

Parcellation of the mouse brain using molecular imaging

Alex Lee

Specific Aims

Parallel technological advances in molecular biology and microscopy have dramatically increased the ease with which experimental biologists can make spatially resolved measurements of large numbers of proteomic or transcriptomic features. These high-fidelity molecular mapping techniques are promising to catalog the enormous diversity of cell types in complex organs like the brain, whose cellular heterogeneity has been recognized since the early work of Ramon y Cajal¹. These new capabilities are the focus of large scale collaborations facilitated by the NIH BRAIN initiative², which aim to precisely elucidate the structural organization of the brain by generating multimodal, cross-species datasets for comparison. The datasets that will result from these projects will be unprecedented in size and complexity. They will require new and innovative computational techniques to organize and analyze, but will transform our understanding of the molecular and cellular organizing principles of the brain.

Two principal focus areas in the neuroscience community's investigation of these relationships are variation in gene expression³ and connectivity^{4,5}, and how they interact in the circuitry of the brain. For example, single-cell RNA sequencing (scRNA-seq) has been enormously impactful for our ability to catalog neuronal variation³, but to date it is unclear how subclasses observed in scRNA-seq correspond to graded connectivity differences, such as those found in the cornu ammonis (CA) of the hippocampus⁶. Reconciling these two perspectives requires an understanding of the cell types that make up a given region of the brain and how they interact in neural circuits.

One approach to understanding the composition of these circuits is to identify subregions detected in large scale datasets such as spatial transcriptomics. In this case a functional subregion can be defined as a distinct pattern of gene expression variation that can be used for categorization of a subgroup of cells. In the larger field of computational biology, a variety of unsupervised and supervised machine learning techniques have been developed for extraction of latent features for spatial clustering or subregion detection in spatial imaging data⁷, but these methods typically target 2D tissue sections, and incorporate assumptions that are not ideal for analysis of 3D volumetric data or neural cell types, such as interactions only on a small spatial scale⁸⁻¹¹. In addition, there are few techniques suited for multimodal data integration of spatial imaging data, with most methods focusing on spatial transcriptomics exclusively.

I propose to address this unmet need by developing and applying new computational methods for machine learning-based spatial subregion detection and multimodal data integration to large scale datasets from the Allen Institute for Brain Science^{4,12,13} and the Broad Institute¹⁴. I will use these techniques to map the structure and organization of these subregions in the mouse brain. I will also establish an interpretable statistical framework to map transcriptomic correlates and drivers of differential connectivity, leveraging the single-cell resolution of emerging spatial transcriptomic datasets. I hypothesize that continuous or graded gradients of gene expression that have been observed in well-characterized regions such as the hippocampus¹⁵ are prevalent throughout the brain and can be more readily identified in spatial imaging data than in non-spatial datasets.

Aim 1: Establish a spatial transcriptomics data analysis pipeline to identify and characterize novel subregions in the mouse brain

First, I will develop a novel interpretable machine learning pipeline for unsupervised and interpretable discovery of functional subregions in 3D spatial transcriptomics data^{12,14}. I will validate this approach by testing whether this method can recover subfields of the hippocampus, specifically dorsal cornu ammonis regions CA1, CA2, and CA3, and the dentate gyrus. These subfields thus far have not been identified in a data-driven way in spatial imaging data. Next, I will characterize whether these subregions are driven by specific cell-types or by continuous variation within cell types using regression analysis. Determination of the principal gene expression spatial patterns in these datasets will pave the way towards understanding the organizing principles of cell type composition in the brain.

Aim 2: Determine transcriptomic and non-neuronal cell-type determinants of differential connectivity across the brain by integrating the Allen Mouse Connectivity Atlas with spatial transcriptomics datasets

First, I will use a predictive regression model for association of connectivity patterns in the Allen Mouse Connectivity Atlas⁴ with single-cell transcriptomes measured in Yao et al.¹² and Langlieb et al.¹⁴. I will use this model to identify whether there are significant genes or gene modules that are predictive of differential connectivity among neurons. In addition I will identify whether this method can recapitulate an existing finding¹⁶ that distinct transcriptomic classes of layer-5 pyramidal tract (PT) neurons are almost exclusively connected to one of either the thalamus or medulla. Systematic identification of the gene expression correlates of connectivity will give insight into the laws by which the brain constructs and constrains neural circuitry.

1 Background

From the first neuroanatomical studies¹, it has been recognized that although cells can be broadly grouped into high-level categories (like neurons, or glia), even visual inspection of silver-stained cells

is sufficient to identify a diverse array of cellular subtypes. Originally, these cells were observed to primarily diverge based on their morphology and connection to other cells, however the development of new molecular techniques has revealed new axes by which these cells vary. Understanding the role of this diversity in facilitating the complex computations the brain performs remains one of the fundamental projects of modern neuroscience research. However, the scope and complexity of this project has increased in the modern 'omics age due to the accessibility of high dimensional molecular measurement tools. In particular, the ubiquity of high-fidelity molecular imaging data provides a unique opportunity to study the correspondence between our current and previous understanding of the organizing principles of brain structure and the patterns that can be identified in new datasets using statistical machine learning techniques.

Large scale molecular atlases such as the Allen Brain Atlas¹³ (ABA) and Allen Mouse Connectivity Atlas⁴ (ACA) have been critical tools to study the molecular organization of the brain. The ABA data uses one-at-a-time ISH to spatially map the expression of several thousand genes in the mouse brain. The ACA is a large scale imaging dataset where both cell-type specific and pan-neuron AAV vectors are used as tracers to map the connectivity of neurons to a source site, focusing on the right hemisphere. The integration of automated high-throughput data collection methods, comprehensive informatics, and fully open data sharing have facilitated an enormous number of studies (as measured by the 3,474 citations of the original Allen Brain Atlas Paper [as of 2023-02-18]). These datasets were integral to the development of the Allen Common Coordinate Framework (CCF)¹⁷, a systematic effort to integrate gene expression, connectivity, transgenic expression, and histology to generate a neuroanatomical consensus structural labeling. This effort brought together a large team of neuroanatomists to develop a consensus spatial parcellation of the brain into known regions based on manual inspection of different data sources. The CCF is a highly important resource for neuroscientists, helping at the level of both hypothesis generation (for example, to identify regions of interest with respect to one or several molecular characteristics) or for hypothesis confirmation (to visualize distribution of certain features identified in an experiment). However, its creation was heavily manual, requiring a consensus of several expert neuroanatomists. Creating a new CCF from emerging datasets using novel technologies would require significant effort: therefore, development of computational techniques to organize and integrate new datasets in the mouse brain would be an important resource for neuroscientists.

New datasets utilizing new modalities are rapidly emerging in systems neuroscience, such as from Yao et al.¹², using MERFISH¹⁸, and Langlieb et al.¹⁴, using Slide-SeqV2¹⁹, which offer spatial single-cell resolution. The opportunity to leverage single-cell level spatial measurements is transformative for the field, as it allows for the characterization of the organized spatial grouping of specific cell types into the components of neural circuits. The ability to analyze cells in their native spatial context is unprecedented, as previous spatially oriented investigations required techniques such as dissection of a particular region for bulk sequencing, dissociation for single-cell sequencing or cell-sorting using a particular cell marker. These techniques have relatively low spatial resolution and throughput. Alternative techniques with high spatial resolution such as in-situ hybridization^{13,20} precluded single-cell level measurement of many genes at a time. The development of techniques for high spatial fidelity measurements across the transcriptome¹⁴ or with subcellular resolution¹² is particularly timely to resolve high degrees of observed heterogeneity within and across cell types discovered by non-spatial scRNA-seq³. For example, graded gene expression gradients within cell types have been observed in investigations of a small number of regions using techniques to target specific

cell populations, particularly in pyramidal cells and in both excitatory and inhibitory neurons in the cortex¹⁵. Development of computational tools to incorporate this cellular heterogeneity and interpretably identify the major subtypes of cells in their spatial contexts will be crucial to understanding neural function²¹.

Multimodal data integration is another key challenge for efforts to understand neural structural organization principles. A comprehensive classification of neuronal types is thought to require characterization involving both molecular (e.g. transcriptomic) and non-molecular modes such as connectivity^{15,21}. Systematic comparison of the cell type makeup and connection configuration of neural circuits would be a crucial step towards resolving function from structure in the brain. At the microscale, pioneering work such as from Chen et al.²² has found that brain regions that have similar cell-type organization are likely to be connected. In a small number (~100) of single cells, an investigation by Economo et al.¹⁶ has displayed a high correspondence between single-cell transcriptomic clusters and projection patterning in layer 5 pyramidal tract neurons. However, particularly with the release of one-hemisphere or whole-brain mouse spatial transcriptomic data, researchers can now systematically investigate the covariation of gene expression patterns and connectivity patterns. Several methods for multimodal data integration have been developed and applied for non-spatial multiomics datasets, such as MOFA²³, LIGER²⁴, although a ridge-regression workflow was shown to be effective²⁵ in modeling gene modules predictive of connectivity. These methods have yet to be tested for single-cell resolution datasets such as the above two, but could yield important insights as to the network organization of the brain.

Application of existing computational methods to integrate and identify major features in spatial transcriptomics datasets is limited by their large scale and sparse spatial structure. An attractive class of methods for spatial pattern discovery and multimodal data integration, for example for spatial transcriptomics data, utilizes Gaussian processes^{8,11}. These models are attractive because of their ability to incorporate prior knowledge on the form of the covariance between data observations in space. However, for 3D data with uneven spatial sampling it becomes harder to specify this relationship. For example, it is not obvious that the spatial correlation along the anterior-posterior axis should match, for example, the dorsal-ventral axis. In addition, Gaussian processes scale with $O(n^3)$ and existing methods cannot easily scale beyond several tens of thousands of cells⁸, making these methods unsuitable for region discovery in 3D whole-brain mouse transcriptomic datasets. An emerging perspective in modeling cells in space uses graph-neural networks⁹. In these models, a spatial cell neighborhood is created and networks can be optimized for supervised cell type prediction or using self-supervised learning (learning to predict characteristics of a hidden cell from its neighbors). However, these networks as implemented currently require construction of a spatial neighborhood of cells via thresholded distance. The construction of this cellular neighborhood graph in the brain is likely to incorporate many ad-hoc decisions; for example, it may not be desirable to have neurons that are spatially close be connected in a cell neighborhood graph, and it may be difficult to assign edge connections between neurons and various types of glia. A promising class of alternative methods are derived from non-negative matrix factorization^{20,26,27} (NMF), which has previously been employed for spatial pattern identification in imaging data but not been adapted to the highly sparse, large scale setting of 3D spatial transcriptomics. A significant benefit of approaches derived from NMF is that it is possible to easily visually interpret the factors.

In order to comprehensively catalog the transcriptomic and connectomic diversity of the mouse

brain, the proposed research will create new computational methods that addresses these limitations. A comprehensive parcellation of transcriptome-defined subregions in the brain will be defined by developing a novel interpretable machine learning algorithm for latent variable discovery and then segmentation of the brain. By then leveraging the single-cell resolution of emerging spatial transcriptomics datasets, I will analyze the specific gene gradients that contribute to these parcels. I will also utilize the single-cell measurements provided in these new datasets to conduct an association analysis of connectivity patterns and transcriptomic variation across the brain, to characterize whether gene expression gradients are attributable to differential connectivity patterns, as has been observed in some regions such as the hippocampus²⁸.

2 Aim 1: Establish a method for the parcellation and analysis of transcriptomic imaging data

My first aim will be to develop and test methods for unsupervised, interpretable discovery of functional subregions in the Langlieb et al.¹⁴ and Yao et al.¹² whole-brain mouse spatial transcriptomics datasets. The size of these datasets (in Yao et al., at $10\mu\text{m}$ resolution, each gene is imaged in a large image of ~ 63 million voxels) requires development of robust software for subregion detection and characterization. Graded gene expression amongst cell types has been characterized in several regions in the brain, particularly in the hippocampus^{15,28}, and provides an attractive target to validate the effectiveness of this approach. Completion of this aim will produce a clear enumeration of the principal gene expression gradients and their spatial location in the mouse brain.

2.1 Aim 1.1: Development and validation of an autoencoder approach for subregion discovery

I will build off of previous work performed by our lab²⁹ and others²⁶ that poses region discovery as a latent variable discovery problem. In both works, an NMF approach is used to identify spatial factors that are optimal for reconstruction. I am a co-author on the former work (Cahill et al.) where we applied this approach (which we call ontology-stability NMF, or osNMF) to the Allen Brain Atlas dataset, however it has yet to be tested and refined on larger datasets such as the above spatial transcriptomics datasets.

An important modeling decision in factor analysis is the choice of inner dimension, or number of factors k . In order to guide this decision, we adopt the stability-driven framework of Wu et al.²⁶, where our algorithm is repeatedly fit on bootstrapped partitions of the data at different choices of k . The stability of the iterations is quantified using an Amari-like measure, which computes a distance between two factorization solutions through the sum of the absolute values of the row and column maxima of their cross-correlation matrix.

In a proof-of-concept analysis on the Yao et al. dataset I applied the osNMF technique. I conducted a sparse stability analysis, screening values of k between 5 and 25. This initial analysis identified an optimal value of 18 factors across the whole brain. Because of the large size of the

dataset, I downsampled the dataset from a coronal-slice resolution of $10\ \mu\text{m}$ to $40\ \mu\text{m}$. After binarizing the patterns to 18 for visualization, it was not clear by visual inspection that osNMF was able to resolve hippocampal subfields Figure 1a.

Because of the inability of osNMF to resolve the desired hippocampal architecture, I introduced two changes. First, I implemented an autoencoder approach using 2-3 layers parameterized by 3D-convolutions instead of linear projection into low dimension, as in NMF. I include a ReLU activation in the last layer prior to a linear projection into native image spatial dimension for reconstruction. By constraining this last layer of neurons’ weights to be strictly positive, and since the last layer activations are also constrained to be positive, we retain the interpretation of a “parts-based” positive representation that ordinary NMF is characterized. When a network with this architecture was trained on downsampled image data (for comparison with osNMF), again with $k = 18$, hippocampal patterns were better resolved@fig-cnn. While the osNMF patterns do not resolve CA1, and CA3, they are delineated in the

In preliminary tests direct application of osNMF towards the Yao et al. dataset does not allow for resolution of subregions in the hippocampus. Both osNMF and

In addition, because of the voxelwise vectorization of the images prior to the NMF fitting, the latent variable discovery step in osNMF does not directly incorporate spatial correlations between the data. Therefore, I will develop a shallow autoencoder based approach to allow for more flexibility in the latent variable discovery process. In order to maintain the interpretability of NMF, I will constrain the last layer of the autoencoder and the previous layer activations to be strictly positive, allowing direct readout of the neuron weights into spatial patterns. I will use this last layer for our stability characterization, as these are the most directly relevant to the selection of functional subregions. Additionally, in order to incorporate spatial correlations into the latent variable computation, I will use a sparse convolutional encoder. This approach reduces the number of parameters in the neural network compared with an equivalent multilayer perceptron-based autoencoder by roughly 7-fold. I also introduce regularization parameters into the loss function to control the spatial correlation of the spatial patterns, empirically finding this increases accuracy of the reconstructions. A preliminary analysis of this architecture’s performance shows (Figure 2) that this new architecture, which we also equip with additional regularization terms, produces factors that allow a coarse regionization of the hippocampus into

This work identifies latent variables that are optimal for reconstruction of imaging data, with the former work (where I am a co-author) applied successfully to the Allen Brain Atlas dataset. By comparing our latent factors with those found in the CCF, we identify Pearson correlation coefficients of up to 0.73 with regions across a hierarchy of resolutions. We also find that our approach is able to resolve gene gradients that sub-parcellate regions such as the isocortex, displaying the utility of our approach in identifying novel gene expression gradients in 3D spatial data.

In order to augment this work to adapt to the very large scale of the whole-brain spatial transcriptomics data, I will introduce a shallow autoencoder approach instead of using an NMF-like matrix factorization. I will incorporate a strict positivity constraint on the last linear layer of the autoencoder to match the framework of NMF, so that I can read out the neuron weights in the last layer to visualize the discovered spatial patterns.

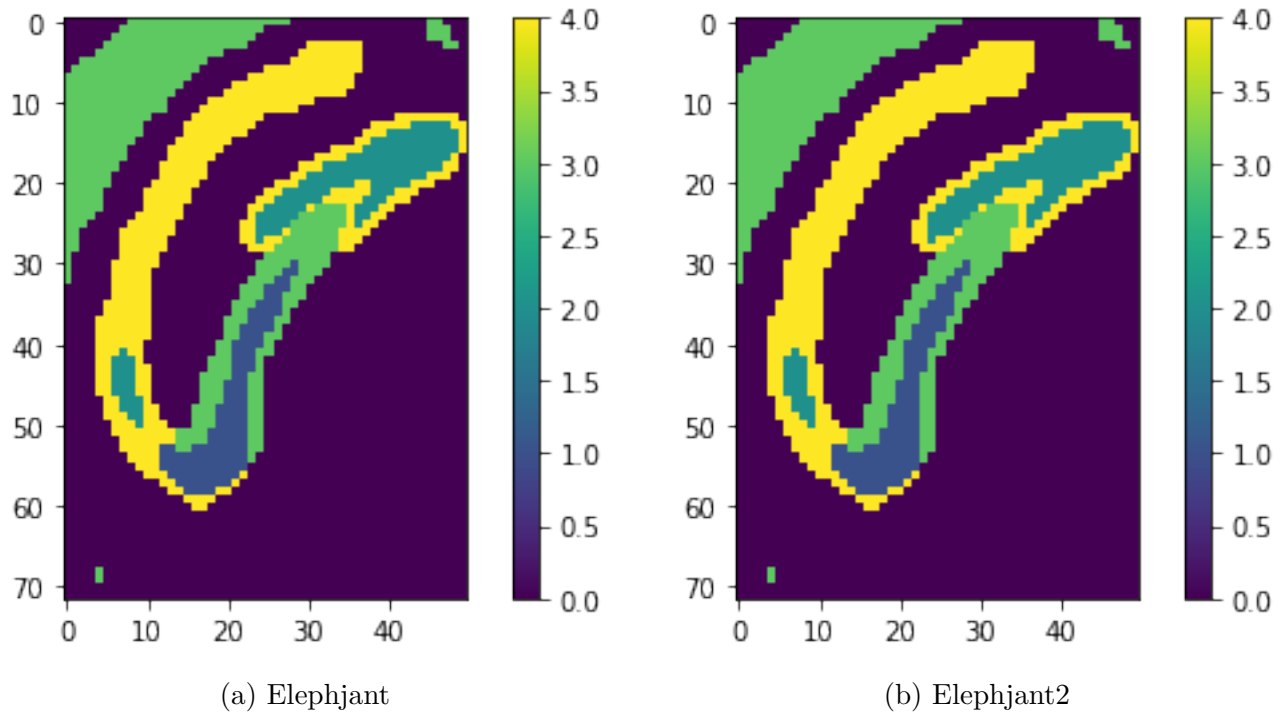


Figure 1: something something something

3 Aim 2: establish an analysis pipeline for the integration of transcriptomic and connectomic imaging data

- Existing techniques for multimodal data integration / subregion discovery focus really on 2D, not suited for brain imaging data
- hypothesis: a small number of gene expression gradients form the hierarchical structure of the mouse brain and can be robustly identified from ST data; these expression gradients also template connectivity patterns across the brain

aim1: create new comp method for data driven atlas creation to identify key gradients in gene expression across brain * prelim work on non spatial transcriptomic test data

aim 1.1 computational method for v large, sparse ST data aim 1.2 analysis of cell type contribution to these gradients

aim 2: identify correlates of differential connectivity in gene expression

aim 2.1 comp method for multimodal region discovery using connectivity and gene expr data aim 2.2 regression analysis of transcriptomic/cell type contributors to differential connectivity

but existing techniques for spatial subregion detection are poorly suited for analysis of large-scale, 3D imaging data. For example, one common class of methods utilizes spatial Gaussian processes⁸ with $O(n^3)$ scaling, which is undesirable for large spatial (3D) imaging datasets. Other classes of

promising methods would require assumptions to be made about the size of the spatial neighborhood for cellular interaction, which is difficult to assess with serial sections⁹.

Single-cell sequencing studies^{3,30,31} in the mouse have made significant progress towards charting cell type gradients as defined by transcriptomic content; likewise, computational efforts have been developed for very fine-grained connectivity mapping and clustering³². Molecular atlases such as the Allen Institute’s Common Coordinate Framework¹⁷ (CCF) atlas are critical to help organize inquiries into these phenomena by facilitating comparisons across studies and spatial scales. One strength of these atlases is their extensive manual curation, starting from the level of image registration and quality-checking towards consensus decisions by expert anatomists to delineate subregions and their boundaries.

1. Cajal, S. R. y. *Texture of the nervous system of man and the vertebrates*. (Springer, 1999). doi:10.1007/978-3-7091-6435-8.
2. NIH BRAIN initiative launches projects to develop cell atlases and molecular tools for cell access. National institute of mental health (NIMH). <https://www.nimh.nih.gov/news/science-news/2022/nih-brain-initiative-launches-projects-to-develop-cell-atlases-and-molecular-tools-for-cell-access> (2022).
3. Tasic, B. *et al.* [Shared and distinct transcriptomic cell types across neocortical areas](#). *Nature* **563**, 72–78 (2018).
4. Oh, S. W. *et al.* [A mesoscale connectome of the mouse brain](#). *Nature* **508**, 207–214 (2014).
5. Sun, Y.-C. *et al.* [Integrating barcoded neuroanatomy with spatial transcriptional profiling enables identification of gene correlates of projections](#). *Nat Neurosci* **24**, 873–885 (2021).
6. Cembrowski, M. S. *et al.* [Dissociable structural and functional hippocampal outputs via distinct subiculum cell classes](#). *Cell* **173**, 1280–1292.e18 (2018).
7. Walker, B. L., Cang, Z., Ren, H., Bourgain-Chang, E. & Nie, Q. [Deciphering tissue structure and function using spatial transcriptomics](#). *Commun Biol* **5**, 1–10 (2022).
8. Townes, F. W. & Engelhardt, B. E. [Nonnegative spatial factorization applied to spatial genomics](#). *Nat Methods* **20**, 229–238 (2023).
9. Brbić, M. *et al.* [Annotation of spatially resolved single-cell data with STELLAR](#). *Nat Methods* **19**, 1411–1418 (2022).
10. Zhao, E. *et al.* [Spatial transcriptomics at subspot resolution with BayesSpace](#). *Nat Biotechnol* **39**, 1375–1384 (2021).
11. Velten, B. *et al.* [Identifying temporal and spatial patterns of variation from multimodal data using MEFISTO](#). *Nat Methods* **19**, 179–186 (2022).
12. Yao, Z. *et al.* A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. (2023) doi:10.1101/2023.03.06.531121.
13. Lein, E. S. *et al.* [Genome-wide atlas of gene expression in the adult mouse brain](#). *Nature* **445**, 168–176 (2007).
14. Langlieb, J. *et al.* The cell type composition of the adult mouse brain revealed by single cell and spatial genomics. (2023) doi:10.1101/2023.03.06.531307.

15. Cembrowski, M. S. & Spruston, N. Heterogeneity within classical cell types is the rule: Lessons from hippocampal pyramidal neurons. *Nat Rev Neurosci* **20**, 193–204 (2019).
16. Economo, M. N. *et al.* Distinct descending motor cortex pathways and their roles in movement. *Nature* **563**, 79–84 (2018).
17. Wang, Q. *et al.* The allen mouse brain common coordinate framework: A 3D reference atlas. *Cell* **181**, 936–953.e20 (2020).
18. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
19. Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
20. Thompson, C. L. *et al.* Genomic anatomy of the hippocampus. *Neuron* **60**, 1010–1021 (2008).
21. Zeng, H. What is a cell type and how to define it? *Cell* **185**, 2739–2755 (2022).
22. Chen, X., Fischer, S., Zhang, A., Gillis, J. & Zador, A. M. Modular cell type organization of cortical areas revealed by in situ sequencing. (2022) doi:10.1101/2022.11.06.515380.
23. Argelaguet, R. *et al.* Multi-omics factor analysis—a framework for unsupervised integration of multi-omics data sets. *Molecular Systems Biology* **14**, e8124 (2018).
24. Welch, J. D. *et al.* Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887.e17 (2019).
25. Timonidis, N., Bakker, R. & Tiesinga, P. Prediction of a cell-class-specific mouse mesoconnectome using gene expression data. *Neuroinformatics* **18**, 611–626 (2020).
26. Wu, S. *et al.* Stability-driven nonnegative matrix factorization to interpret spatial gene expression and build local gene networks. *Proceedings of the National Academy of Sciences* **113**, 4290–4295 (2016).
27. Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).
28. Bienkowski, M. S. *et al.* Integration of gene expression and brain-wide connectivity reveals the multiscale organization of mouse hippocampal networks. *Nat Neurosci* **21**, 1628–1643 (2018).
29. Cahill, R. *et al.* Unsupervised pattern discovery in spatial gene expression atlas reveals mouse brain regions beyond established ontology. (2023) doi:10.1101/2023.03.10.531984.
30. Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014.e22 (2018).
31. Bandler, R. C. *et al.* Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).
32. Knox, J. E. *et al.* High-resolution data-driven model of the mouse connectome. *Netw Neurosci* **3**, 217–236 (2018).