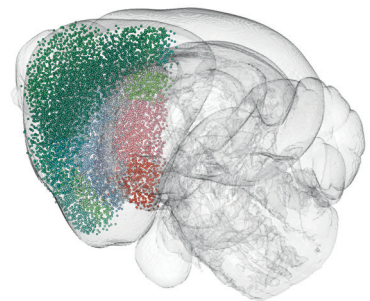
ISH techniques can face issues of off-target hybridization, says Zador. The main ISH challenge is multiplexing, he says. But seqFISH, for example, in which mRNAs are targeted with barcodes and sequential rounds of hybridization, addresses multiplexing better than more conventional single-molecule hybridization approaches. Another technique Zador likes, spatial transcriptomics, was codeveloped by Joakim Lundeberg, who runs the genomics core for SciLifeLab, a collaboration between multiple research insti-

tutions in Sweden. Tissue is transferred onto an array on a glass slide filled with surface-immobilized barcoded probes. RNA is captured, reverse-transcribed and sequenced. “Right now the resolution is not quite cellular,” says Zador, but he hopes the technique will get there.

The current arrays, says Lundeberg, have 100-micrometer resolution, which corresponds to around 5–80 cells per investigated region. The team has recently developed arrays with single-cell resolution and is testing them to ensure robustness. He and colleagues have also founded a company, Spatial Transcriptomics, that offers kits and tissue preparation, RNA sequencing and data visualization as services.

Lundeberg is collaborating with New York Genome Center researchers Hemali Phantani and Richard Bonneau in a project on the neurodegenerative disease amyotrophic lateral sclerosis (ALS). Using two mouse models of the disease, the researchers have applied spatial transcriptomics to monitor and characterize disease progression at four time points. To date they have done transcriptome-wide analysis on 400 tissue sections, says Lundeberg. The early data lend the disease model a spatial dimension: they can identify transcriptional changes before morphological changes emerge, such as neuronal death, an ALS hallmark. “The specific spatial localization of these early transcriptional disease signatures makes us quite confident in the observations,” says Lundeberg.

The method captures all genes equally well using the poly A sequence of the mRNA, says Lundeberg, and that means scientists can quantify and identify coregulated genes within a spatially defined region. Among ALS genes, the team studies “key suspects” and has a hint of genes not previously associated with ALS. They have also started to



Researchers in Sweden are building a mouse brain atlas that superimposes spatial transcriptomic data on neuroanatomy.

K. Meletis, D. Fürth, Karolinska Institutet



SciLifeLab

Spatial data from single-cell research are critical for understanding neurodevelopment, tissue maintenance and disease, says Joakim Lundeberg.

looking at spatially resolved transcriptomic data from the olfactory bulb, among other regions.

Using their spatial transcriptomics approach, Lundeberg and colleagues are building a mouse brain atlas and planning a transcriptome-wide map linked to the canonical structures of the brain. The team has mapped transcriptional events in over 30,000 tissue regions thus far. They present data in a [portal](#) where stained regions are superimposed on neuroanatomy. Machine learning techniques are used to place the ‘spots’ on the brain map and offer a 3D image of the analyzed sections, says Lundeberg.

For labs considering options, it’s Lundeberg’s sense that ISH-based technologies require expertise and more sophisticated, costly equipment than the array-based spatial transcriptomics approach. His method is not trivial but it’s an advantage that single-cell sequencing tools are generally standard in labs. Gene-specific hybridization-based methods are not unlike assays with polyclonal antibodies, he says—even a beautiful signal needs validation. “From this perspective, I think RNA sequencing can offer validation data sets” for ISH-based results, he says. A number of labs are also building ways to computationally connect scRNA-seq data with spatial transcriptomic data as a kind of shortcut to obtain higher resolution data. His group, too, has started applying such techniques.

### Hippocampus dive

Now is the “golden age of transcriptomics,” says Mark Cembrowski, a postdoctoral fellow in the lab of Nelson Spruston at Janelia Research Campus. Tools and resources abound, so it’s a good way to discern func-

tionally important patterns of gene expression in the brain. The lab uses its own data and databases to find patterns in hippocampal RNA-seq and ISH data.

look at the brains of deceased ALS patients to validate these observations. In a separate project, Lundeberg and his team study mouse models of aging and Alzheimer’s disease, which can affect certain brain regions more than others. Alzheimer’s tends to affect odor sensing, for example, which is why he and his team are

tionally important patterns of gene expression in the brain. The lab uses its own data and databases to find patterns in hippocampal RNA-seq and ISH data.

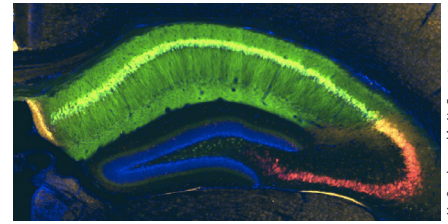
The lab’s focus is on the hippocampus, a part of the mammalian brain involved in many functions including spatial navigation and memory. The team has developed Hipposeq, a [hippocampal atlas](#) based on bulk RNA-seq data of hippocampal principal neurons and cross-validated with data from other approaches such as ISH. Knowing which cells express certain genes in the hippocampus and their location will, says Cembrowski, open the door to many types of experiments that leverage differential expression of these cells and contribute to understanding of the mechanisms of hippocampal processing.

The Spruston lab studies the hippocampus using bulk RNA-seq, ISH, immunohistochemistry and electrophysiology. Lesion experiments and other data indicate that one hippocampal region is more the “cold cognitive processor” and the other handles “hot, emotive processing,” says Cembrowski. Yet the neuronal circuitry in these areas looks the same. Perhaps the cell types in these regions differ or maybe the same cells function differently at different locations. The lab has found transcript diversity in hippocampal cells and determined a gradient of gene expression along the hippocampus’ dorsal-ventral axis.

In contrast to these findings, the lab of Long Cai at the California Institute of Technology and colleagues performed *in situ* transcriptomic profiling of single cells using an amplified version of their single-molecule ISH method, seqFISH, which involves a set of fluorophores and sequential hybridization phases. They detected

not a gradient but rather spatially defined regions in the hippocampus. Cembrowski and Spruston published a paper in response, to which the Cai lab replied<sup>1–3</sup>.

Among other concerns, says Cembrowski, is that many genes used by the Cai lab to characterize the spatial organization were either expressed at



M. Cembrowski, N. Spruston, Janelia Research Campus

The brain’s cells are spatially organized. Shown here, a micrograph of the hippocampus, with its CA1 pyramidal cells in green.

low levels or not found using other methods. When classifying how genes shape cellular function, low-level expression “might not make a big difference,” he says.

Cai believes his approach places gene expression in a spatial context such that brain and other tissues are characterized as a complex mixture of cells, each with a spatial and molecular signature. The team detected a wide dynamic range of gene expression in single cells in hippocampal tissue and regions with distinct cell classes. He is concerned about using bulk RNA-seq for spatial gene expression analysis because it “misses out on all the tissue heterogeneity,” says Cai. Averaging values even over ten cells can erase important signatures. And by multiplexing RNAs in single cells researchers capture correlated expression patterns, which labs can miss when using approaches that assess one gene at a time.

Speaking more generally, Cai suggests that labs consider using scRNA-seq to identify, at high resolution, cell types of interest and then follow up with a multiplexed ISH approach to target these identified neurons.

Despite the different viewpoints, Cembrowski says that emerging approaches in spatial transcriptomics are “incredibly exciting,” and that “something like seqFISH has tremendous potential for neuroscience.” When generating spatial gene expression data, ISH and RNA sequencing complement each other well and each method can address the blind spots in the other. “No method is perfect,” he says, which is why labs benefit from applying different ones. Methods that involve scRNA-seq and its high resolution are “awesome,” he says. But users of these methods, he says, must contend with the as yet unsolved issues of transcript loss when reverse-transcribing RNA to cDNA.

To avoid this loss and to obtain greater depth and statistical power, Cembrowski chose bulk RNA-seq for compiling Hipposeq. Bulk RNA-seq lets labs study



Janelia Research Campus

The hippocampus has a “cold cognitive processor” and another region for “hot, emotive processing,” says Mark Cembrowski.

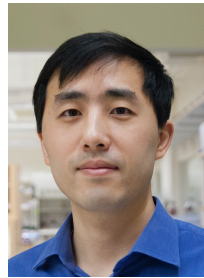
genes that are expressed in graded, non-binary ways and at low levels. ScRNA-seq data offer high resolution but they cut into labs' ability to analyze gene expression in a statistically sound fashion and to group cells into classes. This grouping will enable researchers "to move from molecules, through cells and circuits and ultimately to behavior and function," he says.

The published papers have not settled the issues surrounding hippocampal spatial organization and spatial techniques more generally, says Zeng. "Each party has their respective valid points," she says. And each takes on the scientific questions with different assumptions based on different types of approaches. "I do believe in the presence of a canonical cell taxonomy," she says. "It may be hard to define right now, with individual pieces of the puzzle not put in the right places and not related to each other in the right way."

Eventually, says Zeng, a clear picture will emerge that reveals the underlying

principles that define the relationship of cells in the brain. Gathering data from multiple aspects and correlating them will make it possible to put the pieces of this cellular puzzle together. "Of course, a danger of this approach is that we might also be swamped by all the data and still can't agree with each other, so the debate will continue," she says. In her view,

Labs can use scRNA-seq to identify cell types of interest, then target neurons with a multiplexed *in situ* hybridization approach, says Long Cai.



Lance Hayashida

one possible resolution is a hierarchical approach, which is "identifying the most-agreed-upon core concepts and branching out from there in a multiple-evidence-based manner, leaving the more controversial terminal leaves to further investigation and assessment of their relevance to the functional questions we have in hand."

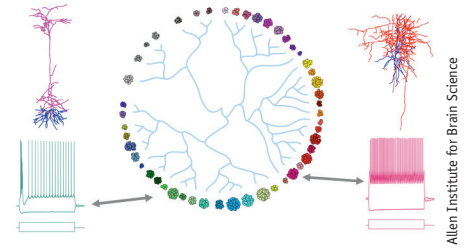
## Cell-typing

Transcriptomic techniques bring a long-standing issue in neurobiology to the foreground: cell-typing. On its own, understanding cell types will likely not suffice for understanding the brain, says Zador, but it certainly will help. Making inferences from a snapshot of a cell's gene expression can, however, be problematic, which is why many scientists use the term 'expression state' when characterizing cells such as neurons.

It is not yet clear if one gene or a small number of genes suffice to determine a neuronal cell class, says Zador. In a comparison of gene expression profiles between groups of neurons, only one of which had been exposed to a stimulus, the cells could differ so much that they might seem to belong to two different classes. "Is that really the right way to think about it?" he asks. In actuality, the cells might not be fundamentally different after all, he says. The concept of cell class in neurobiology is associated with deeper issues. Spatial transcriptomic techniques can help resolve these issues. Their resolution is not necessary for spatial tool-making, he says, but tool utility does depend on the extent to which classes of neurons can be discerned by the expression of one or several genes.

In parsing *in situ* gene expression data, neuroscientists are working on an organ that has yet to be comprehensively profiled in terms of function, connectivity or molecular properties. To explore neuronal function, neurobiologists need "what and where information" to pinpoint cells and to characterize their microenvironment, says Lundberg. The Allen Brain Atlas is an enormous contribution to the field, he says, and the field of single-cell analysis "is providing new opportunities to improve the atlas." ScRNA-seq adds new names to the overall directory of genes and quantitative information, too. Spatial technologies contribute positional information to the dictionary, which converts the dictionary into an atlas. He believes that spatial information emerging from single-cell research will be critical for understanding the context of neurodevelopment, tissue maintenance and disease.

Zeng and Harvard University neuroscientist Joshua Sanes point out that a systematic



Allen Institute for Brain Science

Transcriptional similarity, morphology and physiology criteria can help labs classify neuronal cell types.

categorization of neuronal cell types will not, in itself, lead to enlightenment<sup>4</sup>. "We do however, believe that without it, enlightenment will be unattainable." Cell-type classification will require morphological, physiological, molecular and possibly connective categories, best accomplished by collecting two or more data types from the same cells. Scaling up high-throughput analysis with multiple modalities for cell-type classification in neuroscience is needed and it presents challenges, they say, but the possibilities presented by FISH and *in situ* sequencing applied to tissue are "exciting developments."

Spatial transcriptomics can potentially be an effective way to create comprehensive cell atlases for all sorts of tissues in animals and humans and a variety of functional or dysfunctional states, says Zeng. These resources will enable detailed, anatomically precise comparison across a variety of conditions at the single-cell level, and let scientists correlate cell-type identity, cell states and cell function. What the community stands to gain, she says, is an "unprecedented understanding of the organization and function of the brain by integrating information at gene, cell and network levels."

Vivien Marx is technology editor for *Nature Methods* ([v.marx@us.nature.com](mailto:v.marx@us.nature.com)).

1. Shah, S., Lubeck, E., Zhou, W. & Cai, L. *Neuron* **92**, 342–357 (2016).
2. Cembrowski, M. & Spruston, N. *Neuron* **94**, 747–751 (2017).
3. Shah, S., Lubeck, E., Zhou, W. & Cai, L. *Neuron* **94**, 752–758 (2017).
4. Zeng, H. & Sanes, J.R. *Nat. Rev. Neurosci.* **18**, 530–546 (2017).