TOWARDS LARGE-SCALE NEURAL CIRCUIT MAPPING AND ANALYSIS USING ELECTRON MICROSCOPY

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

STEPHAN URS GERHARD

MSc in Neural Systems and Computation, ETH Zurich

born on 02.05.1984

citizen of Zurich, Switzerland

accepted on the recommendation of

Prof. Dr. Richard Hahnloser, examiner
Dr. Albert Cardona, co-examiner
Prof. Dr. Kevan Martin, co-examiner
Dr. Matthew Cook, co-examiner

2015

ABSTRACT

One of the great scientific challenges of our time is the reverse-engineering of the algorithmic principles operating in nervous systems. Progress in our understanding of these principles requires the elucidation of the structure of neural circuits. The structure of neural circuits can be obtained by mapping the morphology of neurons and annotating their synaptic connections using electron microscopy (EM) methods. Large EM image volumes of neural tissue can be generated routinely at nanometer resolution with automated acquisition methods. However, the size of these volumes and the complexity of neuronal arbors renders the extraction and analysis of neural circuits slow and tedious. This represents a major bottleneck for the analysis of neural circuits.

To tackle this bottleneck, we developed a web-based open-source software for the mapping, analysis, and visualization of neural circuits (www.catmaid.org). This software enables fast, accurate, and collaborative mapping of neural circuits of interest by globally distributed groups of researchers. The software implements a novel iterative, non-redundant circuit mapping approach. This approach was validated by mapping a subset of neurons in a proprio-motor circuit of the *Drosophila melanogaster* larval ventral nerve cord. We compared the reconstruction speed and accuracy from our approach to state-of-the-art, redundant methods. Results yielded similar levels of accuracy at a faster reconstruction speed for our approach. Detailed analyses suggest that cellular neuroanatomy of connectivity of *Drosophila* neurons are decisive for the achieved accuracy. These properties generalize to *Drosophila* neurons at different life stages and cell types, and enable robust and efficient mapping of neural circuits in *Drosophila*. The toolkit is currently applied to map circuits across the phylogenetic tree including mammals.

I applied this novel mapping approach to investigate how synaptic circuits and their properties change between developmental stages using the *Drosophila* nociceptive system as a model. Previous studies suggest that noxious stimulation causes behavioral phenotypes at the late stages of larval development that are absent at early stages. The question thus arises how the underlying synaptic circuits for nociception change across development. Volumetric EM datasets were obtained from several individuals using large-scale, serial-section transmission electron microscopy. Our novel circuit mapping method was applied to reconstruct the class IV multi-dendritic nociceptors and their postsynaptic circuitry at both early and late developmental stages. Changes in synaptic connectivity patterns and morphological properties of neurons were investigated between early and late developmental stages. Results revealed that the general organization and synaptic connectivity of all nociceptive postsynaptic interneurons were preserved from the early to the late stages. Moreover, across developmental stages, interneuron arbors grew considerably and synapse numbers increased 3-4-fold. However, the proportion of the nociceptor inputs relative to the total number of dendritic synaptic inputs remained similar and was cell type-specific. Furthermore, different types of local interneurons receive inputs from different subsets of somatotopically-organized nociceptors, suggesting parallel, specialized pathways for noxious signal processing. The newly identified interneuron types and their genetic driver lines provide a basis for future research to dissect this tractable model system for nociception.

In summary, this work contributes tools and methods towards mapping, analyzing and visualizing large-scale neural circuits derived from volumetric EM and demonstrates their practical applicability. This work hints at the tremendous potential of circuit mapping studies to elucidate the relationship of synaptic circuit maps to neural function and animal behavior.

ZUSAMMENFASSUNG

Die Entdeckung der algorithmischen Prinzipien der Funktion von Nervensystemen ist eine der grossen wissenschaftlichen Herausforderungen unserer Zeit. Die Rekonstruktion der Struktur von neuronalen Schaltkreisen ist erforderlich um Fortschritte in unserem Verständnis dieser Prinzipien zu erlangen. Die Struktur der neuronale Schaltkreise kann durch die Rekonstruktion von Neuronen Morphologien und der Markierung von synaptischen Verbindungen mittels Elektronenmikroskopie (EM) erhalten werden. Das Nanometer-Auflösungsvermögen von EM und automatisierte Erfassungsmethoden erlauben grosse Mengen an 3D-Bilddaten von Nervengewebe routinemäßig und auf grossen Skalen zu erzeugen. Die Extraktion und Analyse neuronaler Schaltkreise ist jedoch langsam und mühsam und stellt daher den Haupt-Engpass dar für die Gehirnrekonstruktion.

In dieser Arbeit haben wir eine neuartige, web-basierte Open-Source-Software-Umgebung für neuronale Schaltkreis Kartierung, Analyse und Visualisierung (catmaid.org) entwickelt. Die Umgebung ermöglicht global verteilten Forschergruppen eine schnelle, genaue und kooperative Kartierung von interessanten, neuronalen Schaltkreisen in sehr grossen 3D-EM-Bilddatensätzen. Als Machbarkeitsstudie haben wir einen Proprio-Motor-Schaltkreis in der *Drosophila melanogaster* Larve kartiert, sowie neuartige Zelltypen identifiziert und eine Vielzahl von Verschaltungsmotiven analysiert. Wir demonstrieren, wie die neuroanatomischen Eigenschaften von *Drosophila* Neuronen unserem Kartiertungsverfahren Robustheit verleiht. In einer Validierungsstudie zeigen wir weiter, dass unser iterative, nicht-redundante Kartierungsansatz ein Mehrfaches schneller ist als aktuelle, redundante Methoden und exakte Schaltpläne generieren kann. Der interaktive Aspekte der Analyse und Visualisierung in Echtzeit, sowie die reibungslose Navigation von Bilddaten und Schaltkreismerkmalen in unserer Umgebung, sind von entscheidender Bedeutung für die Erforschung und das Verständnis von Schaltungsstrukturen in *Drosophila* und möglicherweise auch andere Nervensystem.

Als nächstes haben wir weitere *Drosophila* Larven Schaltkreise auf der synaptischen Ebene kartiert und zwischen Individuen und über Entwicklungsstadien verglichen. Bei den späten Larven Entwicklungsstadien verursacht schädliche Stimulation bestimmte Verhaltensphänomene, die in einem frühen Stadium nicht vorhanden sind. Wir untersuchten die offene Fragestellung, inwiefern sich die zugrunde liegenden synaptischen Verschaltung für Nozizeption über die Entwicklung ändert. Wir haben die class IV Nozizeptoren und ihre nachgeschalteten Interneuronen im zentrale Nervensystem in frühen und späten Entwicklungsstadien mittels grosen Serienschnitten von Transmissionselektronenmikroskopie abgebildet. Wir fanden, dass die allgemeine Organisation und die Verschaltung von 14 identifizierten Interneuronentypen von frühem bis spätem Stadium erhalten bleibt. Überraschenderweise bleit auch nach erheblichen Wachstum von Interneuronen und der 3-4-fache Erhöhung der synaptischen Verbindungen von frühem bis spätem Stadium der Anteil der Nozizeptoren

Typ-spezifische synaptische Eingänge in beiden Stadien erhalten. Darüber hinaus erhalten verschiedene Arten von lokalen Interneuronen einen zelltypspezifische Eingang von den somatotopisch organisierten Nozizeptoren, was auf parallele, potentiell spezialisierte, Leitbahnen für die Transduktion von schädlich Signalen hinweist. Die neu identifizierten Interneuronen können mit genetischen Methoden weiter funktionell analysiert werden und tragen zur Etablierung dieses Model-Systems für die Untersuchung von Nozizeption bei.

Zusammenfassend haben wir Werkzeuge und Methoden für die Kartierung, Analyse und Visualisierung von neuronalen Schaltkreise abgeleitet von grossen, volumetrischen Elektronenmikroskopie Datensätzen beigetragen. Wir verwendeten diese neuen Werkzeuge und Methoden, um verschiedene neuronale Schaltkreise zu extrahieren und haben damit verschiedene neurobiologische Fragen adressiert. Dabei haben wir das enorme Potenzial dieser Werkzeuge für zukünftige Studien gezeigt, die unser Verständnis der Struktur-Funktions-Beziehungen oder für vergleichende Fragestellungen in den Neurowissenschaften, voranzubringen. Zukünftige Studien werden nun weiter die Gemeinsamkeiten und Unterschiede der neuronalen Schaltungsstruktur charakterisieren können in verschiedenen Individuen, Spezies, in verschiedenen Entwicklungsstadien oder in Gehirnen von gesunden und erkrankten Personen.

ACKNOWLEDGMENTS

CONTENTS

INTRODUCTION

1.1 GENERAL INTRODUCTION TO CONNECTOMICS

The drawings of the elaborate arborization of neurons by Ramón y Cajal in the 19th century have captivated the imagination of explorers of the mind-brain connection ever since. Generations of neuroanatomist have followed suit and utilized state-of-the-art technology of their times to visualize and reconstruct neurons in exquisite details in many species and brain regions (Senft 2011). However, the spatial resolution limits of light microscopy prevented investigators to describe synapses - the contact points between neurons. Only with the advent of electron microscopy (EM) in the 1950s, investigators could overcome this limitation and obtain the first images of chemical synapses and their pre- and postsynaptic specializations (Palade 1954).

Recently, advances in automation and scaling of EM data acquisition technology form the basis for the newly emerging field of large-scale cellular EM connectomics (Briggman et al. 2006a; Briggman et al. 2011a). Its major goals are to apply high-throughput, large-scale EM acquisition and data processing methods to elucidate the intricate networks of neurons and their detailed synaptic connectivity. A major motivating force behinds these efforts is that neural circuit wiring diagrams help to constrain models of circuit function, support the generation of new hypotheses, or guide new experimental work and are therefore necessary to advance our understanding of brain function and dynamics (Bargmann et al. 2013).

Neuroscience conceptualizes animal behavior and subjective experience as the result of computations in neural circuits (Sherrington 1906; Carandini 2012). These neural computations and their dynamics are constrained by the underlying circuit structure. In turn, neural computations associated with behavioral learning and memory alter the structural properties of neurons and synapses at cellular and molecular levels (Mayford et al. 2012). In order to understand the detailed interplay between circuit structure and function, the production of structural circuit maps envisioned by EM connectomics will be crucial to inform the interpretation of physiological circuit data (Morgan et al. 2013).

A first deep insight into a neural structure-function correlate was the hypothesis that synaptic vesicles are the structural source of spontaneous miniature pulses observed in neurons with microelectrode recordings (Palay et al. 1955; De Robertis et al. 1955, reviewed in Wells 2005). This was the first time, 60 years ago, that structural EM provided a confirmation for the functional hypotheses of quantal transmitter release. Another 32 years later, a study found a reduction in the number of presynaptic vesicles docked at presynaptic release sites after induction of synaptic depression in *Aplysia* (Bailey et al. 1988) - a first glimpse into a structural correlate of learning. These pioneering studies exemplify the potential to advance our understanding of synaptic and circuit function by studying the ultrastructural details of the nervous

system. Scaling up EM connectomics to larger volumes will enable studies of the structural underpinnings of function at the whole circuit level.

To date, only a few studies exist where connectomic circuit-level information could be directly linked with functional features of neurons and circuits. In a recent study, the asymmetry in specific synaptic connections in retinal starbust amacrine cells have been found to contribute to the directional tuning of neurons (Briggman et al. 2011b; Kim et al. 2014). In another study, the computation of motion anticipation was found to be dependent on the distribution of inhibitory and excitatory synaptic inputs on the dendritic trees of retinal ganglion cells (Johnston et al. 2015). These studies are precursors that exemplify how circuit-level synaptic connectivity information obtained from EM can support and inform a mechanistic understanding of the computations in neurons and circuits (Koch 2004).

Beyond better understanding normal circuit function, EM connectomics may also provide new windows into brain diseases in animal model systems - and eventually in humans. Pathologies in synaptic connectivity have been hypothesized to underlie various neurological and psychiatric conditions. For instance, autism, epilepsy or Rett syndrome are thought to result, at least in part, from the abnormal growth and development of the nervous system. Such neuro-developmental disorders are caused by abnormalities of circuit wiring leading to aberrant neural computations, which then ultimately negatively affect an individual's emotion, cognition, learning abilities, self-control or memories. Hence, one promising outcome of large-scale EM connectomics efforts could be the uncovering of circuit structure signatures of such disorders. With this knowledge, circuit-based intervention strategies could be designed that target specific neuronal cell types. Treatment outcomes can be evaluated as deviations from normal distributions on the level of synaptic properties and connectivity (Karayiorgou et al. 2012). With novel neuronal cell replacement therapies on the horizon, the integration of neurons derived from stem cells into existing circuitry could be carefully evaluated and calibrated (Bargmann 2015).

An important requirement for such highly quantitative EM studies is quantifying the accuracy and variability of synaptic connectivity. Biological synaptic variability arises from developmental differences between genetically identical individuals and from differences arising through experience. In addition to the biological variability, errors in the reconstruction process can introduce spurious variability. For strong connections in the wiring diagram, i.e. connections with many synaptic contacts between neurons or neuron types, the reconstruction error might be negligible, and therefore does not impact the interpretation of circuit information flow. However, weak connections in the wiring diagram can be functionally relevant, and their discovery or proper quantification requires very low error rates in the reconstruction process in order not to mask true biological differences.

A major source of biological variability of circuit connectivity arises from the stochastic nature of neural circuit development. Although genetic programs dictate the rules for the growth of different neuronal cell types and the formation and elimination of synapses, these processes are not deterministic and synaptic connectivity is thus variable. An open question is, for instance, how this circuit variability relates to physiological or behavioral variability observable in populations of individuals.

The study of synaptic variability in genetic model organism at the same age, with the same genotype and with very similar sensory experiential histories are an opportunity to provide first insights into the true biological variability of neural circuits, and the redundancy necessary to ensure their robust functioning. Such studies have also been recognized as one of the main subgoals of the recently launched BRAIN Initiative in the United States (Jorgenson et al. 2015). The species of *Drosophila melanogaster* is an advantageous, tractable genetic model system to study questions of circuit variability, because genotypes, developmental stage and the exposure to stimuli from its environment can be well controlled. Furthermore, its nervous system and neurons are small compared to other systems (see Figure 1.1) which allows to minimize reconstruction errors.
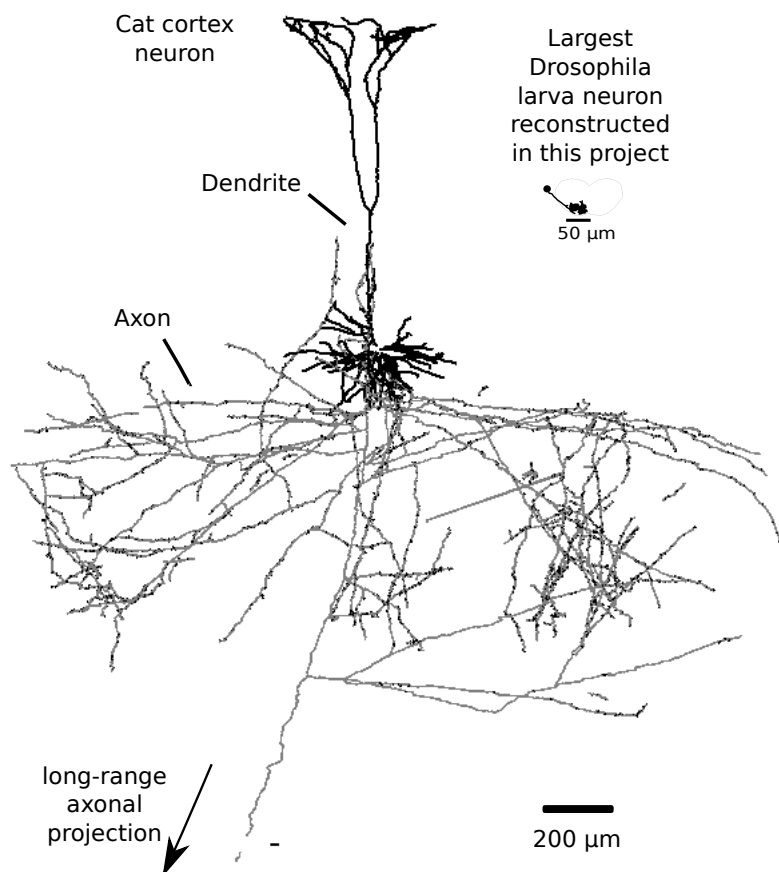
Figure 1.1.: **Size comparison of invertebrate and vertebrate neurons.** Dendritic and axonal arbors of tufted layer 5 pyramidal cell in cat visual cortex V1 (Binzegger et al. 2004). The dendrite of the pyramidal cells receives more than 10'000 inputs. The largest neuron reconstructed in this work is a *Drosophila* larva ventral nerve cord interneuron - a Basin-2 cell - with a total cable length of 1.68 mm with 1600 inputs and over 1000 outputs. For comparison, both neurons are shown at the same scale.

Ultimately, measures of synaptic accuracy and variability will be crucial to evaluate how genetic mutations and environmental factors contribute to deviations in neural circuit structure underlyying behavioral and experiential abnormalities. In this thesis project, we contribute to different aspects of these long-term goals.

In **Chapter 2**, we present a novel web-based framework to map, analyze and visualize neural circuits in large-scale, volumetric EM datasets. We implemented a method for targeted circuit reconstruction that is both fast and accurate by leveraging neuroanatomical knowledge of *Drosophila* neurons. The method delivers sparse connectomes that can answer specific neurobiological questions in short time frames. We validated the method in a small *Drosophila melanogaster* larva proprio-motor circuit. We demonstrate the viability of the entire approach by reconstructing a large proprio-motor *Drosophila* circuit and analyze circuit motifs and cell types that underlies larval locomotion.

In **Chapter 3**, I use the novel method and framework to address basic questions of comparative circuit neuroanatomy in the nociceptive system of *Drosophila* larvae. The advantageous fact that larvae are genetically-identical and share similar life experiences are both factors which reduce confounds when comparing circuits. I evaluate the extent of similarities and differences in the nociceptive circuit architectures both inter-individually, and as larvae develop from early to late stages. I tried to relate these changes in synaptic circuitry with changes in nocifensive behavioral patterns that change in larval development, and uncovered a number of anatomical circuit invariants. Additionally in **Chapter 4**, I performed detailed comparisons of changes and invariants in neuronal and synaptic features in different neuron types between individuals and across developmental stages.

The **Chapter 5** presents work aiming at automating circuit reconstruction workflows. The production of ground truth datasets enabled us to automatically extract highly accurate volumetric information of a variety of structures in EM volumes such as membranes, synapses, mitochondria and glia cells with state-of-the-art classification methods. We used this data to develop a prototype pipeline for large-scale neuron segmentation. Newly developed user interfaces were applied for semi-automated volumetric neuron reconstruction and crowd-sourceable corrections of results from neuron segmentation algorithms.

1.2 HISTORICAL BACKGROUND

In the past century, a series of remarkable inventions and studies lead to the conception of the nervous system as a complex, adaptive signaling network. Early anatomical and functional methods provided a wealth of data about single neurons, and this focus importantly shaped the doctrine to understand nervous systems based on single neurons as conceptual units. Novel, large-scale structural and functional neural circuit mapping methods have started to shift this focus on single neurons to a more comprehensive view of neural networks and groups of co-activated neurons as basic building blocks of our understanding of nervous system function (Yuste 2015).

In this fragmentary historical overview, I highlight a few important historical advances from the early conception of neurons and synaptic connections to our current efforts in trying to map structural networks from large EM volumes. These large-scale networks offer new windows into the complexity of nervous systems, and novel tools for their visualization and analysis promise to transform data into knowledge and insight. The interested reader is also referred to a recent in-depth review on the history of neuron reconstruction methods by Senft 2011.

For the first time, the visualization of individual nerve processes was made possible by the invention of the *black reaction* by Camillo Golgi (Golgi 1873). This reaction allowed investigators to appreciate the intricate arborization patterns of individuals neurons that constitute the nervous system. It marked the beginning of a new era in the anatomical study of the nervous system. The method contributed importantly to the *neuron doctrine*, the conception that the nervous system is made up of discrete individual cells (reviewed in Pannese 2007). Ramón y Cajal was one of the most effective advocates of this doctrine and extensively used the Golgi method to gather evidence. Golgi, the inventor of the method, however, believed in another theory that conceptualized the nervous system as a single continuous entity. It was not until the invention of the electron microscope that these conflicting views could be settled - in favor of Cajal.

Using the Golgi method, labeling of individual neurons in many species and brain regions revealed distinct differences in arbor morphologies, and neuroanatomists soon recognized geometric similarities. Similarities of arbor morphology became soon the first criteria to define *cell types* (e.g. Purkinje cell types in the cerebellum, Figure 1.2). The structural similarities that were discovered did indicate the possibility of shared functional and computational roles. Studies soon discovered common neuronal cell types across brains of different species with remarkable similarities (Figure 1.3).

On a sub-cellular level, distinct structural features of compartments of neuronal arbors further indicated a segregation of function. Parts of the arbors were found to contain thick varicosities and were thought of as sites of contact between neurons. Other parts contained large number of so-called spines, small protrusions extending from the main arbors, which were believed to be contact points between neurons (Figure 1.4). These spines were soon proposed to be points of inputs to a neuron and varicosities as points of outputs.

Figure 1.2.: **Purkinje cells type.** Purkinje cell in the cerebellum drawn by Ramón y Cajal with their intricate morphology. (Image source: Wikipedia)
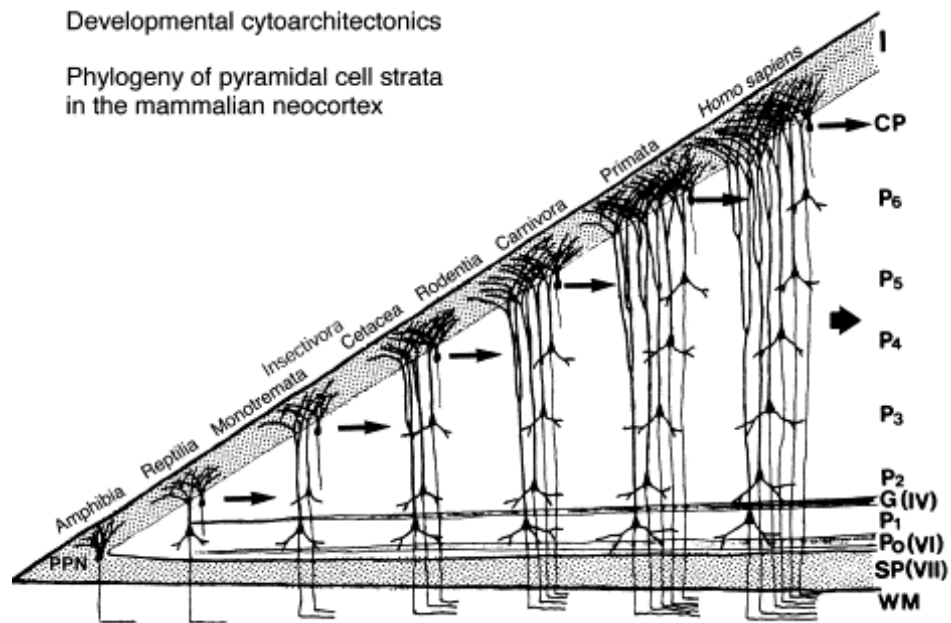
Figure 1.3.: **Evolution of an ancestral cell type - the pyramidal cells**. The evolution of the cerebral cortex of amphibians, reptilians, and mammals and their correspondences with a progressive increase in pyramidal cell strata (layers) (Reuse with permission from Marín-Padilla 1998).

This distinction of neuronal arbors into input and output regions lead William James and other investigators to conceive of the notion of networks of neurons, and the flow of some as yet undefined *nerve-energies* between neurons (Berlucchi 1999). The law of dynamic polarization was clearly proposed by Cajal which states that there is a preferred direction of transmission within neurons (reviewed in DeFelipe 2009). However, the spatial resolution of the light microscopes was a limiting factor to conclusively prove the existence of hypothetical contacts between neurons called synapses and their directionality (Sherrington 1906, illustrated in Figure 1.5). Another round of technological advancement was necessary in the history of connectomics.
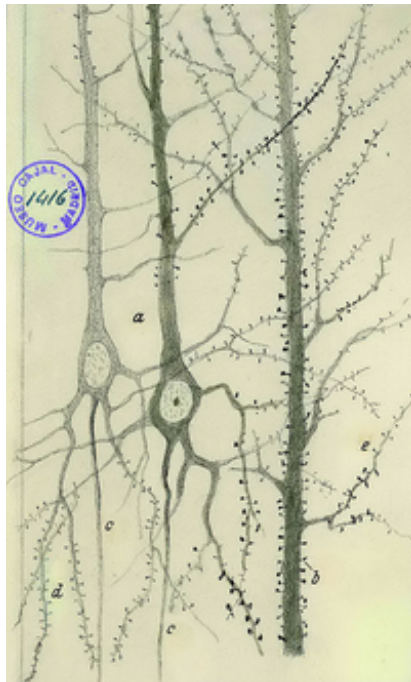
Figure 1.4.: **The discovery of dendritic spines in vertebrate neurons**. Drawing by Ramón y Cajal of pyramidal cell of rabbit cerebral cortex with clear depiction of dendritic spines. Cajal Legacy. Instituto Cajal in Madrid, Spain. (Image source: Scholarpedia)

The invention of the electron microscope (EM) in the early 1930s by Ernst Ruska provided for the first time sufficient spatial resolution to investigate neurons and their hypothetical synaptic contacts at nanoscale resolution. With this resolution, neuroanatomists could provide the first images of synapses and their synaptic cleft at the site of neuron membrane appositions (Palade 1954). These studies unambiguously proved that specialized structures exist as contact points between neurons - the chemical synapses. The discovery of synaptic vesicles as structural mediators of signals at synapses (reviewed in Wells 2005) further confirmed the proposed uni-directional action at these sites (Sherrington 1906). This evidence further corroborates the proposed law of dynamic polarization, now being able to more clearly state the preferred direction of information flow from dendrites to axons (reviewed in DeFelipe 2015).

LM image of a single neuron

no postsynaptic
site suggested
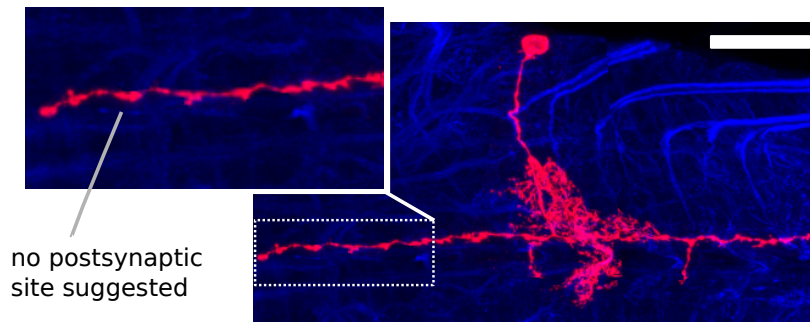
EM Reconstruction of the same neuron

P —— A

presynaptic site
postsynaptic site
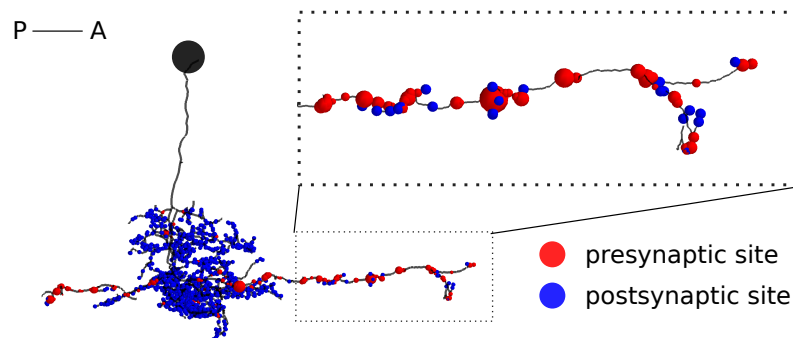
Figure 1.5.: **Postsynaptic sites on axonal arbors are below the resolution limit of light microscopy** Typical confocal stacks with in-plane resolution of 0.3μm can not resolve postsynaptic sites on axonal arbors. Light microscopy image shows dorsal view of a Basin-2 neuron in *Drosophila* larva ventral nerve cord (Image courtesy of Jim Truman). Visibility of varicosities in the light microscope indicates the axonal character of the output arbor. However, any inputs to this axonal arbor are masked by the resolution limit of the light microscope. To resolve these inputs electron microscopic nanometer resolution is necessary. Input sites are directly on axonal varicosities or on thin, filopodia-like processes protruding from the varicosities, and can be resolved by EM.

The stage was now conceptually set for connectomics to conceive the mapping of an entire nervous system. At this point in time, it was clear what the *known unknowns* were. The major technical challenges were essentially how to scale electron microscopy methods to larger volumes and more samples, while preserving the best possible sample quality. To study entire neural networks in whole animal brains, neuroscientists ultimately needed to image sufficiently large volumes that contain complete neuronal morphologies and all its synaptic contacts. A major challenge for such a project was the limited field-of-view of EM methods.

In order to image volumes of neural tissue, samples needed to be cut into series of thin sections, on the order of tens of nanometer in thickness, and imaged individually and manually using transmission electron microscopes (TEM). The reliability of the

cutting and section pickup process depended mainly upon the skills of the electron microscopist, and the throughput was strongly limited due to the manual nature of the task.

Throughput in EM imaging was also limited. First, the microscopes had to be operated manually to image individual portions of the section. Second, thin sections were imaged on film. Subsequently, these films had to be developed one-by-one in the dark room by a laborious manual process. Despite all these challenges, dedication allowed neuroscientists already in the 1970s to reconstruct neurons in 3D from series of 2D micrographs. This projects can be seen as the first connectomes in the visual system of crustacean Daphnia magna (Horridge et al. 1970; Macagno et al. 1973).

A pragmatic approach to overcome the field-of-view limitation for whole nervous system imaging was choosing to work in a system that *fits the view*. This lead to the pioneering and ambitious project conceived by Sydney Brenner in the 1970s to map the complete connectome of the nematode *Caenorhabditis elegans* (Emmons 2015). An attempt was made to use an advanced laboratory computer for the reconstruction of neurons and synapses, but the 64 KB of memory and 22 MB storage space were insufficient to handle the amount of EM data. However, by manually marking neurons with colored pens on large prints of the EM micrographs, the complete connectome was eventually published - after 15 years from the initial start (White et al. 1986).

The following decades saw improvements in volumetric electron microscopy techniques and a revolution in Information Technology (Briggman et al. 2006a). Computers became sufficiently powerful to handle and process large collections of digital images. At the beginning of the 21st century, we have entered an era where EM datasets can be produced at an unprecedented size and the extraction of circuit data from these large volumes has become the main limiting factor.

In anticipation of the future, once the extraction of wiring diagrams from large-scale EM volumes becomes fully automated, we will be facing the next big challenge. How to make sense of all this intricate and complex circuit data? Thinking about the trajectory of this development, Sydney Brenner has envisioned the possibility of a framework which he called CellMap in his Nobel laureate speech (Brenner 2003). This framework should be able to store vast amounts of biological data and help researchers to query and extract knowledge and insights about the organization of nervous systems at molecular, cellular and network levels. Eventually, such a system will also need to incorporate physiological and behavioral data to build predictive models of neural circuit dynamics and function. Importantly, for such a system to work, Brenner emphasized the significance of researchers working as a mutually-supportive community.

In the coming decades, I believe that this collaborative aspect will become increasingly important for large-scale circuit mapping projects, and for neuroscience in general. A new era of a networked, collaborative science enterprise is on the rise, supported by modern web technology that harnesses collective intelligence (Nielsen 2013). Facing the intricate complexity of neurobiological organization at its many different spatial and temporal levels will definitely require new and creative solutions. I hope that this thesis work contributes a step towards this future.

1.3 INTRODUCTION TO CIRCUIT MAPPING

Advances in volumetric electron microscopy can deliver unprecedented image volumes of neural tissue at high spatial resolution (Marx 2013). Although currently, large by the standards of EM connectomics means *fractions of a cube millimeter*, this size is sufficient to completely fit interesting brains such as that of an adult *Drosophila melanogaster* or larval zebrafish with approximately 100'000 neurons. In large-scale EM connectomics projects now underway, the amount of raw image data produced from such brains requires storage capacities of hundreds of terabytes. Our ability to collect data far exceeds our ability to extract neural circuits from these large image volumes by orders of magnitude. This gap is currently the major bottleneck in connectomics and hinders progress in mapping the structure of neural circuits at large scales.

The first, pioneering large-scale connectomic project was done in the nematode *Caenorhabditis elegans*. To this date, the *C. elegans* connectome is the only complete circuit map of an entire animal (White et al. 1986). To map the 302 neurons, 5000 synapses and 600 gap junctions, the pioneering neuroanatomists spent more than 15 years to complete the project (Brenner 2003, reviewed in Emmons 2015). As is still the case today, the majority of the work was and is spent on manually extracting the neurons and synapses from serial images of electron micrographs. However, at the time of the *C. elegans* mapping project this was without any help from computers and only aided by pen technology and perseverance.

In more recent large-scale connectomics projects aiming at mapping complete circuits, the time from the first data acquisition run to a partially completed connectome is much shorter. A connectome of the inner plexiform layer of mouse retina containing 950 neurons was produced using over 20'000 man-hours of manual, redundant skeletonization (Helmstaedter et al. 2013). A map of 600 neurons with over 8600 synaptic edges was produced in the rabbit retina by a team of dedicated annotators over several years (Marc et al. 2014). Also in the mouse retina, the largest crowd-driven connectomics project, EyeWire [1], counts more than 100'000 registered users, with about 100 contributors doing more than half of the work. The produced connectome contained 195 bipolar cell axons and 79 starburst amacrine cells (Kim et al. 2014). In the optic system of *Drosophila melanogaster*, Takemura et al. 2013 reconstructed 379 insect neurons using about 14'000 man-hours worth of proofreading. Direct comparison of these efforts are difficult due to differing workflows and systems. Nonetheless, they clearly demonstrate the scale of the challenge when it comes to circuit mapping and proofreading.

In the coming decades, interest in such datasets is likely to grow, and currently advances in accelerating the speed of data generation are underway and will continue. A recent technological achievement uses a multi-beam scanning EM technology and promises to increase imaging speed several-fold (Eberle et al. 2014) with the possibility to parallelize acquisition of a single sample (Hayworth et al. 2014). With such speed-ups, the cost for acquiring a whole mouse brain was estimated to cost less than $100 million dollar (Marblestone et al. 2013). Such an imaging run would

---

1 www.eyewire.org

produce a dataset of over 60 petabytes (Mikula et al. 2015). With further advances in parallelization and automation in high-throughput EM data collection, the gap between data generation and data analysis will continue to widen further. Even with the best circuit mapping approaches available today, completion of a mouse connectome would presumably take several decades.

What is our contribution to tackle these challenges in large-scale connectomics projects? We developed a software framework that can be installed in labs or institutions and allows researchers to effectively map, analyze and visualize circuits in terabyte-sized EM image datasets. Due to the large size of these datasets, we adopted a client-server architecture where researchers access relevant subregions of the data online for collaborative annotation in their preferred web browser. Annotation data are stored and backed up in a central database. The servers hosting entire raw EM datasets provide efficient access to the image data. These servers can be mirrored to locations around the globe to improve data access speed.

In the framework, we developed and implemented a novel circuit reconstruction methodology that is both fast and delivers accurate neural circuitry. We achieved this by an intelligent combination of software architectural decisions and insights from neuroanatomy, in particular from *Drosophila* neurons, with a pragmatic outlook on delivering circuits of interest in manageable time frames by sparse reconstruction (Costa et al. 2013). The basis of our method uses an iterative and non-redundant approach to circuit mapping which employs high-level features of *Drosophila* neuroanatomy for focused proofreading and quality control. The details of the method are introduced in detail in Chapter 2. Three additional aspects of our framework and method confer advantages for large-scale circuit mapping projects: the focus on circuits of interest guided by specific biological questions, the reduction of redundancy in the reconstruction, and collaboration and parallelization.

First, in our framework brain mapping projects can advance incrementally with opportunities for major results by contributing labs and researchers on the way. Neuroscientist driven by particular research questions need to be able to address them in a reasonable amount of time compatible with the duration of academic projects. We sought to decrease the iteration cycle time from producing a useful circuit-of-interest to informing further experiments and subsequent circuit mapping and analysis. In this incremental approach, a circuit of interest is defined by a set of research questions which delineates what neurons and circuits are mapped with high priority. As the circuit map takes shape and is explored in real-time, further mapping can be confined to specific subsets of neurons, such as specific synaptically-connected pre- and postsynaptic partner neurons that seem most promising for follow-up questions.

With this step-wise procedure, the iteration cycle from mapping to insight is reduced. This approach prevents situations that are the norm in dense mapping workflows where resulting circuits may be interrogated offline only after years of mapping and proofreading a complete EM volume (Seung 2009). In the workflows of the previously mentioned large-scale connectomics projects, this iteration cycle was generally very long - up to several years. Therefore, our framework provides the means for a researcher, lab or institution to keep questions matched to available

mapping resources (Costa et al. 2013), and defines an optimal route to prioritize mapping and crawl the neuronal networks efficiently.

A key requirement for sparse circuit mapping is the availability of specific neurons as targets to start the reconstruction. In a neural system with a fixed set of neurons with stereotyped morphologies and locations, visualization of single neurons imaged with light microscopy is sufficient to uniquely identify and match them to EM reconstructions (Ohyama et al. 2015). Cell body location and major arborization patterns serve as landmarks to quickly find one's neuron of interest in a large EM volumes.

Another promising strategy in less stereotyped systems consist in aligning light-microscopy image stacks from calcium imaging experiments with EM volumes to find matching cell bodies that then serve as seeds for mapping (Bock et al. 2011; Briggman et al. 2011b). Yet another approach is array tomography (Micheva et al. 2007) which is based on correlating light and electron microscopy. Neuron types in EM volumes are distinguished based on different, overlayed fluorescent labels which serve as starting points for whole circuit reconstruction (Oberti et al. 2011). We aimed at designing our framework to support all of these use cases. For instance, landmark detection is facilitated by fast, multi-scale browsing of image stacks and colored, transparent light microscopy images can be overlayed on top of the raw EM images.

Second, our approach to mapping individual neuronal arbors and annotating synapses should proceed as fast as possible and maximally reduce redundant manual work. Manual skeletonization of the center lines of individual neuronal processes is orders of magnitude faster than volumetric pixel labeling (Helmstaedter et al. 2011). Additionally, the skeleton is a compact representation of neurons and a good proxy for overall neuronal shape and a variety of biologically-relevant neuronal features. This representation further facilitates interactive, real-time analysis and visualization of large sets of neurons.

Furthermore, our user interface allows to simultaneously annotate synapses, i.e. all pre- and postsynaptic sites on a neuronal arbors, while skeletonizing neuronal arbors. Many of the currently applied approaches separate synapse annotation from arbor reconstruction and thereby increase the annotation workload. With our approach, the circuit topology and synaptic counts between neurons can be extracted quickly.

Our approach aimed to reduce redundant reconstruction work as much as possible. The state-of-the-art method introduced in Helmstaedter et al. 2011 spends a large fraction of the 20'000 man-hours to redundantly skeletonize 950 neurons with a factor of four on average. The redundancy for individual skeletons can be up to seven. In our reconstruction method, we only require a first initial reconstruction and subsequent review passes, typically less than two, which are faster than initial reconstruction. In Chapter 2, we demonstrate in a validation experiment that our method can achieve similar levels of accuracy to state-of-the-art methods while reducing the total reconstruction time.

Third, we have built our application in such a way that workflows for mapping and analysis are integrated in the same system, and can be performed online, collaboratively and in parallel. To achieve this, we extended the existing open-source web

toolkit CATMAID (Saalfeld et al. 2009). A major advantage of a web-based tool is that collaborative workflows can be implemented naturally through the centralized storage of all annotation data and user management. The central data storage and management has the advantage that multiple researchers can work in parallel on the same dataset. This architecture creates synergies that are difficult to leverage if researchers would work independently in the same datasets.

For instance, when researchers only reconstruct parts of a neurons and then abandon them, these parts can be readily reused by another researcher without duplicated work. Similar reuse is possible if parts of a neuron are mistakenly connected and then cut off. Resulting short fragments can be integrated into the reconstruction of other neurons later.

The annotation of synapses is by convention made so that seed skeleton nodes are placed for all pre- and postsynaptic locations. If the pre- and postsynaptic neurons are reconstructed later, the synapses are checked for correct annotation multiple times. These are examples of how the implementation of our method can naturally support data reuse and improve data quality. Overall, the collaborative virtual space so created allows for the efficient reuse of data and repeated observations improve data quality. In addition to our mapping method, this further speeds up the reconstruction and improves the quality of the resulting circuit maps.

We applied our sparse, targeted circuit mapping strategy to extract the connectome of a proprio-motor circuit in the *Drosophila melanogaster* larval ventral nerve cord underlying locomotion. The circuit mapping was performed collaboratively from two continents (Ashburn, USA and Zurich, Switzerland) from a number of contributors and yielded in over 400 neuronal arbors with over 50'000 synapses. We then use this proprio-motor circuit as a model to demonstrate how detailed circuit analysis and visualization can be performed in our framework. In particular, we elaborate in **Chapter 2** in detail how features of *Drosophila* neuroanatomy make our reconstruction approach reliable and accurate, and demonstrate how interesting biological insights can be gained by the analysis of particular circuit motifs.

1.4 INTRODUCTION TO NOCICEPTIVE CIRCUITRY

In Chapters 3 and 4, my focus is on the neural circuits for nociception in *Drosophila melanogaster*. After building the necessary software framework and developing methods for fast and accurate EM circuit mapping described Chapter 2, I apply them to address comparative questions in the *Drosophila* larva circuits that underlies the transduction of noxious stimuli.

The perception of pain after noxious stimulation is a phenomenon known to all of us. Our understanding of how pain arises made a major advance in 17th century, when René Descartes proposed a mechanistic basis of pain perception in his specificity theory (Figure 1.6). For the first time, he conceived pain as some form of *disturbance* passed down nerve fiber pathways until they reach the brain and cause a perception (Descartes 1664).

Figure 1.6.: **First depiction of a pain pathway in 1664**. Illustration of the specificity theory by René Descartes (Descartes 1664). (Image source: Wikipedia)

The first definition of nerve fibers dedicated to detect noxious stimuli came from physiological experiments of Sherrington (Sherrington 1906, reviewed in Woolf et al. 2007). He termed these fibers *nociceptors*. Their activation by intense, multi-modal stimuli triggers withdrawal and escape behaviors and autonomic responses, thereby signaling a potentially harmful event to the animal that supports its survival.

Since that time, our knowledge about the specific cellular and molecular underpinnings of the pathways and processes underlying pain perception expanded considerably (Basbaum et al. 2009). In vertebrate systems, we know a lot about different types of nociceptive fibers, their sensitivities and the general organization of circuit pathways in the spinal cord responsible for the transmission of pain signals (Kandel et al. 2012). We have also recognized that probably almost all animals display nociceptive-type responses, and most of them have dedicated nociceptors including fish (Sneddon et al. 2003) or invertebrates (Smith 1991). To what extent *pain is felt*

in 'lower' invertebrate species, however, is an open question and one refers to the behavioral escape and withdrawal correlates after noxious stimulation as nocifensive. Investigating the nociceptive circuitry within the central nervous system of these species postsynaptic of dedicated nociceptors might ultimately help to *'decide'* the question to what extent these animals can and do suffer.

To this day, however, a comprehensive, quantitative map of central nociceptive circuitry at the synaptic level in any animal species is lacking. Also for this reason, a complete understanding of central circuit mechanism in normal and pathological pain conditions remains incomplete (Todd 2010). The availability of central synaptic level circuit maps would provide an important basis for further understanding of the functional organization of central nociceptive processing. Due to volume limitations in current EM technology, comprehensive synaptic level mapping is only feasible in a model organism with a sufficiently small central nervous system (CNS).

The small CNS of *Drosophila melanogaster* larvae and its small set of of known polymodal nociceptor, the so-called class IV multi-dendritic sensory neurons (Merritt et al. 1995), provided an unique opportunity to map systematically their postsynaptic interneuron networks (Figure 1.7). In Chapter 3, I describe comprehensively the first-order nociceptive circuit map in the *Drosophila* larva ventral nerve cord. Furthermore, the availability of multiple ventral nerve cord EM volumes provided another unique opportunity to compare the same circuitry between multiple individuals. Larger sample sizes (N>1) serve to validate the mapped circuits and allow to get a first glimpse of the synaptic variability present in a supposedly stereotyped circuit.
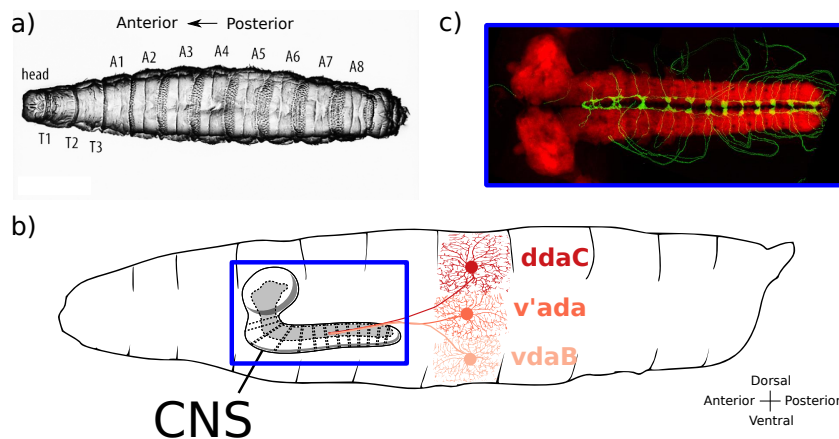
Figure 1.7.: ***Drosophila melanogaster* larvae and nociceptors** a) The larva body consists of a repeated set of segments, three thoracic (T1-T3) and eight abdominal (A1-A8). b) Outline of a larva with the central nervous system (CNS) consisting of the brain lobes and the ventral nerve cord. The three types of polymodal nociceptors tile the body wall in each hemisegment and project to the CNS. c) Dorsal view of the CNS (red) with repeated nociceptive axon terminals and projecting nerves visualized using a genetically-encoded green fluorescent protein (Image courtesy of *balapagos* on flickr).

The availability of invertebrate nociceptive circuit maps might reveal interesting similarities and differences to vertebrate nociceptive circuits. Many similarities exist in nociceptors across the animal kingdom, in particular at the level of receptors (Woolf et al. 2007; Smith et al. 2009). In both *C. elegans* and *Drosophila*, conserved sensory transduction channels and signaling molecules have been found (Tobin et al. 2004). On a circuit level, deep homologies between vertebrate and invertebrate brain organization have been proposed (Strausfeld et al. 2013). It appears conceivable that aspects of nociceptive circuit organization and processing may be common between invertebrate and vertebrate species (see Discussion in Chapter 3). The nociceptive circuit maps introduced in Chapter 3 enable further studies of the general principles of nociceptive function and contribute to establish *Drosophila* larva as a genetically-tractable model for circuit studies of nociception (Im et al. 2012).

Another unique opportunity for comparative circuit neuroanatomy was made possible by the availability of a ventral nerve cord EM volume at a late stage of larval development. With this late stage EM dataset, I aimed at comparing identified interneurons and circuits from the early to the late larval developmental stage. Larvae show withdrawal behaviors to noxious stimuli already early in their development, but the characteristic, cork-screw, escape-rolling maneuver only appears later in development (Sulkowski et al. 2011). Therefore, I asked in Chapter 3 to what degree the first-order circuits postsynaptic to nociceptors are similar and different across stages. Is a remodeling of the wiring diagram at the synaptic level observable? Are the interneuron partners the same or do they change? How does the number of synapses change and is the relative synaptic strength the same as the neurons grow? The unique identity of neurons in the *Drosophila* larva ventral nerve cord makes addressing these questions possible because individual neurons can be matched between individuals based on morphological criteria alone (Figure 1.8).

In Chapter 4, I use a subset of the local interneurons of the nociceptive circuit to quantitatively measure and compare detailed neuronal and synaptic properties at two developmental stages. With the several-fold increase in neuronal size and synaptic numbers across development, neuronal growth needs to fulfill developmental and metabolic constraints. The result of these developmental constraints are, for instance, reflected in a preserved proportion of neuronal arbor length between microtubuli-free and microtubuli-filled processes across different *Drosophila* neurons (see also Chapter 2). I extended these qualitative observations to a quantitative assessment of five different neuronal types, and compared additionally a number of single neuron features between early and late developmental stages.

One goal was to find simple descriptions and models of neuronal scaling across development, and invariant neuronal features between developmental stages that can, for instance, cluster cell types independent of their age. A number of these detailed characterizations of neurons and their compartments can be employed advantageously for constraining automated neuron reconstruction approaches, as well as for making highly accurate, biologically-realistic computational models of neuronal and circuit development.

Early larva stage neurons
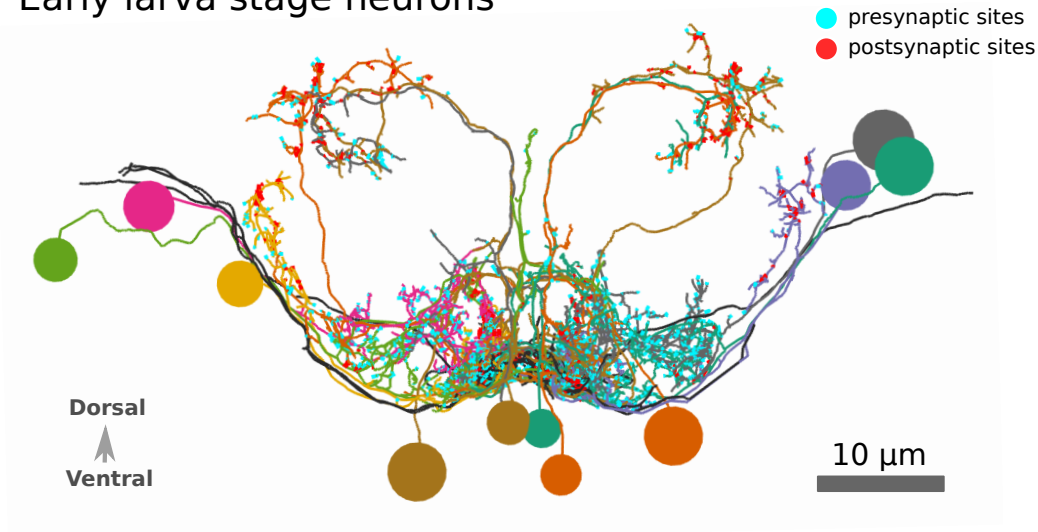
Late larva stage neurons
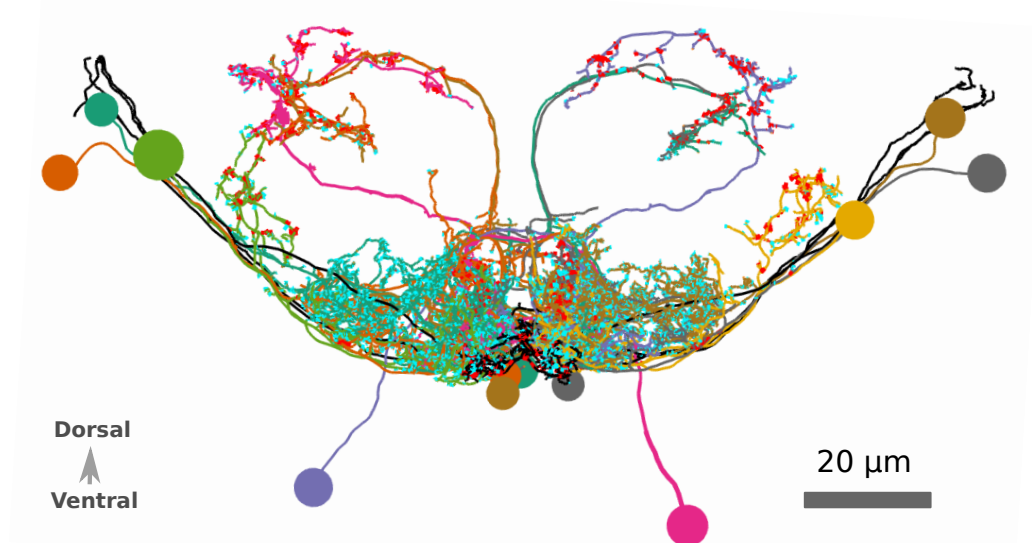
Figure 1.8.: **Visualization of the 3D morphology of local nociceptive interneurons in the ventral nerve cord at early and late developmental stages.** The shown set of interneurons is postsynaptic to primary nociceptors in larva and local to an abdominal segment of the ventral nerve cord. As the larva develops from early to late stage, its ventral nerve cord expands and neurons grow yet maintain their overall shape. Size difference is indicated by scale bars. The same set of neurons can be identified independently in each stage and individual, and therefore matched between different developmental stages.

1.5 TOWARDS SEMI-AUTOMATED VOLUMETRIC CIRCUIT MAPPING

A major future challenge for quantitative neurobiology is to extract meaningful information from *Big Data* (Freeman 2015). In EM connectomics, the anticipated *Big Data* challenges are formidable due to the large volumes of image datasets (Lichtman et al. 2014). Novel data acquisition instrumentation such as multi-beam scanning electron microscopes are now able to produce image data at unprecedented rate of 0.45 Gigapixel/second with 3.8 nm pixel size (Eberle et al. 2015). As the image production rate is anticipated to further increase, associated infrastructure must scale storage capacity and computational processing power. As an example, storing a whole mouse brain at 10 nm isotropic voxel resolution is estimated to require more than 60 petabytes for raw image data storage (Mikula et al. 2015).

In order to extract neural circuit data from such increasingly large 3D volumes, automated image segmentation pipelines need to supplement any human effort. In recent years, deep learning algorithms have dramatically improved image classification results (LeCun et al. 2015; Schmidhuber 2015). These algorithms also deliver state-of-the-art results for classifying EM images (Turaga et al. 2010; Ciresan et al. 2012).

Essentially, raw EM images are processed by a series of trained filters to produce probability maps for different classes of objects present in EM images, such as membranes or mitochondria. The generation of very accurate membrane probability maps is a crucial factor to achieve good post-processing results for automated neuron reconstruction. However, membranes are not homogeneously well preserved in EM images and are cut in different angles, thereby limiting the accuracy of probability maps. The limited accuracy can cause downstream segmentation algorithms to produce reconstructions with topological errors, such as spurious merging of neuronal processes (Funke et al. 2015). Avoiding and correcting such topological errors is an active area of research in the field of automating EM connectomics (Vazquez-Reina et al. 2011). Although improvements in sample preparation and image classification are expected to further reduce pixel error rates, topological errors will be difficult to avoid completely and will continue to require human proofreading (Plaza et al. 2014).

In Chapter 5, I demonstrate the results of a prototype image processing pipeline for large datasets based on an existing neuron reconstruction algorithm (Funke et al. 2012). I also generated a new ground truth dataset (Gerhard et al. 2013) that was applied to train deep convolutional networks for EM image data classification (Ciresan et al. 2012). Using the results of this segmentation pipeline, I developed two user interfaces, one to select volumetric neuron segmentations interactively in CATMAID, and a second interface for correcting topological reconstruction errors resulting from limitations in the segmentation pipeline. These proof-of-concept implementations were intended to demonstrate the feasibility of volumetric segmentation approaches in multi-terabyte sized volumes using the CATMAID framework.

1.6 CONTRIBUTIONS AND COLLABORATION

> "Right now, one of the most popular majors in colleges is computer science. I think that will be good for neurotechnology ... Students will be completely fluent in programming by the time they get to the lab, and I think they'll lead the synthesis between computer science and neuroscience that has to happen." — *Rafael Yuste at Kavli Foundation Roundtable Discussion, 2015.*

In this early phase of the formation of the field of large-scale EM connectomics, every single dataset is the product of the concerted effort of experts, resources and tools. I am extremely grateful to have been able to collaborate with so many excellent people for this thesis project. They contributed their time and expertise ranging from sample preparation, data acquisition and processing to neurobiology, software engineering, computer vision and machine learning to enable and facilitate my work.

I was extremely fortunate to start my PhD project with EM datasets that were already acquired with serial-section transmission electron microscopy at HHMI Janelia Research Campus. *Drosophila melanogaster* larva samples were prepared with high-quality preservation of tissue structure and good membrane contrast for reconstruction. Long series with thousands of ultrathin 40-50 nanometer sections were cut and collected manually with an ultramicrotome to cover extended volumes of the larval neuropil. This diligent work was carried out by Richard Fetter. Image acquisition was then performed over the course of many weeks and months using TEM and STEM at Janelia Research Campus. See Appendix A for details on the data generation and the people and labs involved.

At the same time, the challenge of assembling large amounts of raw 2D images from the microscopes into aligned 3D volumes was solved. Large-scale image alignment algorithms were developed which can montage a large numbers of small 2D image tiles into large 2D sections, on the order of $50'000^2$ pixels. These 2D mosaics then had to be further non-linearly aligned into 3D volumes to make them actually usable for circuit mapping (Saalfeld et al. 2012). These novel algorithms were implemented in the TrakEM2 software in Fiji and performed well after sophisticated parameter tuning (Cardona et al. 2012; Schindelin et al. 2012). I had the opportunity to learn how to use this toolsuite first-hand from their original developer, Stephan Saalfeld and Albert Cardona. I aligned two large EM datasets of *Drosophila* ventral nerve cord at first- and third-instar developmental stages myself (datasets L3 and L1c, see Appendix A). Two other first-instar larva datasets used in this thesis project were aligned by Stephan Saalfeld and Albert Cardona (datasets L1a and L1b).

First as a research assistant and then as a PhD student in Albert Cardona's group at the Institute of Neuroinformatics (INI), I started to work with Mark Longair on extending the CATMAID software with circuit reconstructions tools. The **C**ollaborative **A**nnotation **T**oolkit for **M**assive **A**mounts of **I**mage **D**ata project was initially started by Saalfeld et al. 2009. Casey Schneider-Mizell and Albert Cardona provided a never-ending stream of proposal for useful new features and improvements for the tracing tools. They kept Mark and me very busy and made an *extreme* programming approach a necessity. In the summer of 2012, I could supervise Oliver Uvman as a

Google Summer of Code student. He helped us complete the port of the CATMAID backend codebase from PHP to Python and Django. The International Neuroinformatics Coordinating Facility (INCF) was hosting the project. Meanwhile, the project was growing, and the list of GitHub issues with bugs and feature enhancements was steadily increasing. It was with Tom Kazimiers who joined the lab in Janelia in 2013 as a full-time software engineer, and Albert's investments in refactoring the codebase, that the project could further grow. We were very fortunate to also have Andrew Champion joining in 2014 as a second full-time developer on the project. It has been an extremely instructive for me to learn from and to work with this team of excellent engineers on the tools for neural circuit mapping and analysis, and see the project grow from a small lab effort to a widely used toolsuite[2].

The increased manpower behind the tool development allowed me then to focus on the neuroscience part of my PhD project - to actually apply this new toolkit to map and analyze neural circuits. Neural circuit mapping is laborious and requires many hours of sustained, concentrated attention. And as with many things in science, it often takes longer than expected. I was very fortunate to have Ingrid Andrade, and also Casey Schneider-Mizell and Waleed Osman, help me with the circuit mapping and reviewing work.

I was fortunate to work in the lab of Albert Cardona, and in institutions at both INI and Janelia, who recognize the importance of software engineering as a prerequisite to do novel science and their investment into open source development with tight development cycles with continuous user feedback. It is due to these great collaborators and the supportive environments that I was able to advance so rapidly with my thesis work.

CHAPTER 2 is based on the publication *Quantitative neuroanatomy for connectomics in Drosophila* (in revision) adapted for this thesis format. I share co-first-authorship with lead author Casey Schneider-Mizell. This work is further co-authored by Mark Longair, Tom Kazimiers, Andrew Champion, Frank Midgley, Stephan Saalfeld (software); Feng Li, Maarten F. Zwart (biology); Richard Fetter (EM) under the supervision of Albert Cardona. See also Contribution section at the end of the chapter, the commit log on GitHub[3] and Appendix B for details.

CHAPTER 3 is based in part on neurons published in Ohyama et al. 2015. The third-instar dataset was prepared and imaged by Richard Fetter and collaborators as part of the Janelia FlyEM project (see Appendix A). I was supported in circuit mapping by Ingrid Andrade, Waleed Osman and reviews by Casey Schneider-Mizell and Albert Cardona. The identification of larva interneuron was made possible by utilizing the light-level Neuron Catalog developed in the lab of James Truman at Janelia Research Campus.

CHAPTER 4 is based on my annotation of more than 3'500 microtubule termination points in 36 neurons in three datasets. Their fast annotation was made possible by utilizing the Dendrogram Widget developed by Tom Kazimiers.

---

2 `www.catmaid.org`
3 `github.com/catmaid`

CHAPTER 5 I generated a multi-label ground truth image stack of the L3 dataset (see Appendix A) which was published on the FigShare platform (Gerhard et al. 2013). I collaborated with deep learning experts Dan Ciresan and Alessandro Giusti in the group of Jürgen Schmidhuber at the Dalle Molle Institute for Artificial Intelligence in Lugano. We used the ground truth dataset to train state-of-the-art deep convolutional neural networks and generate multi-label probability maps of the whole L3 dataset. Our collaborator Jan Funke, a computer scientist in the group of Matthew Cook at the Institute of Neuroinformatics in Zurich, used the ground truth and probability maps very productively to advance automated reconstruction and error measures (Funke et al. 2014; Funke et al. 2015). I developed semi-automated neuron reconstruction interfaces and pipeline prototypes with support of Joachim Ott (user interface) who I supervised during the Janelia Undergraduate Program in 2012, and with help of Jan Funke at INI.

APPENDIX D The reimaging of putative septate junctions in the *Drosophila* CNS with multi-tilt TEM was done by Richard Fetter. The CATMAID coordinate to EM grid transformation was helped by Stephan Saalfeld.

2

QUANTITATIVE NEUROANATOMY FOR CONNECTOMICS IN DROSOPHILA

ABSTRACT

Large-scale neuronal circuit mapping using electron microscopy demands laborious proofreading by humans who resolve local ambiguities with larger contextual cues or by reconciling multiple independent reconstructions. We developed a new method that empowers expert neuroanatomists to apply quantitative arbor and network context to proofread and reconstruct neurons and circuits. We implemented our method in the web application CATMAID, supporting a group of collaborators to concurrently reconstruct neurons in the same circuit. We measured the neuroanatomical underpinnings of circuit connectivity in *Drosophila* neurons. We found that across life stages and cell types, synaptic inputs were preferentially located on spine-like microtubule-free branches, "twigs", while synaptic outputs were typically on microtubule-containing "backbone". The differential size and tortuosity of small twigs and rigid backbones was reflected in reconstruction errors, with nearly all errors being omission or truncation of twigs. The combination of redundant twig connectivity and low backbone error rates allows robust mapping of *Drosophila* circuits without time-consuming independent reconstructions. As a demonstration, we mapped a large sensorimotor circuit in the larva. We found anatomical pathways for proprioceptive feedback into motor circuits and applied novel methods of representing neuroanatomical compartments to describe their detailed structure. Our work suggests avenues for incorporating neuroanatomy into machine-learning approaches to connectomics and reveals the largely unknown circuitry of larval locomotion.

2.1 INTRODUCTION

Mapping circuits from electron microscopy (EM) volumes is hard (Helmstaedter 2013). Manually working through large volumes is slow and prone to attentional errors (Kreshuk et al. 2011; Helmstaedter et al. 2011). Combining multiple independent reconstructions of the same neuron can reduce errors (Helmstaedter et al. 2011; Kim et al. 2014) at the cost of multiplying the required labor. Current computational approaches operate only with "local" information, i.e. the EM micrographs and algorithmically detected fine structures such as cell membranes and mitochondria. They are therefore sensitive to noise (Jain et al. 2010), particularly in anisotropic EM data where the smallest neurites may be thinner than the thickness of individual serial sections (Veeraraghavan et al. 2010; Helmstaedter 2013). Machine-generated neuron reconstructions are therefore proof-read by humans (Chklovskii et al. 2010; Haehn et al. 2014).

Experts are able to resolve ambiguities that amateurs and current algorithmic approaches cannot by using large-scale features about neurons to inform decisions made at the level of nanometer-scale image data. In *Drosophila*, where neurons are highly stereotyped, large branches in an EM reconstruction of a given cell can be confirmed by comparing the observed anatomy to that of homologous cells from light microscopy data or other reconstructions (Takemura et al. 2013; Ohyama et al. 2015). This suggests that one way to improve the toolkit for neuron reconstruction and circuit mapping is to facilitate the application of cell- and circuit-level features for disambiguation at noisy locations on EM volumes.

Crucially, different errors do not alter the wiring diagram equally. Missing small dendrites can be acceptable; useful and reproducible wiring diagrams can be created even when omiting 56% of all postsynaptic sites (Takemura et al. 2013) while missing a single large branch hosting all the synapses in one neuropil region could omit connectivity to entire populations of partners. Prioritizing proofreading time towards the most impactful errors improves reconstruction efficiency (Plaza et al. 2012; Kim et al. 2014).

However, to understand the impact of reconstruction errors in *Drosophila*, we need to understand the relationship between circuitry and anatomy. Mesoscale anatomy, in particular the placement of large branches, is well understood as a key component of circuit structure (Zlatić et al. 2003; Zlatić et al. 2009; Wu et al. 2011; Couton et al. 2015). Similarly, the connectivity graph of a stereotyped circuit can itself relate back to anatomy by consideration of the location of the synaptic sites between pairs of neurons. However, relatively little is known about the smallest scales of synaptic connectivity, the distribution of individual synapses on a neuron. Microtubule-free and actin-rich structures have been identified as sites of excitatory input in the adult *Drosophila* visual system (Scott et al. 2003; Leiss et al. 2009a), but questions remain about how ubiquitous these are in the nervous system.

Here, we describe a collection of quantitative anatomical and connectivity features across scales, from fine dendritic branches to multi-neuron graphs, and a method for using them to swiftly and accurately map the *Drosophila* nervous system. We implemented the calculation and visualization of such features on-demand as an extension

of the web-based large image data viewer CATMAID (Saalfeld et al. 2009). Because the detection of high-impact errors can occur concurrently with reconstruction via interactive analysis, our tool removes the need for time-consuming repeated reconstructions (Helmstaedter et al. 2013; Kim et al. 2014) and enables synergy among collaborating expert neuroanatomists. We demonstrate our methods by mapping a sensorimotor circuit in the *Drosophila* larva from proprioceptive sensory neurons to motoneurons.

2.2 RESULTS

Collaborative circuit mapping

We extended the web-based image data viewer CATMAID (Saalfeld et al. 2009) to enable a geographically distributed group of researchers to map neuronal circuitry. A neuron is reconstructed with a skeleton, a directed tree graph with one or more nodes in every cross-section of neurite in an EM volume (Helmstaedter et al. 2011; Cardona et al. 2012). Where possible, we root skeletons at the soma to model the anatomical notions of proximal and distal in the data structure. Synapses (see Supplemental Fig. 2.7) are annotated as a relation from a node on the presynaptic neuron skeleton to an intermediate "connector node" and then to a node of a postsynaptic neuron skeleton. To express the polyadic nature of insect synapses (Meinertzhagen et al. 1991), connector nodes can have multiple postsynaptic "targets", but only one presynaptic "source". Reconstructions are immediately synchronized across all users to avoid duplicate or conflicting work, and to take advantage of existing reconstructions to aid further reconstruction and circuit discovery. For further details, see Supplemental Text.

As a case study of our method, we focused on sensorimotor circuits in an abdominal segment of the first instar *Drosophila* larval central nervous system (CNS) using an EM volume covering one and a half abdominal segments (Supplemental Fig. 2.8; previously used in Ohyama et al. 2015). In total for this work, we reconstructed 425 neuronal arbors spanning 51.8 millimeters of cable, with 24,068 presynaptic and 50,927 postsynaptic relations, (see Supplemental Text for details). Reconstruction time was 469 hours for reconstruction with synapse annotations plus 240 hours for review (see below), for an average rate of $\sim$73 microns of proofread arbor with synapses per hour.

Microtubule-free twigs are the principal site of synaptic input

To be able to use neuronal anatomy to guide circuit reconstruction, it was crucial to better understand the distribution of synaptic input onto *Drosophila* neurons. We started by looking in detail at the relationship between the microtubule cytoskeleton (Supplemental Fig. 2.8) and synaptic inputs in EM reconstructions of neurons from different regions of the nervous system and life stages. For a diverse collection of neurons, we marked all locations where the arbor continued distal to a microtubule-containing process (Fig. 2.1a). We call such a terminal branch a "twig". By definition, all twigs have their base on a microtubule-containing backbone shaft. Following the classification in Leiss *et al.* (Leiss et al. 2009a), a spine is a twig with a maximal depth of less than 3 μm and that is not a presynaptic varicosity (Fig. 2.1b).

We found twigs in all neurons investigated, across multiple CNS regions and life stages of *Drosophila*, and in all cases they were the dominant sites of synaptic input (Fig. 2.1c–g). We first considered motoneurons aCC and RP2 (Landgraf et al. 1997), which have functional and structural similarities to vertebrate neurons (Sánchez-Soriano et al. 2005; Nicolai et al. 2010; Günay et al. 2015). In the first instar CNS, we find aCC and RP2 have numerous twigs, which together host more than 80% of their total number of postsynaptic sites (Fig. 2.1c). We found a similar majority of inputs onto twigs in three hemisegmental pairs of premotor interneurons (Fig. 2.1d) and brain neurons (Ohyama et al. 2015) in the first instar (Fig. 2.1e). We tested whether the observed distribution of postsynaptic sites onto twigs generalizes across larval stages by comparing a somatosensory interneuron in the first instar to its homologue in late third instar (Fig. 2.1f). At both life stages, we find more than 80% of inputs were onto twigs, suggesting that twigs are not a temporary developmental structure. In the adult fly, light microscopy-level analysis of lobula plate tangential cells of the visual system suggests a similar distribution of postsynaptic sites onto twigs (Leiss et al. 2009a; Scott et al. 2003). We annotated EM skeletonizations of medullar Tm3 neurons reconstructed by Takemura et al. 2013 in the adult visual system neuropil and found that nearly all their inputs were onto twigs (Fig. 2.1g). Our findings suggest that microtubule-free twigs are a general feature of *Drosophila* neurons and constitute the primary anatomical location of synaptic input. Spine-like twigs are found in all cell types, but host a variable, typically non-majority, amount of synaptic input (Fig. 2.1c–g). We consider all twigs for the remainder of our analysis.
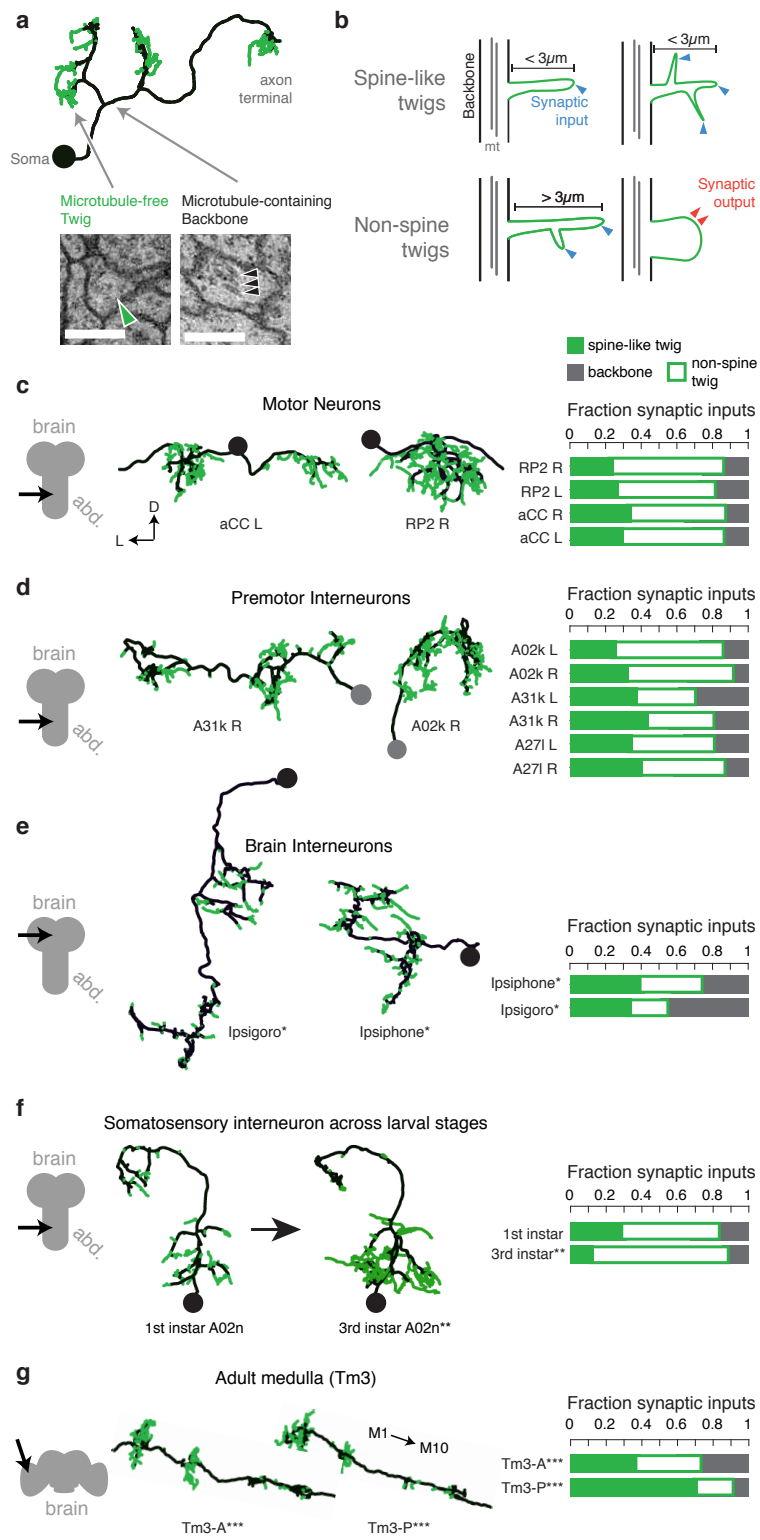
Figure 2.1.: **Twigs, small microtubule-free neurites, are the primary site of input in *Drosophila* neurons.**

Twigs, small microtubule-free neurites, are the primary site of input in *Drosophila* **neurons.** *(continued)* **a)** A typical Drosophila neuron has a contiguous microtubule backbone from the soma to its arbors. Numerous terminal arbors without microtubules ("twigs") branch off the backbone. The presence (black arrows) or absence (green arrows) of microtubules can be seen in EM cross-sections of the neuron. **b)** Twigs less than 3 µm are considered spine-like, while those longer or primarily presynaptic are not. **(c–g)** EM reconstructions (middle) of *Drosophila* neurons from different parts of the nervous system (left) showing backbone (black) and twigs (green). At right, the fraction of all synaptic inputs onto short spine-like twigs, longer twigs, and backbone. Data sets are indicated by marks: no asterisk: 1.5 segment volume. *: Whole CNS volume. **: 3rd instar abdominal segment volume. ***: Adult medulla skeletons and images, generously provided by Janelia FlyEM [9]. Neurons are individually scaled to show complete arbors. **(c)** Motoneurons in 1st instar larva. **(d)** Premotor interneurons of 1st instar larva. **(e)** Interneurons in the brain of the 1st instar larva. **(f)** A somatosensory interneuron cell type across life stages, 1st instar and 3rd instar larvae. **(g)** Tm3 cells in the adult fly medulla.

Distribution of inputs onto motoneuron dendrites

For a given presynaptic partner, a postsynaptic neuron could concentrate its input synapses onto a single region or distribute them widely. The spatial distribution of synaptic inputs has implications for dendritic processing (Polsky et al. 2004), developmental robustness (Couton et al. 2015), and as we show, reconstruction accuracy.

To study the relationship between presynaptic neurons and the anatomical locations of postsynaptic sites, we reconstructed all neurons synaptically connected to motoneurons aCC and RP2 in the third abdominal segment of a first instar larva (Fig. 2.2a-d). A dynamically-generated and interactive table of synaptic connectivity in CATMAID enabled users to systematically trace all connected arbors. We found 198 identifiable neurons (Supplemental Fig. 2.10) and named them according to a developmental lineage-based nomenclature (Ohyama et al. 2015) and classified 107 other arbors spanning the full segment into eight distinct intersegmental bundles (Supplemental Fig. 2.11). Motoneurons each received between 1 and 28 synaptic inputs from individual presynaptic neurons, with a maximum of 7.3% of all inputs coming from a single neuron (Fig. 2.2e). The fraction of synapses accounted for by their presynaptic partners (rank-ordered by number of synapses) is well-fit by an exponential survival function, with a decay indicating that approximately the top 22 presynaptic partners of one motor neuron contribute 63% of all its synaptic inputs (Fig. 2.2f).

We next asked how the synaptic input onto aCC and RP2 is distributed across independent twigs. Most individual twigs were small, with the median twig measuring 1 µm in cable and hosting 1 postsynaptic site. The largest typical twig had 16 µm of cable and 20 postsynaptic sites (Fig. 2.2g, but see Supplemental Text for discussion of three larger outliers). We find that presynaptic neurons connect to between 0 (backbone only) and 13 twigs, with nearly all connections with 3 or more synapses per edge being distributed across multiple twigs (Fig. 2.2h). Similarly, numerically strong edges spanned multiple twigs in the adult visual system Tm3 neurons (Supplemental Fig. 2.12).

Mitochondria and presynaptic sites

We next looked for ultrastructural features related to presynaptic sites in six premotor interneurons, since the motoneurons had no central synaptic output. We found that, in stark contrast to inputs, presynaptic sites were often associated with microtubules (Supplemental Fig. 2.13). Approximately 50% of presynaptic sites are located on the backbone and 90% were within 3 µm of it. Since presynaptic sites are metabolically demanding (Laughlin et al. 1998), we also annotated the location of all mitochondria in the neurons (Supplemental Fig. 2.13). We found that all presynaptic sites were within 3 µm of their closest mitochondrion, though only 47% of neuronal cable was within 3 µm of a mitochondrion. Taken together, this suggests that the internal structure of neurons puts strong limits on the location of presynaptic sites.
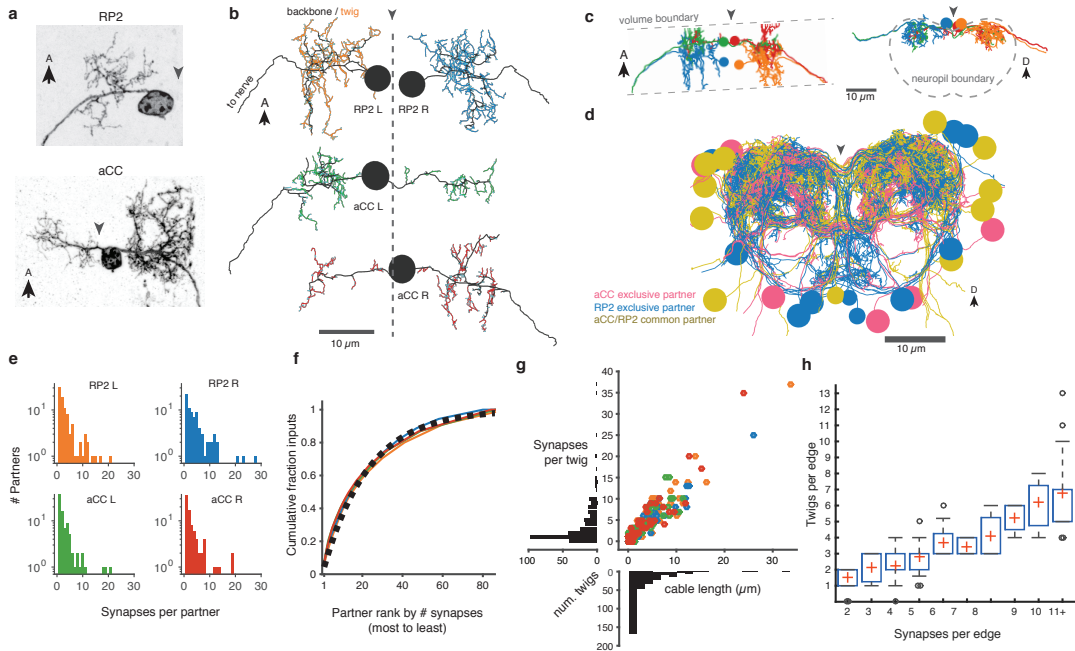
Figure 2.2.: **Twigs are crucial to larval motor circuitry**. **(a)** Dorsal projections of genetically labeled motoneurons RP2 (top, from 1st instar) and aCC (bottom, from 3rd instar). Each cell type has characteristic dendritic arbors. Midline indicated by gray arrowhead. **(b)** EM reconstructions of each of four motoneurons aCC and RP2 in the 1st instar larva match the left and right homologs of aCC and RP2. Backbone is indicated by black, twigs by colors. Midline is shown as dashed line.**(c)** True spatial relationship of the four motoneurons in b), shown dorsally (left) and in cross-section (right). Note that the boundary of the EM volume is limited. **(d)** All arbors presynaptic to aCC and RP2. Colors indicate if neuron is presynaptic to one or both motoneuron cell types. **e)** Histograms of premotor partners connected via number of synapses. **(f)** Colored lines: the cumulative fraction of total inputs as a function of ranked presynaptic partners for each motoneuron are extremely similar. Black dashed line: simultaneous fit for all four motoneurons to $1 - \exp\left(-r/\rho\right)$ for rank r gives $\rho = 22.34$. **(g)** Scatterplot and histogram of the total length and number of synapses on each of the 305 twigs for each of the four motoneurons (colors as previous). **h)** Number of twigs contacted by motoneuron partners as a function of the number of synapses in the connection. Crosses are median, boxes the interquartile range, whiskers the 10th to 90th percentiles. Outliers shown.

Circuitry for proprioceptive feedback into a motor circuit

We next looked at the cell and circuit level for regularities that could inform proof-reading. In the *Drosophila* larva, developmentally homologous neurons are strongly stereotyped (Li et al. 2014), making quantitative analysis of their properties useful for identifying irregularities between homologous cells. Most cell types are represented in the fly nervous system by at least one homologous bilateral pair of individual cells. Bilateral homology suggests both arbor morphology and synaptic wiring are mirrored, up to developmental noise (Ohyama et al. 2015). To let morphology guide proofreading, we developed a collection of neuroanatomical measurements that were independent of absolute location. These metrics, combined with 3d visualization, quickly summarize the structure of complex neurons to help identify and localize inconsistencies (Supplemental Fig. 2.14) which could result from true developmental differences (Supplemental Fig. 2.15) or reconstruction errors.

As a case study, we applied our tools to describe a complete sensorimotor circuit. During forward crawling, a peristaltic wave of muscle contraction travels from posterior to anterior segments (Hughes et al. 2007; Heckscher et al. 2012). Signals from the segmentally repeated proprioceptive neurons dbd have been suggested to act via a polysynaptic inhibitory pathway to stop motoneuron activity after successful contraction of a given segment (Hughes et al. 2007). To find pathways between proprioceptive and motor neurons, we further reconstructed axons for proprioceptive sensory neurons dbd, vbd, dmd1, ddaD, and ddaD (Hughes et al. 2007; Grueber et al. 2007). Because of its implication in proprioceptive feedback (Hughes et al. 2007), we further reconstructed all partners of the left and right dbd (Fig. 2.3a).

Using a graph search within CATMAID, we identified all 1–3 hop pathways from dbd to motoneuron RP2. Comparison of the identifiable intermediate neurons revealed five pairs of homologous neurons with consistent shape, connectivity, and basic morphological quantities (Fig. 2.3b,c). Inconsistencies in any property led to further review to determine if they were due to reconstruction error, true developmental variability, or limitations of the raw data. For example, one strong inconsistency in this network, the connection from A02l to A31k (Fig. 2.3b), was due to the expected synapse locations being outside the imaged volume on one side but not the other (Supplemental Fig. 2.8).

The five pairs of identified neurons could also be matched to light-level images of neurons identified through sparse stochastic labeling (Nern et al. 2015) of neurons within a GAL4 expression pattern (Fig. 2.3c). Of these, two directly premotor interneurons (A27j and A31k) are immunoreactive to anti-GABA (Supplemental Fig. 2.16), whereas the others, all from A02 lineage, are members of the glutamatergic neuron class described in Kohsaka et al. 2014. These novel, putatively inhibitory sensorimotor pathways are well-positioned to mediate a hypothesized "mission accomplished" signal (Hughes et al. 2007). This map also could explain why genetic silencing of A02 neurons was shown to slow peristalsis (Kohsaka et al. 2014), as doing so removes a major channel for proprioceptive feedback.
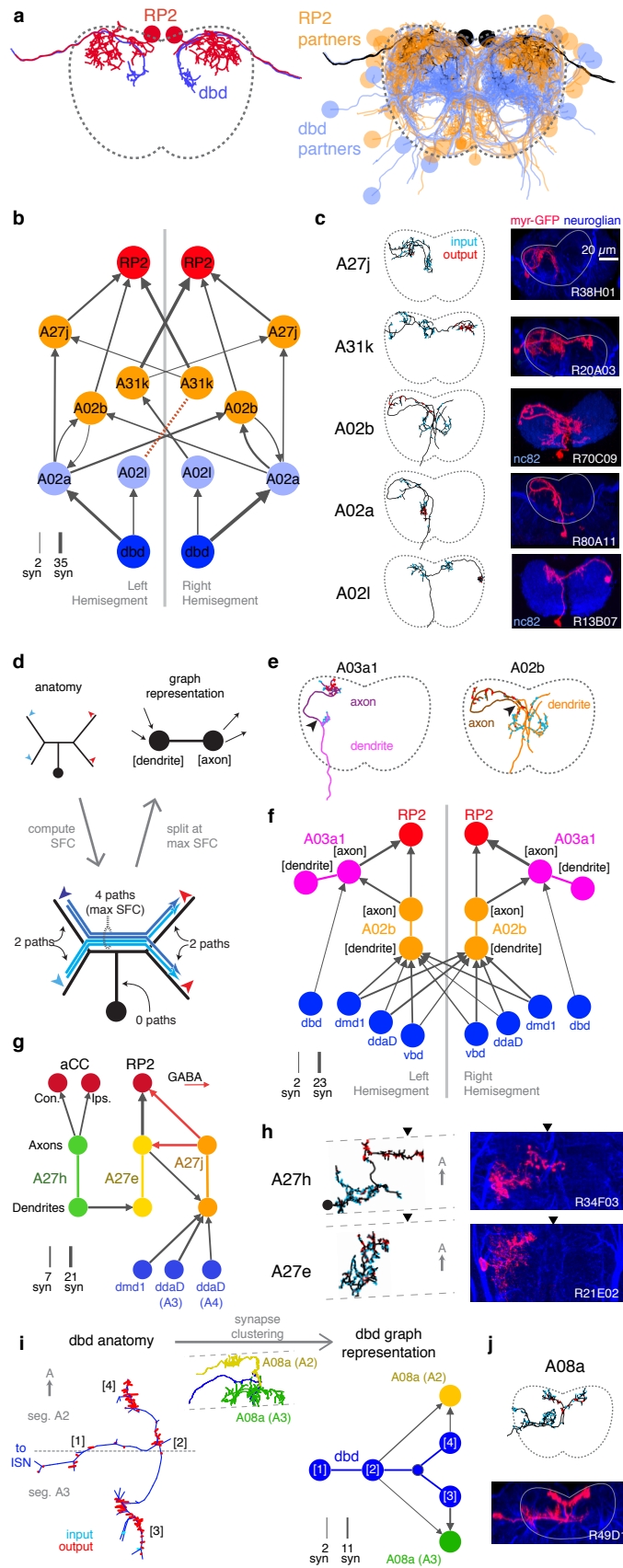
Figure 2.3.: **Wiring diagrams for larval motor circuitry enriched with anatomical detail.**

Wiring diagrams for larval motor circuitry enriched with anatomical detail. *(continued)* **a)** Starting with motoneuron RP2 and proprioceptive sensory neuron dbd (left), we identified all synaptic partners of each (right). **b)** Five symmetric pairs of identified neurons link the two cell types with three or fewer hops. All edges are observed in both the left and right hemisegments, except for a single edge outisde the volume boundary (red dashed line, see Supplemental Fig. 2.8). Line thickness increases with number of synapses (max. and min. values shown). In this and all network diagrams, single synapse edges are not shown for clarity. **c)** All identified cells in EM (left) could be matched to confocal maximum intensity projections of single neurons found in sparsely labeled GAL4 lines (right, see methods for details). For neuroglian staining, an approximate neuropile boundary is shown; for nc82 staining, the blue region is a profile of neuropile. **d)** Cartoon example of splitting neurons using synapse flow centrality (SFC). **e)** Examples of two premotor interneurons split into axonal (darker) and dendritic (lighter) regions with this method. Split point is indicated by the arrowhead. **f)** Splitting interneurons into axonal and dendritic compartments in a proprio-motor circuit reveals stereotypic pre- and post-synaptic connectivity to premotor interneuron A03a1 and differential contributions from proprioceptor dbd relative to other proprioceptors dmd1, ddaD, and vbd. **g)** Splitting interneurons A27j, A27e, and A27h reveals GABAergic pre- and post-synaptic input to a premotor connection, as well as dendro-dendritic coupling between interneurons that connect to synergistic motoneurons aCC and RP2. **h)** Dorsal projections of A27h and A27e from EM (left) and light (right), as in **c**. Midline indicated by arrowheads. **i)** Synapse clustering can represent dense groups of synapses on an arbor as distinct nodes in a graph widget. Here, each of four groups of synapses (numbers in brackets) is represented as a separate node in the graph. The graph demonstrates that posterior and anterior clusters (3 and 4, respectively) target homologous neurons in consecutive segments. **j)** Cross-sectional projection of A08a from EM (top) and light (bottom), as in **c**.

Anatomically enriched wiring diagrams reveal propriomotor circuit motifs

The physiology of synaptic input and output can differ across neuronal compartments. For example, presynaptic inhibition is important for gain control in fly sensory circuits in a fundamentally distinct manner than dendritic inhibition (Clarac et al. 1996). This suggests that connectivity can be stereotyped at the compartmental level and therefore useful for proofreading. We thus sought a graph representation of a circuit that could faithfully distinguish distinct types of connections (Fig. 2.3d-g).

In *Drosophila*, many neuronal cell types have distinct input and output compartments, while a few have entirely intermingled inputs and outputs. Our approach assumes that the neuron can be split into distinct compartments, and at the end checks to see if the split was successful. First we calculate all paths along the skeleton from each of the neuron's input synapses to each of its output synapses and for each node of the skeleton compute the number of centripetal (towards soma) and centrifugal (away from soma) paths that pass through it (Fig. 2.3d). This quantity, which we call "synapse flow centrality" (SFC), is analogous to a synapse-specific version of betweenness centrality (Newman 2010). For most neuronal arbors, we find that the most proximal skeleton node with the highest centrifugal SFC corresponds to an intuitive generalization of the locations of spike initiation zones in known polarized neurons of *Drosophila* (Gouwens et al. 2009; Günay et al. 2015) and other insects (Gabbiani et al. 2002).

We quantify how completely input and output are separated on a neuron with a "segregation index", an entropic measure of the amount of input/output mixing in each compartment, normalized by that of the whole arbor (see Methods and Supplemental Fig. 2.17). A very low segregation index means that pre- and post-synaptic sites are intermingled and an axon/dendrite compartmentalization is inappropriate. Using this approach, we classified all identifiable neurons found in both the left and right hemisegments of the proprio-motor circuitry described above. Of the 3834 synapses between these cells, we found 79% were axo-dendritic (3033), 11% axo-axonic (424), 9% dendro-dendritic (334) and 1% dendro-axonic (43).

We consider two examples of how compartment-enriched graphs add important anatomical detail to small microcircuits. First, we analyzed how different proprioceptive inputs converge onto motoneuron RP2 (Fig. 2.3e–h). By splitting interneuron A02b into axon and dendrite we observed that its dendrites receive bilateral proprioceptive input, while its axon synapses both onto the ipsilateral RP2 and axoaxonically onto its strong premotor partner, A03a1 in both hemisegments (Fig. 2.3f). In contrast, while dbd only connects indirectly with A02b (Fig. 2.3b), it synapses exclusively ipsilaterally and axo-axonicaly onto A03a1 (Fig. 2.3f). This suggests that the role of dbd in modulating motor patterns could be qualitatively different than the other proprioceptive sensory neurons, since its direct pathways are typically longer or involve connections types other than axo-dendritic.

Second, we analyzed interactions between the premotor neurons of aCC and RP2 (Fig. 2.3g,h) . We found that a neuron presynaptic to the aCC motoneuron on both sides targets dendro-dendritically a pre-RP2 neuron (A27h), potentially coordinating the joint excitation of their targets (Fig. 2.3g). We also found a premotor interneuron

(A27e) with reciprocal connections with a GABAergic premotor interneuron (A27j; see Supplemental Fig. 2.16) that receives convergent inputs from dorsal proprioceptive neurons (dmd1, ddaD, ddaE; Fig. 2.3g,h). This suggests that A27j might not only act as an inhibitory premotor input in response to proprioceptive activity, but also have subtler modulatory effects onto other sources of motor input.

Specific connections can also be allocated to specific regions of a neuronal arbor, which could be used to localize proofreading guided by inconsistencies in connectivity. We thus extended the concept of splitting a neuron into two regions to an arbitrary number, by defining a region as a cluster of synapses near each other along the arbor cable (see Methods). As an example, we consider the axon terminal of dbd, which enters at the interface between two segments and extends symmetric arbors towards the anterior and posterior segments (Fig. 2.3i). The synapses form multiple well-separated clusters that we can visualize as a group of nodes, revealing that the anterior and posterior branches synapse onto homologous interneurons (A08a) for their respective segments (Fig. 2.3j). This pattern suggests that each A08a cell gets convergent input from the dbd of two consecutive segments, which could reflect that adjacent pairs of segments move together during locomotion (Heckscher et al. 2012).

Proofreading and error correction

Based on the spectrum of features described above, we developed a two step iterative method of proofreading after an initial reconstruction. An initial systematic review consists in traversing a whole arbor from distal to proximal to follow the expected gradual changes in anatomical properties (e.g. caliber tapering and cytoskeletal changes from microtubule-free to increasing number of microtubules). This review mode helps discover attentional errors or anatomical inconsistencies, such as non-contiguous microtubule cytoskeleton. Next, we reconstruct a second instance of homologous circuitry and use the high-level quantitative anatomical and connectivity measurements to highlight inconsistencies, which are then subjected to focused review (see Supplemental Text for details). This review mode helps ensure that the broad structure of the neuron is consistent and no high-impact large branches are missed.

Importantly, any error correction is performed on the basis of local information, as contained within the EM images (e.g. microtubules, texture, consistency with neighboring neurites, etc. See Supplemental Text). Irregularities in higher level features only serve to guide attention, not determine correctness. The strong stereotypy of *Drosophila* neurons (Cardona et al. 2010; Jenett et al. 2012; Li et al. 2014) enables the use of light microscopy images or contralateral homologs in the EM for overcoming ambiguities in large, microtubule-rich backbone branches, but not for more variable twigs. To do this requires partial annotation of all feasible continuations within the EM volume to ensure that there is only single solution consistent with the stereotypic features (see Supplemental Text). We also annotate ambiguities during reconstruction with a scale from 1 to 5 for further consideration in subsequent quantitative or visual analysis of the arbors and the wiring diagram.

Validation of our iterative circuit mapping method

Our approach to circuit mapping consists of an initial reconstruction, followed by edits by the same or different users during proofreading or incidental discovery of errors during subsequent work. Small arbor pieces, left over from pruning when proofreading other neurons or from explorative tracings, are merged in. We refer to this as "iterative," as compared to consensus methods that combine multiple independent reconstructions (Helmstaedter et al. 2011; Takemura et al. 2013; Kim et al. 2014).

We evaluated the accuracy of our method for *Drosophila* circuits by comparing our results to the those of a consensus approach. We selected six interconnected neurons from the premotor network for independent reconstruction by four individuals. Each individual skeletonized and reviewed his or her reconstructions and then consensus skeletons were computed for each neuron using RESCOP (Helmstaedter et al. 2011). Both methods resulted in extremely similar arbors, although each method found branches not seen in the other (Fig. 2.4a, Supplemental Fig. 2.18). All sites of disagreement between the two methods were validated by an expert to determine a gold-standard morphology. Reconstruction and review of these six neurons in the iterative approach took a total of 26.37 hours, while the redundant method by four people took a total of 107.73 hours, almost exactly four times as long.

Existing consensus approaches only calculate neuronal morphology, not synaptic connectivity. We estimated the connectivity between consensus skeletons by transferring each synapse from each independent skeleton to the consensus, normalized by number of skeletons contributing to the consensus at the location of the synapse. A synapse would have unit weight when annotated at the same location in all independent skeletons.

We found that both methods recover an identical set of edges in the wiring diagram, with similar number of synapses per edge (Fig. 2.4b,c). We next considered the fine differences between consensus skeletons and skeletons reconstructed with our method. The six gold-standard neurons had a total of 1341 input synapses, with 111 on neurites only present in the consensus skeletons, 229 on neurites only in our method's reconstructions, and 1001 in the arbor found by both. We located 91 missed or incomplete branches (false negatives) in our method, 89 in twigs and 2 in backbones; and 7 incorrect continuations (false positives), 6 in twigs and 1 in backbone. False positives added 30 incorrect postsynaptic inputs in total. Individual missed branches were small in size, complexity, and number of synapses (Fig. 2.4e–g), with more than 40 missed or truncated twigs having no impact on connectivity (Fig. 2.4f). The 3 errors in backbones occurred in small distal dendritic shafts containing one single microtubule, resulting in 7 missed and 4 false postsynaptic sites. Error rates for synaptic output were even lower. The gold-standard neurons had a total of 510 presynaptic sites, of which 509 were found by our iterative reconstructions.

Our data suggest that anatomical structure strongly influences the rate of reconstruction errors in our iterative method. Our total error rate is dominated by false negatives and is much higher for twigs ($16.2\,\mu m$/error) than for backbone ($375.8\,\mu m$/error). While attentional errors seemed to dominate missed branches,

data ambiguities were often associated with backbone errors. One backbone false merge happened across two adjacent sections in poor registration with one another, while an erroneous truncation occurred across a section where electron-dense precipitate occluded the neurite and its surrounding area.

Estimating errors in a reconstructed wiring diagram

Neuroanatomy strongly constrains the impact of a typical error on the wiring diagram because, as shown above, the most likely error is to miss a twig and an individual twig hosts few or no synapses.

To estimate the probability of omitting a true edge in the wiring diagram, we analyzed the distribution of synaptic contacts across twigs as a function of the total number of synapses per edge. Edges comprising multiple synaptic contacts were found to be distributed across multiple twigs (Fig. 2.2h). With the RESCOP-based validation we found that our method identified 88% (672/761) of twigs, containing 91.7% of synapses (1230/1341). From these two observations, we estimated the probability of completely missing a true edge as a function of the number of morphological synapses per edge. We found that our method recovers more than 99% of the wiring diagram edges that have at least 3 synapses (Fig. 2.5a), assuming twigs are missed uniformly at random (see Supplemental Fig 2.19).

In the larva, we are primarily interested in the most reliable edges between cell types, as those are most likely to generalize across individual animals. Moreover, we are concerned less about adding extra synapses to true connections and more about adding false edges that would be interpreted as pathways that are not actually present. To estimate the likelihood of introducing a false edge between cell types not just once, but twice (e.g. in a left and right pair of homologs), we simulated adding false twigs to a neuron. The probability of adding a false edge depends both on the probability of adding a false twig (observed false positive error rate: 7 errors in 605 twigs) and the number of nearby but unconnected neurons with presynaptic sites. This will vary depending on the circuit in question. For example, a neuropile with all-to-all connectivity will have no opportunity for false positive edges, while in an array of rigorously separated labeled lines any false positive synapse would be a false positive edge. Further, larger neurons offer more opportunities for false positives than smaller neurons.

For a concrete and realistic example, we consider the motoneuron RP2 (a large neuron). We estimated the number of proximate but unconnected neurons by considering all axons presynaptic to all motoneuron dendritic fields that overlap RP2's dendrites (Fig. 2.5b). We assume that a false positive reconstruction error distributes m synapses across all available axons at random. Even if we assume that m is always among the largest observed ($m = 20$), our model suggests that for the RP2 wiring diagram we can trust symmetric connections of at least 2 synapses (Fig. 2.5c). We further note that the small size of individual twigs and the ability in CATMAID to jump directly to the image data associated with synapses comprising an edge make review of a suspect false positive edge extremely fast, on the order of seconds.
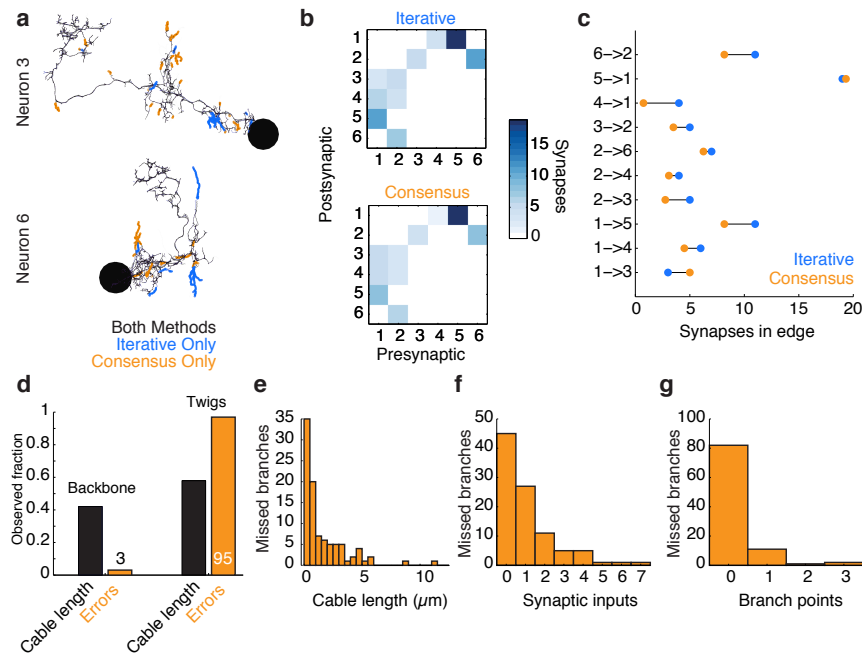
Figure 2.4.: **Our method for iterative reconstruction produces reliable connectivity and anatomy. (a)** Dorsal view of two of six neurons for which we compared our iterative reconstruction method to a RESCOP-generated consensus of four independent reconstructions. Arbor found in both, dark lines; iterative only, blue; consensus only, orange. **(b)** The adjacency matrix produced by our iterative method has an identical set of edges as that of the consensus method, with variability only in the amount of synapses per edge. **(c)** The weights of each edge (the amount of synapses) are similar between methods. **(d)** Point errors in iterative reconstructions are not distributed equally across the cable of neuronal arbors, instead falling overwhelmingly on twigs. **(e-g)** Branches missed by our iterative method but observed in the consensus method are typically very small and lightly connected as seen from histograms of their **(e)** cable length, **(f)** synaptic inputs, and **(g)** number of branch points.
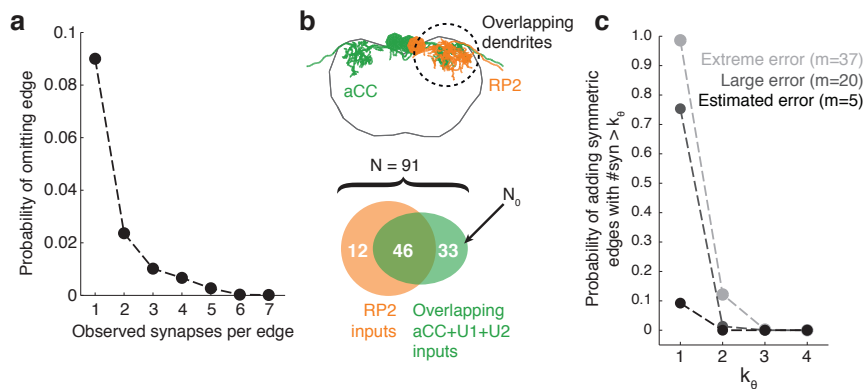
Figure 2.5.: **Estimating errors that affect graph topology. a)** Estimated probability of fully omitting an edge as a function of how many synapses were on the edge based on omitting random twigs with the frequency observed in the validation data. **b)** Cartoon of dendritic overlap between RP2 and aCC, U1, and U2. On average, 91 axons put at least 2 synapses on any motoneuron (denoted N in the false positive estimate model, see text for details), of which 33 are not connected to RP2 (denoted N_0). **c)** Probability that, given a pair of homologous postsynaptic neurons, introducing m false inputs randomly distributed across N presynaptic neurons yields at least one pair of false edges of k_θ or more synapses each. The number of axons were estimated in **b**, and false input counts are shown estimated from the validation data ($m = 5$), as well as if they came from adding a rare but large twig ($m = 20$), and the largest observed twig ($m = 37$).
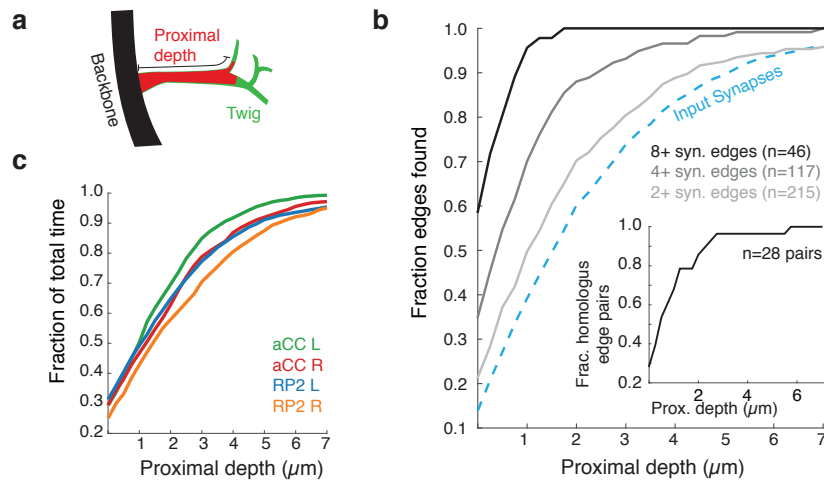
Figure 2.6.: **Proximal regions of twigs reflect final wiring a)** Cartoon of the proximal depth (red) into of a twig (green) measured from from the backbone (black). **b)** The fraction of >2 synapse edges onto aCC and RP2 that would be found when considering only synapses onto the backbone and twigs cropped at a maximum depth. From light to dark gray are those edges whose final measured connectivity has more than 2, 4, and 8 synapses. Blue dashed line indicates fraction of all input synapses. The inset depicts the fraction of pairs of homologous edges from identified neurons (N=28 edge pairs) that would be identified using synapses up to a given depth. **c)** Fraction of total reconstruction time for each of the four motoneurons (see legend) as a function of cropping twigs at a maximum depth. Note that 0 μm depth cropping corresponds to backbone reconstruction only.

Since most errors were of omission and took the form of cropped twigs, we also measured the impact of omitting the distal ends of twigs. Considering again aCC and RP2, we looked at the connectivity observed by considering only synapses located at a given depth into the twig relative to its base on the backbone (Fig. 2.6a). With all twigs cropped to zero depth, only inputs onto the backbone remain. More than 90% of postsynaptic sites lay within 5 μm of the backbone (Fig. 2.6b). We find that the first 2 μm already finds at least two synapses between ∼90% of the most connected partners. The first 4 μm similarly detects ∼90% of all partners with 2 or more synapses and 27/28 pairs of homologous edges (Fig. 2.6b). These results indicate that, given the observed distribution of synapses over multiple twigs (Fig. 2.2h), edges with many synapses are robust to errors of omission such as truncated twigs. Considering the marginal time involved in reconstructing the full extent of twigs (Fig. 2.6c), this robustness could be intentionally exploited towards discovering strong synaptic partners in a time-efficient manner.

2.3 DISCUSSION

Neuroanatomy as the foundation for circuit mapping

Neurons are highly structured cells. A human expert's success at circuit mapping from EM volumes stems from the ability to use this structure and apply cell and circuit-level context to interpret nanometer-scale image data. Here, we presented our approach to circuit mapping in EM by building tools in CATMAID that ease and emphasize the use of high level features concurrent with image-level reconstruction. In addition to using existing measures of neuroanatomy and graph connectivity, we devised novel algorithms to quantify the distribution of synapses across neurons, which we applied to both quantitative neuroanatomy and to enrich wiring diagrams.

Central to our approach is the observation that *Drosophila* neurons contain a contiguous microtubule-rich backbone and numerous small microtubule-free distal twigs. We found that small twigs are the primary site of synaptic input for *Drosophila* neurons and that numerically strong connections between neurons are typically spread across many distinct twigs. If, contrary to observations, neurons were to only contact each other via a single twig that hosts many postsynaptic sites, then this connection would be fragile with respect to developmental noise (Couton et al. 2015). In contrast, backbones define the spatial extent and stereotyped shape of a neuron, and we found that most presynaptic sites are located on or very near the backbone's microtubules and mitochondria. Our findings are consistent with the notion that metabolic needs and microtubule-based trafficking are limiting factors for the structure of synaptic output.

These different biological requirements for different neuronal compartments are reflected in the rate of reconstruction errors. The large calibers and relatively gradual turns associated with microtubules made errors on backbone less frequent by a factor of nearly 20 relative to on smaller and tortuous twigs. However, we propose that the circuit's resilience to developmental noise, achieved in part by connecting via multiple twigs, underlies the resilience of wiring diagrams to the omission of small dendritic branches, the most typical error observed both here and in reconstructions in the fly visual system (Takemura et al. 2013).

Irregularities within a cell type guide review toward small fractions of specific neuronal arbors that could be responsible for a potential error. While reconstructing a neuron, a user can quickly pull up its complete anatomy and connectivity to compare to homologous cells or inspect for irregularities and, crucially, return quickly to the locations in the image data necessary to make the appropriate decisions. We find that this smooth flow from image data to high level features and back to image data — without *post hoc* or offline analysis — is possibly the most important feature in our EM reconstruction technique.

Dispensing with repeated reconstruction without reducing accuracy enables our method to support concurrent neuron reconstruction by many collaborators. This setup prevents duplicated work while ensuring that important locations are visited multiple times. For example, synaptic relations are inspected at least twice in different ways, once each from the pre- and postsynaptic side. The presence of existing

and correct skeletons in complicated areas, such as registration errors between consecutive sections or gaps, reduces the time necessary for resolving possible ambiguities and effectively provides an extra step of proof-reading by not allowing contradictory reconstructions. Further savings originate in the reuse of data, for example exploratory reconstruction of backbones in search of specific neurons or branches pruned during proofreading are merged into the arbor currently being reconstructed. In summary, in a collaborative environment, the more neurons that are reconstructed, the faster new ones can be added, and the fewer errors existing reconstructions will contain.

Automated methods will be necessary to map circuits with more than a few thousand neurons (Helmstaedter 2013), but they require extensive proof-reading (Chklovskii et al. 2010; Plaza et al. 2012; Haehn et al. 2014). Our methods for analysis of arbor morphology, synaptic distribution and circuit structure and reliability, and their application in proof-reading, apply equally to manually and automatically reconstructed neurons. Neuroanatomical measurements suggest mixed strategies for leveraging both automated algorithms and human effort. For example, mitochondria can be reliably located automatically (Lucchi et al. 2011; Funke et al. 2014) which, together with our finding of a distance constraint between mitochondria and presynaptic sites, could assist in automated synapse detection (Kreshuk et al. 2011; Becker et al. 2012; Kreshuk et al. 2014). Similarly, the properties of neuronal backbone and twigs suggest that algorithms for the automatic detection of microtubules in serial section EM would be a profitable source of constraints for automated reconstruction of neurites across consecutive sections (Vazquez-Reina et al. 2011; Funke et al. 2012). While we only considered the relationship between error rate and the presence or absence of microtubules, with the use of automated detection methods it will be important to look at more detailed measures of arbors such as the number of microtubules, curvature, or caliber.

We expect that the fundamental principles described here will hold for regions of the nervous system of the adult fly other than the optic neuropils, although the exact properties, e.g. mean twig length, will likely differ across time and cell type. In contrast, applying our method to vertebrate neurons would require extensive adaptation. Connectivity and morphology are less stereotyped, making it impossible to build strong expectations from a single example of a cell type. However, given careful measurement of cell types, such as the distribution of branches or synapses in a given retinal layer already used in retinal reconstructions (Helmstaedter et al. 2013; Kim et al. 2014; Seung et al. 2015) or automated cell type discovery using anatomy and connectivity (Jonas et al. 2015), similar neuroanatomically-guided reconstruction and proofreading could inform methods for vertebrate connectomics. The key challenge will be to find, as we did here for *Drosophila* neurons, measurable quantities that are most informative for discovering those errors that have the greatest impact on the interpretation of wiring diagrams.

Sensorimotor circuitry

Larval locomotion, like many motor patterns, results from rhythmic activation of motoneurons (Heckscher et al. 2012), but few central components of the underlying premotor circuitry had been identified (Kohsaka et al. 2014; Couton et al. 2015). Our reconstruction of propriomotor circuitry revealed a complex network comprised of numerous cell types, including a subset of those previously described (Kohsaka et al. 2014). We identified a rich collection of local neurons, including neurons anatomically well-suited to provide common drive to synergistic muscles (Schaefer et al. 2010) and thus likely a key motor network components. Using anatomically faithful simplifications of neuronal structure, we found several premotor microcircuits employing dendro-dendritic and axo-axonic synapses in parallel with conventional axo-dendritic synaptic connections. For example we found a GABAergic input pre- and post-synaptic to motoneuron input, a motif also observed in mammalian motor circuits (Fyffe et al. 1984).

Although the motor system is rhythmically active in the absence of sensory input (Suster et al. 2002), proprioceptive sensory feedback is required for natural coordination and timing (Hughes et al. 2007; Song et al. 2007). We found diverse and complex circuitry for relaying proprioceptive information, including GABAergic and glutamatergic neurons directly relaying proprioceptive input to motoneurons. This motif is well-posed to provide an inhibitory "mission accomplished" signal to surpress motoneuron activity after a successful contraction during forward locomotion (Hughes et al. 2007). However, we also observed that many synaptic partners of dbd were themselves presynaptic to neurons downstream of the other proprioceptive axons, suggesting other, more complex roles for proprioceptive feedback in modulating motor activity. Surprisingly, the axon terminals of proprioceptive neurons themselves almost entirely lacked presynaptic input. This suggests that proprioceptive input is privileged by the larval nervous system and not under fast, dynamic modulation by central circuitry (Clarac et al. 1996), unlike proprioceptive afferents in the locust leg (Burrows et al. 1994) and other somatosensory modalities in the larva (Ohyama et al. 2015).

Wiring diagrams have been deemed necessary, yet not sufficient, for understanding neural circuits (Bargmann 2012) and a fast approach for discarding hypotheses of circuit function (Denk et al. 2012; Takemura et al. 2013). The neuronal wiring reconstructed here offers insights into the structure of an insect motor circuit and its control by sensory feedback, and serves as a complementary resource for detailed functional studies. With the circuit mapping tools and methods demonstrated here, fast, accurate, and targeted reconstruction of circuits in *Drosophila* larva (Ohyama et al. 2015) and adult, and other species (e.g. *Platynereis* (Randel et al. 2015)) is possible.

2.4 METHODS

Serial-section Transmission Electron Microscopy

Wild-type *Drosophila* first instar larval central nervous systems were manually dissected by mechanical separation of the anterior tip of the larva from the posterior portion in PBS, and immediately transferred to 2% glutaraldehyde in 0.1 M Nacacodylate, pH 7.4 buffer. Samples were post-fixed in 1% OsO_4 in the same buffer and stained *en bloc* with 1% aqueous uranyl acetate before subsequent dehydration in ethanol and propylene oxide, and embedding in Epon. Serial 45 nm sections were cut with a Leica UC6 ultramicrotome using a Diatome diamond knife, and picked up on Synaptek slot grids with Pioloform support films. Sections were stained with uranyl acetate followed by Sato's lead (Sato 1968). Sections were then imaged at 4.4 nm x 4.4 nm resolution using Leginon (Suloway et al. 2005) to drive an FEI Tecnai 20 transmission electron microscope. The resulting 77,000 image tiles were contrast-corrected, montaged and registered with TrakEM2 (Cardona et al. 2012) using the nonlinear elastic method (Saalfeld et al. 2012). The generated data volume of $22775 \times 18326 \times 462$ voxels corresponds to a volume of $91 \times 73 \times 21 \mu m^3$. The data set covers approximately the posterior half of abdominal segment A2, and a nearly complete abdominal segment A3.

Immunolabelings and light microscopy

Central nervous systems of *Drosophila* larva were dissected in PBG (10% NGS [Normal Goat Serum] in 1% PBS) with tweezers under a scope and fixed with 4% paraformaldehyde in 1% PBS for 30 min, washed 3×10 minutes in PBT (1% Triton-X100 in 1% PBS), blocked for 1 h in PBG, then washed 3×10 minutes in PBGT (1% Triton-X100 in PBG), and incubated with primary antibodies (rabbit anti-GABA: Sigma A2053 at 1/1000; chick anti-GFP: Abcam ab13970 at 1/2000) in PBGT for 24 h at 4^oC on small Eppendorf tubes laid on a gentle horizontal shaker. They were then washed 4×15 min in PBT, and incubated with secondary antibodies (goat antichick 488: Invitrogen, at 1/500; goat anti-rabbit 568: Invitrogen, at 1/500) in PBGT at 4^oC in Eppendorf tubes wrapped in aluminum foil on a horizontal shaker for 24 h, subsequently washed in PBT 4×15 min (wrapped in foil), and mounted on poly-lysine coated glass slides. Then samples were dehydrated by dipping the slides in an alcohol series (30%, 50%, 70%, 90% in distilled water, then twice 100%) and then in 100% xylene (3 times) using Columbia glass jars with slits for slides; then mounted on glass slides in DePeX (Li et al. 2014) using spacer coverslips on the sides. Glass slides were left to dry in a large Petri dish with a lid, wrapped in foil and at 4^oC for 3 days. Imaging was done with a Zeiss 710 confocal laser-scanning microscope. Positive immunoreactivity was confirmed by consistent labeling across multiple GFP-labeled cells per imaged nervous system in two or more nervous systems.

Synapse clustering algorithm

In order to associate synaptic connectivity not to whole neurons, but to regions of neurons, we adopt an approach where we cluster nearby synapses. Mean shift clustering has been shown to be an effective approach to finding synapse clusters in 3d without assuming a particular number of groups a priori (Binzegger et al. 2007). This approach involves convolving synapse locations with a Gaussian kernel to estimate the density of synapses in space. A cluster is then the set of synapses for which, starting at their location, gradient ascent reaches the same density peak. However, locations on one neuron that are close in space can be very far apart along the neuron. Here, instead of considering the density of a neuron's synapses in 3d space, we use a similar procedure to estimate the density of synapses at every point on the arbor (following the cable) and define synapse clusters in the same manner. The only parameter in both approaches is the width of the Gaussian kernel, a physically meaningful parameter.

For these purposes, we define the skeletonization of a neuron to be a graph with a set of nodes $\mathcal{N}$ with locations $\mathbf{x_i}$ for $i \in \mathcal{N}$ and edges $\mathcal{E}$. Because the neuron's graph is tree-like, there is a unique non-overlapping path on the graph between any two points $i, j \in \mathcal{N}$ with distance δ_{ij}. All synapses (both inputs and outputs) associated with the neuron are represented by the set of their associated nodes, $\mathcal{S} \subset \mathcal{N}$, noting that the same node can be associated with multiple synapses and thus appear multiple times in $\mathcal{S}$. For every node in the neuron graph, we compute the synapse density

$$d(i) = \sum_{j \in \mathcal{S}} \exp \left(-\frac{\delta_{ij}^2}{2\lambda^2} \right)$$

where λ is a bandwidth parameter that effectively determines the size of clusters, and presynaptic sites of polyadic synapses are counted as many times as they have postsynaptic partners. Note that due to branches, a single synapse close to a branch point may contribute more total density than one that is very distant, reflecting its greater within-graph proximity to more of the arbor. We then look for all maxima in the synapse density and the basins of attraction that flow to them via gradient ascent (i.e. starting at a given node, moving along the maximally positive difference in density between adjacent nodes). A cluster of synapses is then all synapses associated with nodes found within a single basin of attraction of the density function. For neurons found in the 1st instar larva, with $\approx$500–2000 μm of cable, bandwidths around 8–30 μm provide clusterings that match the subjective description of either "dendritic arbor" or "axon". Smaller bandwidth values result in more granular breakdowns of dendritic and axonal trees (e.g. dbd axons in Fig. 2.3i).

Synapse flow centrality of segments of a neuronal arbor

We define synapse flow centrality (SFC) as the number of possible paths between input synapses and output synapses at each point in the arbor. We compute the flow centrality in three flavors: (1) centrifugal, which counts paths from proximal inputs

to distal outputs; (2) centripetal, which counts paths from distal inputs to proximal outputs; and (3) the sum of both.

We use flow centrality for four purposes. First, to split an arbor into axon and dendrite at the maximum centrifugal SFC, which is a preliminary step for computing the segregation index, for expressing all kinds of edges (e.g. axo-axonic, dendro-dendritic) in the wiring diagram, or for rendering the arbor in 3d with differently colored regions. Second, to quantitatively estimate the cable distance between the axon terminals and dendritic arbor by measuring the amount of cable with the maximum centrifugal SFC value. Third, to measure the cable length of the main dendritic shafts using centripetal SFC, which applies only to insect neurons with at least one output synapse in their dendritic arbor. And fourth, to weigh the color of each skeleton node in a 3d view, providing a characteristic signature of the arbor that enables subjective evaluation of its identity.

Segregation index: a measure of synaptic sign distribution in a neuronal arbor

An ideal, textbook neuron has a purely dendritic tree and a purely axonal tree, that is, one tree that only receives inputs and another that only delivers outputs onto other neurons. In reality, dendro-dendritic and axo-axonic synapses are common. We have observed that homologous neurons (e.g. identifiable neurons in the left and right hemisegments) have a similar synaptic distribution, which differs from that of other neurons. In insects, we find (1) neurons that abide to the canonical model as outlined above—separated input and output trees—such as motor neurons and many types of projection neurons; (2) neurons that consist of a unique tree with mixed inputs and outputs (some of these neurons have been reported as non-spiking (Burrows 1992)); and (3) everything in between.

Having clustered synapses into groups (either by synapse clustering or by splitting the neuron by centrifugal synapse flow centrality), we ask how neuronal inputs and outputs are distributed among the clusters. If the clustering can adequately separate axon from dendrite, then a highly polar neuron will have most of its outputs on the "axon" cluster and most of its inputs on the "dendrite" cluster. Motor neurons in the abdominal segments of the *Drosophila* larva are examples of completely polarized neurons. Conversely, highly non-polar neurons will have inputs and outputs distributed homogeneously throughout their arbor. An example of the latter are non-spiking neurons that perform extremely local computations, such as GABAergic antennal lobe interneurons (Wilson et al. 2005).

A measure of synaptic sign distribution in a neuronal arbor has the potential to distinguish similar yet uniquely different neurons, as well as to suggest broad functional roles of the neuron. Here, we describe an algorithm to quantify the degree of segregation between input and outputs in a neuronal arbor.

For each synapse cluster i on a neuron with N_i synapses, compute the fraction p_i that are postsynaptic. We measure the uniformity of the distribution of inputs and outputs within cluster i by computing its entropy, for which we consider synapses as entities with two possible states: input or output. At the limits, when all synapses of the cluster are either inputs or outputs, its entropy is zero. When half of the synapses

are inputs and the other half are outputs, the entropy is maximal. The contribution of each cluster i to the total entropy is weighted by its fraction of the total synapses (either inputs or outputs).

The entropy of the input/output distribution for each cluster is then

$$S_i = -\left(p_i \log p_i + (1-p_i)\log\left(1-p_i\right)\right).$$

The total entropy for the arbor is then just

$$S = \frac{1}{\sum_i N_i}\sum_i N_i S_i.$$

However, for reference we need to compare this to an unstructured arbor (i.e. non-clustered) with the same total number of inputs and outputs; for this we consider the whole arbor as one cluster, and we compute

$$S_{norm} = \frac{\sum_i p_i N_i}{\sum N_i}\log\left(\frac{\sum_i p_i N_i}{\sum N_i}\right) + (1-\frac{\sum_i p_i N_i}{\sum N_i})\log\left(1-\frac{\sum_i p_i N_i}{\sum N_i}\right)$$

(where $\frac{\sum_i p_i N_i}{\sum N_i}$ is just the total fraction of synapses that are inputs).

We define the segregation index as

$$H = 1 - \frac{S}{S_{norm}}$$

so that $H = 0$ corresponds to a totally unsegregated neuron, while $H = 1$ corresponds to a totally segregated neuron. Note that even a modest amount of mixture (e.g. axo-axonic inputs) corresponds to values near $H = 0.5$–0.6 (Supplemental Fig. 2.17).

Model for estimating false negatives and false positives in the wiring diagram

To estimate the probability of completely missing an edge as a function of the number of synapses in the edge, we combine the twig distribution with the error rates obtained from multi-user reconstruction. We found that our reconstruction identified 672 out of 761 twigs, giving our method a recall rate for complete twigs of $q = 0.88$. Let the distribution of n_b twigs across edges with m synapses be $p(n_b; m)$. Assuming each branch has a probability q of being correctly observed, the probability of not observing a specific connection across all n_b twigs is $(1-q)^{n_b}$. The probability of omitting an edge with m synapses is thus given by

$$P_{loss}(m) = \sum_{n_b=1}^{m} p(n_b; m)(1-q)^{n_b}.$$

In our reconstruction method, we emphasize connections that are found consistently between cells of the same type, typically hemisegmental homologs of a presynaptic and postsynaptic neuron. Using a simple model, we approximate the

likelihood of introducing a symmetric, but false, edge between cell types in our wiring diagram due to reconstruction mistakes. Consider two neurons, $j = 1, 2$, of the same cell type, with the dendrites of each sufficiently close to N axons to physically permit connections. To add an incorrect edge to the connectivity graph and not just reweight an existing one, any added branches must have synapses from otherwise unconnected neurons. Let the number of axons with zero true connectivity be N_0. Assuming symmetry, the number of axons for both neurons should be similar. We then suppose that errors add m synapses to each neuron, with each synapse assigned uniformly at random to axon $i \in (1, 2, ..., N)$, with the final added edge count onto neuron j from axon i given by $k_{i,j}$. For clarity, we order the axons such that $i \leq N_0$ designates an axon with no true connectivity. We then ask what is the probability that both $k_{i,1}, k_{i,2} > k_\theta$ for any $i \leq N_0$.

The parameters of this model will vary depending on the properties of the neuron and neuropil in question. For example, larger neurons will have more opportunities for error than smaller ones, while neurons with more stringent synaptic specificity have more true zero edges than broadly synapsing neurons. To estimate a realistic set of values for the neurons here, we consider the validation data. Because nearly all false positives occur on the terminal arbors, the number of synapses added by error m can be expressed as $m = rL_t\bar{k}$, the product of the rate of incorrect branches per length of twig r, the total length of twigs L_t, and the expected synapses per added twig $\bar{k}$. Based on the independent reconstructions, we estimate r as 6 false positive errors per 1.63×10^3 μm, $\bar{k} = 5$ synapses, and a typical $L_t = 257$ μm, making $m = 5$. Determining N and N_0 is difficult, as it requires knowledge of axons that would not be in the connectivity-driven reconstruction. We estimate reasonable values using the aCC and RP2 network, since aCC dendrites strongly overlap RP2 dendrites, but have several presynaptic neurons not shared with RP2 (Fig. 2.5). In addition to the axons presynaptic to RP2, we find a mean of $N_0 = 36$ inputs exclusive to aCC, so we estimate $N = 87$. We simulated the 10^6 iterations of the model for $k_\theta = 1$–4. To investigate more extreme errors than the ones measured, we also considered $m = 37$ synapses, the largest twin twig observed across all neurons looked at, and $m = 20$ synapses, a more typical value for the largest twig of a single neuron.

Estimating skeleton reconstruction and review time

Skeletons are chimeras, with multiple contributors creating various parts at different points in time. We estimate the total amount of time spent skeletonizing an arbor—including its synapses—by counting the number of 20-second intervals that contain at least one skeleton node or connector node related to the skeleton. This approach is robust to the discontinuity in time and authorship of adjacent skeleton nodes, but tends to overestimate slightly reconstruction time, given the contribution of 20-second intervals for single nodes that were created earlier in time as pre- or postsynaptic placeholder skeletons with a single node, and which were subsequently merged into the growing skeleton. If the latter were each counted as contributing 6 seconds only, reconstruction times per skeleton typically shrink between 15 and 25%.

We estimate the time for the systematic review of a neuron similarly, with the added caveat that parts of the same arbor may have been reviewed more than once. We count the number of minutes for which at least one skeleton node was reviewed, for every contributor that reviewed a fraction of the arbor, and then add up all the minutes of each contributor.

CATMAID software

We rewrote and greatly developed the Collaborative Annotation Toolkit for Massive Amounts of Image Data, CATMAID (Saalfeld et al. 2009) (GPL), to implement our methods for neural circuit reconstruction, visualization and analysis, and with a user and group management system with sophisticated permissions for graded access. The toolkit consists of four parts: (1) the client (a web page), and three types of servers, namely (2) an application server based on the Django web framework (`https://www.djangoproject.com`), (3) one or more image volume servers, and (4) an instance of the relational database PostgreSQL (`http://www.postgresql.org`) with all non-image data, which includes metadata such as the spatial information of skeletons, the location of which types of synaptic relations, the text annotations, timestamps and provenance of every action. The original web client accesses, in constant time, arbitrary fields of view of remote stored image volumes. We have greatly extended this capability to include 3-way views (XY, XZ and ZY) and a number of color overlays for multi-channel data such as light-microscopy images or computed derivative data such as membrane probability maps. Analysis of neurons and circuits is performed primarily in the client using the programming language JavaScript, relying on a large number of open source libraries for numerical processing, data management and visualization (D3.js, Numeric Javascript, Cytoscape.js, three.js, jsNetworkX, Raphaël, jQuery, SVGKit). Offline analysis for validation and probability calculations was performed by custom scripts in MATLAB (Mathworks). Documentation and installation instructions are available at `http://catmaid.org`.

Preparation of EM images for CATMAID

For display in CATMAID, we Gaussian-smoothed montages of registered EM images (sigma=0.7 pixels, sufficient to remove high-frequency noise to increase the effectiveness of JPEG compression without sacrificing signal) and then generated an image pyramid with five zoom levels and diced it to collections of 256×256 pixel tiles (512×512 and larger can work better for fast Internet connections), stored in JPEG format (75% compression and stripped of headers with jpeg-optim). This approach reduced data storage from over 700 to 90 gigabytes, which were served from a fast seek time solid-state hard drive.

Server and database configuration

We setup a single server machine (Intel Xeon X5660 with 12 cores, 48 GB of RAM, 10 Gb network card) running Ubuntu 12.04 to host the PostgreSQL database, the image server and the Django server. LDAP id caching was enabled for best performance. Images were stored on high-performance solid-state drives mounted with noatime flag or as read-only, and served via proxy with in-RAM varnishd for caching. The database was configured with large shared buffers (4 GB) and autovacuum on (naptime: 8642 min; scale factor 0.4; analyze scale factor 0.2; cost delay -1; cost limit -1) for optimal performance. We chose to serve pages with Nginx, running 8 processes, with epoll on, 768 worker connections, disabled logs and gzip on (except for JPEG image tiles) for best performance, and with public caching and no-expire settings for images. Django was run via Gunicorn with python 2.7 using 8 processes.

ACKNOWLEDGEMENTS

CONTRIBUTIONS

AC, CSM designed the study;
AC, RF performed ssTEM;
CSM, AC devised algorithms;
ML, SG, AC, TK, ACh, SS, FM wrote the software;
AC, CSM, SG, ML mapped the circuits;
MZ, FL performed light-microscopy labelings and annotated mitochondria;
CSM, AC, SG analyzed the data;
AC, CSM, SG wrote the paper with input from co-authors.

2.5 SUPPLEMENTAL TEXT

NEURONAL RECONSTRUCTION AND PROOFREADING IN CATMAID

Finding known neurons in vast EM volumes

Finding specific neurons in unreconstructed data demands prior knowledge, for example using image volumes of genetically labeled neurons, and reference markers like neuroglian or fasciclin II tracts (Landgraf et al. 2003) that have anatomical correlates that are conspicuous in EM. The ability to navigate vast volumes at intermediate or low resolution aids in identifying large features such as nuclei, nerves, trachea, or neuropil boundaries, helpful for crossing the resolution gap between light-level microscopy and EM. Although not used in the project described here, to further facilitate finding specific neurons of interest in vast EM volumes, CATMAID supports overlaying other volumetric image data on the EM images.

Given a good guess of the approximate location, finding a neuron of interest involves reconstructing multiple partial backbones (the low-order, microtubule-rich processes), an operation that consumes up to 10 minutes per arbor. The large caliber, the presence of continuous microtubules, and the paucity of synapses on the primary neurites facilitate fast and accurate reconstruction of backbones. In larval *Drosophila*, even partially reconstructed backbones suffice to identify individual neurons by comparing with high resolution single neuron images at light level, given that the lowest-level branches are typically sufficient for unique identifiability of individual neurons. In our experience, the best starting points are stereotyped features like the main branch points, tracts, commissural crossings, or the neuropil entry point of the primary neurite bundle of all sibling neurons of the same lineage (Cardona et al. 2010). Unfinished backbones remain for other contributors to expand or merge into full arbors later, if desired.

Reconstruction of metadata-rich skeleton representations of neurons

We represent neuronal arbors as skeletons. Each skeleton is a collection of nodes, each with a spatial coordinate and additional metadata (authorship, timestamp, radius, text annotations, revision status and synaptic relations). Each node refers to one single other node as its parent, except for the root node which has none, defining a topological tree.

The process of creating a skeleton to represent a neuron is as follows. Upon locating a cross-section of a neuron of interest, a contributor places a seed node in it and proceeds to grow the skeleton of the neuron by browsing to adjacent sections and placing further nodes, which are linked together in a sequential parent-child relationship. The relationship—a skeleton edge—is annotated for perceived accuracy using a confidence value from 1 to 5, with the maximum of 5 the default value. This confidence value is then used in the analytical tools, for example to split neurons at these nodes to consider the skeletons on each side as separate entities. Skeleton edges with reduced confidence are displayed in the 3d viewer as a clickable point

to jump to that location in the image stacks. Upon reaching a branch point, we place a node at the base of each branch, which acts as a placeholder to return to later. When the neurite being annotated ends, we tag the terminal node as complete ("ends"), and resume from the nearest open leaf, which is found automatically (with the 'R' key) or from a table of skeleton nodes. Eventually all open leaves have been explicitly declared finished and the first draft of the arbor is complete. Optionally, a neurite radius can be measured for each skeleton node, with cylinders modeling arbor segments in 3d and spheres modeling the somas.

Each skeleton node may also be tagged with any number of arbitrary text snippets to express metadata. Search tools enable finding tags in a specific skeleton or across the data set. We use text tags on skeleton nodes in two different ways. First, free text can be used to denote structures of interest (e.g. "Golgi apparatus") or as a personal or team communication convention (e.g. "check this synapse"). Second, standardized text tags, accessible via key bindings, are used by the analysis tools to better interpret the annotations. The standardized tags are: "ends", "uncertain end", "uncertain continuation", "not a branch", "soma", "TODO", "mitochondrion", and "microtubules end". The first four mark terminal leaf nodes. In detail:

- "ends" (key binding 'K') signals that the skeleton does not continue beyond the node. An open end (untagged leaf node) signals that a neuronal process needs to be continued, or at least looked at further.

- "uncertain end" (key binding 'U') signals that a leaf node likely continues, but can't be followed further. These are often used when a twig enters an occlusion (e.g. an electron-dense precipitate stain covering over one or more sections), a gap (a set of missing sections), or is ambiguous due to the thickness of a single section being comparable to the thickness of a very small neurite.

- "uncertain continuation" (key binding 'C') signals that a leaf node in all likelihood continues into a large branch (e.g. it contains one or more microtubules), but its continuation is unclear due to occlusions or gaps. These are generally resolved later by reconstructing several candidate continuations and finding the one consistent with fundamental neuroanatomy (e.g. continuation of the microtubules and heuristics such as that neurons have one soma) and homologous neuron morphologies from EM or light microscopy. This tag typically indicates large omissions in the reconstructed neuron.

- "not a branch" (key binding 'N') indicates that the skeletonization involved branches that do not represent the true neuronal topology. A typical situation in which this occurs are large varicosities, such as at sensory axon terminals, which are compact but may have bulges that span multiple sections. An annotator exploring such a varicosity may lay down branches either because this was not initially obvious or to help annotate synapses. This tag signals this, and allows the analytical tools to collapse these artifactual branches, reassigning any synapses to the base of the branch.

- "soma" (key binding 'M') signals a leaf node located at the soma. While we generally reroot the skeletons at the soma, when forgetting to do so the analytical

tools will reroot the skeleton at the node labeled with "soma" for the purposes of the analysis. The built-in skeleton analytics widget will automatically bring attention to this and other issues during the review process.

- "TODO" (key binding 'L') signals the need to revisit a skeleton node without specifying any further as could be done with a custom text tag. Nodes with this tag are indicated in the 3d viewer as a clickable sphere to let the user jump to that location in the image stacks.

- "mitochondrion" signals the presence of a mitochondrion in the immediate proximity of the tagged skeleton node.

- "microtubules end" (key binding 'F') signals that no microtubules are found distal to the tagged skeleton node, relative to the soma.

Synapse annotation

Synapses are modeled as relations between nodes of different skeletons, mediated by a connector node. The presynaptic arbor has a node presynaptic to the connector, and one or more postsynaptic arbors have nodes postsynaptic to the connector. Each relation is annotated with a confidence value from 1 to 5, with 5 being the default and highest confidence. The connector node allows the annotation of synapses as placeholders, expressing that a synapse exists at a defined location even if the partner arbor hasn't been reconstructed yet. The one-to-many structure of the connector models either the polyadic synapses typical of invertebrates or the one-to-one synapses typical of vertebrates.

In *Drosophila* larva, we identify synapses at locations that match the following criteria: 1) a T-bar presynaptic structure; 2) vesicles nearby; 3) distinct synaptic cleft; 4) postsynaptic densities. Synapses of different neuron types can differ in ultrastructural appearance, although cholinergic, glutamatergic and GABAergic synapses appear very similar. Presynaptic sites on serotonergic neurons are far smaller and distinct from the others (Supplemental Fig. 2.7)

Systematic review of a skeleton

Reviewing a skeleton consists in visualizing each of its skeleton nodes in a sequence, and adding or editing nodes and synaptic relations to the skeleton as necessary. For this purpose, we partition the arbor to generate the smallest possible set of the largest possible sequences of continuous nodes to minimize the number of times that the reviewer has to switch to a different arbor path. We sort leaf nodes by path length (in number of intermediate nodes) to the root node in descending order. Starting from the most distal leaf node, we generate a sequence of nodes all the way to the root. Then we pick the second most distal node and generate another sequence of nodes until reaching a branch point that has already been assigned to a sequence, and so on for each remaining leaf node. When done, sequences of nodes are sorted by length. The reviewer then iterates each sequence, automatically

marking each node as reviewed upon visiting it (using 'q' and 'w' key bindings to go forward and backward in a sequence, and 'e' to jump to the beginning of the longest unreviewed sequence). As a visual aid, each node is centered in the screen, facilitating the detection of changes in the contour of the sectioned neurite, as well as drastic shifts of the field of view that indicate an error (e.g. a jump to an adjacent neurite).

The enforcement of a unique directionality and simple one dimensional path—from distal ends towards inner parts of the arbor or the soma—facilitates noticing glaring inconsistencies such as a path starting off large and microtubule-rich, then reducing to small and microtubule-free, then becoming again large and microtubule-rich. In other words, a review approach coherent with the expected tapering out of neurite caliber and cytoskeleton from soma to distal ends adds context that helps the reviewer.

The total fraction of nodes of a skeleton that have been reviewed is indicated in most tools that can display lists of neurons (e.g. selection table, connectivity tables, connectivity graph), as well as a skeleton coloring mode in the 3d viewer. The coloring (green for 100%, red for 0%, and hues from green to red for any value in between) enables evaluating at a glance the current status of revision of e.g. all upstream and downstream synaptic partners of one or more neurons in the connectivity widget, of all neurons in a wiring diagram in the connectivity graph, or the revision status of a specific branch in the 3d viewer. Given that one or more users may review any node of a skeleton, and the different proficiency of each user, the settings of each user enable the definition of a team of trusted reviewers. The coloring of skeletons by review status therefore can be restricted to the user's self reviews, to the union of everyone's reviews, or to the union of all reviews performed by the team of trusted reviewers.

Synapses are effectively reviewed multiple times, given that they are seen from at least two arbors (the pre- and the postsynaptic); more in the case of polyadic synapses, as nearly all synapses in the *Drosophila* larva are. We consider synapses as two elements: the presynaptic relation between a skeleton node and a connector node, and the postsynaptic relation between a skeleton node of a neuron and the connector node. Reviewing the associated skeleton node tacitly marks its part of the synapse as reviewed.

The process of creating a skeleton for a neuronal arbor includes a self-review by the same contributor who authored or cares the most about the specific neuron. Given that neurons participate in many connections with other arbors, their morphology and connectivity are inspected many times in different contexts; when found in disagreement with expectations, individual neurons are then partially or fully reviewed by additional contributors who have reasons to suspect errors. Most of the added time cost is incurred in locating errors, a task partially facilitated by lowering the confidence of skeleton edges deemed potentially incorrect during the initial reconstruction. These low-confidence edges can be inspected first by sorting all skeleton edges by confidence or through highlighting them in the 3d viewer.

To further facilitate systematic review, a "Skeleton Analytics" tool lists in an interactive table some potential issues that must be addressed in a neuron or collection of

neurons. The user can move through the list and immediately jump to the associated region in the image data to determine whether or not the issue describes a genuine error.

- Skeleton nodes tagged with "ends" which are not leaf (terminal) nodes.

- Skeleton nodes tagged with "TODO", "uncertain end" or "uncertain continuation".

- Leaf nodes that are not tagged with neither "ends" nor "not a branch".

- Autapses, in the form of a skeleton node that is postsynaptic to a connector that is presynaptic to the same skeleton. All autapses in the larva detected so far indicated errors in reconstruction, however they may be legitimate in other nervous systems.

- Potentially duplicated synapses: when a neuron synapses twice onto another neuron within a small cable distance it is possible that the synaptic active zone, which spans from 2 to 20 serial sections, has been annotated multiple times erroneously.

- Potentially duplicated postsynaptic relations: when a skeleton receives more than one postsynaptic relation from the same connector, is possible that the extra postsynaptic relations are accidental duplicates rather than e.g. two different dendritic twigs of the same postsynaptic skeleton.

- Lack of a node tagged as "soma", or the root node of the skeleton not corresponding to a node tagged as "soma". The case of multiple somas is generally noticed immediately and addressed without needing special tools.

Neuroanatomy-driven proofreading techniques

Arbor structure

A correctly reconstructed neuronal arbor must be biologically plausible. The distribution of microtubules is a biologically grounded approach to subdivide an arbor for analysis. Additionally microtubules are robust to artifacts in serial section EM (e.g. missing sections and noise), since they span many sections while remaining straight and in a consistent configuration within a neurite (Supplemental Fig. 2.8). Errors in the backbone are readily detected by comparing with homologous neurons, or by the extreme and evident consequences of the reconstruction error, such as dramatic changes in direction (sharp angle without branching), the presence of more than one soma, missing large axonal or dendritic trees, or violations of the self-avoidance of neuronal processes, which is often but not always observed.

While most twigs were short and had few synaptic contacts, we did find three interesting outliers. Each outlier twig was much larger and some included neurites that, upon careful inspection, were in the backbone in their contralateral homolog. This could suggest that parts of the larger twigs could become backbone at a later

developmental time point, or they contained labile microtubuli that were not captured in the EM sample preparation. Furthermore these three outlier twigs all had smooth endoplasmic reticulum at their base and branched very early, suggesting that they are effectively a pair of twin twigs. Considered alone, each half fell within the dimensions of typical large twigs.

Comparing cell types

Neurons of the same cell type share many properties in common. For *Drosophila*, we define cell type as the pair of left and right homologs, symmetric across the midline, as observed from light-microscopy (Li et al. 2014; Vogelstein et al. 2014). Most cell types repeat across multiple consecutive segments, though a few do not. Exceptionally a cell type may consist of more than one pair of cells, or of a single unpaired cell with a bilaterally symmetric arbor. Quantitative analysis of the anatomy, synaptic distribution and connectivity for a group of neurons containing potential pairs of homologs helps with detecting potential issues by comparing homologs with each other and with other types.

Reconstructions of homologous neurons can differ due to true developmental differences (Supplemental Fig. 2.15), errors in EM reconstruction, misidentification of homology, or asymmetries in imaging data (Supplemental Fig. 2.8). To help detect and classify these differences we generate interactive plots of numerous user-selected metrics on demand. If a pair of homologs is consistently more similar to each other than to other neurons for all three kinds of metrics—anatomy, synaptic distribution and connectivity—the likelihood that the pair contains significant errors is low, and therefore other neurons can be prioritized for review.

We consider several anatomical quantifications that are independent of absolute spatial coordinates and orientation, avoiding issues of mirroring and alignment of neurons. To determine which measures are most helpful for identifying homologous neurons, we applied a distance metric learning method (Xing et al. 2002) that scales individual dimensions to minimize Euclidean distance between homologs and maximize distance from other cells. The two most effective measures were total cable length of branches other than the main branch and normalized difference between the number of input and output synapses.

The distribution of synapses over an arbor is characteristic of each cell type. Some cell types have distinct input and output regions with or without dendritic outputs and axonic inputs, while others have arbors with entirely intermingled inputs and outputs (Supplemental Fig. 2.17). We devised a novel metric, the "segregation index", that measures the degree of separation between input and output regions of an arbor. First we calculate the number of paths from input synapses to output synapses that pass through every node on the skeleton, a quantity we call synapse flow centrality (SFC). We found that the point on the arbor with the highest centrifugal SFC (which considers for any skeleton node only distal outputs and proximal inputs) best separates the axon (distal) from the dendrites (proximal).

To detect and analyze neurons with similar network roles, we reduce neuronal arbors and synapses to nodes and edges in a graph. From the adjacency matrix de-

scribing the connectivity, we measure the signal flow and perform graph partitioning via spectral graph analysis (Varshney et al. 2009). We find that neurons of the same cell type can group together, even when the number of neurons is variable such as in the optic lobe of *Drosophila* larva (Simon Sprecher, Ivan Larderet & Albert Cardona, unpublished observations).

Analysis of edges in the wiring diagram

Not all connections in a wiring diagram are equivalent. The synapses that count towards an edge (a connection) have a specific spatial distribution on the postsynaptic arbor. Small errors could change network topology when all synapses of an edge are spatially close or on the same twig. In contrast, when synapses are widely distributed over the postsynaptic arbor small errors will mostly alter the number of synapses per edge. From this intuition we define an edge risk metric as the fraction of inputs eliminated from the postsynaptic arbor by removing the spanning tree between all the postsynaptic sites associated with the edge. Synapses and twigs involved in high-risk edges are thus sensitive to small errors and are subject to focused review.

The neuroanatomical basis for the edge risk measurement is as follows. Dendrites taper out centrifugally, and the radius of the curvature of dendrites shortens with diminishing calibers. Therefore terminal dendrites take relatively sharp turns. We measured the smallest caliber of a distal dendrite at about 40 nm. In combination, the small caliber and a short bending radius of distal dendrites results in that a fraction of terminal dendrites will travel horizontally inside a single 45-nm section, increasing the reconstruction difficulty. Therefore synapses on terminal dendrites are more likely to be false positives, introducing noise into the wiring diagram. Conversely, synapses onto central, large-caliber shafts are relatively uncontroversial. Furthermore we note that a large number of postsynaptic sites distributed over different parts of an arbor are unlikely to be all wrong, whereas when concentrated in a single terminal dendrite, one error would render the wiring diagram edge incorrect. We therefore define the concept of synaptic robustness of an edge (the edge risk) in the wiring diagram as the complement of the fraction of all postsynaptic sites that would be removed when cutting the skeleton of the postsynaptic arbor at the lowest common ancestor segment of all the postsynaptic sites that constitute the edge.

Resolving conflicts in a collaborative setting for editing skeletons

In an environment where multiple contributors simultaneously reconstruct neuronal arbors, eventually an ongoing reconstruction reaches that of another contributor. Attempted edits are resolved according to predefined permission rules for who can edit whose work. These rules are implemented as permissions granted to a contributor to alter another contributor's work. The status of "superuser" enables a trusted expert neuroanatomist to edit at will.

Our system operates at two levels: locked and unlocked skeletons. Skeletons that are deemed complete are locked by the contributor, and by default cannot be edited by others unless they have been granted permission to do so. Unlocked skeletons,

such as partial reconstructions produced when searching for a specific neuron or when pruning away incorrect branches, can be merged or split by others at will. Neurons are unlocked by default and locking is only to be used upon completion, which prevents sudden and unexpected changes in established results. Individual skeleton nodes and their relations to connectors (which express synapses) can only be edited by the original author, or by others that have been explicitly granted permission to edit the contributions of the original author. In case of conflict or insufficient permissions, a notification system delivers the request to the contributor who can review and effect the change. The result of the collated work of multiple contributors can be inspected in 3d, with each node of the skeleton colored according to the identity of the contributor, as well as in the review table.

A circuit mapping strategy to efficiently identify strongly connected partners

As described above, to identify a neuron quickly in the larva, the first few minutes are best spent skeletonizing the largest structures on the backbone and tracing them to the soma. This minimal representation generally suffices to identify the neuronal lineage and the overall span of the arbor. When the correct neuron has been found, reconstructed in full and reviewed, we begin to map its synaptic partners.

To find out the strongly connected partners of a neuron, we use the connectivity table that aggregates all synaptic relationships, whether with fully reconstructed neurons or single-node skeletons used as placeholders to indicate synaptic partners. Starting at each single-node skeleton, we reconstruct the arbor all the way to the soma by choosing, at every branch point, the larger caliber (may require jumping back to the last branch node occasionally), momentarily ignoring the rest of the arbor. This partial reconstruction suffices to obtain a minimum of information about the partner arbor, such as the lineage. Partner neurons that receive more than one synapse from the neuron of interest will quickly accumulate further fractions of their arbors. These preferred partners—those with many synapses with the arbor of interest—can then be selected for full-arbor reconstruction, while the completion of single-synapse partners (of which a neuron has many, and which in the *Drosophila* larva may play a lesser role in understanding the circuit role of a neuron) can be postponed.

Application of prior knowledge to resolve ambiguities

Electron microscopy image volumes of neuropils contain noise. For serial section transmission EM, noise originates during fixation (e.g. broken membranes and reduced extracellular space), serial-sectioning (e.g. folds, cracks, missing sections, thick sections), counter-staining (e.g. precipitated heavy metals, dust particles, or absence of staining due to microscopic air bubbles), and imaging (e.g. locally uneven illumination, tile-wise constant noise originating in improper correction of the camera's dark- and brightfields); for examples see Supplemental Figure 2 in Saalfeld et al., 2012 (Saalfeld et al. 2012).

The most common form of noise consists in missing data either as a partial occlusion of a section, or by the loss of one or more sections. When reconstructing a neuronal arbor, upon reaching an area with missing data (a gap) we use both global and local cues to identify the correct continuation, labeling the skeleton edge that crosses the gap with an appropriate confidence value to express our degree of certainty in the decision. These low-confidence edges enter into the visualization and analytical tools for further evaluation. Generally, the direction, caliber and cytoplasmic characteristics of the neuron and its neighboring neurons suffices to identify the corresponding continuation on the other side of the gap. The larger the gap and smaller the neurite, typically the lower the confidence in the identification of the correct continuation.

Locally, gaps up to 500 nanometers (e.g. 10 serial sections) are crossable using microtubules. The number, direction and spatial arrangement of microtubules in a neurite are constant over lengths of micrometers, making them reliable structures over many sections (Supplemental Fig. 2.8). Similarly, mitochondria take tubular shapes inside neurites, and their sparseness and relatively constant dimensions identify a neurite across consecutive serial sections (Supplemental Fig. 2.8). Other cues can include the smooth endoplasmatic reticulum that lines large and mid-size neurites; the presence of vesicles of a specific kind (e.g. dark, 50-nm diameter neuropeptide vesicles, or clear large unevenly shaped vesicles, or small, packed clear-core cholinergic vesicles, and others); or other distinctive characteristics such as the presence of microtubules on a specific side of the neurite, or membrane-associated structures, or distinctive cytoplasm texture, such as relative darkness compared to neighboring neurites.

Globally, the properties of a neuronal arbor help to identify continuations across gaps. For example, an axonal neurite tends to continue being axonal in nature within the gap-sized span of a few hundred nanometers; same for dendrites. An obvious feature is that differentiated neurons present a single soma; continuations that lead to a second soma are therefore most likely incorrect.

Dynamically generated and annotated wiring diagrams guide circuit reconstruction and highlight errors

Interactive, partial wiring diagrams calculated on demand during neuron reconstruction guide circuit mapping and the identification of errors. Connectivity-dependent coloring schemes highlight desired features of the circuit, sorting neurons into groups.

A common case is the inspection one or more orders of synaptic neighborhoods. Given one or more neurons of interest (such as RP2), we load all synaptic partners into the graph. For small circuits, visual comparisons between the neighborhoods of left and right homologs can identify similar neurons (e.g. by coloring by stereotyped properties such as the ratio of inputs and outputs, or by their graph centrality; see below) and highlights missing or differently connected neurons, prompting focused review. Coloring the circuit graph relative to a central neuron highlights the relative synaptic order of all other neurons.

Given two neurons, an important circuit question is if there are any paths between them and, if so, through what neurons. This can be queried and added to the graph from within the graph widget, with filters for how many synapses an edge must be.

Other coloring modes include by betweenness centrality (Brandes 2001) of the wiring diagram (calculated as a directed graph), which stresses the role of a neuron within a circuit; and by the percentage of review of the neuronal arbor, indicating at a glance the approximate level of completeness within a group.

When reconstructing neuronal arbors with skeletons, the nodes of the skeletons are annotated with a confidence value signifying the degree of certainty in the continuation of the axon or dendrite. We carry on this confidence into the dynamic wiring diagram representation by splitting the skeleton that models a neuronal arbor at the low-confidence edges, resulting in independent graph nodes. The connectivity of these fragments aids in evaluating their impact on the wiring diagram and their potential correctness.

Contributor-centric neuron annotations

With many expert contributors come many points of view on how to describe neurons. Boundaries between multiple possible classifications are ill-defined, and therefore instead of enforcing a controlled ontology, we allow the annotation of any neuron with a number of arbitrary text snippets. These annotations can express a variety of potentially overlapping concepts, from body regions to cell types, gene expression patterns, genetic driver lines and neurotransmitter profiles, among many others.

The combination of multiple annotations on the same neuron enables powerful queries, in particular when further constraining search with the wiring diagram. For example to find the list of neurons annotated with 'GABA', 'left' and 'segment A3' that synapse onto a specific set of neurons present in another, existing list. Constraining even further by contributor, by time of annotation or by review status of the neuron magnifies the power of annotation search.

Annotations also enable the co-existence of multiple nomenclatures for naming neurons. These could be for example by GAL4 line, by developmental grouping (a name composed of region, segment, lineage and birth order), or by gene expression. In CATMAID, many widgets lists neurons by name. These displayed names are customizable, so that each contributor can see his or her own names, even if the neurons in question were created by others. To this end, each contributor chooses a setting for neuron display names among multiple possibilities, including skeleton IDs, own annotations, all annotations, or most usefully, annotations that are themselves annotated with, for example, "GMR GAL4 lines" or "Truman nomenclature" to indicate naming schemes.

Annotations, neurons and contributors are related via the co-annotation of neurons and the contributor authorship of an annotation. To make annotations discoverable, we construct an ever-expanding recursive hierarchical tree structure that starts off with three entries: the list of all annotations, the list of all neurons, and the list of all contributors, with each paginated list reducible by regular expression matching. For each annotation we list five lists: neurons annotated with it, annotations annotated

with it (which act as meta-annotations), annotations that it annotates (acting itself as a meta-annotation), the list of contributors that have used it to annotate an entity (a neuron or an annotation), and the list of co-annotations (other annotations onto the neurons that it annotates). Each annotation, neuron and contributor is expandable, letting the user navigate a graph of relations. For co-annotations, further expansions constrain the listing of neurons to those that share all chosen annotations. For example, starting at annotation 'segment A3', continuing with the co-annotation 'left', and then the co-annotation 'GABA', leads to the listing of all GABAergic neurons on the left hemisegment of abdominal segment A3. Similarly, starting from 'GABA' could lead to 'A3' and 'left' as well, resulting in the same list of neurons. This approach enables the co-existence of many contributor-centric representations of the same neuronal circuits, and therefore, the cross-pollination of information from different fields, e.g. development, genetic tools, anatomy and physiology.

In summary, the flexibility afforded by the annotation system supports from long-term, contributor-centric publication-ready naming schemes that embody established nomenclatures, to single-use lists helping personal data organization or team collaboration.

Quantification of the quality of a contributor's work over time

Collaborative reconstructions require that users be able to trust the work of others. It is therefore important for a project manager to be able to track the work of each contributor. To estimate an individual's speed and quality, we consider only contributions that have been reviewed by others, within a specific time period. We quantify the number of edits performed by the reviewer, in particular splits (cutting away an incorrect branch), merges (appending a missing branch) and the addition or removal of synapses. While speed and quality are independent, we typically see that better contributors are also faster.

After an initial period, lasting anywhere from a couple of days to about two weeks of continuous work, a contributor typically becomes acquainted with the reconstruction task and stops adding erroneous synapses or merging branches from different neurons into one. Remaining errors are typically missing branches or synapses, which are far easier to resolve and have a less significant impact on interpretation of the wiring diagram.

We observe that different areas of the nervous system exhibit profound differences in arbor and synapse morphology, from extensively branching trees in some ventral nerve cord neurons to cloistered self-contacting axons like A02l or in the olfactory lobes (data not shown). Subjectively, contributors that reconstructed neurons in diverse areas of the nervous system experienced a larger variety of shapes and morphologies, which correlated with the acquisition of greater skill.

LARVAL MOTOR SYSTEM CIRCUITRY

The data volume used was described in Ohyama et al. 2015. It is comprised of 462 sections, each 45 nm thick and imaged at 4x4 nm per pixel resolution. It is bounded anteriorly approximately at the intersegmental nerve entry point in segment A2 and posteriorly near the segmental nerve entry of segment A3. Cutting occured approximately $8°$ angle relative to a true transverse, resulting in the left side including a region slightly posterior to the right.

Using their characteristic morphology we identified and reconstructed motor neurons U1, U2, the three VUM motor neurons, aCC, RP5 and RP2 and sensory neurons dbd, dmd1, ddaD, ddaE, and vbd for segment A3. Because dbd, ddaD and ddaE axon terminals also project into anterior and posterior segments we used segmental repetition to identify the projections of these neurons from adjacent segments that participate in the local circuitry of A3. We chose to focus on dbd, aCC, and RP2 and continued to reconstruct all arbors synaptically connected to the pair of sensory axons and two pairs of motor neuron dendrites for these cells in A3.

We found 425 arbors spanning 51.8 millimeters of cable, with a total of 24,068 presynaptic and 50,927 postsynaptic relations. We named each of the arbors and annotated them as an identifiable neuron (198), an ascending or descending projection that spans the imaged volume (107), a neuron spilling over from adjacent segments beyond the imaged volume (84), or an unresolvable fragment (36) (Supplemental Fig. 2.10, Supplemental Fig. 2.11). The 198 identifiable neurons amount to 83% of all cable, 88% of all inputs and 62% of all outputs, with ascending or descending projections contributing 29% of all outputs. Interestingly, they included some of the few larval central neurons previously described, including members of a glutamatergic premotor class (Kohsaka et al. 2014) (named here as various members of the A02 lineage, e.g. A02b) and the two seratonergic cells found in larval nerve cord (Chen et al. 2008) (named here as A26d and A26e) which can modulate locomotor activity (Okusawa et al. 2014). Numerous other previously unreported cell types are included as well.
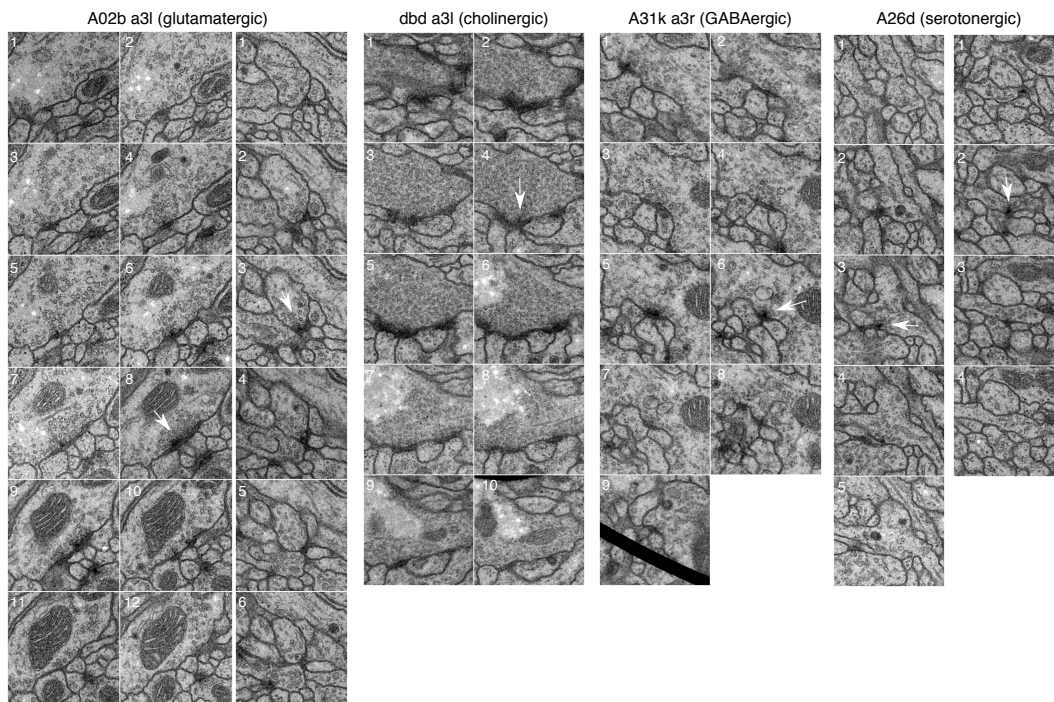
Supplemental Figure 2.7.: **Synapses of neurons with different neurotransmitters**
For all examples, white arrow points at the most prominent slice of the T-bar. Each panel measures 1024 nanometers on the side. **a)** Two examples of synapses in the glutamatergic neuron A02b (a "looper" or "PMSI" neuron; Kohsaka et al. 2014) spanning twelve and six 50 nm sections respectively. Glutamatergic synapses vary considerably in size and number of postsynaptic partners (from 3 to over 15). **b)** An example of a cholinergic synapse (a sensory axon, dbd; Yasuyama and Salvaterra, 1999) spanning 12 sections. **c)** An example of a synapse from a GABA immunoreactive cell type, A31k, spanning 9 sections. The arrowhead annotates the T-bar at panel 6 for the GABAergic synapse. The black thick line crossing panel 9 in A31k is the shadow of a fold in the support film. **d)** Two examples of the typically small, dyadic synapses found in serotonergic neurons like A26d (Chen and Condron, 2008). Synapses in the serotonergic neurons typically span only 2 sections (100 nm) and contact 2 or occasionally 3 postsynaptic partners.
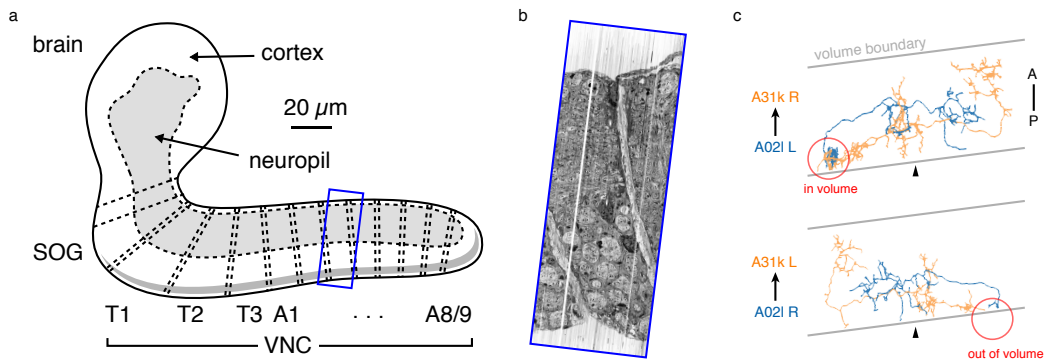
Supplemental Figure 2.8.: **Serial section electron microscopy of abdominal segment 3 of the *Drosophila* larval central nervous system.** Sections are 45 nm thick. **a)** Cartoon depicting the CNS (adapted from Ito et al. 1995) with a blue rectangle indicating the volume sectioned and imaged. Vertical discontinuous lines in the cartoon are neurohemal organs for reference of segmental boundaries. **b)** Sagittal view (side view of the EM series of sections) through the midline, intersecting neurohemal organs (white diagonal broad bands). By reconstruction of the various arbors, including RP2, that have unique branches only present in A3 (data not shown), we were able to correctly localize the series to the A2 and A3 segments. **c)** Sections were cut approximately 8 degrees from transverse. Due to the volume limits, observed circuitry could be asymmetric as in the case of A02l synapsing onto 31k (see Fig. 2.3). The region where the contact occurs (red circle) is present on one side of the data but not the other.
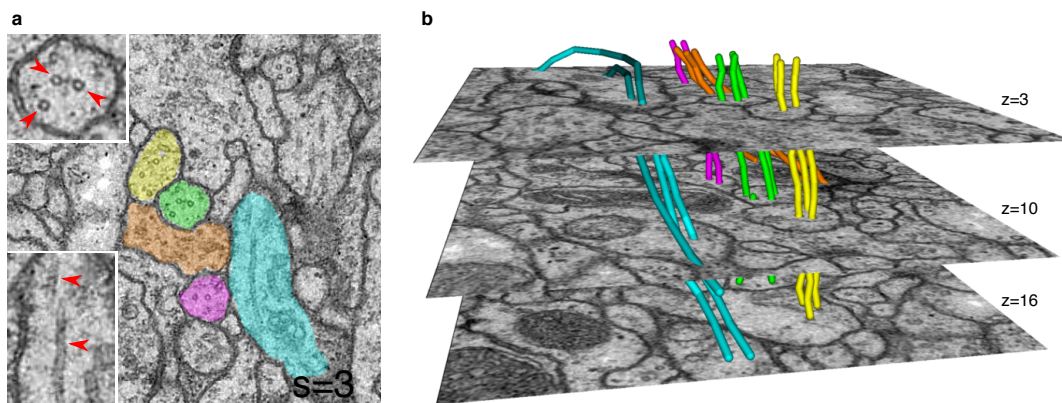
Supplemental Figure 2.9.: **Microtubules are visible in EM and span multiple sections. a)** Defining the backbone of neuronal arbors in the fly, microtubules are visible in EM sections whether cut transverse (top inset, red arrowheads) or obliquely (bottom inset, red arrowheads). **b**) Microtubules in a given neuronal process span several sections (3 shown here; microtubules were traced over 16 sections) and maintain their relative orientations. This slow change across sections makes them a useful guide for crossing small ambiguities in image data. Microtubules are color coded as in the processes in **a** and were traced and viewed in the software TrakEM2.
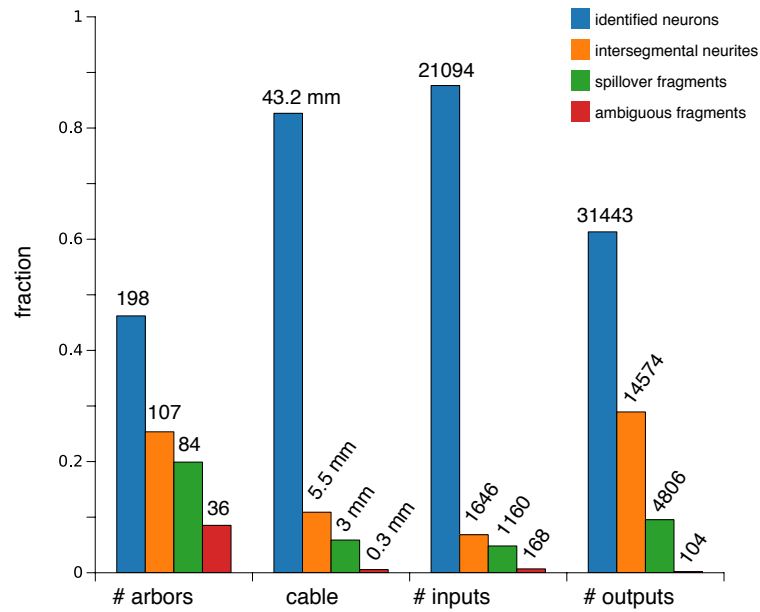
Supplemental Figure 2.10.: **Counts of reconstructed neuronal arbors.** We reconstructed 425 arbors which fall into 4 groups: "identified neurons" (arbors that could be associated with a single neuron name or at least with a lineage of origin, even if the whole arbor is not present within the imaged EM volume), "intersegmental neurites" (neurites that cross the volume from anterior to posterior, the majority of which are an unbranched axon that synapses onto motorneurons), "spillovers" (partial arbors that are not recognizable and which originate in neurons in the anterior or posterior segments) and "ambiguous fragments" (very small arbors with few inputs or outputs and which terminate within the boundaries of the volume; most of them terminate at a 3-section gap between sections 348 and 352). Identified neurons are either synaptic partners of dbd, aCC or RP2, or are any of the other 16 proprioceptive axons (the left and right ddaD in segments 3 and 4; ddaE in segments 2 and 3; dmd1 and vbd in segment 3; and dbd in segments 2 and 4) or 9 motorneurons (the left and right RP5, U1, U2 and the unpaired VUMs in abdominal segment 3). While arbors in the intersegmental group could not be identified, in numerous occasions the left and right homologs are recognizable, given idiosyncratic characteristics (not shown).
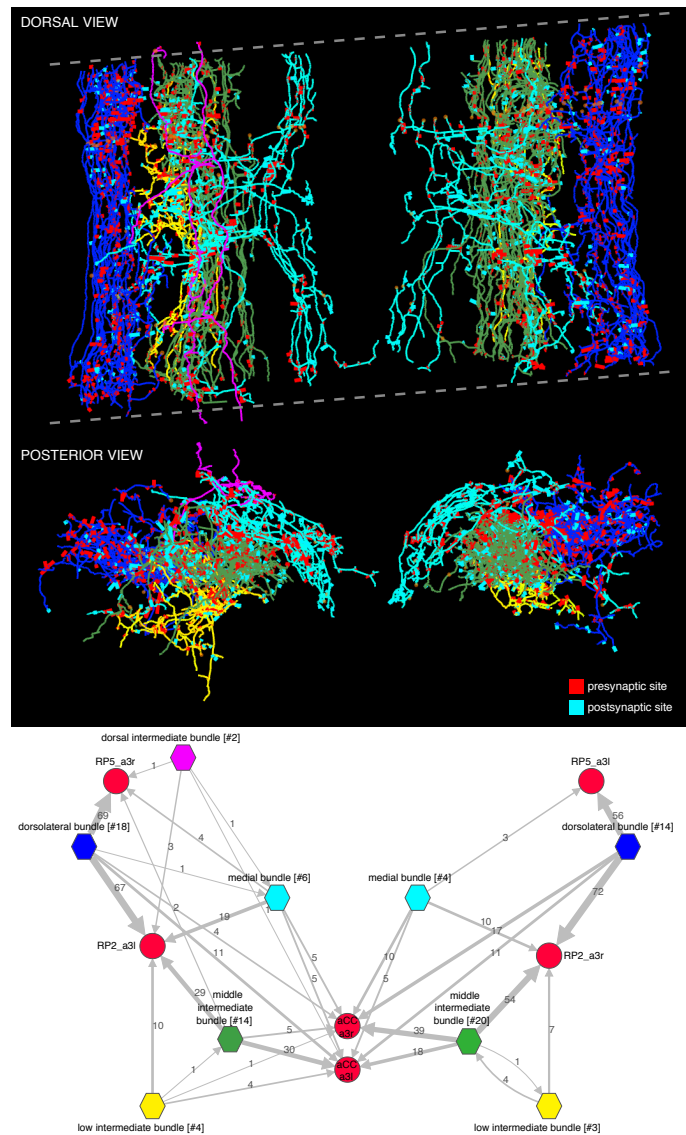
Supplemental Figure 2.11.: **Bundles of premotor axons that run the length of the imaged volume.** These axons originate in neurons whose somas are located in areas of the central nervous system beyond the limits of the imaged EM volume. We distinguish 5 bundles: dorsolateral (blue), low intermediate bundle (yellow), middle intermediate bundle (green), dorsal intermediate bundle (magenta; only present on the left side and weakly connected to motoneurons), and medial bundle (cyan). The dotted lines delimit the imaged volume, which had a tilt of about 8 degrees relative to the transversal plane. Arbors, particularly the most dorsal ones (magenta) may appear outside the plane or short of the imaged limits due to the perspective projection. Below, wiring diagram bundle-wise onto the motoneurons RP2, RP5 and aCC. The name of each bundle includes the number of member neurons in brackets; asymmetries between left and right originate in the fact that some members contribute a single synapse onto motoneurons and may not appear on the other side. Intrabundle edges are not shown. Notice how the only large source of inputs onto RP5 is from the dorsolateral bundle, which also places many synapses onto RP2. Note that we did not reconstruct all synaptic partners of RP5.
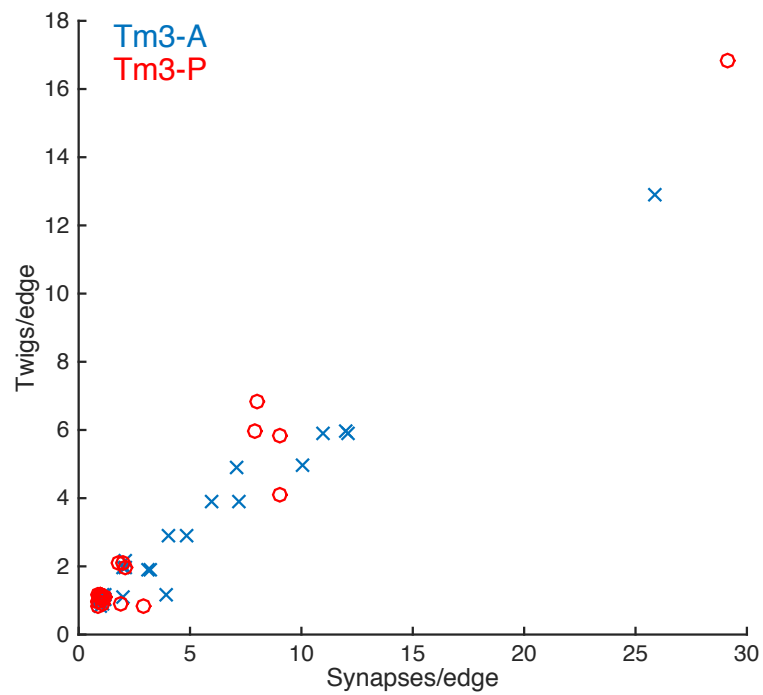
Supplemental Figure 2.12.: **Numerically high synapse edges are distributed over many twigs in adult Tm3 neurons.** Each input onto the Tm3s analyzed in Fig. 2.1 is a data point (see legend). The x-axis is the number of synapses this edge is comprised of, the y-axis is the number of distinct twigs this edge spans. As in the larval motoneuron data, edges with multiple synapses are almost always distributed across multiple twigs. Points are jittered to avoid overlap.
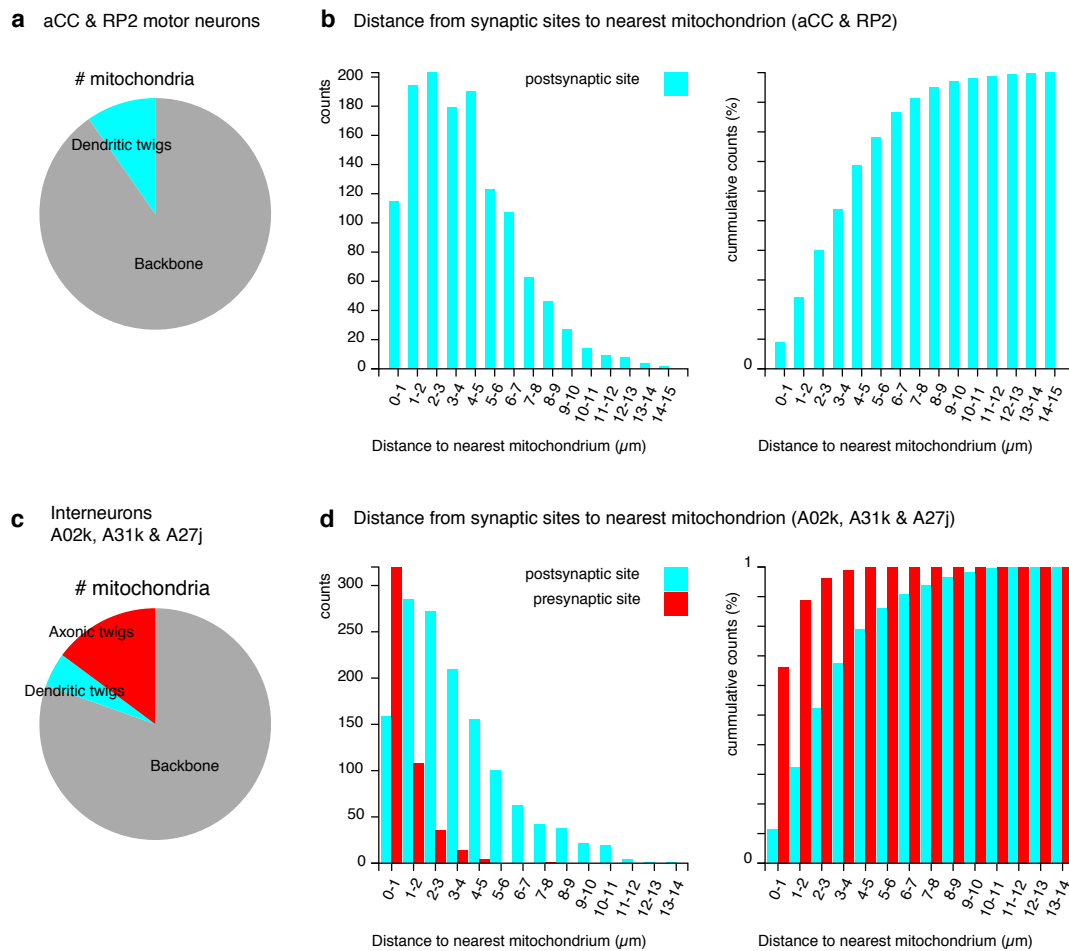
Supplemental Figure 2.13.: **Relationship between mitochondria, synapses and arbor compartments.** Two groups are considered: the left and right aCC and RP2 motoneurons, and the left and right homologs for interneurons A02k, A31k and A27j. **a), c)** Relative content of mitochondria in backbones, dendritic twigs and axonic twigs. **b), d)** Distances from pre- and postsynaptic sites to the nearest mitochondrion. We found that, for both motoneurons and their synaptic partners, 97% of input synapses are evenly distributed within 0–10 μm of a mitochondrion (maximum distance: 15 μm). 96.1% of presynaptic sites on interneurons lay within 3 μm of a mitochondrion (88.6% within 2 μm; maximum distance: 5 μm). In interneurons, the backbone contains 79.2% of all mitochondria, axonic twigs 14.6% and dendritic twigs only 4.6%. motoneuron backbones (the part within the CNS) contain 90.3% of all mitochondria and their twigs the remaining 9.7%.
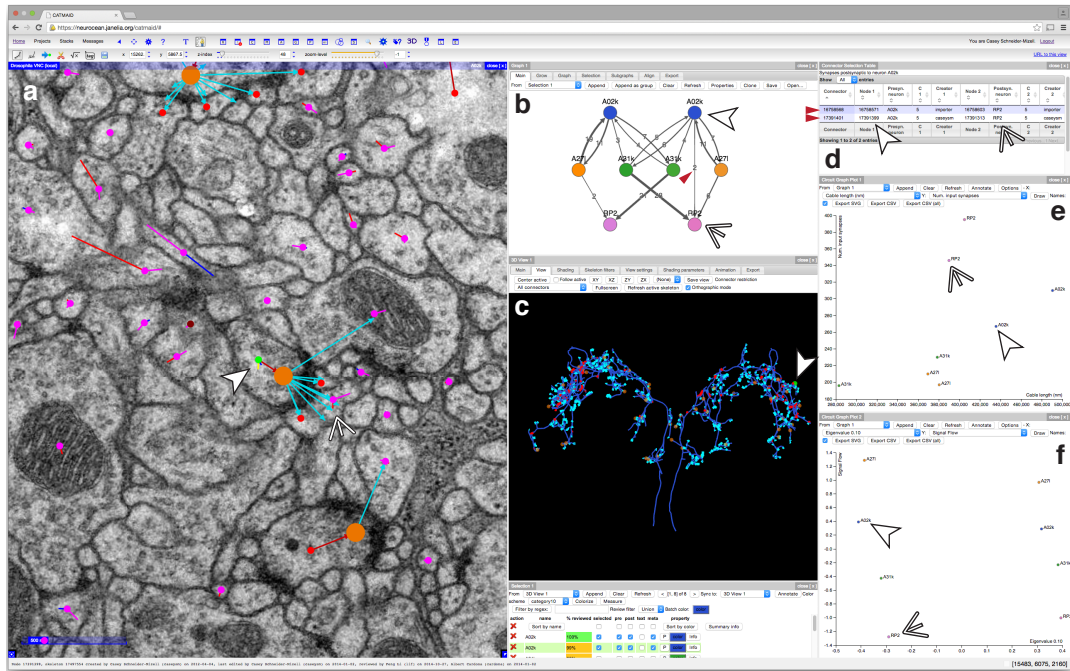
Supplemental Figure 2.14.: **Multiple representations of neurons in CATMAID.** The CATMAID interface is designed to let the user interact with multiple aspects of neurons. Shown here is an annotated screenshot. Other than the image view, all other widgets are pulled up on demand by the user. **a)** An image pane shows the EM data, all reconstructed nodes in the view (purple dots), synapse connector nodes (orange dots), and the active node (green dot, indicated by solid white arrowhead). The current active node belongs to an A02k neuron (name indicated in upper right) and is presynaptic to a node from motoneuron RP2 indicated by the thin arrowhead. **b)** Graph representation of a collection of six neurons, including the selected pair. **c)** Homologous A02k cells shown in a 3d viewer. The active node in the image pane is shown by a green dot in the viewer (indicated by solid arrowhead). Presynaptic and postsynaptic sites are clickable and center the image viewer on that point. **d)** List of synapses between A02k and RP2, represented in the graph widget by an edge (red arrowhead). These sites are clickable, letting the user jump to that location and permit fast reviewing of specific connections. **e–f)** Plots of various properties of the six neurons shown in **b)**. Features that can be calculated on demand include morphological quantities (**e**) and graph theoretic quantities of the network between displayed neurons (**f**). This allows rapid comparison of different properties of homologous neurons and circuits.
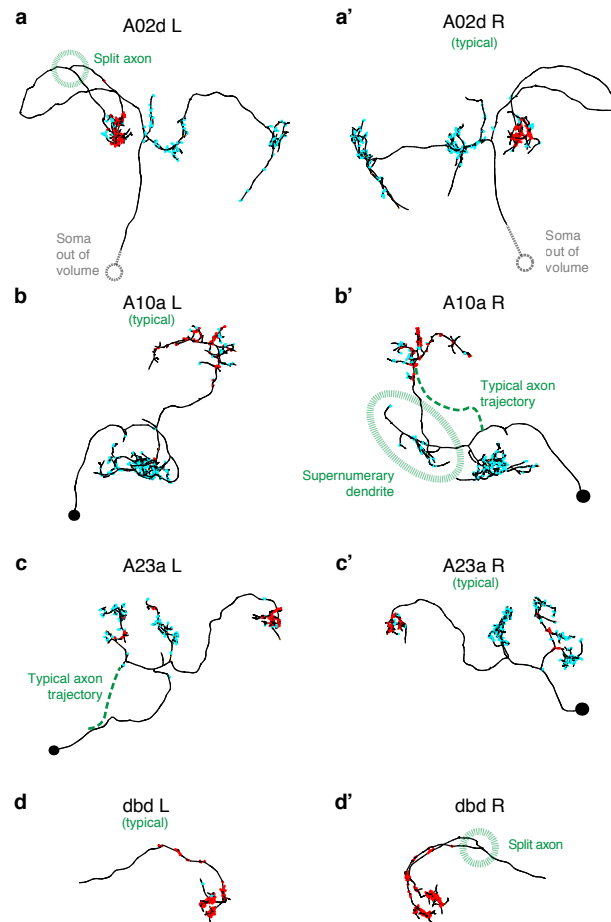
Supplemental Figure 2.15.: **Developmental deviants.** Four pairs of left and right homologs in posterior view, where one of the pairs (canonical) conforms with the arbor shape found in light microscopy (not shown) and the other presents deviations. There are two cases of an early split of the axon (A02d_a3l and dbd_a3r; dotted circle marks the split); normally the split would occur at the proprioceptive domain (where the output synapses are, in red). A10a_a3r presents a misrouted axon that reaches the correct target area (and connects to the same neuron types as the canonical homolog does) but sprouts an ectopic, supernumerary dendrite along the way (which accounts for 14 input synapses out of 178 total). A23a_a3l presents a correct arbor but the path to the soma is different. Connectivity-wise, deviant neurons are very similar to their canonical homologs.
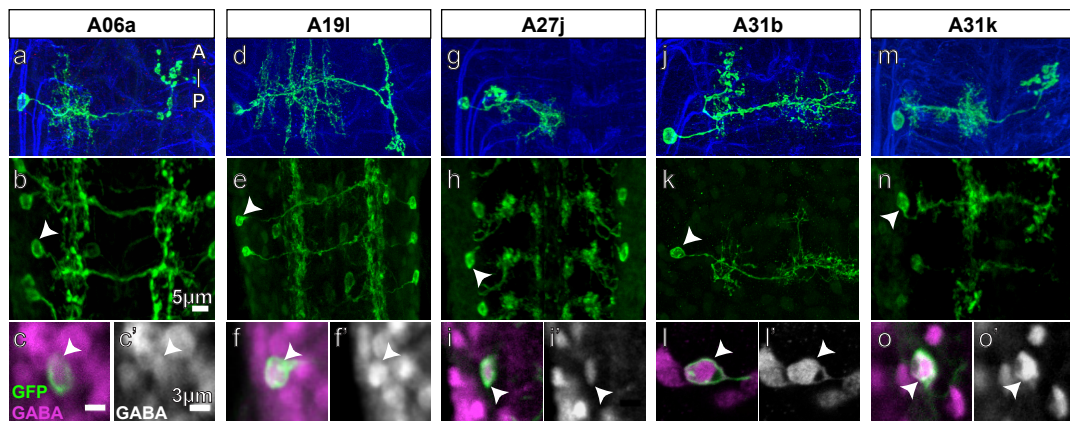
Supplemental Figure 2.16.: **GABAergic interneurons. a-o** Dorsal views. **a, d, g, j, m** Projections of fluorescently labeled single-cell clones of identified neurons (courtesy of James W. Truman, HHMI Janelia Research Campus). **b, e, h, k, n** Dorsal views of projections of parent lines used to generate single-cell clones, expressing myr::GFP. **c, f, i, l, o** Single z-plane at high magnification of cell indicated by arrowhead in (b, e, h, k, n) showing immunoreactivity to anti-GABA (magenta) and GFP (green). **c', f', i', l', o'** Same view as (c, f, i, l, o) only showing the anti-GABA channel in grayscale.

Supplemental Figure 2.17.: **Neurons are distributed throughout the complete range of possible segregation indices.** Top, plot of segregation index vs. cable length for the 39 pairs of identified neurons. Neurons span the full range of the segregation index. Some neurons present completely unsegregated arbors like the serotonergic neurons A26d, with large arbors that present a mixture of inputs and outputs throughout; or compact neurons that are also fully mixed like A27e. Other neurons present intermediate segregation index, which generally takes one of two forms: purely postsynaptic dendrites with mixed axons (i.e. axons that receive a number of inputs) like A08a, and neurons that additionally present a small amount of dendro-dendritic output synapses (e.g. A23a). Finally, some neurons present purely segregated arbors, with dendrites with only postsynaptic sites and axons with only presynaptic sites, like A19l. At bottom left of the chart, the majority of neurons are intersegmental premotoneurons (partial axonic arbors from neurons present in segments not in the imaged EM volume). Note that neurons A27e (both sides) and A08_a3r have their soma outside the imaged EM volume. SI, segregation index; red dots: neurons presynaptic to aCC or RP2; cyan dots: neurons postsynaptic to dbd; violet dots: neurons both postsynaptic to dbd and presynaptic to aCC or RP2; yellow dots: motoneurons, proprioceptive axons or serotonergic neurons not downstream of dbd. The complete set of measurements is available in Supplemental Measurements in CSV format.

Supplemental Figure 2.18.: **Four independent reconstructions of a six neuron circuit. a)** Morphology of the six neurons used for the comparison. The arbors of these neurons were reconstructed independently four times, and used for generating a consensus skeleton for each arbor using the RESCOP method. Branches found only on the consensus and only in the CATMAID approaches are indicated. Neurons 1 and 2 are A02k, 3 and 4 are A31k, and 5 and 6 are A27l. **b)** Graphs of each of the four independent reconstructions. Notice that all four individuals agree on almost all edges with similar amounts of synapses per edge, except in a missing edge for two tracers between A02k_a3l and A31k_a3r.
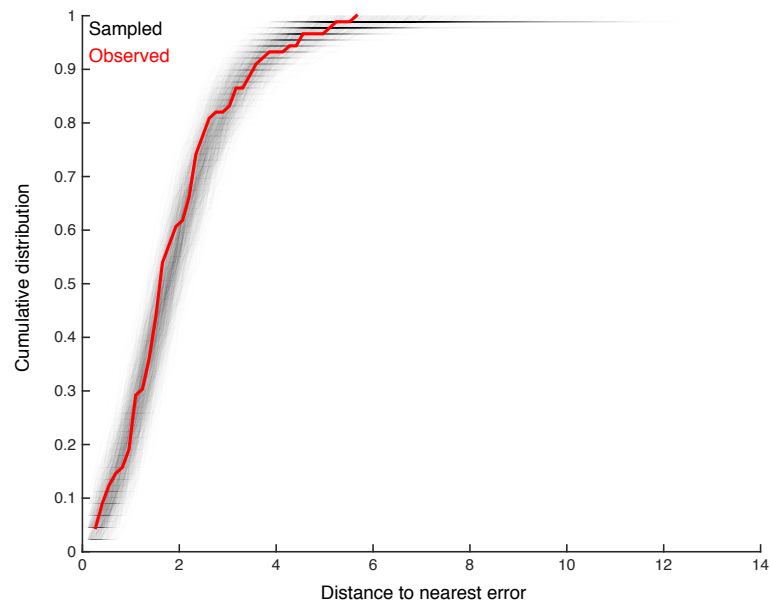
Supplemental Figure 2.19.: **Observed errors are distributed similarly to randomly sampled points.** We look for clustering in the spatial distribution of errors found by comparison with multiple independent reconstructions. For each of the 89 missed branches, we computed the distance to the nearest other error. The cumulative distribution of such distances is shown in red. We compared this distribution to that obtained by randomly sampling 89 nodes from twigs across all six neurons, disallowing two nodes from the same twig. Distributions for each of the 1000 samples are shown in gray. Only 65/1000 differ from the observed distribution in a two-sample Komologorov-Smirnov test with a p-value < 0.05. We thus conclude that spatial clustering of errors is minimal.

3

QUANTITATIVE DIFFERENCES OF WIRING DIAGRAMS ACROSS DEVELOPMENT OF DROSOPHILA

ABSTRACT

How do neural circuits change when their neurons grow in size and complexity? Here, we investigated this question using the *Drosophila melanogaster* as a model system that provides the unique advantage of a nervous system with identifiable neurons. Single neurons are trackable over time and across animals, thus allowing to compare identified circuits througout multiple stages of larval development. We have chosen the nociceptive neural circuit in the larva ventral nerve cord as a model. We used large-scale serial-section transmission electron microscopy to generate EM volumes of the earliest stage (2 hours post hatching first instar) and latest stage (wandering third instar). In all volumes, we reconstructed the identifiable axon terminals of multi-dendritic class IV polymodal nociceptors and mapped all the neurons that are postsynaptic to these sensory axonal terminal within the ventral nerve cord. We found that at both stages, nociceptors connect centrally to sets of local, ascending and descending interneurons. The synaptic organization of the network appears to be established at the earliest stage and persists to the latest stage with 3-4-fold increase in synaptic numbers with only a few differences. Despite cell-type-specific differences in growth, dendrites of postsynaptic interneurons maintain an almost constant proportion of nociceptive inputs throughout development. The circuits of two early stage individuals displayed a high-degree of similarity in morphological and synaptic properties. Additionally, dendritic arbors of local interneurons show preferred synaptic connectivity to different subsets of the somatotopically organized nociceptors. This preference is cell-type-specific and is more refined at late stages, suggesting the emergence of receptive-field-like postsynaptic integration of noxious signals.

3.1 INTRODUCTION

As animals mature and develop, their brain needs to support increasingly complex sensory and motor capabilities. With the increase in body size, most organs grow in size by adding more of their elementary cell types in order to keep fulfilling their functional role. An exception is the brain where the connectivity patterns of neurons, their number, size and complexity may need to change according to new information-processing requirements of the organism (Purves 1988).

Considerable diversity exists across animals in how neural circuits scale in development. One strategy is to increase the size and complexity of neuronal arbors. These structural changes are accompanied by the regulation of formation, elimination and maintenance of synapses between neuron types. For instance, larger postsynaptic arbors need to receive more presynaptic inputs to maintain sufficient synaptic drive to effect larger muscles (Zwart et al. 2013). Another circuit scaling strategy is proliferate neurons of a particular type while the size and complexity of arbors and synaptic numbers may remain similar (Wittenberg et al. 2007). In both strategies, synaptic connectivity patterns need to be modified in development for new behaviors to emerge without compromising existing functional circuitry.

In many invertebrate neural systems, the number of neurons remains fixed throughout larval development. Single neurons derive from stereotyped cell lineages and are uniquely identified in individuals based on their morphology and cell body location (Skeath et al. 2003). These systems, such as *Drosophila* larva used here, provide an unique opportunity to study the changes in synaptic circuits throughout development.

Recent approaches measure synaptic connectivity using novel genetic or viral-based methods to label specific neurons and their putative synapses (Feinberg et al. 2008). Interindividual variability and developmental changes of putative synapses was measured between identified neurons in motor networks of *Drosophila* larva (Couton et al. 2015), and between two identified neurons in *C. elegans* across several days (Desbois et al. 2015). Results of both studies found increased numbers of putative synapses in identified connections of the wiring diagram across larval development. These studies revealed an unexpected degree of interindividual variability for supposedly hard-wired circuits.

However, volumetric electron microscopy (EM) is the only method that can capture the synaptic connectivity of entire neural circuits. So far, we are not aware of any study that mapped the development of synaptic circuits with complete neurons using EM. Recently developed circuit mapping tools for *Drosophila* connectomics enable targeted reconstruction of circuits of interests in very large EM volumes (Schneider-Mizell et al. 2015). Here, we have used serial-section transmission electron microscopy (ssTEM) to generate a volume of *Drosophila* larva ventral nerve cord (VNC) at late, third-instar developmental stage. We further reused existing EM volumes of larval VNC at early, first-instar stages (Ohyama et al. 2015; Schneider-Mizell et al. 2015).

As a model of a developing circuit with identified neurons, we sparsely mapped the identifiable polymodal nociceptors axon terminals (class IV md) and their postsy-

naptic synaptic networks at both early and late developmental stages. The class IV nociceptors have been extensively used as targets for molecular, morphological and behavioral studies (Grueber et al. 2002; Tracey et al. 2003; Hwang et al. 2007; Grueber et al. 2007; Jan et al. 2010). However, little is known about their postsynaptic target interneurons within the central nervous system. We describe here for the first time the complete structure of the circuit and its identified neurons, and then compare the neuroanatomical and connectivity properties of the postsynaptic network across developmental stages. We found that despite considerable quantitative difference in arbor size and synapse counts the overall structure and key properties of the networks are preserved throughout development.

3.2 RESULTS

Wiring diagram of nociceptive neural circuits

We mapped nociceptive neural circuits in three volumetric ssTEM datasets. Two datasets were acquired at first-instar stage (0-24 hr after larval hatching): Dataset L1a covers the whole central nervous system, and dataset L1b covers abdominal segments A2-A3 of the ventral nerve cord (VNC). The third dataset L3 was acquired at the third-instar stage (48-96+ hr after larval hatching) and covers VNC segments A2-A5. We focused the reconstruction of nociceptive circuitry to segment A1 in L1a, and A3 in both L1b and L3 (Supplementary Fig. 3.1).
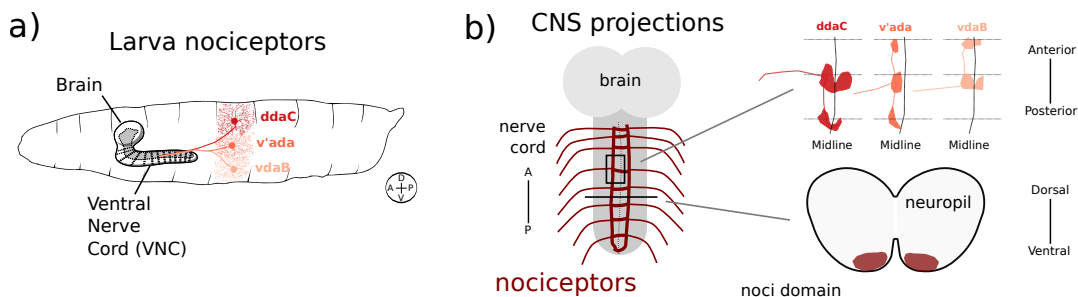
Figure 3.1.: **Nociceptive system of *Drosophila melanogaster* larva** a) Three types of nociceptor sensory cells, the multi-dendritic class IV cells, cover the larva body-wall with their dendritic arborization without overlap. In each hemisegment, nociceptor types are organized somatotopically from dorsal with *ddaC*, to lateral with *v'ada* to ventral with *vdaB*. b) Nociceptors project to the central nervous system to a vento-medial domain in the ventral nerve cord. Body wall segments correspond to segmental subdivisions in the ventral nerve cord. The three types of nociceptors have distinguishable axonal termination profiles which allows to identify them uniquely. Sensory axons of *ddaC* project posteriorly and crosses the midline, *v'ada* projects both anteriorly and posteriorly without crossing the midline, and *vdaB* crosses the midline and projects anteriorly.

We first mapped the axonal terminals of the class IV multi-dendritic nociceptors in each hemisegment and dataset (Figure 3.1 and 3.2). The morphology of each of the three somatotopically-organized nociceptors subtypes has been described peripherally and centrally at light-microscopy levels (Merritt et al. 1995; Grueber

et al. 2007). The described morphology and projection nerves were sufficient to unambiguously distinguish dorsal *ddaC*, lateral *v'ada*, and ventral *vdaB* nociceptors in all datasets (Figure 3.1).

We then mapped the postsynaptic interneurons of the set of six nociceptors in each abdominal segment in each dataset according to the methodology described in Schneider-Mizell et al. 2015. The overall projection patterns of postsynaptic interneurons groups them into three classes: local interneurons extent their dendritic arbors in one or two VNC segments. Ascending projection interneurons integrate inputs from more than two segments and project beyond the local segment. And descending interneurons have distant dendritic arbors beyond the VNC and both pre- and postsynaptic connections to nociceptors in local segments in the VNC. The proportion of nociceptor synaptic outputs diverges with 60% into local, 35% into ascending and 5% into descending pathways (Figure 3.2).

Using the light-microscopy-based neuron catalog of identified ventral nerve cord interneurons (Jim Truman et al., *unpublished*, also see Method in Section 3.4), we identified in each dataset the interneurons of these pathways by matching their distinct arbor morphology, primary neurite bundle, synaptic distribution and cell body location. In total, we identified across all datasets the local interneurons Down-and-back-1 (A09l), Chair-1 (A10a), Basin-2 (A09a), Basin-4 (A09c), Basin-1 (A09b), Pseudolooper-3 (A02m), Pseudolooper-4 (A02n), A01l (Figure 3.3); the ascending interneurons A08n, TePn19, A02o, A09e, A09o, TePn05 (Supplementary Fig. 3.10-3.15); and one type of descending interneuron SelN138 (Supplementary Fig. 3.16).

Comparison of network architecture between developmental stages

We found that the network topology, i.e. the connectivity to identified postsynaptic interneurons, is preserved between early and late stage nociceptive networks. The synaptic connections to local, ascending and descending interneurons in the first-instar dataset L1a could be recovered in third-instar dataset L3 circuitry (Supplementary Fig. 3.6 - 3.9). There was one exception, namely local interneuron A01l in dataset L3 was not found to be connected to nociceptors in the first-instar dataset.

In both first-instar datasets, the same set of postsynaptic interneurons were reproducibly found in dataset L1a (segment A1) and L1b (segment A3) with one exception. Due to the limited nerve cord volume in dataset L1b, the ascending interneuron TePn05 could not be unambiguously identified. Several short intersegmental processes that remained unidentified are candidate fragments for TePn05. Otherwise, the network topology was highly reproducible between the two first-instar stage networks (Supplementary Fig. 3.7,3.8). The cell-type-based connectivity between nociceptor axon terminals and their postsynaptic interneurons appears to be established at an early larval stage and persists to late stages without major changes of the presence of identified connections.

In order to quantitatively characterize the changes of synaptic connectivity in these networks, we confined our further analysis to a set of six local cell types. The dendritic arborization of these local interneurons are fully contained in all three datasets, and thus allow to quantify synaptic counts with nociceptors in an
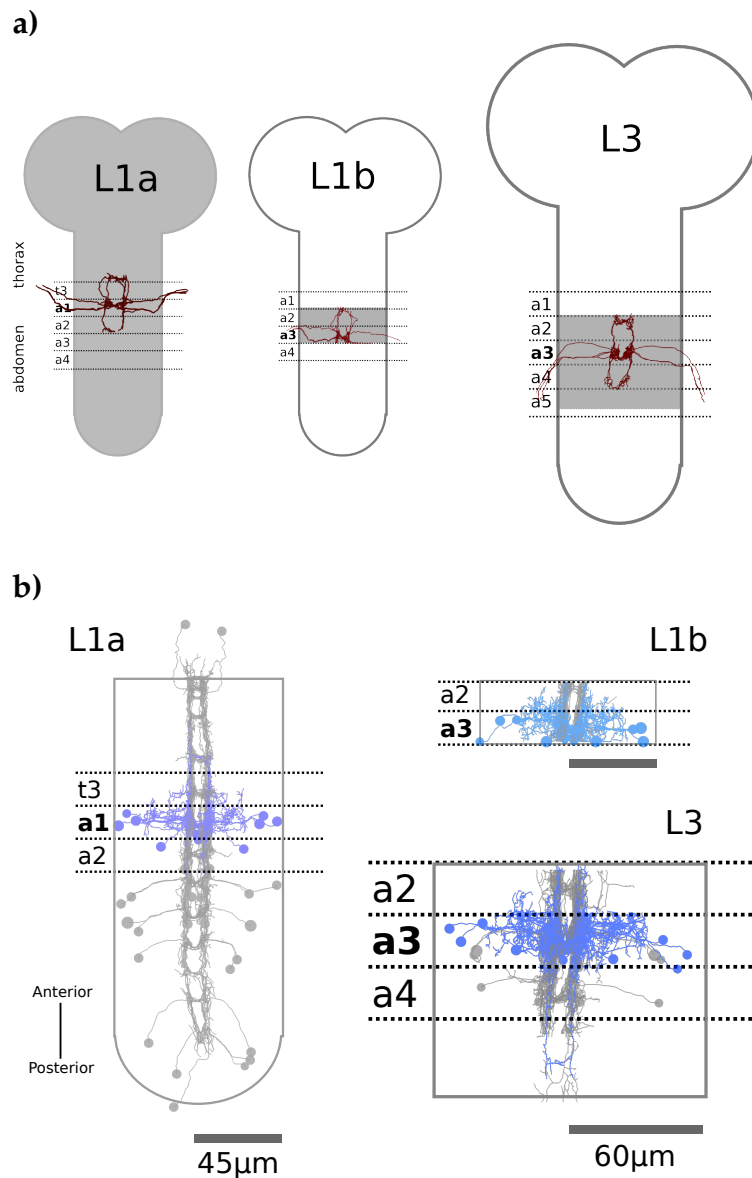
Figure 3.2.: **First-order network mapped in three datsets** a) Reconstructed nociceptor axons in the three datasets. Gray background shows extent of the EM volumes covering the central nervous system. b) Mapped first-order network downstream of nociceptor. Local interneurons in blue, ascending and descending interneurons in grey.

unbiased way (Supplementary Fig. 3.1). We selected Down-and-back-1 (blue), Chair-1 (purple), Basin-2 (bright green), Basin-4 (dark green), Pseudolooper-3 (red) and Pseudolooper-4 (red) (Figure 3.3). The majority of synapses in this small subcircuit are the axo-dendritic connections from nociceptors on which we focus our analysis first. Additional axo-dendritic and dendro-dendritic synaptic connections are present in this subcircuit that are also analyzed in more detail below.
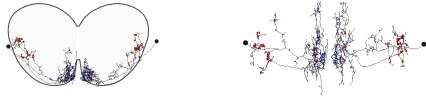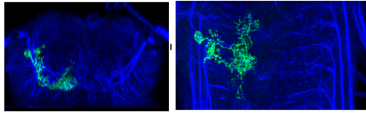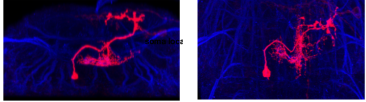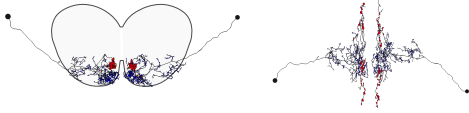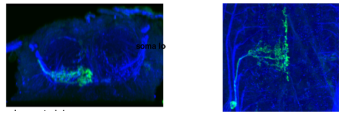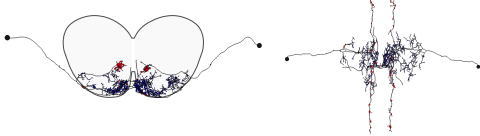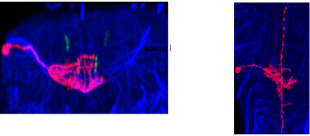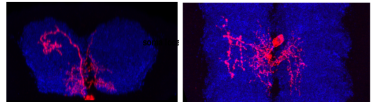
EM Reconstruction	LM Neuron Catalog	GAL4 Lines

Figure 3.3.: **Identified local interneurons connected to nociceptors shown in EM and LM** a) Morphology of the selected five local cell types b) Single neuron morphology based on FLP-out approach for stochastic single-cell labeling (Neuron Catalog of Jim Truman, *unpublished*; Nern et al. 2015). c) GAL4 lines which contain expression of local interneurons (Jenett et al. 2012)

We quantitatively assessed the synaptic connection strength, measured as the absolute number of synapses, of the axo-dendritic connections from nociceptors to local interneurons. Despite the limited sample size, trends in the variability and scaling of synaptic counts are readily apparent. We found some level of variability in the absolute synaptic counts between homologous neurons in all datasets as well as between both first-instar datasets L1a and L1b (Figure 3.4). The observed variability seems to be consistent with the variability observed in identified connections in the larval motor networks at early larval stage (Couton et al. 2015). Our data suggests no drastic reduction of the left-right difference of synaptic counts in homologous neurons in the late stage network consistent across all considered cell types.

The synaptic counts scale from early to late stage across all cell types. The scaling of synaptic count from L1 to L3 is generally about 3-fold, with 50 or less synapses in L1, and 150 or more synapses in L3 (Figure 3.4). With an expection of the Pseudoloopers who receive on the order of 15 synapses from any of the nociceptors at early stage and scale 4-fold to about 60 synapses in L3. This level scaling of synaptic counts

across development is consistent with the scaling in the number of putative synapses found in the study by Couton et al. 2015.
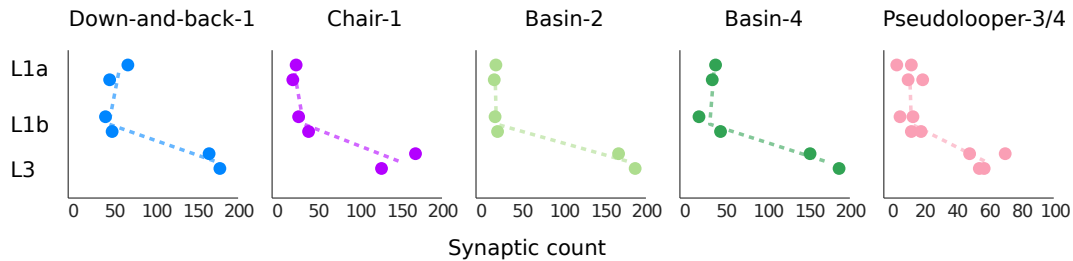
Figure 3.4.: **Absolute synaptic count of axo-dendritic connections from nociceptors to local interneurons** Individual data points represent individual instances of local interneuron types in two first-instar datasets (L1a and L1b) and one third-instar dataset (L3). For better visibility, the upper points denote the left instance, the lower point the right instance for each dataset. Dashed line represents the change in the mean synaptic count for each pair of cells. The more vertical the orientation of the dashed line, the more consistent are the synaptic counts. Counts between L1a and L1b are very similar and scale in number in the L3 dataset.

The local interneurons receive nociceptor inputs not only from the nociceptors of the same segment. The overlap of the dendritic fields with axonal terminals of nociceptors from adjacent segments suggests that these also contribute synapses to the dendritic arbors of local interneurons (Figure 3.1). In order to quantify their contribution, we reconstructed adjacent nociceptors axonal terminal in L1a (thoracic segment T3, abdominal segment A2) and L3 (abdominal segments A2 and A3). The *ddaC* nociceptor has a posterior projection, so *ddaC*s in the anterior segment connect to posterior local interneurons (e.g. *ddaC* in A2 in dataset L3 connects to A3 local interneurons, see Supplementary Fig. 3.6). The *vdaB* nociceptor has anterior projections, and so the *vdaB* in the posterior segment connects to anterior local interneurons (e.g. *vdaB* in A4 in dataset L3 connects to A3 local interneurons, see Supplementary Fig. 3.6). The *v'ada* nociceptors arborize in both anterior and posterior segments.

We found that anterior and posterior nociceptors also connect to local interneurons consistent with their overlap. The axonal arbors of nociceptors in adjacent segments are generally smaller compared to the arbor in the local part of the segment and therefore contain fewer presynaptic sites. Considering these nociceptive inputs to dendritic arbors of local interneurons, we also did not find that the synaptic counts of connections on the left and right sides becomes more symmetric at late stages compared to the early stage (Supplementary Fig. 3.6 - 3.8). The particular overlap and midline crossing patterns of nociceptors causes a particular synaptic innervation pattern to the spatially-confined dendritic arbors of local interneurons. We continue to analyze these patterns from the vantage point of the dendritic arbors.

Comparison of cell-type-specific dendritic convergence patterns

The Basin-2 and Basin-4 neurons have been shown to exhibit a preference in the connectivity from dorsal and ventral nociceptors at early stage (Ohyama et al. 2015). Here, we analyze the relative contribution of nociceptors and their subtypes, for the local and adjacent segments, to inputs of dendritic arbors of local interneurons. Since we have reconstructed the full dendritic arbors and all input synapses, we were able to quantify in detail the proportion of synapses that nociceptors contribute. We extend this analysis here to the six local interneuron cell types at two developmental time points.

We found the relative amount of identified input to be characteristic for each cell-type and to be well preserved between individuals of the same stage (Figure 3.5). Local interneurons such as Down-and-back-1, Chair-1 or the Pseudoloopers receive between 30-50% nociceptor input. Basin-2 and Basin-4 neurons only 10% and 30% respectively. Inputs from local interneurons contributes to 20% of the total dendritic inputs in Chair-1 and Pseudoloopers. All local interneurons also receive inputs from adjacent nociceptors with a smaller proportion than from the local segment, consistent with the smaller overlap described above.

Remarkably, after considerable neuronal arbor growth and formation of synapses from early to late stage, the proportion of dendritic convergence remains very similar at both developmental stages (Figure 3.5). The amount of axo-dendritic and dendro-dendritic between local interneurons decreases slightly and consistently from early to late stage. We found that the maximal difference of the convergence fraction between left and right homologous neurons in L3 is smaller compared to the maximum difference in L1.

This suggests that over development the synaptic count normalized by the total amount of input of the postsynaptic neuron becomes symmetric, and not absolute synaptic count between identified neurons.

Nociceptor subtype-specific targeting of dendritic arbors

We evaluated to what degree the dendrites of local interneurons connect to specific subtypes of nociceptors, and if so, how these patterns change in development. The peripheral dendritic fields of nociceptors tile the whole larval body wall non-overlappingly (Grueber et al. 2002), and the position of a single nociceptor along the dorso-ventral extent of the body wall determines its type (Schrader et al. 2000; Grueber et al. 2007). This somatotopic organization distinguishes dorsal *ddaC*, lateral *v'ada* and ventral *vdaB* nociceptor types.

Type-specific connectivity to postsynaptic local interneurons may thus indicate selectivity to the spatial origin of noxious stimuli. Additionally, both *ddaC* and *vdaB* cross the VNC midline and connect contralaterally to interneurons which might confer signals of the lateralized origin of noxious stimuli. For further analysis, we therefore separated nociceptor-type specific connectivity to local interneurons into their ipsilateral and their contralateral contribution.
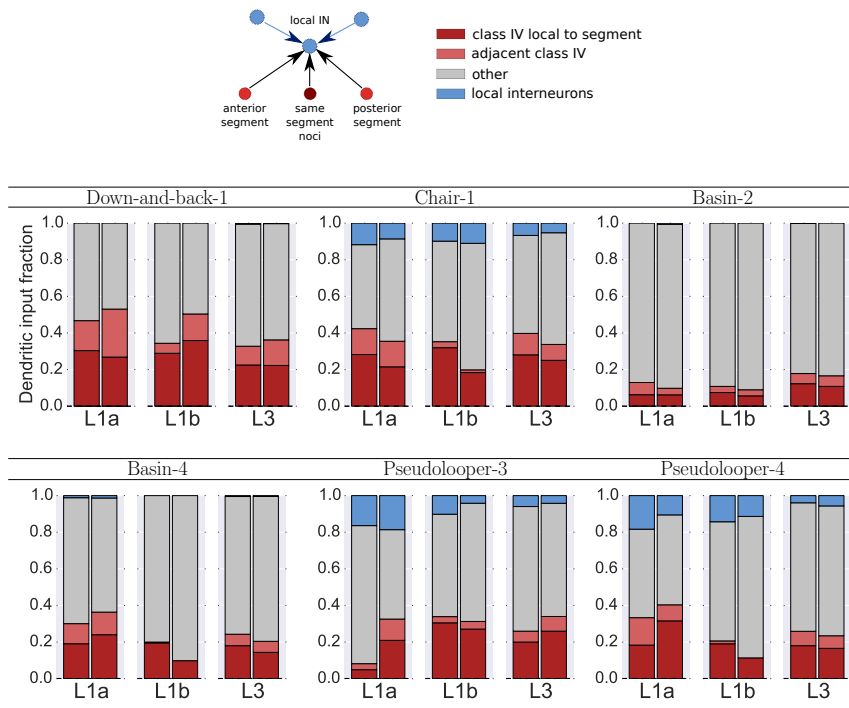
Figure 3.5.: **Convergence of identified inputs onto the dendrites of local interneurons.** The plots show the percentage amount of total dendritic inputs from specificy types of neurons for each local interneuron in the left and right hemisegments of each dataset. Input fraction is from nociceptors of the same, local segments (dark red), from anterior and posterior segments (bright red) or other local interneurons (blue). Despite a large increase in the number of synapses in neurons at the third-instar stage, the relative contribution of synapses from nociceptors remains similar.

We found that each local interneuron type receives preferred inputs from a specific set of nociceptor types at late stage (Figure 3.6). The Down-and-back-1 neurons receive ipsilateral input from all three nociceptor types in similar amounts. Down-and-back-1 dendrites do not cross the midline, so connections with both contralateral *ddaC* and *vdaB* are made through their midline-crossing arbors. No synapses are made with the contralateral *v'ada* which is explained by it not crossing the VNC midline. The Chair-1 neurons receive mostly input from *ddaC* and *vdaB*, with more inputs on the contralateral than on the ipsilateral side (note that the cell body is opposite to the main dendritic arbor). Chair-1 also receives inputs from *v'ada*, but fewer compared to *ddaC* and *vdaB*. The Basin-2 neurons receive mostly from ipsilateral *ddaC* and *v'ada*, and to a lesser degree from contralateral *ddaC*. Basin-2 does not receive inputs from contralateral *v'ada* also because its dendrites do not cross the midline. The Basin-4 neurons receive mostly from ipsilateral *v'ada* and *vdaB*, and due to its midline crossing-dendrite, also from both contralateral *v'ada* and *vdaB* , however to a lesser extent. These biases should be seen as preferences and are exclusive, as for instance both neurons also receive inputs from *vdaB* and *ddaC* respectively. Finally,

both Pseudoloopers receive very symmetric ipsilateral (cell body is on the same side/hemisegment as the class IV axonal terminals) and contralateral (cell body is on the other hemisegment as the class IV axonal terminals) inputs mostly from dorsal *ddaC* and to a lesser degree also from *v'ada* and *vdaB* on both sides.
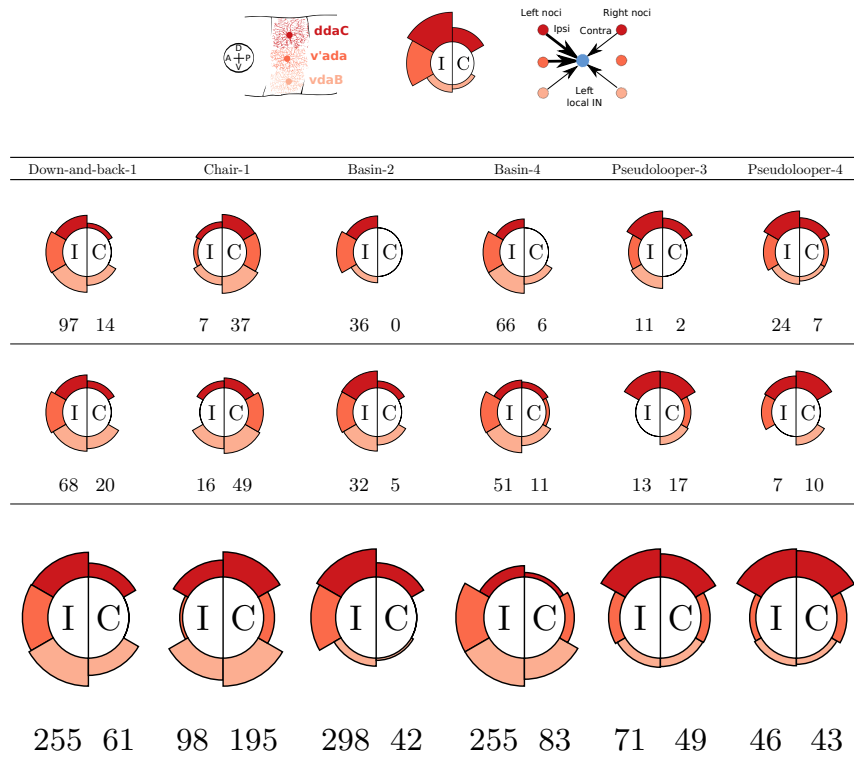
Figure 3.6.: **Dendritic nociceptor-type specific input distribution to local interneurons** The three rows (datasets L1a, L1b, L3) show the input distribution from the three nociceptor types to local interneurons. Distribution from both local interneurons in one abdominal segment are shown summed together. Nociceptor inputs are distinguished into ipsilateral and contralateral depending on the location of the local interneuron cell body. The number below the plot denote the total amount of synapses made from either ipsi- or contralateral nociceptors. The amount of synapses from specific nociceptors (*ddaC*, *v'ada*, *vdaB*) is represented as the area in the plot. Relative contributes are readily apparent. L3 plots are scale in size to reflect the larger amount of total synaptic inputs.

How does this connectivity patterning compare to the first-instar network with fewer nociceptive inputs per dendrite? We may expect that the increased synapse counts at late stage has reduced spurious synapses arising due to developmental noise. Is the signal-to-noise level at an early stage already sufficiently high to distinguish these patterns?

We found that, despite the lower absolute synapse counts at early stage, similar connectivity preferences are discernible in both L1 datasets. Overall, these preferences are consistent with the patterning found in L3. In the L1 datasets, both Basin-2

and Basin-4 neurons exhibit dorso-lateral and ventro-lateral preferences in their nociceptive inputs. Pseudoloopers show a preference to *ddaC*. Down-and-back-1 receive equal amounts inputs from ipsilateral nociceptors. Only Chair-1 is slightly exceptional, because its contralateral inputs are equally distributed across all nociceptors. The lower contribution of synapses from *v'ada* is not seen consistently in the L1 datasets.

Overall, we found a good agreement of synaptic patterning of individuals at the same stage. However, at the lower absolute synaptic counts at early stage, the signal-to-noise level is lower and thus the preferences is not as pronounced. Our data suggests that an increase in signal-to-noise level is possible by averaging synaptic counts for the same neuron types across multiple same-stage datasets. The reproducibility of these patterns in abdominal segment A1 (L1a) and segment A3 (L1b) further suggest that averaging synaptic counts across ventral nerve cord segments in the same individual can also be used to increase signal-to-noise levels. This is also suggested by the small segment-specific variation, at least in abdominal VNC segments, found in the recent study of Couton et al. 2015.

Nociceptor-type-specific connectivity of dendritic compartments

We asked to what extent the nociceptor-specific synaptic inputs are distributed across different dendritic compartments, in particular across backbone and twig. We were interested to see whether the 10%-20% of total synapses on the dendritic backbone is sufficient to exhibit the found connectivity patterning in early and late stages. During the growth of dendritic processes, actin-rich filopodia become dynamically invaded by microtubuli to stabilize neuronal arbors (Hu et al. 2008). At the EM level, we can distinguish microtubuli-filled (backbone) and microtubuli-free, distal dendritic processes (twigs). Recently, we have shown that *Drosophila* neuron's dendrites receive most of their synaptic inputs (80%-90%) on distal twigs (Schneider-Mizell et al. 2015). The remaining inputs are made onto the backbone of the dendrites. Here, we have manually annotated backbone and twig processes of all local interneurons in all datasets.

Figure 3.7 shows the nociceptor-specific connectivity segregated into twig and backbone synapses. Considering backbone synapses at first-instar stage, we find almost no consistency in the patterning (Figure 3.7a). Remarkably, the Basin-2 dendritic arbor backbone with only 3 synapses in L1a and L1b (less than 10% of its total dendritic input) exhibits the characteristic dorso-lateral connection preference. Similarly, both Pseudoloopers dendritic backbones in L1b receive their only 3 synaptic inputs from *ddaC* . For other neuron types, the backbone connectivity in L1 does not clearly recapitulate the characteristic specificity patterns. In L3, however, the backbone connectivity recapitulates the nociceptor-type-specific connectivity patterns very well with its 10% of the total dendritic synapses (Figure 3.7a). Since most of the synaptic inputs are onto twigs, the overall nociceptor-type-specific patterning we found across the whole arbor is recapitulated in twig synapses in L1, as well as in L3 (Figure 3.7b).

Our data suggests that L3 dendritic backbones are sufficient to recapitulate synaptic patterning. Moreover, this indicates that synaptic circuitry confined to only backbones might be a good recapitulation of the complete circuit diagram including synapses made on twigs.

a) Connectivity from nociceptors to dendritic backbones

Down-and-back-1	Chair-1	Basin-2	Basin-4	Pseudolooper-3	Pseudolooper-4
25 1	1 4	3 0	12 2	0 0	1 0
10 3	0 18	3 0	8 1	1 2	2 2
32 11	13 24	42 10	26 5	10 8	1 3
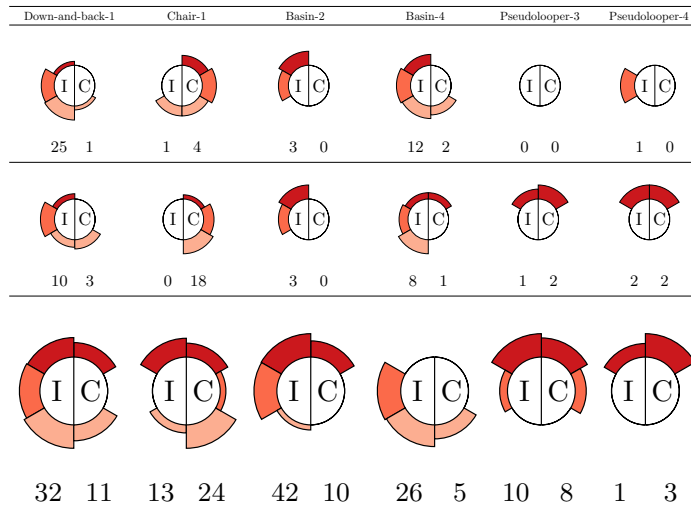
b) Connectivity from nociceptors to dendritic twigs

Down-and-back-1	Chair-1	Basin-2	Basin-4	Pseudolooper-3	Pseudolooper-4
72 13	6 33	33 0	54 4	11 2	23 7
58 17	16 31	29 5	43 10	12 15	5 8
223 50	85 171	256 32	229 78	61 41	45 40
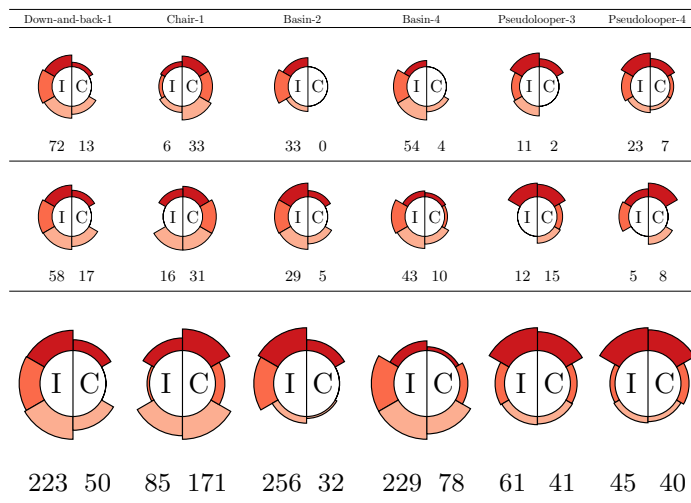
Figure 3.7.: **Dendritic nociceptor-type specific connectivity segregated by backbone and twig inputs** The three rows (L1a, L1b and L3) show the input distribution from the three nociceptor types to local interneurons considering only backbone synapses or twig synapses. a) Synaptic connections exists between nociceptors and local interneuron dendritic backbones for all cell types (except Pseudolooper-3 in L1a). Despite the lower synaptic count, the pattern of selectivity is well preserved in the L3 dataset for all interneurons. The pattern is less obvious in the L1 datasets, and not consistent between L1a and L1b (except for Basin-2). b) Synaptic connections between nociceptor and local interneuron dendritic microtubuli-free processes (twigs). The pattern of nociceptor-type selectivity is better visible in the L1 datasets compared to the backbone, and is very clear in the L3 dataset.

89

Distribution of identified synapses across dendritic arbors

Cell-type-specific distribution of synapses on postsynaptic subarbors have been shown to be important for neuronal function (Branco et al. 2010; Briggman et al. 2011b). Here, we investigated whether nociceptor-type-specific preferences could arise from synapses being preferentially clustered to specific postsynaptic dendritic subarbors. For many neuron types with complex arbor geometry, it is difficult to directly appreciate the cell-type-specific synaptic input distribution from visualizing synapses and their dendritic (skeletonized) morphology in 3D.

We devised a simple visualization method which collapses 3D arbor morphologies to 2D network graphs that preserve arbor topology and points-of-interests such as pre- and postsynaptic locations (see Method in Section 3.4). This 2D presentation of a neuron allows to readily depict the distribution of cell-type-specific synaptic inputs by color-coding respective presynaptic cell-types. For instance, Pseudolooper-3 receives inputs from Basin-2 and all other nociceptors, with a preference in inputs from *ddaC* (Figure 3.8). In this case, two ipsilateral subarbors receive mostly inputs from Basin-2 but no nociceptors, consistent with the spatial segregation of different presynaptic domains. Ipsilateral subarbors with nociceptor inputs show a mixture of nociceptor-types with the expected preference for synapses from *ddaC*. A larger contralateral subarbor, however, extends to both Basin-2 and nociceptor domains and receives mixed inputs without clear segregation.

We found very similar mixing and preference of nociceptor-type inputs by visualizing the 2D graphs of all other local interneurons in all datasets. However, we found no clear subarbor-wise segregation of input types (Supplementary Fig. 3.18 - 3.21). This indicates that postsynaptic neurons do not differentiate distinct presynaptic input types in segregated subarbor levels. The nociceptor-type-specific preferences of postsynaptic cell-types arise from mixing of synaptic input types across whole dendritic arbors and biasing their counts.

a   L3 Pseudolooper-3L

b
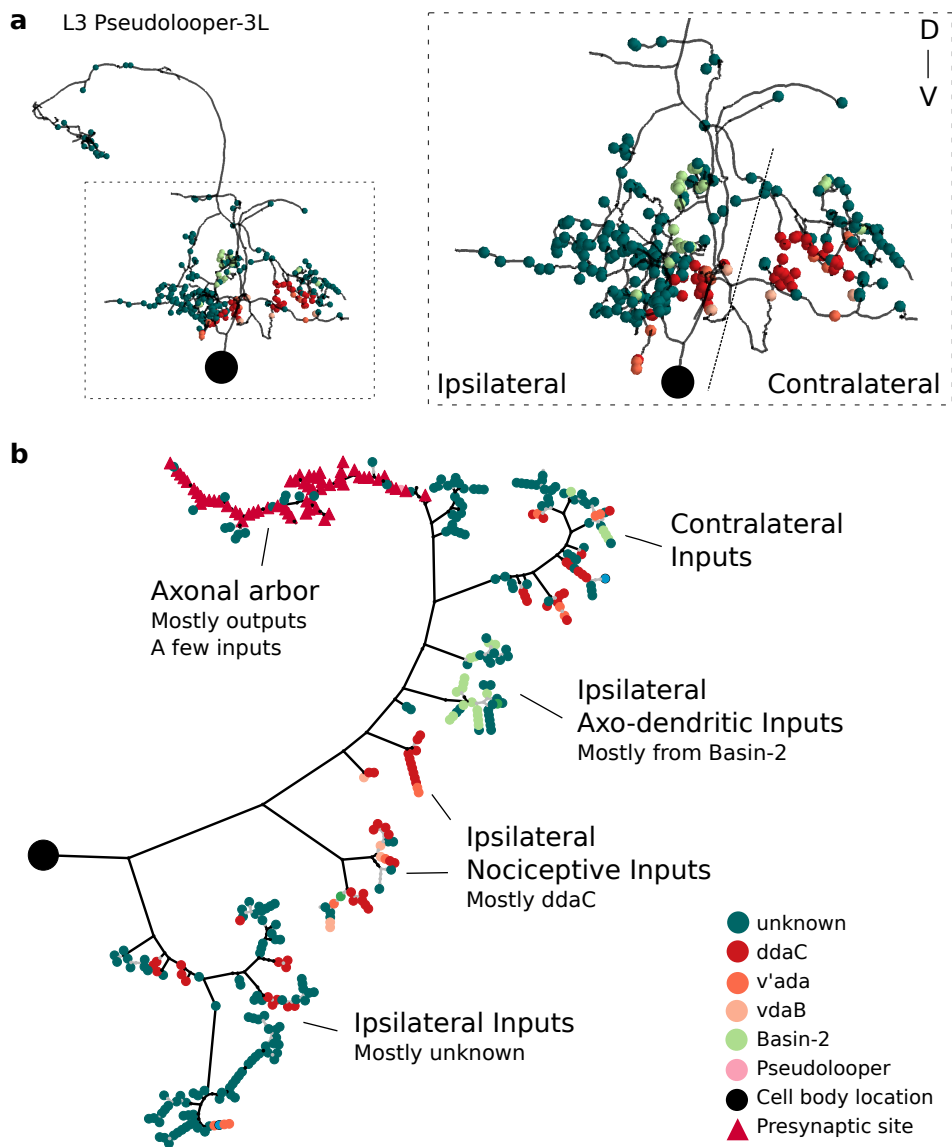
Figure 3.8.: **Arbor morpholgy and synaptic distribution represented topologically**
a) The arbor morphology of Pseudolooper-3L embedded in 3D. Zoomed in
version of the dendritic input field with input sites color-coded by known cell
type. Topological relationships of subarbors are difficult to see in 3D. b) The
same arbor represented as 2D graph. The visualization preserves the topology
of points of interest: input sites (circles) in different colors corresponding to
the input cell type; output sites (triangles). Grey edges denote twigs. Input
sites color-coded by cell type reveal distinct subarbors on ipsi- and contralat-
eral sides. Ordering of major subarbors from close to the cell body towards
the axonal output arbors is readily apparent. This topological representation
might give hints about dendritic integration properties. Edges could be scaled
by electrotonic distance, varying lengths are a consequence of the layouting
algorithm (Cytoscape Organic).

Microcircuit motifs and their development

The relevance of microcircuit motifs has long been recognized to contribute importantly to neural signal processing (Shepherd 1978). We found a few microcircuit motifs in the nociceptive circuitry that we elucidate and compare in the following.

One such motif is the axo-dendritic, feedforward connection from Basin-2 axons to Pseudolooper dendrites. Pseudolooper dendrites receive inputs directly from nociceptors and indirectly though the Basin-2 connection. This feedforward connection is present in both L1 and L3 (Figure 3.9a). In terms of absolute synaptic count, the Basin-2 axon to Pseudolooper dendrite connection increases less than two-fold from L1 to L3. With the increase of Pseudolooper dendritic arbor size and number of synaptic inputs, the relative contribution of the Basin-2 axonal input decreases considerably. The axo-dendritic synapses contribute 15% of a total of 224 dendritic inputs summed across all four Pseudoloopers in L1a, 10% across 211 inputs in L1b, and 4% across 1122 inputs in L3 (Figure 3.5).

We observe that the Basin-2 axonal arbors, unlike nociceptor axonal arbors, do not grow much more in L3 compared to L1. The limited availability of presynaptic axonic sites could explain the reduced potential for novel synapse formation, despite considerable growth of Pseudolooper dendrites. This indicates that differences in scaling of synaptic counts could be explained by differences in scaling of arborizations and the resulting variation in their overlap and potential for synapse formation.

Dendro-dendritic connections have been well-studied in the past decades. They have been found to be involved in microcircuit motifs that mediate processes such as lateral inhibition (Shepherd 2009). We also found two dendro-dendritic connections between identified neurons in the nociceptive network. The Chair-1 dendrites receive dendro-dendritic synaptic input from both Down-and-back-1 and Basin-4 dendrites. Both dendrites host a number of presynaptic sites on their dendritic arbors.

We first visualized the locations of these presynaptic sites using our 2D arbor layouting method to better understand their dendritic distribution (Figure 3.10). In both neurons, we find that presynaptic sites localize to the base of subarbors with exclusive postsynaptic inputs. The localization to the base of subarbors is consistent between left and right homologous neurons in L3, as well as in the homologous neurons in L1 (Supplementary Fig. 3.18, 3.21). The distribution of presynaptic sites suggests local dendritic processing where inputs on dendritic subarbors in the vicinity are integrated and drive local dendritic outputs.
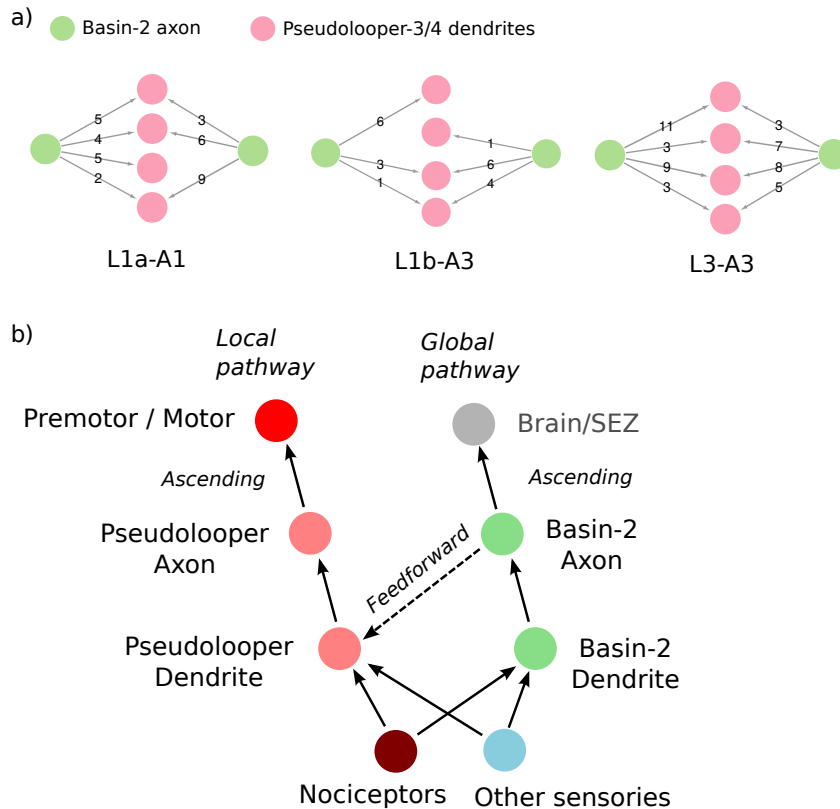
Figure 3.9.: **Axo-dendritic connections from Basin-2 to Pseudoloopers** a) Network diagrams show the axo-dendritic connection from the left and right instances of Basin-2 axons (green) to the Pseudolooper-3 and -4 neuron dendrites (bright red). Synaptic count for each edge is shown. The increase in the synaptic count of any edge from L1 to L3 is less than two-fold. b) The axo-dendritic connection from Basin-2 to Pseudolooper is embedded in a feed-forward connection (dashed lines) that links a global with a local pathway. In Schneider-Mizell et al. 2015, Pseudolooper-3 and -4 were found to target premotor and motor neurons (in particular U1, U2, RP2) within an abdominal ventral nerve cord segment. Pseudoloopers, therefore, establish a direct local pathway from sensory (nociceptive) signals to motor circuitry while also integrating signals, via Basin-2, destined to ascend in a global pathway.

a Down-and-back-1 in L3

▲ Presynaptic sites
● Postsynaptic sites

Left
neuron

Right
neuron

b Basin-4 in L3

Right
neuron

Left
neuron
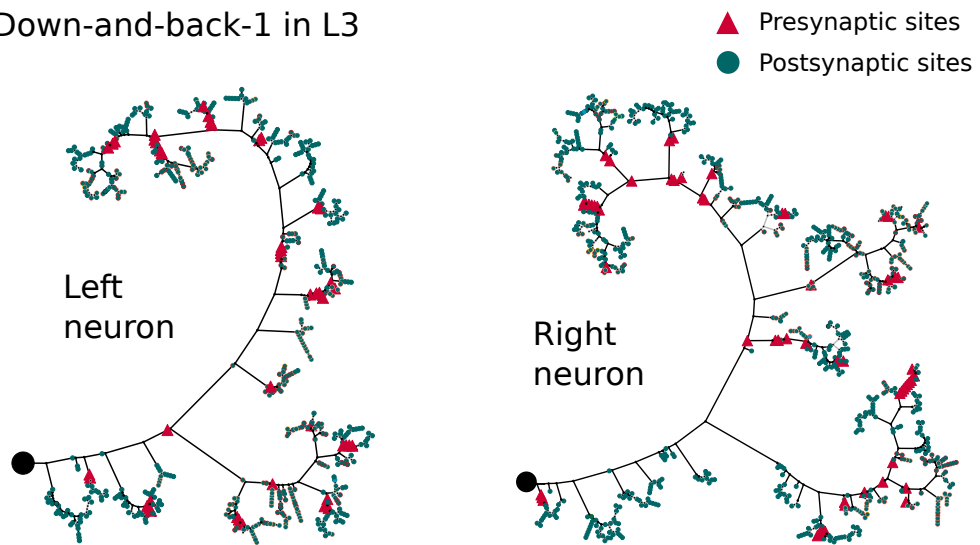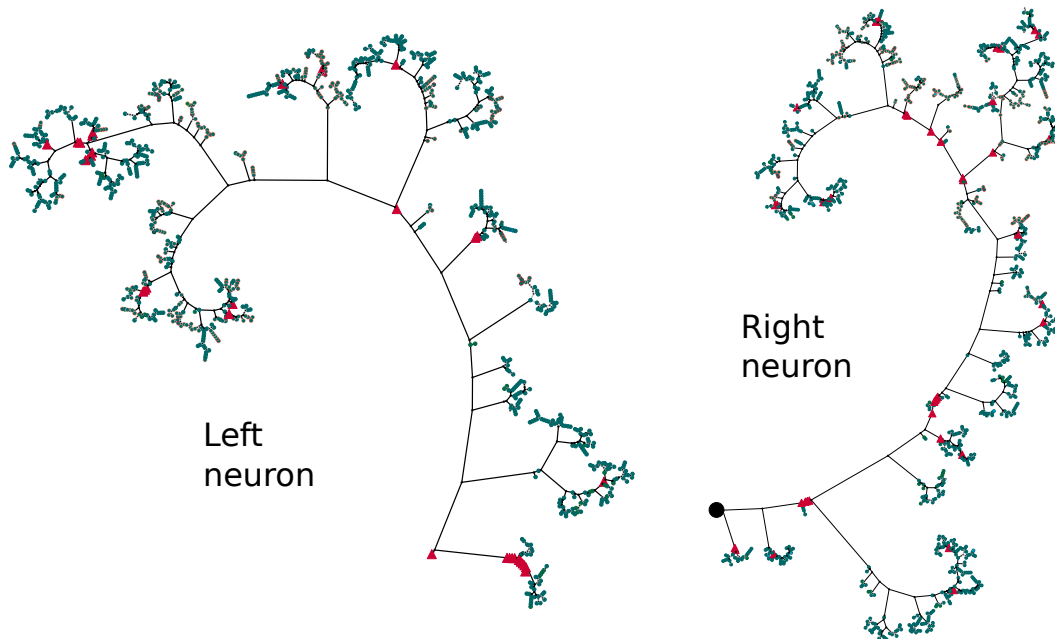
Figure 3.10.: **Distribution of presynaptic sites on two dendritic arbors** Presynaptic
sites are clustered at the base of subarbors on the dendritic arbors of a)
Down-and-back-1 and b) Basin-4 interneurons. The location of these
presynaptic sites at the base of subarbors is reproducible in early stage
larva (data not shown).

The dendro-dendritic connections contribute only a small fraction of the total amount of synapses in the nociceptive network. However, they were reproducibly found in both early and late stage circuits.

From L1 to L3, Down-and-back-1 increases the number of dendritic outputs more than 4-fold, and Basin-4 dendritic outputs increase about 3-fold (Figure 3.11a,b). However, only a small fraction of the total number of dendritic outputs is made from Down-and-back-1 and Basin-4 onto Chair-1. In L3, Down-and-back-1 connects with less than 5%, and Basin-4 with less than 15%, of its dendritic outputs onto Chair-1 dendrites. Both Down-and-back-1 and Basin-4 connections to Chair-1 scale less than two-fold in terms of absolute synaptic counts from L1 to L3 (Figure 3.11c). When accounting for the increased size and postsynaptic counts in the Chair-1 dendrites in L3, this amounts to less than 2% dendritic inputs from Down-and-back-1, and less than 4% from Basin-4.
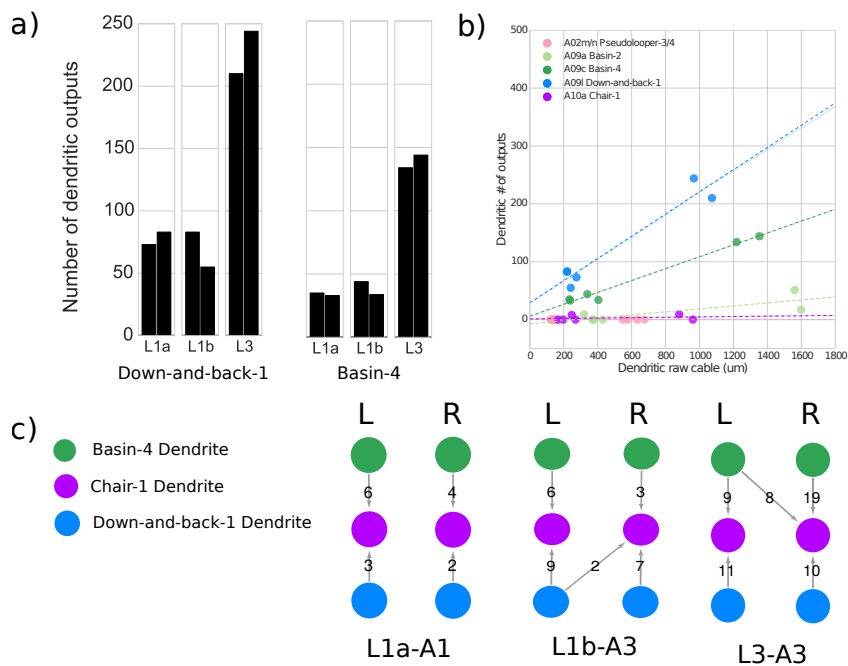
Figure 3.11.: **Dendro-dendritic connections of three local interneurons** a) Total number of dendritic outputs of Down-and-back-1 and Basin-4. The amount is consistent in L1a and L1b, and then increases severalfold in L3. b) Dendritic number of outputs as a function of total dendritic cable length. Down-and-back-1 increases its total cable length less than Basin-4 from L1 to L3. Its increase in the number of dendritic outputs is larger. c) Absolute synaptic counts of dendro-dendritic edges for the left and right instances of the local interneuron. Compared to the total increase in dendritic cable length and synapse number from L1 to L3, the increase in dendro-dendritic synapse count is modest.

Reciprocal connectivity motif

We found a reciprocal connectivity motif between nociceptors and the descending interneuron SelN138 (Supplementary Fig. 3.16). The pair of SelN138 interneurons connect with their descending arbors to all three types of nociceptive axonal terminals along the whole length of the ventral nerve cord. The descending arbor targets nociceptors in each segment and also receive reciprocal input from all types of nociceptors. In addition to these axo-axonic connections, the extensive dendritic arborizations of SelN138 in the subesophageal zone SEZ receives other synaptic inputs (Supplementary Fig. 3.16). The axo-axonic connections are reproducibly found in both early and late stage larvae (Figure 3.12c).

We analyzed how postsynaptic sites distribute on nociceptor axons and found that SelN138 synapses are made across the whole presynaptic axonal arbors without particular clustering in both L1 and L3 datasets (Figure 3.12a). In terms of absolute synaptic counts, we find on average that every synaptic contact made between nociceptor and SelN138 is reciprocated with about two contacts from SelN138. Remarkably, this 2:1 ratio of reciprocal connectivity persists from L1 to L3, with five- to six-fold increase in absolute synaptic counts (Figure 3.12d). A particular feature of SelN138 neurons at presynaptic sites are the co-occurences of small vesicles together with larger, dense-core vesicles (Figure 3.12b). This further suggests a modulatory role of the SelN138 interneuron with regard to nociceptor axon terminals.
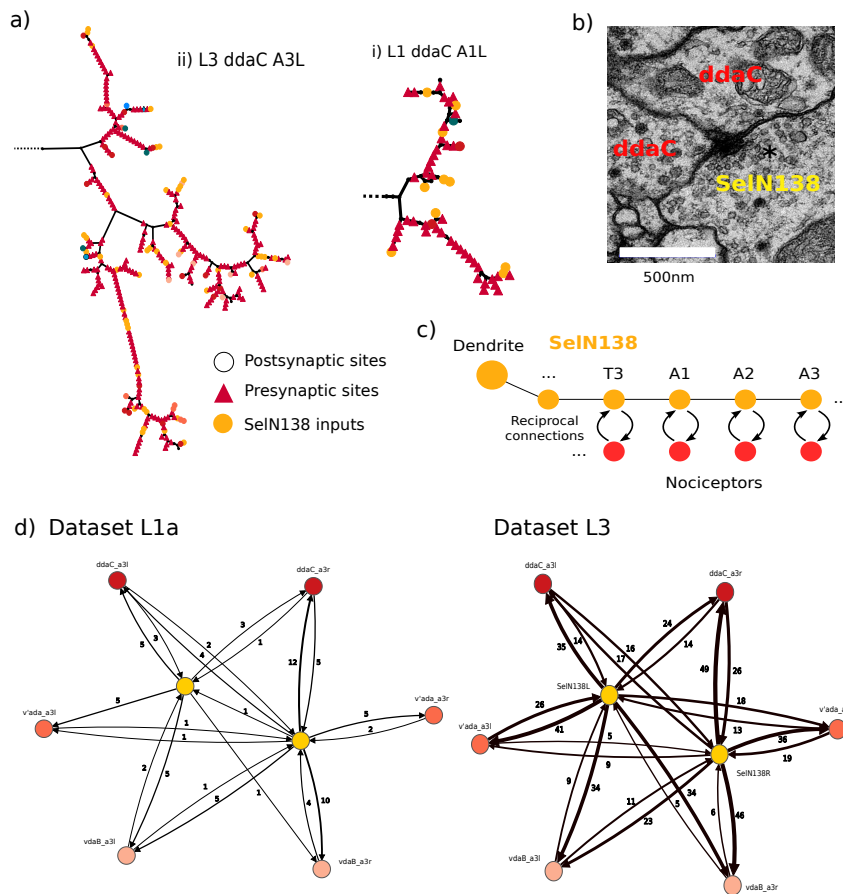
Figure 3.12.: **Nociceptive axon terminals reciprocally connected to descending SelN138 interneuron** a) Sensory axon terminals receive inputs from descending SelN138 interneuron. Topological 2D representation of ddaC shown in L1 and L3. SelN138 postsynaptic sites shown in yellow circle, presynaptic sites shown with red cycles. Other postsynaptic sites shown as circles. b) EM image in L3 depicting presynaptic site (*) of SelN138 interneuron, targeting two *ddaC* neurons. Dense-core vesicles are visible at both presynaptic and postsynaptic locations. c) Schematic diagram depicting reciprocal connectivity of nociceptors with descending interneuron in consecutive segments. The SelN138 has an additional dendrite in the subesophageal zone (SEZ). d) Reciprocal synaptic connectivity between SelN138 and individual nociceptors in datasets L1a and L3.

3.3 DISCUSSION

We have mapped three *Drosophila* larva nociceptive synaptic circuits in the ventral nerve cord and found that the overall circuit architecture is preserved from early to late stage. Two circuits at the same stage exhibit high stereotypy at the early stage with similar numbers of synaptic counts. The polymodal multi-dendritic nociceptors connect at both stages to a set of at least fourteen neuron types segregated into distinct local, ascending and descending pathways. We found neuron-type-specific variation in arbor growth reflected in differences in scaling of synaptic counts of identified network connections. Although synapse counts increases considerably from early to late stage, the proportion of nociceptive inputs converging onto postsynaptic dendrites of local interneurons is well preserved at both considered stages.

Behaviors mediated by the nociceptive circuit change from early, first-instar to late, third-instar developmental stages. These distinct larval nocifensive behaviors were quantitatively assessed using thermal and optogenetic assays across development (Sulkowski et al. 2011). The authors have found that early stage larvae do not perform the stereotypical corkscrew-like roll which is a characteristic escape behavior in late stage larvae. This escape-roll behavior is mediated by polymodal nociceptors through transient noxious thermal or mechanical stimuli (Tracey et al. 2003; Zhong et al. 2010). The lack of rolling behaviors at early stage was hypothesized to be due to imature sensory or motor systems. Our connectivity data suggests that on the sensory side, all major elements of the nociceptive circuit and their relative connection strengths are already in place early in the larval life cycle. From a purely structural standpoint, the nociceptive circuitry should be able to transduce nociceptive signals and activate downstream circuits. Therefore, downstream circuits or motor systems might not be able to generate the stereotyped motor activation patterns required for rolling at an early, first-instar stage. An alternative hypothesis is that the body wall muscles and the neuromuscular junction might to be fully matured to adequately transform motor neuron activation into this behavioral escape response.

The study by Ohyama et al. 2015 has provided evidence that mapping ultrastructural circuitry at an early stage is sufficient to inform the study of late stage behavioral phenotypes. The authors focused on an ascending pathway where nociceptor stimulation increases the likelihood of escape rolling when chordotonal sensory organs are co-activated. The early stage circuit suggested Basin neurons as convergence point of both nociceptor and chordotonal signals in this pathway. Neurogenetic behavioral and optical imaging experiments carried out in late stage larvae were consistent with the early stage circuit maps. Here, we confirmed that the nociceptor targets of this pathway, the Basin-2 and Basin-4 neurons, remain strongly connected in the late stage circuitry. Our data also suggests that early stage circuits at the cell-type level are adequate to imply late stage connectivity patterns, and can thus serve as a proxy to inform functional and behavioral experiments at late stage. Practically, laborious manual circuit mapping can be performed at early stage when neurons have a smaller arbor and fewer synapses to reconstruct.

However, the signal-to-noise level of synaptic counts at early stage can be low and might hamper the interpretation of connection strength between identified neuron

types. For instance, the connection from nociceptors to Basin-1 local interneurons is weak at early stage (< 3 synapses) in relation to its total dendritic input and compared to synaptic counts of other connections in the circuit. Optogenetic activation of nociceptors did not induce a Calcium response in Basin-1 neuron cell bodies at late stage (Ohyama et al. 2015) which could be explained by the low synaptic count at early stage. However, the authors find that co-activation of Basin-1 with Basin-4 increased the likelihood of rolling from 45% to 70% in third-instar larvae. Electrophysiological experiments show that Basin-1 can spike when only nociceptor axonal terminals are activated optogentically in third-instar larvae (Mei Shao, personal communication). Ohyama et al. 2015 suggest a facilitatory, modulatory role of Basin-1 in triggering rolling responses. In our data, we see a ten-fold increase in synaptic count of the nociceptor to Basin-1 connection from early to late stage that might account for this differences. Therefore, caution is advisable when interpreting low synaptic counts at early stages, as they can, as in the case for Basin-1, scale disproportionally strong. This could lead to different interpretations of their roles in the overall wiring diagram.

In Sulkowski et al. 2011, a possible functional role of the nociceptive circuit at early stage was further probed by quantifying distinct nocifensive behaviors using an optogenetic assay. The authors showed that larvae exhibit an aversive, long-lasting withdrawal behavior at early stage when nociceptors were continuously activated at low intensity levels. Therefore, nociceptors not only transduce transient, high-intensity noxious stimuli but also signals that cause avoidance responses on longer time scales. Also, such an ecologically-relevant type behavior has been shown to depend on nociceptors at late stage. Wandering larvae exhibit avoidance of dry surface for which multi-modal nociceptors are necessary (Johnson et al. 2012).

These long-lasting stimuli are hypothesized to increase the firing rate of nociceptors across the whole body wall simultaneously. In our circuit maps, the ascending interneurons would be particularly good candidates to integrate such body-wide aversive signals as they receive nociceptor inputs from each body wall segment. Targets of these ascending interneurons within the SEZ or brain could mediate biases in turn and crawl behavior that could result in avoidance phenotypes observed experimentally. The identification of these interneurons in this study, and the availability of GAL4 lines, allows for further studies to dissect their involvement in such long-lasting avoidance behaviors.

The organization of sensory receptors into topographic maps is fundamental for central processing in all sensory modalities (Weinberg 1997). Here, the non-overlapping tiling of nociceptor dendrites in the body wall provided an unique opportunity to study somatotopic aspects of sensory processing and circuit organization (Grueber et al. 2002). We demonstrated that specific local interneurons selectively receive synaptic input from the somatotopically organized nociceptor subtypes. In analogy to the developmental emergence of topographic maps in many sensory systems (Kaas 1997), we found that the synaptic preference between identified cell types is already present at early stages, and subsequently becomes more refined at the late larval stage. The preference to form connections from distinct

subtypes of nociceptors might have important consequences for central processing of stimuli transduced by class IV multi-dendritic neurons.

For instance, the late refinement of somatotopic organization might coincide with an ecologically-relevant requirement to detect localized noxious stimuli. In their natural habitats, late stage larvae might encounter the localized sting of parasitoid wasps which induces the vigorous escape-roll behavior (Tracey et al. 2003; Hwang et al. 2007). The possibility of directional biases of the escape-roll and other segmentally-specific nocifensive behaviors indicate that nociceptive circuits indeed process the somatotopic origin of noxious stimuli (Robertson et al. 2013; Tracey et al. 2003). Although other sensory modalities contribute to localize stimuli (Ohyama et al. 2015) in downstream circuitry, the segmental, left-right and dorso-ventral location of noxious stimuli sources might also be transduced directly by the activation of specific sets of nociceptors. Thus, the selective sampling of these signals by downstream interneurons might directly contribute to activate circuitry for location-dependent behavioral outputs. The high degree of synaptic preference observed in late stage circuits may further indicate increased functional relevance for such fine-scale discrimination of stimulus location in more exposed, late-stage wandering larvae.

A recent study showed functionally that largely cropped dendritic arbors can still produce robust motor neuron outputs (Ryglewski et al. 2014). This study found deficits in sophisticated motor behaviors that scaled with the degree of dendritic defect. Our analysis of the distribution of synapses on the dendritic backbone and twigs could be interpreted in this functional context. We have demonstrated that synaptic connectivity patterns recovered from synapses made only on the core dendritic backbones exhibits similar connectivity patterns as considering the total dendritic arbor as a whole. Assuming that backbone synapses are sufficient to drive postsynaptic response, the network could still perform adequately without synapses on distal dendritic twigs. In the study, the induced dendritic defects likely remove first these synapses in distal twigs, and thus the robust network output they observed could be due to the backbone connectivity. Since the twig synapses account mostly for the connection preference to specific interneurons, their removal might hamper more sophisticated, fine-grained behaviors.

Decades of work have elucidate a number of mechanism that work synergistically to control neurite guidance and synaptogenesis to produce cell-type specific connectivity patterns (extensively reviewed in Sanes et al. 2009). In *Drosophila*, the molecular diversity of *Dscam* receptors has been implicated for establishing wiring specificity between cell types (Chen et al. 2006; Millard et al. 2010). The patterns of synaptic preference in the nociceptive circuit could be established by similar developmental mechanisms. Due to the absence of a clear spatial segregation of distinct presynaptic axonal terminals of different nociceptor types, postsynaptic dendritic may selectively increase the probability to maintain synapses formed with preferred input partners based on their molecular identity. The cell-cell recognition mechanism at synaptic locations may require the postsynaptic neuron process to detect the identity of the presynaptic cell type via the expression of specific receptors (Zipursky et al. 2010). Subsequently, the probability for synapse maintenance could then be modulated locally by intracellular signaling cascades, and thus give rise to

100

the observed pattern of synaptic preference and specificity. The genetically-tractable nociceptive system could be used to further test the molecular basis of these synaptic specificity-producing processes. The unique opportunity for volumetric EM in this system would allow investigations of the results of genetic alterations on patterns of synaptic and circuit level differences.

In pain processing circuits in vertebrate systems, endogenous modulation of pain perception has been shown to be mediated by pathways originating in the brainstem and descending within the spinal cord (Mason 2012). Descending pathways can act to both facilitate or inhibit transmission of nociceptive signals (Ossipov et al. 2010; Zeilhofer et al. 2012). One mechanism by which inhibition is achieved is through presynaptic inhibition of sensory terminals (Hedwig et al. 1996; Rudomin et al. 1999).

Analogically in the *Drosophila* larvae, we found the SelN138 descending interneuron to be a candidate interneuron that could mediate such presynaptic inhibition. Through the segmentally-local reciprocal connectivity, the activation of nociceptive terminals could be counter-balanced by local feedback inhibition. Additionally, the activation of the SelN138 dendrites in the subesophageal zone could simultaneously inhibit all nociceptors in the VNC to completely inhibit or reduce nociceptive signal transmission.

We speculate that a similar pattern of reciprocal inhibition may occur in the axonal output arbors of local interneurons. In six local interneuron types at both early and late stage, we found that each axonal output arbor receives inputs from unknown presynaptic neurons (Supplementary Fig. 3.23). These presynaptic inputs could modulate the efficacy of axonal outputs, similar to the SelN138 for nociceptors, through an inhibitory action. Further circuit dissection will be necessary to test this hypothesis.

3.4 METHODS

Serial section Transmission Electron Microscopy

The central nervous systems from feeding 3rd instar wild-type larvae were dissected in PBS and immediately transferred to 125 ul of 2% glutaraldehyde in 0.1 M Na cacodylate buffer, pH 7.4 in a 0.5 dram glass vial (Electron Microscopy Sciences, cat. no. 72630-05) on ice. 125 ul of 2% OsO4 in 0.1 M Na-cacodylate buffer, pH 7.4 was then added and briefly mixed immediately before microwave assisted fixation conducted with a Pelco BioWave PRO microwave oven (Ted Pella, Inc.) at 350W, 375 W and 400 W pulses for 30 sec each, separated by 60 sec intervals. Samples were rinsed 3 x 30 sec at 350 W with 0.1 M Na-cacodylate buffer, separated by 60 sec intervals, and post-fixed with 1% OsO4 in 0.1 M Na-cacodylate buffer at 350W, 375W and 400W pulses for 30 sec each, separated by 60 sec pauses. After rinsing with distilled water 3 x 30 sec at 350W with 60 sec pauses between pulses, the samples were stained en bloc with 7.5% uranyl acetate in water overnight at 4 degree C. Samples were then rinsed 3 x 5 min with distilled water, dehydrated in an ethanol series followed by propylene oxide, infiltrated and finally embedded in Epon resin. Serial 50 nm sections were cut using a Diatome diamond knife and a Leica UC6 ultramicrotome, and picked up on Pioloform support films with 2 nm C on Synaptek slot grids. Sections were stained with uranyl acetate followed by Sato's lead (Sato 1968) prior to imaging. An FEI Spirit TEM operated at 80 kV was used to image the serial sections at 2.3 x 2.3 nm pixel resolution using Leginon (Suloway et al. 2005) to produce the L3 dataset.

The sample preparation and imaging procedure for the L1a and L1b datasets were similar and are described in (Schneider-Mizell et al. 2015; Ohyama et al. 2015) respectively. Briefly, dataset L1a covers the whole CNS of a first-instar larva. The pixel resolution is 3.8 x 3.8 nm, section thickness is approximately 50 nm, and a total of 4841 sections (56 lost) were collected, and 187048 images were taken with 4k x 4k pixel dimension. Dataset L1b covers half of the abdominal segment A2 and the segment A3 of the ventral nerve cord (VNC) of a first-instar larva. The pixel resolution is 4.4 x 4.4 nm, section thickness is approximately 50 nm, and a total of 462 sections (4 lost) were collected.

Image montaging and alignment

The approximately 300'000 images of the L3 dataset (4k x 4k pixel for each imaged tile) were montaged and aligned using linear and nonlinear methods described in (Saalfeld et al. 2012) and implemented in TrakEM2 (Cardona et al. 2012). In particular, filters for brightness and contrast correction were applied before montaging (Default min and max, normalized local contrast, enhance contrast). Images were first montaged in a section with two passes of linear montaging, first targeting only a translation transformation (correspondence weight 1.0), and in the second pass targeting an affine transformation. This was followed by an elastic, non-linear montaging pass. For alignment between sections, parameter exploration was performed

on a scaled down substack (scale factor 10) of 5 sections, targeting extraction of approximately 2000+ features, 100+ correspondences and an average displacement of 10 pixel. Linear alignment was applied to all the sections using an affine transformation model. The block matching test plugin in Fiji (Schindelin et al. 2012) was used on 5 adjacent sections to find optimal parameters for the elastic registration pass. Elastic alignment was applied with local smoothness filter approximating an affine local transformation. The resulting aligned image stack was exported to an image tile pyramid with six scale levels for browsing and reconstruction in CATMAID (Schneider-Mizell et al. 2015). The datasets L1a and L1b were aligned with the same methods (Schneider-Mizell et al. 2015; Ohyama et al. 2015) and optimal parameter settings specific to their respective image quality.

Neuronal circuit mapping

The nociceptive neural circuits were reconstructed manually from the aligned EM image volumes using the collaborative circuit reconstruction tools in CATMAID (Schneider-Mizell et al. 2015). Briefly, neurons of interest were reconstructed manually by following their neuronal processes and annotating a 3d skeleton on their centerline. While reconstrucing neuronal processes, presynaptic active zones and postsynaptic partners were annotated following the criteria described in (Schneider-Mizell et al. 2015).

In all three datasets, the identification of the class IV multi-dendritic sensory axons (mdIV) followed the same approach. Central axonal projection patterns of all three mdIV *ddaC*, *v'ada* and *vdaB* in one hemisegment were previously described (Grueber et al. 2007), and their morphology, projecting nerve and target projection zone allows to distinguish them from other classes of sensory axons and into their specific cell-type. Only the dorsal *ddaC* projects in the intersegmental nerve (ISN), whereas *v'ada* and *vdaB* project in the segmental nerve (SN). The projection zone of the mdIVs in a cross-section of the ventral nerve cord is located medially and ventrally. Individual fibers in both nerve bundles were followed and abandoned if their trajectory was inconsistent with their expected termination zone for mdIVs. The texture of sensory axons appears in the EM preparation darker compared to adjacent dendritic processes, revealing bundles of mdIVs. These bundles highlighted potential mdIV candidates form anterior and posterior segments. By following back out to the nerve these candidate mdIVs were readily identified as belonging to the mdIVs. The *vdaB* does cross the midline axis, whereas *v'ada* does not, thus allowing to distinguish those cells which both project in the SN.

The synaptic partners of the complement of the 6 sensory axons in an abdominal segment (three in each hemisegment) were then used to seed the reconstruction of the upstream and downstream interneurons of the mdIVs. Using the connectivity widget in CATMAID, each placeholder skeleton node associated with the mdIVs was used to start seed the reconstruction of a neuron. Generally, these reconstruction terminate in a cell body and the neuron can be identified. However, a fraction of synaptically connected processes leave the EM volume, run into a gap of sections, or have no apparent connection to a larger neuronal arbor. We call these processes

fragments to which a fraction of nociceptors outputs connect (Supplementary Fig. 3.6). These fragments are usually short processes with few postsynaptic sites. They usually do not contain microtubuli, which makes finding the connection point to the processes of a larger neuron difficult.

A number of output connections from nociceptors connect to other nociceptors. The annotation of postsynaptic sites to other nociceptors were consistent with our synapse annotation criteria. However, we have no evidence that these are real synaptic connections in the functional sense. We excluded these synaptic edges in the present analysis, although we annotated such connections in all three datasets.

Identification and naming of interneurons

We classified the full complement of identified neurons synaptically adjacent to the multi-dendritic class IV sensory axons into three categories: a) local interneurons, b) ascending projection neurons, c) descending neurons. Each identified interneuron type is given a name based on the Jim Truman's neuron catalog nomenclature. This neuron catalog is based on labeling of individual cells from a collection of GAL4 lines based on a flp-out method (Jim Truman, Nern et al. 2015). For ventral nerve cord interneurons of the catalog, the name is composed of segment (A: abdominal, T: thoracic), two digit number denoting the cell lineage (e.g. 08), and an enumeration character (e.g. b), for instance *A08b*. In addition to these standardized names, we use a naming scheme for ease of reference that is usually based on an overall shape feature. For in-depth comparative analysis, six local interneuron that are local within a segment were selected: a) A02m: Pseudolooper-3; b) A02n: Pseudolooper-4; c) A10a: Chair-1; d) A09a: Basin-2; e) A09c: Basin-4; f) A09l: Down-and-back-1. The dendritic arborization of these neurons is confined mostly to the segment where their cell body is located. These six interneurons were given extensive review in order to complete the reconstruction of their arbors in non-mdIV dendritic domains as well as their axonal output domains.

Other local, ascending and descending interneurons downstream of nociceptors were reconstructed to the degree where their identity could be determined but not reviewed. Postsynaptic and presynaptic sites were annotated. Interneurons were matched to morphologies in the neuron catalog and named. It was required that both left and right homologous are consistently found in all three datasets. We identified additional local (A09b: Basin-1), ascending (A08n; TePn19; A02o; A09e; A09o; TePn05) and descending (SelN138) interneurons. In dataset L3 but not in L1, we additionally identified local interneuron A01l postsynaptic of nociceptors.

Annotation of neuronal compartments

The quantitative analysis on the subarbor levels requires the annotation of compartment boundaries on nodes of the skeleton morphology. The skeleton representations are directed tree structures with their root nodes at cell body locations for interneurons and at nerve entry location for sensory neurons by convention. Thus, by way of

using the tagging system in CATMAID to label individual skeleton nodes by a string, we can implicitly classify all the skeleton nodes downstream of this tagged node, i.e. a subtree rooted at the tagged node, as belonging to the subcompartment denoted by the string. By combining subtrees demarked using this scheme, we can flexibly segregate any subcompartment that might be of interest for detailed quantitative analysis.

We applied this scheme to demark the dendritic and the axonal parts of the preferred local interneurons. We choose highly stereotyped locations on the arbors that are consistently definable in first- and third-instar interneurons. Generally, the entry point of a neuronal process into the neuropil is a very stereotyped location not impacted by variations of exact cell body location. Furthermore, we found the last postsynaptic site (or branch point with a postsynaptic site) on the primary process between dendritic and axonal arborization to be a good termination point to contain the dendrite, and the location of the first presynaptic site on the very same primary process to begin the axonal part of the neuron.

Furthermore, we segregated neurons into backbone and twigs compartments using the same method. We iterate over the skeleton nodes starting from the leaf nodes and add tags at the boundary locations where microtubuli terminate. This annotation is tremendously simplified by the dendrogram widget in CATMAID that presents an interactive radial layout of the skeleton structure collapsed into branch nodes with associated tags. Subtrees rooted by a selected tag are highlighted in red and allow to readily see what part has already been annotated. The dendrogram allows to interactively navigate to points on the arbor.

Minigraph neuron visualization

The 2D graph visualization of neuronal arbors (Supplementary Fig. 3.17 - 3.21) were generated using Cytoscape Version 3.2.1 and the Organic Layouting. The 3D skeleton morphologies were postprocessed using custom scripts so that only points of interest and the arbor topology were kept. Points of interest are the cell body location, presynaptic and postsynaptic locations, branch and leaf nodes. All the other skeleton nodes were removed, such as continuation nodes, to preserve the arbor topology. Postsynaptic sites are color-coded according to the cell-type of the presynaptic neurons. Color-codes are as used throughout this paper. This visualization method highlights the distribution of cell-type-specific synaptic input across neuronal arborizations. For instance, clusters of nociceptive input on a specific subarbors may correspond to inputs from specific spatial domains, and are readily visible in 2D. Furthermore, presynaptic site locations, for instance at the base of subarbors, are readily visible.

CONTRIBUTIONS

SG, AC designed the study;
RF performed ssTEM;
IA, SG, CSM, AC mapped and reviewed the circuits;
SG analyzed the data;
SG wrote the paper with input from co-authors.

3.5 SUPPLEMENTARY FIGURES
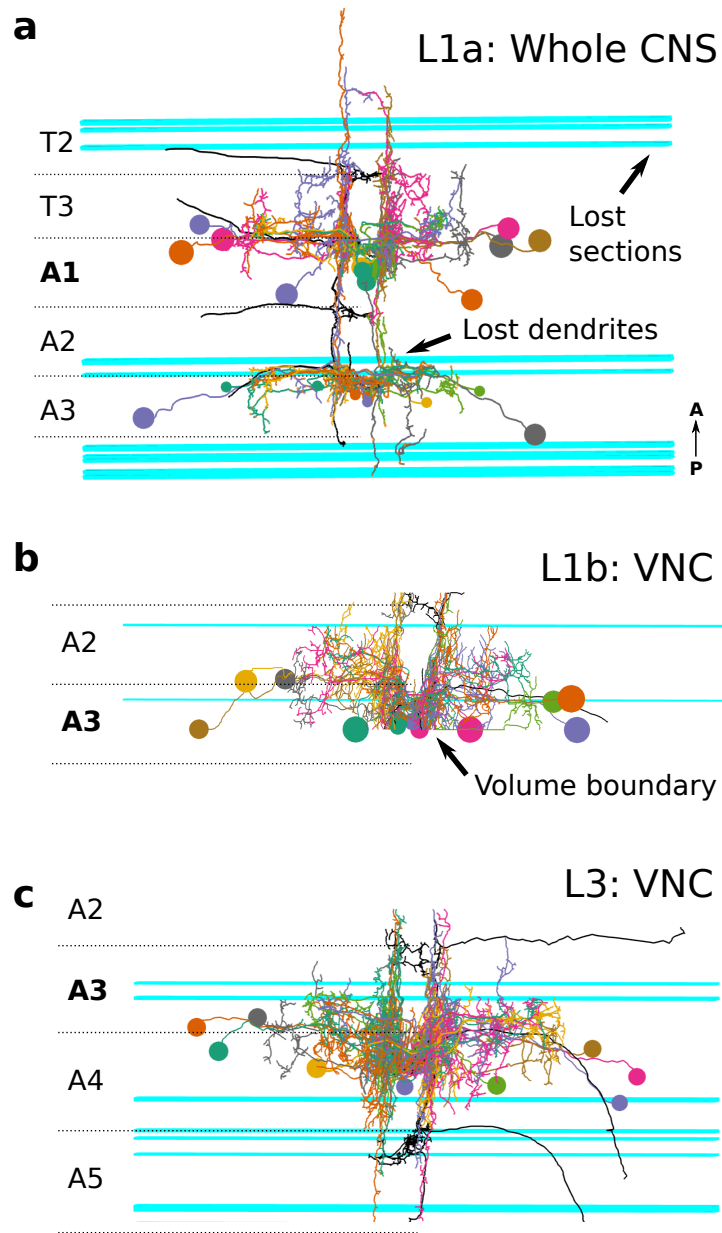
Supplemental Figure 3.1.: **Dataset limitations and their impact for circuit comparison** A set of lost sections in abdominal segment A3 in the whole CNS volume (L1a) limited completeness of reconstruction of local interneuron arbors. Especially, thin dendritic processes are hard to follow across gap large gaps (spanning 10 sections, about 500nm) and are lost. A second first-instar dataset (L1b) containing an almost complete abdominal segment A3 was used to establish similarity of cell-types features and connectivity with neurons in abdominal segment A1 in L1b. Developmental comparison of third-instar neurons in segment A3 was then made with both segment A1 in L1b and A3 in L1b.
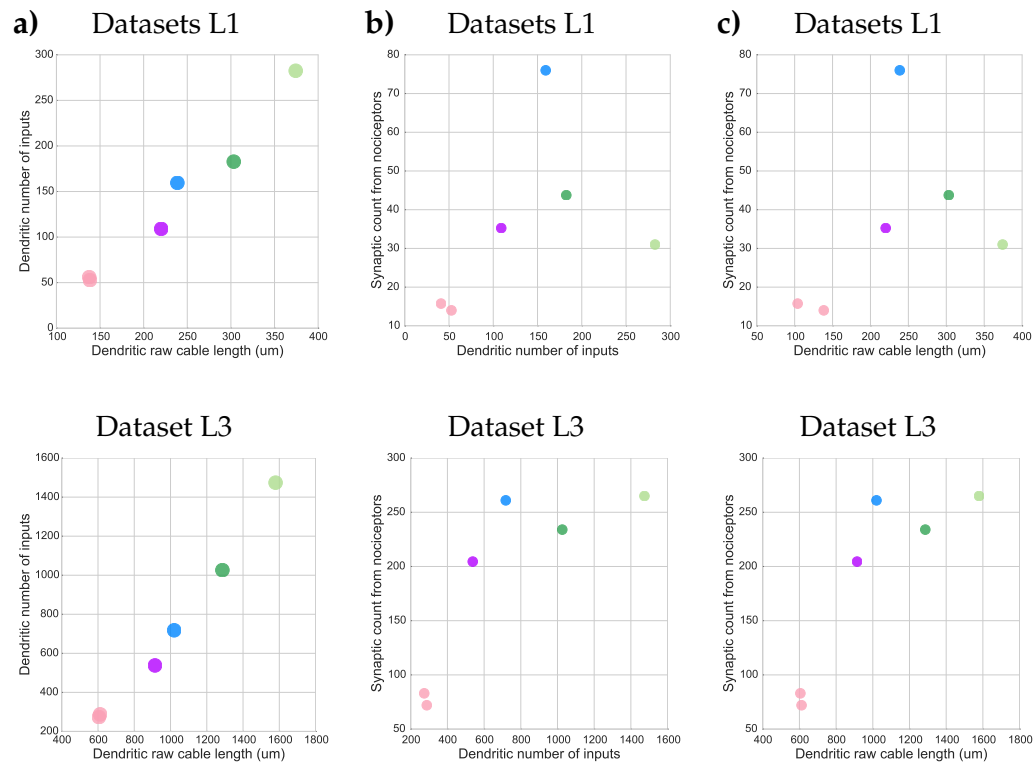
Supplemental Figure 3.2.: **Correlation of dendritic cable length, number of inputs in total and from any nociceptor** Upper panel shows data for first-instar laervae (average of dataset L1a and L1b), and lower panel shows data for third-instar larvae (dataset L3). Color-code corresponds to the different local interneuron cell-types. a) The number of dendritic inputs is linearly correlated with the amount of dendritic cable. As neurons grow from L1 to L3, the total amount of dendritic arbor increases but the ordering of the cell-types remains the same which is indicative of similar growth rates. b) and c) Absolute number of synapses from any nociceptor onto local interneurons as a function of total dendritic inputs and total cable length. The overall pattern stays remarkably similar, with some more variation of Down-and-back-1 (blue) and Basin-2 (bright green).
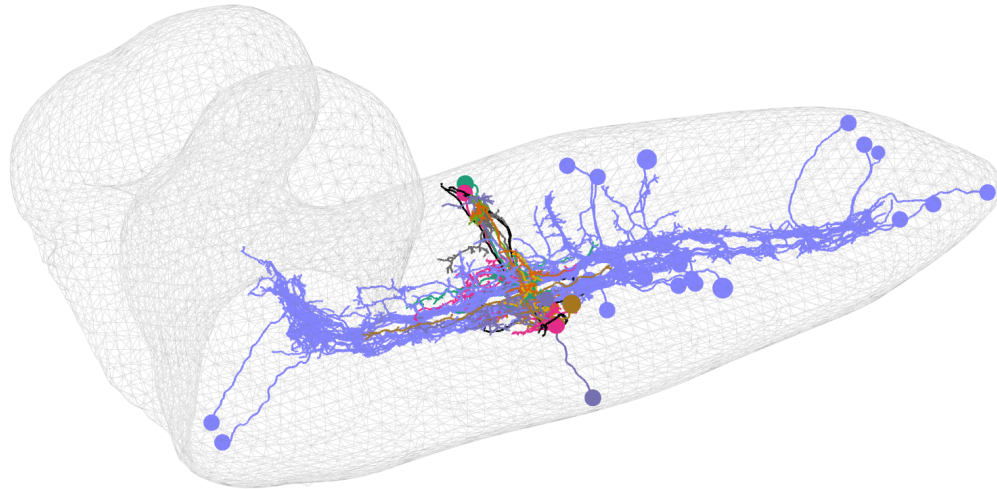
Supplemental Figure 3.3.: **3D visualization of morphology of nociceptors and first-order network in L1a** Central nervous system mesh is shown for context. Local interneurons in abdominal segment A1 are shown in colors. Nociceptors are in black colors. Ascending and descending interneurons in blue.

	L1a	L1b	L3	scale	linear scale
body length early (mm)	0.7	0.7	2.7		
neuropil area at A3 (μm^2)	1000	1000	3375	3.4	1.8
neuropil height at A3 (μm)	25	25	45	1.8	1.8
neuropil width at A3 (μm)	40	40	75	1.9	1.9
VNC width at A3 (μm)	90	70	120	1.5	1.5
VNC height at A3 (μm)	50	45	70	1.5	1.5
intersegmental distance (μm)	15	-	23	1.5	1.5
local volumetric mdIV domain (μm^3)	24	22	184	8.0	2.0

Table 3.1.: **Scaling of anatomical features in first- and third-instar (wandering) larva.**
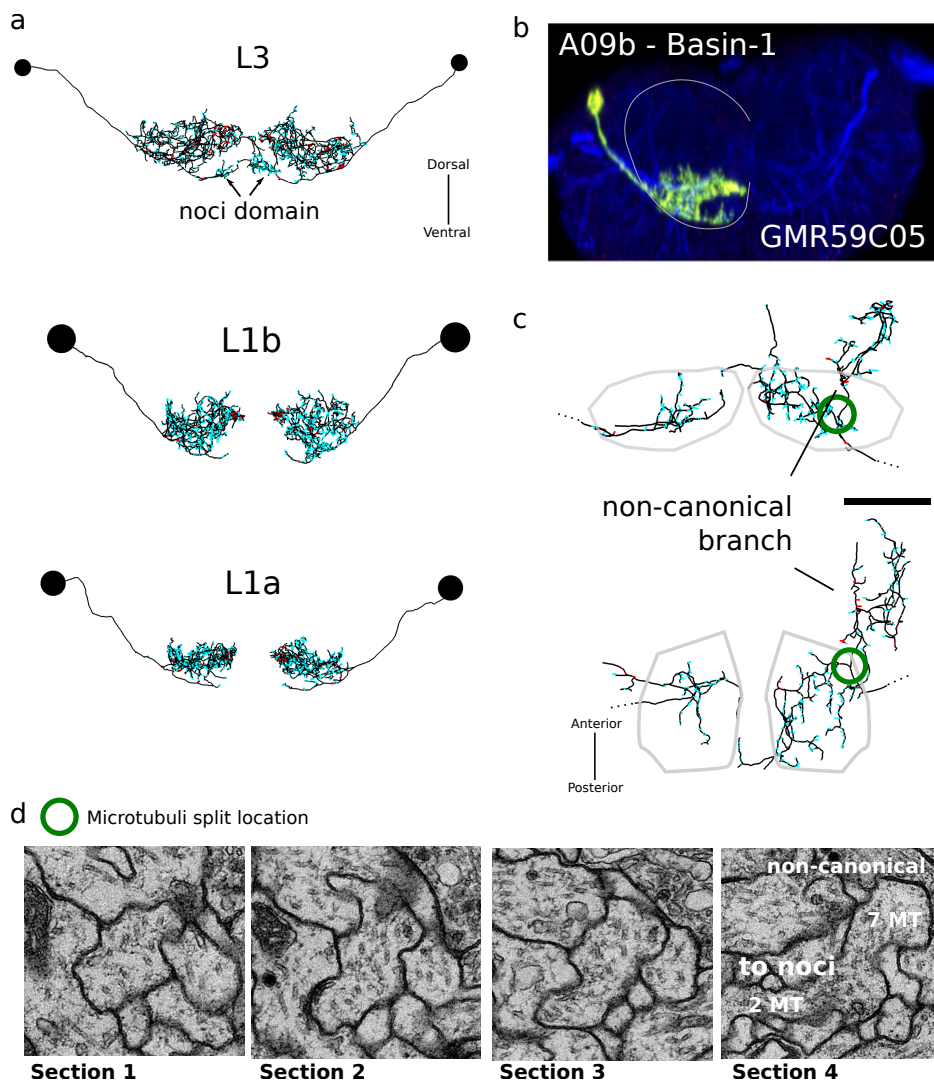
Supplemental Figure 3.4.: **Asymmetry of microtubuli number in the Basin-1 nociceptive subarbor** a) Reconstruction from three EM datasets. Dendritic arborization of Basin-1 cells avoid the nociceptive mdIV domain except for one microtubuli-filled backbone branch. This branch is already present in L1 datasets, and is much more elaborate in the L3 datset. b) Confocal image of a fluorescently labeled Basin-1 neuron shows a weak signal of the nociceptive subarbor in third-instar larvae. c) Stereotyped location of presynaptic site across all 3 datasets. Close-up view within the nociceptive domain of the left and right Basin-1 subarbors in the L3 dataset. The right subarbor contains nine microtubuli, whereas the left subarbor contains only two. The location of a non-canonical microtubuli split for the right Basin-1 is marked with a green circle. d) Series of EM images at the microtubuli split location. The upwards branch contains seven microtubuli, whereas the branch to the nociceptive domain contains only two, like at its contralateral location. The non-canonical branch hosts a number of presynaptic sites in addition to additional postsynaptic sites which are not present on the contralateral side.
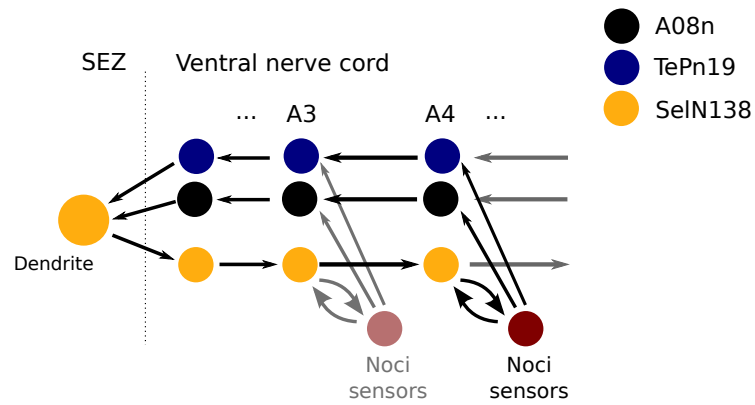
Supplemental Figure 3.5.: **Circuit motif of the descending SelN138 interneuron** Nociceptive signals are integrated across the nerve cord by two types of ascending projection neurons (A08N, TePn19). These two neuron types target the dendritic arbors of the descending neuron SelN138 in the subesophageal zone (SEZ).

	A09l_a3l Down-and-back-1	A09l_a3r Down-and-back-1	A10a_a3l Chair-1	A10a_a3r Chair-1	A09a_a3l Basin-2	A09a_a3r Basin-2	A09c_a3l Basin-4	A09c_a3r Basin-4	A02m_a3l Pseudolooper-3	A02m_a3r Pseudolooper-3	A02n_a3l Pseudolooper-4	A02n_a3r Pseudolooper-4	Sum
ddaC_a2l	36	4	2	15	50	8	2	2	7	6	8	9	149
ddaC_a2r	7	18	5	3	4	39		1	6	12	12	6	113
ddaC_a3l	43	14	12	37	84	19	13	2	26	23	19	17	309
ddaC_a3r	10	40	43	24	24	81	1	6	18	36	27	24	334
ddaC_a4l	8		1	9	4								22
ddaC_a4r		2	1			6						1	10
v'ada_a2l	15			5	9		10	1				1	41
v'ada_a2r		13	3	1		10		4				1	32
v'ada_a3l	40			9	69		55		4	1	1	2	181
v'ada_a3r		47	6	1		62	13	63	5	6	6	1	210
v'ada_a4l	12			9	18		23	1	3		1		67
v'ada_a4r		12	8	1		20	4	18	1	3			67
vdaB_a2l	9												9
vdaB_a2r	5	7						4					16
vdaB_a3l	59	16	23	58	10	1	59	22	1	1		1	251
vdaB_a3r	26	48	43	39		4	46	59	3	3	1	3	275
vdaB_a4l	17	1	4	20	5		17	4				2	70
vdaB_a4r		13	20	7		3	10	28	1		1		83
Sum	287	235	171	238	277	253	253	215	75	91	76	68	2239

Supplemental Figure 3.6.: **Connectivity matrix from local and adjacent nociceptors to local interneuron in L3**

113

	A09l_a1l Down-and-back-1	A09l_a1r Down-and-back-1	A10a_a1l Chair-1	A10a_a1r Chair-1	A09a_a1l Basin-2	A09a_a1r Basin-2	A09c_a1l Basin-4	A09c_a1r Basin-4	A02m_a1l Pseudolooper-3	A02m_a1r Pseudolooper-3	A02n_a1l Pseudolooper-4	A02n_a1r Pseudolooper-4	Sum
ddaC_t3l	14		1	2	8		5		3	1	5	4	43
ddaC_t3r		8	2			5		2				1	18
ddaC_a1l	6	2	2	5	4		5		4	1	7	3	39
ddaC_a1r	1	16	5			7		3	1	2	2	7	44
ddaC_a2l	1			2			1						4
ddaC_a2r		2	2									2	6
ddaC_a2r													0
v'ada t3l	20	1		4	9		8	1	2			1	46
v'ada t3r		16	2			7		11		1		1	38
v'ada_a1l	13		1	6	11		10		2		7	1	51
v'ada_a1r		27	5			11		19					62
v'ada_a2l	1		2				3						6
v'ada_a2r		3						1					4
vdaB_a1l	18	4	2	11	2		15	2	3		3	1	61
vdaB_a1r	7	18	5	2		1	4	14		1			51
vdaB_a2l	7			4	1		4	1		1			18
vdaB_a2r		5	4					3					12
Sum	88	102	33	36	35	31	55	57	15	6	24	21	503

Supplemental Figure 3.7.: **Connectivity matrix from local and adjacent nociceptors to local interneuron in L1a**

	A09l_a3l Down-and-back-1	A09l_a3r Down-and-back-1	A10a_a3l Chair-1	A10a_a3r Chair-1	A09a_a3l Basin-2	A09a_a3r Basin-2	A09c_a3l Basin-4	A09c_a3r Basin-4	A02m_a3l Pseudoloooper-3	A02m_a3r Pseudoloooper-3	A02n_a3l Pseudoloooper-4	A02n_a3r Pseudoloooper-4	Sum
ddaC_a2l	6		3	2	7		1		2	2	1		24
ddaC_a2r		4	1			7							12
ddaC_a3l	10	2	2	4	5	1	2	1	3	2	2	2	36
ddaC_a3r	3	7	6	2	1	8	2	2	11	10	6	2	60
v'ada_a2l	6				1								7
v'ada_a2r		2				3							5
v'ada_a3l	22				5	12	20				1	3	63
v'ada_a3r		7	13				1	4	1				26
vdaB_a2l	4												4
vdaB_a2r	3	1			1								5
vdaB_a3l	9	11	10	15			14	2				1	62
vdaB_a3r	4	13	8	1	3	9	5	9	3		1		56
Sum	67	47	43	29	30	28	45	18	20	15	13	5	360

Supplemental Figure 3.8.: **Connectivity matrix from local and adjacent nociceptors to local interneuron in L1b**

a) Dataset L3

	A08n right	A08n left	A02o_a3r	A02o_a4r	A02o_a4l	A02o_a3l	TePn19 left	TePn19 right	A09e_a3r	A09e_a3l	A09e_a4r	A09e_a4l	A09o_a3r	A09o_a3l	TePn05 right	TePn05 left	TePn05 right	TePn05 left	TePn05 right	TePn05 left	SeIN138 right	SeIN138 left	Sum
ddaC_a3l	33	48	20	9	44	30	48	10	20	14	14	6	13	7	8	8	5	8	5	11	16	14	391
ddaC_a3r	29	33	47	38	16	18	5	30	26	8	6	2	11	5	15	1	8	5	10	2	26	14	355
v'ada_a3l	27	36	1		51	55	58		8	16	5	1	6	4		13		8		5	5	26	325
v'ada_a3r	46	19	65	63				41	19	8	8		13	8	8		16	9	8		19	13	363
vdaB_a3l	16	21	16	9	28	32	15	8	10	13	4		1	1	5	13	3	3	1	3	11	9	222
vdaB_a3r	41	32	31	46	10	11	5	19	14	17	1	1	3	5	18	6	4	2	1	1	6	5	279
Sum	192	189	180	165	149	146	131	108	97	76	38	10	47	30	54	41	36	35	25	22	83	81	1935

b) Dataset L1a

|	A08n right	A08n left	A02o_a2r	A02o_a3r	A02o_a2l	A02o_a3l	TePn19 left	TePn19 right	A09e_a3l	A09e_a3r	A09e_a4l	A09e_a4r	A09o_a3l	A09o_a3r	TePn05 left	TePn05 right	SeIN138 left	SeIN138 right	Sum
ddaC_a1l	8	7			7	8		15	11	9	2	1		3	2		3	2	78
ddaC_a1r	10	13	7	14	1	1	7		6	14	2	5	3			2	2	5	92
v'ada_a1l	2	4			6	8		12	10	1	4			4	2		5	2	60
v'ada_a1r	4	4	8	12			11		3	6	1		3			1	1	7	61
vdaB_a1l	1	2			1	2	1	5	12	4				2	3	1	4		38
vdaB_a1r	5		4	6			4		5	7	2	1			1			4	39
Sum	30	30	19	32	15	19	23	32	47	41	11	7	6	9	8	4	15	20	368

Supplemental Figure 3.9.: **Synaptic connectivity from local nociceptors to identified ascending and descending interneurons**
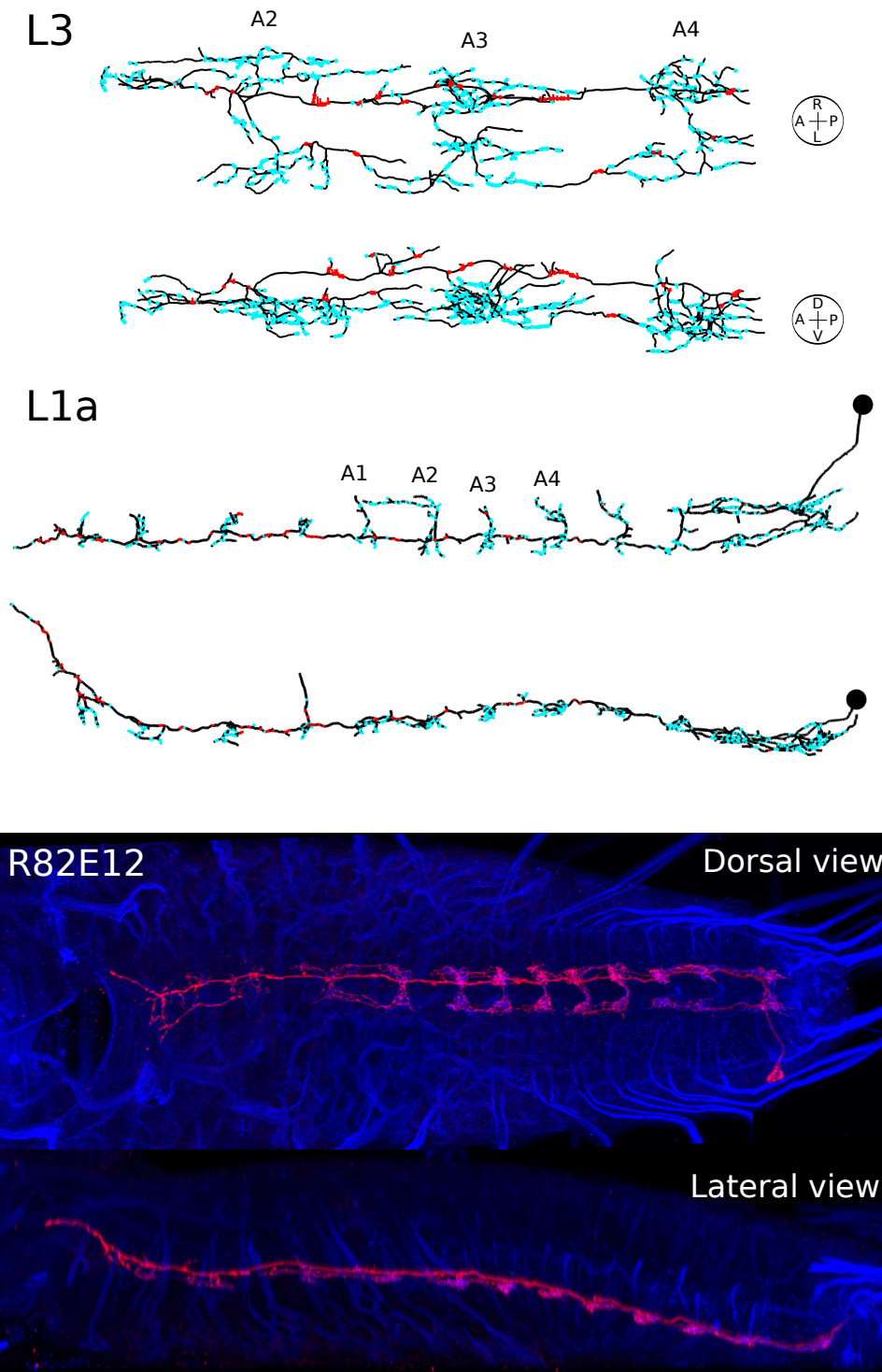
Supplemental Figure 3.10.: **Ascending projection neuron: A08n** The ascending projection neuron is strongly connected to nociceptors. Notice the absence of any presynaptic sites on the dendritic arbor in the noiceptor domain, and presynaptic site clustering between the segments on the main anterior projecting process. The axonal output arbor with presynaptic sites is located most anteriorly in the SEZ. Thick varicosities with presynaptic sites are nicely visible on the projecting process in the dorsal view of a single cell labeled at third-instar stage (*This and the following Neuron Catalog LM images courtesy of Jim Truman, HHMI JRC*)
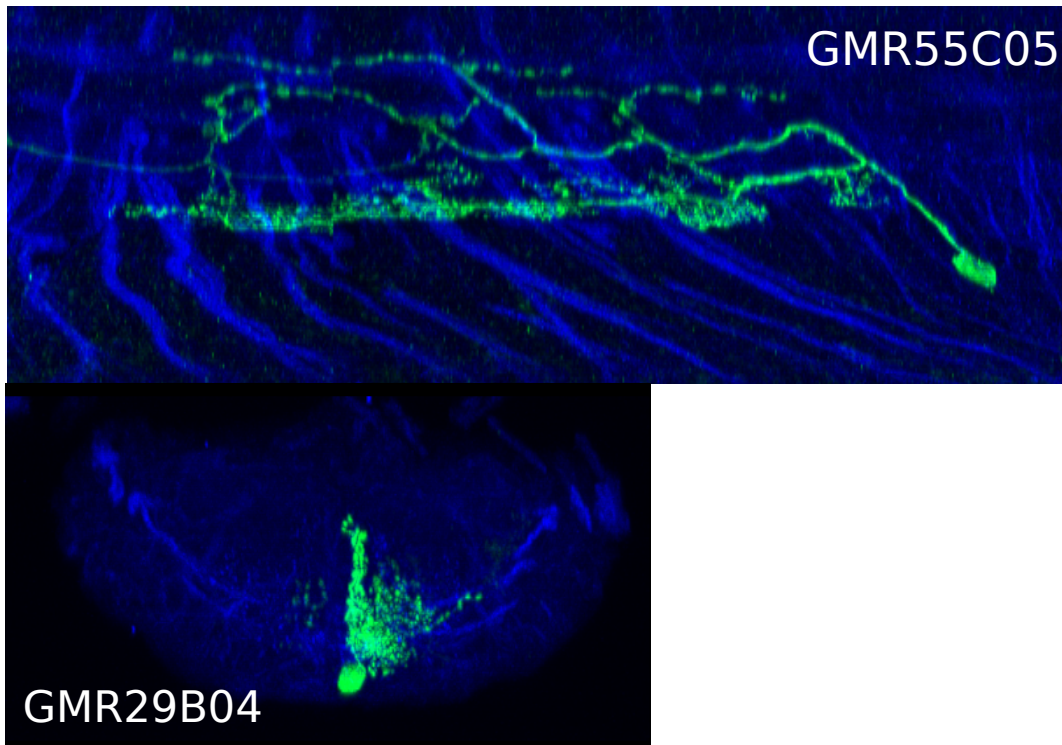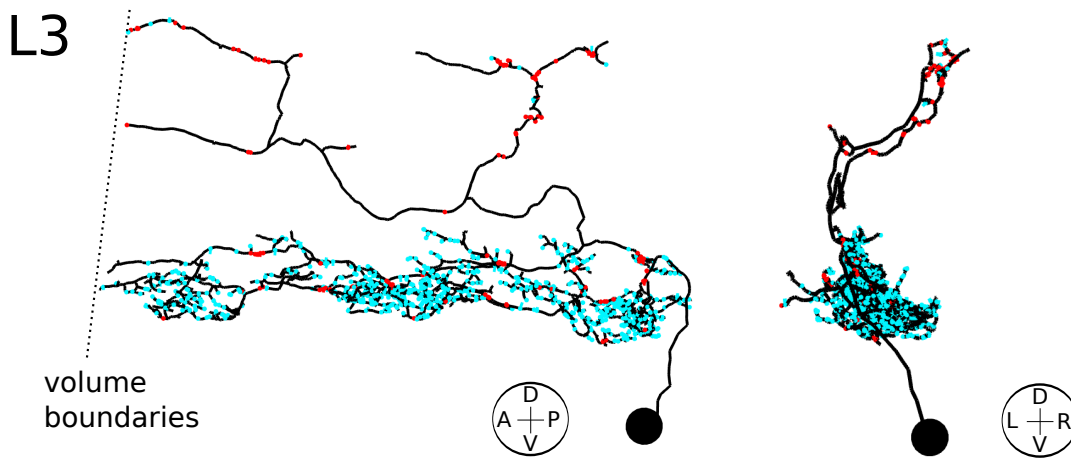
Supplemental Figure 3.11.: **Ascending neuron: A02o**. The Wave-1 cell projects anteriorly on the ipsilateral side without mideline crossing and innervates the class IV domains of multiple ventral nerve cord segments. Unlike other ascending projection neurons, the output arbors project dorsally and spans multiple anterior segments. Notice the presynaptic sites on the dendritic arbors, and postsynaptic sites on the axonal arbors.
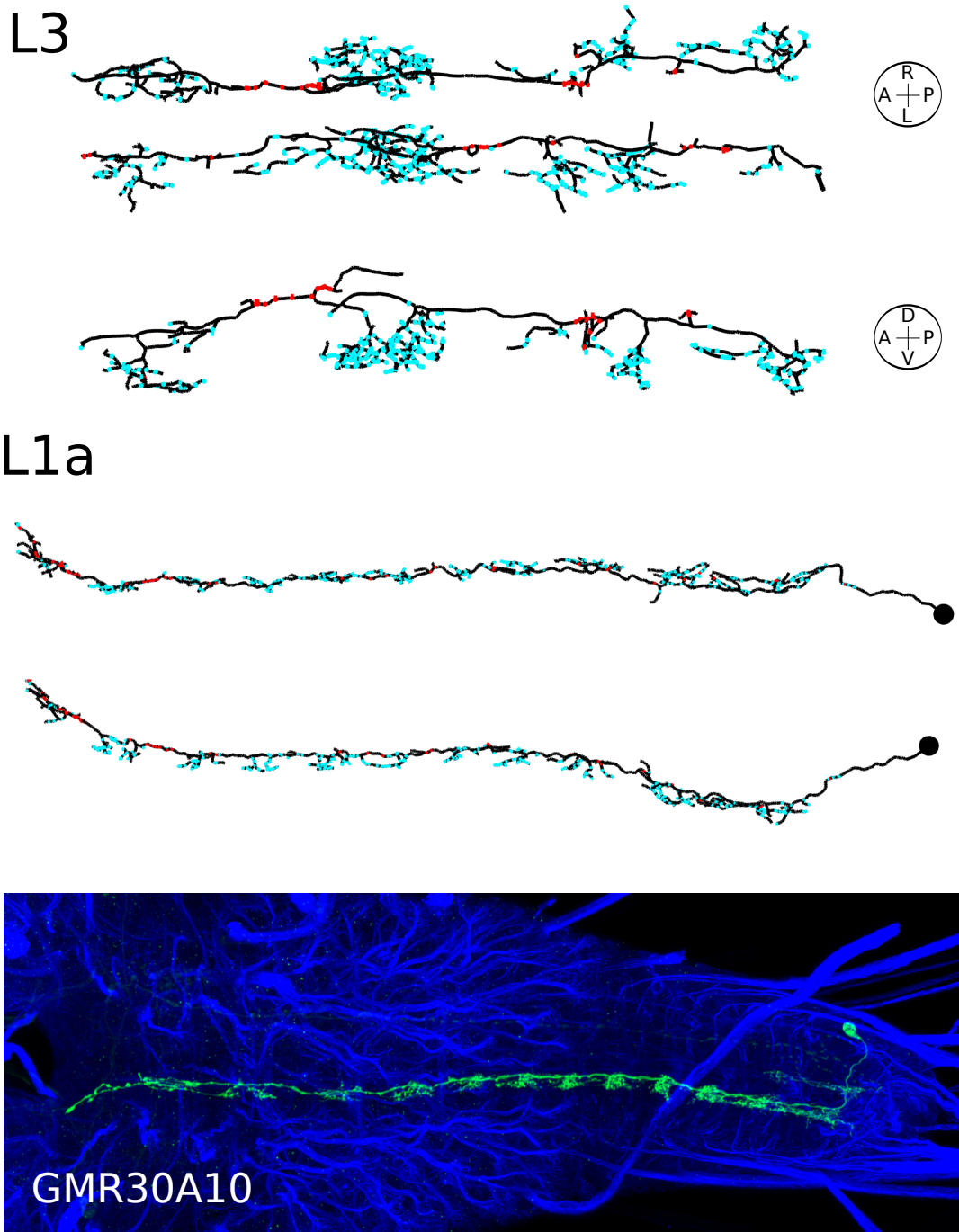
Supplemental Figure 3.12.: **Ascending projection neuron: TePn19** The confined dendritic arborization per segment are nicely visible in L3 and the LM images. In L1a, these arborizations are starting to emerge. Intermixed presynaptic sites along the main anteriorly projecting process, and a terminal axonal branch in the SEZ.
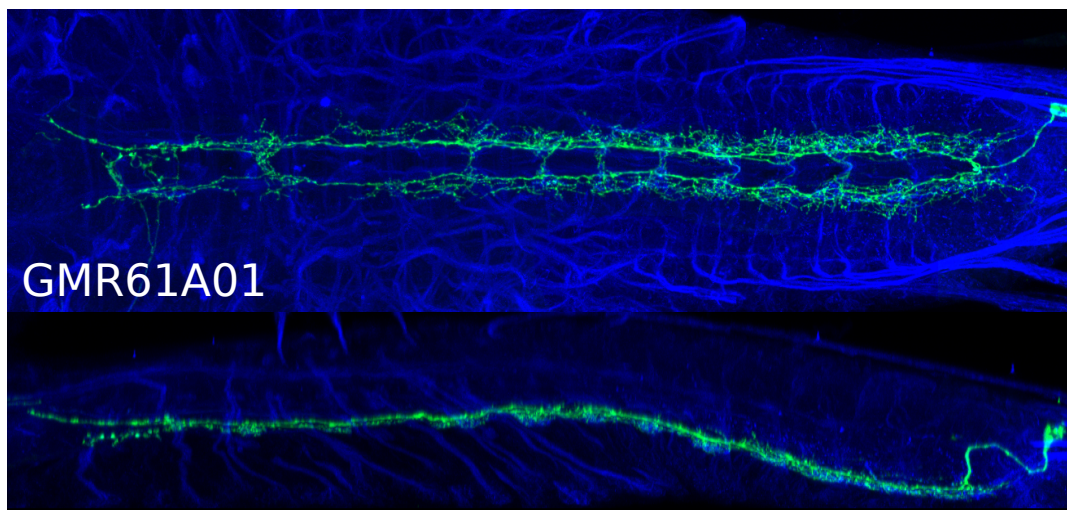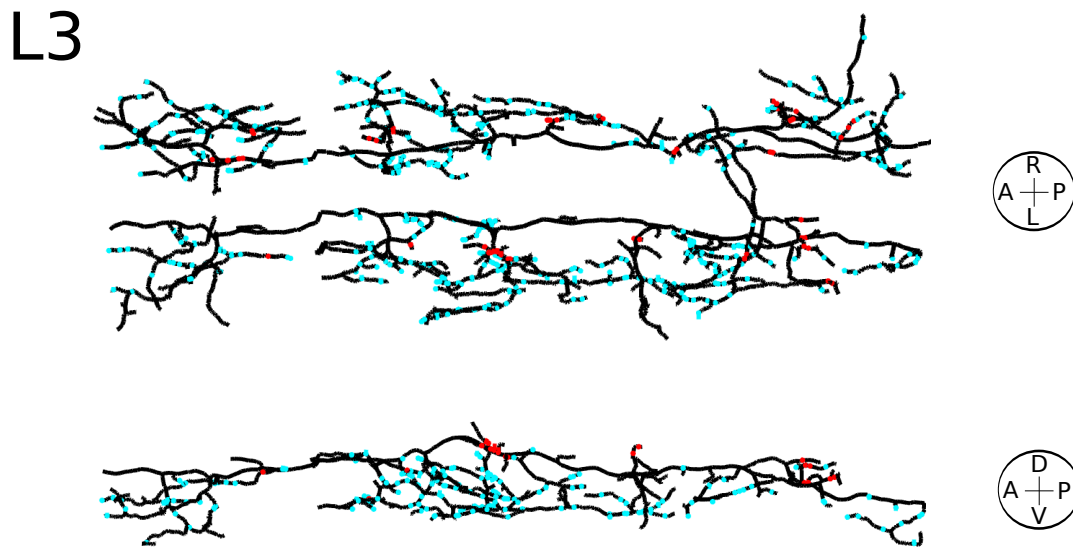
Supplemental Figure 3.13.: **Ascending projection neuron: TePn05** Due to the limited volume size in L3, left and right partial arbors likely belong to the bilaterally projection TePn05 neuron. Notice the mideline crossing of arborization, and distribution of presynaptic sites to intersegmental regions.

L3

L1a
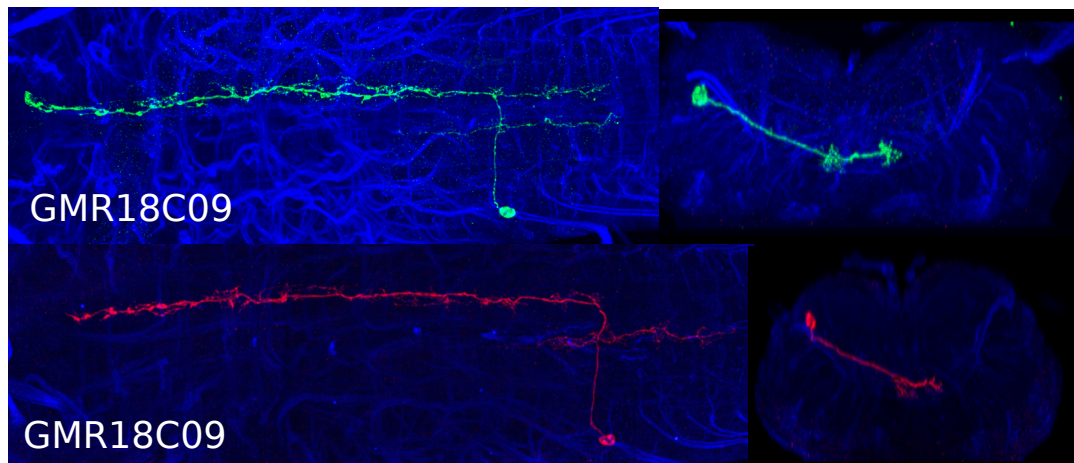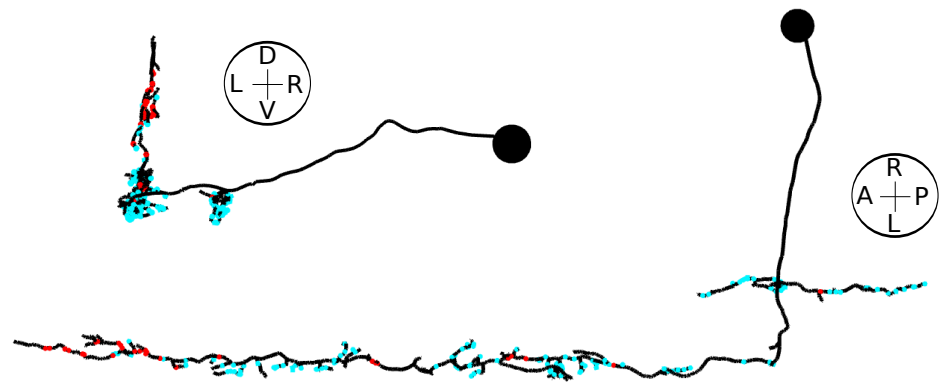
Supplemental Figure 3.14.: **Ascending projection neuron: A09o**

L3

L1a

GMR56H02

GMR52B10

GMR48F12
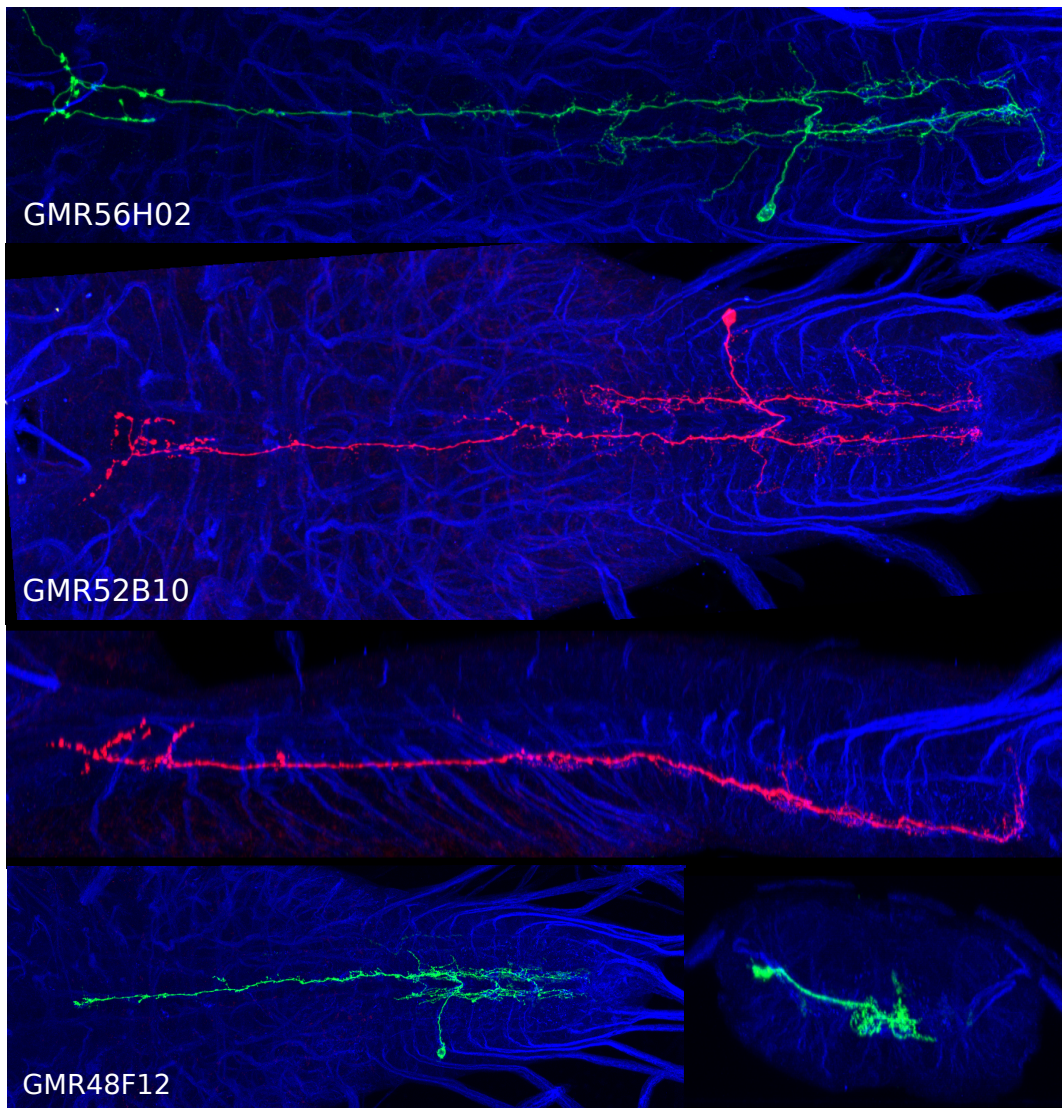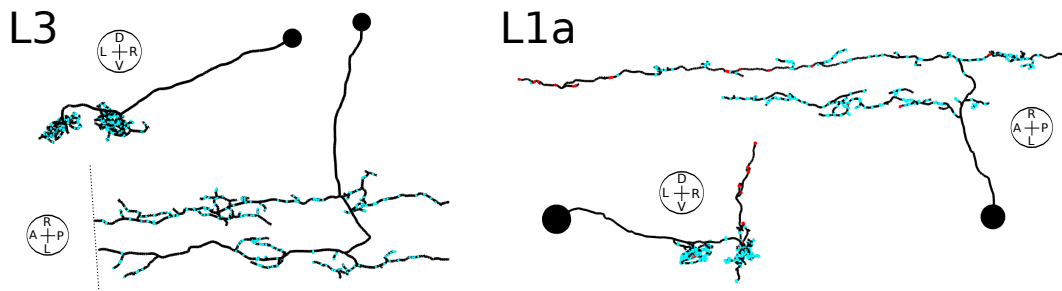
Supplemental Figure 3.15.: **Ascending projection neuron: A09e**
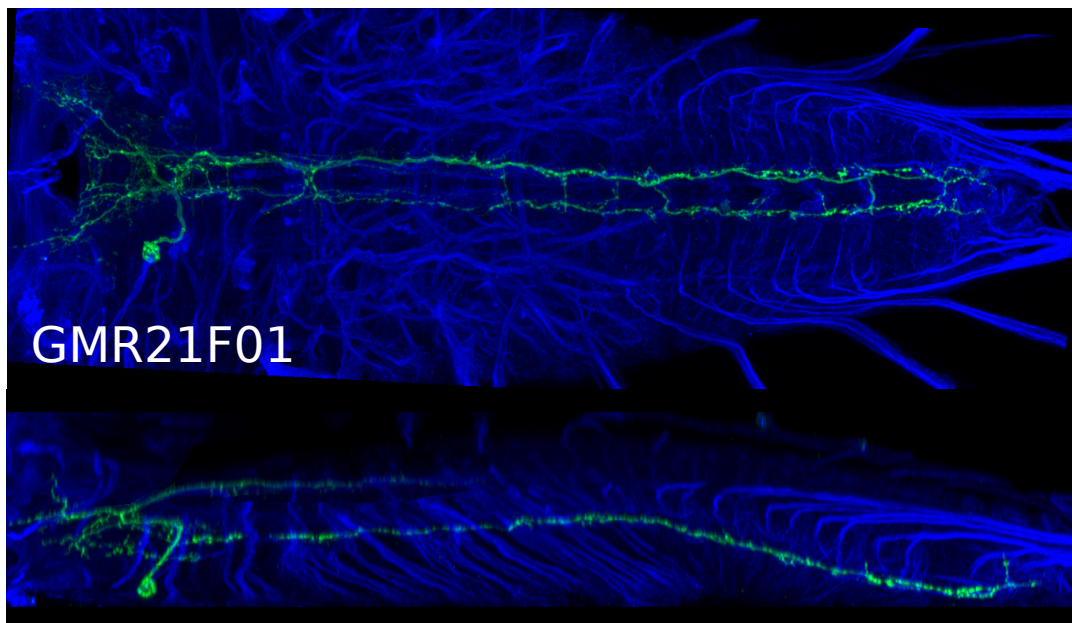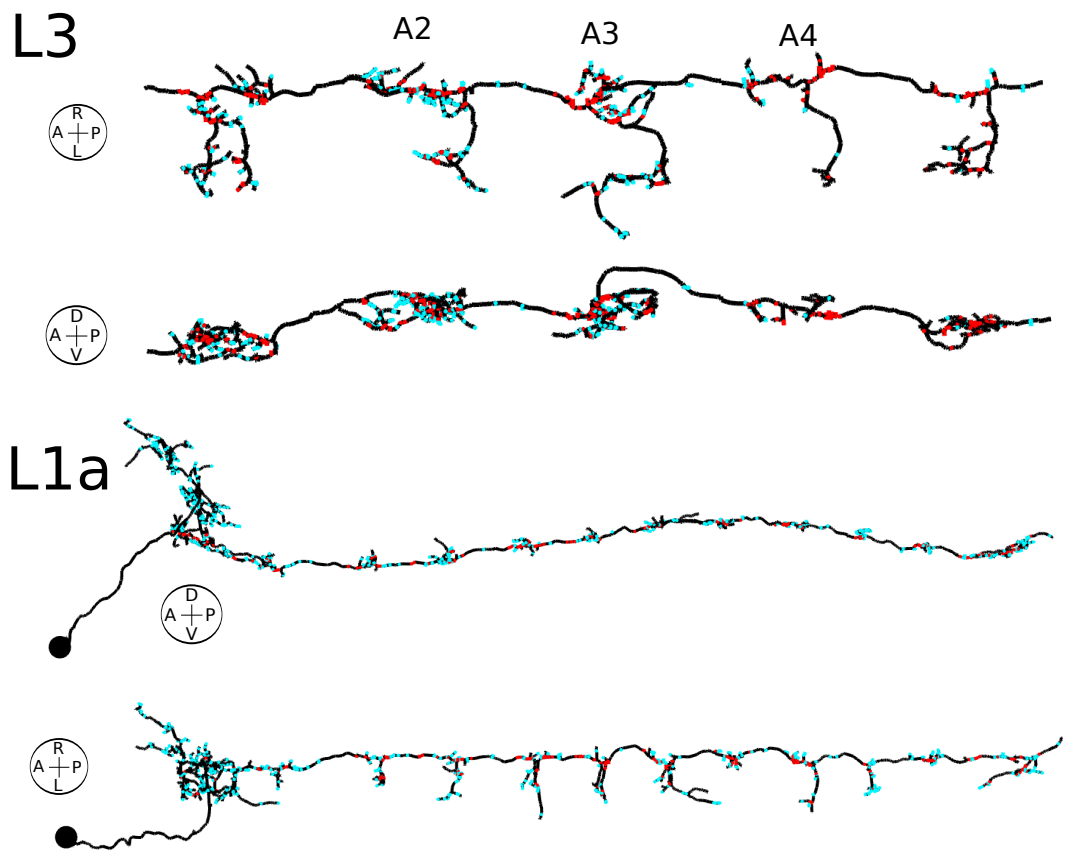
Supplemental Figure 3.16.: **Descending neuron: SelN138**

a) Dataset L1a
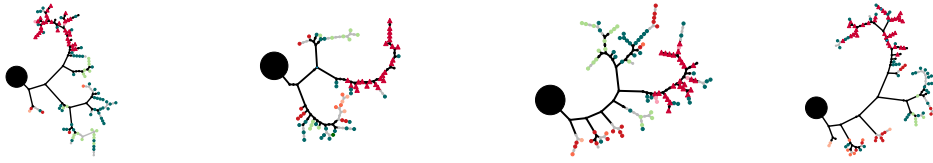
b) Dataset L1b
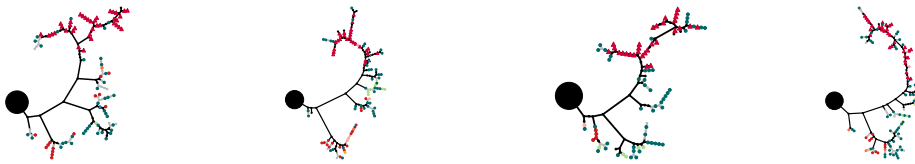
c) Dataset L3
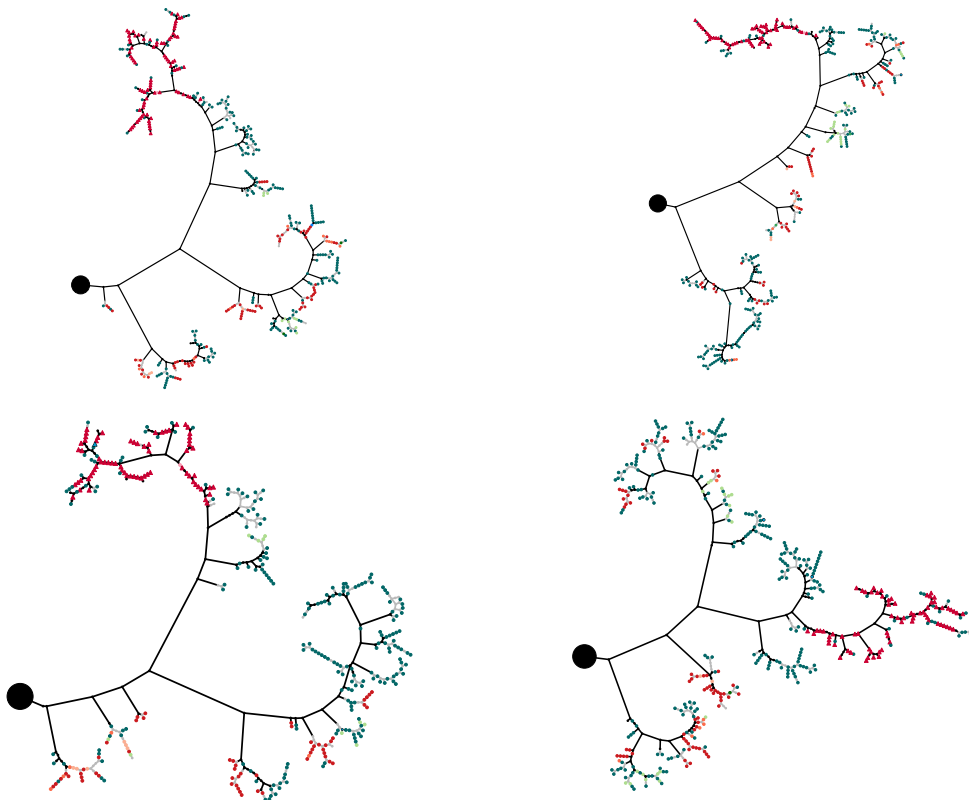
Supplemental Figure 3.17.: **Toplogical representation of Pseudolooper-3 and Pseudolooper-4.** 3D morphologies of neuronal arbors are shown in 2D. Red triangle show presynaptic sites, colored circles show postsynaptic sites. The color-code corresponds to the cell-type-specific colors previously used.

a) Dataset L1a

b) Dataset L1b
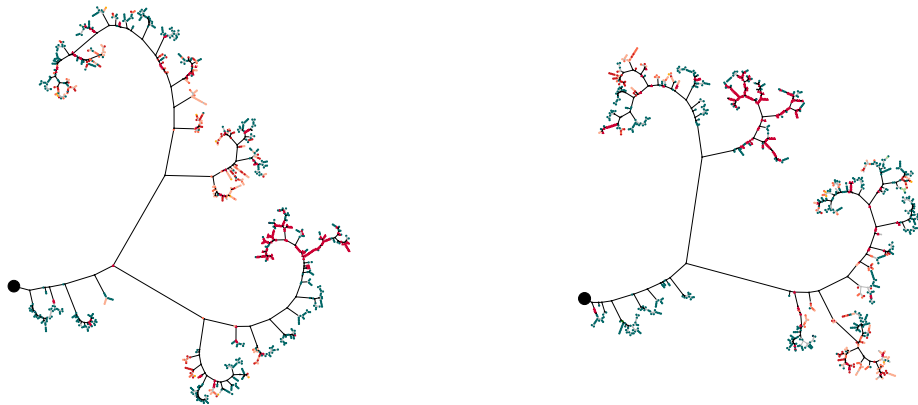
c) Dataset L3

Supplemental Figure 3.18.: **Toplogical representation of Down-and-back-1**

a) Dataset L1a

b) Dataset L1b
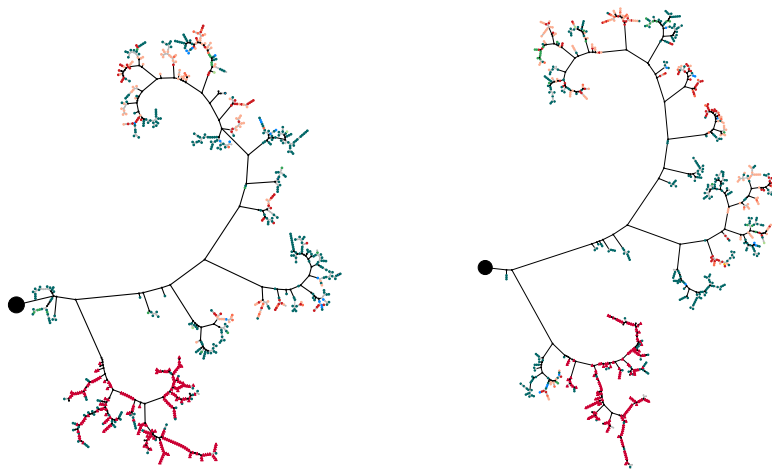
c) Dataset L3

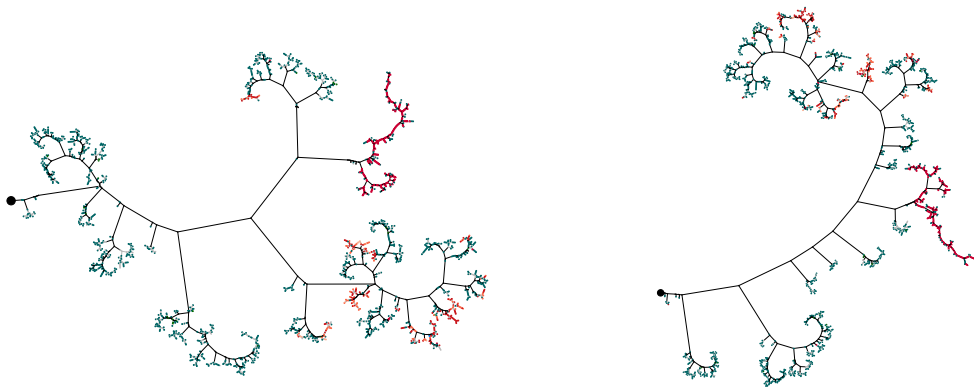Supplemental Figure 3.19.: **Toplogical representation of Chair-1**

a) Dataset L1a
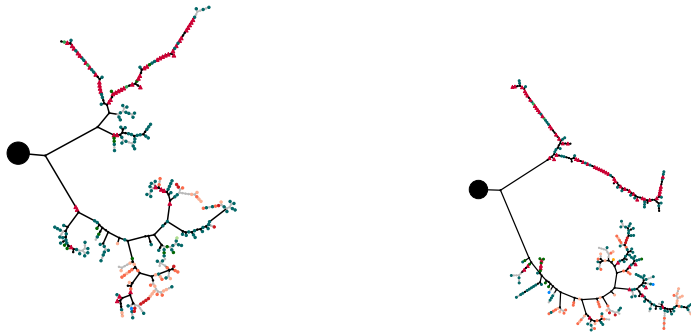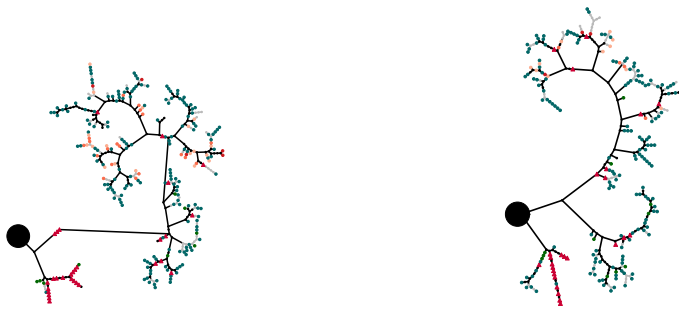

b) Dataset L1b

c) Dataset L3

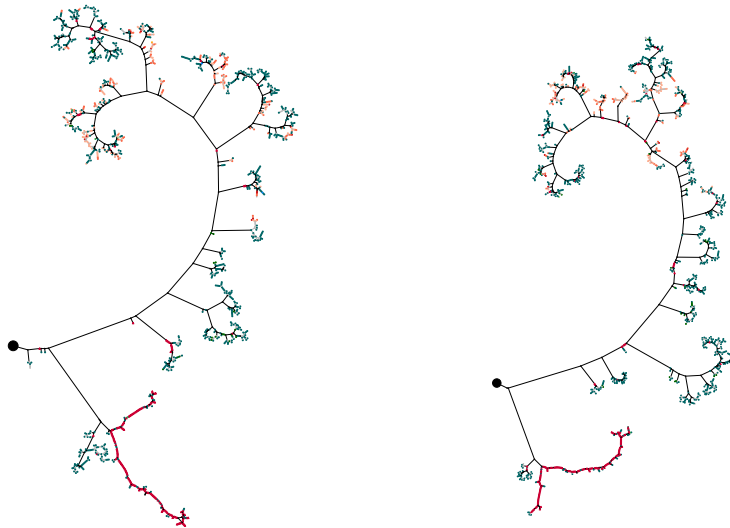Supplemental Figure 3.20.: **Toplogical representation of Basin-2**
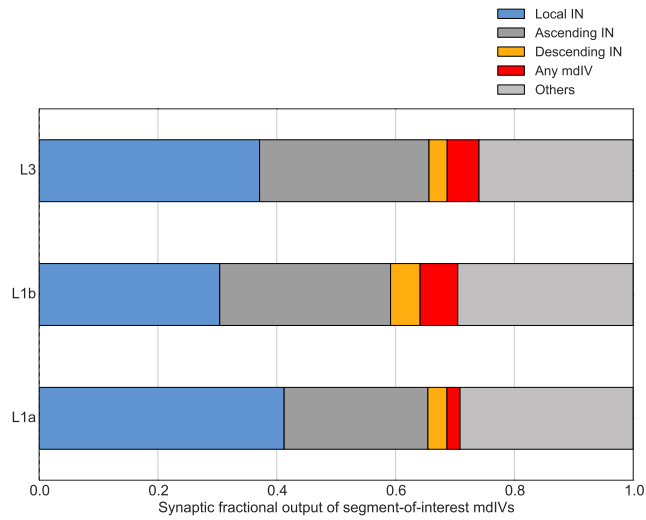
a) Dataset L1a

b) Dataset L1b
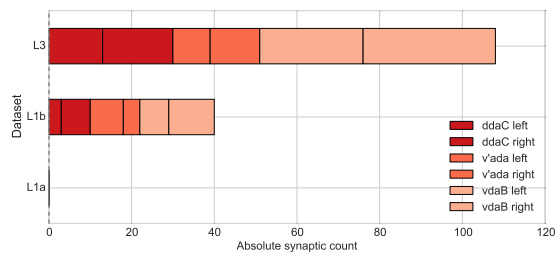
c) Dataset L3

Supplemental Figure 3.21.: **Toplogical representation of Basin-4**

a) Fraction of synaptic outputs from nociceptors

b) Nociceptor - Nociceptor connectivity

Supplemental Figure 3.22.: **Connectivity from nociceptors to different groups of downstream neurons** a) Fraction of synapses from nociceptors to different groups of interneurons (local, ascending, descending), as well as to other nociceptors (mdIV) and fragments (fragments that leave the volume or at a gap, ambiguous fragments). b) Inputs from other nociceptors displays no specificity to subsets of nociceptor types.
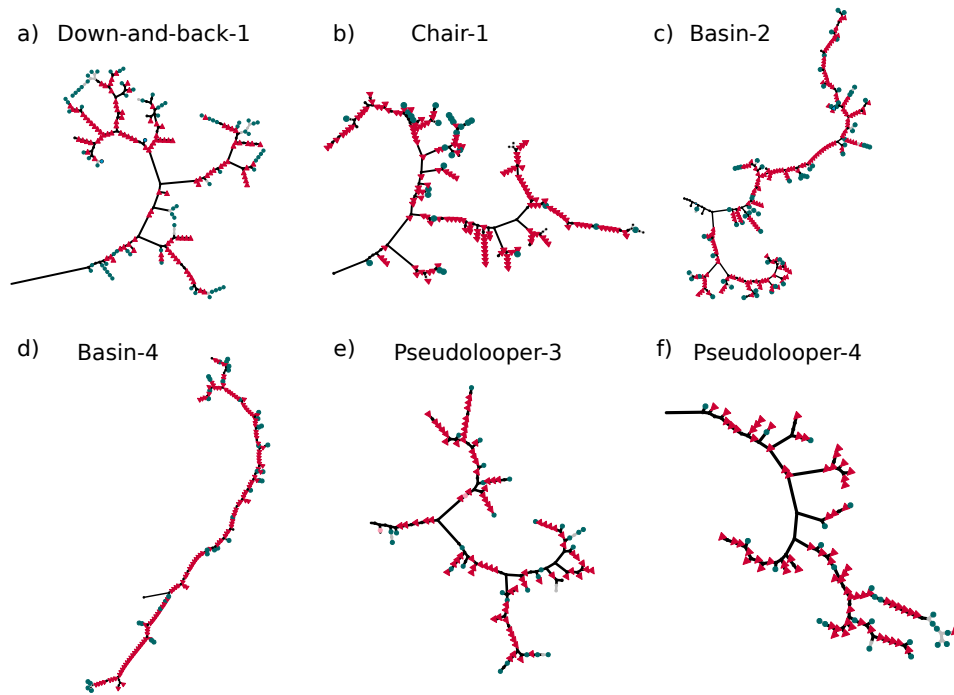
Supplemental Figure 3.23.: **Postsynaptic site distributed on axonal output arbors**
All six local interneuron types show occurences of post-synaptic sites on their axonal output arbors.

SkeletonID	Upstream neuron	Class	Count	Class Sum
1717121	A09a_a3l Basin-2	LocalIN	20	
1771441	A09a_a3r Basin-2	LocalIN	24	
1570360	A09a_a4l Basin-2	LocalIN	6	
405384	A09b_a3l Basin-1	LocalIN	31	
1735814	A09b_a3r Basin-1	LocalIN	38	174
1725730	A09c_a3l Basin-4	LocalIN	16	
1868541	A09c_a3r Basin-4	LocalIN	16	
1591495	A09l_a3l Down-and-back-1	LocalIN	5	
1591487	A09l_a3r Down-and-back-1	LocalIN	13	
1631566	A09l_a4r Down-and-back-1	LocalIN	5	
1629316	ddaC_a2l	mdIV	95	
1606096	ddaC_a2r	mdIV	68	
189519	ddaC_a3l	mdIV	253	
1591472	ddaC_a3r	mdIV	280	
1748751	ddaC_a4l	mdIV	24	
1755633	ddaC_a4r	mdIV	7	
478181	v'ada_a2l	mdIV	22	
184591	v'ada_a2r	mdIV	23	1752
1591464	v'ada_a3l	mdIV	152	
1591477	v'ada_a3r	mdIV	156	
1748756	v'ada_a4l	mdIV	53	
1813877	v'ada_a4r	mdIV	58	
1591468	vdaB_a3l	mdIV	207	
1870224	vdaB_a3r	mdIV	237	
1757231	vdaB_a4l	mdIV	49	
1755250	vdaB_a4r	mdIV	68	
935736	A02o_a3r Wave-1	OtherIN	6	
389760	A02o_a4l Wave-1	OtherIN	9	
1573859	A02o_a4r Wave-1	OtherIN	9	
368291	A09q? classIII-related projection left	OtherIN	51	
407348	A09q? classIII-related projection right	OtherIN	48	204
487018	classIII-related descending sog left	OtherIN	6	
1522894	classIV-related descending sog left	OtherIN	7	
387817	TePn05 classIV-related projection C left	OtherIN	36	
21291	TePn05 classIV-related projection C right	OtherIN	32	
1579257	putative class I	otherMd	27	
39008	putative class II	otherMd	74	
1673960	putative class II	otherMd	72	
37587	putative class II	otherMd	50	
443702	putative class II	otherMd	18	
200465	putative class III	otherMd	123	
1585339	putative class III	otherMd	87	
602968	putative class III	otherMd	85	
1570917	putative class III	otherMd	82	
1682251	putative class III	otherMd	72	900
441941	putative class III	otherMd	55	
42746	putative class III	otherMd	30	
1571883	putative class III	otherMd	11	
1425228	putative md class IV	otherMd	15	
1629549	putative md sensory	otherMd	37	
1580065	putative md sensory	otherMd	26	
1618945	putative md sensory	otherMd	24	
1629555	putative md sensory	otherMd	12	
	Others			2958
	Sum			5988

Supplemental Figure 3.24.: **Upstream neurons of the set of 12 local interneurons in the abdominal segment A3 in third-instar larva**. Upstream neurons other than localIN and mdIV classes are not reviewed and synaptic count numbers are preliminary. However, a relative strong contribution of inputs from other multi-dendritic sensory neurons onto preferred, such as class II and class III, are apparent. These suggest a partially shared central circuitry of local interneurons downstream of nociceptive sensory neurons with other sensory neurons found to be necessary and sufficient for eliciting behavioural touch responses (Tsubouchi et al. 2012).

4

COMPARATIVE CONNECTOMICS OF NEURONAL AND SYNAPTIC PROPERTIES

4.1 ABSTRACT

Over the course of animal development, neural circuits change by virtue of neuronal morphogenesis and synaptogenesis. The growth of neuronal arbors and the formation of new synapses is a carefully regulated process subject to various developmental and metabolic constraints. At any specific time point, the exact arbor geometry and the distribution of synapses results from these constrained, regulated processes. Moreover, the neuronal cytoskeleton composed of microtubules is essential for the stabilization of growing arbor geometries. Using recent advances in volumetric electron microscopy and neural circuit mapping, we mapped whole neurons in the small nervous system of *Drosophila melanogaster* larva and quantified morphological and synaptic properties at the ultrastructural level. In order to compare a variety of properties, we reconstructed and annotated six different neuronal cell types at early and late stages of the larval development lifecycle.

We show that the scaling of dendritic interneuron arbors is consistent with a previously described power law. Our data suggests that the power law relationship also holds for scaling of synaptic inputs per dendritic cable length. We demonstrate how to achieve robust cell-type clustering that is independent of developmental stage by considering a population-based normalization using only the number of synaptic inputs and outputs. We found cell-type-specific differences in arbor growth rate, but a number of morphological and synaptic features remain invariant across development. Dendritic twigs are microtubule free, distal arbor processes that were found to host between 80% and 90% of interneuron dendritic inputs at both stages and across different cell types. The proportion of twig cable to non-twig cables increases from 60% to 70% at late stage consistently for all cell types, while the density of twigs per cable, or the input density on twigs or backbone remains similar. We found the Strahler Index at the cell body location to assume a highly robust value for all considered interneurons, and increase by only one order from early to late stages.

Our dataset provides important constraints for detailed computational models of neuronal development and synaptogenesis. Furthermore, automated circuit reconstruction methods can employ the annotated neurons and synapses as a training and test datasets for automated whole circuit mapping. The quantitative distributions of morphological and synaptic features can further serve to guide proofreading.

4.2 INTRODUCTION

The development of neural circuits is a process of great complexity. The self-assembly of neural circuits is specified by genetic programs and unfolds through various stages influenced by intra- and extracellular environments. Cells proliferate, migrate and differentiate into neurons; neurites extend from the cell body and are specified as axons and dendrites; axons are guided to their target regions where they form synapses with other neuron's dendrites. The growth and properties of dendritic arbors has been extensively described and compared between neuronal cell types (Wittenberg et al. 2007). Derived from such studies, mathematical and computational modeling approaches have contributed to our understanding of how cellular and molecular processes determine global morphological features such as arbor geometry (Ooyen 2011).

Dendritic development involves a complex interplay between intracellular and extracellular signals. Both of these signals are integrated at the dendritic growth cones and modulate growth trajectories and branching events, resulting in specific dendritic shapes and sizes. These processes have been investigated using phenomenological, computational models of growth cone dynamics and their branching probabilities (Van Pelt et al. 1997; Pelt et al. 2002). Branching probabilities are determined by three main factors: a factor that decreases with developmental time, a factor that changes with the growth cone's topological distance to the soma, and a factor that decreases with the total number of growth cones of the neuron. To estimate parameters for such models, most approaches rely on reconstructions of neuronal morphology performed at light microscopic resolution. However, at electron microscopic resolution, intracellular structures of neurons, such as microtubules, can be visualized and mapped in the context of circuit development and neuronal growth.

The major structural components of dendritic morphogenesis are actin and microtubules (Jan et al. 2010). The distal filopodia-like growth cones are rich in actin and are essential to guide neurite growth based on extracellular cues (Smith 1988). Tubulin is produced in the cell body and is transported to growth cones where it is assembled into microtubules (Bamburg et al. 1986). Intracellular calcium concentration influences microtubules assembly and contributes to either neurite elongation or branching (Audesirk et al. 1997; Hely et al. 2001). Microtubules have been shown to invade the actin-based filopodia and thereby stabilize dendritic arbor shape (Hu et al. 2008; Jaworski et al. 2009). In addition to their structural support function, they were also implicated in long-term memory storage (Smythies 2015). Dendritic spines in vertebrate have recently been linked to the filopodia-like, actin-based structures in *Drosophila* (Leiss et al. 2009b). Thus, the study of *Drosophila* dendritic morphogenesis and microtubules compartmentalization could provide important insights into the constraints of neuronal development and their cell type-specific differences.

In this study, we map the morphology, synaptic distribution and microtubuli-skeleton of single, identified neurons of *Drosophila* larvae using volumetric serial section transmission electron microscopy. We reconstructed six pairs of left and right homologous neurons of different type in the *Drosophila* abdominal ventral nerve cord at two developmental stages, both in early, first-instar and late, third-instar

developmental stages. The unique identifiability of single neurons in an individual allowed matching of neurons between developmental stages for comparison. With resolution afforded by EM, we could distinguish the microtubules cytoskeleton of each neuron and annotate their microtubuli termination points in each of 36 interneurons (N=3500). The cytoskeleton annotation allows to extensively characterize a variety of neuronal and synaptic properties across development. We aimed at comparing similarities and differences of these features between developmental stages to uncover details of circuit development with relevance for models of neuronal and circuit development.

4.3 RESULTS

We reconstructed six different neuron types which are downstream of nociceptive sensory axons: Down-and-back-1, Chair-1, Basin-2, Basin-4, Pseudolooper-3 and Pseudolooper-4. In these neurons, all presynaptic and postsynaptic sites were also annotated. Additionally, we annotated different types of compartments in these neurons. Neuronal arbors were segregated into axonal compartments that host mostly presynaptic sites, and dendritic compartments with mostly host postsynaptic sites. A second, orthogonal compartmentalization of *Drosophila* neurons distinguishes neurites into microtubules containing processes (backbone), and distal, microtubules-free processes (twigs). Both axonal and dendritic compartments comprise backbone and twig processes. For each distinct neuronal compartment and their combination (e.g. dendritic backbone, dendritic twigs, etc.), we measured a number of features (Table 4.1).

The body length of larvae increases from less than a milimeter in first-instar (L1) to several milimeter in third-instar (L3) larvae in a period of less than 48 hours. During this time, neurons in the central nervous system grow considerably. We found neuron-type-specific differences in growth rate from L1 to L3 with scaling of total cable length of neurons by 3.75x Pseudolooper-3/4; 3.9x Basin-2; 3.73x Basin-4; 3.47x Down-and-back-1; 3.45x Chair-1. The length of arbor cable increases with the amount of synaptic inputs and outputs of the neurons. Across all neurons, the total amount of inputs scales linearly as a function of total cable length with slope 0.77 (R^2=0.97, p=2.8e-22). The total amount of outputs as a function of cable length scales linearly with slope 0.57 (R^2=0.94, p=2.13e-17) (Fig. 4.1).

For optimal dendritic branching, Cuntz et al. 2012 proposed that dendritic scaling is predicted by a 2/3 power-law that relates total cable length to the number of dendritic branch points. Cuntz et al. 2012 showed that such a power-law scaling indeed holds for a large number of neurons across a large number of species. We tested whether our data is consistent with this power-law scaling for measures of dendritic complexity involving total cable length and the number of branch points. A power-law scaling was already found for *Drosophila* larva peripheral multi-dendritic classIV nociceptor dendrites (Iyer et al. 2013). We also found our data to be consistent with a power-law scaling for dendritic arbors (Fig. 4.2a,b).

We then measured global neuronal properties which are independent of the absolute size of neurons and compared them across development. Previous studies found

feature	description
Raw cable length	The total cable length as the sum of distances of individual sections of the compartment skeleton. Unit is in microns
Nr. of inputs	The total number of postsynaptic sites on the compartment
Nr. of presynaptic sites	Total number of annotated T-bars (vesicles release sites)
Nr. of outputs	Single presynaptic sites target multiple postsynaptic sites. The total number of outputs is the total number of target postsynaptic sites
Nr. of branch points	The total number of branch points of the skeleton
Volume of arbor extent	The volume of the convex hull around all compartment skeleton nodes
Volume of input site extent	The volume of the convex hull around all postsynaptic sites
Volume of presynaptic site extent	The volume of the convex hull around all presynaptic sites
Soma size	The radius of the cell body

Table 4.1.: **Set of features measured for each compartment** For each neuronal compartment, a set of neuroanatomical features were computed.
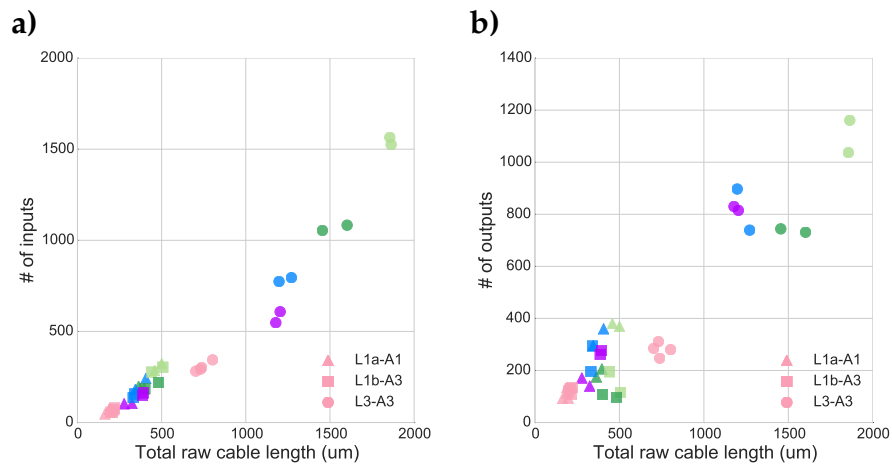
Figure 4.1.: **Scaling of inputs and outputs as a function of total cable length** a,b) Scaling of the left and right instances of local interneuron types. The number of inputs and outputs as a function of total cable length. Pseudolooper-3/4 (bright red); Chair-1 (purple); Down-and-back-1 (blue); Basin-4 (dark green); Basin-2 (bright green).

that the centrifugal order of dendritic segments may influence the probability of dendritic branching (Ooyen 2011). The centrifugal order counts the number of branch
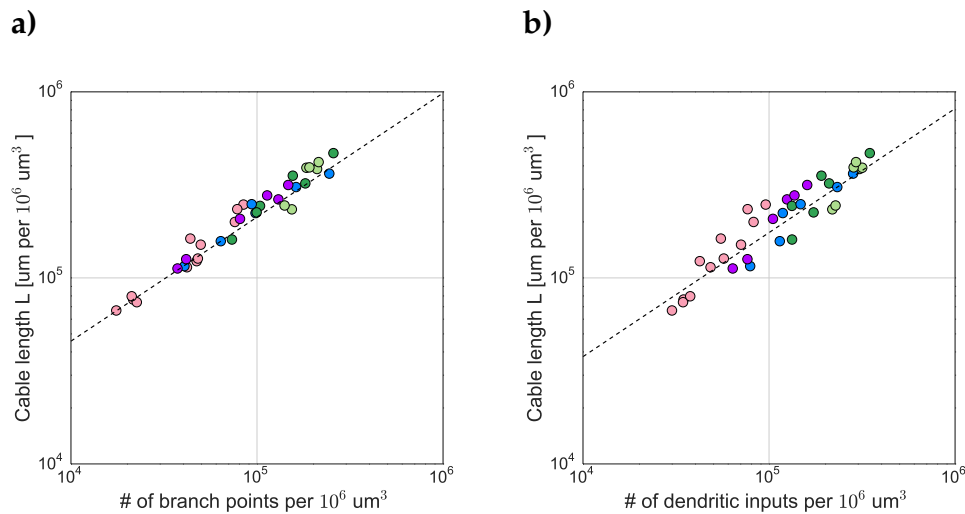
Figure 4.2.: **Neuronal scaling consistent with power-law** Scaling law of neuronal properties across development is consistent with previously described power-law scaling across a large range of neurons in a variety of species (Cuntz et al. 2013). A power-law relationship holds for both a) total cable length and b) the number of dendritic inputs.

points along the path from a dendritic segment to the cell body. The mean centrifugal order has been shown to be related to the mode of growth of tree structures (Pelt et al. 1989) and has also been applied to characterize dendritic arborizations (Uylings et al. 2002). Here, we compute the Strahler number, a similar measure of tree complexity, for neurons at the cell body location. We found this index to be a reliable measure that is highly consistent across the variety of neuron-types with differing morphologies at a specific developmental stage (Fig. 4.3). The Strahler number increases on average by one order, from 6 in L1 to 7 in L3, across the considered population of neurons. For comparison, we also show that the Strahler number computed for six nociceptor sensory axons at the neuropil nerve entry point. The values remain at 5 in both 1st and 3rd instar stages.

Another size-independent measure considers only the input-output structure of a neuron. The ratio of the number of synaptic inputs vs. outputs may be a simple proxy to characterize a neuron type independent of its connectivity to other neurons. By a normalization of dendritic input and axon output counts with the population mean and standard deviation, we found an embedding space that clusters the studied neuron types well. It has the additionally property of clustering neurons independent of their developmental stage (Fig. 4.4a). This clustering becomes more discriminative if we consider the number dendritic outputs instead of axon outputs (Fig. 4.4b). The results suggest that the normalization based on a neuronal population measure can subtract the effect of developmental stage and assignment each neuron to its type by using a simple measures based only on neuronal synaptic input-output counts.

		L1a-A1	L1b-A3	L3-A3
Interneurons	Pseudolooper-3	5	4	6
	Pseudolooper-3	6	6	7
	Pseudolooper-4	6	6	6
	Pseudolooper-4	5	5	6
	Basin-2	6	6	7
	Basin-2	6	6	7
	Basin-4	6	6	7
	Basin-4	6	6	7
	Down-and-back-1	6	6	7
	Down-and-back-1	6	6	7
	Chair-1	6	6	7
	Chair-1	6	6	7
Nociceptors	ddaC	5	5	6
	ddaC	5	5	5
	v'ada	5	5	5
	v'ada	4	5	5
	vdaB	5	5	5
	vdaB	5	5	6

Figure 4.3.: **Strahler number as robust global neuron feature** The Horton-Strahler number is computed at the cell body location for interneuron, and at neuropil entry points for nociceptive sensory axon.
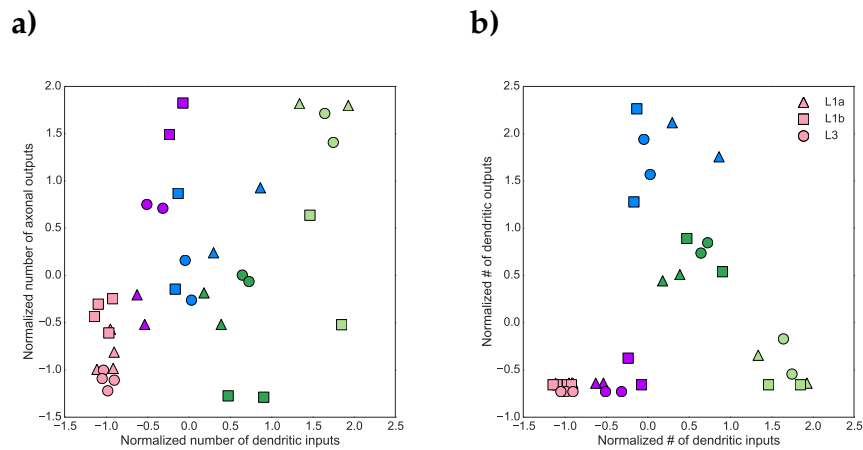
Figure 4.4.: **Clusters of local interneurons based on normalized synaptic input-output relationship** a) A low-dimensional feature space independent of absolute neuronal size segregates cell types in two developmental stages into clusters. The normalization of the number of inputs and outputs is computed by subtracting the population mean divided by the population standard deviation of the respective feature. b) The clusters based on the number of dendritic outputs appears to segregate cell types better than using the number of axonal outputs.

Comparison of twig and backbone features across development

We then characterize the changes of neuronal compartments across development. We recently showed that the amount of backbone and twig cable in *Drosophila* neurons is well preserved between neurons in different regions of the larva nervous systems, as well as in adult *Drosophila* (Schneider-Mizell et al. 2015). The absolute number of twigs increases from L1 to L3 for all neuron types, as well as the absolute amount of cable length, both for backbone and twig compartments (Fig. 4.5a,b). Consistent with the idea of microtubules invasion of twigs during development, we found that the number of branch points in the dendritic backbone increases in development (Fig. 4.5c).

a) Total dendritic twig and backbone cable length
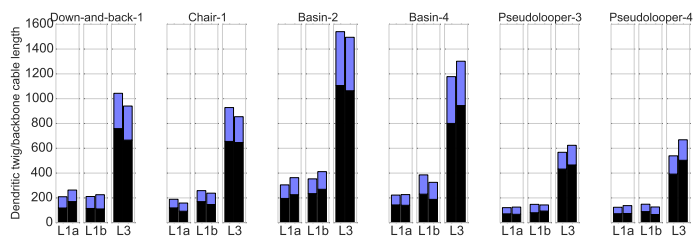
b) Number of dendritic twigs
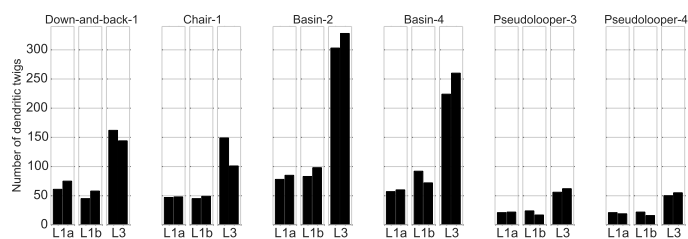
c) Number of branches of dendritic backbone
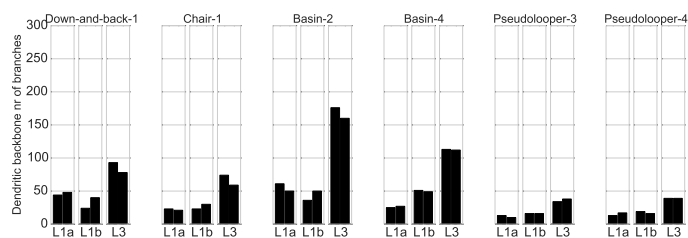
Figure 4.5.: **Absolute neuronal measures with respect to backbone and twig compartments** Left bars denote the left instance of the neuron within one abdominal ventral nerve cord segment, the right bar the right neuron instance. For each cell type, the plots show data from three datasets, from both early (L1a, L1b) and late (L3) developmental stages.

A number of backbone-twig features were found to be invariant across developmental stages and consistent between individuals of the same stage. If the amount of dendritic twig cable is normalized by the total amount of dendritic cable, we found that twig cable contributes about 60% of the total dendritic cable length in L1, and about 70% in L3 to the dendritic arbors. This proportion is consistent across different neuron types (Fig. 4.6a). Considering the number of dendritic twigs per dendritic backbone cable, we found this measure of twig density to vary between neuron types, but to be very consistent across developmental stages (Fig. 4.6b). The majority of dendritic inputs in *Drosophila* neurons was found to be onto twigs (Schneider-Mizell et al. 2015). We found this to be the case also across developmental stages with 90% of dendritic synaptic inputs onto twigs for Pseudolooper-3 and -4, and at about 80% in the other neurons (Fig. 4.6c).

a) Normalized dendritic twig cable length
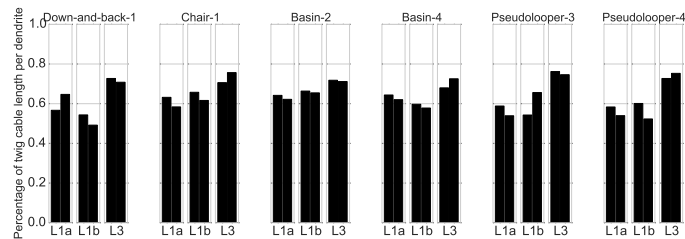
b) Number of dendritic twigs per dendritic backbone
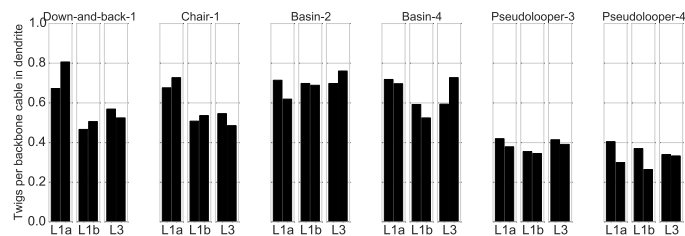
c) Percentage of dendritic inputs into twigs
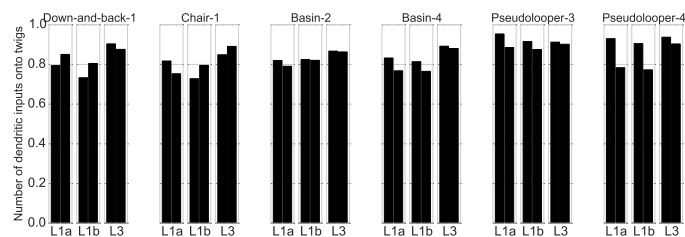
Figure 4.6.: **Normalized backbone and twig measures**

We have shown that the total number of inputs of a neuron scales with its total cable length. If we confine the same analysis to only the dendritic arbor of the neuron, we found that neuron-type-specific differences are more apparent (Fig. 4.7a). For instance, Basin-2 neurons have on average a higher number of inputs per dendritic

cable length than Pseudoloopers. These type-specific differences translate also to synaptic input density differences on twigs and backbones for different neuron types. We found a remarkable constancy in input density onto twigs (Fig. 4.7b). The synaptic input density on the backbone compared is lower to twigs (Fig. 4.7c). This difference can be explained by the fact that most inputs are onto twigs, but the amount of twig cable is about 60% (Fig. 4.6a). The difference in backbone input density are again cell type-specific and preserved across development to a high degree.
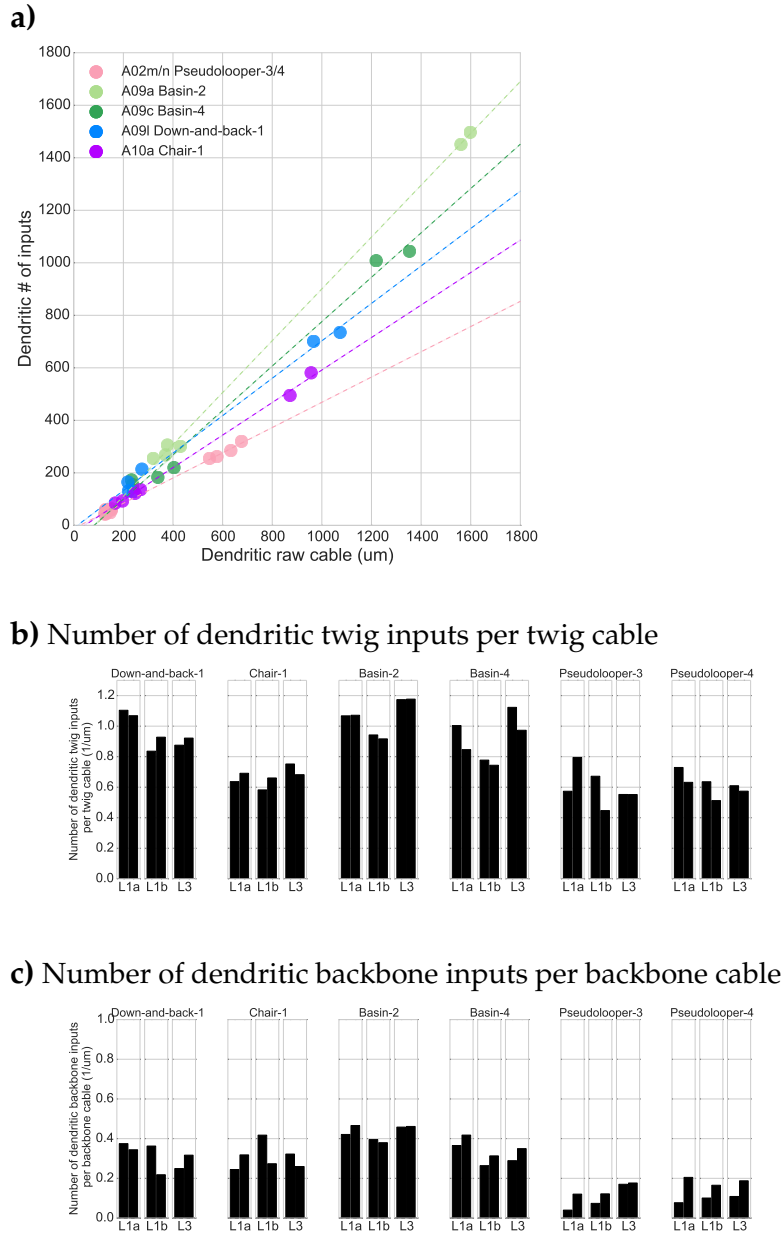
Figure 4.7.: **Constancy of density of dendritic input sites per cable.** a) Cell-type-specific differences in scaling of inputs as a function of total cable length b) Input density only on dendritic twigs. c) Input density only on dendritic backbone. All units are inputs/um cable length.

Quantitative comparison at the individual twig level

Interestingly, we found that the filopodia-like twigs not only occur in dendritic compartments of neurons, but also on the axonic output terminals, although with a lower frequency. Axonic twigs were found in all neuron types considered, and they differ in a number of characteristics. They are shorter, exhibit a lower branching complexity and receive fewer inputs than their dendritic counterparts. Although axonic twigs occurred in all studied interneurons, they are not a general feature of axonic terminals. The nociceptive sensory axon terminals receive substantial amount of inputs, but directly onto the presynaptic varicosities. They exhibit no twig processes. In the following analysis of twig properties, we have focused mostly on the more abundant dendritic twigs.

The question arises whether dendritic twigs are stereotyped entities with shared structural and synaptic features across development. The relationship of the 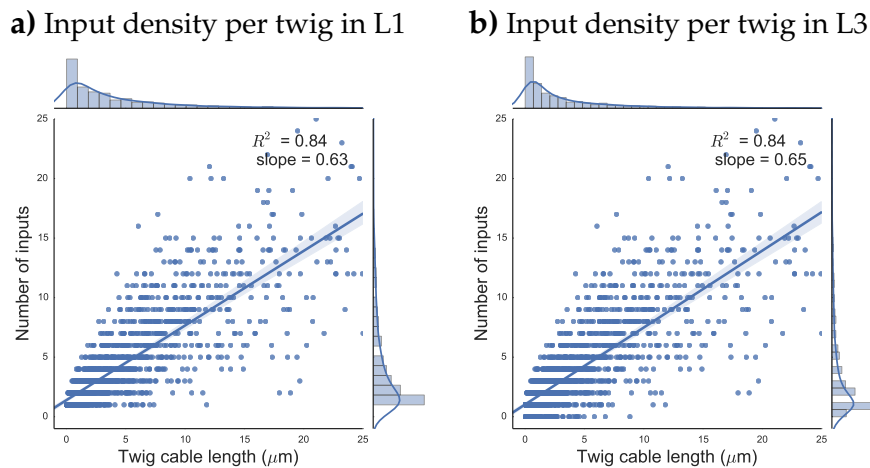number of synaptic inputs to twigs relative to its total cable length was proposed to be stereotyped (Schneider-Mizell et al. 2015). Here, we compared this relationship across all cell types segregated by developmental stage. We found a strong linear correlation between number of inputs and cable length of twigs at both early L1 and late L3 stages (Fig. 4.8).

a) Input density per twig in L1   **b)** Input density per twig in L3

Figure 4.8.: **Linear relationship between number of inputs and twig cable length is preserved between L1 and L3**. a) Each individual points denotes a twig, and the number of synaptic inputs per twig of different sizes is shown for all twigs in L1, and for the twigs in L3 in b).

This linear trend is also consistent for dendritic twigs analyzed independently for different cell types and developmental stages (Fig. 4.9). Differences exist in the slopes of the regression models. The slopes are an indication of average input density on the twigs. For instance, Basin-2 or Basin-4 twigs have on average a higher input density (close to 1 input/μm) than Pseudoloopers with lower input density. In the first-instar datasets (L1a and L1b), most of the twigs have a total cable length smaller than 10 μm, whereas in the third-instar L3 dataset, more larger twigs exist across all cell-types.
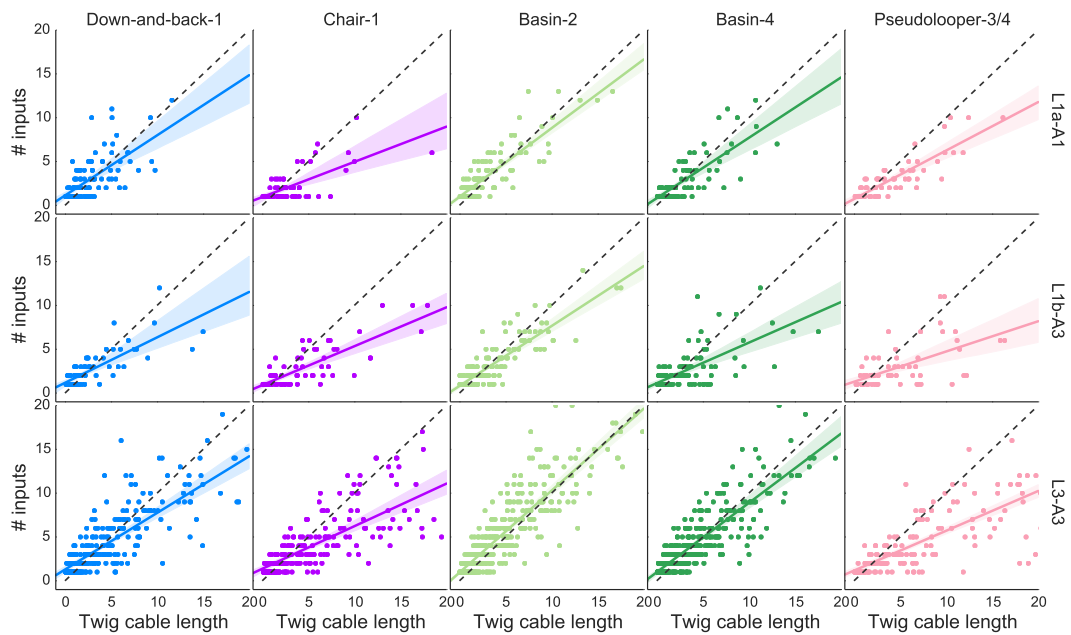
Figure 4.9.: **Cell-type-specific linear regression of number of inputs vs. twig cable length for dendritic twigs**. Confidence interval of 95% plotted with the regression line. Dashed line represents slope of value 1.0 corresponding to 1 twig input/μm cable length.

Next, we compared both absolute and topological twig features as well as characteristic synaptic input distributions between early and late stage interneurons. For each feature, cell type and dataset, we plot the cumulative distribution function (CDF). The shape of the CDF is a characteristic measure of the distribution of a particular feature across all the twigs independent of the absolute number of twigs. We use the two-sample Kolmogorov-Smirnov (KS) statistic to test whether we can reject the null hypothesis that both distributions are identical.

In Schneider-Mizell et al. 2015, the authors analyzed the distance of twig inputs from the twig root node, i.e. the location where the twig connects to the neuron's backbone. They found that for a set of motor and premotor neuron twigs, more than 91.4% of all input synapses lay within 5 um of the twig base and backbone. Here, we found differences between L1 and L3 datasets in the CDF of twig input depth from twig base for all dendritic twigs across cell types considered (Fig. 4.10). In dataset L3, about 80% of inputs lay within 5 um of the twig base, and in the L1 dataset 93.8% of inputs. We can reject the null hypothesis that both distributions are identical (KS statistic: 0.216, p<0.0001), and thus the difference is significant.
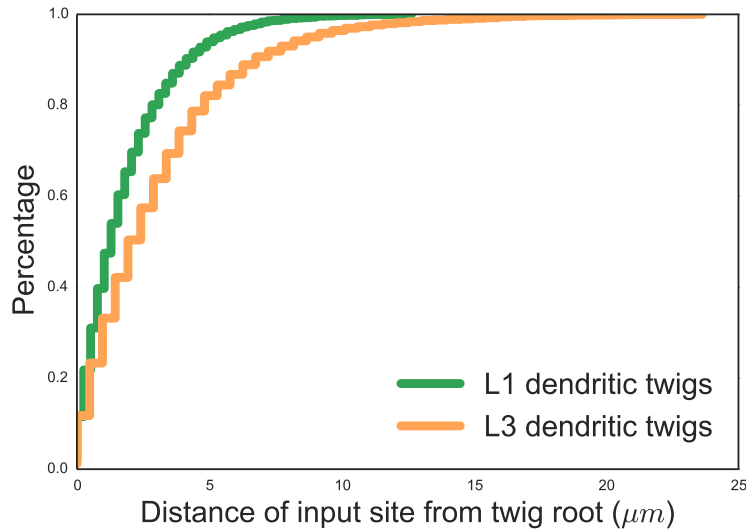
Figure 4.10.: **Cumulative density function for the distance of twig input sites from twig root in L1 and L3**. In L3, approximately 80% of all synaptic inputs are within 5μm of the twig root, whereas for all L1 dendritic twigs, the fraction of inputs within the first 5μm from the twig root is larger.
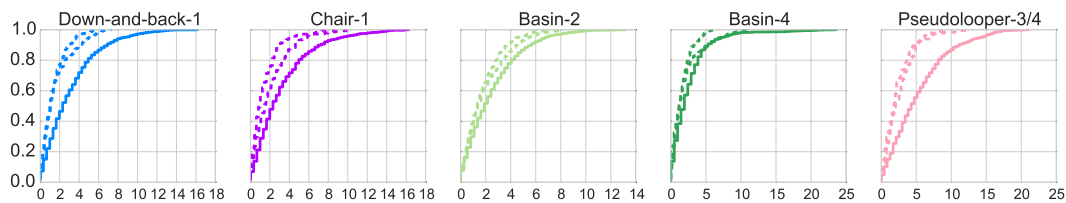
Figure 4.11.: **Comparison of distance of input site from twig root in L1 and L3 segregated by cell type**. Dashed lines correspond to L1a and L1b datasets, continuous line to the L3 dataset.

We found slight differences between cell types in the distance of twig inputs from the twig root (Fig. 4.11). For all cell types, the differences of the distributions between L1 and L3 are significantly different (KS tests with all p<0.001). However, we observe that the differences are not as pronounced for Basin-2 and Basin-4 as for the other cell types.

We further compare other twig features between two developmental stages (Fig. 4.12). In particular we compare the total raw cable length of the twig, the total number of inputs per twig, the Strahler Index at the twig root (a measure of branching complexity) and the input density (the number of inputs per cable per twig). For the total cable length of individual twigs, both Basin-2 and Basin-4 cell types are not significantly different between developmental stages (Fig. 4.12a). For the other cell types, Chair-1, Down-and-back-1 and Pseudoloopers, the total size differs significantly between L1 and L3 (KS test, p<0.001). There exists a larger fraction of larger twigs in L3 than in L1. The number of inputs per twig changes significantly

from L1 to L3 at the 5% level except for Basin-2 ($p<0.03$, $p<0.24$ for Basin-2) (Fig. 4.12b). The Strahler Index at the twig root shows a high degree of similarity across cell types in both L1 and L3 datasets (Fig. 4.12c). Across all cell types, more than 90% of all twig root nodes have Strahler Index 1 or 2. The CDF plots show that in L3, the fraction of twigs with Strahler index 3 at the root node increases slightly. In terms of input density, no significant differences could be found between cell types and between datasets L1 and L3 at the 5% level (Fig. 4.12d).

a) Total raw cable length

b) Number of inputs per twig

c) Strahler Index at Twig Root

d) Input density (inputs per cable)
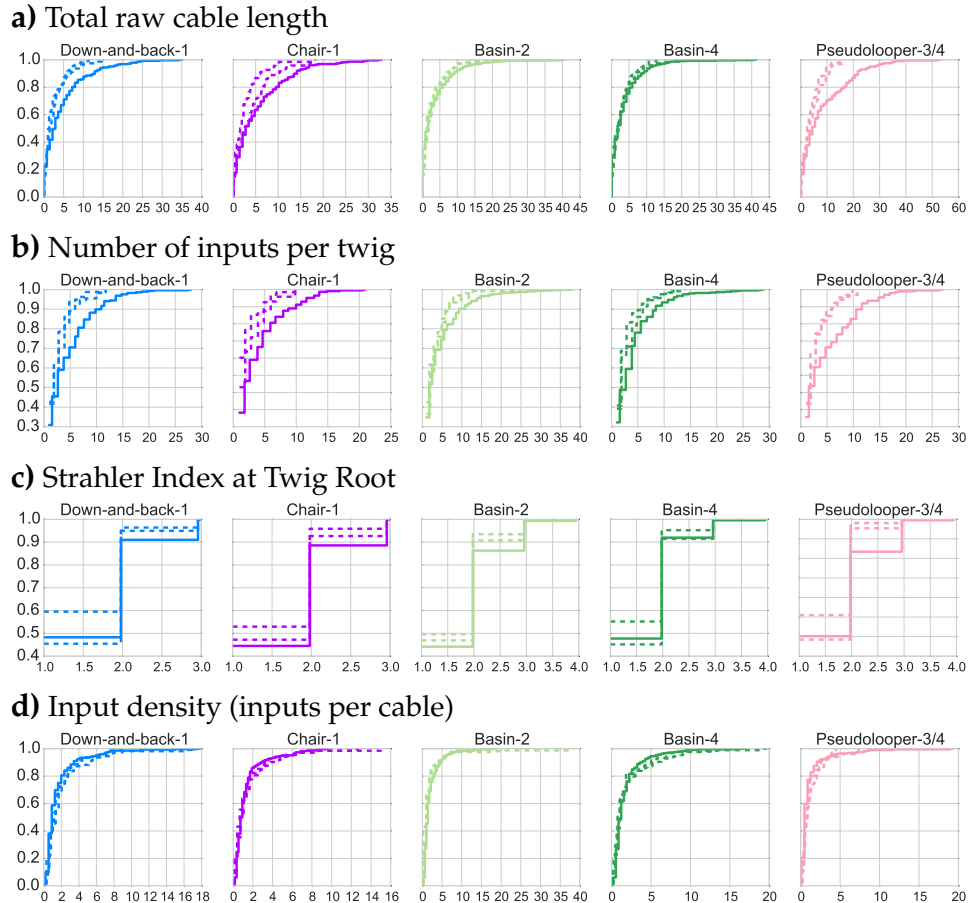
Figure 4.12.: **Comparison of dendritic twig features in L1 and L3** A number of features of dendritic twigs shown as cumulative distribution plots. The dashed lines refer to the first-instar datasets, the solid line to the third-instar dataset.

We found the input density per twig to be highly preserved, yet there is a variation of absolute twig sizes. Patterns in the specific location of inputs might therefore be masked by the different sizes of twigs. We therefore calculated the maximum depth of a twig as the maximum cable distance from twig root to any of its leaf nodes. We then used this twig depth to normalize the distance of input sites from the twig base.

We found, consistent with our experience from reconstructing twigs, that the distal termination location at twig leaf nodes shows a higher number of inputs. In Figure 4.13, we plot the input location normalized by twig depth (between 0 and 1) for all

the twigs separately for dataset L1 and L3. The increase of inputs at distal twig input sites is visible in the plot. In agreement with the constant input density across the rest of the twig arbors, the histogram shows similar levels of inputs across the rest of the twigs. Visualized as a cumulative distribution function, the patterning is consistent for each cell type (Fig. 4.14).
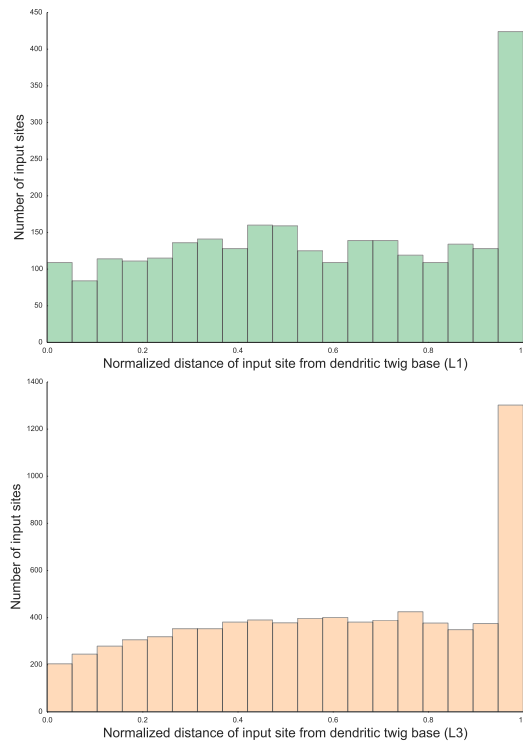
Figure 4.13.: **Input distribution on normalized dendritic twigs in L1 and L3.** When normalizing for the twig length, we observe a close to uniform distribution of input sites and a strong accumulation of input sites at the twig terminal.
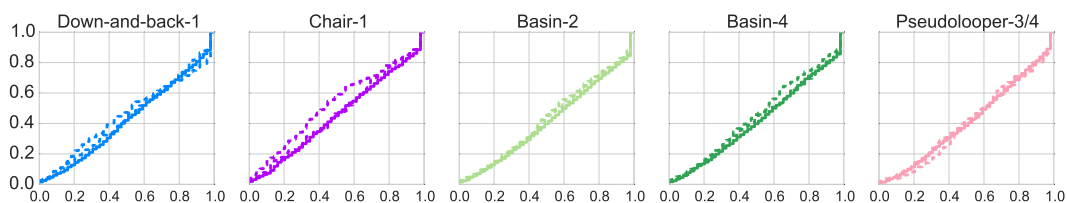
Figure 4.14.: **Comparison of the normalized distance of input site from twig base in L1 and L3 across cell types.** A high degree of similarity is found between first-instar and third-instar larvae.

We have only analyzed twigs with inputs, but we also found a small number of dendritic twigs without inputs. Across all the twigs annotated for all cell types, both early and late stage neurons have zero-input twigs (L1a, L1b, L3 with N=119, N=118, N=143). We found that twigs without inputs are considerably shorter than twigs

with inputs in both L1 and L3 datasets (Fig. 4.15). We observed that Pseudlooper zero-input twigs are in both L1 and L3 datasets larger than zero-input twigs of the other cell types considered.

a) Dendritic twig cable length
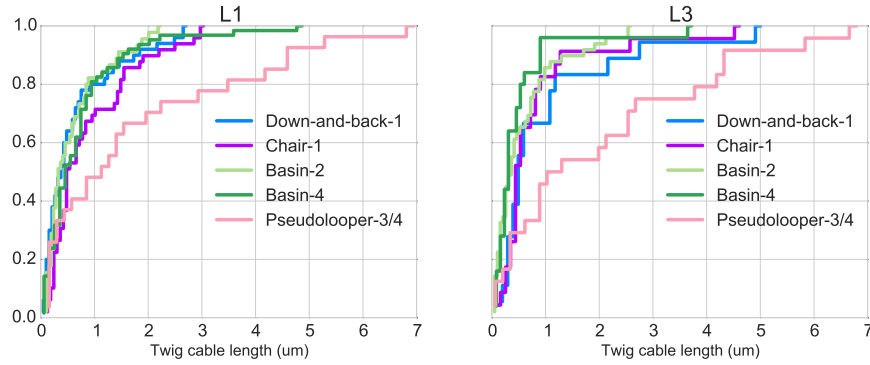
b) Dendritic maximum depth of twig
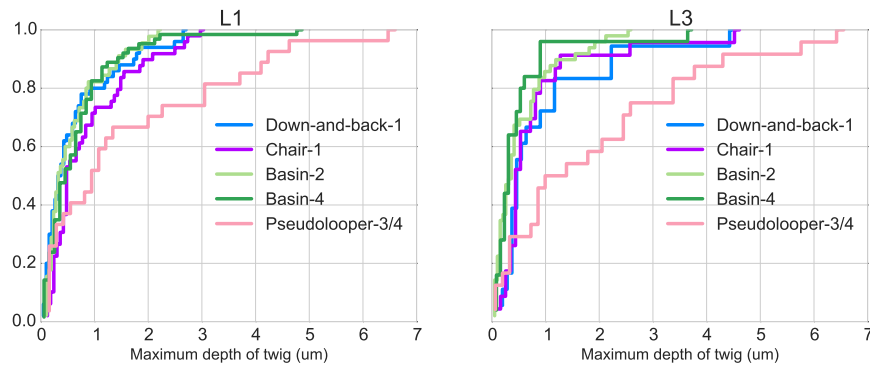
Figure 4.15.: **Cumulative distribution function of twigs without input synapses** a) Twigs without input synapses are considerable shorter than twigs with at least one or more input. Zero-input twig size is not different from L1 and L3. The Pseudolooper cell type shows considerably larger zero-input twigs than the other cell-types considered. b) When considering the maximum twig depth of twigs, the results are consistent with a).

146

What insights can the analogy of dendritic spines in vertebrate neurons and twigs in invertebrate neurons provide? The function of dendritic spines is still debated today (Malanowski et al. 2014), so it is difficult to try to establish exact functional equivalences. However, dendritic spines as well as twig processes are related to a neuron's balance between stability and flexibility. A representation of stability is the neuronal cytoskeleton which is energetically more costly to build and maintain, but is essential to ensure that the overall shape of neurons remains stable over time. This stability in shape confers stability to the overall wiring diagram by maintaining spatial proximity of neuronal processes that can form and maintain synaptic contacts within confined spatial domains. The exact number of synapses and synaptic contacts between neurons, however, might need to change as a consequence of development and experience. It is therefore important to maintain flexibility by which synaptic weights can be changed on faster time scales than is required for building cytoskeletonally-stabilized neuronal processes. The well-known dynamics of dendritic spines that extend and retract on short time scale might thus have an analogon in invertebrate twigs, whose dynamics has not been extensively studied *in vivo*.

A common viewpoint is that dendritic spines allow both biochemical as well as electrical compartmentalization within neurons (Shepherd 1996). Similarly, *Drosophila* twigs present a compartmentalization of neurons that could be biochemically relevant. Intracellular Ca2+ levels have been shown to modulate dendritic growth and the initiation of filopodia protrusions (Mattson et al. 1987; Lau et al. 1999; Duch et al. 2002). Twigs could act as a spatial constraints for the diffusion of Ca2+, thereby locally increase Ca2+ levels and contribute to the modulation and constraining of growth dynamics. Additionally, the observation of the bias of input synapses towards the tips of twigs may play a role in the stabilization and orientation of twig processes. Similarly, in growing zebrafish dendrites, dendritic filopodia have been shown to be stabilized by synapses (Niell et al. 2004). Through this fixation at the tip locations, twigs can remain oriented which could be important for subsequent microtubuli and appropriate dendritic backbone growth.

Our data suggests that dendritic growth proceeds by adding more stereotyped twigs to the dendritic backbone, while keeping the ratio of twig vs. backbone cable at a constant level of 60%-70%. The measurement of other twig features show a high-level of similarity across cell types, for instance with regards to twig size and branching complexity. The developmental process appears to target cell-type specific differences while maintaining an equal density of twigs per backbone arbor.

The morphology and electric properties of dendritic spines have been shown to influence dendritic integration and processing (Harnett et al. 2012). Similarly, *Drosophila* twigs could have a filtering role for synaptic inputs due to their compartmental nature and small neurite process diameter (data not shown). Coincident synaptic inputs could be necessary to activate a twig in order for postsynaptic potentials to propagate to the dendritic backbone for subsequent integration. Spontaneous postsynaptic potentials, however, occurring at random at single postsynaptic sites of

twigs could attenuate rapidly due to the small diameter of twigs. Therefore, twigs could act to filter spontaneous activity and confer increased robustness for backbone integration of synchronized synaptic inputs.

Beyond these biological results, the relationships and distributions we uncovered may provide useful heuristics and biological features to improve automated approaches to neural circuit reconstruction. The deviation from experimentally defined baseline distributions can be used to highlight errors in the results of automated neuron reconstruction algorithms and thus serve as a quality control measure and to guide proofreading. The presented features can be computed efficiently, once information on synaptic locations and microtubuli termination points are available.

As an example, the ratio of twig and backbone cable could be computed for an automatically reconstructed *Drosophila* larva neuron, and its value is expected to be within a confined range of values. Statistically significant deviations from this range indicate whether twig or backbone cable is missing in the reconstruction. Additional parameters, such as the densities of twigs per backbone or inputs per cable can provide further cues for proofreading and to find the locations of missing twigs. The usage of characteristic distributions to provide indications of deviations from expectations are *a posteriori* and allow to filter results and to guide proofreading. These features could also directly be integrated as *prior* knowledge into the optimization procedures of reconstruction algorithms.

Importantly, we have shown that for many of these baseline distributions, variation between cell types exist. Some features are very similar across developmental stage, and some vary linearly with developmental time. These constraints and variations have to be taken into account when utilizing these measures. The distributions may need to be calibrated for specific cell types and specific developmental stages.

As an example, we have considered the Strahler Index computed at the cell body location. We have shown that this index differs between sensory neurons and interneurons, but is highly stereotyped for different neuron types with a variety of shapes. With such robust measures that hold for a larger number of interneurons, they can serve as a simple heuristic to check the quality of a neuron reconstructions once basic parameters are known. Outliers that differ from the expected Strahler Index for a class of neuron types at a particular developmental stage can be highlighted for further proofreading. When automated neuron reconstructions can produce thousands of neurons, automated quality control and outlier finding will be of great utility.

TOWARDS SEMI-AUTOMATED NEURAL CIRCUIT MAPPING

5.1 INTRODUCTION

The challenges of automating the task of neural circuit mapping are not only algorithmic. The computer vision and machine learning communities focus on developing novel algorithms that aim at improving over state-of-the-art results in neuron reconstruction and synapse detection. These algorithms, however, are mostly tested and compared within these communities, and in the majority of cases with only small test dataset. There is no immediate incentive to make these algorithms work at scale for large EM datasets, for different types of EM datasets, or to integrate them into existing frameworks to make them useful to neuroscientists. Therefore, it is crucial to build bridges between research communities and provide sufficient amount of software engineering resources to implement solutions so that novel algorithmic developments can maximally impact the practice of neural circuit mapping.

Here, our motivation was to make a few contributions to these underattended areas. Over the long term, investments in these areas will be an important driver for progress in semi-automated neural circuit mapping at large scales.

PARAMETER TUNING FOR SPECIFIC DATASETS The parameters of reconstruction algorithms need to be tuned for specific EM datasets to accurately evaluate their performance. For this evaluation, it is necessary to train algorithms on ground truth datasets that were produced by experts. Obtaining high-quality ground truth datasets is tedious and laborious, but it is of utmost necessity, as supervised learning algorithms need a large amount of examples to be trained. We have produced such a pixel-based ground truth dataset in one of our high-quality EM volumes that includes labels for different classes of objects (membrane, synapses, mitochondria, glia). We used this to train and tune state-of-the-art image classifiers.

SOFTWARE INFRASTRUCTURE AND PIPELINE Due to the nature of neural circuit data, currently existing reconstruction algorithms can not be trivially parallelized from small to large datasets. Neurons extend their axonal and dendritic processes over large distances in volumes that can not be maintained in memory of normal desktop computers (>1TB). One approach to solve this problem is to split volumes into smaller chunks, segment these chunks individually and merge segmentation results to rebuild whole neurons. To successfully implement such a scalable, distributed system, expertise in distributed systems design, databases and high-performance computing is required. Here, we demonstrate the results of a prototype segmentation pipeline for large EM volumes that use existing, state-of-the-art algorithms.

DESIGN OF USER INTERFACES A segmentation pipeline provides partially correct results of volumetrically reconstructed neurons in an EM volume. These results only become biologically useful to the extent that researchers can inspect and correct them before quantitative analysis. This necessitates the development of appropriate user interfaces for proofreading and error correction. We have developed two user interfaces for this purpose. The first is geared towards researchers who need accurate volumetric segmentations of neurons or other segmented structures. The second interface allows to correct so-called topological errors committed by reconstruction algorithms. This includes false splits or false merges of neuronal processes. The interface was made as simple and easy-to-use as possible with the potential for deployment in a public, crowd-sourced setting with minimal requirements on user training.

5.2 RESULTS

Multi-label pixel classification

A high-quality, sufficiently large ground truth dataset is essential for training computer vision algorithms and validating their performance. The current generation of computer vision algorithms rely on pixel labeled datasets as training data. The manually generated skeletonized reconstructions of neurons can at the moment not be used for training, therefore we needed to generated pixel-based image ground truth directly. We selected a substack of size 1024 x 1024 pixels with 20 sections from the large, high-quality third-instar volume L3 (see Appendix A). We then labeled each pixel in this stack according to the ultrastructural element it represents. In particular, we densely labeled membranes, synapses, mitochondria, cell interior and glia cells (Figure 5.1). We made the dataset available to the computer vision research community by publishing it on FigShare (Gerhard et al. 2013).
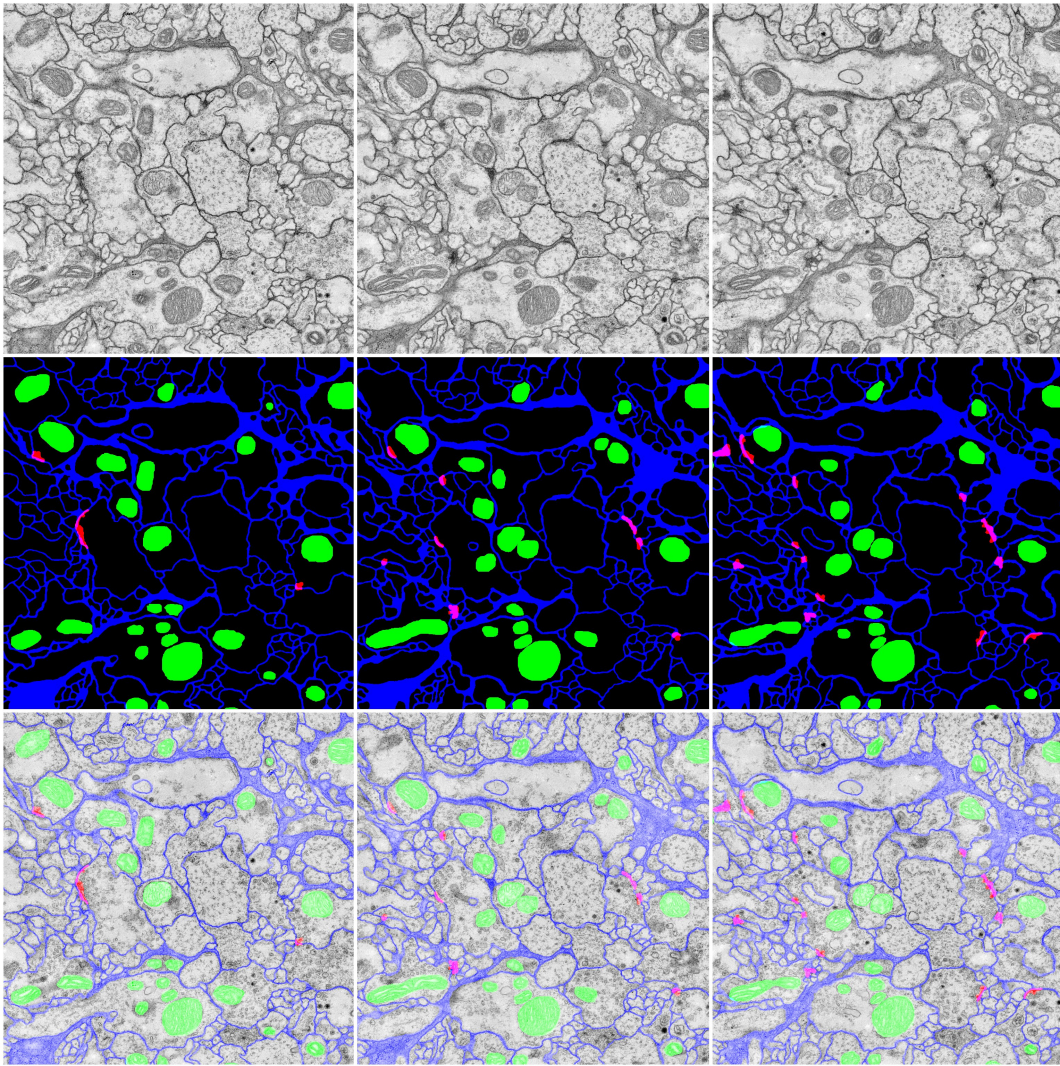
Figure 5.1.: **Ground truth dataset for training**. Three consecutive sections are shown in each column. Raw image data (top row), ground truth labeling (middle row), and ground truth overlayed on top of the raw images (bottom row). The displayed labels are cell interior (black), membranes (blue), synapses (red) and mitochondria (green).

We then collaborated with the winners of the ISBI 2012 Challenge for EM classification [1] to train deep convolutional neural networks. Their winning networks (Ciresan et al. 2012) were extended to work with our multi-label ground truth datasets. Figure 5.2 shows classification results on a test stack for multiple labels. Recent evaluations of the accuracy of neuron reconstructions using these probability maps with novel error measures indicated on the order of 20 false-merge and 10 false-split errors on a 1024 x 1024 pixel image (Funke et al. 2015).
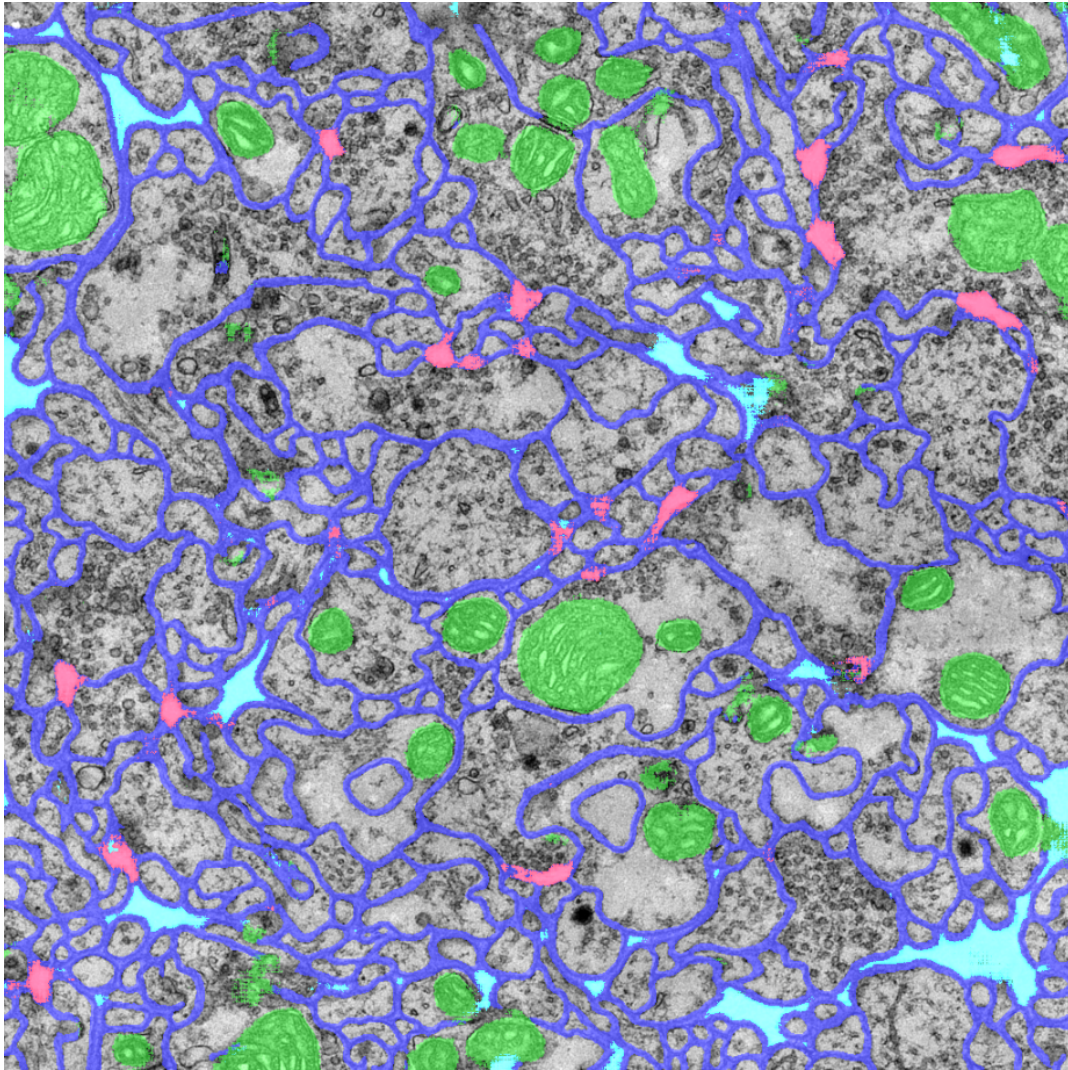
Figure 5.2.: **Teststack2 - Classification results of Convolutional Neural Network**. The results of the CNN from the Schmidhuber group. Multi-label classification including membranes (blue), mitochondria (green), synapses (red) and glia cells (bright blue). For the full series of 20 sections, see Appendix E.

In order to inspect the quality of classification in a larger volumes, we applied their networks to the entire neuropil region of the whole L3 dataset of *Drosophila* larva

---

1 http://brainiac2.mit.edu/isbi_challenge/

ventral nerve cord (Appendix A). We confined image classification to the subregion of left and right neuropils which hosts all the neuronal arborizations and synapses, thereby leaving out the surrounding cortex of cell bodies. This amounted to a volume of 370k Gigapixels (16k x 11k x 2100 sections, 73.6x50.6x105 μm). Exporting the resulting probability maps to CATMAID image stacks allowed us overlay them on the raw image data and visually inspect the classification accuracy in different regions of neuropil (Figure 5.3). We found that the results were qualitatively similar in different regions of the neuropil, for instance in sensory and motor regions.
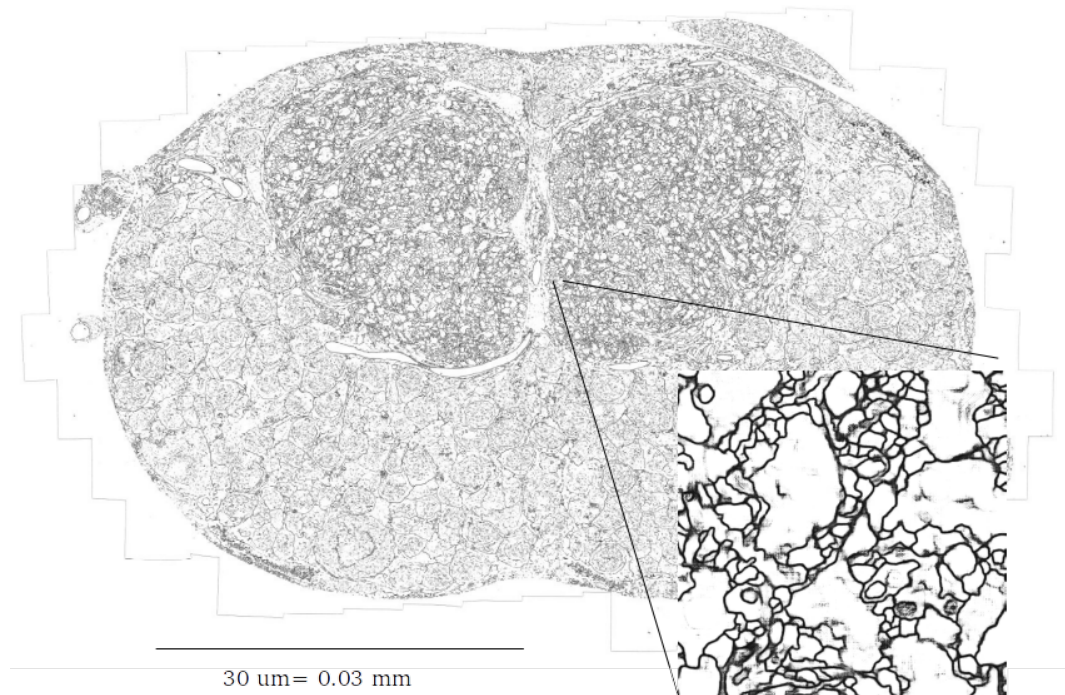

30 um= 0.03 mm

Figure 5.3.: **Membrane probability map of a large EM section of *Drosophila* ventral nerve cord** One EM section of L3 dataset classified with the convolutional neural network by Ciresan et al. 2012 showing the probability of a membrane (probability 1.0, black; probability 0.0, white). Inset shows a zoomed-in version of the membrane map.

We then qualitatively tested the degree with which the trained classifier learned filters that could generalize to other datasets with slightly different tissue properties. Richard Fetter at Janelia Research Campus provided us with a short series of an adulty fly brain preparation with similar tissue contrast as the L3 dataset. We applied the trained networks from the larva dataset to these adult EM images (Figure 5.4). Visual inspection revealed that particular neuropil regions, e.g. large numbers of parallel running fibers with good membrane contrast, were classified well. This indicates that the trained convolutional networks learned filters that can generalize to other datasets. However, training the networks using ground truth data for specific neuropil regions should provide improved classification results.

The generalizability hints at the fact that deep convolutional neural networks learn to discriminate features which are invariant between neuropils. For instance, the diameter of the lipid bilayer of membranes is a conserved quantity across cell types measuring about 4 nanometer. The same membrane composition and thickness in neurons in different types of tissue can account for some of the generalizability and learned invariant features of trained convolutional networks. This tests suggests that deep convolutional neural networks are powerful learning machines to classify different types of EM datasets.
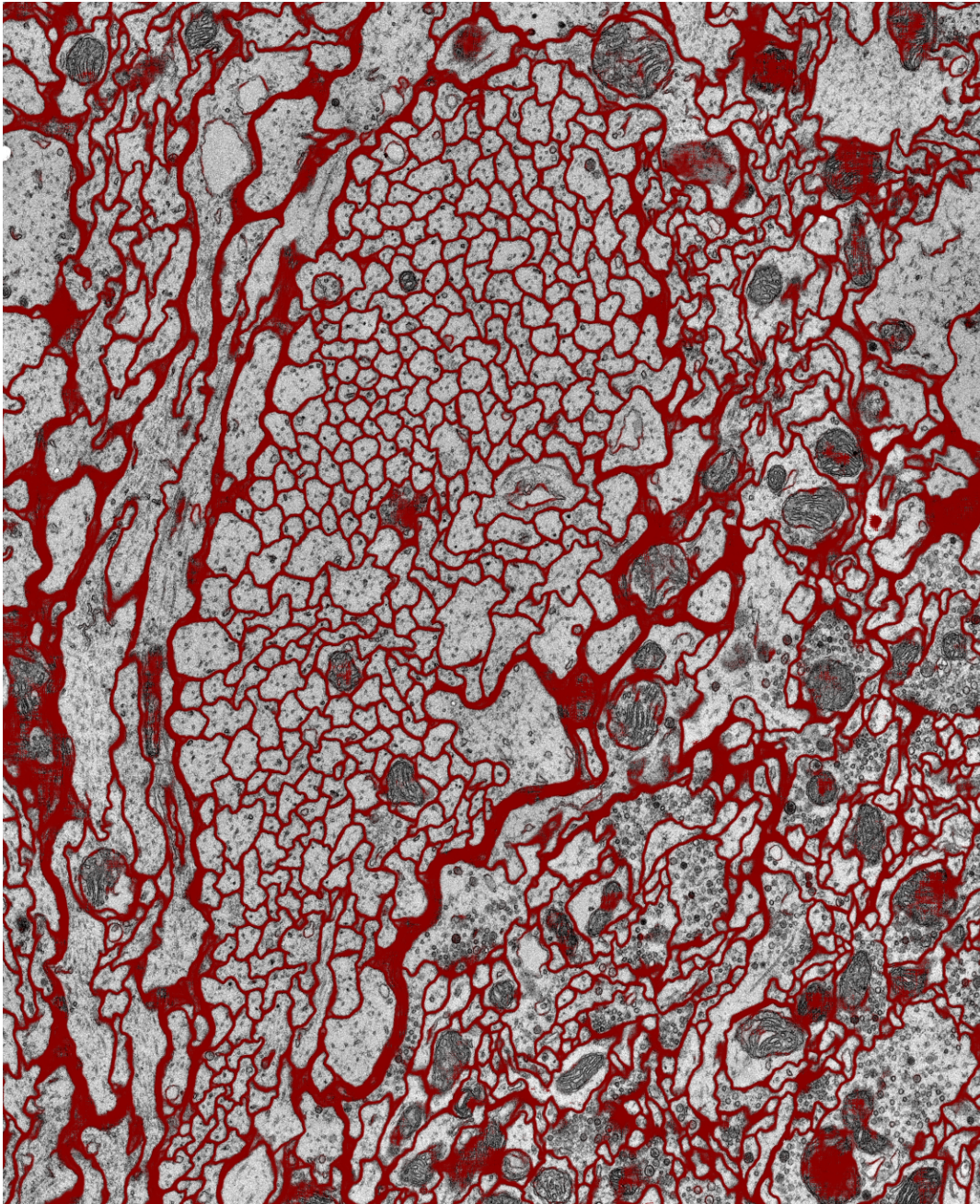
Figure 5.4.: **Membrane classification of adult fly brain tissue**. Membrane classification performed with convolutional neural networks (Ciresan et al. 2012) that were trained on neural tissue of another *Drosophila* larva dataset. The results appear surprisingly good qualitatively . Image courtesy of Richard Fetter, HHMI Janelia Research Campus FlyEM Project.

Neuron segmentation pipeline

Our next goal was to use the generated probability maps for the whole L3 dataset and apply the neuron reconstruction pipeline SOPNET (Funke et al. 2012). The original

implementation of SOPNET [2] relies on the assumption that a complete image stack and results from computations fit into a computer's main memory. The size of the whole L3 dataset precluded running the original SOPNET implementation, so we aimed at modifying the SOPNET neuron segmentation pipeline to process large image volumes on standalone desktop computers. We achieved this by modifying SOPNET to process pairs of large 2D sections that can fit in normal desktop's main memory. The segmentation results were then stored persistently to hard disk and into a database [3].

SOPNET extracts candidate neuron slices (the cross-sectional profile of a neuronal process) from the probability maps, and produces possible pairings between slices in adjacent sections called segments (see the inset of Figure 5.5 for an example neuron slice). These segments have associated feature vectors that are later used to find a good reconstruction in a given volume. So the minimal requirement of our modified SOPNET version demanded to process at least two sections at any given time. The slices and segments data were then inserted into an extended version of the CATMAID data model for persistent storage, and the slice masks for neurons were stored as PNG image on disk. The run of the whole L3 dataset produced approximately 227 million slice candidates and 363 million segments (continuation or branch segments), which resulted in a database with 355 GB that also included segment features and indices for fast data lookup.
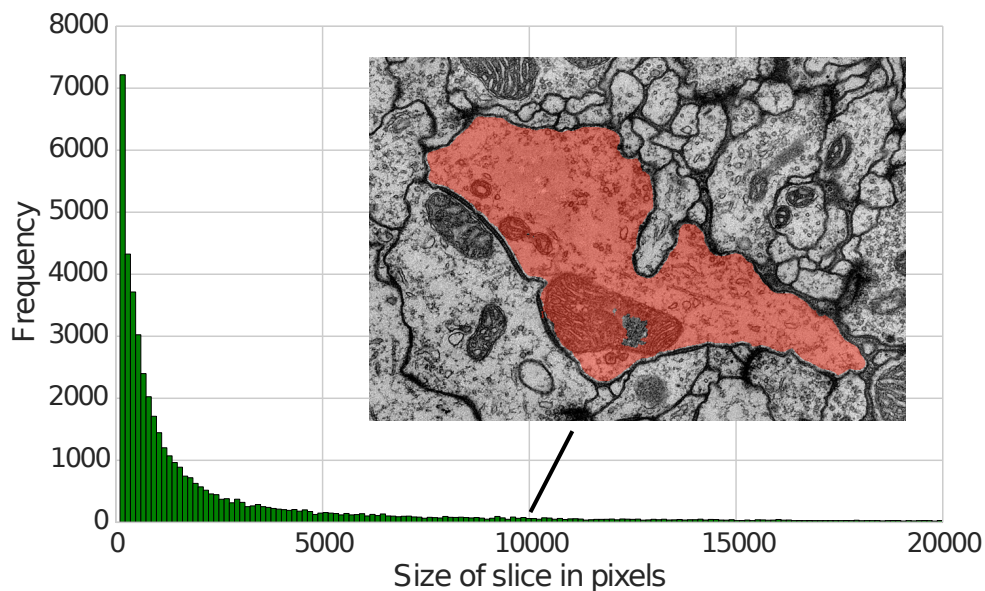

Figure 5.5.: **Histogram of neuron slice sizes in a single section for a 16'000 x 11'000 pixel neuropil region of interest**. Inset shows an example slice of approximately 100k pixels.

---

2 https://github.com/funkey/sopnet

3 Source code of the prototype in the *neurocity* branch on http://github.com/catmaid/CATMAID/

The distribution of slice sizes follows an exponential decay (Figure 5.5). The same distribution was found for several tested sections. The distribution has a long tail, and we found slices up to a total of 500'000 pixels. More than 86% of the slices are less than 2000 pixels. A large fraction of the larger slices are spuriously connected to smaller slices due to the inaccuracy of the membrane probability maps given the selected extraction parameters (Figure 5.6).
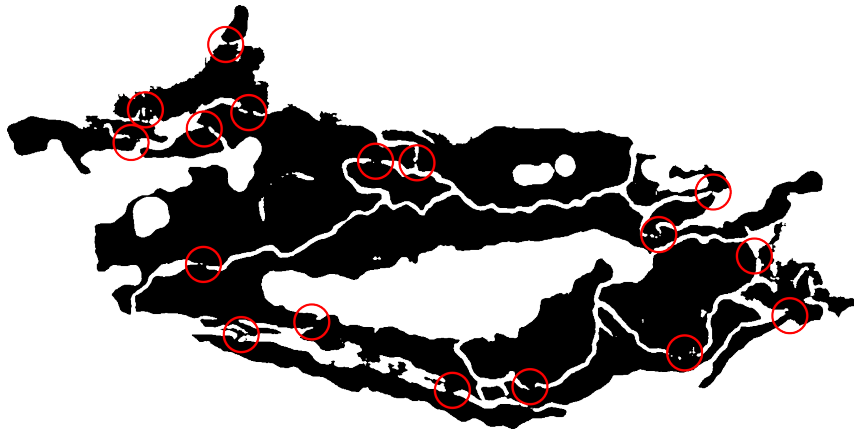

Figure 5.6.: **Spurious large neuron slice resulting from wrongly connected small slices**. The red circle indicate locations where small slices have been wrongly merged to produce a large slice candidate.

Due to artefacts in EM images such as folds or tears, larger section to section variation exists in terms of number of slices and segments extracted. On average, we extracted 107'710 slices per section of size 16'000 x 11'000 pixel (Figure 5.7). If we consider only continuation and branch segments, we extracted on average 172'272 segments per section (Figure 5.8). For each neuron slices, additional end segments on either side are stored in the database. This yields an average of 3.6 segments per individual slice. Note that due to overlap of slices from the extraction procedure, these numbers overestimate the effective number of slices in a section.
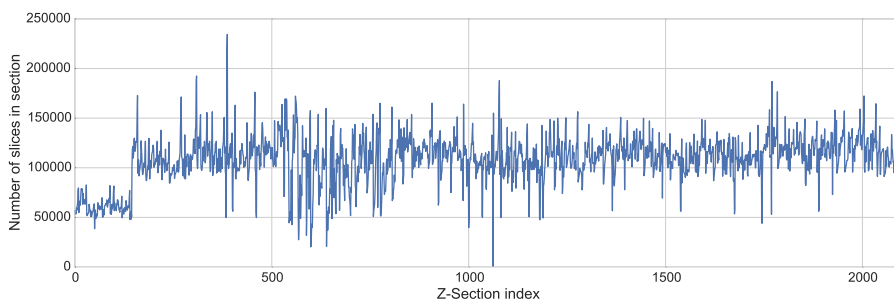

Figure 5.7.: **Number of neuron slices of a single z section of size 16'000 x 11'000 pixel**. Membrane quality, EM artefacts and differences in the actual number of slices in a section lead to variation in the number of extracted slices. The reason for the lower number of slices in the first 144 sections is unknown.
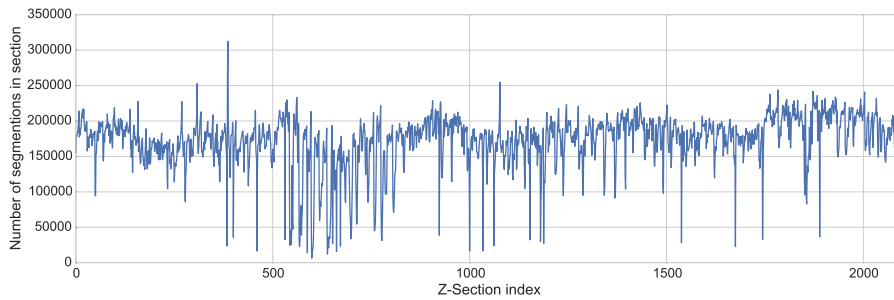
Figure 5.8.: **Number of neuron segments for a pair of z sections of size 16'000 x 11'000 pixel**. An average of approximately 180'000 segments is found throughout the stack. The variation in the number of segments can be due to image noise or artefacts in the samples.

The database containing neuron slices and segments allowed us then to start building user interfaces for volumetric neuron reconstruction. The development of a user interface based on a small test stack would have been possible. However, building an interface with a sufficiently large data volume has important advantages. Also, working directly with a larger dataset can reveal problems earlier in the development process. For instance, some neuron slices are very large on the order of many thousand pixel in width, demanding special treatment to display quickly. The challenge that might arise from following neuronal processes over long distances could be overlooked when only focusing on small test datasets where neuronal processes are very short.

User Interface 1: Volumetric neuron reconstruction in CATMAID

We designed and implemented a user interface for reconstructing single neurons volumetrically. Volumetric segmentation of processes in existing software requires painting each individual pixels or contours that then are filled (e.g. VAST see Kasthuri et al. 2015). A much faster option is possible with the existing database of generated neuron slices. The user just needs to select the correct slices from a small set of possible candidates. In our prototype interface (Figure 5.9), the user can click on an arbitrary location in the image stack to start the reconstruction of a neuron. The neuron slice with the minimal size at this location is displayed, and neuron slice candidates can chosen using the mouse scroll wheel. If the neuron slice fits the expectation, the user can press a key and move to the next section. In the next section, the slice is already displayed that has a minimal cost connected through a segment with the previous slice. This cost was defined in the SOPNET process that extracted the segment features. If the slice is suitable, the user can press a key again to move to the next section and neuron slice, and so forth. Whenever a slice does not match the underlying EM data, another slices from the possible set can be selected. This allows to quickly move through the stack along continuation and branch segments to generate a volumetric segmentation of a neuron. However, in this prototype interface

the reconstruction of a single neuron takes longer than manual skeletonization. It is therefore better suited for targeted volumetric reconstructions of single neurons than for the extraction of the topology of complete circuits.
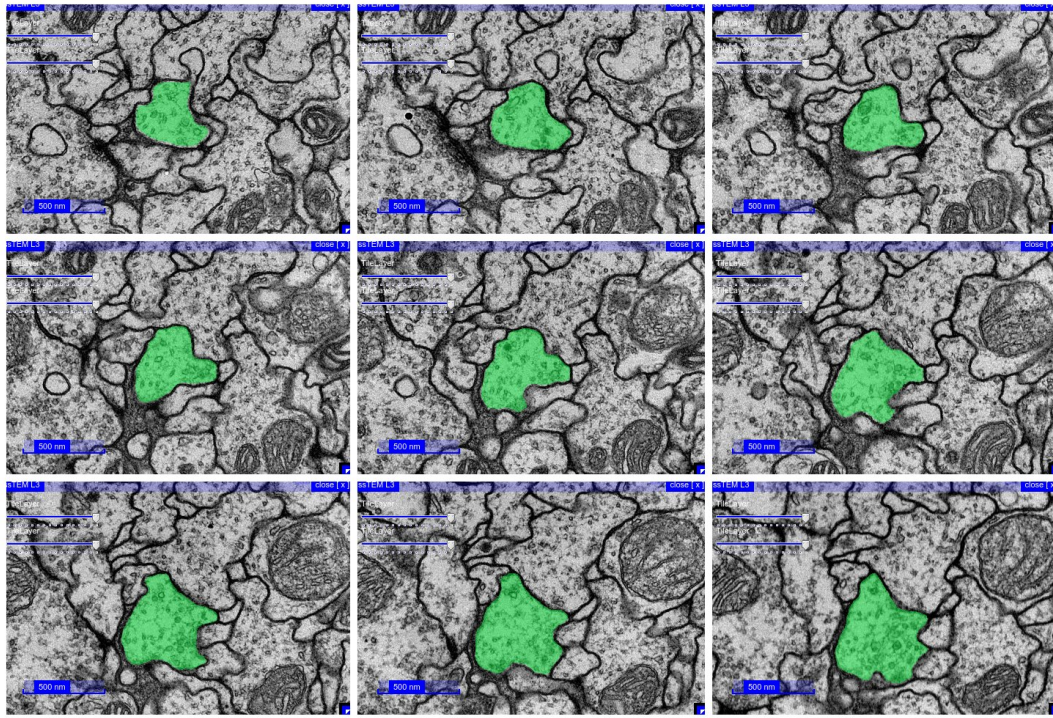

Figure 5.9.: **User interface prototype of interactive volumetric neuron reconstruction interface in CATMAID** From top-left to bottom-right, individual images show consecutive candidate slices (green) of a specific neuron of interests. If individual candidates do not match the expected neuron outline, a different candidate can be selected from a set of extracted candidates manually using the mouse-wheel.

User Interface 2: Crowd-sourceable interface to correct topological errors

We implemented a second user interface to quickly fix topological errors from the reconstruction pipeline (Figure 5.10). The upper row shows the current section, and the lower row the possible continuations in the next section (which could include ends, branches or continuations). The users then clicks on one of the possible images in the lower row to mark it as the correct continuation of the neuron process. It is possible to toggle the slice overlays to better see the raw image data for context, and to discard the decision if there is uncertainty associated with it. A larger number of such decisions can be made in a short time, and the logic behind can be quickly grasped by novice annotators. Importantly, the bandwidth requirements is minimal, and only a small amount of image data has to be transferred to clients, namely a small number of compressed images for the raw image data and the slice overlays.
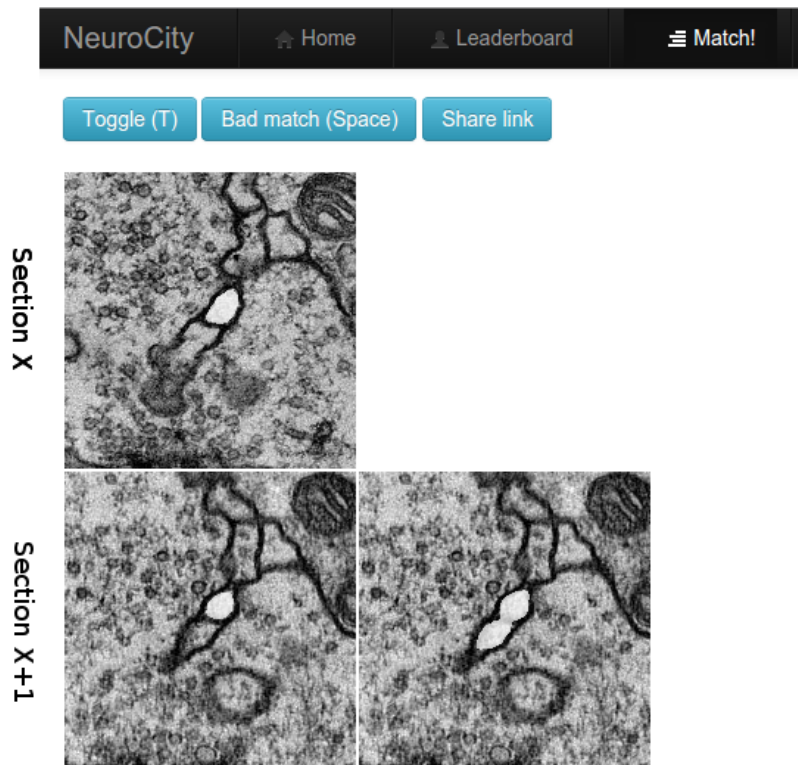
Figure 5.10.: **User interface prototype to link neuron slices between sections** The selection of the correct continuation segment from a small number of alternatives is posisble in the interface. The top row shows a selected neuron slices from one section, and the bottom row displays two alternatives for the continuation of this slice in the consecutive section. Selecting the left slice/image is the correct choice in this case.

Different options exist to provide motivation and incentives to users that may want to contribute topological error corrections or decisions. Beyond the idealistic prospect of contributing to the quest to map the brain, more competitive aspects could be introduced in the workflow. We implemented a simple leader board that shows the number of decisions made in a give time frame (Figure 5.11). Revealing particular types of information about users such as the amount of annotations performed might further incentivize competitive motivation. Such a simple interface might not only work in a public, game-like setting, but could also be employed in a setting where users are payed for making decisions. Users of micro jobbing platforms (such as Amazon Mechanical Turk) may provide a large workforce to contribute to reconstruction via such an interface. Users of such platforms could be monetarily compensated for individual decision, for instance as a function of their expertise and training. Alternatively, dedicated neuron reconstruction services such as Ariadne Service [4] could provide demand-driven expert proofreading man-power for large reconstruction projects.

---

4 e.g. www.ariadne-service.ch

Figure 5.11.: **Leaderboard** Overview of contributors with the number of contributed annotation decisions (votes).

The order of presentation of such segment decisions could be prioritized by different criteria. For instance, high priority could be given to decisions that are part of a neuron of interest in a large volume, thereby effectively guiding proofreading to biologically relevant locations. In another scheme, segments could be selected that are in an ambiguous region of the segment cost function (Klein 2014). Segments with a very high or very low cost are presumably more likely to be classified as correct or wrong segment than segments in the intermediate region. The generation of manually classified segments can serve as ground truth data to implement an online learning scheme at large scales. Such targeted user feedback would permit ongoing improvements to the classifier that selects the best neuron slice and segments.

5.3 DISCUSSION

In the future, a circuit mapping approach along the lines presented could leverage computational segmentation pipelines and most effectively uses human proofreading time in a closed feedback loop. The sparse, targeted circuit mapping approach introduced in Chapter 2 already reduced reconstruction time over redundant approaches. Optimally, such a sparse, targeted approach could be combined with a densely segmented volume containing a limited number of errors, and thus could further speedup reconstruction time. Such a targeted, semi-automated reconstruction approach for proofreading would guide human attention to error locations that are most relevant for a particular circuit of interest.

Attempts have been made in this direction at large scales, but to our knowledge, none explicitly implemented targeted proofreading to circuits of interest (e.g. Plaza et al. 2014; Kaynig et al. 2015). The advantages of keeping the human in a tight feedback loop have been demonstrated by image classification and segmentation tools such as Ilastik (Sommer et al. 2011). By inspecting the classification results and iteratively adding new training data to improve the pixel classifier, the resulting classification can quickly improved to useful segmentations. The importance of immediate feedback with a short delay is crucial to guide new annotations and for evaluating the results.

For large-scale EM image stacks, a similar iterative approach to circuit mapping could be envisioned. Individual proofreading decisions can be used in a feedback loop to immediately and iteratively improve the accuracy of the underlying segmentation algorithms. If used by annotators in parallel, this online learning scheme implemented by such a loop would become beneficial to every single contributor in their self-contained circuit mapping projects. The more people use the system, the more training data becomes available to improve the underlying learning system. With retraining, it can thus quickly converge to the maximal accuracy. With further optimization of the system, human proofreading time can be allocated to resolve only the most difficult errors.

With the system presented here, an assumption is made that the probability maps are sufficiently accurate to extract at least one neuron slices that is in good correspondence with the ground truth. This assumption is not generally true, but with this assumption a major advantage arises. Unlike other systems that require pixel-based splitting of neuron slices or merging them together (e.g. Haehn et al. 2014), the complexity of proofreading decisions could be reduce to only select an appropriate neuron slice. It is thus not required to change the pixel-based geometry of individual slices, an operation which takes longer than a simple selection.

Furthermore, the feedback provided to users after a decision could include the impact of the decision on neuronal morphology or the circuit diagram. For instance, after selecting an alternative neuron slices or a different continuation, the resulting neuronal morphology after this decision could be displayed to the user. This is one example of a number of high-level cues, such as described in the previous Chapters of this thesis, that can be employed to evaluate proofreading decisions. The metrics that implement some of these higher-level features could be displayed alongside the different proofreading decisions to provide additional information. Eventually, such high-level features could be integrated into reconstruction algorithms themselves as hard or soft constraints.

For further development of automated reconstruction algorithms at the circuit level, good training and benchmark datasets are required. The neurons and synapses that we mapped and annotated in Chapter 3 and Chapter 4 could serve as a benchmark dataset. In addition to the pixel-based ground truth provided here for the third-instar EM volume, these skeletons and synapse annotations can be used to train algorithms that aim at reconstructing whole neurons and circuits within a large image volume. For instance, the left set of neurons can be used as training dataset, whereas the right set of neurons can be used to validate and evaluate the performance of the algorithm. Additionally, the mapped neurons can serve as test data to evaluate how much of the topology of a circuit can be reliably recovered with automated methods.

The main focus of automated EM reconstruction algorithm development is on extracting the topology of neural circuits, but a wealth of additional data is present in EM images. Volumetric information about synapses, mitochondria or glia cells add valuable quantitative information to existing wiring diagrams. For instance, the size and distribution of mitochondria is important for presynaptic and postsynaptic synapse function (Li et al. 2004; Stowers et al. 2002). The amount of mitochondria

volume has been implicated in discriminating excitatory and inhibitory neuron due to their differing energetic demands (Kasthuri et al. 2015). Also, the size of presynaptic sites and synaptic contact areas might be a better proxy for synaptic strength than an integer count. Furthermore, volumetric information of neuronal processes is critical information when building biophysically-accurate electrotonic models of single neurons. Glia cells have been difficult to characterize volumetrically at the ultrastructural level due to their intricate, fine-scale structure. Automated morphological measurements could provide interesting new avenues for glial study such as their relation to synaptogenesis (Molofsky et al. 2012). The association of glial cell with synapses has been shown to significantly interact with neural circuits and therefore would be a valuable target to volumetrically reconstruct (Araque et al. 2010; Yoon et al. 2014).

5.4 OUTLOOK

Using some of the experiences gained in this project, we have embarked on building the scalable, distributed segmentation pipeline CATSOP [5]. This new pipeline aims at providing an effective way to scale SOPNET to large image volumes by a block-wise processing approach, and seamlessly integrate into the existing CATMAID framework for semi-automated circuit reconstruction (hence the name CATSOP). The long-term goal is to iteratively replace the manual process of circuit mapping. With the modular architecture of CATSOP, individual modules of the pipeline can be replaced as they improve (e.g. image classification, segment cost computation etc.). An important requirement of this approach is to ensure that resulting segmentations remain compatible with CATMAID so all the existing tools for circuit review, analysis, organization and visualization can be used.

---

[5] github.com/catsop

DISCUSSION

The goals to accomplish cost-effective, high-throughput EM connectome mapping have been considerably challenged and criticized (Morgan et al. 2013). These challenges and criticisms are reminiscent of those that high-throughput full genome sequencing faced at early technological phases. Full genome sequencing was initially perceived as too costly and time-consuming to become practically useful. Incentivized by its utility for basic and clinical research, and its promises for medical applications, new sequencing technologies developed and evolved rapidly. The initial sequencing of the whole human genome got underway in 1990 and took more than 13 years to complete with a cost of more than 3 billion US dollars. In the year 2015, technological advances reduced the cost to sequence a whole human genome to less than 5000 dollars, and the genome map can be completed in a few days (Check-Hayden 2014).

Similarly, the currently ongoing large-scale connectomics projects are very costly, involve a large number of people and take many years to completion (e.g. Takemura et al. 2013). With ongoing technological advances and automation, however, generating large and many EM volumes is becoming more cost-effective and can be performed in an increasingly high-throughput fashion. Improvements involve, for instance, better staining protocols (Mikula et al. 2015; Hua et al. 2015), the development of automated and robust methods for sectioning EM tissue (e.g. Denk et al. 2004, iTome developed in the Hess lab at Janelia Research Campus), parallelization of acquisition by parallel electron beams and camera arrays to increase field-of-view (e.g. Eberle et al. 2015 or Bock et al. 2011 respectively), faster CCD cameras, faster sample stages and a host of algorithms to control and optimize the throughput of these instrumentations.

Additionally, image processing pipelines are required to further process raw EM images into aligned 3D image volumes. To handle the large volumes of data, alignment algorithms need to be parallelized and operate on large compute cluster (e.g. Burns et al. 2013; Scheffer et al. 2013). With these active and ongoing developments, I anticipate a growth trajectory for EM connectomics in the coming decades with similar acceleration potential like genome sequencing technology. A key success factor will be in convincingly demonstrating the scientific, clinical or neurotechnological merits that justify such large-scale efforts to public and private funding bodies.

Recently, a number of objections have been raised that challenged the scientific merit of large-scale connectomics projects (reviewed in Lichtman et al. 2008). From the pioneering project that mapped the connectome of *C. elegans*, we have some insights about the possible limitations and opportunities of connectomic datasets. Clearly, structural maps of neural circuits will not immediately reveal the functioning of neural circuits, as genomes have not revealed completely the working of the interior of cells. They are, however, deemed necessary - although not sufficient for

understanding brain function (Bargmann et al. 2013). As we continue to associate structural with functional circuit data in quantitative behavioral assays, we will be increasingly better at defining the usefulness and limits of connectome datasets.

Integrating connectome data in small model organisms

As a case in point, the availability of the *C. elegans* connectome has certainly contributed significantly to the development of an entire research field (Emmons 2015). It allowed to generate new hypotheses about neuron and circuit function and guide experiments to test, validate or falsify conflicting models. There are all reasons to expect that connectomes of other species and circuits can similarly contribute to generating new hypotheses and guide experiments.

The focus on circuits in a genetically-tractable model species with a small nervous systems, such as *Drosophila*, has many advantages (Simpson 2009). In a recent large-scale study which was facilitated by the contribution of this thesis project exemplifies these advantages (Ohyama et al. 2015). This work integrated large scale screening of behaviors, functional imaging with the manipulation of neural activity, and large-scale electron microscopy and circuit mapping (Dierick et al. 2015). All these techniques needed to be integrated to reveal a more comprehensive picture of the functioning of a subset of neurons in a well-defined behavioral phenotype in *Drosophila*.

The integration of anatomy, physiology and behavior by means of neuroinformatics platforms and pipelines will significantly inform brain theories, computational neuronal models and circuit simulations, and has the potential to drive new neurotechnology developments (Figure 6.1).

Small organisms and circuits are further a perfect testbed to integrate anatomical, physiological and behavioral data. The data from different modalities can often be obtained at the same resolution, thereby minimizing the resolution mismatch between data generation technologies (Meinertzhagen et al. 2012). A mismatch occurs when the resolution of data obtained by different imaging modalities can not be matched. For instance, volumetric EM provides data about the detailed distribution of single synapses on a neuronal arbors and their neuron-neuron connectivity. If physiological data of the activity of neurons is limited in spatial resolutions to clusters of cells, it is not straightforward to match these clusters to single neuron connectivity derived from EM. Similarly, the spatial and temporal resolution of quantitative, behavioral measurements of an organism's activity need to match the resolution of physiological recordings to relate these different data types conclusively, for instance via muscular activation dynamics. In small circuits in small animals, such as *Drosophila*, *C. elegans* or larval zebrafish, the trend is to continuously reduce this mismatch by means of newly emerging large-scale imaging technology (Ahrens et al. 2015). Quantitative, high-throughput pipelines for imaging the temporal dimension (e.g. Amat et al. 2014) or behavioral analysis (e.g. Ohyama et al. 2013) have to be established and integrated with ultrastructural data of synaptic connectivity for a comprehensive understanding of the operation of neural circuits.
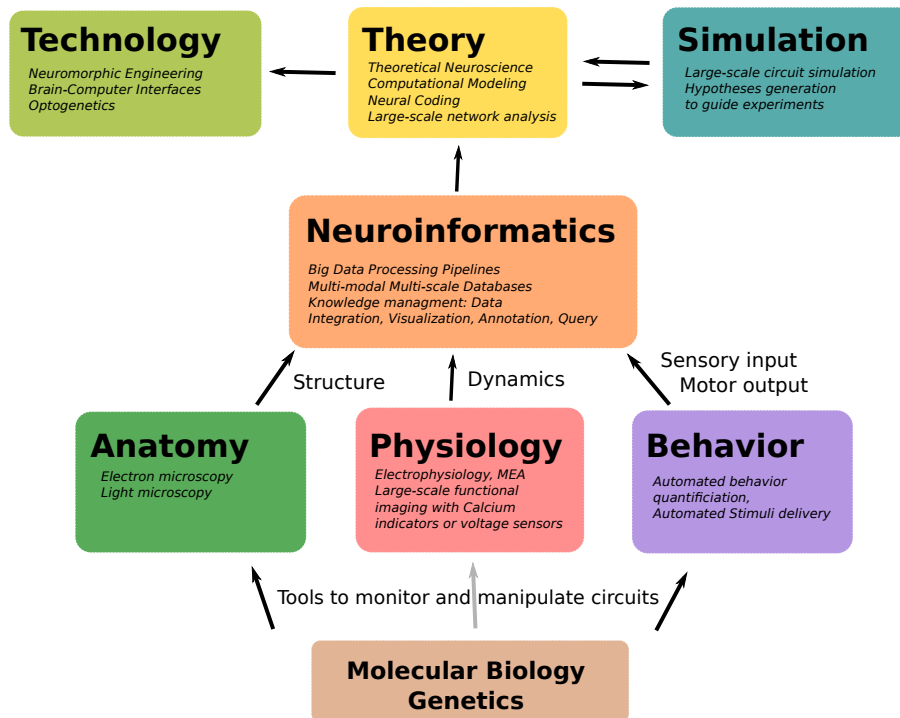
Figure 6.1.: **Neuroscience subdisciplines and their relationships** A comprehensive understanding of neural circuit operation requires knowledge of its anatomy (structure), its physiology (dynamics) and its function (sensory inputs, motor outputs). The genetic toolkit from molecular biology allows to manipulate neuronal cell types to monitor and interfere with activity. In large-scale projects, neuroinformatics tools provide means to process data streams in automated computational pipelines, and analyze, visualize and manage their results. The role of theory is to parsimoniously explain data by building mathematical models, thereby transforming data to knowledge. These models can then be used to implement neural simulations which is useful for testing hypotheses *in silico*. Importantly, simulations might provide a framework to systematically determine what the right questions to ask are, and thereby guiding novel experimental work effectively to validate, extend or falsify models. The theoretical understanding of neural circuit function will facilitate the recapitulation of principles in neuromorphic hardware and allow for building better brain-computer interfaces.

Scaling circuit mapping to large brains and statistical connectomes

The framework and method presented in Chapter 2 represents a particular instance of a neuroinformatics tool for neuroanatomy that naturally bridges small lab efforts (e.g. those interested in mapping small circuits of interest) and large-scale projects (e.g. mapping the nervous system of an entire organism). Vice versa, mapped circuits can be made available to uninvolved labs working in physiology, behavior or in theoretical and computational fields. To achieve this, large-scale projects and

consortia are likely to become increasingly common in neuroscience in the future complementing research in independent, smaller labs.

Scaling up EM circuit mapping approaches to larger circuits and brains, for instance to an entire mammalian brain, will need to overcome even bigger challenges (Lichtman et al. 2014; Mikula et al. 2015). At the moment, it is not yet completely clear to what degree the cost of producing mammalian-brain level EM circuit maps can be justified by its benefits (Marblestone et al. 2013). Therefore, creative alternative approaches are being envisioned, such as molecularly annotated neurons and their connections that are subsequently sequenced (Marblestone et al. 2014). Other approaches using light microscopy may provide circuit mapping strategies from micro-scale to meso-scale level with relatively high throughput (Rah et al. 2015). For such light-based approaches to become useful at large-scales, they need to be developed and deployed in an industrial-scale settings, optimally focusing on a small set of brain regions with the aim of comprehensive mapping.

The Allen Brain Institute pursues such a strategy of comprehensively mapping all the cell types in the visual cortex of the mouse using fully automated pipelines. The resulting datasets are made publicly available[1]. This cell type database contains neuron reconstructions derived from light microscopy images, and includes information on arbor morphology and localization to cortical layers together with associated physiological data. Such neuronal reconstructions can be employed as the basis to predict connectivity information between neurons to build so-called statistical connectomes (Li et al. 2007; Reimann et al. 2015).

Such statistical connectomes have been proposed to be sufficient to build accurate computational models of neural function (Hill et al. 2012; Egger et al. 2014). At the moment, these approaches are the only viable strategy to obtain more realistic representations of neural circuit organization at both local and long-range scales. However, it will be crucial to eventually validate such statistical maps with information obtained from ultrastructural studies using EM (Costa et al. 2013; Morgan et al. 2013; Kasthuri et al. 2015). This is also a long-term strategic goal of the Allen Brain Institute (Koch 2014).

Centralized Infrastructure and Collaborative Connectomics

Depending on the size of the brain under scrutiny, investment cost into volumetric EM data generation infrastructure can be tremendous (Marblestone et al. 2013). Similar to the astronomy or particle physics communities, with its expensive telescopes or particle accelerators, a pooling and centralization of data generation, processing and storage infrastructure could be envisioned for EM connectomics (Burns et al. 2014). Many research labs would tremendously benefit from the availability of such a public connectomics infrastructure. This would allow small labs access to connectomic datasets without requirements to setup a costly infrastructure, workflow and expertise on their own.

---

1 http://celltypes.brain-map.org/

Proposal for such *Collaborative Connectomics* efforts and infrastructures have already been made (Burns et al. 2013; Hayworth et al. 2014). A centralized system with appropriate regulations on data access and incentive structures would provide entire research communities with access to circuit data in their favored model organism or system. In the long-term, such research community resources could significantly contribute in accelerating progress in understanding brain function at the cellular and circuit level.

Building large-scale EM connectomics infrastructure is already underway in several research institutions, such as the Harvard University, the Allen Brain Institute, the Max-Planck Institutes or HHMI Janelia Research Campus. Each institution has so far focused their efforts on different EM acquisition methods in particular model systems and brain regions. As the acquired EM volumes in these efforts are incredibly large, they are likely going to attract researchers that can collaboratively mine their richness for years. Furthermore, once pipelines are setup and work reliably, they can be applied to many different brains.

In the spirit of such a *Collaborative Connectomics* approach, I hope the contributions made in this thesis increased the awareness of the potential impact of mapping and analyzing neural circuit structure collaboratively by researchers world-wide.

6.1 FRAMEWORK FOR CIRCUIT MAPPING, ANALYSIS AND VISUALIZATION

The goals of the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative launched in 2013 in the United States are to develop novel tools for the comprehensive mapping of the nervous system at multiple temporal and spatial scales (Jorgenson et al. 2015). In addition to generating new tools, the importance of creating organizational frameworks to store, distribute and share data in a meaningful and practical way have been recognized and emphasized (Yuste 2015). In line with this, our framework provides a timely contribution to not only map, but also to organize, analyze, and visualize the structure of neural circuits.

Here, we briefly discuss three high-level aspects of this contribution to the emerging field of large-scale EM connectomics, complementing the discussion in Chapter 2: First, how image and circuit data can be shared with other researchers more easily by providing an easy-to-use web interface and a programmable interfaces for data access. Second, how our tool suite can be extended and promotes the integration of novel connectomic data analysis and visualization techniques. Third, how the framework can foster new ways of scientific collaboration and integration, especially between different research communities.

(1) Data sharing and data access

Improving practices for data sharing has advantages for both the scientific community and individual scientists (Foster et al. 2007). For instance, the reproducibility of results can improve by releasing code and data together with the associated publications. Recent years show a trend towards requirements of data release with publication by journals and funding agencies (Nat Neurosci Editorial 2007; Kaye et al. 2009). The genomics community is at the forefront in driving the development of best practices and policies for data sharing. Many of the lessons learned in genomics are very likely transferable to the emerging field of EM connectomics.

The funding bodies in the genomics community generally require sequenced genomes to be made available in centralized databases. These datasets can then be retrieved through standardized web API calls from these databases or browsed through online web interfaces. Our framework already supports a variety of these standardized data access patterns for connectome datasets together with its web-based interface for image and circuit data access (e.g. see Appendix B and Supplementary Text in Chapter 2). The Open Connectome Project[2] is the first effort to make published connectomics datasets available to the public through a variety of interfaces (Burns et al. 2013). The project also uses the CATMAID framework to make terabyte-sized image stacks accessible for browsing and follow-up analysis to the public both for EM data as well as data generated by array tomography (Weiler et al. 2014).

---

2 openconnectomeproject.org

(2) Extensibility, data analysis, annotation and visualization

Analyzing and visualizing connectomic datasets is challenging due to their inherent complexity and size. Experts in statistics, network science and machine learning will need to develop novel analysis methods to extract meaningful patterns from large, mapped neuronal networks (Bullmore et al. 2009). Existing graph-theoretical methods might be adequate to help an individual neuroscientist answer specific questions (Rubinov et al. 2010). It is likely that non-standard types of analyses tailored towards connectomics will require expert knowledge to develop and implement.

Our framework is designed to allow integration of novel analysis methods developed by experts. These methods can be integrated within the framework, for example as a plug-in or widget implemented in the JavaScript programming language (see Appendix B). Subsequently, these contributions can be deployed to large groups of researcher. They can become immediately accessible as part of the framework toolkit once the server hosting the CATMAID instance is updated, and users only need to refresh the tool web page in their web browser. Complex, large-scale network analysis methods, such as network motif finding algorithms or cell-type classification methods (e.g. Jonas et al. 2015), can also execute in the back-end. The framework architecture allows to start long-running computational processes on compute cluster infrastructure. Once the results of these processes are available, they can be presented to users in the web front-end. Therefore, novel connectomic analysis algorithms can be quickly deployed within the framework with benefits for both users and analysis methods developers.

Alternatively, neuron and circuit data can be retrieved to the local computer for analysis. For instance, skeletons representations, synaptic connectivity data or connectivity matrices can be retrieved for selected sets of neurons via a web API. The data can be directly retrieved into interactive programming environments, such as Python[3] or R[4]. In these environments, a large number of libraries are available for scientific computing for follow-up analysis (e.g. Pérez et al. 2011).

For instance, I performed the data analysis for Chapter 3 and Chapter 4 by retrieving nociceptive circuits from early and late stage *Drosophila* from three independent CATMAID projects. The data was then analyzed inside Interactive Python Notebooks[5]. IPython notebooks combine code execution, rich texts and plots (Pérez et al. 2007; Rossant 2013; Rossant 2014), and can provide a basis for building reproducible, collaborative data analysis workflows online (Ragan-Kelley et al. 2013). The IPython notebook infrastructure has been recently extended into the Jupyter project and forms the foundation for a newly emerging paradigm of *interactive computation* and collaborative, computational narratives (Perez 2015). Such 'data analysis notebooks' could provide novel opportunities to share connectomics analysis workflows online (e.g. see associated Python scripts of Jonas et al. 2015) and help to introduce new students to the field.

---

[3] github.com/catmaid/CAT
[4] github.com/jefferis/rcatmaid
[5] ipython.org/notebook.html

Our framework also provides methods to manage and organize neurons and circuits by a flexible annotation system. The importance of standardized nomenclatures and ontologies has been recognized as an important requirement to organize knowledge and improve scientific communication (Rubin et al. 2006). It is pursued already in many model system at both brain region level (e.g. Reiner et al. 2004; Ito et al. 2014) and cellular levels (e.g. Ascoli et al. 2008). Large-scale volumetric EM has the potential to make such neuronal classification systems complete because it can comprehensively map all the neurons within a given neuropil or brain region.

In our framework, the annotation system allows to organize neurons according to user-defined classification schemes. Neuron annotation tags can also express, for instance, information about a neuron's associated genetic line or neurotransmitter profile. With this metadata employed in search queries, connectomes can be explored flexibly via the filtering of subsets of neurons that are then further analyzed (e.g. show all GABAergic interneuron in brain region X and their connections). Eventually, annotations of neurons should be curated to build community-agreed ontologies with unique identifiers that facilitate database interoperability and linking of across data modalities.

Circuit visualization

To advance our understanding of brain networks, the ability to visualize complex datasets will be crucial. For exploratory connectomic data analysis, the visualization of 3D neuronal morphology data with its associated connectivity information is of great importance. Insights from cognitive science help to better understand how to leverage the powerful pattern detection capabilities of the human visual system and build tools that are most appropriate for the tasks (Goldstone et al. 2015). While designing and implementing our suite of visualization tools, we were strongly driven by the needs to look at specific aspects of neuronal morphology and circuits.

Large-scale connectome datasets are very complex and span multiple levels from image pixels to 3D skeletons of neurons to high-level graph representations of networks. Each level has its own challenges for visualization and user interactions. Moreover, for exploratory data analysis it is crucial to be able to transition seamlessly between very different types of data representations. Additionally, due to the large size of these datasets - even when reduced to network graphs - the many elements and relations are difficult to comprehend simultaneously. It is crucial to form piecemeal mental representations of the network complexity. This is facilitated by our tools to selectively slice out subsets of the data, flexibly add and remove irrelevant and relevant parts or collapse sets of elements to simpler elements.

The 3D viewer is a central element for interactively visualizing arbor morphologies and synapses. The 3D representation of neurons as skeletons and the distribution of synaptic input and outputs, together with connectivity information with other neurons, provides important cues to generate hypotheses about their functional roles. A variety of shading and coloring options help to emphasize different aspects of a neuron's morphology (see Supplementary Material in Chapter 2).
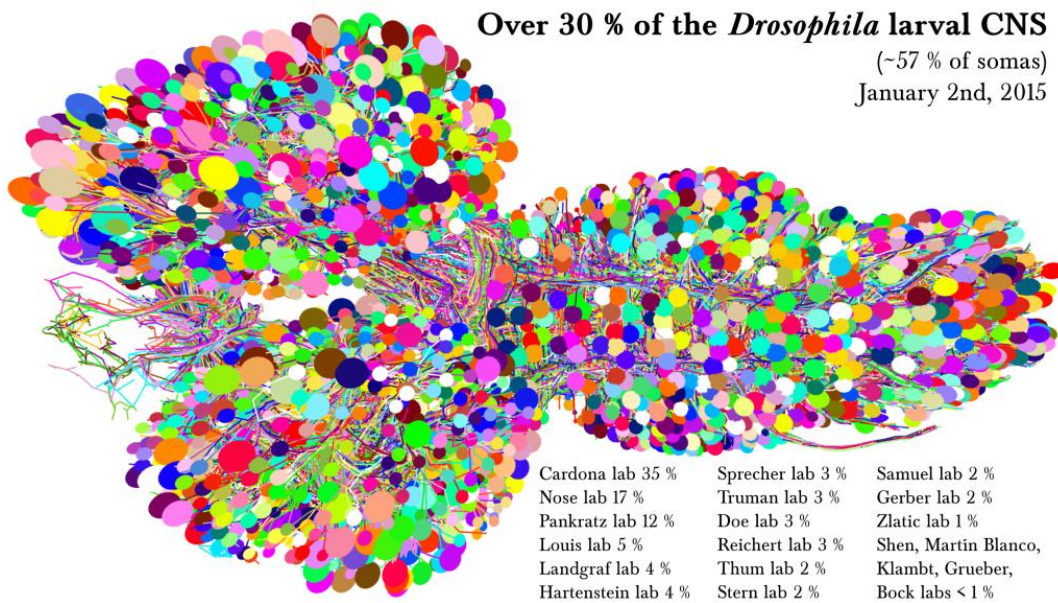
Figure 6.2.: **Wiring diagram of the *Drosophila* larva in CATMAID**. About 30% of the wiring diagram is reconstructed for a serial section Transmission Electron Microscopy dataset of the whole central nervous system of a *Drosophila* larva. Collaborative mapping of neural circuitry is hosted by the Cardona lab at Janelia Research Campus. A number of contributing labs performed circuit mapping locally and remotely using CATMAID in a period of about 2 years. *Image from* `twitter.com/albertcardona`

For a set of neurons shown in 3D, different representations of the same neuron set can be created in other widgets. For instance, a set of neurons visualized in 3D can be used to populate the graph widget that shows the same set of neurons as a 2D network with associated connectivity. Another often used widget is the connectivity widget where all postsynaptic and presynaptic partner neurons can be displayed in tabular form. As these examples show, each widget offers different options to select, group and visualize data. This can create powerful and flexible workflows to interactively explore subsets of large neural circuits in real-time.

Research into visualization of large-scale, multi-scale network graphs continues to be an active and growing field of research. Our architecture is designed in a modular, extensible way to be able to accommodate new developments from this research (e.g. Beyer et al. 2013). New widgets offering novel visualization methods can be implemented independently and added to the framework. With this plugin-like architecture, network visualization researchers can make their results applicable to neuroscience and useful to a large number of users with minimal additional implementation effort.

DISCUSSION

(3) Virtual scientific collaboration

The virtual collaborative space created by our framework promotes novel forms of scientific collaboration. For instance, synergies between research labs are created when neurons only need to be mapped once, but can be analyzed in the context of different research questions. A shared framework thereby reduces the overall mapping efforts and over time, insights of the involvement of particular sets of neurons in different functional or behavioral contexts grows. As subcircuits are mapped in parallel within one EM volume, the resulting circuit maps can be eventually analyzed together, and thereby offer the potential for new insights not obtainable from single subcircuit analysis alone. With this setup, the mapping of distinct subcircuits by individual labs world-wide can converge towards whole connectome maps of an EM volume. In the *Drosophila* larva, such a whole CNS mapping effort based in the Cardona lab at Janelia Research Campus has already generated more than 30% of the complete larva wiring diagram (Figure 6.2).

In such large-scale efforts, it is important to track individual annotation contributions. In our framework, every annotation is time stamped and associated with the users who created or modified it. Therefore, it is possible to recapitulate the contributions made for the initial reconstructions, synapse annotations, reviews, annotation or tagging events. On one hand, this allows for detailed contributor analytics and evaluation of performance and improvements of annotation skill over time. On the other hand, credit can be properly given to individuals and labs who contributed in the mapping efforts.

The tracking of contributions is also important when published circuits have to be revisited. This situation might arise when additionally data needs to be mapped that was originally not annotated because it was not relevant for the addressed questions. It is expected that published circuits will be extended or that issues in an existing circuit map may need additional reviews. In the *C. elegans* connectome, after its first publication (White et al. 1986), investigators have subsequently improved the accuracy, completeness and self-consistency of the maps (Varshney et al. 2009). Functional data that suggests connections between neurons or neurons and muscles that were not found in original EM connectomes (e.g. Albertson et al. 1976), might subsequently be distinguished because of less stringent criteria in annotating chemical synapses, or neuro-muscular contacts (Bhatla et al. 2015). This reevaluation or revisiting requires that the original, raw EM image volumes remain associated and in the same coordinate system as the published circuit maps. Our framework supports this use case naturally, as circuits are mapped from the raw image data and can remain in the same system after public release.

Connectomes are useful to different neuroscience subfields

Maps of neural circuit connectivity are therefore an invaluable resources to advance research in different subfields of neuroscience. Modern studies of the neural basis of behavior in model organism such as *Drosophila* have access to powerful genetic toolkit to monitor and manipulate the neural activity of cell types in behaving

animals (Meinertzhagen et al. 2012). In genetic model systems, hypotheses about the mechanism of neural computations can be generated from synaptic connectivity maps that can then be functionally and behaviorally tested. Circuit maps with identified neural pathways can restrict the space of possible experiments that have to be performed, for instance, by providing candidate cell types involved in certain pathways.

Novel imaging methods can record the activity of populations of neurons at the whole-brain scale (Keller et al. 2015). The correlated neural activity of populations of neurons defines functional circuits. The understanding of neural dynamics in functional circuits could be tremendously informed by the availability of whole-brain scale structural connectivity maps. For instance, the changes in neural activity invoked by behavioral learning are related to anatomical network constraints (Sadtler et al. 2014). Network connectivity can make learning of more familiar patterns easier than unfamiliar ones, because familiar patterns can arise more naturally from already existing connectivity structure.

At the cellular level, single neuron reconstructions from EM have a high degree of anatomical detail and are useful for computational models. Biologically-realistic compartmental models of the morphology of single neurons can be combined with the distribution of specific types of synaptic inputs across their dendritic arborizations (O'Connor et al. 2009). Such detailed models combined with physiological parameters might provide insights into how specific patterns of synaptic inputs interact with the postsynaptic membrane properties to generate neuronal responses.

Neural connectivity data stored in our framework could be accessed by researchers in labs from these different subdisciplines of neuroscience. Researchers do not need to know about the particular process that generated the circuits maps. They can access a project online, search for their favorite neuron types and retrieve the information about their connectivity structure.

Neuronal and circuit data can also be exported in a variety of formats depending on the researcher's needs for further postprocessing. For instance, neuron morphology data can be exported in the standard SWC or NeuroML formats (Gleeson et al. 2010), or network connectivity as matrices in CSV format or other graph formats such as GraphML. Data export options into formats such as NeuroML facilitates data interoperability, for instance to use mapped neuronal morphologies and circuits in computational models that are shared on public platforms such as the *Open Source Brain*[6].

---

6 http://opensourcebrain.org/

Impact of the CATMAID circuit mapping framework

Publicly available information on neural circuitry is expected to have a big impact on neuroscience, neurology, psychiatry and other fields, analogically to the availability of genome datasets to biochemistry, molecular biology and cell biology. Our framework offers convenient means to make structural circuit data available online to researchers around the world. As a proof-of-concept of the versatility of our framework, I have imported several published neural circuit datasets in CATMAID with associated annotations (Figure 6.3). These circuits were mapped using independent EM mapping workflows (Figure 6.3a,b) or by reconstructing neurons from light microscopy images and inferring connectivity (Figure 6.3c).

At the time of writing, two circuit maps have been mapped in our framework and were already published in *Drosophila* larva (Ohyama et al. 2015), as well as in Platynereis, a polychaete worm (Randel et al. 2015). Currently, a number of labs world-wide and a few large-scale projects use CATMAID for neural circuit mapping. This includes the efforts at HHMI Janelia Research Campus to map a whole adult fly brain from a large, serial-section TEM dataset as part of the FlyTEM project, with more than 25 concurrent annotators (August 2015). Another large-scale effort aims at mapping the wiring diagrams of *C. elegans* larvae at different developmental stages using CATMAID[7]. In a two year period, the *C. elegans* mapping project has seen contributions from more than 25 different annotators world-wide. Futhermore, CATMAID is currently used to map circuits in the zebrafish[8] and mouse visual cortex[9].

---

7 Principal Investigators of this project are Mei Zhen - University of Toronto and Lunenfeld-Tanenbaum Research Institute, Toronto, ON, Canada; Aravinthan D.T. Samuel - Center for Brain Science, Harvard University, MA; Jeff Lichtman - Center for Brain Science, Harvard University, MA; Andrew Chisholm - Division of Biological Sciences, University of California, San Diego, CA; David Hall - Albert Einstein College of Medicine, New York, NY.

8 David Hildebrand at Harvard University

9 Clay Reid at Allen Brain Institute; Wei-Chung Allen Lee at Harvard University
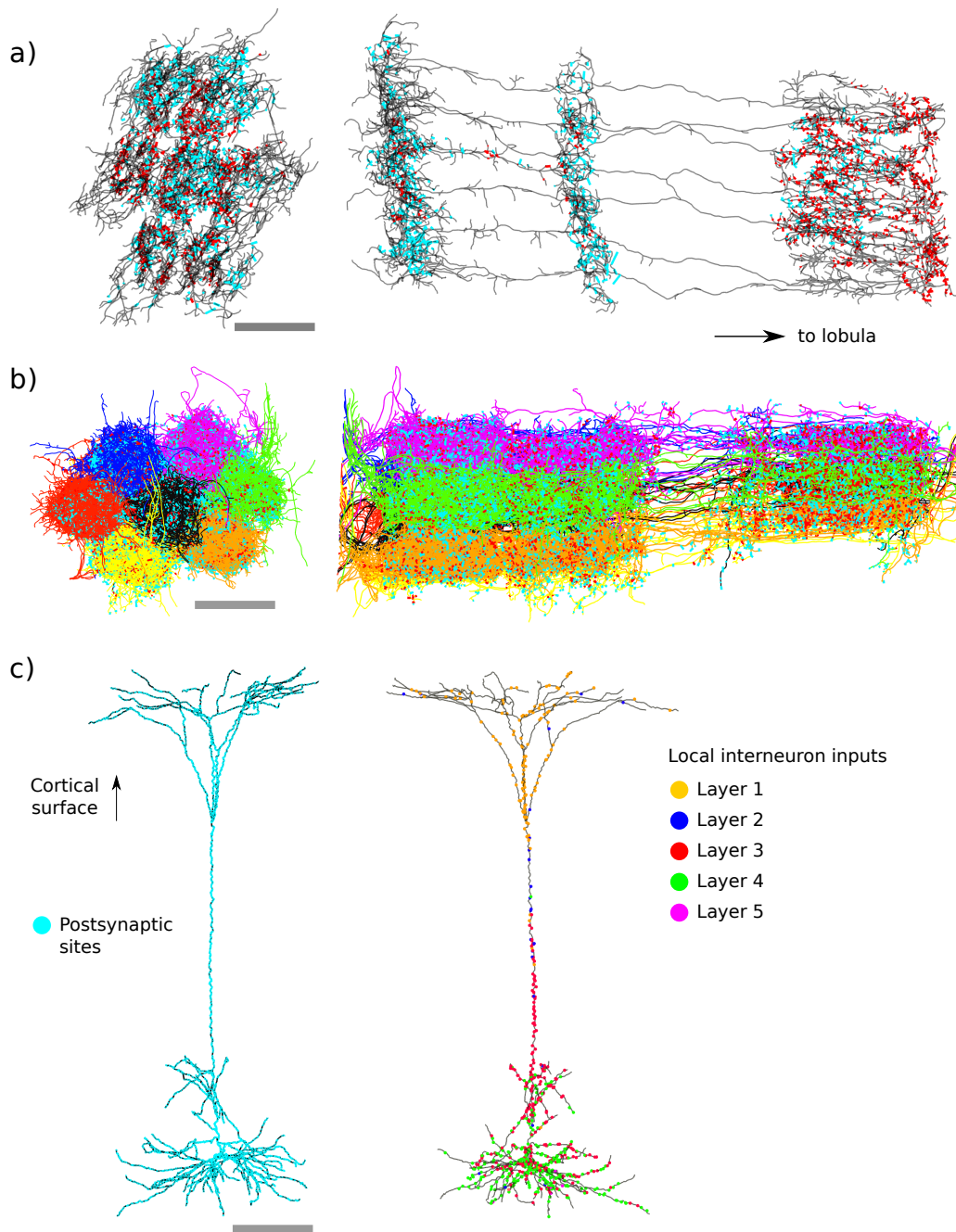
Figure 6.3.: **Neural circuits mapped by different workflows imported into CAT-MAID**. a) Several Mi1 neurons reconstructed from the *Drosophila* adult medulla circuit in Takemura et al. 2013. b) Singular column neurons of the *Drosophila* adult seven-column medulla circuit color-coded. (Data from Janelia Research Campus FlyEM github.com/janelia-flyem/ ConnectomeHackathon2015) c) Dendrite of a layer 5 pyramidal neuron with more than 20'000 postsynaptic sites from Narayanan et al. 2015. Inputs from cortical layer-specific local interneurons are interactively color-coded and visualized in CATMAID. (Data kindly provided by M. Oberlaender and R. Egger). Scale bars, 8 μm a) and b), 200 μm c).

6.2 COMPARATIVE CONNECTOMICS IN DROSOPHILA

The search for invariants in neural circuit structure promises to be a fruitful avenue to find general architectural principles underlying brain function (Helmstaedter 2015). In order to find such invariants, it is necessary to be able to compare neural circuits. With the ability to compare and analyze many circuits, it will become possible to extract general structural patterns such as network motifs. If we can understand how certain network motifs underlie particular canonical neural computation in one model system, it might be possible to understand how these motifs combine in different and more complex circuits (Carandini et al. 2011).

Applying connectomics to small invertebrate model systems, such as *Drosophila*, has important advantages (Meinertzhagen et al. 2012). The smallness and the genetic tractability of these systems allow to monitor and manipulate neural activity with exquisite detail at the level of individual cell types (Simpson 2009). Thereby, small organisms provide the unique opportunity to discover network motifs and reveal neural computational principles. The integration and association of data from anatomical, physiological and behavioral levels should allow to derive principles of circuit operation.

The hope is that principles and insights discovered in small system can inspire and guide research in larger, more complex systems. For instance, the conservation of homologous genes and biochemical pathways has been proven as a useful guide to find homologies, for instance, between invertebrate and vertebrate species (Katz et al. 2013). The evolutionary conservation of molecular pathways can also lead to conserved principles of neural circuit formation, such as in the case of dendritic morphogenesis (Jan et al. 2010). Similarly, the hope on the neural circuit level is that insights gained from studying circuits of invertebrate systems and their operation principles could translate and inspire our understanding of more complex circuit architectures, such as those in vertebrate systems.

In the same fashion, principles of neural circuit development might generalize across a wide range of species. In Chapters 3 and 4, I have focused my efforts to investigate the development of neural circuits at the synaptic level in isogenic *Drosophila melanogaster* larvae. Here, I mapped complete neurons and synaptic circuits in *Drosophila* larvae at several distinct developmental stages by taking advantage of the possibility of EM volumes that cover large parts of its nervous system. Similarly, *C. elegans* is small enough at distinct stages to enable mapping of circuits at the whole animal level. Genetic model organisms such as *Drosophila* larva or *C. elegans* provide thus an important opportunity, not possible in other larger systems, to gain insights into how synaptic circuits self-assemble and develop. Ultimately, such invertebrate studies can contribute importantly to tracing the effects of genetic mutations to the development of the wiring diagram and the resulting consequences for physiology and behavior (Brenner 2003).

Neurodegenerative diseases have presently unknown causes at the synaptic circuit level. Neurite degeneration, for instance, occurs after brain trauma and could underlie the functional loss observed in patients. Key signaling pathways that promote this destructive processes are now investigated in *Drosophila* and could

inform the understanding in vertebrate systems, including human patients (Rooney et al. 2014). Pure neurite degeneration can be studied without recurrence to EM connectomics. However, a better understanding of changes and reorganization at the circuit level will benefit from ultrastructural maps of the wiring diagram, in particular to evaluate treatment methods that target neurite regrowth. Neurons need to regrow and form again synapses with their appropriate partners, which can be tested using EM connectomics.

Also for many neurodevelopmental diseases, there are no good model systems to study how circuits differ in the diseased compared to normal individuals. Dendritic morphogenesis is a key factor in circuit formation, and its regulation by genes has been extensively studied in the past decades (Jan et al. 2010). Mutations in genes may create alterations in normal developmental processes, such as overgrowth or lack of arbor pruning or changes in synaptogenesis. These differences could be discovered with EM level connectivity maps immediately if good baseline datasets are available that describe the normal processes of circuit development. With the high level of detail available from EM micrographs, alterations beyond changes in the synaptic connectivity between cell types could be discovered. The properties studied in Chapters 3 and 4 might serve as a basis for future studies on the development of neural circuits in a genetically-tractable model system.

The rich digital data of detailed neuronal morphologies including synaptic distributions and connectivity obtained in this work could also be reused for computational studies of development. Computational models have tried to recapitulate developmental processes of circuit formation. For instance in tadpole, the connectome of the spinal cord has been produced using models of axonal growth (Borisyuk et al. 2011). The produced *in silico* neural circuit exhibited overall network features that were not preprogrammed in the developmental algorithms and were consistent with experimental data. Another framework aimed at modeling the development of neocortex by defining simple sets of genetic rules for circuit formation (Zubler et al. 2013). Due to the public availability of large datasets of single neuron morphologies [10], researchers interested in computationally modeling neuronal growth could use the realistic neuronal reconstructions to fit the parameters of their models. The sharing of circuit datasets at the EM level will further facilitate such computational research that can also be extended to integrate synaptic information into models. A number of software packages implement algorithms to generate realistic neuronal morphologies *in silico*. The packages could be extended with synaptic level information (Eberhard et al. 2006; Gleeson et al. 2007; Koene et al. 2009; Wolf et al. 2012).

The here reconstructed connectomic datasets and many future circuit maps could be used in a similar way to develop more sophisticated models of circuit formation. These models may reveal the minimal descriptions and parameters necessary to produce the wiring diagrams we observe *in vivo*. The set of different cell types contained in our data could provide insights into cell-type-specific variation in defining parameters of these models, as different neuron types express different developmental programs and thus may follow different growth logic. Such models would not only need to incorporate neuronal morphologies but also specific features

---

10 e.g. http://neuromorpho.org, also see Parekh et al. 2013

of synaptic distributions. As we have shown, the distribution of synapses is not uniform across a *Drosophila* arbor. For instance, most of the synaptic inputs are confined to distal, microtubuli-free twig processes. Information on the microtubuli distribution should be a constraint in such developmental models. Eventually, artificially generated circuits should display physiological and behavioral responses that are not previously encoded in the model. Recent studies in tadpole spinal cord indeed succeeded in validating artificial models with experimental data (Roberts et al. 2014).

I have chosen the nociceptive system of *Drosophila* larvae for this study, because it contains a comparatively small number of neurons for mapping but is still sufficiently large for circuit comparisons. Furthermore, it can provide a basis for future studies that aim at functionally dissecting nociceptive processing mechanism with possible implications for understanding pain perception in other systems. Far from being just a simple, reactive reflex system, the nociceptive circuits in *Drosophila* larva show a variety of hallmarks that one would not expect from a simple sensory system that has to presumably only signal a simple noxious-yes/noxious-no. Unexpectedly, the polymodal nociceptors in larvae not only transduce a variety of noxious stimuli, but have recently been implicated in aspects of proprioception (Gorczyca et al. 2014). Of course, from the careful description and observation of circuit structure and combining knowledge from the literature about this circuit's inputs, one can speculate about the possible functional implication of observed connectivity patterns. In order to make conclusive statements about functional involvements of particular neurons and pathways, however, monitoring and manipulation of neuronal activity in defined cell types with behavioral monitoring is necessary (e.g. Ohyama et al. 2015).

The nociceptive circuit maps presented in this work facilitate a host of avenues for further functional investigations. For this the existence of light-level neuron catalogs of neuronal cell types (for *Drosophila* larva by Jim Truman et al.) enable matching of reconstructed EM neuronal morphologies with neuronal cell types labeled in genetic driver lines (e.g. the GAL4/UAS system by Brand et al. 1993). This is crucial to create the link from synaptic circuit maps to genetic lines that can then be employed in optical imaging, electrophysiological or behavioral experiments. In *Drosophila* larvae, the advantage to have a limited set of cell types with usually only one or a few instantiation of neurons allows high confidence when matching reconstructed EM morphologies to labeled single neurons in confocal stacks at the resolution of light microscopy. With the identification of all downstream partner interneurons of primary nociceptors and their mapping to GAL4 lines, genetic manipulation of the cell types can be used to further dissect the function of different nociceptive pathways (Dierick et al. 2015). *Drosophila* larvae have been proposed as an emerging model system to study basic circuit mechanism of nociception (Im et al. 2012) and the synaptic nociceptive circuit maps allow for a rigorous, targeted approach aimed at a mechanistic, functional understanding.

The availability of high-throughput genome sequencing methods opened up the possibility to obtain genomes of many species as the basis for large comparative phylogenetic studies. Similarly, the field of comparative EM connectomics studies is likely going to grow considerably once EM data generation and connectome mapping becomes more automatized (Hale 2014).

In this thesis, I have singled out and focused on one aspect of the comparative connectomics spectrum, namely comparing neural circuits at different developmental time points. The development of neural circuits, however, is just one out of many possible axes of comparison such as for instance in aging, in health and disease, in learning or between species. The number of comparisons and questions that can be addressed are vast. Importantly, comparative connectomics studies between species have the potential to provide deep insights into similarities and differences in neural network architectures and the evolution of neural computation (Yuste 2008; Katz et al. 2013).

Circuit comparison across species

A pioneering EM study compared neural circuitry in closely related nematode species (Bumbarger et al. 2013). Both nematode species exhibit differences in their feeding behaviors and were found to have marked differences in synaptic connectivity of identifiable, homologous neurons of the pharyngeal system. This finding suggests that synaptic connectivity differences might be responsible for divergent behavioral outputs. However, nothing precludes the possibility that in principle, very similar network architecture can produce divergent behavioral outputs. Such multifunctional circuits can be switched, for instance, by neuromodulatory effects that bias network dynamics into different operating regimes (Briggman et al. 2006b).

A large body of structural and functional data across many related species will be necessary to conclusively state the effective influence of network topology and connection strength on behavioral performance. Storing collections of structural maps of many species and brain regions, for instance in a framework such as CATMAID, would simplify such comparisons tremendously. Akin to the genomics community, where newly sequenced and published genomes are deposited in centralized databases, neural circuit datasets from different species could be made accessible in centralized repositories in a similar fashion. Although the brains of invertebrate and vertebrate species might differ vastly in scale, gross similarities in the basic ground plan of brain architecture have been proposed to be preserved (Strausfeld et al. 2013). Comparing cell types, neural circuits and network motifs between species at different spatial scales, from brain regions to the individual neuron level, would allow for an in-depth study of the evolutionary origins of neural circuits required for neural computations that are common between species (Shinomiya et al. 2015).

Circuit comparison across brain regions

The comparison of neural circuits across brain regions in the same species is another, potentially very fruitful, endeavor. For instance, similarities in the organization of different primary sensory cortical areas have been established, but the extent of uniformity of fine-level details such as synaptic connectivity of existing cell types are unknown (Streissler et al. 2002). Scaling volumetric EM connectomics to larger areas of the cortical sheet might provide sufficient information to reveal invariant circuit motifs and also differences at the synaptic level that could underlie the *atoms of neural computation* (Marcus et al. 2014). Despite known differences, canonical synaptic connectivity motifs, and thus potentially principles of elementary functions could be shared across cortical areas (Douglas et al. 2007). A shared organization would indicate that streams of inputs originating from different sensory modalities in the periphery are processed centrally with fixed sets of computational primitives before propagation to higher association areas. In this context, the comparison of differences in species could provide important insights into the diversity of such computational primitives, if increasingly sophisticated sensory and cognitive capabilities of different species and regions could be reflected in their neural circuit substrate (Marín-Padilla 1998; Elston et al. 2006).

Circuit comparison in learning and memory

Another interesting avenue to pursue for comparative connectomics is investigating the extent of structural changes in synapses and synaptic circuits in the context of learning and memory (e.g. Holtmaat et al. 2009; Seung 2013). Experience, that is the ongoing activity in neural circuits, shapes the underlying circuit structures at molecular, synaptic and network levels. Large-scale EM volumes will provide the opportunity to compare changes of synaptic properties and network level synaptic structure in populations of organism before and after learning. Learning can take a variety of forms, from simple habituation to more complex associate learning tasks such as operant conditioning. Simpler forms of learning, such as habituation, might not have easily discernible signature at the ultrastructural level and might be implemented by changing synaptic efficacy on the molecular level. For more complex forms of learning, such as associative learning, distinct brain structures have been shown to be necessary, such as the mushroom bodies in *Drosophila* (e.g. Séjourné et al. 2011). Such learning processes are thought to modify a circuit's synaptic strength by changing, for instance, properties of individual synapses or the formation of new synapses between sets of neurons (Bailey et al. 2015). Depending on the exact morphological properties (Marrone et al. 2002), these changes could be extracted from large EM volumes, and linked to network function and behavioral phenotypes. However, being able to generate large samples will likely be crucial to find the structural instantiation of memories, and to be able to distinguish them from individual variation arising, for instance, from developmental variability.

The discovery of the structural substrate of long-term memories has so far eluded scientific studies. The novel views offered by volumetric EM connectomics might

be able to provide new avenues for its study. Recently, a hypothesis proposed that long-term memories could be stored in patterns of holes in the perineuronal nets (Tsien 2013). These nets are specialized extracellular structures that envelop neurons and can constrain both the outgrowth of neurites as well as the formation of synapses. Together with an arsenal of genetic methods, volumetric EM could provide the means to map these nets and their relationship with neurons and synapses and help clarify their role in long-term memory storage.

Other proposals for the substrate of long-term memory have been made where volumetric EM may provide data to test their predictions. The cytoskeleton of neurons, consisting of stable, intracellular microtubules, were suggested to provide long-term structural support for the persistence of synaptic connectivity (Smythies 2015). This hypothesis could be tested by using a model circuit where sets of neurons are known to be necessary for long-term memory support. These neurons could be reconstructed from volume EM including their microtubules cytoskeleton in conditions with and without induced long-term memory. Then, their cytoskeletal and synaptic characteristics could be compared, similar to the analyses in Chapter 4.

Circuit comparison in health and disease

With brain disease and mental illness being among the largest cost drivers in health care in the world (Collins et al. 2011), comparative EM connectomics of healthy and diseased circuits is prone to become a major connectomics focus in the coming decades. Overly simplified explanatory models of mental disorders, such as chemical imbalances of the brain in depression, may hamper progress in understanding associated circuit changes to a sufficient level of detail and thereby misguide intervention strategies (Castrén 2005). Finding structural markers in the synaptic connectivity in diseased compared to healthy circuits would facilitate the development of specific, targeted intervention strategies. In model organism, the effect of pharmacological interventions on circuit connectivity, function and behavior could be monitored and treatment outcomes evaluated in a more principled way. Interventions could be evaluated with regard to their effectiveness to transform aberrant connectivity structure back to normal.

Furthermore, the detailed knowledge of aberrant circuit connectivity can inform our understanding of disease etiology by delineating genetic, developmental vs. environmental, experience-based origins. The elucidation of how a particular genotype of an individual might contribute to the development of neural circuitry that is more susceptible to environmental factors, such as stressful life events, leading to mental disorders would be a major medical advance. At present it is unknown to what extent such advances can rely solely, or at all, on structural circuit data obtained from EM, but careful comparative circuit studies with large sample sizes might reveal potential signatures.

However, future research into comparative connectomics will only advance if reliable high-throughput data acquisition and circuit reconstruction pipelines can be built. Pipelines in genomics have reduced the time it takes to sequence a few genes from a whole PhD project to fractions of a second within a few decades and the same

can be achieved for EM connectomics[11]. What then remains are the challenges to make sense of the data, and to spend all of our available brain cycles to analyze and understand the intricate system that is cycling itself.

---

11 A story told by the director of Janelia Research Campus, Dr. Gerry Rubin, about his PhD project in genomics. I am positive to tell a similar story about my PhD project in a not too distant future.

APPENDIX

A

ELECTRON MICROSCOPY DATASETS

The electron microscopy datasets used in this thesis project were generated at HHMI Janelia Research Campus using serial-section Transmission Electron Microscopy (ssTEM). The four datasets cover various portions of the central nervous system (CNS) of *Drosophila melanogaster* larvae at early and late developmental stages. In the following sections, the people and labs are mentioned that funded and were involved in the various aspects of the data generation and alignment process. Additionally, technical details on the individual datasets are given.

The transverse views provide orientation of the CNS coverage of individual datasets and include the brain or portions of the thoracic and abdominal segments of the ventral nerve cord (VNC). The views were generated using the BigDataViewer by referencing the original CATMAID tile sources of the respective image stacks (Pietzsch et al. 2015). The cross-sectional views demonstrate the sample quality and contrast at the highest available magnification. Typically, circuit mapping is performed at the highest magnification level.

A.1 L1A: FIRST-INSTAR WHOLE LARVAL CNS

Funding	HHMI Janelia Research Campus, FlyEM Project
Sample preparation	Richard Fetter
Sectioning/Collection	Richard Fetter
Imaging	Richard Fetter, Shirley Lauchie, Andrea Brothers
Microscope	FEI Spirit BioTWIN TEM and Leginon
Alignment	Albert Cardona, Stephan Saalfeld with TrakEM2
Reference publication	Ohyama et al. 2015
Covering tissue	Brain/SOG/VNC L1
	(The larva was kept and screened for sex-specific DNA and found to be female)

XY dimension	28128 x 31840
Raw tile dimension	4k x 4k pixels
Raw tile size	15.6 µm x 15.6 µm
Number of sections	4841
Number of missing sections	56
XY resolution	3.8 x 3.8 nm
Section thickness	50 nm
CNS volume	874,283 µm^3
Neuropil volume	226,652 µm^3
Neuropil filling fraction	25.9%
Time to image	29 weeks (4597 TEM hrs)
Number of images	187,048
Raw dataset size	5.98 TB (MRC)
Image pyramid size	994 GB (JPG)

Table A.1.: Dataset information
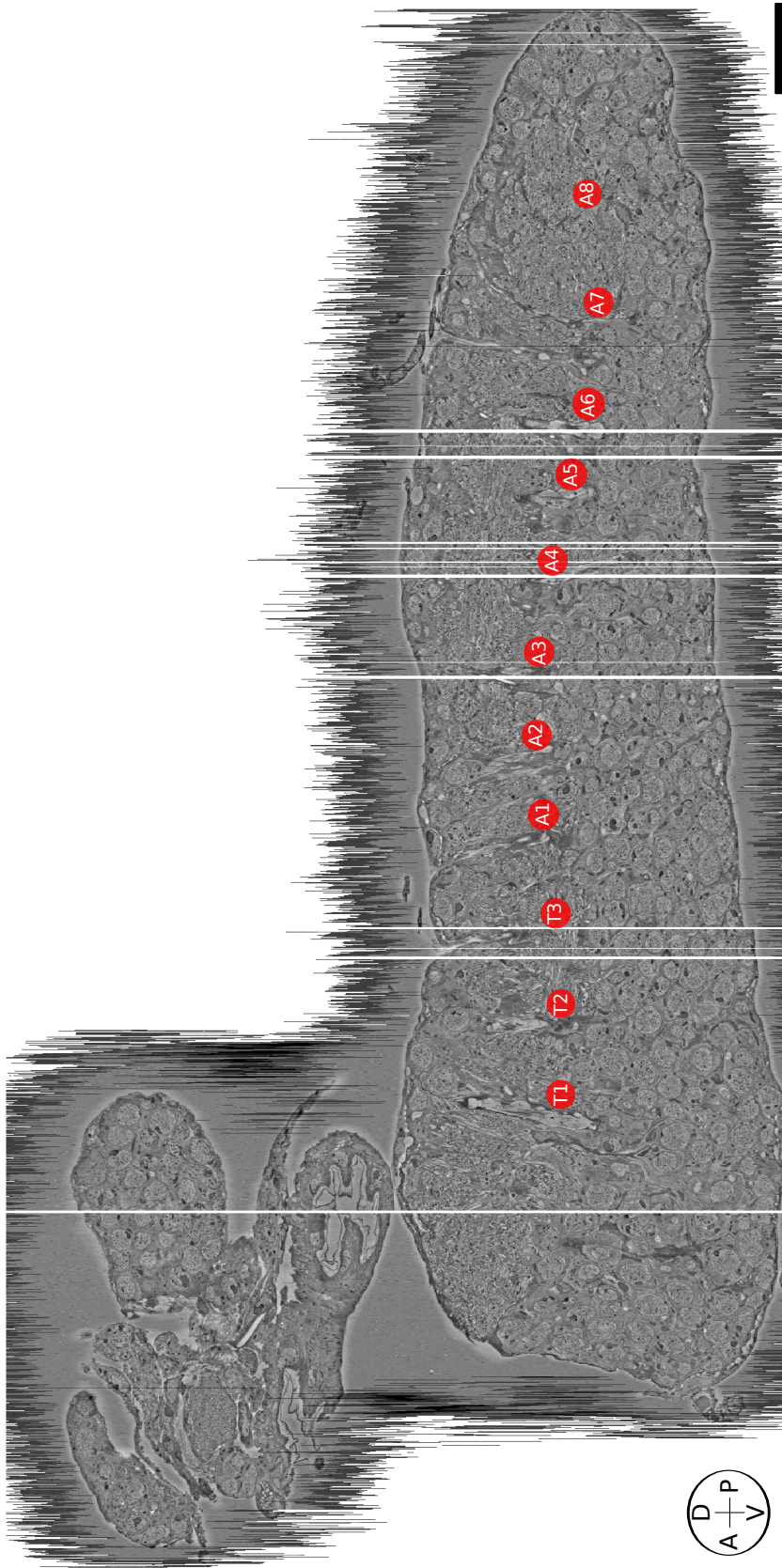
Figure A.1.: **First-instar dataset L1a.** Transverse section of the whole CNS first-instar dataset at the ventral nerve cord midline. VNC commisural locations are marked with their associated thoracic and abdominal segment (red circles). White vertical lines indicate sections that were lost in the sample preparation or imaging session. Section thickness is approximately 50 nanometer. Scale bar 15 μm.

A.2 L1B: FIRST-INSTAR LARVAL VENTRAL NERVE CORD

Funding	HHMI Janelia Research Campus
	Visiting Scientist Program to Albert Cardona
	sponsored by Julie Simpson and Richard Fetter
Sample preparation	Richard Fetter
Sectioning/Collection	Richard Fetter
Imaging	Albert Cardona and Wayne Pereanu
Microscope	FEI T20 TEM
Alignment	Albert Cardona, Stephan Saalfeld with TrakEM2
Reference publication	Schneider-Mizell et al. 2015
Covering tissue	VNC (mid-A2 to A3)

Raw tile dimension	2k x 2k pixels
Raw tile size	8.9 µmx 8.9 µm
XY dimension	22775 x 18326
Number of sections	462
Number of missing sections	4
XY resolution	4.4 nm/pixel (CATMAID:4)
Section thickness	50 nm (CATMAID:45)
Image pyramid size	90 GB (JPG)

Table A.2.: Dataset information
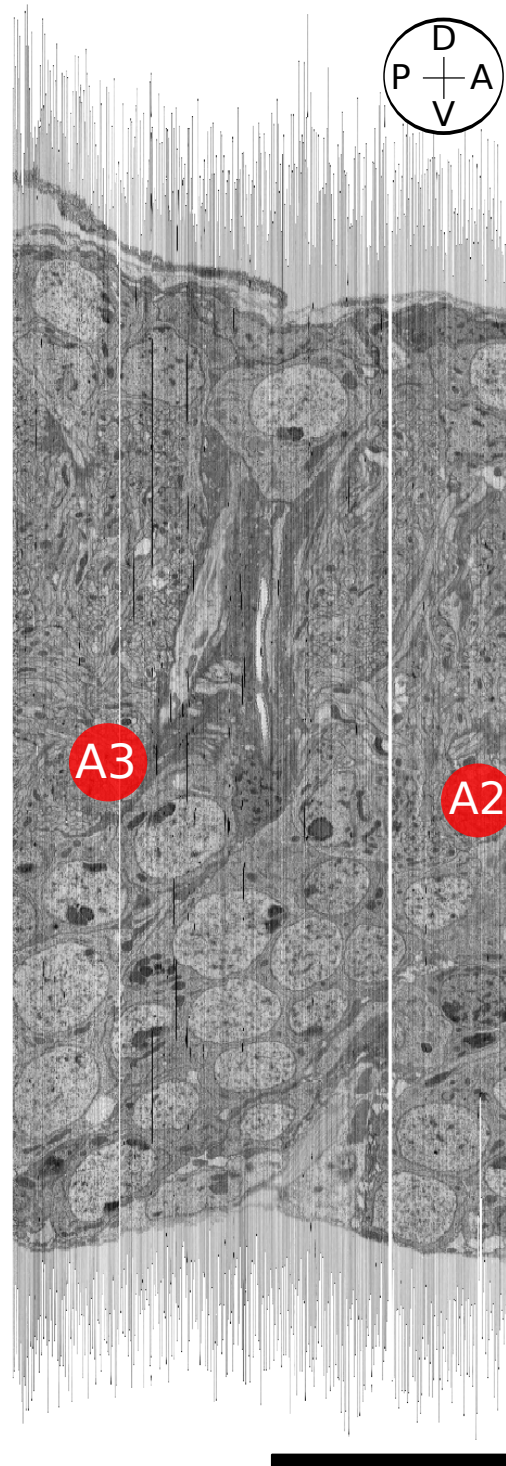
Figure A.2.: **First-instar dataset L1b**. Transverse section through the abdominal ventral nerve cord covering segments A2 and A3. Section thickness is approximately 50 nanometer. Scale bar 10 μm.

A.3 L1C: FIRST-INSTAR LARVAL VENTRAL NERVE CORD

Funding	HHMI Janelia Research Campus
	FlyEM Project and Harald Hess lab
Sample preparation	Richard Fetter
Sectioning/Collection	Richard Fetter
Imaging	Mehdi Bolorizadeh (Hess lab) with STEM
Alignment	Stephan Gerhard, with help of Stephan Saalfeld.
Alignment Method	Elastic alignment with modification
	to deal with large tile sizes in TrakEM2
Reference publication	-
Dataset availability	http://neurocity.janelia.org/catmaidstem
Covering tissue	Ventral Nerve Cord (posterior T3 to anterior A4)

Raw tile dimension	20k x 20k
Raw tile size	80 µmx 80 µm
Tiles per section	1
XY dimension	25344 x 17408
Number of sections	1996
Number of missing sections	71
XY resolution	4 x 4 nm
Section thickness	50 nm
Image pyramid size	826 GB (PNG)

Table A.3.: Dataset information
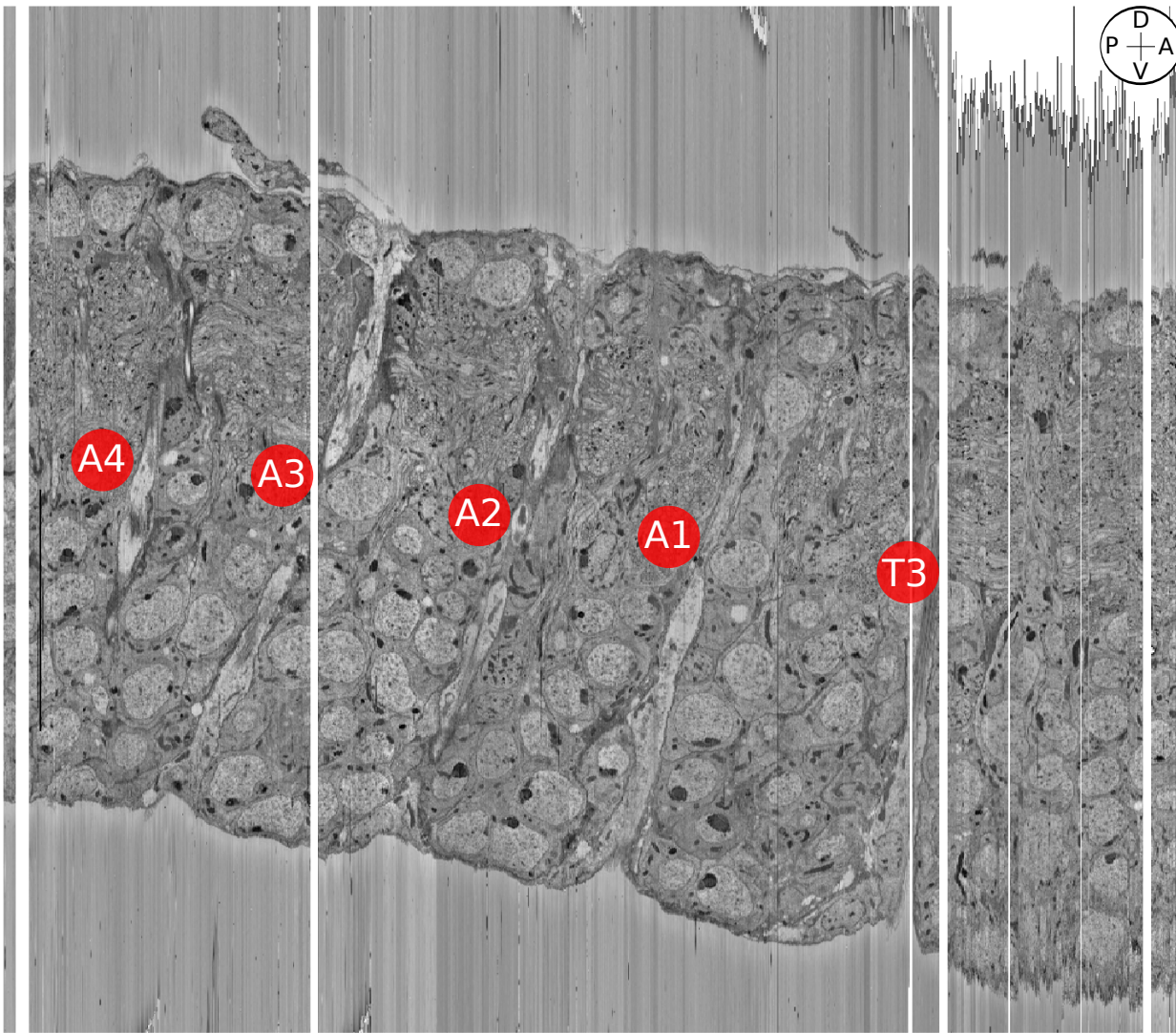
Figure A.3.: **First-instar dataset L1c**. Transverse section through the abdominal ventral nerve cord covering segments from thoracic T3 to abdominal A4. Section thickness is approximately 50 nanometer. Scale bar 15 μm.

A.4 L3: THIRD-INSTAR LARVA VENTRAL NERVE CORD

Funding	HHMI Janelia Research Campus, FlyEM Project
Sample preparation	Richard Fetter
Sectioning/Collection	Richard Fetter
Imaging	Richard Fetter, Shirley Lauchie, Andrea Brothers
Microscope	FEI Spirit BioTWIN TEM and Leginon
Alignment	Stephan Gerhard, with help of Stephan Saalfeld in TrakEM2
Reference publication	Gerhard et al. (in preparation)
Larval stage	L3 (wandering)
Covering tissue	VNC (mid-A2 to mid-A6)

Raw tile dimension	4k x 4k pixels
Raw tile size	9.4 μmx 9.4 μm
XY dimension	61952 x 46592
Number of sections	2156
Number of missing sections	50
XY resolution	2.3 x 2.3 nm
Section thickness	50 nm
Time to image	10 weeks (1642 TEM hrs)
Number of images	336,326
Image pyramid size	Compressed JPEG
	scale level 0: 1.4 TB
	scale level 1: 515 GB
	scale level 2: 103 GB
	scale level 3: 25 GB
	scale level 4: 5.9 GB
	scale level 5: 1.5 GB
	scale level 6: 375 MB
Raw dataset size	10.76 TB (MRC format)

Table A.4.: Dataset information

Figure A.4.: **Third-instar dataset L3**. Transverse section through the abdominal ventral nerve cord covering segments from abdominal A2 to abdominal A6. Section thickness is approximately 50 nanometer. Scale bar 15 μm.

A.5 EM MICROGRAPHS OF FOUR DATASETS
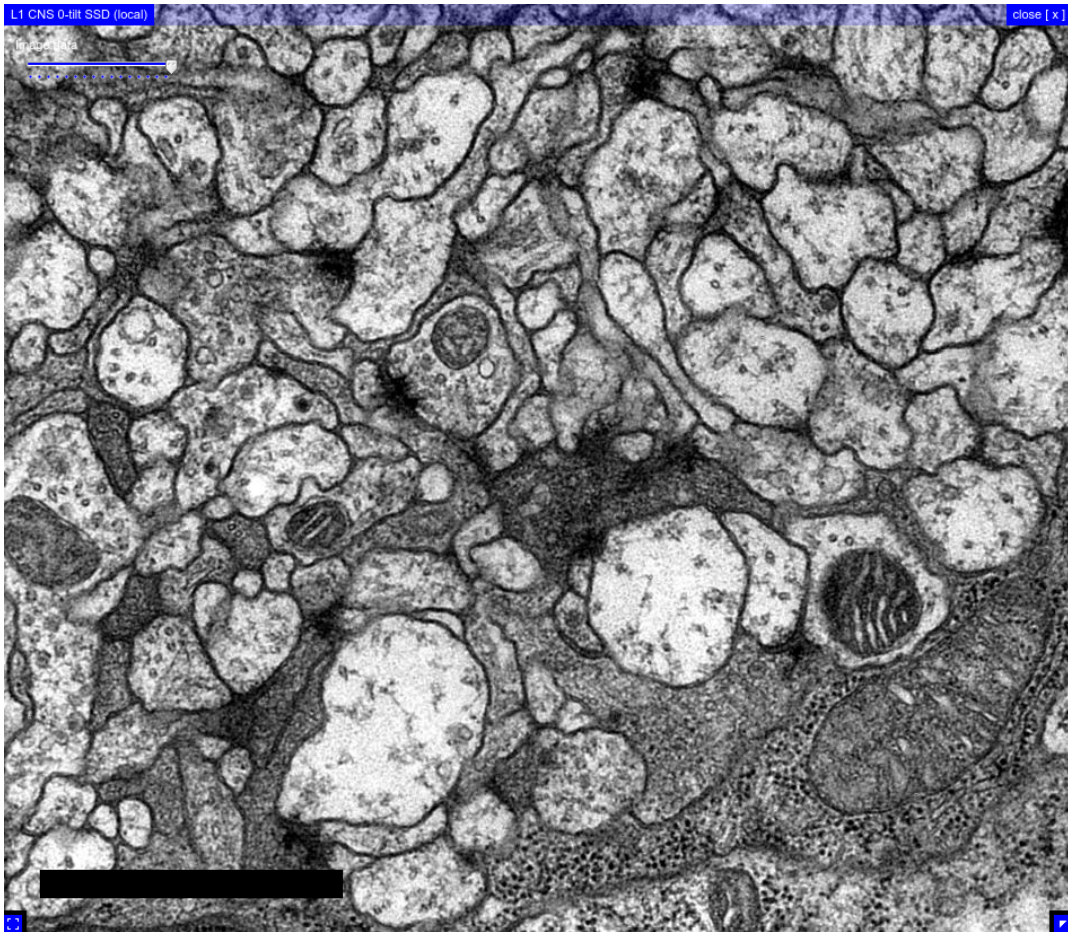
Figure A.5.: **Dataset L1a** View of a cross-section (XY) of the dataset at acquisition magnification. Scale bar 1μm. Pixel resolution 3.8 nm/pixel.
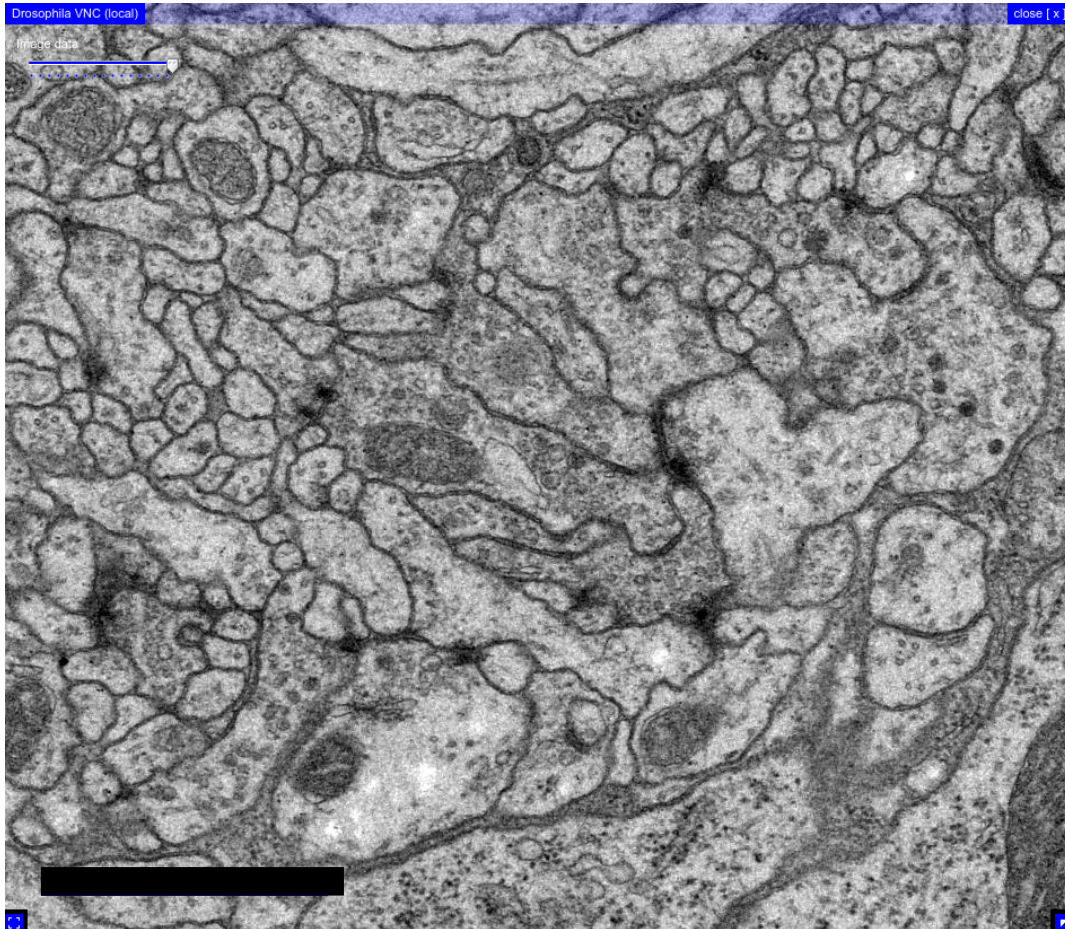
Figure A.6.: **Dataset L1b** View of a cross-section (XY) of the dataset at acquisition magnification. Scale bar 1μm. Pixel resolution 4.4 nm/pixel.
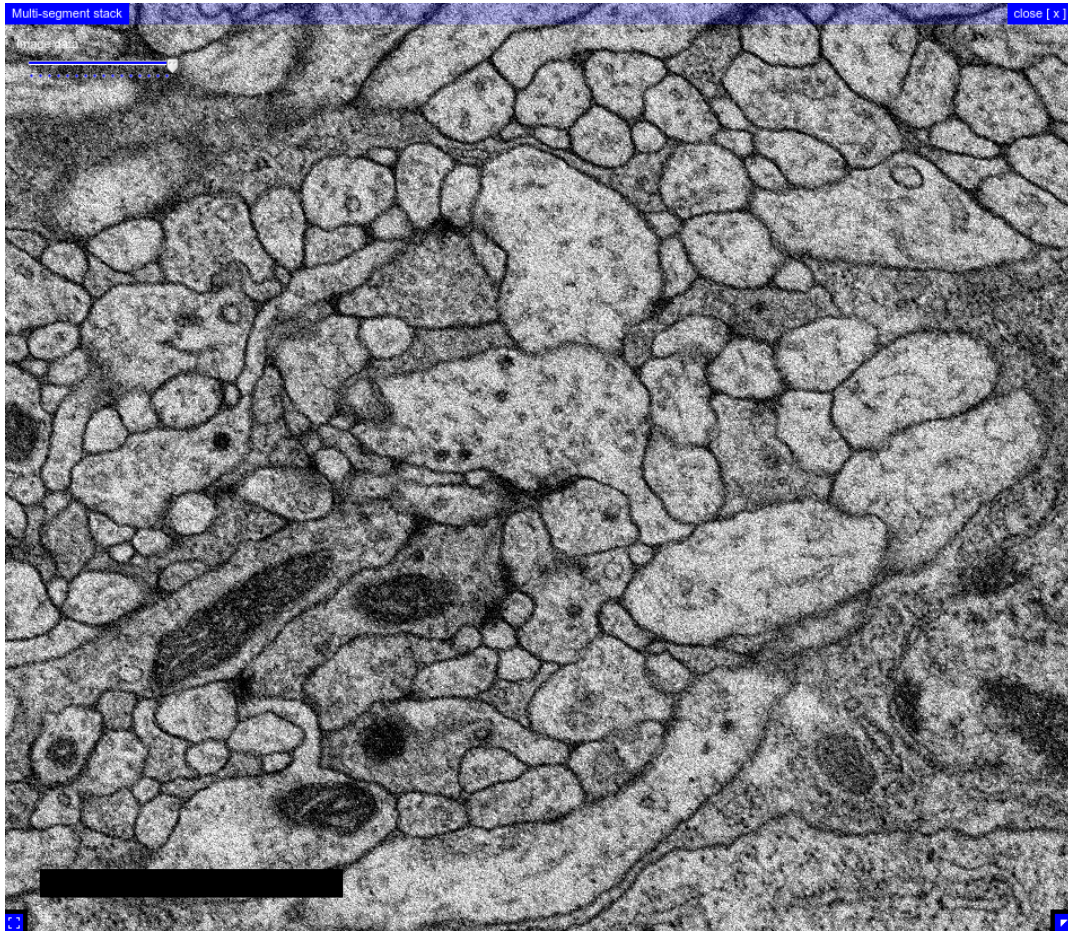
Figure A.7.: **Dataset L1c** View of a cross-section (XY) of the dataset at acquisition magnification. Scale bar 1μm. Pixel resolution 4 nm/pixel.

Figure A.8.: **Dataset L3** View of a cross-section (XY) of the dataset at acquisition magnification. Scale bar 500 nm. Pixel resolution 2.3 nm/pixel.

CATMAID PROJECT

B.1 CONTRIBUTORS
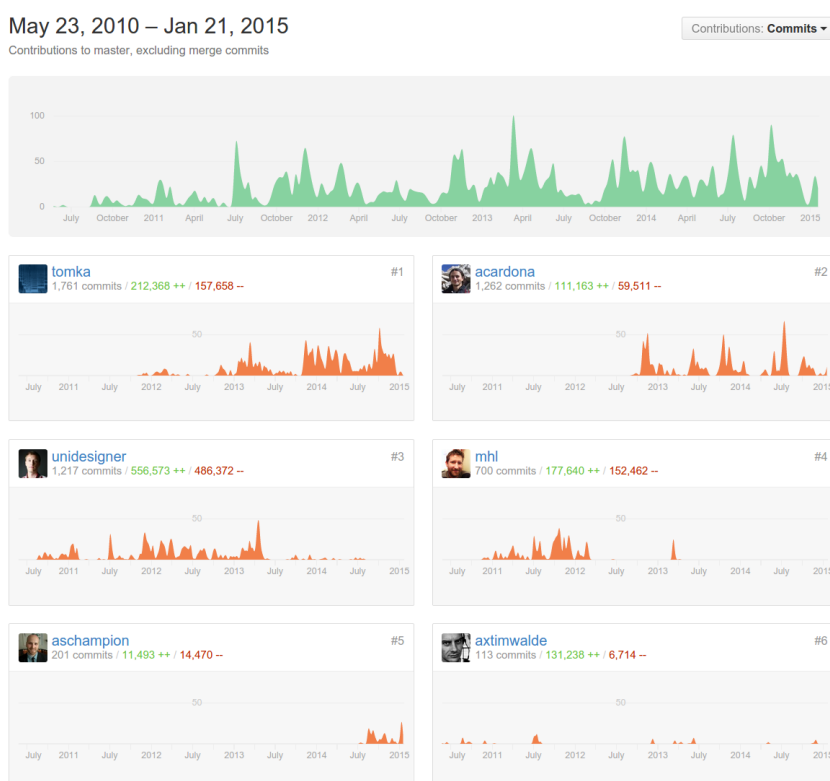
Figure B.1.: **History of contributions as numbers of commits by the six major contributors in a time window of five years:** The development of the neural circuit reconstruction tool in CATMAID started with an initial commit to the repository of the original CATMAID code-base published in Saalfeld et al. 2009 (`http://fly.mpi-cbg.de/˜saalfeld/catmaid/`). This original version implemented a multi-scale browsing interface for 3D and 2D datasets with basic text annotation capability and a cropping tool. Contributors full name: Tom Kazimiers (tomka), Andrew Champion (aschampion), Albert Cardona (acardona), Stephan Gerhard (unidesigner), Mark Longair (mhl), Stephan Saalfeld (axtimwalde).

B.2 TRAFFIC TO CATMAID.ORG
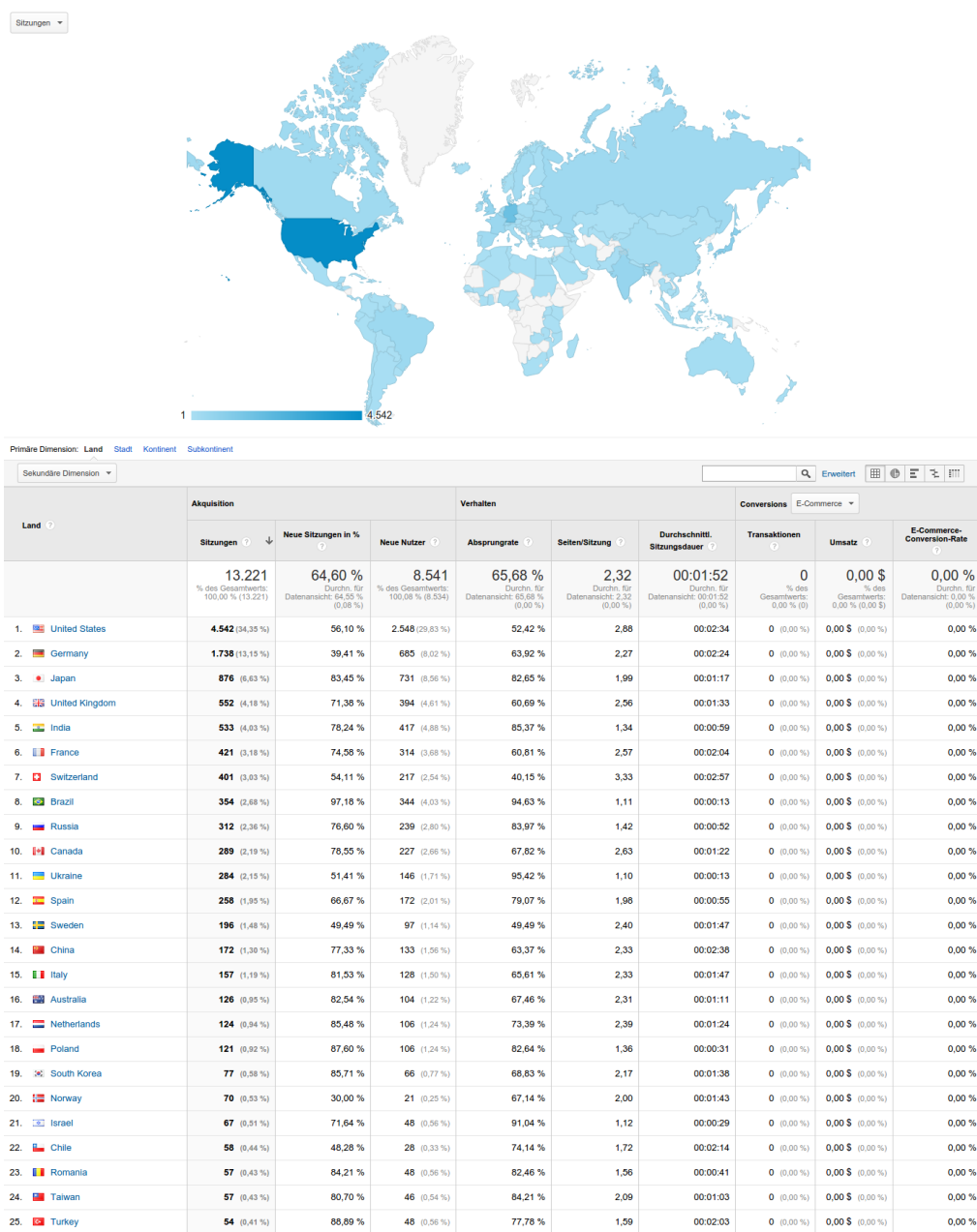
Figure B.2.: **Access statistics for `www.catmaid.org`** Bottom panel shows the top 25 countries that accessed the website since its creation. (*Screenshot from Google Analytics*)

B.3 OPEN-SOURCE TOOLSTACK

backend

nginx	webserver	nginx.org
postgres	relational database	postgresql.org
django	web framework	djangoproject.com
celery	distributed task queue	celeryproject.org
pgmagick	image manipulation	pgmagick.readthedocs.org
travis	continuation integration	travis-ci.org

javascipt frontend libraries

jquery	web page manipulations	jquery.com
datatables	table plug-in for jquery	datatables.net
numeric.js	web matrix computations	numericjs.com
jsnetworkx	graph manipulation	jsnetworkx.org
raphaeljs	scalar vector graphics for overlays	raphaeljs.com
gRaphael	charting library based on raphaeljs	g.raphaeljs.com
colorwheel	interactive colorwheel	jweir.github.io/colorwheel
d3.js	interactive plots and charts	d3js.org
three.js	3d graphics based on WebGL	threejs.org
cytoscapejs	interactive graphs	cytoscape.github.io/cytoscape.js
fabric.js	HTML5 canvas library	fabricjs.com
sylvester	vector, matrix and geometry library	sylvester.jcoglan.com
arbor.js	a graph vizualization toolkit	arborjs.org
filesaver.js	implements the saveAs interface	github.com/eligrey/FileSaver.js

python libraries for client-side analysis and visualization

networkx	graph library	networkx.github.io
scipy and numpy	numerical computations	scipy.org
pandas	data analysis library	pandas.pydata.org
matplotlib	2d plotting library	matplotlib.org
mayavi2	3d visualization	code.enthought.com/projects/mayavi
scikit-learn	machine learning package	scikit-learn.org
scikit-image	2d image manipulation	scikit-image.org
ipython	2d interactive python	ipython.org

Table B.1.: **The CATMAID framework build on a large number of open-source tools and libraries** In the backend, we count on a tool stack that is widely used world-wide in web production systems with millions of users. The development of the CATMAID frontend user interface was tremendously accelerated by using a number of widely-used JavaScript libraries. A number libraries for the Python programming language were used for connectomics data analysis and visualization.

B.4 DATA MODEL
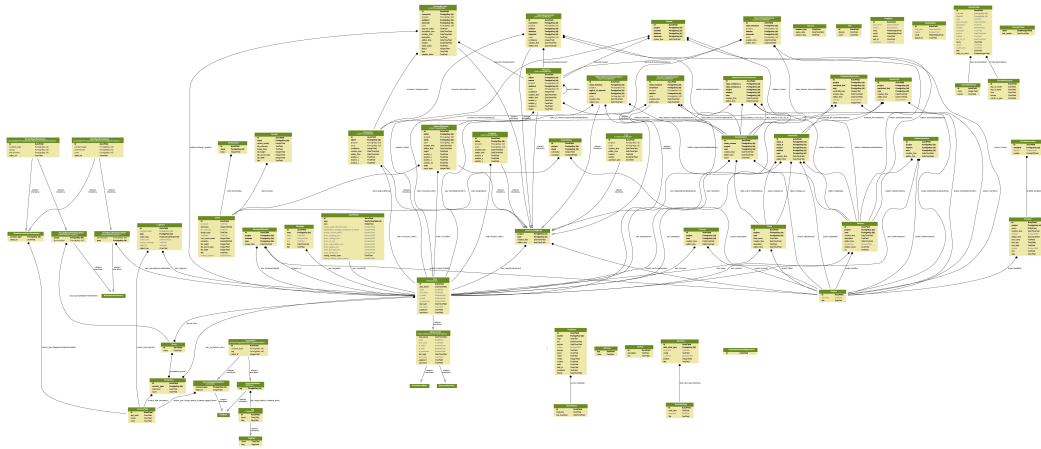
Figure B.3.: **Entity-Relationship diagram of the CATMAID data model**. With more than 70 tables, the CATMAID data model has grown considerably over the years. Models include for instance tables for user managment and permissions, circuit and annotation data, region of interests or project and stack configurations.
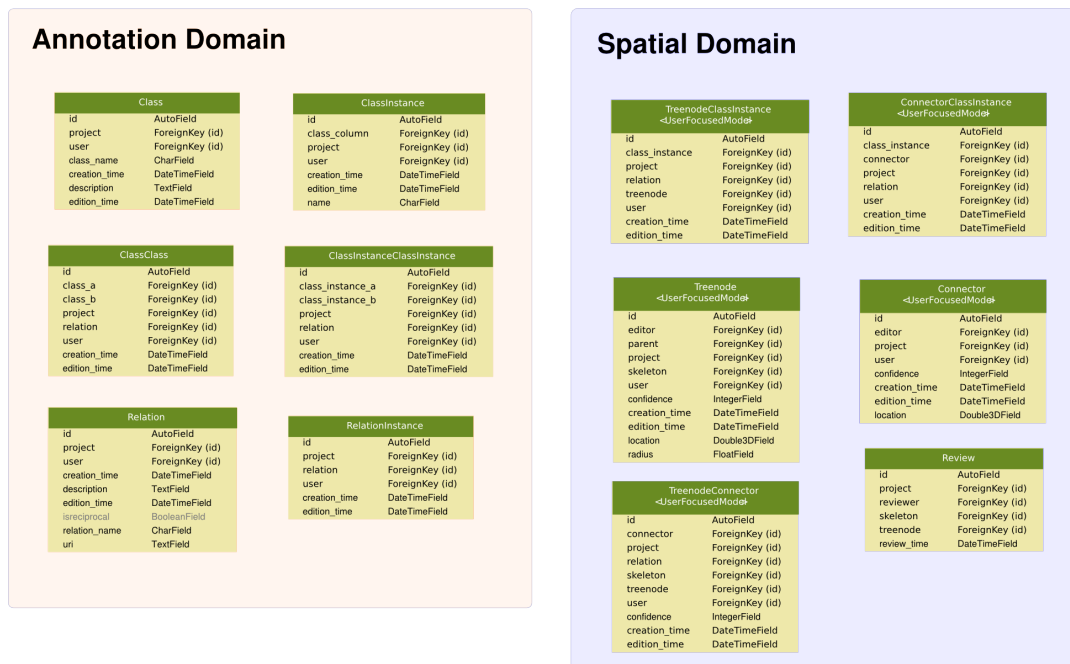
Figure B.4.: **The core neural circuit data models**. Representations for neural circuit data is separated by semantic information about neuron identities, names, annotations and tags, and spatial, geometric information (skeletons for neurons, connectors for synapses), and the relationships between these two domains.

B.5 REMOTE API ACCESS TO CATMAID USING PYTHON

```python
# Create a connection to a CATMAID instance in Python.
from cat.connection import Connection
c = Connection( CATMAID_URL, USERNAME, PASSWORD, CATMAID_PROJECT_ID )
c.login()

# Retrieve a skeleton morphology, its name and
# synapse information
from cat import morphology
skeleton = morphology.get_skeleton(c, SKELETON_ID )

# Update the skeleton graph edges with distance between to skeleton
#  nodes and compute the total cable length of the skeleton
from cat import labeling, features
labeling.update_skeleton_edge_with_distance( skeleton )
print features.get_total_cable_length( skeleton )
```

Figure B.5.: **Example code to interact with a remote CATMAID instance from Python** Web services provide a mechanism to retrieve neural circuit data from the remote CATMAID instances. The example demonstrates how to login into a CATMAID instance, retrieve a skeleton morphology of a neuron given a SKELETON_ID and compute a morphological feature, namely the total arbor cable length. We provide a thin library for authentication and access to CATMAID web server instances (github.com/catmaid/cat). With this approach, neurons and circuit data can be pooled across multiple CATMAID instances in a local IPython shell and custom, offline circuit analysis and visualization can be performed.

B.6 CATMAID WIDGETS

A CATMAID instance is accessed via a web-browser, such as Google Chrome or Firefox, by visiting its URL. After the login and the selection of the image stack, the user is presented with the *Stack Widget* showing the first 2D image of the image stack at low resolution (Figure B.6). In the top toolbar, a number of tools are available, including the the tracing tool for circuit mapping and analysis. Functionality of the tracing tool that supports the mapping, analysis and visualization workflows are encapsulated in individual widgets. Widgets can be resized and arranged individually in the visible window to optimally layout the workspace. In the following figures, a number of widgets and their functionality are briefly described.
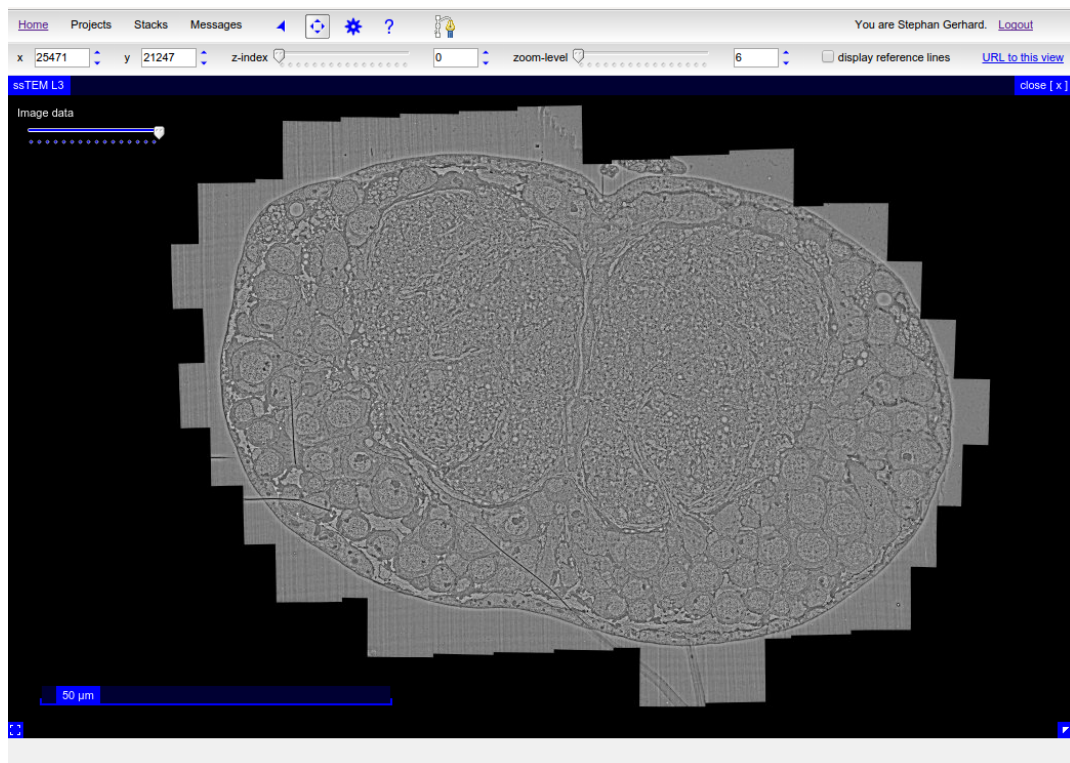
Figure B.6.: **Stack Widget** Large 3D image stacks can be browsed at multiple scale levels. Zoomed out, low resolutions views of the image data are shown here from a cross-section of larva ventral nerve cord. These low resolution views are useful to navigate and orient using gross anatomical landmarks such as cell bodies. The scale bar shows the absolute size of image data displayed at a given scale level (bottom left).
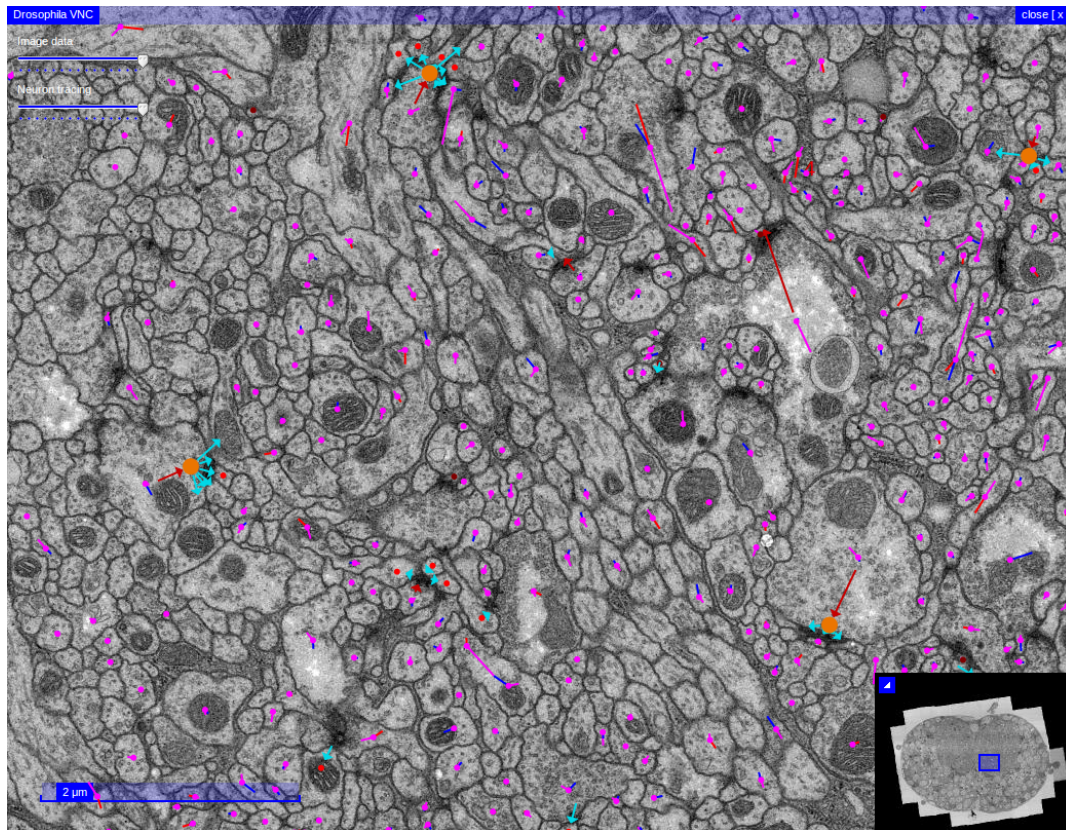
Figure B.7.: **Image Stack With Annotation Overlay** 3D image stack of raw EM data is shown overlayed with reconstructed skeletons (pink nodes) and synapses (orange nodes). Skeletons can be activated in the interactive overlay to create new nodes or move or delete existing nodes. Presynaptic and postsynaptic links to connector nodes (big orange nodes) define connectivity between skeleton nodes at synaptic locations. The widget can show arbitrary number of layers stacked on top of each other with controls for transparency (top-left sliders). The thumbnail overview of the whole 2D plane of the current field-of-view allows quick navigation to distant locations by clicking (bottom-right). Scale bars are displayed in physical units (bottom-left).

type	tags	c	x	y	z	s	r	user	last modified	reviewer
Leaf ▼										
L	not a branch	5	31332.00	19356.00	11520.00	256	-1.0	importer	01-04-2013 15:04	caseysm
L	ends	5	34478.99	20814.51	12510.00	278	-1.0	importer	01-04-2013 15:05	caseysm
L	ends	5	34219.97	20376.21	11115.00	247	-1.0	importer	01-04-2013 15:02	caseysm
L	not a branch	5	40910.85	25627.14	11250.00	250	-1.0	importer	01-04-2013 15:05	caseysm
L	ends	5	42020.73	28669.93	10575.00	235	-1.0	importer	01-04-2013 15:04	caseysm
L	ends	5	41210.84	28543.57	12285.00	273	-1.0	importer	01-04-2013 15:05	caseysm
L	ends	5	39224.75	30257.45	12780.00	284	-1.0	importer	01-04-2013 15:05	caseysm
L	ends	5	39613.86	29502.57	12780.00	284	-1.0	importer	01-04-2013 15:04	caseysm

Figure B.8.: **Skeleton Node Table** Table to display information about skeleton nodes. Rows can be filtered by skeleton node type such as branch, leaf, continuation or root node. Leaf nodes are shown with tags, confidence value, spatial location, section index, radius value, user who created the node, last modification date and users who reviewed the node. Each column can be sorted and rows filtered by specific tags. Double-click on a row navigates to the particular node location in the image stack widget and activates the skeleton node.

connector id	target skeleton ID	x	y	z	s	tags	# nodes in target skeleton	username	target treenode ID
16379846	14687636	32039.74	19768.02	11970	266		2865	importer	14689236
16379846	18601538	31722.84	20131.39	11970	266		3432	importer	14815172
16379864	14687636	34725.59	21248.69	11745	261		2865	importer	14688888
16379864	17242996	34341.43	21444.12	11745	261		4471	importer	15777310
16379864	16137292	34484	21064	11610	258		1018	importer	16137428
16379887	14687636	34274.44	20875.81	11385	253		2865	importer	14688698
16379887	17242996	33880.3	20714.02	11430	254		4471	importer	15777156
16379887	16137292	33980	20620	11340	252		1018	importer	16137368
16379910	14687636	34037.09	20997.46	12195	271		2865	importer	14689036
16379910	18601538	34092	20896	12330	274		3432	importer	14815412
16379910	18601538	34524	21212	12420	276		3432	importer	14815436
16379910	17242996	34511.94	21356.68	12330	274		4471	importer	15777552
16379910	16137292	33988	20752	12285	273		1018	importer	16137294
16379943	14687636	37352	21684	12690	282		2865	importer	14688842
16379943	14687636	38080	21224	12735	283		2865	importer	14688934

Figure B.9.: **Connector Node Table** All connector nodes that are associated with a skeleton (either outgoing or incoming synapses) are displayed. For outgoing connectors, each individual connector can span multiple rows to individual target skeletons with their spatial location, section index, tags, number of nodes in the target skeleton, username of creator and the target skeleton node ID shown.

Figure B.10.: **Review System** Multiple users can review a particular skeleton. The skeleton is split into non-overlapping segments of cable that are reviewed for errors individually. Iteration over the skeleton nodes while reviewing with keyshort short cuts (Q,W: Forward, Backwards; Shift-Q,W: Next unreviewed node on segment; E: next segment). See the supplemental text in Chapter 2 for details.

Figure B.11.: **Notification System** When skeleton nodes are tagged by users, this table notifies the original node creator about the event and the modification can be approved or rejected.

Connectivity 1									close [x]

From | Active skeleton ▼ | All partners ▼ | Append | Clear | Refresh | Sync to: | None ▼ | Open plot

Tables side by side ☑

Selected	Neuron	Upstream Threshold	Downstream Threshold
1. ☑	A02a_a3l Looper-1	1 ▼	1 ▼

☐ Hide single node partners   Reviewed by | All (union) ▼ |

Upstream neuron	syn count	reviewed	node count	select	Downstream neuron	syn count	reviewed	node count	select
▼ ALL (21neurons)	62	82%	30595	☐	▼ ALL (136neurons)	219	82%	50707	☐
dbd_a3l	22	100%	985	☐	A08i2_a3l	19	100%	3901	☐
? class I md	13	100%	771	☐	A02b_a3r Looper-2	12	100%	4986	☐
Bowl-?2_a2r	4	100%	921	☐	A07f2_a3l	12	100%	3912	☐
crescent-1L	2	100%	3709	☐	A27j_a3l	10	99%	3693	☐
ddaD class I md	2	100%	1047	☐	A06?2_a3l #e	8	100%	3408	☐
Drunken-7_a3l	2	97%	2540	☐	IS premotor #cL	7	100%	3162	☐
A02b_a3l Looper-2	2	100%	3944	☐	A07b_a3l	6	100%	3201	☐
A31k_a3r	2	100%	3165	☐	Looper-4_a3r	5	100%	3109	☐
neuron TODO #101200	1	0%	2607	☐	saddle-4L-post	4	0%	2393	☐
A12f_a3r	1	99%	4471	☐	A02b_a3l Looper-2	4	100%	3944	☐

Figure B.12.: **Connectivity Widget** Given a single skeleton or a set of skeletons, all upstream and downstream skeletons are displayed with their respective synapse counts, review status, and number of nodes. Sets of skeletons can be selected to populate other widgets.
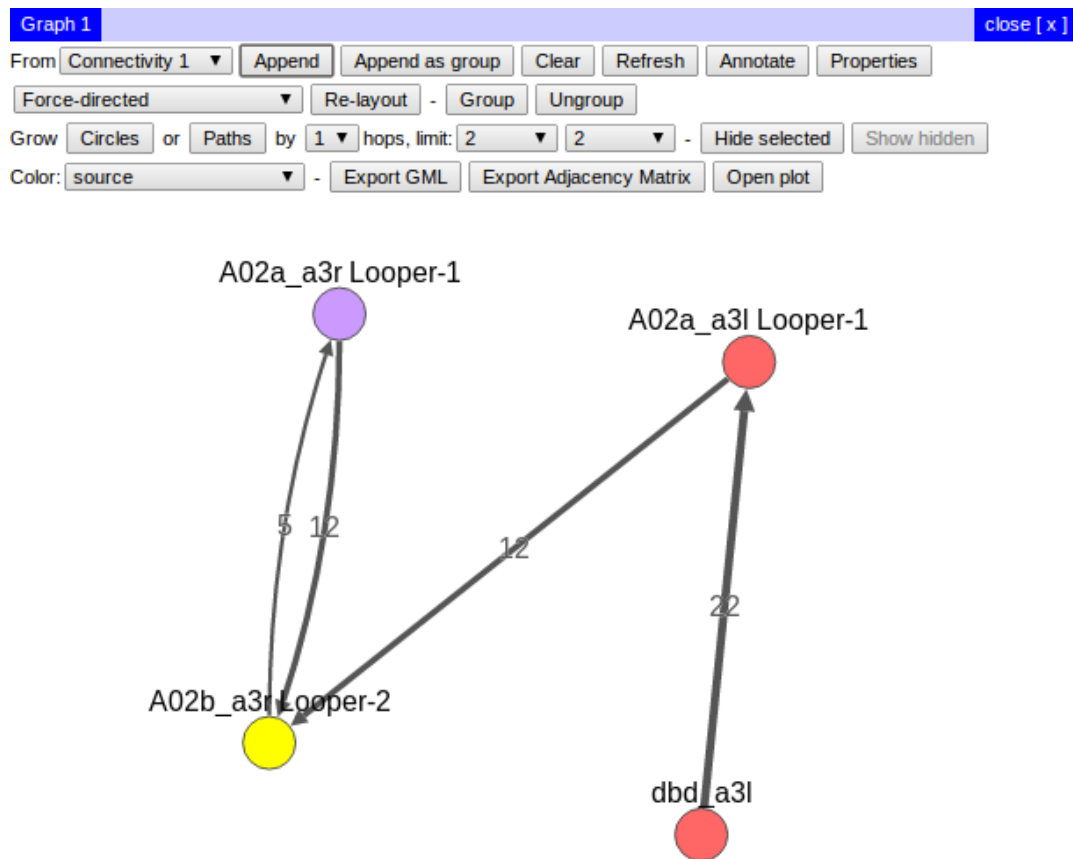
Figure B.13.: **Graph Widget** The widget shows the connectivity between a selected set of skeletons as an interactive network graph. Individual nodes can be moved and their their color changed, or they can be made invisible grouped into supernodes to facilitate understanding. The graph can be grown based on a set of selected nodes via synaptic adjacency. A variety of graph layouting algorithms can be run used to position visible nodes.
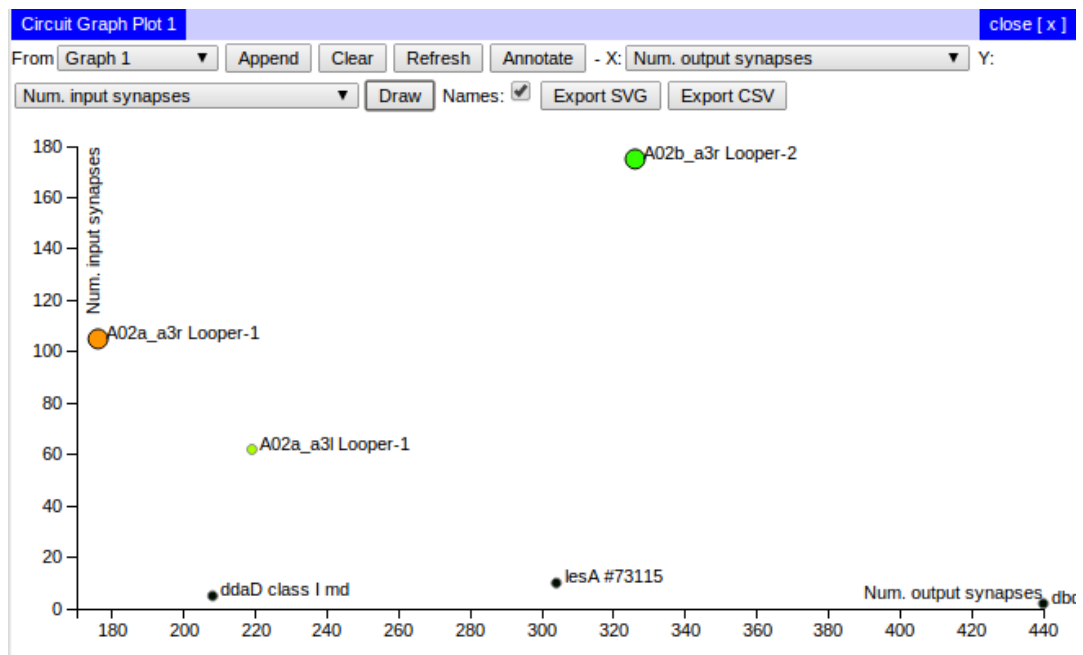
Figure B.14.: **Circuit Graph Plot** Scatter plot of skeleton attributes. The attributes for the x and y axes are interactively specified. Selectable attributes include for instance the number of input or output synapses, difference of input and output synaptic counts or arbor cable length. Network-based attributes include for instance signal flow, eigenvalues of the graph Laplacian, the graph partition or betweenness centrality. The generated plots can be exported to SVG or to a CSV file.
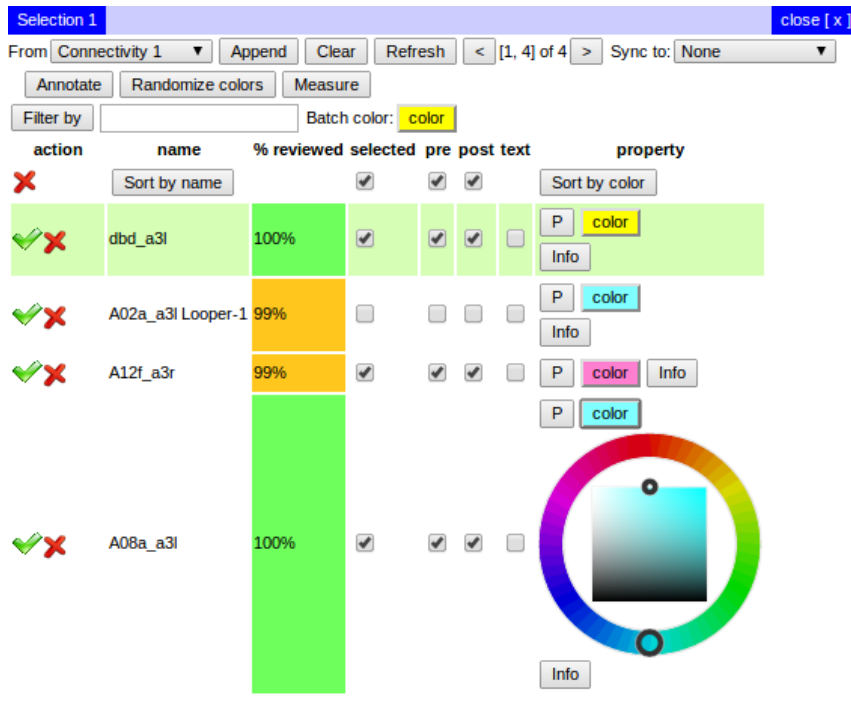
Figure B.15.: **Selection Table** Most of the widgets have a group of skeletons associated. This skeleton selection can be used to populate the content of other widgets. This mechanism allows to flexibly combine sets of skeletons and analyze and visualize them appropriately. In the selection table, skeletons can be colored or annotated. The table can be synchronized, for instance, to the 3D viewer to define which skeletons to show.

Figure B.16.: **Skeleton Measurements Table** The measurements table computes attributes of single skeletons such as cable length, smoothed cable length, the number of synaptic inputs and outputs, or skeleton node counts.

213
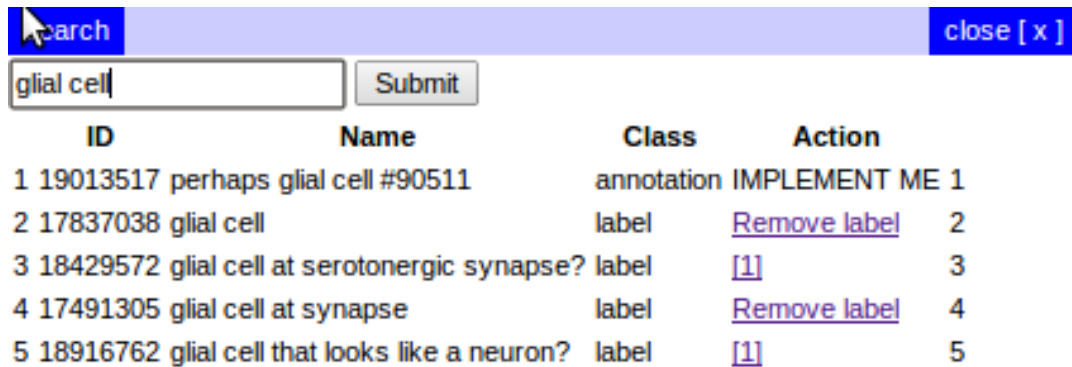
Figure B.17.: **Search Window** In this widget, the current project is queried for any string, such as skeleton node tags or neuron names or annotations. Resulting skeleton tags can be used to directly move to the skeleton node with the associated tag. This facilitates bookmarking of individual locations in large volumes for fast revisiting.
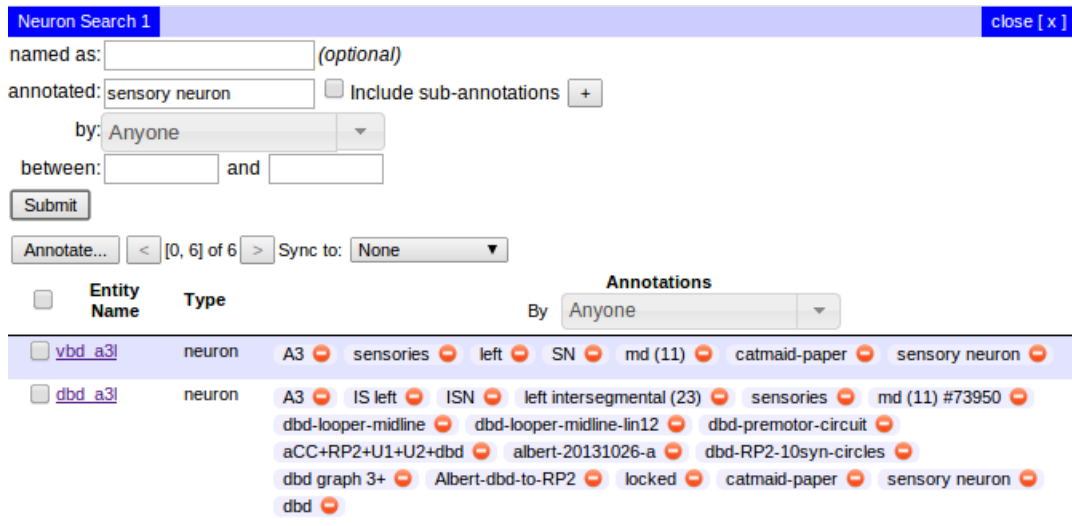
Figure B.18.: **Neuron Search** Neurons can be queried by their name, by a combination of their annotations, by creator and within a given time interval of creation. Intersectional search queries using a combination of these attributes can also be performed. The search result shows neurons along with associated annotations in a table.
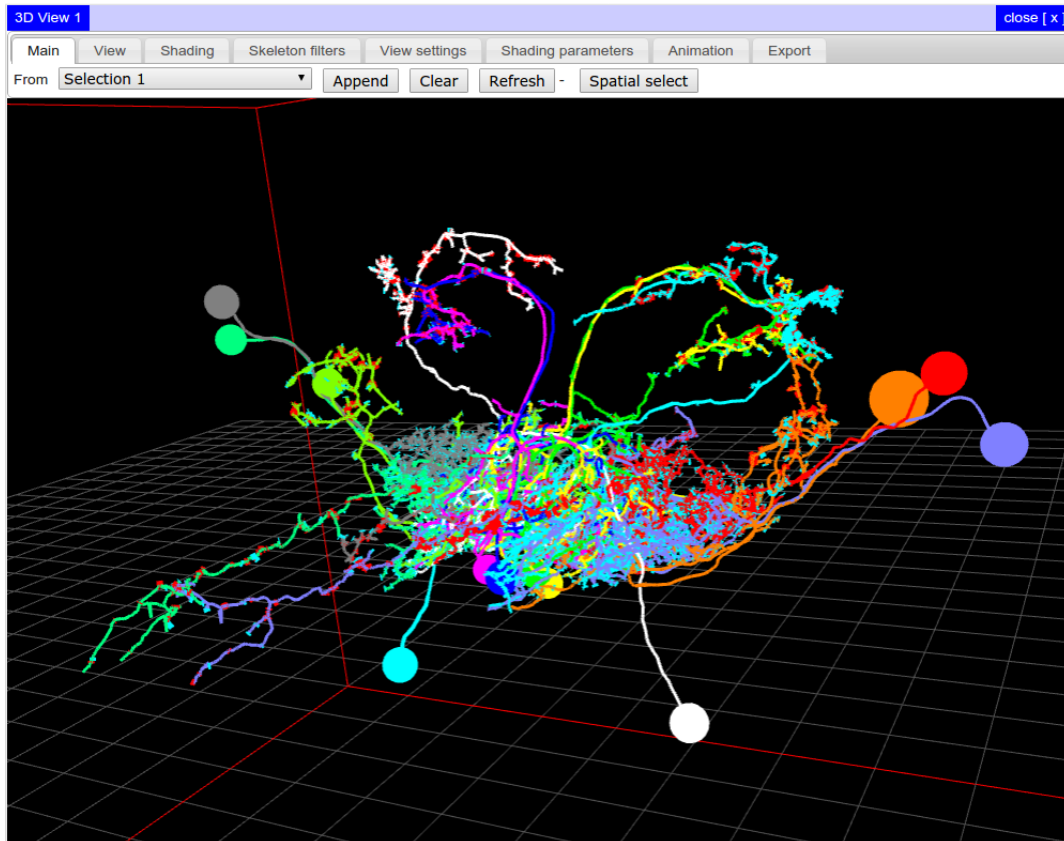
Figure B.19.: **3D Viewer** The skeleton morphology of a set of neurons and their synapses is shown in an interactive 3D window based on WebGL. The skeleton colors and other attributes can be modified individually for each neuron. By selecting individual synapses in the 3D view, one can directly navigate to their location in the 3D image stack window. Various shading and filtering modes, as well as export options including animated videos are available.

Figure B.20.: **Skeleton Analytics Widget** Automated detection of issues ensures quality and consistency of skeletons. For instance, neurons do not normally synapse onto themselves, which is indicated as an autapse issue. See the supplemental text in Chapter 2 for details on other issues.
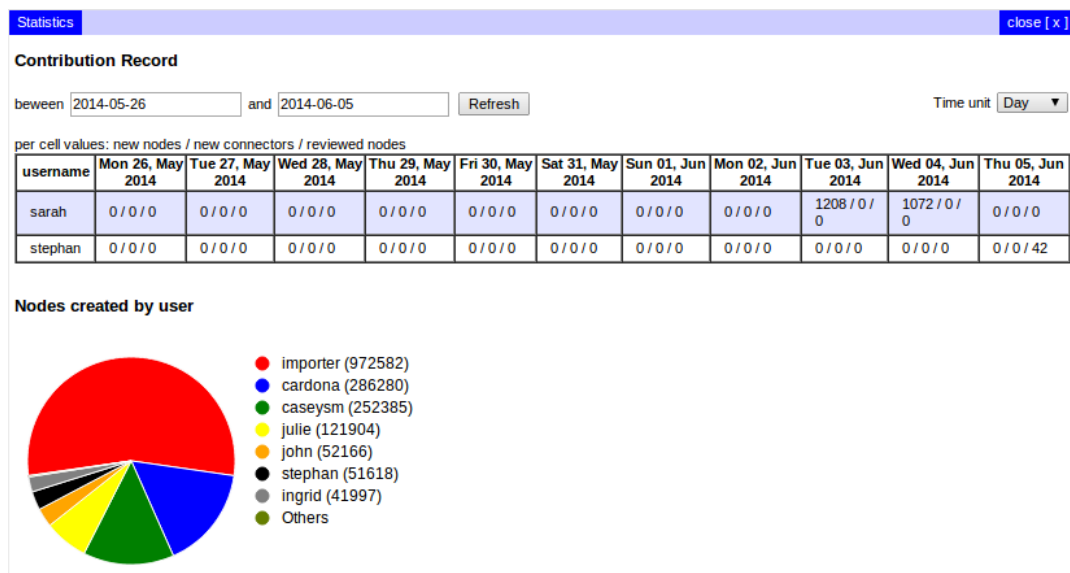
Figure B.21.: **Project Statistics** Contribution of users in terms of newly created nodes and connectors and the number of reviewed nodes are shown. Furthermore, the overall skeleton node contributions of individual users to the project is shown as a pie chart.

Figure B.22.: **Log widget** Every user action (e.g. joining, splitting skeletons, renaming or deleting neurons) is logged and can be queried.
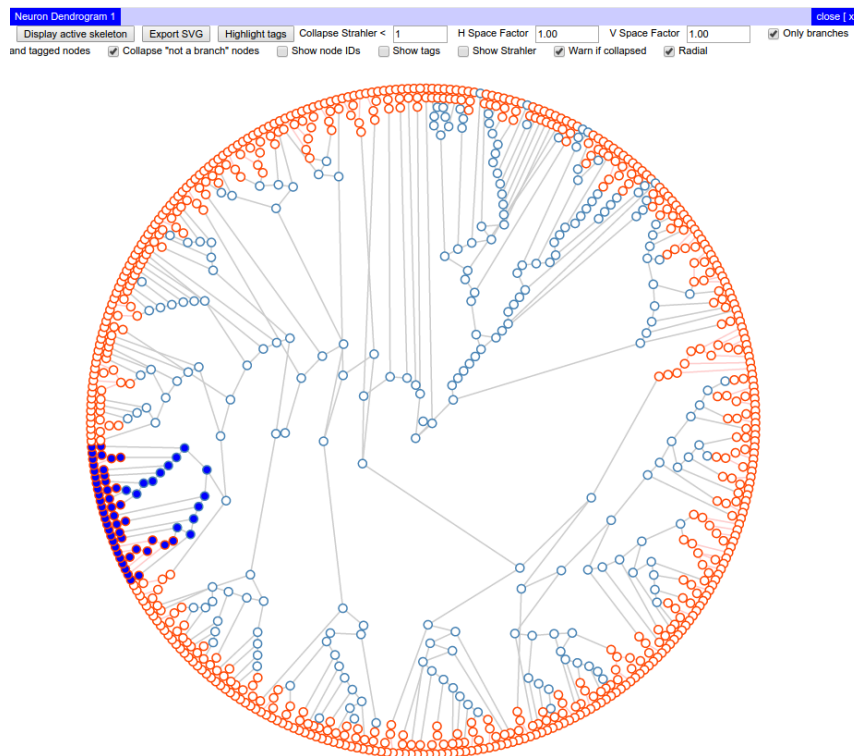
Figure B.23.: **Dendrogram view of left Chair-1 neuron of a first-instar larva** The root node at the center corresponds to the soma location. Red circles denote leaf and branch nodes downstream of the *microtubules ends* tag (which tag should be used for highlighting can choosen). A random node was selected and all the descending nodes downstream of that nodes are shown in blue. (*Widget developed by Tom Kazimiers*)
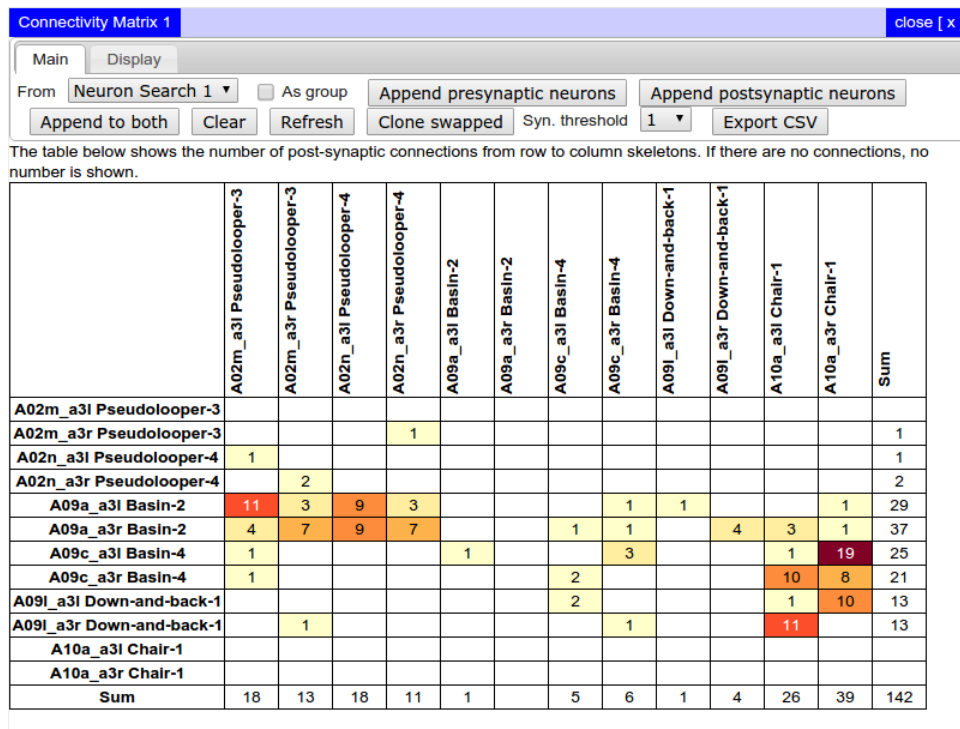
Figure B.24.: **Connectivity Matrix** A connectivity matrix can be generated on demand for a set of presynaptic and postsynaptic neurons. Exporting the matrix to a CSV file allows to further process it outside of CATMAID. (*Widget developed by Tom Kazimiers*)
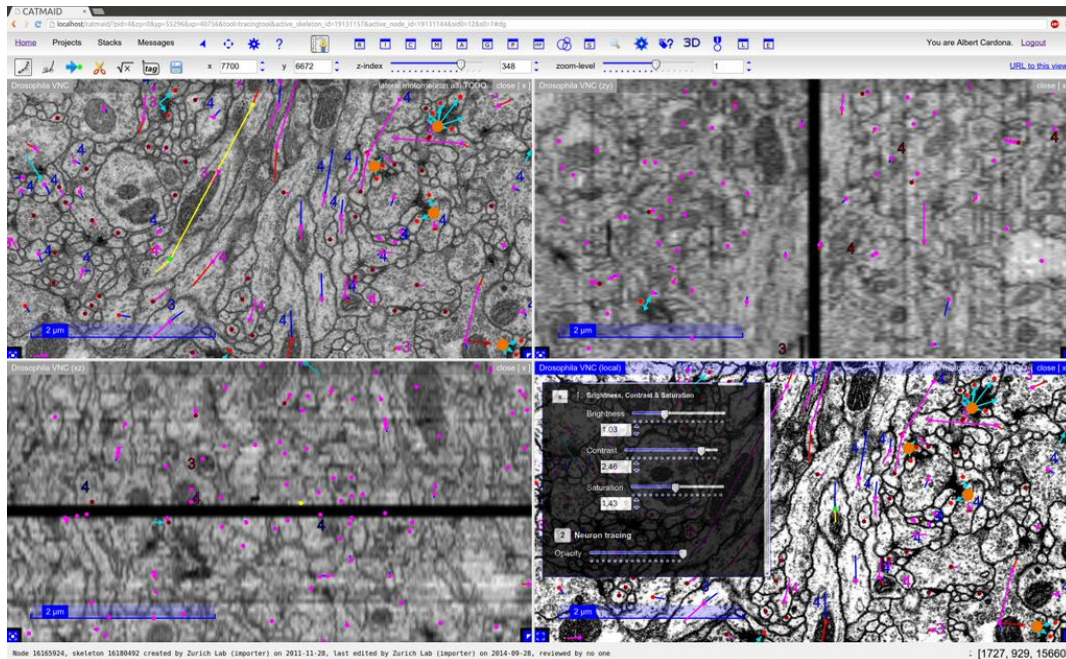
Figure B.25.: **Orthogonal views of EM image data, skeletons and image filters**. A serial-section TEM dataset with anisotropic voxel resolution (4x4x50nm) is shown in CATMAID with 3 orthogonal, synchronized views and overlayed skeletons. Orthoviews of EM data with more isotropic voxel size, e.g. from FIBSEM or SBEM, allow to trace neurons and see synapses from optimal cut planes. The bottom right widget shows the application of image filters in the browser, for instance for on-the-fly contrast and brightness correction, implemented as shaders in WebGL. (Developed by Tom Kazimiers and Andrew Champion. Image source: twitter.com/albertcardona.)

C

C.1 THIRD-INSTAR LOCAL INTERNEURONS OF THE NOCICEPTIVE NETWORK

Neuron	Raw cable (nm)	Smooth cable (nm)	# inputs	# outputs	# presyn	# nodes	# branches	# ends
A02m_a3l Pseudolooper-3	738439	678176	305	247	67	8227	202	210
A02m_a3r Pseudolooper-3	731110	664090	298	311	83	7756	222	228
A02n_a3l Pseudolooper-4	803064	734151	346	280	76	9087	230	237
A02n_a3r Pseudolooper-4	702934	640755	287	285	82	7282	216	223
A09a_a3l Basin-2	1855481	1679959	1601	1037	198	19920	1150	1219
A09a_a3r Basin-2	1863558	1674293	1585	1161	237	19438	1019	1071
A09c_a3l Basin-4	1455267	1321245	1077	744	170	15994	601	619
A09c_a3r Basin-4	1601876	1450079	1101	731	177	17404	650	671
A09l_a3l Down-and-back-1	1271136	1157308	833	739	176	14115	447	458
A09l_a3r Down-and-back-1	1197451	1089692	788	897	184	13159	473	487
A10a_a3l Chair-1	1178205	1061426	555	827	213	13116	384	404
A10a_a3r Chair-1	1204255	1089528	618	815	176	13179	407	415

Table C.1.: **Anatomical properties of local interneurons** Properties of the six local interneuron types downstream of primary nociceptors in the abdominal segment A3 in the third-instar datasets L3.
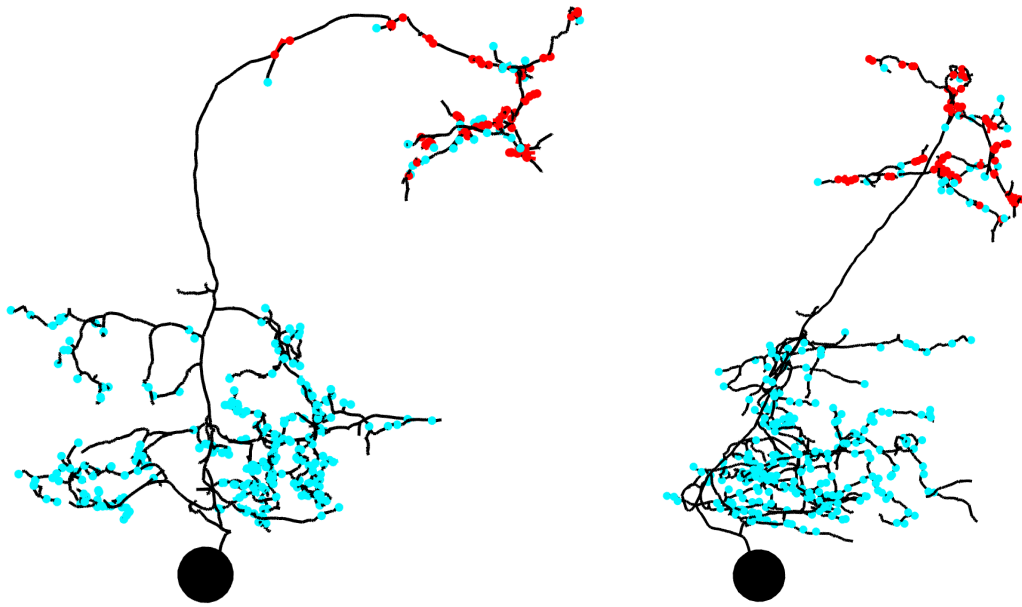
Figure C.1.: **Right Pseudolooper-4 (A02n) in abdominal segment A3**. Posterior view (left) and lateral view (right). Presynaptic sites (red) and postsynaptic sites (bright blue) are shown with spheres.
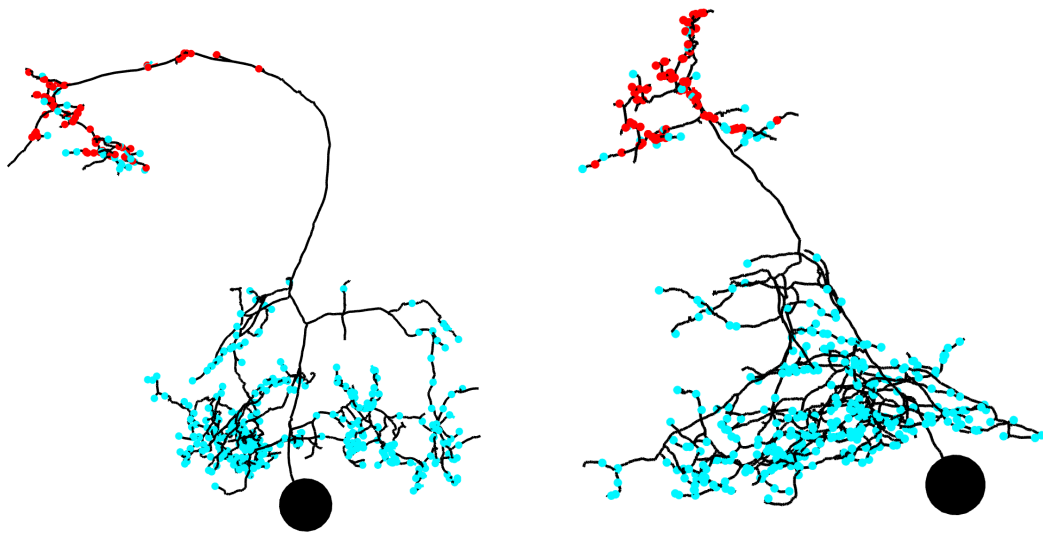
Figure C.2.: **Left Pseudolooper-4 (A02n) in abdominal segment A3**.

Figure C.3.: **Left Pseudolooper-3 (A02m) in abdominal segment A3**.
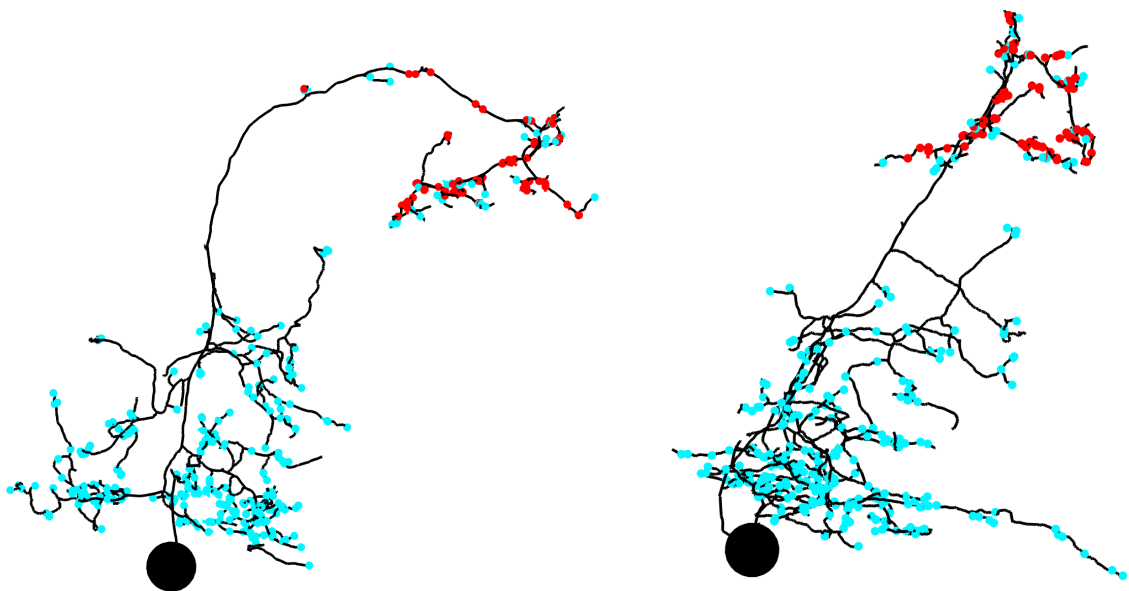
Figure C.4.: **Right Pseudolooper-3 (A02m) in abdominal segment A3**.
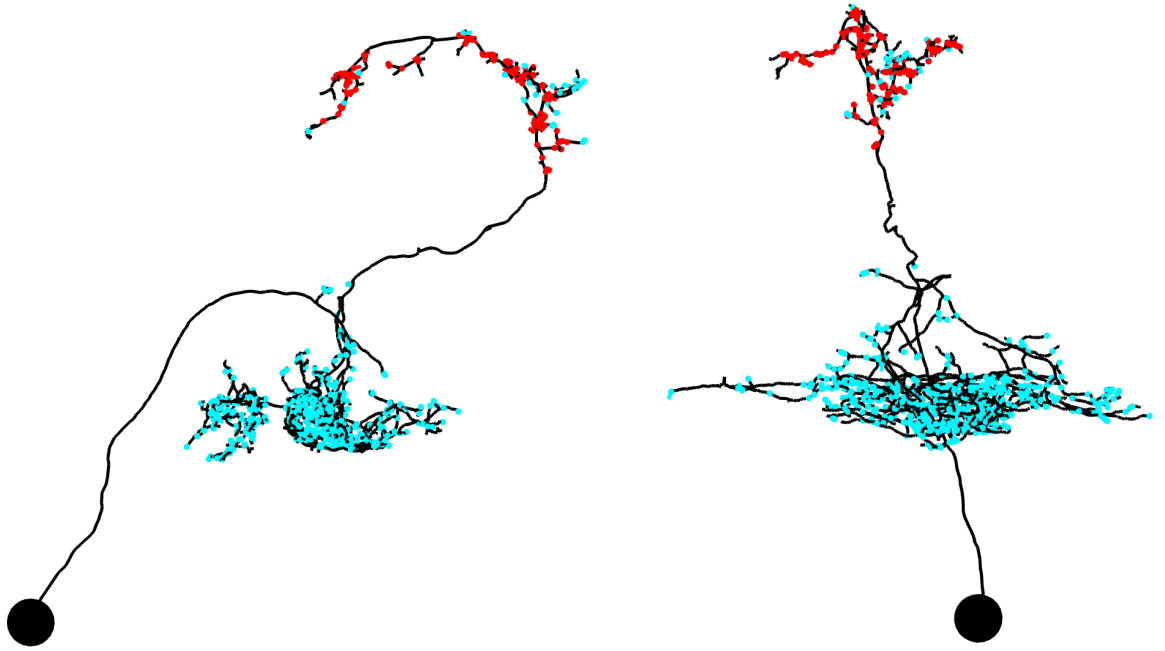
Figure C.5.: **Right Chair-1 (A10a) in abdominal segment A3**.
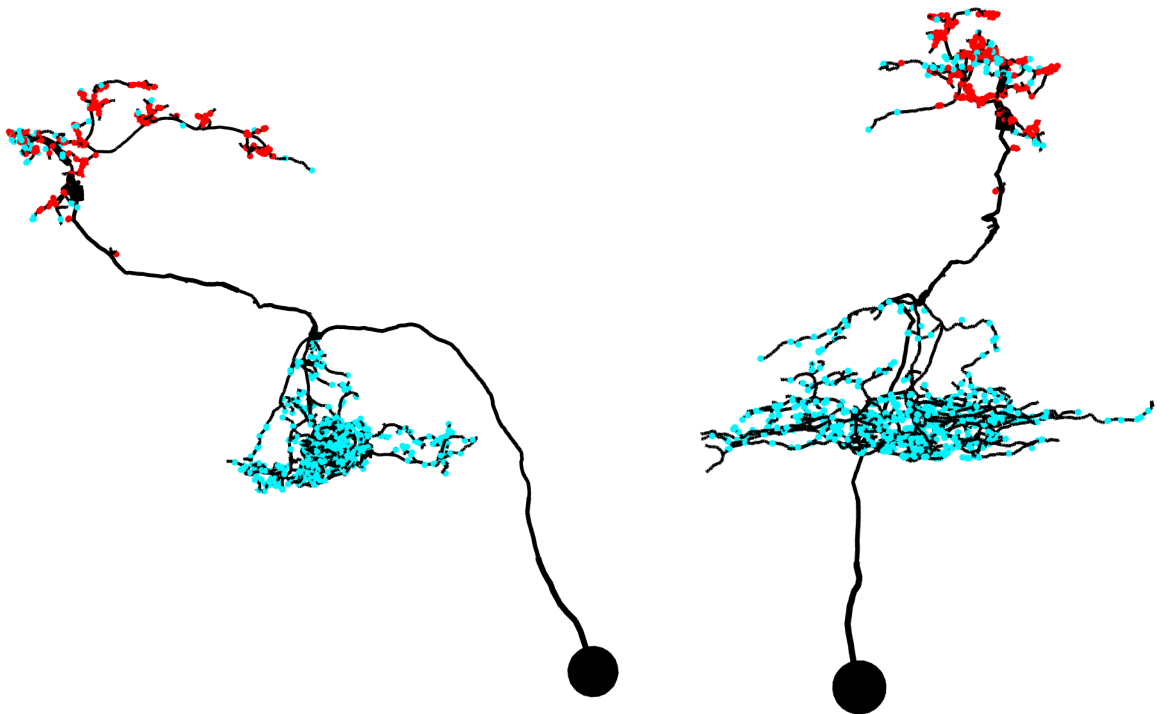
Figure C.6.: **Left Chair-1 (A10a) in abdominal segment A3**.
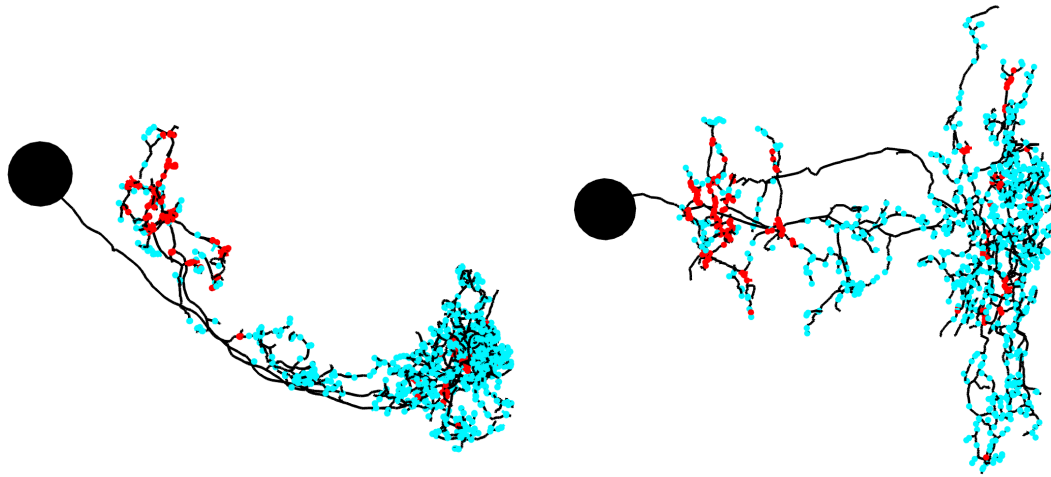
Figure C.7.: **Left Down-and-back-1 (A09l) in abdominal segment A3**.
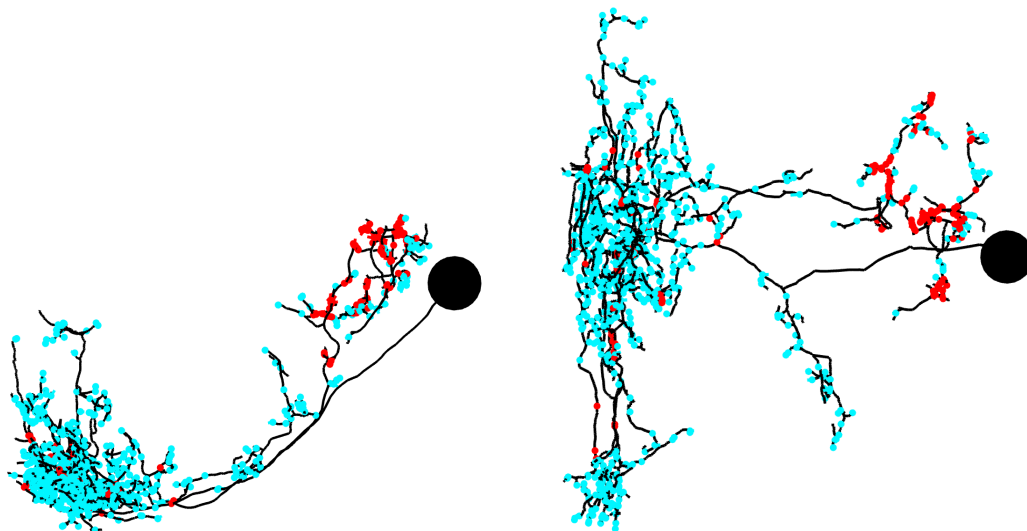
Figure C.8.: **Right Down-and-back-1 (A09l) in abdominal segment A3**.
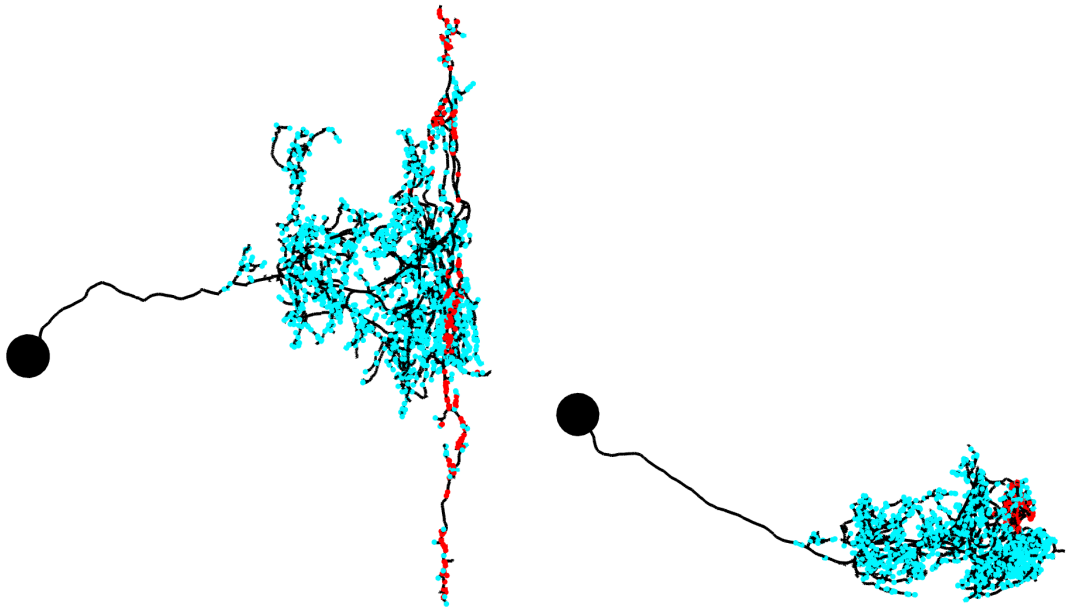
Figure C.9.: **Left Basin-2 (A09a) in abdominal segment A3**.
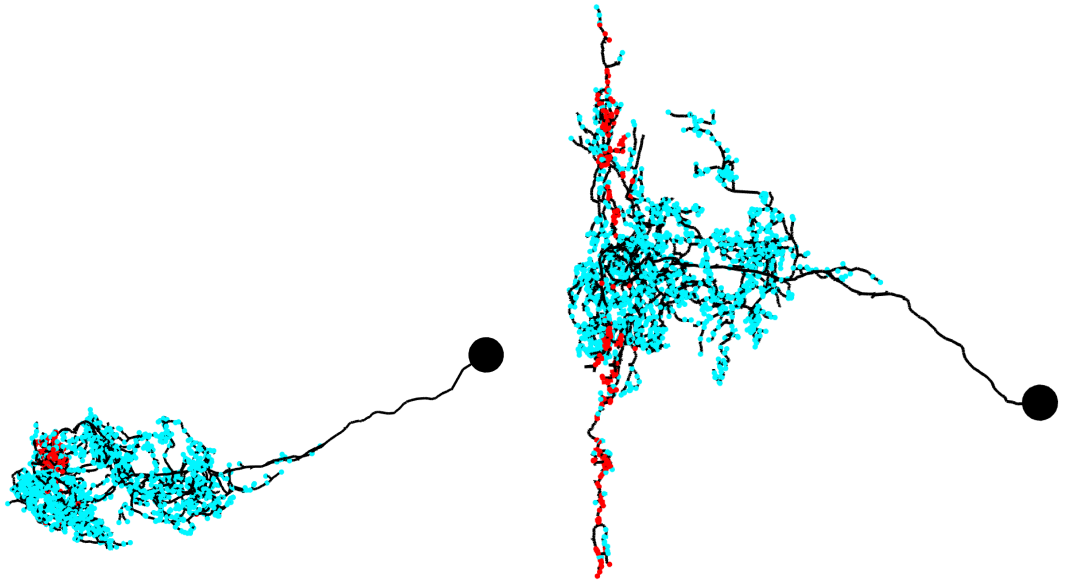
Figure C.10.: **Right Basin-2 (A09a) in abdominal segment A3**.
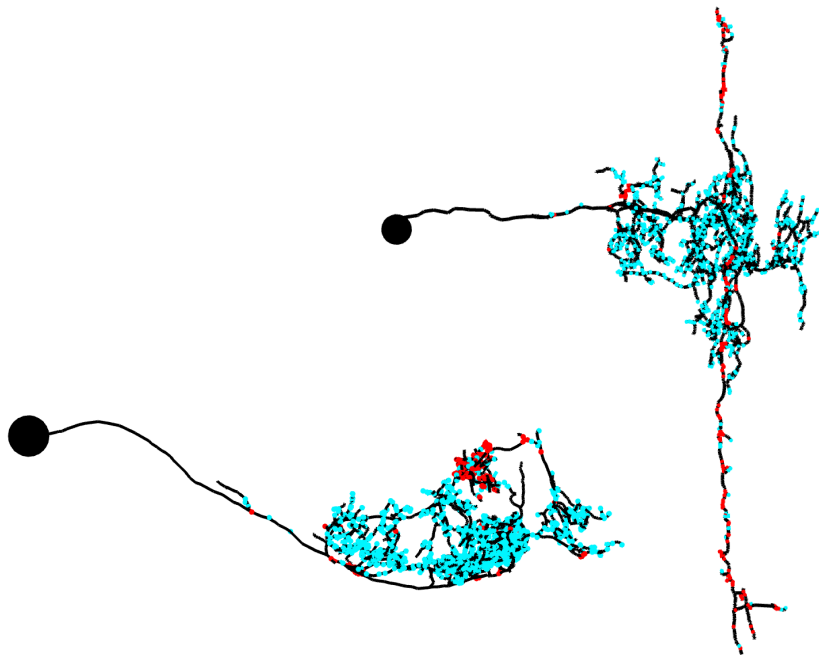
Figure C.11.: **Left Basin-4 (A09c) in abdominal segment A3**.
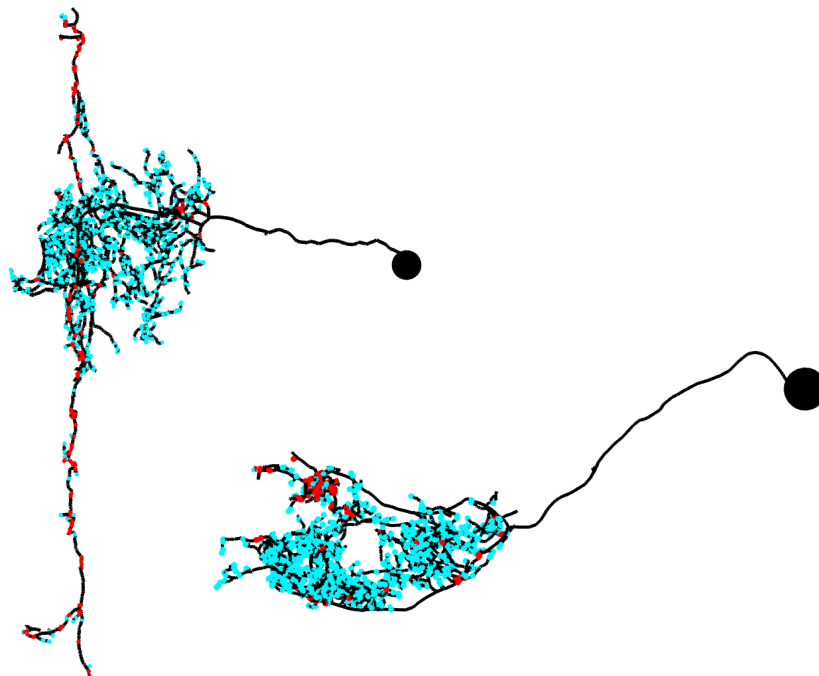
Figure C.12.: **Right Basin-4 (A09c) in abdominal segment A3**.

C.2 PROPRIO-MOTOR INTERNEURON CATALOG
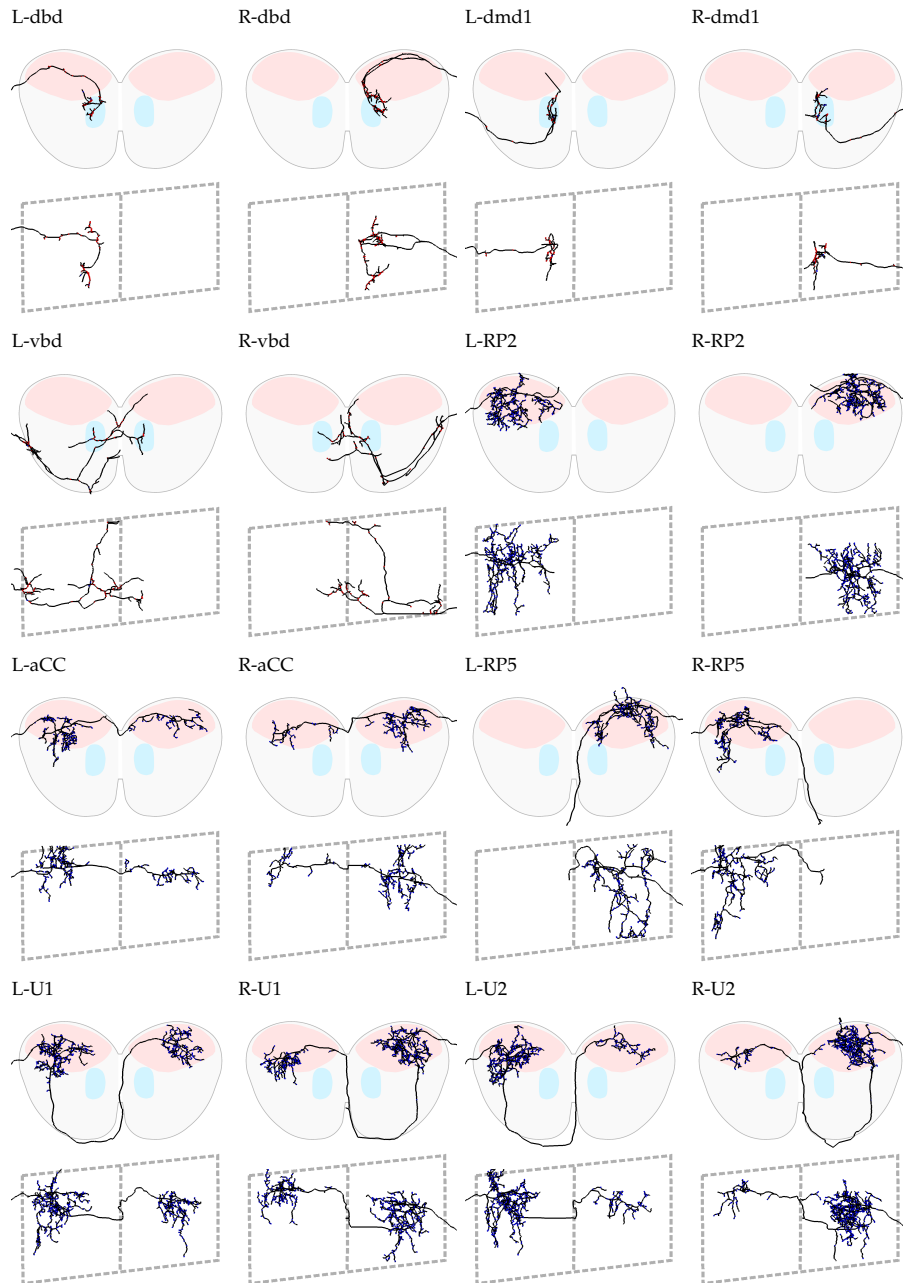
Figure C.13.: **Morphology of identified proprio-motor circuit neurons** A gallery of neurons mapped and identified in Chapter 2. Red points denote presynaptic sites, blue points denote postsynaptic sites. The top view is a cross-sectional, posterior view of the ventral nerve cord. For reference, the red outline denotes the motor area, the blue outline the proprioceptive area. The bottom view is a dorsal view (top is anterior, bottom is posterior). All skeletons were mapped in dataset L1b. Presynaptic sites are in red, postsynaptic sites are in dark blue.

L-A00f R-A00f L-A02a R-A02a

L-A02b R-A02b L-A02d R-A02d

L-A02l R-A02l L-A02m R-A02m

L-A02n R-A02n L-A03a3 R-A03a3
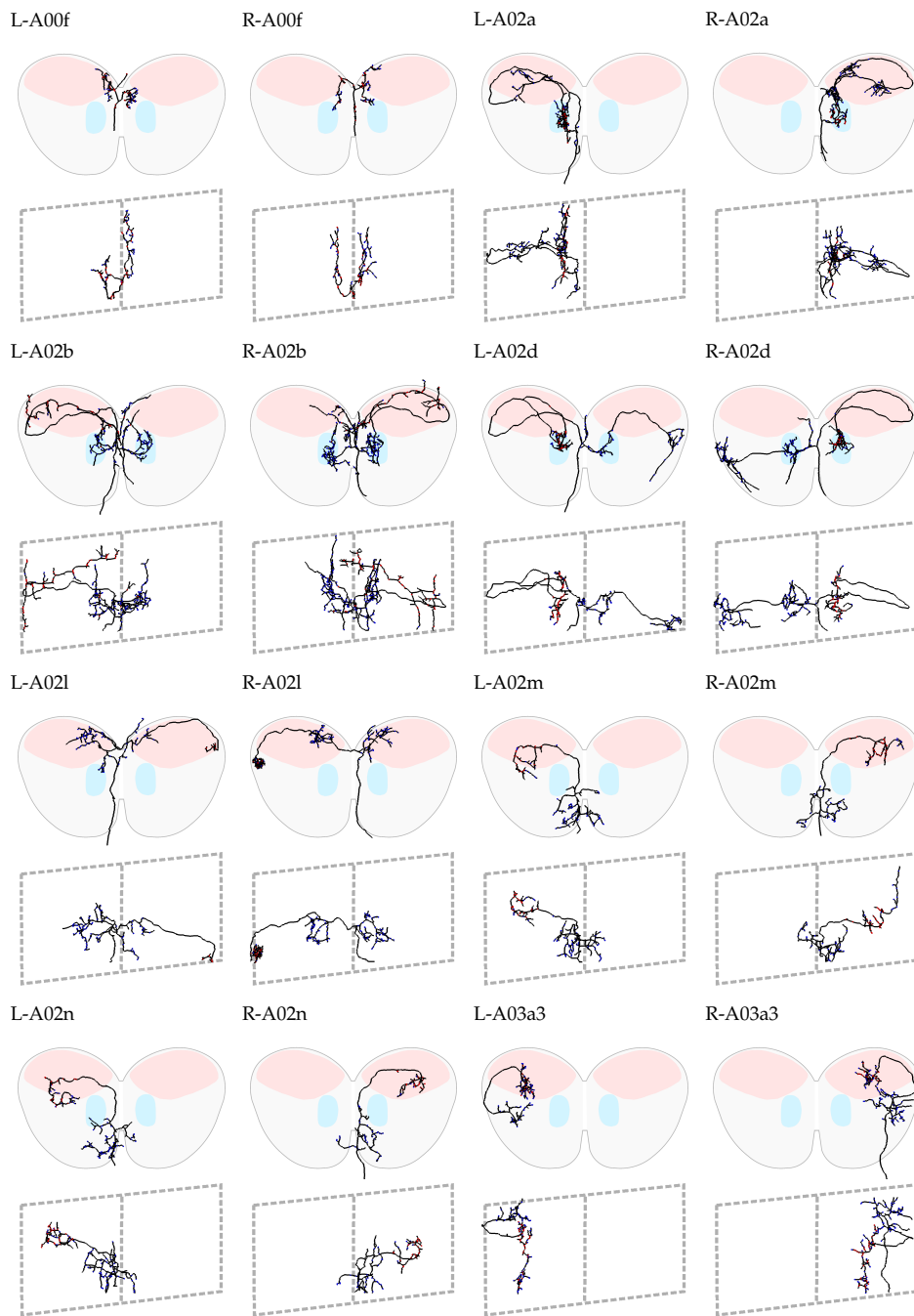
Figure C.14.: **Morphology of identified proprio-motor circuit neurons** (continued)

L-A03x1   R-A03x1   L-A03x2   R-A03x2

L-A05e   R-A05e   L-A06a   R-A06a

L-A07b   R-A07b   L-A07f2   R-A07f2

L-A08a   R-A08a   L-A08e   R-A08e

Figure C.15.: **Morphology of identified proprio-motor circuit neurons** (continued)
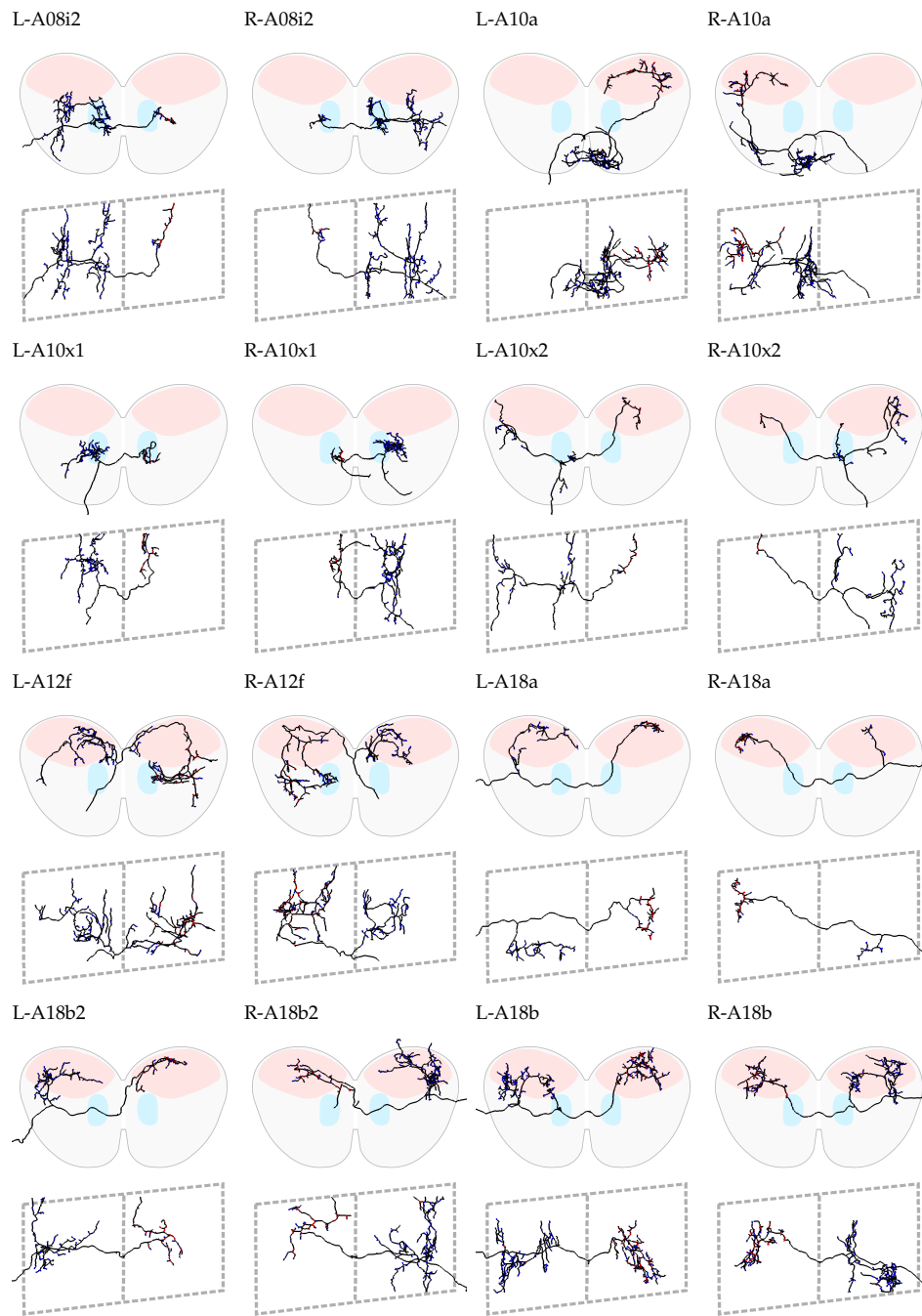
Figure C.16.: **Morphology of identified proprio-motor circuit neurons** (continued)

Figure C.17.: **Morphology of identified proprio-motor circuit neurons** (continued)

D

FURTHER OBSERVATIONS

D.1 SEPTATE JUNCTIONS IN THE DROSOPHILA LARVA CNS

Beyond the well-known chemical synapses and electrical gap junction, other types of intercellular junctions have been known to exist from ultrastructural studies. In particular, a class of junctions called septate junctions has been found in both invertebrate and vertebrate cells (Banerjee et al. 2006). One of their defining characteristic is a ladder-like arrangement along the contact surface between cells. It is believed that their main function is to act as para cellular barriers to restrict the flow of extracellular molecules and to compartmentalize tissue. In *Drosophila* epithelial cells, they are implicated in cell growth and cell shape (Lewin et al. 2006).

In our circuit mapping efforts, we discovered putative septate-junctions in the central nervous system of *Drosophila* larva at both early and late stage, which have not been described previously. The junctions were observed extensively between the primary axonal terminals of sensory neurons, and also between interneuronal processes. Some of these observed junctions presented a ladder-like arrangement in the EM micrographs (Figure D.1). To confirm and further investigate these details, additional high-resolution reimaging was performed with TEM at 49,000x magnification with pixel x,y dimensions of 0.22 nm.

Septate junctions have parallel running membranes in close apposition, and have been previously mistaken to be gap junctions (Hall et al. 1983). We performed reimaging at different tilt-angles to better resolve individual parts of these junctions (Figure D.2 - D.5). Measurements at optimal, perpendicular orientation to the junctional complex revealed that they are approximately 25 nm wide. This size range suggests that these junctions are not gap junctions, which are known to be only 2-3 nm wide (Maeda et al. 2009). This structural measurement provides evidence for the putative existence of septate junctions in the central nervous system of *Drosophila melanogaster* larvae.

The functional implications of these septate-like junctions in the CNS are unknown. We could not find any form of systematic barriers that could constrain the flow of extracellular material, but are sample size is very limited. Our observations suggest that the junctions could act as a structural support to bind together neuronal processes of the same type. We did not further reconstruct the neurons constrained by these junctions. Future mapping studies could elaborate whether sensory or interneurons show preferred junctional coupling if they are of the same type or function. We observed, however, that by the junctional coupling of sets of axons, the surface area where synaptic contacts could be potentially formed is reduced.

The powerful genetic toolkit available in *Drosophila* will allow to interfer with the cell-type-specific expression of constituting junction proteins (Banerjee et al. 2006) to study their effect on neural function and behavior.
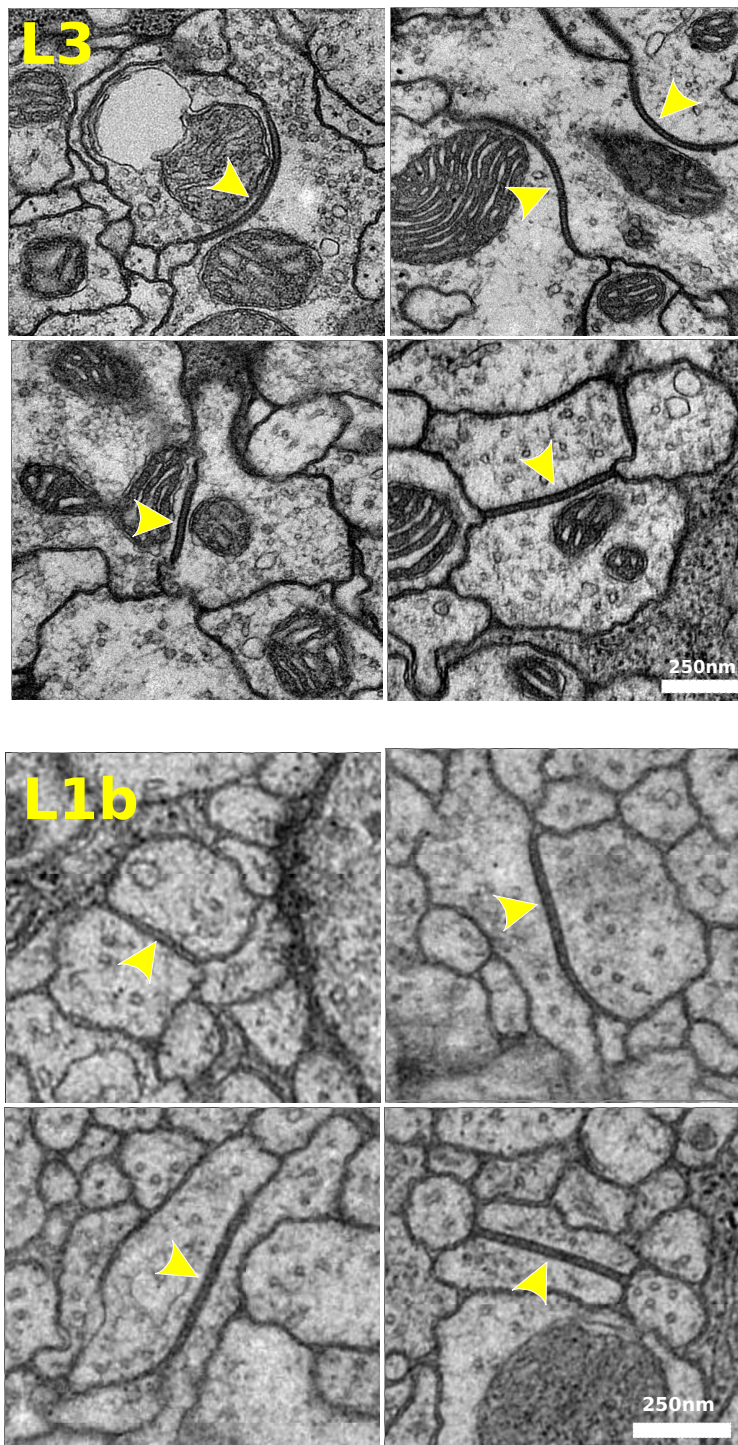
Figure D.1.: **Example of putative septate-junction in *Drosophila* larva CNS**. Examples of putative septate-junctions in third-instar (L3) and first-instar (L1b) datasets. Their characteristic ladder-like arrangement is hardly visible from these EM micrographs. They appear in most of the cases as linear walls between cells with a fixed width. Examples are from sensory nerves, sensory axonal terminal within the CNS and interneurons within the CNS.
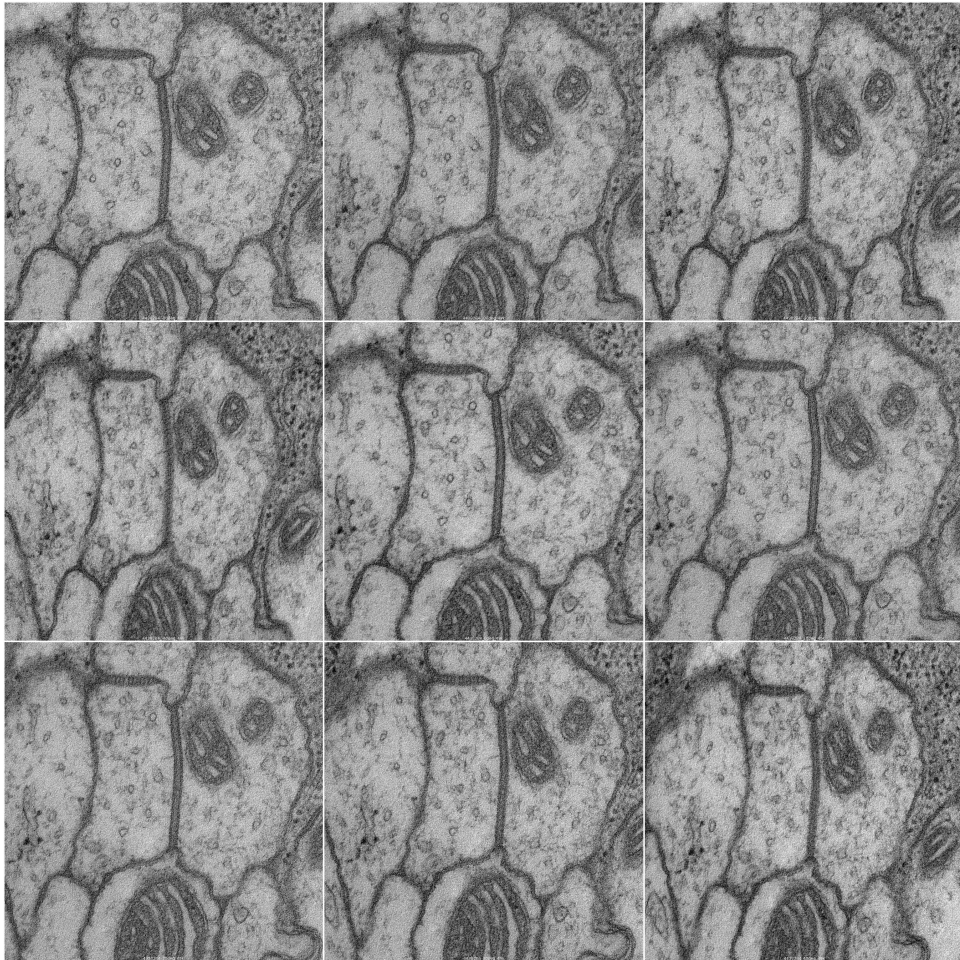
Figure D.2.: **Multi-tilt images of putative third-instar septate junction**. Putative septate junction connecting sensory or motor processes in the nerves of a third-instar larva. Location of the junction is at x=101620.8nm, y=43552.2nm and z=4000nm in the L3 dataset. Tilt angles are -10,-20,-30,-40,0,10,20,30,40 degree from top-left to bottom-right.
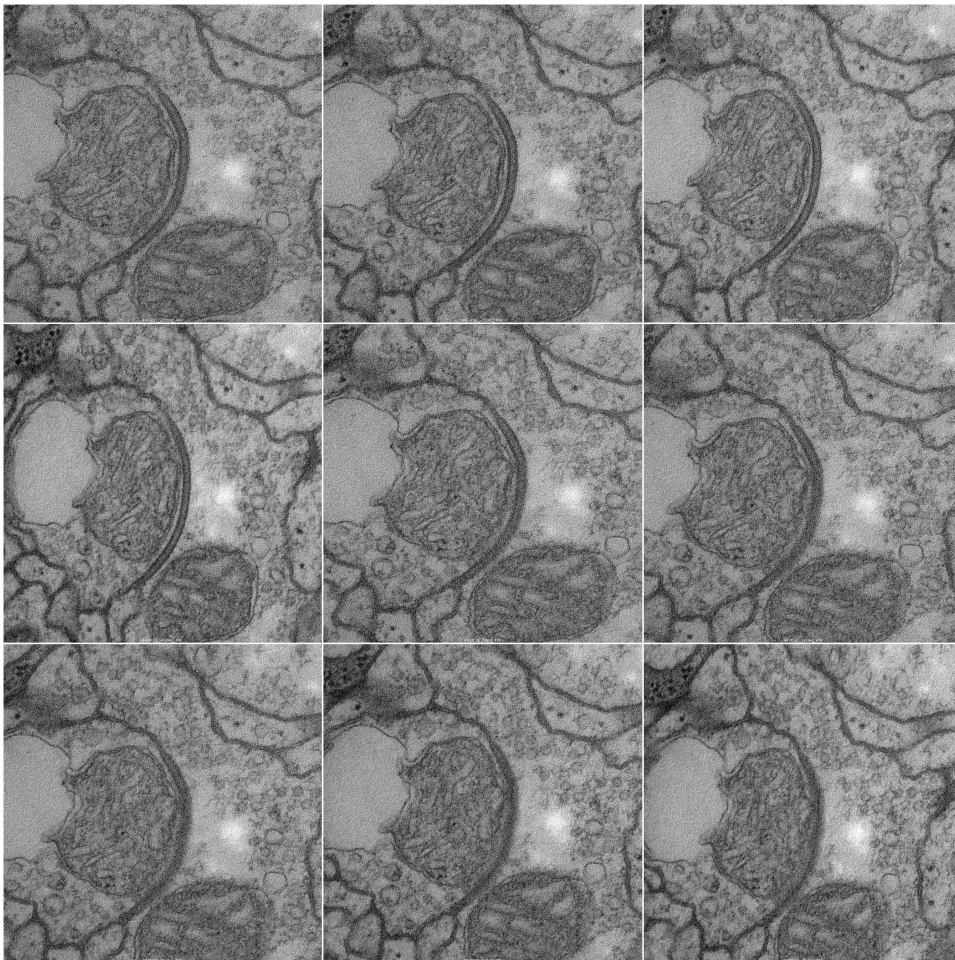
Figure D.3.: **Multi-tilt images of putative third-instar septate junction**. Putative septate junction between axonal arbors in the CNS of a third-instar larva. Location of the junction is at x=77808.5nm, y=35495.5nm and z=61250nm in the L3 dataset. Tilt angles are -10,-20,-30,-40,0,10,20,30,40 degree from top-left to bottom-right.
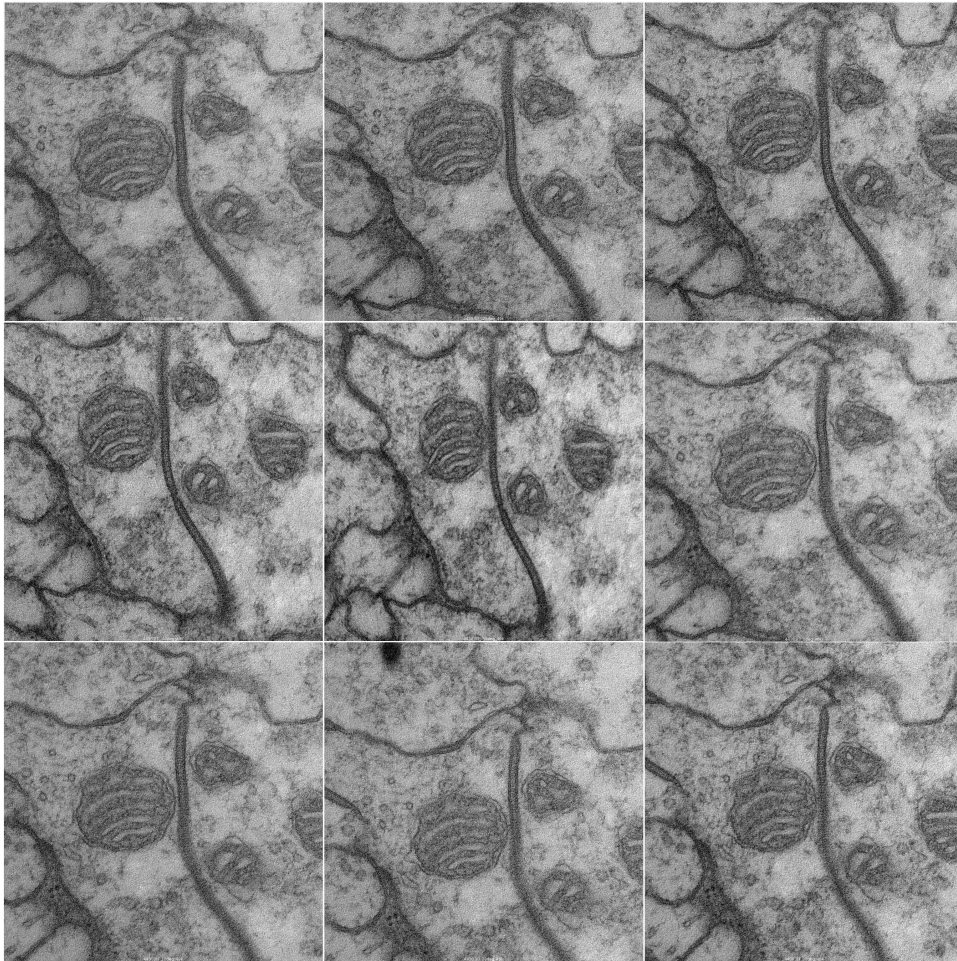
Figure D.4.: **Multi-tilt images of putative third-instar septate junction**. Putative septate junction between axonal arbors in the CNS of a third-instar larva. Location of the junction is at x=85859nm, y=55232nm and z=62150nm in the L3 dataset. Tilt angles are -10,-20,-30,-40,-50,0,10,20,30 degree from top-left to bottom-right.
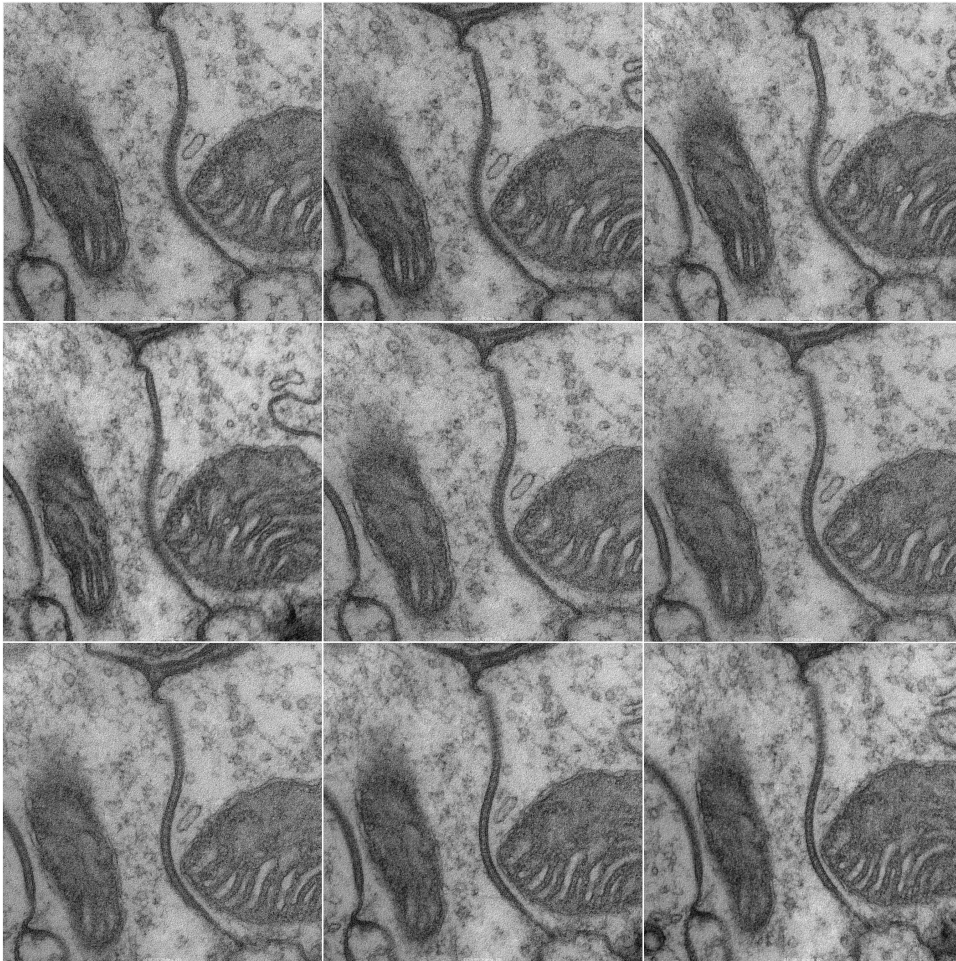
Figure D.5.: **Multi-tilt images of putative third-instar septate junction**. Putative septate junction between axonal arbors in the CNS of a third-instar larva. Location of the junction is at x=87383.025nm, y=51868.78nm and z=62300nm in the L3 dataset. Tilt angles are -10,-20,-30,-40,0,10,20,30,40 degree from top-left to bottom-right.

D.2 DEVELOPMENTAL MISTAKES OF A FEW NEURONS
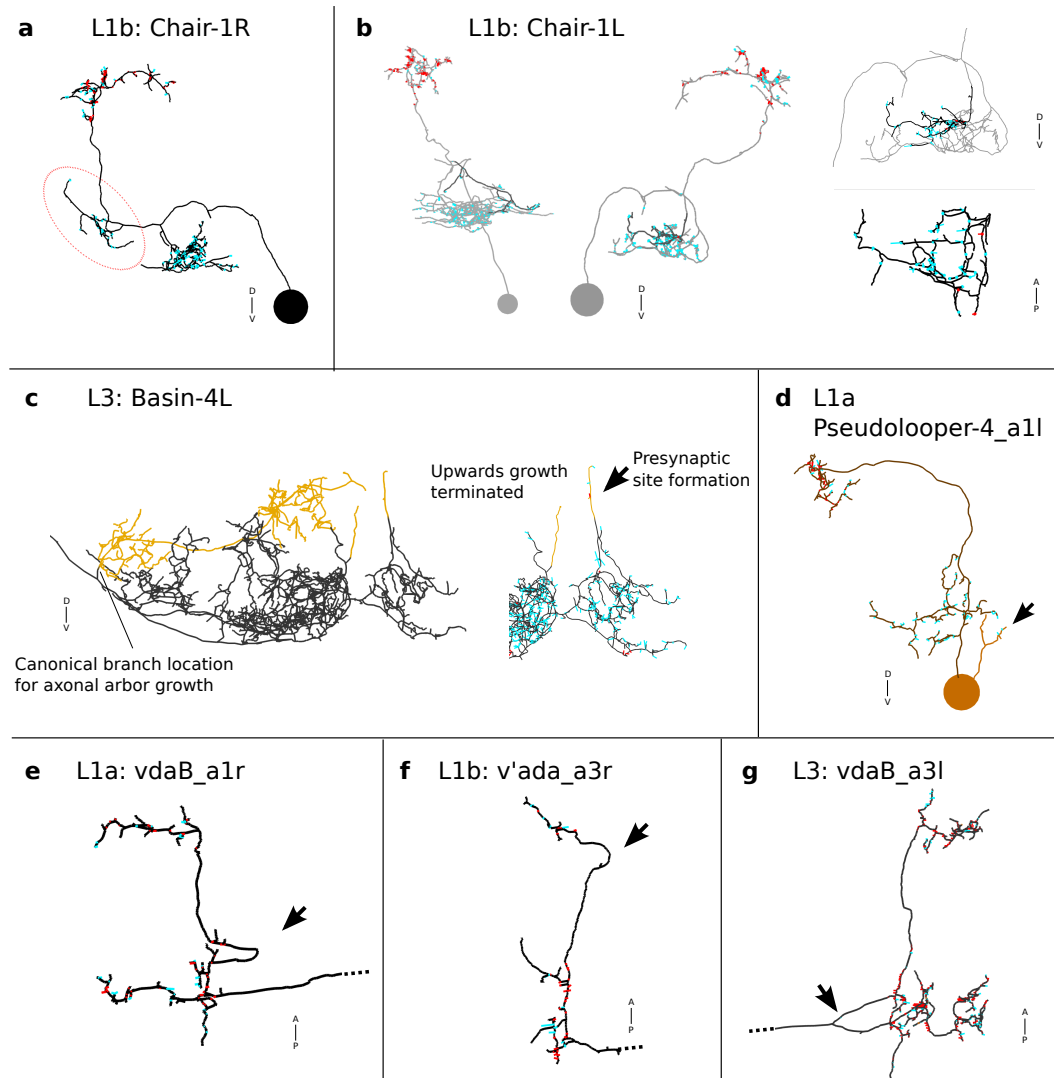
Figure D.6.: **Gallery of non-canonical neuronal morphologies** a) L1b Chair-1: Non-canonical side branch and connectivity with slight deviation from canonical axonal trajectory b) L1b Chair-1: The exceptional occurrence of presynaptic sites on Chair-1 dendrites is explained by a side branch that also contains dense-core vesicles, similar to axonal branches of Chair-1. c) L3: Basin-4L: Two processes growing towards dorsal from a non-canonical part of the dendritic tree. One process terminates early, does not form any synapses and does not contain microtubuli at its final tips. A second upwards process reaches the appropriate height at the same level for correct formation of axonal projection and forms a single presynaptic site. It contains microtubuli up to the presynaptic site. A branch has formed hosting a few postsynaptic input sites. d) Pseudolooper-4 Second upwards branch from cell body with postsynaptic sites e) vdaB axonal detour f) L1b: vada a3r. Deviation from canonical path g) L3 vdaB a3l: early split axon.

CONVOLUTIONAL NEURAL NETWORK CLASSIFICATION RESULTS
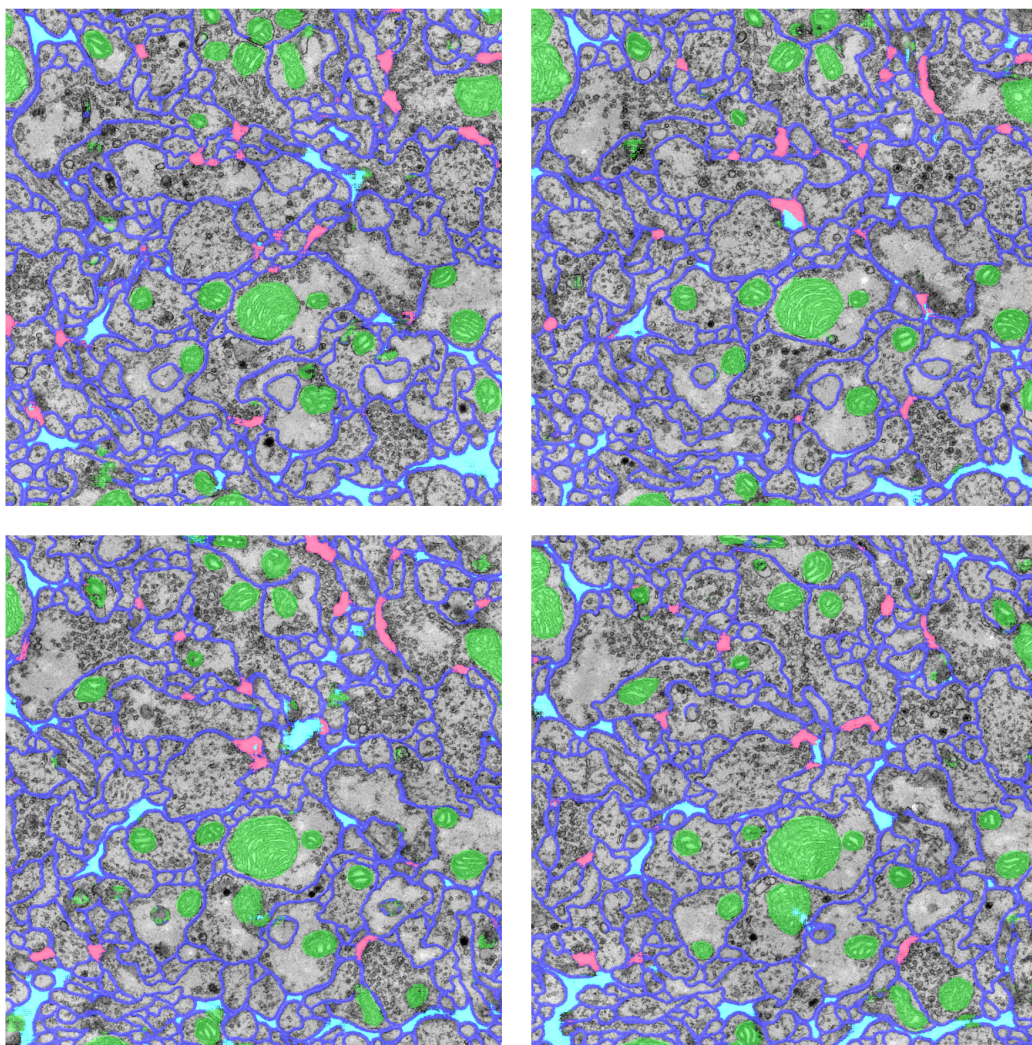
Figure E.1.: **Teststack 2 of the third-instar *Drosophila* larva VNC dataset L3 comprises 20 sections. Sections 1 - 4 are shown.** A single image has XY dimension of 4.7 x 4.7 μm(1024x1024 pixels) and the Z section thickness is 50 nm. The image were classified with an extended version of the convolutional neural network by Ciresan et al. 2012. Membrane in blue, mitochondria in green, synapses in red, glia in bright blue and cell interior is transparent.

Figure E.2.: **Third-instar *Drosophila* larva VNC dataset - Teststack 2 - Sections 5 - 8.**

Figure E.3.: **Third-instar *Drosophila* larva VNC dataset - Teststack 2 - Sections 9 - 12.**

Figure E.4.: **Third-instar *Drosophila* larva VNC dataset - Teststack 2 - Sections 13 - 16.**

Figure E.5.: **Third-instar *Drosophila* larva VNC dataset - Teststack 2 - Sections 17 - 20.**

BIBLIOGRAPHY

Ahrens, Misha B and Florian Engert (2015). "Large-scale imaging in small brains." In: *Current Opinion in Neurobiology* 32, pp. 78–86 (cit. on p. 166).

Albertson, Donna G. and J. N. Thomson (1976). "The Pharynx of Caenorhabditis elegans." In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 275.938, pp. 299–325 (cit. on p. 174).

Amat, Fernando, William Lemon, Daniel P. Mossing, Katie McDole, Yinan Wan, Kristin Branson, Eugene W. Myers, and Philipp J. Keller (2014). "Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data." In: *Nature Methods* 11.9, pp. 951–958 (cit. on p. 166).

Araque, Alfonso and Marta Navarrete (2010). "Glial cells in neuronal network function." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1551, pp. 2375–2381 (cit. on p. 163).

Ascoli, Giorgio A., Lidia Alonso-Nanclares, Stewart A. Anderson, German Barrionuevo, Ruth Benavides-Piccione, Andreas Burkhalter, György Buzsáki, Bruno Cauli, Javier DeFelipe, Alfonso Fairén, Dirk Feldmeyer, Gord Fishell, Yves Fregnac, Tamas F. Freund, Daniel Gardner, Esther P. Gardner, Jesse H. Goldberg, Moritz Helmstaedter, Shaul Hestrin, Fuyuki Karube, Zoltán F. Kisvárday, Bertrand Lambolez, David A. Lewis, Oscar Marin, Henry Markram, Alberto Muñoz, Adam Packer, Carl C. H. Petersen, Kathleen S. Rockland, Jean Rossier, Bernardo Rudy, Peter Somogyi, Jochen F. Staiger, Gabor Tamas, Alex M. Thomson, Maria Toledo-Rodriguez, Yun Wang, David C. West, and Rafael Yuste (2008). "Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex." In: *Nature Reviews Neuroscience* 9.7, pp. 557–568 (cit. on p. 172).

Audesirk, Gerald, Leigh Cabell, and Marcey Kern (1997). "Modulation of neurite branching by protein phosphorylation in cultured rat hippocampal neurons." In: *Developmental Brain Research* 102.2, pp. 247–260 (cit. on p. 132).

Bailey, C. H. and M. Chen (1988). "Morphological basis of short-term habituation in Aplysia." In: *The Journal of Neuroscience* 8.7, pp. 2452–2459 (cit. on p. 1).

Bailey, Craig H., Eric R. Kandel, and Kristen M. Harris (2015). "Structural Components of Synaptic Plasticity and Memory Consolidation." In: *Cold Spring Harbor Perspectives in Biology* 7.7, a021758 (cit. on p. 182).

Bamburg, J. R., D. Bray, and K. Chapman (1986). "Assembly of microtubules at the tip of growing axons." In: *Nature* 321.6072, pp. 788–790 (cit. on p. 132).

Banerjee, Swati, Aurea D. Sousa, and Manzoor A. Bhat (2006). "Organization and function of septate junctions: an evolutionary perspective." In: *Cell Biochemistry and Biophysics* 46.1, pp. 65–77 (cit. on p. 235).

Bargmann, CI (2012). "Beyond the connectome: how neuromodulators shape neural circuits." In: *Bioessays* 34.6, pp. 458–65 (cit. on p. 43).

Bargmann, Cornelia I (2015). "How the new neuroscience will advance medicine." In: *JAMA* 314.3, pp. 221–222 (cit. on p. 2).

Bibliography

Bargmann, Cornelia I and Eve Marder (2013). "From the connectome to brain function." In: *Nature Methods* 10.6, pp. 483–490 (cit. on pp. 1, 166).

Basbaum, Allan I, Diana M Bautista, Grégory Scherrer, and David Julius (2009). "Cellular and molecular mechanisms of pain." In: *Cell* 139.2, pp. 267–284 (cit. on p. 14).

Becker, Carlos, Karim Ali, Graham Knott, and Pascal Fua (2012). "Learning Context Cues for Synapse Segmentation in EM Volumes." In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*. Ed. by Nicholas Ayache, Hervé Delingette, Polina Golland, and Kensaku Mori. Vol. 7510. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 585–592 (cit. on p. 42).

Berlucchi, Giovanni (1999). "Some aspects of the history of the law of dynamic polarization of the neuron. From William James to Sherrington, from Cajal and van Gehuchten to Golgi." In: *Journal of the history of the neurosciences* 8.2, pp. 191–201 (cit. on p. 7).

Beyer, J., M. Hadwiger, A. Al-Awami, Won-Ki Jeong, N. Kasthuri, J. W. Lichtman, and H. Pfister (2013). "Exploring the Connectome: Petascale Volume Visualization of Microscopy Data Streams." In: *IEEE Computer Graphics and Applications* 33.4, pp. 50–61 (cit. on p. 173).

Bhatla, Nikhil, Rita Droste, Steven R. Sando, Anne Huang, and H. Robert Horvitz (2015). "Distinct Neural Circuits Control Rhythm Inhibition and Spitting by the Myogenic Pharynx of C. elegans." In: *Current Biology* (cit. on p. 174).

Binzegger, Tom, Rodney J Douglas, and Kevan A C Martin (2004). "A quantitative map of the circuit of cat primary visual cortex." In: *The Journal of Neuroscience* 24.39, pp. 8441–8453 (cit. on p. 3).

Binzegger, Tom, Rodney J Douglas, and Kevan AC Martin (2007). "Stereotypical bouton clustering of individual neurons in cat primary visual cortex." In: *The Journal of Neuroscience* 27.45, pp. 12242–12254 (cit. on p. 45).

Bock, Davi D, Wei-Chung Allen Lee, Aaron M Kerlin, Mark L Andermann, Greg Hood, Arthur W Wetzel, Sergey Yurgenson, Edward R Soucy, Hyon Suk Kim, and R. Clay Reid (2011). "Network anatomy and in vivo physiology of visual cortical neurons." In: *Nature* 471.7337, pp. 177–182 (cit. on pp. 12, 165).

Borisyuk, Roman, Abul Kalam Al Azad, Deborah Conte, Alan Roberts, and Stephen R Soffe (2011). "Modeling the connectome of a simple spinal cord." In: *Frontiers in neuroinformatics* 5, p. 20 (cit. on p. 179).

Branco, Tiago and Michael Häusser (2010). "The single dendritic branch as a fundamental functional unit in the nervous system." In: *Current opinion in neurobiology* 20.4, pp. 494–502 (cit. on p. 90).

Brand, A H and N Perrimon (1993). "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes." In: *Development (Cambridge, England)* 118.2, pp. 401–415 (cit. on p. 180).

Brandes, Ulrik (2001). "A faster algorithm for betweenness centrality." In: *Journal of Mathematical Sociology* 25.2, pp. 163–77 (cit. on p. 60).

Brenner, Sydney (2003). "Nature's gift to science (Nobel lecture)." In: *Chembiochem: A European Journal of Chemical Biology* 4.8, pp. 683–687 (cit. on pp. 9, 10, 178).

Briggman, Kevin L and Davi D Bock (2011a). "Volume electron microscopy for neuronal circuit reconstruction." In: *Current Opinion in Neurobiology* (cit. on p. 1).

Briggman, Kevin L and Winfried Denk (2006a). "Towards neural circuit reconstruction with volume electron microscopy techniques." In: *Current Opinion in Neurobiology* 16.5, pp. 562–570 (cit. on pp. 1, 9).

Briggman, Kevin L, Moritz Helmstaedter, and Winfried Denk (2011b). "Wiring specificity in the direction-selectivity circuit of the retina." In: *Nature* 471.7337, p. 183 (cit. on pp. 2, 12, 90).

Briggman, Kevin L. and William B. Kristan (2006b). "Imaging Dedicated and Multifunctional Neural Circuits Generating Distinct Behaviors." In: *The Journal of Neuroscience* 26.42, pp. 10925–10933 (cit. on p. 181).

Bullmore, Ed and Olaf Sporns (2009). "Complex brain networks: graph theoretical analysis of structural and functional systems." In: *Nature reviews Neuroscience* 10.3, pp. 186–198 (cit. on p. 171).

Bumbarger, Daniel J., Metta Riebesell, Christian Rödelsperger, and Ralf J. Sommer (2013). "System-wide rewiring underlies behavioral differences in predatory and bacterial-feeding nematodes." In: *Cell* 152.1, pp. 109–119 (cit. on p. 181).

Burns, Randal, William Gray Roncal, Dean Kleissas, Kunal Lillaney, Priya Manavalan, Eric Perlman, Daniel R. Berger, Davi D. Bock, Kwanghun Chung, Logan Grosenick, Narayanan Kasthuri, Nicholas C. Weiler, Karl Deisseroth, Michael Kazhdan, Jeff Lichtman, R. Clay Reid, Stephen J. Smith, Alexander S. Szalay, Joshua T. Vogelstein, and R. Jacob Vogelstein (2013). "The Open Connectome Project Data Cluster: Scalable Analysis and Vision for High-Throughput Neuroscience." In: *Scientific and statistical database management : International Conference, SSDBM* (cit. on pp. 165, 169, 170).

Burns, Randal, Joshua T. Vogelstein, and Alexander S. Szalay (2014). "From Cosmos to Connectomes: The Evolution of Data-Intensive Science." In: *Neuron* 83.6, pp. 1249–1252 (cit. on p. 168).

Burrows, Malcolm (1992). "Local circuits for the control of leg movements in an insect." In: *Trends in neurosciences* 15.6, pp. 226–232 (cit. on p. 46).

Burrows, Malcolm and Thomas Matheson (1994). "A presynaptic gain control mechanism among sensory neurons of a locust leg proprioceptor." In: *The Journal of neuroscience* 14.1, pp. 272–282 (cit. on p. 43).

Carandini, Matteo (2012). "From circuits to behavior: a bridge too far?" In: *Nature Neuroscience* 15.4, pp. 507–509 (cit. on p. 1).

Carandini, Matteo and David J Heeger (2011). "Normalization as a canonical neural computation." In: *Nature reviews Neuroscience* advance online publication (cit. on p. 178).

Cardona, A, S Saalfeld, I Arganda Carreras, W Pereanu, J Schindelin, and V Hartenstein (2010). "Identifying neuronal lineages of *Drosophila* by sequence analysis of axon tracts." In: *J Neurosci* 30.22, pp. 7538–53 (cit. on pp. 35, 51).

Cardona, Albert, Stephan Saalfeld, Johannes Schindelin, Ignacio Arganda-Carreras, Stephan Preibisch, Mark Longair, Pavel Tomancak, Volker Hartenstein, and Rodney J. Douglas (2012). "TrakEM2 Software for Neural Circuit Reconstruction." In: *PLoS ONE* 7.6, e38011 (cit. on pp. 19, 25, 44, 102).

Castrén, Eero (2005). "Is mood chemistry?" In: *Nature Reviews Neuroscience* 6.3, pp. 241–246 (cit. on p. 183).

Check-Hayden, Erika (2014). "Technology: The \$1,000 genome." In: *Nature* 507.7492, pp. 294–295 (cit. on p. 165).

Chen, Brian E, Masahiro Kondo, Amélie Garnier, Fiona L Watson, Roland Püettmann-Holgado, David R Lamar, and Dietmar Schmucker (2006). "The molecular diversity of Dscam is functionally required for neuronal wiring specificity in Drosophila." In: *Cell* 125.3, pp. 607–620 (cit. on p. 100).

Chen, John and Barry G. Condron (2008). "Branch architecture of the fly larval abdominal serotoninergic neurons." In: *Dev Biol* 320.1, pp. 30–38 (cit. on p. 62).

Chklovskii, DB, S Vitaladevuni, and LK Scheffer (2010). "Semi-automated reconstruction of neural circuits using electron microscopy." In: *Current Opinion in Neurobiology* 20.5, pp. 667–75 (cit. on pp. 24, 42).

Ciresan, Dan Claudiu, Alessandro Giusti, Luca Maria Gambardella, and Jürgen Schmidhuber (2012). "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images." In: *NIPS*, to appear (cit. on pp. 18, 152, 153, 155, 243).

Clarac, F and D Cattaert (1996). "Invertebrate presynaptic inhibition and motor control." In: *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale* 112.2 (cit. on pp. 34, 43).

Collins, Pamela Y., Vikram Patel, Sarah S. Joestl, Dana March, Thomas R. Insel, Abdallah S. Daar, Scientific Advisory Board and the Executive Committee of the Grand Challenges on Global Mental Health, Warwick Anderson, Muhammad A. Dhansay, Anthony Phillips, Susan Shurin, Mark Walport, Wendy Ewart, Sir John Savill, Isabel A. Bordin, E. Jane Costello, Maureen Durkin, Christopher Fairburn, Roger I. Glass, Wayne Hall, Yueqin Huang, Steven E. Hyman, Kay Jamison, Sylvia Kaaya, Shitij Kapur, Arthur Kleinman, Adesola Ogunniyi, Angel Otero-Ojeda, Mu-Ming Poo, Vijayalakshmi Ravindranath, Barbara J. Sahakian, Shekhar Saxena, Peter A. Singer, and Dan J. Stein (2011). "Grand challenges in global mental health." In: *Nature* 475.7354, pp. 27–30 (cit. on p. 183).

Costa, Nuno Maçarico da and Kevan A. C. Martin (2013). "Sparse reconstruction of brain circuits: Or, how to survive without a microscopic connectome." In: *NeuroImage*. Mapping the Connectome 80, pp. 27–36 (cit. on pp. 11, 12, 168).

Couton, Louise, Alex S. Mauss, Temur Yunusov, Soeren Diegelmann, Jan Felix Evers, and Matthias Landgraf (2015). "Development of connectivity in a motoneuronal network in Drosophila larvae." In: *Current biology: CB* 25.5, pp. 568–576 (cit. on pp. 24, 29, 41, 43, 78, 82, 83, 87).

Cuntz, Hermann, Friedrich Forstner, Bettina Schnell, Georg Ammer, Shamprasad Varija Raghu, and Alexander Borst (2013). "Preserving Neural Function under Extreme Scaling." In: *PLoS ONE* 8.8, e71540 (cit. on p. 135).

Cuntz, Hermann, Alexandre Mathy, and Michael Haeusser (2012). "A scaling law derived from optimal dendritic wiring." In: *Proceedings of the National Academy of Sciences of the United States of America* 109.27, pp. 11014–11018 (cit. on p. 133).

De Robertis, Eduardo D. P. and H. Stanley Bennett (1955). "Some features of the submicroscopic morphology of synapses in frog and earthworm." In: *The Journal of Biophysical and Biochemical Cytology* 1.1, pp. 47–58 (cit. on p. 1).

DeFelipe, Javier (2009). "Cajal's Place in the History of Neuroscience." In: *Encyclopedia of Neuroscience*. Ed. by Larry R. Squire. Oxford: Academic Press, pp. 497–507 (cit. on p. 7).

DeFelipe, Javier (2015). "The dendritic spine story: an intriguing process of discovery." In: *Frontiers in Neuroanatomy* 9, p. 14 (cit. on p. 7).

Denk, W and H Horstmann (2004). "Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure." In: *PLoS Biology* 2.11, pp. 1900–1909 (cit. on p. 165).

Denk, Winfried, Kevin L Briggman, and Moritz Helmstaedter (2012). "Structural neurobiology: missing link to a mechanistic understanding of neural computation." In: *Nature Reviews Neuroscience* 13.5, pp. 351–358 (cit. on p. 43).

Desbois, Muriel, Steven J. Cook, Scott W. Emmons, and Hannes E. Bülow (2015). "Directional Trans-Synaptic Labeling of Specific Neuronal Connections in Live Animals." In: *Genetics*, genetics.115.177006 (cit. on p. 78).

Descartes, Rene (1664). *Treatise of Man*. Prometheus Books. 184 pp. (cit. on p. 14).

Dierick, Herman A. and Fabrizio Gabbiani (2015). "Drosophila Neurobiology: No Escape from Big Data Science." In: *Current Biology* 25.14, R606–R608 (cit. on pp. 166, 180).

Douglas, Rodney J and Kevan A C Martin (2007). "Mapping the matrix: the ways of neocortex." In: *Neuron* 56.2, pp. 226–238 (cit. on p. 182).

Duch, C. and R. B. Levine (2002). "Changes in Calcium Signaling During Postembryonic Dendritic Growth in Manduca sexta." In: *Journal of Neurophysiology* 87.3, pp. 1415–1425 (cit. on p. 147).

Eberhard, J. P., A. Wanner, and G. Wittum (2006). "NeuGen: A tool for the generation of realistic morphology of cortical neurons and neural networks in 3D." In: *Neurocomputing*. Neural Networks Selected Papers from the 7th Brazilian Symposium on Neural Networks (SBRN '04) 7th Brazilian Symposium on Neural Networks 70.1, pp. 327–342 (cit. on p. 179).

Eberle, A. L., S. Mikula, R. Schalek, J. W. Lichtman, M. L. Knothe Tate, and D. Zeidler (2015). "High-resolution, high-throughput imaging with a multibeam scanning electron microscope." In: *Journal of Microscopy* (cit. on pp. 18, 165).

Eberle, Anna Lena, Olaf Selchow, Marlene Thaler, Dirk Zeidler, and Robert Kirmse (2014). "Mission (im)possible – mapping the brain becomes a reality." In: *Microscopy*, dfu104 (cit. on p. 10).

Egger, Robert, Vincent J. Dercksen, Daniel Udvary, Hans-Christian Hege, and Marcel Oberlaender (2014). "Generation of dense statistical connectomes from sparse morphological data." In: *Frontiers in Neuroanatomy* 8, p. 129 (cit. on p. 168).

Elston, Guy N., Ruth Benavides-Piccione, Alejandra Elston, Bendan Zietsch, Javier Defelipe, Paul Manger, Vivien Casagrande, and Jon H. Kaas (2006). "Specializations of the granular prefrontal cortex of primates: implications for cognitive processing." In: *The Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology* 288.1, pp. 26–35 (cit. on p. 182).

Emmons, Scott W. (2015). "The beginning of connectomics: a commentary on White et al. (1986) 'The structure of the nervous system of the nematode Caenorhabditis elegans'." In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370.1666 (cit. on pp. 9, 10, 166).

Feinberg, Evan H, Miri K Vanhoven, Andres Bendesky, George Wang, Richard D Fetter, Kang Shen, and Cornelia I Bargmann (2008). "GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems." In: *Neuron* 57.3, pp. 353–363 (cit. on p. 78).

Foster, Morris W. and Richard R. Sharp (2007). "Share and share alike: deciding how to distribute the scientific and social benefits of genomic data." In: *Nature Reviews Genetics* 8.8, pp. 633–639 (cit. on p. 170).

Freeman, Jeremy (2015). "Open source tools for large-scale neuroscience." In: *Current Opinion in Neurobiology*. Large-Scale Recording Technology (32) 32, pp. 156–163 (cit. on p. 18).

Funke, Jan, Björn Andres, Fred A. Hamprecht, Albert Cardona, and Matthew Cook (2012). "Efficient Automatic 3D-Reconstruction of Branching Neurons from EM Data." In: *IEEE Conference on Computer Vision and Pattern Recognition* (cit. on pp. 18, 42, 155).

Funke, Jan, Jonas Klein, Albert Cardona, and Matthew Cook (2015). "A Tolerant Edit Distance for Evaluation and Training of Electron Microscopy Reconstruction Algorithms." In: *arXiv:1503.02291 [cs]*. arXiv: `1503.02291` (cit. on pp. 18, 21, 152).

Funke, Jan, Julien N. P. Martel, Stephan Gerhard, Bjoern Andres, Dan C. Cireşan, Alessandro Giusti, Luca M. Gambardella, Jürgen Schmidhuber, Hanspeter Pfister, Albert Cardona, and Matthew Cook (2014). "Candidate Sampling for Neuron Reconstruction from Anisotropic Electron Microscopy Volumes." In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*. Ed. by Polina Golland, Nobuhiko Hata, Christian Barillot, Joachim Hornegger, and Robert Howe. Lecture Notes in Computer Science 8673. Springer International Publishing, pp. 17–24 (cit. on pp. 21, 42).

Fyffe, R E W and A R Light (1984). "The ultrastructure of group Ia afferent fiber synapses in the lumbosacral spinal cord of the cat." In: *Brain research* 300.2, pp. 201–209 (cit. on p. 43).

Gabbiani, Fabrizio, Holger G Krapp, Christof Koch, and Gilles Laurent (2002). "Multiplicative computation in a visual neuron sensitive to looming." In: *Nature* 420.6913, pp. 320–324 (cit. on p. 34).

Gerhard, Stephan, Jan Funke, Julien Martel, Albert Cardona, and Richard Fetter (2013). *Segmented anisotropic ssTEM dataset of neural tissue* (cit. on pp. 18, 21, 150).

Gleeson, Padraig, Sharon Crook, Robert C. Cannon, Michael L. Hines, Guy O. Billings, Matteo Farinella, Thomas M. Morse, Andrew P. Davison, Subhasis Ray, Upinder S. Bhalla, Simon R. Barnes, Yoana D. Dimitrova, and R. Angus Silver (2010). "NeuroML: A Language for Describing Data Driven Models of Neurons and Networks with a High Degree of Biological Detail." In: *PLoS Comput Biol* 6.6, e1000815 (cit. on p. 175).

Gleeson, Padraig, Volker Steuber, and R. Angus Silver (2007). "neuroConstruct: A Tool for Modeling Networks of Neurons in 3D Space." In: *Neuron* 54.2, pp. 219–235 (cit. on p. 179).

Goldstone, Robert L., Franco Pestilli, and Katy Börner (2015). "Self-portraits of the brain: cognitive science, data visualization, and communicating brain structure and function." In: *Trends in Cognitive Sciences* (cit. on p. 172).

Golgi, Camillo (1873). "Sulla struttura della sostanza grigia del cervello." In: *Gazz. Med. Lombarda* 33, pp. 244–246 (cit. on p. 5).

Gorczyca, David A., Susan Younger, Shan Meltzer, Sung Eun Kim, Li Cheng, Wei Song, Hye Young Lee, Lily Yeh Jan, and Yuh Nung Jan (2014). "Identification of Ppk26, a DEG/ENaC Channel Functioning with Ppk1 in a Mutually Dependent Manner to Guide Locomotion Behavior in Drosophila." In: *Cell Reports* 9.4, pp. 1446–1458 (cit. on p. 180).

Gouwens, Nathan W and Rachel I Wilson (2009). "Signal propagation in *Drosophila* central neurons." In: *The Journal of Neuroscience* 29.19, pp. 6239–6249 (cit. on p. 34).

Grueber, Wesley B, Lily Y Jan, and Yuh Nung Jan (2002). "Tiling of the Drosophila epidermis by multidendritic sensory neurons." In: *Development (Cambridge, England)* 129.12, pp. 2867–2878 (cit. on pp. 79, 84, 99).

Grueber, Wesley B, Bing Ye, Chung-Hui Yang, Susan Younger, Kelly Borden, Lily Y Jan, and Yuh-Nung Jan (2007). "Projections of Drosophila multidendritic neurons in the central nervous system: links with peripheral dendrite morphology." In: *Development (Cambridge, England)* 134.1, pp. 55–64 (cit. on pp. 31, 79, 84, 103).

Günay, Cengiz, Fred H Sieling, Logesh Dharmar, Wei-Hsiang Lin, Verena Wolfram, Richard Marley, Richard A Baines, and Astrid A Prinz (2015). "Distal Spike Initiation Zone Location Estimation by Morphological Simulation of Ionic Current Filtering Demonstrated in a Novel Model of an Identified Drosophila Motoneuron." In: *PLoS Comput Biol* 11.5, e1004189 (cit. on pp. 26, 34).

Haehn, Daniel, Johanna Beyer, Hanspeter Pfister, Seymour Knowles-Barley, Narayanan Kasthuri, Jeff Lichtman, and Mike Roberts (2014). "Design and evaluation of interactive proofreading tools for connectomics." In: *IEEE Transactions on Visualization and Computer Graphics* PP (99) (cit. on pp. 24, 42, 162).

Hale, Melina E. (2014). "Mapping Circuits beyond the Models: Integrating Connectomics and Comparative Neuroscience." In: *Neuron* 83.6, pp. 1256–1258 (cit. on p. 181).

Hall, David H., David C. Spray, and Michael V. L. Bennett (1983). "Gap junctions and septate-like junctions between neurons of the opisthobranch mollusc Navanax inermis." In: *Journal of Neurocytology* 12.5, pp. 831–846 (cit. on p. 235).

Harnett, Mark T., Judit K. Makara, Nelson Spruston, William L. Kath, and Jeffrey C. Magee (2012). "Synaptic amplification by dendritic spines enhances input cooperativity." In: *Nature* 491.7425, pp. 599–602 (cit. on p. 147).

Hayworth, Kenneth J., Josh L. Morgan, Richard Schalek, Daniel R. Berger, David G. C. Hildebrand, and Jeff W. Lichtman (2014). "Imaging ATUM ultrathin section libraries with WaferMapper: a multi-scale approach to EM reconstruction of neural circuits." In: *Frontiers in Neural Circuits* 8, p. 68 (cit. on pp. 10, 169).

Bibliography

Heckscher, Ellie S, Shawn R Lockery, and Chris Q Doe (2012). "Characterization of *Drosophila* larval crawling at the level of organism, segment, and somatic body wall musculature." In: *The Journal of Neuroscience* 32.36, pp. 12460–12471 (cit. on pp. 31, 35, 43).

Hedwig, B and M Burrows (1996). "Presynaptic inhibition of sensory neurons during kicking movements in the locust." In: *Journal of neurophysiology* 75.3, pp. 1221–1232 (cit. on p. 101).

Helmstaedter, Moritz (2013). "Cellular-resolution connectomics: challenges of dense neural circuit reconstruction." In: *Nature methods* 10.6, pp. 501–507 (cit. on pp. 24, 42).

Helmstaedter, Moritz (2015). "The Mutual Inspirations of Machine Learning and Neuroscience." In: *Neuron* 86.1, pp. 25–28 (cit. on p. 178).

Helmstaedter, Moritz, Kevin L Briggman, and Winfried Denk (2011). "High-accuracy neurite reconstruction for high-throughput neuroanatomy." In: *Nature Neuroscience* 14, pp. 1081–1088 (cit. on pp. 12, 24, 25, 36).

Helmstaedter, Moritz, Kevin L. Briggman, Srinivas C. Turaga, Viren Jain, H. Sebastian Seung, and Winfried Denk (2013). "Connectomic reconstruction of the inner plexiform layer in the mouse retina." In: *Nature* 500.7461, pp. 168–174 (cit. on pp. 10, 25, 42).

Hely, Tim A, Bruce Graham, and Arjen Van Ooyen (2001). "A Computational Model of Dendrite Elongation and Branching Based on MAP2 Phosphorylation." In: *Journal of Theoretical Biology* 210.3, pp. 375–384 (cit. on p. 132).

Hill, Sean L., Yun Wang, Imad Riachi, Felix Schürmann, and Henry Markram (2012). "Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits." In: *Proceedings of the National Academy of Sciences* 109.42, E2885–E2894 (cit. on p. 168).

Holtmaat, Anthony and Karel Svoboda (2009). "Experience-dependent structural synaptic plasticity in the mammalian brain." In: *Nature Reviews Neuroscience* 10.9, pp. 647–658 (cit. on p. 182).

Horridge, G. A. and I. A. Meinertzhagen (1970). "The Accuracy of the Patterns of Connexions of the First- and Second-Order Neurons of the Visual System of Calliphora." In: *Proceedings of the Royal Society of London B: Biological Sciences* 175.1038, pp. 69–82 (cit. on p. 9).

Hu, Xindao, Chris Viesselmann, Sookin Nam, Elliott Merriam, and Erik W. Dent (2008). "Activity-dependent dynamic microtubule invasion of dendritic spines." In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 28.49, pp. 13094–13105 (cit. on pp. 87, 132).

Hua, Yunfeng, Philip Laserstein, and Moritz Helmstaedter (2015). "Large-volume en-bloc staining for electron microscopy-based connectomics." In: *Nature Communications* 6, p. 7923 (cit. on p. 165).

Hughes, Cynthia L and John B Thomas (2007). "A sensory feedback circuit coordinates muscle activity in *Drosophila*." In: *Mol Cell Neurosci* 35.2, pp. 383–96 (cit. on pp. 31, 43).

Hwang, Richard Y, Lixian Zhong, Yifan Xu, Trevor Johnson, Feng Zhang, Karl Deisseroth, and W Daniel Tracey (2007). "Nociceptive neurons protect Drosophila

larvae from parasitoid wasps." In: *Current Biology: CB* 17.24, pp. 2105–2116 (cit. on pp. 79, 100).

Im, Seol Hee and Michael J. Galko (2012). "Pokes, sunburn, and hot sauce: Drosophila as an emerging model for the biology of nociception." In: *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 241.1, pp. 16–26 (cit. on pp. 16, 180).

Ito, Kei, Kazunori Shinomiya, Masayoshi Ito, J. Douglas Armstrong, George Boyan, Volker Hartenstein, Steffen Harzsch, Martin Heisenberg, Uwe Homberg, Arnim Jenett, Haig Keshishian, Linda L. Restifo, Wolfgang Rössler, Julie H. Simpson, Nicholas J. Strausfeld, Roland Strauss, and Leslie B. Vosshall (2014). "A Systematic Nomenclature for the Insect Brain." In: *Neuron* 81.4, pp. 755–765 (cit. on p. 172).

Iyer, Eswar Prasad R., Srividya Chandramouli Iyer, Luis Sullivan, Dennis Wang, Ramakrishna Meduri, Lacey L. Graybeal, and Daniel N. Cox (2013). "Functional Genomic Analyses of Two Morphologically Distinct Classes of Drosophila Sensory Neurons: Post-Mitotic Roles of Transcription Factors in Dendritic Patterning." In: *PLoS ONE* 8.8, e72434 (cit. on p. 133).

Jain, V, HS Seung, and SC Turaga (2010). "Machines that learn to segment images: a crucial technology for connectomics." In: *Current Opinion in Neurobiology* 20.5, pp. 653–66 (cit. on p. 24).

Jan, Y and L Jan (2010). "Branching out: mechanisms of dendritic arborization." In: *Nature Reviews Neuroscience* 11.5, pp. 316–328 (cit. on pp. 79, 132, 178, 179).

Jaworski, Jacek, Lukas C. Kapitein, Susana Montenegro Gouveia, Bjorn R. Dortland, Phebe S. Wulf, Ilya Grigoriev, Paola Camera, Samantha A. Spangler, Paola Di Stefano, Jeroen Demmers, Harm Krugers, Paola Defilippi, Anna Akhmanova, and Casper C. Hoogenraad (2009). "Dynamic microtubules regulate dendritic spine morphology and synaptic plasticity." In: *Neuron* 61.1, pp. 85–100 (cit. on p. 132).

Jenett, Arnim, Gerald M Rubin, Teri-TB Ngo, David Shepherd, Christine Murphy, Heather Dionne, Barret D Pfeiffer, Amanda Cavallaro, Donald Hall, Jennifer Jeter, et al. (2012). "A GAL4-Driver Line Resource for *Drosophila* Neurobiology." In: *Cell reports* 2.4, pp. 991–1001 (cit. on p. 35).

Johnson, Wayne A. and Justin W. Carder (2012). "Drosophila Nociceptors Mediate Larval Aversion to Dry Surface Environments Utilizing Both the Painless TRP Channel and the DEG/ENaC Subunit, PPK1." In: *PLoS ONE* 7.3, e32878 (cit. on p. 99).

Johnston, Jamie and Leon Lagnado (2015). "General features of the retinal connectome determine the computation of motion anticipation." In: *eLife* 4, e06250 (cit. on p. 2).

Jonas, Eric and Konrad Kording (2015). "Automatic discovery of cell types and microcircuitry from neural connectomics." In: *eLife* 4, e04250 (cit. on pp. 42, 171).

Jorgenson, Lyric A., William T. Newsome, David J. Anderson, Cornelia I. Bargmann, Emery N. Brown, Karl Deisseroth, John P. Donoghue, Kathy L. Hudson, Geoffrey S. F. Ling, Peter R. MacLeish, Eve Marder, Richard A. Normann, Joshua R. Sanes, Mark J. Schnitzer, Terrence J. Sejnowski, David W. Tank, Roger Y. Tsien,

Kamil Ugurbil, and John C. Wingfield (2015). "The BRAIN Initiative: developing technology to catalyse neuroscience discovery." In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 370.1668, p. 20140164 (cit. on pp. 3, 170).

Kaas, Jon H (1997). "Topographic Maps are Fundamental to Sensory Processing." In: *Brain Research Bulletin* 44.2, pp. 107–112 (cit. on p. 99).

Kandel, Eric R., James H. Schwartz, Thomas M. Jessell, Steven A. Siegelbaum, and A. J. Hudspeth, eds. (2012). *Principles of Neural Science, Fifth Edition*. 5th edition. New York: McGraw-Hill Education / Medical. 1760 pp. (cit. on p. 14).

Karayiorgou, Maria, Jonathan Flint, Joseph A. Gogos, and Robert C. Malenka (2012). "The best of times, the worst of times for psychiatric disease." In: *Nature Neuroscience* 15.6, pp. 811–812 (cit. on p. 2).

Kasthuri, Narayanan, Kenneth Jeffrey Hayworth, Daniel Raimund Berger, Richard Lee Schalek, José Angel Conchello, Seymour Knowles-Barley, Dongil Lee, Amelio Vázquez-Reina, Verena Kaynig, Thouis Raymond Jones, Mike Roberts, Josh Lyskowski Morgan, Juan Carlos Tapia, H. Sebastian Seung, William Gray Roncal, Joshua Tzvi Vogelstein, Randal Burns, Daniel Lewis Sussman, Carey Eldin Priebe, Hanspeter Pfister, and Jeff William Lichtman (2015). "Saturated Reconstruction of a Volume of Neocortex." In: *Cell* 162.3, pp. 648–661 (cit. on pp. 158, 163, 168).

Katz, Paul, Sten Grillner, Rachel Wilson, Alexander Borst, Ralph Greenspan, György Buzsáki, Kevan Martin, Eve Marder, William Kristan, Rainer Friedrich, and Dmitri "Mitya" Chklovskii (2013). "Vertebrate versus invertebrate neural circuits." In: *Current Biology* 23.12, R504–R506 (cit. on pp. 178, 181).

Kaye, Jane, Catherine Heeney, Naomi Hawkins, Jantina de Vries, and Paula Boddington (2009). "Data sharing in genomics — re-shaping scientific practice." In: *Nature Reviews Genetics* 10.5, pp. 331–335 (cit. on p. 170).

Kaynig, Verena, Amelio Vazquez-Reina, Seymour Knowles-Barley, Mike Roberts, Thouis Jones, Narayanan Kasthuri, Eric Miller, Jeff Lichtman, and Hanspeter Pfister (2015). "Large-Scale Automatic Reconstruction of Neuronal Processes from Electron Microscopy Images." In: *Medical Image Analysis* 22.1, pp. 77–88 (cit. on p. 161).

Keller, Philipp J. and Misha B. Ahrens (2015). "Visualizing Whole-Brain Activity and Development at the Single-Cell Level Using Light-Sheet Microscopy." In: *Neuron* 85.3, pp. 462–483 (cit. on p. 175).

Kim, Jinseop S., Matthew J. Greene, Aleksandar Zlateski, Kisuk Lee, Mark Richardson, Srinivas C. Turaga, Michael Purcaro, Matthew Balkam, Amy Robinson, Bardia F. Behabadi, Michael Campos, Winfried Denk, H. Sebastian Seung, and The EyeWirers (2014). "Space-time wiring specificity supports direction selectivity in the retina." In: *Nature* 509.7500, pp. 331–336 (cit. on pp. 2, 10, 24, 25, 36, 42).

Klein, Jonas (2014). "Cost Functions for Structured Learning for Anisotropic Neuron Reconstruction." Master Thesis. ETH Zurich (cit. on p. 161).

Koch, Christof (2004). *Biophysics of Computation: Information Processing in Single Neurons*. 1 edition. New York: Oxford University Press. 588 pp. (cit. on p. 2).

Koch, Christof (2014). *Building Brain Observatories: The Ten Year Vision* (cit. on p. 168).

Koene, Randal A., Betty Tijms, Peter van Hees, Frank Postma, Alexander de Ridder, Ger J. A. Ramakers, Jaap van Pelt, and Arjen van Ooyen (2009). "NETMORPH: A Framework for the Stochastic Generation of Large Scale Neuronal Networks With Realistic Neuron Morphologies." In: *Neuroinformatics* 7.3, pp. 195–210 (cit. on p. 179).

Kohsaka, Hiroshi, Etsuko Takasu, Takako Morimoto, and Akinao Nose (2014). "A Group of Segmental Premotor Interneurons Regulates the Speed of Axial Locomotion in *Drosophila* Larvae." In: *Current Biology* (cit. on pp. 31, 43, 62).

Kreshuk, Anna, Ullrich Koethe, Elizabeth Pax, Davi D. Bock, and Fred A. Hamprecht (2014). "Automated detection of synapses in serial section transmission electron microscopy image stacks." In: *PloS One* 9.2, e87351 (cit. on p. 42).

Kreshuk, Anna, Christoph N Straehle, Christoph Sommer, Ullrich Koethe, Marco Cantoni, Graham Knott, and Fred A Hamprecht (2011). "Automated detection and segmentation of synaptic contacts in nearly isotropic serial electron microscopy images." In: *PloS ONE* 6.10, e24899 (cit. on pp. 24, 42).

Landgraf, M, T Bossing, GM Technau, and M Bate (1997). "The origin, location, and projections of the embryonic abdominal motorneurons of *Drosophila*." In: *J Neurosci* 17.24, pp. 9642–55 (cit. on p. 26).

Landgraf, Matthias, Natalia Sánchez Soriano, Gerd M. Technau, Joachim Urban, and Andreas Prokop (2003). "Charting the *Drosophila* neuropile: a strategy for the standardised characterization of genetically amenable neurites." In: *Developmental Biology* 260.1, pp. 207–25 (cit. on p. 51).

Lau, Pak-ming, Robert S. Zucker, and David Bentley (1999). "Induction of Filopodia by Direct Local Elevation of Intracellular Calcium Ion Concentration." In: *The Journal of Cell Biology* 145.6, pp. 1265–1276 (cit. on p. 147).

Laughlin, Simon B, Rob R de Ruyter van Steveninck, and John C Anderson (1998). "The metabolic cost of neural information." In: *Nature neuroscience* 1.1, pp. 36–41 (cit. on p. 29).

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning." In: *Nature* 521.7553, pp. 436–444 (cit. on p. 18).

Leiss, Florian, Claudia Grohand, Nancy J. Butcherand, Ian A. Meinertzhagen, and Gaia Tavosanis (2009a). "Synaptic organization in the adult *Drosophila* mushroom body calyx." In: *J Comp Neurol* 517.6, pp. 808–24 (cit. on pp. 24, 26).

Leiss, Florian, Ewa Koper, Irina Hein, Wernher Fouquet, Jana Lindner, Stephan Sigrist, and Gaia Tavosanis (2009b). "Characterization of dendritic spines in the Drosophila central nervous system." In: *Developmental Neurobiology* 69.4, pp. 221–234 (cit. on p. 132).

Lewin, Benjamin, Lynne Cassimeris, Vishwanath R. Lingappa, and George Plopper (2006). *CELLS*. 1 edition. Sudbury, Mass: Jones & Bartlett Learning. 863 pp. (cit. on p. 235).

Li, H.-H., J. R. Kroll, S. M. Lennox, O. Ogundeyi, J. Jeter, G. Depasquale, and J. W. Truman (2014). "A GAL4 Driver Resource for Developmental and Behavioral Studies on the Larval CNS of *Drosophila*." In: *Cell Reports* 8.3, pp. 897–908 (cit. on pp. 31, 35, 44, 56).

Bibliography

Li, Wen-Chang, Tom Cooke, Bart Sautois, Stephen R Soffe, Roman Borisyuk, and Alan Roberts (2007). "Axon and dendrite geography predict the specificity of synaptic connections in a functioning spinal cord network." In: *Neural Development* 2, p. 17 (cit. on p. 168).

Li, Zheng, Ken-Ichi Okamoto, Yasunori Hayashi, and Morgan Sheng (2004). "The Importance of Dendritic Mitochondria in the Morphogenesis and Plasticity of Spines and Synapses." In: *Cell* 119.6, pp. 873–887 (cit. on p. 162).

Lichtman, Jeff W., Hanspeter Pfister, and Nir Shavit (2014). "The big data challenges of connectomics." In: *Nature Neuroscience* 17.11, pp. 1448–1454 (cit. on pp. 18, 168).

Lichtman, Jeff W and Joshua R Sanes (2008). "Ome sweet ome: what can the genome tell us about the connectome?" In: *Current Opinion in Neurobiology*. Signalling mechanisms 18.3, pp. 346–353 (cit. on p. 165).

Lucchi, A., K. Smith, R. Achanta, G. Knott, and P. Fua (2011). "Supervoxel-Based Segmentation of Mitochondria in EM Image Stacks With Learned Shape Features." In: *IEEE transactions on medical imaging* (cit. on p. 42).

Macagno, E. R., V. Lopresti, and C. Levinthal (1973). "Structure and Development of Neuronal Connections in Isogenic Organisms: Variations and Similarities in the Optic System of Daphnia magna." In: *Proceedings of the National Academy of Sciences of the United States of America* 70.1, pp. 57–61 (cit. on p. 9).

Maeda, Shoji, So Nakagawa, Michihiro Suga, Eiki Yamashita, Atsunori Oshima, Yoshinori Fujiyoshi, and Tomitake Tsukihara (2009). "Structure of the connexin 26 gap junction channel at 3.5 A resolution." In: *Nature* 458.7238, pp. 597–602 (cit. on p. 235).

Malanowski, Sarah and Carl F. Craver (2014). "The spine problem: finding a function for dendritic spines." In: *Frontiers in Neuroanatomy* 8, p. 95 (cit. on p. 147).

Marblestone, Adam H., Evan R. Daugharthy, Reza Kalhor, Ian D. Peikon, Justus M. Kebschull, Seth L. Shipman, Yuriy Mishchenko, David A. Dalrymple, Bradley M. Zamft, Konrad P. Kording, Edward S. Boyden, Anthony M. Zador, and George M. Church (2013). "Conneconomics: The Economics of Large-Scale Neural Connectomics." In: *bioRxiv* (cit. on pp. 10, 168).

Marblestone, Adam H., Evan R. Daugharthy, Reza Kalhor, Ian D. Peikon, Justus M. Kebschull, Seth L. Shipman, Yuriy Mishchenko, Je Hyuk Lee, Konrad P. Kording, Edward S. Boyden, Anthony M. Zador, and George M. Church (2014). "Rosetta Brains: A Strategy for Molecularly-Annotated Connectomics." In: *arXiv:1404.5103 [q-bio]*. arXiv: `1404.5103` (cit. on p. 168).

Marc, Robert E., James R. Anderson, Bryan W. Jones, Crystal L. Sigulinsky, and James S. Lauritzen (2014). "The AII amacrine cell connectome: a dense network hub." In: *Frontiers in Neural Circuits* 8, p. 104 (cit. on p. 10).

Marcus, Gary, Adam Marblestone, and Thomas Dean (2014). "The atoms of neural computation." In: *Science* 346.6209, pp. 551–552 (cit. on p. 182).

Marín-Padilla, M. (1998). "Cajal-Retzius cells and the development of the neocortex." In: *Trends in Neurosciences* 21.2, pp. 64–71 (cit. on pp. 6, 182).

Marrone, Diano F and Ted L Petit (2002). "The role of synaptic morphology in neural plasticity: structural interactions underlying synaptic power." In: *Brain Research Reviews* 38.3, pp. 291–308 (cit. on p. 182).

Marx, Vivien (2013). "Neurobiology: Brain mapping in high resolution." In: *Nature* 503.7474, pp. 147–152 (cit. on p. 10).

Mason, Peggy (2012). "Medullary circuits for nociceptive modulation." In: *Current Opinion in Neurobiology* 22.4, pp. 640–645 (cit. on p. 101).

Mattson, M. P. and S. B. Kater (1987). "Calcium regulation of neurite elongation and growth cone motility." In: *The Journal of Neuroscience* 7.12, pp. 4034–4043 (cit. on p. 147).

Mayford, Mark, Steven A. Siegelbaum, and Eric R. Kandel (2012). "Synapses and Memory Storage." In: *Cold Spring Harbor Perspectives in Biology* 4.6, a005751 (cit. on p. 1).

Meinertzhagen, IA and SD O'Neil (1991). "Synaptic organization of columnar elements in the lamina of the wild type in *Drosophila melanogaster*." In: *J Comp Neurol* 305, pp. 232–63 (cit. on p. 25).

Meinertzhagen, Ian A and Chi-Hon Lee (2012). "The genetic analysis of functional connectomics in Drosophila." In: *Advances in genetics* 80, pp. 99–151 (cit. on pp. 166, 175, 178).

Merritt, D J and P M Whitington (1995). "Central projections of sensory neurons in the Drosophila embryo correlate with sensory modality, soma position, and proneural gene function." In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 15.3, pp. 1755–1767 (cit. on pp. 15, 79).

Micheva, Kristina D. and Stephen J. Smith (2007). "Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits." In: *Neuron* 55.1, pp. 25–36 (cit. on p. 12).

Mikula, Shawn and Winfried Denk (2015). "High-resolution whole-brain staining for electron microscopic circuit reconstruction." In: *Nature Methods* (cit. on pp. 11, 18, 165, 168).

Millard, S Sean, Zhiyuan Lu, S Lawrence Zipursky, and Ian A Meinertzhagen (2010). "Drosophila dscam proteins regulate postsynaptic specificity at multiple-contact synapses." In: *Neuron* 67.5, pp. 761–768 (cit. on p. 100).

Molofsky, Anna V., Robert Krenick, Erik Ullian, Hui-hsin Tsai, Benjamin Deneen, William D. Richardson, Ben A. Barres, and David H. Rowitch (2012). "Astrocytes and disease: a neurodevelopmental perspective." In: *Genes & Development* 26.9, pp. 891–907 (cit. on p. 163).

Morgan, Joshua L. and Jeff W. Lichtman (2013). "Why not connectomics?" In: *Nature Methods* 10.6, pp. 494–500 (cit. on pp. 1, 165, 168).

Narayanan, Rajeevan T., Robert Egger, Andrew S. Johnson, Huibert D. Mansvelder, Bert Sakmann, Christiaan P. J. de Kock, and Marcel Oberlaender (2015). "Beyond Columnar Organization: Cell Type- and Target Layer-Specific Principles of Horizontal Axon Projection Patterns in Rat Vibrissal Cortex." In: *Cerebral Cortex*, bhv053 (cit. on p. 177).

Nat Neurosci Editorial (2007). "Got data?" In: *Nature Neuroscience* 10.8, pp. 931–931 (cit. on p. 170).

Bibliography

Nern, Aljoscha, Barret D. Pfeiffer, and Gerald M. Rubin (2015). "Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system." In: *Proceedings of the National Academy of Sciences*, p. 201506763 (cit. on pp. 31, 104).

Newman, Mark (2010). *Networks: An Introduction*. OUP Oxford (cit. on p. 34).

Nicolai, Laura JJ, Ariane Ramaekers, Tim Raemaekers, Andrzej Drozdzecki, Alex S Mauss, Jiekun Yan, Matthias Landgraf, Wim Annaert, and Bassem A Hassan (2010). "Genetically encoded dendritic marker sheds light on neuronal connectivity in *Drosophila*." In: *Proceedings of the National Academy of Sciences* 107.47, pp. 20553–20558 (cit. on p. 26).

Niell, Cristopher M., Martin P. Meyer, and Stephen J. Smith (2004). "In vivo imaging of synapse formation on a growing dendritic arbor." In: *Nature Neuroscience* 7.3, pp. 254–260 (cit. on p. 147).

Nielsen, Michael (2013). *Reinventing Discovery: The New Era of Networked Science*. Reprint edition. Princeton, N.J.: Princeton University Press. 272 pp. (cit. on p. 9).

Oberti, Daniele, Moritz A. Kirschmann, and Richard H. R. Hahnloser (2011). "Projection Neuron Circuits Resolved Using Correlative Array Tomography." In: *Frontiers in Neuroscience* 5 (cit. on p. 12).

O'Connor, Daniel H., Daniel Huber, and Karel Svoboda (2009). "Reverse engineering the mouse brain." In: *Nature* 461.7266, pp. 923–929 (cit. on p. 175).

Ohyama, Tomoko, Tihana Jovanic, Gennady Denisov, Tam C Dang, Dominik Hoffmann, Rex A Kerr, and Marta Zlatic (2013). "High-throughput analysis of stimulus-evoked behaviors in Drosophila larva reveals multiple modality-specific escape strategies." In: *PloS one* 8.8, e71706 (cit. on p. 166).

Ohyama, Tomoko, Casey M. Schneider-Mizell, Richard D. Fetter, Javier Valdes Aleman, Romain Franconville, Marta Rivera-Alba, Brett D. Mensh, Kristin M. Branson, Julie H. Simpson, James W. Truman, Albert Cardona, and Marta Zlatic (2015). "A multilevel multimodal circuit enhances action selection in Drosophila." In: *Nature* 520.7549, pp. 633–639 (cit. on pp. 12, 20, 24–26, 29, 31, 43, 62, 78, 84, 98–100, 102, 103, 166, 176, 180, 188).

Okusawa, Satoko, Hiroshi Kohsaka, and Akinao Nose (2014). "Serotonin and Downstream Leucokinin Neurons Modulate Larval Turning Behavior in *Drosophila*." In: *The Journal of Neuroscience* 34.7, pp. 2544–2558 (cit. on p. 62).

Ooyen, Arjen van (2011). "Using theoretical models to analyse neural development." In: *Nature Reviews Neuroscience* 12.6, pp. 311–326 (cit. on pp. 132, 134).

Ossipov, Michael H., Gregory O. Dussor, and Frank Porreca (2010). "Central modulation of pain." In: *The Journal of Clinical Investigation* 120.11, pp. 3779–3787 (cit. on p. 101).

Palade, G (1954). "Electron Microscope Observations of interneuronal and neuromuscular synapses." In: *The anatomical record* 118.2, pp. 335–336 (cit. on pp. 1, 7).

Palay, Sanford L. and George E. Palade (1955). "The Fine Structure of Neurons." In: *The Journal of Biophysical and Biochemical Cytology* 1.1, pp. 69–88 (cit. on p. 1).

Pannese, E. (2007). "The contribution of Camillo Golgi to our understanding of the structure of the nervous system." In: *Archives Italiennes De Biologie* 145.2, pp. 111–115 (cit. on p. 5).

Parekh, Ruchi and Giorgio A. Ascoli (2013). "Neuronal Morphology Goes Digital: A Research Hub for Cellular and System Neuroscience." In: *Neuron* 77.6, pp. 1017–1038 (cit. on p. 179).

Pelt, J. van, R. W. H. Verwer, and H. B. M. Uylings (1989). "Centrifugal-order distributions in binary topological trees." In: *Bulletin of Mathematical Biology* 51.4, pp. 511–536 (cit. on p. 135).

Pelt, Jaap van and Harry B M Uylings (2002). "Branching rates and growth functions in the outgrowth of dendritic branching patterns." In: *Network: Computation in Neural Systems* 13.3, pp. 261–281 (cit. on p. 132).

Perez, Fernando (2015). *Project Jupyter: Computational Narratives as the Engine of Collaborative Data Science*. Project Jupyter. URL: `http://blog.jupyter.org/2015/07/07/project-jupyter-computational-narratives-as-the-engine-of-collaborative-data-science/` (visited on 07/21/2015) (cit. on p. 171).

Pérez, Fernando and Brian E. Granger (2007). "IPython: A System for Interactive Scientific Computing." In: *Computing in Science & Engineering* 9.3, pp. 21–29 (cit. on p. 171).

Pérez, Fernando, Brian E. Granger, and John D. Hunter (2011). "Python: An Ecosystem for Scientific Computing." In: *Computing in Science & Engineering* 13.2, pp. 13–21 (cit. on p. 171).

Pietzsch, Tobias, Stephan Saalfeld, Stephan Preibisch, and Pavel Tomancak (2015). "BigDataViewer: visualization and processing for large image data sets." In: *Nature Methods* 12.6, pp. 481–483 (cit. on p. 187).

Plaza, Stephen M, Louis K Scheffer, and Dmitri B Chklovskii (2014). "Toward large-scale connectome reconstructions." In: *Current Opinion in Neurobiology*. Theoretical and computational neuroscience 25, pp. 201–210 (cit. on pp. 18, 161).

Plaza, Stephen M, Louis K Scheffer, and Mathew Saunders (2012). "Minimizing Manual Image Segmentation Turn-Around Time for Neuronal Reconstruction by Embracing Uncertainty." In: *PLoS ONE* 7.9, e44448 (cit. on pp. 24, 42).

Polsky, Alon, Bartlett W Mel, and Jackie Schiller (2004). "Computational subunits in thin dendrites of pyramidal cells." In: *Nature neuroscience* 7.6, pp. 621–627 (cit. on p. 29).

Purves, Dale (1988). *Body and Brain: A Trophic Theory of Neural Connections*. Harvard University Press. 244 pp. (cit. on p. 78).

Ragan-Kelley, Benjamin, William Anton Walters, Daniel McDonald, Justin Riley, Brian E Granger, Antonio Gonzalez, Rob Knight, Fernando Perez, and J Gregory Caporaso (2013). "Collaborative cloud-enabled tools allow rapid, reproducible biological insights." In: *The ISME Journal* 7.3, pp. 461–464 (cit. on p. 171).

Rah, Jong-Cheol, Linqing Feng, Shaul Druckmann, Hojin Lee, and Jinhyun Kim (2015). "From a meso- to micro-scale connectome: array tomography and mGRASP." In: *Frontiers in Neuroanatomy*, p. 78 (cit. on p. 168).

Bibliography

Randel, Nadine, Réza Shahidi, Csaba Verasztó, Luis A. Bezares-Calderón, Steffen Schmidt, and Gáspár Jékely (2015). "Inter-individual stereotypy of the Platynereis larval visual connectome." In: *eLife* 4 (cit. on pp. 43, 176).

Reimann, Michael Wolfgang, Eilif Benjamin Muller, Srikanth Ramaswamy, and Henry Markram (2015). "An Algorithm to Predict the Connectome of Neural Microcircuits." In: *Frontiers in Neural Circuits* 9, p. 28 (cit. on p. 168).

Reiner, Anton, David J Perkel, Laura L Bruce, Ann B Butler, Andras Csillag, Wayne Kuenzel, Loreta Medina, George Paxinos, Toru Shimizu, Georg Striedter, Martin Wild, Gregory F Ball, Sarah Durand, Onur Guetuerkuen, Diana W Lee, Claudio V Mello, Alice Powers, Stephanie A White, Gerald Hough, Lubica Kubikova, Tom V Smulders, Kazuhiro Wada, Jennifer Dugas-Ford, Scott Husband, Keiko Yamamoto, Jing Yu, Connie Siang, and Erich D Jarvis (2004). "The Avian Brain Nomenclature Forum: Terminology for a New Century in Comparative Neuroanatomy." In: *The Journal of comparative neurology* 473, E1–E6 (cit. on p. 172).

Roberts, Alan, Deborah Conte, Mike Hull, Robert Merrison-Hort, Abul Kalam al Azad, Edgar Buhl, Roman Borisyuk, and Stephen R. Soffe (2014). "Can Simple Rules Control Development of a Pioneer Vertebrate Neuronal Network Generating Behavior?" In: *The Journal of Neuroscience* 34.2, pp. 608–621 (cit. on p. 180).

Robertson, Jessica L, Asako Tsubouchi, and W Daniel Tracey (2013). "Larval Defense against Attack from Parasitoid Wasps Requires Nociceptive Neurons." In: *PloS one* 8.10, e78704 (cit. on p. 100).

Rooney, Timothy M. and Marc R. Freeman (2014). "Drosophila models of neuronal injury." In: *ILAR journal / National Research Council, Institute of Laboratory Animal Resources* 54.3, pp. 291–295 (cit. on p. 179).

Rossant, Cyrille (2013). *Learning IPython for Interactive Computing and Data Visualization*. Birmingham: Packt Publishing. 138 pp. (cit. on p. 171).

Rossant, Cyrille (2014). *IPython Interactive Computing and Visualization Cookbook*. Birmingham: Packt Publishing - ebooks Account. 423 pp. (cit. on p. 171).

Rubin, Daniel L., Suzanna E. Lewis, Chris J. Mungall, Sima Misra, Monte Westerfield, Michael Ashburner, Ida Sim, Christopher G. Chute, Solbrig Harold, Margaret-Anne Storey, Barry Smith, John Day-Richter, Natalya F. Noy, and Mark A. Musen (2006). "National Center for Biomedical Ontology: Advancing Biomedicine through Structured Organization of Scientific Knowledge." In: *OMICS: A Journal of Integrative Biology* 10.2, pp. 185–198 (cit. on p. 172).

Rubinov, Mikail and Olaf Sporns (2010). "Complex network measures of brain connectivity: uses and interpretations." In: *NeuroImage* 52.3, pp. 1059–1069 (cit. on p. 171).

Rudomin, P and Robert F Schmidt (1999). "Presynaptic inhibition in the vertebrate spinal cord revisited." In: *Experimental Brain Research* 129.1, pp. 1–37 (cit. on p. 101).

Ryglewski, Stefanie, Dimitrios Kadas, Katie Hutchinson, Natalie Schuetzler, Fernando Vonhoff, and Carsten Duch (2014). "Dendrites are dispensable for basic motoneuron function but essential for fine tuning of behavior." In: *Proceedings of the National Academy of Sciences of the United States of America* 111.50, pp. 18049–18054 (cit. on p. 100).

Saalfeld, S., A. Cardona, V. Hartenstein, and P. Tomancak (2009). "CATMAID: collaborative annotation toolkit for massive amounts of image data." In: *Bioinformatics* 25.15, pp. 1984–1986 (cit. on pp. 13, 19, 25, 49, 201).

Saalfeld, Stephan, Richard Fetter, Albert Cardona, and Pavel Tomancak (2012). "Elastic volume reconstruction from series of ultra-thin microscopy sections." In: *Nature Methods* (cit. on pp. 19, 44, 58, 102).

Sadtler, Patrick T., Kristin M. Quick, Matthew D. Golub, Steven M. Chase, Stephen I. Ryu, Elizabeth C. Tyler-Kabara, Byron M. Yu, and Aaron P. Batista (2014). "Neural constraints on learning." In: *Nature* 512.7515, pp. 423–426 (cit. on p. 175).

Sánchez-Soriano, Natalia, Wolfgang Bottenberg, André Fiala, Ulrike Haessler, Afroditi Kerassoviti, Elisabeth Knust, Robert Löhr, and Andreas Prokop (2005). "Are dendrites in *Drosophila* homologous to vertebrate dendrites?" In: *Developmental biology* 288.1, pp. 126–138 (cit. on p. 26).

Sanes, Joshua R. and Masahito Yamagata (2009). "Many Paths to Synaptic Specificity." In: *Annual Review of Cell and Developmental Biology* 25.1, pp. 161–195 (cit. on p. 100).

Sato, T. (1968). "A modified method for lead staining of thin sections." In: *J Electron Microsc (Tokyo)* 17, pp. 158–9 (cit. on pp. 44, 102).

Schaefer, J E, J W Worrell, and R B Levine (2010). "Role of Intrinsic Properties in Drosophila Motoneuron Recruitment During Fictive Crawling." In: *Nature neuroscience* 104.3, pp. 1257–1266 (cit. on p. 43).

Scheffer, Louis K., Bill Karsh, and Shiv Vitaladevun (2013). "Automated Alignment of Imperfect EM Images for Neural Reconstruction." In: *arXiv:1304.6034 [q-bio]*. arXiv: 1304.6034 (cit. on p. 165).

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona (2012). "Fiji: an open-source platform for biological-image analysis." In: *Nature Methods* 9.7, pp. 676–682 (cit. on pp. 19, 103).

Schmidhuber, Juergen (2015). "Deep Learning in Neural Networks: An Overview." In: *Neural Networks* 61, pp. 85–117. arXiv: 1404.7828 (cit. on p. 18).

Schneider-Mizell, Casey M., Stephan Gerhard, Mark Longair, Tom Kazimiers, Feng Li, Maarten F. Zwart, Andrew Champion, Frank Midgley, Richard Fetter, Stephan Saalfeld, and Albert Cardona (2015). "Quantitative neuroanatomy for connectomics in Drosophila." In: *bioRxiv*, p. 026617 (cit. on pp. 78, 80, 87, 93, 102, 103, 137, 138, 141, 142, 190).

Schrader, S and D J Merritt (2000). "Central projections of Drosophila sensory neurons in the transition from embryo to larva." In: *The Journal of comparative neurology* 425.1, pp. 34–44 (cit. on p. 84).

Scott, Ethan K, John E Reuter, and Liqun Luo (2003). "Small GTPase Cdc42 is required for multiple aspects of dendritic morphogenesis." In: *The Journal of neuroscience* 23.8, pp. 3118–3123 (cit. on pp. 24, 26).

Séjourné, Julien, Pierre-Yves Placais, Yoshinori Aso, Igor Siwanowicz, Séverine Trannoy, Vladimiros Thoma, Stevanus R Tedjakumala, Gerald M Rubin, Paul Tchénio,

Kei Ito, Guillaume Isabel, Hiromu Tanimoto, and Thomas Preat (2011). "Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in Drosophila." In: *Nature neuroscience* (cit. on p. 182).

Senft, Stephen L. (2011). "A Brief History of Neuronal Reconstruction." In: *Neuroinformatics* 9.2, pp. 119–128 (cit. on pp. 1, 5).

Seung, H Sebastian (2009). "Reading the Book of Memory: Sparse Sampling versus Dense Mapping of Connectomes." In: *Neuron* 62.1, pp. 17–29 (cit. on p. 11).

Seung, H Sebastian and Uygar Sümbül (2015). "Neuronal Cell Types and Connectivity: Lessons from the Retina." In: *Neuron* 83.6, pp. 1262–1272 (cit. on p. 42).

Seung, Sebastian (2013). *Connectome: How the Brain's Wiring Makes Us Who We Are*. Boston: Mariner Books. 384 pp. (cit. on p. 182).

Shepherd, G. M. (1978). "Microcircuits in the nervous system." In: *Scientific American* 238.2, pp. 93–103 (cit. on p. 92).

Shepherd, G. M. (1996). "The dendritic spine: a multifunctional integrative unit." In: *Journal of Neurophysiology* 75.6, pp. 2197–2210 (cit. on p. 147).

Shepherd, Gordon M. (2009). "Dendrodendritic synapses: past, present and future." In: *Annals of the New York Academy of Sciences* 1170 (cit. on p. 92).

Sherrington, CS (1906). *The integrative action of the nervous system*. New Haven, CT: Yale University Press (cit. on pp. 1, 7, 14).

Shinomiya, Kazunori, Shin-ya Takemura, Patricia K. Rivlin, Stephen M. Plaza, Louis K. Scheffer, and Ian A. Meinertzhagen (2015). "A common evolutionary origin for the ON- and OFF-edge motion detection pathways of the Drosophila visual system." In: *Frontiers in Neural Circuits*, p. 33 (cit. on p. 181).

Simpson, Julie H (2009). "Mapping and manipulating neural circuits in the fly brain." In: *Advances in genetics* 65, pp. 79–143 (cit. on pp. 166, 178).

Skeath, James B. and Stefan Thor (2003). "Genetic control of Drosophila nerve cord development." In: *Current Opinion in Neurobiology* 13.1, pp. 8–15 (cit. on p. 78).

Smith, Ewan St. John and Gary R. Lewin (2009). "Nociceptors: a phylogenetic view." In: *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 195.12, pp. 1089–1106 (cit. on p. 16).

Smith, Jane Anne (1991). *A Question of Pain in Invertebrates*. URL: http://www.abolitionist.com/darwinian-life/invertebrate-pain.html (visited on 04/25/2012) (cit. on p. 14).

Smith, S. J. (1988). "Neuronal cytomechanics: the actin-based motility of growth cones." In: *Science (New York, N.Y.)* 242.4879, pp. 708–715 (cit. on p. 132).

Smythies, John (2015). "Off the beaten track: the molecular structure of long-term memory: three novel hypotheses—electrical, chemical and anatomical (allosteric)." In: *Frontiers in Integrative Neuroscience* 9, p. 4 (cit. on pp. 132, 183).

Sneddon, L. U, V. A Braithwaite, and M. J Gentle (2003). "Do Fishes Have Nociceptors? Evidence for the Evolution of a Vertebrate Sensory System." In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270.1520, pp. 1115–1121 (cit. on p. 14).

Sommer, C, C Strähle, U Köthe, and F. A. Hamprecht (2011). "Ilastik: Interactive learning and segmentation toolkit." In: Eighth IEEE International Symposium on Biomedical Imaging (ISBI). Proceedings. Pp. 230–233 (cit. on p. 161).

Song, Wei, Maika Onishi, Lily Yeh Jan, and Yuh Nung Jan (2007). "Peripheral multidendritic sensory neurons are necessary for rhythmic locomotion behavior in *Drosophila* larvae." In: *Proceedings of the National Academy of Sciences* 104.12, pp. 5199–5204 (cit. on p. 43).

Stowers, R. Steven, Laura J. Megeath, Jolanta Górska-Andrzejak, Ian A. Meinertzhagen, and Thomas L. Schwarz (2002). "Axonal Transport of Mitochondria to Synapses Depends on Milton, a Novel Drosophila Protein." In: *Neuron* 36.6, pp. 1063–1077 (cit. on p. 162).

Strausfeld, Nicholas J. and Frank Hirth (2013). "Deep homology of arthropod central complex and vertebrate basal ganglia." In: *Science (New York, N.Y.)* 340.6129, pp. 157–161 (cit. on pp. 16, 181).

Streissler, Erich W., Almut Schuz, and Robert Miller (2002). *Cortical Areas: Unity and Diversity*. London ; New York: Routledge Chapman & Hall. 526 pp. (cit. on p. 182).

Sulkowski, Mikolaj J., Mathieu S. Kurosawa, and Daniel N. Cox (2011). "Growing Pains: Development of the Larval Nocifensive Response in Drosophila." In: *The Biological Bulletin* 221.3, pp. 300–306 (cit. on pp. 16, 98, 99).

Suloway, C., J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, Quispe, S. Stagg, C. S. Potter, and B. Carragher (2005). "Automated molecular microscopy: the new Leginon system." In: *Journal of Structural Biology* 151, pp. 41–60 (cit. on pp. 44, 102).

Suster, ML and M Bate (2002). "Embryonic assembly of a central pattern generator without sensory input." In: *Nature* 416.6877, pp. 174–8 (cit. on p. 43).

Takemura, Shin-ya, Arjun Bharioke, Zhiyuan Lu, Aljoscha Nern, Shiv Vitaladevuni, Patricia K. Rivlin, William T. Katz, Donald J. Olbris, Stephen M. Plaza, Philip Winston, Ting Zhao, Jane Anne Horne, Richard D. Fetter, Satoko Takemura, Katerina Blazek, Lei-Ann Chang, Omotara Ogundeyi, Mathew A. Saunders, Victor Shapiro, Christopher Sigmund, Gerald M. Rubin, Louis K. Scheffer, Ian A. Meinertzhagen, and Dmitri B. Chklovskii (2013). "A visual motion detection circuit suggested by Drosophila connectomics." In: *Nature* 500.7461, pp. 175–181 (cit. on pp. 10, 24, 26, 36, 41, 43, 165, 177).

Tobin, David M and Cornelia I Bargmann (2004). "Invertebrate nociception: behaviors, neurons and molecules." In: *Journal of Neurobiology* 61.1, pp. 161–174 (cit. on p. 16).

Todd, Andrew J (2010). "Neuronal circuitry for pain processing in the dorsal horn." In: *Nature reviews Neuroscience* 11.12, pp. 823–836 (cit. on p. 15).

Tracey, Daniel, Rachel Wilson, Gilles Laurent, and Seymour Benzer (2003). "painless, a Drosophila gene essential for nociception." In: *Cell* 113.2, pp. 261–273 (cit. on pp. 79, 98, 100).

Tsien, Roger Y. (2013). "Very long-term memories may be stored in the pattern of holes in the perineuronal net." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.30, pp. 12456–12461 (cit. on p. 183).

Turaga, Srinivas C, Joseph F Murray, Viren Jain, Fabian Roth, Moritz Helmstaedter, Kevin Briggman, Winfried Denk, and H Sebastian Seung (2010). "Convolutional

networks can learn to generate affinity graphs for image segmentation." In: *Neural computation* 22.2, pp. 511–538 (cit. on p. 18).

Uylings, Harry B M and Jaap van Pelt (2002). "Measures for quantifying dendritic arborizations." In: *Network (Bristol, England)* 13.3, pp. 397–414 (cit. on p. 135).

Van Pelt, Jaap, Alexander E. Dityatev, and Harry B.M. Uylings (1997). "Natural variability in the number of dendritic segments: Model-based inferences about branching during neurite outgrowth." In: *The Journal of Comparative Neurology* 387.3, pp. 325–340 (cit. on p. 132).

Varshney, Lav R, Beth L Chen, Eric Paniagua, David H Hall, and Dmitri B Chklovskii (2009). "Structural Properties of the Caenorhabditis elegans Neuronal Network." In: *arXiv* q-bio.NC (cit. on pp. 57, 174).

Vazquez-Reina, Amelio, Daniel Huang, Michael Gelbart, Jeff Lichtman, Eric Miller, and Hanspeter Pfister (2011). "Segmentation Fusion for Connectomics." In: Barcelona, Spain: IEEE (cit. on pp. 18, 42).

Veeraraghavan, A, A Genkin, S Vitaladevuni, L Scheffer, S Xu, H Hess, R Fetter, M Cantoni, G Knott, and D Chklovskii (2010). "Increasing Depth Resolution of Electron Microscopy of Neural Circuits Using Sparse Tomographic Reconstruction." In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (cit. on p. 24).

Vogelstein, Joshua T, Youngser Park, Tomoko Ohyama, Rex A Kerr, James W Truman, Carey E Priebe, and Marta Zlatic (2014). "Discovery of brainwide neural-behavioral maps via multiscale unsupervised structure learning." In: *Science* 344.6182, pp. 386–392 (cit. on p. 56).

Weiler, Nicholas C., Forrest Collman, Joshua T. Vogelstein, Randal Burns, and Stephen J. Smith (2014). "Synaptic molecular imaging in spared and deprived columns of mouse barrel cortex with array tomography." In: *Scientific Data* 1, p. 140046 (cit. on p. 170).

Weinberg, Richard J (1997). "Are Topographic Maps Fundamental to Sensory Processing?" In: *Brain Research Bulletin* 44.2, pp. 113–116 (cit. on p. 99).

Wells, William A. (2005). "The discovery of synaptic vesicles." In: *The Journal of Cell Biology* 168.1, pp. 12–13 (cit. on pp. 1, 7).

White, JG, E Southgate, JN Thomson, and S Brenner (1986). "The structure of the nervous system of the nematode Caenorhabditis elegans." In: *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 314.1165, p. 1 (cit. on pp. 9, 10, 174).

Wilson, RI and G Laurent (2005). "Role of GABAergic inhibition in shaping odor-evoked spatiotemporal patterns in *Drosophila* antennal lobe." In: *J Neurosci* 25, pp. 9069–79 (cit. on p. 46).

Wittenberg, Gayle M and Samuel S.-H Wang (2007). "Evolution and scaling of dendrites." In: *Dendrites*. Ed. by Greg Stuart, Nelson Spruston, and Michael Häusser. Oxford University Press, pp. 42–67 (cit. on pp. 78, 132).

Wolf, S, S Grein, and G Queisser (2012). "Employing NeuGen 2.0 to Automatically Generate Realistic Morphologies of Hippocampal Neurons and Neural Networks in 3D - Springer." In: (cit. on p. 179).

Woolf, Clifford J and Qiufu Ma (2007). "Nociceptors–noxious stimulus detectors." In: *Neuron* 55.3, pp. 353–364 (cit. on pp. 14, 16).

Wu, Z, LB Sweeney, JC Ayoob K Chak, BJ Andreone, T Ohyama, R Kerr, L Luo, M Zlatić, and AL Kolodkin (2011). "A combinatorial semaphorin code instructs the initial steps of sensory circuit assembly in the *Drosophila* CNS." In: *Neuron* 70.2, pp. 281–98 (cit. on p. 24).

Xing, Eric P, Michael I Jordan, Stuart Russell, and Andrew Y Ng (2002). "Distance metric learning with application to clustering with side-information." In: *Advances in neural information processing systems*, pp. 505–512 (cit. on p. 56).

Yoon, Bo-Eun, Junsung Woo, Ye-Eun Chun, Heejung Chun, Seonmi Jo, Jin Young Bae, Heeyoung An, Joo Ok Min, Soo-Jin Oh, Kyung-Seok Han, Hye Yun Kim, Taekeun Kim, Young Soo Kim, Yong Chul Bae, and C. Justin Lee (2014). "Glial GABA, synthesized by monoamine oxidase B, mediates tonic inhibition." In: *The Journal of Physiology* 592 (Pt 22), pp. 4951–4968 (cit. on p. 163).

Yuste, Rafael (2008). "Circuit neuroscience: the road ahead." In: *Frontiers in Neuroscience* 2.1, p. 6 (cit. on p. 181).

Yuste, Rafael (2015). "From the neuron doctrine to neural networks." In: *Nature Reviews Neuroscience* advance online publication (cit. on pp. 5, 170).

Zeilhofer, Hanns Ulrich, Hendrik Wildner, and Gonzalo E. Yevenes (2012). "Fast Synaptic Inhibition in Spinal Sensory Processing and Pain Control." In: *Physiological reviews* 92.1, pp. 193–235 (cit. on p. 101).

Zhong, Lixian, Richard Y. Hwang, and W. Daniel Tracey (2010). "Pickpocket Is a DEG/ENaC Protein Required for Mechanical Nociception in Drosophila Larvae." In: *Current Biology* 20.5, pp. 429–434 (cit. on p. 98).

Zipursky, S. Lawrence and Joshua R. Sanes (2010). "Chemoaffinity Revisited: Dscams, Protocadherins, and Neural Circuit Assembly." In: *Cell* 143.3, pp. 343–353 (cit. on p. 100).

Zlatić, M, M Landgraf, and M Bate (2003). "Genetic specification of axonal arbors: *atonal* regulates *robo3* to position terminal branches in the *Drosophila* nervous system." In: *Neuron* 37, pp. 41–51 (cit. on p. 24).

Zlatić, Marta, Feng Li, Maura Strigini, Wesley Grueber, and Michael Bate (2009). "Positional cues in the *Drosophila* nerve cord: semaphorins pattern the dorso-ventral axis." In: *PLoS Biology* 7.6, e1000135 (cit. on p. 24).

Zubler, Frederic, Andreas Hauri, Sabina Pfister, Roman Bauer, John C. Anderson, Adrian M. Whatley, and Rodney J. Douglas (2013). "Simulating Cortical Development as a Self Constructing Process: A Novel Multi-Scale Approach Combining Molecular and Physical Aspects." In: *PLoS Comput Biol* 9.8, e1003173 (cit. on p. 179).

Zwart, Maarten F, Owen Randlett, Jan Felix Evers, and Matthias Landgraf (2013). "Dendritic growth gated by a steroid hormone receptor underlies increases in activity in the developing Drosophila locomotor system." In: *Proceedings of the National Academy of Sciences of the United States of America* 110.40, E3878–3887 (cit. on p. 78).