Long-Range Thalamic Projections To Neocortex and Basal Ganglia: A Model for the Mesoscopic Connectome

By

Deniz S. Kusefoglu

A MASTER'S THESIS

Presented to the Neuroscience Graduate Program
and the Oregon Health and Science University
School of Medicine

In partial fulfillment of
the requirements of the degree of

Master of Science in Neuroscience
June 2013

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the M.S. dissertation of

DENIZ KUSEFOGLU

has been approved on June 6, 2013

[redacted]

Advisor, Tianyi Mao, Ph.D.

[redacted]

Member and Chair, Gary Westbrook, M.D.

[redacted]

Member, Vivek Unni, M.D., Ph.D.

[redacted]

Member, Claudio Mello, Ph.D.

TABLE OF CONTENTS

Acknowledgements

I'd like to thank my mentor Tianyi Mao for taking me on board, and investing her time and resources with me. Thank you Karel Svoboda, for being nice, and for dreaming up and supporting this project. Thank you Gary Westbrook for many conversations and supportive words.

Thanks to colleagues at Janelia for a true experience. Thank you Julide Bilen for caring. Thank you Malte Kramer for sitting with me.

Thank you Jeannie Hunnicutt for being a role model of optimism and patience in the face of an unknown amount of work. Thank you Brian Long for his work and time closing out this project after me. Thank you Shane Tillo for being a friend.

Thank you Allen Andrews, for something.

Thanks to my parents Selim Kusefoglu and Banu Oney, and my grandmother Necla Oney for being there, and understanding.

I dedicate this thesis to Fehime Kusefoglu, who was a great listener, and had great heart.

May we forgive all hope for a better past!…

Long-Range Thalamic Projections To Neocortex and Basal Ganglia: A Model for the Mesoscopic Connectome

One of the largest subcortical structures in the mammalian brain, the thalamus transmits peripheral sensory inputs to the cortex and striatum and acts as a relay station for motor circuits. Here we describe a method for the comprehensive description of projections emanating from the thalamus, with the goal of using differences in projection patterns to suggest a structural delineation of thalamic nuclei. We used multi-color fluorescent proteins expressed in small subsets of cells in the thalamus, delivered through stereotaxic injections, in order to achieve precise and complete coverage of the volume of the thalamus. Labeled brains were then serially sectioned and imaged. Using Matlab, the image of each slice was registered through rigid body transformations into a 3-dimensional structure, and the volume of each brain was normalized to the population average. Reconstructed brain images were manually segmented into regions of interest in the striatum and cortex. Each injection was scored on the basis of its anatomical connectivity. A method to use this scoring method to parcellate thalamic nuclei is described. Maps of the long-range connectivity of groups of cells, with the mesoscopic connectome method described here, will permit the investigation of complex neural circuits underlying the activity of different brain regions.

Chapter 1. Introduction

The connectome is the description of the circuit architecture of the brain

The nervous system is composed of cells that are structurally and functionally interconnected within a complex circuitry. The fine structure of the long distance anatomical connectivity between neurons underlies the dissemination of information between distant regions of the brain. From the perspective of the behaving organism, the missing link between incoming sensory information and the generation of appropriate behavioral response is a series of transformations in which the signals encoding sensory inputs are relayed to the appropriate computational centers and contrasted with competing sensory information or internal states. Although much work has been done to describe the structural architecture of the nervous system, a complete description of neuronal connections, -the connectome- is unknown for most model organisms. The characterization of the features of the connectome will provide new mechanistic insights into how the disruption of this organization affect brain function in disease states. To address this issue, the experiments outlined in this thesis illustrate a method to describe the projections emanating from the mouse thalamus at a novel completeness, and resolution. This method will constitute a prototype for future whole brain connectivity studies, and be expandable to other brain regions towards the construction of the connectome.

Approaches to the connectome are reflective of the available technologies, and are microscopic or macroscopic in scale

From the early days of the study of the brain, observations on the anatomical structure of brain cells and their connections have yielded important clues to how the brain performs its myriad complex tasks. Approaches to the connectome over time have reflected available technologies in labeling, imaging and data analysis methods. In the early 1900's Santiago Ramón y Cajal used the Golgi method to sparsely label neurons, allowing them to be visualized against the otherwise indistinguishable mass of brain (Ramón y Cajal, 1904). Using light microscopy to magnify his samples, Cajal assembled more than a thousand manual drawings of neurons and their processes. He was able to generate meaningful interpretations of the observed structure, correctly guessing the direction of signal travel in neurites, among other subjects.

More recently, neuronal connectivity has been studied and described at three distinguishable levels of organization: the macroscopic, microscopic and mesoscopic scales. At the macroscopic scale, the brain has been subdivided into anatomically distinguishable macro regions such as the cortex, cerebellum, brain stem, etc. The boundaries of and connections between these macro regions have been investigated through methods such as dissection, histological staining, degeneration methods, and the tracing of projections with tracer substances (Köbbert et al., 2000). These methods have been effective at revealing the most prominent landmarks of macro brain regions, and confirming the existence of a portion of their connections. Most of the image data annotation and quantification has been manual. The Brain Architecture Management System (BAMS http://brancusi.usc.edu/bkms) assembles anatomical connectivity data

emanating from tract tracing literature into a searchable, publicly accessible database (Bota et al., 2005). However, a recent appraisal of the rat brain connectivity data available through BAMS underlined that only about 10% of possible connections have been investigated. (Bohland et al., 2009) The collation of data is also hindered by the lack of consistency inherent to manual image annotation, the lack of a shared anatomical grid or consistent nomenclature between the studies. The multiplicity of tissue preparation methods also make it difficult for the generated data to be directly integrated.

More recently, non-invasive neuroimaging methods such as functional magnetic resonance imaging (fMRI) have been used to study human brain connectivity at a macroscopic, whole brain level. Resting state fMRI (R-fMRI), in which brain activity is recorded while subjects rest quietly, exhibits large-amplitude, low-frequency spontaneous fluctuations in the fMRI signal. Temporal correlations in these fluctuations are interpreted as indicative of functional connectivity between distant brain areas (Biswal et al., 2010). R-fMRI data collected through the cooperation of multiple institutions, is compiled and made freely available under the umbrella of the 1000 functional connectomes project (http://fcon_1000.projects.nitrc.org). The Human Connectome Project (http://www.neuroscienceblueprint.nih.gov/connectome) is another macroscopic anatomy effort currently under way. This project involves the use of diffusion MRI (dMRI), or diffusion tensor imaging (DTI), which is conducted with powerful magnets, and involves tracking the movement of water molecules- exploiting the anisotropy of water diffusion in white matter to delineate the location of fiber tracts and boundaries of brain regions (Hagmann et al., 2008). Macroscale anatomical methods succeed in describing whole brain connectivity, but are limited in their spatial resolution.

At the microscopic scale, the anatomy of single neurons and their synapses are investigated. Electron microscopy (EM) is currently the only approach with sufficient spatial resolution to image multiple axonal processes and dendritic spines over tissue volumes. However, the sheer magnitude of data generated by this process and the complexity of 3D reconstruction of images of ultrathin serial sections (Chklovskii et al., 2010) hampers efforts to scale up this process to large organisms. In fact, nearly thirty years after its publication, the mapping of the cellular connections in *Caenorhabditis elegans* (White et al., 1986) remains the only complete connectome of any organism generated at the microscopic scale. The analysis of this data set has generated important insights into the principles of wiring economy (Chen et al., 2006) and the topology of connections, (Varshney et al., 2009) principles that continue to be investigated in other organisms. The 3D Segmentation of Neurites in EM Images (SNEMI3D) challenge was launched this year to spur interest in the generation of machine learning algorithms to automate the segmentation of neuronal EM data. Although EM reconstructions offer an unparalleled method of visualizing the fine structure of neurites, the number of cells which compose the brain and the lack of efficient automated segmentation methods indicate that a complete connectome at the EM level will remain computationally out of reach for some time. Recent two photon imaging experiments demonstrating the rapid appearance and disappearance of dendritic spines (Bonhoffer and Yuste, 2002), also suggests that neuronal connectivity may be quite dynamic at the microscopic scale.

The mesoscopic approach must achieve a high level of completeness and must be high resolution

In between the macro and micro scales, the mesoscopic approach to brain connectivity is capable of describing the connectome at a novel and unparalleled combination of completeness and resolution. The desirable features of this approach are enumerated here. The mesoscopic approach labels and renders contrast on the projections of **(1) spatially grouped neurons**, on the order of hundreds of micrometers. Spatially grouped labeling can be expected to reveal the canonical and stereotyped circuitry of units of anatomically or functionally distinct neuronal populations, independent of the expected variance in neuron to neuron connections.

It is also important that the method describes the projections of neurons as they travel through and terminate within far away brain regions; imaging on the scale of the **(2) whole brain completeness**. An established whole brain imaging pipeline will produce expectation-independent, complete descriptions of the relationship between labeling origin and projection termination. Whole brain imaging will also allow neuronal trajectories to be reliably reconstructed in their entirety. Structurally differentiating en-passant versus terminating projections will inform the functional interpretation of connectivity data. Whole brain landmarks will facilitate the fitting of connectivity data into a common **(3) anatomical coordinate system**, such that spatial relationships between separate labeling samplings may be revealed, and the topographic relationship between labeling origin and projection distribution may be discerned. A common anatomical coordinate system will allow for data to be integrated across sampling iterations, and registered against existing brain wide gene expression, neuronal cell type distribution, and brain atlas data.

A high resolution in imaging, achieving **(4) individual axon resolution** will

allow the tracing of neurites to their location of termination even in regions where projections are sparse. This mesoscopic approach will aim for a **(5) complete sampling** of each macro brain region of origin. Characterizing the entirety of projections will rule out unlabeled regions as potential projection targets, describing a positive and negative projection space. A level of **(6) redundancy of sampling** in labeling regions will serve to confirm intra-subject regularity of projections. Groups of neurons are being probed, so redundancy in the form of labeling overlap will further subdivide the region being studied. A **(7) Multiplicity of comparable labeling methods** will reveal the topography of the projections within one brain. It is desirable that this method achieve high resolution and completeness whilst generating a **(8) reasonable volume of data** that is within reach of current information technology to analyze, display and render widely available for querying and further analysis.

Finally, it is intended that this method become a template for **(9) high throughput** data generation. Accordingly, the methods of cell labeling and tissue preparation must be relatively approachable, reproducible and error-resistant. It is important that methods and data generation be **(10) scaled up** to generate comparable sets of data for brain regions. The extension of the approach to multiple brain regions is important to reveal whether regions connected in one direction are also connected reciprocally. Complete methods become canons in literature. The Paxinos Atlas (Paxinos and Franklin, 2004) is an example of a complete anatomical description effort that, while laborious, was essential for the establishment of a common anatomical nomenclature among investigators. The Allen Brain Atlas for gene expression (Lein et al., 2007) has shown that the scaling up and repeated application of a process can result in a significant

resource to the neuroscience community.

Viral mediated dye delivery and sectioning in the mouse brain constitute a feasible mesoscopic connectome approach

With these guidelines for a fruitful and approachable method of enquiry established, the mesoscopic approach that achieves these objectives will greatly contribute to the current understanding of the connectome. The method I describe here, involves the use of recombinant adeno-associate viral vectors (AAV) to deliver enhanced green fluorescent protein (eGFP) or the red fluorescent tdTomato genes into groups of neurons. Stereotactically delivered microinjections of AAV are capable of reliably labeling anterograde projections, with virtually no retrograde transport, and low tissue damage (Chamberlin et al., 1998). The fluorescent tracer (eGFP, or tdT) fills the cell body, axon and nerve terminals entirely, yielding a complete description of monosynaptic projections, even to targets that are millimeters away from the cell body. The labeling is robust, and fluorescence is preserved after tissue fixation. The availability of multicolor fluorescent markers enabled me to perform bicolor injections. Bicolor injections performed side by side are able to reveal the topography of projections. Using AAV injections I was able to achieve a systematic grid of labeling in the target region.

The mouse is a natural model organism for the generation of this data set, due to its widespread of use in the neuroscience community, and the great range of data available from single cell activity to behavioral studies, along with the existence of alternate whole brain characterizations, such as The Allen Brain Atlas (ABA) for gene expression in the mouse. The existence of transgenic mouse lines also creates options for

the targeting of specific cell types in future iterations of this method. The size of the mouse brain is sufficiently small that whole brain high resolution imaging still produces tractable amounts of data, approximately half a terabyte per brain under our methods.

Following AAV labeling, in order to image fine structures at high resolution, the brain tissue is fixed and serially sectioned on the cryostat. The technological development that has perhaps been the most instrumental in the feasibility of this method has been breakthroughs in automated virtual microscopy. The Olympus Nanozoomer (Hamamatsu Photonics, Japan) was developed for automated pathology, but is suitable for this method due to its multichannel fluorescence microscopy capability. The Nanozoomer batch processes 100's of slides, mechanically delivering and removing them to and from the microscope. Samples on the slide are detected, and the plane of focus is determined automatically from initial brightfield images. Slides are imaged line by line, while they are accurately moved under the CCD camera- this serves to limit optical aberration along the scan line, and standardizes and simplifies the reconstruction of the whole slide image. The imaging of a slide containing 6-8 sections took approximately one hour. Using the Nanozoomer, we were able to automate the high-throughput acquisition of high resolution, high magnification images of the sections spanning the whole brain.

The thalamus is a well-suited prototype region for the mesoscopic connectome

In this thesis, I describe a method to comprehensively describe the anatomy of projections emanating from the thalamus at a mesoscopic scale. One reason the thalamus is a natural beginning point for this investigation is that the thalamus receives sensory information in a topographically ordered manner, such that similar features of sensory

information are kept anatomically adjacent. The complete description of thalamic projections, as described in this thesis, reveals principles of how the topographic organization of sensory information is retained or altered in higher order brain regions. The principal targets of thalamic projections are the striatum and the cortex, with some overlap in the origins of projections to the two regions the anatomy of these connections have not been comprehensively described.

The study of the organization of the striatum stands to benefit from the description of thalamic projections. The striatum is the principal input nucleus to the basal ganglia, which is critically involved in the generation and execution of motivated behavior. Despite its size, the striatum is striking in its apparent lack of architectonic subdivisions. About 95 % of cells in the striatum are medium spiny neurons (MSNs) which receive direct inputs from the cortex and thalamus, interact within the local striatal microcircuit, and relay to downstream nuclei of the basal ganglia (Steiner and Tseng, 2010). Measures of ensemble MSN activity in awake, behaving animals have demonstrated functional specializations between different regions of the striatum. Given the homogeneity of cell type distribution in the striatum, it must be the distribution of inputs from the thalamus and the cortex, rather than local integrative properties that establish functional differences between regions of the striatum. There is an increase of cell activity in the dorsolateral striatum as performance on a task improves and as behavior becomes highly stereotyped and habitual (Thorn et al., 2010). Habitual behavior features a tight temporal coupling between stimulus and response, suggesting a strong connection between the dorsolateral striatum and sensory and motor areas of the thalamus and cortex. In contrast, a behavior is considered ‘goal-directed’ rather than habitual if it is

closely coupled and responsive to the positive or negative consequences of the action, requiring a broader sense of context. Dorsomedial striatum is preferentially engaged during goal-directed behaviors and during the initial learning stages of a task, suggesting that it is preferentially innervated by 'associative' areas of the thalamus and cortex.

What little is known of the thalamic projection patterns to the striatum is informed by data generated in primates, suggesting that the intralaminar thalamic nuclei comprise the majority of inputs to the striatum. Recent work using retrograde labeling with fluorescent latex microspheres, demonstrates that a much more broad and diverse region of the thalamus projects to the striatum. (Pan et al., 2010) Our preliminary results indicate that virtually all thalamic projections travel in bundles through the striatum on their trajectory towards their termination in the cortex, many also collateralize within the striatum, functionally integrating into the local striatal microcircuit. I discuss an automated image analysis pipeline, able to discriminate between en passant and terminating labeled projections that will be instrumental in the functional interpretation of connectivity data in target regions in general, and specifically in the striatum. With the knowledge that thalamic nuclei are highly specific in the kinds of peripheral sensory or higher order information they relay, the comprehensive description of thalamic projections using the mesoscopic anatomical approach described here will reveal the origin of input to different regions of the striatum, and contribute to a meaningful functional parcellation of the striatum.

Another field that will benefit greatly from the work described here is the study of how thalamic inputs integrate in to the cortical microcircuit. The primary functional unit of the cortex is widely accepted to be the cortical column, with six distinguishable layers

of cells that exhibit stereotyped circuitry and similar receptive fields. According to work done originally in the visual cortex, thalamic input arrives in layer 4, the excitatory cells in layer 4 project to the superficial layers, which in turn project to layer 5. The pyramidal cells in layer 5, either target layer 6 neurons, closing the loop with a projection from layer 6 to the input layer 4, or output to downstream subcortical structures and cortical areas. (Gilbert & Wiesel 1989) Although this basic architecture has been confirmed for a several cortical regions, the sub-parcellation of thalamic nuclei has allowed for the investigation of differences in the laminar profile of thalamic projections to different cortical regions. In particular, rat primary somatosensory cortex receives inputs from the ventral posteromedial (VPM) and posteromedial (POm) nuclei of the thalamus, with different laminar profiles. (Wimmer et al., 2010) A recent Channelrhodopsin-2-assisted circuit mapping (CRACM) study demonstrate two distinct kinds of input to vM1 with distinct laminar profiles. (Hooks et al., 2011) This study did not parcellate the thalamic origin of these inputs at sufficiently high resolution to distinguish individual nuclei, instead referring to these regions as "motor thalamus" – encompassing the anteromedial (AM) ventral anterior (VA) and anterior parts of the ventrolateral (VL) thalamic nuclei- and "sensory thalamus" – encompassing POm and the posterior portion of the VL. A recent single neuron tracing study performed in the rat also distinguished two types of thalamocortical projections emanating from the VA and VL thalamic nuclei: one which projects directly to deep layers of the cortex, and another which collateralizes in the striatum before terminating in superficial layers of the cortex. (Kuramoto et al., 2009) The distribution and origin of thalamocortical inputs to different layers of the cortex in different cortical sub-regions remains an open question with significant implications to

the nature of the computations carried out by cortical microcircuits. The method I describe here attempts to describe thalamic projections comprehensively and analyze the data from the perspective of several cortical regions. This method will yield not only the thalamic innervation of different regions of the striatum as described previously, but also the corresponding cortical destinations of the same projections, revealing the functional coupling between regions of the striatum and the cortex.

In this thesis, I describe and execute a method for describing the anatomy of long range thalamic projections at a mesoscopic scale, with the intention that it may constitute a prototype for high throughput, scaled up connectome projects to come. The method labels spatially grouped neurons with stereotactic injections of AAV, expressing fluorescent indicators in two colors. Dual color, bilateral injections in the thalamus allow for four independent labeling regions in each brain, sampling the mouse thalamus at a high level of completeness and redundancy. I section and image the whole brain with automated microscopy methods capable of visualizing and resolving individual axons. I describe work done towards fitting the data to a common anatomical coordinate system. I discuss approaches to further analyze and represent the anatomical connectivity data in order to render it most useful to the neuroscience community.

Chapter 2. Methods

Experimental protocols were conducted according to National Institutes of Health guidelines for animal research and were approved by the Institutional Animal Care and Use Committee at Janelia Farm Research Campus and Oregon Health & Science University. All experiments were carried out in accordance with biosafety level 1 guidelines.

Anterograde Tracing

For the anterograde tracing methods described here, we used adeno-associated virus (AAV; serotype 2/1) expressing eGFP (www.addgene.com) or tdTomato (a gift from J. Magee) under the transcriptional control of the cytomegalovirus enhancer/chicken b-actin (CAG) promoter.

Stereotactic Injections

Injections were performed on C57BL/6J mice (Charles River) of either sex, at 14-18 days following birth. Animals were deeply anaesthetized for surgery using an isoflurane-oxygen mixture (2% isoflurane) and placed in a custom stereotactic apparatus modified from the David Kopf system. The stereotactic apparatus was engineered to maintain the virus delivery pipette and the skull at a fixed orientation between injections. The skull was stabilized with the ear bars and the mouth piece and oriented flat in the anterior-posterior and left-right directions. Animal body temperature was maintained with a heating pad. A dental drill (Henry-Schein) was used to thin the skull above the injection coordinates, allowing the insertion of a pulled glass pipette (Drummond) (tip diameter:

10 μm -20 μm) beveled sharp at the tip. Injections were carried out by lowering the pipette to the desired coordinate of each injection at a speed of 0.5 mm/s. A 10 nl volume of 1:16 diluted virus solution was then dispensed at a speed of 5 nl/s, using a custom made positive displacement pump attached to a plunger in the pipette. A reliable quantification of viral particle count was not available. I monitored the meniscus between the virus and mineral oil in the pipette to ensure discharge of virus. Following a five minute waiting period, the pipette was first retracted 0.3 mm, paused for 3 minutes, and then retracted at a rate of 0.008 mm/second to the pia surface. These procedures were implemented in order to minimize unwanted cell labeling along the path of the retracting pipette.

Four injections targeting the thalamus were attempted in each animal - one each of the eGFP and tdTomato expressing virus in each hemisphere (**Fig. 1A**). Injections of eGFP and tdTomato were made side by side. The injections were spaced by 0.4 mm's whenever possible, thus generating dual-labeled cells that could be analyzed distinctly from the cells labeled green or red only. (**Fig. 1E**) Coordinates used for the injections were measured from the bregma, and ranged from x: 0.2 mm anterior to 1.65mm posterior to bregma, y: 0 mm to 2 mm lateral to bregma and z: 2.8 mm to 4.2 mm from pia surface, and spaced in order to achieve a systematic grid of injections covering the volume of the thalamus. Bilateral injections could be analyzed distinctly because thalamic projections do not cross the midline. (**Fig. 1F**) Some dual injections were oriented in the medial-lateral direction, some were oriented in the anterior-posterior direction, to reveal the topography of projections. (**Fig. 1E and 1G**)

Tissue Preparation and Sectioning

Animals were sacrificed 14 to 24 days after viral injection, and the brains were fixed by transcardial perfusion with 25 ml phosphate buffered saline (PBS) followed by 50 ml of 4% paraformaldehyde (PFA) in PBS. The brains were removed from the skull and incubated in 4% PFA at 4°C for 24 hours, and maintained in PBS at 4°C. 48 hours prior to sectioning, brains were placed in a 30% sucrose in PBS solution for cryoprotection. Brains were removed from solution and placed in a rectangular mold filled with optimal cutting temperature medium (OCT- Tissue Tech). The longitudinal fissure and the anterior edge of the cerebellum were aligned with the horizontal and vertical edges of the mold in order to minimize variation in cutting angles. A small incision was made along the ventral left side of the brain using a razor blade, in order to aid the discrimination of the orientation of sections on the slide. After 20 minutes in a -80°C freezer, or in a solution of dry ice and ethanol, the OCT embedded brain was separated from the rectangular mold and placed flush with the cutting chuck of the cryostat. The cryostat razor was also aligned to cut perpendicular to the surface of the cutting chuck. Brains were sectioned coronally at a thickness of 50 μm on a Leica 3050 S cryostat (**Fig. 1B**). Slices were collected beginning at the first section containing portions of the cortex (approx. 3.24mm anterior to bregma in the 2008 Allen Mouse Brain Reference Atlas), until sections no longer contained thalamus (approx. 3mm posterior to the bregma in the 2008 Allen Mouse Brain Reference Atlas) whenever possible. Sections outside this range were checked for fluorescence and, when none was found, they were excluded from the data set.

Each OCT embedded brain section was removed and floated in ample PBS to remove the surrounding OCT medium, before being manually placed on Superfrost-Plus microscope slides (FisherBrand). Slides were allowed to dry to facilitate adhesion, and then rehydrated in PBS before mounting with Fluoromount (Sigma) mounting medium, and gold seal cover glass (number 1.5). The edges of slides were sealed with clear nail polish for preservation.

Imaging

All sections were imaged at 20x magnification using the Nanozoomer slide scanner (Hamamatsu), at a uniform exposure level (**Fig. 1C**). The resulting images had a resolution of 0.5 μm/pixel, and a dynamic range of 12 bits. The proprietary data format generated by the nanozoomer was converted to the TIFF format. The exposure level was chosen to keep fluorescence levels in the labeled projection sites within the dynamic range of the camera. This exposure level often caused images of the injection sites to become over-exposed due to an abundance of fluorophores in cell bodies. The overexposure created a halo around injection sites, such that boundaries were not well defined. In order to determine the location and borders of the injection region, additional images were taken of sections which contained labeled cell bodies at a lower illumination using the Axioscope (Zeiss) upright microscope which also had automated image stitching capabilities (**Fig. 1D**). The Axioscope images were taken under 5x magnification with a dynamic range of 8 bits. All images were quantified using Matlab scripts (Mathworks).

Chapter 3. Results

The data set describing thalamic projections

Data collected from 54 mice were used for analysis, and the data set contained 208 separate injections. Sectioning of the brain resulted in 95 to 152 sections, with an average of 120 sections per brain. (**Fig. 2A**) With approximately 0.5 terabytes of data per brain, the complete data set was on the order of 20 terabytes. 8x downsampled versions of images (also output from the Nanozoomer) were used in the analysis of the injection site- with pixel coordinates indicating regions of interest (ROI) translatable from the downsampled images to the original high resolution images.

The data set has an inherently higher resolution in x (body medial-lateral) and y (body dorsal-ventral) than in z (body anterior-posterior), as a result of sectioning the brain at 50μm thickness. Accordingly, each pixel in a downsampled image represents a 4μm x 4μm x 50μm volume of the brain, and each pixel in a high resolution image represents a 0.5μm x 0.5μm x 50μm volume of the brain. High resolution images were able to resolve individual labeled axons. (**Fig. 2B**)

3D alignment and whole-brain normalization of section images

The Nanozoomer images whole slides mounted with 6-10 sections each, so individual section images were first extracted from each slide image through the manual selection of boundary points above and below each section. (**Fig. 3A**) The auto-fluorescence of brain tissue in the green channel was sufficient to discriminate edges of the sections from background. Extracted section images were placed on a 2500x3500

pixel black background in order to standardize the 3D coordinate space in which the brains were placed.

Sectioning brains for the purposes of imaging produces a series of two dimensional images. These images must be assembled into an accurate three dimensional alignment in order to reconstruct their relationship in the intact brain. In order to merge the two dimensional section images into three dimensional alignment, I first performed a rotational transformation followed by two rigid translations in x (body medial-lateral) and y (body dorsal-ventral) on section images. As the brain midline was the most reliably visible and constant feature within section images, each section was first corrected for orientation and center alignment through the manual selection of two points, one point at the top and one at the bottom of the midline (**Fig. 3B**). The top and bottom points were used to rotate images such that the line connecting the points were oriented vertically. Each image was translated in x (body medial-lateral) such that the vertical line was at the center of the background, thereby aligning the midline in three dimensions.

Alignment of images in the y (body dorsal-ventral) direction proved more challenging because there is not a consistent, straight anatomical landmark spanning the whole brain. First, I chose a threshold intensity value to discriminate between pixels falling ‘inside’ and ‘outside’ the section. The auto-fluorescence of unlabeled brain tissue in the green channel was sufficient to allow a wide range of options in the selection of the threshold. The mean y (body dorsal-ventral) coordinate of ‘inside’ pixels was defined as the center of mass (COM) for each section. (**Fig. 3C**) A COM distribution curve used to line up images in the y direction was based on the Paxinos Mouse Brain Atlas. First I described the COM curvature of a sagittal slice lateral 0.36 mm from bregma (figure 104

of the Paxinos Atlas). I calculated the COM for pixels in each 3 pixel (corresponding to section thickness of 50μms) vertical column. I calculated the transformation that would need to be applied to each coronal Paxinos Atlas section COM to reconstruct the sagittal section, and used this transformation to generate an idealized COM distribution curve. (**Fig. 3D**) The first section in each brain data set was placed in a consistent y position, and the COM distribution curve was resampled to generate a translation value for each subsequent section. In this manner, I was able to generate a reasonable alignment of images in the y (body dorsal-ventral) direction. (**Fig. 4**)

With the completion of 3D alignment, the next step towards registration of anatomical data was to correct for variability in brain size. This normalization step involved brain wide isotropic scaling in x and y. I did not normalize images in the z direction, because there is a low sampling rate in the anterior-posterior direction, and the normalization involves resampling and averaging of image data. I calculated the maximum x (width) and y (height) dimensions of each section. (**Fig. 3E**) Each brain wide mean of section maximums was compared to the population-wide mean of the maximums, producing an x and y scale factor for each brain. The effectiveness of this normalization step was confirmed by graphing the number of 'inside' pixels in each brain against the number of image sections before and after normalization. (**Fig. 3F-G**) The linearity of the normalized graph (**Fig. 3G**) indicates that following the normalization step, most of the variability in brain size was restricted to the unscaled, z (body anterior-posterior) direction.

Images of injection sites generated on the Axioscope revealed the boundaries of the injection more accurately, and were aligned to corresponding Nanozoomer section

images through rigid translation via the manual selection of corresponding anatomical landmarks visible in both images. Axioscope images were then aligned and normalized using the same parameters as corresponding Nanozoomer images.

Wherever there was tissue damage that interfered with the alignment steps, a manual best-fit decision was made for placement. Damaged sections were excluded from the generation of normalization parameters.

Thalamus segmentation, registration and description of the injection site

The next step in the analysis was the segmentation of the boundaries of the thalamus, and the determination of the boundaries of injection regions within the thalamus for each brain. The thalamus segmentation was performed manually according to anatomical landmarks labeled in the Paxinos and 2008 Allen Mouse Brain Atlases (**Fig 5A**). In each brain, the most anterior section containing the thalamus was defined to be the first one in which the third ventricle connected with both lateral ventricles. The last section in the anterior-posterior axis containing any portion of the thalamus was defined to be 10 sections posterior to the section where the last remaining connection of the corpus collosum was visible. Variations in the auto-fluorescence of brain tissue, suggesting cell density or myelination patterns, contributed to the segmentation process. The segmented thalamus was expressed as a 3D matrix with dimensions of 2500x3500x (# of sections) filled with 1's in pixels corresponding to the thalamus, and 0s otherwise.

In order to define the boundaries of each injection site, I applied a threshold value of 26, which corresponded to 10% of the maximum intensity value in the 8 bit Axioscope images in the green and red channels. (**Fig. 5B**) This step required manual refinement

because of the disparity in exposure levels between Axioscope images. Injections of 10nl of virus resulted in varying volumes and shapes of labeled regions. Injections labeled a drop shaped region with a diameter sized 579 μm ± 102 μm (calculated for 60 injections). This size corresponds to approximately 1 % of thalamic volume. (1.1 ± 0.1 %). It remains to be described how labeled pixel volumes correspond to the number of labeled cells in the thalamus. The thresholded injection locations were expressed as a 3D matrix with dimensions of 2500x3500x (# of sections) filled with 1's in pixels with labeled cells, and 0s otherwise. (**Fig. 5C**)

Creation of a model thalamus, coverage and redundancy

The data generated here, when viewed in three dimensions and as an ensemble, describes the destination of projections from any specified volume of the thalamus, without the need for further analysis. The power of this data set is dependent on our ability to create a connectivity map, such that the 54 brain volumes and 208 labeled regions may be expressed within a common coordinate space. The common coordinate space for the thalamus, the 'model' thalamus, was created by another member of the lab, Brian Long, and is not included in this thesis. Later work done by my colleagues Brian Long and Jeannie Hunnicutt fit each individual thalamus volume to the model thalamus. Plotting every injection in the model thalamus revealed that 90% of the thalamus is labeled by at least one injection, and 50% of the thalamus is labeled by five or more injections, indicating a high level of completeness and redundancy in the data set.

In further work done on this model by Jeannie Hunnicutt, the Paxinos and 2008 Allen Mouse Brain Atlases were segmented according to the boundaries of thalamic

nuclei indicated in the atlases (Jones et al., 1985). The borders of these nuclei were registered to the coordinates of the model thalamus, effectively parcellating the model thalamus.

Projection site segmentation and quantification

The thalamus projects primarily to the cortex and the striatum. Projections to the cortex travel through the striatum in the myelinated bundles which give the striatum its name. Some projections simply pass through in *bundles* without interacting with the striatal microcircuit, some collateralize en route in a *diffuse* style and make functional connections within the striatum. (**Fig. 6A**) The interdigitation of these projections makes it difficult to reliably distinguish them. Although it is possible to distinguish bundles and diffuse projections by eye, the task of masking each bundle manually is intractable because of the volume of data. The bundles do not run perpendicular to the coronal axis, following a radial route instead, exhibiting different profiles in the anterior-posterior axis. A manually trained ImageJ image analysis module (Advanced Weka Segmentation, Hall et al., 2009) had success distinguishing between bundles and diffuse projections on an image by image basis, but was not reproducible across the entirety of the striatum because of the diversity in sectioning profiles.

Faced with these limitations, I devised an alternate, manual scoring method to describe the projection pattern associated with each injection. Three independent scorers made qualitative assessments on whether there were or were not labeled diffuse projections in the striatum emanating from each injection. Scorers reported a - if there were absolutely no projections exiting the bundles (or if there were no labeled bundles),

and a + if there were any diffuse projections at all in the defined regions. These scores were interpreted as described below in the 'representing connectivity' section.

In the cortical regions we investigated, there were again no anatomical boundaries between the functional sub-regions which could be detected through automated image segmentation. In the absence of an automated image segmentation method, cortical sub-regions were scored manually according to the boundaries defined by the Paxinos Atlas. We chose to analyze projections to the orbital cortex (OC), frontal cortex (FC), anterior cingulate cortex (ACC), and the vibrissal primary motor cortex (vM1). The manual scoring method described for the striatum (scorers reporting -'s and +'s) was applied to each cortical region. (**Fig. 6B-D**)

I went further with the analysis of vM1, in order to describe the layer specificity of projections from the thalamus. Previous work suggested that the two populations of thalamic cells that projected to vM1 could be distinguished by whether or not they exhibited a prominent projection to layer 5b (Hooks et al., 2013). However, the method did not have sufficient spatial resolution to describe the thalamic origin of these two projections. The sections containing vM1 were defined to be three sections: the section where the corpus collosum first connected across the midline and the ones immediately before and after. On each section, vM1 was segmented manually, with some automated assistance. I selected two points on the pia surface: the point at which the cortex folds towards the midline and the point at which the pia surface begins to be parallel to the external capsule. I selected two more points on the external capsule and corpus collosum, The points were joined by straight lines to define the medial and lateral boundaries of vM1. (**Fig. 7**) The pia edge defined the boundary on the dorsal side, and the ROI

boundary was sealed with a small arc around the cingulum. The ROI defined in low resolution images were translated to the high resolution images, and the rest of the analysis was performed in high resolution images. Using the intersection of the lateral and medial edges of the vM1 as the reference point, each ROI pixel was assigned a normalized radial distance (*r*) value between 1 and 0, describing its location between the pia surface and the claustrum. *R* values were used to assign pixels to different layers of the cortex, such that pixels with $r<0.11$ were assigned to layer 1, $0.11<r<0.26$ layer 2/3, $0.26<r>0.4$ layer 5a, $0.4<r>0.78$ layer 5b, $0.78<r>1$ layer6, according to the definition of vM1 layers generated in previous work done in the lab. (Mao et al., 2011) Unlike in the striatum, all the labeled projections in the cortex could be interpreted to be functionally relevant. I summed the green and red fluorescence intensity values of pixels with *r* values in different layers to determine the layer specificity pattern of each projection. The variability in auto-fluorescence across cortical layers was not taken into account in this analysis. vM1 projections were scored to describe the two separate parcellations of the thalamus. The injections which produced no labeling in vM1 received a score of -.One population was defined as those that exhibited a significant amount of fluorescence in layer 5b, and scored with a +. The other population was defined as those that did not project to layer 5b in a significant way. The results of the scoring will be represented as described below.

Representing connectivity: parcellating thalamus from perspective of projections

The data generated by this method is novel in its completeness, resolution- and multidimensionality. The connectivity pattern between the two three dimensional

structures -that of the thalamus and a projection region- is difficult to capture and portray in the two dimensions available for its representation. In addition to basic connectivity, there are additional dimensions of the data that may be represented, such as the strength (density) of projections to a location, or the laminar profile of projections in cortical regions, among others. In order to simplify the representation and render the data more useful for functional interconnectivity interpretations, I formulated a representation scheme which parcellates the thalamus from the perspective of individual projection site queries.

As an example, from the perspective of the ACC, each injection had received a qualitative score of '-' or '+', with '-' representing no projections to ACC whatsoever and '+' representing any amount of labeled axons. The final sub-regions of the thalamus that projects to the ACC was defined by the pixels representing injections in the thalamus that were scored '+', minus the pixels that were scored '-'. The logic of this representation is that, injection regions which receive a score of + for the ACC are necessarily overestimations of the real volume of the thalamus which projects to the ACC. Due to the overlap of injection volumes in the thalamus, subtracting the injections that have been scored as -'s constitutes a refinement, and a closer approximation of the real volume of the thalamus which projects to the ACC. This representation scheme reveals the parcellation of the thalamus from the perspective of each projection region.

Chapter 4. Discussion

It is the nature of this project that results, in the form of a parcellation of the thalamus from the perspective of projection sites, depend on a completed pipeline for analysis, and for the complete data set to be analyzed. In previous chapters, I described the work I did towards a complete analysis. I have registered brains in three dimensions, and normalized them. I segmented the thalamus and determined the location of injection sites. I did not constructed a model thalamus to place each separate brain into a common coordinate space. I scored each injection for its projections to multiple regions, and I described my ideas on how to portray this connectivity data most effectively. In this discussion, I will highlight the innovations presented in this method, discuss limitations inherent to some choices I made, and future extensions of the method that are possible, along with approaches that I tried without success.

Dual color bilateral labeling

The data generated by this method is unique in its completeness and resolution; the principal innovation that made this possible was an efficient labeling and imaging pipeline. Dual color bilateral injections of the thalamus allowed me to keep the animal count low and the data volume within reason. Placing two adjacent injections in the same hemisphere with some overlap revealed the internal topography -the direction of injection to projection site correspondence- of projections between the thalamus and target regions. (**FIG. 1E**) Dual color labeling revealed that some thalamic projections were topographically direct, and others crossed over. (**FIG. 1H-I**) How the topographic

representation of sensory information in the thalamus is transformed as it is conveyed to higher order regions of the brain has important implications for information processing. The redundancy and overlap in labeling, combined with the +/– scoring scheme, allowed us to further refine the analysis and more closely represent the 'real' volume of thalamus that projects to a particular region. At the time we conducted experiments, we were not able to procure viral vectors that expressed fluorophores that could be separated from eGFP and tdTomato under our imaging set up. With further iterations of this method, it may be possible and desirable to use three or even more injections per hemisphere.

High-throughput imaging

The technological development that contributed most to the feasibility of this pipeline was the automated image stitching and high-throughput slide imaging abilities of the Nanozoomer. Without this automation, the labor involved in imaging this data set may have been unapproachable. The uniformity of illumination across the field of view - which is frequently a weak point of automated imaging setups- greatly aided the analysis. It would have been more convenient if I had been able to image the injection sites with a lower exposure under the Nanozoomer, and been able to avoid overexposure. I frequently needed to re-mount slides in order to image under the Axioscope, and the exposure of the Axioscope was not uniform across imaging sessions. The lack of uniformity in Axioscope images complicated image analysis, adding manual steps and lowering confidence in the precision of our detected labeling boundaries.

AAV mediated cell labeling

Stereotactic injections of AAV under the CAG promoter reliably produces labeled volumes small enough to serve the purposes of a mesoscopic connectome. It is possible to reliably target small subcortical volumes using stereotactic coordinates, despite the variability in the location of bregma from which the coordinates are calculated. The positive displacement pump and glass pipette set up I used for injections resulted in a low failure rate. In previous work we demonstrated that the same viral preparation is able to label cortical neurons similarly (Mao et al., 2011). Thus, AAV mediated cell labeling may be used to label spatially targeted cells as part of a whole-brain connectome approach. It will be interesting to extend this method to further brain regions to confirm or negate the existence of reciprocal or convergent projections. Bilateral injections may not be appropriate for regions that project bilaterally, but dual color injections will remain useful.

An advantage of using viral vectors to label cells is that different promoters may be used to label genetically specified subpopulations of cells in future iterations of this method. As molecular markers for functionally distinct regions become known, this method could be used to identify their specific projection patterns as well. In the initial stages of this experiment, I tried a viral vector that expressed GFP or tdTomato under the synapsin promoter in order to selectively label neurons. In my hands, and under the two week expression protocol, I did not see sufficient expression of the fluorophore to image the injection sites or the projection sites under our imaging set up. It may be beneficial to improve the synapsin pipeline in further iterations of this method.

Another method I tried was to mix fluorescent LumaFluor microbeads (LumaFluor Inc) with virus during injections. We had utilized this method previously in the lab. (Mao et al., 2011) LumaFluor microbeads are taken up by axons, and are retrogradely transported to their cell bodies of origin. This method labels the destination of projections from a region, as well as simultaneously indicating the origin of innervations to the labeled region. Unfortunately, in my hands, the beads were not distinguishable from the anterogradely labeled axons. This was mostly because we had only two colors available for the beads, and they were the same as eGFP and tdTomato. Additional colors, and improved bead detection image analysis algorithms might render combined anterograde-retrograde labeling feasible.

One disadvantage of the viral gene delivery method is that there is always some amount of 'tail labeling' along the path of the pipette withdrawal. I tried to eliminate this tail by loading only the minimum amount of virus in the pipette. Unfortunately the remaining mineral oil in the pipette leaked as it was being withdrawn, and led to cell death at the injection site. I was able to minimize tail labeling by withdrawing the pipette very slowly and with intermittent stops.

Animals: age, sex, handedness

One of the early decisions we made when planning experiments was the age of animals at viral injection. We anticipated an extension of this project would involve injections with an alternate viral vector: that expressed the Channelrhodopsin-2 (ChR2) gene tagged with a fluorescent protein under the same promoter. Under this protocol, the labeled projections could be visualized similarly, but would have the added benefit of

demonstrating or negating functional connectivity, as we had done in previously published work (Mao et al., 2011). Indeed, further work done in the lab has verified that diffuse projections in the striatum are functionally integrated while activating striatal bundles causes no activation in the striatum. We decided that the optimal age for slice physiology was P28, and with the two weeks required for viral expression, we made the decision to perform injections at around P14. Unfortunately, at P28, the mouse is considered an adolescent, and it is unknown how the connections between the thalamus and the cortex and the striatum varies with age. Both the Paxinos Mouse Brain Atlas and the Allen Mouse Brain Atlas characterize adult mice. Using adult atlas coordinates to define thalamic nuclei boundaries may not be entirely appropriate. Our data set may have been more immediately comparable to other whole brain data sets had we chosen to image adult mouse brains. On the other hand, the use of adolescent animals is pervasive in the literature, particularly in electrophysiological experiments where slice health is related to age. Frequently, the age of the animal is considered an obstacle that cannot be overcome, and is barely noted. As knowledge of brain connectivity continues to improve, it is certain that a map of the mouse brain in adolescence will remain valuable.

I did not choose to differentiate between the left and right side of the brain, and I did not consider differences in sex. Further iterations of this method with larger data sets, may explore distinguishing between the two sides and the two sexes to investigate possible differences.

Cryostat sectioning and manual mounting of free floating slices

One of the challenges I faced during the analysis stage was the process of detecting and correcting for tissue deformations caused by the mounting process. Cryostat sectioning itself produced sections that were undamaged and consistent in thickness. The mounting step involved floating each section in PBS and placing the floating section on a semi-submerged glass slide manually using a brush. During placement, tissue is regularly sheared and deformed. It then proved very difficult to fix tissue deformations in the analysis stages. Block face imaging involves coupling the sectioning device to a multi-photon microscope, allowing sections to be imaged in intact tissue before they are sectioned. This method greatly reduces deformations and places the generated image in a pre-aligned 3D space, removing the need for the alignment steps described in the results section. Future iterations of this method may benefit from deploying block face imaging in order to reduce the uncertainty and time expenditure of the analysis step.

I did not make any efforts to correct for possible variations of cutting angle, even though it was clear that some sections were visibly cut at an angle. The detection and correction of the cutting angle may vastly improve the accuracy of mapping each thalamus on to the model thalamus and subsequent injection site localization.

Automated image analysis

The lack of readily available automated image analysis methods complicated the extraction of information from this dataset. At every step, though we attempted to automate the analysis steps, we had to settle for manual segmentation. Historically, the

vast majority of anatomical studies have been annotated and interpreted manually. Manual selection is a solution to segmentation, but it is non-reproducible, and biased in ways unknown; it is not a sustainable solution to anatomical studies at this scale. A single comprehensive algorithmic solution to reconstruct and segment serially sectioned mouse brains is needed. Until an algorithm for reconstruction and segmentation is available, there will be considerable barriers to scaling up the method described here.

Chapter 5. Conclusions

In this thesis, I outlined a pipeline for data collection, and approaches towards the analysis of a mesoscopic connectome prototype project. I have discussed that further analysis, and the model thalamus will be necessary to describe the underlying architecture of the thalamocortical and thalamostriatal system. In the introduction, I discussed the ideal characteristics of a mesoscopic connectome method. I argue that bilateral dual color AAV cell labeling of (1) spatially grouped neurons, imaged at (2) whole brain completeness, is a novel and appropriate mesoscopic connectome method, able to achieve (7) multiplicity of comparable labeling methods, (5) completeness and (6) redundancy of sampling, it is possible to place the data in a common (3) anatomical coordinate system, image at (4) individual axon resolution, while generating a (8) reasonable amount of data. The method is (9) high-throughput and amenable to be (10) scaled up to investigate connectivity in the whole brain.

It is an exciting time for brain wide anatomical connectivity projects. Large Institutions such as the Allen Brain Institute are standardizing the collection and assembly of brain wide connectivity data, and there are calls for increased funding towards these efforts. Indeed the promise of brain wide connectivity projects is far-reaching and substantial. Here I will describe the various questions which stand to benefit from the connectome project.

The study of disease and plasticity at the mesoscale

The establishment of the baseline for brain wide connectivity in healthy brains will reveal how structural connectivity may be altered in the disease state. Circuit diagrams of cell subtypes, will reveal how the degeneration of certain subtypes may be responsible for brain wide and specific disease deficits. It will also be possible to describe how brain wide connectivity varies across genetically different populations, with age, or among the sexes.

That there is a high degree of plasticity at the microscopic level of individual dendritic spines has been established. It is unknown whether there is also a significant level of plasticity at the mesoscopic scale as a response to experience (Oberlaender et al., 2012) or injury. The brain wide connectome will set a baseline for the description of plasticity at the mesoscopic scale.

Neuronal modeling, and complexity

The empirical data generated by the connectome will serve to constrain neuronal network models. Already, studies of cortical circuit organization have inspired biologically accurate modeling of inter-cortical computation on supercomputers. (Markram, 2006) The abstract network properties extracted from biological structure will reveal whether brain networks represent a singularity of organizational complexity or are reducible to known complexity categories.

Functional connectivity and anatomical connectivity

An important extension of the connectome will be to describe how anatomical connectivity in the form of axo-dendritic overlap corresponds to and serves to indicate

functional coupling between cells and between brain regions. In previous work, we have used ChR-2 expressing viral vectors to investigate this relationship in isolated regions (Mao et al., 2011). Functional interpretation of anatomical connectivity will benefit from the investigation of this relationship.

Construction of the connectome

The connectome must be a compression of the raw data, which extracts the most relevant information regarding the principles of organization. It need not be exact replica of the connectional anatomy down to the finest ramifications of neurites or individual synaptic boutons. How the connectome is compiled, how it is reduced, and the manner in which it is made available to the scientific community will determine how useful and informative it will be.

In order to create the most useful and integrative connectome database, the imaging data and the reduced form of the data must be freely available as digital brain atlases. It will be useful if a pipeline for the collection of data and its segmentation and quantification become standardized and open source. This standardization will ensure that whole brain data produced to reveal different aspects of neuroanatomy, or constructed from multiple perspectives may be combined in a common coordinate framework. With the creation of large data repositories, I foresee the separation of the data generation effort from the data mining, and knowledge discovery processes. The ability for separate groups to specialize in one or the other of the aspects of this data will increase efficiency by reducing redundancy in the field.

Standardized high-throughput labeling and imaging pipelines are becoming easy to achieve. But there is a need for advances in large data and workflow management, image recognition, automated segmentation, in order to place these data into mineable data repositories of greatest benefit to the neuroscience community.

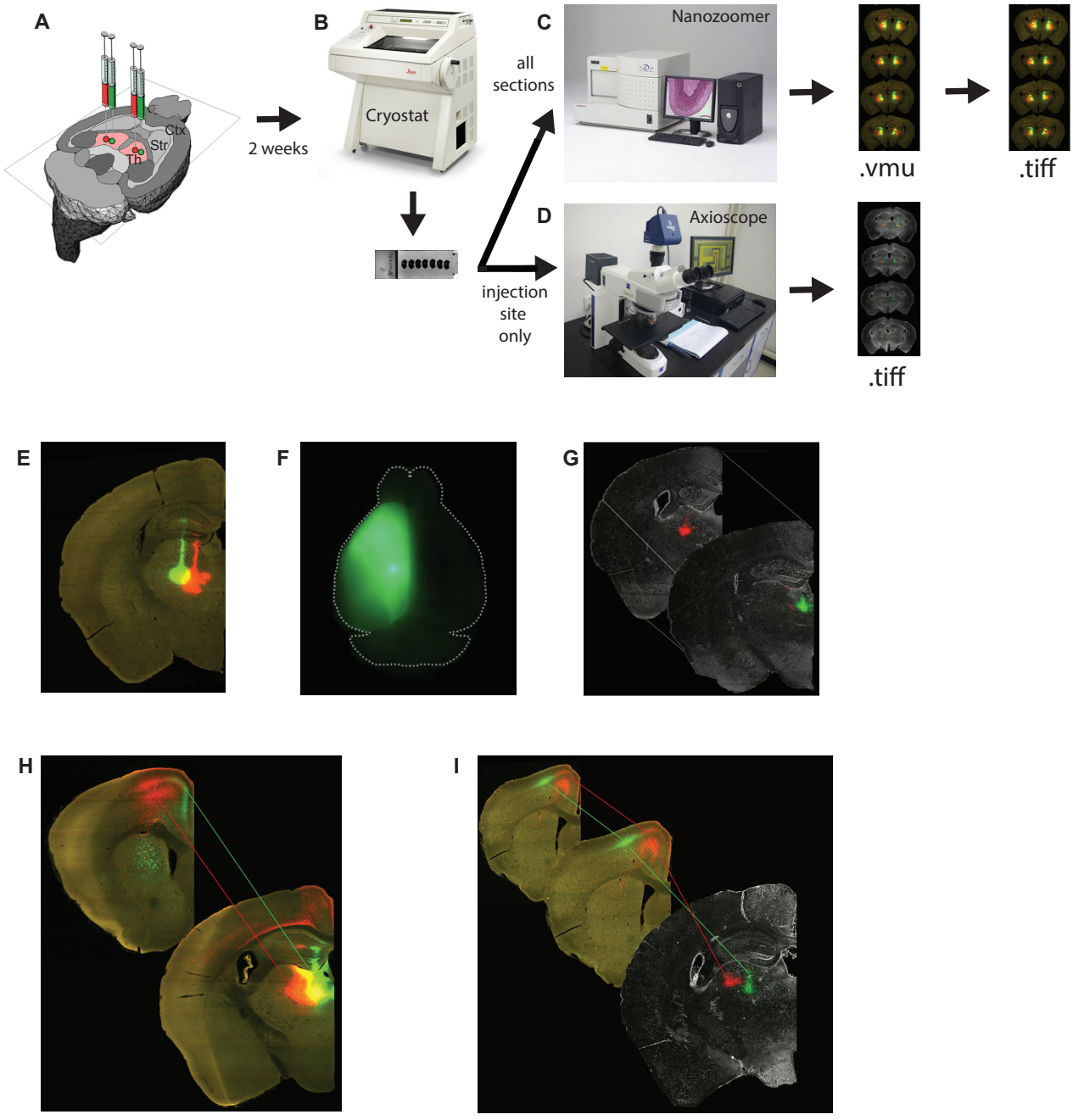

Figure 1. Whole brain connectivity data acquisition pipeline. **A**. Each animal received bilateral, dual-color stereotaxic injections of AAV expressing GFP and tDT in the thalamus. **B**. Brains were sectioned on the cryostat at 50 μm thickness, and placed on glass slides. **C**. All sections were first imaged on the Nanozoomer. Notice the over exposure of injection sites in nanozoomer images. Images were converted to the .tiff format for analysis **D**. Sections containing labeled cell bodies were re-imaged on the Axioscope. **E**. Some overlap in injection sites resulted in dually labeled cells. **F**. Thalamic projections do not cross the midline, so injections can be analyzed separately. **G**. Some dual color injections were oriented in the anterior-posterior direction relative to each other. **H**. Some dual color injections oriented in the medial-lateral (ML) direction revealed topographically *direct* thalamic projections **I**. Some ML injections revealed *crossed* projections.

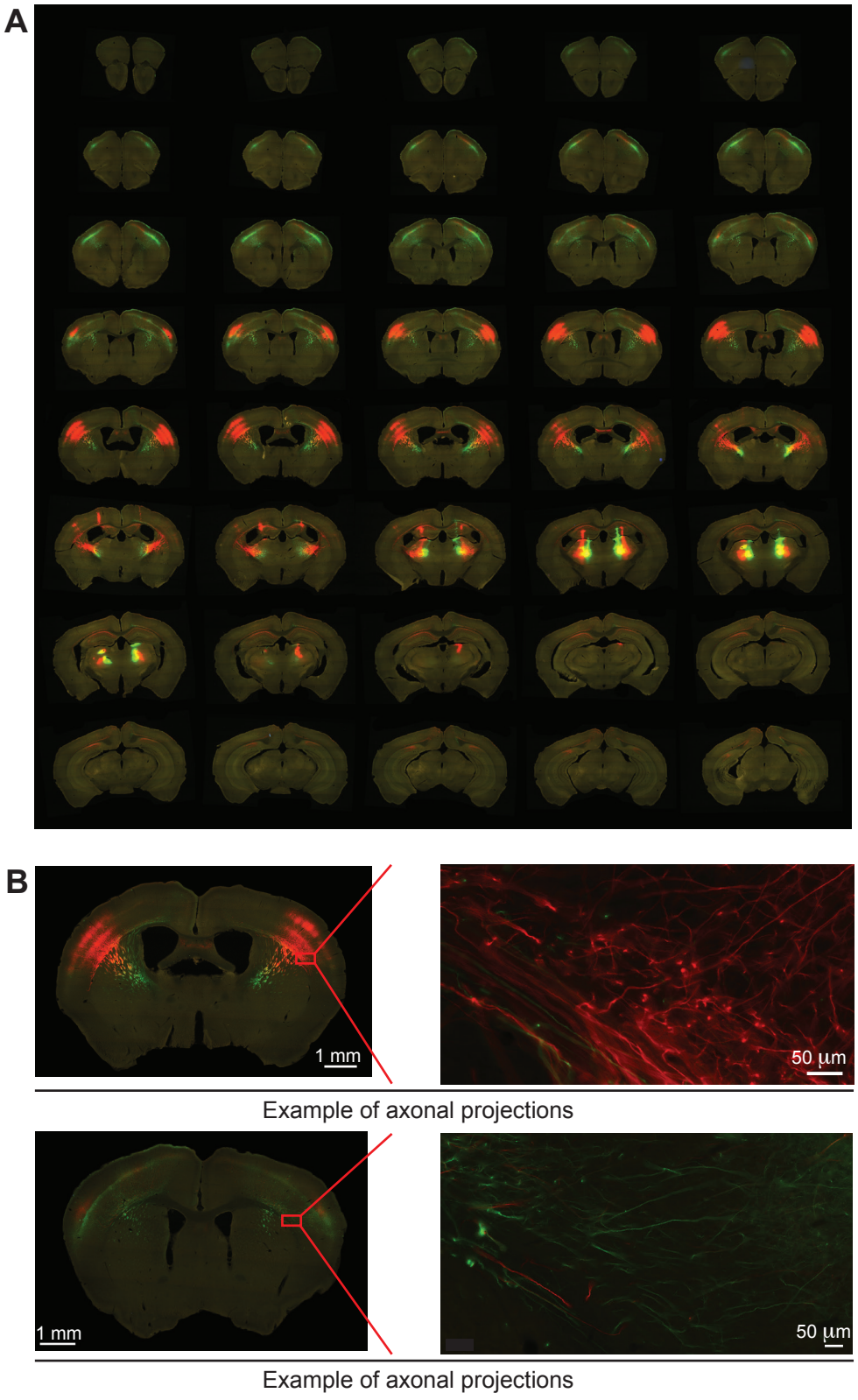

Figure 2. The data set describing thalamic projections. A. Sections were collected from the most anterior section which contained any portion of the cortex, to the most posterior section containing any part of the thalamus. Every third section from one brain is shown here. **B**. Image resolution was sufficient to image individual axons in both the green and red channels.

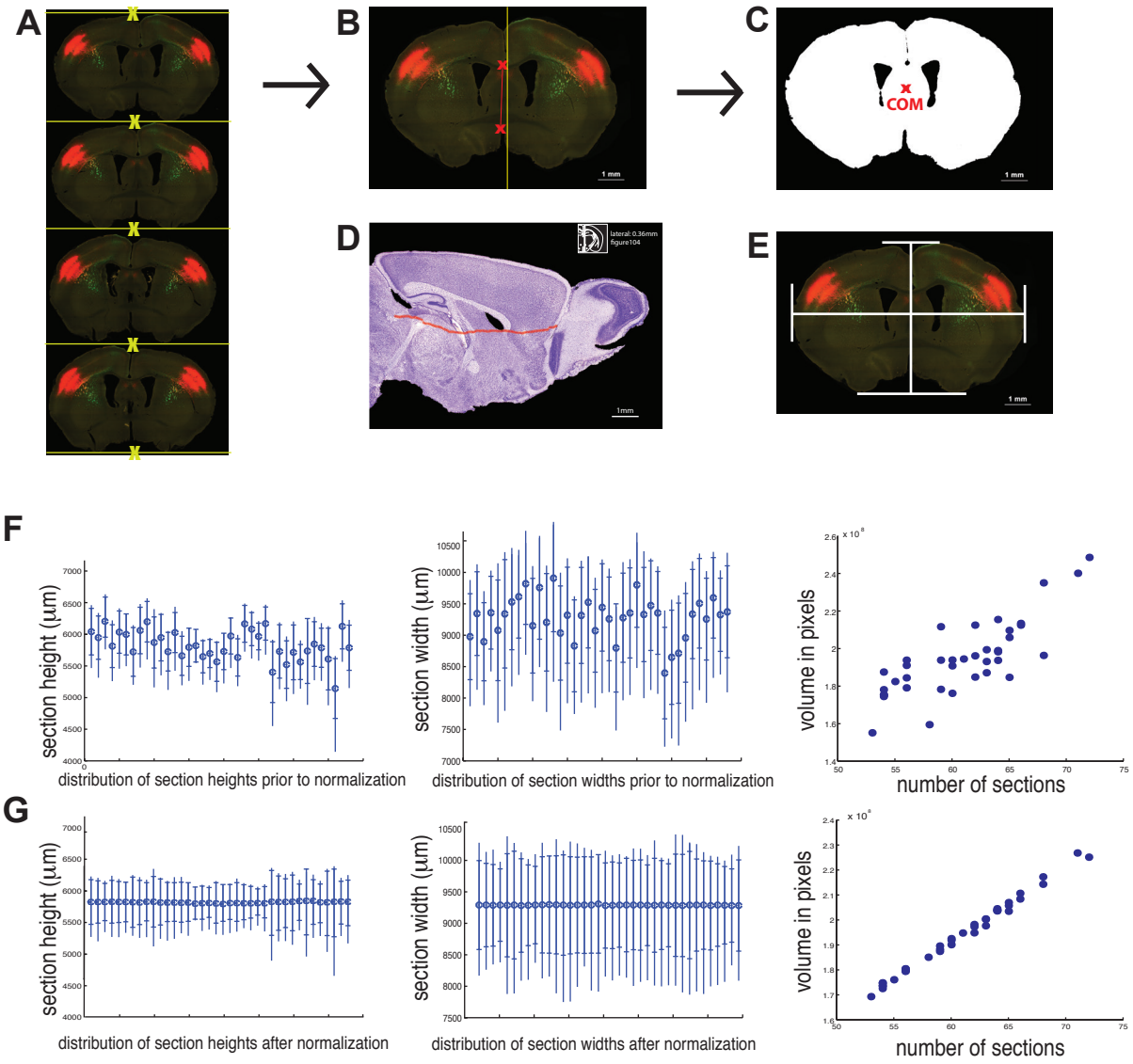

Figure 3. Alignment and normalization pipeline. A. I manually selected boundary points above and below each section to extract them from whole slide images. **B**. I selected points at the top and bottom of the midline and used the points to rotate and align section images in the x (medial-lateral) direction. **C**. I thresholded images to determine 'inside' and 'outside' pixels, and determined the mean x and y coordinate, the center of mass (COM), of each section. **D**. I calculated the COM distribution curve in a Paxinos Mouse Brain Atlas sagittal section, and used this curve to align section COMs in the y (dorsal-ventral) direction. **E**. I calculated the maximum x (height) and y (width) dimensions of each section. **F**. Section heigths and widths prior to normalization. **G**. Section heights and widths after normalization. The volume in pixels vs. number of sections after normalization is linear, indicating that variability in size has been restricted to the non-normalized z (anterior-posterior) direction.

s !

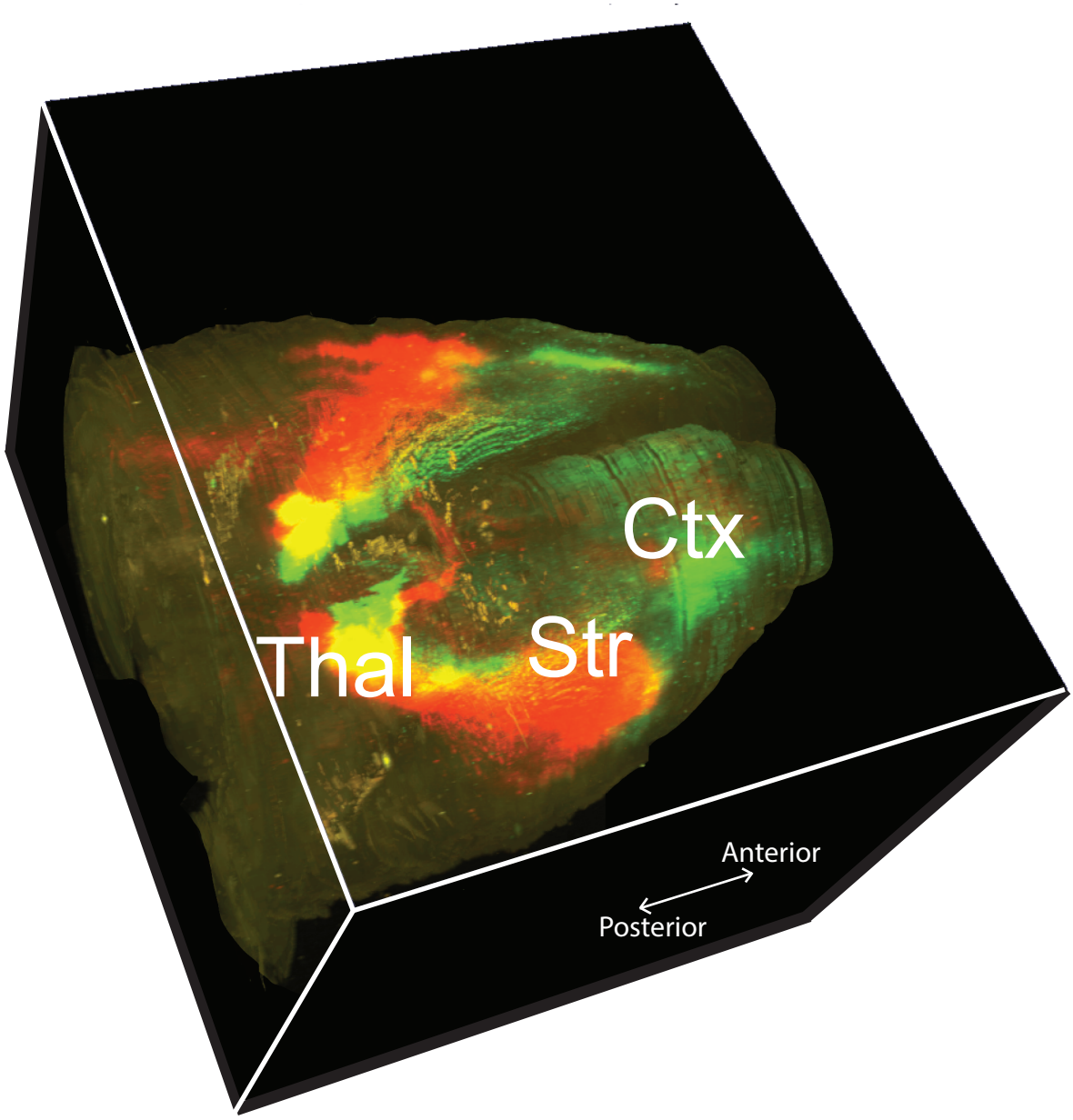

Figure 4. The 3-dimensionally aligned and normalized brain. I was able to achieve a 3-dimensional reconstruction of brain sections.

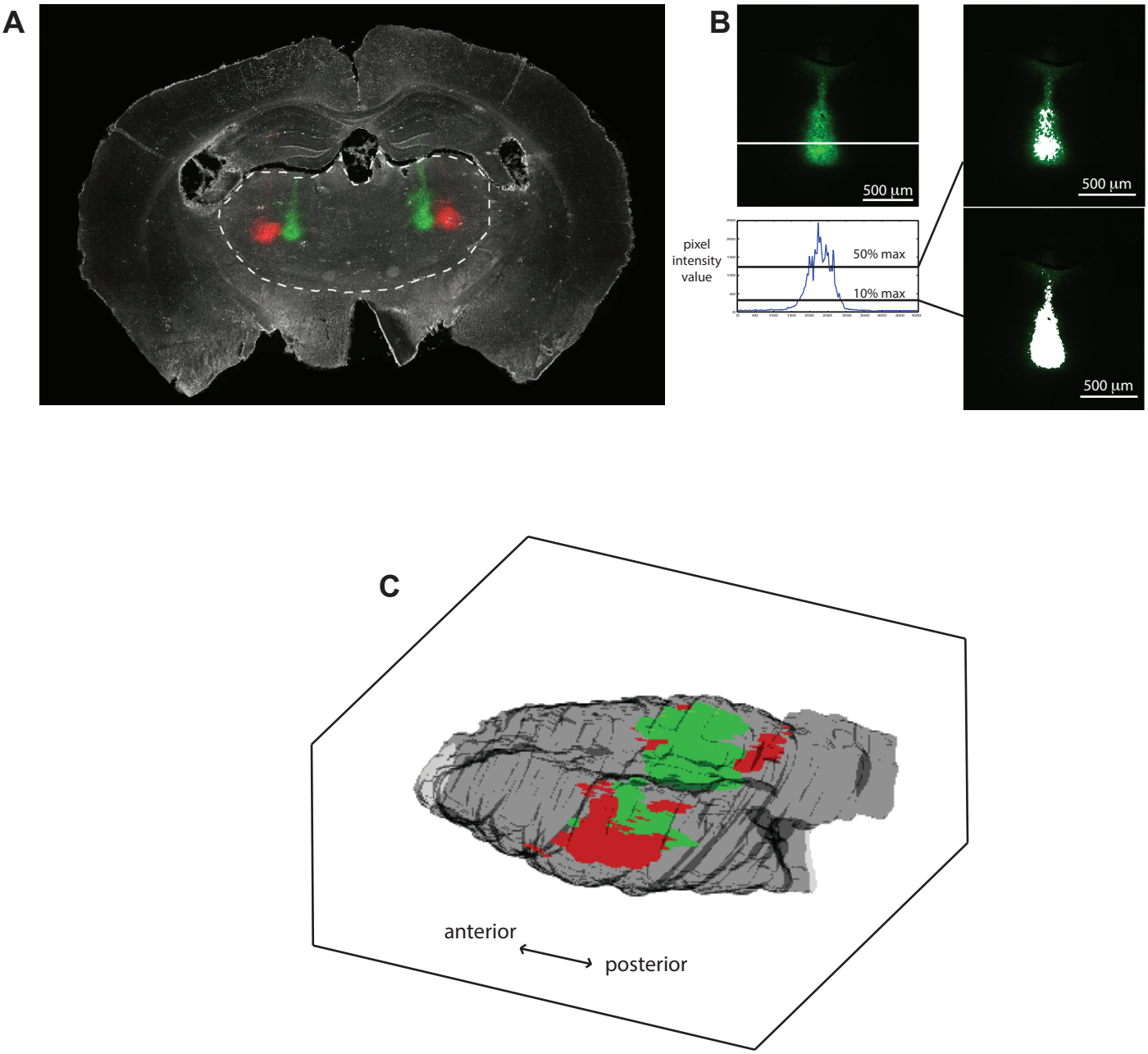

Figure 5. Thalamus segmentation and determination of injection site boundary. **A**. I manually segmented the borders of the thalamus according to anatomical landmarks labeled in the Paxinos and 2008 Allen Mouse Brain Atlases. **B**. I thresholded Axioscope images at 10 % of the maximum intensity to determine the boundaries of labeled cells. **C**. The thalamus segmentations and thresholded injection sites are expressed in the same 3D coordinate space.

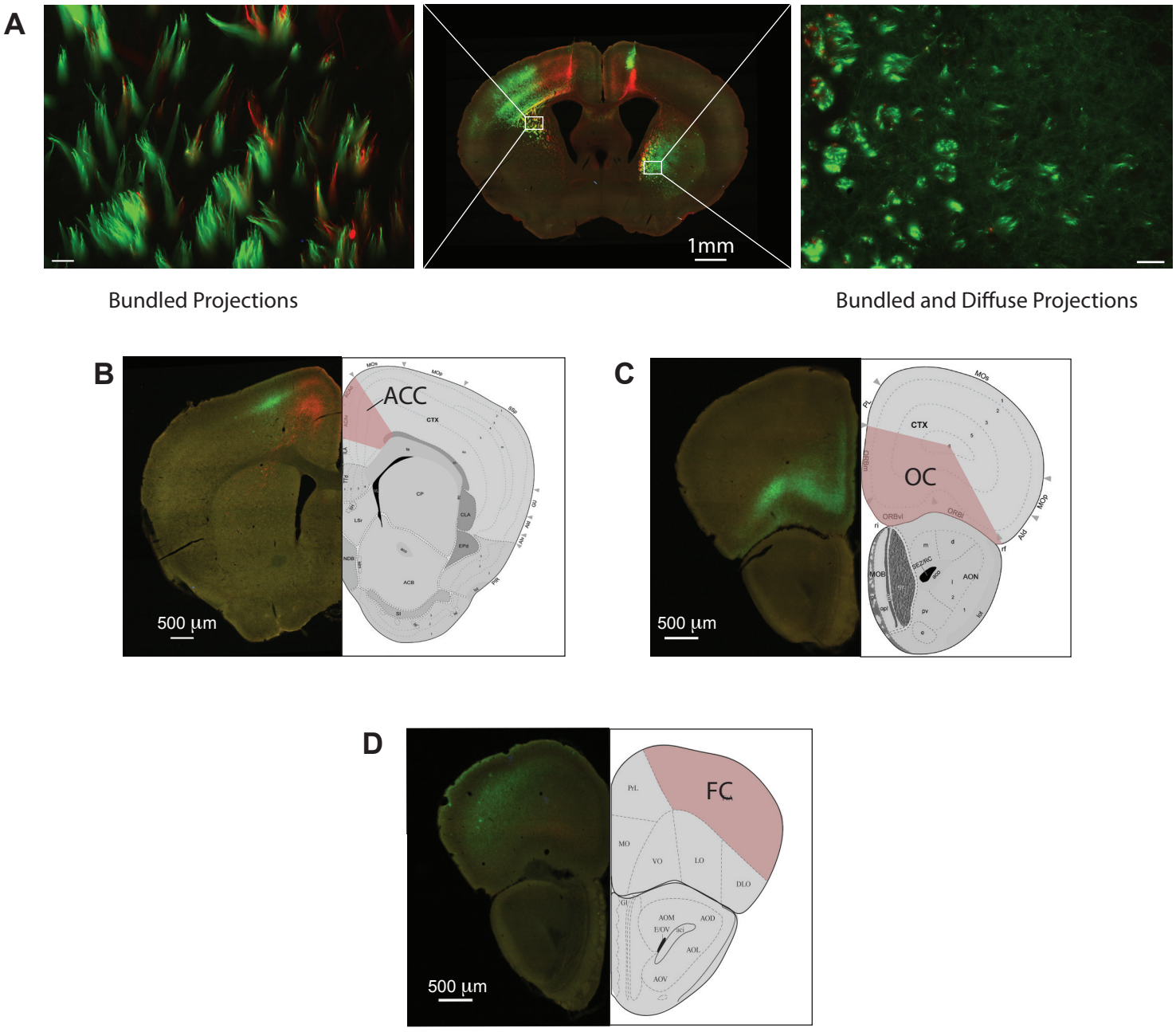

Figure 6. Labeling in the striatum and cortex. A. Thalamic projections appear in the striatum in bundles or as diffuse projections. The two groups are easily discernible by eye, but their interdigitation makes it difficult to discern them using automated image detection tools. On the left side image, both the red and green injections receive a negative score in the striatum. On the right side, the green injection receives a positive score due to the presence of diffuse projections. **B**. Cortical regions were identified according to boundaries in the Allen Brain or Paxinos Atlases, and scored based on the existence of labeled axons. The red injection receives a positive score in the anterior cingulate cortex. **C**. The green injection receives a positive score in the orbital cortex (OC) . **D**. The green injection receives a positive score in the frontal cortex (FC).

(Fig. 6A) !

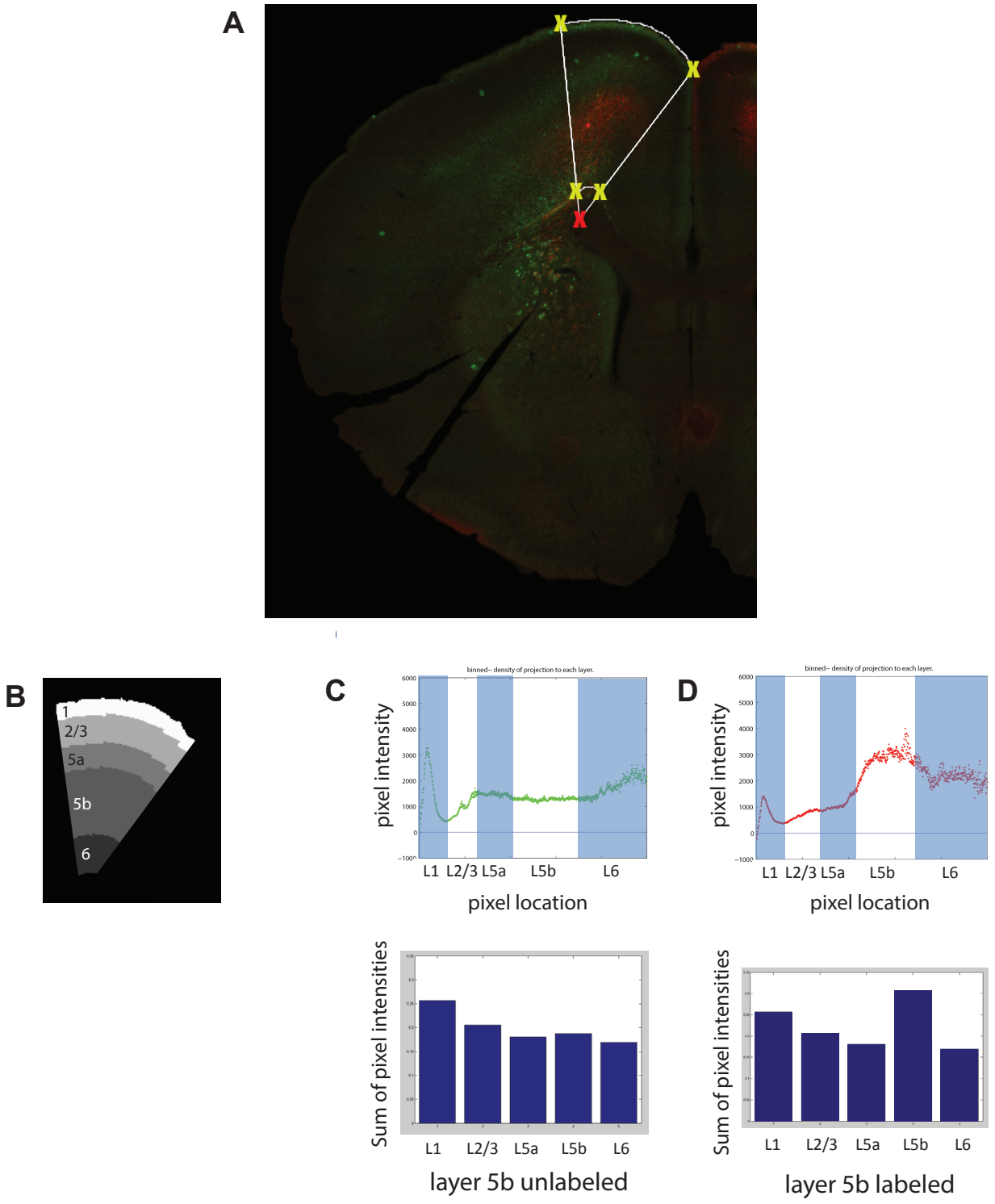

Figure 7. Layer specificity of projections in vM1. A. The vM1 region of interest is defined in a semi-automated fashion. I manually select boundary points indicated in yellow. Each pixel within the ROI is assigned a distance (*r*) value between 0-1, based on its position within the ROI calculated from the intersection of the two sides of the ROI. **B**. I use the *r* values to assign each pixel to different layers of the cortex. **C**. Layer specificity of projections is quantified by adding the pixel intensity values of all pixels in each layer. The green injection was scored '-' for labeling in layer 5b. D. The red injection was scored + for labeling in layer 5b.

References:

Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., ... & Windischberger, C. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, *107*(10), 4734-4739.

Bohland, J. W., Wu, C., Barbas, H., Bokil, H., Bota, M., Breiter, H. C., ... & Mitra, P. P. (2009). A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS computational biology*, *5*(3), e1000334.

Bonhoeffer, T., & Yuste, R. (2002). Spine motility. Phenomenology, mechanisms, and function. *Neuron*, *35*(6), 1019.

Bota, M., Dong, H. W., & Swanson, L. W. (2005). Brain architecture management system. *Neuroinformatics*, *3*(1), 15-47.

Bürgel, U., Amunts, K., Hoemke, L., Mohlberg, H., Gilsbach, J. M., & Zilles, K. (2006). White matter fiber tracts of the human brain: three-dimensional mapping at microscopic resolution, topography and intersubject variability.*Neuroimage*, *29*(4), 1092-1105.

Chamberlin, N. L., Du, B., de Lacalle, S., & Saper, C. B. (1998). Recombinant adeno-associated virus vector: use for transgene expression and anterograde tract tracing in the CNS. *Brain research*, *793*(1), 169-175.

Chen, B. L., Hall, D. H., & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(12), 4723-4728.

Chklovskii, D. B., Vitaladevuni, S., & Scheffer, L. K. (2010). Semi-automated reconstruction of neural circuits using electron microscopy. *Current opinion in neurobiology*, *20*(5), 667-675.

Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *The Journal of Neuroscience*, *9*(7), 2432-2442.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS biology*, *6*(7), e159.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10-18.

Hooks, B. M., Mao, T., Gutnisky, D. A., Yamawaki, N., Svoboda, K., & Shepherd, G. M. (2013). Organization of Cortical and Thalamic Input to Pyramidal Neurons in Mouse Motor Cortex. *The Journal of Neuroscience*,*33*(2), 748-760.

Jones, E. G., Steriade, M., & McCormick, D. (1985). *The thalamus* (p. 935). New York: Plenum Press.

Köbbert, C., Apps, R., Bechmann, I., Lanciego, J. L., Mey, J., & Thanos, S. (2000). Current concepts in neuroanatomical tracing. *Progress in neurobiology*, *62*(4), 327-351.

Kuramoto, E., Furuta, T., Nakamura, K. C., Unzai, T., Hioki, H., & Kaneko, T. (2009). Two types of thalamocortical projections from the motor thalamic nuclei of the rat: a single neuron-tracing study using viral vectors. *Cerebral Cortex*, *19*(9), 2065-2077.

Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., ... & Kuan, C. L. (2006). Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, *445*(7124), 168-176.

Mao, T., Kusefoglu, D., Hooks, B. M., Huber, D., Petreanu, L., & Svoboda, K. (2011). Long-range neuronal circuits underlying the interaction between sensory and motor cortex. *Neuron*, *72*(1), 111-123.

Markram, H. (2006). The blue brain project. *Nature Reviews Neuroscience*,*7*(2), 153-160.

Crick, F., & Jones, E. (1993). Backwardness of human neuroanatomy.*Nature*, *361*(6408), 109.

Oberlaender, M., Ramirez, A., & Bruno, R. M. (2012). Sensory experience restructures thalamocortical axons during adulthood. *Neuron*, *74*(4), 648-655.

Pan, W. X., Mao, T., & Dudman, J. T. (2010). Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. *Frontiers in neuroanatomy*, *4*.

Paxinos, G., & Franklin, K. B. (2004). *The mouse brain in stereotaxic coordinates*. Gulf Professional Publishing.

Ramón, S., & Junquera, C. (1904). *Textura del sistema nervioso del hombre y de los vertebrados*

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein.*Nature biotechnology*, *22*(12), 1567-1572.

Steiner, H. & Tseng, K. Y. in *Handbook of basal ganglia structure and function* (Academic Press, 2010).

Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*, *66*(5), 781-795.

Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., & Chklovskii, D. B. (2011). Structural properties of the Caenorhabditis elegans neuronal network. *PLoS computational biology*, *7*(2), e1001066.

White, J. G., Southgate, E., Thomson, J. N., & Brenner, S. (1986). The structure of the nervous system of the nematode Caenorhabditis elegans.*Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *314*(1165), 1-340.

Wimmer, V. C., Bruno, R. M., de Kock, C. P., Kuner, T., & Sakmann, B. (2010). Dimensions of a projection column and architecture of VPM and POm axons in rat vibrissal cortex. *Cerebral cortex*, *20*(10), 2265-2276.