

**Development of Light Microscopy Based Approaches in Neurocartography**

by

Fred Shen

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Neuroscience)  
in the University of Michigan  
2020

Doctoral Committee:

Assistant Professor Dawen Cai, Chair  
Associate Professor Sara Aton  
Assistant Professor Bo Duan  
Assistant Professor Ken Kwan

Fred Shen

fshen@umich.edu

ORCID iD: 0000-0002-2345-780X

© Fred Shen 2020

### **Dedication**

This dissertation is dedicated to my girlfriend, Lorraine, and my parents, Bo and Hong, who have given me unconditional love and unwavering support.

## Acknowledgements

First, I want to thank my mentor Dawen Cai for his support in graduate school. He has given me the freedom to explore my diverse interests and pushed me to achieve more than I thought was possible. I want to thank all the Cai lab members (Ye, Nigel, Logan, Erica, Jimmy, Danny, Xingyu, Chris, Kyle, Alexis) that have come and gone over the past 4 years. You all have been a pleasure to work with and I appreciate your generous help and assistance on my projects. More specifically I want to first acknowledge Douglas Roossien, who has been both a mentor and a friend over the past four years. Thank you for teaching me lab skills, giving me countless advice, and being a kind and steady source of support. I was sad to see you leave the lab, but excited to see the novel directions your new lab explores at Ball State! Next, I want to acknowledge three different research technicians that I have sequentially worked with across the years: HJ Sekhon, Yimeng Zhao, and Meg Harrington. Thank you HJ for teaching me biochemistry and for contributing to the work presented in this dissertation. I very much enjoyed collaborating on projects, bouncing crazy ideas off each other late at night, and dreaming up new tools for neuroscience. Thank you Yimeng for working on a biochemistry side project with me that was fun to develop together and troubleshoot. Thank you Meg for the past year of assistance in helping me wrap my projects up. It was nice to work so in sync that you could already predict what I was going to ask you to do. Next, I want to acknowledge my committee members: Ken Kwan, Sara Aton, and Bo Duan. Thank you so much for your helpful comments and suggestions across all my committee meetings. I grew so much as a scientist through my interactions with you. I would not be here today without the support of all the administrative staff from MSTP, NGP, and CDB over the years. I want to acknowledge Ron Koenig for his

mentorship and guidance throughout the last 6 years. You gave me wonderful and wise advice during some of my toughest times in graduate school, and I wish you nothing but the best in retirement. I am indebted to Justine Hein, Ellen Elkin, Hilikka Ketola, and Liz Bowman for helping take care of administrative details behind the scenes so that I could focus on being a student. I also want to thank Audrey Seasholtz, Ed Steunkel, Valerie Smith, Rachel Harbach, and Les Satin from NGP and Deontae Hooks, Brooke Lorigan, Brittany Longeway and Kristen Hug from CDB for your administrative assistance. In addition, I am tremendously grateful to my friends I have made in graduate school (Ahmed, Tony, Dalia, Sarah, Lexi, Shannon, Dan, Dipika, Danny, Lauren, among others) for all the laughs we have shared. Finally, I want to thank Andy Lieberman for taking me under his wing as a JCI Insight Scholar over the past few months. It has been an amazing experience reviewing papers with you and learning how the peer review process goes from a journal's perspective.

## Table of Contents

<b>Dedication</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>x</b>
<b>Abstract</b>	<b>xii</b>
<b>Chapter 1. Introduction</b>	<b>1</b>
1.1 Dissertation Structure	1
1.2 Studying neuroanatomy at multiple scales	1
1.2.1 Mesoscale neurocartography	3
1.2.2 Microscale neurocartography	3
1.2.3 Nanoscale neurocartography	4
1.3 Spatially dense labeling generates more biological insights	5
1.3.1 Higher optical resolutions enable denser labeling	6
1.3.2 Increased color diversity using Brainbow permits denser labeling	6
<b>Chapter 2. Optimizing Expansion Microscopy Protocols</b>	<b>10</b>
2.1 Limitations of histology imaging	10
2.2 Expansion microscopy permits super-resolution imaging using routine light microscopy.	11
2.3 Expansion microscopy principles and workflow	12
2.4 Results	13
2.4.1 Optimizing acryl NHS-ester crosslinker concentration	13
2.4.2 Optimizing gel monomer density	14

2.4.3 Optimizing protein extraction for post-gelation immunostaining	15
2.4.4 Homogenous antibody labeling can be achieved in thick tissue	15
2.5 Discussion	16
2.5.1 Summary	16
2.5.2 Expansion microscopy parameters need to be adapted for the priorities in each experiment	16
2.5.3 Expanding the whole mouse brain	17
2.6 Methods	18
2.7 Acknowledgements	19
<b>Chapter 3. Development of New Brainbow Viral Tools</b>	<b>25</b>
3.1 Introduction	25
3.1.1 The fluorescent protein revolution	25
3.1.2 Sparse fluorescent protein labeling in the brain	26
3.1.3 Dense fluorescent protein labeling using Brainbow	26
3.1.4 Analog coding in Brainbow follows the RGB color model	27
3.1.5 Digital coding in Brainbow	28
3.1.6 Brainbow color outcomes and balance	29
3.1.7 Brainbow viral tools	29
3.1.7.1 'On/off' Brainbow AAVs	29
3.1.7.2 'Always on' Brainbow AAVs	30
3.2 Results	30
3.2.1 'On/off' design outperforms 'always on' design in both analog and digital color coding	30
3.2.2 Extending the 'on/off' Brainbow AAV toolbox	32
3.2.2.1 Promoter choice	32
3.2.2.2 Twelve antigenically distinct new fluorescent proteins	32
3.2.2.3 Simultaneous membrane and cytoplasmic expression	33
3.2.2.4 Systemic delivery using PHP.eB capsid AAVs	33
3.2.3 Simulating Brainbow AAV barcoding in mouse brains	34

3.3 Discussion	35
3.3.1 Summary	35
3.3.2 Developing multiplexed readout strategies	35
3.3.3 Future fluorescent proteins to pursue	36
3.3.4 Subcellular labeling strategy to generate more barcodes	36
3.3.5 Developing a new image processing pipeline	37
3.3.6 Brainbow AAVs offer several advantages over transgenic mice	38
3.3.7 Extending Brainbow to neurons beyond the brain	39
3.4 Methods	39
3.5 Acknowledgements	42
<b>Chapter 4. Light Microscopy Based Approach for Mapping Connectivity with Molecular Specificity</b>	<b>51</b>
4.1 Introduction	52
4.2 Results	53
4.3 Discussion	59
4.4 Methods	62
4.5 Acknowledgements	66
4.6 Author contribution	67
<b>Chapter 5. Discussion and Future Directions</b>	<b>91</b>
5.1 Summary of contributions	91
5.2. Platform for high throughput microscale neurocartography	91
5.2.1 Motivation	91
5.2.2 Specimen preparation	92
5.2.3 Light imaging	93
5.2.4 Image processing	93
5.2.5 Adopting automatic segmentation algorithms to reduce manual tracing	94
5.2.6 ‘Traceless’ mapping of axon projections at single neuron resolution	94
5.2.7 Morphometric analysis	95
5.2.8 Open access and well annotated pipelines	95



5.2.9 Biological applications	96
5.3 Spectral connectomics can be applied in a bottom up approach to understand higher order brain functions	96
<b>References</b>	<b>98</b>

### **List of Tables**

Table 3.1 Table of fluorescent proteins arranged by order, family, and genus.	49
Table 3.2 Matrix of protein sequence homologies.	50
Table 4.1 Experimental Conditions	88
Table 4.2 Primary Antibodies	89
Table 4.3 Secondary Antibodies	90

## List of Figures

Figure 1.1 Spatial scales of neurocartography	8
Figure 1.2 Dense neuron reconstruction is facilitated by increasing color diversity and/or increasing imaging resolution	9
Figure 2.1 Expansion microscopy achieves super-resolution through physical magnification of the sample.	20
Figure 2.2 Schematic of two complementary expansion microscopy workflows.	21
Figure 2.3 Effects of NHS-ester concentration on post-gelation immunostaining expansion microscopy protocols.	22
Figure 2.4 Effect of different monomer concentrations on immunostaining.	23
Figure 2.5 Post-gelation immunostaining expansion microscopy protocol enables homogenous immunostaining in thick tissue.	24
Figure 3.1 Analog versus digital generation of Brainbow color barcodes.	43
Figure 3.2 Modeling 'on/off' versus 'always on' viral designs.	44
Figure 3.3 Six new Brainbow AAV vectors with simultaneous membrane and cytoplasmic targeting.	46
Figure 3.4 Brainbow AAVs packaged in PHP.eB capsids can be injected intravenously to systemically infect neurons across the brain.	47
Figure 3.5 Effect of color barcoding capacity on labeling different neural population sizes.	48
Figure 4.1 miriEx robustly preserves antigens across multiple rounds of immunostaining, imaging, and stripping.	68
Figure 4.2 miriEx can be combined with Brainbow to simultaneously profile molecular expression and neuron morphologies.	69
Figure 4.3 Putative synaptic connections can be defined between Brainbow neurons.	71
Figure 4.4 Molecularly specific analyses can be performed on spectral connectomics data.	73
Figure 4.5 Putative excitatory and inhibitory synapses of aspiny inhibitory neurons can be quantified.	75

Supplemental Figure 4.1 Heat and SDS are efficient at eluting antibody from miriEx gel-tissue hybrids.	77
Supplemental Figure 4.2 miriEx gel-tissue hybrid expansion.	78
Supplemental Figure 4.3 Image registration pipeline between different rounds of miriEx.	79
Supplemental Figure 4.4 miriEx enables multiplexed immunostaining in formalin fixed human brain tissue.	80
Supplemental Figure 4.5 miriEx enables highly multiplexed molecular profiling.	81
Supplemental Figure 4.6 Morphology quantification of 53 inhibitory neurons traced in the BLA.	83
Supplemental Figure 4.7 Homer1 and Gephyrin can be paired with Bassoon immunostaining to represent the majority of excitatory and inhibitory synapses.	84
Supplemental Figure 4.8 AAV-PHP.eB Brainbow expression in PV-Cre mice is highly sensitive and specific.	85
Supplemental Figure 4.9 Brainbow FPs, synaptic machinery, and molecular markers can be interrogated in the same piece of tissue.	86
Supplemental Figure 4.10 Number of unique axon colors can be identified for k-means clustering.	87

## Abstract

The process of mapping neuroanatomy at multiple scales is defined as neurocartography and has the ultimate goal of revealing a complete wiring diagram of synaptic connections. Neurocartography is an important pursuit because brain function is intricately linked to neuroanatomy, similar to how the function of a protein depends on its structure. Here, we developed new light microscopy approaches for neurocartography at both microscale and nanoscale levels to map single neuron morphologies and synaptic connectivity respectively. Spatially sparse labeling of neurons has been necessary when studying neuron morphologies to minimize overlap and avoid ambiguities during reconstructions. We developed two strategies to overcome this limitation and achieve spatially dense labeling. First, we used a multicolor genetic labeling (Brainbow) approach to stochastically express fluorescent proteins in a spatially dense population of neurons, facilitating reconstruction of single neuron morphologies. We extend the Brainbow viral toolbox by 1) introducing 12 fluorescent proteins in the form of 6 new viruses, 2) using a membrane and cytoplasmic dual labeling strategy, and 3) adopting the AAV.PHP.eB capsid to systemically induce expression and better control color diversity. Second, we used expansion microscopy to increase confocal imaging resolution by physically magnifying brain samples. We developed a multi-round immunostaining Expansion Microscopy (miriEx) protocol that enables multiplexed protein detection at multiple imaging resolutions. We then combined Brainbow with miriEx to simultaneously map morphology, molecular markers, and connectivity in the same brain section. We define the derivation of these properties from hyperspectral fluorescent channels as spectral connectomics, a light microscopy based approach towards mapping neuroanatomy and connectivity with molecular specificity. We applied our multimodal profiling strategy to directly link inhibitory neuron cell types with their network morphologies.

Furthermore, we showed that correlative Brainbow and endogenous synaptic machinery immunostaining can be used to define putative synaptic connections between spectrally unique neurons, as well as map putative inhibitory and excitatory inputs. We envision that spectral connectomics can be applied routinely in neurobiology labs to gain insights into normal and pathophysiological neuroanatomy across multiple animals and time points. We hope that the light microscopy approaches developed in this dissertation will facilitate the extraction of new biological insights from neurocartography.

# Chapter 1. Introduction

## 1.1 Dissertation Structure

This dissertation is organized in 5 major chapters. In chapter one I will motivate the importance of mapping neuroanatomy and briefly review the state of neuroanatomy research using spatial scales as a guide. In chapter two, I will discuss our development and optimization of expansion microscopy protocols. In chapter three, I will review Brainbow technology and describe the rationale and design of new Brainbow virus tools. Chapter 3 includes data that is being prepared in a manuscript for submission. In chapter four, I will demonstrate how I combined expansion microscopy and Brainbow to develop a light microscopy based approach for mapping connectivity with molecular specificity called spectral connectomics. Chapter four represents a manuscript that was submitted and is currently under review. In chapter five, I will summarize my dissertation research and describe future directions.

## 1.2 Studying neuroanatomy at multiple scales

Our brains are the most complex organ in our bodies, acting as a computational device that transforms sensory input into motor output. A fundamental discovery that marked the birth of modern neuroscience came from studying neuroanatomy. At the turn of the 19th century, Santiago Ramón y Cajal used golgi staining to sparsely label neurons and made extensive drawings of the structures he saw. Cajal observed that unlike other cells in the body, neurons had polar morphologies consisting of elaborate dendritic and axonal arbors. This led to a hypothesis that neural information flow was also polar and transmitted from dendrite to soma to axon, which has proven true in our modern understanding that neurons are the principle units of the brain, wired together to perform different computations. Cajal's work is a wonderful example of the importance of neuroanatomy and that knowledge of how our brains function can stem from understanding the underlying structure.

One hundred years later, we now recognize that one of the major goals in neuroscience is to map a comprehensive wiring diagram of how each neuron connects to one another, or a connectome. In 1986, Sydney Brenner and colleagues published a landmark paper detailing the full wiring diagram of 302 neurons in *C. Elegans*<sup>1</sup>. This important resource has influenced future studies of *C. Elegans* neurobiology by providing a structural map to constrain and generate new hypotheses. Although the utility of mapping connectomes has been debated, one intuitive analogy is to think about a connectome as a map of all the roads in the country. It does not tell you *where* the cars, or neural signals, are driving, but it does tell you all the *possible routes* cars can take and is fundamental to understanding neuropsychiatric disorders that may *disrupt* how cars reach their destinations.

Compared to *C. Elegans*, the human brain represents an exponential challenge, containing 100 billion neurons, 100 thousand miles of neuron arbors, and 100 trillion to 1 quadrillion synapses<sup>2</sup>. The sheer scale of mapping the human connectome is extremely daunting, and to simplify the problem, neuroscientists have turned to mapping the connectome of smaller model animals. The mouse brain is more feasible, but still challenging. It is three orders of magnitude smaller, containing ~80 million neurons and 1 trillion synapses. Consequently, in our quest to map the mouse connectome, we have sought to simplify the problem even further and map the neuroanatomy of the mouse brain at several larger scales to enhance our understanding of the underlying connectivity. Neuroscientists have not always uniformly followed the precise definition of connectome. For example, connectome has been used to refer to the connections between human brain regions as mapped by diffusor tensor imaging<sup>2</sup>, or connections between different mouse brain areas as mapped by axon projections<sup>3</sup>. For the rest of this dissertation, I will use the more precise definition of connectome as a comprehensive wiring diagram defined at the synaptic level, and instead refer to the *multiple scales of mapping neuroanatomy* using a broader term: neurocartography. Below, I will briefly describe neurocartography research at three different spatial scales - mesoscale, microscale, and nanoscale - to conceptualize how new insights are being integrated to solve the mouse connectome (**Fig. 1.1**).



### 1.2.1 Mesoscale neurocartography

Mesoscale neurocartography focuses on mapping the connectivity between different brain regions. The mouse brain can be split into anatomically, neurochemically, and functionally distinct areas such as motor cortex, sensory cortex, visual cortex, etc. These different regions have a rich history of literature detailing their circuitry and how they affect sensory input and behavioral output. As a result, one natural question that emerges is how do these different brain regions connect and transmit information to one another? Hongkui Zeng's group at the Allen Institute and Hong Wei Dong's group at USC have approached this question by injecting anterograde tracers in different brain regions, followed by systematic imaging of serial sections and quantification of axonal projection patterns to produce a connectivity matrix between identified brain regions<sup>3-5</sup>. Anterograde tracers refer to substances that fill neural axon arbors, such as fluorescent proteins, PHA-L, and dextran amines. The key assumption in their work is that the presence of axons in target areas implies synaptic connectivity between the two regions. Although likely true, follow up studies using methods like Channelrhodopsin assisted circuit mapping<sup>6</sup> are needed to confirm functional synaptic connectivity. While examining *output* projections between region A and its downstream target regions can be used to determine connectivity, other studies adopted a complementary approach. They used rabies virus monosynaptic tracing to map the *inputs* between region A and the upstream regions that innervate it<sup>7,8</sup>. The brainwide connectivity maps generated using these complementary studies can then be used to generate hypotheses about which regions are part of circuits based on potential connectivity and worthy of further investigation.

### 1.2.2 Microscale neurocartography

While mesoscale neurocartography focuses on mapping the structure of a *population* of neurons, microscale neurocartography focuses on mapping the structure of *single neurons*. Individual neuron morphologies, including subcellular features like dendritic spines and axonal boutons, are important to study because they can be a structural readout of maladaptive changes seen in neuropsychiatric disorders<sup>9,10</sup>. They can also help us chart the circuit organization of a local region. For example, the dendritic and axonal projection pattern of a neocortical neuron across the 6 cortical layers indicates the neuron's role in transmitting information between different layers. Knowing the dendritic and axonal fields of multiple neurons

in the same region can also help generate hypotheses about their connectivity. Peters' rule states that the degree of spatial overlap between axons and dendrites correlates with potential connectivity<sup>11</sup>. Finally, single neuron morphologies can be used to classify neuronal cell types. For instance, molecular layer interneurons in the cerebellum can be divided into basket cells that innervate the perisomatic region of Purkinje cells or stellate cells that innervate the dendrites of Purkinje cells. Single neuron morphologies are typically studied by labeling a sparse population of neurons with a contrast agent. For example, following patch clamp recording, we can fill the cell with biotin and perform histological analysis of the neuron's morphology. However, three distinct challenges remain. First, neuron arbors can be enormous and difficult to capture in their entirety. The majority of studies use less than 300  $\mu\text{m}$  thick tissue sections, meaning they only sample the proximal portion of a neuron's morphology. Second, the tortuous nature of axons along with their small diameters (hundreds of nanometers in some portions) make full axonal arbors difficult to reconstruct<sup>12</sup>. Third, sparse labeling strategies, which are necessary to limit ambiguity between neighboring neurons, are inherently low throughput and laborious.

### 1.2.3 Nanoscale neurocartography

Nanoscale neurocartography focuses on mapping the synaptic connections between neurons. Chemical synapses are typically located at the junctions between axons and dendrites and are composed of structures hundreds of nanometers in size. On the presynaptic side, the active zone is composed of proteins that mediate neurotransmitter release. On the postsynaptic side, a dense scaffold of proteins known as the postsynaptic density (PSD) regulates neurotransmitter receptors at the synapse. Mapping synapses and the subcellular organelles that accompany them is an important endeavor because they represent the fundamental connection point between neurons and are a critical link between how a neuron's cell biology can alter its physiological properties and subsequently its role in the circuit. Nanoscale neurocartography studies have mostly relied on electron microscopy (EM) because of the small size of synaptic structures and diffraction limited resolution of light microscopy. The combination of close membrane appositions and visualization of active zones and PSDs allows us to define synaptic connectivity between neurons from an ultrastructural perspective<sup>13</sup>. An alternative strategy involves using super-resolution light microscopy (srLM) techniques, such as STORM<sup>14</sup> or STED<sup>15</sup>, to visualize synaptic structures with nanoscale resolution. Although the optical

resolution is worse than EM, super resolution light microscopy techniques have the advantage of preserving molecular content<sup>16</sup>. EM relies on a heavy metal stain to produce contrast, mostly labeling membranes and organelles. As a result, specific protein information is lost. In contrast, super resolution light microscopy techniques can use antibody labeling to probe for specific proteins and reveal the molecular content of the synapse.

While mesoscale, microscale, and nanoscale neurocartography can yield distinct research objectives, they actually are quite intertwined. For example, if we fully map the mouse connectome at the nanoscale level, then that dataset would also contain every neuron's morphology and complete our knowledge of microscale connectivity. Subsequently If we knew the axon projections for every neuron in a brain region at the microscale level, we could also obtain the mesoscale connectivity between different brain regions as well. As a result, mesoscale neurocartography can be thought of as a subset of microscale neurocartography, with both being a subset of nanoscale neurocartography.

### 1.3 Spatially dense labeling generates more biological insights

My dissertation research has primarily focused on developing light microscopy based approaches in both microscale and nanoscale neurocartography. In the realm of microscale and nanoscale neurocartography, we have identified that sparse labeling is a major technological hurdle because it is low throughput and prevents us from studying multiple neuron morphologies and how they interact spatially. Achieving dense labeling of neurons that is still amenable to single cell reconstructions would be an important tool in a neurocartographer's arsenal. It would increase the throughput of experiments by allowing us to reconstruct hundreds of neurons in one tissue sample rather than just a few, as well as lower the number of animals sacrificed. We would also be able to measure how different neurons interact in the same sample, i.e. ask what spatial relationships exist between specific neuron cell types? Do they tile and avoid each other, or do they overlap and sample the same volume? Dense labeling would also set the stage for integrating nanoscale cartography by allowing us to sample a population of neurons and examine their synaptic connectivity.

### 1.3.1 Higher optical resolutions enable denser labeling

One method to increase labeling density is to increase the optical imaging resolution (**Fig. 1.2 b-d**). Low optical resolution (especially axial) hinders the reconstruction, or “traceability”, of neurons by convolving the neurites in space. For example, 1  $\mu\text{m}$  XYZ imaging resolution would be insufficient to resolve two axons coursing close to one another, separated by 500 nm in space, leading to a tracing ambiguity where we would not know which axon belonged to which neuron. Increasing the optical imaging resolution would allow us to distinguish the two axons and trace them from their respective cell bodies accurately and reliably. Another example would be to think of EM nanoscale cartography studies as the furthest point along the axis, where extremely high imaging resolution on the order of  $\sim 10$  nm allows us to achieve saturated labeling of every neuron’s membrane and reconstruct their morphologies unambiguously.

Increasing imaging resolutions in routine light microscopy methods, like confocal, can be accomplished by using higher numerical aperture (NA) objectives. A high NA (1.4) objective can achieve lateral resolutions up to 200 nm and axial resolutions up to 500 nm in confocal microscopy, which is sufficient to reconstruct most of a neuron’s morphology except for small axons. To break the diffraction limit and achieve higher imaging resolutions, super-resolution techniques like STORM or expansion microscopy are required<sup>14,17</sup>. In the latter method, higher imaging resolutions result from physical magnification of the sample. As a result, no special imaging equipment is required and expansion microscopy is compatible with routine light microscopy methods.

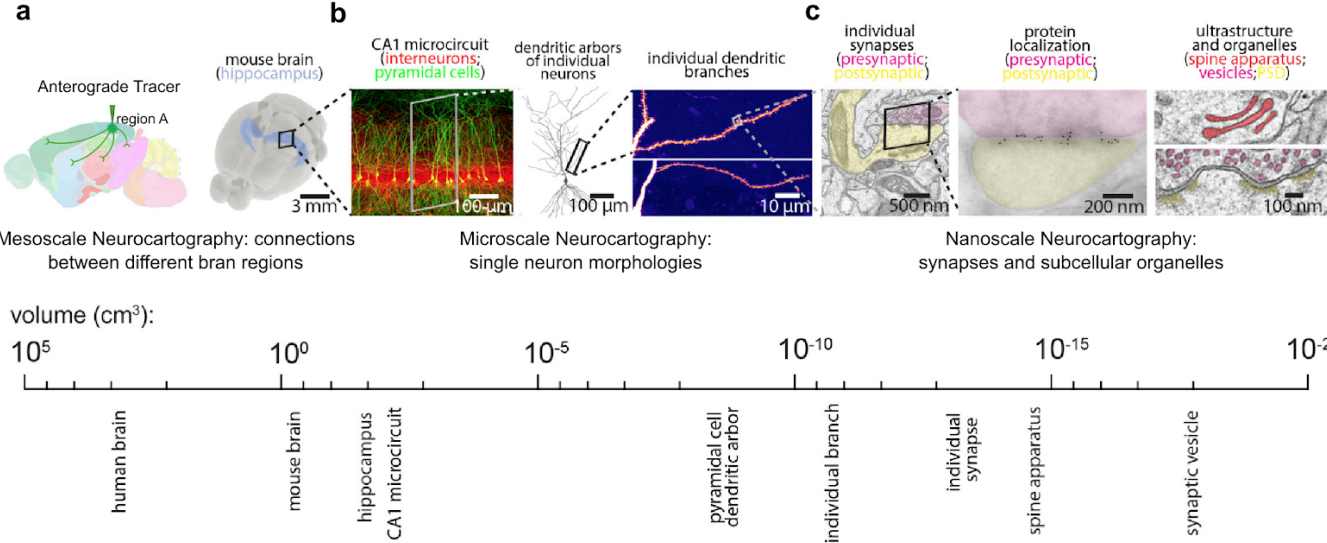
### 1.3.2 Increased color diversity using Brainbow permits denser labeling

Another strategy to increase the density of labeling is to use Brainbow, a genetic labeling method to stochastically express multiple fluorescent proteins in each neuron<sup>18</sup> (**Fig. 1.2c**). Typically, a single fluorophore is used as a contrast agent in sparse labeling studies of neuron morphologies. In 2007, Livet and colleagues pioneered a new genetic tool, Brainbow, that expressed random combinations of three fluorophores to label neurons<sup>18</sup>. Combinatorial expression of the three fluorophores yielded unique color expression in neurons. Denser labeling was now compatible with single neuron reconstructions because even when neighboring neural arbors overlapped, they could be distinguished unambiguously based on

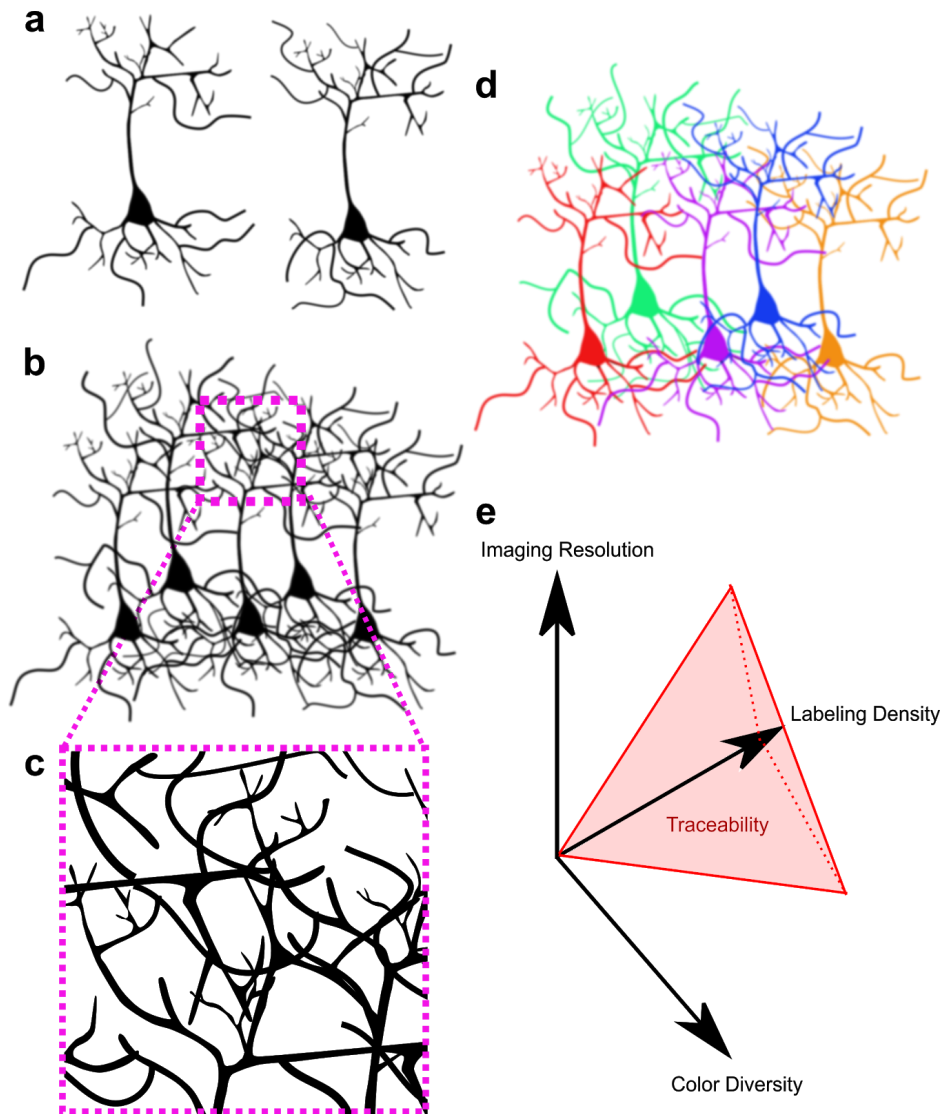
color. Therefore, a Brainbow strategy of achieving dense labeling is constrained by color diversity, a metric that factors in the number of color outcomes and probability of their expression. Low color diversity, as seen in the Rgbow mouse<sup>19</sup>, limits the density of labeling because neighboring neurons will have a high probability of being the same color, otherwise known as the color collision rate. Increasing the color diversity, such as by using viral Brainbow tools described in Chapter 4, increases the density of labeling because color collision rates are lowered.

Thus, imaging resolution and color diversity represent two important parameters that affect the “traceability” of neurons (**Fig. 1.2e**), and can be adjusted relative to the biological question being asked. For instance, increasing the color diversity of a neural population can be paired with lowering imaging resolution to reach the same traceability as before. Chapter 2 of this dissertation will discuss my efforts to increase the optical imaging resolution through optimizing expansion microscopy protocols. Chapter 3 will describe my approach to increase color diversity by developing new Brainbow virus tools. Chapter 4 will illustrate how I combined these tools together and integrated microscale and nanoscale neurocartography to define synaptic connectivity between densely labeled neurons.

# 1.3 Figures



**Figure 1.1 Spatial scales of neurocartography.** (a) Mapping connectivity throughout the mouse brain can occur at several different scales. Mesoscale neurocartography is primarily focused on mapping the structure of a population of neurons, i.e. the connectivity matrix between neuroanatomically/functionally distinct areas. This can be accomplished through anterograde tracer injections in region A, and then mapping axonal projections to downstream areas. (b) Microscale neurocartography is primarily focused on investigating the structure of single neurons. This can be accomplished through labeling single neurons with a contrast agent to visualize and reconstruct their axonal and dendritic arbors. (c) Nanoscale neurocartography is primarily focused on elucidating the structure of synapses and their subcellular organelles. This can be accomplished through electron microscopy imaging of synaptic structures to measure connectivity between single neurons. Adapted from Zeng et al. (2017)<sup>20</sup> and Bloss et al. (2019)<sup>21</sup>.



**Figure 1.2 Dense neuron reconstruction is facilitated by increasing color diversity and/or increasing imaging resolution.** (a) Example of sparsely labeled neurons, resulting in minimal overlap of neurite arbors, letting us trace the morphology accurately. (b) Example of densely labeled neurons. The overlap between neighboring neuron neurite arbors makes it difficult to distinguish which branch belongs to which neuron and hinders accurate tracing. (c) Example of how increased imaging resolution can also facilitate more accurate reconstruction. The magenta box corresponds to the inset seen in b, but at 4x greater imaging resolution, which decreases the overlap between neurites in volumetric space and increases our ability to trace each neuron. (d) Example of how Brainbow can facilitate more accurate reconstructions in the same dense population of neurons. Neighboring neuron morphologies can now be distinguished based on color. (e) Summary graph showing the relationship between labeling density, imaging resolution, and color diversity. The red pyramid represents “traceable” datasets given proper optimization of the three parameters.

# Chapter 2. Optimizing Expansion Microscopy Protocols

## 2.1 Limitations of histological imaging

*In situ* examinations of brain samples have been traditionally limited to thin (<100  $\mu\text{m}$ ) tissue sections. Imaging deeper in tissue results in blurry images caused by light scattering from refractive index mismatches. Although many neuroscientists are still able to answer important questions that do not require deep imaging, neurocartographers are limited when a neuron's structure can be much greater than 100  $\mu\text{m}$ . To address this problem, various clearing protocols have been developed to make the tissue more transparent and reduce light scattering. The four distinct strategies are organic solvent based clearing, simple immersion in a refractive index matching solution, hyperhydration based clearing, and hydrogel based clearing<sup>22</sup>. I will not elaborate on the former three, and instead focus on the last strategy for the rest of this chapter. The story of hydrogel based clearing began in 1981 when Hausen and Dreyer demonstrated that *Xenopus* embryos and tadpoles could be embedded in a polyacrylamide gel<sup>23</sup>. In 2013, Chung and colleagues employed a similar strategy (CLARITY) to embed brain samples in a polyacrylamide hydrogel scaffold<sup>24</sup>. The idea was that the hydrogel would preserve proteins and other biomolecules, while lipids were removed by SDS delipidation to reduce light scattering. Combined with refractive index matching, the end result was a transparent brain sample that could be imaged deeply with minimal scattering.

Diffraction limited light microscopy is another barrier in neurocartography. Standard light microscopy methods, such as confocal, have a maximum resolution of ~200-300 nm. Although sufficient for microscale neurocartography, nanoscale structures like synapses, cytoskeleton filaments, and other subcellular organelles remain difficult to study. Super-resolution light microscopy methods have been successfully applied to fill this gap. For example, Sigal and



colleagues developed a platform using STORM to investigate inhibitory synaptic input fields of retinal neurons<sup>25</sup>. They immunostained retinal tissue, embedded them in resin, made 70 nm serial sections, imaged them using STORM, and stitched the slices together for volumetric reconstruction. Compared to electron microscopy, Sigal et al.'s approach has the advantage of identifying specific synaptic proteins with immunostaining. For example, glycine and GABA receptors were separately labeled and mapped across a neuron's dendrites. Although electron microscopy offers greater resolution, the authors achieved 20 nm lateral and 70 nm axial resolution using STORM, which is sufficient for mapping many synaptic structures. The biggest limiting factors were complicated specimen preparation and long imaging time (~3 days for a ~100 x 100 x 20  $\mu\text{m}^3$  volume).

## 2.2 Expansion microscopy permits super-resolution imaging using routine light microscopy.

In 2015, Chen, Tilberg, and Boyden built off the idea of embedding brain sections in hydrogels and established the concept of expansion microscopy<sup>26</sup>. They introduced sodium acrylate into the hydrogel monomer recipe, a critical polymer found in diapers and artificial snow that allows the gel to absorb water and expand up to 4.5 fold. The expansion process effectively dilutes the refractive index of the tissue to that of water, rendering it completely transparent with minimal light scattering during imaging. Optical magnification through ~4x physical expansion of brain samples enabled Chen et al. to break the diffraction limit and achieve super-resolution using standard light microscopy. For example, confocal imaging of 4x expanded samples at 400 nm lateral resolution would result in an effective pre-expansion lateral resolution of 100 nm. Importantly, the gel-tissue hybrids were found to expand isotropically and preserve spatial relationships between biomolecules. Other super-resolution methods, like STED or STORM, rely on sharpening the point spread function or superimposing the centroid-fitted point spread functions of individual fluorophores to increase resolution respectively. In contrast, expansion microscopy does not shape point spread functions, but instead physically expands the distance between them to increase effective resolution (**Fig. 2.1**). Although STED/STORM initially had the resolution advantage, recent expansion microscopy protocols have further modified the hydrogel recipe to achieve greater expansion factors (up to 10x) to match STED/STORM's resolution<sup>27</sup>. Expansion microscopy has the added advantage of being compatible with high speed and large volume imaging modalities, like lightsheet microscopy<sup>16</sup>. As a result, expansion

microscopy is an ideal tool for neurocartographers to map nanoscale structures across large neuron volumes.

## 2.3 Expansion microscopy principles and workflow

Two expansion microscopy workflows have since emerged that I will briefly describe below (**Fig. 2.2**). For the first workflow, thin (<100  $\mu\text{m}$ ) brain samples first undergo standard immunohistochemistry protocols. Following immunostaining, the sample is treated with a N-Hydroxysuccinimide (NHS) ester that selectively adds acryl groups at a protein's primary amines. Next, the sample is embedded into a hydrogel composed of acrylamide, sodium acrylate, and bis-acrylamide monomers. Each protein's acryl groups serve as anchoring points that crosslink the protein into the hydrogel scaffold during polymerization. The gel-tissue hybrid is then mechanically homogenized using a nonselective protease (Proteinase K) to extract proteins and enable isotropic expansion. Some of the dye conjugated secondary antibodies are lost in this step (presumably the ones that were not well crosslinked) and fluorescence quenching can also occur, but enough signal remains for useful imaging. For example, cyanine dyes like Alexa 647 are incompatible with proExM because the fluorescence is lost<sup>28</sup>. The sample will now be ~2 fold expanded in 1xPBS, and almost completely transparent. By dialyzing in water, the sample can expand 4-4.5 fold, enabling confocal imaging resolutions up to 50 nm laterally.

The second workflow uses most of the same steps, but rearranges them such that immunostaining is performed *after* gelation compared to *before*. Here, the brain sample is first treated with NHS ester to modify proteins with acryl groups, and then embedded in the hydrogel. Instead of Proteinase K digestion, the sample undergoes SDS and high heat based protein denaturation and protein extraction. The reasoning is that Proteinase K digests most proteins and destroys epitopes, limiting the effectiveness of immunostaining. This explains why in the first workflow, antibody labeling is performed before Proteinase K digestion. In contrast, SDS/heat treatment preserves epitopes and is compatible with post-gelation immunostaining<sup>29</sup>. The caveat is that the antibodies have to recognize denatured epitopes, which is common as many antibodies are simultaneously validated for both immunohistochemistry and western blotting applications<sup>29</sup>.

## 2.4 Results

Our goal was to image thick brain sections to capture entire neuron morphologies, and with sufficient resolution to map nanoscale structures. To meet this goal, we optimized an expansion microscopy protocol by exploring the parameter space of three main steps: acryl NHS-ester crosslinking, gel monomer composition, and protein extraction for isotropic expansion. The end result was an optimized post-gelation immunostaining protocol called miriEx (**m**ulti-**r**ound **i**mmunostaining **E**xpansion microscopy) that is described more fully in Chapter 4<sup>30</sup>. Instead, in the rest of this chapter I will describe our experiences optimizing different aspects of the protocol and demonstrate that it can be applied to thick brain sections.

### 2.4.1 Optimizing acryl NHS-ester crosslinker concentration

We empirically found that for the post-gelation immunostaining expansion microscopy workflow, the concentration of acryl NHS-ester (Aax) is an important variable (**Fig. 2.3**). Too little Aax minimizes crosslinking of proteins with the hydrogel scaffold, leading to an overall decreased preservation of proteins. Although this results in a more transparent and expanded sample, the signal from immunostaining is decreased since there is less antigen for antibodies to bind to. In contrast, too much Aax increases crosslinking of proteins with the hydrogel and maximizes preservation of proteins. The result is a less expanded and more opaque gel-tissue sample because increased crosslinking prevents the gel from expanding. Another negative is that the excessive modification of proteins with acryl groups can mask epitope sites and reduce antibody binding.

Consequently, optimizing the Aax concentration is important for obtaining good expansion and immunostaining results. Although the optimal Aax concentration varies for each protein, we have empirically found that values from 1 to 5 mM tend to be sufficient (data not shown). We can control the degree of Aax action not only by varying the concentration, but also through the buffer used. Different groups have demonstrated that a 2-(N-morpholino)ethanesulfonic acid (MES) based saline buffer at pH 6 results in less efficient crosslinking compared to 1xPBS at pH 7<sup>28</sup>. In contrast, a sodium bicarbonate buffer at pH 8.5 results in more efficient crosslinking<sup>27</sup>.

## 2.4.2 Optimizing gel monomer density

Another important variable that affects the quality of immunostaining is the hydrogel monomer composition. The hydrogel is composed of three polymerizable reagents: sodium acrylate, acrylamide, and bis-acrylamide. The former two form long polymer strands, while bis-acrylamide serves as a crosslinker to weave the strands into a hydrogel matrix. Varying the monomer recipe can affect the stability of the gel, expansion factor, and efficacy of immunostaining. We and others have found that lowering bis-acrylamide leads to higher expansion factors, but less stable gels<sup>31,32</sup>. Increasing the proportion of sodium acrylate can also lead to higher expansion factors and less stable gels to a smaller degree, while increasing total density of acrylamide and sodium acrylate result in more stable gels. Taken as a whole, different monomer recipes can be used to balance gel stability and expansion factor depending on the imaging resolution required. A large unknown in the field, however, is the effect of different gel recipes on immunostaining.

To fill this gap, we processed 100  $\mu\text{m}$  GFP expressing brain sections with a post-gelation immunostaining expansion microscopy protocol and tried 4 different monomer recipes. The first recipe was 5.3% sodium acrylate, 4% acrylamide, and 0.1% bis-acrylamide, and represents the formula used in miriEx. The second recipe was 10% sodium acrylate, 30% acrylamide, and 0.1% bis-acrylamide, and represents the formula used by Ku et al. in their MAP protocol<sup>29</sup>. The third recipe was 20% sodium acrylate, 10% acrylamide, and 0.1% bis-acrylamide, and approximates the formula used by Gambarotto et al. in their U-ExM protocol<sup>33</sup>. The fourth recipe was similar, but with 0.01% bis-acrylamide to test the effects of lowering bis-acrylamide. We performed post-gelation immunostaining with an anti-GFP antibody and observed striking differences in the results. The miriEx gel recipe initially yielded higher signal than the rest, which is partially explained by the fact that miriEx gel had lower signal dilution from smaller volumetric expansion compared to the other gels that expanded more (**Fig. 2.4a**). However, after normalizing for different expansion sizes, we found that miriEx and low bis-acrylamide U-ExM gels had similarly high signal levels with the other two gels significantly lower. Our interpretation is that increased monomer density limits the pore size of the gel matrix and inhibits antibody diffusion. This is especially evident when comparing 0.1% and 0.01% bis-acrylamide U-ExM recipes, in which decreased gel crosslinking increased the normalized fluorescent signal. In further support of this model, we found that MAP and U-ExM recipes had dimmer neurons in the

center of the tissue compared to the surfaces, indicating that antibody was not able to diffuse completely. In contrast, miriEx and 0.01% bis-acrylamide U-ExM recipes had similarly bright neurons in the middle and surface of the tissue, indicating more homogenous antibody penetration.

### 2.4.3 Optimizing protein extraction for post-gelation immunostaining

To homogeneously label thick brain samples, we chose to use a post-gelation immunostaining expansion microscopy protocol. Achieving uniform labeling in thick brain samples using standard immunohistochemistry (pre-gelation immunostaining) is difficult because antibodies diffuse slowly at 4°C and can be depleted and trapped in a dense antigen matrix. Post-gelation immunostaining mitigates these barriers by generating faster antibody diffusion at higher temperatures and a more porous antigen matrix. To elaborate, the gel-tissue hybrid is very stable, allowing the sample to survive higher temperatures at which normal tissue samples would degrade. The SDS denaturation step also removes lipids that hinder antibody diffusion and extracts proteins for a more porous antigen matrix<sup>24,34</sup>. Also contributing is the fact that the gel-tissue hybrid is partially expanded during immunostaining, which leads to molecular decrowding and exposure of new epitope sites<sup>35</sup>.

We first tried to replace Proteinase K with other more selective proteases in an effort to preserve proteins for post-gelation immunostaining. We tried collagenase to selectively digest collagens in the extracellular matrix and trypsin to selectively digest the carboxyl side of basic amino acids, but neither of them worked. Although one or two antibodies worked in post-gelation immunostaining, the majority of proteins were still lost during digestion across the panel of antibodies we tested (data not shown). Instead, we found that SDS/heat based denaturation and extraction of proteins, as first described by Ku et al., was most effective for preserving proteins and successful post-gelation immunostaining<sup>29</sup>.

### 2.4.4 Homogenous antibody labeling can be achieved in thick tissue

Next, we wanted to demonstrate that post-gelation immunostaining could be applied to homogeneously label thick brain sections. 500 µm GFP expressing brain sections were processed with miriEx and subjected to 5 days of primary antibody and 3 days of secondary antibody staining shaking at 37 °C. We found that 1 day of antibody staining is necessary per

100  $\mu\text{m}$  of tissue. We decreased secondary staining time by using F(ab')<sub>2</sub> antibodies that are smaller and diffuse faster. We found strong immunostaining signal at the top, middle, and bottom of a 500  $\mu\text{m}$  deep (pre-expansion size) imaging volume, indicating homogeneous diffusion of antibody (**Fig. 2.5**).

## 2.5 Discussion

### 2.5.1 Summary

In summary, we described our experiences in optimizing expansion microscopy parameters for advanced histological analysis of brain sections. We adopted a post-gelation immunostaining expansion microscopy protocol for compatibility with thick brain sections. We found that concentration of acryl NHS-ester and gel monomer compositions were important variables that affect gel expansion factors and immunostaining results. A range of acryl NHS-ester concentrations needs to be tested for each protein target of interest to find the optimal concentration that preserves proteins, but avoids masking epitopes. Less dense monomer recipes, such as that of miriEx, result in higher signal and better antibody diffusion throughout the tissue. SDS/heat based extraction of proteins outperforms Proteinase K, trypsin, and collagenase for preservation of proteins for post-gelation immunostaining.

### 2.5.2 Expansion microscopy parameters need to be adapted for the priorities in each experiment

One of the important lessons we learned is that there is not a 'one size fits all' expansion microscopy protocol that can be applied universally. Instead, each parameter needs to be adapted to the goal of the experiment because many trade offs exist. For instance, if the top priority of the experiment is to probe the cytoskeleton in primary neuron cultures with the highest resolution, then a useful protocol may be one that favors the most expansion. An example could be pre-gelation immunostaining combined with proteinase K digestion and a large expansion factor gel recipe (0.01% bis-acrylamide U-ExM or X10). If the goal of the experiment is to immunolabel 1 mm brain sections for low abundance proteins, then a suitable protocol may involve higher acryl NHS-ester concentrations to try to preserve as much protein as possible combined with post-gelation immunostaining.

In addition, the gel's final expansion size is another important parameter to tune based on how bright the immunostaining is. The gel's size can be controlled by the salt concentration in the water. Pure water results in maximum linear expansion, while 1xPBS typically results in half the maximum linear expansion. Imaging gel-tissue hybrids that are only partially expanded has the advantage of brighter signal because of lower volumetric dilution compared to fully expanded samples. Subsequently, we recommend that the gel's final expansion size be tuned to the minimal acceptable imaging resolution in each experiment to maximize antibody signal.

### 2.5.3 Expanding the whole mouse brain

An important goal in the future is extending expansion microscopy protocols to the whole mouse brain. Doing so would enable probing of biomolecules and mapping neural structures in an intact brain at scalable resolutions. I have identified three obstacles for which solutions must be found. First, we need to find a method to diffuse acryl NHS-ester throughout the brain and modify proteins with acryl groups uniformly. NHS-esters hydrolyze quickly in aqueous solutions and label the outsides of the tissue better than the core<sup>28</sup>. Although not as much of a problem with thinner brain sections, the challenge with a whole brain is that the outer brain regions would be better labeled compared to the inner regions. Two potential solutions are delivering acryl NHS-esters systemically through blood vessels using perfusion, or employing a SWITCH strategy where the NHS-ester is inactivated during diffusion and then activated after it has permeated the brain<sup>36</sup>. The second obstacle is developing whole brain immunostaining methods. Again, the problem is that antibodies are slow to passively diffuse throughout an entire brain and often take weeks. Nanobodies are an exciting alternative as they passively diffuse an order of magnitude faster<sup>37</sup>. Another option is to use stochastic electrotransport to actively diffuse antibodies throughout the tissue quicker<sup>38</sup>. The third obstacle is to develop light microscopy strategies to accommodate imaging an expanded brain. ~4x expansion would yield mouse brain dimensions of ~4.8 cm x ~3.2 cm x ~2.4 cm, a challenging volume that will require an innovative approach to image.

## 2.6 Methods

### Expansion microscopy protocol

#### *miriEx*

Acrylic acid N-hydroxysuccinimide ester (AAX, Sigma A8060) was prepared by dissolving in N,N-Dimethylformamide to 125 mM. Tissue samples were incubated with 1-5 mM Aax in a MBS buffer (100 mM MES, 150 mM NaCl, pH 6) with 0.1% Triton X-100 shaking overnight at 4 °C. The next day, the samples were washed 3x with 1xTBS (Bio-rad 1706435) for 1 hour each to quench the reaction. The samples were then incubated in monomer solution (5.3% Sodium Acrylate, 4% Acrylamide, 0.1% Bis-Acrylamide, 0.5% VA-044, in 1x PBS) shaking overnight at 4 °C. The next day, the samples were gelled for 2.5 hours at 37 °C in a humidity chamber by cover slipping the sample surrounded by monomer solution in a gelling chamber. The gel-tissue hybrids were then carefully cut out and denatured at 70 °C overnight in denaturing buffer (200mM SDS dissolved in 1xTBS). The next day, denaturing buffer was washed out by incubating the samples 4x for 2 hours each in 0.1% PBST (1x PBS with 0.1% Triton X-100) shaking at 50 °C.

For the different gel recipes in **Fig. 2.4**, various monomer concentrations were substituted but the rest of the protocol remained the same.

#### *Calculating expansion factors*

1.5 mm circular tissue sections were cut from 100 µm fluorescent GFP brain sections to be used for testing 4 different gel recipes. DAPI was used to visualize the entire gel-tissue circle after histology processing was complete. Before imaging the samples (mounted in 1x PBS), an overview picture of DAPI was taken. The diameter of the expanded gel-tissue circle was measured and divided by the original diameter (1.5 mm) to calculate the expansion status in 1xPBS.

#### *Pro-ExM workflow*

Protocol was followed exactly as described by Asano and Gao et al<sup>39</sup>.

### Immunohistochemistry of gelled tissue sections



Blocking and permeabilization of gel-tissue hybrids were skipped as they were found to minimally decrease background. Gel-tissue samples were incubated with primary antibody diluted in 0.1% PBST w/ 0.02% azide (PBSTz) at 37 °C for 1 day per 100 µm thickness. Gel-tissue samples were washed 3x with 0.1% PBST w/ 0.02% Azide at RT for 1 hour each, and then incubated with secondary antibody diluted in 0.1% PBSTz at either 37°C for 1 day per µm thickness . All 4 gel recipes in **Figure 2.4** were immunostained with the exact same protocol (1:500 Biorad Sheep-antiGFP primary antibody for 1d at 37°C, 1:500 Jackson Immunoresearch Donkey-Anti-Sheep-Cy3 for 1d at 37°C). DAPI was also added during secondary antibody immunostaining.

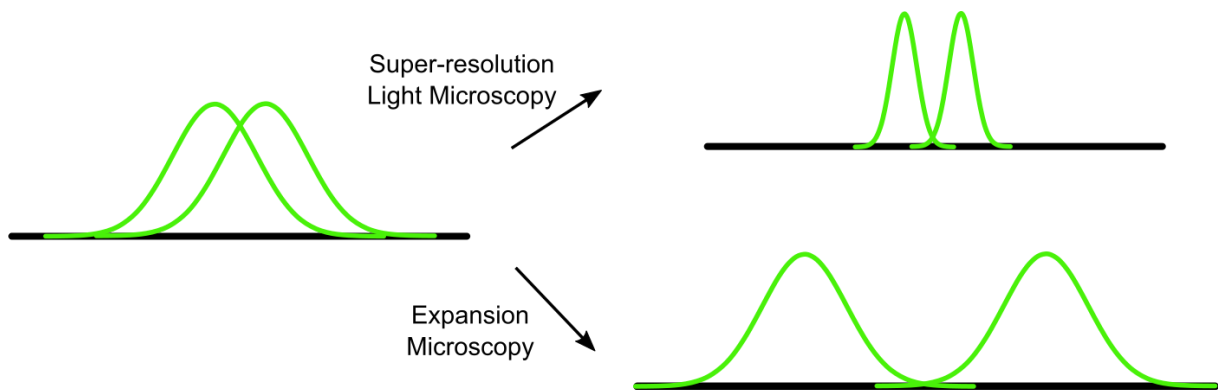
### **Gel expansion and fluorescence microscopy**

Gel-tissue hybrids were incubated in 1xPBS and then mounted in Poly-L-Lysine coated 6 cm dishes (Corning 354517) and submerged in 1xPBS to prevent drifting. All confocal LM was performed using an upright Zeiss LSM780. Water immersion objectives (10X/NA0.4 or 20x/NA1.0) were directly lowered into the PBS solution over the sample for imaging.

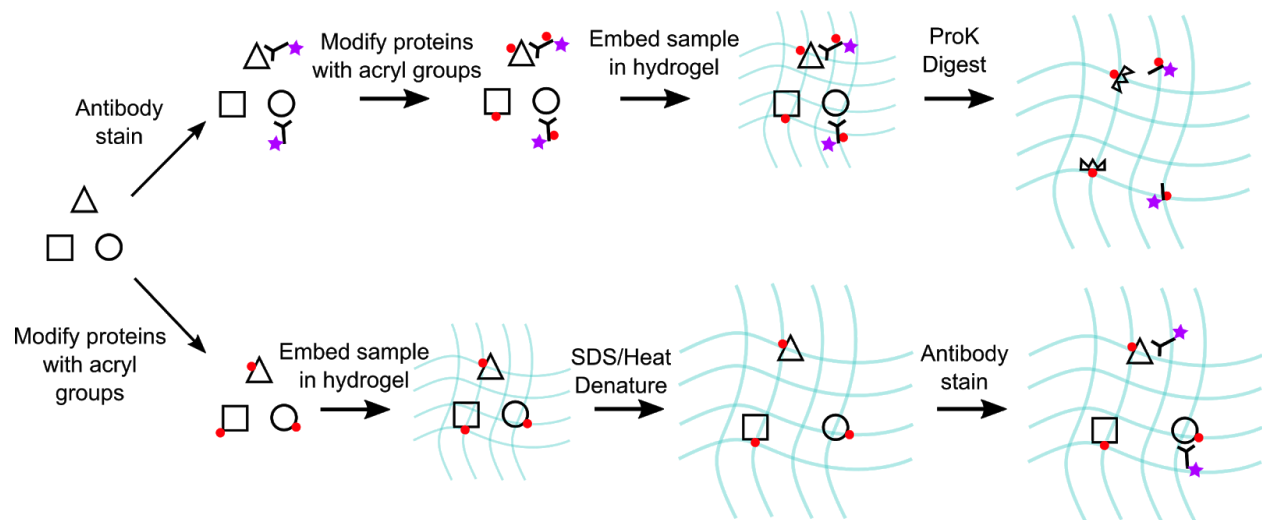
## **2.7 Acknowledgements**

I want to thank HJ Sekhon for insightful discussions about NHS-ester and gel monomer biochemistry. I also want to acknowledge Meg Harrington for helping generate the data for **Figure 2.4**. Finally, I want to thank Alexis Vega for helping with experiments aimed at testing alternative proteases.

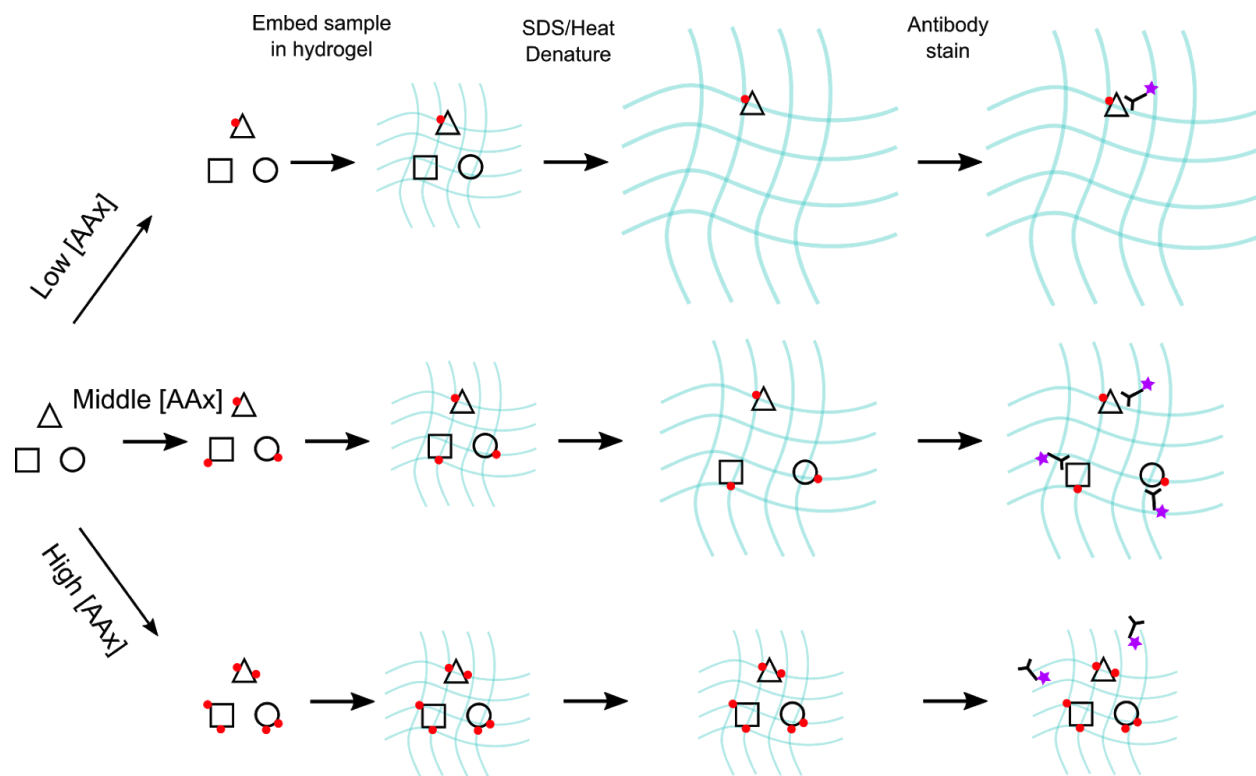
## 2.8 Figures



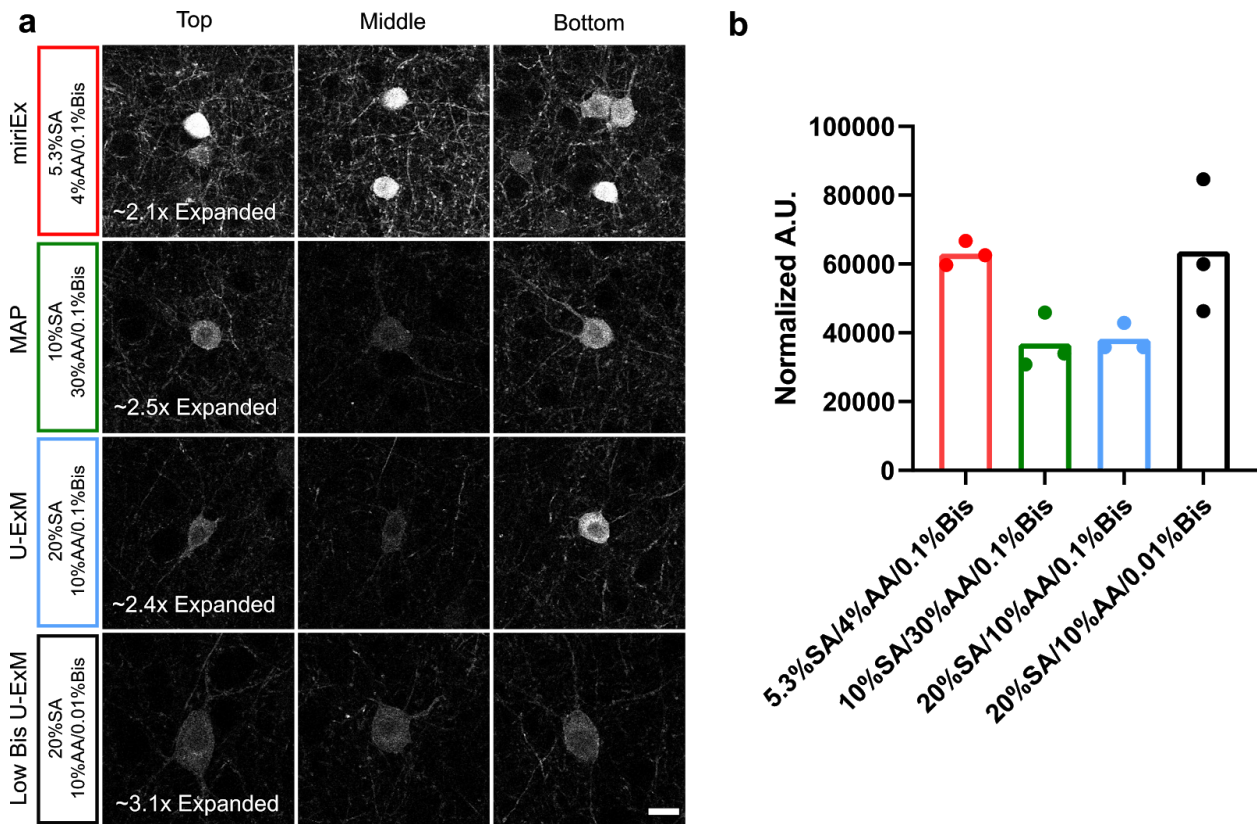
**Figure 2.1 Expansion microscopy achieves super-resolution through physical magnification of the sample.** (a) Example of two point spread functions (PSF) overlapping in space that cannot be resolved by diffraction limited light microscopy. Super-resolution light microscopy techniques, including STED, SIM, and STORM, achieve increased imaging resolution by sharpening the PSFs. Alternatively, expansion microscopy achieves increased imaging resolution by physically separating the PSFs through expansion of the sample. The former strategy requires specialized microscopes, specific sample preparations, and additional image processing. In contrast, the latter strategy is compatible with routine light imaging methods, such as confocal, widefield, and lightsheet microscopy.



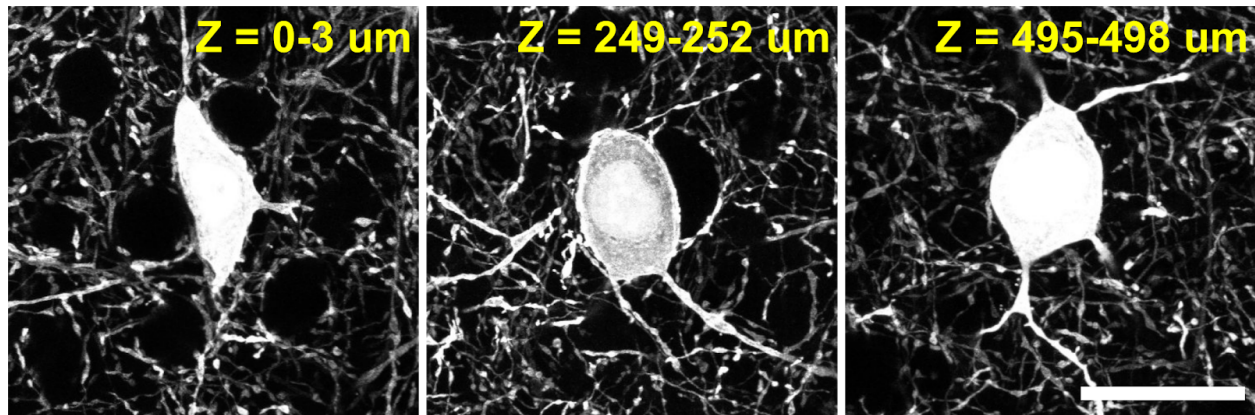
**Figure 2.2 Schematic of two complementary expansion microscopy workflows.** For the top (pre-gelation immunostaining) workflow, standard immunostaining is first performed on the sample (proteins represented as shapes, fluorescent dyes represented as purple stars). The sample is then modified with acryl groups (red dots) using NHS-ester biochemistry and subsequently embedded in an expandable hydrogel (cyan matrix). Proteinase-K is used to homogenize and loosen the tissue by digesting proteins and allowing the sample to expand. For the bottom (post-gelation immunostaining) workflow, proteins are first modified with acryl groups and then embedded in an expandable hydrogel. This gel-tissue hybrid is next subjected to a combination of SDS and high heat, denaturing proteins and allowing the sample to expand. Immunostaining can then be performed on the expanded sample.



**Figure 2.3 Effects of NHS-ester concentration on post-gelation immunostaining expansion microscopy protocols.** A schematic is shown for the effects of treating tissue samples with low, medium, and high concentration of acryl NHS-ester (Aax). The shapes (squares, triangles, circles) represent different proteins, and red dots represent acryl groups which serve as anchoring points in the hydrogel (blue wavy lines). Decreasing Aax concentrations results in larger expansion factors. Increasing Aax concentrations can result in masking of epitope sites. Optimized AAX concentrations are necessary for best immunostaining results.



**Figure 2.4 Effect of different monomer concentrations on immunostaining.** (a) 4 different monomer recipes of varying sodium acrylate (SA), acrylamide (AA), and bis-acrylamide (Bis) concentrations were used in a post-gelation immunostaining expansion microscopy protocol with 100  $\mu\text{m}$  fluorescent tissue sections ( $n=3$  tissue sections). Shown are representative images for GFP immunostaining at the top, middle, and bottom of the gel-tissue hybrid using the same laser power and display settings. The expansion factor of each gel was measured, averaged across 3 tissue sections, and displayed in the first column. (b) Quantification of average soma intensities after normalizing for varying gel expansion sizes ( $n=3$  samples per condition). Scale bar: 20  $\mu\text{m}$ .



**Figure 2.5 Post-gelation immunostaining expansion microscopy protocol enables homogenous immunostaining in thick tissue.** 500  $\mu\text{m}$  fluorescent tissue was processed with a post-gelation immunostaining expansion microscopy protocol probing for GFP. Shown are small 3  $\mu\text{m}$  maximum projections at the top, middle, and bottom of the tissue to demonstrate homogenous labeling. Scale bar: 25  $\mu\text{m}$ .

# Chapter 3. Development of New Brainbow Viral Tools

## 3.1 Introduction

### 3.1.1 The fluorescent protein revolution

In the 1960s, Osamu Shimomura collected many bioluminescent jellyfish, known as *Aequorea victoria*, from the Puget Sound in the Pacific ocean. He was interested in investigating what was causing the jellyfish's ring to glow with a faint green light. Shimomura was able to isolate two proteins that were responsible for the fluorescence: aequorin and GFP<sup>40</sup>. Aequorin binds calcium ions and glows blue, while GFP absorbs that blue light to emit green fluorescence. Two decades later, Douglas Prasher cloned GFP and his collaborator, Martin Chalfie, expressed it in *E. Coli* and *C. Elegans* for the first time<sup>41</sup>. Roger Tsien then further improved GFP by introducing various mutations to make it more useful for biologists. To commemorate the importance of their work, the 2008 Nobel Prize in Chemistry was awarded to the trio of Shimomura, Chalfie, and Tsien.

GFP has revolutionized biology by giving scientists the ability to see within a cell. It can be fused to proteins to track where they are, report transcription by being placed downstream of a promoter, and operate as a biosensor among many other uses. Serendipitously, EGFP was found to have several favorable properties that make it one of the best performing fluorescent proteins today: low aggregation, little cytotoxicity, protease resistance, photostability, etc<sup>42</sup>. Since the discovery of GFP, researchers have isolated fluorescent proteins from a wide range of organisms, such as coral, lancelets, and hydrozoa, and introduced a wide range of mutations to extend the fluorescence palette (**Table 3.1**).

### 3.1.2 Sparse fluorescent protein labeling in the brain

Following their introduction, many research groups quickly adopted fluorescent proteins to label neurons in the brain. Importantly, widespread GFP expression in neurons was shown to have minimal toxic effects<sup>19</sup>. Sparse labeling strategies were developed to track single neuron morphologies, neural migrations, and lineage trees<sup>43</sup>. For example, Feng and colleagues presented a suite of transgenic mice, each with a different subpopulation of neurons labeled with one of four GFP spectral variants<sup>44</sup>. Certain lines, such as Thy1-YFP-H, are still useful today as they have strong fluorescent protein expression in a sparse set of neurons and can be used to study how perturbations affect dendritic and axonal morphologies both *in situ* and *in vivo*<sup>45,46</sup>. They also are often employed in whole brain clearing and imaging studies because they have widespread and beautifully discernible axonal projections, dendritic arbors, and cell bodies<sup>24</sup>. To gain even sparser labeling in mice, three common strategies that have emerged are intrinsically sparse Cre-driver lines, tamoxifen inducible Cre-driver lines, and MADM<sup>47</sup>. The latter two are especially useful to label single neural clones for lineage tracing in neurodevelopmental studies.

### 3.1.3 Dense fluorescent protein labeling using Brainbow

Brainbow was introduced in 2007 by Livet, Lichtman, and Sanes and naturally built off the earlier Thy1-XFP work from the same group years ago. While Feng et al.'s work focused on generating transgenic mice that expressed *single* GFP variants in a sparse population of neurons, Livet et al. generated transgenic mice capable of *multiplicative* expression of GFP and dsRed spectral variants in a dense population of neurons. Through combinatorial expression of multiple fluorescent proteins, unique colors could be generated to label neighboring neurons in order to distinguish their respective morphologies. One of the key technological advances was employing fluorescent protein variants to take advantage of the spectral space. For example, Brainbow1.1 mice use YFP and CFP variants of GFP from jellyfish, and a RFP variant of dsRed from coral. This system solely relies on native fluorescence and being able to separate the three fluorescent proteins using multispectral light microscopy methods. It is incompatible with antibody amplification of fluorescence because GFP antibodies cross react with CFP and YFP.



A general principle of fluorescent protein immunostaining is that fluorescent proteins from the same organism, including their variants, share homologous protein sequences and tend to be recognized with the same antibody. In contrast, fluorescent proteins from different organisms represent different antigens and can be recognized with orthogonal antibodies. Cai and colleagues used this principle to improve upon the initial Brainbow design by replacing CFP, YFP, and RFP with fluorescent proteins from different organisms: EGFP, mOrange2, and mKate2<sup>48</sup>. The resulting Brainbow3.2 mice can have their fluorescent protein expression be amplified through immunohistochemistry to produce higher contrast images.

### 3.1.4 Analog coding in Brainbow follows the RGB color model

The initial Brainbow tools developed by Livet and colleagues generated color diversity for analog analyses. The design was heavily influenced by the RGB color model, in which red, green, and blue light are combinatorially combined to produce a wide array of colors (**Fig. 3.1a**). The RGB color model is especially prevalent in electronic devices, where each pixel in the screen is composed of a separate red, green, and blue light source. The combination of the three, along with intensity variations from each light source, gives rise to the colored pixels on our devices' screens. Analogously, Brainbow neurons can be thought of as pixels, where three distinct fluorescent proteins represent red, green, and blue light sources. The different intensities of these "RGB" fluorescent proteins can be combined to produce a broad array of colors across a neural population.

As a result, an essential part of this Brainbow design is generating distinct RGB fluorescent protein intensities to maximize color outcomes. The intensity variation of each fluorescent protein comes from both intrinsic expression variability between neurons and different amounts of fluorescent protein that are turned on for expression. To further explain the latter point, multiple copies of the transgene cassette expressing the fluorescent proteins were inserted into the mouse genome. As a result, different amounts of the same fluorescent protein can be expressed, ranging from 1 to  $n$  cassettes, to result in varying intensities. The question then becomes: how can we express different levels of fluorescent proteins *between neurons*?

Livet and colleagues address this question by elegantly employing site specific recombinases to randomize expression between neurons<sup>18</sup>. Cre is an example of such an enzyme that binds to

specific DNA sequences known as loxP sites. Depending on the orientation of the loxP sites, Cre can either excise or invert the DNA between them. Livet and colleagues turn on different amounts of fluorescent protein by flanking all three fluorescent proteins with three sets of orthogonal loxP sites and letting Cre randomly choose which pair to recombine. This results in stochastic expression of one fluorescent protein from each cassette, and when totaled across multiple cassettes, generates diverse color outcomes.

### 3.1.5 Digital coding in Brainbow

Analog Brainbow, where stochastic expression of three “RGB” fluorescent proteins generates different colors, is an important step forward, but is fundamentally limited by human perception. For example, although 8 bit RGB space can generate  $256^3 = 16,777,216$  colors, most of them cannot be perceived by humans. In fact, there are only 7,500 names in the *Color Names Dictionary* published by the US Department of Commerce<sup>49</sup>. This number is still probably an overestimation of possible Brainbow colors because there are less than 100 colors one can select from RGB color space that have enough perceptual contrast for an average person to easily distinguish<sup>50</sup> (**Fig. 3.1b**). As a result, visualizing analog Brainbow imaging datasets on RGB computer screens probably has an estimated ceiling of ~100 colors.

To circumvent this limitation, one can replace RGB analog with digital color coding in Brainbow. The idea is to transform the continuous variation of fluorescent protein intensities into discrete units to label each neuron (**Fig. 3.1c**). A recent study by Veling and colleagues described the simplest form of this as Bitbow<sup>51</sup>, where the ‘on’ or ‘off’ expression of a fluorescent protein is coded as a ‘1’ or ‘0’. Using 5 fluorescent proteins, they were able to generate  $2^5 - 1 = 31$  distinct barcodes to label each neuron. The advantage of digital Brainbow systems is that the color, i.e. barcoding, capacity scales exponentially with more fluorescent proteins and can easily surpass the 100 color limit of RGB analog Brainbow systems. Moreover, digital brainbow systems are more error tolerant because variations in fluorescent protein intensity along the neuron have less of an effect on the barcoding outcome. In contrast, variations in fluorescent protein intensity along a neuron can significantly change the RGB color of the neuron and cause problems with downstream analysis.

Although Bitbow is the simplest implementation of digital Brainbow analyses, one can arbitrarily divide the intensities of each fluorescent protein into discrete levels. The advantage would be scaling the barcoding capacity to  $n_{\text{levels}}^{\text{\#Fluorescent Proteins}} - 1$ . For example, three intensity levels (i.e. trits) and 5 fluorescent proteins would scale to  $3^5 - 1 = 242$  barcodes. The downside of more intensity levels is that it reduces error tolerance and may be difficult to establish proper biological ‘cutoffs’.

### 3.1.6 Brainbow color outcomes and balance

Two metrics that determine the usefulness of a Brainbow system are color outcomes and balance. Color outcomes is defined as the maximum number of color possibilities that can be achieved, while color balance is defined as the distribution of probabilities for achieving each of the color outcomes. Color diversity is a metric that factors in both color outcomes and balance. For example, a system that generates hundreds of color outcomes, but in which a few colors are much more probable than the rest (poor color balance) is not as useful and has low color diversity because many of the neurons will be labeled the same color. An ideal Brainbow system generates both high color outcomes and an optimized color balance, resulting from when the probability of each color outcome is equal.

### 3.1.7 Brainbow viral tools

#### 3.1.7.1 On/Off Brainbow AAVs

In 2013, Cai and Colleagues introduced the first set of Brainbow Adeno-associated virus (AAVs) tools. AAVs are an attractive way to deliver a small (<5 kB) genetic payload into neurons because they have a low inflammatory response, are relatively non toxic, and can maintain long periods of expression. Researchers have spatial control over expression by choosing the quantity of virus and injection site within the brain. The process of AAV infection leads to stochastic viral copy numbers in each neuron to generate color diversity, similar to how Brainbow mice have multiple fluorescent protein cassettes. Cai and colleagues developed an ‘on/off’ expression construct where two different fluorescent proteins were inverted and flanked by different loxP66 and loxP71 mutants<sup>52</sup>. Upon Cre recombination, one of three options was possible: expression of fluorescent protein #1, expression of fluorescent protein #2, or no expression, i.e. a dark state (**Fig. 3.2a**). Two AAVs were designed; one expressed membrane

mCherry and TFP, while the other one expressed membrane EYFP and tagBFP. These 4 fluorescent proteins were chosen to have separatable native fluorescence spectra and be antigenically distinct for antibody amplification.

### 3.7.1.2 'Always On' Brainbow AAVs

A similar, but subtly different Brainbow AAV design was introduced by Sakaguchi and colleagues as Tetbow<sup>53</sup>, and by Chan and colleagues as VAST<sup>54</sup> (vector-assisted spectral tracing). Their convergent approach was to use a cocktail of three different AAVs, each expressing a fluorescent protein under the control of a tetracycline responsive element (TRE) promoter (**Fig. 3.2b**). Tetbow uses mTurquoise2, EYFP, and TdTomato, while VAST uses mNeonGreen, mTurquoise2, and mRuby. The latter set is compatible with immunostaining while the former set is not since both mTurquoise2 and EYFP are GFP variants. Importantly, when a cell is infected with AAV, the expression of the virus is 'always on' since there is not a stochastic process to turn it off like with the 'on/off' design. The TRE promoter was chosen for two reasons. The first is to serve as a very strong promoter to boost expression of the fluorescent proteins, even at low virus copy numbers. Combined with a tetracycline transactivator (tTA) synthetic transcription factor, TRE promoters can drive higher expression levels than commonly used CAG promoters<sup>55</sup>. The second reason for using the tTA-TRE system is that labeling density can be controlled by introducing tTA as a separate virus and titrating its concentration. For example, low titers of tTA AAV can be injected to infect a sparse population of neurons and induce fluorescent protein expression<sup>54</sup>.

## 3.2 Results

### 3.2.1 'On/Off' design outperforms 'always on' design in both analog and digital color coding

To compare the two strategies, I modeled the effects of an 'on/off' versus 'always on' virus design on both analog and digital Brainbow color encoding. To simplify the problem, I only considered two fluorescent proteins: red and green. The 'on/off' design was composed of just one virus that had three expression outcomes of equal probability: red, green, or dark (**Fig. 3.2a**). The 'always on' virus design consisted of two individual red and green viruses that constitutively expressed their respective colors (**Fig. 3.1b**). I used a poisson distribution to

model the infection rate of a population of 1000 virtual neurons<sup>53</sup>. **Fig. 3.1c** shows each neuron's red-green color outcome given an average copy number of 1, 5, or 20 expression cassettes per cell per virus. As expected, the 'on/off' design produced strikingly different color distributions than 'always on', with the color distribution shifted towards the dark state. Higher copy numbers of virus per cell resulted in increased color outcomes, but also a more biased color distribution as most cells converged towards equal ratios of red and green. To quantify this loss in color diversity, I randomly chose 100 pairs of neurons from the simulation and plotted their red-green color distance as a function of average copy number/cell/virus (**Fig. 3.2d**). The resulting plot demonstrates that the 'on/off' design leads to better color separation on average compared with the 'always on' design.

To compare the two designs for digital instead of analog Brainbow coding, I plotted the fraction of dark neurons for a single color as a function of average copy number/cell/virus (**Fig. 3.2e**). I will assume a 'trit' digital coding scheme ('0', '1', or '2') is feasible to use with Brainbow AAVs. Trit coding was chosen because it seemed like a good match empirically for the distribution of fluorescent protein intensities (dim and bright) generated by Brainbow AAVs. Bit coding has a smaller barcode capacity and ignores useful fluorescent intensity variations, while quaternary or higher coding may have difficulty establishing biological 'cut-offs' for fluorescence intensity. For ideal 'trit' barcode distributions,  $\frac{1}{3}$  of the neurons would be dark ('0'),  $\frac{1}{3}$  would be dim ('1'), and  $\frac{1}{3}$  would be bright ('2'). The cutoff point between dim and bright is arbitrary and can be set so that the empirical distribution matches the ideal probabilities. Therefore, a more important metric is the fraction of 'dark' neurons. For example, in a poisson distribution of 3 average viral copies/cell in the 'on/off' design,  $\sim\frac{1}{3}$  of the neurons are dark,  $\sim\frac{1}{3}$  of the neurons express 1 copy of the virus, and  $\sim\frac{1}{3}$  of the neurons express 2 or more copies of the virus. As a result, neurons expressing 1 copy/cell would be coded as '1' and neurons expressing 2 or more copies/cell would be coded as '2'. I observed that for the 'always on' design, an average copy number of  $\sim 1$  would result in  $\sim 33\%$  of the neurons being dark. In contrast, for the 'on off' design, an average copy number of  $\sim 3$  would yield the same result. Again, the 'on/off' design outperforms the 'always on' design because of a higher tolerance for different virus copy numbers. To put into practical terms, there is a small window of average viral copy numbers (0-1.5) that would be difficult to experimentally optimize in the 'always on' design, assuming up to 10% error tolerance for the dark fraction. In contrast, the 'on/off' design is experimentally easier with a larger window of average viral copy numbers (2.5-4.5) that would be useful for trit coding.

## 3.2.2 Extending the 'on/off' Brainbow AAV toolbox

### 3.2.2.1 Promoter choice

Based on the modeling described above, we decided to build novel Brainbow AAV constructs using the 'on/off' design because of better analog and digital color coding capacity. We chose to use the pan-neuronal human synapsin 1 promoter because it was smaller than the EF1alpha promoter used in Cai et al., and would therefore accommodate larger fluorescent proteins described below. The downside is that we lost flexibility to label non-neuronal populations, such as glia. In addition, we chose hSyn1 over TRE to avoid toxicity concerns from overexpression in the tTA-TRE system<sup>53</sup>. Both Sakaguchi et al. and Chan et al. reported potential toxicity that was ameliorated by carefully controlling the tTA dosage and length of expression. We were willing to sacrifice brighter native fluorescence for healthier neurons because we planned to rely on immunostaining to amplify the signal of our fluorescent proteins.

### 3.2.2.2 Twelve antigenically distinct fluorescent proteins

We identified 12 fluorescent proteins that were distinct antigens for immunostaining: mCherry, TFP, EYFP, tagBFP, mKO2, mNeonGreen, phiYFP, smFLAG, smMYC, smHA, smV5, and smOllas. The latter 5 are known as spaghetti monster fluorescent proteins and were introduced by Viswanathan and colleagues<sup>56</sup>. These fluorescent proteins are composed of multiple epitope tags that were strategically inserted in a super-folder GFP scaffold, which acts as a vessel for displaying these epitope tags to be visualized with immunostaining. To preserve the spectral bandwidth of native fluorescence, we mutated the chromophores of mKO2, mNeonGreen, and phiYFP to render them dark. The 5 spaghetti monster fluorescent proteins were previously mutated to be dark as well<sup>56</sup>. Our intention was to maximize flexibility so that the native fluorescence of the original four fluorescent proteins (mCherry, TFP, EYFP, and tagBFP) can be sufficient if additional fluorescent proteins are not required for the experiment. We created new chicken, rabbit, and rat antibodies against the 7 non spaghetti monster fluorescent proteins, and validated they are sensitive, specific, and have minimal cross reactivity (data not shown). Combined with commercial epitope tag antibodies, we were able to perform immunostaining for all 12 fluorescent proteins. Although it would seem that our sheep anti-EGFP antibody would cross react with the spaghetti monster GFP scaffolds, preliminary data indicates that they can actually be distinguished separately *in situ*. The GFP antibody may be recognizing the spaghetti

monster scaffold with lower efficiency due to epitope masking, allowing us to separate EYFP from spaghetti monster expression.

### 3.2.2.3 Simultaneous membrane and cytoplasmic expression

The original 'on/off' Brainbow AAVs use farnesylation tags for all 4 fluorescent proteins (mCherry, TFP, tagBFP, EYFP) to target them to the membrane. Membrane targeted fluorescence has several advantages over standard cytoplasmic expression. First, membrane fluorescent proteins better label fine subcellular features, such as dendritic spines and axonal boutons, because cytoplasmic fluorescent proteins may not diffuse as well into these small compartments<sup>48</sup>. Second, they result in less saturated and more balanced imaging. Cytoplasmic imaging tends to result in intrinsically large intensity ranges, where somas and large dendrites accumulate more fluorescent protein and saturate faster than the rest of the neuron. Consequently, this large dynamic range is difficult to capture using a single laser power during imaging and intensity information can be lost. Membrane expression results in more balanced expression between the different neural compartments because the amount of plasma membrane is more consistent across the neuron. The disadvantage of membrane versus cytoplasmic fluorescent protein expression is that membrane expression is dimmer, somas can be difficult to discern, and segmentation of neural membranes is more difficult computationally. To combine the advantages of both membrane and cytoplasmic expression, we employed a novel design that allows fluorescent proteins to be simultaneously targeted to both the plasma membrane and cytoplasm. We added this new membrane and cytoplasmic tag to the 12 fluorescent proteins previously described, inserted them into the 'on/off' design, and created 6 new viral vectors (**Fig. 3.3**).

### 3.2.2.4 Systemic delivery using PHP.eB capsid AAVs

The original 'on/off' Brainbow AAVs from Cai et al. were packaged in serotype 2/9 and meant to be used for stereotaxic injections in Cre driver lines to label cell type specific populations. Although this gives us spatial and molecular control over Brainbow fluorescent protein expression, there are several challenges. First, the viral copy number per cell is hard to control, as neurons further away from the injection site are infected less frequently. In fact, neurons close to the injection site tend to become saturated with viruses (up to 100 copies) and appear white, which limits color diversity as shown in **Fig. 3.2c**. Neurons at the edge of viral diffusion are sparsely infected, which also results in simple color outcomes and limited color diversity.

Somewhere in between are neurons infected with the optimal amount of virus that results in maximal color diversity. Subsequently, targeting small brain regions with the optimal amount of viral copy numbers is difficult because the injection would have to be 'off center' to transduce optimal infectivity in the region of interest. Second, stereotaxic injections of AAVs are targeted by definition and cannot be used as a global reporter for Cre activity in the way that transgenic Brainbow mice can.

To overcome these challenges, we packaged our 6 new Brainbow AAVs in PHP.eB serotype<sup>54</sup> so that they can cross the blood brain barrier and be delivered systemically. Using a retro-orbital injection of  $10^{11}$  to  $10^{12}$  genomes per virus in different Cre lines gave us widespread infection across the brain in cell type specific populations (**Fig 3.4**). Importantly, systemic delivery is thought to yield a poisson distribution of virus infection<sup>57</sup>, giving us tighter control of viral copy number that can be tuned by adjusting viral titers.

### 3.2.3 Simulating Brainbow AAV barcoding in mouse brains

I modeled the density of neuron labeling that could be achieved with our new 'on/off' Brainbow AAVs to gauge the potential usefulness of the tool. To simplify the problem, I ignored long range projection neurons and assumed that all neurons project locally in a 0.5 mm radius sphere that contains all of their axon and dendritic arbors. Using density estimates from literature, I also assumed that a  $1 \text{ mm}^3$  volume of brain tissue contains  $\sim 100,000$  neurons<sup>58</sup>. Next, I asked, *how many neurons potentially overlap arbors and need to be uniquely labeled for unambiguous reconstruction?* Against intuition, not all the neurons in the brain need to be uniquely labeled because spatially distant neurons do not overlap and can be ignored. Instead, the answer can be modeled by aligning neuron spheres side by side. Only neurons within a 1 mm radius sphere have the potential to overlap neurites (**Fig. 3.5a**). Consequently, there are  $\sim 400,000$  neurons in this spherical volume that we must uniquely label to unambiguously reconstruct their individual morphologies. If we decrease the labeling density and are only interested in 1% of the neurons, then there are only  $\sim 4,000$  neurons that must be distinguished and uniquely labeled.

Next, I assumed that we have the capacity to densely label all neurons through systemic injection of a 6 AAV cocktail that expresses 12 fluorescent proteins. If we use a trit color barcoding scheme, then we have  $3^{12} - 1 = 531,440$  unique barcodes. Assuming an even



probability distribution of barcodes, I simulated the fraction of neurons that would have unique color barcodes that do not collide, and repeated the simulation using 4 and 8 fluorescent protein trit barcoding schemes. The results encouragingly suggest that a 12 fluorescent protein trit barcoding scheme is sufficient to uniquely label ~70% of *all* neurons in a 1 mm sphere. If we decrease the labeling density to ~10%, such as by using a sparser Cre-driver line, we can uniquely label over 90% of all the neurons. This is a massive improvement over the original 'on/off' Brainbow AAV system (4 fluorescent protein trit system), in which only 20% of neurons are uniquely barcoded at 0.1% labeling density and almost no neurons are uniquely barcoded at 1% or greater labeling densities.

## 3.3 Discussion

### 3.3.1 Summary

In this chapter I described the rationale and design of new Brainbow AAV vectors that can be used for neurocartography. I demonstrated that an 'on/off' design outperforms an 'always on' design through modeling viral infectivity and color expression. Consequently, we inserted 12 antigenically unique fluorescent proteins into 6 'on/off' Cre dependent vectors for stochastic membrane and cytoplasmic expression in neurons. We packaged AAVs in PHP.eB serotype for systemic infection across the brain using intravenous delivery. Compared to the Brainbow AAVs described by Cai et al., our new system has three orders of magnitude more color outcomes from using more fluorescent proteins, and a better balanced color distribution through systemic delivery of virus. By simulating color barcode collisions, we estimate that we should be able to easily study single neuron morphologies at 10% density, represented by labeling all cortical parvalbumin interneurons using PV-Cre mice. We even have the potential to reach 100% density by labeling *all* neurons using Syn1-Cre mice.

### 3.3.2 Developing multiplexed readout strategies

One challenge with using a 12 fluorescent protein Brainbow system will be optimizing multiplexed readout strategies. Limited visible light bandwidth and number of host animal species restricts standard immunostaining methods to detecting 3-4 targets at a time. As a result, highly multiplexed readout methods will have to be employed. For example, we can explore DNA barcoded antibody readout methods, such as CODEX<sup>59</sup>, SABER<sup>60</sup>, and PRISM<sup>61</sup>.

Alternatively, we can try methods based on multiple rounds of immunostaining, imaging, and antibody stripping, such as SHIELD<sup>62</sup>, array tomography<sup>63</sup>, and miriEx<sup>30</sup>. The last method (miriEx) is a multi-round immunostaining expansion microscopy protocol that I developed and further describe in Chapter 4.

### 3.3.3 Future fluorescent proteins to pursue

The 6 new AAVs introduced in this chapter use fluorescent proteins originating from 7 distinct marine animals. A quick literature search reveals that additional fluorescent proteins from unique organisms have been discovered and can be incorporated in future Brainbow AAVs. I propose that the next 4 fluorescent proteins to add be dendra2, mKeima, mPapaya, and turboGFP, as they have been the most validated in literature and share the least amount of sequence homology<sup>42,64,65</sup> (**Table 1-2**). Dendra2 is a commonly used photoconvertible fluorescent protein that changes its emission spectra upon induction with violet light<sup>66</sup>. Although this property of Dendra2 is potentially interesting, we would mutate all 4 new fluorescent protein's chromophores to render them dark, and only use them as antigens for immunostaining. Commercial antibodies already exist for all four fluorescent proteins, but more will have to be produced in different host animals to increase flexibility during multiplexed immunostaining. All antibodies used need to be rigorously validated for minimal cross reactivity across the panel of 16 fluorescent proteins.

Extending the simulations in **Fig. 3.5**, approximately  $10^9$  color barcodes are needed to uniquely label the entire mouse brain and its 80 million neurons. This means that 21 fluorescent proteins are needed in a trit coding Brainbow system to generate a sufficient amount of barcodes. We can reach this number by adding a combination of unused fluorescent proteins (mAzamiGreen, mEos, mMiCy, mKikGR) and chemical tags (SNAP, HALO, and CLIP). While the latter are not fluorescent proteins, they can be used in the same way by detecting their expression using dye conjugated ligands<sup>53</sup>.

### 3.3.4 Subcellular labeling strategy to generate more barcodes

Instead of adding more antigenically unique fluorescent proteins to increase the barcoding capacity, I propose an additional strategy based on labeling distinct subcellular organelles prevalent in neurites. Mitochondria are widespread throughout a neuron's dendrites and axons

to provide energy to synapses and represent an attractive labeling target<sup>67,68</sup>. We could stochastically label mitochondria with the same set of fluorescent proteins used to label the plasma membrane. With the help of expansion microscopy, we should be able to spatially distinguish mitochondria from the plasma membrane, even in the smallest of neurites<sup>69</sup>. As a result, labeling an additional subcellular compartment would further generate color barcodes. Two sets of AAVs, both stochastically expressing the same 4 fluorescent proteins (mCherry, TFP, tagBFP, EYFP), but separately targeted to the plasma membrane and mitochondria, would result in  $(3^4-1) \times (3^4) = 6320$  barcodes (assuming that membrane expression is necessary to localize mitochondria accurately). Although the barcodes generated would be slightly less than an 8 fluorescent protein system, the advantage is that only 4 fluorescent proteins need to be detected, which will simplify the amount of multiplexed immunostaining and imaging required.

### 3.3.5 Developing a new image processing pipeline

Developing the genetic tools to properly label neurons with unique barcodes is an important step, but also requires new approaches in computational processing to maximize utility by the neuroscience community. The current workflow involves only a few basic steps. The channels in a raw image are first normalized to one another using a histogram matching algorithm, and also corrected for laser misalignments and chromatic aberration. Then, nTracer, a Fiji/ImageJ plugin, is used to load the 3 channel dataset in RGB colors and trace each neuron. Below, I will describe additional processing steps that can be performed on the data.

First, denoising algorithms can be applied to the imaging data to increase the signal to noise. Two promising ones that are worth future investigation are BM4D and CANDLE. BM4D uses grouping and collaborative filtering in a transformed domain to denoise 3D image datasets<sup>70</sup>. CANDLE is designed for low signal to noise conditions when imaging deep in a sample, and uses a non-local means filter to adapt the amount of smoothing across the dataset<sup>71</sup>. Denoising algorithms would be helpful in extracting signal from dim neurons in the dataset. Next, we can borrow and apply imputation algorithms from bioinformatics that are commonly used. Imputation is the process of replacing missing data points with estimated values, and is used in bioinformatics to fill in missing sequencing results to extract more biological insights. The same process would be useful for Brainbow imaging data because there can be gaps in membrane fluorescent protein expression from a variety of factors. Assuming we could group all the pixels

associated with a neuron's morphology, imputation would allow us to fill in the gaps along the neuron. Following imputation, the image needs to be thresholded into discrete levels for digital barcoding. For trit coding, the image histogram could be binned into three levels representing '0', '1', and '2' expression levels for each fluorescent protein channel. This thresholding step would further serve to reduce intensity fluctuations along the neuron. Instead of thresholding to a predetermined number of levels, an interesting idea would be to calculate the natural number of bins each channel's histogram can be divided into and let that adaptively determine the digital barcoding scheme. For example, channel 1 may separate better into 3 bins while channel 2 separates better into 4 bins.

Finally, for tracing purposes, we need a way to transform the digital barcodes into RGB space for visualization on a computer monitor. In other words, the 12 dimensions of fluorescent protein expression needs to be reduced to 3 RGB dimensions for computer display. I propose to use well described dimensionality reduction algorithms like PCA or UMAP to accomplish this. For example, we could find the first three principal components with PCA, map each pixel onto this space, and display it by treating the three principal components as RGB. UMAP could also be similarly used, and has the added advantage of spreading clusters of pixels further apart within RGB space because it is nonlinear. An alternative method would be to randomly choose 3 of the 12 fluorescent protein channels for display during tracing. When a color ambiguity is encountered, switching to another random subset of 3 channels will potentially yield different colors that allow the ambiguity to be resolved. This method would allow us to reshuffle the RGB values of neurons to overcome the limited number of colors that can be distinguished.

### 3.3.6 Brainbow AAVs offer several advantages over transgenic mice

A viral Brainbow approach has several disadvantages and advantages compared to a transgenic mice approach. Brainbow mice have the advantage of being more useful for lineage tracing and neurodevelopmental studies that require labeling at embryonic time points. AAVs cannot be used because the earliest time point of injection and expression is postnatal. However, future capsids that enable delivery of AAVs across the placenta to infect embryonic brains would erase this advantage. Brainbow mice also have the advantage of being a brain wide reporter for Cre expression compared to stereotaxic injections of Brainbow AAV that are spatially restricted. However, packaging our new Brainbow AAVs with PHP.eB capsids for

systemic delivery allows us to achieve efficient brainwide expression to match that seen in Brainbow mice. Brainbow AAVs offer more flexibility and control over the number and color diversity of fluorescent protein expression compared to transgenic mice. For example, we can choose which Brainbow AAVs to include in the virus cocktail depending on the fluorescent proteins desired, and adjust virus titers to control the copy numbers of virus expressing in each cell. Furthermore, a 12 fluorescent protein system would be difficult to implement in transgenic mice. The majority of transgenes are under 20 kB in size, but a 12 fluorescent protein cassette would be larger and have to be split up and inserted through multiple steps.

### 3.3.7 Extending Brainbow to neurons beyond the brain

Chan and colleagues showed that AAV.PHP.eB capsids transduce not only cells in the brain, but also in the spinal cord as well<sup>54</sup>. Consequently, our newly developed Brainbow PHP.eB AAVs can be injected intravenously to map the morphologies of densely labeled spinal cord neurons. If precise labeling of a specific vertebral level is not required, systemic injection offers a much easier route of delivery compared to intraspinal injection. Chan et al. also developed a AAV-PHP.S capsid to efficiently transduce peripheral nervous system cells including dorsal root ganglion, cardiac, and enteric neurons using intravenous injection. Furthermore, the same group recently developed a multiplexed virus capsid evolution platform to discover even more capsid variants with distinct tropisms for neurons, astrocytes, and endothelial cells<sup>72</sup>. Future work can be directed towards packaging Brainbow AAVs in PHP.S and other novel capsids to label these specific cell populations systemically.

## 3.4 Methods

### **Mouse Lines**

All experiments were carried out in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Our study protocol was reviewed and approved by the University of Michigan Institutional Animal Care & Use Committee. The transgenic mice used in this study were: PV-Cre (Jackson stock no. 008069) and VGAT-Cre (Jackson stock no. 016962). P28 to P56 male and female mice were used for virus injections.

### **Brainbow AAV injections**

Brainbow AAV vectors from Cai et al. were obtained from Addgene as AAV-2/9 or from the University of Michigan vector core as AAV-PHP.eB. Newly generated Brainbow AAVs (dual membrane and cytoplasmic expression) were obtained from the University of Michigan vector core as AAV-PHP.eB. For stereotaxic injections, transgenic mice were anesthetized continuously with isoflurane and mounted on a stereotaxic frame. Brainbow AAVs were mixed together to reach a concentration of 1E12 gc/mL individually, and injected at 100nL/min at specific AP, ML, and DV coordinates depending on the target. Afterward, the pipette was left in place for 5 min. for the virus to diffuse, before slowly retracting out of the brain. We waited 3-4 weeks for virus expression before perfusing the animals. For systemic delivery of PHP.eB AAVs, 1E11-1E12 gc for each virus was mixed and injected into the retro-orbital sinus. We waited 3-4 weeks for virus expression before perfusing the animals.

### **Mouse perfusion and tissue sectioning**

Mice were anesthetized with tribromoethanol and perfused transcardially with ice-cold 1xPBS, followed by 4% PFA in 1xPBS. The brains were dissected and postfixed in 4% PFA in 1xPBS overnight shaking at 4 °C. The next day, brains were washed in 1xPBS before slicing on a vibratome (Leica VT1000s) at 100 µm or 200 µm thickness.

### **Immunohistochemistry**

For immunostaining, tissue sections were first blocked and permeabilized in StartingBlock (Thermo 37538) with 1% Triton X-100 shaking overnight at 4 °C. Tissues were washed the next day 3x with 1xPBS for 1 hour each. Tissues were then incubated with primary antibodies diluted in 0.5% PBST for 3 days at 4 °C. After washing 3x with 0.5% PBST for 1 hour each at RT, tissue sections were then incubated with secondary antibody diluted in 0.5% PBST for 2 days at 4 °C. Tissues were then mounted with Vectashield (Vectorlabs H-1000) and imaged.

### **Image processing**

Specific image processing steps for each imaging dataset are listed in **Table S1**, and are briefly explained below.

Stitching of multi-tile datasets was performed using the BigStitcher<sup>41</sup> ImageJ/Fiji plugin. The dataset was loaded and converted to HDF5 format, and the tiles were arranged in the order they were imaged with 10% overlap. The Stitching Wizard was used to calculate pairwise shifts

using phase correlation, verify links, and undergo global optimization. Affine refinement was then performed with a high threshold. The resulting interest points were used for non-rigid refinement during advanced image fusion.

Chromatic aberration was corrected using the Detection of Molecules (DoM) ImageJ/Fiji plugin. To calibrate the DoM plugin, 0.5  $\mu\text{m}$  TetraSpeck fluorescent beads were mounted on a slide and imaged on a Zeiss LSM780 confocal microscope. More specifically, we calibrated the orange channel (540-600 nm) excited with a 543 laser and far-red channel (630-700 nm) excited with a 633 laser to the green channel (480-540 nm) excited with a 488 laser. Calibration was performed for both the 10x and 20x objectives.

Histogram matching was done to normalize intensity between z-slices in image stacks using the nTracer Align-Master ImageJ/Fiji plugin. A high signal to noise z-slice (usually at the top or bottom of the stack) was chosen as the reference slice for which the rest of the stack was normalized to. All three Brainbow FP channels were histogram matched to the same reference slice.

nTracer, an ImageJ/Fiji plugin, was used to trace somas, dendrites, and axons of Brainbow labeled neurons.

### **Comparing ‘on/off’ and ‘always on’ virus designs through simulations**

A poisson distribution was used to model the copy number/cell/virus of 1000 virtual neurons given a mean viral infectivity. For the ‘on/off’ model, equal probabilities of red, green, or dark expression were assumed and the outcome was randomly chosen for each infection. Red green color distance was calculated using the euclidean distance of a random pair of neurons chosen from the total population (repeated 100 times). Fraction dark was calculated as the fraction of neurons that lacked single color expression.

### **Calculating fraction of unique barcodes**

Assuming an even probability of all color barcodes,  $n$  neurons, and  $b$  barcodes, the collision number was calculated using the equation:

$$collisions = n - b + b\left(\frac{b-1}{b}\right)^n$$

The fraction of unique color barcodes was calculated by using the equation:

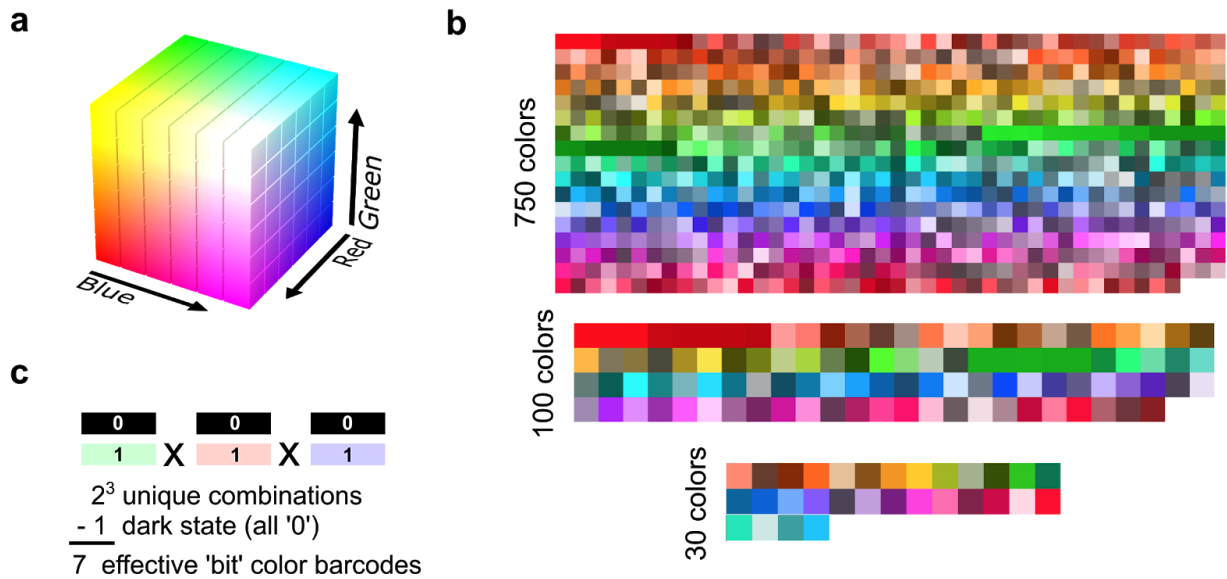
$$f = 1 - collisions/neurons$$

## 3.5 Acknowledgements

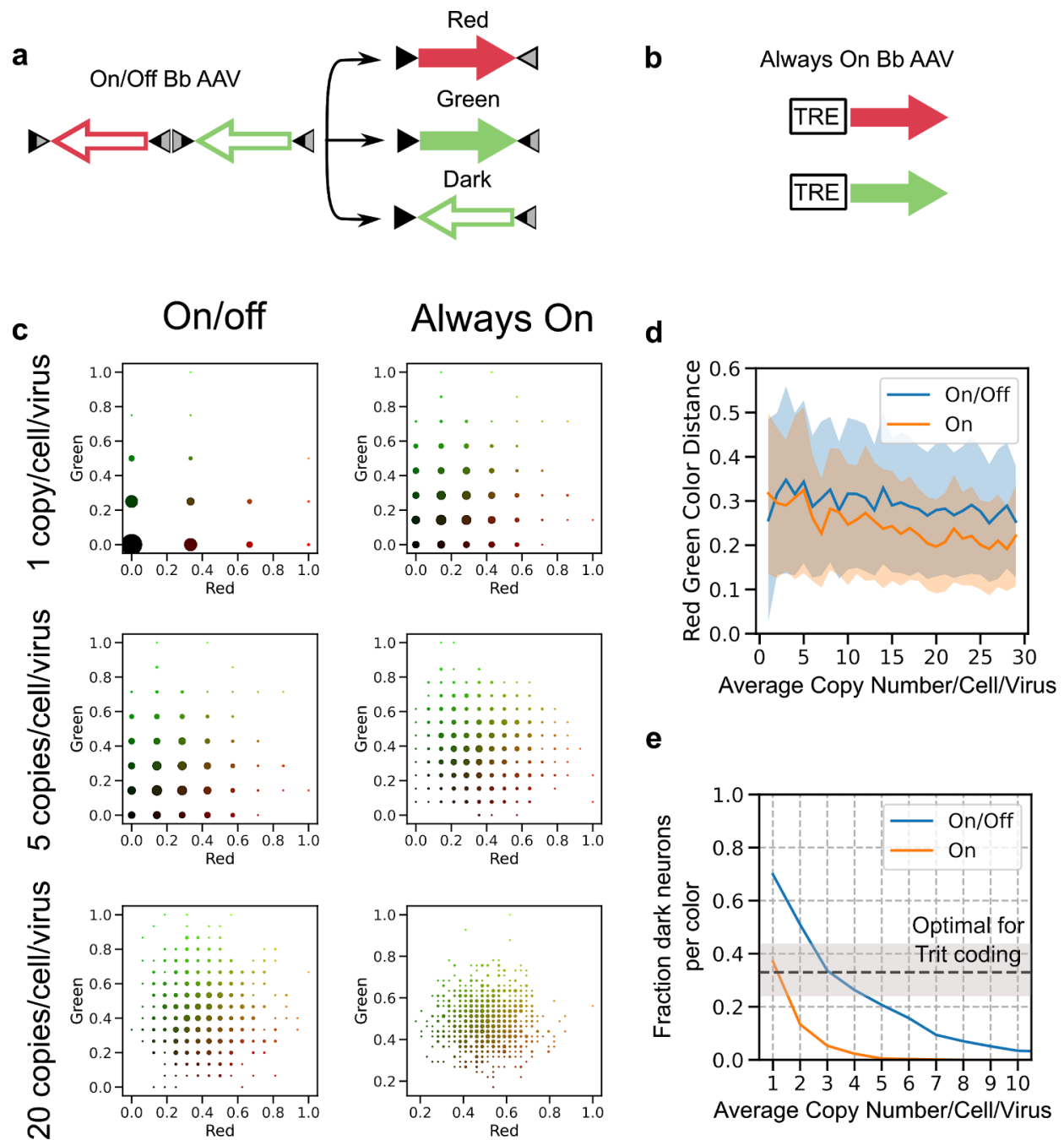
I want to thank HJ Sekhon for developing the dual membrane and cytoplasmic design. HJ Sekhon and Yimeng Zhao purified fluorescent proteins and created new rabbit, rat, and chicken antibodies against mKO2, neonGreen, tagYFP, and phiYFP. I want to acknowledge Meg Harrington for helping clone and create the new Brainbow AAVs. In addition, Meg Harrington helped create the matrix of fluorescent protein sequence similarities. Finally, I want to thank the University of Michigan Vector core for helping package AAV.PHP.eB viruses.



## 3.6 Figures

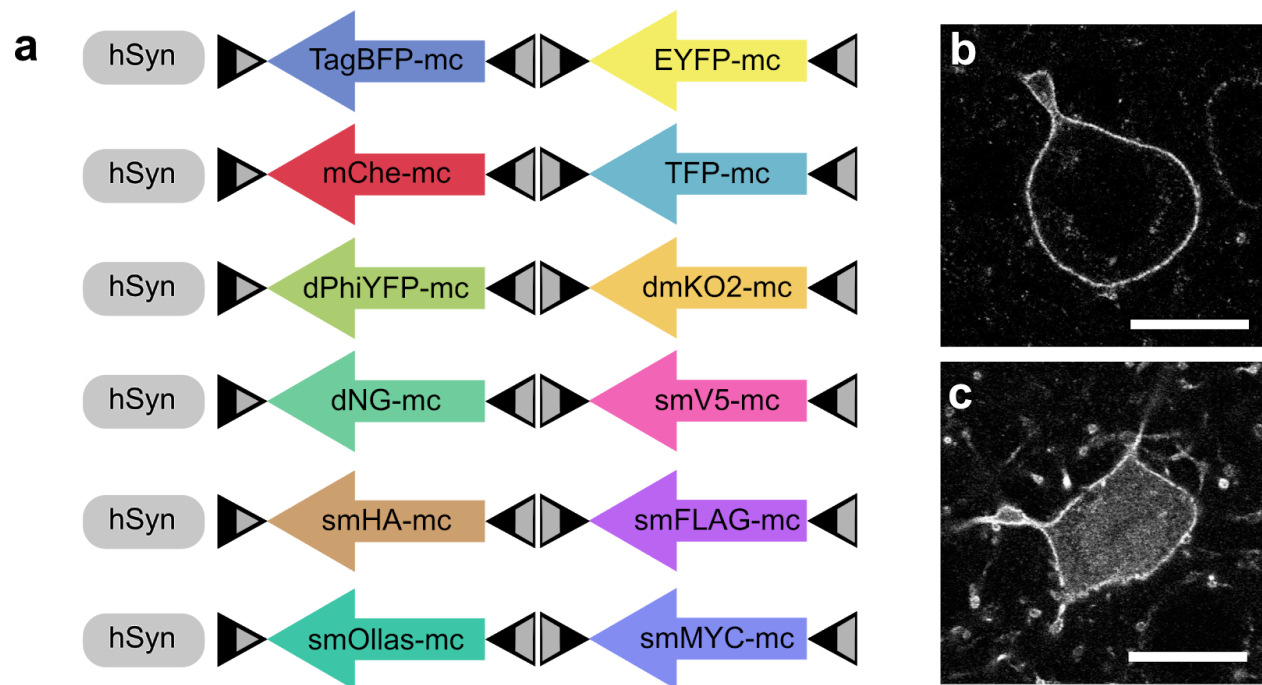


**Figure 3.1 Analog versus digital generation of Brainbow color barcodes.** (a) Standard RGB color model where combining red, green, and blue intensities (0-255) generates distinct colors. (b) Example color palettes containing 750, 100, and 30 colors that were generated with an emphasis on being visually distinguishable (<http://phrogz.net/css/distinct-colors.html>). (c) Example of digital Brainbow encoding where each fluorescent protein channel is divided into discrete bins. Color barcodes are generated by multiplying the number of discrete states across each channel.

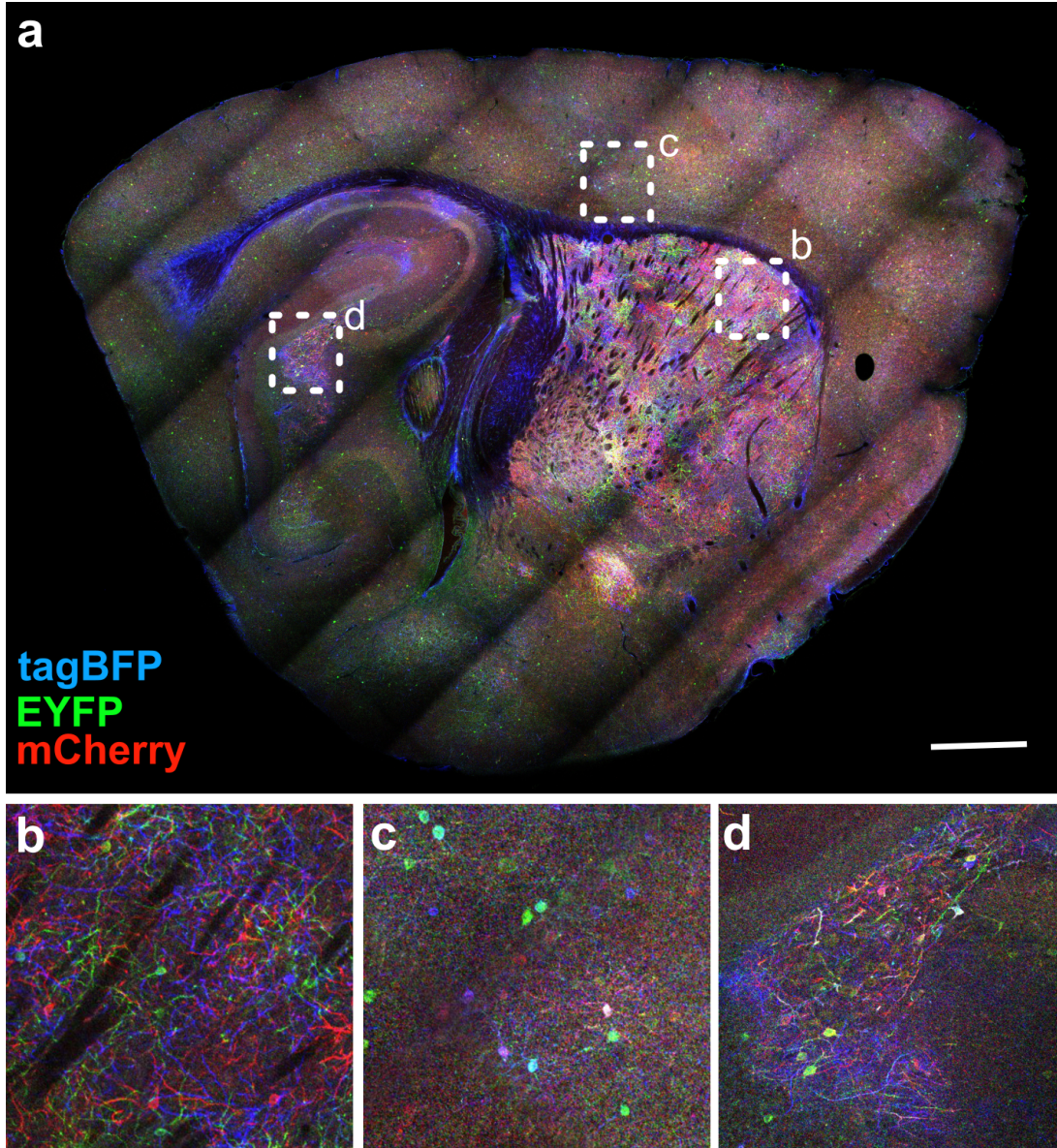


**Figure 3.2 Modeling ‘on/off’ versus ‘always on’ viral designs.** (a) ‘On/off’ design (from Cai et al.) where a single virus containing red and green fluorescent proteins can stochastically result in three states upon Cre recombination: red, green, or dark. (b) ‘Always on’ design where an infected copy of the virus always results in expression of the fluorescent protein. Two individual red and green viruses are needed to create color diversity. (c) Simulation of the distribution of all color outcomes given 1,5, or 20 copies/cell/virus in 1000 virtual neurons. The size of each dot correlates with the number of neurons expressing that color outcome. (d) Average red-green euclidean color distance plotted as a function of average copy number/cell/virus with standard deviation indicated by the shaded area. (e) For just 1 color (i.e. red), the fraction of dark neurons per color was plotted as a function of average copy number/cell/virus to compare the two viral

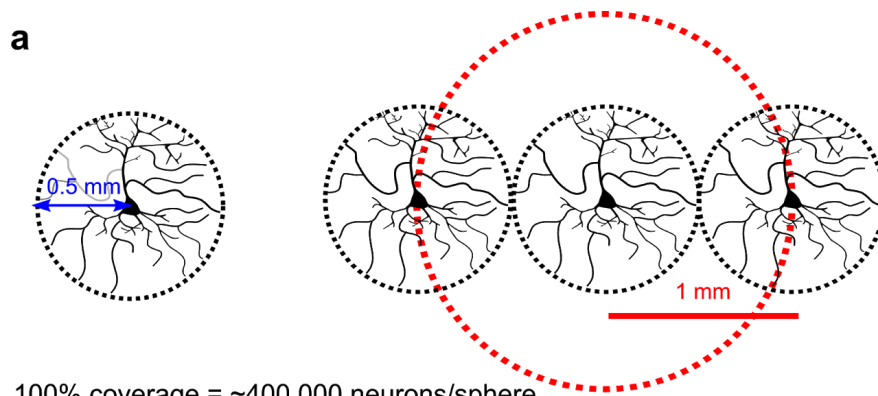
designs. A dashed line is drawn at 0.33 to simulate the optimal dark probability for trit coding. The grey shaded area represents a 10% tolerance.



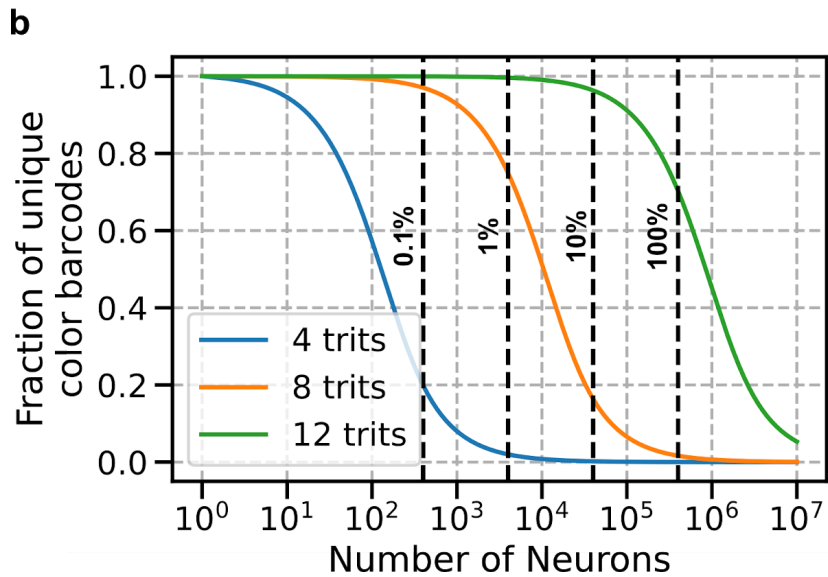
**Figure 3.3 Six new Brainbow AAV vectors with simultaneous membrane and cytoplasmic targeting.** (a) 12 fluorescent proteins were inserted in the 'on/off' design from Cai et al., but with the EF1alpha promoter replaced with Hsyn1. 'd' and 'sm' refers to dark and spaghetti monster fluorescent proteins with mutated chromophores to remove fluorescence. (b) Example confocal slice of membrane only targeted EYFP expression. (c) Example confocal slice of dual membrane and cytoplasmic targeted EYFP expression. Scale bars: (b-c) 15  $\mu$ m.



**Figure 3.4 Brainbow AAVs packaged in PHP.eB capsids can be injected intravenously to systemically infect neurons across the brain.** (a) Sagittal section of VGAT-cre mouse retro-orbitally injected with Brainbow AAVs from Cai et al. packaged in PHP.eB serotype. (b-d) Zoomed images of square insets in a, representing striatum, sensory cortex, and hippocampus respectively. Scale bar: 1 mm.



100% coverage = ~400,000 neurons/sphere  
 10% coverage = ~40,000 neurons/sphere  
 1% coverage = ~4,000 neurons/sphere  
 0.1% coverage = ~400 neurons/sphere



**Figure 3.5 Effect of color barcoding capacity on labeling different neural population sizes.** (a) Simple model to calculate the number of neurons that have potentially overlapping arbors, assuming that 1) dendritic and axonal arbors innervate a 0.5 mm radius sphere and 2) there are 100,000 neurons per 1 mm<sup>3</sup> volume. Doing so, we see that we only need to take into account the neurons that lie within a 1 mm radius sphere, which is an estimated ~400,000 neurons total. If we want to study a smaller fraction, such as interneurons, that are 10% of the whole population, then there are 40,000 neurons that need to be labeled. (b) Simulating the fraction of unique color barcodes given different digital Brainbow encoding schemes. Shown are three curves representing 4, 8, and 12 fluorescent proteins that are encoded as trits. The vertical dashed lines represent the estimated number of neurons that need to be labeled uniquely as calculated from a.

### 3.7 Tables

Fluorescent Protein	Genus	Family	Order
tagBFP	Entacmaea	Actiniidae	Actinaria
mTFP	Clavularia	Clavulariidae	Alcyonacea
dendra2	Dendronephthya	Nephtheidae	Alcyonacea
mCherry	Discosoma	Discosomidae	Corrallimorpharia
mEos	Lobophyllia	Lobophylliidae	Scleractinia
mAzamiGreen	Galaxea	Euphylliidae	Scleractinia
mKeima	Montipora	Acroporidae	Scleractinia
mMiCy	Acropora	Acroporidae	Scleractinia
mKikGR	Favia	Faviidae	Scleractinia
mKO2	Fungia	Fungiidae	Scleractinia
mPapaya	Zoanthus	Zoanthidae	Zoantharia
turboGFP	Pontellina	Pontellidae	Calanoida
EYFP	Aequorea	Aequoreidae	Leptothecata
phiYFP	Phialidium	Campanulariidae	Leptothecata
tagYFP	Aequorea	Aequoreidae	Leptothecata
neonGreen	Branchiostoma	Branchiostomidae	Amphioxiformes

**Table 3.1** Table of fluorescent proteins arranged by order, family, and genus.

	tagBFP	mTFP	dendra2	mCherry	mEos	mAzamiGreen	mKeima	mMiCy	mKikGR	mPapaya	turboGFP	EYFP	phiYFP	tagYFP	mKO2	neonGreen
tagBFP	100%	47.00%	49.80%	55.50%	47.00%	49.10%	51.60%	41.40%	45.70%	39.00%	27.60%	28.20%	30.60%	25.70%	43.50%	20.70%
mTFP	47.00%	100%	72.00%	56.40%	66.50%	72%	52.50%	51.00%	63.50%	47.30%	28.90%	31.80%	35.00%	30.20%	49.00%	32.30%
dendra2	49.80%	72.00%	100%	55.60%	71.30%	75.80%	60.00%	52.50%	67.00%	50.90%	27.40%	33.80%	36.60%	32.40%	51.50%	28.90%
mCherry	55.50%	56.40%	55.60%	100%	49.10%	53.70%	62.40%	50.00%	51.40%	47.30%	32%	30.90%	34.40%	27.90%	48.50%	32.60%
mEos	47.00%	66.50%	71.30%	49.10%	100%	70.8%	52.30%	51.80%	75.20%	49.50%	25.30%	34.80%	36.60%	34.10%	51.80%	29.80%
mAzamiGreen	49.10%	72%	75.80%	53.70%	70.8%	100%	58.2%	55.20%	73.90%	47.20%	28.30%	32.60%	30.60%	32.00%	56.10%	28.60%
mKeima	51.60%	52.50%	60.00%	62.40%	52.30%	58.2%	100%	51.00%	54.10%	49.70%	31.70%	32%	31.30%	30.50%	53.80%	28.20%
mMiCy	41.40%	51.00%	52.50%	50.00%	51.80%	55.20%	51.00%	100%	50.90%	49.60%	32.10%	31.90%	34.10%	33.30%	54.70%	33.90%
mKikGR	45.70%	63.50%	67.00%	51.40%	75.20%	73.90%	54.10%	50.90%	100%	52.00%	26.90%	32.60%	30.60%	31.00%	55.70%	30.70%
mPapaya	39.00%	47.30%	50.90%	47.30%	49.50%	47.20%	49.70%	49.60%	52.00%	100%	27.90%	26.30%	31.70%	36.70%	46.30%	30.60%
turboGFP	27.60%	28.90%	27.40%	32%	25.30%	28.30%	31.70%	32.10%	26.90%	27.90%	100%	26.40%	25.70%	25.30%	28.90%	34.40%
EYFP	28.20%	31.80%	33.80%	30.90%	34.80%	32.60%	32%	31.90%	32.60%	26.30%	26.40%	100%	52.60%	79.50%	34.10%	31.20%
phiYFP	30.60%	35.00%	36.60%	34.40%	36.60%	30.60%	31.30%	34.10%	30.60%	31.70%	25.70%	52.60%	100%	52.20%	30.60%	25.60%
tagYFP	25.70%	30.20%	32.40%	27.90%	34.10%	32.00%	30.50%	33.30%	31.00%	36.70%	25.30%	79.50%	52.20%	100%	31.70%	30.80%
mKO2	43.50%	49.00%	51.50%	48.50%	51.80%	56.10%	53.80%	54.70%	55.70%	46.30%	28.90%	34.10%	30.60%	31.70%	100%	30.40%
neonGreen	20.70%	32.30%	28.90%	32.60%	29.80%	28.60%	28.20%	33.90%	30.70%	30.60%	34.40%	31.20%	25.60%	30.80%	30.40%	100%

**Table 3.2 Matrix of protein sequence homologies.** Values > 70% are highlighted in red and signal increased concern for cross reactive antibodies because of similar sequences. For example, antibodies generated against tagYFP were found to be cross reactive with EYFP (data not shown). Sequences were obtained from fpbase.org.



## Chapter 4. Light Microscopy Based Approach for Mapping Connectivity with Molecular Specificity<sup>1</sup>

---

<sup>1</sup> This chapter represents the following manuscript:  
Shen, F. et al. Light microscopy based approach for mapping connectivity with molecular specificity.  
bioRxiv (2020). <https://doi.org/10.1101/2020.02.24.963538>

## 4.1 Introduction

Mammalian brains are extraordinarily complex pieces of circuitry, composed of trillions of connections between diverse cell types. Understanding these wiring patterns is a fundamental piece of the puzzle towards understanding how our brains work. A comprehensive wiring diagram of neuronal connections is called a connectome, while the pursuit of a connectome is known as connectomics. Neurons are connected to each other primarily through chemical synapses. They are located where presynaptic and postsynaptic neurons touch, and are composed of structures spanning hundreds of nanometers in size. The presynaptic compartment contains a region that mediates neurotransmitter release called the active zone<sup>73</sup>. The postsynaptic compartment contains a region that is protein dense called the postsynaptic density (PSD), which is comprised of neurotransmitter receptors, scaffolding proteins, and signaling molecules<sup>74</sup>.

Due to their small size, neuroscientists have relied on electron microscopy (EM) to observe chemical synapses and map connections between neurons. EM is a powerful tool that offers unparalleled nanometer resolution<sup>75</sup>, sufficient to observe synaptic structures. Recent works have used EM to define the adult *Drosophila* hemibrain connectome<sup>76</sup>, as well as curate a 92.6 x 94.8 x 61.8  $\mu\text{m}^3$  connectome from mouse somatosensory cortex<sup>77</sup>. Despite the recent technological and biological advances with EM, several challenges remain. Molecular information is lost as proteins can only be rarely identified with EM alone. Long imaging timespans generate large datasets that require demanding computational processing and analysis. All of these challenges have placed the use of EM for connectomics beyond the reach of common neurobiology labs. Moreover, because of the aforementioned difficulties in scaling the technology, the use of EM for connectomics is not currently suitable for high throughput experimentation. Given that neuroanatomy can vary between animals and substantially change in disease states, a new approach is needed for connectivity analysis that considers molecular information and is easily scalable.

In contrast to EM, light microscopy (LM) can generate specific molecular information through immunostaining, but lacks the nanoscale resolution to map synapses. Super-resolution light microscopy (srLM) techniques, such as STORM<sup>14,25,78</sup> or STED<sup>79</sup>, possess nanoscale resolution

and molecular specificity, but are not suitable for thick brain volumes because of slow imaging speeds, photobleaching, and optical distortions. The recent development of expansion microscopy<sup>17,28,80</sup>, which grants super-resolution to routine LM imaging through isotropic, physical magnification of hydrogel-tissue hybrids, offers an alternative path forward that can combine molecular specificity, nanoscale resolution, and rapid LM in thick brain volumes<sup>16</sup>.

Here, we describe a strategy to obtain high throughput morphology measurements of densely labeled neurons, with integrated molecular and connectivity information for multimodal analysis. We combine a multi-color genetic labeling tool (Brainbow) with a **multi-round immunostaining Expansion microscopy (miriEx)** strategy to simultaneously profile single neuron morphologies, molecular marker expression, and connectivity in the same brain section. We define the derivation of these properties from hyperspectral fluorescent channels as spectral connectomics, a LM based approach towards mapping neuroanatomy and connectivity with molecular specificity.

## 4.2 Results

Spectral connectomics is based on the ability to acquire multichannel LM datasets at nanoscale resolution that encode neuron morphologies, cell type profiles, and connectivity. Thus, we needed to develop a strategy for 1) dense labeling of a neuronal population with the ability to unambiguously trace dendrites and axons, 2) multiplexed readout of important protein markers, and 3) multiscale imaging that can range from nanoscale resolution for resolving synapses, to microscale resolution for resolving cell type markers. To do so, we first optimized an expansion microscopy protocol for multi-round immunostaining (miriEx) that let us generate multichannel LM datasets at multiple resolutions. We then added Brainbow to stochastically express fluorescent proteins (FPs) in a cell-type-specific population of neurons<sup>18,48</sup>. The combination of miriEx and Brainbow allowed us to curate hyperspectral, multiscale LM datasets that contain information about molecular markers, neuron morphologies, and synaptic machinery (**Fig. 1a**).

We start by describing the development of miriEx for multiplexed immunostaining. Probing multiple proteins using traditional immunohistochemistry (IHC) is typically limited by host animal species of primary antibodies (most tend to be either mouse or rabbit) and visible light bandwidth, such that detecting more than 4 targets becomes difficult. Recent strategies, such as

Immuno-SABER<sup>81</sup>, PRISM<sup>82</sup>, and CODEX<sup>83</sup> have been developed to overcome these limitations using antibody-DNA barcoding and read out strategies. Alternatively, tissue sections can undergo multiple rounds of routine antibody staining, imaging, stripping, and restaining, as seen in array tomography<sup>63</sup>, CLARITY<sup>84</sup>, MAP<sup>29</sup>, and SHIELD<sup>62</sup>. We adopted the latter strategy and optimized a protein crosslinking protocol to anchor antigens into an expandable hydrogel. Specifically, we used acrylic acid *N*-hydroxysuccinimide ester to modify proteins with acryl groups so they can be crosslinked and polymerized into an expandable hydrogel. We replaced the Proteinase K digestion step from standard expansion microscopy protocols<sup>28,80</sup> with SDS/heat based denaturation because Proteinase K destroys endogenous proteins. In contrast, SDS/heat treatment preserves endogenous proteins<sup>29</sup> and is compatible with post-gelation immunostaining. Furthermore, because our denaturation method is similar to that of SDS-PAGE western blots, we found that antibodies already validated for western blots usually work with miriEx (**Table 4.2**). We also found that the same SDS/heat treatment can efficiently strip antibodies after each round of probing (**sFig. 4.1**). We validated that gel-tissue hybrids in miriEx expand ~2x in 1x PBS and ~4x in 0.001x PBS respectively (**sFig. 4.2**).

We then demonstrated that miriEx robustly preserves antigens across multiple rounds of immunostaining and stripping. Conceptually, different rounds of staining can be used to probe different neuronal properties, such as molecular markers, morphology, and/or synaptic markers (**Fig. 1a**). We probed 5 different rabbit antibodies across 7 rounds of imaging in one basolateral amygdala (BLA) sample (**Fig. 1b**), imaging DAPI in each round to use as a fiduciary channel for registration (**sFig. 4.3**). We found that the same parvalbumin (PV) antibody showed similar staining patterns across multiple rounds with high signal remaining in neurites, indicating that antigen signal can be reliably retrieved, even in later rounds of miriEx. To extend the application of multi-round immunostaining, we showed that miriEx works with formalin-fixed human tissue (**sFig. 4.4**). We also demonstrated that through multispectral imaging for 3-4 antigens in each round, we could achieve highly multiplexed profiling of 15 different targets in the same piece of mouse striatum tissue (**sFig. 4.5**).

In the BLA multi-round immunostaining experiment described above, we observed that although the majority of amygdala PV neurons co-expressed calbindin (Calb), some were calbindin negative. Past studies indicate that these PV+/Calb- neurons are axo-axonic cells that specifically innervate the axon initial segment<sup>85</sup>, while PV+/Calb+ neurons represent basket cells

that innervate the perisomatic region<sup>86</sup>. Somatostatin (SOM) also marks another broad interneuron subtype in the amygdala, with both Calb positive and negative co-expression<sup>87</sup>. As a result, we chose this system to demonstrate the ability to interface molecular marker information with morphology analysis using miriEx.

To differentiate intermingled neurons *in situ*, we used Brainbow, a technique that relies on the stochastic expression of fluorescent proteins (FPs) to label neighboring neurons in unique colors<sup>18,48</sup>. Importantly, 1) the FPs used are distinct antigens, allowing their signal to be amplified through miriEx immunostaining, 2) the FPs are membrane targeted, which enables better labeling of neuron subcellular morphology<sup>48</sup>, and 3) dense labeling can be achieved to study morphology in a more high throughput manner compared to other techniques that rely on sparse labeling. PV-Cre and SOM-Cre double transgenic animals were generated to allow genetic access to two broad interneuron types. Brainbow AAVs 2/9 were stereotaxically injected in the basolateral amygdala (BLA), and 200  $\mu\text{m}$  sections of tissue were processed with miriEx (**Fig. 2a**). In round 1, three different molecular markers (Calb, PV, and SOM) were probed to define 4 molecular cell types: PV, PV/Calb, SOM, SOM/Calb (**Fig. 2b-e**). In round 2, three Brainbow FPs were immunostained to reveal morphology (**Fig. 2f-i**). DAPI was co-stained as a fiduciary channel for registration of the two rounds. Both rounds of imaging took place with the gel-tissue hybrid expanded  $\sim 2\times$  in 1xPBS, giving us an effective imaging resolution of  $\sim 150 \times 150 \times 350 \text{ nm}^3$ . Brainbow AAVs labeled 2 out of 4 (50%) PV neurons, 20 out of 31 (70%) PV/Calb neurons, 7 out of 7 (100%) SOM neurons, and 24 out of 28 (86%) SOM/Calb neurons within a  $590 \times 404 \times 160 \mu\text{m}^3$  volume. Following identification of each Brainbow neuron by its molecular subtype, we reconstructed its dendritic morphology using nTracer<sup>88</sup>, an ImageJ/Fiji plugin for tracing multispectral datasets (**Fig. 2g-h**). A total of 53 neurons in the imaged volume was reconstructed across all 4 molecular subtypes and various morphology parameters were analyzed (**Fig. 2l-m, sFig. 4.6**). In agreement with past findings, PV expressing neurons appear to have more complex branching patterns compared with SOM expressing neurons, despite having similar dendritic lengths<sup>25,26</sup>. This type of dense reconstruction enables study of how different neuronal cell types interact with each other anatomically, making it a powerful tool for network morphology analysis.

While understanding the projection patterns of axons and dendrites is an important aspect of mapping neuroanatomy, another central objective is to understand how neurons connect.

Chemical synapses are hundreds of nanometers in size and serve as bridges of communication between neurons<sup>78</sup>. Measuring synapses using conventional light microscopy (LM) techniques is difficult as the distance between synaptic structures and neuronal boundaries can be smaller than the diffraction limit<sup>25</sup>. Recently, expansion microscopy has been shown to be a viable strategy for resolving synaptic structures and assigning them to neurons<sup>16,89,90</sup>. Consequently, we combined Brainbow with miriEx to measure synaptic structures located at the junctions between different neurons in an effort to define connectivity using LM.

We first confirmed that miriEx was compatible with endogenous synaptic machinery immunostaining by probing for Gephyrin (inhibitory PSD), Homer1 (excitatory PSD), and Bassoon (presynaptic active zone) in layer 4 somatosensory cortex (**sFig. 4.7**). ~4x expansion of the sample gave us an effective imaging resolution of ~70 x 70 x 200 nm<sup>3</sup> using confocal microscopy. We found inhibitory Gephyrin-Bassoon synapse pairs to be less common (21%) than excitatory Homer1-Bassoon synapse pairs (77%). Importantly, only ~2% of the Bassoon puncta were not paired with Gephyrin or Homer1, which were mutually exclusive (**sFig. 4.7**). This gave us confidence that the Gephyrin and Homer1 antibodies we used mark inhibitory and excitatory synapses almost completely. We also showed that over 90% of inhibitory and excitatory synapses are greater than 300 nm away from their closest neighbor, allowing us to reliably distinguish neighboring synapses at our imaging resolution (**sFig. 4.7**).

We then packaged Brainbow in AAV-PHP.eB<sup>54</sup> serotype to efficiently transduce neurons systemically across the brain via intravenous injection. We retro-orbitally injected AAV-PHP.eB Brainbow in PV-Cre mice and found that we had near complete coverage of PV neurons in somatosensory cortex (**sFig. 4.8**). 100% of the Brainbow labeled neurons were positive for PV immunostaining indicating that our labeling strategy was both highly sensitive and specific (**sFig. 4.8**). 100  $\mu$ m sections of somatosensory cortex were processed with miriEx. In round 1, three Brainbow FPs were stained and the sample was expanded and layer 4 was imaged (**Fig. 3a-b**). In round 2, a presynaptic marker (Bassoon), inhibitory postsynaptic marker (Gephyrin), and EYFP were stained, and the sample was again expanded and imaged in layer 4 (**Fig. 3c**). We observed that Bassoon-Gephyrin pairs could be resolved and were located at axosomatic and axodendritic contact points between Brainbow labeled PV neurons (**Fig. 3d-e**). Measuring the line profile of these putative synapses revealed that Gephyrin, the post-synaptic Brainbow membrane, and Bassoon were arranged in the expected order (**Fig. 3h-i**). The distance

between Gephyrin and Bassoon puncta was between 100-200 nm and matched previous reports<sup>78,89</sup>. The synaptic cleft between pre- and post-synaptic Brainbow membranes is tens of nanometers wide and could not be resolved at this resolution, but the trio of Gephyrin-Brainbow-Bassoon signal still gave us confidence to define putative inhibitory synapses between Brainbow labeled PV neurons. Moreover, our approach towards spectral connectomics allowed us to eliminate false positives that would not be possible with diffraction-limited LM. Neuronal contacts have previously been used as correlates for synaptic sites<sup>88,91</sup>, but we show an example of a close apposition lacking synaptic machinery that would be falsely classified as a connection (**Fig. 3g, k**). Another type of error that could arise is the wrongful assignment of a physically close axon that merely passes through and doesn't form a synapse (**Fig. 3f, j**). Given that axons can be small caliber with bouton sizes that are hundreds of nanometers<sup>12</sup>, diffraction limited LM may confuse neighboring axons, leading to synapse assignment mistakes.

Following the first two rounds of miriEx that probed Brainbow and synaptic markers, a third round of miriEx was performed to probe SOM, adding information about molecular cell types to the dataset (**sFig. 4.9**). By doing so, we show the ability to map PV connectivity onto a variety of cell types. However, we note that it remains difficult to accurately assign synapses without both the presynaptic and postsynaptic Brainbow membranes. Without a postsynaptic SOM neuron membrane label, it is challenging to tell if the PV axon synapses directly on the soma or on an unlabeled, small caliber dendrite sandwiched in between. As a result, we did not attempt to analyze PV to SOM connectivity.

After validating that putative inhibitory synapses could be identified between Brainbow labeled PV neurons, we set out to trace the axons and dendrites of 8 PV neurons whose somas were located inside a  $\sim 100 \times 100 \times 60 \mu\text{m}^3$  imaging volume (**Fig. 4a**). We then traced every Brainbow labeled PV axon that innervated these 8 PV neurons, and annotated all the putative inhibitory synapses (Gephyrin-Brainbow-Bassoon trios) that we could identify. 189 axons were traced and 422 molecularly specific (PV-PV) putative synapses were defined (**Fig. 4b**). First we analyzed the connections between these 8 PV neurons by plotting their connectivity matrix (**Fig. 4c**). We observed that neuron 37 innervated 4 other PV neurons, matching previous reports of local PV-PV connectivity<sup>92</sup>. Interestingly, while local PV-PV connection is common in our dataset, we did not find local reciprocal inhibition between two PV neurons. Next we added 189 traced PV

axons, whose somas were not located in this volume, to the same connectivity matrix (**Fig. 4d**). We found multiple examples of presynaptic PV axons that innervate more than one postsynaptic PV neuron. However, many of the dimly labeled PV axons could only be reliably traced for a short distance, leaving the full extent of their synaptic connections uncovered. This likely skewed the connectivity matrix towards under-representing PV axons that synapse onto multiple postsynaptic PV neurons. Compared to traditional monochromatic labeling, Brainbow labeling provides spectral information to identify the source of innervating axons. For example, we can be confident that two spectrally unique axons come from different PV neurons even without tracing them back to their somas. Some colored axons are more difficult to interpret as they could come from different neurons that happen to have similar colors, or they could originate from the same neuron branch *outside* our imaging volume. Nevertheless, we can still use color to estimate the upper and lower bounds when asking the convergence question of *how many unique PV neurons innervate one PV neuron*. The upper bound is determined by the number of innervating axons, while the lower bound is determined by the number of unique colors these axons possess (**Fig. 4e**). To determine the number of unique colors, we plotted their RGB color on a ternary plot and used an elbow plot to conservatively estimate the number of k-means color clusters (**sFig. 4.10**). Within our dataset volume, the average upper and lower bound of unique PV neurons converging onto one PV neuron are estimated as  $27.5 \pm 12.6$  and  $10.9 \pm 4.5$  respectively ( $n = 8$ , mean  $\pm$  standard deviation).

Next we shifted our attention from the distribution of axons to the distribution of synapses. We analyzed the distribution of PV-PV inhibitory synapses by splitting them into two spatial compartments: somatic and dendritic (**Fig. 4f**). We plotted the number of soma or dendrite targeting axons based on the number of putative synapses they provide. Cortical PV neurons are known as basket cells and tend to innervate the perisomatic region of other neurons<sup>33</sup>. While we confirmed the existence of putative PV-PV axosomatic synapses, a larger number was actually axodendritic. The soma targeting axons most commonly formed two putative synapses, but in contrast, the majority of the dendrite targeting axons formed one putative synapse. **Fig. 3d** shows an example of a single axonal bouton providing two putative axosomatic synapses. We then counted the total number of putative inhibitory synapses on the soma and plotted it alongside those that were annotated as PV-PV (**Fig. 4g**). We observed that approximately  $33.3 \pm 16.5$  % of the somatic inhibitory synapses were PV-PV ( $n = 8$ , mean  $\pm$  standard deviation). Because inhibitory synapses are plastic and can dynamically remodel<sup>93-95</sup>, we wanted to



determine if there were any differences in synapse size, a known correlate for synapse strength<sup>96,97</sup>, between the 8 postsynaptic PV neurons. We looked at the distribution of Gephyrin volume for all PV-PV synapses for each of the 8 postsynaptic PV neurons and found they were largely consistent with no major differences (**Fig. 4h**). Next, we analyzed individual postsynaptic PV neurons and asked whether axons that provide more synapses are correlated with increased PSD sizes. Again, we observed PSD sizes were consistent and invariant of the number of synapses an axon formed (**sFig. 4.11**).

Previous studies performing *in situ* synapse measurements tend to focus on excitatory neurons because dendritic spines are commonly used as a proxy for excitatory input<sup>98</sup>. Changes in spine size and density are often used as a correlate to structural synaptic plasticity<sup>98</sup>. Studies of inhibitory neuron synapses are more challenging and underrepresented since many of them are aspiny with no obvious morphological correlate. To meet this challenge, we simultaneously labeled aspiny BLA PV neurons with Brainbow, excitatory (Homer1), and inhibitory (Gephyrin) PSD markers (**Fig. 5a**). In round 1, three Brainbow FPs were immunostained and the sample was expanded and imaged (**Fig. 5b**). In round 2, Homer1 and Gephyrin were immunostained along with EYFP as a fiduciary channel (**Fig. 5c**). Expansion of the sample gave us the resolution to optically resolve individual synaptic puncta along the Brainbow membrane (**Fig. 5d-i**). We reconstructed the dendritic morphology of 5 PV cells in a 220 x 220 x 85  $\mu\text{m}^3$  imaging volume, and annotated all the excitatory ( $915 \pm 274$ , mean  $\pm$  standard deviation,  $n=5$ ) and inhibitory ( $409 \pm 141$ , mean  $\pm$  standard deviation,  $n=5$ ) PSDs to create a synaptic input map for each cell (**Fig. 5j-k**). We propose that these synaptic input maps may be another useful metric to distinguish cell types. For example, we intriguingly observed that three out of five PV neurons possess a skewed distribution of excitatory vs. inhibitory inputs, while the other two PV neurons possess a more balanced distribution (**Fig. 5l**). The excitatory vs. inhibitory input ratio fundamentally influences a neuron's role in the circuit, and is another measure of connectivity that can be used to distinguish neuronal subtypes.

## 4.3 Discussion

In summary, we present a LM based approach for connectivity analysis with molecular specificity, termed spectral connectomics, developed by combining Brainbow and miriEx

technologies. We demonstrated multimodal measurements of morphology, molecular markers, and connectivity in a single mammalian brain section. We showed that miriEx supports robust preservation of antigens to enable multiple immunostaining, imaging, and stripping cycles. This allowed us to reconstruct the morphology and specify the molecular subtype of 53 different BLA interneurons in the same brain section. Adding endogenous immunostaining of pre and postsynaptic pairs enabled us to define putative synaptic connections between molecularly specified neurons. We traced 8 postsynaptic PV neurons, 189 innervating PV axons, and 422 PV-PV inhibitory synapses, and performed connectomic analyses similar to previous EM studies. Finally, we showed that Brainbow could be combined with endogenous PSD immunostaining to quantify inhibitory and excitatory synaptic input onto individual aspiny interneurons.

miriEx enhances the potential of using Brainbow for mapping neuroanatomy. Depending on the labeling density, diffraction-limited LM may be insufficient for tracing small caliber Brainbow neurites<sup>8,26</sup>. Expansion microscopy allowed us to resolve these neurites effectively for morphology and connectivity analysis. The membrane targeted FPs also proved to be an useful landmark when probing synaptic machinery. Both pre and postsynaptic membranes were required along with immunostaining for pre and postsynaptic machinery to identify putative synapses. In addition, membrane targeted FPs are shown to be better than cytosolic FPs, which may not diffuse as well, for depicting small subcellular structures<sup>25</sup>. Two challenges that remain with combining Brainbow and miriEx are the limited color diversity and dilution of membrane FP signal. Throughout our experiments, we immunostained for only three out of four possible FPs. We can already observe tens of easily distinguishable color barcodes, which can be exponentially increased in the future by probing more FPs. For example, 5 FPs were used in the recently developed Bitbow, a “digital” format of Brainbow<sup>51</sup>. Expansion of the sample is important for increasing imaging resolution, but comes at the cost of diluting antibody signal. For instance, ~4x expansion results in a ~64 fold reduction of antibody signal. As a result, axons can become dim and challenging to trace long distances. That said, we expect a more photon efficient volumetric imaging modality, such as light sheet microscopy, will yield better signal to noise under the same labeling conditions than confocal microscopy used in our study. Moving forward, it will be important to optimize FP expression along axonal membranes, as well as explore alternative signal amplification technologies, such as immuno-SABER<sup>81</sup>.

A critical component of our spectral connectomics strategy is molecular specificity, which comes from two sources: endogenous protein immunostaining and choice of Cre-driver line. The ability to probe endogenous molecular markers with miriEx increases the flexibility of studying multiple neuronal cell types simultaneously. For example, future work can focus on using a broad transgenic line (e.g. GAD2-Cre) that can be molecularly specified by additional immunostaining (PV, SOM, VIP, etc). Furthermore, miriEx immunostaining can be extended to investigate the molecular content of synapses. Bassoon, Gephyrin, and Homer1 are only a subset of the synaptic machinery that can be probed, which includes neurotransmitter transporters, receptors, gap junctions, and ion channels. The rich library of Cre-driver lines can give us a genetic handle on studying specific cell types. Labeling a subset from the global population of neurons also helps to constrain and focus our connectivity analysis. For example, we only quantified and analyzed 422 PV-PV synapses out of ~60,000 inhibitory synapses in our dataset (~0.7% of the total). Because we asked molecularly specific, targeted questions of PV-PV connectivity, we reduced the burden of tracing neurons and annotating synapses to a scale that was manageable by manual curation (~40 human work hours).

Moreover, many Cre-driver lines are actually composed of different neuronal subtypes<sup>99</sup>. Our definitions of cell types are constantly changing with the introduction of new technologies. We have progressed from using unimodal criteria (e.g. morphology, physiology, or molecular expression) to using a combination of these properties in a multimodal fashion<sup>100</sup>. Recent studies have demonstrated the power of this multimodal approach to curate and refine definitions of inhibitory and excitatory neuronal cell types in visual cortex<sup>101,102</sup>. Similarly, our multimodal strategy for measuring morphology, molecular markers, and connectivity can be also applied towards refining cell type definitions across the brain. More specifically, integrating connectivity information will enhance our understanding of how input/output properties inform cell type definitions. In addition, spectral connectomics is compatible with techniques that measure physiological properties, such as patch clamp electrophysiology and *in vivo* calcium recordings, for more comprehensive analysis.

Finally, we envision spectral connectomics to supplement EM as an alternative strategy for mapping connectivity that is accessible to neurobiology labs. miriEx and Brainbow use commercially available, “off the shelf”, reagents, and LM is a prevalent imaging modality across many labs. Data sizes are an order of magnitude smaller compared to EM datasets despite

containing multiple channels (~20 GB vs ~200 GB for a similar sized volume<sup>77</sup>). Targeted analysis makes it possible to map connectivity in a reasonable amount of time, and application of automated reconstruction and synapse segmentation methods will further simplify this process. EM based connectomics currently remains the gold standard for generating complete reference connectomes of animal brains, but we hope that spectral connectomics can be employed for validating molecularly specific circuit motifs and testing hypotheses of how neuroanatomy dynamically changes with perturbations. We imagine that in the future, spectral connectomics can be scaled up to generate comprehensive connectomes. This likely will come from developing new Brainbow designs to globally label neurons<sup>54</sup>, whole brain expansion microscopy protocols, whole brain LM methods<sup>103,104</sup>, whole brain immunolabeling techniques<sup>38,105</sup>, and new computational pipelines for automated neuron reconstruction<sup>106</sup> and synapse segmentation<sup>107</sup>. Although challenging, many of these puzzle pieces are already being developed by the community. When combined in the spectral connectomics framework, they can fit together to reveal a mammalian connectome.

## 4.4 Methods

### Mouse Lines

All experiments were carried out in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Our study protocol was reviewed and approved by the University of Michigan Institutional Animal Care & Use Committee. The transgenic mice used in this study were: PV-Cre (Jackson stock no. 008069), SOM-Cre (Jackson stock no. 013044), VGAT-Cre (Jackson stock no. 016962), and Thy1-YFP-H (Jackson stock no. 003782). P28 to P56 male and female mice were used for virus injections.

### Brainbow AAV injections

Brainbow3.0 AAV-2/9 and AAV-PhP.EB were obtained from Addgene and University of Michigan vector core, respectively. Transgenic mice (PV-Cre/SOM-Cre or PV-Cre) were anesthetized continuously with isoflurane and mounted on a stereotaxic frame. TagBFP-EYFP and mCherry-TFP virus were mixed together to reach a concentration of 1E12 gc/mL individually, of which 500 nL was injected at 100 nL/min using a capillary pipette backfilled with mineral oil at +3.5 ML and -1.7 AP relative to bregma, 2.7 DV from pia surface. Afterward, the

pipette was left in place for 5 min. for the virus to diffuse, before slowly retracting out of the brain. We waited 3-4 weeks for virus expression before perfusing the animals. Brainbow3.0 AAV-PhP.EB was used for retro-orbital injection to systematically label neurons throughout the brain. 50  $\mu$ L of mixed virus ( $1 \times 10^{12}$  gc total each for TagBFP-EYFP and mCherry-TFP) was injected into the retro-orbital sinus. We waited 3-4 weeks for virus expression before perfusing the animals.

### **Mouse perfusion and tissue sectioning**

Mice were anesthetized with tribromoethanol and perfused transcardially with ice-cold 1xPBS, followed by 4% PFA in 1xPBS. The brains were dissected and postfixed in 4% PFA in 1xPBS overnight shaking at 4 °C. The next day, brains were washed in 1xPBS before slicing on a vibratome (Leica VT1000s) at 100  $\mu$ m or 200  $\mu$ m thickness.

### **miriEx protocol**

Acrylic acid N-hydroxysuccinimide ester (AAX, Sigma A8060) was prepared by dissolving in N,N-Dimethylformamide to 125 mM. Tissue samples were incubated with 1-5 mM Aax in a MBS buffer (100 mM MES, 150 mM NaCl, pH 6) with 0.1% Triton X-100 shaking overnight at 4 °C. The next day, the samples were washed 3x with 1xTBS (Bio-rad 1706435) for 1 hour each to quench the reaction. The samples were then incubated in monomer solution (5.3% Sodium Acrylate, 4% Acrylamide, 0.1% Bis-Acrylamide, 0.5% VA-044, in 1x PBS) shaking overnight at 4 °C. The next day, the samples were gelled for 2.5 hours at 37 °C in a humidity chamber by cover slipping the sample surrounded by monomer solution in a gelling chamber. The gel-tissue hybrids were then carefully cut out and denatured at 70 °C overnight in denaturing buffer (200mM SDS dissolved in 1xTBS). The next day, denaturing buffer was washed out by incubating the samples 4x for 2 hours each in 0.1% PBST (1x PBS with 0.1% Triton X-100) shaking at 50 °C.

### **Immunohistochemistry of regular tissue sections**

For immunostaining of non-gelled samples, tissue sections were first blocked and permeabilized in StartingBlock (Thermo 37538) with 1% Triton X-100 shaking overnight at 4 °C. Tissues were washed the next day 3x with 1xPBS for 1 hour each. Tissues were then incubated with primary antibodies diluted in 0.5% PBST for 3 days at 4 °C. After washing 3x with 0.5% PBST for 1 hour each at RT, tissue sections were then incubated with secondary antibody

diluted in 0.5% PBST for 2 days at 4 °C. Tissues were then mounted with Vectashield (Vectorlabs H-1000) and imaged. Primary and secondary antibody choice and concentrations can be found in **Table 4.1**.

### **Immunohistochemistry of gelled tissue sections**

Blocking and permeabilization of gel-tissue hybrids were skipped as they were found to minimally decrease background. Gel-tissue samples were incubated with primary antibody diluted in 0.1% PBST w/ 0.02% azide (PBSTz) at either RT or 37 °C. They were washed 3x with 0.1% PBST w/ 0.02% Azide at RT for 1 hour each, and then incubated with secondary antibody diluted in 0.1% PBSTz at either RT or 37°C. Primary and secondary antibody choice and concentrations can be found in **Table 4.1**. Vendor information can be found in **Table 4.2 and 4.3**.

### **Gel expansion and fluorescence microscopy**

miriEx gel-tissue hybrids were incubated in 0.001x PBS three times for 45 min. each for saturated expansion to ~4x. They were then mounted in Poly-L-Lysine coated 6 cm dishes (Corning 354517) and submerged in 0.001x PBS. All confocal LM was performed using an upright Zeiss LSM780. Water immersion objectives (10X/NA0.4 or 20x/NA1.0) were directly lowered into the PBS solution over the sample for imaging. More imaging details are presented in **Table 4.1**.

### **Antibody elution**

miriEx gel-tissue hybrids that previously have undergone immunostaining, expansion, and imaging were shrunken by washing in 1xPBS 3x for 1 hour each. They were then put in denaturing buffer (200 mM SDS dissolved in 1xTBS) at 70 °C overnight. Samples were washed the next day 4x for 2 hours each in 0.1% PBST shaking at 50 °C before undergoing the next round of immunostaining.

### **Human brain sample processing with miriEx**

Formalin-fixed human brain samples were obtained from the University of Michigan Brain Bank. A 1 cm slab of sensory cortex was macro dissected and washed in 1x PBS at 4 °C overnight. The sample was then sectioned into 100 µm slices before processing with miriEx.

## **Image preprocessing**

Specific image processing steps for each imaging dataset are listed in **Table 4.1**, and are briefly explained below.

Stitching of multi-tile datasets was performed using the BigStitcher<sup>41</sup> ImageJ/Fiji plugin. The dataset was loaded and converted to HDF5 format, and the tiles were arranged in the order they were imaged with 10% overlap. The Stitching Wizard was used to calculate pairwise shifts using phase correlation, verify links, and undergo global optimization. Affine refinement was then performed with a high threshold. The resulting interest points were used for non-rigid refinement during advanced image fusion.

Chromatic aberration was corrected using the Detection of Molecules (DoM) ImageJ/Fiji plugin. To calibrate the DoM plugin, 0.5  $\mu\text{m}$  TetraSpeck fluorescent beads were mounted on a slide and imaged on a Zeiss LSM780 confocal microscope. More specifically, we calibrated the orange channel (540-600 nm) excited with a 543 laser and far-red channel (630-700 nm) excited with a 633 laser to the green channel (480-540 nm) excited with a 488 laser. Calibration was performed for both the 10x and 20x objectives.

Histogram matching was done to normalize intensity between z-slices in image stacks using the nTracer Align-Master ImageJ/Fiji plugin. A high signal to noise z-slice (usually at the top or bottom of the stack) was chosen as the reference slice for which the rest of the stack was normalized to. All three Brainbow FP channels were histogram matched to the same reference slice.

## **Image registration and alignment between rounds.**

Following data preprocessing as described above, the fiducial marker channels from different rounds were loaded into ImageJ/Fiji Big Warp<sup>108</sup> plugin for rough, initial alignment. After this, the fiducial marker channels were registered through a B-spline transformation using Elastix<sup>109,110</sup>. The resulting transformation was applied to each individual channel to create a merged image hyperstack.

## **Morphology reconstruction, synapse identification, and analysis**

nTracer, an ImageJ/Fiji plugin, was used to trace somas, dendrites, and axons of Brainbow labeled neurons. The morphology reconstructions were exported in SWC format, and a custom Python script was used to render skeleton visualizations in TIFF format from SWC files. Putative synapses were identified and manually marked using ROI manager in ImageJ/Fiji. The list of X,

Y, Z synapse coordinates was saved to a CSV file and linked with their parent neurons by adding an additional data column in the SWC file marking synaptic locations along the dendrite or axon. Blender 2.81 (Blender Foundation; [www.blender.org](http://www.blender.org)) or 3Dscript<sup>111</sup> was used to generate movies. Morphology features (i.e. number of stems, bifurcations, branches, etc.) were calculated by importing SWC files into Vaa3D's Global Neuron Feature plugin<sup>112</sup>. Sholl analysis was performed using ImageJ/Fiji Sholl Analysis plugin. Automatic segmentation of Gephyrin volumes was achieved through a custom Python script by first applying an Otsu thresholding step, followed by watershed segmentation of Gephyrin puncta. Plots were made in GraphPad Prism 8 or through Matplotlib Python scripts.

### **miriEx expansion measurement**

100  $\mu$ m Thy1-YFP-H sections were mounted in Vectashield and imaged using confocal microscopy. The sections were then washed in 1xPBS to remove the vectashield and were processed with miriEx. An anti-GFP antibody was used to label native YFP signal. After miriEx, the samples were imaged in 1xPBS in the same location as before. They were then further expanded in 0.001x PBS (3x washes for 45 min. each) and imaged again. Elastix<sup>109</sup> was used to calculate an affine transformation between different expansion states, and the X, Y, and Z scaling factors were averaged to measure the expansion factor.

### **Data Availability**

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### **Code Availability**

Custom code for analysis of image processing and data analysis is available from the corresponding author upon reasonable request.

## **4.5 Acknowledgements**

We thank Nigel Michki and Douglas Roossien for providing comments, editing, and discussion of the manuscript. We thank Harsimranjit Sekhon for insightful discussion. We thank Andy Lieberman for providing formalin-fixed human brain tissue. We thank Bo Duan and Lorraine Horwitz for providing SOM-Cre mice and genotyping assistance. FYS acknowledges support by

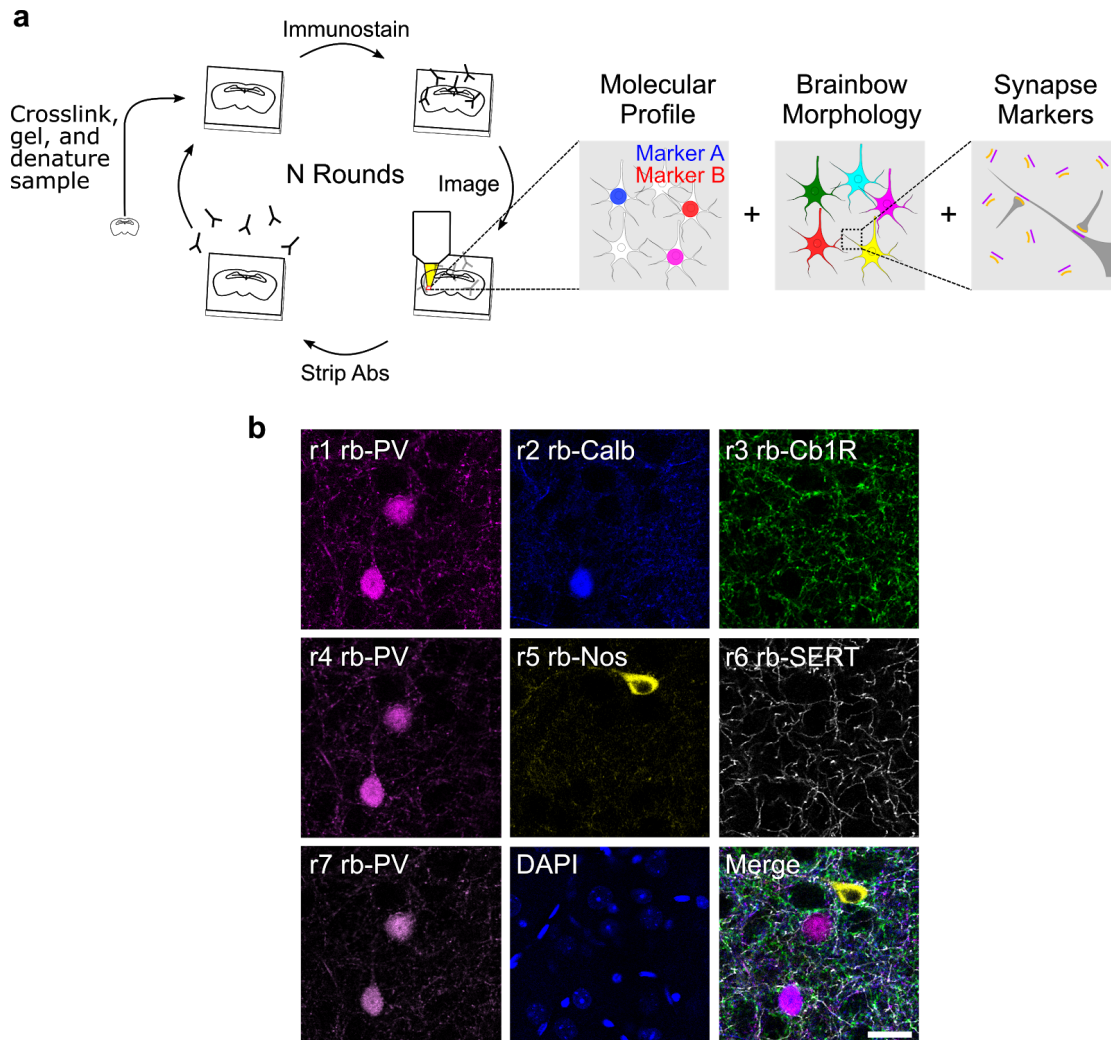


National Institutes of Health (NIH) 1F31NS11184701. DC acknowledges support by NIH 1UF1NS107659, 1R01MH110932, and 1RF1MH120005, and National Science Foundation NeuroNex-1707316. ESB acknowledges support by Lisa Yang, John Doerr, the Open Philanthropy Project, and Schmidt Futures, and NIH 1R01NS102727, 1R01EB024261, and 1R01MH110932.

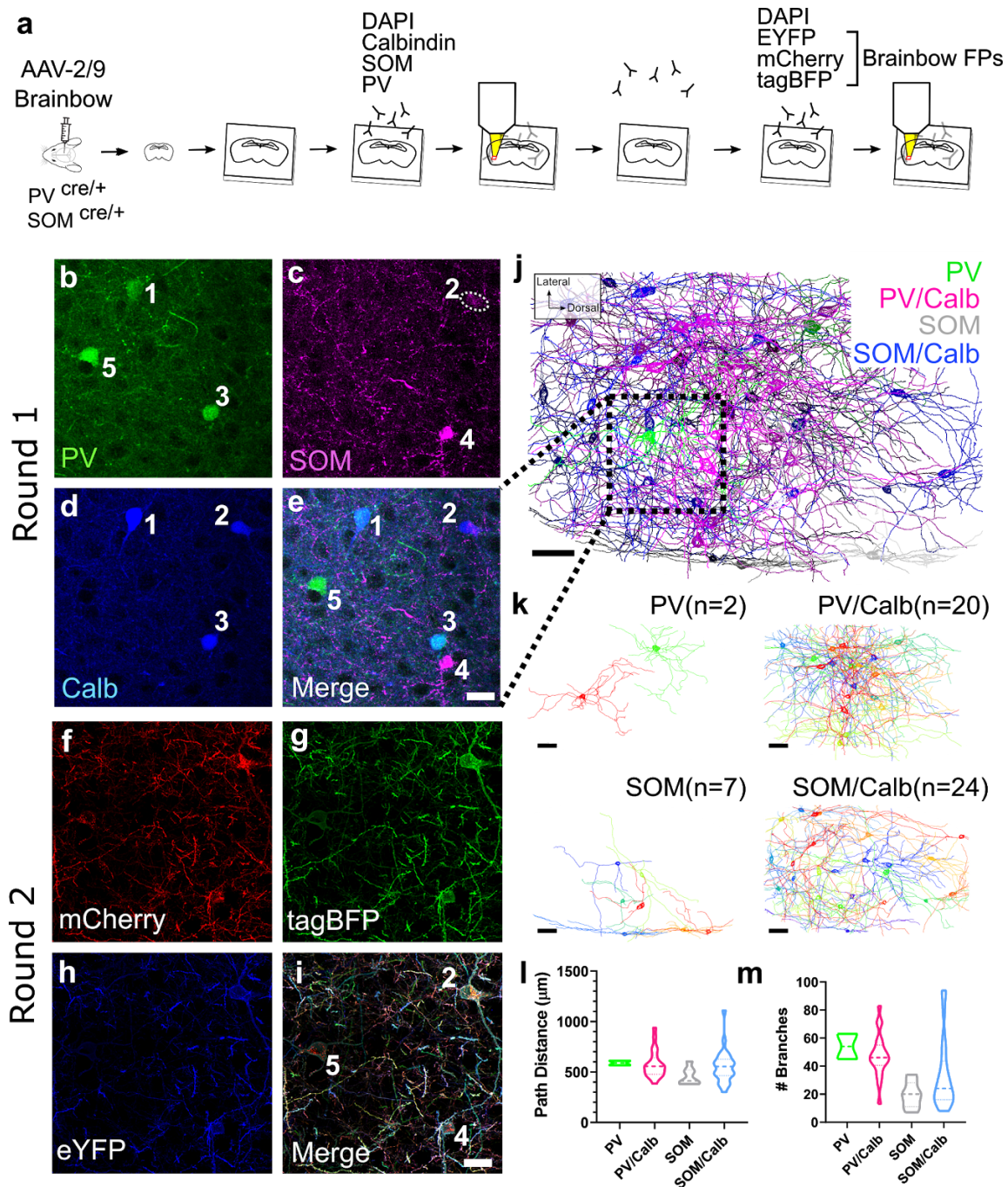
## 4.6 Author contribution

FYS and DC conceived the project, designed the experiments, and wrote the manuscript with input from all authors. FYS performed and analyzed all miriEx and Brainbow experiments. HPJC and FYS traced the 53 BLA PV and SOM neurons. LAW wrote code for automatic segmentation of Gephyrin puncta and helped create the BLA molecular subtypes video. MMH and FYS quantified Gephyrin-Bassoon and Homer1-Bassoon pairs, and PV and Brainbow overlap. ESB supported ExM development. DC initiated and supervised the project.

## 4.7 Figures

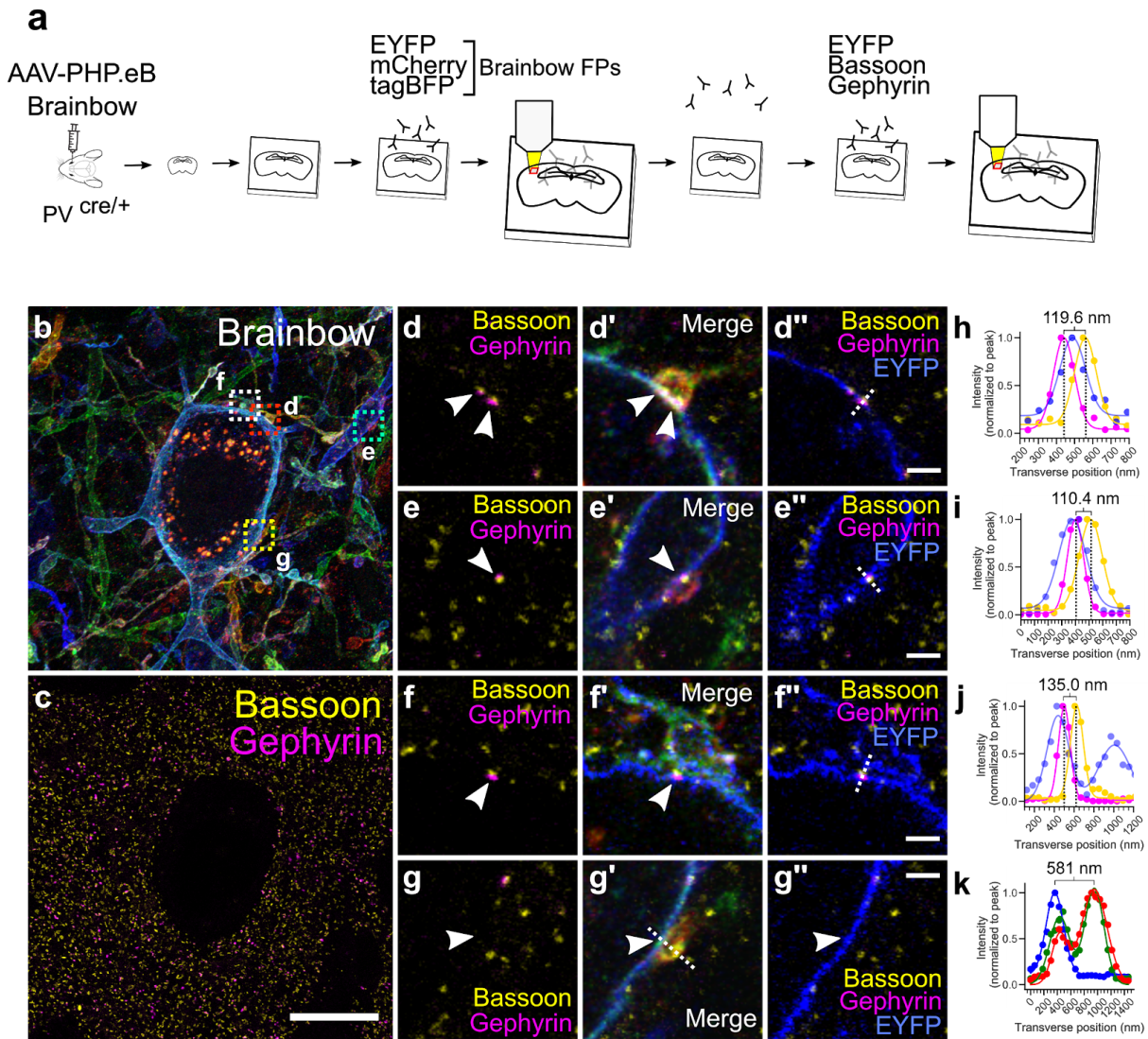


**Figure 4.1 miriEx robustly preserves antigens across multiple rounds of immunostaining, imaging, and stripping.** (a) miriEx strategy to simultaneously measure molecular profile, morphology, and/or connectivity. Tissue samples are embedded in a hydrogel to create expandable gel-tissue hybrids. The sample can then undergo multiple rounds of immunostaining, imaging, and stripping to measure various neuron properties. We then register and merge the different rounds of imaging to correlate the results. (b) 5 distinct rabbit antibodies were used across 7 sequential rounds of immunostaining, with PV re-probed in rounds 4 and 7 to demonstrate retention of antigen. Abs, antibodies; PV, parvalbumin; Calb, calbindin; Cb1R, cannabinoid receptor type 1; Nos, nitric oxide synthase; SERT, serotonin transporter. Scale bar: (b) 25  $\mu\text{m}$  (pre-expansion size).



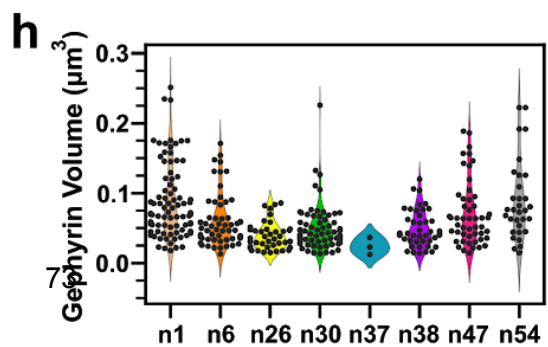
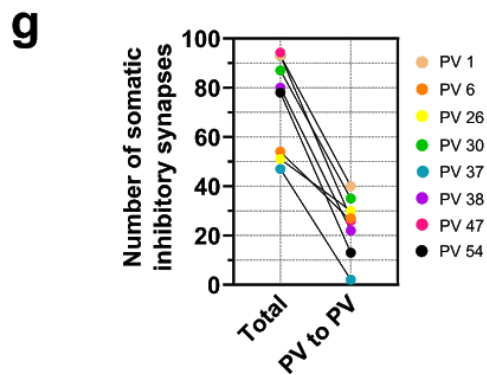
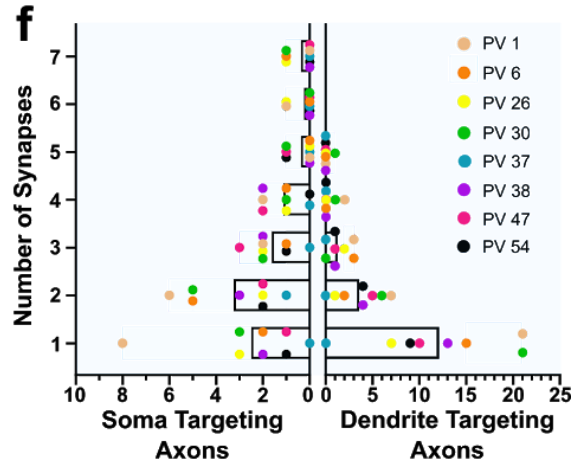
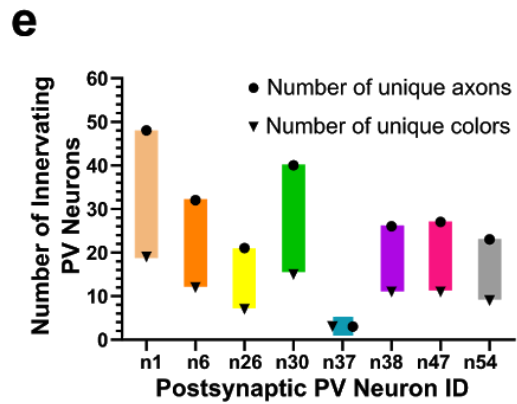
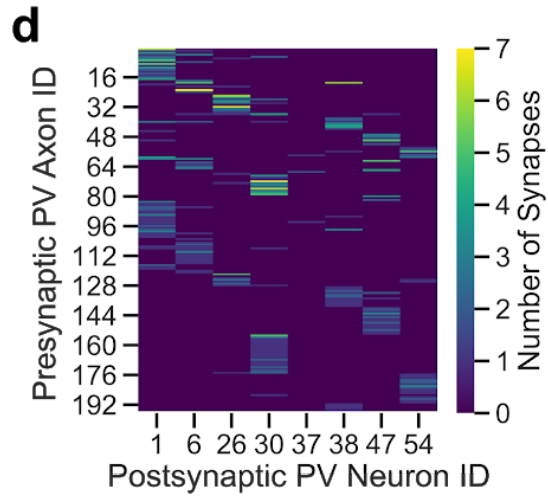
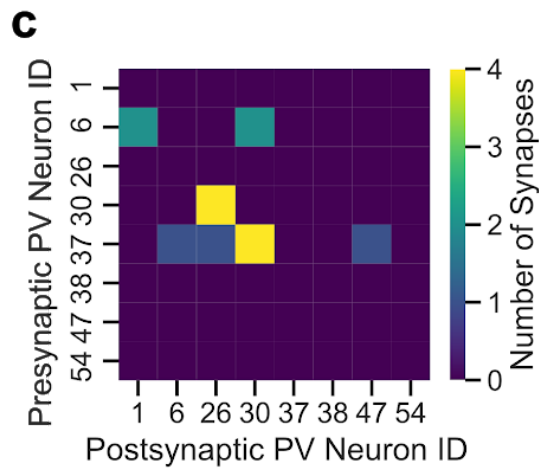
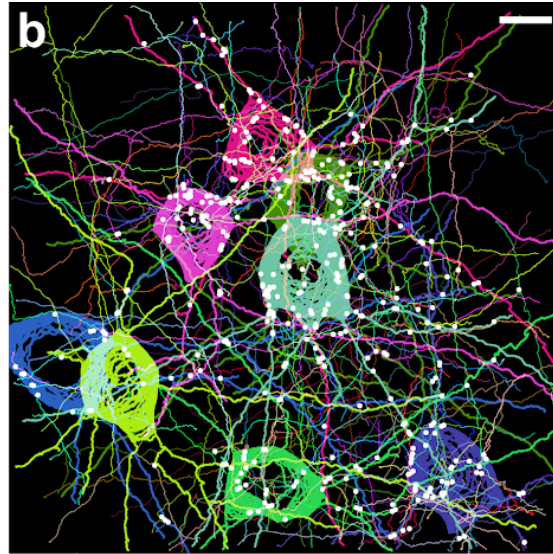
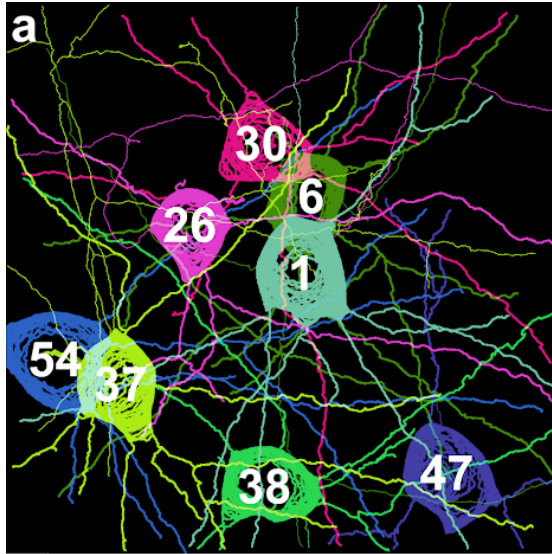
**Figure 4.2 miriEx can be combined with Brainbow to simultaneously profile molecular expression and neuron morphologies.** (a) Experiment design: Brainbow and molecular markers are imaged across two rounds of immunostaining using DAPI channel for registration. (b-e) MIP showing the molecular markers (PV, SOM, Calb) imaged in round 1. Four distinct molecular subtypes could be identified: PV, PV/Calb, SOM, SOM/Calb. (f-i) MIP showing the Brainbow channels imaged in round 2. (j) nTracer reconstruction of 53 neurons encompassing the 4 subtypes previously identified in a  $590 \times 404 \times 160 \mu\text{m}^3$  volume. The red square represents the field of view seen in b-i. (k) Individual nTracer reconstructions based on molecular subtype. (l) Total path distance plotted for each of the molecular subtypes. (m)

Number of dendritic branches plotted for each of the molecular subtypes. Violin-plot: bold dashed line, median; dashed line, upper and lower quartiles. MIP, maximum intensity projection; PV, parvalbumin; Calb, calbindin; SOM, somatostatin. Scale bars: (**e,i**) 25  $\mu\text{m}$  (pre-expansion size). (**j,k**) 50  $\mu\text{m}$  (pre-expansion size).



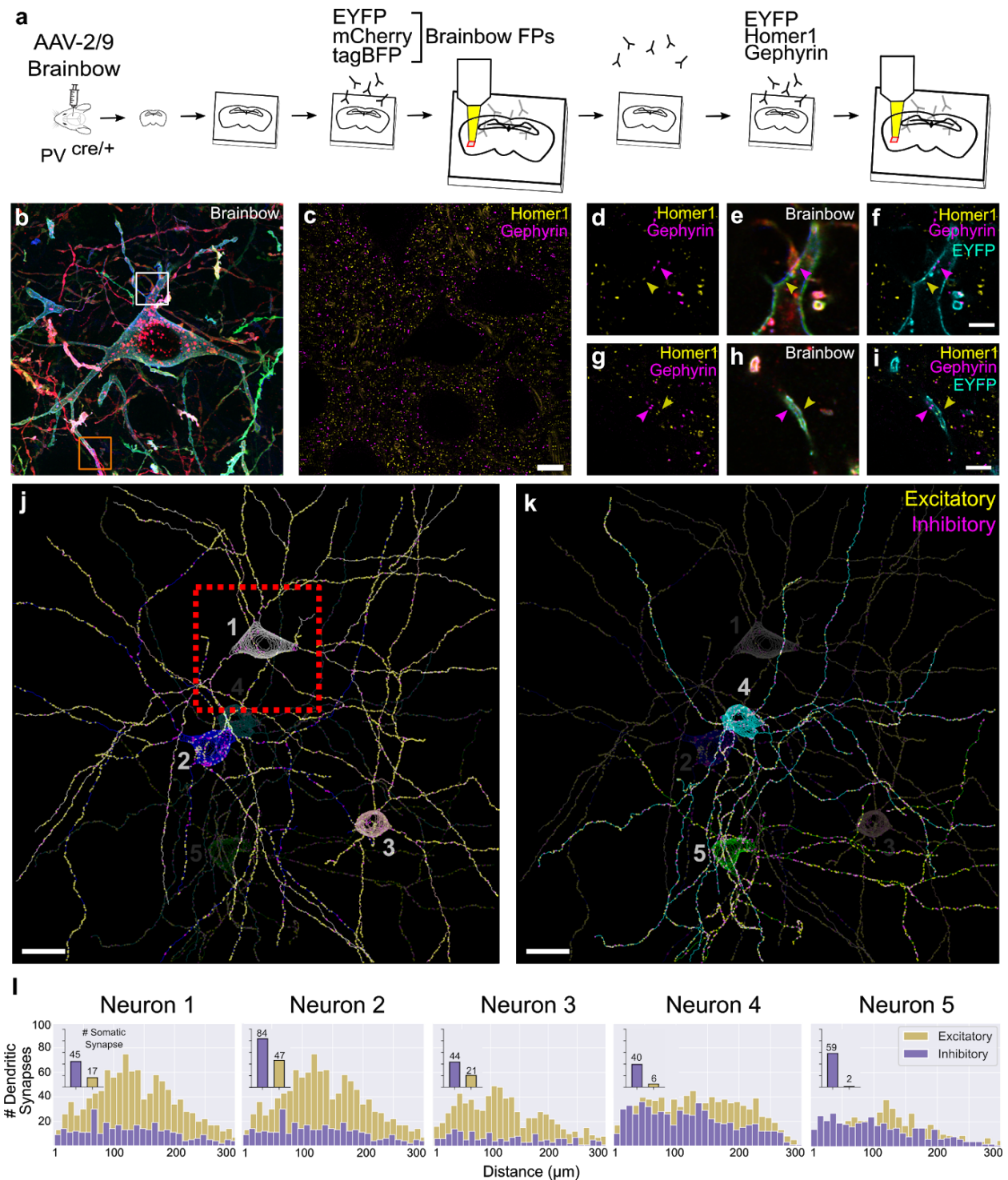
**Figure 4.3 Putative synaptic connections can be defined between Brainbow neurons.** (a) Experimental design: Brainbow FPs and endogenous synaptic markers are imaged across two rounds of immunostaining using the EYFP channel for registration. (b) MIP showing the Brainbow channels imaged in round 1. (c) MIP showing the Bassoon and Gephyrin puncta imaged in round 2. (d-d'') Single slice zoomed inset of red square in b. Example of two axo-somatic synapses where the orange axon contacts the blue soma. Two Gephyrin-Bassoon pairs (white arrows) are shown within the axonal bouton. (e-e'') Single slice zoomed inset of cyan square in b. Example of an axo-dendritic synapse where the red axon contacts the blue dendrite. White arrow points to the Gephyrin-Bassoon pair. (f-f'') Single slice zoomed inset of white square shown in b. Example of an axon that is physically close, but doesn't actually form a synapse. White arrow points to the Gephyrin-Bassoon pair. (g-g'') Single slice zoomed inset of yellow square in b. Example of an axo-somatic contact that lacks Gephyrin-Bassoon machinery (white arrow) and is not a true synapse. (h-k) Normalized line intensity profiles of dotted lines drawn in d-g respectively. Distances between Gephyrin and Bassoon peaks or between two Brainbow cell membranes were measured and shown

respectively for **h-j** or **k**. Scale bars: (**c**) 10  $\mu\text{m}$  (pre-expansion size). (**d-g**) 1  $\mu\text{m}$  (pre-expansion size).



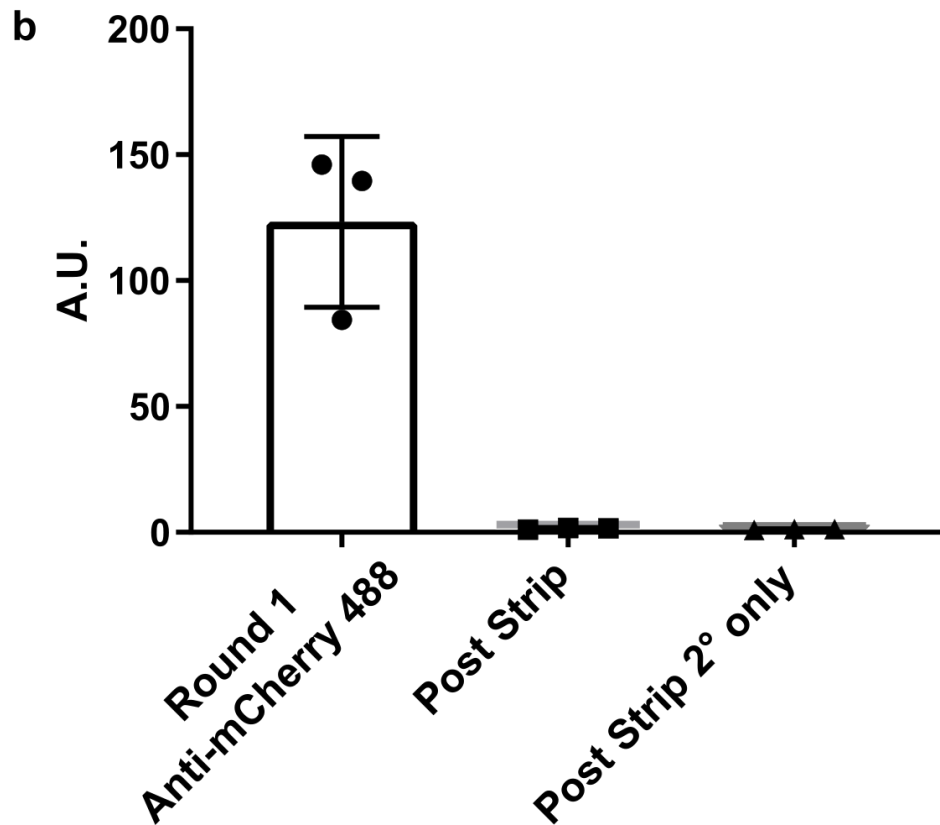
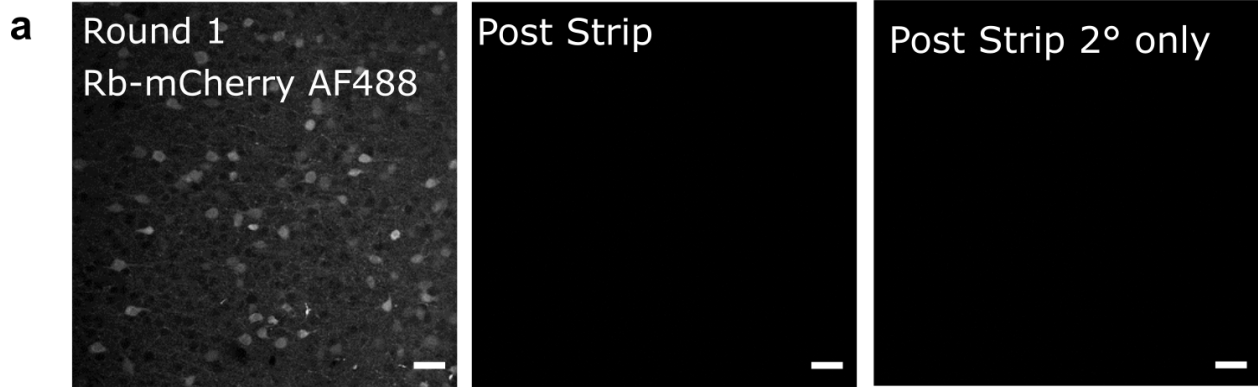
**Figure 4.4 Molecularly specific analyses can be performed on spectral connectomics data.** (a) MIP of 8 Brainbow labeled PV neurons reconstructed with nTracer. Thick neurites represent dendrites; thin neurites represent axons. The neuron IDs are overlaid. (b) Same as a, plus all of the other 189 innervating axons overlaid and 422 identified synapses marked with white circles. (c) Connectivity matrix between the 8 reconstructed PV neurons. (d) Connectivity matrix as in c, plus the 189 innervating axons. (e) Plot of the maximum and minimum number of unique presynaptic PV neurons that innervate each of the 8 postsynaptic PV neurons. Circles represent the number of spatially distinct PV axons that formed synapses, while triangles represent the number of unique colors that could be identified from the innervating axons. (f) Distribution of the number of soma targeting and dendrite targeting axons as a function of the number of synapses they form. Each of the 8 afferent PV neurons are color coded and plotted with the average represented by the bar graph. (g) Total number of somatic inhibitory synapses were plotted relative to the PV to PV subset for each of the 8 postsynaptic PV neurons. (h) Inhibitory PSD (Gephyrin) size plotted for all the PV to PV synapses found on each of the 8 postsynaptic PV neurons. Scale bars: 10  $\mu\text{m}$  (pre-expansion size).



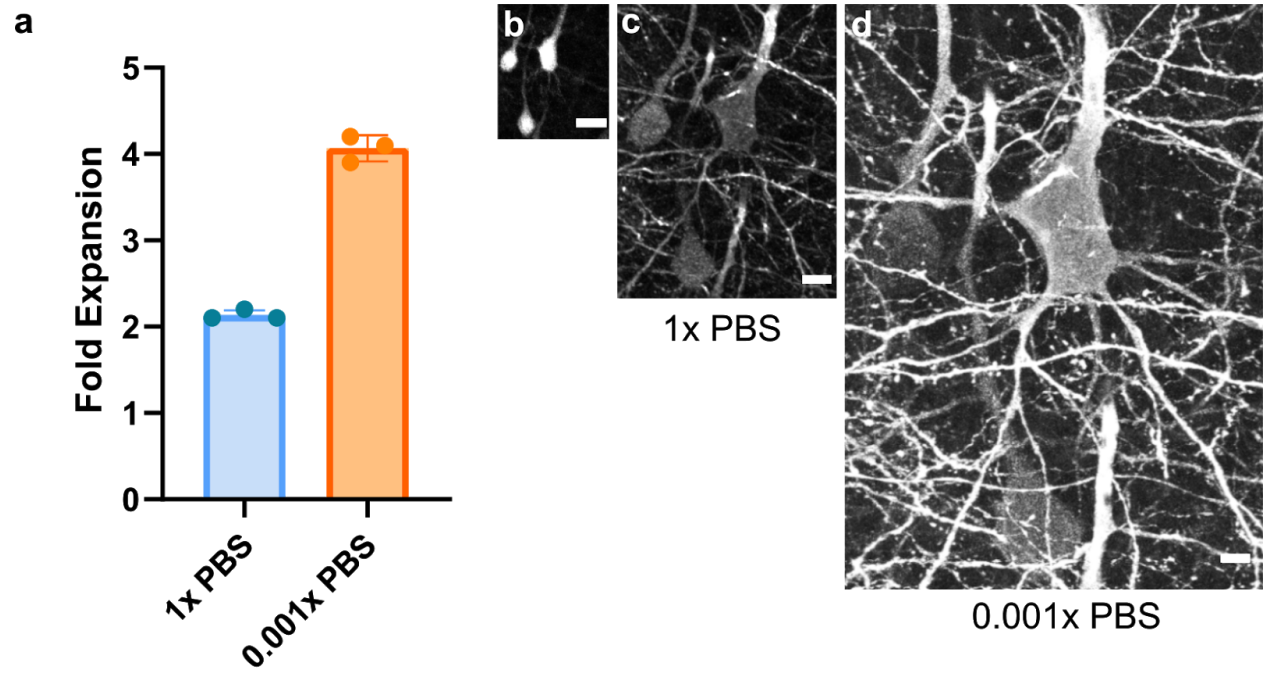


**Figure 4.5 Putative excitatory and inhibitory synapses of aspiny inhibitory neurons can be quantified.** (a) Experiment design: Brainbow and endogenous PSD markers are imaged across two rounds of immunostaining using the EYFP channel for registration. (b) MIP of the three Brainbow channels imaged in round 1. (c) Single slice of excitatory (Homer1) and inhibitory (Gephyrin) PSDs imaged in round 2. (d-i) Zoomed in synaptic marker + Brainbow single slice images of the white square (top row) and orange square (bottom row) respectively.

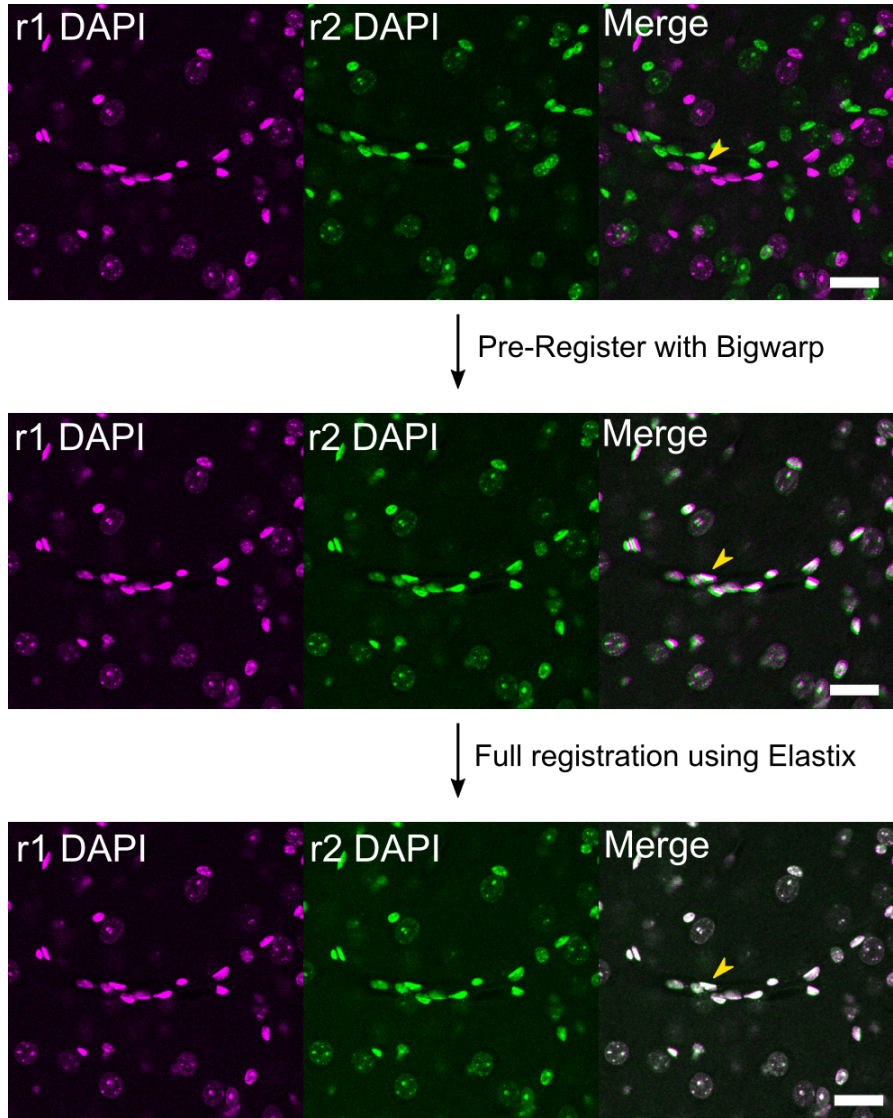
Yellow and purple arrows point to excitatory and inhibitory synaptic puncta respectively. **(j-k)** nTracer reconstruction of 5 neurons' morphology and putative synaptic inputs. The red box represents the field of view seen in **b** and **c**. **(l)** Histogram of the number of putative dendritic synapses as a function of dendritic length away from the soma for each neuron. The inset bar graph represents the number of putative somatic synapses for each neuron. Scale bars: **(c)** 10  $\mu\text{m}$  (pre-expansion size), **(f,i)** 2.5  $\mu\text{m}$  (pre-expansion size), **(j,k)** 20  $\mu\text{m}$  (pre-expansion size).



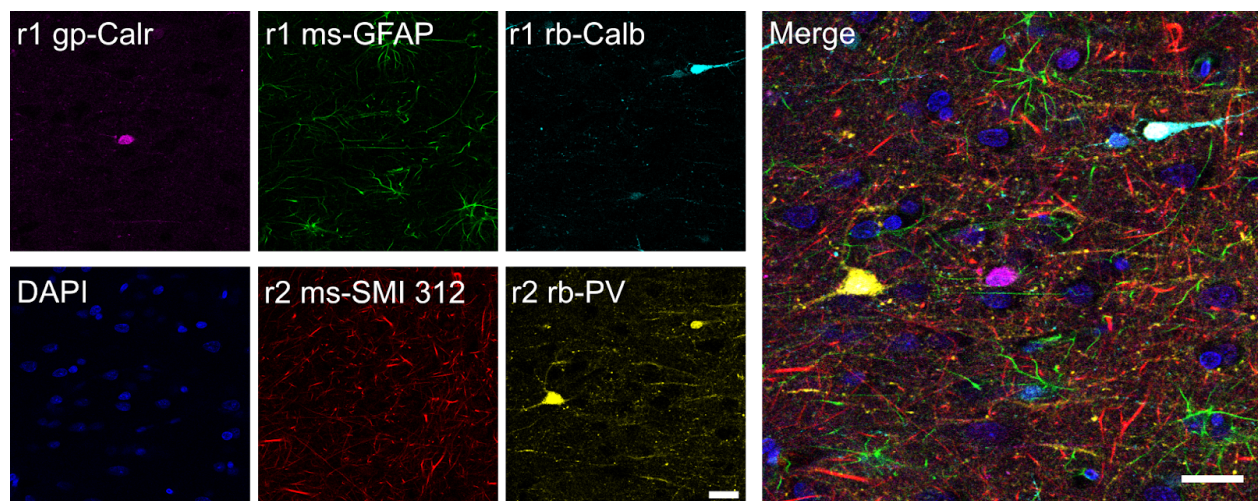
**Supplemental Figure 4.1 Heat and SDS are efficient at eluting antibody from miriEx gel-tissue hybrids.** (a) VGAT-Cre x Ai14 tissue sections were processed with miriEx and TdTomato+ neurons were immunostained with rb-mCherry antibody. The samples were then stripped using heat/SDS and re-imaged. The samples were then stained with secondary antibody only to check that the primary antibody was stripped off. (b) Quantification of fluorescence intensity seen in a (n=3 samples). Scale bars: (a) 50  $\mu$ m (pre-expansion size).



**Supplemental Figure 4.2 miriEx gel-tissue hybrid expansion.** (a) miriEx gels expand approximately ~2 fold in 1xPBS, and ~4 fold in 0.001xPBS (n=3 samples). (b-d) Same field of view imaged in three conditions: before miriEx processing, after miriEx processing in 1x PBS, and after miriEx processing in 0.001x PBS respectively. Scale bars: (b-d) 20um.

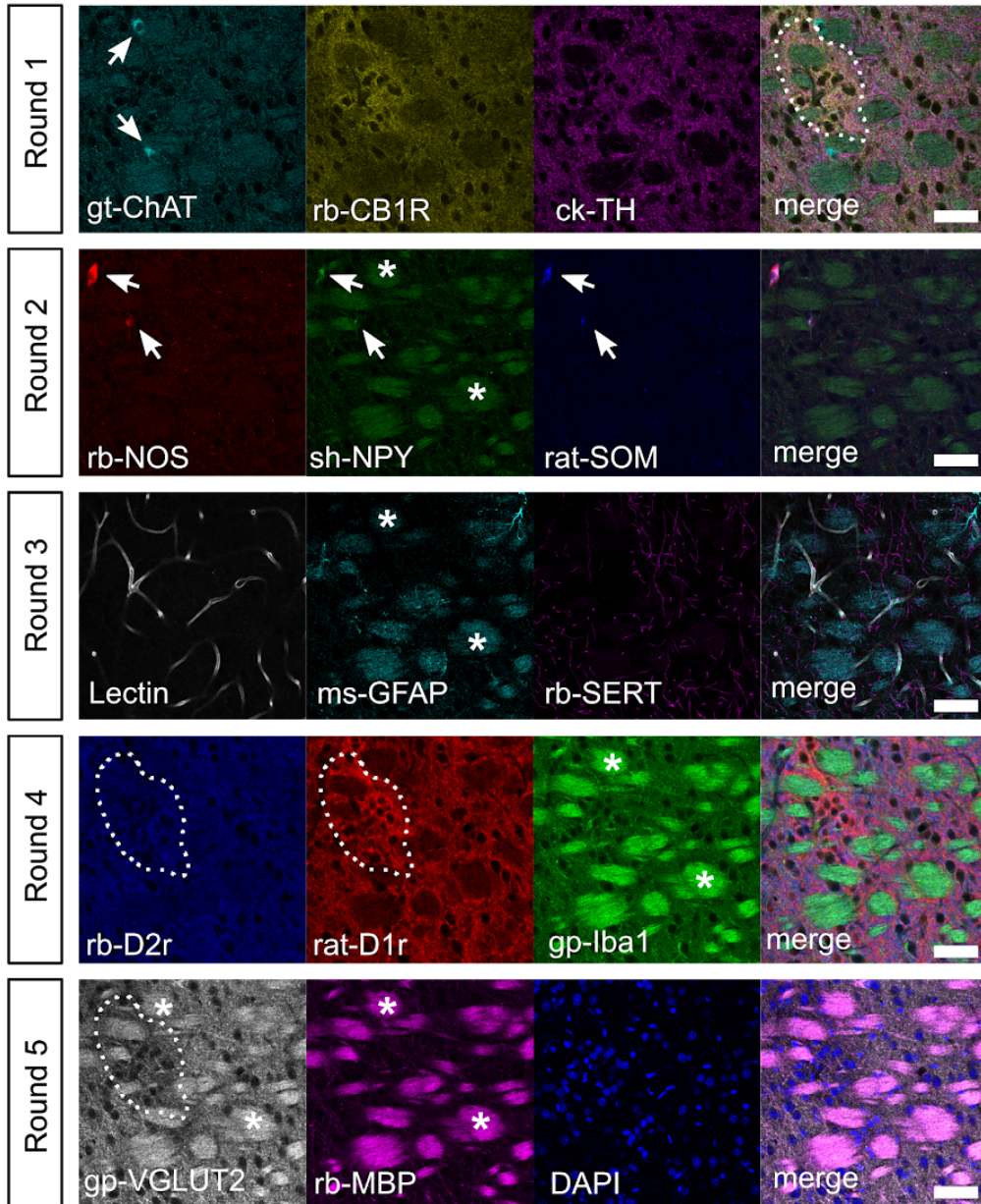


**Supplemental Figure 4.3 Image registration pipeline between different rounds of miriEx.**  
Scale bars: 25  $\mu\text{m}$  (pre-expansion size).

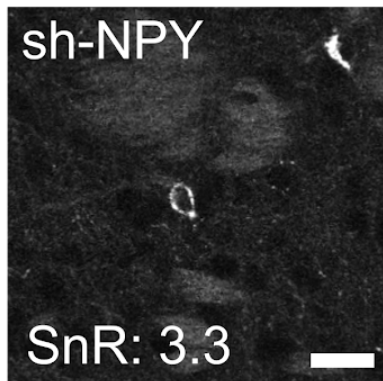


**Supplemental Figure 4.4 miriEx enables multiplexed immunostaining in formalin fixed human brain tissue.** Calr, Calretinin; PV, parvalbumin; Calb, calbindin; GFAP, glial fibrillary acidic protein. Scale bars: 25  $\mu\text{m}$  (pre-expansion size).

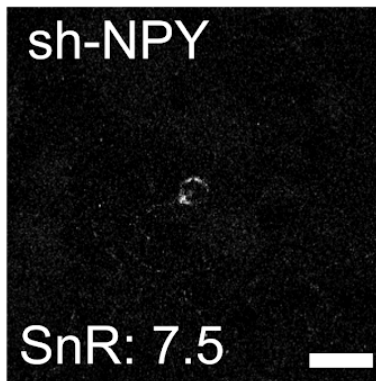
**a**



**b**

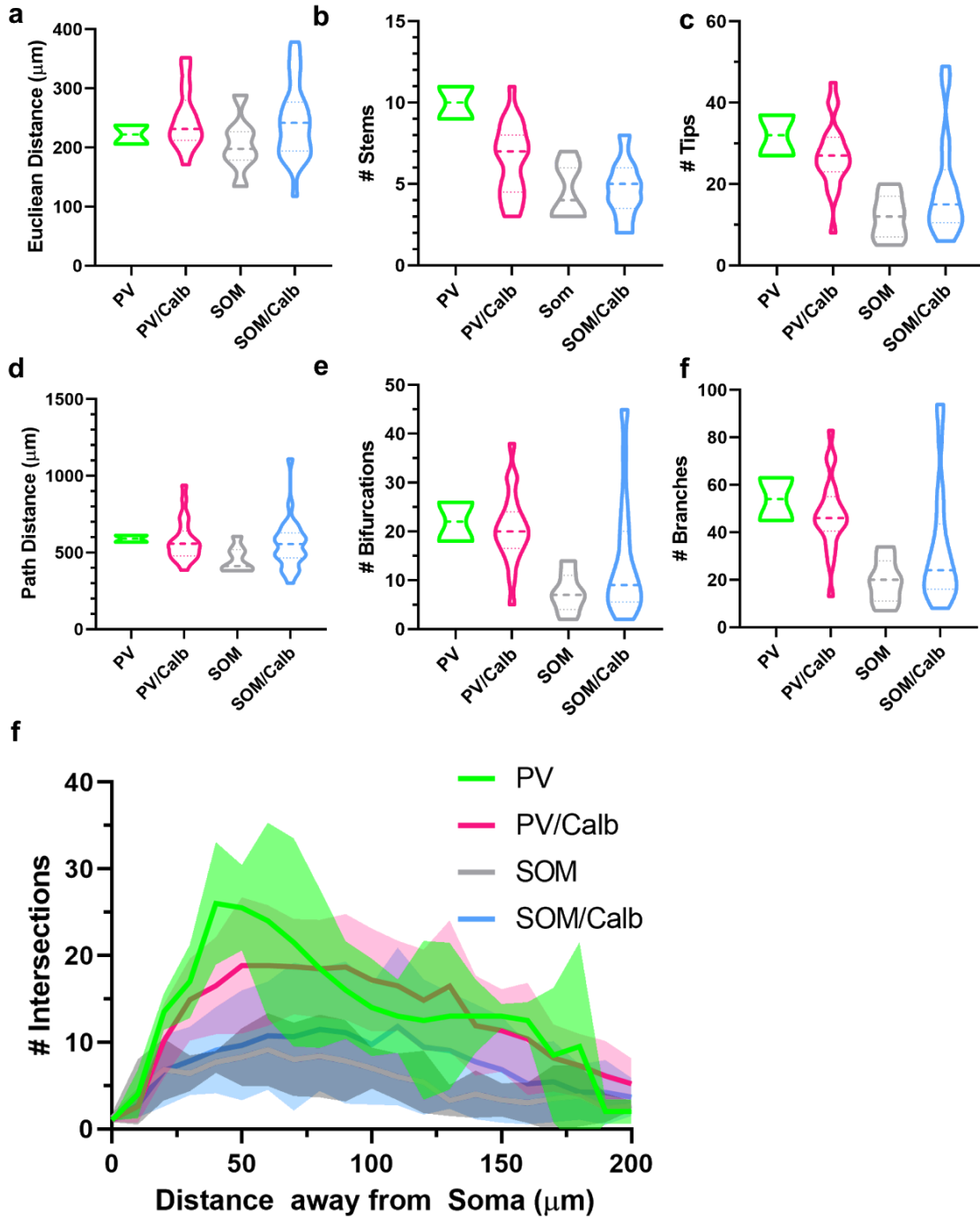


**c**

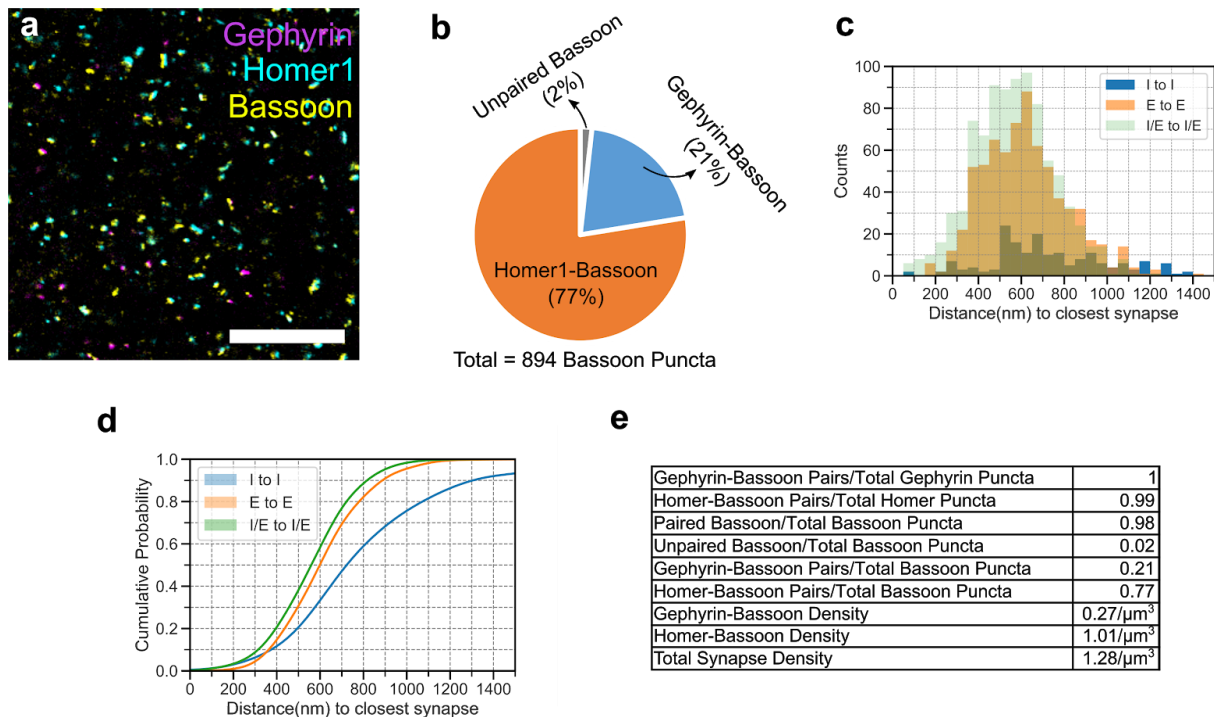


**Supplemental Figure 4.5 miriEx enables highly multiplexed molecular profiling.** (a) A total of 15 different antigens across 5 rounds were probed and imaged with confocal microscopy. DAPI was stained in each round to use as a fiduciary channel for registration. CB1R expression marks patches, a histochemically distinct compartment within the striatum (dashed white outline, round 1). NOS, NPY, and SOM expression largely overlap (round 2). D1r expression is higher in striatal patches, while D2r and VGLUT2 expression is higher in the surrounding matrix compartment (round 4). White arrows point to cell bodies with positive molecular marker expression. Asterisks point to myelin bundles immunostained by myelin binding protein (MBP, round 5), typically seen in coronal striatal sections. However, myelin bundles can also be non-specifically stained by secondary antibodies as high background in other rounds. (b) An example of NPY immunostaining with high background in striatum. (c) Demonstration that lowering the secondary antibody concentration two fold from 1:500 to 1:1000 decreases background and increases signal to noise (SnR). ChAT, choline acetyltransferase; CB1R, cannabinoid receptor type 1; TH, tyrosine hydroxylase; NOS, nitric oxide synthase; NPY, neuropeptide Y; SOM, somatostatin; GFAP, glial fibrillary acidic protein; SERT, serotonin transporter; D1r, dopamine receptor D1; D2r, dopamine receptor D2; Iba1, ionized calcium binding adaptor molecule 1; VGLUT2, vesicular glutamate transporter 2; MBP, myelin basic protein; Scale bars: (a) 50  $\mu\text{m}$  (pre-expansion size). (b-c) 25  $\mu\text{m}$  (pre-expansion size).

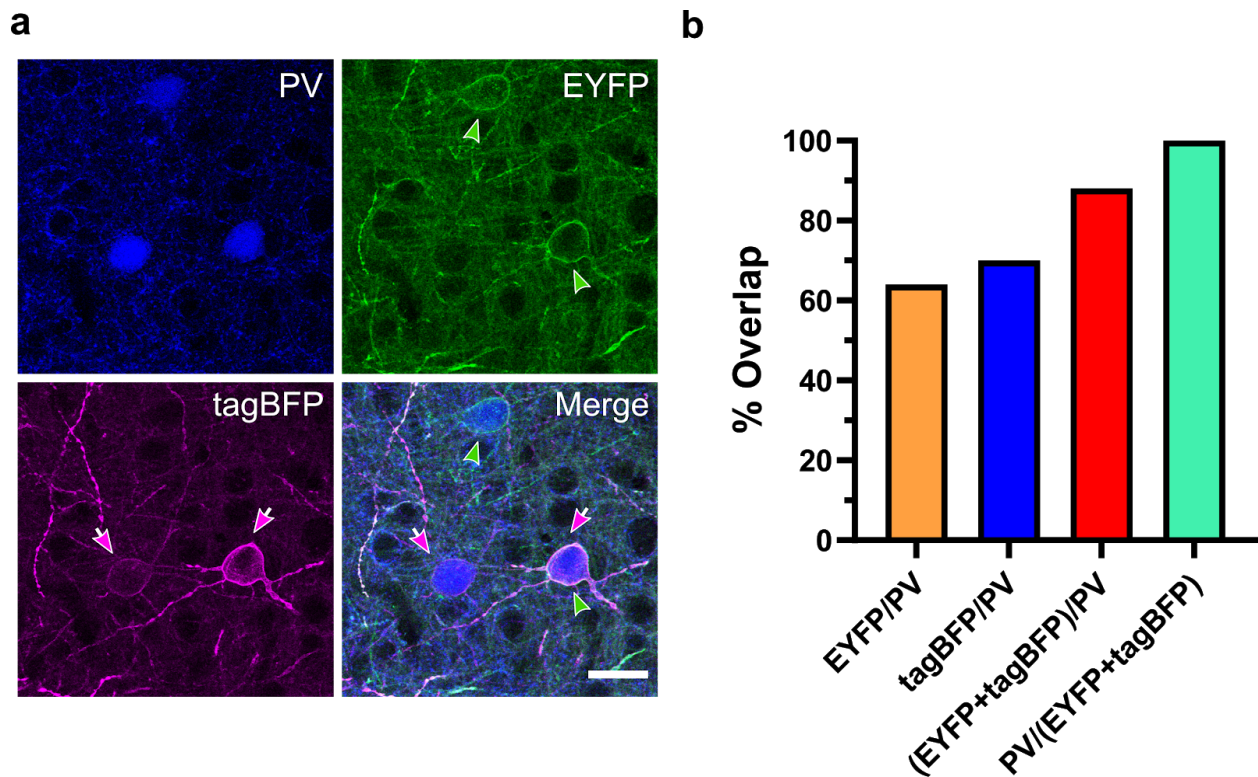




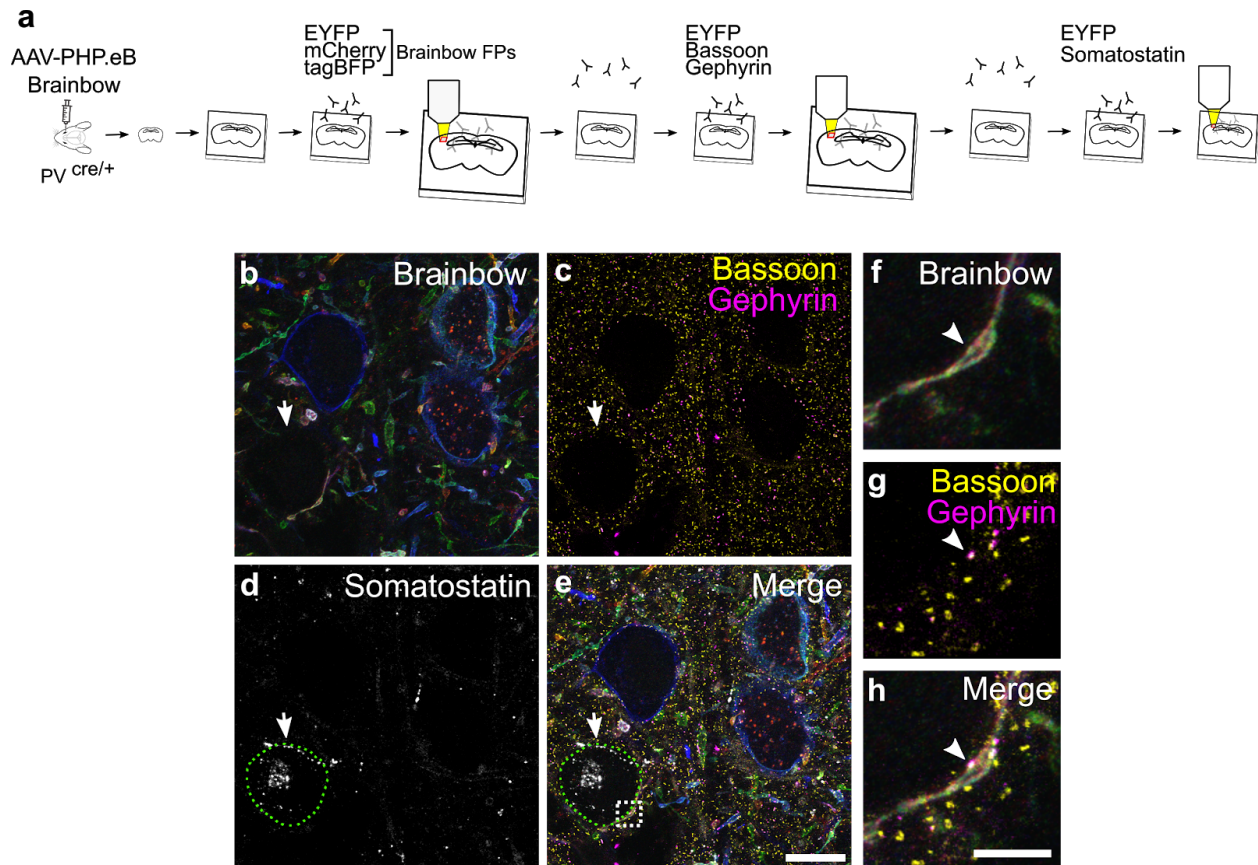
**Supplemental Figure 4.6 Morphology quantification of 53 inhibitory neurons traced in the BLA.** (a-f) Various morphometric parameters were calculated using Vaa3D global neuron feature for each of the 4 molecular subtypes. (g) 3D-Sholl analysis for each of the 4 molecular subtypes.



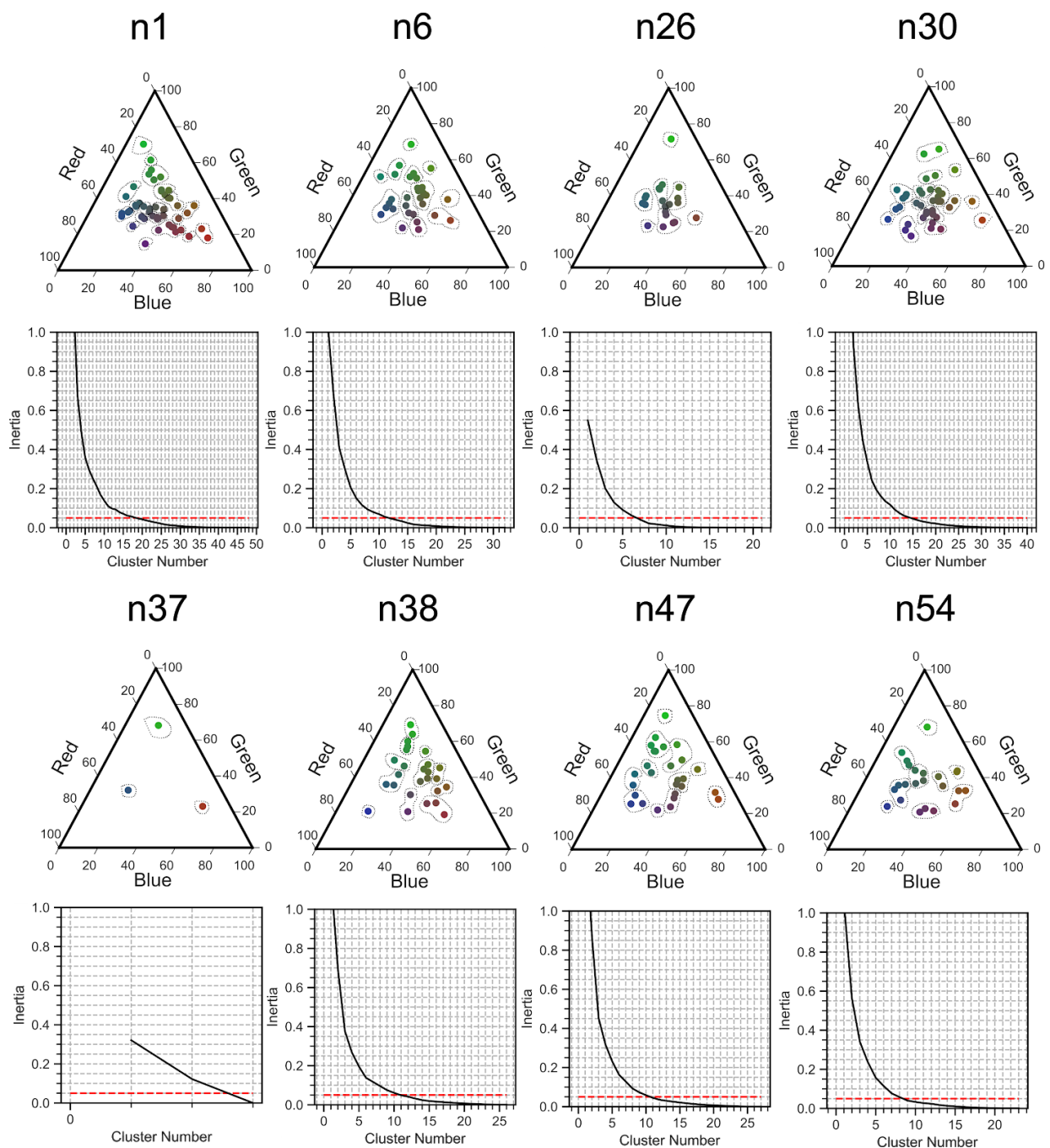
**Supplemental Figure 4.7 Homer1 and Gephyrin can be paired with Bassoon immunostaining to represent the majority of excitatory and inhibitory synapses.** (a) 100  $\mu\text{m}$  somatosensory cortex tissue was processed with miriEx and triple immunostained for Gephyrin, Homer1, and Bassoon. The sample was expanded  $\sim 4\text{x}$  and imaged. (b) All the synaptic pairs were manually annotated within a  $15 \times 15 \times 3 \mu\text{m}^3$  volume, yielding 706 excitatory synapses (Homer1-Bassoon) and 188 inhibitory synapses (Gephyrin-Bassoon). Pie chart shows that the majority of Bassoon puncta are paired with either Gephyrin or Homer1 in a mutually exclusive manner. (c) The distance between inhibitory (I) and excitatory (E) synapses to their nearest neighbor is plotted. (d) Cumulative probability plot of the histogram shown in c. (e) Table of other ratio and density measurements. Scale bars: 5 nm (pre-expansion size).



**Supplemental Figure 4.8 AAV-PHP.eB Brainbow expression in PV-Cre mice is highly sensitive and specific.** (a) 100  $\mu$ m somatosensory cortex tissue was processed with miriEx and triple immunostained for PV, EYFP, and tagBFP. (b) PV neurons across all 6 layers of somatosensory cortex were analyzed for co-expression of PV and 2 out of 4 Brainbow FPs. 88% of all PV neurons (n=50, identified through immunostaining) were positive for either EYFP or tagBFP. 100% of all Brainbow labeled neurons (n=44) were positive for PV immunostaining. PV, parvalbumin; FP, fluorescent protein. Scale bars: 50  $\mu$ m (pre-expansion size).



**Supplemental Figure 4.9 Brainbow FPs, synaptic machinery, and molecular markers can be interrogated in the same piece of tissue.** (a) Experimental design: Brainbow FPs, endogenous synaptic markers (Bassoon, Gephyrin), and cell type markers (SOM) are imaged across three rounds of immunostaining using the EYFP channel for registration. (b) ~750 nm MIP showing the Brainbow channels imaged in r1. (c) ~750 nm MIP showing the Bassoon and Gephyrin puncta imaged in r2. (d) ~750 nm MIP showing a Somatostatin positive soma. (e) Merged MIP of all three rounds. White arrow points to an “empty hole” that is later defined as a SOM positive soma outlined in a dashed green circle. (f-h) Single slice zoomed inset of white square shown in e; the arrow points to a putative synaptic connection between the PV axon and soma of the SOM neuron. FP, fluorescent protein; MIP, maximum intensity projection; SOM, somatostatin. Scale bars: (e) 10  $\mu$ m (pre-expansion size). (h) 3  $\mu$ m (pre-expansion size).



**Supplemental Figure 4.10 Number of unique axon colors can be identified for k-means clustering.** For each neuron, axon color identities were extracted through nTracer and plotted on a RGB color plot. An elbow plot was used to find the optimal number of k for k-means clustering. This value was conservatively chosen to be where inertia < 0.05. The dashed lines represent the k-means color cluster assignments.

## 4.8 Tables

Figure	Sample	[Aax]	Primary Ab	Primary incubation time and temperature	Secondary Ab	Secondary incubation time and temperature	Imaging	Processing: stitching, chromatic aberration, histogram matching
Fig 1	100 µm BLA from WT mouse	5 mM	r1: rb-PV (1:500) r2: rb-Calbindin (1:500) r3: rb-CB1R (1:333) r4: rb-PV (1:500) r5: rb-NOS (1:500) r6: rb-SERT (1:500) r7: rb-PV (1:500)	r1+r7: 1d at 37c	r1+r7: rb-AF488 (1:500), DAPI	r1+r7: 1d at 37c	r1+r7: Confocal, 10x Objective (0.4 NA), imaged in 1xPBS	Yes (3x2 tiles), No, No
Fig 2	200 µm BLA section from PV-Cre/Som-Cre stereotactically injected with AAV2/9 Brainbow in BLA	1 mM	r1: rb-Calbindin (1:500), rat-SOMec(1:100), gp-PV (1:500) r2: rb-mCherry (1:500), sh-GFP (1:500), gp-tagBFP (1:500)	r1: 3d at RT r2: 2d at 37c	r1: rat-AF488(1:500), rb-AF647 (1:500), gp-Cy3 (1:500) r2: rat-AF488(1:500), rb-AF647 (1:500), gp-Cy3 (1:500)	r1: 2d at RT r2: 2d at 37c	r1: Confocal, 10x objective (0.4 NA), imaged in 1x PBS r2: Confocal, 20x Objective (1.0 NA), imaged in 1x PBS	Yes (3x2 tiles), Yes, Yes
Fig 3, sFig1	100 µm S1 Cortex section from PV-Cre retro-orbitally injected with PIP-eB Brainbow	1 mM	r1: rb-mCherry (1:500), sh-GFP(1:500), gp-tagBFP (1:500) r2: sh-GFP (1:500), gp-Bassoon (1:500), ms-GephyrinSC (1:100) r3: rat-SOMec (1:100), sh-GFP (1:500)	r1: 2d at RT r2: 2d at 37c r3: 2d at 37c	r1: rb-AF488 (1:500), sh-Cy3 (1:500), gp-AF647 (1:500) r2: ms-AF488 (1:500), sh-Cy3 (1:500), gp-AF647 (1:500) r3: sh-Cy3 (1:500), rat-AF647 (1:500)	r1: 1d at RT r2: 1d at 37c r3: 1d at 37c	r1: confocal, 20x objective (1.0 NA), imaged in 0.01x PBS r2: confocal, 20x objective (1.0 NA), imaged in 0.01x PBS r3: confocal, 20x objective (1.0 NA), imaged in 1x PBS	No, Yes, Yes
Fig 5	100 µm S1 cortex section from PV-Cre stereotactically injected with AAV2/9 Brainbow in BLA	1 mM	r1: rb-mCherry (1:500), sh-GFP (1:500), gp-tagBFP (1:500) r2: gp-Bassoon (1:500), ms-GephyrinSC (1:100), sh-GFP (1:500)	r1: 2d at 37c r2: 1d at 37c	r1: gp-AF647 (1:500), ms-AF488 (1:500), sh-Cy3 (1:500) r2: rb-AF488 (1:500), gp-AF647 (1:500), sh-Cy3 (1:500)	r1+r2: 1d at 37c	r1: Confocal, 20x Objective (1.0 NA), imaged in 0.01x PBS r2: Confocal, 20x Objective (1.0 NA), imaged in 0.001x PBS	Yes (2x2 tile), Yes, Yes
sFig 1	100 µm VGAT-Cre x Ai14 S1 cortex	1 mM	r1: rb-mCherry (1:500)	r1: 1d at 37c	r1: rb-AF488 (1:500) post strip: rb-AF647 (1:500)	r1/post strip: 1d at 37c	Confocal, 10x objective(0.4 NA) , imaged in 1x PBS	No, No, No
sFig 2	100 µm Thy1-YFP S1 cortex	1 mM	sh-GFP (1:500)	1 day at 37c	sh-AF594 (1:500)	1d at 37c	Confocal, 10x Objective (0.4 NA), imaged in Vectashield, 1xPBS, 0.001x PBS	No, No, No
sFig 4	100 µm human sensory cortex	5 mM	r1: rb-Calbindin (1:500), gp-Calretinin (1:500), ms-GFAP (1:500) r2: ms-SM312 (1:333), rb-PV(1:500)	r1+r2: 1d at 37c	r1: rb-Cy3 (1:500), gp-AF488 (1:500), ms-AF647 (1:500), DAPI r2: ms-AF488 (1:500), rb-AF647 (1:500), DAPI	r1+r2: 1d at 37c	r1+r2: Confocal, 10x objective (0.4 NA), imaged in 1xPBS	Yes (2x1 tile), No, No
sFig 5	100 µm WT dorsolateral striatum	5mM	r1: gp-ChAT (1:500), rb-CB1R(1:500), ck-TH (1:500) r2: rb-NOS (1:500), sh-NPY(1:500), rat-SOM (1:100) r3: Lectin (1:1000), ms-GFAP (1:500), rb-SERT (1:500) r4: rb-D2R (1:500), rat-D1R(1:500), gp-Iba1(1:300) r5: gp-VGLUT2 (1:500), rb-MBP (1:500)	r1+r5: 1d at 37c	r1: ck-AF488 (1:500), rb-Cy3 (1:500), gp-AF647 (1:500), DAPI r2: sh-AF488 (1:500), rb-Cy3 (1:500), rat-AF647 (1:500), DAPI r3: ms-AF488 (1:500), rb-AF647 (1:500), DAPI r4: gp-AF488 (1:500), rb-Cy3 (1:500), rat-AF647 (1:500), DAPI r5: gp-AF488 (1:500), rb-Cy3 (1:500)	r1+r5: 1d at 37c	r1+r5: Confocal, 10x objective (NA 0.4), imaged in 1x PBS	Yes (2x2 tile), No, No
sFig 7	100 µm WT mouse	1 mM	rb-Homer1a (1:500), ms-GephyrinSC (1:100), gp-Bassoon (1:500)	2d at 37c	ms-AF488 (1:500), rb-Cy3 (1:500), gp-AF647 (1:500)	1d at 37c	Confocal, 20x objective (1.0 NA), imaged in 0.001x PBS	No, Yes, Yes
sFig 8	100 µm S1 Cortex section from PV-Cre retro-orbitally injected with PIP-eB Brainbow (same brain as Fig.3)	1 mM	gp-tagBFP(1:500), rb-PV (1:500), sh-GFP (1:500)	1d at 37c	gp-AF647 (1:500), sh-Cy3 (1:500), rb-AF488 (1:500)	1d at 37c	Confocal, 10x objective (0.4 NA), imaged in 1xPBS	Yes (2x2 tile), Yes, Yes

**Table 4.1. Experimental Conditions**

Target	Vendor	Catalog Number	Host Species
VGAT	Synaptic Systems	131 011	Ms
Calbindin	Synaptic Systems	214 002	Rb
PV	Synaptic Systems	195 004	Gp
PV	Abcam	ab32895	Gt
Homer1	Synaptic Systems	160 002	Rb
Gephyrin	Synaptic Systems	147 111	Ms
Bassoon	Synaptic Systems	141 004	Gp
CamKII	Abcam	ab22609	Ms
TH	Abcam	ab76442	Ck
NOS	Sigma	n2780	Rb
NPY	Millipore	ab1583	Sh
Somatostatin	Millipore	mab354	Rat
VIP	Immunostar	20077	Rb
Calretinin	Synaptic Systems	214 104	Gp
SMI-312	Biologend	837904	Ms
SERT	Synaptic Systems	340 003	Rb
CB1R	Synaptic Systems	258 003	Rb
D2R	Synaptic Systems	376 203	Rb
D1R	Sigma	D2944	Rat
GFAP	Dako Agilent	Z033401	Rb
Vglut2	Synaptic Systems	135 404	Gp
IBA1	Synaptic Systems	234 004	Gp
MBP	Synaptic Systems	295 002	Rb
Lectin	Vector Labs	DL-1177	
GAD67	Millipore	mab5406	Ms
SomatostatinSC	Santa Cruz	YC7	Rat
GephyrinSC	Santa Cruz	G6	Ms
PV	Abcam	ab11427	Rb
NeuN	Millipore	mab377	Ms
Homer1	Synaptic Systems	160 006	Ck
GFAP	Sigma	G3893	Ms
CTIP2	Abcam	ab18465	Rat
tagBFP	Cai Lab custom made	NA	Gp
mCherry	Cai Lab custom made	NA	Rb
GFP	Biorad	4745-1051	Sh

**Table 4.2. Primary Antibodies**

<b>Target</b>	<b>Vendor</b>	<b>Catalog Number</b>	<b>Host Species</b>
Rb-AF488	Jackson ImmunoResearch	711-545-152	Dk
Rb-Cy3	Jackson ImmunoResearch	711-166-152	Dk
Rb-AF647	Jackson ImmunoResearch	711-605-152	Dk
Gp-AF488	Jackson ImmunoResearch	706-545-148	Dk
Gp-Cy3	Jackson ImmunoResearch	706-166-148	Dk
Gp-AF647	Jackson ImmunoResearch	706-606-148	Dk
Sh-AF488	Jackson ImmunoResearch	713-546-147	Dk
Sh-Cy3	Jackson ImmunoResearch	713-166-147	Dk
Rat-AF488	Jackson ImmunoResearch	712-545-150	Dk
Rat-AF647	Jackson ImmunoResearch	712-606-153	Dk
Ck-AF488	Jackson ImmunoResearch	703-546-155	Dk
Ck-AF647	Jackson ImmunoResearch	703-606-155	Dk
Ms-AF488	Jackson ImmunoResearch	715-546-151	Dk
Ms-Cy3	Jackson ImmunoResearch	715-166-151	Dk
Ms-AF647	Jackson ImmunoResearch	715-606-151	Dk

**Table 4.3. Secondary Antibodies**



# Chapter 5. Discussion and Future Directions

## 5.1 Summary of contributions

In this dissertation, we developed new light microscopy based approaches for microscale and nanoscale neurocartography. Chapter 2 describes the optimization and development of an expansion microscopy protocol called miriEx. Chapter 3 describes the rationale and design of 6 new Brainbow viruses expressing 12 different fluorescent proteins. Chapter 4 describes how we integrated Brainbow and miriEx tools to enable mapping of synaptic connectivity in a molecularly specific population of neurons. This chapter describes extensions and future directions of my work.

## 5.2. Platform for high throughput microscale neurocartography

### 5.2.1 Motivation

Single cell technologies have revolutionized our understanding of the brain. For example, single cell RNA sequencing has deepened our understanding of neural cell types because of two big advantages - single neuron resolution, and high throughput capability. Over the last 10 years, neuroscientists have probed the transcriptomes of thousands of cells, generating more data in a single experiment than decades of *in situ* RNA experiments that came before. The key to single cell RNA sequencing's maturity and success is 1) high throughput sampling of single cells with enough sensitivity (~20% mRNA capture efficiency<sup>113</sup>) to extract biological insights, 2) established computational analysis pipelines, and 3) open access to shared datasets. The popularity of the commercial 10x Genomics Chromium platform points to the desire of neurobiology labs to generate high throughput, single cell sequencing data in an easy and straightforward way.

Methods of studying single neuron morphologies, comparatively, have lagged behind. The biggest challenge has been that most methods rely on sparse labeling of neurons in order to unambiguously reconstruct their structures, restricting the throughput to just a few neurons per sample. For example, in recent patch-seq experiments performed by Gouwens and colleagues to obtain multimodal profiles of neurons, morphology reconstructions were the most difficult to obtain (350 morphology profiles relative to ~3700 transcriptomic and electrophysiological profiles)<sup>102</sup>. In addition, only recently has technology progressed to the point that neuroscientists at the Allen Institute and Janelia Farms have been able to map the whole brain morphology of hundreds to thousands of single projection neurons<sup>114–116</sup>. Although an impressive technical achievement, only tens of neurons were reconstructed from each brain, meaning that multiple brains were imaged, reconstructed, and pooled for subsequent analysis. A high throughput platform would greatly accelerate this process and amplify the amount of biological insights gained. An analogy is how in the early days of single cell RNA sequencing, neurons were manually picked into wells for subsequent processing, limiting the throughput to hundreds of cells. High throughput methods, like Drop-Seq<sup>117</sup> and 10x Chromium, completely changed the landscape by allowing users to profile up to two orders of magnitude more cells. The benefits were cheaper experiments, increased statistical power, and higher coverage of rare cell types. Many of the same benefits would also apply to high throughput single neuron morphology studies. Profiling more neurons will lead to more accurate morphology cell type clustering and reveal rare subtypes. It can uncover anatomical relationships between neurons and be combined with perturbations to study how neuropsychiatric disease models impact neuroanatomy. Furthermore, increased morphology profiles will inform biophysical computational models of neurons. Below I will discuss how future work can incorporate the tools I have introduced in the previous chapters into a robust platform for high throughput microscale neurocartography.

### 5.2.2 Specimen preparation

We propose that our newly developed Brainbow AAVs could be applied to study interneuron morphologies across the brain. This system is more tractable because interneurons have local, smaller innervation territories compared to long range projection neurons. In addition, smaller imaging volumes are required, generating manageable dataset sizes on a gigabyte scale that could be processed without using a separate set of tools to handle big imaging data on a

terabyte scale. These tractable datasets could be used to build and optimize the computational pipeline. Once the workflow is established, everything could be scaled up to map larger imaging volumes containing projection neurons. miriEx can be used to perform multiplexed immunostaining in thick brain sections to readout all 12 fluorescent proteins. However, antibody choice and order would have to be optimized along with acryl NHS-ester concentration to maximize signal across many rounds of immunostaining.

### 5.2.3 Light imaging

We propose using lightsheet microscopy to image expanded Brainbow brain sections. The two big advantages over confocal microscopy are faster imaging speeds and decreased photobleaching, which are essential when imaging large volumes in millimeter thick fluorescent brain sections. Gao and colleagues demonstrated the power of using a lattice lightsheet microscope to map a mouse cortical column at nanoscale resolution<sup>16</sup>. Compared to other tissue clearing protocols that rely on refractive index matching using potentially harsh reagents to prepare the sample for imaging<sup>27</sup>, expanded brain sections are more transparent and can be directly mounted in aqueous solutions. In addition, long working distance objectives are necessary to cover the depth of the tissue, and typically have lower numerical apertures and optical resolutions, especially axially. In fact, the majority of thick tissue imaging studies use optical imaging resolutions greater than 1  $\mu\text{m}$ . Expansion microscopy mitigates the trade off between thick tissue imaging and optical resolution by boosting the effective imaging resolution of lower numerical aperture objectives. Consequently, the combination of lightsheet microscopy and expansion microscopy offers a nice balance between high imaging resolution and fast imaging speeds. The exact choice of imaging resolution would have to be optimized. As discussed at the end of Chapter 1, the color diversity and density of neurons labeled will influence the necessary imaging resolution for reconstructing neurons.

### 5.2.4 Image processing

See Chapter 3.3.5 for detailed discussion of proposed image processing steps.

### 5.2.5 Adopting automatic segmentation algorithms to reduce manual tracing

The reconstruction of Brainbow neurons has become a bottleneck for high throughput studies because current methods, like nTracer, involve laborious human inputs. *The most pressing need is for reliable automatic segmentation algorithms for multispectral Brainbow data.*

The EM connectomics field recognized manual tracing as an important obstacle, and developed several solutions that we can borrow and take inspiration from. Machine-learning approaches have advanced in accuracy to the point that they greatly speed up the reconstruction process. For example, Januszewski and colleagues developed a flood filling approach that uses convolutional neural networks to accurately segment EM data and obtained an average error free neurite length of 1.1 mm<sup>106</sup>. The convergent strategy used by many groups is to shift the burden of tracing neurites to machine-learning based segmentation algorithms, and have human annotators focus instead on proofreading and correcting errors<sup>77</sup>. Doing so improves the speed of reconstructing neurons ten fold. Similarly, we should seek to adapt these algorithms to Brainbow light microscopy data so neurites can be segmented and joined into a complete reconstruction. nTracer will still be useful for manually tracing portions of the data to be used as ground truth for classifier training.

### 5.2.6 ‘Traceless’ mapping of axon projections at single neuron resolution

An alternative strategy to reduce the burden of manual reconstruction is to take advantage of the vast amount of Brainbow color barcodes. Given a scenario where the number of color barcodes is significantly greater than the number of neurons labeled, we would have statistical confidence that most of the neurons are uniquely labeled with low collision frequency. If true, we could look at the projection fields and quantify the number of unique axonal color barcodes to assess how many different neurons innervate the target area. For example, we could label dopamine neurons in the midbrain (~10,000/hemisphere<sup>118</sup>) with Brainbow and map their axon projections in different striatal regions. Previous approaches in literature involved reconstructing the entire axonal arbor of sparsely labeled dopaminergic neurons<sup>119</sup>. Our proposed *traceless* approach would instead focus on measuring the axonal color barcode distribution in the striatum and the somatic color barcode distribution in the midbrain. Assuming most of the barcodes are

unique, we would be able to assign each axonal projection pattern to its parent neuron and avoid manual tracing.

### 5.2.7 Morphometric analysis

Different groups have established ~100 different morphological parameters that can be extracted from the structure of neurons<sup>120</sup>. These include soma shape, dendrite length, axon branch number, etc., and can be used to compare neurons from different conditions, as well as classify them into morphological cell types. For example, morphological parameters can be one phenotype of how neurons structurally change in neuropsychiatric disease models. They can also be pooled across multiple neurons and used to define morphological cell types similar to how single cell RNA sequencing methods use dimensionality reduction methods to cluster and define transcriptomic cell types<sup>114</sup>. In addition, the acquisition of multiple single neuron morphologies from a densely labeled imaging volume offers an opportunity to define new parameters that quantify the spatial interactions between them. For instance, the degree of overlap between neighboring neurons' innervation territories can be measured. The spatial interactions of densely labeled neurons helps define the network anatomy of the broader circuit. Being able to measure these spatial interactions in a high throughput manner would yield new insights into neural cell types and connectivity principles.

### 5.2.8 Open access and well annotated pipelines

It will be important to be able to share the raw data, processed data, and morphology reconstructions with the neuroscience community so that others can validate it and contribute new findings. Open access to our data will be important for groups that lack advanced lightsheet microscopes and/or technical expertise, but still want to study cell type specific single neuron morphologies. The Allen Institute is a good example of how building dedicated web interfaces to disseminate data can amplify the impact of the research on the community. For example, researchers have used and cited the Allen *in situ* atlas of gene expression over 3600 times<sup>121</sup>. In addition, all developed software should be well documented and publically available. The hope is that transparent methods would foster the adoption of our platform in other labs.

### 5.2.9 Biological applications

Given that we successfully establish a high throughput platform for mapping single neuron morphologies, what kind of biological questions can be addressed? An atlas of single neuron morphologies could be created to classify neurons into cell types based on shape. Sex and mouse strain specific neuron morphology differences could also be quantified and included. An open question is how well do morphological definitions of cell type align with transcriptomic definitions? Reconciling the two will be important in our endeavor to create an accurate “periodic table of cell types” for the brain. In addition, high throughput studies of morphology could be combined with behavioral screens to study individual variation. The ability to study large numbers of neurons from a single animal would allow us to correlate neuroanatomy and behavioral individuality. For instance, Linneweber and colleagues recently demonstrated that innate, stochastic differences in the wiring of the visual system led to variable line walking behavior in flies<sup>122</sup>.

## 5.3 Spectral connectomics can be applied in a bottom up approach to understand higher order brain functions

In Chapter 4, we refer to the spectral connectome as a light microscopy derived map of neuron morphologies, molecular identities, and synaptic connectivity. Spectral connectomics can be used to iteratively understand the cell types in a given brain region, the circuits they form, and their functional properties in relation to behavior. We can consider an example brain region for how this iterative mapping can be performed. The first step would be leveraging single cell RNA sequencing to map the transcriptomic cell types within the region. Once we understood the main molecular cell types, we could identify which subsets to study using either specific Cre mice and/or recently developed Cre AAVs that use special enhancers and promoters to drive cell type specific expression of Brainbow AAVs<sup>123</sup>. If a broad labeling strategy is used to express Brainbow in a variety of cell types, individual molecular subtypes could be identified by probing molecular markers at either the protein or RNA level. Developing new expansion microscopy methods to combine Brainbow fluorescent protein immunostaining with *in situ* hybridization to probe RNA will be tremendously useful for this type of analysis. Next, the connectivity between different neural subtypes could be mapped by probing for endogenous synaptic machinery. This

process can be augmented by future computational methods that segment neuron morphologies, synapses, and automatically identify connected partners. The overall goal is to obtain a complete structural map of all the different molecular subtypes within the brain region at both mesoscale and nanoscale resolutions. Such a map will reveal connectivity motifs and how molecular subtypes are wired together into circuits. Furthermore, physiological properties could be layered on top of this map by using *in vivo* calcium or voltage imaging of cell type specific populations during behavior. The full potential of spectral connectomics is achieved when combined with single cell RNA sequencing and *in vivo* recording of neural activity to precisely define each neuron's morphology, connectivity, transcriptome, and functional profile in a multimodal fashion. This type of bottom up approach, where we fully describe the myriad properties of single neurons, will inform our understanding of higher order processes like sensory processing, attention, memory, and consciousness.

## References

1. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **314**, 1–340 (1986).
2. Herculano-Houzel, S. The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. *Front. Hum. Neurosci.* **3**, (2009).
3. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
4. Hintiryan, H. *et al.* Connectivity characterization of the mouse basolateral amygdalar complex. <http://biorxiv.org/lookup/doi/10.1101/807743> (2019) doi:10.1101/807743.
5. Zingg, B. *et al.* Neural networks of the mouse neocortex. *Cell* **156**, 1096–1111 (2014).
6. Petreanu, L., Huber, D., Sobczyk, A. & Svoboda, K. Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. *Nat. Neurosci.* **10**, 663–668 (2007).
7. Beier, K. T. *et al.* Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. *Cell* **162**, 622–634 (2015).
8. Wall, N. R., De La Parra, M., Callaway, E. M. & Kreitzer, A. C. Differential Innervation of Direct- and Indirect-Pathway Striatal Projection Neurons. *Neuron* **79**, 347–360 (2013).
9. Penzes, P., Cahill, M. E., Jones, K. A., VanLeeuwen, J.-E. & Woolfrey, K. M. Dendritic spine pathology in neuropsychiatric disorders. *Nat. Neurosci.* **14**, 285–293 (2011).
10. Lozovaya, N. *et al.* GABAergic inhibition in dual-transmission cholinergic and GABAergic striatal interneurons is abolished in Parkinson disease. *Nat. Commun.* **9**, 1–14 (2018).
11. Rees, C. L., Moradi, K. & Ascoli, G. A. Weighing the Evidence in Peters' Rule: Does Neuronal Morphology Predict Connectivity? *Trends Neurosci.* **40**, 63–71 (2017).
12. Chéreau, R., Saraceno, G. E., Angibaud, J., Cattaert, D. & Nägerl, U. V. Superresolution imaging reveals activity-dependent plasticity of axon morphology linked to changes in action potential conduction velocity. *Proc. Natl. Acad. Sci.* **114**, 1401–1406 (2017).
13. Kasthuri, N. *et al.* Saturated Reconstruction of a Volume of Neocortex. *Cell* **162**, 648–661 (2015).



14. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796 (2006).
15. Huang, B., Bates, M. & Zhuang, X. Super resolution fluorescence microscopy. *Annu. Rev. Biochem.* **78**, 993–1016 (2009).
16. Gao, R. *et al.* Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution. *Science* **363**, eaau8302 (2019).
17. Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* **347**, 543–548 (2015).
18. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56–62 (2007).
19. He, M. *et al.* Strategies and tools for combinatorial targeting of GABAergic neurons in mouse cerebral cortex. *Neuron* **91**, 1228–1243 (2016).
20. Zeng, H. Mesoscale Connectomics. *Curr. Opin. Neurobiol.* **50**, 154–162 (2018).
21. Bloss, E. B. & Hunt, D. L. Revealing the Synaptic Hodology of Mammalian Neural Circuits With Multiscale Neurocartography. *Front. Neuroinformatics* **13**, (2019).
22. Richardson, D. S. & Lichtman, J. W. Clarifying Tissue Clearing. *Cell* **162**, 246–257 (2015).
23. Hausen, P. & Dreyer, C. The Use of Polyacrylamide as an Embedding Medium for Immunohistochemical Studies of Embryonic Tissues. *Stain Technol.* **56**, 287–293 (1981).
24. Chung, K. *et al.* Structural and molecular interrogation of intact biological systems. *Nature* **497**, 332–337 (2013).
25. Sigal, Y. M., Speer, C. M., Babcock, H. P. & Zhuang, X. Mapping Synaptic Input Fields of Neurons with Super-Resolution Imaging. *Cell* **163**, 493–505 (2015).
26. Chen, F. *et al.* Nanoscale imaging of RNA with expansion microscopy. *Nat. Methods* **13**, 679–684 (2016).
27. Truckenbrodt, S., Sommer, C., Rizzoli, S. O. & Danzl, J. G. A practical guide to optimization in X10 expansion microscopy. *Nat. Protoc.* **14**, 832–863 (2019).
28. Tillberg, P. W. *et al.* Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* **34**, 987–992 (2016).
29. Ku, T. *et al.* Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nat. Biotechnol.* **34**, 973–981 (2016).
30. Shen, F. Y. *et al.* Light microscopy based approach for mapping connectivity with molecular

specificity. *bioRxiv* 2020.02.24.963538 (2020) doi:10.1101/2020.02.24.963538.

31. Hörl, D. *et al.* BigStitcher: reconstructing high-resolution image datasets of cleared and expanded samples. *Nat. Methods* **16**, 870–874 (2019).
32. Hu, Y. *et al.* Improving resolving ability of expansion microscopy by varying crosslinker concentration. *Chem. Commun.* **56**, 4176–4179 (2020).
33. Gambarotto, D. *et al.* Imaging cellular ultrastructures using expansion microscopy (U-ExM). *Nat. Methods* **16**, 71–74 (2019).
34. Yang, B. *et al.* Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158**, 945–958 (2014).
35. M'Saad, O. & Bewersdorf, J. *Light microscopy of proteins in their ultrastructural context.* <http://biorxiv.org/lookup/doi/10.1101/2020.03.13.989756> (2020) doi:10.1101/2020.03.13.989756.
36. Murray, E. *et al.* Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell* **163**, 1500–1514 (2015).
37. Cai, R. *et al.* Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull–meninges connections. *Nat. Neurosci.* **22**, 317–327 (2019).
38. Yun, D. H. *et al.* Ultrafast immunostaining of organ-scale tissues for scalable proteomic phenotyping. *bioRxiv* 660373 (2019) doi:10.1101/660373.
39. Expansion Microscopy: Protocols for Imaging Proteins and RNA in Cells and Tissues - Asano - 2018 - Current Protocols in Cell Biology - Wiley Online Library. <https://currentprotocols.onlinelibrary.wiley.com/doi/10.1002/cpcb.56>.
40. Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J. Cell. Comp. Physiol.* **59**, 223–239 (1962).
41. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805 (1994).
42. Cranfill, P. J. *et al.* Quantitative assessment of fluorescent proteins. *Nat. Methods* **13**, 557–562 (2016).
43. Luo, L. Fly MARCM and Mouse MADM: Genetic Methods of Labeling and Manipulating Single Neurons. *Brain Res. Rev.* **55**, 220–227 (2007).
44. Feng, G. *et al.* Imaging Neuronal Subsets in Transgenic Mice Expressing Multiple Spectral Variants of GFP. *Neuron* **28**, 41–51 (2000).
45. Trachtenberg, J. T. *et al.* Long-term in vivo imaging of experience-dependent synaptic

- plasticity in adult cortex. *Nature* **420**, 788–794 (2002).
46. Carter, L. M. *et al.* The Yellow Fluorescent Protein (YFP-H) Mouse Reveals Neuroprotection as a Novel Mechanism Underlying Chondroitinase ABC-Mediated Repair after Spinal Cord Injury. *J. Neurosci.* **28**, 14107–14120 (2008).
  47. Zong, H., Espinosa, J. S., Su, H. H., Muzumdar, M. D. & Luo, L. Mosaic Analysis with Double Markers in Mice. *Cell* **121**, 479–492 (2005).
  48. Cai, D., Cohen, K. B., Luo, T., Lichtman, J. W. & Sanes, J. R. Improved tools for the Brainbow toolbox. *Nat. Methods* **10**, 540–547 (2013).
  49. Kelly, K. L. & Judd, D. B. *Color: universal language and dictionary of names*. (U.S. Dept. of Commerce, National Bureau of Standards : For sale by the Supt. of Docs., U.S. Govt. Print. Off., 1976).
  50. Green-Armytage, P. A Colour Alphabet and the Limits of Colour Coding. 23 (2010).
  51. Veling, M. W. *et al.* Identification of Neuronal Lineages in the Drosophila Peripheral Nervous System with a “Digital” Multi-spectral Lineage Tracing System. *Cell Rep.* **29**, 3303-3312.e3 (2019).
  52. Missirlis, P. I., Smailus, D. E. & Holt, R. A. A high-throughput screen identifying sequence and promiscuity characteristics of the loxP spacer region in Cre-mediated recombination. *BMC Genomics* **7**, 73 (2006).
  53. Sakaguchi, R., Leiwe, M. N. & Imai, T. Bright multicolor labeling of neuronal circuits with fluorescent proteins and chemical tags. *eLife* **7**, e40350 (2018).
  54. Chan, K. Y. *et al.* Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017).
  55. Daigle, T. L. *et al.* A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. *Cell* **174**, 465-480.e22 (2018).
  56. Viswanathan, S. *et al.* High-performance probes for light and electron microscopy. *Nat. Methods* **12**, 568–576 (2015).
  57. Prasad, K.-M. R., Xu, Y., Yang, Z., Acton, S. T. & French, B. A. Robust cardiomyocyte-specific gene expression following systemic injection of AAV: in vivo gene delivery follows a Poisson distribution. *Gene Ther.* **18**, 43–52 (2011).
  58. Keller, D., Erö, C. & Markram, H. Cell Densities in the Mouse Brain: A Systematic Review. *Front. Neuroanat.* **12**, (2018).
  59. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).

60. Saka, S. K. *et al.* Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol.* **37**, 1080–1090 (2019).
61. Guo, S.-M. *et al.* Multiplexed and high-throughput neuronal fluorescence imaging with diffusible probes. *Nat. Commun.* **10**, 4377 (2019).
62. Park, Y.-G. *et al.* Protection of tissue physicochemical properties using polyfunctional crosslinkers. *Nat. Biotechnol.* **37**, 73–83 (2019).
63. Micheva, K. D. & Smith, S. J. Array Tomography: A New Tool for Imaging the Molecular Architecture and Ultrastructure of Neural Circuits. *Neuron* **55**, 25–36 (2007).
64. Malik, A. M. *et al.* Matrin 3-dependent neurotoxicity is modified by nucleic acid binding and nucleocytoplasmic localization. *eLife* **7**, e35977 (2018).
65. Sun, N. *et al.* Measuring in vivo mitophagy. *Mol. Cell* **60**, 685–696 (2015).
66. Gurskaya, N. G. *et al.* Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nat. Biotechnol.* **24**, 461–465 (2006).
67. Delgado, T. *et al.* Comparing 3D ultrastructure of presynaptic and postsynaptic mitochondria. *Biol. Open* **8**, (2019).
68. Frontiers | Presynaptic Boutons That Contain Mitochondria Are More Stable | Frontiers in Synaptic Neuroscience. <https://www.frontiersin.org/articles/10.3389/fnsyn.2019.00037/full>.
69. Zhang, Q., Lee, W.-C. A., Paul, D. L. & Ginty, D. D. Multiplexed peroxidase-based electron microscopy labeling enables simultaneous visualization of multiple cell types. *Nat. Neurosci.* **22**, 828–839 (2019).
70. Maggioni, M., Katkovnik, V., Egiazarian, K. & Foi, A. A Nonlocal Transform-Domain Filter for Volumetric Data Denoising and Reconstruction. 15.
71. Coupé, P., Munz, M., Manjón, J. V., Ruthazer, E. S. & Louis Collins, D. A CANDLE for a deeper in vivo insight. *Med. Image Anal.* **16**, 849–864 (2012).
72. Ravindra Kumar, S. *et al.* Multiplexed Cre-dependent selection yields systemic AAVs for targeting distinct brain cell types. *Nat. Methods* 1–10 (2020) doi:10.1038/s41592-020-0799-7.
73. Südhof, T. C. The Presynaptic Active Zone. *Neuron* **75**, 11–25 (2012).
74. Sheng, M. & Kim, E. The Postsynaptic Organization of Synapses. *Cold Spring Harb. Perspect. Biol.* **3**, (2011).
75. Xu, C. S., Pang, S., Hayworth, K. J. & Hess, H. F. Enabling FIB-SEM Systems for Large Volume Connectomics and Cell Biology. *bioRxiv* 852863 (2019) doi:10.1101/852863.

76. A Connectome of the Adult *Drosophila* Central Brain | bioRxiv.  
<https://www.biorxiv.org/content/10.1101/2020.01.21.911859v1>.
77. Motta, A. *et al.* Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science* **366**, (2019).
78. Dani, A., Huang, B., Bergan, J., Dulac, C. & Zhuang, X. Super-resolution Imaging of Chemical Synapses in the Brain. *Neuron* **68**, 843–856 (2010).
79. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).
80. Chozinski, T. J. *et al.* Expansion microscopy with conventional antibodies and fluorescent proteins. *Nat. Methods* **13**, 485–488 (2016).
81. Saka, S. K. *et al.* Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol.* **37**, 1080–1090 (2019).
82. Guo, S.-M. *et al.* Multiplexed and high-throughput neuronal fluorescence imaging with diffusible probes. *Nat. Commun.* **10**, 4377 (2019).
83. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968–981.e15 (2018).
84. Chung, K. *et al.* Structural and molecular interrogation of intact biological systems. *Nature* **497**, 332–337 (2013).
85. Vereczki, V. K. *et al.* Synaptic Organization of Perisomatic GABAergic Inputs onto the Principal Cells of the Mouse Basolateral Amygdala. *Front. Neuroanat.* **10**, (2016).
86. McDonald, A. J. & Betette, R. L. Parvalbumin-containing neurons in the rat basolateral amygdala: morphology and co-localization of Calbindin-D28k. *Neuroscience* **102**, 413–425 (2001).
87. McDonald, A. J. & Mascagni, F. Immunohistochemical characterization of somatostatin containing interneurons in the rat basolateral amygdala. *Brain Res.* **943**, 237–244 (2002).
88. Roossien, D. H. *et al.* Multispectral tracing in densely labeled mouse brain with nTracer. *Bioinformatics* **35**, 3544–3546 (2019).
89. Chang, J.-B. *et al.* Iterative expansion microscopy. *Nat. Methods* **14**, 593–599 (2017).
90. Lee, K.-S., Vandemark, K., Mezey, D., Shultz, N. & Fitzpatrick, D. Functional Synaptic Architecture of Callosal Inputs in Mouse Primary Visual Cortex. *Neuron* **101**, 421–428.e5 (2019).
91. Holler-Rickauer, S., Köstinger, G., Martin, K. A. C., Schuhknecht, G. F. P. & Stratford, K. J. *Structure and function of a neocortical synapse*.  
<http://biorxiv.org/lookup/doi/10.1101/2019.12.13.875971> (2019)

doi:10.1101/2019.12.13.875971.

92. Jiang, X. *et al.* Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, (2015).
93. Villa, K. L. *et al.* Inhibitory Synapses Are Repeatedly Assembled and Removed at Persistent Sites In Vivo. *Neuron* **90**, 662–664 (2016).
94. Flores, C. E. & Méndez, P. Shaping inhibition: activity dependent structural plasticity of GABAergic synapses. *Front. Cell. Neurosci.* **8**, (2014).
95. Crosby, K. C. *et al.* Nanoscale Subsynaptic Domains Underlie the Organization of the Inhibitory Synapse. *Cell Rep.* **26**, 3284-3297.e3 (2019).
96. Nusser, Z., Hájos, N., Somogyi, P. & Mody, I. Increased number of synaptic GABA(A) receptors underlies potentiation at hippocampal inhibitory synapses. *Nature* **395**, 172–177 (1998).
97. Lim, R., Alvarez, F. J. & Walmsley, B. Quantal size is correlated with receptor cluster area at glycinergic synapses in the rat brainstem. *J. Physiol.* **516**, 505–512 (1999).
98. Yuste, R. & Bonhoeffer, T. Morphological Changes in Dendritic Spines Associated with Long-Term Synaptic Plasticity. *Annu. Rev. Neurosci.* **24**, 1071–1089 (2001).
99. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
100. Yuste, R. *et al.* A community-based transcriptomics classification and nomenclature of neocortical cell types. *ArXiv190903083 Q-Bio* (2019).
101. Gouwens, N. W. *et al.* Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat. Neurosci.* **22**, 1182–1195 (2019).
102. Gouwens, N. W. *et al.* Toward an integrated classification of neuronal cell types: morphoelectric and transcriptomic characterization of individual GABAergic cortical neurons. *bioRxiv* 2020.02.03.932244 (2020) doi:10.1101/2020.02.03.932244.
103. Li, A. *et al.* Micro-Optical Sectioning Tomography to Obtain a High-Resolution Atlas of the Mouse Brain. *Science* **330**, 1404–1408 (2010).
104. Hillman, E. M. C., Voleti, V., Li, W. & Yu, H. Light-Sheet Microscopy in Neuroscience. *Annu. Rev. Neurosci.* **42**, 295–313 (2019).
105. Renier, N. *et al.* iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume Imaging. *Cell* **159**, 896–910 (2014).
106. Januszewski, M. *et al.* High-precision automated reconstruction of neurons with flood-filling networks. *Nat. Methods* **15**, 605–610 (2018).

107. Huang, G. B., Scheffer, L. K. & Plaza, S. M. Fully-Automatic Synapse Prediction and Validation on a Large Data Set. *Front. Neural Circuits* **12**, (2018).
108. Bogovic, J. A., Hanslovsky, P., Wong, A. & Saalfeld, S. Robust registration of calcium images by learned contrast synthesis. in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* 1123–1126 (2016). doi:10.1109/ISBI.2016.7493463.
109. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. P. W. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Trans. Med. Imaging* **29**, 196–205 (2010).
110. Shamonin, D. P. *et al.* Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer's Disease. *Front. Neuroinformatics* **7**, (2014).
111. Schmid, B. *et al.* 3Dscript: animating 3D/4D microscopy data using a natural-language-based syntax. *Nat. Methods* **16**, 278–280 (2019).
112. Peng, H., Ruan, Z., Long, F., Simpson, J. H. & Myers, E. W. V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nat. Biotechnol.* **28**, 348–353 (2010).
113. Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* **50**, 1–14 (2018).
114. Wang, Y. *et al.* Complete single neuron reconstruction reveals morphological diversity in molecularly defined claustral and cortical neuron types. *bioRxiv* 675280 (2019) doi:10.1101/675280.
115. Winnubst, J. *et al.* Reconstruction of 1,000 Projection Neurons Reveals New Cell Types and Organization of Long-Range Connectivity in the Mouse Brain. *Cell* **179**, 268–281.e13 (2019).
116. Economo, M. N. *et al.* A platform for brain-wide imaging and reconstruction of individual neurons. *eLife* **5**, e10566 (2016).
117. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).
118. Nelson, E. L., Liang, C. L., Sinton, C. M. & German, D. C. Midbrain dopaminergic neurons in the mouse: computer-assisted mapping. *J. Comp. Neurol.* **369**, 361–371 (1996).
119. Matsuda, W. *et al.* Single Nigrostriatal Dopaminergic Neurons Form Widely Spread and Highly Dense Axonal Arborizations in the Neostriatum. *J. Neurosci.* **29**, 444–453 (2009).
120. Scorcioni, R., Polavaram, S. & Ascoli, G. A. L-Measure: a web-accessible tool for the analysis, comparison and search of digital reconstructions of neuronal morphologies. *Nat. Protoc.* **3**, 866–876 (2008).

121. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
122. Linneweber, G. A. *et al.* A neurodevelopmental origin of behavioral individuality in the *Drosophila* visual system. *Science* **367**, 1112–1119 (2020).
123. Graybuck, L. T. *et al.* Enhancer viruses and a transgenic platform for combinatorial cell subclass-specific labeling. *bioRxiv* 525014 (2020) doi:10.1101/525014.