Steering Committee

Lou Scheffer, Harald Hess, Richard Fetter, Ian Meinertzhagen, Stephen Plaza, Reed George, Saul Kravitz, Ulrike Heberlein, Gerry Rubin

Research Staff

Visiting Scientists: Ian Meinertzhagen, Zhiyuan Lu, Kazunori Shinomiya Research Specialists: Pat Rivlin (Proofreader manager), Shinya Takemura, Stuart Berg (50%) Scientific Computing: Donald Olbris, Lowell Umayam, Rob Svirskas, Charlotte Weaver FlyEM Software and Algorithms: Bill Katz, Ting Zhao, Gary Huang Program manager: Stephen Plaza Proofreaders: Roxanne Aniceto, Lei-Ann Chang, Marielbis Garcia, Shirley Lauchie, Omotara Ogundeyi, Aya Shinomiya, Christopher Sigmund, Satoko Takemura, Julie Tran Dalhousie proofreading team: Jane Anne Horne (50%) and 2FTE proofreaders.

Collaborating Scientists

Mitya Chklovskii (Arjun Bharioke), Lou Scheffer (Stephen Plaza, Toufig Parag), Vijay Samalam, Harald Hess (Shan Xu, Ken Hayworth), Richard Fetter, Fred Hamprecht, Graham Knott

Funding Duration: 2 years

FlyEM Project Proposal

TABLE OF CONTENTS

1. Project Overview	. 1
2. Background	. 1
3. Project Progress	2
3.1. Progress in Reconstruction	2
3.2. Progress in Imaging	3
4. Project Research Plan	6
4.1. Specific Aims	6
4.2. Imaging	6
4.3. Sample Preparation	6
4.4. Identifying Gap Junctions in Connectomes	7
4.5. Continued Circuit Analysis of the Deep Optic Neuropiles, Medulla, Lobula and Lobula	
Plate	9
4.5.1. Synapse Morphometry of the Medulla 9	
4.5.2. Dense Reconstruction of the a-lobe of the Mushroom Body 10	
4.5.3. Dense Reconstruction of an Antennal Lobe Glomerulus 11	
4.5.4. Dense Reconstruction of the Central Complex 12	
4.6. Moving from Local to Global Reconstructions of the Drosophila Brain 1	12
4.7. Better Reconstruction Software 1	13
4.8. Cooperation with Other Labs 1	16
5. References 1	18
6. CVs and Websites of Project Members	20

1. Project Overview

The strategic objective of the FlyEM project team is to develop a fully detailed, cellular- and synaptic-resolution map of the central nervous system of *Drosophila melanogaster*, at both larval and adult stages. Simply having this "wiring-diagram" is necessary but not sufficient to understand how the fly's nervous system functions. We are, however, confident that the wiring diagram will be a foundational tool, necessary to develop that greater understanding, in much the same way that genomic sequence information has proved essential in enabling and accelerating studies of genetics, development and molecular and cellular biology.

In addition to their shorter-term applications to the neurobiology of *Drosophila*, in the longer term we expect that the imaging and computational techniques developed by this team will become applicable to ever larger problems in functional neurobiology, such as those posed by vertebrate nervous systems. Towards this end, our project complements and cooperates with other projects at Janelia, such as automated TEM, other reconstruction and tracing efforts, and the common software needed to make these projects affordable and practical. These projects, taken together, spearhead our vision of the role Janelia will play in developing technologies that open new fields of enquiry in the brain sciences.

2. Background

The FlyEM project was launched about five years ago building on prior work from the labs of Dmitri Chklovskii and Richard Fetter, as well as of Ian Meinertzhagen and other Janelia Farm Visiting Scientists. In the first few years we successfully reconstructed an array of lamina cartridges and a medulla column using a newly developed automated reconstruction pipeline for transmission electron microscopy (TEM) images. We also imaged and experimented with many other data sets, including a complete first instar larval brain sectioned and imaged at five tilts (a project now being continued in the Cardona lab).

The medulla reconstruction demonstrated both the feasibility and utility of a connectome at the synapse level, each of which had been doubted. It enabled the identification of the elementary motion detector circuit in the medulla, surmounting a 50-year old barrier in the analysis of motion sensing pathways. Nonetheless, although a huge advance over the purely manual methods previously used, these reconstructions demonstrated two major limitations to further advances. First, the z-axis resolution of TEM imaging, which is determined by an irreducibly minimal section thickness, is nevertheless an order of magnitude poorer than x,y resolution available by TEM, and insufficient to trace all the neurites in the neuropile through consecutive sections. Second, the rate-limiting step of this pipeline occurs when human proofreaders exhaustively scan segmented image stacks to verify the correctness of segmentation.

In order to scale up our reconstruction effort we therefore needed to improve image resolution to obtain more nearly isotropic voxel dimensions, and also to accelerate our proofreading rate. After considering several alternatives, we converged on using Focused Ion Beam Scanning Electron Microscopy (FIB-SEM) to achieve isotropic imaging resolution. The quality of images produced by FIB-SEM along with improvements in automated segmentation algorithms, new work flows, and new analysis software, enabled us to scale up the pipeline speed to reconstruct circuits of larger dimensions to greater completeness than were possible with TEM.

3. Project Progress

The past two years saw progress on many fronts – FIB imaging and reconstruction of seven columns of the medulla, and development of a next generation of imaging and reconstruction technology. These efforts are not independent; it was largely the problems encountered in the previous single-medulla column and the first instar projects that drove the development of this new technology.

Very briefly, the progress can be summarized as follows:

- We finished the single column reconstruction, including tracing into the lobular plate.
- We developed new software and techniques, and reconstructed seven medulla columns from FIB data. We continue to analyze connections to the neighboring columns.
- We dramatically increased the reliability of FIB imaging to the point where large samples are routinely imaged without loss.
- Using this more reliable FIB, we imaged several other samples of biologically interesting regions of increasingly larger size. These include samples including antenna lobes, a sample including portions of the medulla, lobula, and lobular plate, and many samples to test new staining and preparation methods.
- We developed the hot-knife technique to allow us to image larger volumes than possible by FIB alone.
- We acquired and used an Xradia CT unit to test preparation, verify samples, and aid in targeting selected regions within samples.

Each of these developments is discussed in more detail below.

3.1. Progress in Reconstruction

Although it was well underway at the last project review, we completed and published the single column reconstruction (1). This reconstruction, and publication, followed a somewhat unexpected path. While reconstructing M1-M10 of the medulla, we realized it would be very helpful to follow the axons of the T4 cells into the lobula plate. This would enable us to correlate the asymmetry in the arbors and anatomical receptive fields of the T4 input neurons with the direction sensitive layer in the lobula plate innervated by the terminal of the T4s. We were very

fortunate that in sectioning further, we could indeed collect images of T4 axons that could be traced into the lobula plate far enough to find the terminals of the T4s whose dendrites we had reconstructed in the medulla. We aligned and sparsely traced these axons, making possible the analysis of the motion detection circuit, thus closing in on a 50-year objective to identify the biological implementation of the Hassenstein-Reichardt elementary motion detector circuit. This experience illustrates the importance of combining dense reconstruction of a volume with sparse extension into neighboring neuropiles.

In addition to finishing the single column reconstruction, we used new software, combined with the new FIB imaging, to reconstruct a seven-column sample of the medulla. Consideration of the numbers of medulla cell types, and the average numbers of cells per column, led us to believe that seven columns should include representatives from most of the medulla's different cell types, as well as seven representatives of those types found in every column. Furthermore, the Rubin lab and FlyLight have already developed sparse lines for most of the cell types in medulla, and other groups at Janelia already examine in detail their response to visual stimuli, using behavioral screens, calcium imaging, and electrophysiology, and identify cell-specific RNA transcripts as part of the NeuroSeq project. Thus the medulla was and remains an ideal case for collaboration – the FlyLight and FlyEM can use each other's data as both a map and a cross check, NeuroSeq can identify likely modes of synaptic transmission, and the sparse lines already isolated can be used to drive expression in any cell type to facilitate testing of hypotheses about function derived from the connectome as well as to enable purification of cell-type-specific RNA samples for analysis by the NeuroSeq project.

New and extensively modified software was used for this seven-column reconstruction. This included methodology changes (identifying synapses first) and improvements to all other phases of reconstruction. These changes include new alignment algorithms, better segmentation, improved work flows concentrating on particular tasks, the ability to divide tasks across different proofreaders and geographical locations, and new analysis methods and techniques. The net result has been a roughly five-fold improvement in the speed of proofreading per effort expended.

3.2. Progress in Imaging

The demands of dense reconstruction, for reducing the proof reading burden, and accessing relevant biological targets have all driven different aspects of the technology used to acquire images. Specifically, to reliably identify the smaller processes near fly synapses we need to have 10nm or less pixel sizes, given that some features such as vesicles and the finest neurites can be as small as 30 nm. This applies to all dimensions, but the z dimension is the most challenging. Ultramicrotome sections are limited to 30 nm thickness, and blockface microscopy using a diamond knife can achieve 20 nm reliably, but FIB-SEM is the most mature technology to achieve better than 10 nm resolution. The weakness of FIB-SEM is that imaging rates are slower and it was initially thought that imaging volumes larger than (30 micron)³ were not possible. In

order to access biologically relevant sizes, however, we needed to image large self-contained neural circuits such as the optic lobe or mushroom body, which are on the order of a $(100 \text{ micron})^3$ volume, and we wanted at least a conceptual path forward to imaging the entire fly brain. This was far beyond the state of the art two years ago and overcoming these limits has been an important effort and recent accomplishment.

The volume imaged is given by: Volume = VoxelSize x VoxelRate x Duration. The voxel size is set as just described to $(8-10 \text{ nm})^3$. The voxel rate is likewise empirically determined by the reconstruction difficulty for different signal to noise ratios (SNRs). In our case of fly tissue we have found that a minimum signal to noise ratio of ~six is required. This SNR is in turn determined by the primary beam current, sample staining strength, and the physics of the electron scattering and their detection. We have just achieved a 3x improvement in acquisition rate by switching to the Zeiss Merlin SEM which can sustain a primary electron current of close to 8 nanoamps, versus 2 nanoamps of the standard SEM, with only tolerable compromise of x,y beam resolution.

The most significant improvements have been made in the last item of the expression, the duration of seamless acquisition. Initially we could acquire data for only a few days before an uncontrolled termination event, usually a FIB column failure. Now we routinely run for two to three months and only stop when the sample imaging has been completed. This required addressing a variety of interrupt issues: ion source reheat, utility failure (water, power, air, and temperature fluctuation), and microscope failure (focus, electrical, software, vacuum). With improvements and backups for the existing utilities, and with the transition to a new lab space with special environmental and power backup, we have decreased the frequency of these problems. Close monitoring of major FIB-SEM parameters – beam current, focus, and so on – enables us to shut down safely in the case of many remaining failure events. Finally a feedback scheme that controls the milling ion beam enables us to seamlessly restart without losing milling control that otherwise would result in a loss of 100 nm of sample thickness. The seamless restarting capability effectively removes any fixed volume limit due to interruptions.

As an example of improved capabilities, the FIB-SEM data set shown in Fig. 1, covers a major cross section of the medulla and full cross section of the lobula and lobula plate. It was taken over a three-month period and illustrates the size and quality of neuropile that can now be imaged. Several similar sized volumes of complete antennal lobes have also been imaged on a routine basis. Such images are typically taken of several samples to find the best possible stain, contrast, synapse clarity, and membrane integrity to minimize the large proofreading time investment. Two FIB-SEM machines, the original Zeiss NVision and the new FEI/Merlin (with 2-3 x throughput) are now in production mode for the various targets discussed below. In the second year from now we should be acquiring even larger volumes with more extended projections. Much of this will take advantage of the "Hot Knife" technology (see below) if the brain regions that need to be traced more span more than 100 microns, as in the case of the central complex or a more complete thickness of the lobula.



Figure 1. Medulla-Lobula-Lobular Plate sample

To increase the yield of the FIB-SEM data and to screen samples against internal defects such as non-uniform staining or cracks that could waste valuable FIB-SEM time, we have acquired an Xradia Versa for X-ray tomography of samples. This capability also allows us to define more closely the location and general structure of the brain to find a target of interest and optimize the trimming of the block to that target, e.g. the alpha lobe of the mushroom body. For this we use software to visualize the X-ray tomogram and align it to the 'Standard Model' of the fly brain obtained from confocal imaging.

The final constraint of FIB-SEM is the limited milling depth of roughly 100 microns, which prevents study of deep neuropil regions and processes. A "Hot Knife" technique, developed by Ken Hayworth, overcomes this barrier by cutting 20-100 micron thick sections. Each of these slabs is appropriate for FIB-SEM, and has near perfect surfaces that can be seamlessly stitched together in reconstruction. Numerous tissues, both mouse and fly, have been explored under different fixations, such as chemical and high-pressure freeze, and with different epoxy embeddings, Epon or Durcupan. A large cut surface data set across a full lobula has been generated to perfect any registration issues and test the scaling of this approach. This "Hot Knife" capability will be useful to access structures such as the complete mushroom body and even following some projections or the central complex. If a few remaining problems, such as tears near epithelial sheaths for HPF samples, can be addressed, we are hopeful this approach can be extended to an entire fly brain with further hardware scaling.

4. Project Research Plan

4.1. Specific Aims

- Imaging and imaging improvement (FY15-16 effort: 11%)
- Sample preparation, including both production samples and process improvement (FY15-16 effort: 7%)
- Identify, Image, and Assign Gap Junctions (FY15-16 effort: 3.5%)
- Circuit reconstruction (FY15-16 effort: 55%)
- New connectome analysis tools (FY15-16 effort: 3.5%)
- Software Infrastructure Data store and morphology analysis (FY15-16 effort: 7%)
- Improve reconstruction software (FY15-16 effort: 12.5%)

4.2. Imaging

We currently have two FIB-SEM machines, the original Zeiss NVision and the new FEI/Merlin, in which the Janelia group of Harold Hess integrated an ion gun from FEI with a Merlin SEM from Zeiss. The new machine has 2-3 x greater throughput, and will be in production mode for the various targets discussed below. With the Xradia Versa-based ROI targeting and this imaging capacity we are well prepared to acquire further images of many different targeted regions to support fly brain project over the next year. In a second year from now we should be acquiring larger volumes with more extended projections. Much of this will take advantage of the "Hot Knife" technology if the biological structures need to be traced more than 100 microns, as is the case in the central complex or a more complete thickness of the lobula.

We will also re-examine the tradeoff between imaging and reconstruction time. Imaging at higher resolution and better signal to noise is slower, but makes reconstruction easier and faster, and the overall tradeoff may be favorable. Previous use of this tradeoff was limited by the need to finish the sample before FIB failure, but with the recent improvements in reliability this will be re-visited.

4.3. Sample Preparation

The remaining path to a higher imaging rate is the heavy metal staining strength, of both the plasma membrane and synapses. Fly neural tissue presents many unique challenges and the staining strength of the preferred high pressure freezing/freeze-substitution prepared tissue, remains modest compared to what can be obtained in vertebrate tissue.

Given the potential to help reduce the heavy proof reading burden, there will be a continued effort to improve sample preparation methods. Explorations by Zhiyuan Lu, Richard Fetter and Graham Knott are ongoing to enhance preservation of extra cellular space, increase membrane staining, maintain pre- and post-synaptic density, achieve good contrast between these and the

cytosol, minimize membrane holes, and extend the quality over large volumes without allowing cracks to form. The "Hot Knife" method adds further demands and its compatibility with various protocols is summarized in Table 1. The final barrier to full fly brain sectioning is that HPF samples do not section well across an epithelial sheath. Interestingly, this is not an issue with chemical fixed samples. Alternative resins, infiltration protocols, and other strategies are being evaluated to address these remaining issues.

Overview of what works and what doesn't:



Table 1: Progress in staining, sample preparation, and hot-knife. Protocols are named according to the labs

 developing them – Knott, Lu, Fetter, and Mikula.

Richard Fetter and Loren Looger are working to make a more aldehyde resistant form of HRP in order to develop a reagent capable of making an electron dense signal with greatly enhanced ultrastructure compared with current HRP reagents. The fortified HRP will then be coupled to specific subcellular localization tags and driven by cell specific promoters, to provide an easy method to identify distal neurites in specific neuropils without having to reconstruct the entire cell(s), in preparations sufficient for connectome-quality imaging.

4.4. Identifying Gap Junctions in Connectomes

While our progress in generating a synaptic connectome may have been impressively successful, we have not systematically recorded sites of putative electrical transmission at gap junctions (GJs) between neurons. GJs in protostomes are encoded by a family of innexin channel proteins and *Drosophila* has eight innexin genes (2). Those in *C. elegans* have been well identified on structural grounds (3, 4), but the neurons of *C. elegans* are simple and tubular whereas those of *Drosophila* are highly branched, and candidate gap junctions exhibit far fewer clear

ultrastructural features than chemical synapses. Yet GJs can be equally influential in circuit dynamics (5). A major objective now is therefore to make progress with documenting the numbers and locations of gap junctions. These have been well documented at few sites, notably the lamina (6) and giant fibre terminals (7). We will develop criteria to identify the presence, location and numbers of gap junctions between identified neurons in specific circuits, and integrate these with data on dye coupling and innexin transcript expression, to develop and validate structural criteria for GJs at sites more generally. Current descriptions indicate that GJs are marked by areas of close membrane apposition having increased membrane density and linearity, sometimes accompanied by vesicles (as observed in the giant fibre).

We will use the lamina to develop methods for identifying gap junctions. There, GJs, form close appositions between photoreceptor terminals R1-R6. Each R-cell terminal forms about 15 such GJ contacts in two populations, one extending throughout the lamina that does not require the innexin gene shakB, and a second at distal levels, which does (8). Each contact has a diameter of about 250 nm and a cleft about 4 nm. Dye coupling from an injected to a non-injected R cell terminal excludes Lucifer Yellow (MW 443Da), but works with Biocytin (MW 372Da), so innexin pores must be permeable to small ions and Biocytin (Shaw, pers. comm.). GJs are formed by the Innexin channel protein family in *Drosophila (ogre, inx2-inx7, shakB)*, of which only *shakB, ogre*, and *inx2* have alleles that clearly express in the adult *Drosophila* visual system (9). While these numbers may be small, and some expression may be glial, heteromerisation (channels comprising different subunits) is common, leading to intercellular channels of homotypic (two hemichannels identical) or heterotypic (two hemichannels of differing molecular composition) composition (10).

Several approaches are possible to identify gap junctions. Two alternative possible methods, Mini-SOG (11) or immuno-EM of innexin epitopes, although highly touted, seem to us less promising. Although we will explore these possibilities (mini-SOG with Ng, at the Cambridge LMB) and innexin immuno-EM with Bauer, in Bonn, both methods would require a specific reagent for each innexin protein, and thus in turn successive preparations for all eight innexin genes.

Two additional approaches will also be tried. Most of the problems associated with finding GJs by ultrastructural criteria lie in knowing where to look. First we will develop Innexin-GFP fusions, or use the similar MiMIC based technology (12) in collaboration with the Card lab, and use these to identify using light microscopy where we may expect to see GJs between identified neurons. We will then use such data to refine our search images for GJs at EM level, especially the presence of membrane densities that might be diagnostic. As needed we will also develop HRP or other peroxidase fixation resistant variants that are being developed by Richard Fetter and Loren Looger to label GJs directly at EM level. Second, Richard Fetter has developed new sample preparation methods that preserve more of the extracellular space and thus make more obvious any sites of close apposition between neurons.

4.5. Continued Circuit Analysis of the Deep Optic Neuropiles, Medulla, Lobula and Lobula Plate.

Using a FIB image stack, we will continue to identify motion circuits arising from T4 and T5 cells. These local direction-selective motion sensing cells are of four subtypes, each with a terminal in one of the four directionally tuned strata of the lobula plate, where it provides input to well characterized wide-field lobula plate tangential cells (LPTCs). There the terminals directly connect to LPTC dendrites in the same stratum via cholinergic synapses to provide preferred direction excitation. T4 and T5 cells with opposite tuning terminate in the adjacent layer and provide feed-forward null direction input to the same LPTC via putative GABAergic or glutamatergic inhibitory neurons (13). Local microcircuits for these interactions have yet to be identified in detail and FIB imaging is well suited to uncovering them.

Anatomical receptive fields have been identified for T4 cell dendrites in the proximal medulla (1) and, less completely, for T5 cells in the distal medulla (14), but both reports are still incomplete. We will therefore reconstruct additional T-cells and their inputs from identified medulla cells to identify more closely the receptive field subcomponents contributed by each transmedulla (Tm) cell input, with particular focus on the range and variation in their vector angles between T-cells, and to trace T5 cells to lobula plate strata. This and the lobula plate microcircuits would be the final steps in identifying candidate motion-sensing pathways for fly vision.

The T-cell tracing will be undertaken by Kazunori Shinomiya, a new postdoctoral recruit from the Meinertzhagen lab who will arrive in September 2014. Kazunori will also analyze circuits in the deeper strata of the lobula. These are completely unknown to science, and involve the further processing for color (15) and higher-order visual features. Columnar neurons in the lobula segregate and project to a group of discrete optic glomeruli in the lateral protocerebrum (16). Eleven glomeruli in the posterior ventral, and seven in the posterior region of the lateral procerebrum each receive exclusive input from a single class of lobula columnar neuron (LCn). Comprehensive light level anatomy of the various LC cell types, including the generation of cell-type specific lines, exists from the efforts of the Rubin lab and FlyLight. Several of these LC cell types induce interesting behavior when activated, such as escape or backing up (Wu and Rubin, in prep), and the Reiser lab is characterizing their responses to visual stimuli using electrophysiology and imaging. Uncovering their inputs from Tm cells and lobula microcircuits, will help us achieve a major objective of FlyEM, to document the circuits between photoreceptor input and behavioral output.

4.5.1. Synapse Morphometry of the Medulla

Drosophila synapses can be identified by the presence of a T-shaped pre-synaptic density called a T-bar ribbon (17). Fly synapses are polyadic; post-synaptic partners can be identified by proximity and the presence of post-synaptic density (PSD). As a first step towards understanding synapse diversity across the medulla, we will use our FIB and TEM datasets to extract a number of morphometric parameters including: number of T-bars per synapse, minimum spacing between synapses, synapse size, and number of post-synaptic partners per T-bar. Comparisons will be made to photoreceptor synapses in the lamina and the larval NMJ (18, 19). This morphometric analysis will provide benchmarks for examining other regions of the fly brain.

4.5.2. Dense Reconstruction of the α -lobe of the Mushroom Body

The mushroom body is a high-order sensory integrative center of the fly's brain required for olfactory behavior (e.g., (20, 21)) that provides a substrate for learning and memory in the fly (22, 23). They are also critical for complex activities such as olfactory discrimination learning, courtship conditioning memory, context generalization in visual learning, and the control of walking (22, 24). Some information is available for the input region, or calyx, to which projection neurons from the antennal lobe ascend, forming large terminals that provide input to the tiny claw-shaped dendrites of intrinsic neurons called Kenyon cells (25, 26). The axons of the ~2,500 Kenyon cells then fasciculate to form a tract, the stalk or peduncle, before entering the lobes. Three specific classes of Kenyon cells constitute the mushroom body lobes: α/β , α'/β' , and γ neurons, the α -lobes receiving input from α and α' Kenyon cells. These lobes have the following advantages as initial targets for dense reconstruction: (1) terminals of their Kenyon cells are less branched than those in other lobes, making reconstructions easier; (2) their volume is smaller than the medial lobes; and (3) their cell composition is known at the light level and there are only a small number of cell types present in addition to the Kenyon cells ((27); Aso et al., in prep).

The $\alpha 3$, $\alpha 2$, $\alpha' 3$ and $\alpha' 2$ divisions occupy a volume of 30 x 30 x 50µm and we will image a 40 x 40 x 70 µm. This volume is well suited to dense reconstruction from a FIB stack. The library of cell types and GAL4 driver lines from FlyLight is essentially complete for the mushroom body (Fig 2). We plan to complete the dense tracing in the first year after the stack currently being imaged is collected. This volume is, we believe, sufficient to recognize neuron types identified from FlyLight from their segments reconstructed from FIB. We hope to be able to determine a number of facts about the local circuit structure such as: (1) what is the frequency of Kenyon cell to Kenyon cell synapses; (2) what fraction of the Kenyon cells make synapses with an output neuron; and (3) who are the targets of the dopaminergic modulatory neurons. We will recruit a postdoctoral fellow to undertake this work.



Figure 2. The full set of dopaminergic input cells and output cells innervating the $\alpha 3$, $\alpha' 3$, $\alpha 2$ and $\alpha' 2$ compartments of the MB lobes. These renderings from confocal microscope images illustrate the small number of extrinsic cells and cell types present. The number in parenthesis indicates the number of cells of each cell type. The total cell complement is small: there are only 6 types of output neurons (9 cells; all cholinergic) and three types of dopaminergic neurons (6 cells; 3 left and 3 right); the complete arbors of these cells in the MB lobes are confined to these compartments. Other broadly arborizing cells have some processes in these compartments: four SIF amide neurons, one GABA intrinsic neuron (MB-APL), one serotonin intrinsic neuron (MB-DPM), the terminals of two types of MB output neurons from other compartments (3 cells in total; 2 GABAergic and 1 glutamatergic), and the input axons from Kenyon cells (five types and ~1400 cells in total).

4.5.3. Dense Reconstruction of an Antennal Lobe Glomerulus

We will analyze the connectivity of one or more antennal lobe glomeruli, from the olfactory receptor neuron axons of the antennal nerve to their projection neuron targets that relay to the mushroom body and higher olfactory centers. The project scientists will be internal to the group – Ian Meinertzhagen and Pat Rivlin. We propose to work on the VI, DM2, or DM6 glomeruli, and characterize their synapses and neuronal arbors, matching these to published images. From comparisons with other insect species, we anticipate that primary olfactory receptor neuron afferent terminals will be presynaptic at sites with T-bar ribbons, but it may be that cellular branching patterns will not be structurally determinate, as they are in the optic lobe, and that cell types may therefore be identified less discretely. We have two existing image stacks, obtained

with different sample preparations, and will choose one glomerulus for detailed examination. (We could also acquire a new image set, if neither preparation is satisfactory, but from initial inspection the two existing samples look promising.) For example, glomerulus DM2 (28) has a volume of $6,000\mu m^3$ (29), roughly a 22 μm sphere, compatible with FIB analysis at a moderate scale. Its projection neurons exhibit classical conditioning (30), and a later option would be to analyze the structural basis for this conditioning, probably via local interneuron synapses, from FIB series taken before and after conditioning.

4.5.4. Dense Reconstruction of the Central Complex

The central complex is a region of the *Drosophila* brain concerned with locomotor behavior and spatial orientation memory, populated by neurons with delicately branched arbors that exhibit a high degree of morphological stereotypy. In these characteristics, they resemble the optic neuropiles, probably for the same reason that circuits in both signal at high temporal resolution. Various reasons to reconstruct those circuits prompt our choice. First, it is a brain region that underlies specific behaviors, such as locomotor behavior (31) and spatial orientation memory (32) that have attracted quantitative analyses. In addition, it is the main focus of the Jayaraman lab, which is undertaking functional imaging and electrophysiology on central complex cells in flies performing various behavioral tasks. In addition, the Rubin lab in concert with FlyLight is developing reagents at the level of light-microscopic anatomy, together with clean reporter lines suitable for analysis by NeuroSeq. Isolation and characterization of lines for the central complex is progressing very well under intense effort and should be complete in about a year, in good time for our proposed launch of an EM project. In addition several Janelia labs are screening these lines for specific behaviors: the Card Lab, Reiser Lab, Rubin Lab, Simpson Lab, and the Branson lab.

Dense reconstruction from a FIB image stack will be undertaken by Tanya Wolff who is a world expert on central complex cell types (and is currently performing the light level anatomical work), and thus ideally qualified to undertake and direct proof-reading of an image stack to be produced. Given the delicacy of the cells, we currently favor an analysis from a FIB stack to be collected during the first year, and given the size of the target and the widespread projection of relay neurons to different brain regions, we will complement this analysis with sparse tracing using the hot knife approach to span consecutive slices.

4.6. Moving from Local to Global Reconstructions of the Drosophila Brain

Until now, the fly's brain has been mostly treated to systems-based approaches applied to single neuropiles. The next challenge will be to extend our reconstructions using sparse proofreading to trace neurons that project between neuropiles. In a sense our work on the optic neuropiles has initiated this more global approach, but these are connected by a tract comprising axons from many cells of the same type that project topographically, and we must now move to tracing the projections of many different cell types that project through different diverging tracts. These tracts have been already identified rather comprehensively by light microscopy (e.g. (33, 34)).

Our analysis will depend upon a mixture of dense and sparse proofreadings of biological targets already identified above, the exact blend reflecting the number of different cell types and the range and diversity of their projections. Examples to be tackled include projection neurons to and from the mushroom body α -lobe, extrinsic neurons of the mushroom body calyx, and projection neurons of the antennal lobe glomerulus and central complex. The morphology of these neurons at the light level is well established through the work of many labs in the field, as well as the work of the FlyLight project. For long-range projections we will use the hot knife approach.

Starting with dense proofreading of the lobula and lobula plate, and the α -lobe of the mushroom body, we will undertake sparse tracing between neuropiles, aided if necessary by HRP labeling, existing maps of axon tracts, light-level morphology of the relevant neurons, and ultimately prepare for whole-brain reconstructions. This analysis will result in a seamless mixture of dense and sparse proof-read volumes that include targets for which we already have preliminary information (e.g. the lobula), to ones for which EM analysis will be totally novel (e.g. the central complex).

4.7. Better Reconstruction Software

The most fundamental problem we face is that reconstructions are still time-consuming, even with the improvements we have made so far. Although we completed the seven-column reconstruction in roughly the same time as the previous single column, it is still just a tiny fraction of the fly brain. Addressing this will require both scaling up proofreading and better software.

With improvements in segmentation, synapses are becoming a bottleneck. In our original TEM flow from 4 years ago, synapse and post-synaptic density (PSD) identification took only 10-20% of the time, and was performed completely manually. With improved images due to FIB-SEM, and improved segmentation, the time spent correcting segmentation has decreased dramatically, and synapse and PSD identification is now almost half the manual effort. We have made efforts to help the proofreaders with automation of this task, but this step still needs more attention. To give a rough scale to this problem, in our seven-column reconstruction there were about 50,000 T-bars (pre-synaptic sites) and 310,000 post-synaptic sites. Comparing this to the 120,000 minutes in a year of eight-hour days shows why this portion of the task alone took 2.2 person-years. Since this was only 7/800 of the full medulla, which itself is small fraction of the brain, it is clear this task needs attention. Even worse, at a density of one synapse per cubic micron, a cubic mm of brain (which the TEM effort at Janelia is not far from imaging) contains about 1,000,000,000 synapses. At one per minute, that would require 7,000 person years for this task alone. Clearly better automation is needed.

Our previous efforts highlight the following reconstruction challenges (in no particular order): (1) the intractability of manually annotating synapses; (2) fixing errors (both those that are manual and those arising from image segmentation) that result in the false body merging; (3)

coordinating segmentation-driven reconstruction among multiple proofreaders; (4) determining the completeness of the reconstruction; (5) the need for tools to quickly analyze and mine data from a connectome; and (6) focusing proofreading efforts to minimize the number of operations required to transform segmented data to an actual connectome. We suspect that future requirements to identify gap junctions and other ultrastructural features will demand specialized object recognition algorithms and proofreading/analysis software. Furthermore, continual improvements to boundary prediction and segmentation remain essential.

To this end, FlyEM proposes the following solutions:

We will dedicate at least two-three FTE to focus on segmentation and object recognition problems. In particular, we believe that it is possible to produce a synapse detector with accuracy that is 'close enough' to human performance, eliminating (or greatly reducing) the need for manual annotation. Based on the seven column reconstruction, if synapse detection can be automated it could eliminate roughly 30-50% of the reconstruction time. Perhaps, more importantly, such automation could be used to guide segmentation algorithms. We propose to develop new, connectomics-guided segmentation strategies that will exploit more relevant metrics. We also plan to continue fundamental research into EM image segmentation and look for synergies between deep learning techniques and other machine learning approaches that might require less training data. Attention will be given to algorithms that not only provide correct segmentation, but do so confidently (35). Through accurate confidence intervals, the proofreading efforts can be focused appropriately.

We believe that using biologically salient information such as synapses is critical to enhance the effectiveness of segmentation and subsequent proofreading. Therefore, last year we also exploited high-level biological prior knowledge through cell-shape matching and morphological analyses (Zhao, unpublished). By automatically skeletonizing our volumetrically reconstructed neurons, we successfully clustered similar cell types together and determined morphological attributes useful in connectome analysis. We will extend this work to be used as a quality control check for the correctness and completeness of a reconstructed neuron by matching to light microscopic data and to automatically create potential cell type taxonomies based on morphology.

At least half an FTE will be dedicated to ensure the smooth operation of our proofreading tools that will be used by several proofreaders for 100s of hours. We have learned that the smallest software inconveniences can translate into thousands of extraneous mouse clicks. In particular, we have observed tremendous challenges with "body splitting." Reconstruction strategies that use image segmentation typically over-segment the dataset because body composition (uniting divided profiles) is easier than body decomposition (splitting falsely merged profiles). Despite our efforts, image segmentation in the seven-column reconstruction still resulted in false merging in some cases, sometimes due to membrane holes from image preparation, and sometimes because proofreaders falsely linked bodies. Therefore, we plan to make significant user interface

and algorithmic advances to better support decomposition transformations. One possibility is exploiting the carving capabilities in Ilastik (36).

From a human resource and workflow perspective, we propose to improve both the accessibility and specialization required in our reconstruction methods. There are some tracing operations that require very experienced proofreaders or postdocs, such as analyzing the biological structure of a neuron reconstructed to near completion. Some actions require less expertise, such as a yes/no decision on whether two regions are connected or not. We plan to ensure that domain experts or experienced proofreaders focus their efforts on only the hardest aspects of the reconstructions and those with the most impact. At the same time, we will try to simplify other proofreading tasks, thereby making them more accessible to a larger resource pool. Furthermore, we will extend the well-defined dense proofreading workflows developed in the seven-column reconstruction to sparse proofreading (currently done in a more ad-hoc manner) to help solve specific biological questions.

As our reconstruction efforts become larger, both in physical size and number of people, we require an appropriate data infrastructure to orchestrate reconstruction efforts. To this end, we started work on a Big Data solution, DVID (distributed, versioned, image, datastore) (Katz, unpublished). The goal of DVID is to facilitate collaboration and coordination between different proofreaders and provide a common interface to our image data. Through DVID, we should be able to support concurrent reconstruction efforts throughout the world. This will overcome a current limitation - we now support concurrency through explicit partitioning and often have only a single proofreader handle the integration of these disjoint partitions. Furthermore, the provenance of the proofreaders' actions is not explicitly tracked in many cases making it hard to determine the origin of modifications. A tool like the current version of CATMAID is inadequate in our domain, since it neither supports versioning of data, nor handles our segmentation-driven methodology. The coordination of segmentation-driven proofreading is a harder problem than coordinating skeletonization efforts. The benefits of having a distributed and versioned datastore for reconstruction go beyond the ability to incorporate proofreaders at distributed sites. The datastore API will form the foundation for an ecosystem that different EM reconstruction teams can exploit. We hope that such an ecosystem will facilitate tool reuse among teams. We also plan to use this environment to help enable segmentation competitions. In addition to FlyEM, both CATMAID and Ilastik developers have already explored the DVID interface successfully in sample applications.

Based on the seven-column reconstruction, we have identified several high priority targets for improvement that we are working on now, and will continue to work on over the next work period. These include machine learning improvements: connectome-based segmentation and fully-automated synapse detectors. Quick training is very important, particularly since we are experimenting with sample preparation and working on new areas of the brain. Uncertainty

estimation (including our Heidelberg collaboration) is a key to reducing the human effort required for proofreading. Toufiq Parag has been concentrating on these issues.

It has also become clear from the seven column connectome that just reporting the raw connectome is not enough, and that a new generation of analysis tools will be needed to understand and analyze larger connectomes. This includes operations such as comparing neurons of the same cell type, integrating NeuroSeq (and other sources of transmitter and receptor types) data, visualizing receptive fields (where the inputs are physical maps), comparison of closely related circuits (such as columns with pale and yellow photoreceptors in the medulla), cell and connection statistics and correlations, and so on. This additional analysis software is in development for the seven-column reconstruction but will remain an ongoing project. Some portions of this work will need to be re-done for other portions of the brain. For example the medulla is divided into layers, and the software assigns features to layers both to help with user understanding, and for compatibility with the existing literature. However, the central brain is divided differently, in a variety of less stereotyped and less well understood ways; for this reason, the comparison with light level data will be particularly helpful.

4.8. Cooperation with Other Labs

Cooperation with other labs is a critical requirement to augment and interpret the connectome. This is because the neurotransmitters used and the receptor types are not identifiable from EM. This means the strength, kinetics, and even the sign of synapses are not determined by conventional EM alone. Furthermore, gap junctions are difficult (at best) to see in the EM preparations used for reconstruction. Therefore the intended use of a connectome, to determine function or principles of neural circuit operation, requires additional information.

The NeuroSeq project is currently determining how much RNA of each gene is expressed in each cell type. Each neurotransmitter has an associated synthesis pathway, and genes regulating that pathway are expected to be highly expressed in all cells using that neurotransmitter, providing an unambiguous determination of transmitter type. This expectation has been verified in the lamina, using cases where the transmitter was already known or reasonably anticipated. Cooperation between NeuroSeq and EM works both ways, since EM provides the number of presynaptic sites for each cell type. An expression level that is proportional to the number of presynaptic sites, and not just high, provides even stronger evidence that the proposed transmitter type is correct.

NeuroSeq may also be helpful for determining receptor types, though there are several difficulties here. Many cells can receive more than one transmitter, and Neuroseq will not show which receptor is used where. Furthermore, receptors genes are expressed at much lower levels than transmitters, leading to greater ambiguity and potential problems with identification. Compounding these problems, there is speculation (Spitzer, N., UCSD) that most cells express almost all receptor genes at some low level.

NeuroSeq has already analyzed the cell types of the lamina, and is now working on a few selected cell types of the medulla. Once the technical bugs are worked out, the preferred order of sequencing (from FlyEM point of view) would be to finish the medulla (where two connectomes are already available), then the mushroom body, then the central complex. The necessary specific GAL4 drivers to support this effort are either ready or being generated by FlyLight in collaboration with the Rubin Lab. It would make excellent sense to have at least one common representative on the FlyEM and NeuroSeq steering committees.

We will also depend heavily on continued cooperation with FlyLight and other groups at Janelia, in particular biological experts to help with reconstruction and the behavioral analysis groups who form our 'customers' and help interpret results. This interaction has been strong in the optic lobe, mostly with Aljoscha Nern of the Rubin lab/FlyLight for anatomy and Michael Reiser for behavior. Also, we expect Kasunori Shinomiya will arrive from Dalhousie as a local lobula expert. We will need similar coordinated efforts in other brain areas for best results. We believe we are in good shape here. The mushroom body lines are essentially complete, with Yoshi Aso (Rubin Lab) a local biological expert. Lines for the central complex should be complete within the year due to the efforts of Tanya Wolff/Rubin lab and FlyLight, and Vivek Jayaraman and his lab are focused on physiology, functional imaging and circuit mechanisms of this structure.

Other groups, both inside and outside Janelia, are working on optical methods that could help verify and extend our results. Larry Zipursky of UCLA is working on independent verification of synapse counts and may provide a way to label gap junctions independent of attempts at Janelia. The Lavis lab at Janelia has developed methods for uncaging a fluorescent dye, using a genetically encoded orthogonal enzyme, so that it becomes small enough to go through a gap junction (37). He is actively trying to adapt this system to the *Drosophila* brain. Ryan Williamson of the Card Lab here at Janelia is also working on optical methods (based on MiMIC, (12)) of identifying gap junctions. In the optic lobes, it seems likely that the combination of gap junctions into the connectome without needing to visualize them in EM. (We are running computational experiments now to see if this is possible.) The outcome strongly depends on the fraction of cells within each area that contain gap junctions, which is currently unknown but may be determinable from NeuroSeq). In other less stereotyped regions of the brain this is more difficult, so the ideal situation would still be to see gap junctions in EM if at all possible.

5. References

- 1. Takemura SY, *et al.* (2013) A visual motion detection circuit suggested by Drosophila connectomics. *Nature* 500(7461):175-181.
- 2. Phelan P (2005) Innexins: members of an evolutionarily conserved family of gapjunction proteins. *Biochimica et biophysica acta* 1711(2):225-245.
- 3. Jarrell TA, *et al.* (2012) The connectome of a decision-making neural network. *Science* (*New York, N.Y.*) 337(6093):437-444.
- 4. White JG, Southgate E, Thomson JN, & Brenner S (1986) The structure of the nervous system of the nematode Caenorhabditis elegans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 314(1165):1-340.
- 5. Rela L & Szczupak L (2004) Gap junctions: their importance for the dynamics of neural circuits. *Molecular neurobiology* 30(3):341-357.
- 6. Shaw SR (1984) Asymmetric distribution of gap junctions amongst identified photoreceptor axons of Lucilia cuprina (Diptera). *Journal of cell science* 66:65-80.
- 7. Blagburn JM, Alexopoulos H, Davies JA, & Bacon JP (1999) Null mutation in shaking-B eliminates electrical, but not chemical, synapses in the Drosophila giant fiber system: a structural study. *The Journal of comparative neurology* 404(4):449-458.
- 8. Shimohigashi M & Meinertzhagen IA (1998) The shaking B gene in Drosophila regulates the number of gap junctions between photoreceptor terminals in the lamina. *Journal of neurobiology* 35(1):105-117.
- 9. Stebbings LA, *et al.* (2002) Gap junctions in Drosophila: developmental expression of the entire innexin gene family. *Mechanisms of development* 113(2):197-205.
- 10. Lehmann C, *et al.* (2006) Heteromerization of innexin gap junction proteins regulates epithelial tissue organization in Drosophila. *Molecular biology of the cell* 17(4):1676-1685.
- 11. Shu X, *et al.* (2011) A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS biology* 9(4):e1001041.
- 12. Venken KJ, *et al.* (2011) MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. *Nature Methods* 8(9):737-743.
- 13. Mauss AS, Meier M, Serbe E, & Borst A (2014) Optogenetic and pharmacologic dissection of feedforward inhibition in Drosophila motion vision. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 34(6):2254-2263.
- 14. Shinomiya K, *et al.* (2014) Candidate Neural Substrates for Off-Edge Motion Detection in Drosophila. *Current biology : CB*.
- 15. Otsuna H, Shinomiya K, & Ito K (2014) Parallel neural pathways in higher visual centers of the Drosophila brain that mediate wavelength-specific behavior. *Frontiers in neural circuits* 8:8.
- 16. Otsuna H & Ito K (2006) Systematic analysis of the visual projection neurons of Drosophila melanogaster. I. Lobula-specific pathways. *The Journal of comparative neurology* 497(6):928-958.
- 17. Prokop A, Meinertzhagen, I. A (2006) Development and structure of synaptic contacts in Drosophila. *In Seminars in cell & developmental biology* 17(1):20-30.
- 18. Meinertzhagen IA (1996) Ultrastructure and quantification of synapses in the insect nervous system. *Journal of neuroscience methods* 69(1):59-73.

- 19. Meinertzhagen IA, and X. Hu (1996) Evidence for site selection during synaptogenesis: The surface distribution of synaptic sites in photoreceptor terminals of the flies Musca and Drosophila. *Cellular and molecular neurobiology* 16(6):677-698.
- 20. Davis RL (1993) Mushroom bodies and drosophila learning. *Neuron* 11(1):1-14.
- 21. Heisenberg M, Borst A, Wagner S, & Byers D (1985) Drosophila mushroom body mutants are deficient in olfactory learning. *Journal of neurogenetics* 2(1):1-30.
- 22. Heisenberg M (2003) Mushroom body memoir: from maps to models. *Nature reviews*. *Neuroscience* 4(4):266-275.
- 23. Keene AC & Waddell S (2007) Drosophila olfactory memory: single genes to complex neural circuits. *Nature reviews. Neuroscience* 8(5):341-354.
- 24. Zars T, Fischer M, Schulz R, & Heisenberg M (2000) Localization of a short-term memory in Drosophila. *Science (New York, N.Y.)* 288(5466):672-675.
- 25. Butcher NJ, Friedrich, A. B., Lu, Z., Tanimoto, H., & Meinertzhagen, I. A. (2012) Different classes of input and output neurons reveal new features in microglomeruli of the adult Drosophila mushroom body calyx. *Journal of Comparative Neurology* 10(520):2185-2201.
- 26. Yasuyama K, Meinertzhagen IA, & Schurmann FW (2002) Synaptic organization of the mushroom body calyx in Drosophila melanogaster. *The Journal of comparative neurology* 445(3):211-226.
- 27. Tanaka NK, Ito K, & Stopfer M (2009) Odor-evoked neural oscillations in Drosophila are mediated by widely branching interneurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29(26):8595-8603.
- 28. Laissue PP, *et al.* (1999) Three-dimensional reconstruction of the antennal lobe in Drosophila melanogaster. *The Journal of comparative neurology* 405(4):543-552.
- 29. Devaud JM, Acebes A, & Ferrus A (2001) Odor exposure causes central adaptation and morphological changes in selected olfactory glomeruli in Drosophila. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 21(16):6274-6282.
- 30. Yu D, Ponomarev A, & Davis RL (2004) Altered representation of the spatial code for odors after olfactory classical conditioning; memory trace formation by synaptic recruitment. *Neuron* 42(3):437-449.
- 31. Strauss R (2002) The central complex and the genetic dissection of locomotor behaviour. *Current opinion in neurobiology* 12(6):633-638.
- 32. Neuser K, Triphan T, Mronz M, Poeck B, & Strauss R (2008) Analysis of a spatial orientation memory in Drosophila. *Nature* 453(7199):1244-1247.
- 33. Chiang AS, *et al.* (2011) Three-dimensional reconstruction of brain-wide wiring networks in Drosophila at single-cell resolution. *Current biology : CB* 21(1):1-11.
- 34. Yu HH, *et al.* (2013) Clonal development and organization of the adult Drosophila central brain. *Current biology : CB* 23(8):633-643.
- 35. Plaza SM, Scheffer LK, & Saunders M (2012) Minimizing manual image segmentation turn-around time for neuronal reconstruction by embracing uncertainty. *PloS one* 7(9):e44448.
- 36. Sommer C, Straehle, C., Kothe, U., & Hamprecht, F. A (2011) Interactive learning and segmentation toolkit. *IEEE*:230-233.
- 37. Tian L, *et al.* (2012) Selective esterase-ester pair for targeting small molecules with cellular specificity. *Proceedings of the National Academy of Sciences of the United States of America* 109(13):4756-4761.

6. CVs and Websites of Project Members

CVs for the following people are included in the Appendix:

- Gary Huang: Image segmentation and object recognition algorithms.
- William Katz: Algorithm designer, visualization specialist, and software architect of Big Data solutions.
- Zhiyuan Lu: Image preparation specialist for FIB-SEM.
- Toufiq Parag: Image training and segmentation; reconstruction software solutions.
- Stephen Plaza: Program manager, lead algorithm and software developer.
- Patricia Rivlin: Chief biologist and manager of proofreader task force.
- Kazunori Shinomiya: Expert biologist to trace neurons particularly in the lobula.
- Shinya Takemura: Expert biologist that will help in tracing neurons particularly in the medulla and help coordinate tracing between proofreaders.
- Ting Zhao: Morphological analysis software and quality-control algorithms in reconstruction.

Webpages for Steering Committee Members:

Harald Hess	http://janelia.org/lab/hess-lab http://www.hhmi.org/scientists/harald-f-hess
Richard Fetter	http://janelia.org/people/scientist/richard-fetter
Ulrike Heberlein	http://janelia.org/lab/heberlein-lab http://www.hhmi.org/scientists/ulrike-heberlein
Ian Meinertzhagen	http://www.dal.ca/faculty/science/psychology_neuroscience/faculty- staff/our-faculty/ian-meinertzhagen.html
Gerald Rubin	http://janelia.org/lab/rubin-lab http://www.hhmi.org/scientists/gerald-m-rubin
Louis Scheffer	http://www.janelia.org/lab/scheffer-lab http://www.hhmi.org/scientists/louis-k-scheffer
Reed George	http://www.janelia.org/people/administration/reed-george
Saul Kravitz	http://janelia.org/people/research-resources-staff/saul-kravitz
Stephen Plaza	http://www.janelia.org/team-project/fly-em http://www.janelia.org/people/scientist/stephen-plaza