

Manuscript Details

Manuscript number	COIS_2017_98_R1
Title	Learning from connectomics on the fly
Short title	Learning from connectomics on the fly
Article type	Review article

Abstract

Parallels between invertebrates and vertebrates in nervous system development, organisation and circuits are powerful reasons to use insects to study the mechanistic basis of behaviour. The last few years have seen the generation in *Drosophila melanogaster* of very large light microscopy data sets, genetic driver lines and tools to report or manipulate neural activity. These resources in conjunction with computational tools are enabling large scale characterisation of neuronal types and their functional properties. These are complemented by 3D electron microscopy, providing synaptic resolution data. A whole brain connectome of the fly larva is approaching completion based on manual reconstruction of EM data. An adult whole brain dataset is already publicly available and focussed reconstruction is under way, but its 40x greater volume would require ~ 500-5000 person-years of manual labour. Nevertheless rapid technical improvements in imaging and especially automated segmentation will likely deliver a complete adult connectome in the next 5 years. To enhance our understanding of the circuit basis of behaviour, light and electron microscopy outputs must be integrated with functional and physiological information into comprehensive databases. We review presently available data, tools and opportunities in *Drosophila*. We then consider the limits and potential of future progress and how this may impact neuroscience in rich model systems provided by larger insects and vertebrates.

Corresponding Author	Gregory Jefferis
Order of Authors	Philipp Schlegel, Marta Costa, Gregory Jefferis

Submission Files Included in this PDF

File Name [File Type]

HIGHLIGHTS.docx [Highlights]

MANUSCRIPT.docx [Manuscript File]

Fig1.pdf [Figure]

Fig2.pdf [Figure]

CONFLICT_OF_INTEREST.pdf [Conflict of Interest]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

- Whole brain EM connectomics will revolutionise neuroscience
- Insects are at the forefront of this revolution
- Mapping between light and EM image data can integrate anatomy and function
- Quantitative definitions of cell type will promote experiments and communication
- Tools and resources that integrate big neuroscience data are critical for neurobiologists

Learning from connectomics on the fly

Authors

Philipp Schlegel^{1*}, Marta Costa^{1*} and Gregory S.X.E. Jefferis^{2,1}

Affiliations

¹*Drosophila* Connectomics Group, Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ UK

²Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK

*Equal Contributions

Corresponding authors

Gregory Jefferis: jefferis@mrc-lmb.cam.ac.uk

Marta Costa: mmc46@cam.ac.uk

Abstract

Parallels between invertebrates and vertebrates in nervous system development, organisation and circuits are powerful reasons to use insects to study the mechanistic basis of behaviour. The last few years have seen the generation in *Drosophila melanogaster* of very large light microscopy data sets, genetic driver lines and tools to report or manipulate neural activity. These resources in conjunction with computational tools are enabling large scale characterisation of neuronal types and their functional properties. These are complemented by 3D electron microscopy, providing synaptic resolution data. A whole brain connectome of the fly larva is approaching completion based on manual reconstruction of EM data. An adult whole brain dataset is already publicly available and focussed reconstruction is under way, but its 40x greater volume would require ~ 500-5000 person-years of manual labour. Nevertheless rapid technical improvements in imaging and especially automated segmentation will likely deliver a complete adult connectome in the next 5 years. To enhance our understanding of the circuit basis of behaviour, light and electron microscopy outputs must be integrated with functional and physiological information into comprehensive databases. We review presently available data, tools and opportunities in *Drosophila*. We then consider the limits and potential of future progress and how this may impact neuroscience in rich model systems provided by larger insects and vertebrates.

1 Introduction

Insects provide ideal model systems to obtain mechanistic understanding of the brain from the molecular basis of development, through circuit organisation and function to behaviour. There is a long history of studying invertebrate nervous systems that goes back beyond Cajal [1–3]. Insect brains encode efficient neural solutions matched to the animal's neuroethological context [4]. Experimental advantages include accessible preparations, the numerical simplicity of the brain and the existence of identified cells, that can be studied from one animal to the next.

We believe that basic principles of insect nervous systems are shared with other species including vertebrates, due to shared computational challenges and deep conservation in basic neuronal hardware and developmental genetics [5]. Indeed there are now numerous examples of both shared principles of circuit organisation as well as specific cases of circuit mechanisms. For example the shared glomerular organisation of the first olfactory relay in insects and mammals has long been recognised [6], while more recently it has been proposed that higher olfactory processing and the insect mushroom bodies might share organisational principles with learning centres in mammalian brains [7,8]. Pioneering studies in insects (especially *Calliphora*) have had a historic impact on the study of motion vision; however recent studies in *Drosophila* have revealed striking examples of shared circuit mechanisms including the separation of ON and OFF motion pathways and the existence of cells tuned to four cardinal directions [9]. Whether these parallels are due to evolutionary conservation or convergence remains uncertain, but in some cases homologous genes argue for conservation. For example at the interface between the nervous system and metabolism, pathways with conserved peptide/hormone signals control analogous body systems e.g. Leptin/Upd2 [10], NPY/NPF [10,11] or NMU/hugin [12].

We believe that connectomics, computational neuroanatomy and molecular genetics combined with the traditional strengths of invertebrate preparations should enable the neurobiology of *Drosophila*, in particular, to have a major impact on the whole of neuroscience over the next 5-10 years. This will be the focus of our review. However we believe strongly that many of the technological advantages we discuss are already, or will increasingly become, accessible in other insects and in vertebrates.

2 History of Brain Maps

An essential enabling step in obtaining a mechanistic understanding of behaviour is to describe the neurons involved and their potential connections. The Golgi method as applied by Cajal [1] allowed for the first time the observation of the intricate morphology of neurons and glia. More importantly, it revealed how varied neuronal morphologies can be, and how structure and function might be linked.

Comprehensive studies of cell types within a brain region have been critical drivers of experimental work. Electron microscopy can in theory reveal all the neurons within a brain, but the labour involved meant that for many years the complete reconstruction of the *C. elegans* nervous system [13] was a one-off. Therefore, the only widely applied methods have depended on sparse labelling, which stochastically labels subsets of cells [14]. Recently, the combination of molecular genetic labelling and high resolution confocal microscopy [15] has proven particularly effective in *Drosophila*, generating maps with single cell resolution of various brain regions and even spanning the whole brain [16–20]. 3D image registration enables direct integration and comparison of high resolution confocal microscopy data from different experiments [19,21,22]; this also allows successive additions to existing maps. Furthermore, functional data from targeted neuron manipulation and physiology can enrich these mostly anatomical brain maps, making it possible to develop and test circuit hypothesis [23–25].

Huge amounts of light microscopy-based neuronal image data are now available for the fruit fly brain. This raises new challenges. Neuroscientists need user-friendly tools to visualise and query these data and especially approaches to integrate across datasets and link neuroanatomical data with data on

neuron function, behaviour, etc. These are very significant practical issues when trying to understand complex brain data and we discuss them in detail below. However, there are major constraints in the resolution of light microscopy data when generating circuit maps. The labelling of the very fine and intricate neurites of single neurons is difficult, and small processes are often missing [26,27]. Crucially, however, connectivity between neurons can only be inferred, relying on overlap [16,19,20,28] or molecular proximity labelling systems like GRASP [29,30]. Recent advances that aim to obtain synaptic resolution with light microscopy by physically expanding the specimen up to 20x [31] in combination with approaches to label synaptic specialisations may eventually be fruitful. For the time being, this approach is likely to increase the density at which sparsely labelled specimens can be resolved. However, this will not allow us to achieve a full reconstruction of neurons and their connections.

3 Challenges and Rewards in Connectomics

Mapping of neural circuits at synaptic level still requires electron-microscopy (EM), due to its ability to resolve nanometer scale synaptic contacts. Comprehensive synapse-level brain maps remain rare: to date, only two connectomes, *C. elegans* [13] and *Ciona intestinalis* [32], exist and the cost-benefit ratio of such maps is controversial. The complete wiring diagram of *C. elegans*' 302 neurons has been available for more than 30 years, but its potential to explain behaviour was initially limited (reviewed in [33,34]. In brief, the lack of information beyond neural connectivity (e.g. neurotransmitter identities, neuronal activity or behavioural significance of individual neurons) made it difficult to understand more than a few basic modules of the worm's connectome. This called the cost-benefit of the then enormous investments to acquire the map into question. Since then, huge technical and much conceptual progress has been made and today generation of another *C. elegans* connectome would be a matter of weeks, not years. Here, we will lay out reasons why we believe the time is ripe for more such maps.

Given their relative small size and neuron number, it appears likely that insect brains will among the next comprehensive connectomes (Figure 1A). However, generation of synaptic resolution connectomes still requires the allocation of considerable resources over extended periods of time. This long-term nature of large EM projects clashes with the short-lived "publish-or-perish" attitude in modern science. Encouragingly however, a number of large-scale projects have produced partial connectomes in vertebrates [35,36] and invertebrates [26,27,37,38] in recent years. The latter are part of efforts to map the brains of larval and adult *Drosophila*, respectively (Box 1 - Insect EM datasets).

The larval EM project, hosted by Dr. Albert Cardona (Janelia Research Campus), is arguably the most successful connectomics effort to date. Begun in 2012, it has involved >15 different labs and > 50 researchers. In the last 2 years multiple high-profile publications, have presented substantial advances on a range of topics: neuron morphology [39], olfaction [40], learning and memory [37], sensory integration [41], decision making [42], neurotransmission [12] and motor systems [43,44]. Multiple factors contribute to the success of this effort: a whole central nervous system EM volume was acquired, allowing reconstruction of all neurons within a single dataset. Instead of a monolithic approach aiming at completion from the beginning, reconstruction was tackled in a modular, circuit-focused manner by splitting up efforts across a large network of collaborating labs with differing interests. These labs brought in necessary expertise in a range of different fields (olfaction, motor systems, learning and memory, etc.) and ensured that reconstruction was from the start biology-guided. This, together with an in-house catalogue of single cell morphologies (obtained from enhancer driver screens), helping to identify EM skeletons and link them to the existing literature, as well as the use of cell-type selective genetic reagents to address circuit function [41], were key initial factors in the project's success.

The same key points are likely to also hold true for even larger ongoing projects such as the adult *Drosophila* EM project. The first complete volume acquired by the group of Davi Bock at Janelia over ~ 16 months is about 45 times larger and has 10 times more neurons than the larva [26]. The image data

for this adult female whole brain volume are already publicly available (see Box 1) and at least 10 groups are already carrying out collaborative reconstruction. Our *Drosophila* Connectomics Group in Cambridge in collaboration with groups at Janelia and Oxford, is focussing on the olfactory-memory circuits, especially the mushroom body. We are already generating hugely valuable data by focussing efforts on tracing of neurons of interest. Our experience of manual reconstruction in this volume is that one can make sudden breakthroughs from just a few days or weeks of work, such as the discovery of previously unknown connections for neurons of interest. Comprehensive reconstruction on the other hand is extremely slow. We have seen that ~4 person-years of tracing has resulted in about 0.7m of neuronal cable (roughly 0.1mm/hour). We calculate based on [40] that *Drosophila* neuropil contains at least 10 km of arbour per mm³; we therefore estimate that the adult brain contains about 100m of cable to reconstruct. This serves to underscore that most of the time (and money) associated with connectomics is in the annotation of data – tracing out neurons and synapses – rather than initial acquisition of serial EM sections [45]. Figure 1B provides some estimates of these quantities in relation to brains of interest. These numbers represent idealised reconstruction times and do not take overhead (e.g. for planning and analysis) into account which we estimate to at least double the time required.

In the long term, annotation must be automated to make connectomics scale. Automated segmentation using machine learning has the potential to speed up data annotation by orders of magnitude [46–48]. However at the moment, it requires higher-resolution, isotropic data to be more effective than manual reconstruction by humans [27,38]. This has important implications when selecting an EM technique for data acquisition. Briefly, ssTEM offers superior x/y resolution and acquisition speed whereas SBEM and FIB-SEM sacrifice speed for improved z resolution (Box 1, reviewed in [49,50]). Ultimately, the method of choice depends on two key factors: the z resolution required to resolve fine dendritic processes and the absolute dimensions of the sample of interest: at current speeds, acquisition times for 1 mm³ (the size of a honey bee brain [51], Figure 1A) range between 400 years on a single FIB-SEM [38], ~30 years on an automated camera array ssTEM [26,38]. Multibeam SEMs could theoretically reduce this to 6 months [52]. Assuming a cable density of about 4.6km/mm³ (mouse cortex, [53,54]) and a manual reconstruction speed of about 0.1mm/h, annotation of a 1 mm³ volume would take 22,115 person-years to complete.

This sample calculation simply illustrates the current realistic limits in connectomics research and is not to be mistaken for pessimism for the future. In fact, a considerable number of densely (all neurons) and sparsely (only circuits of interest) reconstructed EM volumes have been published in recent years (Figure 1C). We expect many connectomic studies focussed on particular brain regions to be produced in the next 1-2 years using the two whole brain EM datasets already available for the *Drosophila* larva and adult. However by combining higher resolution FIB-SEM and advances in machine learning, it is quite feasible that we will obtain a whole adult brain connectome in the next 5 years. Whether dealing with sparse, focussed or complete connectomes, data integration poses major practical challenges to biological interpretation. This is our next topic.

4 Towards Comprehensive Databases

Numerous terabyte scale *Drosophila* neuroanatomical datasets have been generated in the last few years using light microscopy. These datasets fall into two major classes: single cell morphologies (23,000 images: FlyCircuit [20,55], that are sufficient to define a neuron type, or genetic driver lines (13,500 images: Janelia GAL4 and LexA lines [56,57], and Dickson VT lines (B. Dickson, personal communication)) (Box 1) that allow researchers to manipulate specific sets of neurons. Although these represent large efforts, we expect many more of these datasets to be produced in the coming years, as the mapping of the estimated 100,000 neurons of the adult fly brain gathers pace. The generation of more datasets using different methods will also help to overcome the labelling bias (i.e., some individual neurons or neurons types are much more frequently labelled) observed with the current ones. The use of intersectional labelling systems [58], in particular, which allow the targeting of a small number of neurons, will both benefit from and contribute to this increase [23].

In order for the potential advantages of these datasets to be realised, it is necessary to develop user-friendly tools that allow cross-querying and linking to other relevant data (morphology, function and expression for example). In genomic datasets, the organising unit is normally the gene; for brain mapping efforts, we suggest that neuron type is the best fit. Although there is no clear consensus definition of a neuron type [59], operationally this refers to a single or homogeneous group of neurons (from a few to hundreds) marked by features that are reproducibly found from one brain to the next. In practise, scientists seem to cope relatively well without formal definitions of neural type – certainly in the case of identified neurons present at the level of one per brain hemisphere. Recently, efficient neural similarity and search tools have been developed for single traced neurons [60–62]. We also note that there has been significant progress on the more difficult problem of matching traced neurons against 3D images potentially containing many labelled neurons [62,63]. In particular the NBLAST tool [62] has been validated for the morphological definition of cell types on a database of 16,000 single *Drosophila* neurons [20,62]; such approaches may eventually automate classification [64]. In the near future, genome-wide gene expression profiles from single cells [65] will inform the current morphology-focused classification of neuron types.

As mentioned earlier, integrating 3D data can now routinely be achieved with image registration [19,21,22]; the spatial stereotypy of insect brains (reported at 2-3 μm per axis in *Drosophila* [19]) makes this particularly effective. However, even large data providers do not always make their data or metadata available in a format that can be easily shared, integrated, or linked to other data. The challenge, is then, one of resources. Most neuroscientists cannot yet carry out large scale image registration themselves, even though the whole community would benefit from processed data of this sort. In the *Drosophila* neuroscience field, these challenges are being met by the development of the Virtual Fly Brain web resource (VFB) [66] (Box 1). VFB integrates and links neuromorphological data to curated information from the literature, providing tools to easily search and visualise the data. Other fly and non-fly databases that specialise in anatomical data (Allen Brain Atlas [67]; NeuroMorpho.org [68] and Insect Brain database; Box 1), have also been developed in the last few years to cope with this challenge.

The availability of EM connectomics data (Figure 2A) presents a new data integration challenge. One key issue is simply how to match a given EM tracing with light level/genetic driver imagery (Figure 2B). The larval *Drosophila* project has relied extensively on a collection of confocal images of single cell morphologies as an intermediate. However this dataset is indexed by the pattern recognition ability of a human expert: there is (presently) no automated query process because of the absence of co-registered data [41]. In the adult fly where the number of cell types is likely to be significantly greater, new tools have been developed that use image registration techniques to move EM data onto a light microscopy template brain or vice versa (e.g. elm/elmr, Box 1). This means EM data can be compared to any of the light microscopy datasets available. For example, this enables an EM neuron to be searched in seconds with NBLAST [62] against a database of single cell morphologies or driver lines (Figure 2C-G). Our team finds this functionality incredibly helpful in our tracing work focussed on the adult olfactory system/mushroom body.

Tools to analyse connectivity information – the significant advantage of EM data – in large datasets, as the adult fly brain, are still in their early development. It is not only the number of synapses that is relevant to interpret the circuit output, but also their location relative to their partners' morphology e.g. on dendritic or axonal arbours. Although reconstruction tools such as CATMAID [39,69] (Box 1) provide support for interactive network analysis, this is limited to a relatively small number of interactions (number of neurons or paths on a circuit).

Neuroscientists naturally need to test functional predictions from connectomics results. One apparently simple issue is to determine the significance of different numbers of synaptic inputs (pairs of neurons in the adult dataset can have from 1 to >400 connections). At a higher level, the ability to model the output

of complex networks will become ever more important [70,71]. Furthermore, approaches to link behavioural outputs to the activity of specific neuron types at scale [72] will be very significant. To achieve accurate predictions, however, these models should integrate information on morphology and connectivity (synapse number and location), physiology, neurotransmitter and gene expression. In order to facilitate the work of experimental neuroscientists, these data must be curated and integrated into comprehensive databases, following the example of the approach taken by genomic research [73].

5 Future Perspectives

The future development of connectomics (and specifically EM connectomics) as an approach to understand brain and behaviour will depend strongly on technological advances. When the human genome project began in 1990 it seemed an almost impossible task; 13 years later it was complete at an aggregate cost of 3 billion dollars. Now genomes can be sequenced in hours for <\$1000. Will connectomics see the same huge speed-ups? And what will the consequences be?

Present estimates for the effort to complete manual tracing and review of the adult brain of *Drosophila* are in the range 500-2000 person-years (implying a \$50-200M project). To bring this into a plausible range, manual labour must be reduced by a factor of 10-100. This means that the accuracy and completeness of automated results must result in 10-100x less user interaction. If correcting a mistake takes as long as making a manual tracing decision (it probably takes rather longer) then this means that tracing annotation must aim to be >99% accurate. A related issue is whether all results must be manually proof-read. Even if reviewing correct results is much faster than manual tracing this could severely limit the attainable speed-up. This implies that quality control must become increasingly automated. One general approach will be to ensure that automated software can reliably quantify its uncertainty. For whole brain data, quantitative approaches based on left-right symmetry of morphology or connectivity may be one possibility. We suspect that a key issue will be identifying homologous neurons or small groups of neurons and ensuring that they show consistent numbers, morphology and connectivity. Comparison across more than one specimen will clearly also be desirable.

If automated tracing can indeed result in a >10-100x speedup, then imaging speed again becomes an issue. A regular FIB-SEM instrument is the highest resolution tool available but is relatively slow. In theory, new SEM instruments containing >100 parallel scanning beams could speed up imaging by two orders of magnitude [45]. However, for technical reasons, these devices are not currently available with integrated FIB ion beam abrasion columns and therefore can only be used with other cutting methods, likely to have considerably poorer Z slice thickness (≥ 20 nm). This, in itself, may have a huge impact on the performance of automated reconstruction methods for neurites as fine as those in *Drosophila*, but may be acceptable for other species. In the long term, we should also bear in mind the real possibility that a different approach will supplant volume EM in its ability to deliver connectomic information.

We believe that recent efforts to map neural circuits are just the first wave of synapse-resolution connectomes to come. It is entirely plausible that in the future connectomic analyses will be just as ordinary to neuroscientists as sequencing is to present-day geneticists. Assuming that we are indeed able to obtain connectomes routinely, how will they change the science we do? Work already mentioned in the *Drosophila* larva has confirmed the value of using connectivity information to inform our understanding and select new experiments. We would also emphasise the value of comprehensive morphological reconstruction from EM – this frequently reveals previously unknown neuronal classes or enables a census of all neurons of a particular class. We have already highlighted quantitative approaches to morphology and data integration that should streamline the use of connectomes in the study of circuits and behaviour.

Comparative work across specimens will open up new possibilities. First it will be possible to examine variation across individuals. This may have origins in sex differences, in learning during an animal's life

or in random developmental processes. Having complete, high resolution reference connectomes it may be possible to use lower resolution imaging that is not guaranteed to reveal fine dendrites and synaptic connectivity. Likewise, it may be possible to image quite small volumes in conjunction with genetically targeted EM visible markers [74,75]. Either of these possibilities may be useful in combination with functional imaging protocols prior to fixation and preparation for EM.

The *Drosophila* larva appears likely to yield the next connectome of an entire central nervous system, but we naturally anticipate maps for other insects. We see huge value in whole brain EM volumes that would support the kind of focussed but interlocking studies typical of the *Drosophila* larva even before very highly automated reconstruction was possible. One possible area of interest would be to look at other *Drosophilids* with interesting neuroethology/ecology. Other key next targets could include insects of great significance to man such as the honey bee or mosquito. We suggest that a mosquito effort (e.g. *A. gambiae*) is already worth planning given that it is in the size range of *D. melanogaster* [76], whereas order of magnitude technology advances would be required to bring the honeybee into range. In the long term, we see great potential for the fusion of neuroethology, ecology and evo-devo studies through the application of neural circuit studies informed by connectomics to diverse insect models.

Acknowledgements

We thank A. Bates, A. Champion, M. Landgraf and C. Schneider-Mizell for comments on this manuscript. We thank D. Bock, S. Lauritzen and colleagues for early access to the FAFB adult brain dataset and A. Bates, F. Love, Z. Mitrevica, and R. Roberts, for EM tracing and analysis of neurons shown in figures and Y. Aso for sharing light microscopy image data. We also thank M. Ahrens and T. Kawashima for information on the larval zebrafish brain. This work was supported by the Wellcome Trust (203261/Z/16/Z), an ERC Consolidator grant (649111) and core support from the MRC (MC-U105188491) to G.S.X.E.J.

References

1. Cajal SR y.: *Textura del sistema nerviosa del hombre y vertebrados*. Nicolás Moya [English edition (1995): *Histology of the nervous system of man and vertebrates*. Oxford University Press; 1899–1904.
2. Dujardin F: **Mémoire sur le système nerveux des insectes**. *Ann Sci Nat Zool* 1850, **14**:195–206.
3. Kenyon FC: **The Meaning and Structure of the So-Called “Mushroom Bodies” of the Hexapod Brain**. *Am Nat* 1896, **30**:643–650.
4. Sterling P, Laughlin S: *Principles of Neural Design*. MIT Press; 2015.
5. Katz P, Grillner S, Wilson R, Borst A, Greenspan R, Buzsáki G, Martin K, Marder E, Kristan W, Friedrich R, et al.: **Vertebrate versus invertebrate neural circuits**. *Curr Biol* 2013, **23**:R504–6.
6. Strausfeld NJ, Hildebrand JG: **Olfactory systems: common design, uncommon origins?** *Curr Opin Neurobiol* 1999, **9**:634–639.
7. Farris SM: **Are mushroom bodies cerebellum-like structures?** *Arthropod Struct Dev* 2011, **40**:368–379.
8. Stevens CF: **What the fly’s nose tells the fly's brain**. *Proc Natl Acad Sci U S A* 2015, **112**:9460–9465.
9. Borst A, Helmstaedter M: **Common circuit design in fly and mammalian motion vision**. *Nat Neurosci* 2015, **18**:1067–1076.

10. Rajan A, Perrimon N: **Drosophila cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion.** *Cell* 2012, **151**:123–137.
11. Wang Y, Pu Y, Shen P: **Neuropeptide-gated perception of appetitive olfactory inputs in Drosophila larvae.** *Cell Rep* 2013, **3**:820–830.
12. Schlegel P, Texada MJ, Miroschnikow A, Schoofs A, Hückesfeld S, Peters M, Schneider-Mizell CM, Lacin H, Li F, Fetter RD, et al.: **Synaptic transmission parallels neuromodulation in a central food-intake circuit.** *Elife* 2016, **5**:e16799.

* Neural networks operate with various modes of transmissions. This is one of the few examples in which synaptic connectivity has been studied in conjunction with peptidergic transmission.
13. White JG, Southgate E, Thomson JN, Brenner S: **The structure of the nervous system of the nematode Caenorhabditis elegans.** *Philos Trans R Soc Lond B Biol Sci* 1986, **314**:1–340.
14. Fischbach K-F, Dittrich APM: **The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure.** *Cell Tissue Res* 1989, **258**.
15. Jefferis GSXE, Livet J: **Sparse and combinatorial neuron labelling.** *Curr Opin Neurobiol* 2012, **22**:101–110.
16. Wolff T, Iyer NA, Rubin GM: **Neuroarchitecture and neuroanatomy of the Drosophila central complex: A GAL4-based dissection of protocerebral bridge neurons and circuits.** *J Comp Neurol* 2015, **523**:997–1037.
17. Matsuo E, Seki H, Asai T, Morimoto T, Miyakawa H, Ito K, Kamikouchi A: **Organization of projection neurons and local neurons of the primary auditory center in the fruit fly Drosophila melanogaster.** *J Comp Neurol* 2016, **524**:1099–1164.
18. Chin A-L, Lin C-Y, Fu T-F, Dickson BJ, Chiang A-S: **Diversity and wiring variability of visual local neurons in the Drosophila medulla M6 stratum.** *J Comp Neurol* 2014, **522**:3795–3816.
19. Jefferis GSXE, Potter CJ, Chan AM, Marin EC, Rohlifing T, Maurer CR Jr, Luo L: **Comprehensive maps of Drosophila higher olfactory centers: spatially segregated fruit and pheromone representation.** *Cell* 2007, **128**:1187–1203.
20. Chiang A-S, Lin C-Y, Chuang C-C, Chang H-M, Hsieh C-H, Yeh C-W, Shih C-T, Wu J-J, Wang G-T, Chen Y-C, et al.: **Three-dimensional reconstruction of brain-wide wiring networks in Drosophila at single-cell resolution.** *Curr Biol* 2011, **21**:1–11.
21. Manton JD, Ostrovsky AD, Goetz L, Costa M, Rohlifing T, Gregory S X: **Combining genome-scale Drosophila 3D neuroanatomical data by bridging template brains.** *bioRxiv* 2014, doi:10.1101/006353.

* This study introduces open source tools and data that enable all the large light microscopy datasets of Drosophila data to be placed in the same co-registered image space. The same approaches are being used to bridge EM reconstructions and light-level image data in order to e.g. identify genetic reagents.
22. Rohlifing T, Maurer CR Jr: **Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees.** *IEEE Trans Inf Technol Biomed* 2003, **7**:16–25.
23. Aso Y, Sitaraman D, Ichinose T, Kaun KR, Vogt K, Belliart-Guérin G, Plaçais P-Y, Robie AA, Yamagata N, Schnaitmann C, et al.: **Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila.** *Elife* 2014, **3**:e04580.
24. Wu M, Nern A, Williamson WR, Morimoto MM, Reiser MB, Card GM, Rubin GM: **Visual projection neurons in the Drosophila lobula link feature detection to distinct behavioral programs.** *Elife* 2016, **5**:e21022.

25. Aso Y, Rubin GM: **Dopaminergic neurons write and update memories with cell-type-specific rules.** *Elife* 2016, **5**:e16135.
26. Zheng Z, Scott Lauritzen J, Perlman E, Robinson CG, Nichols M, Milkie D, Torrens O, Price J, Fisher CB, Sharifi N, et al.: **A Complete Electron Microscopy Volume Of The Brain Of Adult *Drosophila melanogaster*.** *bioRxiv* 2017, doi:10.1101/140905.

** This study represents a substantial technical advance in acquisition of EM data applied to the generation of a whole brain volume of an adult female fly.
27. Takemura S-Y, Nern A, Chklovskii DB, Scheffer LK, Rubin GM, Meinertzhagen IA: **The comprehensive connectome of a neural substrate for “ON” motion detection in *Drosophila*.** *Elife* 2017, **6**:e24394.
28. Lin C-Y, Chuang C-C, Hua T-E, Chen C-C, Dickson BJ, Greenspan RJ, Chiang A-S: **A comprehensive wiring diagram of the protocerebral bridge for visual information processing in the *Drosophila* brain.** *Cell Rep* 2013, **3**:1739–1753.
29. Feinberg EH, Vanhoven MK, Bendesky A, Wang G, Fetter RD, Shen K, Bargmann CI: **GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems.** *Neuron* 2008, **57**:353–363.
30. Li Y, Guo A, Li H: **CRASP: CFP reconstitution across synaptic partners.** *Biochem Biophys Res Commun* 2016, **469**:352–356.
31. Chang J-B, Chen F, Yoon Y-G, Jung EE, Babcock H, Kang JS, Asano S, Suk H-J, Pak N, Tillberg PW, et al.: **Iterative expansion microscopy.** *Nat Methods* 2017, **14**:593–599.
32. Ryan K, Lu Z, Meinertzhagen IA: **The CNS connectome of a tadpole larva of *Ciona intestinalis* (L.) highlights sidedness in the brain of a chordate sibling.** *Elife* 2016, **5**:e16962.
33. Schafer WR: **Deciphering the neural and molecular mechanisms of *C. elegans* behavior.** *Curr Biol* 2005, **15**:R723–9.
34. Emmons SW: **The beginning of connectomics: a commentary on White et al. (1986) “The structure of the nervous system of the nematode *Caenorhabditis elegans*.”** *Philos Trans R Soc Lond B Biol Sci* 2015, **370**.
35. Helmstaedter M, Briggman KL, Turaga SC, Jain V, Seung HS, Denk W: **Connectomic reconstruction of the inner plexiform layer in the mouse retina.** *Nature* 2013, **500**:168–174.
36. Kim JS, Greene MJ, Zlateski A, Lee K, Richardson M, Turaga SC, Purcaro M, Balkam M, Robinson A, Behabadi BF, et al.: **Space–time wiring specificity supports direction selectivity in the retina.** *Nature* 2014, **509**:331–336.
37. Eichler K, Li F, Litwin-Kumar A, Park Y, Andrade I, Schneider-Mizell CM, Saumweber T, Huser A, Eschbach C, Gerber B, et al.: **The complete connectome of a learning and memory centre in an insect brain.** *Nature* 2017, **548**:175–182.

* This study demonstrates how comprehensive maps of the larval mushroom bodies and their principal inputs and outputs can provide new insights into otherwise well studied systems.
38. Takemura S-Y, Aso Y, Hige T, Wong A, Lu Z, Xu CS, Rivlin PK, Hess H, Zhao T, Parag T, et al.: **A connectome of a learning and memory center in the adult *Drosophila* brain.** *Elife* 2017, **6**:e26975.

** On a technical level, this study illustrates how isotropic FIB-SEM data can already be used to efficiently perform automated segmentation on restricted brain areas of interest.
39. Schneider-Mizell CM, Gerhard S, Longair M, Kazimiers T, Li F, Zwart MF, Champion A, Midgley FM, Fetter RD, Saalfeld S, et al.: **Quantitative neuroanatomy for connectomics in *Drosophila*.** *Elife* 2016, **5**:e12059.

** This publication provides tools and algorithms to analyse connectomic data and investigates the impact of EM reconstruction errors.

40. Berck ME, Khandelwal A, Claus L, Hernandez-Nunez L, Si G, Tabone CJ, Li F, Truman JW, Fetter RD, Louis M, et al.: **The wiring diagram of a glomerular olfactory system**. *Elife* 2016, **5**.

* Comprehensive description of antennal lobe circuits. This publication demonstrates how connectomic data can advance knowledge and generate new hypotheses even in highly studied systems.

41. Ohyama T, Schneider-Mizell CM, Fetter RD, Aleman JV, Franconville R, Rivera-Alba M, Mensh BD, Branson KM, Simpson JH, Truman JW, et al.: **A multilevel multimodal circuit enhances action selection in *Drosophila***. *Nature* 2015, **520**:633–639.

** This study introduced the larval *Drosophila* volume and showed how it can be combined with functional and behavioural analysis to understand the basis of multimodal integration thought to underlie the larval escape response to parasitoid wasps.

42. Jovanic T, Schneider-Mizell CM, Shao M, Masson J-B, Denisov G, Fetter RD, Mensh BD, Truman JW, Cardona A, Zlatic M: **Competitive Disinhibition Mediates Behavioral Choice and Sequences in *Drosophila***. *Cell* 2016, **167**:858–870.e19.
43. Fushiki A, Zwart MF, Kohsaka H, Fetter RD, Cardona A, Nose A: **A circuit mechanism for the propagation of waves of muscle contraction in *Drosophila***. *Elife* 2016, **5**:e13253.
44. Zwart MF, Pulver SR, Truman JW, Fushiki A, Fetter RD, Cardona A, Landgraf M: **Selective Inhibition Mediates the Sequential Recruitment of Motor Pools**. *Neuron* 2016, **91**:615–628.
45. Helmstaedter M: **Cellular-resolution connectomics: challenges of dense neural circuit reconstruction**. *Nat Methods* 2013, **10**:501–507.
46. Januszewski M, Maitin-Shepard J, Li P, Kornfeld J, Denk W, Jain V: **Flood-Filling Networks**. *arXiv* 2016,
47. Staffler B, Berning M, Boergens KM, Gour A, van der Smagt P, Helmstaedter M: **SynEM: Automated synapse detection for connectomics**. *bioRxiv* 2017, doi:10.1101/099994.
48. Funke J, Tschopp FD, Grisaitis W, Singh C, Saalfeld S, Turaga SC: **A Deep Structured Learning Approach Towards Automating Connectome Reconstruction from 3D Electron Micrographs**. *arXiv* 2017,
49. Lichtman JW, Pfister H, Shavit N: **The big data challenges of connectomics**. *Nat Neurosci* 2014, **17**:1448–1454.
50. Kubota Y: **New developments in electron microscopy for serial image acquisition of neuronal profiles**. *Microscopy* 2015, **64**:27–36.
51. Menzel R, Giurfa M: **Cognitive architecture of a mini-brain: the honeybee**. *Trends Cogn Sci* 2001, **5**:62–71.
52. Schalek R, Lee D, Kasthuri N, Peleg A, Jones T, Kaynig V, Haehn D, Pfister H, Cox D, Lichtman JW: **Imaging a 1 mm³ Volume of Rat Cortex Using a MultiBeam SEM**. *Microsc Microanal* 2016, **22**:582–583.
53. Dorkenwald S, Schubert PJ, Killinger MF, Urban G, Mikula S, Svara F, Kornfeld J: **Automated synaptic connectivity inference for volume electron microscopy**. *Nat Methods* 2017, **14**:435–442.
54. Braitenberg V, Schüz A: *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer; 1998.
55. Shih C-T, Sporns O, Yuan S-L, Su T-S, Lin Y-J, Chuang C-C, Wang T-Y, Lo C-C, Greenspan RJ, Chiang A-S: **Connectomics-based analysis of information flow in the *Drosophila* brain**. *Curr*

Biol 2015, **25**:1249–1258.

56. Jenett A, Rubin GM, Ngo T-TB, Shepherd D, Murphy C, Dionne H, Pfeiffer BD, Cavallaro A, Hall D, Jeter J, et al.: **A GAL4-driver line resource for Drosophila neurobiology.** *Cell Rep* 2012, **2**:991–1001.
 57. Pfeiffer BD, Ngo T-TB, Hibbard KL, Murphy C, Jenett A, Truman JW, Rubin GM: **Refinement of tools for targeted gene expression in Drosophila.** *Genetics* 2010, **186**:735–755.
 58. Luan H, Peabody NC, Vinson CR, White BH: **Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression.** *Neuron* 2006, **52**:425–436.
 59. Zeng H, Sanes JR: **Neuronal cell-type classification: challenges, opportunities and the path forward.** *Nat Rev Neurosci* 2017, doi:10.1038/nrn.2017.85.
 60. Cardona A, Saalfeld S, Arganda I, Pereanu W, Schindelin J, Hartenstein V: **Identifying neuronal lineages of Drosophila by sequence analysis of axon tracts.** *J Neurosci* 2010, **30**:7538–7553.
 61. Wan Y, Long F, Qu L, Xiao H, Hawrylycz M, Myers EW, Peng H: **BlastNeuron for Automated Comparison, Retrieval and Clustering of 3D Neuron Morphologies.** *Neuroinformatics* 2015, **13**:487–499.
 62. Costa M, Manton JD, Ostrovsky AD, Prohaska S, Jefferis GSXE: **NBLAST: Rapid, Sensitive Comparison of Neuronal Structure and Construction of Neuron Family Databases.** *Neuron* 2016, **91**:293–311.
- ** This study presents an automated tool (NBLAST) to compare neuronal morphology and uses it to generate a first estimate of the total number of cell types in the fly brain (>1000) based on a database of 16000 co-registered single cells.
63. Ganglberger F, Schulze F, Tirian L, Novikov A, Dickson B, Bühler K, Langs G: **Structure-Based Neuron Retrieval Across Drosophila Brains.** *Neuroinformatics* 2014, **12**:423–434.
 64. Armañanzas R, Ascoli GA: **Towards the automatic classification of neurons.** *Trends Neurosci* 2015, **38**:307–318.
 65. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al.: **Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets.** *Cell* 2015, **161**:1202–1214.
 66. Milyaev N, Osumi-Sutherland D, Reeve S, Burton N, Baldock RA, Armstrong JD: **The Virtual Fly Brain browser and query interface.** *Bioinformatics* 2012, **28**:411–415.
 67. Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, et al.: **Genome-wide atlas of gene expression in the adult mouse brain.** *Nature* 2007, **445**:168–176.
 68. Nanda S, Allaham MM, Bergamino M, Polavaram S, Armañanzas R, Ascoli GA, Parekh R: **Doubling up on the fly: NeuroMorpho.Org Meets Big Data.** *Neuroinformatics* 2015, **13**:127–129.
 69. Saalfeld S, Cardona A, Hartenstein V, Tomancak P: **CATMAID: collaborative annotation toolkit for massive amounts of image data.** *Bioinformatics* 2009, **25**:1984–1986.
 70. Bassett DS, Sporns O: **Network neuroscience.** *Nat Neurosci* 2017, **20**:353–364.
 71. Gornet J, Scheffer LK: **Simulating Extracted Connectomes.** *bioRxiv* 2017, doi:10.1101/177113.
 72. Robie AA, Hirokawa J, Edwards AW, Umayam LA, Lee A, Phillips ML, Card GM, Korff W, Rubin GM, Simpson JH, et al.: **Mapping the Neural Substrates of Behavior.** *Cell* 2017, **170**:393–406.e28.

* High-throughput mapping of neural populations to behaviour. These kinds of data will help be

essential for interpreting connectomes on a large-scale.

73. Thorisson GA, Muilu J, Brookes AJ: **Genotype–phenotype databases: challenges and solutions for the post-genomic era.** *Nat Rev Genet* 2009, **10**:9–18.
74. Joesch M, Mankus D, Yamagata M, Shahbazi A, Schalek R, Suissa-Peleg A, Meister M, Lichtman JW, Scheirer WJ, Sanes JR: **Reconstruction of genetically identified neurons imaged by serial-section electron microscopy.** *Elife* 2016, **5**:e15015.
75. Ng J, Browning A, Lechner L, Terada M, Howard G, Jefferis G: **Genetically targeted 3D visualisation of Drosophila neurons under Electron Microscopy and X-Ray Microscopy using miniSOG.** *Scientific reports* 2016, **6**.
76. Riabinina O, Task D, Marr E, Lin C-C, Alford R, O'Brochta DA, Potter CJ: **Organization of olfactory centres in the malaria mosquito *Anopheles gambiae*.** *Nat Commun* 2016, **7**:13010.
77. Briggman KL, Bock DD: **Volume electron microscopy for neuronal circuit reconstruction.** *Curr Opin Neurobiol* 2012, **22**:154–161.
78. Boergens KM, Berning M, Bocklisch T, Bräunlein D, Drawitsch F, Frohnhofen J, Herold T, Otto P, Rzepka N, Werkmeister T, et al.: **webKnossos: efficient online 3D data annotation for connectomics.** *Nat Methods* 2017, **14**:691–694.
79. Bruckner S, Solteszova V, Groller ME, Hladuvka J, Buhler K, Yu JY, Dickson BJ: **BrainGazer - Visual Queries for Neurobiology Research.** *IEEE Trans Vis Comput Graph* 2009, **15**:1497–1504.
80. Ukani NH, Yeh C-H, Tomkins A, Zhou Y, Florescu D, Ortiz CL, Huang Y-C, Wang C-T, Richmond P, Lo C-C, et al.: **The Fruit Fly Brain Observatory: from structure to function.** *bioRxiv* 2016, doi:10.1101/092288.
81. Clark DA, Mitra PP, Wang SS: **Scalable architecture in mammalian brains.** *Nature* 2001, **411**:189–193.
82. Ahrens MB, Orger MB, Robson DN, Li JM, Keller PJ: **Whole-brain functional imaging at cellular resolution using light-sheet microscopy.** *Nat Methods* 2013, **10**:413–420.
83. Badea A, Ali-Sharief AA, Johnson GA: **Morphometric analysis of the C57BL/6J mouse brain.** *Neuroimage* 2007, **37**:683–693.
84. Azevedo FAC, Carvalho LRB, Grinberg LT, Farfel JM, Ferretti REL, Leite REP, Filho WJ, Lent R, Herculano-Houzel S: **Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain.** *J Comp Neurol* 2009, **513**:532–541.
85. Herculano-Houzel S, Mota B, Lent R: **Cellular scaling rules for rodent brains.** *Proc Natl Acad Sci U S A* 2006, **103**:12138–12143.
86. Larderet I, Fritsch P, Gendre N, Maier L, Fetter RD, Schneider-Mizell C, Truman J, Zlatic M, Cardona A, Sprecher S: **Organization Of The Drosophila Larval Visual Circuit.** *eLife* 2017, **6**:e28387.
87. Randel N, Asadulina A, Bezares-Calderón LA, Verasztó C, Williams EA, Conzelmann M, Shahidi R, Jékely G: **Neuronal connectome of a sensory-motor circuit for visual navigation.** *Elife* 2014, **3**:e02730.
88. Kornfeld J, Benezra SE, Narayanan RT, Svava F, Egger R, Oberlaender M, Denk W, Long MA: **EM connectomics reveals axonal target variation in a sequence-generating network.** *Elife* 2017, **6**:e24364.
89. Wanner AA, Genoud C, Masudi T, Siksou L, Friedrich RW: **Dense EM-based reconstruction of the interglomerular projectome in the zebrafish olfactory bulb.** *Nat Neurosci* 2016, **19**:816–825.

Box 1- Current Methods and Resources

Methods

ssTEM: serial section Transmission Electron Microscopy. This technique relies on manually cutting thin serial sections prior to imaging. It offers the highest acquisition speed when combined with camera arrays and automated high speed sample movement/exchange; resolution is better than FIB-SEM in x/y (4nm), but worse in z (40nm). Image data has to be realigned into a cohesive volume [26,50,77].

FIB-SEM: Focused Ion Beam milling Scanning Electron Microscopy. This technique combines a SEM and gallium ion source. During image acquisition, layers are sequentially shaved off by an ion beam. FIB-SEM has the highest z-resolution currently possible (5nm) and allows generation of isotropic data at the cost of greatly reduced acquisition speed. Images do not need to be realigned but sample is lost during acquisition [50,77].

SBEM: serial block-face scanning electron microscopy. Combination of SEM and in-chamber microtome. Fast cutting and imaging times at cost of x/y resolution [26,50,77].

CATMAID: Collaborative Annotation Toolkit for Massive Amounts of Data. Web interface that allows the navigation and collaborative annotation of 3D biological image datasets. The interface allows the reconstruction of neurons from ssTEM volumes, browsing and initial analysis of the neurons' morphology and connectivity [39,69].

Knossos/WebKnossos: Used to visualise EM data and support skeleton as well as volume reconstruction. WebKnossos (<https://www.webknossos.org>) is the web-based implementation of Knossos (<https://knossostool.org>) [78].

elm+elmr: Fiji/ImageJ plugin to allow co-visualisation of Electron-Light Microscopy data (<https://github.com/saalfeldlab/elm>) allowing iterative identification of landmarks in two image volumes, defining a transformation between them. Building on this, elmr is an R package that moves neuronal tracing and other data between light and EM volumes described in Zheng et al. [26] (<https://github.com/jefferis/elmr>). It interacts with CATMAID and the Neuroanatomy Toolbox (<https://github.com/jefferis/nat>) package, allowing more sophisticated analysis.

Resources

Large light microscopy datasets:

- FlyCircuit (<http://www.flycircuit.tw/>): around 23,000 single neuron images obtained by MARCM [20,55].
- FlyLight (<http://flweb.janelia.org/cgi-bin/flew.cgi>): around 3,500 GAL4 and near 1500 LexA driver lines [56,57].
- Dickson VT lines (<https://braingazer.org/>): around 8,500 GAL4 lines. braingazer.org provides advanced querying tools to search for neurons of interest, potential connectivity and similarity [79].

Most of these datasets provide expression data per neuropil. These datasets are available for academic reuse, but may have limitations for commercial purposes.

Unfortunately, datasets are often released with unclear licenses. This is an issue that data producers should address to maximise the use of their data.

Insect EM datasets:

- **FAFB** (<http://www.temca2data.org/>): Adult *Drosophila* brain of female adult fly comprising 100,000 neurons [26].
- **Larval *Drosophila*** (<https://neurodata.io/data/>): CNS of first instar larva comprising 10-12,000 neurons [41].

Anatomical data:

Virtual Fly Brain (VFB) (<http://www.virtualflybrain.org/>): Integrates *Drosophila* neural information, including published data (anatomy, innervation, connectivity, expression, etc) and image datasets. Users

can query the data in a variety of ways, view and compare the images on a 3D browser, find similar neurons using NBLAST [62], and download the image data [66].

Fruit Fly Brain Observatory (FBBO) (<http://fruitflybrain.org/>): web resource focused on network simulation. It integrates existing data types including morphology, connectivity and physiology, to generate circuit models [80].

Insect Brain Database (<https://insectbraindb.org/>): web inventory of single neurons and 3D brain models in various insect species including honeybee and monarch butterfly. Users are able to submit data.

NeuroMorpho.org (<http://neuromorpho.org/>): Inventory of digitally reconstructed neurons, from many different species, including *Drosophila*. Different datasets are not cross-registered. Users are able to submit and download data [68].

Allen Brain Atlas (<http://www.brain-map.org/>): web resource that generates and aggregates neuronal data for mostly mouse and human. It includes large scale gene expression and characterisation of tissues, a cell type inventory of visual cortex neurons based on activity and anatomy and mesoscale connectivity data [67].

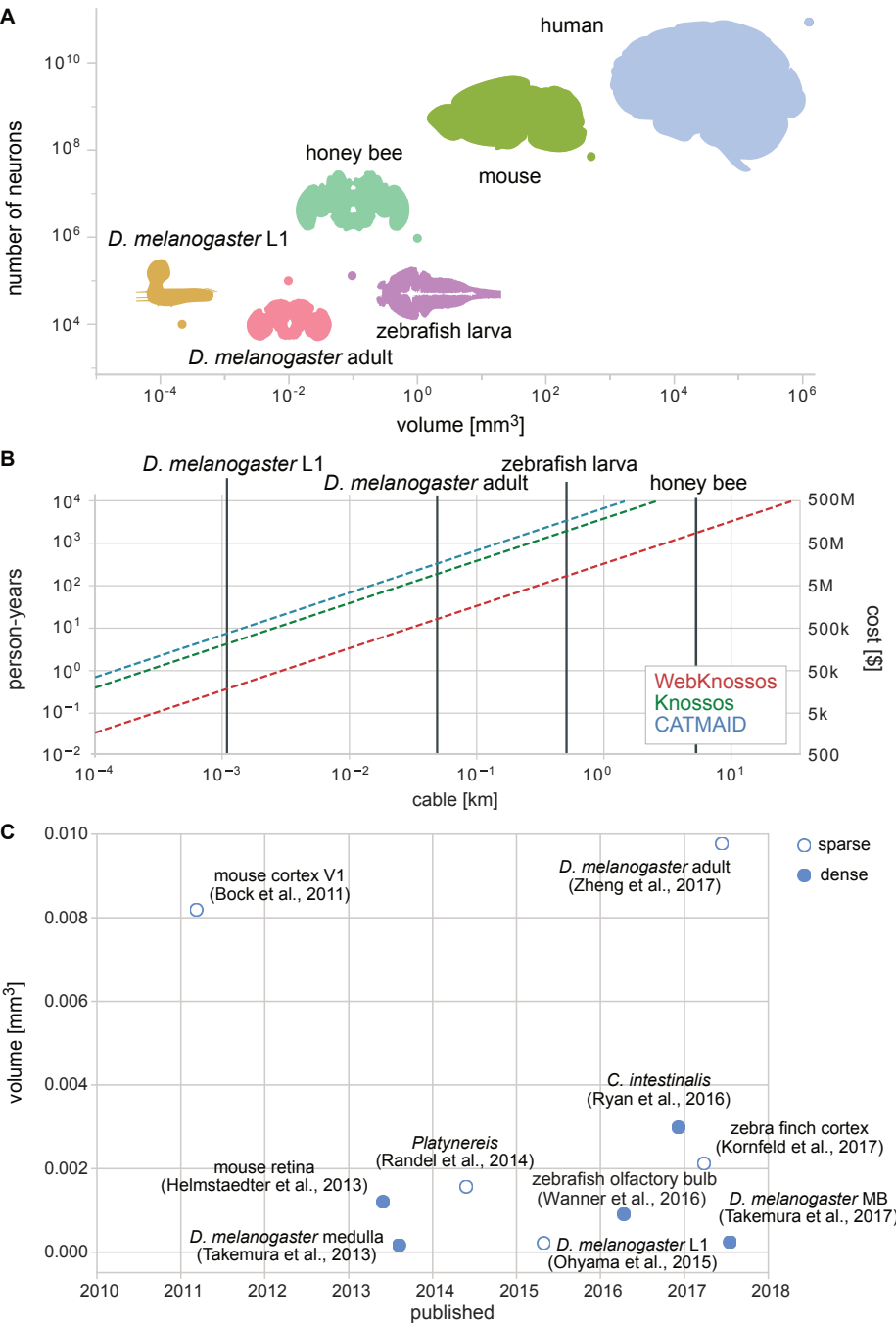
Figure 1 - Size and cost of connectomes

- A) Comparison of brain volume and neuron number for some widely used model systems. Insect brains are well suited for comprehensive mapping due to their relatively small size and lower neuron number. References: [51,81–85]; zebrafish larva (6 dpf) neuron count: 130,000 (T. Kawashima and M. Ahrens, personal communication).
- B) Estimated time and cost to generate whole brain connectomes using state-of-the-art manual annotation tools. An adult *Drosophila* connectome is feasible as a multi-lab collaboration whereas a larval connectome might even be within reach of a single lab effort. Calculations are based on assumed 5km cable/mm³ and \$50,000/person-year, and do not include overhead from planning and analysis. Reconstruction speeds for different tools are based on [35,40,78,86].
- C) Various EM dataset have been published in recent years. Reconstruction method (dense/sparse) refers to whether all or a subset of neurons in given volume were reconstructed. References not already mentioned in the text: *Platynereis* [87]; zebra finch [87,88]; zebrafish [89].

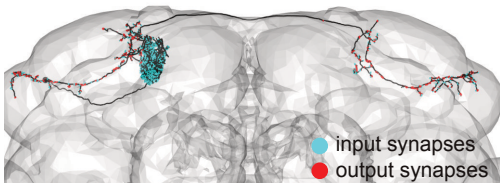
Figure 2 - Using light level tools to analyse EM data

- A) Single identified mushroom body output neuron $\alpha 2sc$ (MBON- $\alpha 2sc$) reconstructed in the FAFB EM dataset [26], plotted in the light level template (FCWB). Input synapses in cyan, output ones in red, brain template in grey.
- B) Comparison between two images of the MBON- $\alpha 2sc$: one obtained by light level microscopy (LM, brown) [23] and another traced in the FAFB EM dataset (EM, green). There is one MBON- $\alpha 2sc$ neuron per brain hemisphere, and both are shown. The left EM neuron has not been fully traced; the difference is noticeable in its dendritic domain.
- C) Flow-chart showing how existing light level tools can aid identification of neurons traced in EM. First, a traced neuron can be imported into R and transformed onto a light level template, FCWB, using `elm` and `elmr` (Box 1). Then, neurons are mirrored to the left hemisphere (if needed) using a mirroring registration [21]. After this, the neuron can be searched against a light level database of ~16,000 single neuron images (Flycircuit) [20] (all on the left hemisphere), using NBLAST, to find its best match based on morphology and position [62]. Around half of the FlyCircuit dataset have been assigned a neuron type (<http://www.virtualflybrain.org/>), allowing users to identify their traced neuron.
- D) Local lateral horn neurons of the AV1 type (n=59) traced in the FAFB EM dataset. The neurons were fetched from CATMAID and transformed onto the light level template FCWB.

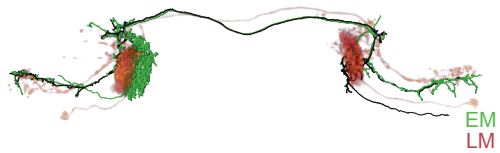
- D') Morphological clustering of AV1 neurons based on NBLAST scores. The NBLAST similarity scores for the neurons shown in D (each one against all others) were calculated and used to perform hierarchical clustering. The resulting dendrogram was cut at a height that produced 5 different groups (1-5). This approach can be used to reveal groups of similar neurons, and helps to identify types and subtypes. The inset shows the neurons coloured according to the clustering group, highlighting that morphologically similar neurons are part of the same group. Neurons on the right hemisphere were mirrored to the left before calculating the scores.
- E) Plot of the neurons in group 3 (n=10) from D'. This group includes neurons on the left and right hemispheres, suggesting they might be homologous neurons of the same subtype.
- F) EM neuron (EM, green) and its light level best NBLAST hit (LM, brown) from the Flycircuit dataset. The EM neuron used was from group 3 (E). The approach followed is detailed in C.
- G) Group 3 neurons (E) (mirrored to the right, in green) and its best NBLAST hit against a database of genetic driver lines [56] (R66B12, in grey and brown) [62]. The full driver line expression is shown in grey, and the regions that match with group 3 neurons are shown in brown.



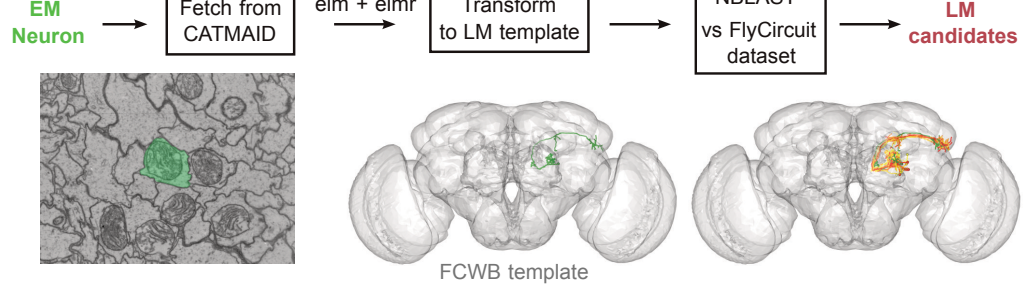
MBON- α 2sc



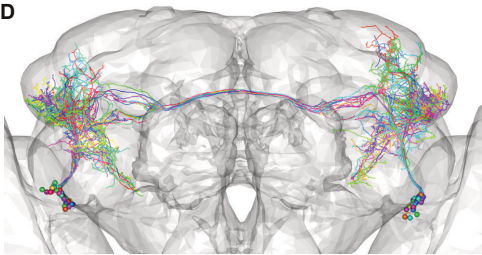
B



C



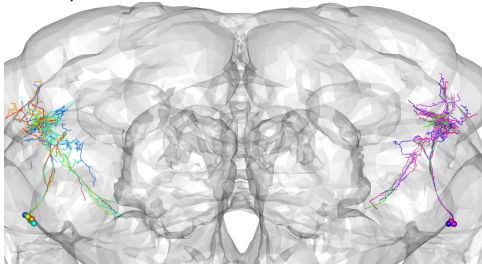
D



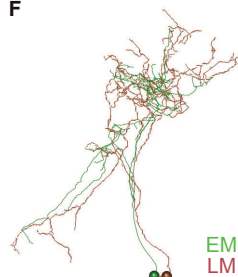
D'



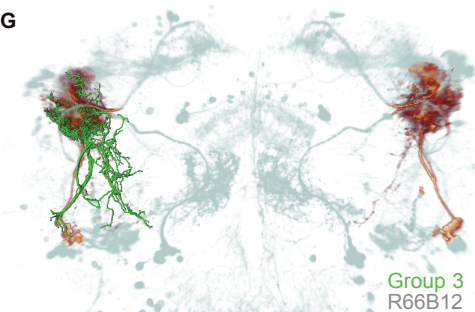
E Group 3



F



G



Conflict of interest

None declared.