

Scaling Up 3D Imaging, Analysis, and Culture  
of Complex Brain Models

by

Justin M. Swaney

B.S. Chemical Engineering  
University of Wisconsin-Madison, 2014

SUBMITTED TO THE DEPARTMENT OF CHEMICAL ENGINEERING IN  
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING  
AT THE  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

DECEMBER 2019

© 2019 Massachusetts Institute of Technology. All rights reserved.

Signature of Author .....

Department of Chemical Engineering  
December 11, 2019

Certified by.....

Kwanghun Chung  
Professor of Chemical Engineering  
Thesis Supervisor

Accepted by .....

Patrick Doyle  
Professor of Chemical Engineering  
Chairman, Committee for Graduate Students

# Scaling Up 3D Imaging, Analysis, and Culture of Complex Brain Models

by

Justin M. Swaney

Submitted to the Department of Chemical Engineering on December 11th, 2019  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Chemical Engineering

## ABSTRACT

The brain is the most complex human organ, containing components from the nanometer scale to the centimeter scale. However, many experimental techniques in neuroscience have been optimized for small brain models. This thesis summarizes a body of work aimed at scaling up 3D imaging, analysis, and tissue culture techniques for large-scale brain models. We present a technique termed SWITCH that inhibits probe binding to allow for diffusion without the formation of a reaction front. To improve imaging resolution, we present a tissue expansion technique called MAP that physically magnifies tissue samples for super-resolution imaging with conventional fluorescence microscopes. Using these tools to achieve volumetric imaging of large-scale brain models generates petabyte-scale data, for which we present horizontally scalable image processing pipelines for analysis of intact mouse brains, marmoset brain samples, and cerebral organoids. The mouse brain pipeline allows region-based statistical analysis of protein expression and cell counts. An efficient single-cell non-rigid coregistration algorithm for multiplexed volumetric fluorescence imaging based on matching corresponding nuclei between imaging rounds is presented. A multiscale phenotyping pipeline allows single-cell, cytoarchitectural, and morphological analyses to be combined into a hyperdimensional statistical analysis of cerebral organoids. We use this pipeline to show phenotypic changes due to neurodevelopment, Zika virus infection, and changes in organoid culture protocols. Current cerebral organoid cultures are limited by nutrient transport since they lack a vascular system. To address this issue *in vitro*, we fabricated synthetic vasculature by two-photon photopolymerization of polyethylene glycol-based resins. Printed micro-vessels were less than 100  $\mu\text{m}$  in outer diameter, permeable to biomolecules through engineered pore structures, and biocompatible. Perfusion of vascularized cerebral organoids cultured for 30 days resulted in the expected neuronal differentiation as well as integration of the vascular network. Future studies can use and build on these technical advances to further our understanding of the brain through the use of large-scale brain models.

Thesis Supervisor: Kwanghun Chung

Title: Professor of Chemical Engineering

# Acknowledgements

I would like to acknowledge all of my colleagues at MIT for their continued support throughout various research projects. Although challenging at times, we managed to accomplish a great deal and learn even more together over the last five years. I have no doubt that those individuals that I have had the privilege to work with at MIT will continue to impact the world in positive ways.

I would also like to thank my students, whether through my role as a teaching assistant or as a undergraduate research supervisor, for teaching me. You may not have realized that I was also learning from the experience, but it was a constant reminder of how important it is to give back to the next generation of problem-solvers. If we can provide them with the tools, then there is no problem too large for the curiosity of a young scientist.

This work never would have been possible without my family. I am grateful for my fiancé, Kassi, who has been with me side-by-side as an MIT PhD student since the beginning. We made it through grad school together, and I'm excited to share our hardships and victories in the future. My brother, Jason, and his wife, Shanon, always gave me a place to stay whenever I went home to Wisconsin during my PhD. My sister, Amy, and niece, Lexi, also welcomed me home when I visited. My dad, Rich, has supported me throughout my entire life and academic career.

Lastly, I'd like to dedicate this thesis to my mother, Margaret Swaney, to whom I owe everything. She has been and will always be my guiding light.

# Table of Contents

<b>Abstract</b>	<b>2</b>
<b>Acknowledgements</b>	<b>3</b>
<b>List of figures</b>	<b>7</b>
<b>List of tables</b>	<b>12</b>
<b>Abbreviations</b>	<b>13</b>
<b>1 Introduction</b>	<b>15</b>
1.1 Multiscale intrinsic complexity of the brain . . . . .	15
1.2 Trends toward more complex brain models . . . . .	16
1.3 Thesis aims . . . . .	18
1.4 References . . . . .	20
<b>2 Simple, scalable proteomic imaging for high-dimensional profiling of intact systems</b>	<b>23</b>
2.1 Summary . . . . .	23
2.2 Introduction . . . . .	24
2.3 Results . . . . .	25
2.4 Discussion . . . . .	41
2.5 Experimental Procedures . . . . .	44
2.6 Supplemental Information . . . . .	46
2.7 Author Contributions . . . . .	46
2.8 Acknowledgments . . . . .	46
2.9 References . . . . .	47
<b>3 Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues</b>	<b>51</b>
3.1 Summary . . . . .	51
3.2 Introduction . . . . .	52
3.3 Results . . . . .	53

3.4	Discussion . . . . .	65
3.5	Acknowledgements . . . . .	67
3.6	Author Contributions . . . . .	67
3.7	Competing Financial Interests . . . . .	67
3.8	Methods . . . . .	68
3.9	References . . . . .	75
<b>4</b>	<b>Scalable image processing techniques for quantitative analysis of volumetric biological images from light-sheet microscopy</b>	<b>78</b>
4.1	Abstract . . . . .	78
4.2	Introduction . . . . .	79
4.3	Materials . . . . .	91
4.4	Procedure . . . . .	94
4.5	Anticipated Results . . . . .	94
4.6	Acknowledgements . . . . .	94
4.7	Author Contributions . . . . .	95
4.8	Competing Interests . . . . .	95
4.9	References . . . . .	96
<b>5</b>	<b>Nuclei-based 3D coregistration for multiplexed whole-brain imaging at single-cell resolution</b>	<b>98</b>
5.1	Abstract . . . . .	98
5.2	Introduction . . . . .	99
5.3	Results . . . . .	100
5.4	Discussion . . . . .	106
5.5	Methods . . . . .	108
5.6	References . . . . .	111
<b>6</b>	<b>3D Imaging and Multiscale Analysis of Intact Human Cerebral Organoids</b>	<b>113</b>
6.1	Abstract . . . . .	113
6.2	Introduction . . . . .	114
6.3	Results . . . . .	115
6.4	Discussion . . . . .	130
6.5	Methods . . . . .	131
6.6	References . . . . .	138
<b>7</b>	<b>Vascularization of cerebral organoids using two-photon stereolithography</b>	<b>141</b>
7.1	Abstract . . . . .	141

7.2	Introduction	142
7.3	Results	144
7.4	Discussion	153
7.5	Methods	154
7.6	References	157
<b>8</b>	<b>Conclusion</b>	<b>159</b>
8.1	Thesis summary	159
8.2	Future work	160
<b>Appendix 1: Curriculum Vitae</b>		<b>162</b>
<b>Appendix 2: Image processing protocol</b>		<b>165</b>
Part A:	Destriping LSFM images using pystripe	166
Part B:	Stitching LSFM stacks using TSV	168
Part C:	Atlas alignment with manual refinement using nuggt	172
<b>Appendix 3: Supplemental Information</b>		<b>183</b>
Chapter 2	Supplemental Information	184
Chapter 3	Supplemental Information	201
Chapter 4	Supplemental Information	209
Chapter 6	Supplemental Information	210

# List of figures

<b>Figure 2.1 .....</b>	26
Synchronizing Dialdehyde-Tissue-Gel Formation Enables Scalable Tissue Preservation	
<b>Figure 2.2 .....</b>	30
SWITCH and Co-registration Algorithms Enable Highly Multiplexed Imaging at Single-Cell Resolution	
<b>Figure 2.3 .....</b>	32
SWITCH Enables Proteomic Imaging and High-Dimensional Quantitative Phenotyping of Human Clinical Samples	
<b>Figure 2.4 .....</b>	35
SWITCH Enables Simple, Rapid, and Scalable Tissue Clearing	
<b>Figure 2.5 .....</b>	37
SWITCH Enables Visualization and Quantitative Analysis of Entire Myelinated Fiber Tracts	
<b>Figure 2.6 .....</b>	40
SWITCH Increases Uniformity of Antibody Labeling in Thick Tissues	
<b>Figure 3.1 .....</b>	54
Magnified and accessible 3D proteome library of whole intact organs	
<b>Figure 3.2 .....</b>	56
Comparison of multiscale architectures before and after MAP processing	

<b>Figure 3.3</b>	58
Super-resolution imaging showing the 3D proteome library and subcellular details of MAP-processed intact tissue	
<b>Figure 3.4</b>	60
Multiplexed staining of MAP-processed tissue	
<b>Figure 3.5</b>	62
Intercellular connectivity and its reconstruction at single-fiber resolution in MAP-processed tissue	
<b>Figure 3.6</b>	64
Immunolabeling and imaging of thick MAP-processed tissues	
<b>Figure 4.1</b>	81
Overall image processing pipeline for whole brain analysis	
<b>Figure 4.2</b>	85
Destriping of light-sheet microscopy images using pystripe	
<b>Figure 4.3</b>	87
Stitching and illumination correction of light-sheet microscopy images using TSV	
<b>Figure 4.4</b>	88
Atlas alignment and region-based fluorescence quantification using nuggt	
<b>Figure 5.1</b>	101
Volumetric image preprocessing and compression	
<b>Figure 5.2</b>	102
Coarse registration based on morphological alignment	
<b>Figure 5.3</b>	104
Nuclei matching and thin-plate spline warping	

<b>Figure 5.4</b>	105
Demonstration of 3D coregistration of a whole mouse hemisphere at single-cell resolution	
<b>Figure 5.5</b>	107
Demonstration of 3D coregistration of a large block of marmoset visual cortex at single-cell resolution	
<b>Figure 6.1</b>	117
Pipeline for intact organoid analysis	
<b>Figure 6.2</b>	119
Single cell detection and analysis	
<b>Figure 6.3</b>	121
Ventricle segmentation and cytoarchitecture phenotyping	
<b>Figure 6.4</b>	123
Whole organoid analysis for unbiased quantitative studies	
<b>Figure 6.5</b>	125
Hyper-dimensional analysis of multiscale changes during cerebral organoid maturation	
<b>Figure 6.6</b>	127
Hyper-dimensional analysis of multiscale changes caused by Zika virus infection	
<b>Figure 6.7</b>	129
Hyper-dimensional analysis of multiscale changes in different culture protocols	
<b>Figure 7.1</b>	145
Construction of a high-resolution two-photon SLA printer	
<b>Figure 7.2</b>	146

Two-photon SLA of micro-vessels from biocompatible materials

**Figure 7.3** ..... 148

Tiling strategy for printing large-scale vascular systems

**Figure 7.4** ..... 149

Direct printing of microporous structures improves mass transport from printed reservoirs

**Figure 7.5** ..... 150

Modulation of microfluidic transport properties by direct fabrication of engineered pore structures

**Figure 7.6** ..... 151

Integration of vasculature into cerebral organoids shows mature neurons micro-vessel integration

**Figure 8.1** ..... 186

SWITCH Antibody Compatibility

**Figure 8.2** ..... 187

Multi-color Imaging of SWITCH-Processed Samples

**Figure 8.3** ..... 188

Autocorrelation Analysis of Myelinated Fibers

**Figure 8.4** ..... 189

Autocorrelation Analysis of Bulk Myelinated Fibers

**Figure 8.5** ..... 190

GA-Induced Fluorescence

**Figure 8.6** ..... 191

mRNA FISH on Thermally Cleared Samples

**Figure 8.7** ..... 205

Expansion of various organs with MAP

**Figure 8.8** ..... 206

Validation of commercial antibodies targeting cell-type markers in MAP-processed tissues

**Figure 8.9** ..... 207

Validation of commercial antibodies targeting neurofilament markers in MAP-processed tissues

**Figure 8.10** ..... 218

Optimization of aldehyde fixation for organoids

**Figure 8.11** ..... 219

Optimization of SHIELD protocol for organoids

**Figure 8.12** ..... 220

Nuclei detection and segmentation accuracy

**Figure 8.13** ..... 221

Ventricle segmentation and U-Net model validation

# List of tables

<b>Table 8.1</b> .....	184
Statistics of All Combinations of Six Cell Markers	
<b>Table 8.2</b> .....	201
Antibody summary for the MAP paper	
<b>Table 8.3</b> .....	208
Comparison of technologies for intact tissue processing and imaging	
<b>Table 8.4</b> .....	210
Multiscale organoid features used in the SCOUT hyperdimensional statistical analyses	

# Abbreviations

<b>3D</b>	three-dimensional
<b>AA</b>	acrylamide
<b>ABA</b>	Allen Brain Atlas
<b>BAC</b>	bacterial artificial chromosome
<b>BSA</b>	bovine serum albumin
<b>CB</b>	calbindin
<b>ChAT</b>	choline acetyltransferase
<b>CLARITY</b>	Clear Lipid-exchanged Acrylamide-Hybridized Rigid Imaging/Immunostaining/In Situ Hybridization-Compatible Tissue-Hydrogel
<b>CR</b>	calretinin
<b>DTI</b>	Diffusion Tensor Imaging
<b>eFLASH</b>	electrophoretically driven Fast Labeling using Affinity Sweeping in Hydrogel
<b>ECC</b>	error-correcting code memory
<b>EGDGE</b>	ethyleneglycol diglycidyl ether
<b>EX-313</b>	glycerolpolyglycidyl ether
<b>FFT</b>	fast Fourier transform
<b>FISH</b>	fluorescence in situ hybridization
<b>GA</b>	glutaraldehyde
<b>GABA</b>	gamma-aminobutyric acid
<b>GE21</b>	1,4-butanediol diglycidyl ether
<b>GE23</b>	dipropylene glycol diglycidyl ether
<b>GFAP</b>	glial fibrillary acidic protein
<b>(e)GFP</b>	(enhanced) green fluorescent protein
<b>GPU</b>	graphics processing unit
<b>HCl</b>	hydrochloric acid
<b>IEG</b>	Immediate Early Gene
<b>IHC</b>	immunohistochemistry
<b>LA-ICP-MS</b>	laser-ablation inductively coupled plasma

	mass spectrometry
<b>LSFM</b>	light-sheet fluorescence microscopy
<b>MALDI-MS</b>	matrix-assisted laser desorption ionization mass spectrometry
<b>MAP</b>	Magnified Analysis of Proteome
<b>mRNA</b>	messenger ribonucleic acid
<b>NA</b>	numerical aperture
<b>NaDC</b>	sodium deoxycholate
<b>NF-M</b>	neurofilament medium subunit
<b>NPY</b>	neuropeptide Y
<b>NeuN</b>	neuronal nuclei
<b>O/N</b>	overnight
<b>PBS(T)</b>	phosphate-buffered saline (with Triton X-100)
<b>PV</b>	parvalbumin
<b>PFA</b>	paraformaldehyde
<b>RI</b>	refractive index
<b>RT</b>	room temperature
<b>SDS</b>	sodium dodecyl sulfate
<b>SE</b>	stochastic electrotransport
<b>SHIELD</b>	Stabilization to Harsh Conditions via Intramolecular Epoxide Linkages to Prevent Degradation
<b>SOX2</b>	(sex determining region Y)-box 2
<b>SST</b>	somatostatin
<b>SWITCH</b>	System-Wide Control of Interaction Time and Kinetics of Chemicals
<b>TBR1</b>	T-box brain transcription factor 1
<b>TH</b>	tyrosine hydroxylase

# Chapter 1

## Introduction

### 1.1 Multiscale intrinsic complexity of the brain

The human brain is a complex biological system with components spanning multiple length scales, including synapses, neurons, and various brain regions. Synapses can transfer signals via neurotransmitters and form local connections between neurons (Südhof & Malenka 2008). Neurons can project over long distances, allowing these local synaptic connections to be used to construct a large-scale intercellular network (Knösche & Tittgemeyer 2011). Glia support neuronal function and influence synaptic remodeling (Araque & Navarrete 2010; Stogsdill & Eroglu 2017), and microglia patrol the parenchyma to clear debris and respond to inflammatory signals (Lenz & Nelson 2018). Although the human brain consists of approximately 100 billion neurons, this number pales in comparison to over 100 trillion synapses within the same brain (Herculano-Houzel 2009). This combinatorial explosion of intercellular interactions in the brain is simultaneously what gives rise its emergent properties and what makes it so difficult for neuroscientists to understand (Bassett & Gazzaniga 2011; Postle 2006).

The pioneering work of Ramón y Cajal depicted neural structures with unprecedeted detail and revealed the existence of different types of neurons based on morphology (Garcia-Lopez et al. 2010). Since then, it has become clear that the brain is no exception to the intimate link between structure and function seen throughout biology (Batista-García-Ramó & Fernández-Verdecia 2018). Above the single-cell scale, interesting patterns in cellular organization, or cytoarchitecture, begin to emerge. Different cytoarchitectural appearances seen within Nissl-stained brain sections serve as the basis for defining more macroscopic brain regions (Erö et al. 2018). These regions are often defined hierarchically, with the coarsest partitioning being the lobes of the brain.

In this description of the brain at various length scales, two broad strategies for studying overall brain function begin to emerge naturally. One is a bottom-up approach, in which cellular details such as electrophysiological properties, gene expression levels, and subcellular structures are emphasized. These details are often used to sub-categorize cells in the hopes that given the functional roles of these subtypes, we may be able to understand how they are wired together (Lake et al. 2016). That is, bottom-up approaches based on cellular sub-typing are analogous to trying to understand how a circuit board works by first identifying its electrical components. Another strategy is a top-down approach, in which the functional associations of each brain region, anatomical variations, and large-scale patterns in electrical activity (also known as “brain waves”) are emphasized (Lopes da Silva 1991). By taking a more holistic view of the brain, top-down approaches hope to address how the brain works as a complete system through the interaction of different brain regions. However, studying the brain as a complete system has proven to be experimentally challenging due to its scale and multiscale intrinsic complexity.

Recent advances in brain atlases allow neuroscience researchers to use bottom-up and top-down strategies in tandem, which has been very useful (Nowinski 2017). By defining a common coordinate system for a typical brain, information from multiple brain samples can be effectively pooled into an aggregate dataset of protein expression, gene expression, and electrophysiological properties via alignment to the anatomical reference brain. The Allen Brain Atlas contains such aggregate datasets and has acted as a catalyst allowing neuroscientists to ask exciting new questions (Jones et al. 2009). Although the atlas approach can provide aggregate phenotypic information in the form of population averages, it is still limited in its ability to describe a single brain sample.

The complete network of neural connections in the central nervous system is often referred to as the *connectome* (Toga et al. 2012). In order to observe the connectome, new technological advances are needed to fully capture the multiscale intrinsic complexity of a single brain. Although atlas approaches may be able to elucidate consistent region-to-region projections using sparse labeling techniques such as viral labeling, a complete picture of the connectome remains elusive (Oh et al. 2014). Such a detailed depiction of the entire biological system would bridge the gap between bottom-up and top-down approaches for a single brain and be arguably as profound as the original drawings from Ramón y Cajal.

## 1.2 Trends toward more complex brain models

In neuroscience, researchers often use model systems as proxies for the human brain, and each brain model has its own set of advantages and disadvantages in a given biological

context. Researchers typically use brain models that have been previously demonstrated to be useful to study the neurological phenomena or complex phenotypes of interest. This is because validating new biological models of the brain is quite difficult, and it is even often unclear how well established models approximate what occurs in humans.

Some of the simplest models of the brain are 2D cell cultures of human neurons (Gordon et al. 2013). These models allow easy access for monitoring the entire culture system throughout experimentation. Such 2D cell cultures have been widely adopted for drug screening purposes due to their high-throughput and scalability (Sharma et al. 2012). However, lead compounds identified in primary cell drug screens often do not succeed in clinical trials, suggesting that there is more to the biological story than what 2D cell cultures are telling us (Hefti 2008).

By far the most commonly used brain model used in experimental neuroscience research is the mouse brain (Ellenbroek & Youn 2016). Due to its low cost, amenability to genetic manipulation, and relatively small size, mouse models are a convenient model of the human brain. Neuroscientists have used mouse models to study a wide range of topics, including human disease pathologies (Leung & Jia 2016), mechanisms of learning and memory (Han et al. 2013), and the behavioral underpinnings of social interaction (Ko 2017). However, a complete view of the mouse brain model is more difficult to obtain than with 2D cell cultures.

Recently, cerebral organoids have emerged as a more complex in vitro brain model (Di Lullo & Kriegstein 2017). Cerebral organoids are derived from human pluripotent stem cells and, therefore, may be more relevant to human neurodevelopment than mouse models (Qian et al. 2019). Cerebral organoids are just a single manifestation of a larger trend of in vitro biological modeling with organoids, and similar models have been developed for the kidney (Nishinakamura 2019), liver (Wu et al. 2019), and intestine (Wallach & Bayrer 2017). While cerebral organoids and 2D neuronal cultures both contain human cells, cerebral organoids are 3D tissue cultures, with extracellular matrix material and a more diverse set of cell types, including neural progenitors, neurons, and astrocytes (Lancaster et al. 2013). This complexity is added intentionally, presumably in the hopes that cerebral organoids will be able to model more complex phenotypes accurately. If this proves to be true, then the experimental possibilities of such an in vitro model are vast.

Another model that has become more commonplace in neuroscience research is the *Callicebus jacchus*—also known as the common marmoset—brain (Kishi et al. 2014). Marmosets are a relatively small non-human primate that is phylogenetically more closely related to humans. Therefore, there has been recent interest in using marmoset models for studying more human-like behaviors rather than mouse models (Miller et al. 2016).

These recent trends toward more complex and larger brain models are undeniable, and they originate from some of the inherent limitations of the brain models that are currently available to neuroscientists. However, as these trends continue toward the larger scale, neuroscientists will be hard-pressed to adapt the technologies and experimental techniques that were developed for 2D cell culture and mouse models. Scaling up the techniques used in experimental neuroscience to larger, more complex brain models is the focus of this thesis.

### 1.3 Thesis aims

Aim 1 - Create methods for staining and imaging whole-brain scale models

Building upon tissue clearing technologies such as CLARITY (Chung, J. Wallace, S.-Y. Kim, Kalyanasundaram, Aaron S Andelman, et al. 2013), new methods for preserving, clearing, magnifying, and staining whole-brain scale tissues are presented. First, the SWITCH technique is introduced to decouple the transport and reaction kinetics of fixative molecules and probes (Murray et al. 2015). Then an expansion-based technique called MAP is introduced to improve the imaging resolution of conventional fluorescence microscopy down to synaptic resolution (Ku et al. 2016).

Aim 2 - Develop computational tools for analyzing whole-brain imaging data

Drawing from open-source tools in the Python community, a scalable image processing pipeline is presented with a detailed protocol for single-cell analysis of whole-brain scale volumetric images (Swaney et al. 2019). Then, an efficient 3D coregistration algorithm is presented that matches corresponding nuclei between multiple rounds of imaging of the same brain. This single-cell coregistration algorithm may enable multiplexed fluorescence imaging of a single whole-brain sample. Finally, a multiscale phenotyping pipeline is presented for understanding the complexity in cerebral organoids as a new brain model. This pipeline is used to measure multiscale phenotypic changes in response to maturation, Zika virus infection, and different culture protocols.

### Aim 3 - Scale up cerebral organoid culture by addressing nutrient transport limitations

To increase the scale to which 3D in vitro brain models such as cerebral organoids can be cultured, an organoid vascularization strategy using two-photon stereolithography is presented. A detailed design for a high-resolution two-photon 3D printer is provided that can achieve 4  $\mu\text{m}$  fabrication resolution and directly print engineered pore structures. Vasculature printed from biocompatible resins were shown to integrate with cerebral organoids cultured with and without perfusion with limited cytotoxicity. Future work may build on this vascularization to create new models of the blood-brain barrier or to perturb cerebral organoids using the integrated micro-vessels.

## 1.4 References

- Araque, A. & Navarrete, M., 2010. Glial cells in neuronal network function. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1551), pp.2375–81.
- Bassett, D.S. & Gazzaniga, M.S., 2011. Understanding complexity in the human brain. *Trends in cognitive sciences*, 15(5), pp.200–9.
- Batista-García-Ramó, K. & Fernández-Verdecia, C.I., 2018. What We Know About the Brain Structure-Function Relationship. *Behavioral sciences (Basel, Switzerland)*, 8(4).
- Chung, K. et al., 2013. Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449), pp.332–7.
- Di Lullo, E. & Kriegstein, A.R., 2017. The use of brain organoids to investigate neural development and disease. *Nature Reviews Neuroscience*, 18(10), pp.573–584.
- Ellenbroek, B. & Youn, J., 2016. Rodent models in neuroscience research: is it a rat race? *Disease models & mechanisms*, 9(10), pp.1079–1087.
- Erö, C. et al., 2018. A Cell Atlas for the Mouse Brain. *Frontiers in neuroinformatics*, 12, p.84.
- Garcia-Lopez, P., Garcia-Marin, V. & Freire, M., 2010. The histological slides and drawings of cajal. *Frontiers in neuroanatomy*, 4, p.9.
- Gordon, J., Amini, S. & White, M.K., 2013. General overview of neuronal cell culture. *Methods in molecular biology (Clifton, N.J.)*, 1078, pp.1–8.
- Han, X. et al., 2013. Forebrain engraftment by human glial progenitor cells enhances synaptic plasticity and learning in adult mice. *Cell stem cell*, 12(3), pp.342–53.
- Hefti, F.F., 2008. Requirements for a lead compound to become a clinical candidate. *BMC neuroscience*, 9 Suppl 3(Suppl 3), p.S7.
- Herculano-Houzel, S., 2009. The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in human neuroscience*, 3, p.31.
- Jones, A.R., Overly, C.C. & Sunkin, S.M., 2009. The Allen Brain Atlas: 5 years and beyond. *Nature Reviews Neuroscience*, 10(11), pp.821–828.
- Kishi, N. et al., 2014. Common marmoset as a new model animal for neuroscience research and genome editing technology. *Development, Growth & Differentiation*, 56(1), pp.53–62.

Knösche, T.R. & Tittgemeyer, M., 2011. The role of long-range connectivity for the characterization of the functional-anatomical organization of the cortex. *Frontiers in systems neuroscience*, 5, p.58.

Ko, J., 2017. Neuroanatomical Substrates of Rodent Social Behavior: The Medial Prefrontal Cortex and Its Projection Patterns. *Frontiers in neural circuits*, 11, p.41.

Ku, T. et al., 2016. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nature biotechnology*, 34(9), pp.973–81.

Lake, B.B. et al., 2016. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science (New York, N.Y.)*, 352(6293), pp.1586–90.

Lancaster, M.A. et al., 2013. Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467), pp.373–379.

Lenz, K.M. & Nelson, L.H., 2018. Microglia and Beyond: Innate Immune Cells As Regulators of Brain Development and Behavioral Function. *Frontiers in immunology*, 9, p.698.

Leung, C. & Jia, Z., 2016. Mouse Genetic Models of Human Brain Disorders. *Frontiers in genetics*, 7, p.40.

Lopes da Silva, F., 1991. Neural mechanisms underlying brain waves: from neural membranes to networks. *Electroencephalography and Clinical Neurophysiology*, 79(2), pp.81–93.

Miller, C.T. et al., 2016. Marmosets: A Neuroscientific Model of Human Social Behavior. *Neuron*, 90(2), pp.219–33.

Murray, E. et al., 2015. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6), pp.1500–14.

Nishinakamura, R., 2019. Human kidney organoids: progress and remaining challenges. *Nature Reviews Nephrology*, 15(10), pp.613–624.

Nowinski, W.L., 2017. Human brain atlasing: past, present and future. *The Neuroradiology Journal*, 30(6), pp.504–519.

Oh, S.W. et al., 2014. A mesoscale connectome of the mouse brain. *Nature*, 508(7495), pp.207–14.

Postle, B.R., 2006. Working memory as an emergent property of the mind and brain. *Neuroscience*, 139(1), pp.23–38.

- Qian, X., Song, H. & Ming, G.-L., 2019. Brain organoids: advances, applications and challenges. *Development (Cambridge, England)*, 146(8).
- Sharma, P. et al., 2012. High-throughput screening in primary neurons. *Methods in enzymology*, 506, pp.331–60.
- Stogsdill, J.A. & Eroglu, C., 2017. The interplay between neurons and glia in synapse development and plasticity. *Current opinion in neurobiology*, 42, pp.1–8.
- Südhof, T.C. & Malenka, R.C., 2008. Understanding synapses: past, present, and future. *Neuron*, 60(3), pp.469–76.
- Swaney, J. et al., 2019. Scalable image processing techniques for quantitative analysis of volumetric biological images from light-sheet microscopy. *bioRxiv*.
- Toga, A.W. et al., 2012. Mapping the human connectome. *Neurosurgery*, 71(1), pp.1–5.
- Wallach, T.E. & Bayrer, J.R., 2017. Intestinal Organoids: New Frontiers in the Study of Intestinal Disease and Physiology. *Journal of pediatric gastroenterology and nutrition*, 64(2), pp.180–185.
- Wu, L.-J. et al., 2019. Organoids of liver diseases: From bench to bedside. *World journal of gastroenterology*, 25(16), pp.1913–1927.

# Chapter 2

## Simple, scalable proteomic imaging for high-dimensional profiling of intact systems

Evan Murray\*, Jae Hun Cho\*, Daniel Goodwin\*, Taeyun Ku\*, **Justin Swaney\***, Sung-Yon Kim, Heejin Choi, Young-Gyun Park, Jeong-Yoon Park, Austin Hubbert, Margaret McCue, Sara Vassallo, Naveed Bakh, Matthew P. Frosch, Van J. Wedeeng, H. Sebastian Seung, and Kwanghun Chung<sup>1</sup>

*Reproduced from DOI: <https://doi.org/10.1016/j.cell.2015.11.025>*

### 2.1 Summary

Combined measurement of diverse molecular and anatomical traits that span multiple levels remains a major challenge in biology. Here, we introduce a simple method that enables proteomic imaging for scalable, integrated, high-dimensional phenotyping of both animal tissues and human clinical samples. This method, termed SWITCH, uniformly secures tissue architecture, native biomolecules, and antigenicity across an entire system by synchronizing the tissue preservation reaction. The heat- and chemical-resistant nature of the resulting framework permits multiple rounds ( $>20$ ) of relabeling. We have performed 22 rounds of labeling of a single tissue with precise co-registration of multiple datasets. Furthermore, SWITCH synchronizes labeling reactions to improve probe penetration depth and uniformity of staining. With SWITCH, we performed combinatorial protein expression profiling of the human cortex and also interrogated the geometric struc-

---

<sup>1\*</sup> indicates co-first authorship.

ture of the fiber pathways in mouse brains. Such integrated high-dimensional information may accelerate our understanding of biological systems at multiple levels.

## 2.2 Introduction

Biological systems are comprised of vast numbers of molecules, cell types, and intricate tissue organizations (Alivisatos et al. 2013; Kasthuri et al. 2015; Yuste 2015). Understanding the complex interactions of these components is essential for many fields of biology and often requires high-dimensional information across many scales. Although it is desirable to obtain such information from the same tissue due to large individual variations, combined measurement of many molecular and anatomical traits remains an unmet goal in biology despite the remarkable success of current pioneering methods, such as array tomography (Rah et al. 2013).

Rapidly evolving tissue-clearing techniques may enable multiplexed labeling and imaging of intact samples using light microscopy (Chung, J. Wallace, S.-Y. Kim, Kalyanasundaram, Aaron S. Andalman, et al. 2013; Richardson & Lichtman 2015; Susaki et al. 2014). For instance, the CLARITY technique has demonstrated three rounds of immunostaining of mouse brain tissue (Chung, J. Wallace, S.-Y. Kim, Kalyanasundaram, Aaron S. Andalman, et al. 2013). However, we have noticed that the polyacrylamide-based framework loses structural integrity upon repeated exposure to the elution condition. Recent reports also suggest that preservation of antigenicity in the CLARITY method may not be optimal (Renier et al. 2014). Furthermore, the necessary tissue-gel hybridization step requires delivery of charged thermal initiators with limited diffusivity and stability. This necessity imposes a limit on the tissue size that can be processed without the use of transcardial perfusion.

We set our goal to develop a simple, scalable, and generalizable tissue-processing method for proteomic imaging of intact biological systems. To achieve this, we created SWITCH (system-wide control of interaction time and kinetics of chemicals), which tightly controls a broad range of chemical reactions in tissue processing via a set of buffers: a SWITCH-On buffer that facilitates chemical reactions between exogenous chemicals and endogenous biomolecules, and a SWITCH-Off buffer that suppresses the reactions. SWITCH-mediated fixation transforms tissue into a heat- and chemical-resistant hybrid while preserving tissue architecture, native molecules, and their antigenicity to a degree suitable for multiplexed proteomic imaging. The hybrids can be rapidly cleared at high temperature without damage. The method does not require perfusion and is thus applicable to both animal and large human samples. In molecular labeling of the processed samples, SWITCH controls probe-target binding kinetics to improve probe penetration depth and

the uniformity of molecular labeling. This method is simple, passive, and does not require any special equipment or reagents.

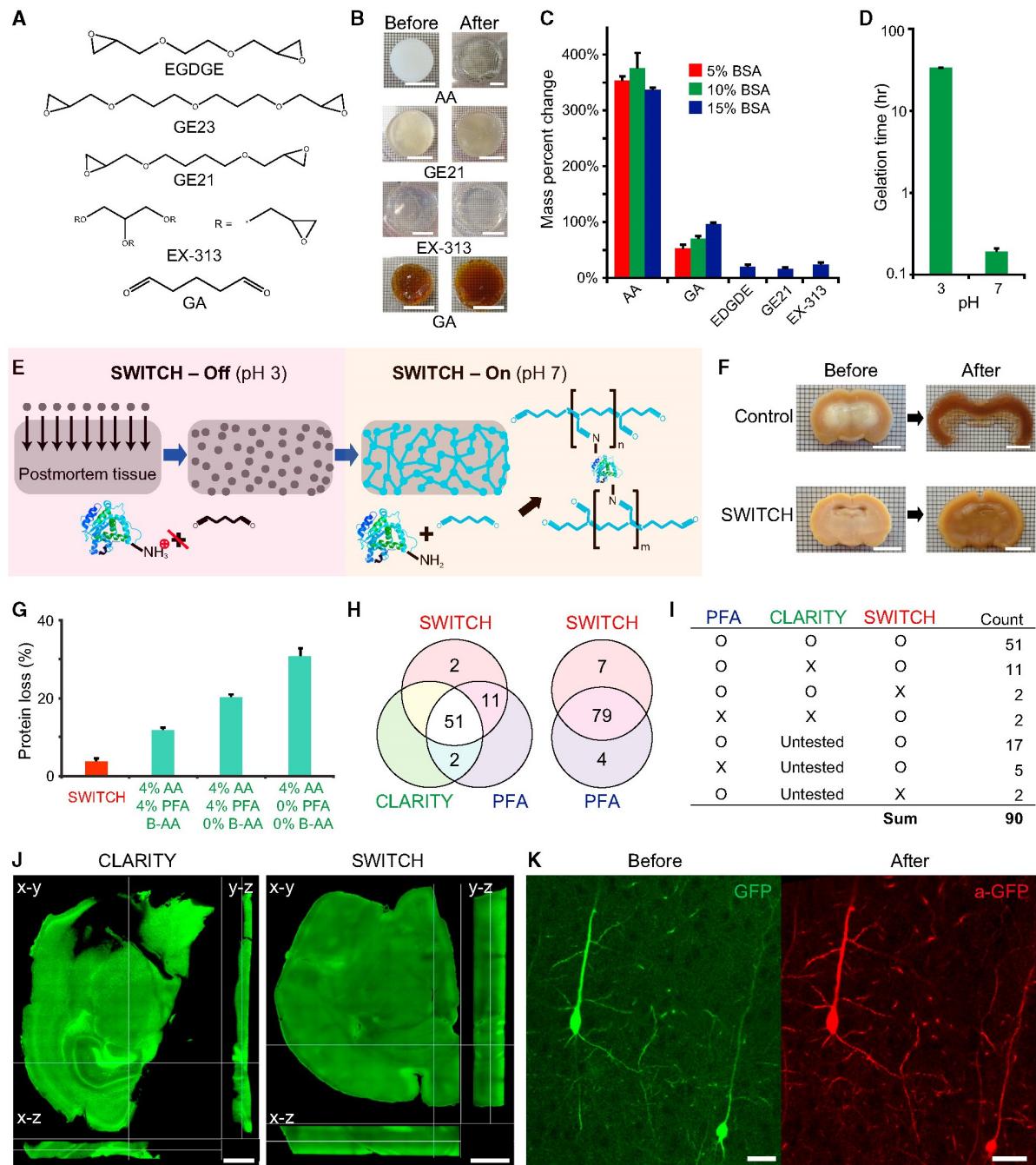
Using SWITCH, we demonstrated that a minimum of 22 rounds of molecular labeling of a banked postmortem human tissue with precise co-registration of multiple datasets at single-cell resolution is possible. We also demonstrated extraction of a wide range of system variables, such as various cell types and microvasculature from a single sample. In summary, we have developed simple tissue processing methods and a volumetric co-registration algorithm that can be readily adopted by most laboratories for scalable proteomic imaging of intact biological systems.

## 2.3 Results

### 2.3.1 Synchronizing Dialdehyde-Tissue-Gel Formation Enables Scalable Tissue Preservation

First, we sought to develop a way to transform animal and human samples into a mechanically and chemically stable form for multiplexed imaging. We hypothesized that small, non-ionic, multifunctional crosslinkers might satisfy two key requirements for such a transformation: (1) rapid penetration without the use of perfusion and (2) a high degree of molecular crosslinking to improve sample durability (Sung et al. 1996). Among many options, we chose to evaluate the following owing to their small size and high water solubility (Figure 2.1A): ethyleneglycol diglycidyl ether (EGDGE), dipropylene glycol diglycidyl ether (GE23), 1,4-butanediol diglycidyl ether (GE21), glycerolpolyglycidyl ether (EX-313), and glutaraldehyde (GA).

We found that all of these chemicals except GE23 formed a solid gel upon incubation with 15% bovine serum albumin (BSA), indicating the formation of a crosslinked network (Figure 2.1B). We examined the stability of the gels along with poly-acrylamide (AA)-BSA gels by measuring the change in their volume after incubation in a 200 mM SDS solution heated to 80C (elution condition). AA-BSA gels swelled and became fragile after exposure to the harsh condition (Figures 2.1B and 2.1C), whereas multifunctional fixative-BSA gels maintained their structural integrity. In particular, GA-BSA gels showed minimal volume change at a wide range of BSA and GA concentrations, whereas others only gelled at high protein concentrations (Figure 2.1C). This result indicates that multifunctional fixatives alone might be sufficient to form a stable matrix that can withstand the harsh elution condition. However, because the average protein content throughout mouse brain samples is around 10% and may be lower within certain regions, we decided that GA



**Figure 2.1:** Synchronizing Dialdehyde-Tissue-Gel Formation Enables Scalable Tissue Preservation.

(A) Chemical structures of various multifunctional fixatives. (B) Crosslinked protein gels before and after exposure to the elution condition. Scale bars, 10 mm. Polyacrylamide (AA) gel swelled and became fragile, whereas multifunctional fixative gels remained intact with minimal expansion. (C) Mass percent change of crosslinked protein gels after exposure to the harsh condition. EDGDE, GE21, and EX-313 were incapable of forming gels at low BSA concentration. Error bars show mean  $\pm$  SD. (D) The gelation time for protein gels crosslinked with GA is nearly 200-fold higher at pH 3 than it is at neutral pH at 4C. Error bars show mean  $\pm$  SD. (E) Schematic diagram illustrating the process of scalable and uniform tissue-gel formation without perfusion using SWITCH. GA molecules diffuse into an intact tissue without reacting with biomolecules in pH 3 buffer (SWITCH-Off step). When GA is uniformly dispersed throughout the tissue, the sample is moved to pH 7 buffer (SWITCH-On step) to initiate global gelation/fixation and achieve uniform tissue preservation. (F) Coronal slices from the middle of whole rat brains passively fixed with (bottom) or without (top) SWITCH. After fixation, the middle coronal slices were cut and incubated in the elution condition for 1 hr. The core of the control slice completely disintegrated, whereas the SWITCH-processed slices remained intact. Scale bars, 6 mm. (G) Only 3proteins are lost in SWITCH-processed brain tissues as opposed to 10AA-based methods. Error bars show mean  $\pm$  SD. (H and I) Antigenicity of proteins is well preserved throughout the clearing process in SWITCH. Of the antibodies tested, 86 of 90 are compatible with SWITCH. (J and K) SWITCH-mediated fixation maximally preserves macroscopic (J) and microscopic (K) structures throughout the elution process. (J) Cross-sectional images of 1-mm-thick mouse coronal slices after exposure to the elution condition. The CLARITY-processed tissue shows significant tissue deformation and collapse, whereas the SWITCH-processed tissue is highly uniform with no signs of macroscopic deformation. Z-step size, 20  $\mu$ m; 10x, 0.3 NA, water-immersion objective. Scale bars, 1 mm. (K) GFP-expressing neurons in the cortex of Thy-1-EGFP mouse brain before and after exposure to the elution condition and anti-GFP staining. 25x, 0.95 NA, water-immersion objective. Scale bars, 30  $\mu$ m.

is the crosslinker most likely to form a uniform framework throughout all regions of a sample.

Next, we asked whether GA can rapidly penetrate tissue to form a uniform tissue-gel without the use of perfusion, which is required for processing most human clinical samples. We incubated a non-fixed whole adult rat brain in PBS containing 1% GA for 2 days and characterized the GA penetration depth and gel formation. Although the small size of GA should make it highly mobile, only the outer layer of the brain was fixed (Figure 2.1F). When a coronal slice from the middle of the brain was exposed to the elution condition, the core of the tissue completely disintegrated, indicating that no gel matrix had formed in the center of the brain (Figure 2.1F). Limited GA penetration has significantly hampered its use in preserving large postmortem tissues (Hopwood 1967). We suspect that rapid reaction of GA with native biomolecules within the outer layer of the brain may cause depletion of GA molecules before they can reach the core.

To overcome this issue, we sought to control the reaction kinetics of GA and biomolecules throughout the system using the SWITCH approach to achieve uniform tissue preservation. We noted that the GA reaction rate is pH-dependent (Hopwood 1970). Indeed, when we titrated solutions of GA and BSA to pH 3, GA-BSA gel formation time increased by nearly 200-fold (Figure 2.1D). Using this pH dependence, we were able to disperse GA uniformly throughout a sample by switching off the crosslinking reaction with a low-pH buffer (Figure 2.1E, left). After 2 days of incubation at low pH, we switched on sample-wide GA-tissue crosslinking by shifting the pH of the sample to a neutral pH (Figure 2.1E,

right). Using this passive buffer-switching approach, we were able to achieve complete GA penetration and uniform gel formation throughout the entire rat brain (Figure 2.1F).

### 2.3.2 Dialdehyde-Tissue-Gel Preserves Structural and Molecular Information Effectively

We next asked whether the GA-tissue-gel has mechanical and chemical properties desirable for multiplexing-based proteomic imaging. Proteomic imaging requires (1) high preservation of endogenous biomolecules and their antigenicity, (2) high structural integrity, and (3) minimal tissue damage during repeated cycles of destaining, labeling, and imaging processes.

We first tested whether endogenous biomolecules are well preserved by measuring protein loss after clearing (see Supplemental Experimental Procedures). We found that control tissues lost an average of 30%–40% protein and AA-tissue-gel lost 10%–20%, but GA-tissue-gel slices lost only 3%–5% of their protein content (Figure 2.1G).

We next asked whether antigenicity of the retained biomolecules is well preserved. We tested 90 antibodies, targeting biomolecules of different sizes (single amino acid to proteins) and subcellular localizations (membrane bound, cytoplasm, nucleus, synapses). Surprisingly, 86 of 90 antibodies were compatible with GA-tissue-gel (Figures 2.1H, and 2.1I, and S1; Table S1). Note that even small molecules, such as dopamine, which are not typically compatible with PFA-fixation, were observable in GA-tissue-gel after the complete removal of lipid bilayers (Figure S1). These biomolecules were stable against heat and chemical treatment, and their antigenicity was well preserved after exposure to elution conditions.

Good structural preservation is essential for resolving protein location with high precision and for studying molecular inter relationships. To characterize the macroscale structural preservation of the samples, we cleared 1-mm-thick tissue blocks using the elution condition and visualized their structural deformation (Figure 2.1J). The PFA-only tissue completely disintegrated. Even the AA-tissue-gel exhibited large deformations overall. GA-tissue-gel, however, showed no signs of structural damage throughout the entirety of the sample.

We next examined structural preservation on a microscopic scale. We imaged GFP-expressing neurons in the cortex of a PFA-fixed 1-mm-thick Thy1-EGFP M line block (Figure 2.1K). We then SWITCH-processed the tissue, cleared it using the harsh elution condition, stained it against GFP, and imaged the same neurons. As shown in Figure 2.1K, the microscopic morphology of the neurons was well preserved throughout the entire

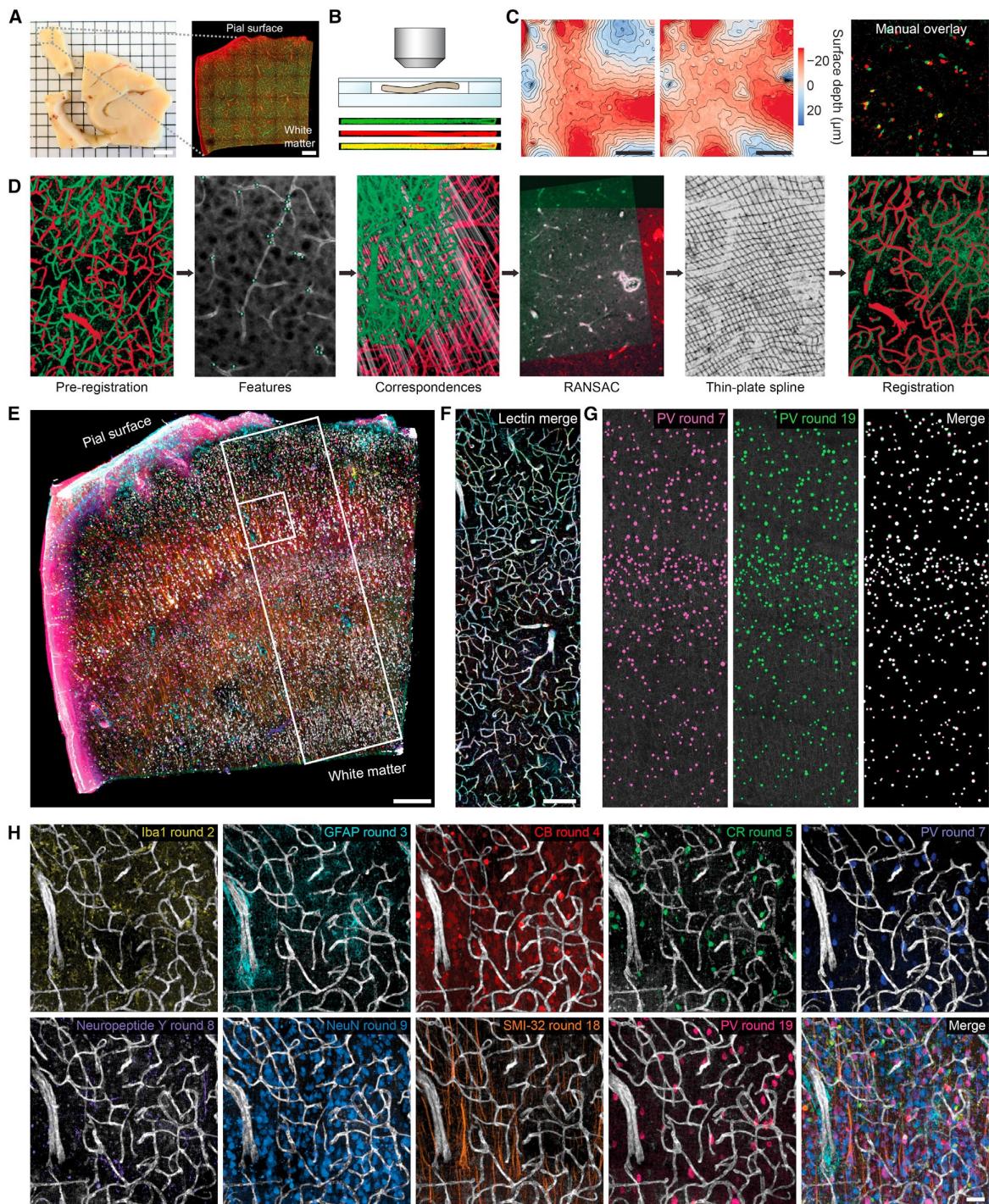
process. These results show GA-tissue-gel may be ideal for highly multiplexed structural and molecular phenotyping.

### 2.3.3 SWITCH and Robust Computational Algorithms Enable Highly Multiplexed Imaging at Single-Cell Resolution

Interrogating the three-dimensional (3D) distribution of molecules, cells, and the overall tissue organization requires precise co-registration of multiple volume images. We first asked if simple manual overlay of two datasets allows precise co-registration. As a stringent test, we used datasets from multi-round imaging of a SWITCH-processed 100- $\mu\text{m}$ -thick human brain slice (100  $\mu\text{m}$ , 200  $\mu\text{m}$ , 200  $\mu\text{m}$ ) (Figure 2.2A). The high aspect ratio of such tissues makes it more prone to physical warping, which renders co-registration particularly challenging. We first stained the tissue using DAPI and anti-parvalbumin (PV) antibody. The slice was then enclosed in a space larger than the tissue to exaggerate possible tissue deformation in the mounting process (Figure 2.2B). After imaging, the sample was exposed to the elution condition overnight (O/N) to completely remove imaged probes. We then restained the tissue using the same probes and repeated the imaging process. Note that only GA-tissue-gels could maintain their integrity against the elution treatment. Both AA-tissue-gels and PFA-fixed samples deteriorated rapidly in the same condition.

As predicted, a large degree of tissue warping in the mounting process (Figure 2.2C) made manual overlay insufficient for the task of interrogating a tissue across multiple staining rounds. To achieve precise co-registration of volume images in the presence of such high-degree warping, we custom-designed a robust computational software based on a feature-detection approach that was ideal for our experimental procedure (Figure 2.2D). Each staining round contained one fluorescence channel devoted to a lectin stain because the morphology of blood vessels creates distinctive keypoints that computer vision algorithms are well suited to identify. With the keypoints, the algorithm warps the tissue in a physically plausible manner into the correct position (see Supplemental Experimental Procedures).

As a stringent test of the algorithm, we used the same SWITCH-processed human sample with the high aspect ratio (Figure 2.2A). For each round, the sample was stained with DAPI, lectin, and one antibody to label a target protein. Although at least three antibodies can be used for each round in addition to lectin and DAPI (Figure S2), we chose to use one antibody for each round to eliminate any possible cross-talk between channels. After acquiring images, we destained the sample and began the next round of labeling. We repeated the above procedure 22 times using markers for various cell types (Figure



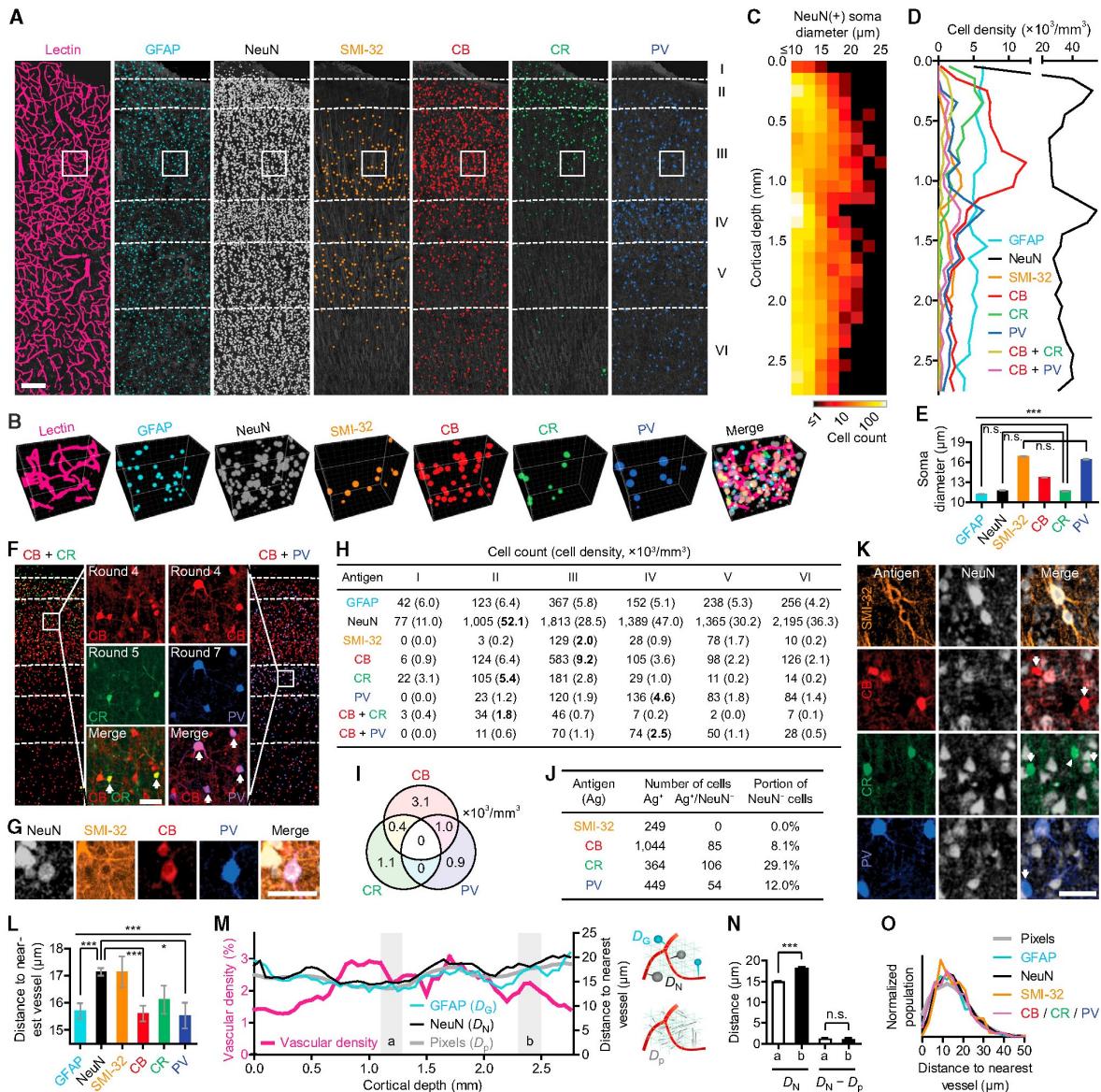
**Figure 2.2:** SWITCH and Co-registration Algorithms Enable Highly Multiplexed Imaging at Single-Cell Resolution.

(A) The left image shows formalin-fixed postmortem human brain tissue (visual association cortex, Brodmann area 18). The right image shows a 100- $\mu$ m section of this brain tissue after SWITCH processing. Scale bars, 5 mm (left), 300  $\mu$ m (right). (B) Natural warping of the sample during imaging was enabled by mounting within a chamber space larger than the size of the sample. Representative crosssections of the sample after several rounds of imaging are shown. Sample thickness, 100  $\mu$ m. (C) Surface contour maps showing warping of the sample between imaging rounds. Attempted manual overlay of two PV datasets shows that sample warping is too severe for single-cell registration without computational correction. Scale bars, 50  $\mu$ m. (D) A flow diagram depicting the sequence of events for automated co-registration of datasets. (E) Fully co-registered image showing an overlay of 9 rounds of immunostaining. A total of 22 rounds of staining of the same tissue was achieved. R#2 (Iba1), R#3 (GFAP), R#4 (calbindin, CB), R#5 (calretinin, CR), R#7 (PV), R#8 (Neuropeptide Y), R#9 (NeuN), R#18 (SMI-32), and R#19 (PV) were used for co-registration and subsequent quantitative analysis (see Figure 3). The boxed regions indicate the ROI's shown in panels. (F–H) Scale bar, 300  $\mu$ m. (F) Vasculature labeling from 9 rounds of staining after co-registration. Scale bar, 200  $\mu$ m. (G) PV cell counts between rounds 7 and 19. After 12 rounds of imaging, 99PV+ cells were again detected and shown to overlay after co-registration of the datasets. (H) Images of individual channels with corresponding vasculature labeling. Scale bar, 50  $\mu$ m.

2.2H; Table 1). Staining was not successful in every round due to the use of non-validated antibodies, sub-optimal staining conditions, or human error, all of which often occur in general laboratory settings and can result in the loss of important samples. However, a SWITCH-processed sample is free from this issue as the tissue can be washed and reused repeatedly.

We were able to successfully co-register all nine datasets with successful staining (Figures 2.2E and 2.2H; Movie S1). We asked whether changes in the sample might be occurring between staining rounds. To test this, we repeated staining with anti-PV antibodies in rounds 7 and 19 and co-registered the resulting datasets. Even when separated by 12 rounds of labeling, we were able to achieve single-cell accuracy of registration with 99% agreement between the two rounds (Figure 2.2G).

We next performed joint statistical analysis of the integrated cross-talk-free dataset to extract diverse phenotypic information from human brain (Figure 2.3). We included lectin, GFAP, NeuN, SMI-32, and three calcium-binding protein channels—calbindin (CB), calretinin (CR), and PV—in the quantitative analysis. First, we used semi-automated algorithms to identify blood vessels and cells expressing the target antigens (Figures 2.3A and 2.3B) and extract their spatial (x, y, z coordinates) and morphological (e.g., cell soma size) information. Density and size profiles of NeuN-positive cells (Figures 2.3C and 2.3D) enabled us to define the cortical layers (Figure 2.3A) according to established criteria (De Sousa et al. 2010). NeuN+ density was high in cortical layers II and IV, with characteristic small cells (NeuN in Figures 2.3A, 2.3C, 2.3D, and 2.3H). Large NeuN+ neurons were concentrated in layers III and V. A portion of these were large pyramidal neurons positive for SMI-32 (Figures 3A, 3E, and 3H). CB+, CR+, and PV+ cells also showed distinct distribution patterns along the cortical axis (Figures 2.3A and 2.3D), in agreement with previous studies (Defelipe et al. 1999; Leuba et al. 1998).



**Figure 2.3:** SWITCH Enables Proteomic Imaging and High-Dimensional Quantitative Phenotyping of Human Clinical Samples.

(A) ROI from Figure 2.2E showing semi-automatically detected locations and sizes of blood vessels (lectin) and diverse cell types (GFAP+, NeuN+, SMI-32+, CB+, CR+, PV+) in human visual cortex. The identified objects are overlaid on maximum intensity-projections of raw images of the corresponding channels (dark gray). Dashed lines divide cortical layers I-VI. (B) 3D rendering of the boxed region in (A) (200  $\mu$ m wide x 200  $\mu$ m high x 104  $\mu$ m deep) showing identified cells and blood vessels. (C) A heat map of the soma size distribution of NeuN+ cells, showing bimodal peaks at cortical layers III and V. (D) Density profiles of various cell types. (E) Comparison of cell sizes among different types of cells. One-way ANOVA was performed (\*\*p < 0.001; N = 1,176, 7,835, 249, 1,044, 364 and 449 for each column). Post hoc tests were mostly p < 0.001 except for three non-significant (n.s.) cases. (F) Distribution of neurons expressing various subsets of calcium-binding proteins in the human visual cortex. Raw images in the middle columns show CB+/CR+ or CB+/PV+ neurons (arrows). (G) A representative NeuN+/SMI-32+/CB+/PV+ cell. (H) Cell counts and densities in different cortical layers. Cortical layers with the highest density for each neuronal channel are highlighted. (I) Cell densities for combinatorial co-expression of three interneuronal markers. (J) Statistics for NeuN- neurons. (K) Representative images showing NeuN-/CB+, NeuN-/CR+, and NeuN-/PV+ cells (arrows). The arrowhead indicates a CR+ cell with low NeuN immunoreactivity. (L) Comparison of cell-to-nearest vessel distances along cortical depth as measured from cell centroids to vascular boundaries. Post hoc tests following one-way ANOVA (p < 0.001; N = 935, 4,101, 210, 817, 265 and 331 for each column) were mostly n.s. except for three cases displayed. \*p < 0.05. (M) Vascular density and distance-to-nearest vessel profiles of GFAP+ or NeuN+ cells along cortical depth. Mean distances from NeuN+ ( $D_N$ ) and GFAP+ ( $D_G$ ) cells and all extravascular pixels ( $D_p$ ) are calculated and plotted. Diagrams illustrate the calculation of the three distances. (N) Cell-to-nearest vessel distances from NeuN+ cells in two regions—a (n = 570) and b (n = 445) in (M)—before ( $D_N$ ) and after ( $D_N - D_p$ ) correction. (O) Distribution profile of extravascular pixel- or cell-to-nearest vessel distances showing similar patterns. Three interneuronal markers are plotted together. Error bars are shown with mean  $\pm$  SEM. Scale bars, 200  $\mu$ m (A), 50  $\mu$ m (F, G, and K).

We next performed unbiased combinatorial expression profiling with the six cell-type specific proteins (GFAP, NeuN, SMI-32, CB, CR, PV). Among 63 possible combinations, 16 were found (Table S2). We identified sub-populations of CB+/CR+ and CB+/PV+ cells, but no CR+/PV+ or CB+/CR+/PV+, cells (Figures 2.3F, 2.3H, and 2.3I), in agreement with a previous report regarding mouse visual cortex (Gonchar et al. 2008). Interestingly, we observed that a significant portion of the CB, CR, and PV-positive neurons do not express detectable levels of NeuN, a widely used pan-neuronal marker (Figures 2.3J and 2.3K). In particular, a majority of CR+ cells showed very weak (Figure 2.3K, arrowhead) or no NeuN immunoreactivity (29.1%), whereas all SMI-32+ cells (Figures 2.3J and 2.3K) were NeuN positive. These results suggest that NeuN expression may be neuronal-type-specific in adult human visual association cortex. We also found a small number of CB+ cells and PV+ cells co-expressing SMI-32, a widely used pyramidal neuronal marker (Table S2) (Campbell & Morrison 1989). Five CB+/PV+ cells were identified as quadruple-positive (NeuN+/SMI-32+/CB+/PV+) cells (Figure 2.3G). All of the CB+ cells and PV+ cells co-expressing SMI-32 were localized in cortical layers III and IV. These results demonstrate the power of SWITCH as a tool for 3D proteomic profiling of intact biological samples at single-cell resolution.

Structural relationships between vasculature and brain cells have been a topic of interest in a broad range of basic and clinical research. Many previous studies obtained the cell-to-vessel distance from 2D images or small tissue volumes, which may hinder precise measurement of such 3D properties. Moreover, in many studies, separate measurements

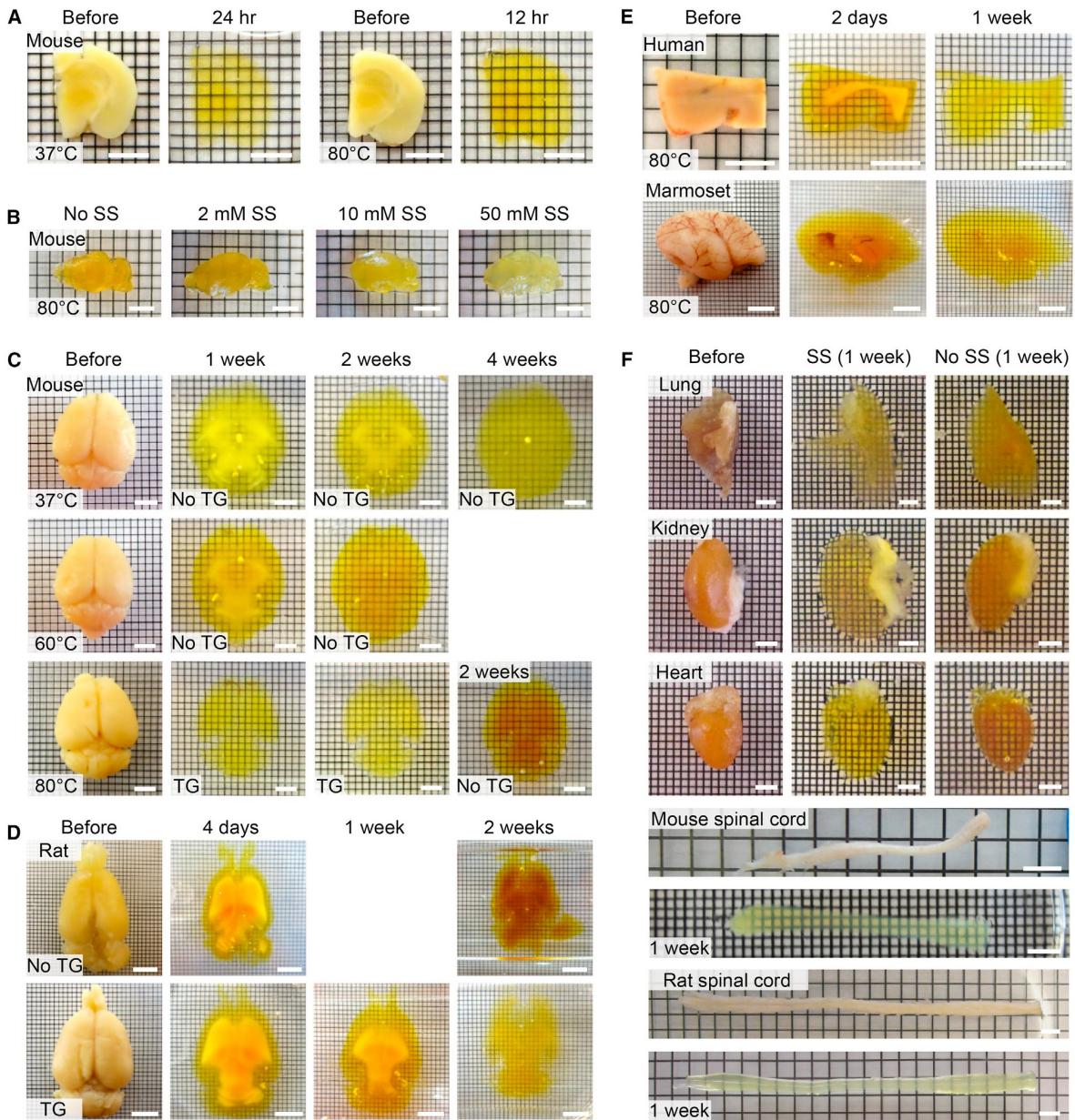
from different tissues needed to be compared without considering individual variabilities in local vasculature geometry. There has been no direct comparison of 3D cell-to-vessel distance among diverse cell types within the same intact tissue.

Using the proteomic imaging capability of SWITCH, for the first time, we were able to directly measure cell-to-vessel distances for six different cell types within a single intact tissue (Figures 2.3L–2.3O). As expected (McCaslin et al. 2011), GFAP+ astrocytes had a shorter mean distance than NeuN+ neurons (Figure 2.3L). CB+ and PV+ cells were also more closely localized near blood vessels than NeuN+ cells, but the difference was relatively small. Figure 3M shows that vascular density is not uniform along the cortex. However, the extravascular pixel-to-vessel distance ( $D_p$ ), which we defined as a reference parameter to reflect the effect of the 3D vascular geometry (Figure 2.3M, right), did not show an inverse relationship with vascular density. This result may suggest that 3D vessel geometry is an important parameter to be considered in understanding a given vascular environment. In fact, cell-to-vessel distance profiles of many cell types closely followed the  $D_p$  profile (GFAP+,  $D_G$ , and NeuN+,  $D_N$ , shown in Figure 2.3M). In particular, when  $D_p$  was subtracted from cell-to-vessel distances ( $D_x$ ) to cancel the influence of vascular geometric variation,  $D_x - D_p$  turns out to be very consistent throughout cortical depth (Figure 2.3N). We further examined the distance distribution profiles for all cell types (Figure 2.3O). All profiles showed similar characteristic curves, which can be seen when objects are randomly located in a 3D space (Manzo et al. 2014). We could not observe any cell-type-specific distribution profile or bi- or multi-modal distribution pattern in this sample. Together, these data demonstrate that SWITCH can be used for high-dimensional quantitative phenotyping of human clinical samples.

### 2.3.4 SWITCH Enables Simple, Rapid, and Scalable Tissue-Clearing

To extend the multiplexed imaging capability of the SWITCH method to large systems, we developed a simple and rapid clearing method. We hypothesized that key steps in detergent-mediated lipid removal, such as permeation of SDS through membranes, might be strongly enhanced by increasing temperature (Keller et al. 2006), and SWITCH-processed samples may endure prolonged incubation at elevated temperatures. Indeed, thermal energy drastically increased the passive clearing speed of SWITCH-processed samples without noticeable tissue damage (Figure 2.4A). We achieved passive clearing of a whole adult mouse brain within 4 days at 80C (versus 4 weeks at 37C) (Figure 2.4C).

Upon prolonged exposure to high temperatures, however, samples developed a brownish hue (Friedman 1996), which may interfere with imaging at certain wavelengths (Figures 2.4B–2.4D and 2.4F). We found that reducing agents, such as sodium sulfite and 1-



**Figure 2.4:** SWIITCH Enables Simple, Rapid, and Scalable Tissue Clearing. (A) Images of 1-mm coronal blocks of an adult mouse brain hemisphere before and after clearing at 37C for 24 hr or 80C for 12 hr. The lipid-extracted tissues were refractive index (RI)-matched (see SI for details). Scale bars, 3 mm. (B) Images of mouse brain hemispheres lipid-extracted at 80C for 10 days with 200 mM SDS containing 0–50 mM sodium sulfite (SS) as an anti-browning agent. Note that the tissues were not RI-matched. Scale bars, 6 mm. (C) Images of intact adult mouse brains cleared at 37C (top) and 60C (middle) and 80C (bottom) with and without 1-thioglycerol (TG). Browning in hightemperature clearing was effectively prevented by TG. Scale bars, 3 mm. (D) High-temperature (80C) clearing of whole rat brain with and without TG. Scale bars, 6 mm. (E) Clearing of human and marmoset samples at 80C. Scale bars, 6 mm. (F) Rapid clearing of various organs at 80C with and without 50 mM SS. Cleared rat spinal cord is not RI-matched. Scale bars, 3 mm.

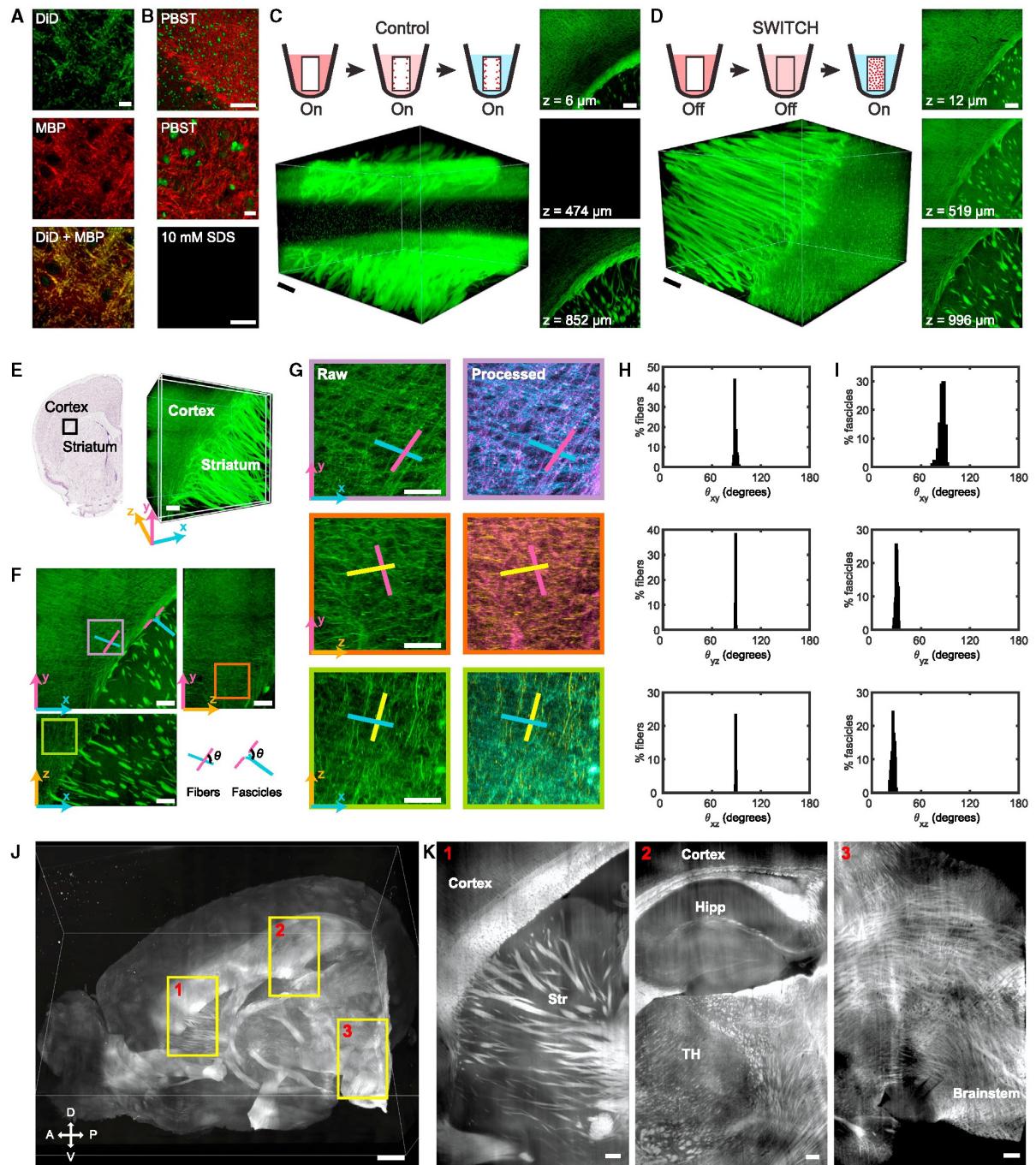
thioglycerol, effectively mitigate tissue browning during thermal clearing (Figures 2.4B–2.4D). Using thermal clearing with the reducing agents, we successfully cleared intact adult rat brains (2 weeks) as well as human (1 week) and marmoset samples (1 week), demonstrating the versatility and scalability of the method (Figures 2.4D and 2.4E). Clearing of various rodent organs was also demonstrated with lung, kidney, heart, liver, and spinal cord (Figure 2.4F). The efficacy of sodium sulfite as an anti-browning agent was seen across all tissues.

### 2.3.5 SWITCH Enables Visualization and Quantitative Analysis of Entire Myelinated Fiber Tracts

We also sought to apply SWITCH to characterizing myelinated fiber pathways in the brain. Visualizing and analyzing neural fibers with high-resolution light microscopy can provide valuable insights into many studies (Thomas et al. 2014; Zuccaro & Arlotta 2013), such as validating diffusion tensor imaging (DTI) and understanding the organizing principles of brain connectivity. Furthermore, quantitative analysis of myelinated fibers in 3D may benefit clinical studies and development of novel treatments for many demyelinating diseases (Steinman 1999), such as multiple sclerosis and transverse myelitis. However, current methods for myelinated fiber visualization require either genetic labeling or a large amount of costly antibodies, limiting their utility to animal tissues or small clinical samples.

We discovered that a subset of lipids preserved in SWITCH-processed tissues (Hopwood 1972; Rozemond 1969) allows lipophilic dyes to selectively visualize lipid-rich membranes (Schlessinger et al. 1977). In particular, we found that long-chain dialkylcarbocyanines robustly stain myelinated axons (Figure 2.5A). However, when we attempted to label an intact tissue using conventional methods, we could not achieve dye penetration deeper than 100  $\mu\text{m}$  because dye molecules were depleted as they rapidly associated with abundant targets in the outer layer (Figure 2.5C).

We hypothesized that SWITCH may enable rapid and uniform labeling of intact tissues by synchronizing the labeling reaction globally. We first screened a range of chemicals for controlling the binding kinetics of the lipophilic dye and discovered that 10 mM SDS effectively inhibits staining (Figure 2.5B). This result indicates that buffers containing 10 mM SDS might have a potential to be used as a “SWITCH-Off” buffer. Using an approach analogous to SWITCH-mediated GA fixation, we thought it might be possible to allow dye molecules to disperse uniformly throughout a sample in the SWITCH-Off buffer and then activate global probe-target binding with the SWITCH-On buffer (Figure 2.5D).



**Figure 2.5:** SWITCH Enables Visualization and Quantitative Analysis of Entire Myelinated Fiber Tracts.

(A) DiD and MBP staining on a SWITCH-processed mouse brain slice showing complete overlap between DiD and MBP. Scale bar, 10  $\mu$ m. (B) DiD staining with PBST or with PBS + 10 mM SDS buffer. DiD staining is completely inhibited in PBS + 10 mM SDS buffer. Green, syto16; red, DiD; scale bars, 100  $\mu$ m (top, bottom), 10  $\mu$ m (middle). (C) DiD staining of a 1-mm-thick mouse coronal block using PBST for 1.5 days at 37C. Only tissue surface is labeled. Scale bar, 200  $\mu$ m. (D) DiD staining of a 1-mm-thick mouse coronal block using SWITCH. The sample was first incubated in DiD, 10 mM SDS containing PBS buffer for 24 hr, then moved to PBST and incubated for 0.5 day at 37C. The whole sample is uniformly labeled. Scale bar, 200  $\mu$ m. (E) Volume image of a 1-mm-thick mouse brain coronal slice stained with DiD to visualize myelinated tracts acquired using a confocal microscope. The volume contains both the striatum and the cortex. Scale bar, 200  $\mu$ m. (F) Maximum intensity projection of the subvolume (illustrated in white in the volume image in (E)) shows fascicles from the striatum diverging at the corpus callosum and fibers near that area in the cortex forming a grid pattern. Scale bar, 200  $\mu$ m. (G) Enlarged images of the selected regions of interest in (F) shows the fibers in the cortex arranged in a grid pattern. Fibers are colorized based on orientation. Scale bar, 100  $\mu$ m. (H) Analysis of all the fibers in the entire volume shows that most fibers make an 89° intersection in xy and yz and an 88° intersection in xz. (I) Analysis of all the fascicles in the entire volume shows that they make an 87° turn in xy, a 26° turn in yz, and a 30° turn in xz. (J) Volume image of a mouse brain hemisphere stained with DiD to visualize myelinated tracts acquired using a custom-built light-sheet microscope. Scale bar, 1 mm. (K) Representative images showing individual fibers and fascicles in three different brain regions in (J). Str, striatum; Hipp, hippocampus; TH, thalamus. Scale bars, 200  $\mu$ m.

To test this approach, we first incubated a 1-mm-thick mouse brain block in PBST containing 10 mM SDS and lipophilic dyes for 24 hr at 37C (SWITCH-Off step). Then, we moved the tissue to PBST and incubated it for 3 hr at 37C (SWITCH-On step). The result was strikingly uniform labeling of all the myelinated axons within the sample (Figure 2.5D). Myelinated fibers were clearly visible throughout the depth while the control tissue showed signal only from the surface (Figure 2.5C).

We leveraged this fiber visualization capability to investigate how fibers and fascicles are organized in a mouse brain. Previous research has shown that fibers may be organized in 3D grids (Wedge et al. 2012). However, the structure of all of the individual fibers has not yet been studied at the microscopic resolutions and macroscopic scales necessary to visualize their 3D organization. To that end, we obtained a volume image of labeled myelinated fibers in a SWITCH-processed mouse brain coronal slice spanning from the cortex to the striatum (Figure 2.5E; Movies S2 and S3). This volume shows three main orientations of the fibers organized in a cubic grid: one radially projecting from the corpus callosum and two parallel to the corpus callosum. These three orientations are all orthogonal to one another (Figure 2.5F; Movie S3). The volume also shows fascicles that radiate from the striatum and diverge, almost at right angles, at the corpus callosum (Figure 2.5E; Movie S3). To quantify this finding in an unbiased manner, we determined the orientation of each of the fibers present in the volume and calculated the angles at which these fibers would intersect (Figure 2.5G). In all three dimensions, the fibers indeed oriented themselves approximately orthogonally to each other (Figure 2.5H). We used a similar approach to examine the fascicle orientations and found that they diverge almost orthogonally with respect to the corpus callosum in one of the axes (Figure 2.5I). These results are corroborated by the autocorrelation results (Figure S3 and S4). This finding

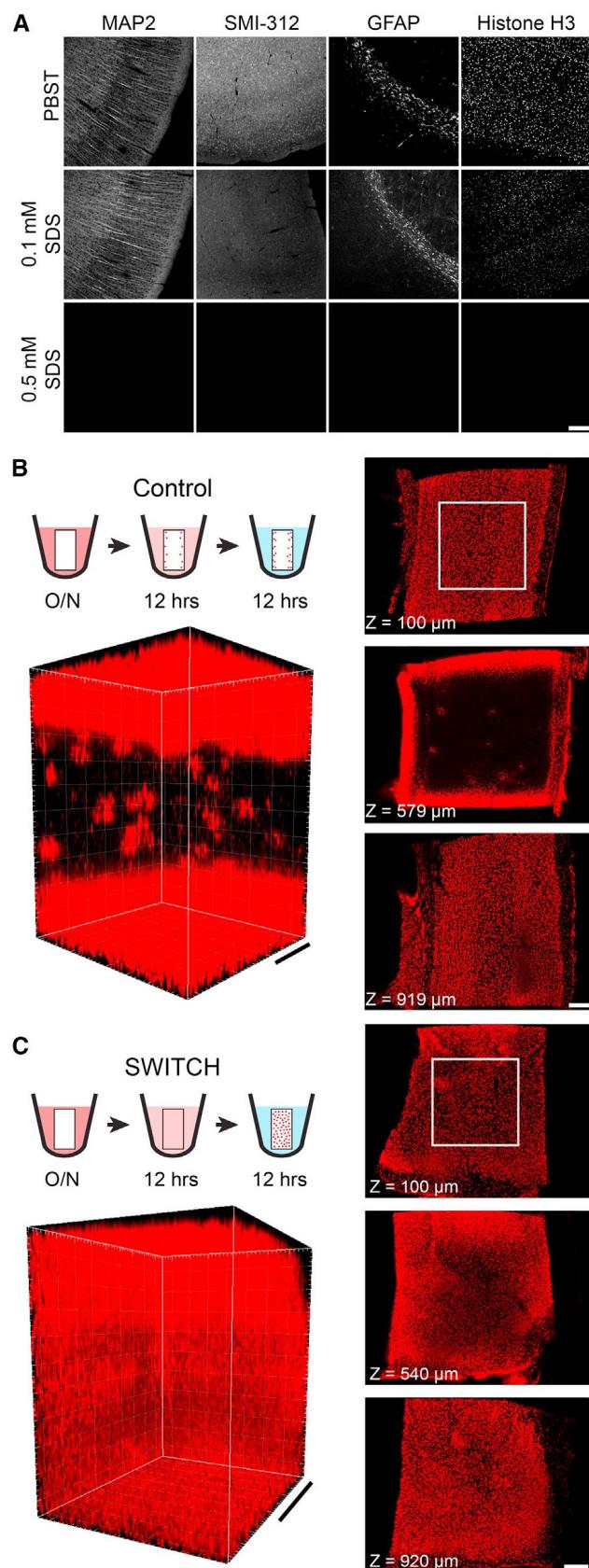
was made possible by the high-resolution and large-volume visualization capability of our method. A low-resolution approach would overlook the individual fibers while a low-volume approach would be unable to capture the entire connectional anatomy.

We then tested whether this application of SWITCH could be scaled to larger tissues. We applied the SWITCH approach for labeling an intact mouse hemisphere, but with 4 days of incubation in PBST containing 10 mM SDS and lipophilic dyes (SWITCH-Off step) and 1 day in PBST (SWITCH-On step). We imaged this larger volume using a custom-built, high-speed light-sheet microscope (Tomer et al. 2012; Tomer et al. 2014) within 2 hr and observed uniform labeling of all myelinated fibers across the entire tissue (Figure 2.5J; Movie S4). As demonstrated, the SWITCH-labeling approach is scalable to organ-scale tissues. Just by scaling the incubation time with respect to the tissue size, we were able to label the whole tissue. The cost of the dye molecules used for labeling the hemisphere was less than one dollar. We also demonstrated that this approach can be used for visualizing myelinated fibers in spinal cords (Movie S5). These results show that the SWITCH-labeling method can be used to uniformly label tissues ranging from a 1-mm-thick block to an entire hemisphere for quantitative analysis.

### 2.3.6 SWITCH Enables Scalable and Uniform Antibody Labeling

We then asked whether SWITCH-mediated labeling could be applied to the use of antibodies. We hypothesized that SDS could again be used as an effective inhibitor of antibody-antigen binding in small concentrations. Indeed, when we assayed for antibody labeling at various concentrations of SDS, we found that 0.5 to 1.0 mM was a high enough concentration to inhibit binding for many antibodies (Figure 2.6A).

Based on the results of our binding assay, we chose PBS containing 0.5 mM SDS as a SWITCH-Off buffer and PBST as a SWITCH-On buffer. We hypothesized that, because very little antibody-antigen binding is occurring in the SWITCH-Off condition, antibodies would effectively be able to diffuse to equilibrium throughout the sample more rapidly than in PBST, in which antibodies are rapidly depleted at the surface (Figure 2.6B). To test this, we attempted to label 1-mm-thick mouse brain blocks using anti-histone H3 antibodies. We labeled one sample using a 12 hr SWITCH-Off/12 hr SWITCH-On cycle and another using a standard immunohistochemistry protocol with 12 hr of primary antibody incubation in PBST followed by a 12 hr wash. For the SWITCH-On step, antibodies were not added to PBST. The result was a large increase in penetration depth and overall signal uniformity in the SWITCH sample relative to the control (Figures 2.6B and 2.6C; Movie S6).



**Figure 2.6:** SWITCH Increases Uniformity of Antibody Labeling in Thick Tissues.

(A) Antibody staining of cleared 100- $\mu$ m mouse brain sections in PBST and various concentrations of SDS in PBS. SDS effectively inhibits antibody-antigen binding in a concentration-dependent manner. Scale bar, 200  $\mu$ m. (B and C) Histone H3 staining of 1-mm-thick mouse cerebral cortex blocks in PBST (B) and using SWITCH (C). Control sample was incubated in antibody-containing PBST for 12 hr then washed for 12 hr. SWITCH sample was incubated in antibody-containing SWITCH-Off solution for 12 hr then washed in SWITCH-On solution for 12 hr. Sections from the top, middle, and bottom of the blocks are shown. 3D renderings were generated from the ROIs shown. SWITCH sample showed vast increase in uniformity of labeling compared to control. Scale bars, 150  $\mu$ m (B, left), 200  $\mu$ m (others).

## 2.4 Discussion

We have developed SWITCH, a simple method that enables scalable proteomic imaging of intact systems without requiring any specialized equipment or reagents. SWITCH is complementary to many pioneering technologies, each of which has its own unique advantages. For example, matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) and laser-ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) allow visualization of a large subset of proteins and other biomolecules without a priori knowledge of targets. Recent advances in imaging mass spectrometry combined with immunohistochemistry (IHC) have significantly improved resolution (Angelo et al. 2014; Giesen et al. 2014), which was limited in MALDI-MS and LA-ICP-MS. This approach remarkably demonstrated analysis of more than 100 targets at subcellular resolution.

Multiplexing strategies for IHC that rely on iterative staining and elution have been developed. Among several pioneering techniques is array tomography, which involves cutting a tissue sample into tens or hundreds of nanometer-thick sections for staining and imaging (Micheva et al. 2010). These sections can be repeatedly washed and stained for probing different proteins. This powerful method yields subcellular resolution images of a small volume of tissue with fairly high multiplexing capability. Although these advanced technologies enable new approaches in studying complex biological systems, these methods require specialized equipment and are, therefore, difficult to implement in most labs.

With the aim of developing a simple and scalable method for proteomic imaging of both large animal and human samples, we first needed to devise the SWITCH method for controlling a broad range of chemical reactions in tissue processing to achieve uniform sample treatment regardless of tissue size and type. SWITCH dynamically modulates chemical reaction kinetics to synchronize the reaction time between molecules throughout the system. This strategy enables all endogenous molecular targets in a large intact tissue to experience similar reaction conditions (time and concentration). As a result, large tissues can be uniformly processed.

The SWITCH approach takes advantage of the way certain chemicals can be reversibly and rapidly changed by simply modulating their surrounding environment. For instance,

in the GA-tissue-gelling step, we were able to decrease the rate of GA-biomolecule crosslinking by two orders of magnitude by using pH 3 buffer, because primary amine groups in endogenous biomolecules are protonated at low pH and the resulting charged amine cannot react with GA (Hopwood 1967). This pH-dependent reactivity means that after uniformly dispersing GA in a tissue at low pH, we can “switch-on” inactivated amine groups by changing the amine’s surrounding environment to a neutral-pH buffer. At neutral pH, charged amine groups are rapidly deprotonated and become reactive. In the case of human samples or animal samples that were previously PFA-fixed for a different purpose, this simple strategy enables all the endogenous biomolecules in a large intact tissue to simultaneously experience similar GA-fixation/gelling condition. PFA-fixed tissues can withstand treatment at low pH while GA molecules are introduced. In the case of non-fixed samples, we recommend that they first be fixed with PFA before exposure to acidic conditions. If perfusion is possible, it is the recommended method of sample preservation.

Uniform GA-tissue-gel formation is a crucial first step toward our goal. Fixation of large samples via traditional immersion is unlikely to uniformly preserve them because highly reactive GA molecules are depleted within the outer layers of a sample. This presents a significant problem for iterative staining-based methods that rely on the removal of imaged probes using harsh elution conditions, because non-uniform preservation results in non-uniform loss of structure and molecules throughout the process. As demonstrated, our pH-SWITCH strategy ensures exceptionally uniform preservation of biological tissues that cannot be perfused (e.g., banked human clinical samples), meeting the requirements of proteomic imaging and quantitative phenotyping.

It has been noted that fixation with GA results in an increase in broad spectrum autofluorescence. While this autofluorescence has been low enough to allow quantitative analysis, it could be problematic in visualizing targets with low copy number. We investigated the use of sodium borohydride as a method of reducing autofluorescence, but found that the tissue damage resulting from this incubation procedure offset any benefits obtained from the modest decrease in autofluorescence that we were able to observe (Figure S5).

The use of reducing agents has allowed us to eliminate the issue of tissue browning during high-temperature clearing, but we also observed that excessive use of these chemicals may cause gradual tissue weakening. This is likely due to the reduction of disulfide linkages that maintain the tertiary structure of proteins within a sample, resulting in increased protein denaturation. Protein denaturation may lead to reduced sample antigenicity, but we have not found this to be an issue when using conservative amounts of reducing agents. Additionally, due to the instability of mRNA at elevated temperatures, this method of rapid clearing is not compatible with methods that require the preservation of mRNA

(Figure S6).

Multiplexed imaging requires software to warp each experiment into a common coordinate system despite the subtle physical differences between each staining round. Variance can come in the form of rigid body changes (rotation, translation, and scale), illumination artifacts, stain quality, and tissue degradation. We observed that a feature-based algorithm gives maximum robustness across these sources of variance at the cost of increased computational requirements—a reasonable trade given the declining costs of such resources. To simplify the process, gross rigid alignments (i.e., rotating the tissue 180 degrees) are still best handled by human eye before the data are passed to the algorithm to achieve the cellular-scale registration.

SWITCH can provide a reliable way to obtain integrated high-dimensional information from intact biological samples. Using the cross-talk-free dataset, we successfully performed non-biased combinatorial expression analysis of a single human clinical tissue to unequivocally identify diverse cell-types based on their distinct protein expression patterns. Our quantitative analysis shows that CR+/PV+ cells do not exist within the examined volume of the human V2 cortex. The same finding was reported in mouse visual cortex (Gonchar et al. 2008), but such co-expression patterns among calcium-binding proteins may differ among brain regions and between individuals and species (Anelli & Heckman 2006), which, therefore, calls for more comprehensive large-scale investigation.

We observed many NeuN-negative interneurons. NeuN, a neuron-specific RNA-binding protein known as Rbfox3 protein (Kim et al. 2009), has been widely used as a pan-neuronal marker for statistical analysis of many types of mature neurons (Baleriola et al. 2014; Pickrell et al. 2015). Only a few types of neurons are exceptions, such as cerebellar Purkinje cells, olfactory bulb mitral cells, and retinal photoreceptor cells (Mullen et al. 1992). However, even though we applied strict criteria to prevent weak NeuN+ cells from being identified as NeuN- cells, substantial portions of CB+, CR+, and PV+ neurons were still NeuN- while all SMI+ neurons were NeuN+. This result is supported by a recent report that some CR+ are not NeuN+, and CR and NeuN immunoreactivities have a negative correlation in the avian brainstem (Bloom et al. 2014). Likewise, in our experiments on human visual association cortex, cells with strong immunoreactivity against calcium-binding protein markers were frequently negative or very weakly positive for NeuN. These findings, together with a series of exceptional reports such as those on NeuN+ cultured astrocytes (Darlington et al. 2008) and GFAP+ neuron-like cells (Oka et al. 2015), indicate that classical cell-type markers, particularly NeuN, may need to be used more carefully in light of their selectivity and function.

The SWITCH method has the potential to modulate a wide range of probe-target binding reactions. Probe-target interactions are governed by a multiplicity of non-covalent bonds

such as hydrogen bonds, electrostatic forces, van der Waals bonds, and hydrophobic interactions (Mian et al. 1991). These weak forces can be effectively controlled by changing the surrounding chemical environment (e.g., ionic strength, pH, chemical additive, and temperature) (Kamata et al. 1996). For instance, we discovered that the addition of SDS alone, in different concentrations, can completely inhibit lipophilic dye-target and antibody-antigen binding reactions.

The SWITCH method's unique uniform-labeling capability enables quantitative analysis of large tissues that was previously only possible for thin tissue sections. Quantitative analysis relies heavily on signal intensity and SNR. Non-uniform or heterogeneous labeling would prohibit or, even worse, bias the analysis. While post hoc image processing methods could correct for small gradients in labeling (or imaging), large gradients caused by non-uniform labeling, where the surface of the tissue is saturated while the core is mostly unlabeled, would preclude image recovery. If the labeling is heterogeneous, the resulting data would be heavily biased, and no image processing methods could salvage such data in a fair way. This is why quantitative analysis of non-uniformly labeled tissues is a great challenge. However, tissues labeled using SWITCH exhibit uniform signal intensity and SNR throughout the tissue. Such a clear dataset lends itself well to quantitative analysis.

Although SWITCH enables processing of large samples, the speed of labeling is still fundamentally limited by passive diffusion. This is not of concern for smaller samples or even single-round investigation of large samples, but multiplexed imaging of large samples becomes impractical as a result, potentially taking months or years to collect the range of desired data. Recently developed methods of stochastic electrotransport (Kim et al. 2015) could potentially be combined with SWITCH to facilitate these experiments.

Together with its simplicity, scalability, and broad applicability, our data suggest that SWITCH provides access to high-dimensional multi-scale information that may help to understand health and disease from molecules to cells to entire systems.

## 2.5 Experimental Procedures

### 2.5.1 Mice

Young adult male and female C57BL/6 and Thy1-eGFP-M mice were housed in a reverse 12 hr light/dark cycle with unrestricted access to food and water. All experimental protocols were approved by the MIT Institutional Animal Care and Use Committee and Division of Comparative Medicine and were in accordance with guidelines from the National Institute of NIH.

### 2.5.2 SWITCH-Mediated Tissue Preservation

PFA-fixed human samples were washed in a solution consisting of 50% PBS titrated to pH 3 using HCl, 25% 0.1 M HCl, and 25% 0.1 M potassium hydrogenphthalate (KHP). This wash solution was then replaced with fresh solution with the addition of 4%–10% GA. The samples were then incubated in this pH 3 solution at 4C for 2 days with gentle shaking. The solution was then replaced with PBS with the addition of 1%–4% GA and the sample was again allowed to incubate for 2 days at 4C and 2–7 hr at 37C with gentle shaking. The sample was then washed in PBS at room temperature (RT) for 1 day with gentle shaking. After washing, reactive GA within the sample was inactivated by incubation in a solution consisting of 4% glycine and 4% acetamide for 1 day at 37C with gentle shaking. Finally, the sample was washed for 1 day in PBS at RT with gentle shaking.

### 2.5.3 Passive Clearing with Thermal Energy

Aqueous clearing solution containing 200 mM SDS, 10 mM lithium hydroxide, 40 mM boric acid, and a variable amount of anti-browning agent (i.e., 0–50 mM sodium sulfite or 0%–0.5% w/v 1-thioglycerol) was titrated to pH 9 using sodium hydroxide before use. Samples were incubated at 60–80C until clear using Easy-Passive (EP-1001; Live Cell Instrument) or a water bath.

### 2.5.4 Sample Delabeling

Imaged samples were delabeled in clearing solution at 60–80C (elution condition) for 1–2 days for large samples and O/N for thin samples.

### 2.5.5 SWITCH-Mediated Fluorescent Labeling

Samples were incubated in SWITCH-Off solution (0.5 or 10 mM SDS in PBS) O/N with gentle shaking at 37C and transferred to a fresh volume of SWITCH-Off solution (containing molecular probes) just enough to cover the sample. Samples were incubated at 37C with gentle shaking and times were scaled with sample size. Samples were then transferred to a large volume of PBST (SWITCH-On) and incubated at 37C with gentle shaking.

## **2.6 Supplemental Information**

Supplemental Information includes Supplemental Experimental Procedures, six figures, two tables, and six movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.11.025> as well as in the appendix.

## **2.7 Author Contributions**

K.C. conceived the SWITCH idea. E.M., J.H.C, D.G., T.K., J.S., and K.C. designed the experiments. E.M. designed and performed multiplexed staining and imaging experiments. D.G. and H.S.S. developed and implemented the registration algorithm. T.K. performed all analysis related to the human visual cortex sample with input from K.C., M.P.F., and H.S.S. J.H.C. performed all analysis related to DiD datasets. E.M., N.B., and S.-Y.K. performed antibody compatibility testing. E.M., J.H.C, N.B., A.H., and J.S. performed testing of cross-linked protein gels with various fixatives. E.M. performed protein loss and structural preservation experiments. E.M., J.S., and N.B. designed and performed experiments relating to thermal clearing and use of anti-browning agents. E.M. and M.M. performed DiD staining experiments. E.M. and Y.-G.P. performed antibody SWITCH experiments. H.C. designed and built the light-sheet microscope, and J.H.C. developed the software. J.-Y.P. provided spinal cord samples. Y.-G.P. performed mRNA experiments. M.P.F. provided human samples. H.S.S., V.J.W., and M.P.F. provided helpful discussion regarding the manuscript. S.V. provided experimental support. E.M., J.H.C., D.G., T.K., J.S., S.-Y.K., V.J.W., and K.C. wrote the manuscript. K.C. supervised all aspects of the work.

## **2.8 Acknowledgments**

We thank the entire Chung laboratory for support and helpful discussions. S.-Y.K. was supported by the Simons Postdoctoral Fellowship and the Life Sciences Research Foundation. K.C. was supported by Burroughs Wellcome Fund Career Awards at the Scientific Interface, the Searle Scholars Program, the Michael J. Fox Foundation, DARPA, the JPB Foundation (PIIF and PNDRF), and NIH (1-U01-NS090473-01). M.P.F. was supported in part by the Massachusetts Alzheimer Disease Research Center (5 P50 AG005134). Resources that may help enable general users to establish the methodology are freely available online. ([www.chunglabresources.org](http://www.chunglabresources.org)).

## 2.9 References

- Alivisatos, A.P. et al., 2013. The brain activity map. *Science*, 339(6125), pp.1284–1285.
- Anelli, R. & Heckman, C.J., 2006. The calcium binding proteins calbindin, parvalbumin, and calretinin have specific patterns of expression in the gray matter of cat spinal cord. *Journal of Neurocytology*, 34(6), pp.369–385.
- Angelo, M. et al., 2014. Multiplexed ion beam imaging of human breast tumors. *Nature Medicine*, 20(4), pp.436–442. Available at: <http://dx.doi.org/10.1038/nm.3488>.
- Baleriola, J. et al., 2014. Axonally synthesized ATF4 transmits a neurodegenerative signal across brain regions. *Cell*, 158(5), pp.1159–1172. Available at: <http://dx.doi.org/10.1016/j.cell.2014.07.001>.
- Bloom, S., Williams, A. & MacLeod, K.M., 2014. Heterogeneous calretinin expression in the avian cochlear nucleus angularis. *JARO - Journal of the Association for Research in Otolaryngology*, 15(4), pp.603–620.
- Campbell, M.J. & Morrison, J.H., 1989. Monoclonal antibody to neurofilament protein (SMI-32) labels a subpopulation of pyramidal neurons in the human and monkey neocortex. *Journal of Comparative Neurology*, 282(2), pp.191–205.
- Chung, K. et al., 2013. Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449), pp.332–7.
- Chung, K. et al., 2013. Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449), pp.332–337. Available at: <http://www.nature.com/doifinder/10.1038/nature12107>.
- Darlington, P.J. et al., 2008. Widespread immunoreactivity for neuronal nuclei in cultured human and rodent astrocytes. *Journal of Neurochemistry*, 104(5), pp.1201–1209.
- Defelipe, J. et al., 1999. Distribution and patterns of connectivity of interneurons containing calbindin, calretinin, and parvalbumin in visual areas of the occipital and temporal lobes of the macaque monkey. *Journal of Comparative Neurology*, 412(3), pp.515–526.
- De Sousa, A.A. et al., 2010. Comparative cytoarchitectural analyses of striate and extrastriate areas in hominoids. *Cerebral Cortex*, 20(4), pp.966–981.
- Friedman, M., 1996. Food browning and its prevention: An overview. *Journal of Agricultural and Food Chemistry*, 44(3).
- Giesen, C. et al., 2014. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods*, 11(4), pp.417–422.

Gonchar, Y., Wang, Q. & Burkhalter, A., 2008. Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. *Frontiers in Neuroanatomy*, 1(3).

Hopwood, D., 1967. Some aspects of fixation with glutaraldehyde. A biochemical and histochemical comparison of the effects of formaldehyde and glutaraldehyde fixation on various enzymes and glycogen, with a note on penetration of glutaraldehyde into liver. *Journal of anatomy*, 101(Pt 1), pp.83–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6047703>.

Hopwood, D., 1972. Theoretical and practical aspects of glutaraldehyde fixation. *The Histochemical Journal*, 4(4), pp.267–303.

Hopwood, D., 1970. The reactions between formaldehyde, glutaraldehyde and osmium tetroxide, and their fixation effects on bovine serum albumin and on tissue blocks. *Histochemistry*, 24(1), pp.50–64.

Kamata, N. et al., 1996. Comparison of pH and ionic strength dependence of interactions between monoclonal antibodies and bovine  $\beta$ -lactoglobulin. *Bioscience, Biotechnology and Biochemistry*, 60(1), pp.25–29.

Kasthuri, N. et al., 2015. Saturated Reconstruction of a Volume of Neocortex. *Cell*, 162(3), pp.648–661.

Keller, S., Heerklotz, H. & Blume, A., 2006. Monitoring lipid membrane translocation of sodium dodecyl sulfate by isothermal titration calorimetry. *Journal of the American Chemical Society*, 128(4), pp.1279–1286.

Kim, K.K., Adelstein, R.S. & Kawamoto, S., 2009. Identification of neuronal nuclei (NeuN) as Fox-3, a new member of the Fox-1 gene family of splicing factors. *Journal of Biological Chemistry*, 284(45), pp.31052–31061.

Kim, S.-Y. et al., 2015. Stochastic electrotransport selectively enhances the transport of highly electromobile molecules. *Proceedings of the National Academy of Sciences*, 112(46), pp.E6274–E6283. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1510133112>.

Leuba, G., Kraftsik, R. & Saini, K., 1998. Quantitative distribution of parvalbumin, calretinin, and calbindin D- 28k immunoreactive neurons in the visual cortex of normal and Alzheimer cases. *Experimental Neurology*, 152(2), pp.278–291.

Manzo, C. et al., 2014. PSF decomposition of nanoscopy images via Bayesian analysis unravels distinct molecular organization of the cell membrane. *Scientific Reports*, 4(April).

- McCaslin, A.F.H. et al., 2011. In vivo 3D morphology of astrocyte-vasculature interactions in the somatosensory cortex: Implications for neurovascular coupling. *Journal of Cerebral Blood Flow and Metabolism*, 31(3), pp.795–806.
- Mian, I.S., Bradwell, A.R. & Olson, A.J., 1991. Structure, function and properties of antibody binding sites. *Journal of Molecular Biology*, 217(1), pp.133–151.
- Micheva, K.D. et al., 2010. Single-synapse analysis of a diverse synapse population: Proteomic imaging methods and markers. *Neuron*, 68(4), pp.639–653. Available at: <http://dx.doi.org/10.1016/j.neuron.2010.09.024>.
- Mullen, R.J., Buck, C.R. & Smith, A.M., 1992. NeuN, a neuronal specific nuclear protein in vertebrates. *Development*, 116(1), pp.201–211.
- Oka, Y., Ye, M. & Zuker, C.S., 2015. Thirst driving and suppressing signals encoded by distinct neural populations in the brain. *Nature*, 520(7547), pp.349–352.
- Pickrell, A.M. et al., 2015. Endogenous Parkin Preserves Dopaminergic Substantia Nigra Neurons following Mitochondrial DNA Mutagenic Stress. *Neuron*, 87(2), pp.371–381. Available at: <http://dx.doi.org/10.1016/j.neuron.2015.06.034>.
- Rah, J.C. et al., 2013. Thalamocortical input onto layer 5 pyramidal neurons measured using quantitative large-scale array tomography. *Frontiers in Neural Circuits*, 7(NOV), pp.1–16.
- Renier, N. et al., 2014. iDISCO: A simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell*, 159(4), pp.896–910. Available at: <http://dx.doi.org/10.1016/j.cell.2014.10.010>.
- Richardson, D.S. & Lichtman, J.W., 2015. Clarifying Tissue Clearing. *Cell*, 162(2), pp.246–257.
- Roozemond, C., 1969. The effect of fixation with formaldehyde and glutaraldehyde on the composition of phospholipids extractable from rat hypothalamus., 17(7).
- Schlessinger, J. et al., 1977. Lateral transport of a lipid probe and labeled proteins on a cell membrane. *Science*, 195(4275), pp.307–309.
- Steinman, L., 1999. Assessment of animal models for MS and demyelinating disease in the design of rational therapy. *Neuron*, 24(3), pp.511–514.
- Sung, H.W. et al., 1996. Cross-linking characteristics of biological tissues fixed with monofunctional or multifunctional epoxy compounds. *Biomaterials*, 17(14), pp.1405–1410.

Susaki, E.A. et al., 2014. Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell*, 157(3), pp.726–739. Available at: <http://dx.doi.org/10.1016/j.cell.2014.03.042>.

Thomas, C. et al., 2014. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proceedings of the National Academy of Sciences of the United States of America*, 111(46), pp.16574–16579.

Tomer, R. et al., 2012. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods*, 9(7), pp.755–763. Available at: <http://dx.doi.org/10.1038/nmeth.2062>.

Tomer, R. et al., 2014. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nature Protocols*, 9(7), pp.1682–1697.

Wedgeen, V.J. et al., 2012. The Geometric Structure of the Brain Fiber Pathways. *Science*, 335(March), pp.1628–1634.

Yuste, R., 2015. From the neuron doctrine to neural networks. *Nature Reviews Neuroscience*, 16(8), pp.487–497. Available at: <http://dx.doi.org/10.1038/nrn3962>.

Zuccaro, E. & Arlotta, P., 2013. The quest for myelin in the adult brain. *Nature Cell Biology*, 15(6), pp.572–575. Available at: <http://dx.doi.org/10.1038/ncb2750>.

# Chapter 3

## Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues

Taeyun Ku\*, **Justin Swaney\***, Jeong-Yoon Park\*, Alexander Albanese, Evan Murray, Jae Hun Cho, Young-Gyun Park, Vamsi Mangena, Jiapei Chen, and Kwanghun Chung<sup>1</sup>

*Reproduced from DOI: <https://doi.org/10.1038/nbt.3641>*

### 3.1 Summary

The biology of multicellular organisms is coordinated across multiple size scales, from the subnanoscale of molecules to the macroscale, tissue-wide interconnectivity of cell populations. Here we introduce a method for super-resolution imaging of the multiscale organization of intact tissues. The method, called magnified analysis of the proteome (MAP), linearly expands entire organs fourfold while preserving their overall architecture and three-dimensional proteome organization. MAP is based on the observation that preventing crosslinking within and between endogenous proteins during hydrogel-tissue hybridization allows for natural expansion upon protein denaturation and dissociation. The expanded tissue preserves its protein content, its fine subcellular details, and its organ-scale intercellular connectivity. We use off-the-shelf antibodies for multiple rounds of immunolabeling and imaging of a tissue's magnified proteome, and our experiments

---

<sup>1</sup>\* indicates co-first authorship.

demonstrate a success rate of 82% (100/122 antibodies tested). We show that specimen size can be reversibly modulated to image both inter-regional connections and fine synaptic architectures in the mouse brain.

## 3.2 Introduction

Biological systems such as the mammalian brain consist of thousands of distinct cell types forming highly interconnected functional networks (White et al. 1986; DeFelipe 2010; Helmstaedter 2013; Craddock et al. 2013; Oh et al. 2014). Understanding how these diverse cells interact to generate system-level responses is essential for many fields of biology. Deciphering these complex interactions requires tools that can simultaneously characterize the molecular identity and fine sub-cellular architectures of individual cells as well as their system-level connectivity, because these properties together determine many cell functions. To date, technologies provide only a subset of the required multilevel information.

Proteomic imaging techniques can potentially provide multilevel information in both animals and human samples (Phizicky et al. 2003; Hanash 2004; Agato et al. 2004; Berglund et al. 2008; Uhlen et al. 2010). The proteome is an ideal substrate for the integrated analysis of functional components owing to proteins' unmatched diversity, functional roles, and distinct subcellular localization at single-molecule precision. The existing large antibody libraries (close to 100,000 antibodies), once rigorously validated, could enable the detection of more than 70% of the human proteome and corresponding cellular architectures (Uhlen, M et al 2010).

For example, multiplexed proteomic imaging techniques (e.g., array tomography and SWITCH) can provide molecular details for individual cells and may allow reconstruction of their surrounding tissue environment (Micheva & Smith 2007; Kay et al. 2013; Murray et al. 2015). Emerging intact tissue clearing approaches preserve proteins and the continuity of neural fibers, which may enable reconstruction of immunolabeled neural architectures (Dodd et al. 2007; Richardson & Lichtman 2015). Super-resolution imaging of immunolabeled thin tissue sections has been successfully used to characterize minute subcellular structures (e.g., synapses) (Rust et al. 2006; Betzig et al. 2006; Hess et al. 2006; Maglione & Sigrist 2013). In addition to conventional super-resolution approaches, Chen et al. (2015) have recently demonstrated that protease digestion of a hydrogel-tissue hybrid homogenizes its mechanical characteristics and allows approximately fourfold linear expansion of the hybrid. Using this approach, termed expansion microscopy (ExM), they demonstrated super-resolution imaging of thin tissue sections with custom-made antibody probes and diffraction-limited microscopes (Chen et al. 2015).

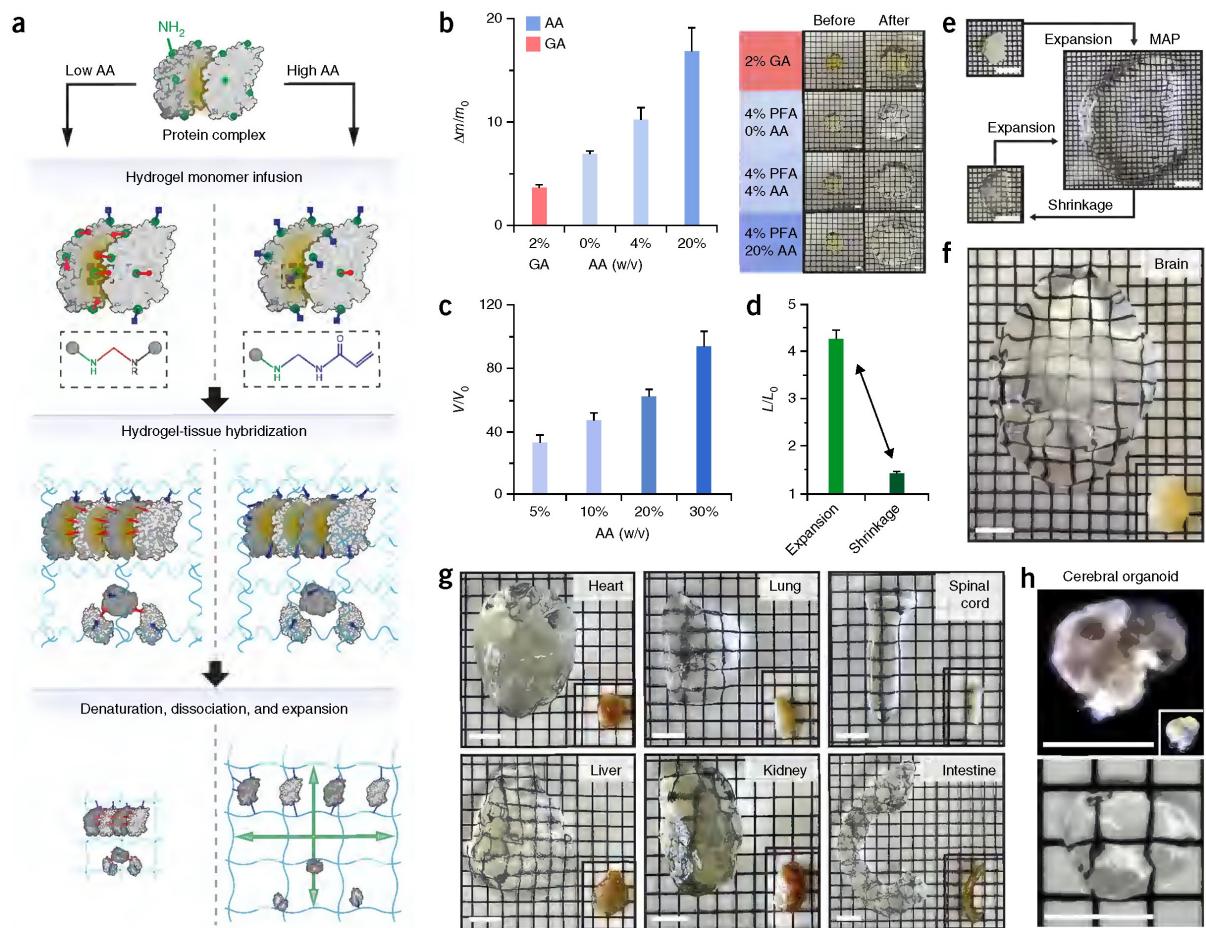
Despite the success of these new technologies, integrated reconstruction of the fine subcellular architectures, molecular details, and intercellular connectivity of diverse cell types in large-scale biological systems remains an unmet goal in biology. For instance, the protease digestion step in ExM causes a loss of proteins, which limits the number of protein structures that can be imaged in the same sample. In addition, intercellular connectivity is largely lost in ExM as it requires tissue sectioning to allow immunolabeling of an unprocessed tissue with limited permeability. By contrast, emerging whole-tissue clearing methods preserve neural connectivity and endogenous proteins, but accurate reconstruction of neural architectures remains challenging owing to the low resolution of diffraction-limited microscopy.

Here we present MAP, a method that enables multiscale proteomic imaging of intact biological systems for combined extraction of the molecular identity, subcellular architectures, and intercellular connectivity of diverse cell types within a single tissue. MAP preserves both the three-dimensional (3D) proteome content and organization and the organ-wide cellular connectivity within an intact tissue-hydrogel hybrid, while rendering it reversibly size-adjustable up to four- to fivefold for multiresolution imaging.

### 3.3 Results

#### 3.3.1 Organ-scale tissue expansion without loss of proteins

The key to MAP is to prevent intra- and interprotein crosslinking during the hydrogel-tissue hybridization step (Chung, J. Wallace, S.-Y. Kim, Kalyanasundaram, Aaron S Andelman, et al. 2013), and then to dissociate and denature proteins to allow natural expansion of the hybrid (Fig. 3.1a). We hypothesized that a high concentration of acrylamide monomers might effectively prevent protein crosslinking by quenching reactive methylols formed by the protein-formaldehyde reaction. If the acrylamide monomer concentration is low, the reactive methylols would react with amide groups within the same protein or adjacent proteins to form methylene bridges (Fig. 3.1a, left column) (Puchtler & Meloan 1985). Such intra- and interprotein crosslinking would prevent dissociation of protein complexes and limit subsequent tissue expansion (Sung et al. 1996). With increased acrylamide concentration (Fig. 3.1a, right column), methylols might preferentially react with excess acrylamide monomers, effectively reducing interprotein crosslinking while maximally tethering individual proteins to an expandable hydrogel mesh. Dissociation and denaturation of protein complexes with heat and anionic surfactant then facilitates natural expansion of the hydrogel-tissue hybrid while preserving overall spatial organization of the proteome.



**Figure 3.1:** Magnified and accessible 3D proteome library of whole intact organs. (a) Comparison of hydrogel-tissue hybridization chemistry, subsequent protein denaturation, and hybrid expansion between low (left column) and high (right column) concentrations of acrylamide (AA; blue) infusion. High-concentration AA prevents intra- and interprotein crosslinking during the hydrogel-tissue hybridization step by quenching reactive methylols formed from amine residues (green) reacting with formaldehyde (red). Dissociation and denaturation of non-crosslinked protein complexes allows natural expansion of the hybrid. In low-concentration AA (left column), the reactive methylols react with available amine or amide groups within the same protein or adjacent proteins and form methylene bridges (red line). Such intra- and interprotein crosslinking prevents complete denaturation and dissociation of the proteins, and thereby limits subsequent tissue expansion. (b) Expansion after denaturation of BSA tissue phantom post-fixed in several AA concentrations. Representative images of tissue phantoms before and after expansion are shown. One-way ANOVA was performed (\*\*P < 0.001). Error bars show s.d. (n = 3). Scale bars, 5 mm. (c) Expansion of 1-mm-thick coronal mouse brain slices perfused with several AA concentrations. One-way ANOVA was performed (P < 0.001). V<sub>0</sub>, volume before expansion; V, volume after expansion. Error bars show mean ± s.d. (n = 6). (d) Average diameter of 1-mm-thick slices relative to the initial diameter after expansion and subsequent shrinkage. Slices were allowed 24 h for both expansion and shrinkage (n = 6). L<sub>0</sub>, mean length of original brain; L, mean length after expansion and shrinkage. (e) Representative photos showing expansion and shrinkage of a 1-mm-thick coronal block. Starting from the top left, the original brain section, the expanded state, and the shrunken state are shown. (f,g) MAP applied to a whole mouse brain and other organs. The entire process from perfusion to full expansion took 7 d. Compared with the original organ size (pictures at right bottom corner), the final expansion of the brain showed a more than fourfold increase in length. (h) MAP applied to cerebral organoid. Top, darkfield images; bottom, brightfield image. Scale bars (e–h), 10 mm.

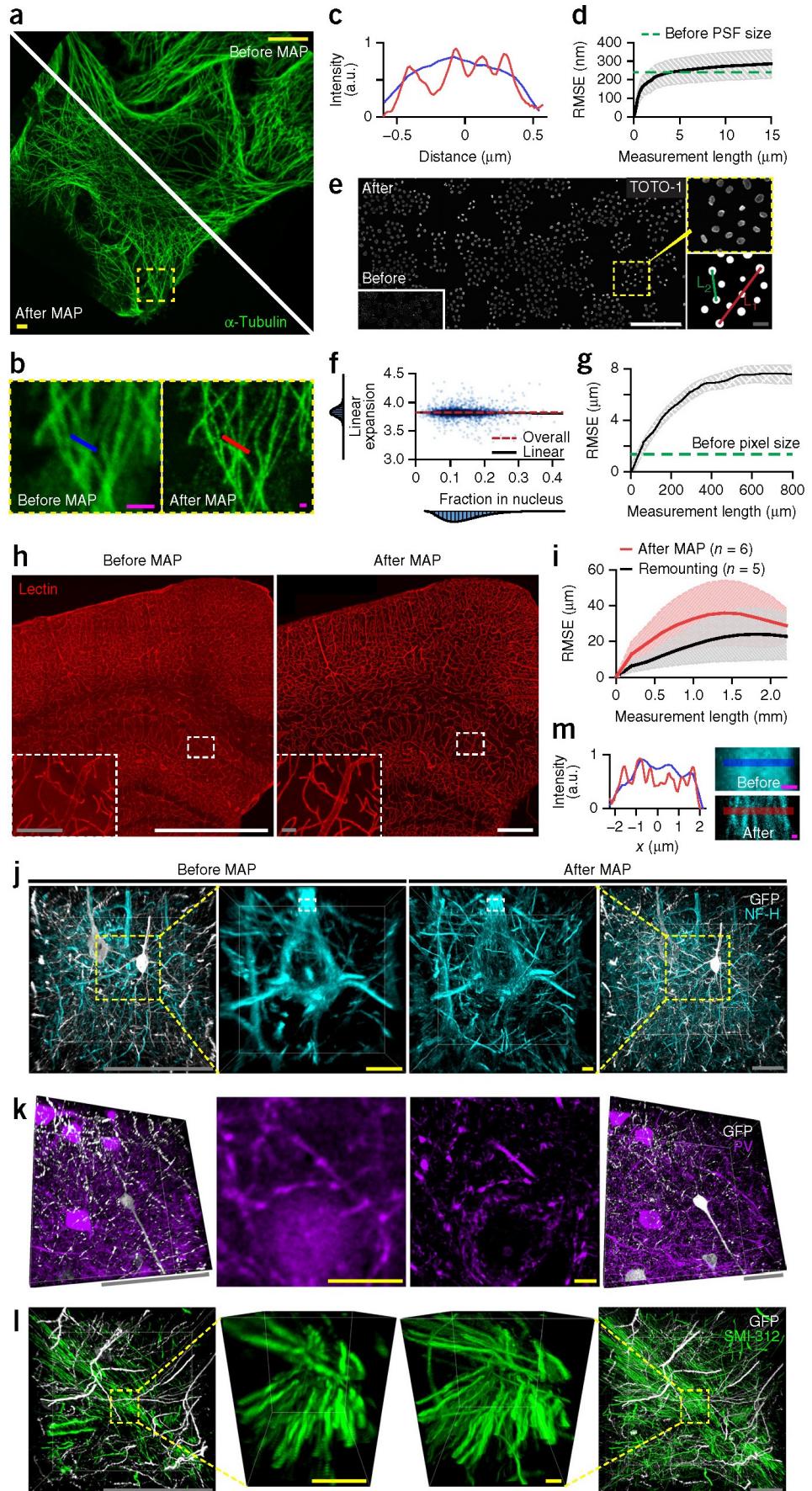
To test our hypothesis, we first measured the effect of acrylamide concentration on tissue expansion. We prepared albumin-containing tissue phantoms and post-fixed them in 4% paraformaldehyde (PFA) with different concentrations (0–20%) of acrylamide. We incubated the tissue phantoms in detergent solution at 95 °C for 1 h to denature and disrupt protein aggregates. As expected, phantoms fixed in higher concentrations of acrylamide showed higher degrees of expansion in water (Fig. 3.1b). We observed a similar trend in mouse brain tissues that were perfused with different concentrations of acrylamide, polymerized, denatured, and expanded (Fig. 3.1c). Thus, we used high concentrations of acrylamide in order to maximize expansion during MAP in all subsequent experiments.

Using this approach, we achieved a fourfold linear expansion of a whole mouse brain within 7 d without protease treatment (Fig. 3.1f). Tissue expansion was reversible and tunable using buffers with different salt concentrations and osmolarities (Fig. 3.1d,e). This method is applicable to other organs including heart, lung, spinal cord, liver, intestine, and kidney, and also cerebral organoids (Fig. 3.1g,h and Supplementary Fig. 1).

### 3.3.2 Preservation of multiscale architectures

We next asked whether MAP retains multiscale structural information and enables super-resolution imaging with diffraction-limited microscopes. To estimate the amount of distortion incurred from expansion, we imaged gel-embedded cultured cells before and after MAP processing (Fig. 3.2a). At the subcellular scale, MAP expansion improved visualization of microtubules and allowed imaging of single tubular structures (Fig. 3.2b,c). The estimated distortion error (root-mean-square error, RMSE) was less than 3% of measured length at both the subcellular scale (Fig. 3.2d) and the multicellular scale (Fig. 3.2g). The degree of gel expansion was not a function of cell density, as indicated by the consistent local expansion within differently populated cell clusters (Fig. 3.2e,f).

To test whether MAP preserves multiscale tissue architectures, we imaged a 100- $\mu$ m-thick mouse brain block labeled with lectin before and after MAP (Fig. 3.2h). The distortion analysis showed less than 4% error (Fig. 3.2i), which was at most a twofold increase, compared to the inevitable distortion from sample mounting for imaging (Fig. 3.2i, remounting). Both the intra-vascular space with low protein concentration and the extra-vascular space with high protein concentration showed similar degrees of expansion (Fig. 3.2h, insets), suggesting that the effect of protein concentration and cell density on gel expansion is minimal. Diffraction-limited microscopy revealed fine 3D details of immunolabeled cells such as cytoskeletal filament structures (Fig. 3.2j,k,m) and better-resolved axonal fibers (Fig. 3.2l) after the MAP process. These results together indicate that MAP preserved multiscale structural information of biological samples and enabled



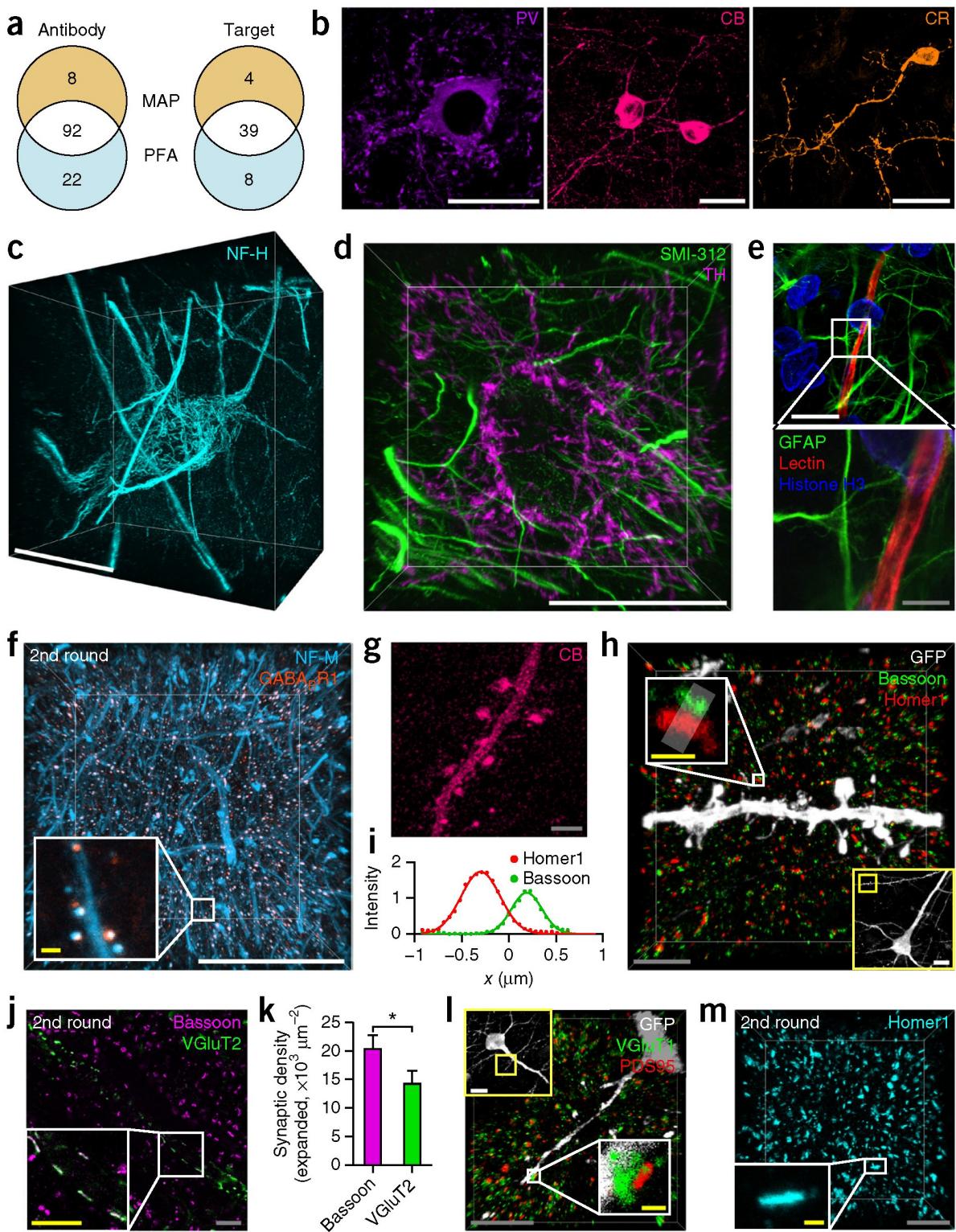
**Figure 3.2:** Comparison of multiscale architectures before and after MAP processing.

(a) HeLa cells stained for alpha-tubulin and imaged before and after MAP processing. (b) Fine tubulin structures imaged before and after expansion, respectively. (c) Intensity profiles along the paths indicated in b based on the pre-expansion size. (d) Root-mean-squared measurement error (RMSE) of tubulin images before and after MAP (black line, mean; filled area, s.d.; n = 4 samples). (e) Morphological image processing of gel-embedded HeLa cells before and after MAP processing. TOTO-1 was used to label nucleic acids. L1 and L2 represent the connecting line segments between two randomly chosen nuclei. (f) Local pairwise expansion ratios versus the fraction of the connecting line segment inside of nuclei (black line, overall expansion ratio (3.82); red dashed line, standard linear regression;  $m = 0.04 \pm 0.02$ ,  $b = 3.83 \pm 0.01$ ). (g) RMSE of TOTO-1 images before and after MAP, same as in d (n = 4 samples). (h) 100- m-thick mouse brain slice showing blood vessels stained with lectin before and after MAP processing. (i) RMSE of lectin images same as in d. Errors measured after MAP processing (n = 6 samples) are compared with errors measured between two rounds of embedding and imaging of the sample before MAP (n = 5 samples). (j-l) 100- m-thick mouse brain slices stained with anti-GFP antibodies and three other antibodies before and after MAP processing. Magnified 3D-rendered images in middle columns show the same structures imaged with the same microscopic resolution near optical sampling limit. Magnified 2D images showing parvalbumin (PV) ultrastructure in middle panels in k were taken from other samples. NF-H, neurofilament heavy unit. (m) Intensity plot (left) along the lines in zoom-in single sections (right) of the insets in j. The length scale is in terms of distance before MAP. Images were obtained using a 10 $\times$ , 0.3 NA dry objective (e) and a 63 $\times$ , 1.30 NA glycerol-immersion objective (a and b) with single-photon (1p) excitation by 488, 568, and 594 nm, and a 20 $\times$ , 0.95 NA water-immersion objective (h and j-m) at the same resolution below optical sampling limit with either 1p excitation by 488 and 594 nm or 2p excitation by 780 nm. Scale bars, 1 mm (white; e and h), 100 m (gray; inset in e, and j-l), 10 m (yellow; a, and insets in j-l), and 1 m (magenta; b and m).

super-resolution imaging with diffraction-limited microscopes, achieving ~60-nm lateral resolution.

### 3.3.3 Labeling proteome with conventional antibodies

We next asked if MAP preserves the 3D proteome composition and organization, and enables super-resolution imaging of fine subcellular architectures using commercially available antibodies. Antibody targets were selected from a wide range of antigens including membrane proteins, cytoplasmic proteins, nuclear proteins, neurofilament proteins, and synaptic proteins to survey the overall proteomic landscape (Fig. 3.3 and Supplementary Table 1). To validate the antibodies, we tested whether antibody staining of MAP-processed tissues shows comparable structures to PFA-fixed tissues, and antibodies from different host species targeting the same protein, if available, provide concordant images (Supplementary Figs. 2 and 3). We found that 100 of 122 antibodies were compatible with MAP-processed samples, and 43 of 51 target molecules were successfully visualized (Fig. 3.3, Supplementary Table 1, and Supplementary Figs. 2 and 3). This high success rate indicates that loss of epitopes by protein denaturation was minimal. Moreover, some antibodies (e.g., calbindin and calretinin) showed negative staining when different antibodies targeting the same endogenous protein showed strong signal (Fig. 3.3b and Supplementary Table 1). This result demonstrates that MAP preserves the tissue proteome with minimal loss of specific epitopes and that the library can be effectively probed using off-the-shelf antibodies without any modification.

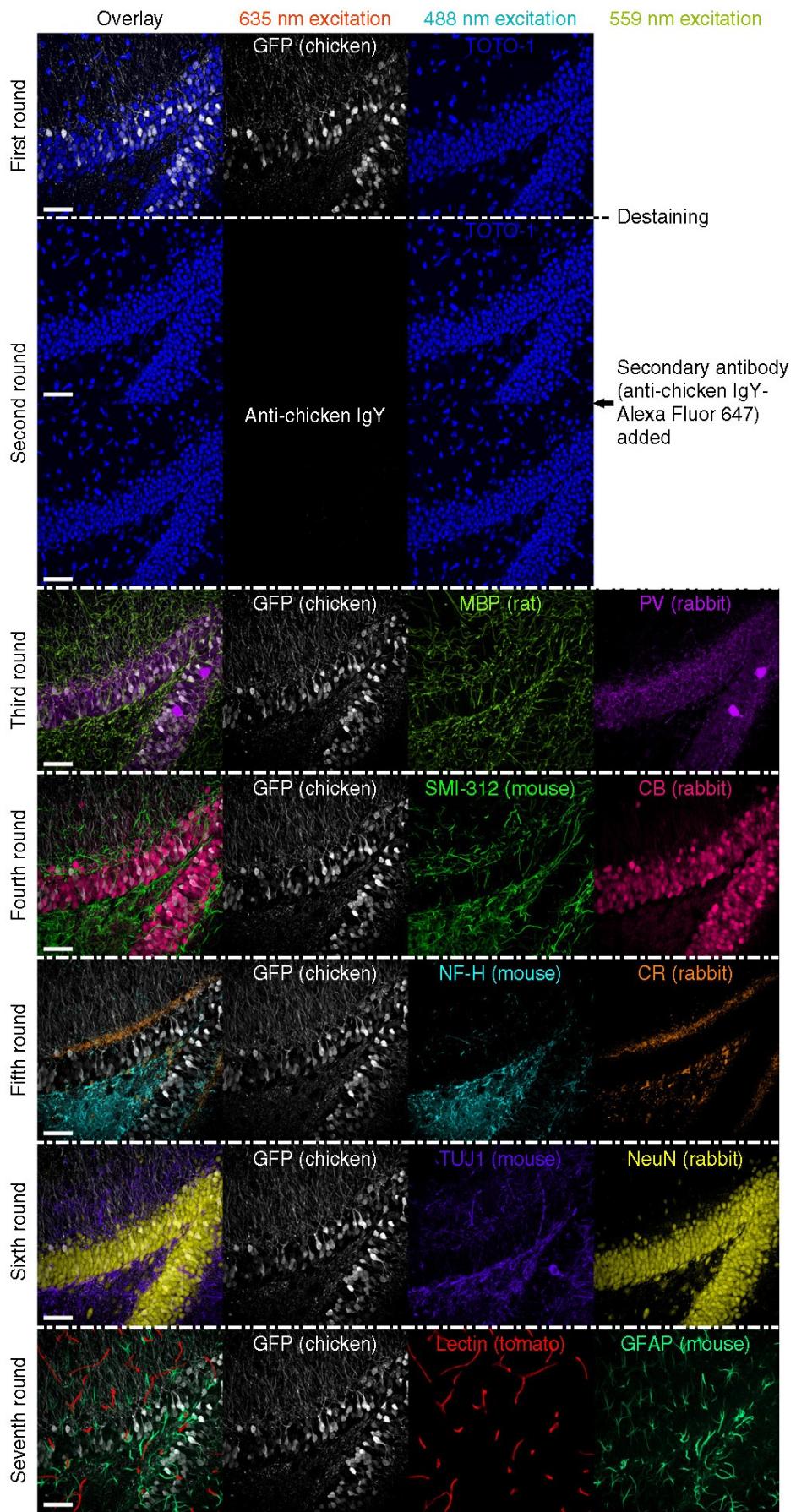


**Figure 3.3:** Super-resolution imaging showing the 3D proteome library and subcellular details of MAP-processed intact tissue.

(a) A total of 122 antibodies were tested on MAP-processed mouse brain sections and PFA-fixed control tissues (Supplementary Table 1). One hundred of 122 antibodies were compatible with MAP, and 43 of 51 target molecules were successfully labeled. (b) Representative images showing morphological details of parvalbumin (PV)-, calbindin (CB)-, and calretinin (CR)-positive cortical neurons. (c,d) 3D rendering of fine cytoskeletal structures visualized by neurofilament protein staining: NF-H in a cortical neuron (c), and SMI-312, a pan-axonal marker, and tyrosine hydroxylase (TH) in a subcortical neuron (d). (e) Structural relationship between astrocytes and capillaries visualized with glial fibrillary acidic protein (GFAP), lectin, and histone H3 in a subcortical region. (f) 3D rendering of a cortical sample stained with neurofilament medium unit (NF-M) and GABAB receptor subunit-1 antibodies. The inset shows synaptic neurofilaments co-localized with the gabaergic post-synaptic proteins. (g) Maximum intensity projection image showing dendritic spines visualized with a cell-type-specific marker, CB. (h) Synaptic structure resolved with a pre-synaptic marker, bassoon, and a post-synaptic marker, homer1, in a cortical region. Yellow box shows a GFP-positive neuron whose dendrite was imaged. The inset (white box) highlights the elliptical structures of pre- and post-synaptic proteins distributed at a synaptic junction. (i) Intensity plot along the axis of the synapse in h (rectangular region, expanded) showing the separation between pre- and post-synaptic distribution profiles. The lines represent the Gaussian curve fitting. (j,l,m) Additional synaptic markers successfully visualizing molecular synapses in MAP-processed tissues. (j) A glutamatergic pre-synaptic marker (VGluT2) co-localized with bassoon in a cortical region (inset). (k) Comparison of the number of synapses between bassoon and VGluT2 in j ( $n = 3$  images). A two-tailed paired t test was performed (\* $P < 0.05$ ). Error bars show mean  $\pm$  s.e.m. (l) The distribution of glutamatergic pre- (VGluT1) and post-synaptic (PSD95) markers in a cortical region. (m) 3D rendering showing the distribution of homer1 clusters in a subcortical region. The images in f, j, and m were obtained from samples stained in a second round. The images were obtained with a 25 $\times$ , 0.95 NA water-immersion objective (b), a 40 $\times$ , 1.25 NA oil-immersion objective (e), and a 63 $\times$ , 1.30 NA glycerol-immersion objective (c, d, f–h, j, l, and m) with 1p excitation (488, 568, 594, and 647 nm). Scale bars, 50  $\mu$ m (white; b–f and low magnification images in h and l), 10  $\mu$ m (gray; inset in e, and h, j, l, and m), and 1  $\mu$ m (yellow; insets in f, h, j, l, and m).

### 3.3.4 Super-resolution imaging of 3D subcellular architectures

Preservation of both nanoscopic structures and the proteome in MAP enables super-resolution imaging of a broad range of fine subcellular architectures. Antibodies targeting filament proteins successfully visualized details of the cytoskeletal networks of various cell types (Fig. 3.3c–e and Supplementary Videos 1–4). No apparent filament fragmentations were observed in the magnified tissues. Labeling of glial fibrillary acidic protein (GFAP) enabled visualization the fine foot-process structures surrounding blood vessels (Fig. 3.3e and Supplementary Video 3) (Kosaka & Hama 1986; Khakh & Sofroniew 2015). Neurofilament medium subunit (NF-M) enabled visualization of both processes and synapses (Yuan et al. 2015), a subset of which co-localized with the postsynaptic marker, GABA B R1 (the C-terminal domain of the GABA B receptor subunit-1) (Fig. 3.3f and Supplementary Video 4). Labeling of calcium-binding proteins (parvalbumin, calbindin, and calretinin) robustly enabled visualization of morphological details of positive cells (Fig. 3.3b). A small subset of calbindin-positive fibers showed expression of calbindin in their dendritic spines (Fig. 3.3g). Many synaptic protein markers were compatible with MAP (Fig. 3.3f,h,j,l,m and Supplementary Table 1). Staining of synaptic proteins clearly visualized distinct and well-separated elliptical disk-shaped clusters of pre- and post-synaptic proteins (Fig. 3.3h,l,m and Supplementary Video 5), enabling their quantitative analysis (Fig. 3.3k).



**Figure 3.4:** Multiplexed staining of MAP-processed tissue.

Repeated staining and destaining of a MAP-processed tissue. A 100- m-thick mouse brain slice including an anterior hippocampal region was used. The tissue in a PBS solution with 0.1 intensity projection of a 64- m thickness. Maximum intensity projection images are displayed. Host species of primary antibodies are shown in the parentheses. Break lines indicate destaining. GFP was stained and imaged at 635-nm excitation channel to show consistent labeling of the same antigen. 488- and 559-nm channels were used to show multiplexed labeling of diverse antigens and complete elution of imaged antibodies used in previous rounds. The same imaging and post-processing conditions were used for 635-nm channel images throughout the first and second rounds. MBP, myelin basic protein; TUJ1, neuron-specific class III beta-tubulin. Scale bars, 200 m.

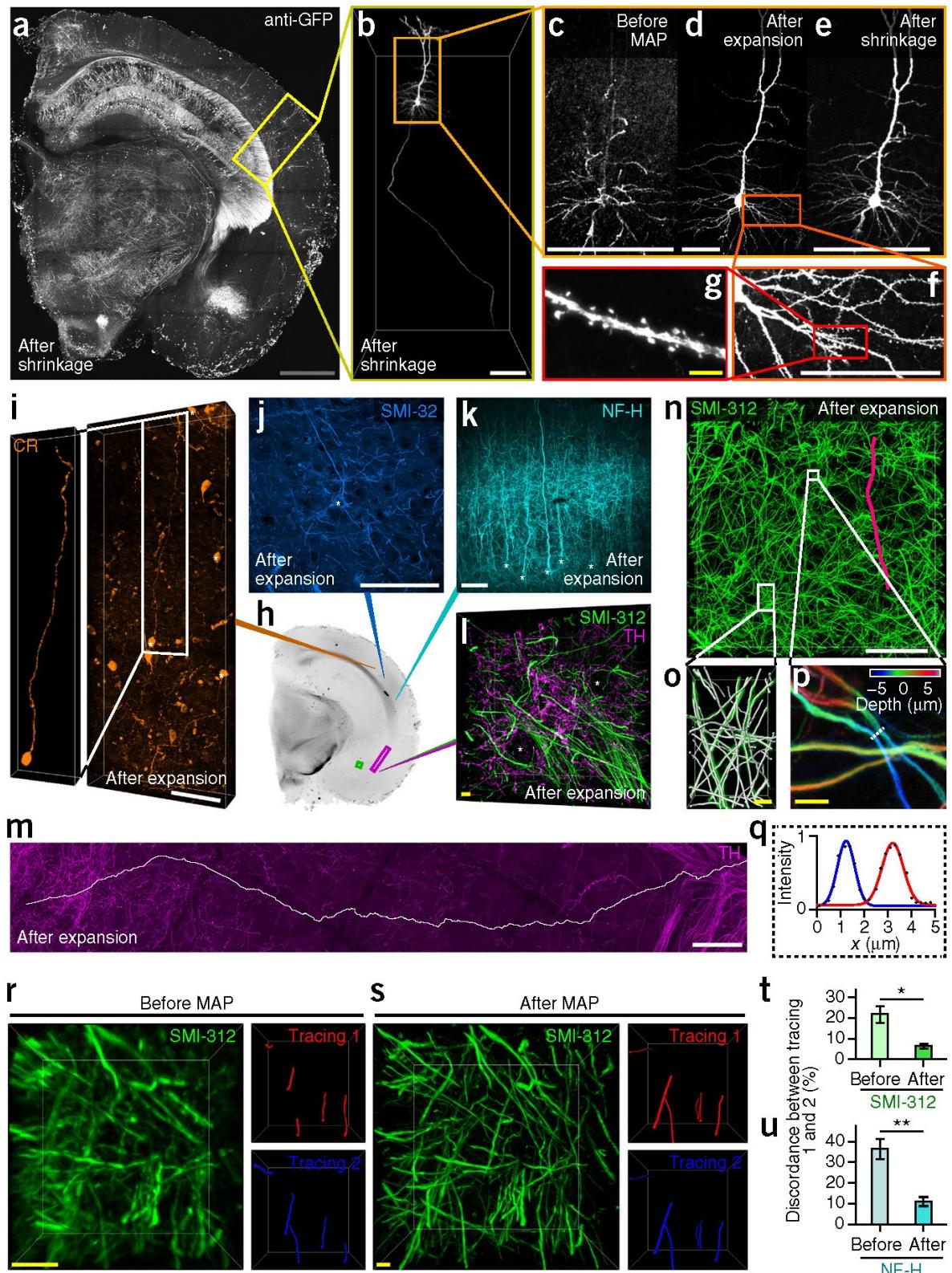
### 3.3.5 Highly multiplexed immunolabeling

Another potential advantage of the MAP technology is that it enables highly multiplexed labeling and imaging of the magnified 3D proteome within a single tissue. We successfully performed seven rounds of immunostaining of a MAP-processed 100- m-thick mouse brain tissue with no obvious signs of tissue damage (Fig. 3.4). Outstanding mechanical stability of the MAP-processed sample enabled repeated manual handling of the tissue with tools that are widely used in biology laboratories. The images after destaining of the imaged anti-GFP antibody (first round) and after solely adding secondary antibodies targeting the eluted anti-GFP antibody (second round) showed little signal, suggesting that the destaining process effectively eliminates antibodies. For the following five rounds of labeling, we continued to use anti-GFP as a landmark in addition to two other antibodies for each round. The consistent GFP signals show that the same antigen can be repeatedly labeled without loss of antigenicity. Successful visualization of ten other targets demonstrated that MAP facilitated exploration of diverse proteins, structures, and cell types within a single tissue.

### 3.3.6 MAPing neural projections

We next explored the potential utility of MAP for mapping inter-areal neural connectivity. If MAP could preserve continuity of neural fibers within magnified, intact brain tissue along with its 3D proteome, highly specific antibodies would enable visualization of a subset of neural projections connecting different brain regions. The sparse labeling and enhanced spatial resolution offered by MAP may allow more accurate reconstruction of the labeled projections. In addition, highly multiplexed proteomic imaging may enable reconstruction of many different cell types with integrated molecular and fine morphological details in a single tissue.

To explore this possibility, we first asked if the continuity of neural processes is preserved within a magnified sample. We expanded a 0.5-mm-thick Thy1-eGFP mouse brain coronal block and then imaged GFP-labeled neurons and their projections (Feng et al. 2000). MAP preserved the continuity of the GFP-expressing projections spanning a large tissue

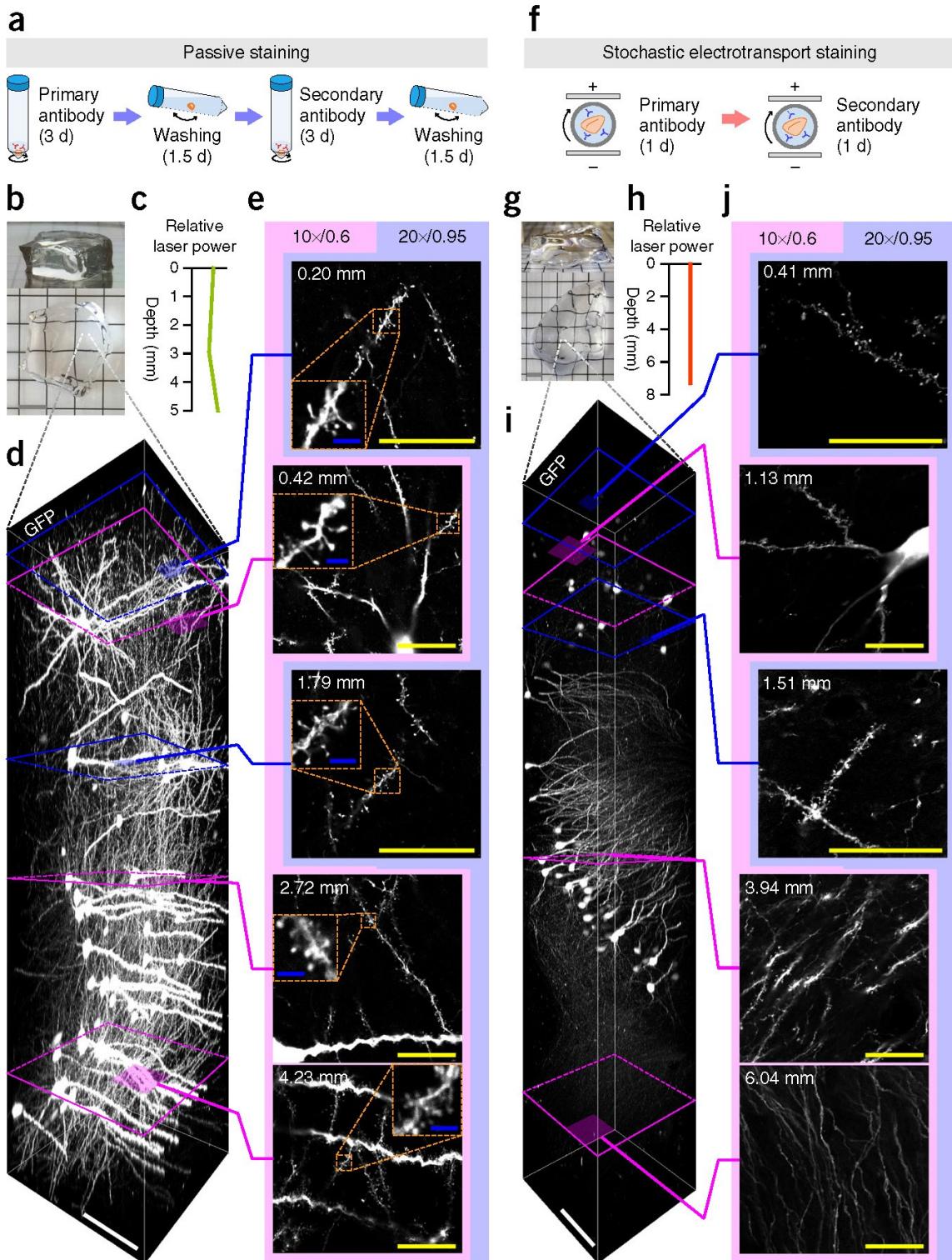


**Figure 3.5:** Intercellular connectivity and its reconstruction at single-fiber resolution in MAP-processed tissue.

(a–g) Connectivity and morphology of cells are preserved in the MAP-processed tissue. (a) A 500- m-thick slice of a mouse brain hemisphere stained with anti-GFP antibody and imaged after expansion and shrinkage. (b) The entire neuronal territory and axon fiber of a cortical pyramidal neuron included within the yellow-boxed region in the left panel. The volume was 3D reconstructed with background removal. (c–e) Images of the same neuron in b obtained with a 10 $\times$ , 0.3 NA objective (c,d) and a 10 $\times$ , 0.6 NA objective (e), showing its morphological details before MAP, after MAP expansion, and after MAP shrinkage. (f,g) Fine subcellular structures preserved and resolved in the expansion states: dendritic network (f) and dendritic spines (g). (h–l) Neuronal fibers and cell body (asterisks) morphology visualized using various markers (CR, SMI-32, NF-H, SMI-312, and TH) in different cortical (i–k) and subcortical (l) regions of a 1-mm-thick hemisphere slice (h). The inset in i shows the isolated neuronal cell body and its axon with background removal. (m) Long-range tracing (white line, longer than 3 mm) of an inter-regional TH neurofilament within the magenta box in h. (n) Subvolume within a square region in h (green box) containing dense SMI-312-positive fibers. The red line shows a tracing result across the volume. (o) One of the densest subregions in the left inset in n visualizing all individual fibers resolved and traced. (p) A dense region in the right inset in n showing closely crossing fibers. Depth-color coding of the fibers clearly shows good separation along the z-axis. (q) The intensities of the two fibers closely located in a similar plane in p are plotted and fitted to Gaussian curves to show their clear separation. (r,s) Comparison of manual tracing results of a 100- m-thick hippocampal sample stained with SMI-312 before (r) and after (s) MAP processing. Apparently detectable fibers were traced by two individuals. Discordant traces are selectively displayed for before MAP image (r), and the same traces, which resulted in concordant tracing after expansion, are shown for after MAP image (s). (t,u) Comparison of discordance ratios between manual tracing results before and after MAP processing. Two markers (SMI-312 and NF-H) were used with four regions of interest for each marker. Two-tailed paired t tests were performed (\*\*P < 0.01, \*P < 0.05). Error bars show mean  $\pm$  s.e.m. Images were obtained using a 10 $\times$ , 0.6 NA CLARITY-optimized objective (a–c, e), a 10 $\times$ , 0.3 NA water-immersion objective (d), a 20 $\times$ , 0.95 NA water-immersion objective (m), a 63 $\times$ , 1.30 NA glycerol-immersion objective (l and n–p), and a 25 $\times$ , 0.95 NA water-immersion objective (f, g, i–k, r, and s) with 1p excitation by 488, 568, and 594 nm. Scale bars, 1 mm (gray; a), 200 m (white; b–f, i–k, m, n), and 10 m (yellow; g, l, o, and p, r, and s).

volume (Fig. 3.5a–f) as well as their fine morphological details (e.g., dendritic spines) (Fig. 3.5g). Immunolabeling of various cytoplasmic proteins including neurofilament proteins (e.g., NF-H and SMI-32), calcium-binding proteins, and metabolic enzymes (e.g., tyrosine hydroxylase) confirmed that MAP preserved the continuity of neural fibers of various cell types (Fig. 3.5h–n and Supplementary Videos 6–8). The preserved continuity enabled detection of single fibers selected in the immunostained tissue volumes (Fig. 3.5m–q and Supplementary Videos 7 and 8).

To test if MAP enables more accurate tracing of densely packed fibers, we imaged neurofilament-stained samples before and after MAP processing using a high numerical aperture (NA; 0.95) water-immersion objective. Two individuals not involved in image acquisition traced fibers within the sample volume (Fig. 3.5r,s). When we compared the concordance between the two tracing results, the discordant rate was significantly lower after MAP (Fig. 3.5t,u; SMI-312, n = 4, P = 0.020; NF-H, n = 4, P = 0.010), even though MAP enabled the tracers to detect more fibers (total numbers of traced fibers were 160 before MAP and 214 after MAP). This result demonstrates that MAP indeed enabled more accurate reconstruction of immunolabeled neural fibers.



**Figure 3.6:** Immunolabeling and imaging of thick MAP-processed tissues.

(a–e) Passive staining of a 1-mm-thick (5 mm after expansion) dorsolateral mouse brain slice with anti-GFP antibodies. (a) The tissue was incubated for 3 d with primary antibody and 3 d with secondary antibody, and expanded in DI water. (f–j) Stochastic electrotransport<sup>28</sup> staining of a 2-mm-thick (8 mm after expansion) mouse hemisphere slice with anti-GFP antibodies. (f) The tissue was stained for 1 d with both primary and secondary antibodies. (b,g) Lateral (top panels) and top (bottom panels) views of the sample. White boxes indicate the regions of interest for z-stack images. The cortical (b) and hippocampal (g) regions were chosen at least 2.5- and 4-mm far from the lateral tissue boundary, respectively, to exclude the chance of dominant staining by lateral diffusion at a middle section. (c,h) Laser power used along depth to obtain a z-stack image. (d,i) Longitudinal tile-scan images along z-axis with 1p excitation and a 10 $\times$ , 0.6 NA, 8-mm WD objective. (e,j) Sectional images at different depths either with a 10 $\times$ , 0.6 NA, 8-mm WD (magenta) or a 20 $\times$ , 0.95 NA, 2-mm WD (blue) objectives. The same image acquisition parameters were used for sectional images with each objective. Insets in e are zoom-in images showing dendritic spines. Scale bars, 5 mm (gray; b and g), 500  $\mu$ m (white; d and i), 100  $\mu$ m (yellow; e and j), and 10  $\mu$ m (blue; insets in e).

### 3.3.7 Immunolabeling and imaging of mm-thick tissues

Reconstruction of individual neurons requires labeling and imaging of thick brain tissues because nerve fibers can extend across a large volume. To test whether MAP is applicable to large-scale brain tissues, we expanded a 1-mm-thick mouse brain block (5-mm thick after expansion, Fig. 3.6b) and passively stained it with anti-GFP antibody (Fig. 3.6a). We then imaged the sample using both a high NA (0.95), short working distance (WD) (2 mm) water-immersion objective and a low NA (0.6), long WD (8 mm) CLARITY objective. Both objectives showed fine dendritic spines of GFP-expressing neurons throughout the entire volume (Fig. 3.6e). Imaging of the sample did not require depth-dependent modulation of the laser power, indicating that staining was uniform and that signal attenuation by light scattering was negligible (Fig. 3.6c,d). MAP is also compatible with stochastic electrotransport (Kim et al. 2015), a method that enables rapid tissue labeling. Using stochastic electrotransport, we were able to label an 8-mm-thick expanded tissue uniformly within only 2 d (Fig. 3.6f–j). The sample was highly transparent. We were able to image the 8-mm-thick sample successfully with the same laser power up to the working distance of the objective (Fig. 3.6h).

## 3.4 Discussion

As a step toward organ-scale reconstruction of diverse cell types and their surrounding environment and to advance the study of complex system-level interactions, we developed a simple and scalable method that enables preservation, reversible expansion, and imaging of the 3D proteome organization within an intact tissue. We discovered that the protein content of a whole organ can be preserved and magnified by preventing intra- and interprotein crosslinking, then denaturing and dissociating the protein complexes to allow natural expansion of hydrogel-tissue hybrids. The expanded hybrids secure both fine

subcellular architectures and organ-scale cellular connectivity. These multiscale properties can be directly imaged by using off-the-shelf antibodies to label the structures' constituent proteins. The high success rate that we achieved with commercial antibodies might be because synthetic peptides or denatured protein fragments are commonly used for conventional antibody production (Uhlen & Ponten 2005). An existing large antibody library, once validated, can be used without any modification. This technique is easy to implement, and it does not require any special equipment or chemicals.

MAP enables repeated interrogation of the same expanded specimen. This is possible because the preserved epitopes, which have already reached complete denaturation by harsh treatment (95 °C, 200 mM sodium dodecyl sulfate (SDS)) for expansion, do not undergo any substantial modification during the milder antibody elution step (70 °C, 200 mM SDS). We performed seven rounds of staining of the same tissue without any signs of tissue damage. However, the practical limit of this approach must be carefully examined.

MAP has the potential to enable scalable super-resolution imaging of large-scale samples. Unlike other methods, thick tissues can be labeled on a practical time scale because antibody-labeling occurs after complete lipid removal and tissue permeabilization. The great transparency of the MAP-processed samples allows high-resolution imaging of the physically expanded tissue with minimal loss of resolution. Currently, 1.0 NA, 2.5-mm WD water-immersion objectives are best suited for MAP. Although the 1.0 NA, 8-mm WD CLARITY objective has the longest WD, it does not provide high-quality images because it is optimized for immersion media with a high refractive index (RI). Further development of high-NA, long-WD water objectives or the development of high-RI immersion media that are compatible with MAP would extend the utility of MAP. Another challenge in the MAP approach is the dilution of fluorescent signals that accompanies physical volumetric expansion. Fourfold linear expansion decreases signal density by 64-fold. Therefore, much higher laser power is required, which in turn causes photobleaching. Future studies will need to explore the compatibility of signal amplification techniques with MAP.

The reversible modulation of physical sample size that is attained with MAP enables multiscale proteomic imaging of a single tissue to capture both system-scale properties and fine local details in a practical way. For example, following sample shrinkage, global projection patterns of labeled neurons may be imaged using high-speed microscopy techniques and long WD objectives (e.g., 25 $\times$ , 1.0 NA, 8-mm WD objective, currently available) (Tomer et al. 2014). After imaging, the same tissue can be expanded for super-resolution imaging of regions of interest. The expanded tissue may need to be sliced before imaging to meet the objective's WD limit. A microscope with a built-in vibrating-blade microtome could also be used for whole-mount imaging to avoid loss of connectivity information (Ragan et al. 2012; Economo et al. 2016; Albanese & Chung 2016). This

approach may allow us to drastically reduce the imaging time and costs associated with data storage and handling of expanded samples.

The advantages of MAP (Supplementary Table 2) may allow combined extraction of rich molecular details, minute subcellular architectures, and cellular connectivity from diverse cell types within a single tissue. Together with its simplicity and broad applicability, MAP may complement existing methods and enable new approaches in the study of complex biological systems.

### **3.5 Acknowledgements**

The authors thank the entire Chung laboratory for support and helpful discussions. K.C. was supported by Burroughs Wellcome Fund Career Awards at the Scientific Interface, the Searle Scholars Program, Packard award in Science and Engineering, JPB Foundation (PIIF and PNDRF) and NIH (1-U01-NS090473-01). Resources that may help enable general users to establish the methodology are freely available online (<http://www.chunglabresources.org>). K.C. is a co-founder of LifeCanvas Technologies, a startup that aims to help the research community adopt technologies developed by the Chung Laboratory.

### **3.6 Author Contributions**

T.K., J.S., J.-Y.P., and K.C. designed the experiments and wrote the paper with input from other authors. T.K. stained and imaged mouse samples. J.S. performed the gel and cell experiments. T.K. and J.S. analyzed the data. J.-Y.P. prepared mouse tissues. J.-Y.P. and V.M. processed mouse MAP samples. A.A. performed the cell and organoid experiments. E.M., Y.-G.P., and T.K. performed the antibody validation test. J.H.C. performed stochastic electrotransport staining. Y.-G.P. and T.K. obtained synaptic images. J.-Y.P., V.M., T.K., and J.S. performed tracing. J.C. performed the gel experiment. K.C. supervised all aspects of the work.

### **3.7 Competing Financial Interests**

The authors declare competing financial interests: details are available in the online version of the paper.

## 3.8 Methods

### 3.8.1 BSA hydrogel denaturation and expansion

Stock solutions of 40% BSA, 40% acrylamide (AA), 32% paraformaldehyde (PFA), and 1% VA-044 were made and kept on ice throughout the experiment. A 10-mL solution of 4% BSA, 4% AA, 4% PFA, 0.1% bis-acrylamide (BA), and 0.1% VA-044 was made in phosphate-buffered saline (PBS). The solution was polymerized under vacuum at 37 °C for 2 h, and the resulting albumin-containing tissue phantom was sectioned. Individual tissue phantom sections were washed in excess PBS with shaking for 12 h. After washing, four phantoms were placed in 10 mL of each of the following PBS solutions for 4 h at 37 °C: 2% glutaraldehyde (GA), 4% PFA, 4% AA with 4% PFA, and 20% AA with 4% PFA. Tissue phantoms were washed similarly, massed, photographed, and then incubated in a 200 mM SDS solution with 50 mM sodium sulfite for 1 h at 95 °C. Phantoms were washed again and incubated in deionized (DI) water for 12 h. After expansion, phantoms were massed and photographed.

### 3.8.2 General MAP protocol

- i. Perfusion and hydrogel embedding. Thy1-eGFP-M mice (6–8 weeks old, male and female) were housed in a reverse 12-h light/dark cycle with unrestricted access to food and water. All experimental protocols were approved by the MIT Institutional Animal Care and Use Committee and the Division of Comparative Medicine and were in accordance with guidelines from the US National Institutes of Health. After anesthesia, the mice were first washed transcardially with a mixture of 2–5% AA, 0–0.05% BA, 0–0.8% sodium acrylate (SA), and PBS, followed by perfusion with a mixture of 4% PFA, 30% AA, 0.05–0.1% BA, 10% SA, 0.1% VA-044 or V-50, and PBS. The perfusion solution could be slightly turbid. Upper transparent solution was used after centrifugation during 3 min with 1,000g. All solution was protected from light and kept on ice before perfusion. Control samples were perfused first with PBS and then with 4% PFA and PBS. The brain and other organs (heart, lung, liver, intestine, kidney, and spinal cord) were harvested and incubated in 20–40 mL of the same fixative solution at 4 °C for 2–3 d and then for 1–3 d at room temperature (RT) with gentle shaking to ensure uniform chemical diffusion and reaction throughout the sample. Following the diffusion and fixation steps, hydrogel-tissue hybridization was performed *in situ* by incubating the tissues using Easy-Gel (LifeCanvas Technologies) with nitrogen gas at 45 °C for 2 h.

- ii. Tissue denaturation. Hydrogel-embedded tissues were incubated overnight in a solution of 200 mM SDS, 200 mM NaCl, and 50 mM Tris in DI water (pH titrated to 9.0) at 37 °C with gentle shaking. The samples were then incubated at 70 °C for 0–50 h and 95 °C for 1–24 h depending on their size using EasyClear (LifeCanvas Technologies). Whole organs were incubated for 24–48 h at 70 °C followed by 12–24 h at 95 °C. 1-mm-thick brain slices were incubated for 5 h at 70 °C and then at 95 °C for 1 h.
- iii. Expansion. Denatured tissues were incubated in 40–100 mL DI water at RT for 12–48 h with gentle shaking. During DI water incubation, the solution was changed every 3–5 h.

### 3.8.3 Expansion according to various AA and SA concentrations

After anesthesia, mice were washed transcardially with 4% AA, 0.05% BA, 0.8% SA, 0.1% VA-044, and PBS and then with 4% PFA, 0.05% BA, 0.1% VA-044, one of four AA and SA combinations (5% AA + 0.8% SA, 10% AA + 1.7% SA, 20% AA + 3.3% SA, and 30% AA + 5% SA) in PBS. Tissues were incubated in 20 mL of the same fixative solution at 4 °C for 2 d and 5 h at RT with gentle shaking. After hydrogel-tissue hybridization the samples were incubated in denaturation solution at 70 °C for 5 h and 95 °C for 1 h with gentle shaking. Denatured tissues were incubated in 40 mL DI water at RT for 12 h with gentle shaking. During DI water incubation, the solution was changed every 3–5 h. Fiji (National Institutes of Health) was used to measure the size of the expanded samples (Schindelin et al. 2012).

### 3.8.4 Shrinkage and RI matching

A customized RI matching solution was made by dissolving 50 g diatrizoic acid, 40 g N-methyl-d-glucamine, and 55 g iodixanol per 100 mL PBS (Kim, S.-Y. et al. 2015). The RI was targeted to 1.47. This solution was used for both shrinkage and imaging. Depending on sample size, samples were incubated in 1–10 mL solution at RT for 2–24 h with gentle shaking. The solution was changed every 1–12 h. For imaging tissues before MAP processing, samples were incubated in 1–10 mL of this solution without PBS at RT with gentle shaking for 2–5 h before imaging.

### 3.8.5 MAP processing of cerebral organoid

Cerebral organoids were made from stem cells following a previously described protocol (Lancaster & Knoblich 2014). Organoids were initially fixed in 4% PFA for 15 min, incubated in a mixture of 4% PFA, 30% AA, 0.1% BA, 10% SA, 0.1% V-50, and PBS for 24 h at 4 °C, followed by 24 h, at RT. Hydrogel embedding, tissue denaturation, and expansion were processed similarly to “General MAP protocol.”

### 3.8.6 Cultured cell experiment

For tubulin imaging in HeLa cells, 8-mm round glass coverslips were coated in 0.1% gelatin in ultrapure water (Millipore). Coverslips were placed in a 48-well plate and seeded with 50,000 HeLa cells overnight. To obtain comparable images before and after MAP processing, cells were washed, fixed with 3% PFA + 0.1% GA in PBS for 10 min, and switched to a solution of 4% PFA, 30% AA in PBS for 8 h at 37 °C. Cells were then placed in 0.1% sodium borohydride for 7 min at RT then incubated in 100 mM glycine for 10 min at RT. Cells were washed and stained with anti-tubulin (Abcam, ab6160), Alexa Fluor 594-conjugated secondary (Abcam, ab150152) antibodies and TOTO-1 (Thermo Fisher Scientific). Cells were mounted in 2,2 -thiodiethanol (Sigma) and imaged with a 63×, 1.3 NA glycerol-immersion objective with the Leica microscope system. Cells were washed extensively and embedded into a MAP hybrid polymer by addition of 20 L of Cell-MAP solution (20% AA, 7% SA, 0.1% BA, 0.5% TEMED, 0.5% ammonium persulfate in PBS). Ammonium persulfate was added last from a freshly prepared 5% stock solution. Cell-MAP solution was quickly added to the coverslip and left to polymerize for 4–5 min. Gels were peeled off the coverslip using forceps, washed extensively and denatured for 30 min in denaturation buffer at 95 °C. Cell-MAP gels were washed extensively, restained with anti-tubulin antibody and TOTO-1 and reimaged.

### 3.8.7 Immunostaining of brain tissue

For typical staining, MAP-processed 100- to 500- m-thick mouse brain coronal slices were incubated with primary antibodies (typical dilution, 1:100) in PBS with 1% (wt/vol) Triton X-100 (PBST) at 37 °C for 8–16 h, followed by washing at 37 °C for 1–2 h in PBST three times. The tissue was then incubated with secondary antibodies (typical dilution, 1:100) in PBST at 37 °C for 6–16 h, followed by washing at 37 °C for 1–2 h in PBST three times. For antibody validation of a given antibody, 100- m-thick PFA-fixed control and MAP-processed samples were stained with the same titer of primary and, if necessary, secondary antibodies overnight in PBST. See Supplementary Table 1 for the

list of antibodies used. To destain for multiplexed labeling, samples were incubated in a denaturation solution 6–16 h at 70 °C, and washed with PBST at 37 °C for 1–2 h three times.

### 3.8.8 Mounting and imaging

Samples were mounted on a slide glass. Blu-Tack adhesive was applied on the Petri dish or the slide glass, and samples were covered with a glass-bottom Willco dish. The space between the bottom material and the Willco dish around the sample was filled with either shrinkage solution or DI water according to the sample immersion medium. Large expanded samples were additionally placed on a 120-mm-diameter Petri dish, and the dish was filled with DI water. Expanded or shrunk samples were stabilized for at least 1 h before imaging. Samples were imaged with either the Olympus FV1200MPE microscope system or the Leica TCS SP8 microscope system. A 10×, 0.6 NA CLARITY-optimized objective (XLPLN10XSVMP; 8.0-mm WD) was used with the Olympus system to obtain wide-field images of shrunk samples and z-stack images of large samples. The images of MAP-processed samples were obtained with a 10×, 0.3 NA water-immersion objective, a 20×, 0.95 NA water-immersion objective, and a 40×, 1.25 NA oil-immersion objective with the Olympus system, or a 10×, 0.3 NA water-immersion objective, a 25×, 0.95 NA water-immersion objective, and 63×, 1.30 NA glycerol-immersion objective with the Leica system. Single-photon confocal laser scanning imaging was performed with 405-, 488-, 559-, and 635-nm lasers (Olympus) or a white-light laser (Leica). Mai Tai DeepSee (Spectra-Physics) was used for multi-photon excitation with 780-nm wavelength. The images were visualized and analyzed with Fiji or Imaris (Bitplane).

### 3.8.9 Large tissue staining

1- and 2-mm-thick mouse brain coronal slices were prepared by “General MAP protocol,” and expanded. A 1-mm-thick slice was chopped to about 3 mm × 3 mm × 1 mm (dimensions before MAP). The sample was stained passively with Alexa Fluor 594-conjugated rabbit anti-GFP antibody (Life Technologies, A21312) for 3 d and Alexa Fluor 594-conjugated donkey anti-rabbit IgG antibody (Abcam, ab150072) for 3 d. 20 L of antibody was used in 500 L PBST, and was washed with 40 mL PBST for 1.5 d (three times in total) for each antibody. We used stochastic electrotransport 28 to stain a 2-mm-thick slice with Alexa Fluor 647-conjugated rabbit anti-GFP antibody (Life Technologies, A31852) and Alexa Fluor 647-conjugated donkey anti-rabbit IgG antibody (Abcam, ab181347). For each antibody, we first electrotransported stochastically a solution containing 20 L of antibody in 4 mL of 0.6 M N-cyclohexyl-3-aminopropanesulfonic

acid (CAPS), 0.2 M Tris, 100 mM NaCl, 20 mM SDS, 1% BSA with an electrophoresis buffer containing 0.3 M CAPS, 0.2 M Tris, 20 mM SDS, 30% sorbitol for 21 h. It was then electrotransported stochastically in a solution containing 4 mL of 0.3 M CAPS, 0.2 M Tris, 20 mM SDS, 30% sorbitol, 1% BSA, and 1% Triton X-100 with an electrophoresis buffer containing 0.04 M Tris, 0.01 M phosphate, 30% sorbitol for 8 h. These two steps complete the stochastic electrotransport labeling. The stained sample was expanded in a solution containing 0.01% (wt/vol) heparin sodium porcine mucosa (Sigma, SRE0027) and 1% Triton X-100 in DI water, and then imaged.

### 3.8.10 Cell distortion analysis

Before and after MAP processing tubulin images were first registered using a scaled-rotation transformation in Fiji. Non-rigid invertible B-spline registration was performed with an  $8 \times 8$  control point grid in bUnwarpJ. Vectors of different length were subjected to the resulting nonlinear transformation, and the input-output difference norm was sorted based on the input vector length and then averaged by root-mean-square. For morphological image processing of nuclei, TOTO-1 images before and after MAP-processing were first registered using a scaled-rotation transformation in Fiji. The registered images were segmented by thresholding and converted to circular particles with equivalent average radii calculated from the “Analyze Particles” function in Fiji. Matched pairs of cells from the before and after images were randomly chosen, and the expansion ratio was calculated from the ratio of the connecting line segment lengths. The fraction in nuclei was obtained by summing the intensity profile along the connecting line segment and averaging the sums from the before and after binary masks.

### 3.8.11 Tissue distortion analysis

After anesthesia, mice were first washed transcardially with 2% AA in PBS followed by perfusion with 4% PFA and 30% AA in PBS. Thy1-eGFP-M mouse brains were harvested and incubated in 20 mL of the same fixative solution at 4 °C overnight and at 37 °C for 3 h. Brains were sectioned to 100- m-thick coronal slices with a vibrating microtome. Slices were stained and imaged to obtain “before MAP” images, and then incubated in a solution containing 4% PFA, 30% AA, 0.1% BA, 10% SA, and 0.1% V-50 in PBS at RT for 8 h. Hydrogel-tissue hybridization was performed *in situ* by incubating the tissues with nitrogen gas at 45 °C for 2 h. Hydrogel-embedded tissues were incubated in denaturation solution at 37 °C for 1 h and 95 °C for 0.5–1 h. Samples were then stained with the same markers and imaged to obtain “after MAP” images.

To quantify distortion errors, regions of  $\sim 3$  mm  $\times$  2.5 mm in size that included cortex and hippocampus were stained with DyLight 594-conjugated lectin (Vector Laboratories, DL-1177) before (8 L in 200 L PBST for up to 8 samples) and after (2 L in 200 L PBST for each sample) MAP processing. Samples were incubated in RI matching solution after staining and mounted and imaged before MAP processing. Five samples were repeated for incubation, mounting and imaging to measure mounting errors. After MAP processing, six samples were stained, expanded in DI water, and imaged. Keypoints of the vasculature in volumetric images were detected and matched between two image sets with a MATLAB code implementing the 3D Harris Corner Detector and 3D SIFT algorithm as described previously (Murray et al. 2015). Using custom-built graphical user interface software developed with Delphi XE4 (Embarcadero Technologies), redundant keypoints closely located to each other and keypoints at tissue margins were removed. Tissue sizes were estimated by the area defined by a convex hull encompassing all keypoints, and the expansion ratio was calculated as the ratio between two squared roots of the areas. The correspondence information was used to generate a regularly spaced deformation mesh using a 3D thin plate spline code written by Yang (<http://www.mathworks.com/matlabcentral/fileexchange/37576-3d-thin-plate-spline-warping-function>). Lengths between each pair of grid points were calculated in both pre-MAP and post-MAP images, considering the expansion ratio. The difference between the two lengths was measured as a distortion error. After averaging the squared errors for each measurement length, the square root of the averages was collected from the samples to obtain statistical values of error.

### 3.8.12 Neurofilament tracing

Tracing of individual neurons was performed using either Fiji or Imaris. For manual tracing of a long tyrosine hydroxylase fiber, a representative fiber was chosen during confocal imaging acquisition and traced by moving the motorized stage and adjusting the z-level. During the tracing, ambiguous crossovers were resolved by obtaining high-magnification subvolume images. After tile-scanning, the target fiber was re-identified from the image volume using Fiji and marked to be displayed in a two-dimensional plane. For the semi-automatic tracing of SMI-312 fibers in a dense region, the entire 800 m  $\times$  800 m  $\times$  150 m (expanded) data set was loaded along with a filament tool into a ‘Surpass’ instance using Imaris. An autopath calculation was performed using a single starting point as indicated in Supplementary Video 8. The fiber endpoint was designated by selecting the portion of the fiber exiting the imaged volume. The fiber representation was changed to a cone representation to visualize the filament diameter as well as the tracing path. The tracing fidelity was confirmed by inspection. To trace multiple fibers

in a dense region, the full data set was cropped to the region indicated in Figure 3.50 using the ‘3D crop’ tool, and similar autopath calculations were performed for each fiber. For traceability comparison between before and after MAP, we imaged the same tissues before and after MAP processing at the same resolution near the optical sampling limit with the Leica system and a  $25\times$ , 0.95 NA water-immersion objective. Two individuals not involved in imaging acquisition performed manual tracing using autodepth assistance in Imaris, and discordance ratios between manual tracing results before and after MAP processing were calculated.

### 3.8.13 Synaptic and fiber intensity profiles

The regions of interest were imported into Fiji. Lines were drawn perpendicular to the synaptic junction or near the fibers of interest and intensity profiles were obtained with adjusting the line width. For synaptic intensity profiles, two Gaussian distributions were fit to the distinct peaks by simultaneous minimization of the sum of squared residuals. Synaptic densities for bassoon and VGluT2 were calculated from three non-overlapping xy-images ( $235\text{ }\mu\text{m} \times 235\text{ }\mu\text{m}$ ) of expanded samples. The images were segmented by thresholding, and individual synaptic structures for each channel were counted using the “Analyze Particles” function in Fiji. The synaptic densities were calculated based on the frame area, and the s.d. was calculated from the three replicates.

### 3.9 References

- Agato, C., Uhlén, M. & Hober, S., 2004. Genome-based proteomics. *Electrophoresis*, 25(9), pp.1280–1288.
- Albanese, A. & Chung, K., 2016. Whole-brain imaging reaches new heights (and lengths). *eLife*, 5, pp.1–3.
- Berglund, L. et al., 2008. A gene-centric human protein atlas for expression profiles based on antibodies. *Molecular and Cellular Proteomics*, 7(10), pp.2019–2027.
- Betzig, E. et al., 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793), pp.1642–1645.
- Chen, F., Tillberg, P.W. & Boyden, E.S., 2015. Expansion microscopy. *Science*, 347(6221), pp.543–548.
- Chung, K. et al., 2013. Structural and molecular interrogation of intact biological systems. *Nature*, 497(7449), pp.332–7.
- Craddock, R.C. et al., 2013. Imaging human connectomes at the macroscale. *Nature Methods*, 10(6), pp.524–539.
- DeFelipe, J., 2010. From the connectome to the synaptome: An epic love story. *Science*, 330(6008), pp.1198–1201.
- Dodt, H.U. et al., 2007. Ultramicroscopy: Three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature Methods*, 4(4), pp.331–336.
- Economou, M.N. et al., 2016. A platform for brain-wide imaging and reconstruction of individual neurons. *eLife*, 5(JANUARY2016), pp.1–22.
- Feng, G. et al., 2000. Imaging Neuronal Subsets Neurotechnique. *Neuron*, 28, pp.41–51.
- Hanash, S., 2004. HUPO initiatives relevant to clinical proteomics. *Molecular and Cellular Proteomics*, 3(4), pp.298–301.
- Helmstaedter, M., 2013. Cellular-resolution connectomics: Challenges of dense neural circuit reconstruction. *Nature Methods*, 10(6), pp.501–507.
- Hess, S.T., Girirajan, T.P.K. & Mason, M.D., 2006. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11), pp.4258–4272. Available at: <http://dx.doi.org/10.1529/biophysj.106.091116>.
- Kay, K.R. et al., 2013. Studying synapses in human brain with array tomography and electron microscopy. *Nature Protocols*, 8(7), pp.1366–1380. Available at: <http://dx.doi.org/10.1038/nprot.2013.071>.

org/10.1038/nprot.2013.078.

Khakh, B.S. & Sofroniew, M.V., 2015. Diversity of astrocyte functions and phenotypes in neural circuits. *Nature Neuroscience*, 18(7), pp.942–952.

Kim, S.-Y. et al., 2015. Stochastic electrotransport selectively enhances the transport of highly electromobile molecules. *Proceedings of the National Academy of Sciences*, 112(46), pp.E6274–E6283. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1510133112>.

Kosaka, T. & Hama, K., 1986. Three-dimensional structure of astrocytes in the rat dentate gyrus. *Journal of Comparative Neurology*, 249(2), pp.242–260.

Lancaster, M.A. & Knoblich, J.A., 2014. Generation of cerebral organoids from human pluripotent stem cells. *Nature Protocols*, 9(10), pp.2329–2340.

Maglione, M. & Sigrist, S.J., 2013. Seeing the forest tree by tree: Super-resolution light microscopy meets the neurosciences. *Nature Neuroscience*, 16(7), pp.790–797. Available at: <http://dx.doi.org/10.1038/nn.3403>.

Micheva, K.D. & Smith, S.J., 2007. Array Tomography: A New Tool for Imaging the Molecular Architecture and Ultrastructure of Neural Circuits. *Neuron*, 55(1), pp.25–36.

Murray, E. et al., 2015. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6), pp.1500–14.

Oh, S.W. et al., 2014. A mesoscale connectome of the mouse brain. *Nature*, 508(7495), pp.207–14.

Phizicky, E. et al., 2003. Protein analysis on a proteomic scale. *Nature*, 422(6928), pp.208–215.

Puchtler, H. & Meloan, S.N., 1985. On the chemistry of formaldehyde fixation and its effects on immunohistochemical reactions. *Histochemistry*, 82(3), pp.201–204.

Ragan, T. et al., 2012. Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nature Methods*, 9(3), pp.255–258.

Richardson, D.S. & Lichtman, J.W., 2015. Clarifying Tissue Clearing. *Cell*, 162(2), pp.246–257.

Rust, M.J., Bates, M. & Zhuang, X., 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Springer Series in Chemical Physics*, 7(10), p.2006.

Schindelin, J. et al., 2012. Fiji: An open-source platform for biological-image analysis.

*Nature Methods*, 9(7), pp.676–682.

Sung, H.W. et al., 1996. Cross-linking characteristics of biological tissues fixed with monofunctional or multifunctional epoxy compounds. *Biomaterials*, 17(14), pp.1405–1410.

Tomer, R. et al., 2014. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nature Protocols*, 9(7), pp.1682–1697.

Uhlen, M. et al., 2010. Towards a knowledge-based Human Protein Atlas. *Nature Biotechnology*, 28(12), pp.1248–1250.

Uhlen, M. & Ponten, F., 2005. Antibody-based proteomics for human tissue profiling. *Molecular and Cellular Proteomics*, 4(4), pp.384–393.

White, J.G. et al., 1986. The structure of the nervous system of the nematode *Caenorhabditis Elegans*. *Phil. Trans. R. Soc. Lond.*, 340, pp.1–340.

Yuan, A. et al., 2015. Neurofilament subunits are integral components of synapses and modulate neurotransmission and behavior in vivo. *Molecular Psychiatry*, 20(8), pp.986–994.

# Chapter 4

## Scalable image processing techniques for quantitative analysis of volumetric biological images from light-sheet microscopy

**Justin Swaney\***, Lee Kamentsky\*, Nicholas Evans, Katherine Xie, Young-Gyun Park,  
Gabrielle Drummond, Dae Hee Yun, Kwanghun Chung<sup>1</sup>

*Reproduced from DOI: <https://doi.org/10.1101/576595>*

### 4.1 Abstract

Here we describe an image processing pipeline for quantitative analysis of terabyte-scale volumetric images of SHIELD-processed mouse brains imaged with light-sheet microscopy. The pipeline utilizes open-source packages for destriping, stitching, and atlas alignment that are optimized for parallel processing. The destriping step removes stripe artifacts, corrects uneven illumination, and offers over 100x speed improvements compared to previously reported algorithms. The stitching module builds upon Terastitcher to create a single volumetric image quickly from individual image stacks with parallel processing enabled by default. The atlas alignment module provides an interactive web-based interface that automatically calculates an initial alignment to a reference image which can be manually refined. The atlas alignment module also provides summary statistics of fluorescence for each brain region as well as region segmentations for visualization. The

---

<sup>1</sup>\* indicates co-first authorship.

expected runtime of our pipeline on a whole mouse brain hemisphere is 1-2 d depending on the available computational resources and the dataset size.

## 4.2 Introduction

Light-sheet fluorescence microscopy (LSFM) is an optical sectioning technique that provides high-speed acquisition of high resolution images. Affordable open-access systems have promoted adoption of LSFM (Pitrone et al. 2013). As a result, LSFM has become commonplace in the study of complex biological systems (Tomer et al. 2012; Stefaniuk et al. 2016; Murakami et al. 2018; Power & Huisken 2017). However, the high-throughput acquisition offered by LSFM can quickly generate terabytes of image data, posing challenges in data storage, processing, and visualization. These challenges must be addressed in order to perform the quantitative analyses needed to answer the complex biological question at hand.

SHIELD is a tissue transformation technique that preserves endogenous biomolecules for imaging within intact biological systems (Park et al. 2018). SHIELD retains fluorescent protein signals through the clearing process and is compatible with stochastic electrotransport staining, allowing visualization and quantification of fluorescence signals throughout the entire brain (Kim et al. 2015). When SHIELD-processed tissues are imaged using LSFM, entire organs such as the mouse brain can be imaged at single-cell resolution in just 2 hours, offering more data than was previously available to answer new biological questions.

Here we present detailed protocols for quantifying fluorescence signals in each brain region of SHIELD-processed mouse brain LSFM datasets. The pipeline is composed of modules for image desstriping, stitching, and atlas alignment. Each module can either be used independently or in combination to perform region-based statistical analyses of fluorescence within intact mouse brainsamples. The pipeline and all its dependencies have been packaged into a single Docker container, allowing for simple installation and cross-platform use. The ease of deployment offered by Docker makes our image processing pipeline more accessible to researchers without much programming experience. We also provide a dataset of a whole mouse brain hemisphere for users to test our pipeline.

### 4.2.1 Development of the protocol

In order to analyze large-scale volumetric images acquired using LSFM, research labs typically create their own image processing pipelines (Amat et al. 2015; Stegmaier et

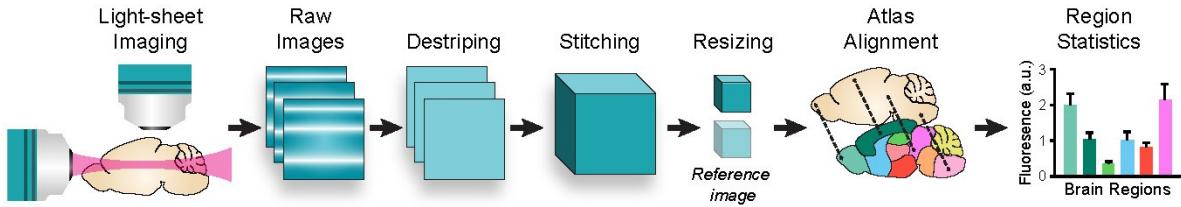
al. 2016; Vogelstein et al. 2018). These image processing pipelines are designed to solve specific problems in applying LSFM to the study of complex biological systems. Real-time cell tracking systems have been reported to study the dynamics of embryogenesis in *D. melanogaster*. The cell tracking pipeline relies on optimized CUDA programming to achieve real-time performance. Several computational pipelines geared toward processing time-lapse images of *D. melanogaster* assume that each time point image is smaller than the amount of available memory. In contrast, LSFM of whole mammalian organs often generates individual volumetric images that are larger than the amount of available memory.

Recently, LSFM images of a whole mouse brain have been used to create a single-cell mouse brain atlas (Murakami et al. 2018). The pipeline consisted of a heterogeneous mix of MATLAB, Python, and C++ software as well as expensive computer hardware, including a dedicated image processing server equipped with four NVIDIA graphics processing units (GPU). In order to handle individual volumetric images that are larger than the amount of available memory, images were processed slice-by-slice for cell detection and rescaled to a manageable size for atlas alignment. Although computationally impressive, such tools often require a great deal of programming expertise or access to proprietary software. As a result, the current large-scale image processing pipelines may be inaccessible to non-experts, and there is a need for large-scale image processing tools for researchers focused on biological questions rather than computational challenges.

The protocols presented here are designed to be easy to setup and applicable to users without much experience in setting up complex development environments. Since some users may only want to use part of our pipeline, the protocols are partitioned into three computational modules: first, our image desstriping for removing streaks and performing flat-field correction in raw LSFM images; second, stitching for creating a single 3D image from the individual 2D images; and third, semi-automatic atlas alignment for segmenting brain regions and quantifying fluorescence (Fig. 4.1). Our protocols have been tested on images of SHIELD-processed mouse brain hemispheres acquired with an axially-swept light-sheet microscope.

#### 4.2.1.1 Development of the desstriping module

Stripe artifacts are commonplace in images acquired with LSFM due to irregularities in the refractive index (RI) of the sample (Salili et al. 2018). This RI mismatch can be compensated for using an immersion medium that has a similar RI to that of the sample (Murray et al. 2015). However, the material properties of biological tissues, including the RI, are generally not uniform throughout, making some degree of RI mismatch inevitable.



**Figure 4.1:** Overall image processing pipeline for whole brain analysis. Raw images from LSFM of a SHIELD-processed mouse hemisphere are destriped and corrected for uneven illumination. Destriped images are then stitched into a multichannel volumetric image, which is resampled to match a reference atlas. Point correspondences from an automatic alignment procedure are manually refined to obtain a region segmentation for the full-resolution, stitched image. The region segmentation is then used to quantify mean fluorescence in each brain region.

RI mismatch usually results in uneven illumination patterns due to optical aberrations that disrupt the incident light.

Current strategies for image destriping are either based on optical filtering or digital filtering (Fehrenbach et al. 2012; Münch et al. 2009; Liang et al. 2016). Optical filtering strategies attempt to compensate for RI mismatch during imaging, effectively removing the stripe artifacts from the source. However, these methods may disrupt the axial resolution of the LSFM system in the process and may not be applicable to large biological samples. In contrast, digital filtering strategies attempt to remove the stripe artifacts after acquisition by exploiting the noise characteristics induced by the optical aberrations. Since digital destriping methods are implemented as image filters, they can be applied more generally to any images with stripe artifacts.

Previous digital destriping methods have included hybrid wavelet-FFT filters, variational removal of stationary noise (VSNR), and multidirectional filters using the contourlet transform (MDSR). Although VSNR and MDSR have shown superior destriping performance, they are prohibitively slow for applying to whole-brain datasets. The hybrid wavelet-FFT filter is the fastest destriping method of these, but its implementation requires a MATLAB license to use and is single-threaded.

In order to provide a fast, open-source destriping solution, we implemented a new digital destriping tool called `pystripe`. `Pystripe` is a Python implementation of the previously reported hybrid wavelet-FFT destriping method with parallel processing support and other improvements. `Pystripe` uses open-source tools instead of proprietary software such as MATLAB. As in the hybrid wavelet-FFT approach, the amount of filtering in `pystripe` can be tuned using a bandwidth parameter. `Pystripe` also adds support for a dual-band filtering mode where the background and foreground of the images can be filtered with separate bandwidths.

#### 4.2.1.2 Development of the stitching module

Imaging large samples with LSFM involves acquiring partially overlapping image stacks which can be stitched together into a single image stack. Several open-source stitching packages are available (Schindelin et al. 2012; Bria & Iannello 2012). Terastitcher has been widely adopted for stitching large volumetric images acquired with LSFM. However, the Terastitcher merging step executes within a single thread by default, resulting in longer execution times than necessary. It should be noted that the Terastitcher team provides a parallelized version of Terastitcher based on message passing interface (MPI) upon request, but we found implementing our own merging step based on the multiprocessing module in Python to be more straightforward than managing MPI.

To address these shortcomings, we created the TSV (Terastitcher Volume) module, which implements the Terastitcher merging step in Python with support for lossless TIFF compression and parallel processing. TSV uses the stack displacements computed from Terastitcher to create a memory-mapped array representing the entire image volume. Multiple workers use this memory-mapped array to convert individual images into a single stack, providing faster overall execution. Each worker merges images together and then saves the result using lossless TIFF compression, resulting in lower overall dataset sizes.

#### 4.2.1.3 Development of the atlas alignment module

In order to segment whole-brain LSFM images into different brain regions, the stitched dataset must be registered to a reference atlas, such as the Allen Mouse Brain Atlas (ABA) (Lein et al. 2007). The ABA consists of an averaged anatomical reference image of autofluorescence and the corresponding region segmentation image. The ABA also contains tools for registering 3D reconstructions from histological sections to the atlas. However, research labs have resorted to custom atlas alignment methods for LFSM images of intact brain samples (Murakami et al. 2018).

Elastix is an open-source medical image registration library that is widely used for non-rigid atlas alignment (Klein et al. 2010). Elastix performs non-rigid atlas alignment by maximizing the mutual information between source and reference images. Elastix was found to have the highest mutual-information benchmark scores in image registration of cleared brain samples among five freely-available software packages (Nazib et al. 2018). The global optimization of mutual information is difficult to scale to whole-brain LSFM datasets since the entire dataset cannot be stored in memory. Following previous work on atlas alignment, we address this issue by rescaling the source image to be a similar size compared to the reference atlas, which is a more manageable size. We use the alignment

computed from Elastix to generate a set of approximate point correspondences which can manually refined.

To visualize the atlas alignment and edit the approximate point correspondences, we created an interactive web-based registration tool called nuggt (NeUroGlancer Ground Truth). Nuggt is built on an open-source visualization package called Neuroglancer. Nuggt does not modify the underlying Neuroglancer code but rather wraps it into a convenient package for interactive LSFM visualization and atlas alignment. Nuggt displays the source and reference images side-by-side along with the point correspondences overlaid on each image. Using nuggt, the point correspondences can be edited and adjusted to improve the atlas alignment. The source image can also be warped while editing the point correspondences, providing rapid visual feedback of the atlas alignment accuracy.

#### 4.2.1.4 Development of the Docker image

One of the main challenges in adopting a new computational pipeline is obtaining the dependencies and recreating the runtime environment that was intended by the developers. In order to simplify the installation of our pipeline, we packaged our pipeline and all of its dependencies into a single Docker container by creating a Dockerfile describing our runtime environment. Docker provides a consistent, light-weight virtual Linux environment on all major operating systems, and our container has been successfully tested on Windows, Mac, and Linux. Docker can be installed from the Docker website, and our container can be downloaded from Docker Hub. By containerizing our pipeline and using web-based visualization, our modules can either be run locally or on dedicated image processing servers which can be accessed from other clients.

#### 4.2.2 Applications of the method

The overall image processing pipeline described in this protocol is designed to calculate fluorescence summary statistics from whole-mouse brain images acquired with LSFM on a per-region basis. Our pipeline has been used to quantify mRuby2 and EGFP fluorescence of virally labeled neurons and presynaptic terminals in SHIELD-processed mouse brain hemispheres 6. Thus, the overall pipeline may be applied in systems neuroscience to quantify fluorescent reporters in cleared samples from mouse models. However, the individual modules that comprise the overall pipeline can also be used independently.

Pystripe can be applied to any images corrupted with horizontal or vertical stripe artifacts. We restricted pystripe to filtering horizontal or vertical stripes because the illumination beam path in most LSFM systems is aligned with the camera detector. Pystripe can,

therefore, also be used with multi-view LSFM systems that rotate the sample rather than change the orientation of the illumination beam path (Preibisch et al. 2014). Pystripe also includes the ability to provide a reference flat-field image for illumination correction of vignetting and other stationary artifacts.

TSV can be used to merge an array of partially overlapping image stacks saved in Terastitcher hierachal format into a single image stack. The memory-mapped array used for stitching is also useful for retrieving sub-volumes of image data. TSV also includes optional utilities for partitioning the stitched image into smaller, uniformly shaped chunks for custom parallel processing. The stitched images can be stored in Neuroglancer precomputed format and served via HTTP, allowing for efficient visualization of whole-brain LSFM datasets at full resolution either locally or over the web.

Nuggt has been used to register SHIELD-processed mouse brain hemisphere datasets to the ABA. Our alignment protocol can be used to register a pair of 3D volumes with mutual information between them and a gradient of mutual information in the initial overlap that can be followed to modify the alignment. In our experience, images that have 90% overlap and rotations of a few degrees can be aligned using this method. The automatic alignment can be refined by adding manually-placed correspondences. The atlas alignment module also provides utilities for calculating image statistics for each brain region given the aligned atlas segmentation.

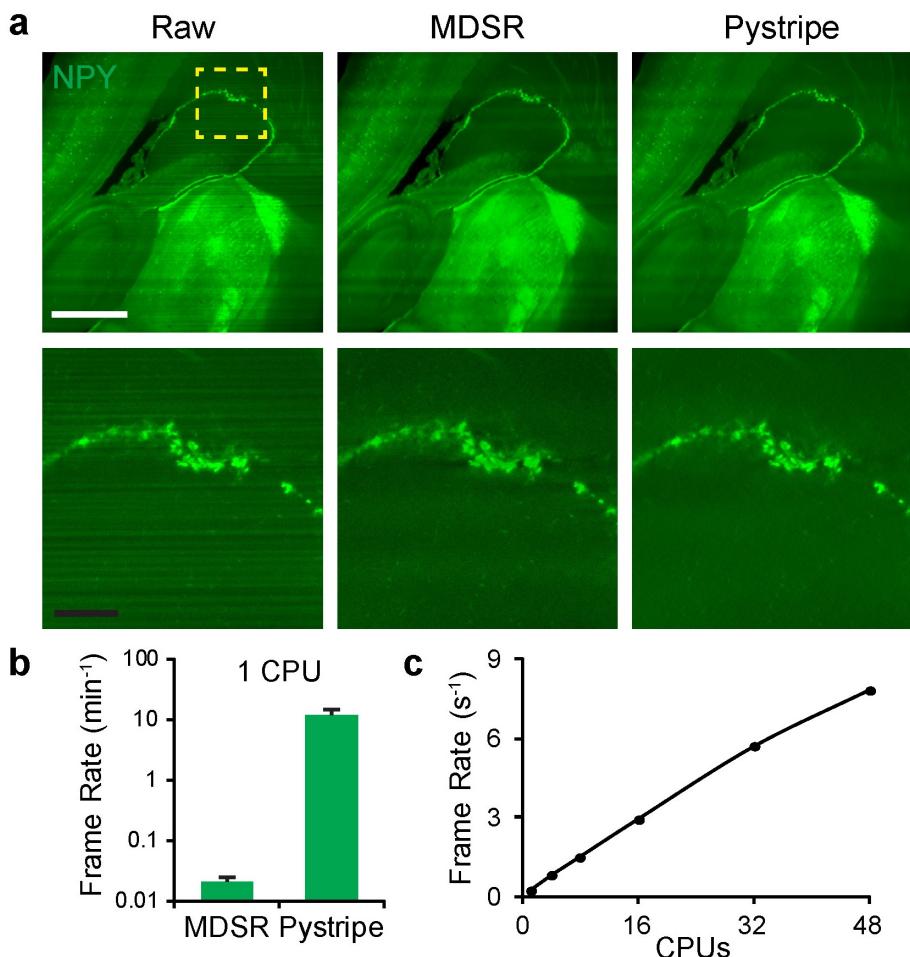
### 4.2.3 Comparison with other methods

Many standard solutions exist for similar tasks addressed in our image processing pipeline. In this section, we compare the methods used in our protocol to existing methods in the context of whole brain LSFM image analysis.

#### 4.2.3.1 Destriping

The previously reported digital destriping algorithm MDSR has achieved state of the art destriping performance on LSFM images. MDSR relies on the contourlet transform to perform energy compaction of striping artifacts in arbitrary orientations, whereas pystripe uses the discrete wavelet transform to remove either horizontal or vertical striping artifacts.

When comparing the resulting images from MDSR and pystripe, similar filtering performances are observed from both methods on our test images (Fig. 4.2a). This suggests that the contourlet transform does not drastically improve energy compaction of the



**Figure 4.2:** Destriping of light-sheet microscopy images using pystripe. (a) Comparison of destriping results from MDSR and pystripe on LSFM images of a SHIELD-processed mouse hemisphere stained for Neuropeptide Y (NPY). Scale bars, 1 mm (white) and 200 m (black). (b) Average single-core destriping speed for MDSR and pystripe on 2048 x 2048 images ( $n = 10$ , error bars indicate standard deviation). (c) Scaling of destriping speed using pystripe with parallel processing on 2048 x 2048 images.

stripe artifacts compared to the discrete wavelet transform when the stripes are oriented horizontally. When comparing the average execution speed for destriping using MDSR and `pystripe` on a single core, MDSR takes over 30 min per frame, while `pystripe` takes only 5 seconds (Fig. 4.2b). Using multiple cores, the destriping frame rate for `pystripe` increases linearly with the number of cores, reaching 8 frames per second with 48 cores (Fig. 4.2c, Supplementary Video 1).

`Pystripe` also allows the user to provide an optional reference image for flat-field correction. Ideally, the reference flat would be calculated retroactively from the imaging data, but in practice a single flat for each channel of a particular imaging system still provides significant improvement. The example dataset includes a reference flat for illumination correction during the destriping step. By performing the illumination correction in `pystripe`, reading and writing the whole dataset multiple times can be avoided.

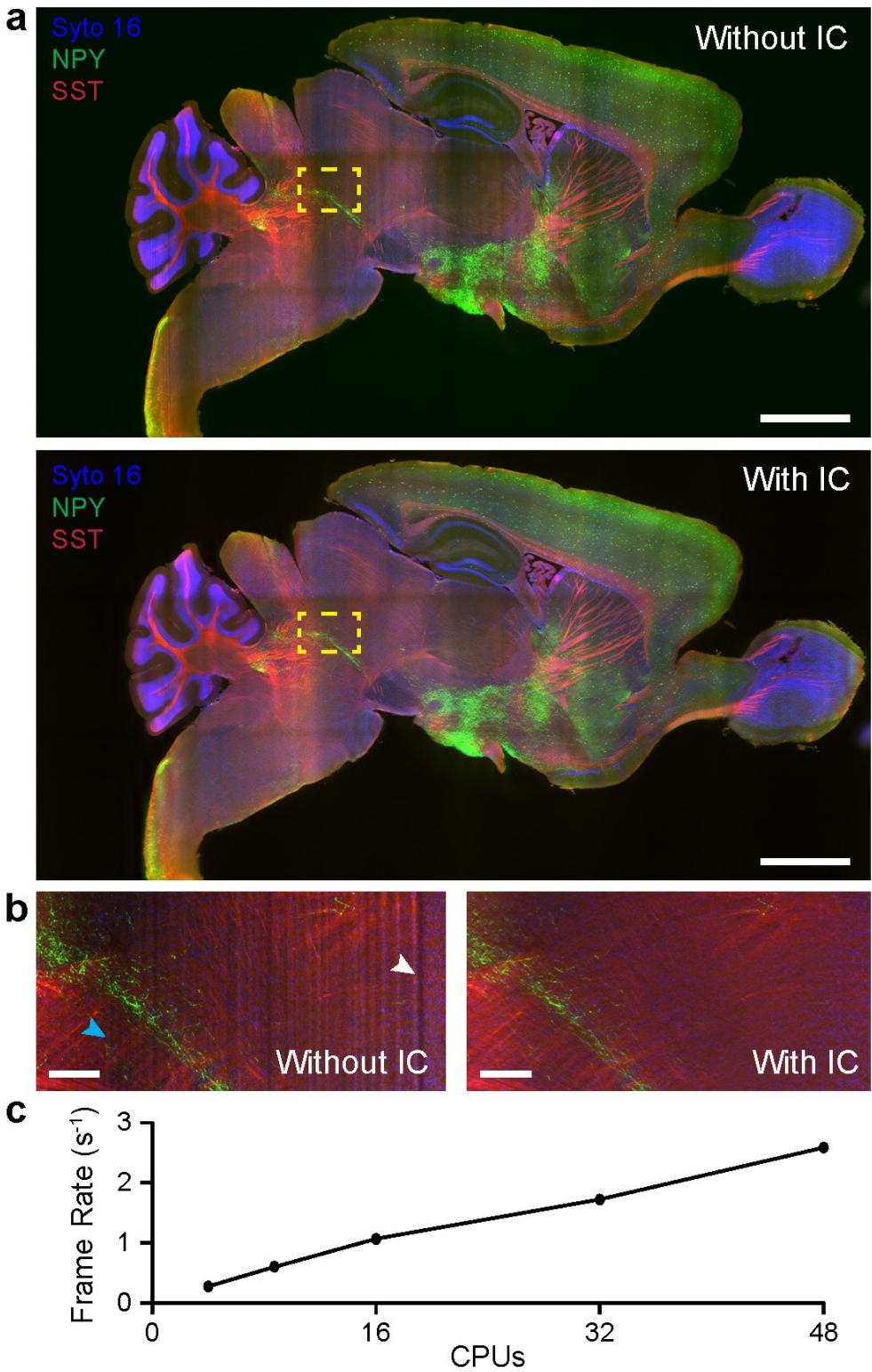
#### 4.2.3.2 Stitching

Building on Terastitcher, TSV allows fast merging of stacks saved in Terastitcher hierarchical format. TSV obtains similar stitching quality as Terastitcher since it uses the same stack displacements and blending functions. Using TSV, a whole mouse hemisphere dataset was stitched with and without illumination correction and destriping using `pystripe` (Fig. 4.3a). Moderate vignetting effects were visible in the stitched original images at the intersections between adjacent stacks. These tiling artifacts were effectively reduced using flat-field correction in `pystripe`. Together, `pystripe` and TSV generate volumetric images that are ready for quantification by removing shadow and tiling artifacts before stitching (Fig. 4.3b, Supplementary Video 2). When comparing the stitching speed of TSV using multiple cores, the stitching frame rate increases linearly, reaching 2.6 frames per second with 48 cores (Fig. 4.3c).

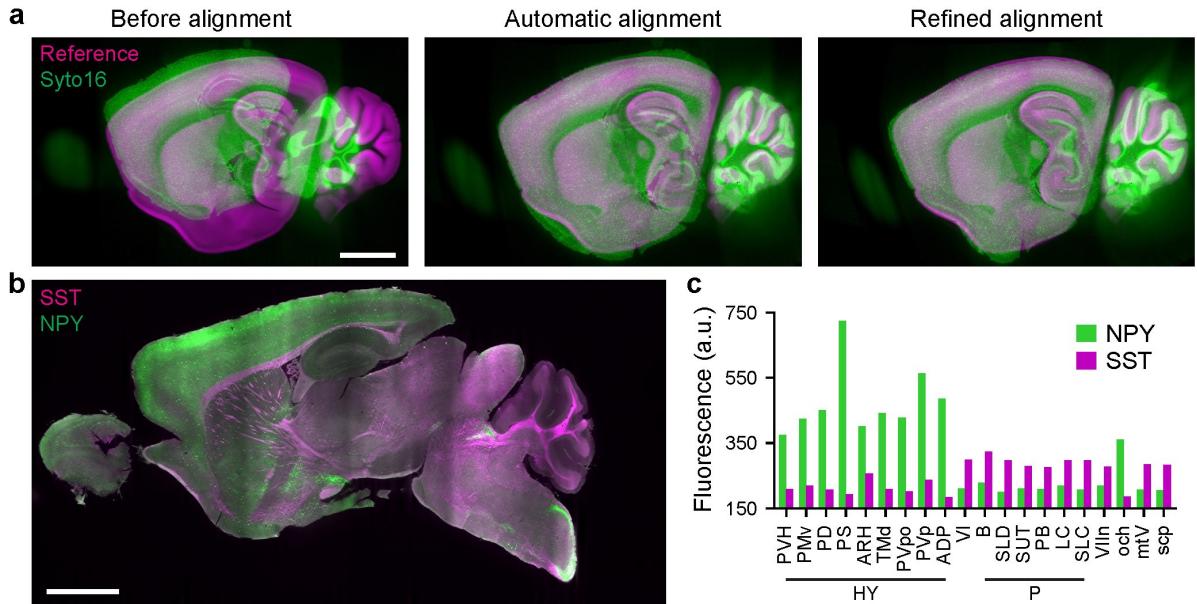
#### 4.2.3.3 Atlas Alignment

Our hybrid automated atlas alignment method with manual refinement differs from wholly automated methods in that the alignment can be improved to the desired degree of accuracy via addition and modification of correspondences between the two volumes to be aligned (Fig. 4.4a). Tools for manual refinement are generally not used in combination with Elastix because integrating the transformations across multiple registration tools can be challenging. To the best of our knowledge, there are no web-based tools for interactive atlas alignment currently available.

After aligning a source autofluorescence or nuclear stain image to the reference image,



**Figure 4.3:** Stitching and illumination correction of light-sheet microscopy images using TSV. (a) Comparison of stitching results with and without destriping and illumination correction (IC) performed with pystripe on LSFM images of a SHIELD-processed mouse hemisphere. Scale bar, 2 mm. (b) Comparison of a region of interest with and without IC. Without IC, both uneven illumination (cyan arrow) and stripe artifacts (white arrow) corrupt the images. (c) Scaling of stitching speed using TSV with parallel processing on 2048 x 2048 images.



**Figure 4.4:** Atlas alignment and region-based fluorescence quantification using nuggt. (a) Comparison of atlas alignment of a syto 16 LSFM image to the reference image in the ABA before alignment, after automatic alignment, and after manual refinement. Scale bar, 2 mm. (b) Overlay of NPY and SST in the whole hemisphere example dataset (c) Mean fluorescence for the top 10 regions for both NPY and SST calculated after manual refinement. All other regions from the ABA are omitted for clarity.

other channels can be aligned to the atlas using the same calculated alignment. For example, neuropeptide Y (NPY) and somatostatin (SST) expression are included with syto 16 in separate channels of the provided example data (Fig. 4.4b). Using the alignment calculated by registering the syto 16 channel and the reference from the ABA, the mean fluorescence intensity of NPY and SST in each brain region can be calculated (Fig. 4.4c).

#### 4.2.4 Experimental design

All software modules are available from Github at

<http://www.github.com/chunglabmit/shield-2018>

as well as from Docker hub at

<https://hub.docker.com/r/chunglabmit/shield-2018>

We also provide example LSFM images of a SHIELD-processed mouse hemisphere dataset, which is available from

<http://leviathan-chunglab.mit.edu/nature-protocols-2019>

In order to adapt our image processing pipeline to other experimental situations, users should first complete our protocol using the provided example dataset. This dataset includes raw LSFM images as well as our intermediate results for users to compare and checkpoint their results throughout the pipeline. We also include a downsampled version of the full example dataset for users that would like to try our protocol on more modest computational hardware.

Young adult (2–4 months; median age 3 months) C57BL/6 mice were housed in a 12 h light/dark cycle with unrestricted access to food and water. To generate the example dataset, a single mouse brain was SHIELD-processed and stained with syto 16 and antibodies targeting NPY and SST using stochastic electrotransport. The mouse brain sample was cut along the mid-sagittal plane and includes the olfactory bulb and the cerebellum. The stained hemisphere was then incubated in a RI-matching solution and imaged using a custom axially-swept LSFM system equipped with a 3.6x/0.2 objective (Special Optics). The resulting voxel width and depth are 1.8 m and 2.0 m, respectively. Although only one animal was involved in preparing the example dataset, our image processing protocol has been tested on over 15 intact mouse brain hemispheres from separate animals. All experimental protocols were approved by the MIT Institutional Animal Care and Use Committee and the Division of Comparative Medicine and were in accordance with guidelines from the National Institute of Health. All experiments using mice were conducted in strict adherence to the ethical regulations of MIT Institutional Animal Care and Use Committee and the Division of Comparative Medicine.

Our protocols have been developed for images of SHIELD-processed mouse brain tissues sectioned along the midsagittal plane with or without the olfactory bulb and cerebellum excised. The images must be acquired in a geometry that allows a transformation of axes (flipping, transposition) and cropping that bring the images into rough alignment with the atlas. In our experience, acquisition of either autofluorescence or a nuclear stain such as syto 16 provides enough mutual information for alignment with the reference volume for the ABA. Images were processed using only the techniques described in the protocol, and figures were prepared using linear lookup tables with adjustment of the minimum and maximum display range.

#### 4.2.5 Expertise needed to implement the protocol

Some minimal computer skills are needed to install Docker and navigate using the command line. If you are unfamiliar with Docker, important introductory information about Docker can be found at <https://docs.docker.com/get-started/>. Our protocol relies on several Docker features for volume sharing and network access, so a basic understanding

of Docker is a prerequisite. Since Docker and most of our software is used from the command line, a basic understanding of how to use a terminal on the host operating system is also required.

#### 4.2.6 Limitations

Our overall image processing pipeline currently has been tested for analyzing mouse brain hemisphere datasets using the ABA autofluorescence reference volume. Therefore, we cannot guarantee that our overall pipeline will work in all cases for other species and atlases.

As previously mentioned, `pystripe` cannot remove striping artifacts that are not horizontally or vertically oriented within raw images. When processing images with very bright signals, the hybrid wavelet-FFT filter may introduce some ringing artifacts to the destriped images. The dual-band mode can mitigate these ringing artifacts, but in some cases, these artifacts may be undesirable. The user can elect to reduce the filter bandwidth or skip the destriping step in such cases.

When analyzing mouse brain hemispheres, severe tissue deformation due to improper sample preparation or handling may result in poor atlas alignment since the initial non-rigid alignment may not converge to a global optimum in such cases. Our atlas alignment protocol can be applied to mouse brain hemispheres with the olfactory bulb and cerebellum excised; however, this excision must be mirrored in the anatomical reference volume. This involves cropping the anatomical reference volume to match the source image, which is somewhat ad hoc. We provide referencevolumes from the ABA for use in our Docker container along with reference volumes with the olfactory bulb, cerebellum, or both excised.

Processing whole-brain LSFM datasets is computationally expensive. We provide minimum system requirements that are recommended for processing our example mouse hemisphere dataset. However, we also provide a downsampled version of the original dataset for users without immediate access to computer hardware that meet the minimum system requirements.

## 4.3 Materials

### 4.3.1 Equipment

#### 4.3.1.1 Example Data

All Example data is available from our laboratory servers<sup>2</sup> (<http://leviathan-chunglab.mit.edu/nature-protocols-2019>), including:

- Raw data, comprising the raw LSFM images and a flat reference image from the provided mouse brain dataset.
- Destriped data, comprising the destriped LSFM images from the provided dataset.
- Stitched data, comprising the stitched destriped LSFM images from the provided dataset.
- Alignment data, containing the mean fluorescence in each region and intermediate results obtained during atlas alignment.
- Downsampled data, comprising our whole brain raw data downsampled and zipped for testing our pipeline on more modest computer hardware.

#### 4.3.1.2 Computer Equipment

- For benchmarks, the computational pipeline was deployed on a workstation (TWS-1686525, Exxact) running Ubuntu 16.04 LTS on a 1 TB solid-state drive (Samsung) with two 24-core processors (Intel Xeon Platinum 8168), 768 GB of ECC memory, and a single 16 GB NVIDIA GPU (Tesla P100).
- Recommended minimum system requirements for processing the full-resolution example LSFM dataset. A computer with enough hard drive space to store the raw image data as well as the intermediate processing results is required. Our total example data is approximately 3 TB, suggesting that at least 4 TB of extra space is required. We recommend the following minimum system requirements in this case:
  - 8-core processor
  - 64 GB of memory
  - 256 GB solid-state drive with at least 32 GB available
  - 4 TB HDD available for data storage
- Recommended minimum system requirements for the processing the downsampled example LSFM dataset:

---

<sup>2</sup>Visit [www.chunglab.org](http://www.chunglab.org) and contact the Chung Lab directly in case of a 404 error

- 2-core processor
- 16 GB of memory
- 128 GB solid-state drive with at least 32 GB available
- Software requirements.
  - Docker is recommended to run our software, but expert users can also directly install our software using the pip Python package manager from GitHub. If using our preconfigured Docker image to run our pipeline, Docker must be installed locally. Docker offers a free version of their software called Docker Community Edition.
  - We recommend using FIJI for inspecting images throughout our pipeline. FIJI can be obtained from <https://fiji.sc/>

### 4.3.2 Equipment Setup

#### 4.3.2.1 Docker setup

1. Create a Docker ID to gain access to Docker software and Docker Hub using your preferred email address at <https://hub.docker.com/>. You may need to verify your email address.
2. Download Docker Community Edition for your particular operating system at <https://www.docker.com/products/docker-engine>.
3. Follow the installation instructions and the provided commands to verify that Docker has been installed correctly.
4. With the Docker deamon running, open a terminal and run the following command to install our preconfigured Docker image:

```
docker pull chunglabmit/shield-2018
```

This command will download our software including all of the dependencies as well as commonly used resources needed for atlas alignment with the ABA. Note that we refer to the computer that is running Docker as the “host” and an instance of our Docker image as the “container”.

#### 4.3.2.2 Downloading full resolution example data

This step requires 4 TB of available hard drive space if the whole dataset and intermediate results are downloaded. The raw data is approximately for each channel is approximately

560 GB. Proceed to *downloading downsampled example data* section if your computer does not have this much available space.

1. Create a folder named “data” to contain all of the example data on your machine and note its full path.
2. Start the Docker container with the data folder mounted by entering the following command into the command line:

```
docker run -it -v path_to_data:/data chunglabmit/shield-2018
```

where “path\_to\_data” should be replaced with the full path to the data folder on the host. The command prompt should indicate that you are now the root user inside a running Docker container (as opposed to your usual username on the host). We refer to the command prompt inside a running container as the “Docker terminal window”.

3. Download the example data needed for a particular stage in the protocol by entering the following command(s) into the Docker terminal window. The data should begin to appear in the data folder that was created. Note that each download may take several hours.

- Raw data (needed for destriping)

```
wget -P /data -r --no-parent -Nh --cut-dirs 1 -R "index.html*" \
http://leviathan-chunglab.mit.edu/nature-protocols-2019/raw_data/
```

- Destriped data (needed for stitching)

```
wget -P /data -r --no-parent -nH --cut-dirs 1 -R "index.html*" \
http://leviathan-chunglab.mit.edu/nature-protocols-2019/destriped_data/
```

- Stitched data (needed for atlas alignment)

```
wget -P /data -r --no-parent -nH --cut-dirs 1 -R "index.html*" \
http://leviathan-chunglab.mit.edu/nature-protocols-2019/stitched_data/
```

- Alignment data (for comparison of results)

```
wget -P /data -r --no-parent -nH --cut-dirs 1 -R "index.html*" \
http://leviathan-chunglab.mit.edu/nature-protocols-2019/atlas/
```

#### 4.3.2.3 Downloading downsampled example data

1. Navigate to <http://leviathan-chunglab.mit.edu/nature-protocols-2019/> in your browser.
2. Click the `downsampled_data.zip` link and choose to save the file instead of opening it.
3. Unzip the `downsampled_data.zip` file to a new folder called “data” and make note of the full path of this folder on the host.

## 4.4 Procedure

*See Appendix for step-by-step protocol*

## 4.5 Anticipated Results

Upon successful completion of our image processing pipeline, our protocol will yield corrected and stitched multichannel volumetric images of a whole mouse brain hemisphere as well as an alignment with a provided atlas for brain region segmentation. The atlas alignment is used to create spreadsheets in CSV format containing the volume, total fluorescence, and mean fluorescence of each brain region and each channel. These results can be used to create visualizations summarizing the fluorescence in each brain region (Fig. 4c). The reported regions volumes are in voxel units, so the physical volumes will depend on the voxel dimensions used during imaging.

## 4.6 Acknowledgements

We thank the entire Chung laboratory for support and discussions. K.C. was supported by the Burroughs Wellcome Fund Career Awards at the Scientific Interface, Searle Scholars Program, Packard award in Science and Engineering, NARSAD Young Investigator Award, and the McKnight Foundation Technology Award. This work was supported by the JPB Foundation (PIIF and PNDRF), the NCSOFT Cultural Foundation, the Institute for Basic Science IBS-R026-D1, and the NIH (1-DP2-ES027992, U01MH117072)

## **4.7 Author Contributions**

J.S. worked on `pystripe` and prepared the main figures. L.K. worked on `TSV` and `nuggt`. K.X., Y.G.P. and D.H.Y prepared the example brain sample, and N.E. acquired the example data. G.D., K.X. and N.E. tested the protocol and provided feedback. J.S. and L.K. wrote the paper together, and all authors reviewed the manuscript. K.C. supervised the project.

## **4.8 Competing Interests**

The authors declare no competing interests.

## 4.9 References

- Amat, F. et al., 2015. Efficient processing and analysis of large-scale light-sheet microscopy data. *Nature Protocols*, 10(11), pp.1679–1696.
- Bria, A. & Iannello, G., 2012. TeraStitcher - A tool for fast automatic 3D-stitching of teravoxel-sized microscopy images. *BMC Bioinformatics*, 13(1).
- Fehrenbach, J. et al., 2012. Variational algorithms to remove stationary noise: Application to SPIM imaging.
- Kim, S.-Y. et al., 2015. Stochastic electrotransport selectively enhances the transport of highly electromobile molecules. *Proceedings of the National Academy of Sciences*, 112(46), pp.E6274–E6283. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1510133112>.
- Klein, S. et al., 2010. Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1), pp.196–205.
- Lein, E.S. et al., 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 445(7124), pp.168–176.
- Liang, X. et al., 2016. Stripe artifact elimination based on nonsubsampled contourlet transform for light sheet fluorescence microscopy. *Journal of Biomedical Optics*, 21(10), p.106005. Available at: <http://biomedicaloptics.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JBO.21.10.106005>.
- Murakami, T.C. et al., 2018. A three-dimensional single-cell-resolution whole-brain atlas using CUBIC-X expansion microscopy and tissue clearing. *Nature Neuroscience*, 21(4), pp.625–637. Available at: <http://dx.doi.org/10.1038/s41593-018-0109-1>.
- Murray, E. et al., 2015. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6), pp.1500–14.
- Münch, B. et al., 2009. Stripe and ring artifact removal with combined wavelet-Fourier filtering. *EMPA Activities*, 17(2009-2010 EMPA ACTIVITIES), pp.34–35.
- Nazib, A. et al., 2018. Performance of Image Registration Tools on High-Resolution 3D Brain Images. Available at: <https://arxiv.org/abs/1807.04917>.
- Park, Y.-g. et al., 2018. Protection of tissue physicochemical properties using polyfunctional., (December).
- Pitrone, P.G. et al., 2013. OpenSPIM: an open-access light-sheet microscopy platform. *Nature Methods*, 10(7), pp.598–599. Available at: <http://www.nature.com/doifinder/10>.

1038/nmeth.2507.

Power, R.M. & Huisken, J., 2017. A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nature Methods*, 14(4), pp.360–373. Available at: <http://arxiv.org/abs/nmeth.4224>.

Preibisch, S. et al., 2014. Efficient bayesian-based multiview deconvolution. *Nature Methods*, 11(6), pp.645–648. Available at: <http://arxiv.org/abs/1308.0730>.

Salili, S.M., Harrington, M. & Durian, D.J., 2018. Note: Eliminating stripe artifacts in light-sheet fluorescence imaging. *Review of Scientific Instruments*, 89(3), pp.43–45.

Schindelin, J. et al., 2012. Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), pp.676–682.

Stefaniuk, M. et al., 2016. Light-sheet microscopy imaging of a whole cleared rat brain with Thy1-GFP transgene. *Scientific Reports*, 6, pp.1–9.

Stegmaier, J. et al., 2016. Real-Time Three-Dimensional Cell Segmentation in Large-Scale Microscopy Data of Developing Embryos. *Developmental Cell*, 36(2), pp.225–240. Available at: <http://dx.doi.org/10.1016/j.devcel.2015.12.028>.

Tomer, R. et al., 2012. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods*, 9(7), pp.755–763. Available at: <http://dx.doi.org/10.1038/nmeth.2062>.

Vogelstein, J.T. et al., 2018. A community-developed open-source computational ecosystem for big neuro data. *Nature methods*, 15(November), pp.846–847.

# Chapter 5

## Nuclei-based 3D coregistration for multiplexed whole-brain imaging at single-cell resolution

**Justin Swaney, Dae Hee Yun, Nicholas Evans, Katherine Xie, Lee Kamentsky, Kwanghun Chung**

### 5.1 Abstract

Fluorescence imaging is a workhorse technology for several experimental studies in biology. However, confocal and light-sheet microscopy are limited in the number of channels that can be imaged simultaneously due to spectral and optical constraints from conventional fluorophores. To address this issue, we propose a multiround imaging strategy in which a single tissue sample is stained for multiple targets in each round. In order to align these multiple rounds of imaging into single coordinate system, we present an efficient 3D coregistration algorithm that identifies corresponding nuclei between imaging rounds to achieve a nonrigid alignment that is accurate at the single-cell level. We demonstrate our coregistration algorithm on a mouse brain hemisphere and a large block of marmoset visual cortex. The resulting image alignment contains an average displacement of 4  $\mu\text{m}$  between corresponding nuclei. Using GPU-accelerated computation, we demonstrate that warping the entire volumetric images from a mouse brain hemisphere and marmoset visual cortex is possible within 6 hours on an NVIDIA P100 GPU. Future work may use this nuclei-based 3D coregistration algorithm to achieve multiplexed fluorescence microscopy and perform more exploratory image analysis at the single-cell level.

## 5.2 Introduction

Fluorescence imaging techniques such as confocal microscopy and light-sheet fluorescence microscopy (LSFM) are commonly used in neuroscience and biology (Power & Huisken 2017). In combination with immunohistochemistry (IHC), fluorescence microscopy can provide spatial protein expression information using antibodies that specifically bind to proteins of interest. In combination with *in situ* hybridization, fluorescence microscopy can provide spatial transcription information using complimentary oligonucleotide probes that bind to RNA sequences of interest (Edsgård et al. 2018). Recent advances in clearing and staining such as SWITCH, MAP, and eFLASH have extended these valuable applications of fluorescence microscopy to large intact tissues such as the mouse, marmoset, and human brain (Murray et al. 2015; Ku et al. 2016; Yun et al. 2019).

Although these breakthrough technological advances have scaled up the effective tissue size that can be studied using conventional antibodies and fluorescence imaging, they are still inherently limited by the number of target proteins that can be imaged simultaneously. In contrast, single-cell RNA sequencing techniques allow for highly multiplexed quantification of transcripts, which is useful in exploratory analyses and for hypothesis generation (Hwang et al. 2018). This difference in the number of target molecules that can be quantified in fluorescence microscopy and RNA sequencing is not due to a lack of antibodies for protein targets; many more antibodies are available than can be imaged simultaneously. Instead, the limiting factor is the number of imaging channels available in the microscope, which is typically limited to between 3 and 6.

The channel limitation of fluorescence microscopy is intimately linked to the fluorophores used in IHC. The mostly commonly used fluorophores are Alexa Fluor dyes, which all have spectral widths of approximately 70 nm as measured by the full-width at half maximum (FWHM) of the emission spectra (Panchuk-Voloshina et al. 1999). Given that the typical range of detectors used in confocal microscopy is between 400 and 800 nm, only 3 to 6 Alexa Fluor dyes would be resolvable without specialized spectral overlap compensation. Without similar breakthroughs in optics or fluorophore chemistry, this limitation is inevitable for simultaneous multicolor fluorescence imaging.

Another approach to multiplexed fluorescence imaging is to image different proteins in multiple rounds of staining. Staining different proteins in separate mouse brains is a common strategy to study population averages given a reference atlas such as the Allen Brain Atlas (Lein et al. 2007). In the Allen Brain Atlas, multiple brains are aligned to a common coordinate system to allow aggregation of protein expression information across many brain samples for each brain region. Although powerful for enabling neuroscience research, these atlas-based approaches do not provide multiplexed protein expression

information for a single brain sample. Extracting expression patterns for more proteins in a single sample becomes more important when dealing with precious samples from non-human primates or human clinical samples.

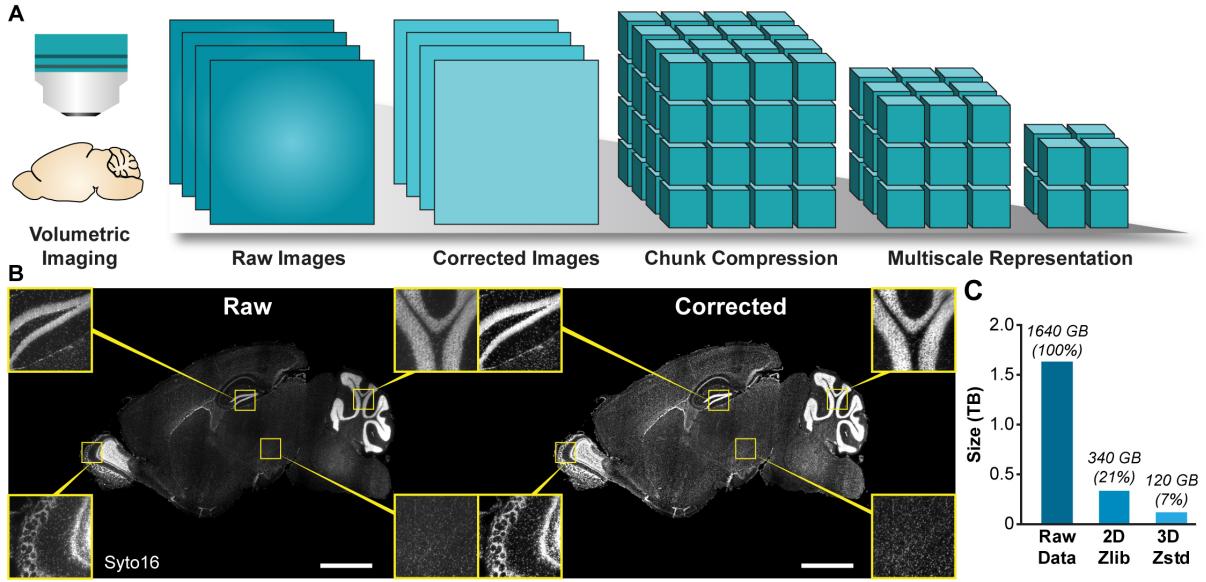
Multiround staining and imaging of a single brain sample has been demonstrated using tissue clearing techniques for over 20 rounds (Murray et al. 2015). Multiple staining rounds of a thin-section of human visual cortex was coregistered using a nonrigid 3D SIFT keypoint matching procedure based on vasculature staining in each round (Lowe 2004). However, the 3D SIFT feature extraction and thin-plate spline (TPS) calculations are not scalable to the whole-brain scale (Bookstein 1989). Here, we present an efficient algorithm for 3D nonrigid coregistration of multiround staining at single-cell resolution that can be applied to whole mouse brains and large marmoset brain samples. Our algorithm matches nuclei centroids across rounds so that the alignment quality can be quantified in terms of single-cell correspondences. As a proof of concept, we apply our algorithm to an intact mouse brain hemisphere as well as a large section of a marmoset brain.

## 5.3 Results

### 5.3.1 Volumetric image preprocessing and compression

In order to develop an algorithm for single-cell coregistration, we first needed to acquire two rounds of imaging for a single mouse brain hemisphere. We SHIELD-processed a mouse brain hemisphere and performed nuclear staining using Syto16. Then, we imaged the entire hemisphere using LSFM before and after remounting with manual tissue handling between imaging rounds.

Multiround imaging of intact mouse brain hemispheres quickly generates terabytes of data. To allow efficient access to the image data and parallel processing, we used the Zarr Python package to partition the image data into more manageable chunks each 64 voxels in width, height, and depth (Figure 5.1a). The Zarr package is open-source and provides lossless compression using the Z-standard compression algorithm. We performed illumination correction using Pystripe (Swaney et al. 2019), stitching using Terastitcher and TSV, and contrast-limited adaptive histogram equalization (CLAHE) before converting to the chunked Zarr format (Figure 5.1a,b). Applying the Z-standard compression algorithm on 3D chunks provided a 2.8x higher compression ratio than the more conventional Z-lib compression on individual 2D TIFF images (Figure 5.1c). After removing the background by thresholding, the uncompressed data from a single round of imaging



**Figure 5.1:** Volumetric image preprocessing and compression. (A) Schematic illustrating the image preprocessing for our pipeline. Volumetric images are acquired as stacks of raw 2D images each with multiple channels. Each image is first corrected for illumination variations across the field of view using local histogram equalization, and streaks are removed via a hybrid wavelet-FFT filter. The corrected images are broken into small chunks for downstream parallel processing, and a downsampling pyramid allows for efficient visualization over multiple length scales. (B) A representative slice of a whole mouse brain hemisphere stained with syto16 before and after preprocessing. Scale bar, 2 mm (C) The storage requirements for a whole hemisphere dataset with no compression, 2D Zlib, and 3D Zstandard lossless compression.

that was originally 1.6 TB was effectively reduced to just 120 GB with no loss in image fidelity within the tissue.

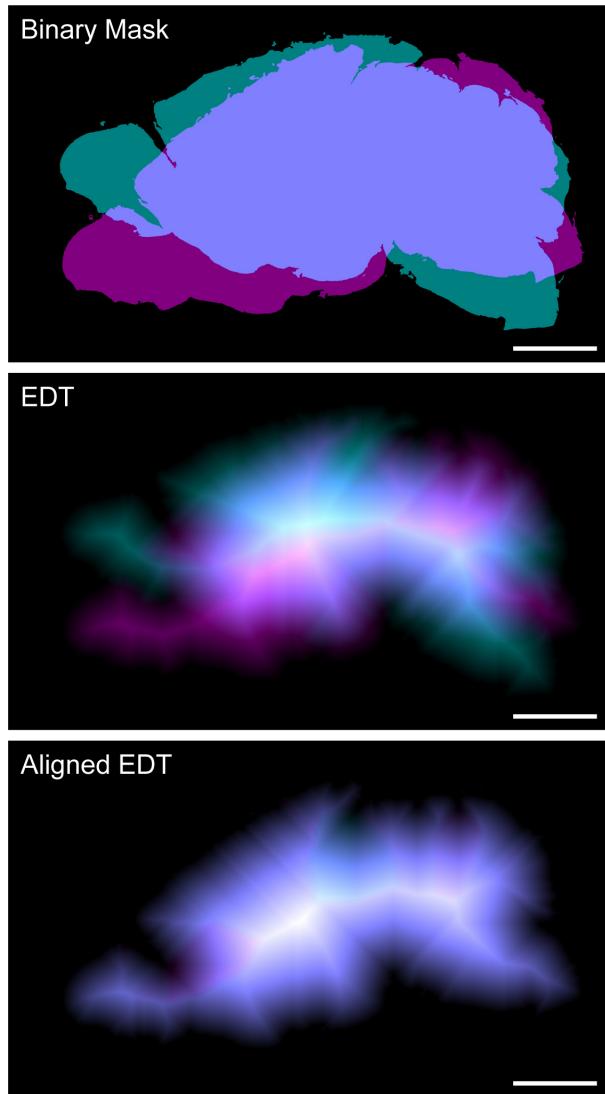
### 5.3.2 Coarse alignment of intact mouse hemisphere

Given that the tissue sample may be oriented arbitrarily in each round of imaging, matching any given nucleus to itself between rounds would require searching through all nuclei in the hemisphere (approximately 30 million nuclei). This “needle-in-a-haystack” problem may be sub-modularized if the approximate locations of the matching nuclei are known *a priori*. If the tissue sample can be roughly aligned between rounds, then the search for a matching nucleus can be restricted to a much smaller set of nuclei within a search radius (typically thousands of cells).

To perform coarse alignment, we first downsampled the original image data 16x isotropically. Downsampling the original resolution images makes iterative, intensity-based image registration techniques computationally feasible. We then sought to align the overall shape of the sample between rounds rather than match similar image intensities directly due to differences in illumination and stitching between rounds. Morphological alignment was performed based on a binary mask outlining the overall tissue shape (Figure 5.2 top). Assuming a rigid deformation model between rounds, we attempted to use L-BFGS to

minimize the squared error between the fixed image binary mask and the moving image binary mask. However, the binary values in the image masks did not provide a reliable gradient signal for the L-BFGS optimizer, and the coarse alignment consistently failed in this case.

By performing the euclidean distance transform (EDT) on the binary masks, a more reliable gradient signal can be provided to the L-BFGS optimizer (Figure 5.2 middle). By minimizing the square error between the EDT of the fixed and moving image binary masks, the parameters of a rigid transformation that roughly align the two rounds can be obtained (Figure 5.2 bottom). These parameters can be converted to the original image resolution and then be applied to nuclei centroids in the fixed image. The result is an approximate alignment between nuclei detections in the fixed and moving images.



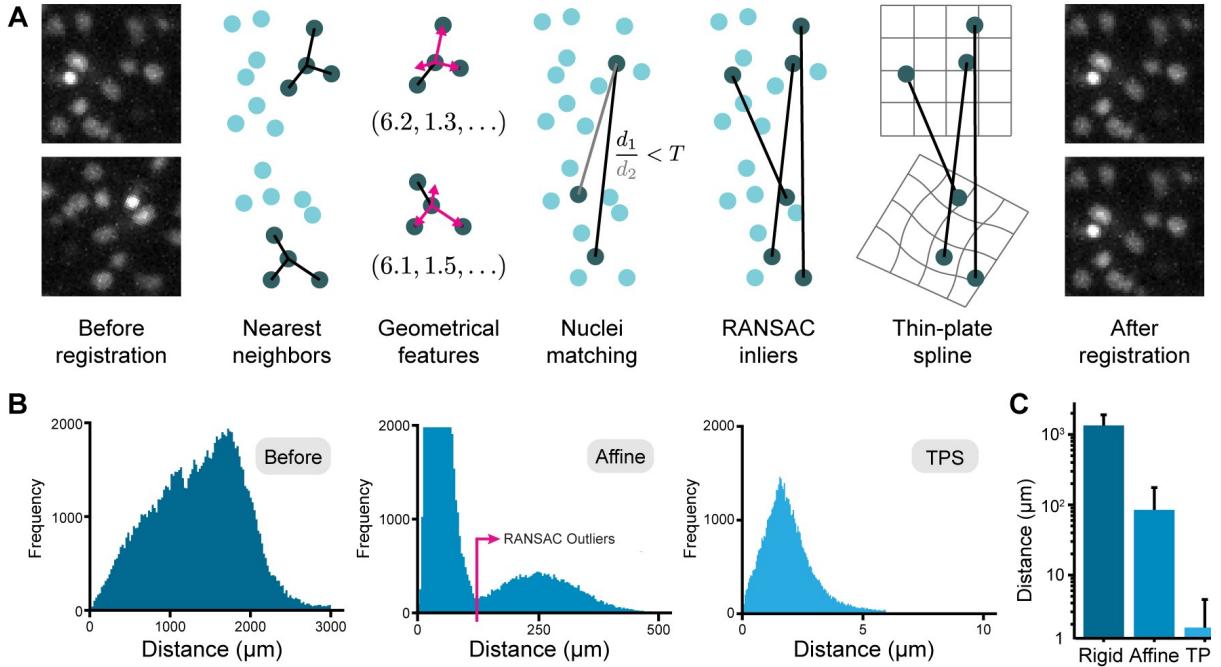
**Figure 5.2:** Coarse registration based on morphological alignment. Coarse registration using a rigid transformation to align the euclidean distance transform (EDT) of brain masks. Performing coarse registration prior to nuclei matching allows for more robust nuclei matching and faster execution times. The EDT provides a gradient signal for intensity-based alignment by minimizing the squared error between the fixed and moving images.

### 5.3.3 Matching nuclei based on geometric features

After estimation of the a rigid transformation for coarse alignment, we detected all nuclei in the mouse brain hemisphere using a blob detection strategy. We first performed gaussian blurring to remove noise from the nuclear stain images, and then we detected all local maxima in the smoothed images. This nuclei detection strategy was applied to both rounds of imaging running in parallel on each image chunk. After extracting nuclei centroids from each image, the single-cell coregistration is reduced to a massive point cloud registration problem. By applying the rigid transformation to the nuclei in the fixed image, the nuclei can be brought into approximate alignment.

In order to match approximately aligned nuclei, each nucleus is described by a translation and rotation invariant feature vector that is constructed using geometrical hashing of the nuclei point cloud (Preibisch et al. 2010). Briefly, each point cloud is represented as a KD-tree and stored in shared memory for access by multiple CPU workers. Each worker queries the KD-tree for the three nearest nuclei in the same image and computes the vectors between the query point and its neighbors (Figure 5.3a). These vectors are used in a Gram-Schmidt process to construct an orthonormal basis that is defined based on the local arrangement of nuclei and is, therefore, invariant to sample rotations and translations. The coordinates of the three nearest nuclei in this new basis are computed using QR decomposition. The nonzero entries of the upper triangular matrix  $R$  are flattened into a final feature vector describing each nucleus.

Given these geometrical features, each nucleus is then putatively matched to the nearest neighbor in feature-space within a small search radius in the coarsely-aligned point cloud. These putative matches are only considered correct if the second nearest-neighbor is clearly not a match as judged by the distance in feature-space (Lowe, D. 2004). After matching nuclei, an affine alignment is estimated using RANSAC that serves as the baseline transformation for the non-rigid coregistration (Figure 5.3b). That is, the overall deformation model is the sum of an affine transformation and a non-linear TPS transformation. The TPS transformation is estimated using the remaining correspondences that are not rejected by RANSAC. Specifically, three separate TPS interpolators are fit to the x, y, and z displacements of the affine-transformed correspondences. The average displacement between correspondences drops dramatically from the affine transformation to the TPS transformation (Figure 5.3c). This shows that there is significant non-linear deformations present between imaging rounds.

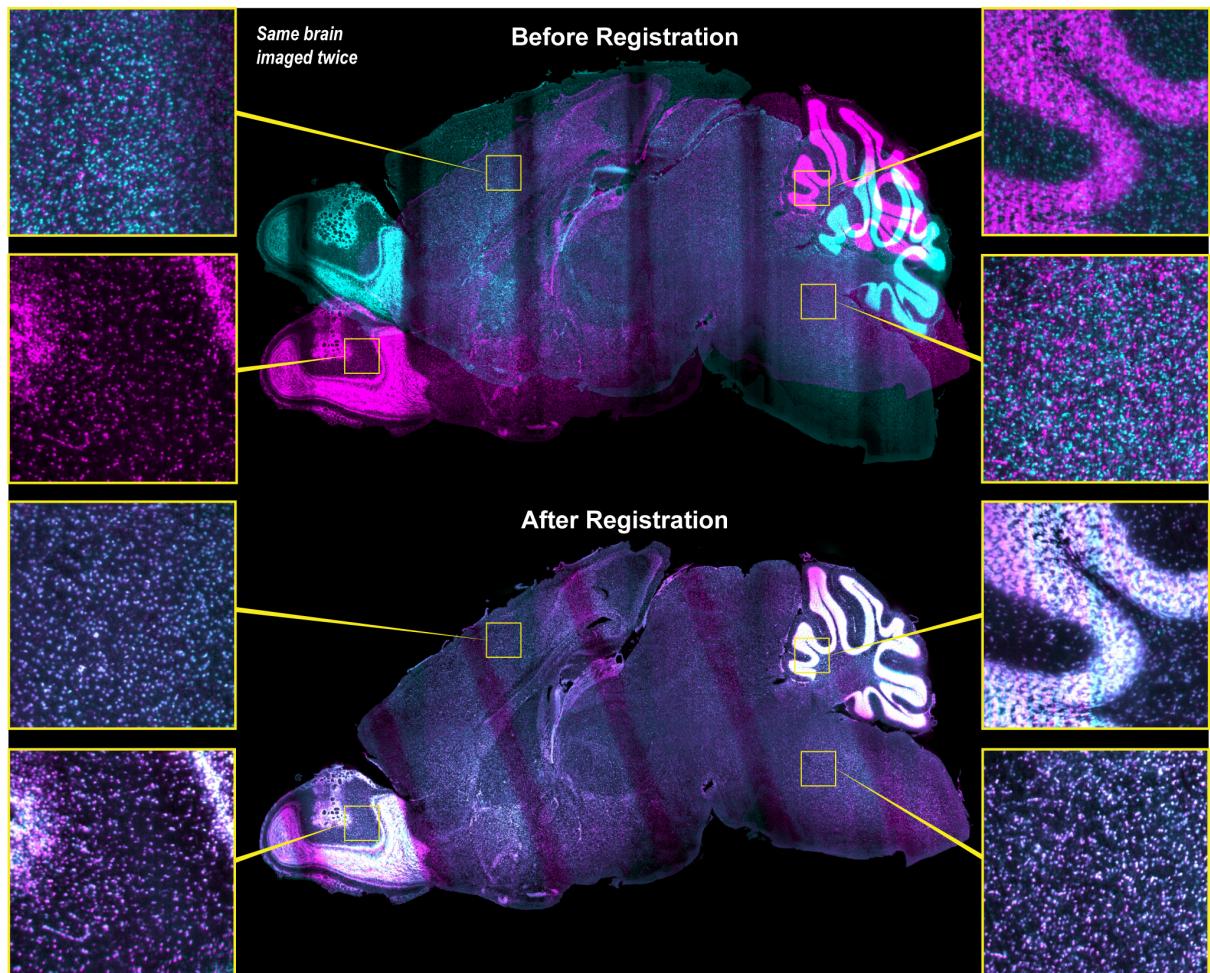


**Figure 5.3:** Nuclei matching and thin-plane spline warping. (A) Illustration of the nuclei-based 3D nonrigid registration procedure. Neighboring nuclei centroids are used to construct local nucleus features that are invariant to translation and rotation. Nuclei are matched to their nearest neighbor if the second nearest neighbor is sufficiently far away in feature space. Nuclei coordinates are warped using a thin-plane spline (TPS) interpolator fit to matching nuclei coordinates. (B, C) Distribution of distances and average distances between matching nuclei before and after affine and nonrigid registration.

### 5.3.4 Single-cell 3D coregistration of intact mouse brain hemisphere

Since the full TPS transformation is computationally intractable when applied to the full resolution image data, we instead approximate the TPS transformation by linear interpolation of a warped grid of control points. The main motivation for this is that there are readily available linear interpolation functions available with GPU-acceleration in packages such as Pytorch (Paszke et al. 2017). By warping a regular grid of 100 x 100 x 100 control points evenly spaced throughout the fixed image domain using the full TPS transformation, the TPS transformation can be approximated by linear interpolation of the result. By transforming the fixed image domain onto the moving image, each voxel in the fixed image can be filled in by interpolation of the moving image.

We successfully implemented this approximate TPS transformation using Pytorch. The coregistration process takes approximately 6 hours on an NVIDIA Tesla P100 GPU for a whole mouse brain hemisphere (Figure 5.4). The resulting alignment shows that corresponding nuclei overlap across the whole hemisphere, including regions with densely-packed nuclei. From a random sample of 500 nuclei throughout the brain, our nuclei matching was 97% accurate, resulting in an average distance of 4  $\mu\text{m}$  between matches after alignment. Although the images presented in Figure 5.4 do not make it apparent,



**Figure 5.4:** Demonstration of 3D coregistration of a whole mouse hemisphere at single-cell resolution. Multiple regions are highlighted to show single-cell accuracy of the alignment globally. Regions with densely-packed nuclei such as the olfactory bulb and cerebellum show corresponding nuclei overlapping. Scale bar, 2 mm.

we emphasize that this coregistration is performed 3D space.

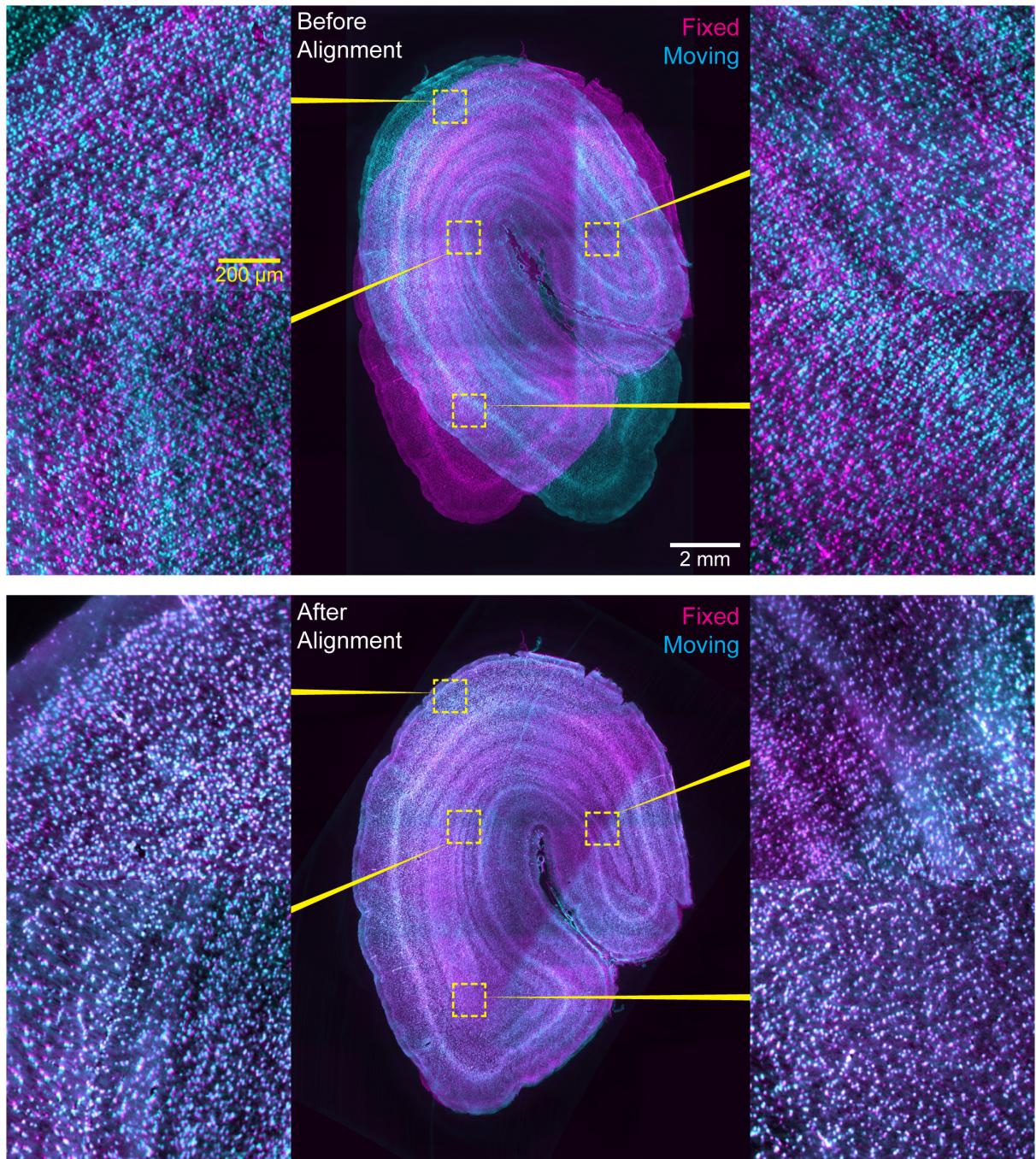
### 5.3.5 Single-cell 3D coregistration of large marmoset brain block

In order to test how robust our coregistration algorithm was against the de-staining and re-staining process as well as in other tissue types, we performed multiround staining and imaging of a large tissue block from the visual cortex of a marmoset. A large block of marmoset brain tissue was SHIELD-processed and stained with Syto16 nuclear dye using eFLASH. After LSFM imaging, the marmoset tissue sample was de-stained using SDS detergent and heat treatment at 56C. The same tissue sample was then re-stained with Syto16 using eFLASH and imaged using LSFM in a second round.

Using a similar strategy to the mouse brain hemisphere for coarse alignment, nuclei detection and matching, and GPU-accelerated warping, we were able to coregister the two rounds of staining and imaging of the marmoset brain sample (Figure 5.5). Before alignment, the tissue is not aligned between imaging rounds. However, the Syto16 staining quality appeared similar between rounds with the exception of some differences in illumination as mentioned previously. The morpholoical coarse alignment was able to approximately align the two rounds despite these illumination differences. After alignment, the quality of alignment across the entire tissue sample is shown by overlapping staining of corresponding nuclei. This demonstration shows that our nuclei-based 3D coregistration is quite general, and it is not limited to mouse brain hemispheres.

## 5.4 Discussion

To the best of our knowledge, these results mark the first demonstration of 3D coregistration of whole-brain scale images at single-cell resolution. Although a proof-of-concept, our results are promising given the resulting alignment quality, speed of execution, and general applicability to multiple biological models such as mouse and marmoset. However, more research is required to show the limitations in the number of rounds that can be combined beyond just two rounds of staining and imaging. Furthermore, the actual value of such 3D coregistration is in aligning other channels than just the nuclear staining channel. We do not anticipate that aligning other channels using the same transformation estimated from the nuclear staining channel would be a problem given the efficiency of our GPU-accelerated image warping. However, showing a single volumetric image with dozens of protein channels remains an impressive feat to be demonstrated in future research.



**Figure 5.5:** Demonstration of 3D coregistration of a large block of marmoset visual cortex at single-cell resolution. After coarse registration based on morphological alignment, nuclei were successfully matched in marmoset visual cortex.

Several improvements to our current approach for single-cell 3D coregistration are possible. For example, the nuclei matching and TPS warping estimation only occur once, but the non-linear TPS warping would change the geometric features that the nuclei matching is based on. Given that nuclei are much closer together than the spacing of our regular grid, we are implicitly assuming that the tissue deformations are locally linear. The success of our coregistration algorithm suggests that this assumption is acceptable for slight tissue deformations incurred during de-staining, re-staining, and manual handling. However, it is unclear if reducing the search radius after a first-pass of nuclei matching and TPS warping could increase the number of correspondences and the overall quality of alignment by an appreciable amount. Another potential improvement would be to allow for semi-automatic coarse alignment with the option for manual correction. For example, the Nuggt tool, which draws on Neuroglancer, would be suitable for manual refinement of the morphological coarse alignment described here (Swaney et al. 2019). Integrating commonly used intensity-based alignment tools such as Elastix into this workflow could make the overall coregistration pipeline more robust (Klein et al. 2010).

Once obtained, a high-dimensional volumetric image of multiple proteins within a single brain sample could be used for more unbiased and exploratory analyses and hypothesis generation. Rather than selecting a brain region and a handful of markers to study in a given experiment, neuroscience researchers may be able to visualize and quantify the expression of many proteins within their samples in a more complete brain-wide manner. This powerful approach would be valuable for studying complex neurological phenotypes which are poorly understood. Multiplexed imaging at the whole-brain scale will also enable spatial analyses that have never before been possible. Since each staining round comes from the same tissue, intricate spatial analyses will be possible at the single-cell level rather than just the region level.

## 5.5 Methods

### 5.5.1 Tissue processing, staining, and imaging

Mouse and marmoset brain samples were prepared according to methods described in the SHIELD and eFLASH papers (Park et al. 2019; Yun et al. 2019). The SHIELD tissue preservation technique was used to fix and clear brain tissue samples, and the eFLASH labeling technique was used to label tissue samples with Syto16 nuclear dye. Stained samples were then incubated in refractive index matching solution (PROTOS), which has been described previously (Yun et al. 2019).

After refractive index matching, stained samples were imaged using a axially-swept light-sheet fluorescence microscope (LifeCanvas SmartSPIM). All tissue samples were imaged using the custom 4x/0.3 NA long working-distance objective (LifeCanvas). The images from the LSFM system were saved in raw binary format.

### 5.5.2 Image preprocessing

Raw images were first destriped and stitched according to previously published methods (Swaney et al. 2019). The stitched images were contrast enhanced using CLAHE from the scikit-image Python package on each optical section before converting to Zarr format. The chunked Zarr format was used to construct a multiscale representation by generating image pyramids down to 32x isotropic downsampling.

### 5.5.3 Coarse morphological alignment

Preprocessed images were downsampled 16x isotropically and then thresholded below the nuclear stain auto-fluorescence to obtain a foreground binary mask. The EDT was obtained using the *distance\_transform\_edt* function in the Scipy package. An objective function for the L-BFGS optimizer was constructed given the parameters to a rigid transformation of the fixed image onto the moving image domain. The fixed image was transformed onto the moving image since the final warping process involves interpolating the moving image data. L-BFGS optimizer was iteratively attempted using the *basinhopping* global optimization function in the Scipy package. The resulting rigid transformation parameters were converted to the full image resolution by taking the downsampling factor into account for the translation.

### 5.5.4 Nuclei detection and geometric hashing

Nuclei were detected in each image by first gaussian blurring 3D image chunks to remove noise. these image chunks were “ghosted” (including some voxels from neighboring chunks) to avoid edge effected in applying the gaussian kernel. After blurring, the *peak\_local\_max* function from scikit-image was used to detect nuclei centroids.

After generating nuclei point clouds, a KD-tree was built for efficient querying of nearest neighbors using the NearestNeighbor object in the scikit-learn package. Geometric features for each nucleus were computed in parallel by querying the 3 nearest nuclei in the point cloud. These neighbors were used to construct and orthonormal basis through using the Gram-Schmidt process considering the 3rd nearest neighbor as the z axis. QR

decomposition was used to compute the coordinates of the nearest nuclei in this new basis, and the upper triangular portion of the matrix  $R$  was flattened into a feature vector containing 6 values describing the relative arrangement of nuclei around each nucleus.

### 5.5.5 Nuclei matching and RANSAC affine estimation

The nuclei point cloud from the fixed image was first transformed into approximate alignment with the nuclei point cloud from the moving image using the rigid transformation from the coarse morphological registration. The KD-tree used to compute geometric features for the nuclei in the moving image was reused to query all nuclei that fall within a 150  $\mu\text{m}$  search radius. A candidate nucleus in this search radius was considered a match if the euclidean distance between the geometric features was less than a “similarity” threshold and the ratio between that distance and the next nearest neighbor in feature space was less than a “prominence” threshold,  $T$ .

After matching nuclei between fixed and moving images, an affine transformation was estimated using the *RANSACRegressor* class in the scikit-learn package by providing an affine transformation as the base model. The resulting affine transformation was used as the basis for computing non-linear deformations from, and any outlier identified by RANSAC were omitted from the list of corresponding nuclei.

### 5.5.6 GPU-accelerated warping of whole-brain volumetric images

A regular grid of 100 x 100 x 100 points was constructed over the fixed image domain, and all points were transformed through the full affine-TPS transformation. The resulting set of coordinates were linearly interpolated using the GPU-accelerated *grid\_sample* function from the Pytorch package. After interpolating the exact affine-TPS displacements, a bounding box for the projection of each chunk in the fixed images on the moving image was defined and used to query the moving image data for only those voxel needed for interpolation. Finally, the moving image data was interpolated at the coordinates obtained from displacements.

## 5.6 References

- Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), pp.567–585. Available at: <http://ieeexplore.ieee.org/document/24792/>.
- Edsgård, D., Johnsson, P. & Sandberg, R., 2018. Identification of spatial expression trends in single-cell gene expression data. *Nature Methods*, 15(5), pp.339–342. Available at: <http://www.nature.com/articles/nmeth.4634>.
- Hwang, B., Lee, J.H. & Bang, D., 2018. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental & Molecular Medicine*, 50(8), p.96. Available at: <http://www.nature.com/articles/s12276-018-0071-8>.
- Klein, S. et al., 2010. Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1), pp.196–205.
- Ku, T. et al., 2016. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nature biotechnology*, 34(9), pp.973–81.
- Lein, E.S. et al., 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 445(7124), pp.168–176.
- Lowe, D.G., 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), pp.91–110. Available at: <http://link.springer.com/10.1023/B:VISI.0000029664.99615.94>.
- Murray, E. et al., 2015. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6), pp.1500–14.
- Panchuk-Voloshina, N. et al., 1999. Alexa Dyes, a Series of New Fluorescent Dyes that Yield Exceptionally Bright, Photostable Conjugates. *Journal of Histochemistry & Cytochemistry*, 47(9), pp.1179–1188. Available at: <http://journals.sagepub.com/doi/10.1177/002215549904700910>.
- Paszke, A. et al., 2017. Automatic differentiation in PyTorch. *NIPS*, (Nips), pp.1–4.
- Power, R.M. & Huisken, J., 2017. A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nature Methods*, 14(4), pp.360–373. Available at: <http://arxiv.org/abs/nmeth.4224>.
- Preibisch, S. et al., 2010. Software for bead-based registration of selective plane illumination microscopy data. *Nature Methods*, 7(6), pp.418–419. Available at: <http://>

//www.nature.com/articles/nmeth0610-418.

Swaney, J. et al., 2019. Scalable image processing techniques for quantitative analysis of volumetric biological images from light-sheet microscopy. *bioRxiv*.

Yun, D.H. et al., 2019. Ultrafast immunostaining of organ-scale tissues for scalable proteomic phenotyping. *bioRxiv*.

# Chapter 6

## 3D Imaging and Multiscale Analysis of Intact Human Cerebral Organoids

Alexandre Albanese\*, **Justin Swaney\***, Dae Hee Yun, Nick Evans, Jenna Antonucci-Johnson, Vincent Pham, Chloe Delepine, Mriganka Sur, Lee Gehrke, Kwanghun Chung<sup>1</sup>

### 6.1 Abstract

Cerebral organoids grown from patient-derived stem cells self-pattern into cytoarchitectures that mimic the developing human cortex. These 3D culture models offer an unprecedented chance to study human cortical brain development as neural progenitors surrounding ventricles mature into cortical neurons and glial cells. Analysis of cells and cytoarchitecture currently requires dissociation or sectioning and the loss of spatial information can bias data interpretation. Here, we present a pipeline for volumetric fluorescence imaging and multiscale analysis of intact cerebral organoids. Our integrated technology platform can rapidly clear, label, and image a large number of intact organoids. The SCOUT pipeline (Single-cell and Cytoarchitectural analysis of Organoids using Unbiased Techniques) applies automated algorithms and neural networks to extract hundreds of features characterizing molecular, cellular, contextual, cytoarchitectural, and morphological features from three dimensional image datasets in a quantitative and unbiased manner. Organoid processing from cell culture to microscope takes less than ten days and imaging with light-sheet fluorescent microscopy requires approximately 15 minutes per sample. SCOUT analysis pipeline automates single-cell segmentation at 90% accuracy, ventricle segmentation at 97% accuracy and utilizes spatial information for unprecedented data analysis. Segmented ventricles were used as the seed point for automated

---

<sup>1</sup>\* indicates co-first authorship.

analysis of radial cell patterning. High-dimensional analysis of 42 organoids, 4591 ventricles and over 95 million cells provided a quantitative overview of the multiscale changes associated with maturation, Zika virus infection, and different culture protocols. SCOUT quantified significant shifts in organoid size, ventricle volume, cell populations and cytoarchitecture clusters, providing the first detailed list of tissue-wide phenotypic changes. SCOUT represents the first high-throughput unbiased quantification of 3D tissue models and establishes the necessary groundwork for multiscale morphological phenotyping. This technological breakthrough improves analysis of complex three dimensional culture systems and provides more accurate models of biology and disease.

## 6.2 Introduction

Embryonic and inducible pluripotent stem cells can grow into virtually any cell type from the body. Two-dimensional stem cell differentiation protocols have enabled the study human development, biology and disease. Recently, organoids have emerged from the recent development of three-dimensional differentiation protocols that drive the self-patterning of differentiated cells into cytoarchitectures resembling tissue subregions. Recently, several protocols have achieved the production of intestinal (Sato et al. 2009; Gjorevski et al. 2016), retinal (Eiraku et al. 2011), kidney (Takasato et al. 2015; Homan et al. 2019), vasculature (Wimmer et al. 2019) and brain organoids (Kadoshima et al. 2013; Lancaster et al. 2013; Yoon et al. 2019; Velasco et al. 2019) amongst others. Each organoid contains a mixture of cell types, intercellular interactions and extracellular microenvironments that can replicate a tissue's natural environment. The structural complexity of these organoids provides novel in vitro models to study the intricate physiological and pathological phenomena not replicated in conventional two-dimension cell culture models.

Cerebral organoids allow us to study early human neurodevelopment from neuroepithelium formation, neural migration and formation of rudimentary networks. Cerebral organoids mimic cortical development by forming ventricles lined with radial glia progenitors that differentiate into cortical neurons. (Renner et al. 2017; Qian et al. 2016) Cerebral organoids recapitulate genetic and epigenetic features of prenatal human brain development (Camp et al. 2015; Luo et al. 2016; Quadrato et al. 2017) and have been used to identify cytoarchitectural changes brought on by microcephaly (Lancaster et al. 2013), lissencephaly (Bershteyn et al. 2017), Rett syndrome (Mellios et al. 2018), Autism (Mariani et al. 2015), and Zika virus infection (Watanabe et al. 2017). These studies are important for characterizing the impact of disease on early cortical development. However, most published work has relied on the manual analysis of histological tissue sections to characterize cytoarchitecture and cell populations. The loss of spatial

information is problematic considering the unique features of individual organoids and can lead to biased data interpretation.

Cerebral organoid heterogeneity produces a quantifiable variability at both inter- and intra-batch levels (Quadrato et al. 2017). Many groups have sought to remedy cerebral organoid heterogeneity through optimization of differentiation protocols (Yoon et al. 2019; Velasco et al. 2019; Qian et al. 2016). However, a potentially unresolvable source of heterogeneity is the independent asynchronous development of multiple neuroepithelial units (neural rosettes, ventricles) each producing their own cells and morphogenic gradients (Knight et al. 2018). The random configuration of ventricles in each organoid does not recreate the embryo’s physiological conditions where a single neuroepithelial tube matures into the entire central nervous system. The organoid’s random configuration of ventricles also makes it especially difficult to interpret local cell populations and cytoarchitectures with incomplete knowledge of spatial context. Unbiased and quantitative organoid phenotyping requires new technologies to achieve whole-tissue analysis capable of combining antibody staining and multiscale contextual analysis of single cells.

Here, we present a novel pipeline for multiscale analysis of intact cerebral organoids. We adapted the SHIELD protocol for whole-tissue clearing (Park et al. 2018) and eFLASH technology for accelerated antibody staining (Yun et al. 2019; Kim et al. 2015; Murray et al. 2015) to generate optically-transparent fluorescently-labeled whole-organoids. We then developed an analytical pipeline for Single-cell and Cytoarchitecture analysis of Organoids using Unbiased Techniques (SCOUT) for quantitative analysis of single cells, ventricles, and cytoarchitectures. Holistic organoid analysis ensures unbiased quantification of these heterogeneous tissues without any loss of spatial context or rare events. SCOUT enables quantification of previously inaccessible features and decreases the variance of conventional features prone to histological sampling biases. For comparative studies, we developed a framework for unbiased hyper-dimensional analysis based on 276 features describing the volumetric data of individual organoids. These features provide a framework to assess the correlation of features at different size scales and provide a quantitative overview of phenotypic changes.

## 6.3 Results

### 6.3.1 Whole-organoid staining

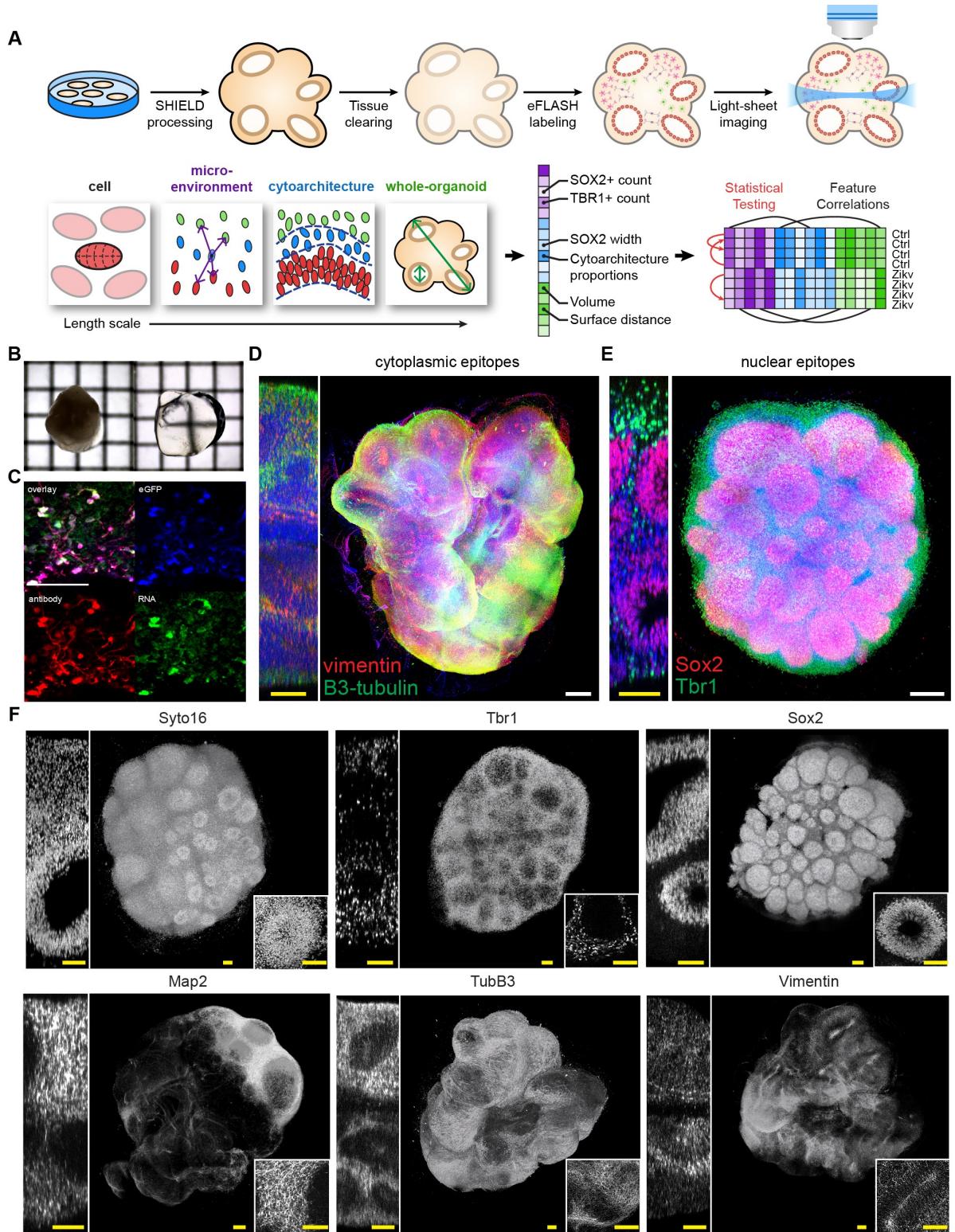
To achieve whole organoid antibody staining and imaging (Figure 6.1a), we adapted the SHIELD (Park et al. 2018) poly-epoxide tissue reinforcement protocol to preserve

biomolecules during detergent and heat-based de-lipidation. Multiple epoxide groups in the polyglycerol 3-polyglycidyl ether molecule (P3PE) react with available amines in the tissue to form stabilizing inter-and intra-molecular covalent bonds. We adapted the original SHIELD protocol to prevent over-fixation in our samples and maintained a near-neutral pH during de-lipidation to ensure bright antibody staining of transcription factors (Figures 8.10, 8.11). After immersion in dPROTOS optical clearing solution, cerebral organoids were optically transparent (Figure 6.1b). Microscopy of SHIELD-cleared organoid tissue sections confirm the preservation of eGFP endogenous fluorescence, epitopes, and mRNA (Figure 6.1c). For whole-organoid staining, we employed eFLASH technology where a rotating electric field accelerates transport of fluorescent probes into the tissue for homogenous antibody staining. The eFLASH protocol was adapted for cerebral organoids to achieve simultaneous staining of 8-10 organoids in 35h using 3-15 µg of primary antibody (see Methods). We achieved whole-organoid staining of nuclear and cytoplasmic proteins (Figure 6.1d-f). Radial glia in whole d35 organoids were stained with vimentin, lipid binding protein, nestin, and SOX2. Neurons were stained with 3-tubulin, MAP2, and TBR1.

As an initial proof-of-concept, we developed our SCOUT pipeline using a nuclear dye and two antibodies imaged with light-sheet fluorescent microscopy (LSFM) for high-throughput organoid phenotyping. We used anti-SOX2 to label radial glial progenitors and anti-TBR1 to label the early post-mitotic neurons in organoids at day 35. However, analysis is adaptable to other combination of experiment-specific markers. We chose to label transcription factors to ensure the straightforward identification of dye-labeled nuclei with antibody staining to identify progenitor and neuronal cells. The position of SOX2 and TBR1 cells relative to the ventricle enables the identification of basic cytoarchitectural patterning in cerebral organoids. Based on previous work, these markers enable identification of a SOX2+ ventricular zone and an early preplate-like layer of TBR1+ neurons (Qian et al. 2016; Watanabe et al. 2017).

### 6.3.2 SCOUT analysis of single cells

Using LSFM with a 10X objective (see Methods), three-channel data acquisition at 0.6 x 0.6 x 2.0µm voxel size took on average ~15 min per organoid. The 3D datasets enabled single-cell nucleus identification and segmentation throughout the organoid (Figure 6.2a). We applied curvature-based seeded watershed on nuclear dye images (Atta-Fosu et al. 2016) and achieved an accuracy of 90% (Figure 8.12). Next, we co-localized segmented nuclei with SOX2 and TBR1 channels for molecular phenotyping of individual cells (Figure 6.2b). Since expression of these two markers is mutually exclusive, we identified three populations in individual organoids: SOX2+ progenitors, TBR1+ neurons, and double



**Figure 6.1:** Pipeline for intact organoid analysis (A) Cerebral organoids are grown from stem cells, fixed in 4% PFA, then post-fixed with SHIELD poly-epoxide crosslinker. Organoids are then cleared in SDS solution at 55°C, rinsed, then labelled with antibodies using eFLASH4 for rapid uniform staining. Samples are imaged by light-sheet microscopy for high-throughput (15 min per organoid) quantitative microscopy. (B) Quantitative analysis applies automated algorithms and deep learning to measure multiscale features. We applied this pipeline for unbiased high-dimensional phenotyping of different experimental models. (C) After de-lipidation and immersion in refractive index matching solution (dPROTOS), organoids are optically transparent and can be imaged with standard confocal microscopy. Grid = 1 mm (D) SHIELD preserves endogenous fluorescence, mRNA and protein epitopes.

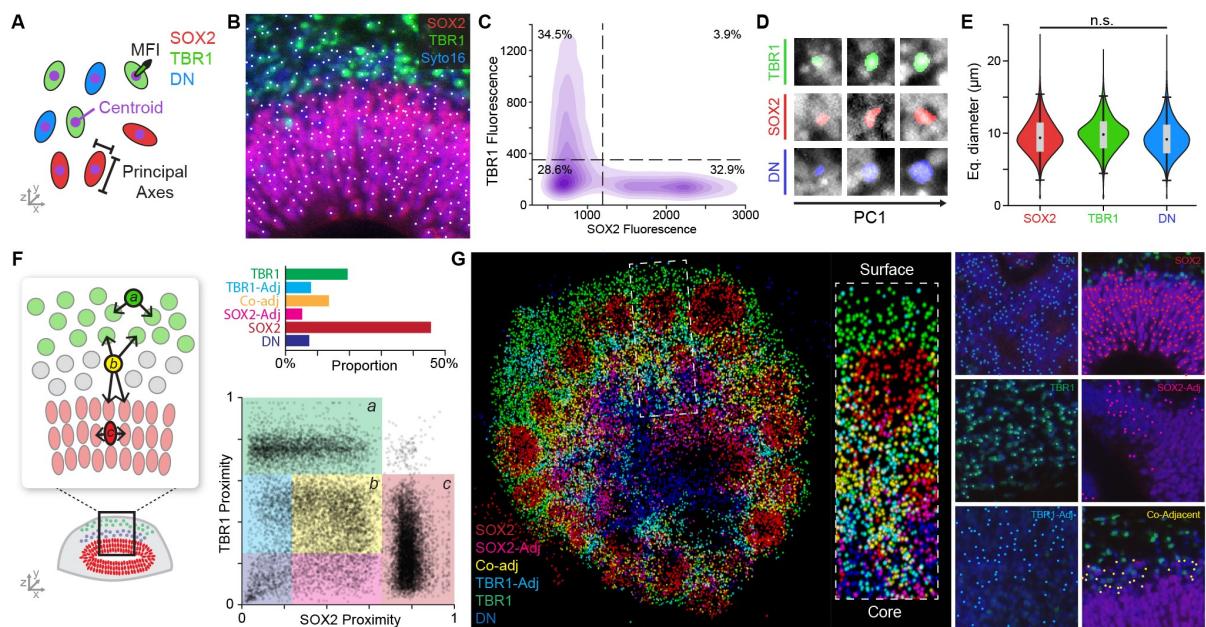
(E) 3D renders of day 35 cerebral organoid stained using cytoplasmic or nuclear epitopes. Left: blue is Syto16 nuclear dye, green is 3-tubulin (neuronal marker) and red is vimentin (glial marker). Right: blue is Syto16, green is Tbr1 (deep layer neurons) and red is Sox2 (progenitor cells /radial glia) (F) Day 35 whole-organoid staining of different antibodies. Scale bars (yellow: 100  $\mu$ m, white: 200  $\mu$ m)

negative (DN) cells. Morphological analysis of individual nuclei shows a  $\sim$ 10 $\mu$ m average diameter for all three cell types, although we detect instances of different volumes for each population (Figure 6.2c,d).

Our datasets include the position and marker expression for every cell inside these highly patterned cerebral organoids (Lancaster et al. 2013; Qian et al. 2016; Watanabe et al. 2017). This allows us to examine the spatial context of individual cells. To explore the value of a cell's spatial context, we combined antibody-based molecular phenotyping data with a “positional proximity score” based on each cell's distance to its nearest SOX2+ and TBR1+ cells. We designed this score to reflect each cell's proximity to the SOX2-rich ventricular zones and neuron-rich TBR1+ preplate-like regions (Figure 6.2f,g). Spatial context quantification reveals that SOX2+ and TBR1+ cells had the highest proximity scores ( $>0.65$ ) to themselves. In other words, SOX2-adjacent cells are  $>99.6\%$  SOX2+ and TBR1-adjacent cells are  $>98.6\%$  TBR1+ cells. The lowest proximity scores ( $<0.15$ ) occurred for DN-adjacent DN cells at the core of the organoid away from the ventricles. Proximity analysis detected three intermediate populations (0.15-0.65): SOX2-adjacent DN cells in the ventricular zone, TBR1-adjacent DN in neuron-rich regions and SOX2/TBR1 co-adjacent DN cells. We hypothesized that the location of these DN subpopulations likely correlates with their identity given that a TBR1-adjacent DN cell is most likely a post-mitotic TBR1- neurons and the infrequent SOX2-adjacent DN cells may represent an early population of intermediate progenitors migrating to form a sub-ventricular zone. This analysis demonstrates the combination of molecular phenotyping and spatial context enables cellular subcategorization in highly patterned tissues.

### 6.3.3 SCOUT analysis of ventricles and radial organization

Next, we sought to characterize ventricles and the radial cell patterning they produced (Figure 6.3). First, we identified the total number and general morphology of ventricles inside our organoids. We adapted U-Net (Ronneberger et al. 2015), a convolutional neural network, to detect SOX2-lined ventricles based on 9 manually segmented (Yushkevich et al. 2006) datasets containing 7596 nuclear dye images (Figure 8.13). We performed manual segmentation using the SOX2 antibody channel, but trained the convolutional neural network on nuclear dye images to potentially eliminate the necessity of Sox2-antibody in future studies. Automated ventricle segmentation by U-Net achieved a Dice



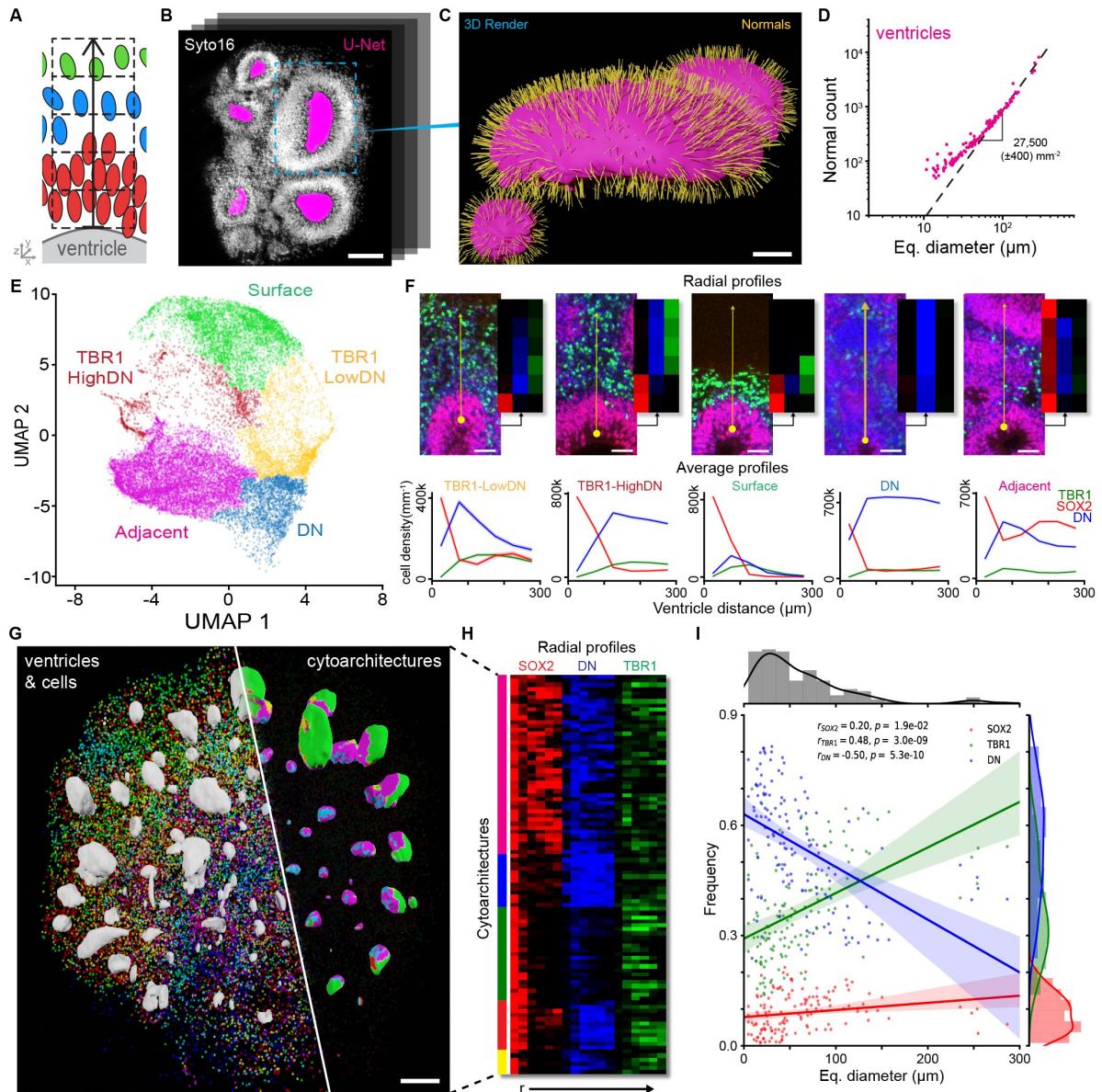
**Figure 6.2:** Single cell detection and analysis (A) Scheme of single cell morphological characterization. (B) Demonstration of automated nuclei detection in 3D datasets. Syto16 in blue, Sox2 in red, Tbr1 in green. Scale bar = 100  $\mu\text{m}$ . (C) Expression of Tbr1 and Sox2 in individual nuclei. Dot plot shows gating parameters and population frequency. (D) Representative images of segmented nuclei for each cell type showing a range of morphological features sorted by principal component analysis. (E) Morphological analysis of individual cells shows a consistent equivalent diameter  $\sim 10 \mu\text{m}$  for all cell types. (F) Scheme of contextual analysis showing the proximity of each cell to the nearest SOX2 and TBR1 cells. Bar graph shows the results of contextual analysis shown below with “context gates” to define six distinct populations. Gates captured  $>99\%$  of all cells. (G) Middle optical section of a day 35 organoid dataset showing detected cells colored according their spatial context-based subcategorization. Inset shows zoomed view of dotted rectangle region. Right subpanels show instances of six different subpopulations identified by SCOUT.

coefficient of 97.2% on a 20% hold-out test set and provided morphological analysis of the three dimensional ventricles (Figure 6.3b, Figure 8.13).

Ventricles are an important anatomical landmark in the heterogeneous landscape of a cerebral organoid. They provide the origin of radial cytoarchitecture orientation that occurs when progenitor glia cells produce post-mitotic neurons. We quantified radial organization of cell populations around ventricles by generating “virtual biopsies” 50 $\mu$ m in diameter and 300 $\mu$ m long perpendicular to the ventricle’s surface (Figure 6.3a-c). Each biopsy captures SOX2, TBR1, and DN cell counts within 50  $\mu$ m-tall stacked cylindrical regions. We generated thousands of virtual biopsies uniformly distributed across the surface of all ventricles for unbiased quantification of radial cytoarchitectures in the organoid. As expected, the number of biopsies generated per ventricle was proportional to its surface area (Figure 6.3d). Performing unsupervised hierarchical clustering after UMAP embedding to reduce hyper-dimensionality (McInnes et al. 2018) revealed five distinct cytoarchitectures: TBR1+DN-low, TBR1+DN-high, surface, “DN only”, and “adjacent ventricles” (Figure 6.3e,f). When we mapped cytoarchitecture clusters onto the surface of the ventricles, we observed a strong positional correlation within the organoid (Figure 6.3g). As expected, surface cytoarchitecture appears at the surface organoid due to the abrupt disappearance of cells beyond the organoid’s border. We also detected “adjacent ventricles” cytoarchitecture on ventricle surfaces facing towards nearby ventricles. A survey of a hundred random biopsies shows a spectrum of different cell profiles and we explored whether ventricle size may affect cell frequency (Figure 6.3h, i). Quantification of cell type frequency in each ventricle’s virtual biopsies reveals a strong correlation between ventricle size and cell type frequency. Correlation analysis revealed that ventricle size was positively correlated with TBR1 cell frequency and negatively correlated with DN frequency in our organoids. These observations suggest that larger ventricles may be farther along in their development program or at positions within the organoid that replicate physiological conditions more accurately. Given the heterogeneity of organoids, conventional quantification of the “average cytoarchitecture” based on tissue sections is prone to noise and misinterpretation. SCOUT, on the other hand, provides independent analysis of multiple cytoarchitectures and ventricles to provide unbiased full-spectrum analysis of cell patterning.

#### 6.3.4 SCOUT whole-organoid analysis and correlation

SCOUT provides the ability to measure absolute frequency of cell populations in individual organoids. We could therefore quantify the accuracy of conventional histological analysis based on two-dimensional tissue sections. First, we compared how three different 100 $\mu$ m thick “pseudo-sections” generated from our 3D datasets of day 35 organoids

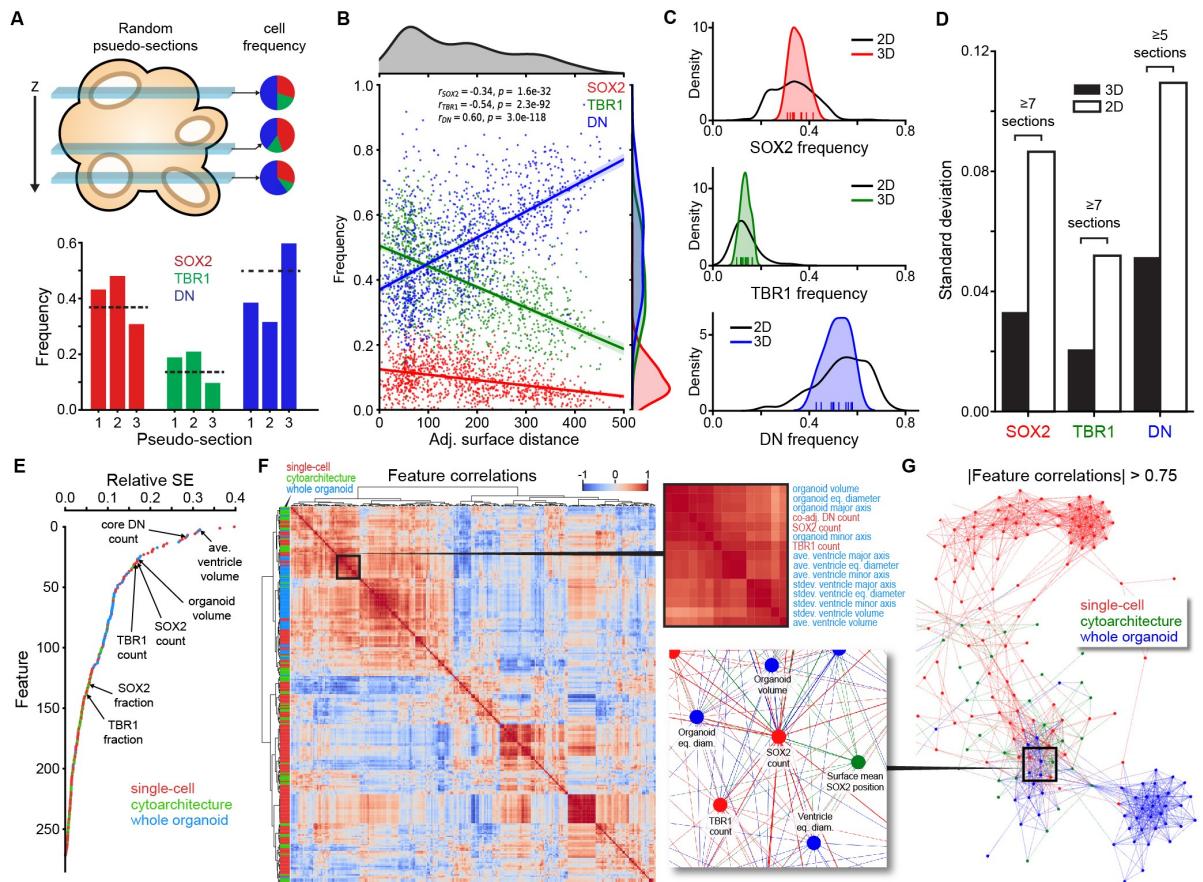


**Figure 6.3:** Ventricle segmentation and cytoarchitecture phenotyping. (A) Scheme of automated cytoarchitecture analysis. We quantified radial organization of cell populations around ventricles using “virtual biopsies” 50  $\mu\text{m}$  in diameter and 300  $\mu\text{m}$  high perpendicular to the ventricle surface. (B) Demonstration of automated ventricle segmentation using U-Net convolutional neural network. Representative optical section of a volumetric dataset with detected ventricles in magenta. (C) 3D render of ventricle highlighted in panel B with normal used to orient virtual biopsies shown in yellow. (D) Graph showing that the total number of normal per ventricle depends on the ventricle’s surface area. (E) UMAP embedding of detected cytoarchitectures in a single organoid color-coded according to each cluster. (F) Representative image and average profile plot of individual cytoarchitecture clusters showing the radial distribution of SOX2 (red), double negatives (blue) and TBR1 (green) cells. (G) 3D render of segmented cells and ventricles from a day 35 organoid. On the left side ventricles are white and six cell populations are colored according to the index in Figure 2G: SOX2 in red, SOX2-adjacent in magenta, co-adjacent in yellow, TBR1-adjacent in cyan, TBR1 in green and core DN in blue. On the right, we mapped the detected cytoarchitectures on the surface of rendered ventricles using the colors in panel F. (H) Three-channel heat map from 100 random cytoarchitectures. Each row shows the number of cells detected in all six 50  $\mu\text{m}$  increments moving away from the ventricle surface. Intensity of red, blue and green represent SOX2, DN and TBR1, respectively. (I) The frequency of SOX2, DN and TBR1 cells detected in a ventricle’s virtual biopsies correlates with the ventricle equivalent diameter. Strongest correlation occurs for increased DN and increased TBR1 in larger ventricles.

predicted whole-tissue cell type frequencies (Figure 6.4a). These results show that individual slices consistently miss the SOX2, TBR1 and DN whole-organoid “true frequency” and reveal an intra-organoid heterogeneity biased by the pseudo-section’s position within the organoid. To address the possibility of positional bias within individual organoids, we quantified the correlation between cell type frequencies for each ventricle and their distance from the surface (Figure 6.4b). This analysis reveals a significant positional bias for each cell type. TBR1 cells are predominantly at the surface of the organoid and decrease with depth. Inversely, DN cell frequency increases as a ventricle moves away from the surface increases and reaches ~80% for ventricles near the organoid’s core. This mirrors our findings correlating ventricle size with TBR1 frequency (Figure 6.3i) since larger ventricles tend to be at the organoid’s surface in our organoids. These trends highlight that an organoid’s histological section will possess cell frequency biases dependent on its overall position in the tissue.

To compare how this sampling-biased heterogeneity compares with biological inter-organoid variability, we compared the statistical distribution of cell type frequencies in 10 organoids with the statistical distribution of possible estimates using 100 $\mu$ m thick “pseudo-sections” in these same organoids (Figure 6.4c,d). Variability caused by uniform pseudo-section sampling was greater than the biological inter-organoid variability. Theoretically, to reduce the heterogeneity of histological sampling to match the inter-organoid variability in cell type frequency, more than five to seven histological sections would be necessary per organoid. This represents the number of 100 $\mu$ m sections that would be required to equate the standard deviation of the sample mean with that of the inter-organoid distribution over each cell type frequency.

To understand inter-organoid differences, we combined all of our multiscale analyses into a set of 276 descriptors capturing single cell, cytoarchitecture and whole-tissue features in a single organoid (Table 8.4). Comparative analysis among day 35 organoid replicates reveals unequal variance in specific features. When looking at the relative standard error of these features, we see a wide range of heterogeneity (Figure 6.4e). Some of the most variable features include average ventricle volume, total count of DN cells, and organoid volume. Most consistent features are the single cell morphology descriptors, which are virtually identical in all replicates. Interestingly, the frequency of SOX2 and TBR1 cells was more consistent than their absolute number. Given the large variability of certain features, we computed pairwise correlations between multiscale features to see if certain features in single-cell features could predict whole-tissue topography (Figure 6.4f). The most interesting multiscale correlation occurs around SOX2 and TBR cell counts. A network analysis of pairwise correlations ( $r > 0.75$ ) reveals a network of correlated multiscale features where SOX2 count and organoid volume were the most central nodes (degree of 30) impacting ventricle size, TBR1 count, co-adjacent DN counts and average



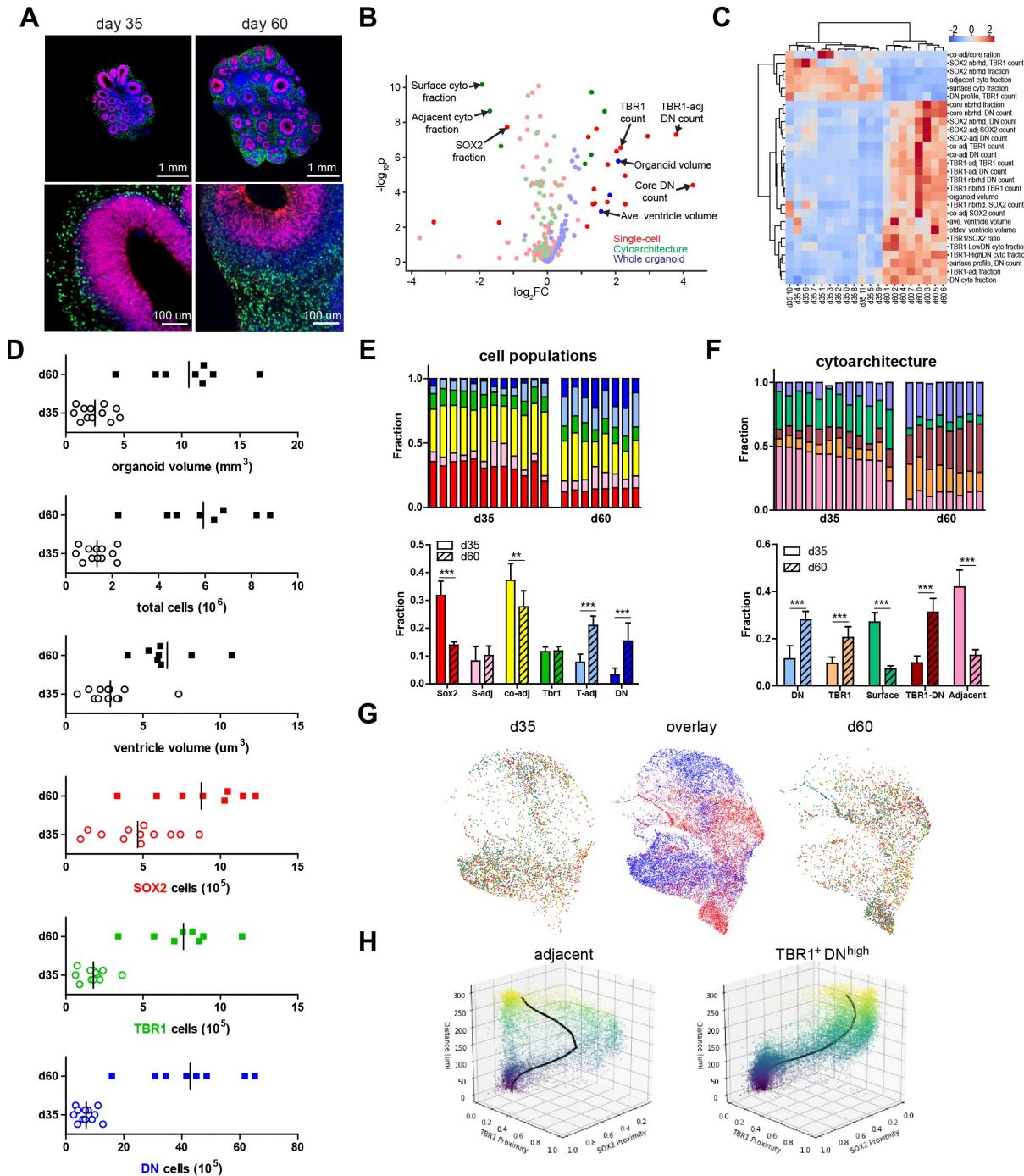
**Figure 6.4:** Whole organoid analysis for unbiased quantitative studies. (A) Scheme (top) and analysis of three different 100  $\mu\text{m}$  pseudo-sections from a 3D dataset. Pie charts (right) show the variable depth-dependent cell frequency for individual slices. Bottom shows the estimated cell frequency for each cell type with the actual whole-organoid frequency (dotted line). (B) Analysis of cell frequencies for 1187 independent ventricles pooled from 12 day 35 organoids. Cell frequency for each organoid was determined by combining the counts of cells detected in all “virtual biopsies” used for cytoarchitecture analysis. (C) Comparison of pseudo-slice heterogeneity with biological inter-organoid variability. Each histogram shows the distribution of 10,000 pseudo-sections 100  $\mu\text{m}$  thick for 12 organoids (black) versus the distribution of whole-organoid frequency (colored histogram, ticks show independent organoid values). (D) Comparison of cell frequency standard deviation for the pseudo-sections sampling variability (2D) versus whole-organoid biological variability (3D). The number of pseudo-sections required to match biological variability is shown for each cell population. (E) Comparison of the relative standard error for 276 multiscale features in 12 d35 organoid replicates where each dot is colored based on to the scale of its analysis. (F) Heat map of Pearson’s correlation coefficient investigating the relationship between the 276 multiscale features and their variation in 12 d35 organoids. Right shows the cropped region where we see a combination of single-cell (red) and whole-organoid (blue) features where  $r > 0.88$ . (G) Network of feature correlation when the absolute Pearson’s correlation coefficient is  $>0.75$ . Cropped region shows the same multiscale correlation as the heat map in panel F. The SOX2 cell count and organoid volume were the most central nodes in this cluster, both having a degree of 30.

SOX2 distance from the surface (Figure 4G). The analyses show that SCOUT can reduce positional bias-related variance, quantify biological heterogeneity in replicates, and can provide the correlation of multiscale features.

### 6.3.5 SCOUT analysis of organoid maturation

Next, we used SCOUT for comparative analysis between day 35 and day 60 cerebral organoids to detect maturation-related multiscale changes (Figure 6.5). We compared individual features of each time point using independent two-tailed t-tests and detected 28 significant ( $p<0.01\%$ ) changes we rank-ordered according to largest fold-change (Figure 6.5b,c). Day 60 organoids showed a four-fold increase in total volume and a two-fold increase in ventricle size. As expected, growth produced a large increase in cell numbers (19X DN, 4X TBR, 2X SOX2; Figure 6.5d). Non-uniform cell expansion caused a 56% reduction in SOX2 cell frequency and two-fold increase in the Tbr1/Sox2 ratio. We suspect the low expansion of progenitor cells is caused by our addition of BDNF at day 40 of organoid culture, which drives neurogenesis and suppresses progenitor expansion (Fukumitsu et al. 2006).

Cell type frequency was relatively consistent in replicate organoids, supporting the importance of whole-organoid analysis to reduce histology-related variability (Figure 6.5e). In d60 cerebral organoids, as progenitor cells generated radially organized cell populations, a visible layer of DN cells appears between SOX2 and TBR1 regions (Figure 6.5a). This new cell layer reduced the proximity of SOX2 and TBR1 cells to one another (43% decrease for SOX2, 34% decrease for TBR1). Maturation also produced a 7X increase in “TBR1-adjacent DN cells” frequency and the emergence of a “Tbr1+DNhi” cytoarchitecture (31.2%) characterized by abundant DN cells appearing after the SOX2 ventricular zone and continuing into the TBR1 region (Figure 6.5f,g). In these cytoarchitectures, DN cells likely represent a combination of intermediate progenitors and non-TBR1 neurons that emerge over the course of organoid maturation. These new cell populations increased the inter-ventricle and ventricle-to-surface distances causing a 70% reduction in “Adjacent” and “Surface” cytoarchitectures, respectively. As organoids mature, they replaced these early cytoarchitectures with “TBR1+DNhi” (31.2%) and “TBR1+DN-low” (28%), highlighting a general increase in TBR1+ cells around the maturing ventricles. Another important change in cytoarchitecture is the doubling of “DN only” cytoarchitectures from 14% to 28% as a consequence of the 14-fold increase in “core DN” cells. The “DN only” cytoarchitectures are usually oriented towards the organoid core where no ventricles or cell markers are detected. This region grows in d60 organoids since the reduced surface-to-volume ratio in these large organoids likely interferes with the transport of oxygen, nutrients and growth factors.

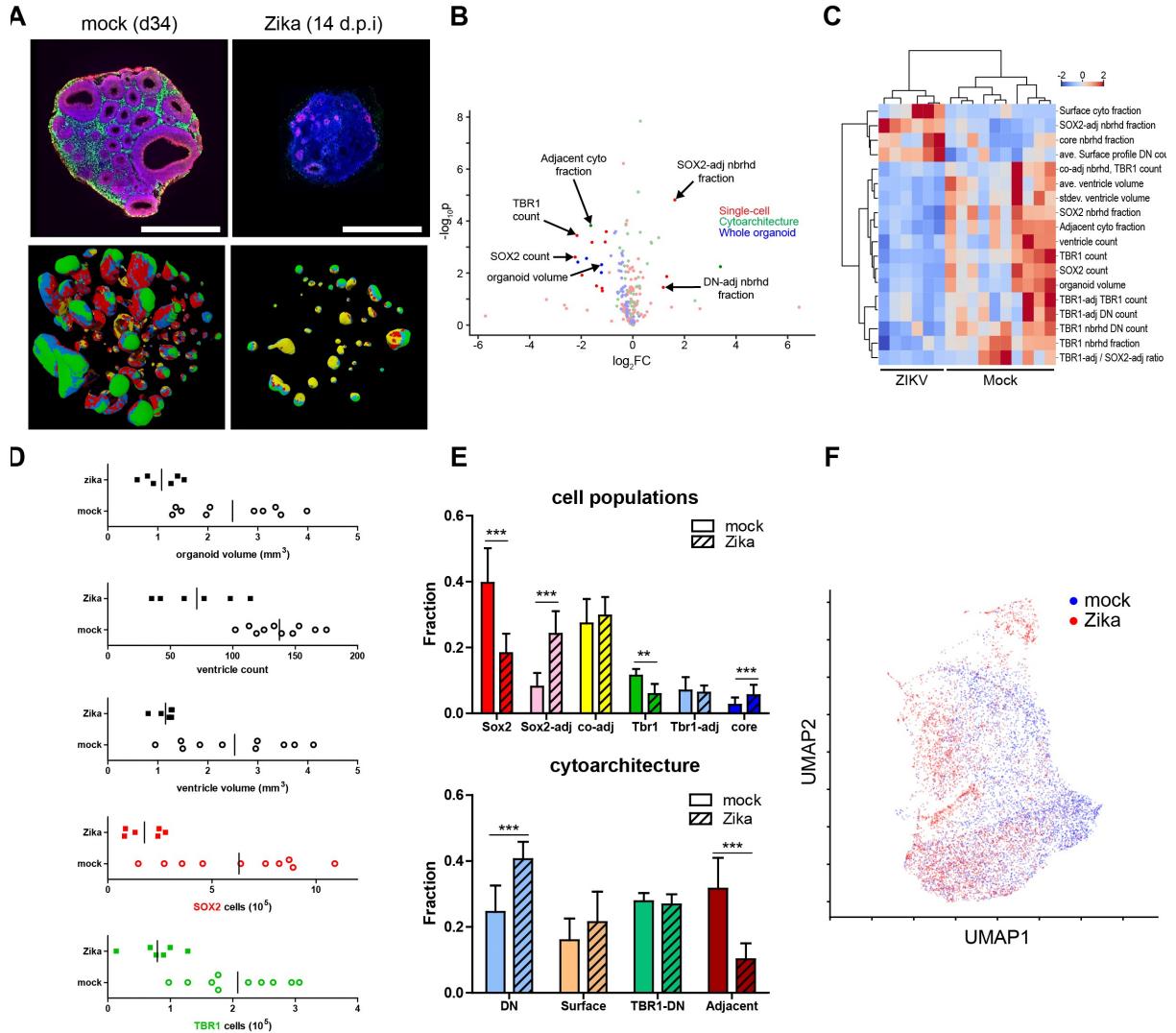


**Figure 6.5:** Hyper-dimensional analysis of multiscale changes during cerebral organoid maturation. (A) Representative image of day 35 and 60 organoids taken from volumetric datasets. (B) Volcano plot highlighting unbiased detection of multiscale differences with dots colored according to scale of analysis. (C) Heat map outlining major differences between twelve d35 vs. eight d60 organoids. (D) Dot plots showing organoid volume, total cells, average ventricle volume, and total counts of antibody-labelled subsets. (E) Cell subpopulation frequencies for independent replicates (top) and average values (bottom). (F) Quantitative analysis of cytoarchitecture frequencies for independent replicates (top) and average values (bottom). (\*\*\*)  $p < 0.001$ , (\*\*)  $p < 0.01$  (G) UMAP embedding of the cytoarchitectures detected in d35 and d60 organoids. Left shows the distribution of cytoarchitectures in d35 organoids where each color represents a different organoid. Middle UMAP shows cytoarchitectures according to age of organoids (d35 blue and d60 red). Right UMAP shows the distribution of cytoarchitectures in the d60 organoids where each color represents a different organoid. (H) Analysis of cell proximity to SOX2 and TBR1 in “adjacent” (left) and “TBR1, DN-high” (right) as a function of the distance from the ventricle surface.

Although organoids appear heterogeneous when comparing random cross sections, full-spectrum cytoarchitecture reveals a relatively consistent distribution of radial organization in our biological replicates. To assess the consistency of specific cytoarchitecture clusters, we “gated” on the “Adjacent” and “TBR1+DNhi” to compare SOX2 and TBR1 proximity of individual cells relative to their distance from the ventricle surface (Figure 6.5h). This analysis confirms that cells found within the “Adjacent” cytoarchitectures consistently show a slight increase in TBR1 proximity at ~75  $\mu$ m then return to high SOX2 proximity. In contrast, the more mature TBR1+DNhi regions show a one-time transition from SOX2-proximity to TBR1-proximity ~150 $\mu$ m away from the ventricle surface.

### 6.3.6 Comparative analysis of Zika infection

After investigating maturation-induced changes, we quantified how cerebral organoids were altered by Zika virus infection (Figure 6.6). Previous studies have used cerebral organoids to monitor the tropism of Zika virus infection and how it produces microcephaly in newborns (Cugola et al. 2016). Initial studies suggest that infection of neural progenitors causes a reduced output of mature neurons and decreased size of brain tissue. Using a clinical isolate from Puerto Rico, we infected organoids at day 21 with 0.1 MOI and waited 14 days post-infection to assess the multiscale impact of viral infection. SCOUT-analysis detected 18 significant differences ( $p<0.01$ ; Figure 6.6b,c). Notably, we detected a ~50% reduction in organoid size, ventricle volume and total ventricle counts (Figure 6.6d). At the cellular level, we detected a 75% total decrease and 50% frequency decrease of both SOX2+ and TBR1+ cells (Figure 6.6e). The combined loss of ventricles and SOX2 cells lead to a 67% decrease (31% to 10%) in the “Adjacent”cytoarchitecture, while TBR1 loss caused the 24 to 40% increase in “DN only”cytoarchitectures. Zika-infection produced an aggressive phenotype that limited organoid size, ventricle growth, and the expansion of SOX2 and TBR1 cells (Figure 6.6f). Given the importance of SOX2 cell number in the emergence of higher order features (Figure 6.4f) it is unsurprising that the significant loss of neural progenitors produced an important decrease in the complexity of tissue topography and cell patterning. Quantification of the Zika infection phenotype is consistent with previous reports and may mirror the mechanisms underlying viral microcephaly in newborns. Unfortunately, it remains unclear whether the loss of TBR1+ neurons is due to disrupted progenitor differentiation or neuron-specific cell death. However, SCOUT can help discern between these two possibilities in future experiments.

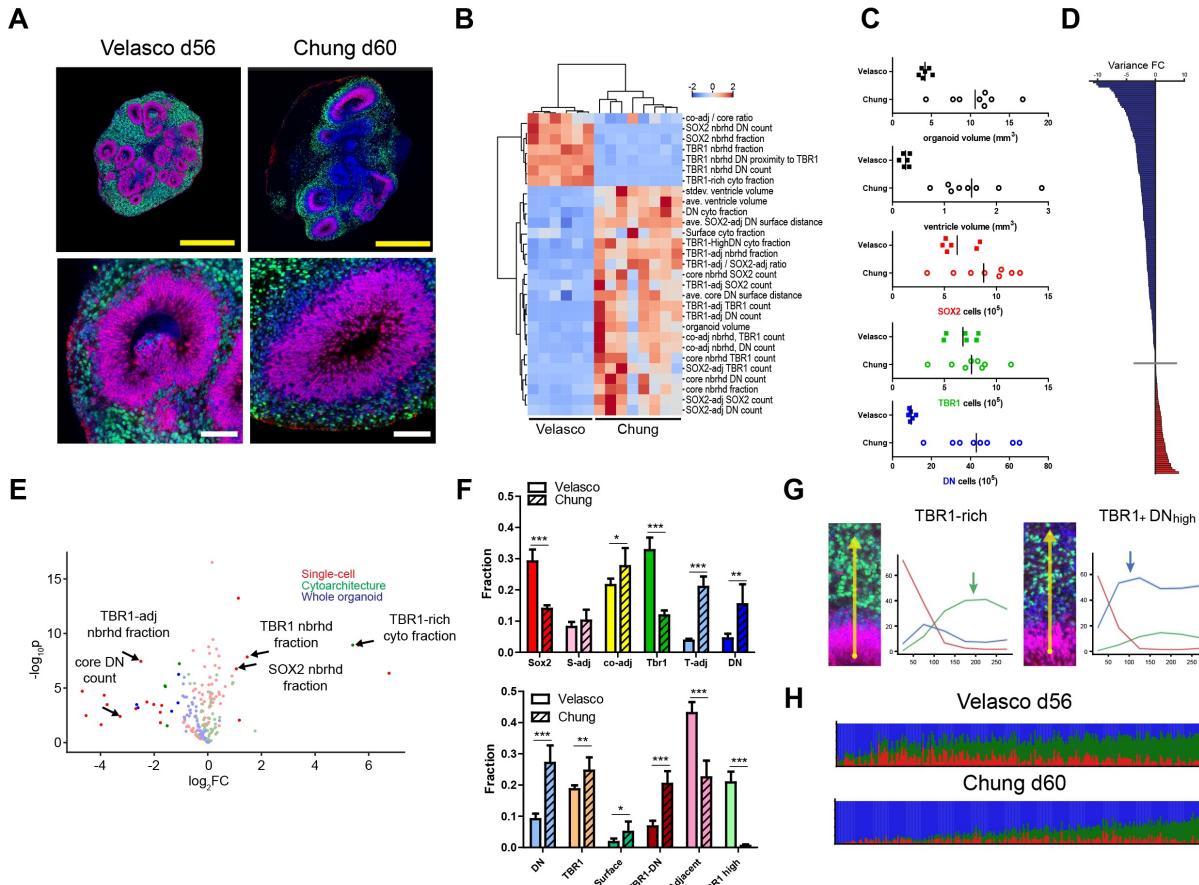


**Figure 6.6:** Hyper-dimensional analysis of multiscale changes caused by Zika virus infection. (A) Top: representative image of age-matched day 34 mock and Zika virus-infected organoid (14 days post-infection) taken from volumetric datasets (scale bar 1 mm). Bottom: rendering of segmented ventricles from each group with their surfaces colored according to detected cytoarchitecture. (B) Volcano plot highlighting unbiased detection of multiscale differences with dots colored according to scale of analysis. (C) Heat map outlining major differences between ten mock vs. six Zika-infected organoids. (D) Dot plots showing organoid volume, ventricle count, ventricle volume, and total counts SOX2 and TBR1 cells. (E) Comparison of cell subpopulation frequencies (top) and cytoarchitecture frequency (bottom). (\*\*\*)  $p < 0.001$ , \*\*  $p < 0.01$  (F) UMAP embedding of the cytoarchitectures detected in mock (blue) and Zika-infected organoids (red) showing a shift in cytoarchitecture.

### 6.3.7 Comparative analysis of different cerebral organoid protocols

High-dimensional phenotyping of multiple whole organoids provides one a strategy to overcome the inter-organoid heterogeneity (Quadrato et al. 2017). However, a recent study (Velasco et al. 2019) by Velasco et al. addressed biological heterogeneity by comparing different brain organoid culture protocols. Their patterned dorsal forebrain organoids showed an impressive organoid-to-organoid consistency in morphology and cellular composition, assessed by single-cell RNA sequencing. We compared day 56 “Velasco organoids” with our day 60 organoids, cultured according to a modified version of the Lancaster protocol (Figure 6.7). We applied dual SMAD inhibition for the first half of neuroepithelium induction, added BDNF after day 40 and used Petri dishes on a shaker instead of a spinning bioreactor (Lancaster & Knoblich 2014). Comparative analysis detects 29 different features whose  $p < 0.01$  (Figure 6.7b,e). Most noticeably, the Velasco organoids and their ventricles were smaller, showing an average 60% and 84% decrease, respectively (Figure 6.7c). At the cellular level, Velasco organoids show a 2X higher frequency of SOX2 cells and 3X increase in TBR1 frequency (Figure 6.7f). The proximity between the SOX2 and TBR1 cells also increased twofold due to the absence DN cells between the progenitor zone and TBR1 cells (Figure 6.7a,g). DN cells were less frequent overall, which produced a 70% reduction in core DN cells and an 80% reduction in TBR1-adjacent DN cells.

Cytoarchitectures were vastly different in these two protocols. The Velasco organoid’s abundant TBR1 cells produced the appearance of a new “TBR1-rich cluster” (~21% of total biopsies), where TBR1 counts are ~35% higher and DN counts are 70% lower compared to the “TBR1+DN-high” cytoarchitecture favored in our day 60 organoids (Figure 6.7g,h). Velasco organoids also show more “Adjacent” cytoarchitectures, consistent with smaller organoids. The 66% reduction in “DN only” cytoarchitecture reflects less “core cells” which can be attributed to smaller organoid size or the improved tissue transport when using a spinning bioreactor. Interestingly, Velasco organoids possessed better feature consistency across multiscale features. We quantified the variance of all 276 features and compared the fold-change in both organoid groups. The Velasco organoids show decreased variance in ~70% of features (Figure 6.7d). The dramatic differences in these two groups was expected given their different culture protocols. However, this proof-of-concept provides the foundation for establishing quantitative unbiased assessment of different culture protocols. These finding suggest that a combination of holistic analysis and protocol optimization can address the widely reported problem of organoid heterogeneity.



**Figure 6.7:** Hyper-dimensional analysis of multiscale changes in different culture protocols. (A) Representative image of day 56 Velasco and our day 60 organoids taken from volumetric datasets. (yellow bar = 1 mm; white bar = 100  $\mu$ m) (B) Heat map outlining major differences between different organoid culture protocols. (C) Dot plots showing organoid volume, average ventricle volume, and total counts of cell types. (D) Comparative analysis of the fold-change in the relative standard error of quantified features. Dotted line shows zero, bars reflect better homogeneity in Velasco (blue) organoids and our organoids (red). Graph shows 72% of features (141/195) were more consistent in Velasco organoids (E) Volcano plot highlighting unbiased detection of multiscale differences with dots colored according to scale of analysis. (F) Comparison of cell subpopulation frequencies (top) and cytoarchitecture frequency (bottom). (\*\*\*)  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ . (G) Representative image and average profile plot of individual cytoarchitecture clusters showing the radial distribution of SOX2 (red), double negatives (blue) and TBR1 (green) cells in the Velasco-specific TBR1-rich cytoarchitecture and the day 60-specific Tbr1+DN-high cytoarchitecture. Arrows show the higher TBR1 frequency (green) and DN band between SOX2 and TBR1 (blue). (H) Frequency of SOX2 (red), TBR1 (green) and DN (blue) for 250 random ventricles pooled from all organoids. Cell frequency was determined by considering the detected cells all of a ventricle's virtual biopsies.

## 6.4 Discussion

SCOUT provides a much-needed pipeline for the phenotypic description and comparative analysis of three-dimensional tissue culture models. Here, we provided the first attempt for unbiased quantification of multiscale changes in different cerebral organoid groups. SCOUT can be applied to generate robust statistically-powered hypotheses through the volumetric analysis of numerous intact organoids. This initial set of studies confirms the possibility of quantifying maturation-induced growth, Zika virus-induced damage, and protocol-specific improvements in cerebral organoids. We provided the first ever quantification of ventricle size, cytoarchitecture distribution and multiscale cellular analysis in well-studied models.

Our current list of features and multiscale correlation will continue to evolve as we gain more experience quantifying three-dimensional cellular biology in the developing brain. In this study, we combined some new concepts such as cell proximity analysis with some widely-used general strategies for characterization of radially-organized cell populations. Analysis was based on SOX2 and TBR1 because we were interested in the radial organization of progenitors and mature neurons in relatively young organoids. However, these antibodies could be substituted to address the biological question being investigated. Additional challenges moving forward include the possibility of increasing the number of markers to provide a richer phenotypic analysis of organoids. Several strategies can increase the number of detectable markers in future studies such as barcoding of primary antibodies, multi-round tissue staining, or the addition of more laser lines and filter sets to the microscope. Another challenge is the inability to disentangle correlation and causation in our current study. However, this can be addressed with strategic experimental design that utilizes time-point analysis and fluorescent reporter systems to catch emergence of new features over time.

In its current state, SCOUT provides the first attempt at holistic cerebral organoid characterization. Although organoids are an exciting *in vitro* model for human brain development, their heterogeneity can be especially problematic when using two-dimensional tissue sections for cellular and morphological analysis. Whole organoid analysis produced some variance, but trends were generally consistent and we were capable of detecting numerous differences in each experiment. Combining SCOUT with improved culture protocols such as the Velasco organoids will establish consistent three-dimensional models where analysis can discriminate multiscale changes with impressive signal-to-noise. This will hopefully empower the discovery of mechanisms underlying the biology and dysfunction of human brain development.

## 6.5 Methods

### 6.5.1 Cerebral Organoid Culture

Organoids were cultured according to the protocol by Lancaster et al. (2014) using the SC101A iPS cell line (Systems Biosciences). The iPSC cells were cultured on Matrigel-coated (Corning) plates using mTeSR medium (Stemcell Technologies), passaged using ReLeSR at 80% confluency, and organoids were made before the 15th passage of initial cells. For organoids, single cells were detached with Accutase (Stemcell Technologies) when iPS cells were at 60-80% confluency. We seeded 9,000 cells in ultra-low attachment round bottom 96-well plates in hESC medium32 with 4 ng/mL bFGF (Peprotech) 50  $\mu$ M Y-27632 Rock-inhibitor (Tocris) for the first 4 days then without for an additional two days. At day 6, organoids were transferred to ultra-low attachment 24-well plates in neural induction medium (Lancaster et al. 2014) with addition of SMAD inhibitors (Mellios et al. 2018) 10 $\mu$ M SB-431542 and 1 $\mu$ M dorsomorphin (Tocris). At day 9, neural induction medium was replaced with fresh medium without the SMAD inhibitors. At day 12, organoids were embedded in 15 $\mu$ L growth-factor reduced Matrigel droplets (Corning). We cultured 12 organoids per 60mm Petri dish for suspension culture in 5mL cerebral organoid differentiation medium (Lancaster et al. 2014) without vitamin A for 4 days. At day 16, medium was replaced with cerebral organoid differentiation medium with vitamin A (Lancaster et al. 2014) and placed on shaker at 85 rpm. Medium was replaced twice per week and at day 40, we added 14 ng/mL BDNF to organoid medium (Quandrato et al. 2017).

### 6.5.2 SHIELD Sample Preparation

Organoids were rinsed once with PBS, then fixed with freshly prepared 4% PFA in PBS (EM grade, Electron Microscopy Sciences) at room temperature for 30 minutes on a shaker. Organoids were rinsed three times in PBS, transferred into ice cold 2% polyglycerol 3-polyglycidylether (wt./v) in 0.1M phosphate buffer pH 7.2 and incubated for two days at 4°C. Organoids were subsequently transferred into pre-warmed 0.1M sodium carbonate buffer (pH 10) and incubated at 37°C for 24h. Organoids were washed extensively with PBS for 8h, cleared in 0.2M SDS + 50 mM phosphate buffer pH 7.3 for 48h at 55°C while shaking in EasyClear system (LifeCanvas), and washed extensively in PBST (PBS, 0.1% Triton X-100, 0.02% sodium azide) for 24h.

### 6.5.3 eFLASH Sample Preparation

SHIELD-processed and cleared organoids were stained using an adapted version of the eFLASH protocol (Yun et al. 2019). We incubated organoids in eFLASH sample buffer (0.3M Tris, 0.2M CAPS, 20% D-sorbitol, 1 mM SDS, 2% bovine serum albumin, 0.02% sodium azide) overnight at room temperature. Organoids were then placed in the Smart-Label system (Lifecanvas) with 1.4 mL sample buffer in the sample cup. For SCOUT pipeline, we added 6 $\mu$ L Syto16 (1 mM solution, ThermoFischer #S7578), 15 $\mu$ g goat anti-SOX2 antibody (R&D Systems #AF2018), 10 $\mu$ g Fab fragment anti-goat IgG Alexa Fluor 594 (Jackson ImmunoResearch #805-587-008), 30 $\mu$ g rabbit anti-TBR1 Alexa Fluor 647 (Cell Signaling Technology #45664S). System was filled with OFF main buffer (0.3M Trizma base, 0.2M CAPS, 20% D-sorbitol, 1 mM SDS) and we applied 90V at 0.01 rpm for 16h. The following day, we added 0.2mL 20% Triton X-100 solution to the sample chamber, drained the system, rinsed the system with deionized water, and filled it with ON main buffer (10mM phosphoric acid, 60mM Trizma base, 20% D-sorbitol). We applied 90V at 0.01 rpm for 21h, removed samples and rinsed them in PBST overnight. To increase antibody signal and minimize signal loss in the immersion medium, we fixed antibody-stained samples in freshly prepared PBS + 4% PFA for 12h at room temperature.

### 6.5.4 Whole-Organoid Imaging

Prior to imaging, organoids were equilibrated in Protos-based immersion medium (Murray et al. 2015; Yun et al. 2019) in two steps. First, we incubated organoids in a 1:1 mix of Protos and PBS for 4h. Then, we replaced the solution with Protos immersion medium for at least 6h. Images in Figure 1 were acquired using the Leica TCS SP8 laser-scanning confocal microscope with a white light laser source for excitations at 488, 594 and 647 nm using a 20X 0.5-NA water-immersion objective (Leica #15506147, HCX APO L 20x/0.50 W U-V-I). For rapid volumetric imaging in subsequent figures, Protos-immersed samples, were mounted in a 1.5% agarose block prepared in Protos. We were able to mount 6-8 organoids per block at a time. After agarose polymerization, the block was equilibrated in 25 mL Protos medium overnight. Samples were mounted and imaged with a SmartSPIM axially swept light-sheet microscope (LifeCanvas Technologies) equipped with three lasers (488nm, 561nm, 642nm) and a 10x objective (Olympus XLPLN10XSVMP, 0.6NA, 8mm WD, lateral resolution 0.66 $\mu$ m in XY). We imaged samples at a 0.6 x 0.6 x 2  $\mu$ m voxel size.

### 6.5.5 Preprocessing LSFM images of cerebral organoids

Cerebral organoid images from the LSFM system (LifeCanvas, SmartSPIM) originally stored in an uncompressed binary format were first destriped and stitched according to a previously reported image processing pipeline (Swaney et al. 2019). This pipeline generates z-slice images with lossless compression that are used to compute image histograms for each channel. After normalizing each channel to the 99th percentile of the histogram, each channel is then partitioned into (64, 64, 1) voxel chunks using the Zarr Python package. This chunk-compressed representation allows for parallel processing of each image data chunk.

### 6.5.6 3D nuclei detection using curvature-based seeded watershed

To detect all nuclei in volumetric nuclear stain images, a curvature-based nuclei filtering strategy was developed. An image filter based on the eigenvalues of the Shape Operator (also known as the Weingarten matrix) was created by defining a probability distribution over the image curvature and intensity as evidence for a nucleus centroid. This nucleus probability map highlights nuclei even in densely-packed regions such as the ventricular zone in cerebral organoids. The probability map was computed for each image chunk after gaussian smoothing to remove noise. Nuclei centroids were extracted from these probability maps through local maxima detection.

To segment each nucleus, the detected nuclei centroids were used as seed points for performing a watershed segmentation of a nucleus foreground mask. This nucleus foreground mask is obtained through binarization of the nucleus probability map. Binarization of the nuclei probability map was accomplished through graph cuts segmentation with a Poisson image intensity prior and constant pairwise interaction penalty.

### 6.5.7 Nuclei detection accuracy measurement

To assess the accuracy of our nuclei detection strategy, nuclei centroids were hand-annotated using a previously reported image alignment tool called Nuggt. Over 1,000 nuclei centroids were hand-annotated in organoid subregions including the ventricular zone and subventricular zone. Detected nuclei were considered true positives (TP) if they fell within 5 voxels of a ground-truth centroid and false positives (FP) if not. Since multiple detections may lie within 5 voxels of a given ground-truth centroid, double-counting TPs was avoided by matching ground truth and detected centroids using the Hungarian method for solving the linear sum assignment problem. The result of this

matching procedure are lists of all TPs, FP, false negatives (FN), and true negatives (TN). These accuracy statistics were used to compute the overall accuracy and F1-score of the nuclei detection strategy. Our curvature-based seeded watershed algorithm was benchmarked against the difference of gaussian (DoG) and Laplacian of guassian (LoG) blob detection algorithms implemented in the scikit-image Python package.

### 6.5.8 Cellular subcategorization using in situ cytometry and spatial proximity analysis

Protein expression of transcription factors was measured by sampling the SOX2 and TBR1 antibody staining in a 3-voxel diameter spherical ball centered around the nuceli centroids. The mean fluorescence intensity within each sphere was computed for each channel and gated to define SOX2-/+ and TBR1-/+ cell populations. The result of this gating strategy is a labeled point cloud of nuclei centorids.

Spatial proximity analysis of the labeled nuclei point cloud was accomplished by first constructing a KD-tree representation of the point cloud for efficient querying of nearest nuclei. The spatial proximity to the nearest  $k$  cells of type  $t$  was calcualted using the following formula for each detected nucleus  $i$ :

$$P_i^{(t)} = \prod_{n=1}^k \frac{1}{1 + d_{i,n}/\sigma^{(t)}}$$

where  $d_{i,n}$  is the distance between the  $i$ -th nucleus and the  $n$ -th nearest nucleus and  $\sigma^{(t)}$  is a reference distance that controls the proximity bandwidth (how close a neighboring nucleus must be to be considered in close proximity).

### 6.5.9 Automatic 3D ventricle segmentation using U-Net

To train a U-Net model for ventricle segmentation, downsampled volumetric images of SOX2 staining were semi-automatically segmented using ITK-Snap. In total, 9 whole organoids were segmented, including 7600 binary 2D images. These images were combined with the corresponding syto16 images at the same resolution and split with a 20% hold-out test set before moving on to model training and validation. The remaining training set was used with ten-fold cross-validation to tune model hyperparameters and the overall model architecture.

The U-Net model was implemented in Keras and slightly modifed from the original architecture. Since our images were higher resolution than what was used in the original

U-Net paper, we added two layers before and after the U-Net bottleneck to increase the receptive field of the model. This modified U-Net model was trained using a hybrid loss containing a weighted binary crossentropy (WBCE) term and a Dice coefficient loss term. The WBCE term was weighted at 90% to the Dice coefficient loss term's 10%. In our experience, the WBCE term helps the model converge to sensible ventricle segmentations due to a simpler gradient signal during training. However, the Dice coefficient loss is required to compensate for the high degree of class imbalance in any given training image.

The test accuracy was assessed after all training and validation steps were complete by computing the Dice coefficient for all test images. The receiver operating characteristic curve was constructed using a random sample of 100 images from the test set to speed up computations. These test images were segmented and thresholded at linearly spaced probability values between 0 and 1. For each threshold, the TP, FP, and FN rates were computed from corresponding pixels in the predictions and ground truth images. These rates were used to compute the final precision and recall values as well as the maximum F1-score and area under curve (AUC).

### 6.5.10 3D cytoarchitectural analysis of cerebral organoids

Binary ventricle segmentations were converted into a 3D mesh using the marching cubes algorithm from the scikit-image Python package. The resulting mesh contained vertices, faces, and normal vectors uniformly distributed over the ventricle surfaces. The normal vectors and vertices were used to query which nuclei centroids were within a 50  $\mu\text{m}$  diameter and 300  $\mu\text{m}$  tall cylindrical volume around the surface normal. These cells were bin counted for each cell type over size 50  $\mu\text{m}$  intervals to construct the final radial cell profiles.

To determine cytoarchitectural types, 5000 radial cell profiles were sampled from each organoid in a given dataset (i.e. from both d35 and d60 organoids in the maturation analysis). These radial cell profiles were flattened into vectors and concatenated into a matrix of cytoarchitecture observations and features. This matrix was provided to UMAP to visualize the distribution of cytoarchitectures in 2D and perform dimensionality reduction before clustering. After performing UMAP embedding, the UMAP model was saved for future analysis. The UMAP embedded cytoarchitectures were then grouped using hierarchical clustering with a euclidean distance metric and average linkage method. The cluster labels were saved and used as training data in a nearest neighbor cytoarchitecture classifier. Using the saved UMAP model and the pretrained nearest neighbor cytoarchitecture classifier, all profiles in each organoid were classified efficiently.

The resulting cytoarchitectural labels were used to compute average profiles for each clusters as well as for coloring the faces of the ventricles in a 3D render. This 3D render of ventricles pseudocolored by cytoarchitecture was generated in Blender.

### 6.5.11 Hyperdimensional statistical testing for comparative organoid studies

To perform statistical testing on all multiscale features, we used independent two-tailed t-tests to obtain significance values for each feature. To reduce the amount of false positives due to multiple comparisons, we also thresholded the significant phenotypic changes at a two-fold change in mean. This thresholding restricted the detected hits to those with large changes in mean so that those significant differences that are due to small sample standard deviations are removed.

### 6.5.12 Pairwise correlation analysis of multiscale organoid features

For 12 d35 organoids, the Pearson correlation coefficient and associated P-value was computed for all pairs of organoid features. These correlation coefficients were stored in a matrix and biclustered using hierarchical clustering.

To construct a network from these pairwise correlations, the absolute value of the correlation coefficients were thresholded at 0.75 to remove edges in the network with smaller weights. This thresholded matrix of pairwise correlation coefficients was used as an adjacency map in the NetworkX Python package. The resulting network was then visualized in an interactive plot using the pyvis Python package.

### 6.5.13 Estimation of intra-organoid section variability using psuedosections

To compare the intra-organoid variability of sections to the inter-organoid variablity in 3D analysis, the nuclei centroids and cell-type labels from 10 d35 organoids were used in both 2D and 3D analyses. Pseudosections were 100  $\mu\text{m}$  virtual sections of the same underlying 3D dataset of nuclei centroids and cell-type labels that align with the XY plane and fall within the Z limited of the organoid. 10,000 pseudosections were sampled from each organoid and used to compute the overall (normalized) cell frquency for SOX2, TBR1, and DN cells. Distributions of these psuedosection cell-type frequencies were comapred to the distributions of cell-type frequencies obtained by comparing the 3D

orgnaoid datasets directly. The central limit theorem was used to compute the sample size of psuedosections needed to match the variance of the inter-organoid distribution.

## 6.6 References

- Atta-Fosu, T. et al., 2016. 3D Clumped Cell Segmentation Using Curvature Based Seeded Watershed. *Journal of Imaging*, 2(4), p.31. Available at: <http://www.mdpi.com/2313-433X/2/4/31>.
- Bershteyn, M. et al., 2017. Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell stem cell*, 20(4), pp.435–449.e4.
- Camp, J.G. et al., 2015. Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proceedings of the National Academy of Sciences of the United States of America*, 112(51), pp.15672–7.
- Cugola, F.R. et al., 2016. The Brazilian Zika virus strain causes birth defects in experimental models. *Nature*, 534(7606), pp.267–271. Available at: <http://www.nature.com/articles/nature18296>.
- Eiraku, M. et al., 2011. Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature*, 472(7341), pp.51–56. Available at: <http://www.nature.com/articles/nature09941>.
- Fukumitsu, H. et al., 2006. Brain-derived neurotrophic factor participates in determination of neuronal laminar fate in the developing mouse cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(51), pp.13218–30.
- Gjorevski, N. et al., 2016. Designer matrices for intestinal stem cell and organoid culture. *Nature*, 539(7630), pp.560–564. Available at: <http://www.nature.com/articles/nature20168>.
- Homan, K.A. et al., 2019. Flow-enhanced vascularization and maturation of kidney organoids in vitro. *Nature Methods*, 16(3), pp.255–262. Available at: <http://www.nature.com/articles/s41592-019-0325-y>.
- Kadoshima, T. et al., 2013. Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. *Proceedings of the National Academy of Sciences of the United States of America*, 110(50), pp.20284–9.
- Kim, S.-Y. et al., 2015. Stochastic electrotransport selectively enhances the transport of highly electromobile molecules. *Proceedings of the National Academy of Sciences*, 112(46), pp.E6274–E6283. Available at: <http://www.pnas.org/lookup/doi/10.1073/pnas.1510133112>.
- Knight, G.T. et al., 2018. Engineering induction of singular neural rosette emergence

within hPSC-derived tissues. *eLife*, 7. Available at: <https://elifesciences.org/articles/37549>.

Lancaster, M.A. & Knoblich, J.A., 2014. Generation of cerebral organoids from human pluripotent stem cells. *Nature Protocols*, 9(10), pp.2329–2340.

Lancaster, M.A. et al., 2013. Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467), pp.373–379.

Luo, C. et al., 2016. Cerebral Organoids Recapitulate Epigenomic Signatures of the Human Fetal Brain. *Cell reports*, 17(12), pp.3369–3384.

Mariani, J. et al., 2015. FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell*, 162(2), pp.375–390.

McInnes, L., Healy, J. & Melville, J., 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Available at: <http://arxiv.org/abs/1802.03426>.

Mellios, N. et al., 2018. MeCP2-regulated miRNAs control early human neurogenesis through differential effects on ERK and AKT signaling. *Molecular Psychiatry*, 23(4), pp.1051–1065.

Murray, E. et al., 2015. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6), pp.1500–14.

Park, Y.-g. et al., 2018. Protection of tissue physicochemical properties using polyfunctional., (December).

Qian, X. et al., 2016. Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell*, 165(5), pp.1238–1254.

Quadrato, G. et al., 2017. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature*, 545(7652), pp.48–53. Available at: <http://www.nature.com/articles/nature22047>.

Renner, M. et al., 2017. Self-organized developmental patterning and differentiation in cerebral organoids. *The EMBO Journal*, 36(10), pp.1316–1329. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.15252/embj.201694700>.

Ronneberger, O., Fischer, P. & Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available at: <http://arxiv.org/abs/1505.04597>.

Sato, T. et al., 2009. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 459(7244), pp.262–265. Available at: <http://www.nature.com/articles/nature07935>.

Swaney, J. et al., 2019. Scalable image processing techniques for quantitative analysis of volumetric biological images from light-sheet microscopy. *bioRxiv*.

Takasato, M. et al., 2015. Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature*, 526(7574), pp.564–568. Available at: <http://www.nature.com/articles/nature15695>.

Velasco, S. et al., 2019. Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature*, 570(7762), pp.523–527. Available at: <http://www.nature.com/articles/s41586-019-1289-x>.

Watanabe, M. et al., 2017. Self-Organized Cerebral Organoids with Human-Specific Features Predict Effective Drugs to Combat Zika Virus Infection. *Cell reports*, 21(2), pp.517–532.

Wimmer, R.A. et al., 2019. Human blood vessel organoids as a model of diabetic vasculopathy. *Nature*, 565(7740), pp.505–510. Available at: <http://www.nature.com/articles/s41586-018-0858-8>.

Yoon, S.-J. et al., 2019. Reliability of human cortical organoid generation. *Nature Methods*, 16(1), pp.75–78. Available at: <http://www.nature.com/articles/s41592-018-0255-0>.

Yun, D.H. et al., 2019. Ultrafast immunostaining of organ-scale tissues for scalable proteomic phenotyping. *bioRxiv*.

Yushkevich, P.A. et al., 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3), pp.1116–1128.

# Chapter 7

## Vascularization of cerebral organoids using two-photon stereolithography

**Justin Swaney\***, Alexandre Albanese\*, Srinu Pujari\*, Nicholas Evans, Dae Hee Yun, Webster Guan, Kwanghun Chung<sup>1</sup>

### 7.1 Abstract

Organoids derived from induced pluripotent stem cells (iPSC) enable the production of patient-specific 3D tissues in vitro. Cerebral organoids mimic human brain development and can grow beyond 4 mm in diameter with intricate, self-assembled internal structures that resemble human corticogenesis. These large organoids require long-term culture in spinning bioreactors to maximize the transport of oxygen and nutrients, but cell death and hypoxia still occur away from the organoid surface. Fetal development resolves the restricted transport in large tissues by the integration of blood vessels in the first trimester in utero. To address this issue in vitro, we fabricated synthetic vasculature by high-resolution photopolymerization of polyethylene glycol (PEG) based monomers using two-photon stereolithography (SLA). Micro-vessels were 100 um in outer diameter, durable yet flexible, and permeable to bio-molecules in a tunable manner. Cerebral organoid culture for 30 days with perfusion was accomplished with a polydimethylsiloxane-based fluidic culture system and resulted in vascular network integration with no indication of material cytotoxicity. Future studies can use our printed vasculature to create and perturb new models of human tissue.

---

<sup>1</sup>\* indicates co-first authorship.

## 7.2 Introduction

Organoids are three-dimensional (3D) biological models wherein patient-derived induced pluripotent stem cells (iPSCs) or embryonic stem cells (ESCs) are differentiated into specific cell types (Fatehullah et al. 2016). These cells self-assemble due to cues from the provided extracellular matrix (ECM) material and culture media. Depending on the cues provided, various kinds of organoids can be created, including renal, intestinal, and neural (Takasato et al. 2015; Chen et al. 2017; Lancaster et al. 2013). To create a cerebral organoid, iPSCs or ESCs are formed into a small spheres called an embryoid bodies (EBs). The EBs are cultured in neural induction media and then seeded into matrigel to form cerebral organoids. Cerebral organoids are cultured for several weeks in a stirred bioreactor to allow for growth and maturation. After three months, cerebral organoids can reach diameters of up to 4 mm. Cerebral organoids offer several advantages over other biological models of the human brain. Being derived from stem cells, organoids can model human development *in vitro*. Cerebral organoids have been used to study developmental neurological disorders such as microcephaly, autism, and Rett syndrome (Mariani et al. 2015; Mellios et al. 2017). Since the iPSCs are patient-derived, organoids also offer personalized therapeutic potential in disease modeling and drug development (Lancaster & Knoblich 2014). Since organoids offer more realistic disease modeling, they could bridge the gap between primary drug screens and animal trials (Pampaloni et al. 2007). For this reason, pharmaceutical companies are investigating the use of organoid models for lead optimization. Despite these recent successes, neural organoid are still an incomplete model of the human brain.

One major shortcoming is the omission of a vascular system, which is needed to provide nutrients uniformly throughout large tissues. As a result, neural organoids are limited to approximately 4 mm in diameter as all nutrients must diffuse in from the organoid surface (McMurtrey 2014). As the nutrients diffuse into the organoid, they are simultaneously consumed via metabolic processes within the organoid. This nutrient reaction-diffusion is non-physiological, and it may hinder the ability for neural organoids to model healthy brain development (Raja et al. 2016). *In vivo*, nutrient delivery is solved by the vascular system, an incredibly sophisticated and dynamic transport system. Capillaries can be as small as 10  $\mu\text{m}$  in diameter, and the average vessel-to-neuron distance is only 20  $\mu\text{m}$  in the human brain (Murray et al. 2015). These characteristics make recreating the vascular system an extremely difficult engineering problem.

Vascularization of *in vitro* cultures has been studied extensively in tissue engineering, and both cell-based and synthetic approaches have been proposed (Lovett et al. 2009). Cell-based approaches such as extravasation rely on the ability of endothelial cells to form capillary networks when subjected to shear forces due to flow (Jeon et al. 2015). The re-

sulting capillaries are approximately 10  $\mu\text{m}$  in diameter, but the overall network geometry is not controlled. Synthetic methods such as 3D bio-printing extrude cell-laden inks into 3D vascular structures (Kolesky et al. 2016). In this case, the printed vascular network geometry is controlled, but the resulting capillaries are typically 500  $\mu\text{m}$  in diameter due to the fabrication resolution of the printer. In both types of vascularization, however, there is usually no characterization of the changes in transport phenomena involved with incorporating the vascular system. Instead, biological assays showing increased cell proliferation and metabolic activity are used as proxies for improvements in nutrient transport. Synthetic vascularization strategies such as 3D bio-printing offer no way of modulating the nutrient transport properties of the resulting vascular system, so measuring transport would only serve to highlight this lack of flexibility.

Two-photon stereolithography (SLA) offers much higher resolution than other 3D bio-printing methods (Sun & Kawata 2004; Kim & Lee 2010). Two-photon SLA relies on the non-linear dependence of the photon absorption rate on light intensity to achieve localized polymerization at the focal spot of a microscope objective. At the focal spot, photoinitiators are excited to their singlet state through two-photon absorption. The excited photoinitiators produce radicals which participate in an initiation reaction of a radical polymerization. By scanning the focal spot around in the focal plane, a single cross section of a 3D object can be created. A 3D object is generated by lifting the objective to stack the printed cross sections sequentially. Feature resolutions down to 100 nm have been reported using two-photon stereolithography (Sun et al. 2004). At such high resolution, microporous vessel walls could be fabricated directly, effectively decoupling the mechanical properties of the material from the overall transport properties of the vascular system. This decoupling of mechanical and transport concerns may allow materials that are well suited for stereolithography to still provide adequate nutrient transport into the cerebral organoid.

Here, the progress toward vascularized cerebral organoids using two-photon SLA is presented. A two-photon SLA printer was built and a biocompatible resin was developed for the fabrication of micro-vessels. The high fabrication resolution enabled direct printing of engineered pore structures to modulate effective transport properties with continuous perfusion from an external pump. Organoids with printed vasculature that were cultured with perfusion showed mature neurons and micro-vessel integration but hindered organoid growth when compared to controls. Future work may focus on creating a more robust long-term perfusion culture apparatus for the vascularized cerebral organoids with an integrated live-imaging system.

## 7.3 Results

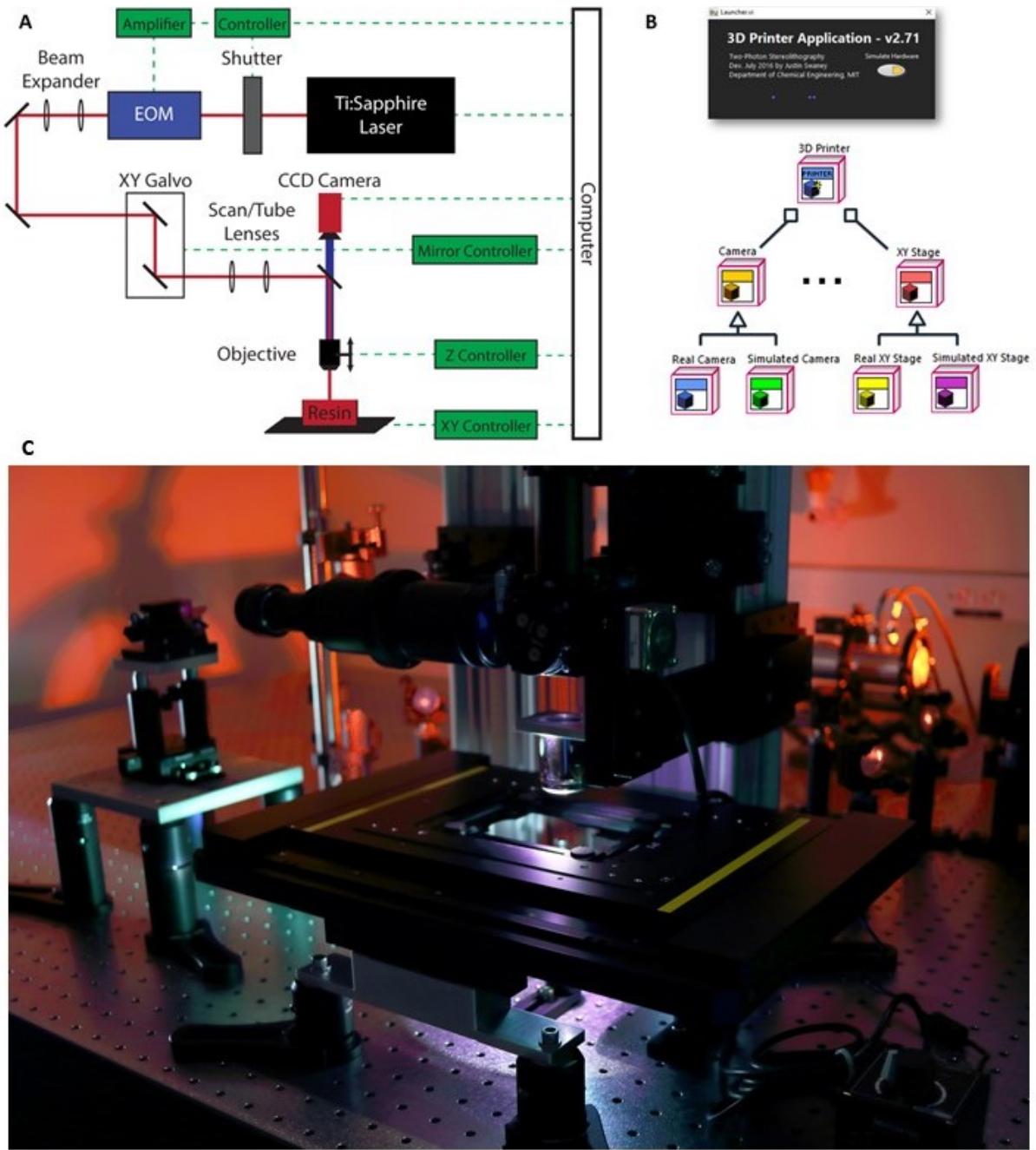
### 7.3.1 Design of a high-resolution two-photon SLA 3D printer

To fabricate synthetic vascular structures with micron-scale resolution, we designed and built a custom two-photon SLA 3D printer (Figure 7.1). Using similar optics to a confocal two-photon microscope, a Ti:Sapphire near-infrared (NIR) femtosecond pulsed laser beam is first directed through an electro-optic modulator (EOM) to control the laser intensity over time (Figure 7.1a). The control signal for the EOM is provided as an analog voltage signal from a control computer. Next, the beam is expanded to an appropriate size for a pair of 7 mm galvanometer (galvo) scanning mirrors. The mid-plane of the XY-galvos is placed conjugate to the scan lens to minimize the spot movement during scanning. By mounting the XY-galvos close together and positioning the focal plane of the scanning mirror directly between the mirrors, the resulting deflection in the projected beam is shared equally in the X and Y directions.

Then, a *4f* scanning system further expands the beam to slightly overfill the back aperture of the printing objective. The NIR laser is directed through the objective using a dichroic mirror with a 700 nm cutoff wavelength mounted in a filter cube. This dichroic mirror reflects the incoming NIR laser but allows light from an illumination LED mounted underneath the XY stage to pass through to a CCD camera. The camera is mounted directly above the objective with a tube lens and chromatic filter to protect the CCD from the NIR laser. Since the deflection of the scanned spot is small, the z-position of the objective can be adjusted without affecting the laser power delivered into the resin. To accomplish this z-positioning, the objective is mounted on a vertically-oriented motorized stage with a 25 mm travel range. The laser is finally focused at the working distance of the objective, which is approximately 8 mm away from the final lens.

Since the two-photon absorption process can excite only a small point in the resin in the focal plane of the objective, the resin within the 8 mm distance between the lens and the focal plane is unaffected by the NIR laser. This simplifies any surface effects encountered in UV-based photo-polymerization due to laser bleed-through and oxygen inhibition.

In order to control the hardware in the two-photon SLA system, a custom LabVIEW program was developed (Figure 7.1b). The control software was implemented using object-oriented design patterns and is based on the actor framework, which allows each physical piece of hardware to be controlled asynchronously through a message passing interface. By creating an abstract base class for each actor in the two-photon SLA system, interfaces for each constituent hardware component were defined. Using this strategy pattern allowed for plug-and-play hardware compatibility while maintaining the asynchronous

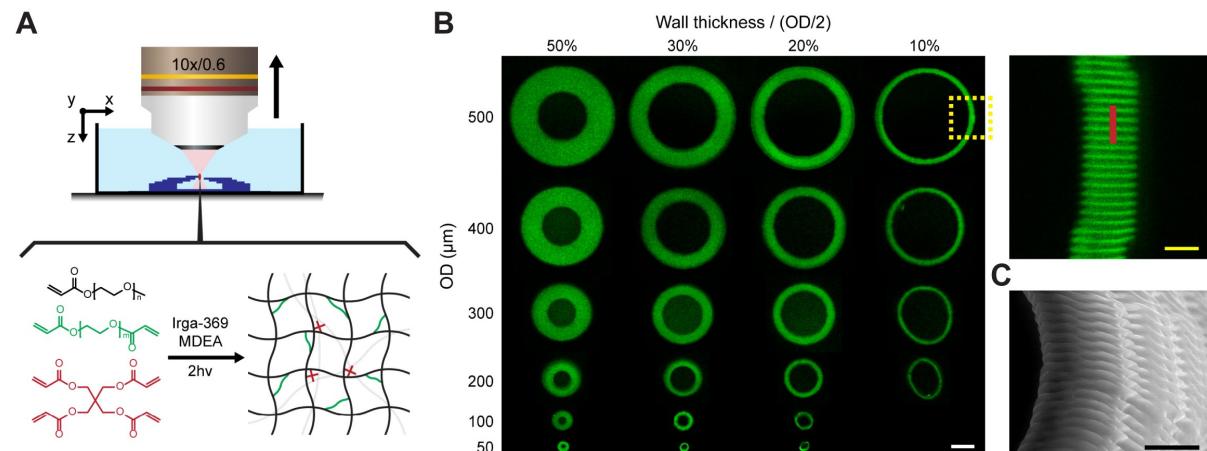


**Figure 7.1:** Construction of a high-resolution two-photon SLA printer. (A) Diagram of laser beam path in the two-photon SLA system. Similar to a confocal point-scanning microscope, galvanometers (galvos) deflect the laser in the XY plane of the objective. A high-frequency electro-optic modulator (EOM) modulates the laser power during galvo scanning. (B) Diagram of the high-level object-oriented software architecture for the two-photon SLA system. (C) Photograph showing the physical two-photon SLA system.

nature of the actor framework. It also allowed for software development without physical hardware connected to the computer because individual hardware actors could be mocked by implementing the interfaces in the hardware abstraction layer using trivial methods.

### 7.3.2 Two-photon SLA enables microfabrication of biocompatible materials

Since a pulsed near-infrared (NIR) laser is required for two-photon absorption, a series of commercially available photoinititators were screened for compatibility with this light source. We found that isopropylthioxanthone (ITX) and Irgacure 369 were both excitable under NIR irradiation and triggered photopolymerization of neat poly(ethylene glycol) diacrylate (PEGDA). When used together, ITX and Irgacure 369 show higher polymerization rates than they do individually due to the effect that ITX has as a sensitizer for Irgacure 369. We also found that adding N-methyldiethanolamin (MDEA) as an amine synergist further increased the rate of polymerization, presumably due to its reactivity with Norrish type II photoinititators such as excited ITX as well as its ability to scavenge singlet oxygen. The most effective two-photon compatible photoinitiation system contained 0.1% ITX, 0.5% Irgacure 369, and 4% MDEA (Figure 7.2a).



**Figure 7.2:** Two-photon SLA of micro-vessels from biocompatible materials. A) Diagram of the print process and polymerization chemistry at the focal spot. (B) Cross sections of printed micro-vessels of various wall thicknesses and diameters. (C) ESEM image of printed micro-vessel walls. Scale bar, 20  $\mu\text{m}$ .

Using this two-photon optimized photoinitiation system, we developed a resin for 3D printing biocompatible materials with micron resolution. We found that adding pentaeryritol tetra-acrylate (PETA) as a crosslinker greatly improved the mechanical stability of our material, but it also led to clogging of printed channels. Since clogging was due to over-polymerization in non-irradiated regions, we added 0.2% 4-methoxyphenol (MEHQ) as a resin stabilizer and polymerization inhibitor. This low concentration of MEHQ was not

enough to completely stop photopolymerization upon NIR irradiation, but it was effective at preventing unwanted polymer growth away from the focal spot during printing. A monomer formulation optimized for 3D printing microfluidic structures containing 10% PETA and 90% PEGDA575 was used in the following experiments.

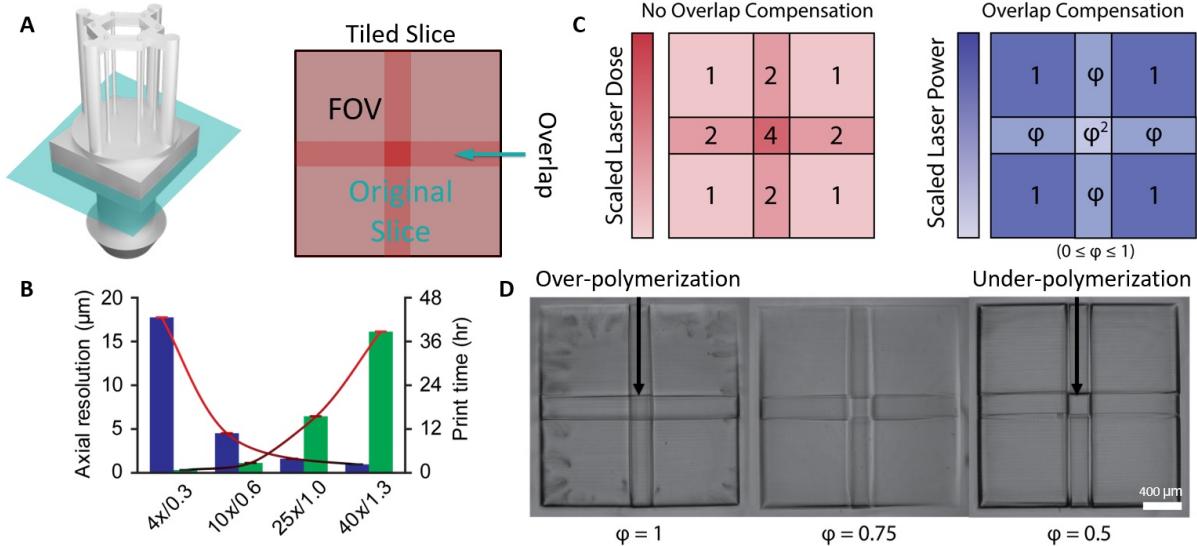
To show micron-resolution printing of unclogged tube structures, we printed a series of fluorescent tubes with a variety of diameters and wall thicknesses using a 10x/0.6 objective. The smallest tubes that we printed were 50  $\mu\text{m}$  in outer diameter (Figure 7.2b). A line scanning pattern from the print process was observable in the walls of the printed tubes, suggesting that some microstructure existed due to the fabrication process. To observe this microstructure in more detail, these printed tubes were imaged using environmental scanning electron microscopy (ESEM). The ESEM images revealed a similar line scanning pattern stacked into layers (Figure 7.2c). Each line of polymer was approximately 4  $\mu\text{m}$  thick and 20  $\mu\text{m}$  tall, corresponding to the fabrication resolution of the printer and resin. This fabrication resolution is approximately two orders of magnitude smaller than the fabrication resolution of 3D bio-printing.

### 7.3.3 Printing large-scale vascular systems by SLA tiling

With a 10x objective, the effective field-of-view (FOV) is approximately 1 mm wide. In order to print 3D models that are wider than a single FOV, a tiling strategy is required (Figure 7.3a). To ensure that adjacent tiles are covalently linked, these print tiles must overlap slightly. Furthermore, since arbitrary 3D geometries could be given to the printer, all XY tiles in a Z slice must be completed before moving on to the next Z slice. However, empty tiles and even blank lines can be skipped to speed up the overall print times.

Using higher magnification objectives would further improve the fabrication resolution, but it would make printing larger structures prohibitively slow (Figure 7.3b). For example, although a 40x/1.3 NA objective would have a theoretical axial resolution near 1  $\mu\text{m}$ , the number of XY tiles increases quadratically with the reduced FOV width. Conversely, although a 4x/0.3 NA objective would reduce the number of XY tiles needed to print larger vascular structures, the axial resolution would be much worse with such a low numerical aperture. For these reasons, we chose to use a 10x/0.6 NA objective for fabrication.

Simply overlapping adjacent tiles is an acceptable strategy for stitching volumetric images together since fluorophores simply emit photons in response to excitation. However, photo-initiators produce radial species that trigger local polymerization, which is an irreversible chemical change in the resin. If overlapping regions are repeatedly subjected to irradiation, then over-polymerization in those regions may cause internal stresses and



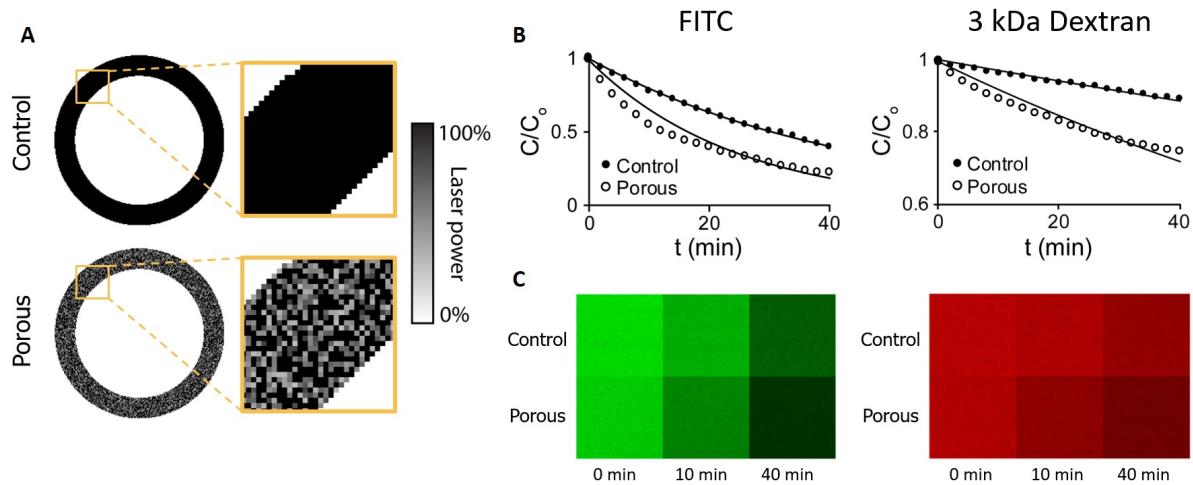
**Figure 7.3:** Tiling strategy for printing large-scale vascular systems. (A) Printing a 3D model wider than a single field-of-view (FOV) requires tiling each XY slice with overlap. (B) Bar graph showing the trade off between objective magnification and print time. Higher magnification objectives typically have higher optical resolutions but smaller FOVs, requiring more XY tiling to print the same 3D model. (C, D) Diagram and images of overlap compensation using laser power scaling. Overlapping regions over-polymerize due to repeated laser exposure. By reducing the laser point in the overlapping regions using the EOM, adjacent tiles can be effectively fused.

mechanical failures in the resulting printed structure. To compensate for this repeated irradiation in overlapping regions, we use an overlap compensation strategy where the laser power is reduced using the EOM in the overlapping regions by a reduction factor  $\phi$  each time a voxel is irradiated more than once (Figure 7.3c). Through trial and error, the optimal setting for  $\phi$  was found to be 0.75 for our optical system and resin (Figure 7.3d).

### 7.3.4 Direct printing of engineered pore structures allows modulation of transport properties

After characterizing the fabrication resolution of our printer, we tried to directly fabricate porous structures. To this end, we created porous and control 3D models for cylindrical reservoirs with the same dimensions and wall thickness (Figure 7.4a). The porous model was created by using random laser powers in the reservoir wall. After filling these reservoirs with 5 µL of FITC or 3 kDa-dextran solution and surrounding them with water, the pseudosteady diffusion of the fluorescent molecules from the reservoir was monitored using a fluorescence microscope (Figure 7.4b,c). The fluorescence intensity within the porous reservoir decreased faster than the control reservoir for both FITC and the 3 kDa-dextran. This demonstrates that the fabrication resolution of our two-photon SLA printer is sufficient to directly print porous structures and modulate transport properties

of printed materials.

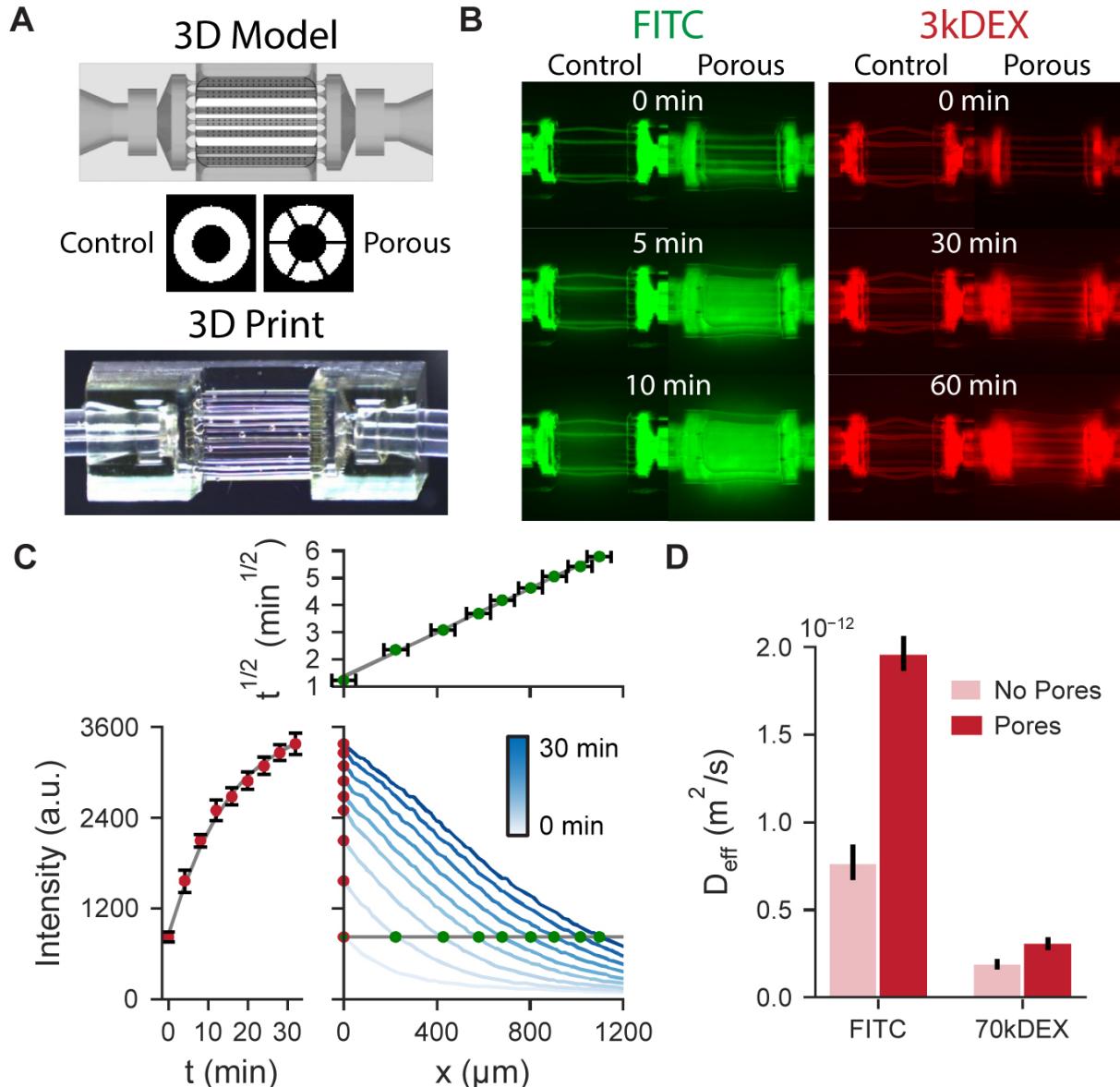


**Figure 7.4:** Direct printing of microporous structures improves mass transport from printed reservoirs. (A) Optical sections of control and porous 3D models of cylindrical media reservoirs. The porous reservoir has a subset of voxels that were exposed to random laser powers between 0% and 100%. (B) Normalized fluorescence intensity of FITC and 3kDa-Dextran conjugated with Texas Red inside the reservoirs over time when surrounded by water. (C) Example images from different time points in (B).

Printing pore structures directly may allow tuning of the transport properties independently from the resin formulation. In order to monitor transport of solutes from printed vessels with perfusion, inlet and outlet connections to an external pump are required. To this end, we designed and printed compression fittings that interface the printed vasculature with 1 mm soft silicone tubing (Figure 7.5a). The compression fittings require minimal force to insert tubing, and the tubing is soft so that the printed micro-vessels are not subjected to large stresses when connected from both sides. Also, the optimized resin produces a material that is rigid enough to stand upright during fabrication, but also flexible enough to accommodate the strain associated with direct tubing connections.

We compared the transport of FITC and a 3kDa dextran from micro-vessels with and without straight pores using time-lapse fluorescence imaging (Figure 7.5b). Each pore was 4  $\mu\text{m}$  wide and 20  $\mu\text{m}$  tall oriented radially in the micro-vessel walls. Pores were included at 18 equally spaced axial locations along the micro-vessels, which were 2.5 mm long (150  $\mu\text{m}$  axial spacing). From the time-lapse images, radial concentration profiles were obtained and used to calculate the effective diffusion coefficients in the micro-vessel wall as well as in the external matrigel (Figure 7.5c). For both FITC and 3kDEX, a diffusion front was observed with a time-dependent boundary condition at the micro-vessel surface. The boundary condition concentration is related to lumen concentration through the conservation equation in the vessel wall, which is a porous medium. We approximated the time scale of diffusion in this porous medium by fitting a first order model to the observed boundary condition. A more elaborate model of the micro-vessel wall would be insightful if more pore structures were tested. However, we found this

simple model to be adequate for quantifying the improved transport properties measured in the micro-vessel wall.

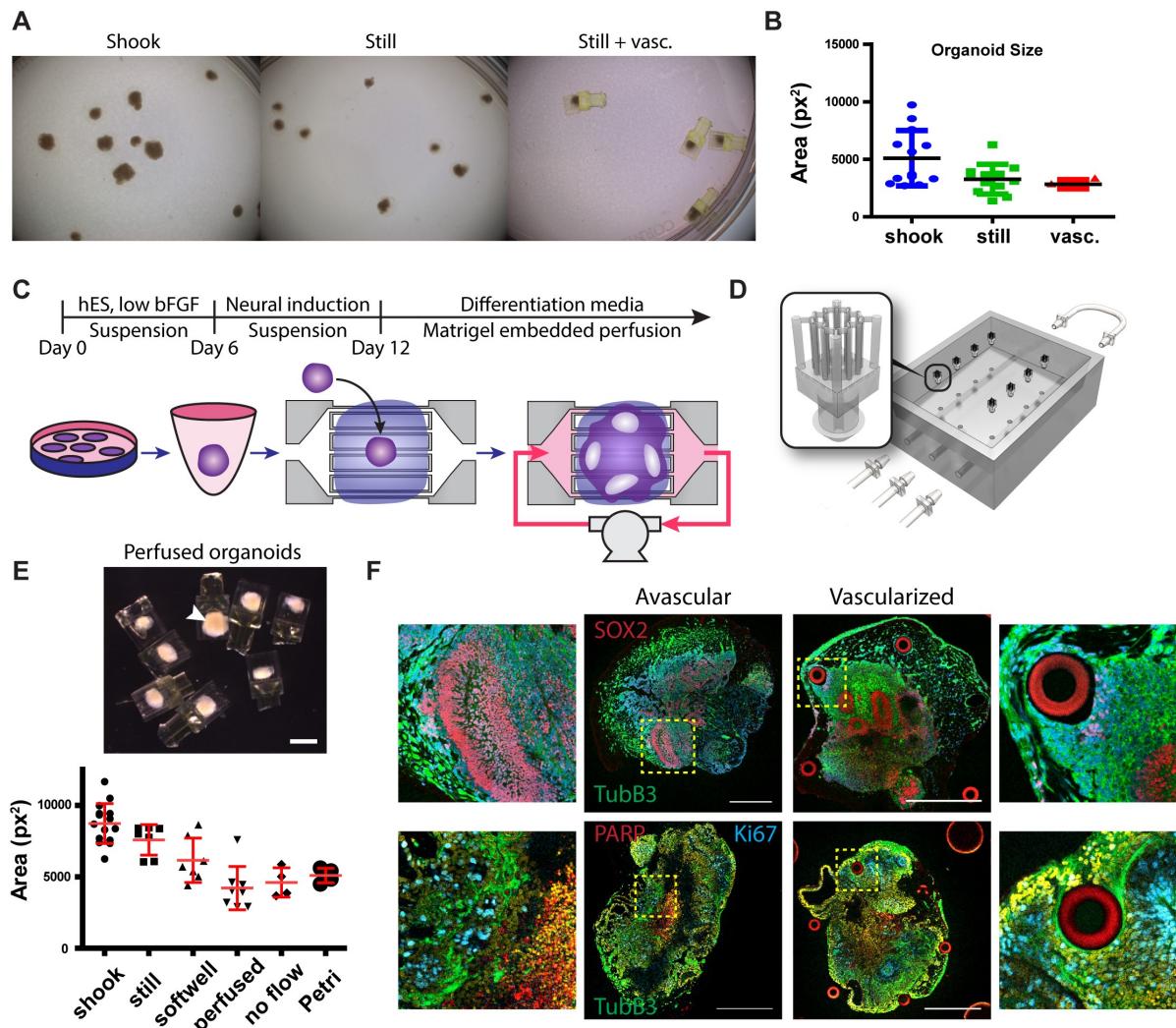


**Figure 7.5:** Modulation of microfluidic transport properties by direct fabrication of engineered pore structures. A) 3D model of transport system with inlet and outlet compression fittings. B) Image of printed transport system. C) Time-lapse images of FITC and 3kDa Dextrans diffusion from the transport system with 100  $\mu\text{L}/\text{min}$  perfusion. D) Quantification of transport properties within the micro-vessel wall with and without printed pores.

### 7.3.5 Integration of printed vasculature shows limited cytotoxicity

After showing that our printed vasculature could deliver biomolecules with tunable transport properties, we wanted to test the compatibility of our printed vasculature with neural organoid culture even without media perfusion. In order to minimize deviations from standard neural organoid culture protocols, we chose to embed the organoid with

the printed vasculature at the time the EB is usually embedded into matrigel. Instead of embedding into a matrigel droplet, the EB was first loaded onto the device in a designated opening, and then matrigel was added around the EB and vasculature. The design of the designated opening for EB embedding is nontrivial since capillary forces dominate how the liquid containing the EB will behave at the millimeter scale. Another complication is that by embedding at the EB stage, we preclude our ability to integrate any micro-vessels into the core of the organoid. This problem could be addressed by embedding the vasculature into the organoid at an earlier stage, but this further deviates the culturing protocol from avascular control organoids. After embedding, the vascularized organoids were cultured in a petri dish without perfusion and compared to a vascular controls with and without shaking.



**Figure 7.6:** Integration of vasculature into cerebral organoids shows mature neurons and micro-vessel integration. A) Schematic of vascularized organoid culture with perfusion. B) Render of softwell culture devices and vasculature equipped with barb fittings. C) Vascularized and avascular organoids grown with and without shaking. D) Quantification of organoid size with and without shaking. (E) Immunostaining of PARP, TubB3, AcTub, and Ki-67 in vascularized organoids.

After culturing for 30 days, we observed a decrease in the size of the vascularized organoids

as well as the controls cultured without shaking (Figure 7.6a,b). This is in agreement with the hypothesis that organoid culture is transport limited, since shaking can influence convective mass transport at the surface of the organoid. The vascularized organoids were not significantly smaller than the controls cultured without shaking, indicating the material may not what is causing the restricted growth. A closer look at the control organoids cultured with shaking revealed a bimodal distribution of organoid sizes, where some organoids were as small as those cultured without shaking and other were much larger. This suggests that some heterogeneity between neural organoids may be due to inconsistent nutrient transport during culture.

### 7.3.6 Perfusion culture of vascularized organoids showed limited apoptosis and mature neurons

Next we wanted to culture our vascularized organoids with media perfusion in an attempt to rescue the restricted growth without shaking. We developed a series of culture devices called softwells made from polydimethylsiloxane (PDMS). These devices were equipped with a single inlet and outlet for media circulation through multiple vasculature devices in parallel (Figure 7.6c,d). For long-term perfusion culture, vasculature devices were printed with barb fittings designed to connect to the PDMS directly through a 2 mm punched hole. After embedding, the vascularized organoids were placed into an incubator alongside an external pump for media recirculation. All experiments used a flow rate of 25  $\mu\text{L}/\text{min}$  per organoid.

We compared our vascularized organoids to avascular controls cultured in petri dishes as well as in softwells (Figure 7.6e). This experiment was designed to ensure that the softwell itself was nontoxic to the organoids. After a month of culture, however, a similar size analysis revealed a significant drop in avascular organoid size when cultured in the softwell, indicating that the PDMS may have interfered with the organoid culture. The perfused organoids were significantly smaller than the control organoids cultured without shaking in a petri dish. However, the perfused organoids were not significantly smaller than the avascular organoids cultured within the same softwell. Upon inspection of the perfused organoids, one large outlier exists, suggesting that the perfused organoids may have not received equal media perfusion. Together, these results suggest that the PDMS softwell restricts organoid growth and that there may be some issue with the softwell devices in delivering equal media to all organoids. We are currently working to remove PDMS from the culture system and instead use a single 3D printed vascular device to perfuse multiple organoids. If unequal perfusion is not an issue in the softwell devices, then material passivation is an option to help rescue organoid growth.

We then stained the vascularized organoids for markers of apoptosis (PARP), proliferation (Ki-67), and mature neurons (TubB3) to verify that the neural organoid differentiated properly (Figure 7.6f). Vascularized organoids showed characteristic ventricular zones and positive expression of mature neural markers, such as tubulin-beta3. The micro-vessels were surrounded by cells and neuronal projections, and no increase in apoptosis was observed near the micro-vessels as measured by PARP expression. Instead, most apoptotic cells were near the core of the organoids and away from the micro-vessels, consistent with the hypothesis that there is insufficient nutrient delivery away from the organoid surface. This suggests that the printed material and vasculature geometry is not preventing the expected differentiation and maturation within vascularized neural organoids.

## 7.4 Discussion

Several commercial 3D printers use vector rather than raster scanning for fabrication. Raster scanning was used in our two-photon SLA system to facilitate the tiling strategy. With raster scanning, the entire 3D object can be represented as an array data structure. Array data structures are easily indexed, which enables partitioning of the 3D object into smaller subarrays for each tile. In vector scanning, 3D objects are typically represented using G-code, which encodes the fabrication process in terms of parameterized tool paths. However, this data structure is not easily partitioned into parametrized tool paths within subregions. The entire G-code parameterization could be used by scanning the XY stage, but this process would be very slow relative to galvo-based laser scanning.

The representation of 3D objects as arrays is also helpful in printing engineered pore structures. The array data structure allows for modulation of the laser power for each voxel, each of which is just 4  $\mu\text{m}$  wide. However, turning the laser power down randomly inadvertently leads to less overall polymerization and weaker structures. Instead, pores can be printed through the entirety of the wall deliberately as in Figure 7.5. Such micro-pores would allow the overall object to remain mostly polymerized, but reveal openings through which biomolecules may easily diffuse. An interesting implication of having these micro-pores is the fact that they would be of similar size to the organoid cells. It may be possible for cells to migrate from the vasculature into the organoid. This could be used to model angiogenesis, introduce microglia, and study metastasis all within the exciting *in vitro* framework of cerebral organoids.

The microfluidic interface and culture system is what ultimately acts as an interface between the macroscopic and microscopic length scales. There are approximately four orders of magnitude between the fabrication resolution of our printer (4  $\mu\text{m}$ ) and the

overall size of the tools used to create organoids (1-10 cm). When faced with a similar problem in microfluidic applications, researchers typically use PDMS. However, there is still room for improvement in the integration of these conventional microfluidics with culture tools that enable prolonged perfusion culture. Specifically the culture systems needs to be pared down to simplify the interpretation of experimental results. Various passivation methods and quality control steps should be tested in attempt to increase overall organoid growth during perfusion culture.

Vascularization of cerebral organoids has been demonstrated using two-photon SLA to print biocompatible micro-vessels that are 100  $\mu\text{m}$  in diameter. The fabrication resolution of 4  $\mu\text{m}$  enabled direct printing of engineered pore structures to increase the permeability of the printed material to large bio-molecules. Prolonged culture with media perfusion for one month was accomplished with a PDMS-based fluidic culture system and resulted in vascular network integration. More work needs to be done to increase organoid growth in perfusion culture as well as in characterizing the vascularized organoids to avascular controls. Future work may also adapt the printed vascular systems with engineered pore structures in live-imaging and co-culture organoid systems.

## 7.5 Methods

### 7.5.1 Construction of two-photon SLA printer

A Ti:Sapphire laser (Vision-S, Coherent) was used to deliver 75 fs pulses of 800 nm light at a repetition rate of 80 MHz. The beam was attenuated with an EOM (M302RM, ConOptics) and then enlarged with a Keplerian beam expander. Following a periscope, the beam entered the pupil of a two-axis galvanometer system (6215H 7 mm, Cambridge Technology) for raster scanning. The focal plane of the scan lens (LSM04-BB, Thor Labs) was positioned between the galvanometers in order to minimize the vignetting effect during scanning. A tube lens relayed the scanning beam to the objective back aperture. The objective was mounted on a single-axis stage with 1" travel range (MP-285, Sutter Instrument). The sample was placed on a two-axis nanopositioning piezoelectric stage with a 135 mm x 85 mm travel range (M-687.UN, Physik Instrumente). A 700 nm shortpass dichoric mirror and laser filter were mounted before a CCD inspection camera (Prosilica GE1050, Allied). An LED with condensing optics and variable power LED driver (M625L3-C1 and LEDD1B, Thor Labs) was positioned beneath the sample stage for brightfield illumination.

The system integration was implemented in LabVIEW using the Actor Framework (AF).

Each piece of hardware was defined as a LabVIEW class which inherits from the abstract actor class. Each actor inherits the functionality of a queue-driven state machine where methods are called using a message-passing paradigm between actors. This allowed for asynchronous control of individual subsystems as well as coordination of several pieces of hardware simultaneously. It also enabled true parallel execution of each actor rather than simple context switching between threads. Since the AF utilizes object-oriented programming, a hardware abstraction layer was defined to allow for a more dynamic runtime behavior. This made the program more accommodating to future hardware changes, and it allowed for the use of dummy hardware actors for debugging or when the physical hardware was not connected to the controlling computer.

### 7.5.2 Measuring the fabrication resolution with ESEM

A 6 x 4 array of 300  $\mu\text{m}$  tall micro-vessels were printed together on a printed 400  $\mu\text{m}$  thick substrate. The micro-vessels were printed with 0.01% FITC acrylate in order to visualize the formed polymer with confocal microscopy. The printed micro-vessels were placed in the chamber of an ESEM and imaged under 50% humidity to prevent drying. The resulting images were analyzed in ImageJ to determine the fabrication resolution.

### 7.5.3 Rapid prototyping of PDMS molds for creating microfluidic culture systems

3D models for the PDMS molds were designed in SOLIDWORKS 2016. These models were fabricated on an Ultimaker 3 printer out of polylactic acid (PLA) and used ejector pins for injection molding as channel molds (McMaster). Sylgard 184 (Dow) mixed at a 15:1 curing ratio was poured directly into the mold. The PDMS prepolymer was degassed for 30 minutes before curing at 60C overnight. Holes for vasculature connections were punched using 2 mm disposable biopsy punches (Integra 33-31-P/25).

### 7.5.4 Perfusion culture of vascularized neural organoids

The microfluidic culture system was connected to silicone tubing and purged with culture media. The printed vasculatures were loaded into the microfluidic culture system, and then an EB was pipetted into each of the printed vasculatures. Matrigel was added to the EBs positioned within the printed vasculatures, and the microfluidic culture system was connected to a peristaltic pump (Ismatec 4 channel Reglo ICC) inside an incubator with 5% CO<sub>2</sub>. The vascularized organoids were cultured for 30 days using media from the

Lancaster 2013 protocol with media perfusion at 25  $\mu$ L/min per organoid. The culture media was replaced every 5 days.

## 7.6 References

- Chen, Y.-w. et al., 2017. A three-dimensional model of human lung development and disease from pluripotent stem cells., 19(5).
- Fatehullah, A., Hui Tan, S. & Barker, N., 2016. Organoids as an in vitro model of human development and disease. *Nature Publishing Group*, 18.
- Jeon, J.S. et al., 2015. Human 3D vascularized organotypic microfluidic assays to study breast cancer cell extravasation., 112(7).
- Kim, R.H. & Lee, K.-S., 2010. 3D Stereolithography by Using Two-Photon Photopolymerization. *Macromolecular Symposia*, 298(1), pp.25–33. Available at: <http://doi.wiley.com/10.1002/masy.201000026>.
- Kolesky, D.B. et al., 2016. Three-dimensional bioprinting of thick vascularized tissues.
- Lancaster, M.A. & Knoblich, J.A., 2014. Generation of cerebral organoids from human pluripotent stem cells. *Nature Protocols*, 9(10), pp.2329–2340.
- Lancaster, M.A. et al., 2013. Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467), pp.373–379.
- Lovett, M. et al., 2009. Vascularization Strategies for Tissue Engineering., 15(3).
- Mariani, J. et al., 2015. FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell*, 162(2), pp.375–390.
- McMurtrey, R.J., 2014. Analytic Models of Oxygen and Nutrient Diffusion, Metabolism Dynamics, and Architecture Optimization in 3D tissue Constructs with Applications and Insights in Cerebral Organoids. *Igarss 2014*, 22(1), pp.1–5. Available at: <http://arxiv.org/abs/arXiv:1011.1669v3>.
- Mellios, N. et al., 2017. MeCP2-regulated miRNAs control early human neurogenesis through differential effects on ERK and AKT signaling., (July 2016), pp.1–15.
- Murray, E. et al., 2015. Simple, Scalable Proteomic Imaging for High-Dimensional Profiling of Intact Systems. *Cell*, 163(6), pp.1500–14.
- Pampaloni, F., Reynaud, E.G. & Stelzer, E.H.K., 2007. The third dimension bridges the gap between cell culture and live tissue. *Nature Reviews Molecular Cell Biology*, 8(10), pp.839–845. Available at: <http://www.nature.com/articles/nrm2236>.
- Raja, W.K. et al., 2016. Self-Organizing 3D Human Neural Tissue Derived from Induced Pluripotent Stem Cells Recapitulate Alzheimer’s Disease Phenotypes., pp.1–18.

Sun, H.B. & Kawata, S., 2004. Two-photon photopolymerization and 3D lithographic microfabrication. *Advances in Polymer Science*, 170, pp.169–273.

Takasato, M. et al., 2015. Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature*, 526(7574), pp.564–568. Available at: <http://www.nature.com/articles/nature15695>.

# Chapter 8

## Conclusion

### 8.1 Thesis summary

This thesis addresses several major challenges associated with scaling up experimental techniques used in neuroscience research. These advances enable neuroscientists to answer long-standing biological questions in new ways by lowering the barrier for staining and imaging, providing scalable image processing tools for analysis, and engineering new culture systems for the study of large-scale brain models.

By building from tissue clearing technologies such as CLARITY, we developed the SWITCH method for controlling reaction kinetics of fixative molecules and molecular probes when processing large tissue samples. The robust GA fixation used in SWITCH also enabled over 20 rounds of antibody staining within the same human clinical brain sample. Then, resolution limitations associated with conventional fluorescence microscopy techniques were addressed by developing MAP, a technique for physical and tunable expansion of hydrogel-embedded whole-brain scale tissue samples. MAP allows synaptic resolution imaging with confocal microscopy and is compatible with long working distance objectives for imaging thick tissue samples.

To address the computational challenges associated with whole-brain imaging, we present a scalable image processing pipeline for extracting meaningful biological information from such whole-brain volumetric images and use this pipeline to study mouse and marmoset tissue at single-cell resolution. An efficient algorithm for nonrigid coregistration of whole-brain scale volumetric images based on nuclei matching is also presented. The nuclei-based 3D coregistration algorithm is shown to be compatible with mouse and marmoset models and allows warping of an entire mouse brain hemisphere within 6 hours.

Building on these computational analysis tools, a multiscale phenotyping pipeline called

SCOUT is presented for understanding single-cell, cytoarchitectural, and whole-tissue scale changes in cerebral organoids models. An automatic 3D ventricle segmentation based solely on nuclear staining using deep learning is presented. Using SCOUT, the multiscale phenotypic changes associated with maturation, Zika virus infection, and culture protocol variations were quantified. In this analysis, the distance from the organoid surface was repeatedly identified as an important factor in predicting orgnaoid cytoarchitecture. This supports the hypothesis that cerebral organoids are affected by nutrient transport limitations.

To address these nutrient transport limitations, an orgnaoid vascularization strategy is presented using two-photon stereolithography to directly print biocompatible and porous vascular systems. The two-photon SLA printer achieved 4  $\mu\text{m}$  lateral fabrication resolution and was able to directly print engineered pore structures. The porous vascular systems showed improved effective nutrient transport with and without perfusion. Printed vascular systems were successfully integrated into cerebral organoids with limited cytotoxicity and undisrupted neuronal differentiation. This suggests that the synthetic material is biocompatible and suitable for organoid culture.

## 8.2 Future work

The technological contributions of this thesis may enable many future biological applications. These future applications may draw from individual aims of the thesis or collectively depending on the context. For example, neuroscientists interested in visualizing brain-wide protein expression with antibody staining may adopt the SWITCH technique. If quantification of these volumetric images is required for comparative studies, then SWITCH may be used in combination with the scalable image processing pipeline. By combining SWITCH and our scalable image processing pipeline, neuroscientists will be able to study their brain models quantitatively without dissecting predetermined brain regions, which may help to understand the brain as a complete biological system.

Using MAP, neuroscientists will be able to more easily quantify synaptic details using confocal microscopy. Several labs are currently using MAP for such analyses already. Interestingly, MAP offers tunable tissue expansion, which may be used to adjust the effective imaging resolution as needed in large tissues. There is an inherent trade off between the size of the tissue and the image data size, so the ability to modulate the tissue size as needed may prove valuable when applying MAP to larger tissues where the final dataset size may become intractable.

The single-cell 3D coregistration algorithm presented in this thesis may enable multi-

plexed fluorescence imaging of a single whole-brain sample. Such a detailed description of a single brain would be invaluable to understanding the multiscale intrinsic complexity of the brain. By combining SWITCH, our scalable image processing pipeline, and the nuclei-based 3D coregistration algorithm, a multiround staining and imaging approach may provide the most complete image of protein expression within a single tissue to date. In the same way that multimodal brain atlases are valuable tools to the neuroscience community, such a hyperdimensional volumetric image of the entire brain would be an asset to future studies. Finally, the coregistration algorithm presented here may be an indespensible tool to extract as much phenotypic information when dealing with precious samples such as human clinical brain samples.

Future work based on SCOUT may continue to utilize deep learning and high-throughput imaging technologies to deliver on the promise of organoids as powerful new in vitro biological models. Specifically, live-imaging modalities combined with real-time ventricle segmentation may elucidate the mechanisms by which cerebral organoids form ventricles and neural rosettes. This detailed view of cerebral orgnaoids with immediate phenotypic feedback could enable rapid protocol optimization as well as detailed perturbation analysis with drugs and morphogens.

The organoid vascularization strategy demonstrated here may serve as a proof-of-concept for future studies that utilize two-photon stereolithography for 3D printing biocompatible culture devices. The high resolution offered by two-photon SLA allows direct printing of porous materials, which is valuable in more contexts than just neuroscience research. However, these engineered pore structures may allow new models of the blood-brain barrier to be developed with separate media for the lumen and extracellular space. A natural next step for the printed vasculature is to combine vascularization with orgnaoid “assembloid” approaches, potentially giving rise to larger functional tissue cultures in vitro. By delivering nutrients, drugs, or biological material into 3D tissue cultures, the printed vasculature may provide researchers new ways to support, monitor, and perturb future in vitro brain models.

# **Appendix 1: Curriculum Vitae**

# Justin M. Swaney

jswaney@mit.edu  
jmswaney.com

[linkedin.com/in/jmswaney](https://www.linkedin.com/in/jmswaney)

## EDUCATION

---

<b>Massachusetts Institute of Technology</b> , Cambridge, MA <i>PhD in Chemical Engineering, Minor in Computer Science</i> <i>Thesis title: Scaling Up 3D Imaging, Analysis, and Culture of Complex Brain Models</i>	December 2019 ( <i>expected</i> )
<b>University of Wisconsin-Madison</b> , Madison, WI <i>Bachelor of Science in Chemical Engineering, Certificate in Mathematics</i>	May 2014 GPA: 3.99/4.0

## RESEARCH EXPERIENCE

---

<b>Samply LLC</b> , Walworth, WI <i>Founder and CEO</i>	May 2019 – Present
<b>Kwanghun Chung Lab, MIT</b> , Cambridge, MA <i>Graduate Researcher</i>	September 2014 – Present
<ul style="list-style-type: none"><li>Led development of a visual sample library explorer for music producers based on a novel deep audio embedding network trained by semi-supervised sample classification in Tensorflow.</li><li>Performed demonstrations and poster presentations of Samply at the MIT Computing Exposition.</li></ul>	

<b>Regina Murphy Lab, UW-Madison</b> , Madison, WI <i>Undergraduate Researcher</i>	June 2013 – May 2014
<ul style="list-style-type: none"><li>Applied Gustafson-Kessel Fuzzy clustering to the tracking analysis of amyloid beta nanoparticles to quantify aggregation dynamics in Alzheimer's disease.</li></ul>	

## TEACHING EXPERIENCE

---

<b>Academy for Advanced Research and Development</b> , Cambridge, MA <i>Assistant Professor (Part-time)</i>	March 2017 – Present
<ul style="list-style-type: none"><li>Supervised STEM-oriented high-school and undergraduate students in individual research projects in data science and machine learning.</li><li>Directed summer courses on self-driving cars using Raspberry Pi's and on applications of data science in healthcare including tumor segmentation in MRI images using Python.</li><li>Provided editorial assistance for student publications and recommendations for college applications.</li></ul>	

<b>UW-Madison Chemical Engineering Department</b> , Madison, WI <i>Teaching Assistant (Part-time)</i>	December 2013 – June 2014
<ul style="list-style-type: none"><li>Taught MATLAB by live coding solutions to example problems in the context of chemical engineering.</li><li>Demonstrated and explained the algorithms behind algebraic and differential equation solvers as well as non-linear optimization tools in MATLAB.</li></ul>	

## SKILLS and TECHNIQUES

---

Python	Git	Docker	Flask / Django
MATLAB	Linux / Unix	TDD / Travis CI	iPSC & organoid culture
LabVIEW / DAQmx	Tensorflow / Pytorch	SLURM	Immunohistochemistry
Javascript	Closure(script)	Accumulo / MySQL	Light-sheet microscopy
C / C++	AWS / Google Cloud	Spark	CLARITY

## PUBLICATIONS

---

- Swaney, J.\***, Pujari, S.\* , Albanese, A.\* , Yun, D.H., Evans, N.B., Chung, K. "Vascularization of neural organoids using two-photon stereolithography", (*in preparation*).
- Albanese, A.\* , **Swaney, J.\***, Yun, D.H., Evans, N.B., Antonucci-Johnson, J., Gehrke, L., Chung, K. "3D Imaging and High Content Morphological Analysis of Intact Human Cerebral Organoids", (*in preparation*).
- Yun, D.H.\* , Park, Y.G.\* , Cho, J.H.\* , Kamentsky, L., Evans, N.B., Albanese, A., Xie, K., **Swaney, J.**, Sohn, C.H., Tian, Y., Zhang, Q., Drummond, G., Guan, W., DiNapoli, N., Choi, H., Jung, H.Y., Ruelas, L., Feng, G., Chung, K. "Ultrafast immunostaining of organ-scale tissues for scalable proteomic phenotyping", *bioRxiv* (submitted to *Nature Biotechnology*), 2019, doi:10.1101/660373.
- Swaney, J.\***, Kamentsky, L.\* , Evans, N.B., Xie, K., Park, Y.G., Drummond, G., Yun, D.H., Chung, K. (2019) "Scalable image processing techniques for quantitative analysis of volumetric biological images from light-sheet microscopy", *bioRxiv*, doi:10.1101/576595.
- Ku, T.\* , **Swaney, J.\***, Park, J.Y.\* , Albanese, A., Murray, E., Cho, J.H., Park, Y.G., Mangena, V., Chen, J., Chung, K. (2016) "Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues", *Nature Biotechnology*, doi:10.1038/nbt.3641.
- Murray, E.\* , Cho, J.H.\* , Goodwin, D.\* , Ku, T.\* , **Swaney, J.\***, Kim, S.Y., Choi, H., Park, Y.G., Park, J.Y., Hubbert, A., McCue, M., Vassallo, S., Bakh, N., Frosch, M.P., Wedeeng, V.J., Seung, H.S., Chung, K. (2015) "Simple, scalable proteomic imaging for high-dimensional profiling of intact systems", *Cell*, 163, 6, 1500–1514. doi:10.1016/j.cell.2015.11.025
- (\* indicates co-first author)

## POSTERS and PRESENTATIONS

---

- Poster** "Non-Parametric Hyperdimensional Analysis of Multiscale Phenotypic Factors in Intact Human Cerebral Organoids", Society for Neuroscience, October, 2019, Chicago, IL.
- Poster** "Non-Parametric Hyperdimensional Analysis of Multiscale Phenotypic Factors in Intact Human Cerebral Organoids", Cell Symposia: Engineering Organoids and Organoids, August, 2019, San Diego, CA.
- Presenter** "Simple, scalable proteomic imaging for high-dimensional profiling of intact systems", Korea Advanced Institute of Science and Technology, January, 2018, Daejeon, South Korea.
- Presenter** "Vascularization of neural organoids using two-photon stereolithography", Massachusetts Institute of Technology, October, 2016, Cambridge, MA.
- Poster** "Simple, scalable proteomic imaging for high-dimensional profiling of intact systems", Society for Neuroscience, November, 2016, San Diego, CA.
- Poster** "Kinetic modeling of amyloid beta aggregation using nanoparticle tracking analysis", Undergraduate Research Symposium, December, 2013, Madison, WI.

## HONORS and AWARDS

---

- 1<sup>st</sup> Place – Machine Learning Across Disciplines Challenge, MIT College of Computing. 2019
- Outstanding Seminar Award, MIT Chemical Engineering Department. 2018
- Individual Accomplishment Award, MIT Chemical Engineering Department. 2018
- Barbara J. Weedon Fellowship, MIT Picower Institute 2016-2017
- Lemelson Engineering Presidential Fellowship, MIT 2015
- Chancellor's Scholarship, UW-Madison 2010-2014 (Class Co-chair: 2010)
- Dean's Honor List, UW-Madison 2010-2014

## **Appendix 2: Image processing protocol**

## Part A: Destriping LSFM images using `pystripe`

*TIMING: 30 min setup (excluding download time), 1-4 h unattended computer time (depending on data size)*

### Container setup

- i) Download the raw data (see ‘Downloading full resolution example data’ in the MATERIALS section) or move pre-existing raw data to a new folder called “data” and make note of the folder path on the host. Skip this step if using the downsampled example data.

**CRITICAL STEP** Pystripe only accepts TIFF and RAW file formats. If using the full-resolution example data, the data folder should be created on a device with at least 4 TB of available space.

### ? TROUBLESHOOTING

- ii) Inspect the raw images using FIJI. ? TROUBLESHOOTING
- iii) Open a terminal and start the Docker container with the data folder from the host mounted inside the container using the following command:

```
docker run -it -v path_to_data:/data chunglabmit/shield-2018
```

where “path\_to\_data” should be replaced with the full path of the data folder on the host. Note that the command prompt will change to the root user indicating that the prompt is now running interactively from within the container.

**CRITICAL STEP** The semantics for mounting a volume to share data with the container are to specify a source path on the host and a target path inside the container. The syntax for expressing this at the command line is “-v path\_on\_host:path\_in\_container”. Note that path\_in\_container is a Unix-style path since the container is a Linux virtual machine.

**CRITICAL STEP** Add quotes around the full path if it contains any spaces. ? TROUBLESHOOTING

- iv) Verify that the data folder is mounted correctly by running the following command from within the Docker terminal window.

```
ls /data
```

This command should list the contents of the shared data folder mounted in the container at the root path.

## Destriping raw images from whole-brain LSF

- v) Run the following command to get the latest version of pystripe:

```
git -C /shield-2018/pystripe pull
```

Inspect the help message from pystripe by entering the following command from within the container.

```
pystripe --help
```

Instructions for using pystripe from the command line should print. Note from the instructions that an input folder and filter bandwidth are required and that pystripe will default to using all CPU cores available to the container.

- vi) Destripe the raw data one channel at a time by entering the following commands within the Docker terminal window for each channel:

```
pystripe -i /data/raw_data/channel -o /data/destriped_data/channel -s1 sigma1 -s2  
sigma2 -w wavelet -c compression -x crossover -f flat.tif -d dark
```

where “channel” represents the name of the folder containing images of the current channel, “sigma1” and “sigma2” represent the dual-band filter bandwidths in pixels (higher gives more filtering), “wavelet” is the mother wavelet name, “compression” is the amount of lossless TIFF compression, and “crossover” is the intensity range used to switch between filter bands. The arguments “flat” and “dark” are optional inputs for applying illumination correction. Including illumination correction in pystripe removes the need to load the images into memory multiple times.

If using the full-resolution example data, then enter the following command.

```
pystripe -i /data/raw_data/Ex_0_Em_0 -o /data/destriped_data/Ex_0_Em_0 -s1 256 -s2  
256 -w db10 -c 3 -x 10 --flat flat.tif --dark 100
```

If using the downsampled example data, then enter the following command.

```
pystripe -i /data/raw_data/Ex_0_Em_0 -o /data/destriped_data/Ex_0_Em_0 -s1 32 -s2  
32 -w db10 -c 3 -x 10 --flat flat_downsampled.tif --dark 100
```

A progress bar showing the destriping progress for that channel will appear, and the destriped images will appear in the data folder, which is also available from the host machine.

Inspect some of the resulting destriped images and the corresponding raw images in FIJI. Note that pystripe maintains the raw data folder structure in its output.

- vii) Repeat steps A) vi-vii for all other channels.
- viii) Shutdown the Docker container by typing “exit” in the Docker terminal window. The terminal will return to its original state before entering the Docker container.

**CRITICAL STEP** Since Docker creates new containers each time they are started, no files inside the Docker container will be saved when restarting the container except for files in shared folders on the host, such as the mounted data folder.

**PAUSE POINT** The destriped data folder can be archived for processing at a later date.

## Part B: Stitching LSFM stacks using TSV

*TIMING: 45 min setup (excluding download time), 4-12 h unattended computer time (depending on data size)*

### Container setup

- i) Download the destriped data (see ‘Downloading full resolution example data’ in the MATERIALS section) or use the destriped data from the previous step. If using the downsampled data, complete part A and use the resulting destriped data for part B.

**CRITICAL STEP** TSV only accepts TIFF and RAW file formats saved in Terastitcher hierarchical format. If using the full resolution example data, the data folder should be on a device with at least 4 TB of available space.

- ii) Open a terminal and start the Docker container with the data folder from the host mounted inside the container using the following command:

```
docker run -it -v path_to_data:/data chunglabmit/shield-2018
```

where “path\_to\_data” should be replaced with the full path of the data folder on the host. Note that the command prompt will change to the root user indicating that the prompt is now running interactively from within the container.

**CRITICAL STEP** The semantics for mounting a volume to share data with the container are to specify a source path on the host and a target path inside the container. The syntax for expressing this at the command line is “-v path\_on\_host:path\_in\_container”. Note that path\_in\_container is a Unix-style path since the container is a Linux virtual machine.

**CRITICAL STEP** Add quotes around the full path if it contains any spaces. Stitching images from whole-brain LSFM

iii) Run the following command to get the latest version of TSV:

```
git -C /shield-2018/tsv pull
```

iv) Use Terastitcher to create a stitching project file based on a single channel by entering the following command in the Docker terminal window.

```
terastitcher -1 --volin=/data/destriped_data/channel_master --ref1=H --ref2=V \
--ref3=D --vx11=x --vx12=y --vx13=z --sparse_data \
--projout=/data/destriped_data/channel_master/xml_import.xml
```

where “channel\_master” represents the name of the folder containing images of the channel used for calculating stack displacements and x, y, and z are the physical voxel dimensions in micron. If using the full-resolution example data, then enter the following command:

```
terastitcher -1 --volin=/data/destriped_data/Ex_1_Em_1 --ref1=H --ref2=V \
--ref3=D --vx11=1.8 --vx12=1.8 --vx13=2 --sparse_data \
--projout=/data/destriped_data/Ex_1_Em_1/xml_import.xml
```

If using the downsampled example data, then enter the following command:

```
terastitcher -1 --volin=/data/destriped_data/Ex_1_Em_1 --ref1=H --ref2=V \
--ref3=D --vx11=14.4 --vx12=14.4 --vx13=16 --sparse_data \
--projout=/data/destriped_data/Ex_1_Em_1/xml_import.xml \
```

**CRITICAL STEP** The voxel dimensions must match the voxel dimensions for the LSFM system in microns.

**? TROUBLESHOOTING**

- v) Calculate the stack displacements using Terastitcher by entering the following command in the Docker terminal window:

```
terastitcher -2 --projin=/data/desriped_data/channel_master/xml_import.xml \
--projout=/data/desriped_data/channel_master/xml_displacement.xml
```

If using the example data, enter the following command:

```
terastitcher -2 --projin=/data/desriped_data/Ex_1_Em_1/xml_import.xml \
--projout=/data/desriped_data/Ex_1_Em_1/xml_displacement.xml
```

**? TROUBLESHOOTING**

- vi) Generate a Terastitcher project file by entering the following command in the Docker terminal window:

```
terastitcher -3 --projin=/data/desriped_data/channel_master/xml_displacement.xml \
--projout=/data/desriped_data/channel_master/xml_displproj.xml
```

If using the example data, enter the following command:

```
terastitcher -3 --projin=/data/desriped_data/Ex_1_Em_1/xml_displacement.xml \
--projout=/data/desriped_data/Ex_1_Em_1/xml_displproj.xml
```

- vii) Use TSV to generate stitched images by entering the following command in the Docker terminal window:

```
tsv-convert-2D-tif --xml-path /data/desriped_data/channel_master/xml_displproj.xml \
--output-pattern /data/stitched_data/channel_master/"{z:04d}.tiff" \
--compression compression --ignore-z-offsets
```

where “/data/stitched\_data/channel\_master” is a new folder that will be created for the stitched images from the master channel, and “compression” is the amount of lossless TIFF compression to use. If using the example data, enter the following command:

```
tsv-convert-2D-tif --xml-path /data/destriped_data/Ex_1_Em_1/xml_displproj.xml \
--output-pattern /data/stitched_data/Ex_1_Em_1_master/"{z:04d}.tiff" \
--compression 4 --ignore-z-offsets
```

A progress bar will appear showing the stitching progress. Note that TSV will default to using all processor cores available to the container.

- viii) Inspect the stitched images in FIJI. TSV generates TIFF images which can be imported directly into FIJI.
- ix) If there are no more channels to stitched, then proceed to part C of the protocol.  
If there are more channels to be stitched, then the following command will stitch other channels using the displacements calculated from the master channel:

```
tsv-convert-2D-tif --xml-path /data/destriped_data/channel_master/xml_displproj.xml \
--output-pattern /data/stitched_data/channel/"{z:04d}.tiff" --compression compression \
--ignore-z-offsets --input /data/destriped_data/channel
```

Where “channel\_master” represents the folder name of the channel used for calculating stack displacements, “channel” represents folder of images from another channel to be stitched using the previously computed displacements, and the other arguments are as described before. Repeat this command for all other channels to be stitched.

If using the example data, then enter the following commands:

```
tsv-convert-2D-tif --xml-path /data/destriped_data/Ex_1_Em_1/xml_displproj.xml \
--output-pattern /data/stitched_data/Ex_0_Em_0/"{z:04d}.tiff" --compression 4 \
--ignore-zoffsets --input /data/destriped_data/Ex_0_Em_0
```

```
tsv-convert-2D-tif --xml-path /data/destriped_data/Ex_1_Em_1/xml_displproj.xml \
--output-pattern /data/stitched_data/Ex_2_Em_2/"{z:04d}.tiff" --compression 4 \
--ignore-zoffsets --input /data/destriped_data/Ex_2_Em_2
```

- x) Close the Docker container by typing “exit” into the Docker terminal window and pressing “Enter”.

PAUSE POINT The stitched data folder can be archived for processing at a later date.

## Part C: Atlas alignment with manual refinement using nuggt

*TIMING: 1-2 hr setup (excluding download time), 6-12 h unattended computer time (depending on data set size)*

### Container setup and Neuroglancer basics

- i) Download the stitched data (see ‘Downloading full resolution example data’ in the MATERIALS section) or use the stitched data from the previous step. If using the downsampled data, complete part B and use the resulting stitched data for part C.
- ii) For Linux users, open a terminal and start the Docker container with the data folder from the host mounted inside the container using the following command:

```
docker run -it --expose 8999 -p 8999:8999/tcp --network host \
-v path_to_data:/data --shm-size 1g chunglabmit/shield-2018
```

For Windows and Mac users, enter the following command:

```
docker run -it --expose 8999 -p 8999:8999/tcp -v path_to_data:/data \
--shm-size 1g chunglabmit/shield-2018
```

**CRITICAL STEP** Add quotes around the full path if it contains any spaces.

- iii) The following files are included in the “/allen-mouse-brain-atlas” folder inside the Docker container:
  - autofluorescence\_25\_half\_sagittal.tif - a 3D image of the autofluorescence channel for the mid-sagittal sectioning of the Allen Mouse Brain atlas reference (to be used in place of “/reference/reference.tiff”).
  - annotation\_25\_half\_sagittal.tif - a 3D segmentation of the Allen Mouse Brain reference (to be used in place of “/reference/segmentation.tiff”).
  - coords\_25\_half\_sagittal.json - the coordinates of key points on the reference (to be used in place of “/reference/points.json”)
  - AllBrainRegions.csv - a mapping of region ID numbers in the segmentation to names of regions.

Use nuggt to inspect the included mouse brain reference atlas by entering the following command into the Docker terminal window:

```
nuggt-display --port 8999 [--ip-address 0.0.0.0] \
--segmentation /allen-mouse-brainatlas/annotation_25_half_sagittal.tif \
--points /allen-mouse-brainatlas/coords_25_half_sagittal.json \
/allen-mouse-brainatlas/autofluorescence_25_half_sagittal.tif reference gray
```

where “–ip-address 0.0.0.0” should be included by Windows and Mac users. The nuggt-display application will print a URL similar to “<http://127.0.0.1:8999/v/...>” for Linux users in the Docker terminal window. For Windows and Mac users, the nuggt-display application will print a URL similar to “[http://\(container-id\):8999/v/...](http://(container-id):8999/v/...)” in the Docker terminal window, and the container ID should be replaced with “localhost” in the next step.

- iv) Open the URL in your browser to see the Neuroglancer user interface. The browser should display four panels. The top left one is the XY view, the top right is the XZ view, the bottom right one is the YZ view, and the bottom left one is a 3D view. The current cursor position in image coordinates is displayed at the top of the Neuroglancer user interface. A list of keyboard shortcuts is available if you type “h” while using Neuroglancer.[? TROUBLESHOOTING](#)

In Neuroglancer, you can move the stack by pressing the left mouse button with the cursor in any of the three orthogonal view panels and dragging. You can scroll through the depth of the stack by turning the mouse wheel, and you can zoom in and out by pressing the “ctrl” key and turning the mouse wheel. Each layer in Neuroglancer listed along the top of the user interface can be hidden by clicking the layer name. Segmentation layers can be used to highlight specific brain regions by double-clicking regions in any view panel.

- v) Exit the nuggt-display application by holding “ctrl” and pressing “C” in the Docker terminal window.

### Create a points file for a custom reference atlas

This step is optional if you are aligning a sagittal image of a mouse brain hemisphere to the Allen Mouse Brain Atlas (which is the case when using the example data). If you are using the example data, proceed to step C xii.

- vi) Locate either a 3D TIFF file of the reference image (e.g. “path\_to\_reference/reference.tiff”) or a folder of 2D TIFF files named in strict ascending alphabetical order (e.g. “path\_to\_reference/images/img\_0000.tiff”) on the host.

**CRITICAL STEP** You must prepare an accompanying segmentation of this reference (e.g. “path\_to\_reference/segmentation.tiff”) where each pixel of the segmentation has a value corresponding to the region in which the pixel lies. You must also have a mapping of region number to the region’s name in a format similar to the file “/allen-mouse-brainatlas/ AllBrainRegions.csv”.

- vii) Open a terminal and start the Docker container with the reference image(s) mounted using the following command:

```
docker run -it --expose 8999 -p 8999:8999/tcp [--network host] \
-v path_to_reference:/reference -v path_to_data:/data chunglabmit/shield-2018
```

where “path\_to\_reference” and “path\_to\_data” represent the full paths of the directories containing the reference image(s) and the images to be aligned to the reference on the host. The “–network host” argument should only be included by Linux users.

**CRITICAL STEP** Add quotes around the full path if it contains any spaces.

- viii) If you have a 3D reference TIFF file, type the following command into the Docker terminal window:

```
nuggt --port 8999 --image /reference/reference.tiff --output /reference/points.json
```

where “reference.tiff” should be replaced with the name of your reference image. If you have a folder of 2D TIFF files, type:

```
nuggt --port 8999 --image "/reference/images/*.tiff" --output /reference/points.json
```

**CRITICAL STEP** Use “\*.tif” instead of “\*.tiff” if that is the extension of your image files.

- ix) The nuggt application will display a line similar to “Editing viewer: <http://127.0.0.1:8999/v/...>” in the Docker terminal window. Windows and Mac users should replace the container ID in the URL with “localhost”. Open the URL in your browser to see the Neuroglancer user interface.

- x) Add fiducial points by placing the cursor where you want the point to be, holding the “shift” key and pressing “A”. You can delete a point by placing the cursor over it, holding the shift key and pressing “D”. You should annotate the image using easily identifiable locations like the dentate gyrus as well as the perimeter of the image (see Supplementary Fig. 1).? TROUBLESHOOTING
- xi) Hold the “shift” key and press “S” to save the annotations. Bring up the Docker terminal window again, hold “ctrl” and press “C” to exit from the nuggt application.

**PAUSE POINT** The points.json file can be saved and reused for future atlas alignment tasks using the same reference image(s).

### **Perform automatic alignment**

- xii) Determine whether the images to be aligned with the reference atlas needs to be flipped or rotated by inspecting the stitched and reference images.

If aligning to the provided sagittal mouse brain atlas, each z-plane of the stitched images should be approximately a sagittal section with the olfactory bulb at the top of the image and the cortex to the right (see Supplementary Fig. 1). The z-plane corresponding to the medial sagittal section should be the last image in the stack (the one that is the last in alphabetical order if 2D planes are used). If the cortex is at the left, you will need to “flip-X”. If the olfactory bulb is at the bottom, you will need to “flip-Y”. If the medial sagittal section is at  $Z = 0$ , then you will need to “flip-Z”. Note that “flipping” in this case means reversing the index of an image dimension, which corresponds to a reflection of the image in that dimension. ? TROUBLESHOOTING

If aligning the example data to the provided mouse brain atlas, then the Ex\_0\_Em\_0 channel (containing syto 16 nuclear stain) should be used for atlas alignment and flipped in the X and Z dimensions (see Supplementary Fig. 2).

- xiii) Determine whether the channel to be aligned with the reference atlas needs to be cropped.

Typically, the reference image extends to the left, right, top and bottom with no background margins, but the stitched images to be aligned have a margin to the left, right, top or bottom. If these margins are large, they will prevent the automatic alignment from succeeding; thus, the image to be aligned must be cropped to match when performing the automated alignment. Note the X-start and X-end, Y-start and Y-end, and Z-start

and Z-end (if whole planes need to be cropped) coordinates of the image after cropping using FIJI to examine the images.

CRITICAL STEP These cropping coordinates apply to the image before flipping.

If using the full resolution example data, then use the following crop coordinates:

- X-start, X-stop: None
- Y-start, Y-stop: 0, 9800
- Z-start, Z-stop: None

If using the downsampled example data, then use the following crop coordinates:

- X-start, X-stop: None
- Y-start, Y-stop: 0, 1225
- Z-start, Z-stop: None

- xiv) Run the rescale-image-for-alignment program to rescale, flip, and crop the channel used for atlas alignment. If step (xiii) determined that the image needs X-flipping, use the `--flipx` switch, and similarly for Y and Z. Likewise, if step (xiii) determined the image needs Xcropping, use the `--clip-x` argument, and similarly for Y and Z. For example, if an image needs to be flipped in X and Z and cropped in Y at the bottom of the image at coordinate 9100, the resulting command to be typed into the Docker terminal window is:

```
rescale-image-for-alignment --input "/data/stitched_data/channel_alignment/*.tiff" \  
--atlas-file /reference/reference.tiff --output \  
/data/downsampled_flip-x_flip-z_clip-y-0-9100.tiff --flip-x --flip-z --clip-y 0,9100
```

where “channel\_alignment” is the folder containing the stitched images of the channel used for alignment. If using the full resolution example data with the provided reference atlas, then type:

```
rescale-image-for-alignment --input "/data/stitched_data/Ex_0_Em_0/*.tiff" \  
--atlas-file /allen-mouse-brain-atlas/autofluorescence_25_half_sagittal.tif \  
--output /data/downsampled_flip-x_flip-z_clip-y-0-9800.tiff \  
--flip-x --flip-z --clip-y 0,9800
```

If using the downsampled example data with the provided reference atlas, then type:

```
rescale-image-for-alignment --input "/data/stitched_data/Ex_0_Em_0/*.tiff" \  
--atlas-file /allen-mouse-brain-atlas/autofluorescence_25_half_sagittal.tif \  
--output /data/downsampled_flip-x_flip-z_clip-y-0-1225.tiff \  
--flip-x --flip-z --clip-y 0,1225
```

```
--atlas-file /allen-mouse-brain-atlas/autofluorescence_25_half_sagittal.tif \
--output /data/downsampled_flip-x_flip-z_clip-y-0-1225.tiff \
--flip-x --flip-z --clip-y 0,1225
```

CRITICAL STEP Record the values for the –flip-x, –flip-y, –flip-z –clip-x, –clip-y and –clip-z switches in the output image name. These parameters will be used again in step C xxiv.

CRITICAL STEP If using the example data, make sure that Ex\_0\_Em\_0, the syto 16 channel, is rescaled and used for atlas alignment.

- xv) Align the rescaled image to the reference using Elastix. If using a custom reference image in the same example scenario presented in step C xiv, then the following command would be used:

```
sitk-align --moving-file /data/downsampled_flip-x_flip-z_clip-y-0-9100.tiff \
--fixed-file /reference/reference.tiff --fixed-point-file /reference/points.json \
--xyz --alignment-point-file /data/alignment.json
```

CRITICAL STEP Substitute the name of your reference image for “reference.tiff” above and substitute the name you used as the –moving-file in step (C)(xiii) for “downsampled.tiff” above.

If using the full resolution example data with the provided reference atlas, then type:

```
sitk-align --moving-file /data/downsampled_flip-x_flip-z_clip-y-0-9800.tiff \
--fixed-file /allen-mouse-brain-atlas/autofluorescence_25_half_sagittal.tif \
--fixed-point-file /allenmouse-brain-atlas/coords_25_half_sagittal.json \
--xyz --alignment-point-file /data/alignment.json
```

If using the downsampled example data with the provided reference atlas, then type:

```
sitk-align --moving-file /data/downsampled_flip-x_flip-z_clip-y-0-1225.tiff \
--fixed-file /allen-mouse-brain-atlas/autofluorescence_25_half_sagittal.tif \
--fixed-point-file /allenmouse-brain-atlas/coords_25_half_sagittal.json \
--xyz --alignment-point-file /data/alignment.json
```

## ? TROUBLESHOOTING

### Refine the alignment manually

- xvi) Use nuggt-align to manually refine the automatic alignment. If using a custom reference image in the same example scenario presented in step C xiv, then the following command would be used:

```
nuggt-align --port 8999 [--ip-address 0.0.0.0] \
--reference-image /reference/reference.tiff \
--moving-image /data/downsampled_flip-x_flip-z_clip-y-0-9100.tiff \
--points /data/alignment.json
```

where “–ip-address 0.0.0.0” should be included by Windows and Mac users.

**CRITICAL STEP** Substitute the name of your reference image for “reference.tiff” above and substitute the name you used as the –moving-file in step (C)(xiii) for “downsampled.tiff” above.

If aligning the full resolution example data to the provided reference atlas, then enter the following command in the Docker terminal window:

```
nuggt-align --port 8999 [--ip-address 0.0.0.0] --reference-image \
/allen-mouse-brainatlas/ autofluorescence_25_half_sagittal.tif \
--moving-image downsampled_flip-x_flipz_ clip-y-0-9800.tiff \
--points /data/alignment.json
```

If aligning the downsampled example data to the provided reference atlas, then enter the following command in the Docker terminal window:

```
nuggt-align --port 8999 [--ip-address 0.0.0.0] --reference-image \
/allen-mouse-brain-atlas/autofluorescence_25_half_sagittal.tif \
--moving-image /data/downsampled_flipx_ flip-z_clip-y-0-1225.tiff \
--points /data/alignment.json
```

- xvii) There are two links that are displayed by nuggt-align in the Docker terminal window. One is the link to a Neuroglander webpage displaying the reference image and one is a link to the image that is to be aligned. Copy each of the links and open them in separate windows in your web browser. Windows and Mac users should change the container ID in the URL to “localhost” (see step C iii).
- xviii) Refine the locations of each of the fiducial points in the image. This is done by selecting the reference image browser window (e.g. by clicking in it), holding down the “shift” key and typing “N”. A red point will appear in both the reference and moving image browsers. You should adjust the point in the moving image browser to correspond with its location in the reference image if it is placed incorrectly. This

is done by holding down the “ctrl” key and clicking on the location in the reference image where the point should be. The point can be readjusted multiple times until you are satisfied with the location. After this is done, hold the shift key down and type, “D” (for “done”) and hold the shift key down and type “N” to move to the next fiducial point. Repeat the adjustment procedure until all of the fiducial points have been adjusted. ? TROUBLESHOOTING

- xix) Hold the shift key down and press “S” to save the adjusted alignment.

**CRITICAL STEP** The location of the refined fiducial points are not saved automatically. The refined atlas alignment progress will be lost if the adjusted alignment points are not saved.

- xx) Hold the shift key down and type “W” to warp the image to be aligned to the reference image.

**CRITICAL STEP** Wait for the message, “Warping complete”, to be displayed before doing anything further.

- xxi) Add additional points to refine the warping further. Add a point to the reference image by holding the “ctrl” key down and pressing the mouse button while hovering over the location of a discrepancy. Focus the alignment image to this location by holding down the “shift” key and pressing “T” (for “translate”). In the alignment image Neuroglancer webpage, adjust the location of the point by holding down the “ctrl” key and pressing the left mouse button, then hold the shift key down and press “D” (for “Done”).

**CRITICAL STEP** Adding duplicate points in the same location will cause the warping to fail. To avoid this problem, it is helpful to warp the moving image after adding each point to verify that the warping is successful.? TROUBLESHOOTING

- xxii) Hold the shift key down and press “S” to save the adjusted alignment.

**CRITICAL STEP** The location of the refined fiducial points are not saved automatically. The refined atlas alignment progress will be lost if the adjusted alignment points are not saved.

- xxiii) In the Docker terminal window, hold down the “ctrl” key and press “C” to exit the nuggt-align program.

- xxiv) Rescale the alignment to the size of your original image. If using a custom reference image in the same example scenario presented in step C xiv, then the following command would be used:

```
rescale-alignment-file --stack "/data/stitched_data/channel_alignment/*.tiff" \
--alignmentimage /data/downsampled_flip-x_flip-z_clip-y-0-9100.tiff \
--input /data/alignment.json --output /data/rescaled-alignment.json \
--flip-x --flip-z --clip-y 0,9100
```

**CRITICAL STEP** replace “–flip-x –flip-z –clip-y 0,9100” above with the values used in step C xiv

If aligning the full resolution example data to the provided reference atlas, then enter the following command in the Docker terminal window:

```
rescale-alignment-file --stack "/data/stitched_data/Ex_0_Em_0/*.tiff" \
--alignment-image /data/downsampled_flip-x_flip-z_clip-y-0-9800.tiff --input \
/data/alignment.json --output /data/rescaled-alignment.json --flip-x --flip-z \
--clip-y 0,9800
```

If aligning the downsampled example data to the provided reference atlas, then enter the following command in the Docker terminal window:

```
rescale-alignment-file --stack "/data/stitched_data/Ex_0_Em_0/*.tiff" \
--alignment-image /data/downsampled_flip-x_flip-z_clip-y-0-1225.tiff \
--input /data/alignment.json --output /data/rescaled-alignment.json \
--flip-x --flip-z --clip-y 0,1225
```

**PAUSE POINT** The alignment.json and rescaled-alignment.json files can be saved for continuing the analysis at a later date.

### **Compute the total and mean intensity per region**

- xxv) Use the rescaled alignment file to compute fluorescence statistics of each channel for all brain regions. If using a custom reference image, then enter the following command into the Docker terminal window:

```
calculate-intensity-in-regions --input "/data/stitched_data/channel/*.tiff" \
--alignment /data/rescaled-alignment.json \
--reference-segmentation /reference/segmentation.tiff \
--brain-regions-csv /reference/brain-regions.csv \
--output /data/results-level-7.csv --level 7
```

where “channel” represents the folder of the current channel to be analyzed.

CRITICAL STEP the “–level” argument picks the granularity of the segmentation description from 7 (finest granularity) to 1 (coarsest granularity). You can specify multiple levels to be calculated at the same time at little additional computational cost.

If aligning the full resolution example data to the provided reference atlas, then enter the following command:

```
calculate-intensity-in-regions --input "/data/stitched_data/Ex_1_Em_1_master/*.tiff" \
--alignment /data/rescaled-alignment.json \
--reference-segmentation /allen-mouse-brainatlas/annotation_25_half_sagittal.tif \
--brain-regions-csv /allen-mouse-brainatlas/AllBrainRegions.csv \
--output /data/results-level-3.csv --level 3 \
--output /data/results-level-4.csv --level 4 \
--output /data/results-level-5.csv --level 5 \
--output /data/results-level-6.csv --level 6 \
--output /data/results-level-7.csv --level 7 --shrink 2 -ncores N
```

where N is the number of cores to use for calculating fluorescence statistics. Note that using more cores causes the memory usage to increase as well. If memory becomes a limiting factor, then limit the number of cores used or increase the shrink factor if acceptable.

If aligning the downsampled example data to the provided reference atlas, then enter the following command:

```
calculate-intensity-in-regions --input "/data/stitched_data/Ex_1_Em_1_master/*.tiff" \
--alignment /data/rescaled-alignment.json \
--reference-segmentation /allen-mouse-brainatlas/annotation_25_half_sagittal.tif \
--brain-regions-csv /allen-mouse-brainatlas/AllBrainRegions.csv \
--output /data/results-level-3.csv --level 3 \
--output /data/results-level-4.csv --level 4 \
--output /data/results-level-5.csv --level 5 \
--output /data/results-level-6.csv --level 6 \
--output /data/results-level-7.csv --level 7
```

This command can be repeated for all other channels to be analyzed. The results file that is created contains a spreadsheet of the region names, the volume of each region in voxels, the sum of all intensities in each region (total\_intensity) and the average intensity of the voxels in each region (mean\_intensity).

## Timing

For processing the full resolution example data:

Stage A, 30 min for setup and 2-4 h of unattended computer time

Stage B, 45 min for setup and 4-12 h of unattended computer time

Stage C, 1-2 h for setup and manual refinement and 4-12 h of unattended computer time

For processing the downsampled example data:

Stage A, 30 min for setup and 10 min of unattended computer time

Stage B, 45 min for setup and 20 min of unattended computer time

Stage C, 1-2 h for setup and manual refinement and 30 min of unattended computer time

## Troubleshooting

*A troubleshooting table is available in the online version of the protocol, which is available at <https://www.biorxiv.org/content/10.1101/576595v1>*

## **Appendix 3: Supplemental Information**

## Chapter 2 Supplemental Information

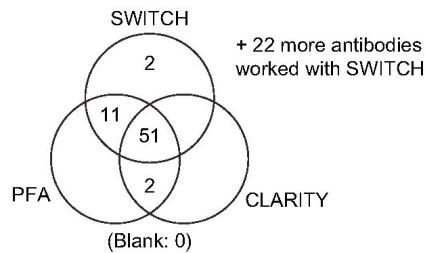
Supplemental videos are available online at [https://www.cell.com/cell/fulltext/S0092-8674\(15\)01505-6](https://www.cell.com/cell/fulltext/S0092-8674(15)01505-6) or from <http://www.chunglab.org/publications>

**Table 8.1:** Statistics of All Combinations of Six Cell Markers, Related to Figure 2.3.

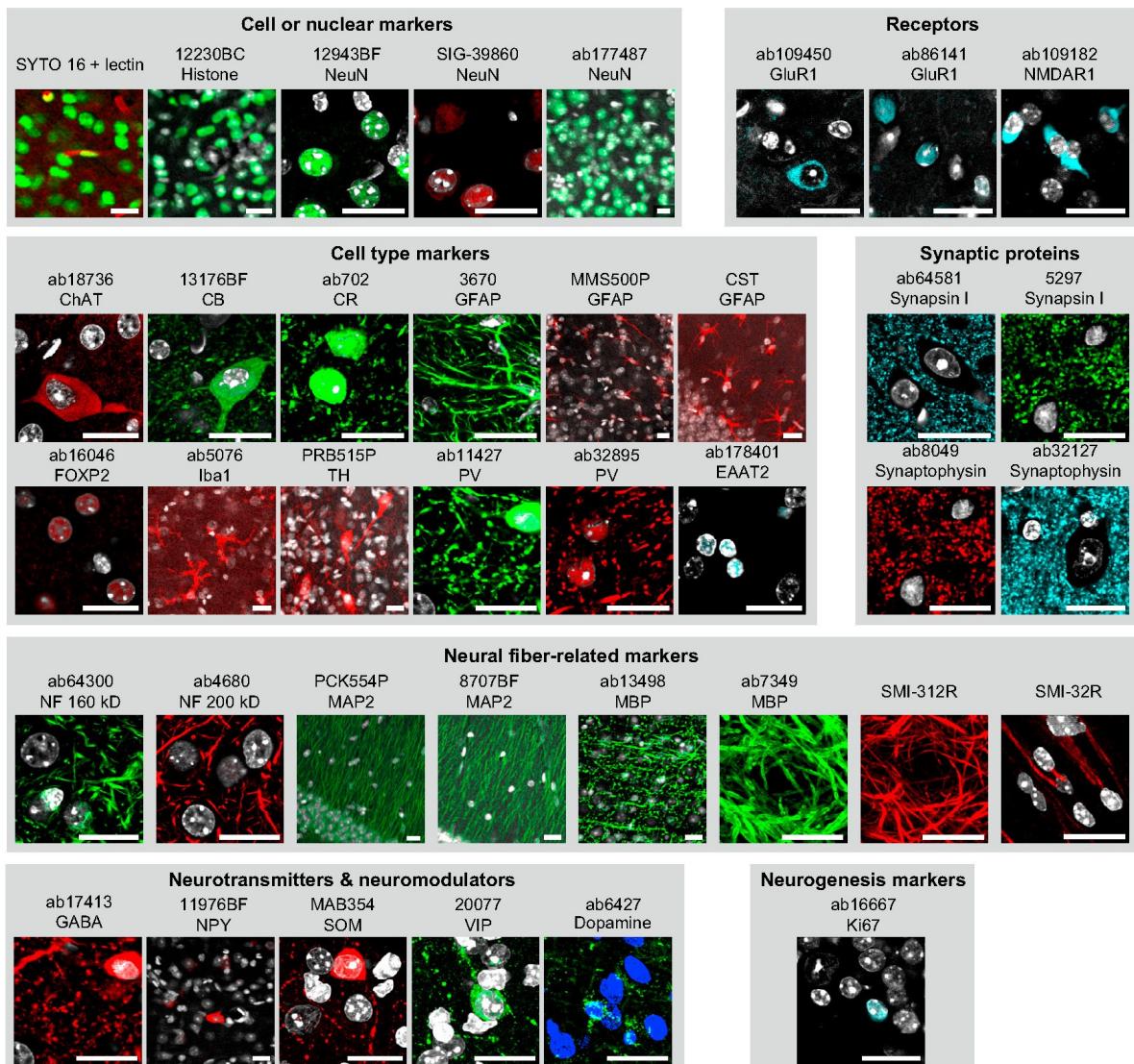
GFAP	NeuN	SMI-32	CB	CR	PV	I	II	III	IV	V	VI
—	—	—	—	—	+	0	3	11	5	1	7
—	—	—	—	+	—	4	15	42	7	5	1
—	—	—	—	+	+	0	0	0	0	0	0
—	—	—	+	—	—	2	0	14	1	3	8
—	—	—	+	—	+	0	2	4	5	1	15
—	—	—	+	+	—	1	11	12	3	1	3
—	—	—	+	+	+	0	0	0	0	0	0
—	—	+	—	—	—	0	0	0	0	0	0
—	—	+	—	—	+	0	0	0	0	0	0
—	—	+	—	+	—	0	0	0	0	0	0
—	—	+	+	—	—	0	0	0	0	0	0
—	—	+	+	—	+	0	0	0	0	0	0
—	—	+	+	—	—	0	0	0	0	0	0
—	—	+	+	+	+	0	0	0	0	0	0
—	+	—	—	—	—	59	828	1,007	1,197	1,158	2,030
—	+	—	—	—	+	0	8	38	57	32	49
—	+	—	—	+	—	15	56	93	15	4	6
—	+	—	—	+	+	0	0	0	0	0	0
—	+	—	—	+	—	1	77	448	22	43	83
—	+	—	—	+	—	0	10	64	66	49	13
—	+	—	+	+	—	2	23	34	4	1	4
—	+	—	+	+	+	0	0	0	0	0	0
—	+	+	—	—	—	0	3	122	24	78	10
—	+	+	—	—	+	0	0	1	0	0	0
—	+	+	—	+	—	0	0	0	0	0	0
—	+	+	—	+	+	0	0	0	0	0	0
—	+	+	—	—	—	0	0	4	1	0	0
—	+	+	+	—	+	0	0	2	3	0	0
—	+	+	+	+	—	0	0	0	0	0	0
—	+	+	+	+	+	0	0	0	0	0	0



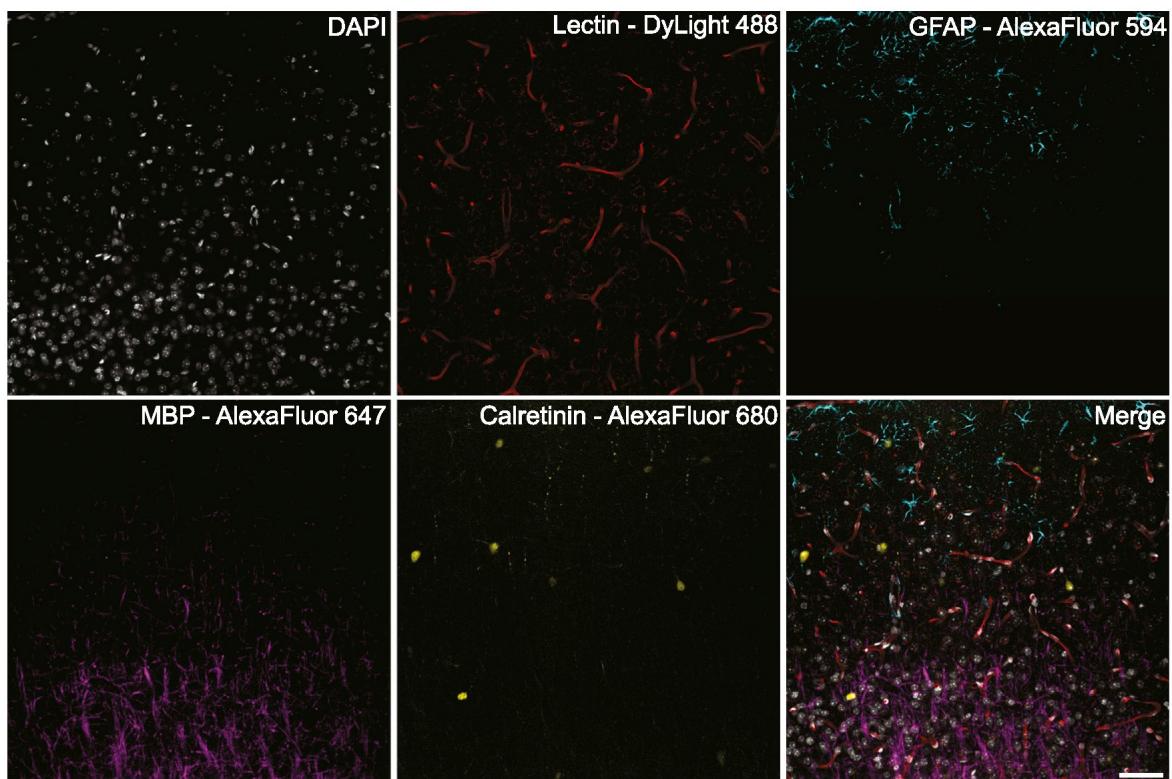
**Summary of immunostaining compatibility**  
(Total 90 antibodies)



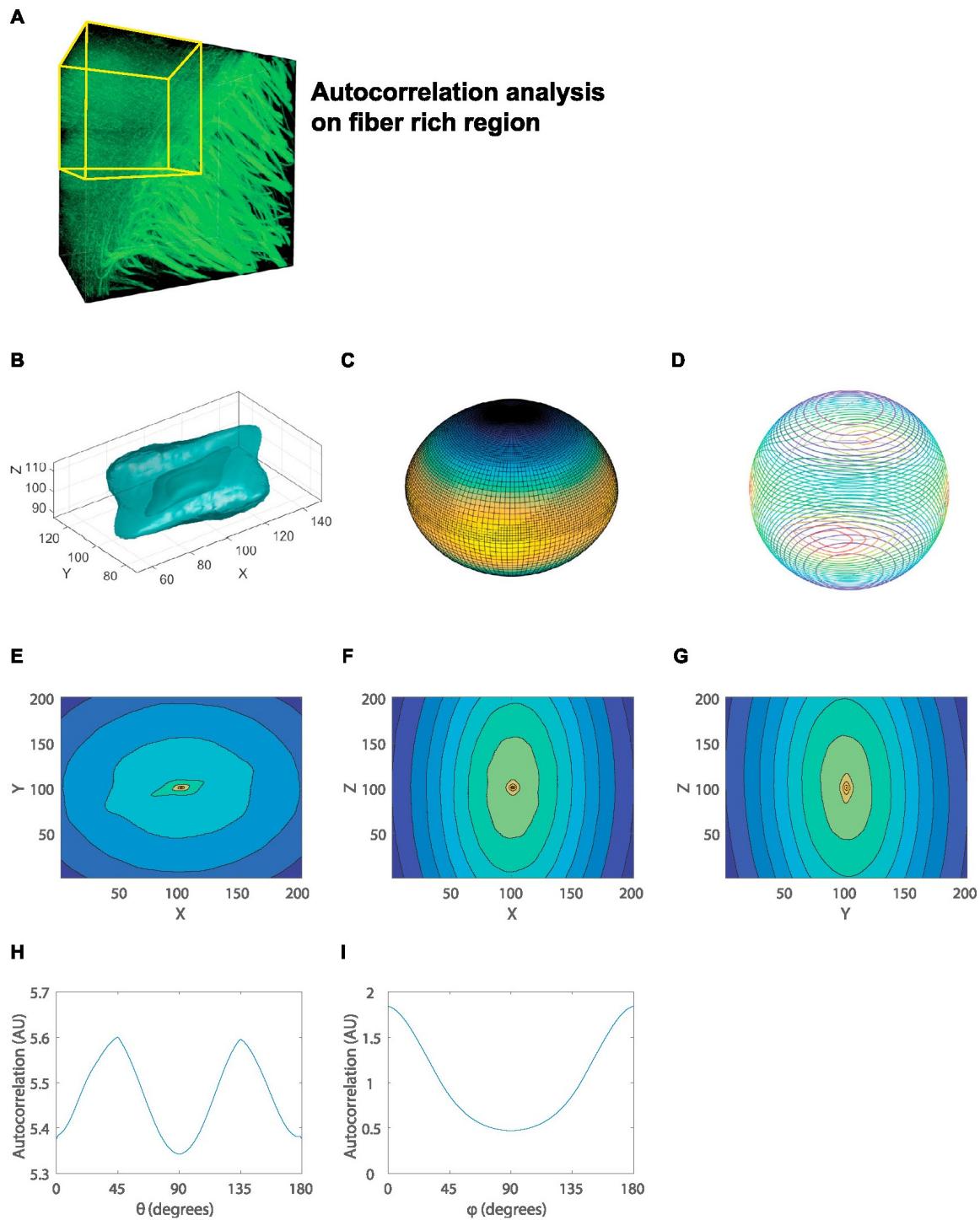
**Examples of successful staining results with SWITCH**



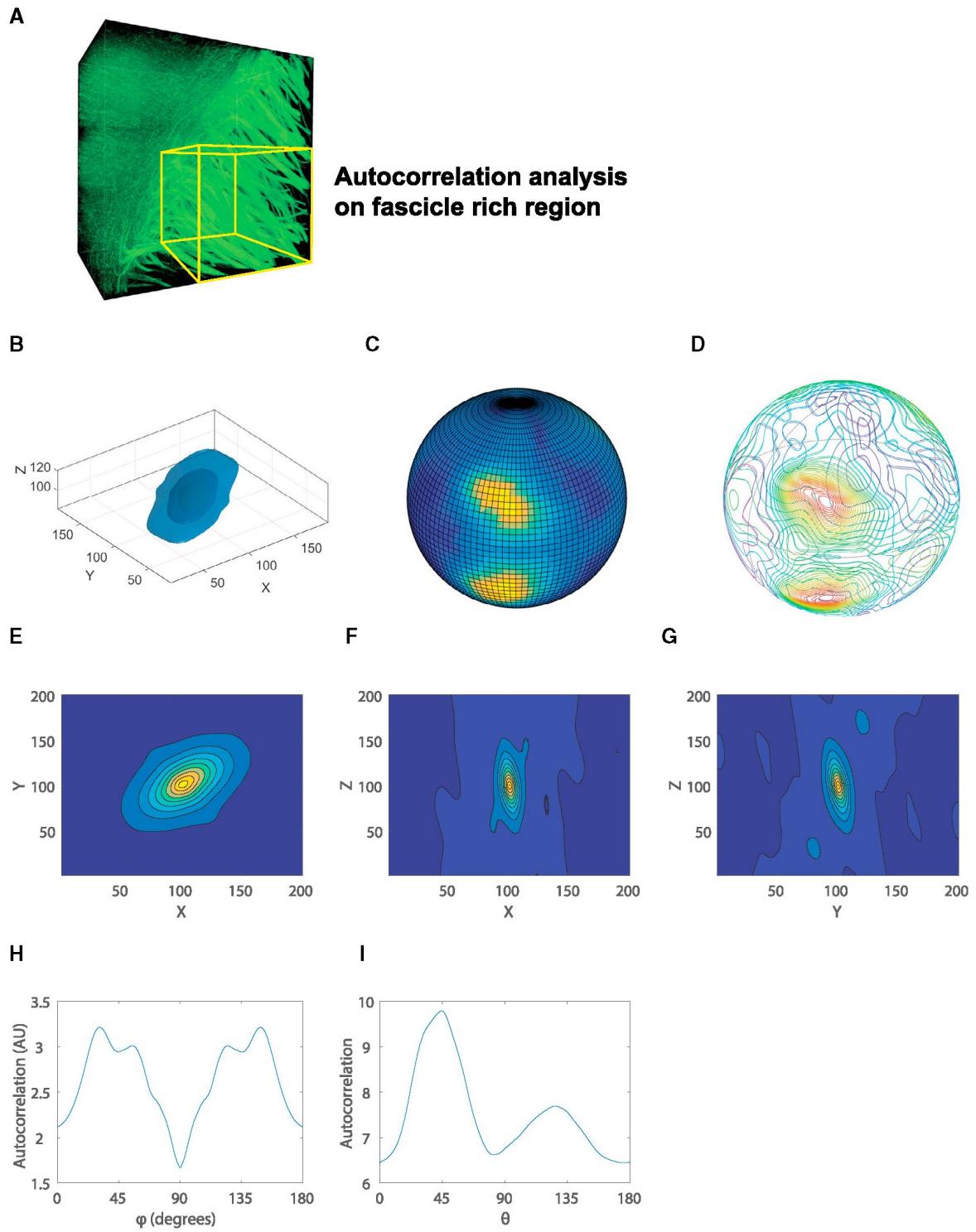
**Figure 8.1:** SWITCH Antibody Compatibility, Related to Figure 2.1. A total of 90 antibodies were tested on SWITCH-processed sections and PFA-fixed control tissues, and a subset of these were also tested on CLARITY-processed sections. A summary of the results are displayed in Venn diagram form, and representative images from staining of SWITCH-processed samples using antibodies specific for various types of targets are shown. All images were obtained with either a 10x, 0.3 NA water-immersion objective or 40x, 1.25 NA oilimmersion objective. Scale bars, 20  $\mu$ m.



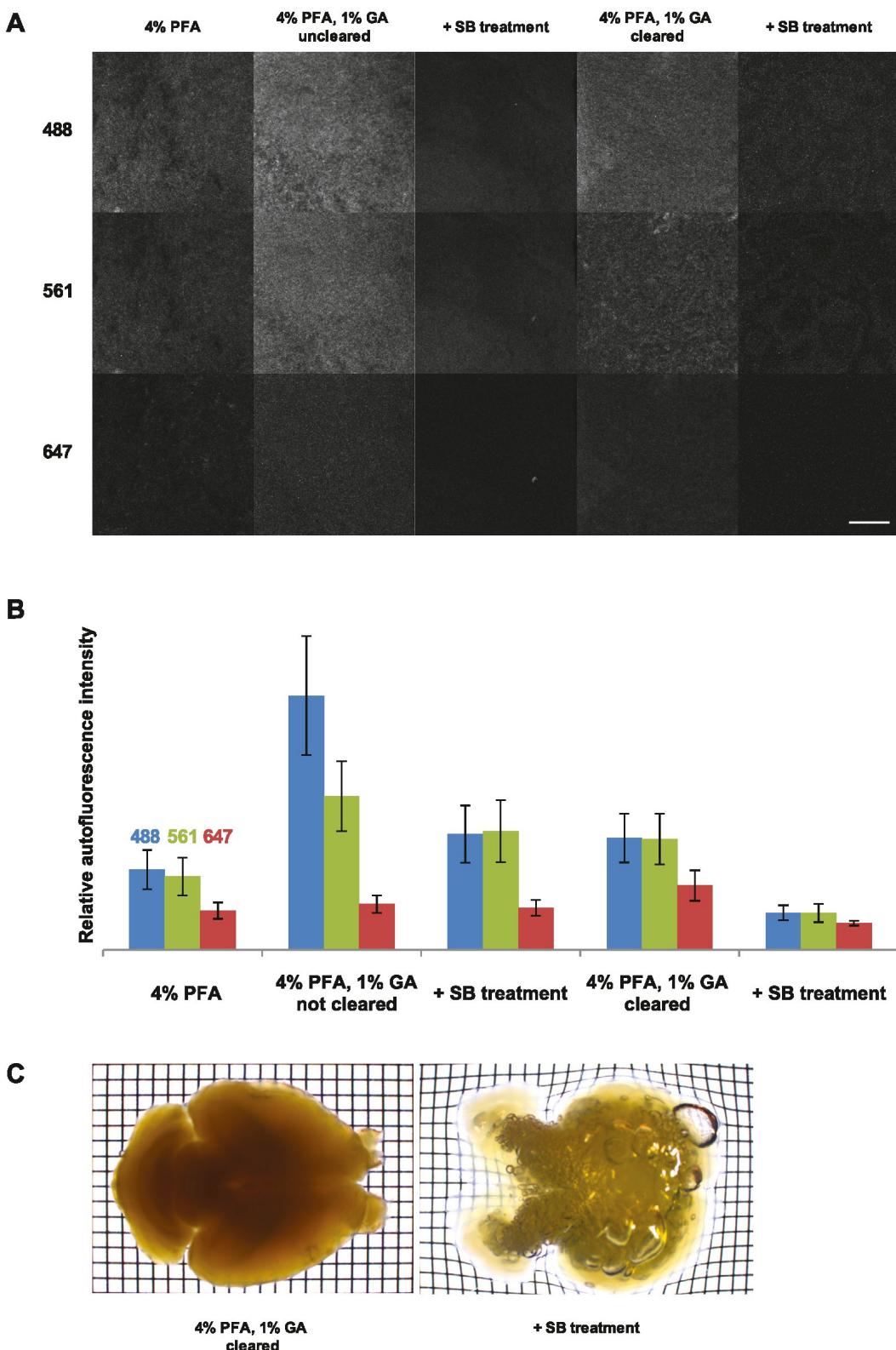
**Figure 8.2:** Multi-color Imaging of SWITCH-Processed Samples, Related to Figure 2.2. A 100- $\mu$ m section of mouse tissue was processed using SWITCH and stained with DAPI, lectin, anti-GFAP, anti-MBP, and anti-CR. Representative images are shown with the fluorophores used for labeling. Leica TCS SP8 equipped with a white-light laser and spectral detection system was used for imaging. 25x, 0.95 NA, water-immersion objective. Scale bar, 50  $\mu$ m.



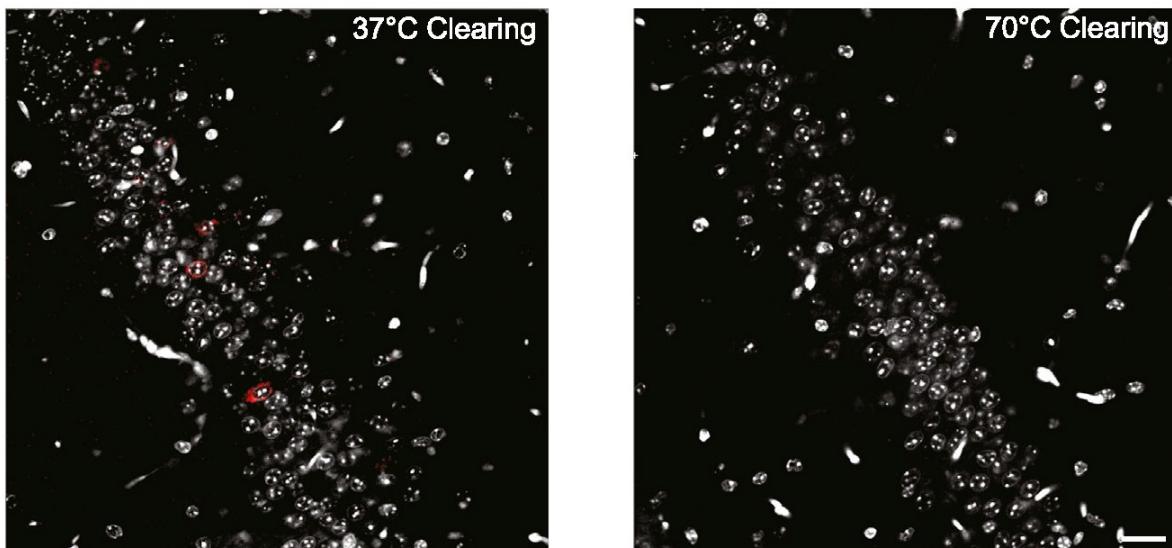
**Figure 8.3:** Autocorrelation Analysis of Myelinated Fibers, Related to Figure 2.5. Autocorrelation shows all the fiber orientations in a more non-biased manner. Autocorrelation was calculated using MATLAB by applying the Fourier convolution theorem with 3D FFT and a periodic boundary condition. (A) Subvolume used for the analysis of the fiber orientations. (B) Autocorrelation isosurfaces show orientation  $F_i$  angle. The autocorrelation result mirrors the finite difference gradient approach in that it showed a cubic grid pattern for the fibers.



**Figure 8.4:** Autocorrelation Analysis of Bulk Myelinated Fibers, Related to Figure 2.5. (A) Subvolume used for the analysis of the fascicle orientations. (B–I) Same as that of the fiber analysis (Figure 8.3) but for fascicles. This autocorrelation result also agrees with the finite difference gradient approach in that it showed an orthogonal divergence of the fascicles. AU, arbitrary unit.



**Figure 8.5:** GA-Induced Fluorescence, Related to Figure 2.1. Aldehyde fixatives can react with tissue components to generate fluorescent products. (A and B) Background fluorescence is worse in GA than PFA and is especially pronounced in the green (488) to yellow (561) regions. Sodium borohydride (SB) treatment can reduce these fluorescent products. N = 3. Scale bar, 100  $\mu$ m. (C) SB treatment results in tissue damage and bubble formation within the sample. Grid spacing, 1 mm.



**Figure 8.6:** mRNA FISH on Thermally Cleared Samples, Related to Figure 2.4. SWITCH-processed samples were cleared at 37C and 70C before being subjected to an mRNA FISH protocol to label c-fos (red). Nuclei are stained with Hoechst (gray). Samples cleared at 37C showed successful mRNA labeling, but samples cleared at 70C contained no detectable mRNA. Scale bar, 100  $\mu$ m.

## Supplemental Experimental Procedures

### Mice

Young adult male and female C57BL/6 and Thy1-eGFP-M mice were housed in a reverse 12-hr light/dark cycle with unrestricted access to food and water. All experimental protocols were approved by the MIT Institutional Animal Care and Use Committee and Division of Comparative Medicine and were in accordance with guidelines from the National Institute of Health.

### Perfusion

Mice were transcardially perfused with ice-cold PBS and a solution consisting of 4% PFA and 1– 4% GA in PBS. Brain tissues were harvested and incubated in the same fixative solution at 4°C for 2–3 days and 2–7 hr at 37°C with gentle shaking to allow for uniform fixation throughout the sample. This incubation time is critical for mitigating the effects of variable perfusion quality and promotes uniform structural and molecular preservation of the sample throughout the SWITCH process, as samples that are not ideally preserved will experience greater loss of biomolecules and a greater degree of sample deformation.

## SWITCH-mediated Tissue Preservation

PFA-fixed human samples were washed in a solution consisting of 50% PBS titrated to pH 3 using HCl, 25% 0.1 M HCl, and 25% 0.1 M potassium hydrogen phthalate (KHP). This wash solution was then replaced with fresh solution with the addition of 4–10% GA. The samples were then incubated in this pH 3 solution at 4°C for 2 days with gentle shaking. The acidic pH of this solution greatly slows down the reaction speed of aldehyde fixatives. The solution was then replaced with PBS with the addition of 1–4% GA and the sample was again allowed to incubate for 2–3 days at 4°C and 2–7 hr at 37°C with gentle shaking. The sample was then washed in PBS at room temperature (RT) for 1 day with gentle shaking. After washing, reactive GA within the sample was inactivated by incubation in a solution consisting of 4% glycine and 4% acetamide for 1 day at 37°C with gentle shaking. Finally, the sample was washed for 1 day in PBS at RT with gentle shaking.

## GA Gelation Time Experiment

All reagents and containers were first cooled to 4°C and handled on ice. For pH 7 gels, 10 mL of a PBS solution containing 10% BSA was made prior to GA injection. For pH 3, 10 mL of a 0.1 M KHP buffer was titrated to pH 3 with HCl. Once the pH 3 and 7 BSA solutions were prepared, GA was added to a final concentration of 4% and a timer was started. Gelation was judged by inverting the tube and inspecting for fluid flow. The time required to form rigid gels was recorded in 3 replicates of the pH 3 and 7 condition.

## Gel Denaturation Experiment

For acrylamide (AA) gels, 10 mL of a 4% PFA, 5–15% bovine serum albumin (BSA), 4% AA, and 0.25% VA-044 solution was prepared in PBS. The polymerization was carried out under vacuum for 2 hr at 37°C in a 15 mL tube. For GA gels, 10 mL of a 5–15% BSA and 1% GA solution was prepared in PBS and allowed to gel at room temperature for 2 hr in a 15 mL tube. For the epoxide gels, 10 mL of a 5–15% BSA and 15% epoxide (i.e., EX-313, GE31, GE22) solution was prepared in 0.1 M carbonate buffer at pH 9 and allowed to gel at 37°C for 8 hr in a 15 mL tube. Each gel was extracted from the tubes and cut into approximately 5-mm-thick disks. The disks were then washed overnight in PBST to remove any unreacted reagents. After washing, each disk was massed and photographed. The disks were then transferred to an 80°C water bath and incubated overnight. The disks were removed, massed, and photographed again.

## Sodium Borohydride Treatment

Sodium borohydride (SB) buffer was made immediately before use by making a 1 mg/ml SB solution with PBS. For 100- $\mu$ m sections, the tissue was treated 3 times, 10 minutes each. For whole brain, the tissue was treated 3 times, 3 hr each. For both, incubation was done at room temperature without shaking.

## Passive Clearing with Thermal Energy

Aqueous clearing solution containing 200 mM SDS, 10 mM lithium hydroxide, 40 mM boric acid, and a variable amount of anti-browning agent (i.e., 0–50 mM sodium sulfite or 0–0.5% (w/v) 1-thioglycerol) was titrated to pH 9 using sodium hydroxide before use. For high temperature clearing, samples were incubated in 40 mL of clearing solution for 8 hr at RT with gentle shaking to allow for the anti-browning agent to diffuse throughout the tissue. Samples were then transferred to a water bath set at 60–80°C. The clearing buffer was replaced if any noticeable color was observed in solution at any point during clearing. To remove the remaining antibrowning agent and SDS after clearing, the samples were washed at 37°C for 24 hr in 40 mL PBST containing 0.02% sodium azide as a preservative.

## mRNA FISH

Mice were perfused with ice-cold PBS and then with fixative (4% PFA and 1% GA in PBS). Brains were incubated in the fixative for one day at 4 °C and then 6 hr at RT for post-fixation. Coronal sections were prepared with a vibratome and sections were inactivated at RT for 6 hr, followed by tissue clearing under 37 °C or 70 °C. All solutions were prepared by using diethylpyrocarbonate (DEPC)-treated water. Digoxigenin- and 2,4-dinitrophenol (DNP)-labeled fos cRNA probes were detected using horseradish peroxidase-conjugated antibodies. FISH signals were visualized using a tyramide amplification kit (Perkin Elmer).

## Refractive Index Matching

A customized refractive index (RI)-matching solution was made by dissolving 50 g dia-trizoic acid, 40 g n-methyl-d-glucamine, and 55 g iodixanol per 100 mL water. Cleared samples were incubated in 10 mL of this solution at RT with gentle shaking for 2 days prior to imaging, replacing the solution after the first day. The listed components and their proportions were chosen to adjust the pH and RI for ideal optical clearing (basic pH with RI near 1.47) as well as optimize the osmolarity of the solution to reverse the

sample expansion observed after clearing. The contrast agents, diatrizoic acid and iodixanol, significantly affect the RI of the solution, while n-methyl-dglucamine is used to adjust the pH to more basic values. All components were considered when optimizing for osmolarity. RI was measured using an Abbemat WR/MW automatic multiwavelength refractometer.

### Mounting and Imaging

To facilitate the use of long working distance immersion objectives, samples were mounted between a slide glass and a glass-bottom Willco dish. Blu-Tack adhesive was rolled into a cylindrical shape of a thickness slightly greater than that of the sample and was placed in a circular orientation on the slide glass with a small opening to allow addition of immersion medium after chamber construction. The sample was placed within the Blu-Tack circle and the Willco dish was secured onto the adhesive, pressing just firmly enough to make slight contact with the sample. This contact prevents the sample from moving during the imaging process, but minimizes sample deformation. For multiplexed staining experiments, contact was not made with the sample. Taking care to avoid introduction of bubbles, RI-matching solution was injected to fill the void space, and the opening was then closed using fast-curing epoxy glue. Three microscope systems were used for the experiments in this study:

- i. Olympus two-photon microscope system (FV1200MPE) equipped with a 25 $\times$  CLARITYoptimized objective (prototype; NA, 1.0; WD, 8.0 mm), a 10 $\times$  CLARITY-optimized objective (XLPLN10XSVMP; NA, 0.6; WD, 8.0 mm), a 10 $\times$  water-immersion objective (NA, 0.30; WD, 3.6 mm), and a 40 $\times$  oil-immersion objective (UPLSAPO40XS; NA, 1.25; WD, 0.3 mm). 405, 488, and 635 nm 1p lasers were used;
- ii. Leica TCS SP8 microscope system equipped with a white-light laser, spectral detection system, a 20 $\times$  water-immersion objective (NA, 0.50; WD, 3.5 mm) and a 25 $\times$  waterimmersion objective (NA, 0.95; WD, 2.4 mm).
- iii. Custom-made light-sheet microscope equipped with 10 $\times$  CLARITY-optimized objective (modified from Tomer et al., 2012). Sample datasets were visualized with IMARIS (Bitplane).

## Sample Delabeling

Imaged samples were delabeled in clearing solution at 60–80°C (elution condition) for 1–2 days for large samples and overnight (O/N) for thin samples. Sulfites were added for large samples to prevent browning during the extended delabeling process.

## Protein Loss Assay

Mouse brain samples were prepared using various preservation methods, hemisected, and then cut into 1-mm sections. The collection of sections from each hemisphere was massed and then placed into 5 mL of 200 mM SDS clearing solution. The samples were incubated at 37°C with gentle shaking for 2 weeks. A small aliquot was taken from each tube and analyzed using the Bio-Rad DC protein assay kit to quantify the degree of protein loss from the samples.

## Microstructure Preservation Assay

A Thy1-eGFP-M mouse was perfused with SWITCH fixative solution and the sample was cut into 1-mm sections. The eGFP expression on the surface of a section was imaged using confocal microscopy and then sample was then subjected to clearing under harsh conditions (200 mM SDS, 80°C) for 1 day. The sample was then labeled using anti-eGFP antibodies and the same region was imaged again using confocal microscopy.

## Macrostructure Preservation Assay

Samples processed using SWITCH and CLARITY were cut in 1-mm sections and subjected to clearing under harsh conditions for 1 day. The samples were then mounted in a chamber larger than the size of the sample to prevent compression and imaged using confocal microscopy. RI-matching solution was used to facilitate imaging of the sample. Cross-sections from the samples were visualized using IMARIS.

## Human Tissue Samples

Samples of fixed autopsy tissue were obtained from the Neuropathology Core of the Massachusetts Alzheimer Disease Research Center. Tissue was collected and banked in accordance with approval from the local Institutional Review Board. All samples studied came from subjects without evidence of neurologic disease on clinical grounds at the time

of death and without evidence of significant disease processes upon full neuropathologic examination.

### Multiplexed Labeling of Thin Samples

A human clinical sample containing visual cortex was obtained and processed using SWITCH. After fixation, 100- m sections were obtained from the sample and cleared under harsh conditions. In each round, the sample was labeled with DAPI, DyLight 488-conjugated lectin, and a variable antibody using standard immunolabeling procedures. The sample was mounted in a chamber larger than the size of the sample, imaged using confocal microscopy, and then delabeled under harsh conditions O/N. RI-matching solution was used to facilitate imaging of the sample.

### Co-Registration of Multiplexed SWITCH Experiments

To register the set of SWITCH experiments from a single tissue, one of the experimental rounds for a tissue is arbitrarily chosen as the fixed reference to which the rest of the experiments will be registered. Each experiment has one fluorescence channel dedicated to lectin, which allows the software to identify distinctive points in the vasculature in order to achieve the fine morphological adjustments across the tissue volume. We used a 3D Harris Corner detector (Harris and Stephens, 1988) to find those keypoints and a 3D modification of the SIFT descriptor (Lowe, 2004; Scovanner et al., 2007) to calculate correspondences. For robustness, we instituted a variation of RANSAC (Fischler and Bolles, 1981) to test affine transformations on local subvolumes, confirming that keypoint correspondences between experiments were legitimate. Finally, with the validated keypoints, a thin plate spline interpolation (Bookstein, 1989) was implemented to warp the tissue in a physically plausible manner.

Each experiment was processed individually by eye before using software to calculate the registration: The size of the interrogated tissues, approximately 1 mm<sup>3</sup> in volume with 1.09 m × 1.09 m × 1.99 m resolution, required imaging subvolumes that were stitched together using the Leica or Olympus microscope software. The resulting tissue volume, a 4-dimensional object for the three spatial coordinates and the fluorescence channel, was examined by human eye using FIJI (Schindelin et al., 2012) to ensure image quality and a common orientation across experiments. Additionally, a rectangular crop was made around the tissue to remove unnecessary, blank voxels. Each individual experiment was then processed using a MATLAB computational pipeline developed for SWITCH and shared online via Github (<https://github.com/dgoodwin208/Registration>).

The registration pipeline has five primary steps based on the Lectin channel of each experiment. First, the image volume is partitioned into 25 subvolumes for parallelization of work and robustness checking in later steps. Each subvolume then identifies distinguishable keypoints using a 3D Harris Corner Detector and uses a 3D modification of SIFT written by Scovanner (<http://www.cs.ucf.edu/~pscovann/>) to create a descriptor vector associated with the keypoint. Note that we calculated keypoints and descriptors at multiple scale levels, achieved by convolution with Gaussian kernels of progressive size, to ensure a sufficient degree of scale invariance to the detected descriptors to successfully find matches despite differences in microscope setups across experiments. The calculations for keypoints and descriptors was often calculated in parallel across subvolumes to save time.

The third step is that each subvolume searches for corresponding keypoints in the appropriate subvolume in the reference experiment using the SIFT metric for measuring similarity between descriptors. To accomplish this we used the open-source VLFeat (<http://www.vlfeat.org>) implementation of SIFT matching algorithm. The fourth step is that the keypoint correspondences are validated via calculating affine transformations of random subsets of 4 corresponding keypoints, and the number of inliers of the resulting transformation assessed by a voxel distance threshold of 3 pixels. Each time a pair of corresponding points is counted as an inlier, it receives a vote, and after the order of 106 affine transformations, correspondence pairs with at least 80% of the votes of the highest voted pair are kept as legitimate correspondences. Finally, the validated correspondences are used to calculate a thin plate spline for the entire volume using an opensource TPS implementation written by Yang, Foong and Ong (<http://www.mathworks.com/matlabcentral/fileexchange/47409-glmdtps-registrationmethod/> content/GLMD\_Demo/src/TPS3D.m), resulting in a highly accurate warp to match the morphology of the reference experiment.

## Semi-automatic Identification of Cells and Blood Vessels

Image volumes were displayed and analyzed using custom-built graphical user interface software developed with Delphi XE4 (Embarcadero Technologies). Each image section was preprocessed to correct the inhomogeneous illumination at each image tile. In detail, we subtracted the mean intensity of a  $100 \times 100$  m<sup>2</sup> window centered to each pixel from its intensity to uniformize the background intensity. A different algorithm was devised specifically for each marker to semiautomatically detect the centroid location and soma size of all cellular objects and vascular pixels. In general, a spherical soma volume was isolated according to the best contrast between intrasomal pixels and background pixels by increasing the size of concentric spheres, and the soma size was determined as the

spherical diameter. After this automatic detection process, we corrected misidentified cell bodies manually, and the portion of the correction was less than 10%. We applied normalization of foreground signal and a Gaussian filter to the NeuN channel prior to the analysis. SMI-32+ cells were fully recognized manually according to their characteristic feature of the soma connected to a vertically oriented fiber with a large nuclear shadow, and the determination of their coordinates and soma sizes was aided by an automation module of the software. Each section in the lectin channel was converted to a vascular pixel mask image according to a customized threshold, and unconnected small clusters of pixels were removed.

### Quantitative Analysis of the Co-registered Image Channels

We used a series of custom-built software developed with Delphi XE4 for the quantitative analysis. For co-expression analysis of two or more markers, an initial decision was made by checking whether the centroids of cells in each channel fell within a 5- m distance. This classification was then manually verified with a quick review software tool. Especially, NeuN– neurons were carefully reviewed, and any weak NeuN signal that changed synchronously with the other marker signal was identified as NeuN+. Cell density, vascular density, and cell- or pixel-to-vessel distance along cortical depth were obtained from their average of 50- or 100- m window with a 50- or 100- m interval, and data points of less than 10 objects were excluded from plotting. Cell and vascular densities were corrected to exclude the dead volume outside the tissue in the ROI. The cell-tovessel distance was calculated as a distance from the centroid location of the cell to the nearest vascular pixel. Cells or pixels positioned at the sections containing any incomplete vascular information ( $z < 24$  m or  $z > 80$  m) were excluded from the analysis of distance to nearest vessel. The distribution profile of cell-to-vessel distance was obtained with a 3- m interval.

### SWITCH-mediated Myelinated Fiber Labeling

To create a DiD solution for myelinated fiber labeling, 1 mg of DiD powder was dissolved in 200 L of a solution consisting of 10 mM SDS in PBS (SWITCH-Off). For 1-mm mouse sections, samples were incubated in SWITCH-Off solution O/N with gentle shaking at 37°C. The solution was replaced with a volume of fresh SWITCH-Off buffer that was sufficient to cover the sample, and 1 L of the DiD solution was added. The sample was allowed to incubate for 1 day at 37°C with gentle shaking, at which point the sample was moved to a large volume of PBST (SWITCHOn) for 1 day at 37°C with gentle shaking. The sample was imaged using confocal microscopy. RI-matching solution was used to facilitate imaging of the sample. For mouse hemispheres, the sample was incubated in

SWITCH-Off solution at 37°C O/N with gentle shaking and then transferred to a volume of fresh SWITCH-Off solution sufficient to cover the sample, at which point 2 L of the DiD solution was added. The sample was incubated in this solution at 37°C for 4 days with gentle shaking and then moved to a large volume of SWITCH-On solution for 1 day at 37°C with gentle shaking. RI-matching solution was used to facilitate imaging of the sample.

### Orientation Analysis of Myelinated Fibers

Analysis was performed on planar images in the xy, yz, and xz planes using OrientationJ (<http://bigwww.epfl.ch/demo/orientation/>). Specifically, OrientationJ was used to calculate the preferred orientation of each pixel (ranging from  $-90^\circ$  to  $90^\circ$ ) using the corresponding finite difference gradient. This generates planar images whose pixel values correspond to the angular component in that plane (i.e., xy planar image contains xy). After obtaining this orientational information in xy, yz, and xz for all slices, the separate components of the orientation (i.e., xy contains x and y components of the orientation) are added together to yield three-dimensional orientation vectors. The orientation vectors represent the orientations of the fibers and the fascicles. These orientation vectors can be binned according to their angles to yield information about how the fibers and the fascicles are distributed in terms of their orientation. This information can then be used to predict what angle of intersection these fibers make. Specifically, each peak in the histogram is identified and the subpopulation is estimated based on the FWHM. These are then assigned to either fibers or fascicles based on observation (i.e., in the xy plane, the image shows that fibers make a vertical/horizontal grid while the fascicles make more of a diagonal/divergent diversion; this means that the peaks near 0 and  $90^\circ$  (which is equivalent to  $-90^\circ$ ) correspond to the fibers, and the peaks near  $-45^\circ$  and  $45^\circ$  correspond to the fascicles). After obtaining the total populations of the entire volume, the intersections are estimated by subtracting the corresponding two peaks and then scaling that result by the FWHM. Then, assuming that all fibers and fascicles have similar pixel counts, the fraction of fibers making certain intersections can be determined.

### Autocorrelation Analysis of Myelinated Fibers

The analysis using the finite difference gradient is a nonlinear process that may introduce error to the analysis. The error is compounded by the fact that the z-resolution of the volume is almost three times lower than the x- and y-resolutions. A more accurate approach would be to use autocorrelation. (The finite difference gradient acts as a high pass filter for the autocorrelation.) Autocorrelation would show all the peak distributions

in a more non-biased manner. As such, we calculated the autocorrelation in the volume image using MATLAB. Specifically, we used the Fourier convolution theorem with 3DFFT and a periodic boundary condition to calculate the autocorrelation of the volume image filtered with a Gaussian window; then, we transformed the resulting autocorrelation data in Cartesian coordinates to spherical coordinates and integrated out the radial component to visualize the data.

### SWITCH-mediated Antibody Labeling

Samples were first equilibrated in a large volume of SWITCH-Off solution (0.5 mM SDS in PBS). Samples were then moved to a volume a SWITCH-Off solution just large enough to cover the sample and containing 20 L of antibody solution (for histone H3 staining of 1-mm-thick tissue blocks.) Care should be taken to ensure that the final concentration of SDS in the SWITCH-Off solution is appropriate after the addition of antibody solution. The amount of antibody solution necessary will depend on the target identity. The samples were incubated in this antibody solution for 12 hr at 37°C with gentle shaking. Samples were then transferred to 10 mL of SWITCH-On solution (PBST) and were washed for 12 hr at 37°C with gentle shaking.

## Chapter 3 Supplemental Information

Supplemental videos are available online at <https://www.nature.com/articles/nbt.3641> or from <http://www.chunglab.org/publications>

**Table 8.2:** Antibody summary.

Target	Vendor	Catalog #	Host	Clonality	Target size	PFA	MAP
Calbindin	Abcam <sup>1</sup>	ab11426	Rb <sup>2</sup>	P <sup>3</sup>	Protein	O	O
Calbindin	CST	13176	Rb	M	Protein	O	O
Calbindin	Abcam	ab82812	Ms	M	Protein	O	X
Calretinin	Abcam	ab702	Rb	P	Protein	O	O
Calretinin	Abcam	ab133316	Rb	M	Protein	O	X
CaMKIIa	Abcam	ab22609	Ms	M	Protein	O	O
ChAT	Aves	CAT	Ch	P	Protein	O	O
GAD2	CST	5843	Rb	M	Protein	O	O
GAD65/2	BioLegend	844502	Ms	M	Protein	O	O
GAD65/2	BioLegend	844503	Ms	M	Protein	O	O
GAD65/67	Millipore	AB1511	Rb	P	Protein	O	O
GAD67	Millipore	MAB5406	Ms	M	Protein	O	O
GFAP	Abcam	ab48050	Rb	P	Protein	O	O
GFAP	Aves	GFAP	Ch	P	Protein	O	O
GFAP	BioLegend	835301	Ms	M	Protein	O	O
GFAP	BioLegend	837201	Ms	M	Protein	O	O
GFAP, A488 <sup>4</sup>	CST	3655	Ms	M	Protein	O	O
GFAP, A594 <sup>5</sup>	CST	8152	Ms	M	Protein	O	O
Iba1	Abcam	ab107159	Gt	P	Protein	O	X
Iba1	Abcam	ab5076	Gt	P	Protein	O	X
Iba1	Novus	NB100-1028	Gt	P	Protein	O	X
Iba1	Wako	019-19741	Rb	P	Protein	O	O
NeuN	Abcam	ab104225	Rb	P	Protein	O	O
NeuN	Abcam	ab177487	Rb	M	Protein	O	O
NeuN	BioLegend	834501	Ms	M	Protein	O	O
NeuN	CST	12943	Rb	M	Protein	O	O
NeuN	CST	24307	Rb	M	Protein	O	O
NeuN	Millipore	MAB377	Ms	M	Protein	O	O
Parvalbumin	Abcam	ab11427	Rb	P	Protein	O	O

<sup>1</sup>CST, Cell Signaling Technology; SYSY, Synaptic Systems; LT, Life Technologies.

<sup>2</sup>Rb, rabbit; Ms, mouse; Gt, goat; Ch, chicken; Rt, rat; GP, guinea pig.

<sup>3</sup>M, monoclonal; P, polyclonal.

<sup>4</sup>A488, Alexa Fluor 488 conjugate.

<sup>5</sup>A594, Alexa Fluor 594 conjugate.

Target	Vendor	Catalog #	Host	Clonality	Target size	PFA	MAP
Parvalbumin	Abcam	ab32895	Gt	P	Protein	O	O
Somatostatin	Millipore	MAB354	Rt	M	Protein	O	O
Tau	Aves	TAU	Ch	P	Protein	X	O
TH	Abcam	ab112	Rb	P	Protein	O	O
TH	Abcam	ab134461	Ch	P	Protein	O	O
TH	Aves	TYH	Ch	P	Protein	O	O
TH	BioLegend	818001	Ms	M	Protein	O	O
CNPase	Aves	CNP	Ch	P	Protein	X	O
MAP2	Abcam	ab32454	Rb	P	Protein	O	O
MAP2	Abcam	ab5392	Ch	P	Protein	O	X
MAP2	Aves	MAP	Ch	P	Protein	O	O
MAP2	BioLegend	822501	Ch	P	Protein	O	X
MAP2	BioLegend	801801	Ms	M	Protein	O	O
MAP2	BioLegend	840601	Rb	P	Protein	O	O
MAP2	CST	8707	Rb	M	Protein	O	O
MBP	Abcam	ab134018	Ch	P	Protein	O	O
MBP	Abcam	ab7349	Rt	M	Protein	O	O
MBP	Aves	MBP	Ch	P	Protein	O	O
MBP	BioLegend	808401	Ms	M	Protein	O	O
MBP	BioLegend	836501	Ms	M	Protein	O	O
MBP	CST	78896	Rb	M	Protein	O	O
NF-H	Aves	NFH	Ch	P	Protein	O	O
NF-H	BioLegend	835801	Ms	M	Protein	O	O
NF-H	CST	2836	Ms	M	Protein	O	O
NF-H NP	BioLegend	801701	Ms	M	Protein	O	O
NF-H P	BioLegend	801601	Ms	M	Protein	O	O
NF-H P	BioLegend	835501	Ms	M	Protein	O	O
NF-H P	BioLegend	835701	Ms	M	Protein	O	X
NF-H/NF-M HP	BioLegend	835602	Ms	M	Protein	O	O
NF-H/NF-M HP	BioLegend	835604	Ms	M	Protein	O	O
NF-H/NF-M P	BioLegend	837702	Ms	M	Protein	O	X
NF-H/NF-M P	BioLegend	837704	Ms	M	Protein	O	O
NF-M	Abcam	ab64300	Rb	P	Protein	O	O
NF-M	Aves	NFM	Ch	P	Protein	O	X
NF-M	CST	2838	Ms	M	Protein	O	X
NF-L	Aves	NFL	Ch	P	Protein	O	O
NF-L	BioLegend	845802	Ms	M	Protein	O	O
NF-L	BioLegend	846002	Ms	M	Protein	O	O
NF-L	CST	2837	Rb	M	Protein	O	O
NF pan-axon	BioLegend	837904	Ms	M	Protein	O	O

Target	Vendor	Catalog #	Host	Clonality	Target size	PFA	MAP
NF pan-neuro	BioLegend	837801	Ms	M	Protein	O	X
Pan-actin	CST	8456	Rb	M	Protein	O	O
PLP	Aves	PLP	Ch	P	Protein	O	O
SMI-312	BioLegend	837902	Ms	M	Protein	O	O
SMI-32	BioLegend	801704	Ms	M	Protein	O	O
Tubulin	Abcam	ab6160	Rt	M	Protein	O	O
Tubulin	Abcam	ab7291	Ms	M	Protein	O	O
Tubulin	CST	5063	Rb	M	Protein	O	O
Tubulin	CST	7634	Rb	M	Protein	O	X
Tubulin 3	Aves	TUJ	Ch	P	Protein	O	O
Tubulin 3	BioLegend	801201	Ms	M	Protein	O	O
Tubulin 3	BioLegend	801202	Ms	M	Protein	O	O
Tubulin 3	BioLegend	802001	Rb	P	Protein	O	O
Tubulin 3	BioLegend	845501	Rb	M	Protein	O	X
Neuropeptide Y	CST	11976	Rb	M	Peptides	O	O
Bassoon	Abcam	ab82958	Ms	M	Protein	O	O
Bassoon	SYSY	141003	Rb	P	Protein	O	O
GABABR1	Millipore	AB2256	GP	P	Protein	O	O
GluR1	Abcam	ab86141	Rb	P	Protein	X	O
GluR1	BioLegend	819801	Ms	M	Protein	O	X
GluR2/3	Millipore	AB1506	Rb	P	Protein	O	O
Homer1	SYSY	160003	Rb	P	Protein	O	O
mGluR1	Aves	ER1	Ch	P	Protein	O	X
mGluR2	Aves	ER2	Ch	P	Protein	O	X
mGluR5	Aves	ER5	Ch	P	Protein	X	O
nAchR- 7	BioLegend	838401	Rt	M	Protein	X	O
Neuroligin-1	BioLegend	819001	Ms	M	Protein	O	X
Neuroligin-3	BioLegend	822001	Ms	M	Protein	O	X
PSD95	BioLegend	810401	Ms	M	Protein	X	O
PSD95	CST	2507	Rb	P	Protein	X	O
PSD95	CST	3450	Rb	M	Protein	O	O
PSD95	CST	3409	Rb	M	Protein	X	O
PSD95	NeuroMab	75-028	Ms	M	Protein	O	O
Synapsin I	Abcam	ab64581	Rb	P	Protein	O	O
Synapsin I	CST	5297	Rb	M	Protein	O	O
Synaptophysin	Abcam	ab52636	Rb	M	Protein	O	O
Synaptophysin	CST	5461	Rb	M	Protein	O	O
Synaptotagmin	Aves	STG	Ch	P	Protein	O	O
SYNPR	Abcam	ab175224	Rb	M	Protein	O	O
VGluT1	Abcam	ab104898	Rb	P	Protein	O	O

Target	Vendor	Catalog #	Host	Clonality	Target size	PFA	MAP
VGluT1	BioLegend	821301	Ms	M	Protein	O	O
VGluT2	Abcam	ab101760	Gt	P	Protein	O	O
E-Cadherin	BD	610181	Ms	M	Protein	O	X
N-cadherin	BD	610920	Ms	M	Protein	O	X
GAP	Aves	GAP43	Ch	P	Protein	O	X
GFP	Aves	GFP-1020	Ch	P	Protein	O	O
GFP	CST	2956	Rb	M	Protein	O	O
GFP	LT	A10262	Ch	P	Protein	O	O
GFP, A594 <sup>6</sup>	LT	A21312	Rb	P	Protein	O	O
GFP, A647 <sup>7</sup>	LT	A31852	Rb	P	Protein	O	O
Lectin, D488 <sup>8</sup>	Vector	DL-1174	N/A	N/A	Protein	O	O
Lectin, D594 <sup>9</sup>	Vector	DL-1177	N/A	N/A	Protein	O	O
Histone H3, A647 <sup>10</sup>	CST	12230	Rb	M	Protein	O	O
DAPI	LT	D1306	N/A	N/A	DNA	O	O
TOTO-1	LT	T3600	N/A	N/A	DNA	O	O

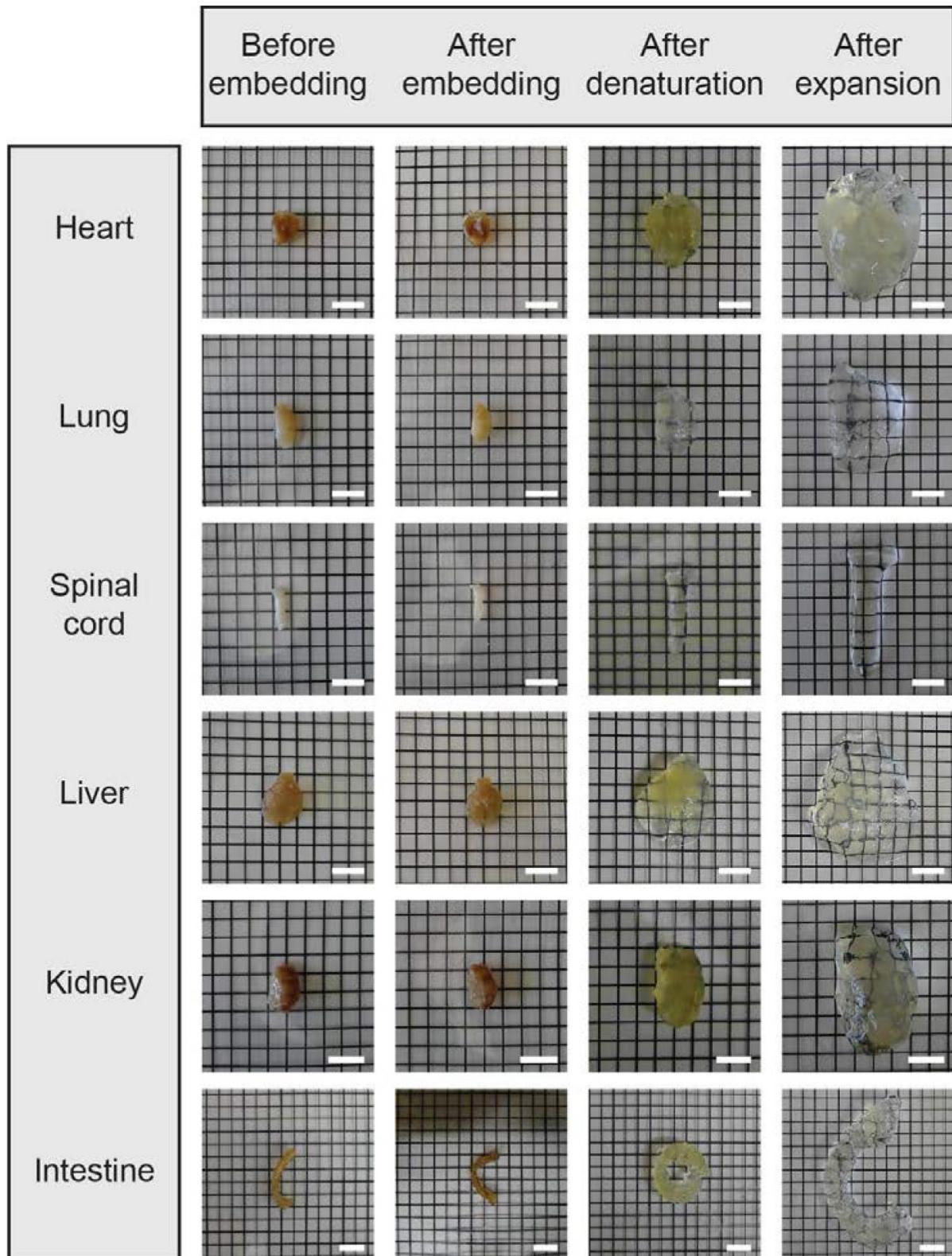
<sup>6</sup>A594, Alexa Fluor 594 conjugate.

<sup>7</sup>A647, Alexa Fluor 647 conjugate.

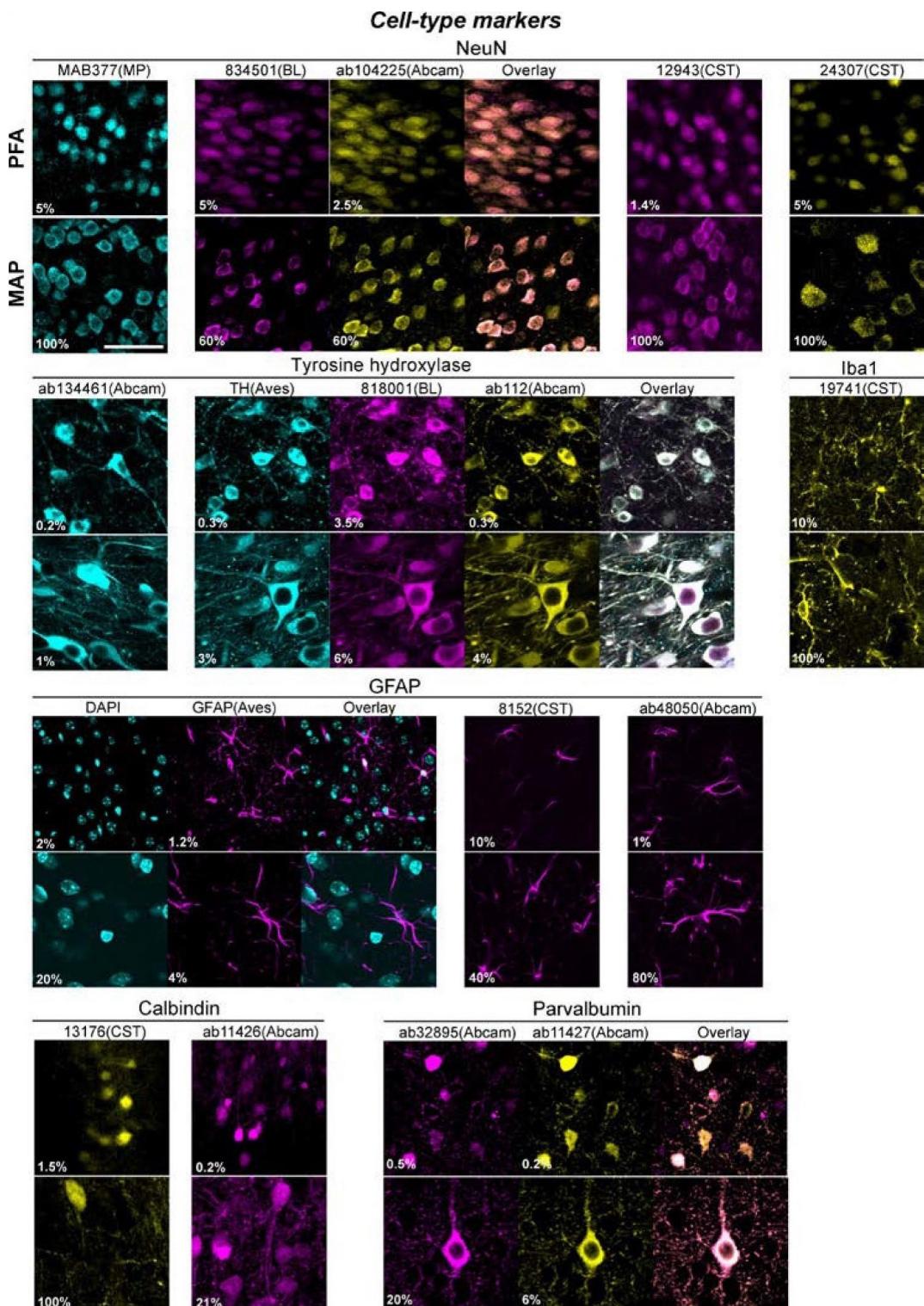
<sup>8</sup>D488, DyLight 488 conjugate.

<sup>9</sup>D594, DyLight 594 conjugate.

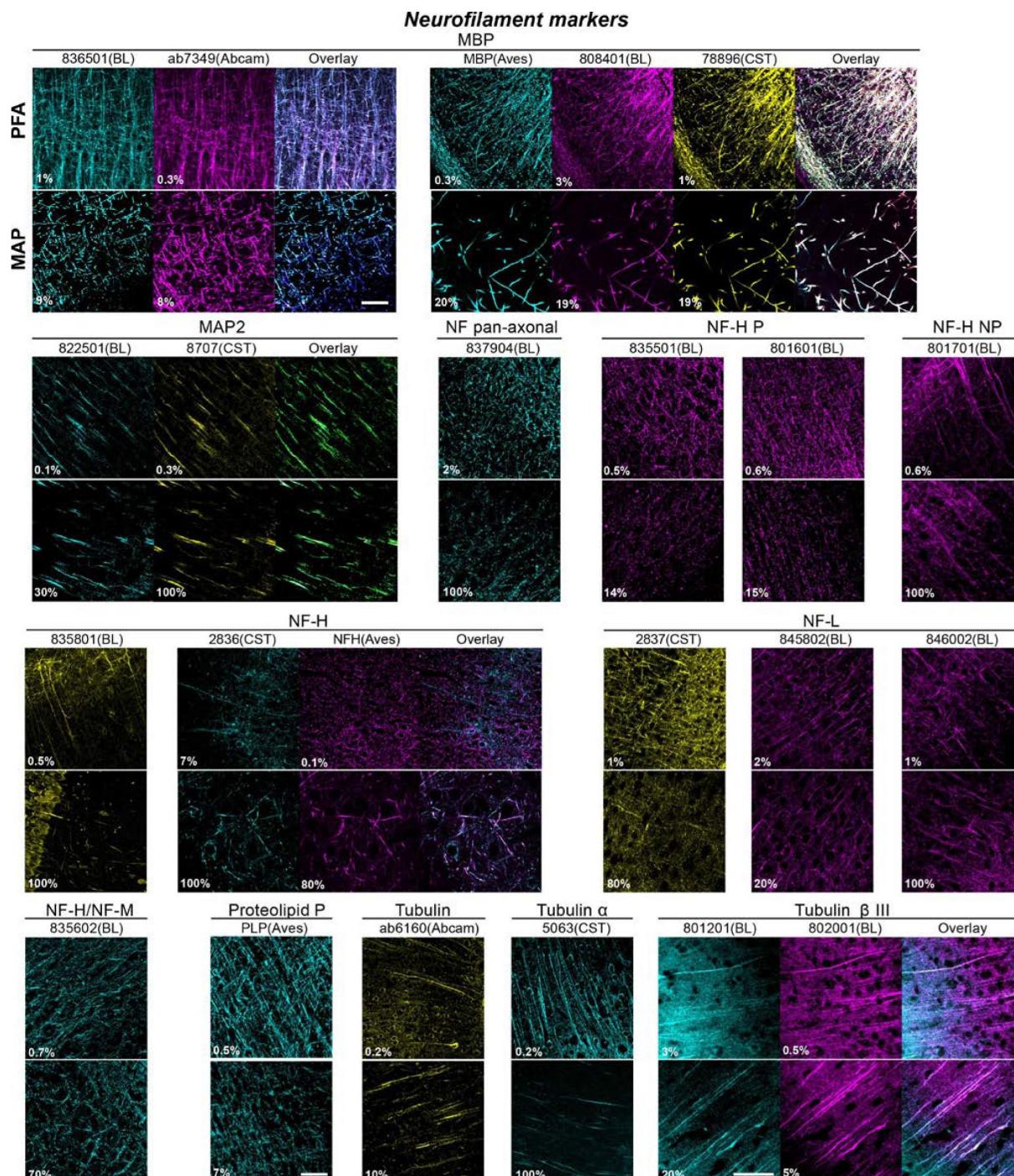
<sup>10</sup>A647, Alexa Fluor 647 conjugate.



**Figure 8.7:** Expansion of various organs with MAP. Photographs of intact organs being MAP processed. Organs were harvested after perfusion using PBS solution containing 4% PFA, 30% AA, 0.05% BA, 5% SA, and 0.1% VA-044. After allowing 2 days for chemical diffusion at 4°C, hydrogel-tissue hybridization was performed at 50°C for 2 h. Hydrogel-embedded organs were incubated in a 200 mM SDS and 50 mM sodium sulfite PBS solution for at least 24 h at 70°C and 12 h at 95°C. Denatured tissues were incubated in 100 ml DI water at room temperature for at least 36 h with gentle shaking. Scale bars, 10 mm.



**Figure 8.8:** Validation of commercial antibodies targeting cell-type markers in MAP-processed tissues. Fluorescence images from various commercial antibodies targeting cell-type markers tested in both MAP and control samples. Control and MAP samples were sectioned to 100- $\mu$ m thickness and then stained after denaturation. Primary incubation was performed for 12 h at 37°C with gentle shaking followed by a two-step wash of 2 h each in PBST. Secondary incubation was performed for 6 h at 37°C, followed by a 2-h wash in PBST, 30 min in 1:50,000 DAPI solution, then another 2-h wash in PBST. To determine specificity, antibodies targeting the same antigen were tested simultaneously in a single tissue using separate color channels when possible. Images were acquired with our Olympus confocal microscope with the following settings: 550 HV, 10 s pixel-1 dwell time, 1,024  $\times$  1,024 resolution, 0% offset, and laser power sufficient to nearly saturate signals. A 20 $\times$ , 0.95 NA water-immersion objective was used. MP, Millipore; BL, BioLegend, CST, Cell Signaling Technology. Scale bars, 20  $\mu$ m.



**Figure 8.9:** Validation of commercial antibodies targeting neurofilament markers in MAP-processed tissues. Fluorescence images from various commercial antibodies targeting neurofilament markers tested in both MAP and control samples. Images were obtained using the same method as Supplementary Figure 2. Scale bars, 20  $\mu$ m.

**Table 8.3:** Comparison of technologies for intact tissue processing and imaging.

	<b>MAP</b>	<b>ExM</b>	<b>CLARITY</b>	<b>SWITCH</b>	<b>Scale</b>
Type	Polymer-tissue hybrid	Polymer-tissue hybrid	Polymer-tissue hybrid	Multifunctional fixative-tissue-gel	Tissue
Initiation system	Thermal initiator	Persulfate / TEMED	Thermal initiator	N/A	N/A
Expansion rate	4-fold	4-fold	<2-fold	<1.5-fold	<2-fold
Homogenization	Protein denaturation and dissociation	Proteinase K digestion	N/A	N/A	N/A
Protein preservation	Yes	No	Yes	Yes	Yes
Staining step	After gel embedding and lipid removal	Before gel embedding	After gel embedding and lipid removal	After gel formation and lipid removal	After lipid removal
Applicable to whole organ	Yes	No	Yes	Yes	Yes
Protein labeling	Conventional antibodies	Custom trifunctional probes	Conventional Antibodies	Conventional Antibodies	Conventional Antibodies
Multiround immuno-labeling	Yes	No	Limited	Yes	No
Requires antibody labeling	Yes	No	Yes	Yes	Yes

## **Chapter 4 Supplemental Information**

Supplemental videos are available online at <https://www.biorxiv.org/content/10.1101/576595v1>

## Chapter 6 Supplemental Information

**Table 8.4:** Multiscale organoid features used in the SCOUT hyperdimensional statistical analyses

SCOUT Feature	Scale
DN nbrhd, sox2 count	Single-cell
DN nbrhd, tbr1 count	Single-cell
DN nbrhd, dn count	Single-cell
SOX2 nbrhd, sox2 count	Single-cell
SOX2 nbrhd, tbr1 count	Single-cell
SOX2 nbrhd, dn count	Single-cell
TBR1 nbrhd, sox2 count	Single-cell
TBR1 nbrhd, tbr1 count	Single-cell
TBR1 nbrhd, dn count	Single-cell
DP nbrhd, sox2 count	Single-cell
DP nbrhd, tbr1 count	Single-cell
DP nbrhd, dn count	Single-cell
MidTBR1 nbrhd, sox2 count	Single-cell
MidTBR1 nbrhd, tbr1 count	Single-cell
MidTBR1 nbrhd, dn count	Single-cell
MidSOX2 nbrhd, sox2 count	Single-cell
MidSOX2 nbrhd, tbr1 count	Single-cell
MidSOX2 nbrhd, dn count	Single-cell
MidInter nbrhd, sox2 count	Single-cell
MidInter nbrhd, tbr1 count	Single-cell
MidInter nbrhd, dn count	Single-cell
DN nbrhd fraction	Single-cell
SOX2 nbrhd fraction	Single-cell
TBR1 nbrhd fraction	Single-cell
DP nbrhd fraction	Single-cell
MidTBR1 nbrhd fraction	Single-cell
MidSOX2 nbrhd fraction	Single-cell
MidInter nbrhd fraction	Single-cell
TBR1 / SOX2 ratio	Single-cell
MidTBR1 / MidSOX2 ratio	Single-cell
MidInter / DN ratio	Single-cell
DN nbrhd, eq diam mean	Single-cell
DN nbrhd, eq diam stdev	Single-cell
SOX2 nbrhd, eq diam mean	Single-cell

SCOUT Feature	Scale
SOX2 nbrhd, eq diam stdev	Single-cell
TBR1 nbrhd, eq diam mean	Single-cell
TBR1 nbrhd, eq diam stdev	Single-cell
DP nbrhd, eq diam mean	Single-cell
DP nbrhd, eq diam stdev	Single-cell
MidTBR1 nbrhd, eq diam mean	Single-cell
MidTBR1 nbrhd, eq diam stdev	Single-cell
MidSOX2 nbrhd, eq diam mean	Single-cell
MidSOX2 nbrhd, eq diam stdev	Single-cell
MidInter nbrhd, eq diam mean	Single-cell
MidInter nbrhd, eq diam stdev	Single-cell
DN nbrhd, major axis mean	Single-cell
DN nbrhd, major axis stdev	Single-cell
SOX2 nbrhd, major axis mean	Single-cell
SOX2 nbrhd, major axis stdev	Single-cell
TBR1 nbrhd, major axis mean	Single-cell
TBR1 nbrhd, major axis stdev	Single-cell
DP nbrhd, major axis mean	Single-cell
DP nbrhd, major axis stdev	Single-cell
MidTBR1 nbrhd, major axis mean	Single-cell
MidTBR1 nbrhd, major axis stdev	Single-cell
MidSOX2 nbrhd, major axis mean	Single-cell
MidSOX2 nbrhd, major axis stdev	Single-cell
MidInter nbrhd, major axis mean	Single-cell
MidInter nbrhd, major axis stdev	Single-cell
DN nbrhd, axis ratio mean	Single-cell
DN nbrhd, axis ratio stdev	Single-cell
SOX2 nbrhd, axis ratio mean	Single-cell
SOX2 nbrhd, axis ratio stdev	Single-cell
TBR1 nbrhd, axis ratio mean	Single-cell
TBR1 nbrhd, axis ratio stdev	Single-cell
DP nbrhd, axis ratio mean	Single-cell
DP nbrhd, axis ratio stdev	Single-cell
MidTBR1 nbrhd, axis ratio mean	Single-cell
MidTBR1 nbrhd, axis ratio stdev	Single-cell
MidSOX2 nbrhd, axis ratio mean	Single-cell
MidSOX2 nbrhd, axis ratio stdev	Single-cell

SCOUT Feature	Scale
MidInter nbrhd, axis ratio mean	Single-cell
MidInter nbrhd, axis ratio stdev	Single-cell
DN nbrhd, sox2 proximity to SOX2 mean	Single-cell
DN nbrhd, sox2 proximity to SOX2 stdev	Single-cell
DN nbrhd, sox2 proximity to TBR1 mean	Single-cell
DN nbrhd, sox2 proximity to TBR1 stdev	Single-cell
DN nbrhd, tbr1 proximity to SOX2 mean	Single-cell
DN nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
DN nbrhd, tbr1 proximity to TBR1 mean	Single-cell
DN nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
DN nbrhd, dn proximity to SOX2 mean	Single-cell
DN nbrhd, dn proximity to SOX2 stdev	Single-cell
DN nbrhd, dn proximity to TBR1 mean	Single-cell
DN nbrhd, dn proximity to TBR1 stdev	Single-cell
SOX2 nbrhd, sox2 proximity to SOX2 mean	Single-cell
SOX2 nbrhd, sox2 proximity to SOX2 stdev	Single-cell
SOX2 nbrhd, sox2 proximity to TBR1 mean	Single-cell
SOX2 nbrhd, sox2 proximity to TBR1 stdev	Single-cell
SOX2 nbrhd, tbr1 proximity to SOX2 mean	Single-cell
SOX2 nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
SOX2 nbrhd, tbr1 proximity to TBR1 mean	Single-cell
SOX2 nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
SOX2 nbrhd, dn proximity to SOX2 mean	Single-cell
SOX2 nbrhd, dn proximity to SOX2 stdev	Single-cell
SOX2 nbrhd, dn proximity to TBR1 mean	Single-cell
SOX2 nbrhd, dn proximity to TBR1 stdev	Single-cell
TBR1 nbrhd, sox2 proximity to SOX2 mean	Single-cell
TBR1 nbrhd, sox2 proximity to SOX2 stdev	Single-cell
TBR1 nbrhd, sox2 proximity to TBR1 mean	Single-cell
TBR1 nbrhd, sox2 proximity to TBR1 stdev	Single-cell
TBR1 nbrhd, tbr1 proximity to SOX2 mean	Single-cell
TBR1 nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
TBR1 nbrhd, tbr1 proximity to TBR1 mean	Single-cell
TBR1 nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
TBR1 nbrhd, dn proximity to SOX2 mean	Single-cell
TBR1 nbrhd, dn proximity to SOX2 stdev	Single-cell
TBR1 nbrhd, dn proximity to TBR1 mean	Single-cell

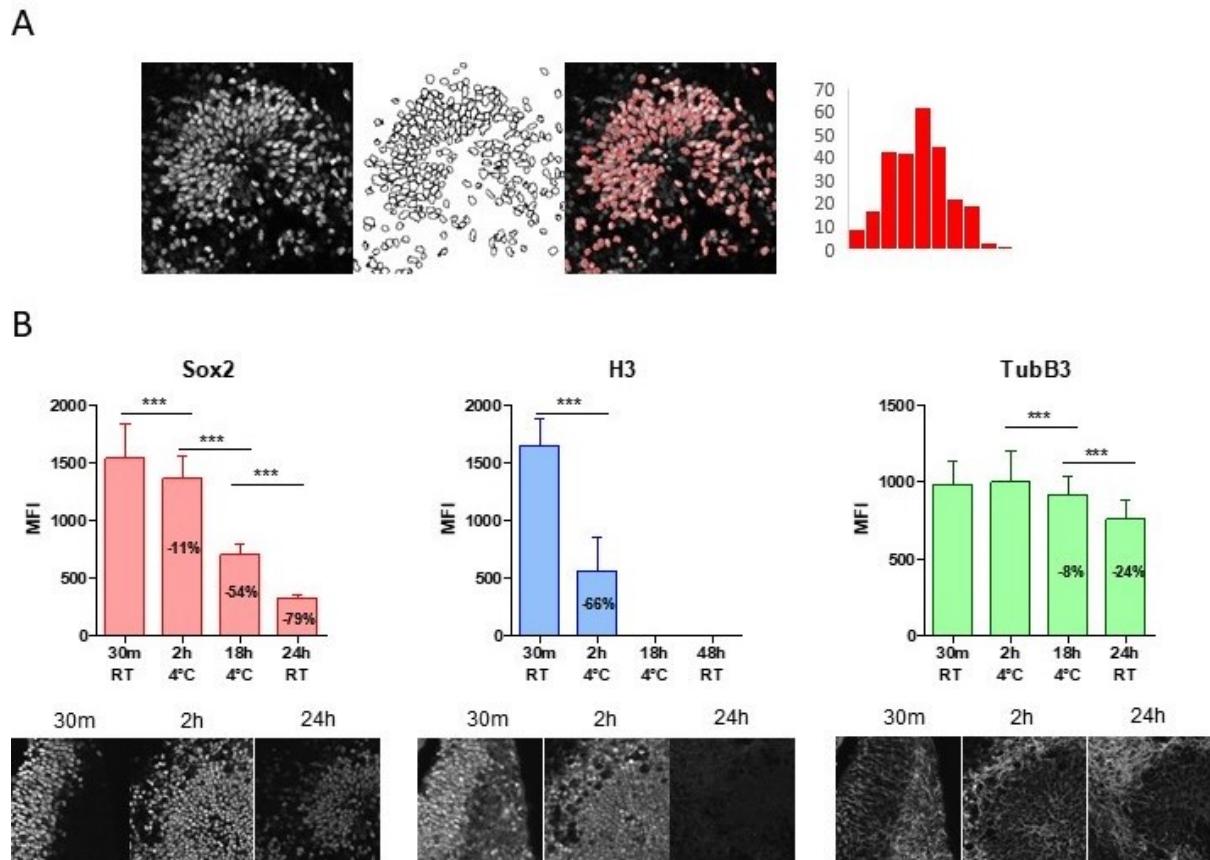
SCOUT Feature	Scale
TBR1 nbrhd, dn proximity to TBR1 stdev	Single-cell
DP nbrhd, sox2 proximity to SOX2 mean	Single-cell
DP nbrhd, sox2 proximity to SOX2 stdev	Single-cell
DP nbrhd, sox2 proximity to TBR1 mean	Single-cell
DP nbrhd, sox2 proximity to TBR1 stdev	Single-cell
DP nbrhd, tbr1 proximity to SOX2 mean	Single-cell
DP nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
DP nbrhd, tbr1 proximity to TBR1 mean	Single-cell
DP nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
DP nbrhd, dn proximity to SOX2 mean	Single-cell
DP nbrhd, dn proximity to SOX2 stdev	Single-cell
DP nbrhd, dn proximity to TBR1 mean	Single-cell
DP nbrhd, dn proximity to TBR1 stdev	Single-cell
MidTBR1 nbrhd, sox2 proximity to SOX2 mean	Single-cell
MidTBR1 nbrhd, sox2 proximity to SOX2 stdev	Single-cell
MidTBR1 nbrhd, sox2 proximity to TBR1 mean	Single-cell
MidTBR1 nbrhd, sox2 proximity to TBR1 stdev	Single-cell
MidTBR1 nbrhd, tbr1 proximity to SOX2 mean	Single-cell
MidTBR1 nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
MidTBR1 nbrhd, tbr1 proximity to TBR1 mean	Single-cell
MidTBR1 nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
MidTBR1 nbrhd, dn proximity to SOX2 mean	Single-cell
MidTBR1 nbrhd, dn proximity to SOX2 stdev	Single-cell
MidTBR1 nbrhd, dn proximity to TBR1 mean	Single-cell
MidTBR1 nbrhd, dn proximity to TBR1 stdev	Single-cell
MidSOX2 nbrhd, sox2 proximity to SOX2 mean	Single-cell
MidSOX2 nbrhd, sox2 proximity to SOX2 stdev	Single-cell
MidSOX2 nbrhd, sox2 proximity to TBR1 mean	Single-cell
MidSOX2 nbrhd, sox2 proximity to TBR1 stdev	Single-cell
MidSOX2 nbrhd, tbr1 proximity to SOX2 mean	Single-cell
MidSOX2 nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
MidSOX2 nbrhd, tbr1 proximity to TBR1 mean	Single-cell
MidSOX2 nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
MidSOX2 nbrhd, dn proximity to SOX2 mean	Single-cell
MidSOX2 nbrhd, dn proximity to SOX2 stdev	Single-cell
MidSOX2 nbrhd, dn proximity to TBR1 mean	Single-cell
MidSOX2 nbrhd, dn proximity to TBR1 stdev	Single-cell

SCOUT Feature	Scale
MidInter nbrhd, sox2 proximity to SOX2 mean	Single-cell
MidInter nbrhd, sox2 proximity to SOX2 stdev	Single-cell
MidInter nbrhd, sox2 proximity to TBR1 mean	Single-cell
MidInter nbrhd, sox2 proximity to TBR1 stdev	Single-cell
MidInter nbrhd, tbr1 proximity to SOX2 mean	Single-cell
MidInter nbrhd, tbr1 proximity to SOX2 stdev	Single-cell
MidInter nbrhd, tbr1 proximity to TBR1 mean	Single-cell
MidInter nbrhd, tbr1 proximity to TBR1 stdev	Single-cell
MidInter nbrhd, dn proximity to SOX2 mean	Single-cell
MidInter nbrhd, dn proximity to SOX2 stdev	Single-cell
MidInter nbrhd, dn proximity to TBR1 mean	Single-cell
MidInter nbrhd, dn proximity to TBR1 stdev	Single-cell
TBR1-LowDN Cytoarchitecture fraction	Cytoarchitecture
TBR1-HighDN Cytoarchitecture fraction	Cytoarchitecture
Surface Cytoarchitecture fraction	Cytoarchitecture
Artifacts Cytoarchitecture fraction	Cytoarchitecture
DN Cytoarchitecture fraction	Cytoarchitecture
Adjacent Cytoarchitecture fraction	Cytoarchitecture
ave. TBR1-LowDN profile, sox2 count	Cytoarchitecture
ave. TBR1-LowDN profile, tbr1 count	Cytoarchitecture
ave. TBR1-LowDN profile, dn count	Cytoarchitecture
ave. TBR1-HighDN profile, sox2 count	Cytoarchitecture
ave. TBR1-HighDN profile, tbr1 count	Cytoarchitecture
ave. TBR1-HighDN profile, dn count	Cytoarchitecture
ave. Surface profile, sox2 count	Cytoarchitecture
ave. Surface profile, tbr1 count	Cytoarchitecture
ave. Surface profile, dn count	Cytoarchitecture
ave. Artifacts profile, sox2 count	Cytoarchitecture
ave. Artifacts profile, tbr1 count	Cytoarchitecture
ave. Artifacts profile, dn count	Cytoarchitecture
ave. DN profile, sox2 count	Cytoarchitecture
ave. DN profile, tbr1 count	Cytoarchitecture
ave. DN profile, dn count	Cytoarchitecture
ave. Adjacent profile, sox2 count	Cytoarchitecture
ave. Adjacent profile, tbr1 count	Cytoarchitecture
ave. Adjacent profile, dn count	Cytoarchitecture
ave. TBR1-LowDN profile, sox2 mean position	Cytoarchitecture

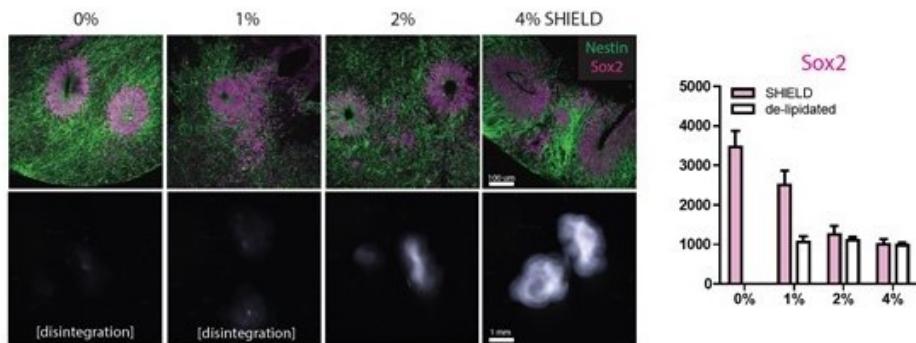
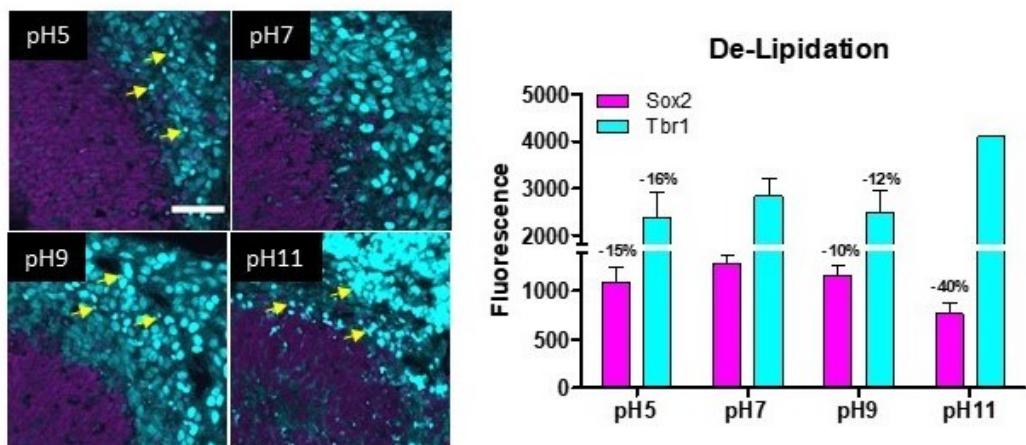
SCOUT Feature	Scale
ave. TBR1-LowDN profile, tbr1 mean position	Cytoarchitecture
ave. TBR1-LowDN profile, dn mean position	Cytoarchitecture
ave. TBR1-HighDN profile, sox2 mean position	Cytoarchitecture
ave. TBR1-HighDN profile, tbr1 mean position	Cytoarchitecture
ave. TBR1-HighDN profile, dn mean position	Cytoarchitecture
ave. Surface profile, sox2 mean position	Cytoarchitecture
ave. Surface profile, tbr1 mean position	Cytoarchitecture
ave. Surface profile, dn mean position	Cytoarchitecture
ave. Artifacts profile, sox2 mean position	Cytoarchitecture
ave. Artifacts profile, tbr1 mean position	Cytoarchitecture
ave. Artifacts profile, dn mean position	Cytoarchitecture
ave. DN profile, sox2 mean position	Cytoarchitecture
ave. DN profile, tbr1 mean position	Cytoarchitecture
ave. DN profile, dn mean position	Cytoarchitecture
ave. Adjacent profile, sox2 mean position	Cytoarchitecture
ave. Adjacent profile, tbr1 mean position	Cytoarchitecture
ave. Adjacent profile, dn mean position	Cytoarchitecture
ave. TBR1-LowDN profile, sox2 profile stdev	Cytoarchitecture
ave. TBR1-LowDN profile, tbr1 profile stdev	Cytoarchitecture
ave. TBR1-LowDN profile, dn profile stdev	Cytoarchitecture
ave. TBR1-HighDN profile, sox2 profile stdev	Cytoarchitecture
ave. TBR1-HighDN profile, tbr1 profile stdev	Cytoarchitecture
ave. TBR1-HighDN profile, dn profile stdev	Cytoarchitecture
ave. Surface profile, sox2 profile stdev	Cytoarchitecture
ave. Surface profile, tbr1 profile stdev	Cytoarchitecture
ave. Surface profile, dn profile stdev	Cytoarchitecture
ave. Artifacts profile, sox2 profile stdev	Cytoarchitecture
ave. Artifacts profile, tbr1 profile stdev	Cytoarchitecture
ave. Artifacts profile, dn profile stdev	Cytoarchitecture
ave. DN profile, sox2 profile stdev	Cytoarchitecture
ave. DN profile, tbr1 profile stdev	Cytoarchitecture
ave. DN profile, dn profile stdev	Cytoarchitecture
ave. Adjacent profile, sox2 profile stdev	Cytoarchitecture
ave. Adjacent profile, tbr1 profile stdev	Cytoarchitecture
ave. Adjacent profile, dn profile stdev	Cytoarchitecture
organoid volume (mm <sup>3</sup> )	Whole-organoid
organoid equivalent diameter (mm)	Whole-organoid

SCOUT Feature	Scale
organoid major axis (mm)	Whole-organoid
organoid minor axis (mm)	Whole-organoid
organoid axis ratio	Whole-organoid
ventricle count	Whole-organoid
ventricle volume mean (um3)	Whole-organoid
ventricle volume stdev (um3)	Whole-organoid
ventricle equivalent diameter mean (um)	Whole-organoid
ventricle equivalent diameter stdev (um)	Whole-organoid
ventricle major axis mean (um)	Whole-organoid
ventricle major axis stdev (um)	Whole-organoid
ventricle minor axis mean (um)	Whole-organoid
ventricle minor axis stdev (um)	Whole-organoid
ventricle axis ratio mean	Whole-organoid
ventricle axis ratio stdev	Whole-organoid
DN nbrhd, sox2 surface distance mean (um)	Whole-organoid
DN nbrhd, sox2 surface distance stdev (um)	Whole-organoid
DN nbrhd, tbr1 surface distance mean (um)	Whole-organoid
DN nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
DN nbrhd, dn surface distance mean (um)	Whole-organoid
DN nbrhd, dn surface distance stdev (um)	Whole-organoid
SOX2 nbrhd, sox2 surface distance mean (um)	Whole-organoid
SOX2 nbrhd, sox2 surface distance stdev (um)	Whole-organoid
SOX2 nbrhd, tbr1 surface distance mean (um)	Whole-organoid
SOX2 nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
SOX2 nbrhd, dn surface distance mean (um)	Whole-organoid
SOX2 nbrhd, dn surface distance stdev (um)	Whole-organoid
TBR1 nbrhd, sox2 surface distance mean (um)	Whole-organoid
TBR1 nbrhd, sox2 surface distance stdev (um)	Whole-organoid
TBR1 nbrhd, tbr1 surface distance mean (um)	Whole-organoid
TBR1 nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
TBR1 nbrhd, dn surface distance mean (um)	Whole-organoid
TBR1 nbrhd, dn surface distance stdev (um)	Whole-organoid
DP nbrhd, sox2 surface distance mean (um)	Whole-organoid
DP nbrhd, sox2 surface distance stdev (um)	Whole-organoid
DP nbrhd, tbr1 surface distance mean (um)	Whole-organoid
DP nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
DP nbrhd, dn surface distance mean (um)	Whole-organoid

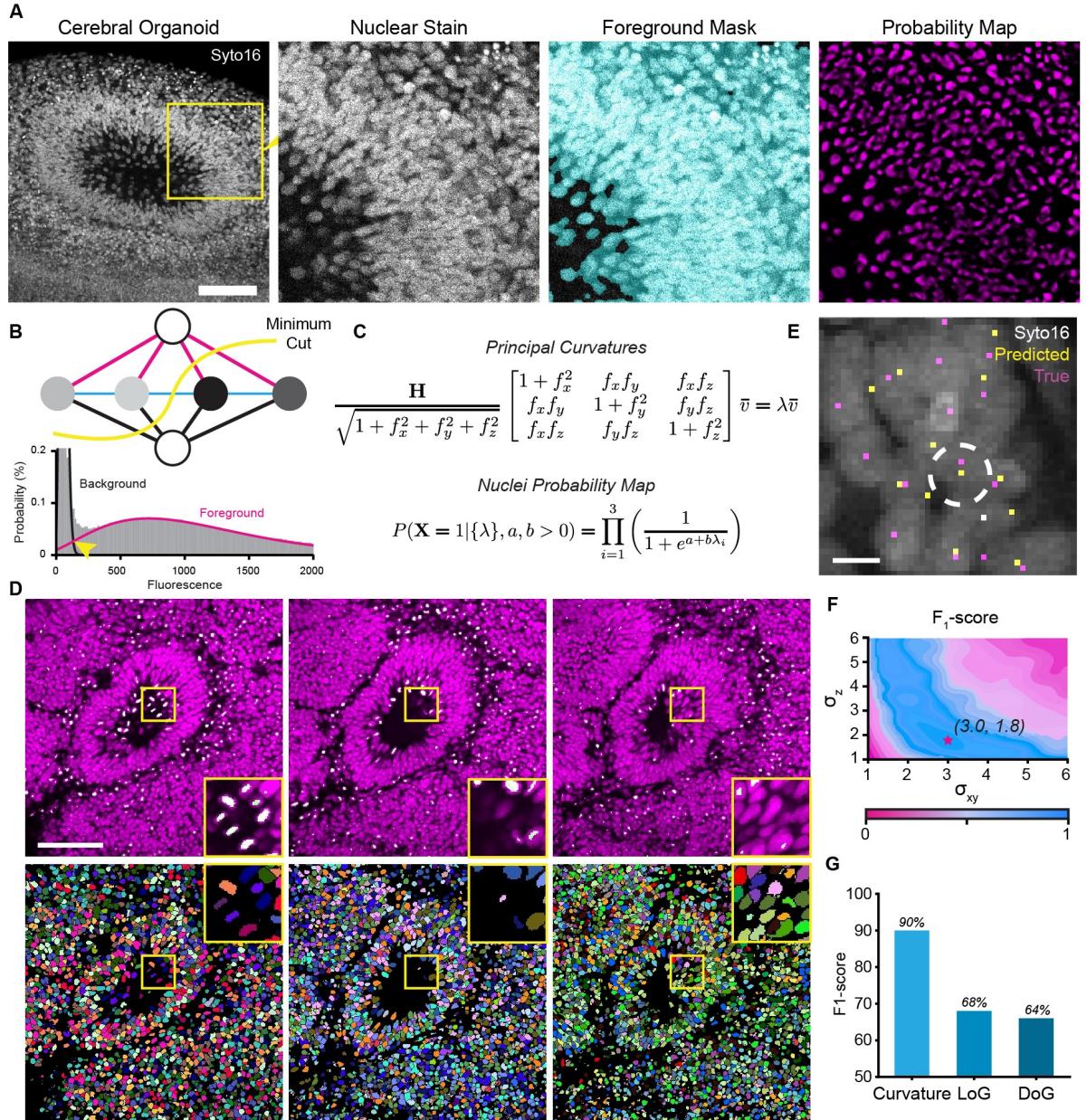
SCOUT Feature	Scale
DP nbrhd, dn surface distance stdev (um)	Whole-organoid
MidTBR1 nbrhd, sox2 surface distance mean (um)	Whole-organoid
MidTBR1 nbrhd, sox2 surface distance stdev (um)	Whole-organoid
MidTBR1 nbrhd, tbr1 surface distance mean (um)	Whole-organoid
MidTBR1 nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
MidTBR1 nbrhd, dn surface distance mean (um)	Whole-organoid
MidTBR1 nbrhd, dn surface distance stdev (um)	Whole-organoid
MidSOX2 nbrhd, sox2 surface distance mean (um)	Whole-organoid
MidSOX2 nbrhd, sox2 surface distance stdev (um)	Whole-organoid
MidSOX2 nbrhd, tbr1 surface distance mean (um)	Whole-organoid
MidSOX2 nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
MidSOX2 nbrhd, dn surface distance mean (um)	Whole-organoid
MidSOX2 nbrhd, dn surface distance stdev (um)	Whole-organoid
MidInter nbrhd, sox2 surface distance mean (um)	Whole-organoid
MidInter nbrhd, tbr1 surface distance mean (um)	Whole-organoid
MidInter nbrhd, tbr1 surface distance stdev (um)	Whole-organoid
MidInter nbrhd, dn surface distance mean (um)	Whole-organoid
MidInter nbrhd, dn surface distance stdev (um)	Whole-organoid



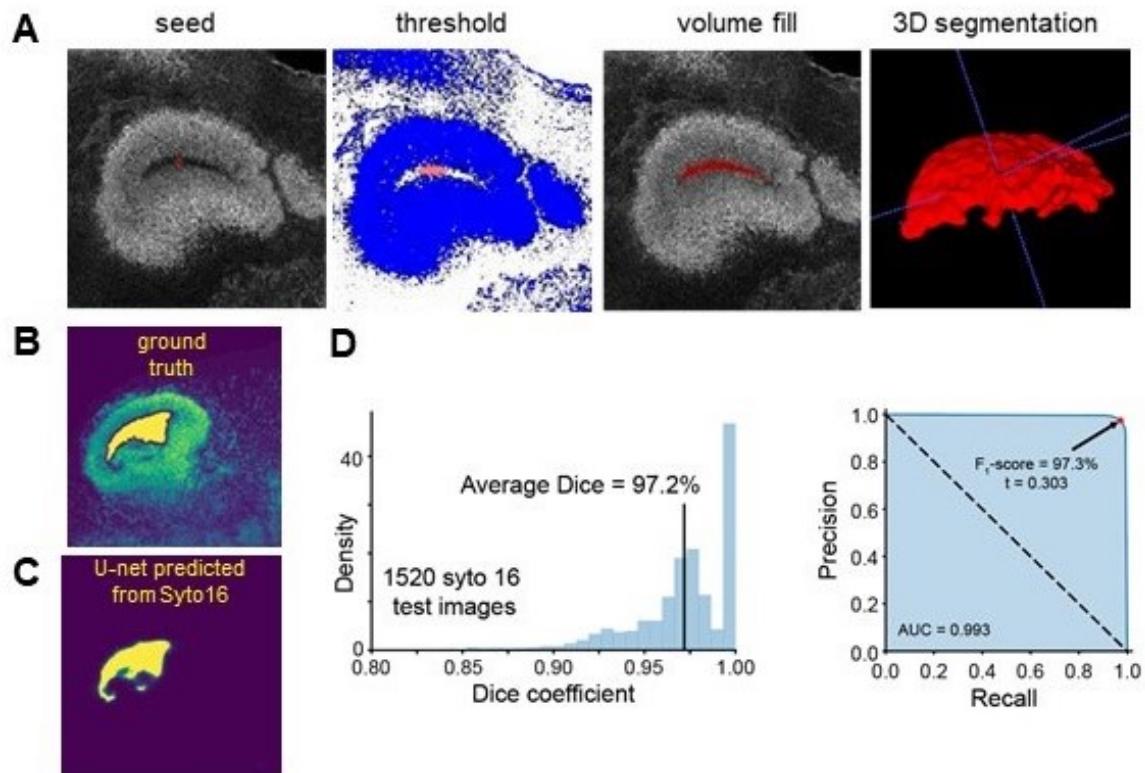
**Figure 8.10:** Optimization of aldehyde fixation for organoids: (A) Methodology for quantifying mean fluorescence intensity for individual cells. ROI regions were generated in FIJI by (i) auto-threshold of images, (ii) analyze particles and watershed (iii) generate mask to analyze greyscale image.(B) Evaluation of organoid fixation on antibody staining using protocols widely used in the literature. Data shows that nuclear epitopes Sox2 and Histone H3 are sensitive to over-fixation and show dramatic signal decrease. We selected a fixation in PBS + 4% PFA for 30 min at room temperature.

**A****B**

**Figure 8.11:** Optimization of SHIELD protocol for organoids: (A) Organoids fixed with 4% PFA in PBS for 30 min at room temperature. Organoids were then incubated in a solution of polyepoxy compound polyglycerol-3-polyglycidyl ether (P3PE) in 0.1M phosphate buffer pH7 at 4°C. Samples were incubated in this solution for 48h to ensure tissue saturation with P3PE, then solution was removed and replaced with pre-warmed 0.1M sodium carbonate buffer (pH10) and incubated at 37°C to activate covalent bonds between epoxide groups and available amines in the tissue. Samples were then cleared in 0.2M SDS, 50 mM boric acid pH9.5 at 55°C for 48h. Samples were washed extensively and stained with anti-Sox2 (Millipore AB5603) and anti-Nestin (Millipore MAB5326). Data shows that below 2% SHIELD (poly-epoxide) samples disintegrated and became fragile. We selected 2% since it preserved samples and maintained anti-Sox2 signal. (B) Optimization of clearing solution shows that at low and high pH, antibody staining is reduced for Sox2 (CST # 5067S) and Tbr1 (CST #32518S), we also saw the appearance of non-specific Tbr1+ spots next to positive cells at high pH. We selected a 0.2M SDS solution at pH7 using 50 mM phosphate buffers.



**Figure 8.12:** Nuclei detection and segmentation accuracy: (A) Nuclei detection and semantic segmentation of a syto16 stained cerebral organoid. The nuclei seed points are used in a 3D watershed segmentation to partition the foreground mask among all the nuclei detected as peaks in the probability map. Scale bar, 100  $\mu$ m (B) Illustration of a minimum cut on an undirected graph used for nuclei semantic segmentation. Log-likelihood penalties were constructed using a mixture of Poisson distributions estimated from the image histogram. (C) Analytical definitions of the shape operator and logistic regression function used to calculate nuclei principal curvatures and probability maps, respectively. (D) 2D slices and nuclei segmentations for a syto16-stained cerebral organoid. Scale bar, 100  $\mu$ m. (E) Nuclei detection accuracy assessed by matching with manually annotated nuclei centroids. Scale bar, 10  $\mu$ m. (F) Tuning hyperparameters for gaussian filtering by maximizing  $F_1$  score. (G) Test set  $F_1$  scores from our technique compared to Laplacian (LoG) and Difference (DoG) of Gaussian blob detectors.



**Figure 8.13:** Ventricle segmentation and U-Net model validation: (A) Example workflow for semi-automatic active-contour segmentation of ventricles using ITK-Snap. A spherical seed volume is iteratively dilated within the masked region obtained from thresholding the SOX2 image. After dilation, the ventricle completely filled and the segmentation can be rendered in 3D. (B) Overlay of a 2D syto16 image with the associated ventricle segmentation from the ground-truth 3D segmentation. (C) Example ventricle probability map output from our U-Net model. These 2D probability maps contain floating point values between 0 and 1. (D) Test performance of ventricle segmentation U-Net based on a 20% hold-out test set (1520 images). The Dice coefficient was computed for each test image resulting in an average of 97.2%. By sweeping the probability threshold used to convert the U-Net output probability maps into binary masks from 0 to 1, the model precision and recall were calculated to construct the ROC curve.