COMMENTARY

PERSPECTIVES

HYPOTHESIS

Unbiased discovery of neuronal architectures

Comparative whole-brain and single-cell analyses identify neurons orchestrating neurological functions

Jordan W. Squair[1,2,3], Michael A. Skinnider[4,5], Gregoire Courtine[1,2,3]

Neuronal architectures comprise synaptically connected neurons distributed throughout the central nervous system, the coordinated activities of which orchestrate neurological functions ranging from breathing to movement and cognition. Disentangling these neuronal architectures and how they are disrupted in disease is a fundamental goal of neuroscience. Historically, this challenge has been addressed with a reductionist framework that translated hypotheses into the interrogation of discrete neuronal subpopulations based on a priori expectations. The advent of high-throughput methodologies, including whole–central nervous system imaging in rodent models and single-cell transcriptomic readouts, now enable the visualization and characterization of neuronal subpopulations throughout the central nervous system. Increases in scale further enable comparative experimental designs that can be navigated with computational frameworks. These advances augur a new era wherein neuronal architectures implicated in diverse neurological functions, yet obscured by the complexity of the central nervous system, can be exposed without bias and interrogated with genetically guided experimental manipulations.

Over the past two decades, methodological advances have opened up the possibility to manipulate and record from specific neuronal subpopulations in mammals (*1*). Methodologies such as optogenetics, chemogenetics, calcium imaging, and recombinase systems have cemented a reductionist framework into the core identity of neuroscience. However, a consequence of this reductionism is that enigmatic yet potentially essential neuronal architectures have remained concealed by the vast complexity of the central nervous system. Whereas the study of complex neuronal architectures has historically been more tractable in invertebrates, such as flies (*Drosophila*), or lower vertebrates, such as zebrafish (*Danio rerio*), technological advances over the past 5 years have augmented the scale and resolution with which enigmatic neuronal architectures can be discovered and dissected in mammals such as mice (*Mus musculus*). Among the most exciting of these advances have been the emergence of whole–nervous system imaging and large-scale single-cell biology.

Neurological functions emerge from the coordinated activity of neuronal architectures that are distributed throughout the nervous system. Consequently, identifying the anatomical locations of these architectures is the first step involved in understanding the neuronal underpinning of a neurological function and necessitates methodologies that enable whole–central nervous system visualization of the neurons involved in the neurological function under study. Attempts to chart these architectures in mammals originally involved painstakingly low-throughput endeavors based on tissue sectioning; labeling of markers of neuronal activity, such as the protein cFos; and section-by-section imaging to reconstruct an imperfect image of neurons and their axonal projections. The response to this limitation was the development of tissue-clearing methodologies that rendered spinal cord and brain tissues of rodents optically transparent while maintaining the three-dimensional cellular integrity of the entire central nervous system (*2, 3*). These methodologies triggered the optimization of light-sheet microscopy to enable the visualization and reconstruction of the intricate three-dimensional neuronal architectures nested within these cleared tissue samples in exquisite detail.

...voxel-level statistics can nominate the most perturbed regions of the brain and spinal cord...

Parallel advances in computational analyses of these large-scale datasets now permit true brainwide perturbation biology. Libraries of brains labeled for the expression of activity-dependent genes during standardized behavioral tasks or tagged with fluorescent proteins to expose neuronal connectivity can be automatically segmented and then registered to common coordinate frameworks, such as the Allen Brain Atlas (*4*). Within these unified coordinate systems, voxel-level statistics can nominate the most perturbed regions of the brain and spinal cord and thus implicate them in the regulation of the studied neurological function (*4, 5*). This framework has yielded unexpected discoveries. For example, a whole-brain functional atlas in mice recently identified the xiphoid nucleus of the thalamus as a critical regulator of food-seeking behavior in response to cold exposure, although this structure had not previously been associated with feeding (*6*). Another example came from a whole-brain atlas of neurons active during the recovery of walking after incomplete spinal cord injury in mice. Notably, this atlas nominated the lateral hypothalamus, which has typically been invoked in motivated and food-seeking behaviors but not walking, as a key region of the brain to explain this recovery (*5*). This discovery guided the development of deep-brain stimulation therapies delivered in the lateral hypothalamus that enhanced the recovery of walking in people with incomplete spinal cord injury.

Although whole-brain atlases provide an exquisite lens to establish the anatomical location of neuronal subpopulations implicated in specific neurological functions, these technologies presently remain unable to resolve the molecular identity of these subpopulations, as multiplexed labeling of mRNA transcripts can only reliably

interrogate a limited number of marker genes. This limitation in the molecular resolution of these anatomical surveys contrasts with the increasing realization that the neuronal subpopulations nested within any given region of the nervous system exhibit remarkable transcriptional diversity (*7*).

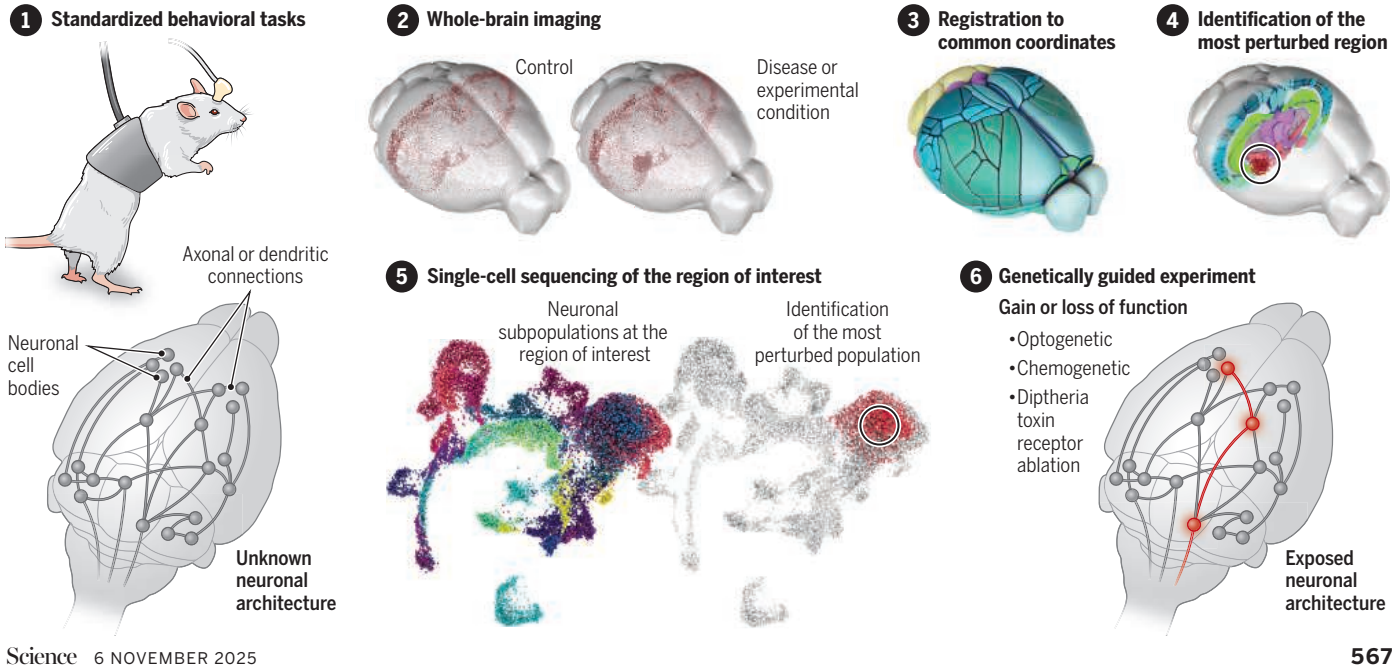
Single-cell and single-nucleus transcriptomics now enable measurement of the complete set of transcriptional programs expressed by individual neurons. Spatial transcriptomics further allows for these transcriptional programs to be localized within the cytoarchitecture of neural tissues. Thus, these technologies provide the molecular resolution that is necessary to interrogate the diversity of the neuronal subpopulations located in the regions of the central nervous system that are prioritized by whole-brain anatomical and functional atlases. Early deployment of these approaches in the nervous system focused on establishing taxonomies of neuronal subpopulations in selected regions of the nervous system or on aligning transcriptomic signatures with predefined morphological or electrophysiological classifications (*8, 9*). With increases in the scale of single-cell and spatial transcriptomic technologies, it became possible to expand the measurement of transcriptomic signatures from individual neuronal subpopulations across multiple experimental conditions or disease models (*10*). Because the full compendium of neuronal subpopulations is measured simultaneously across all the experimental conditions, the need for a priori selection of a specific neuronal subpopulation is not necessary.

Drawing biological inferences from the data generated by comparative single-cell transcriptomics experiments necessitates analytical methods that can address the biological questions motivating these experiments. Initial efforts to identify neuronal subpopulations underlying the studied neurological functions utilized analytical paradigms that relied on the expression of immediate early genes as a proxy for neuronal activity or quantified the total number of differentially expressed genes between experimental conditions in each neuronal subpopulation as a proxy for the magnitude of the transcriptional response within those neurons. However, it rapidly became apparent that such quantifications are poised to generate both false positives and false negatives (*11*). This statistical challenge spurred the development of more tailored computational methods. Several such methods aim to identify changes in the relative proportions of neuronal subpopulations or other cell types (*12*). Another framework to identify neuronal subpopulations involved in a disease or biological perturbation is based on the assumption that neuronal subpopulations that respond to a perturbation should be more separable, within the multidimensional space of single-cell measurements, than less-responsive subpopulations. In turn, the degree of this separability could be quantified by the accuracy with which a given neuronal subpopulation can be classified based on its transcriptome-wide gene expression. These concepts were distilled into a machine-learning framework referred to as cell-type prioritization (*9, 13*). Applied in mouse models, this approach exposed the neuronal subpopulation that was necessary and sufficient for the recovery of walking after paralysis (*9, 14*) and identified the perturbation of neurons from the hypothalamic and periventricular neuronal niches in the aging brain (*15*). Although powerful, the validity of this approach is contingent on appropriate experimental design, as certain types of batch effects and small sample sizes can confound valid inferences.

With the appropriate methodologies to study brainwide and single-cell perturbation biology in rodents now available, the question arises on how to discover previously unrecognized neuronal architectures linked to neurological functions. Applying these methodologies is first contingent on the establishment of rodent models and standardized behavioral tasks that recapitulate the disease or neurological function being studied. Although the expansion of animal models for neurological diseases or behaviors has greatly advanced with increasingly complex transgenic models, there remain neurological diseases and behaviors for which identification

Framework for the identification of neuronal architectures

A proposed framework for the unbiased identification of neuronal architectures involved in behaviors or diseases begins with a standardized behavioral task in an animal model (**1**). The location of neurons active during this task is established through visualization of an activity marker across the whole brain, using tissue clearing and light-sheet microscopy (**2**). These data are registered to a common set of coordinates to enable the location of the active neurons to be identified (**3** and **4**). Subsequent experiments use single-cell sequencing to characterize the neuronal subpopulations in this region of interest and identify the population most perturbed by the behavioral task (**5**). Genetic modifications targeting that population can then be used to esablish a causal role in the behavior and to learn more about the connections and computational characteristics of the identified neurons (**6**).



of a relevant animal model is challenging or controversial. After the establishment of an appropriate animal model and one or more relevant behavioral tasks, the next step involves the visualization of neurons labeled for activity-dependent markers, such as cFos, over the entire brain or spinal cord across multiple biological replicates and experimental conditions. This library of brains and spinal cords can be registered to a common coordinate system, such as the Allen Brain Atlas. Within these unified coordinate systems, these datasets can be interrogated with statistical and machine-learning methods to prioritize the anatomical regions involved in the production of the specific neurological function or behavior (*4*, *5*).

The next step consists of subsequent experiments that profile the prioritized regions with single-cell and spatial transcriptomics across experimental conditions to uncover the identity of the neuronal subpopulations implicated in the biological responses within the nominated regions of the brain or spinal cord, leveraging appropriate statistical or machine-learning methodologies. The prioritized neuronal subpopulations then can be interrogated with genetically guided silencing and activating manipulations, connectome visualization (including neuronal subpopulation–specific electron microscopy imaging), and neuronal subpopulation–specific calcium transient or single-unit recordings to uncover the causal role and computational logic underlying the operation of each subpopulation embedded in the architecture. This methodological framework is poised to uncover the neuronal architectures underlying complex neurological functions in rodent models (see the figure). Some of these steps are technically challenging, experimentally costly, or both. Yet ongoing innovations of clearing and imaging technologies as well as decreases in the costs associated with single-cell experiments increasingly open the door to translating this strategy from rodents to larger mammals.

The technologies necessary to accelerate the discovery of neuronal architectures associated with complex neurological diseases and functions are progressing at a fast pace. The next frontier likely lies in the eventual marriage of whole-brain imaging with single-cell and spatial transcriptomics into a single methodology that will measure the entire transcriptomes of every neuronal subtype, both at steady state and in response to perturbation, to discover neuronal architectures across the entire intact central nervous system. □

REFERENCES AND NOTES

1. C. K. Kim, A. Adhikari, K. Deisseroth, *Nat. Rev. Neurosci.* **18**, 222 (2017).
2. K. Chung *et al.*, *Nature* **497**, 332 (2013).
3. N. Renier *et al.*, *Cell* **165**, 1789 (2016).
4. A. Attarpour *et al.*, *Nat. Methods* **22**, 600 (2025).
5. N. Cho *et al.*, *Nat. Med.* **30**, 3676 (2024).
6. N. K. Lal *et al.*, *Nature* **621**, 138 (2023).
7. M. Zhang *et al.*, *Nature* **624**, 343 (2023).
8. A. Zeisel *et al.*, *Cell* **174**, 999 (2018).
9. F. Scala *et al.*, *Nature* **598**, 144 (2021).
10. C. Kathe *et al.*, *Nature* **611**, 540 (2022).
11. J. W. Squair *et al.*, *Nat. Commun.* **12**, 5692 (2021).
12. E. Dann, N. C. Henderson, S. A. Teichmann, M. D. Morgan, J. C. Marioni, *Nat. Biotechnol.* **40**, 245 (2022).
13. M. A. Skinnider *et al.*, *Nat. Biotechnol.* **39**, 30 (2021).
14. M. A. Skinnider *et al.*, *Nature* **631**, 150 (2024).
15. K. Jin *et al.*, *Nature* **638**, 182 (2025).

ACKNOWLEDGMENTS

The authors acknowledge the Defitech Foundation for support. M.A.S. is a member of the Rutgers Cancer Institute of New Jersey.

10.1126/science.adx6741

[1]NeuroRestore, CHUV/UNIL/EPFL, Lausanne, Switzerland. [2]Department of Neurosurgery, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland. [3] Neuro-X Institute, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. [4] Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [5] Ludwig Institute for Cancer Research, Princeton University, Princeton, NJ, USA. Email: gregoire.courtine@epfl.ch

MOLECULAR BIOLOGY

Cell signaling meets gene transcription

Receptor tyrosine kinases directly regulate RNA polymerase II in the nucleus

Richard Young

Cells rely on the selective transcription of genes to maintain their identity and adapt to changing conditions (*1*, *2*). Proteins called transcription factors regulate selective gene transcription in the nucleus by binding to DNA regulatory elements and recruiting RNA polymerase II (Pol II), which then transcribes the genes. Cells can respond to changes in their extracellular environment by using plasma membrane receptors that bind various ligands, which then send signals to the nucleus. For decades, these receptors were thought to use signaling intermediates that act on transcription factors, which then bind gene regulatory elements and regulate Pol II activity. On page 594 of this issue, Dabas *et al.* (*3*) report a change to this canonical view, revealing that signaling molecules of the receptor tyrosine kinase (RTK) family can also act directly on a key regulatory domain of Pol II within the nucleus.

...conserved kinase activity is converted into highly selective transcriptional outcomes.

RTKs play important roles in cell biology because they recognize extracellular signals for growth, metabolism, differentiation, and survival (*4*, *5*). RTKs autophosphorylate upon ligand binding, which creates docking sites for adaptor proteins and initiates signaling cascades that target transcription factors in the nucleus. In the canonical model of gene regulation, this process alters the activity or location of the transcription factors, which regulate specific genes (see the figure).

Diverse RTKs have been detected in the nuclei of human cells under physiological and disease conditions (*6*–*10*). For some RTKs, an intracellular domain segment of the protein is generated at the cell surface and translocates into the nucleus; for others, the intact receptor manages to find its way into the nucleus (*6*). Nuclear RTKs have been reported to associate with gene regulatory regions and with Pol II itself (*8*, *10*), but the mechanism by which they regulate transcription was unclear. Dabas *et al.* explored the possibility that RTKs directly phosphorylate tyrosine in the key regulatory domain of Pol II—the C-terminal domain (CTD) of the large subunit of this enzyme.

The human Pol II CTD consists of 52 repeats of a sequence of seven amino acids, or a heptad, whose phosphorylation states regulate various stages of transcription and RNA processing (*11*, *12*). During the first step of transcription, when Pol II is recruited to specific DNA regions