Intact Drosophila Whole Brain Cellular Quantitation reveals Sexual Dimorphism

Wei Jiao[1], Gard Spreemann [1†], Evelyne Ruchti[1], Soumya Banerjee[1], Ying Shi[1], R. Steven Stowers[2], Kathryn Hess[1] and Brian D. McCabe[1*].

1. Brain Mind Institute, EPFL - Swiss Federal Institute of Technology Lausanne, Switzerland.
2. Department of Microbiology and Cell Biology, Montana State University, Bozeman, U.S.A

[†] Current address: Telenor Research, Fornebu, Norway

*To whom correspondence should be addressed: brian.mccabe@epfl.ch

ABSTRACT

Establishing with precision the quantity and identity of the cell types of the brain is a prerequisite for a detailed compendium of gene and protein expression in the central nervous system. Currently however, strict quantitation of cell numbers has been achieved only for the nervous system of *C.elegans*. Here we describe the development of a synergistic pipeline of molecular genetic, imaging, and computational technologies designed to allow high-throughput, precise quantitation with cellular resolution of reporters of gene expression in intact whole tissues with complex cellular constitutions such as the brain. We have deployed the approach to determine with exactitude the number of functional neurons and glia in the entire intact *Drosophila* larval brain, revealing fewer neurons and many more glial cells than previously estimated. Moreover, we discover an unexpected divergence between the sexes at this juvenile developmental stage, with female brains having significantly more neurons than males. Topological analysis of our data establishes that this sexual dimorphism extends to deeper features of brain organisation. Our methodology enables robust and accurate quantification of the number and positioning of cells within intact organs, facilitating sophisticated analysis of cellular identity, diversity, and expression characteristics.

INTRODUCTION

Establishing the precise numbers of cell types in the brain is essential to create organ-wide catalogues of neuronal cell-types and gene expression (Lent et al., 2012; Devor et al., 2013). However, apart from the nervous system of the nematode *Caenorhabditis elegans* (302 neurons, 56 glia) (White et al., 1986), the exact numbers of cells within the central nervous system (CNS) of model organisms or that of humans is currently unknown, with estimates, including those based upon extrapolation from direct quantification of brain sub-regions, varying widely (Silbereis et al., 2016; Keller et al., 2018; von Bartheld et al., 2016).

Studies of the central nervous system (CNS) of *Drosophila melanogaster*, which in scale and behavioural repertoire has been viewed as intermediate between nematodes and rodents (Bellen et al., 2010; Alivisatos et al., 2012), currently include large-scale efforts to establish both a neuronal connectome and cell atlas (Scheffer and Meinertzhagen, 2019; Allen et al., 2020; Li et al., 2021). Nonetheless, the precise numbers of active neurons and glia in either the smaller larval or larger adult *Drosophila* brain remain unknown, with estimates ranging from 12,000 to 15,000 neurons in the larval brain (Scott et al., 2001; Meinertzhagen, 2018; Eschbach and Zlatic, 2020) and 100,000 to 199,000 neurons in the adult brain (Simpson, 2009; Chiang et al., 2011; Kaiser, 2015; Scheffer and Meinertzhagen, 2019; Raji and Potter, 2021). The number of glial cells in the larval brain has been estimated as perhaps 10% of the number of neurons (Kremer et al., 2017; Meinertzhagen, 2018; Raji and Potter, 2021). Given the large diversity of ranges proposed, precise quantification of the numbers of neurons and glia in the larval brain, which enable the wide compendium of *Drosophila* larval behaviours(Gerber et al., 2009; Neckameyer and Bhatt, 2016; Eschbach and Zlatic, 2020; Louis, 2020; Gowda et al., 2021), would seem a desirable and achievable goal.

Complicating this aspiration, in addition to the general problem of separating and quantifying primary cell types such as neurons and glia, are two specific confounding factors that limit simple total cell quantification approaches in the *Drosophila* larval brain. Firstly, encompassed within and

surrounding the larval CNS are dividing neuroblasts, which will give rise to adult neurons (Doe, 2017). Relatedly, imbedded within the larval brain are substantial numbers of immature adult neurons, observed from electron micrograph reconstructions as having few or no dendrites and axons that terminate in filopodia lacking synapses (Eichler et al., 2017). These immature neurons are unlikely to contribute to larval brain function and are excluded when considering larval neuronal circuit architecture (Eichler et al., 2017; Scheffer and Meinertzhagen, 2019). It has been suggested that only a small fraction of the total number of larval CNS cells may actually contribute to larval brain function (Ravenscroft et al., 2020).

Here we have sought to develop a synergistic molecular genetic, imaging, and computational pipeline designed *de novo* to allow automated neuron, glia, or other gene expression features to be precisely quantitated with cellular resolution in intact whole brains. Central to the approach are high signal-to-noise gene expression reporters that produce a punctate, nucleus-localised output facilitating downstream automated computational measurement and analysis. Exploiting multiple genetic reagents designed to selectively identify only functional neurons with active synaptic protein expression, we identify substantially fewer larval neurons than previously estimated in the *Drosophila* larval brain, and in addition, substantially more glia. We also discover a previously unsuspected sexual dimorphism in the numbers of both cell types at larval stages. The generation of whole brain point clouds from our data enabled us to apply the tools of topological data analysis to summarize brains in terms of multi-scale topological structures. Utilization of these topological summaries in a support vector machine also supports that sexual dimorphism extends to deeper features of brain organisation. Finally, we applied our pipeline to quantitate the whole brain expression of the *Drosophila* family of voltage-gated potassium channels which revealed divergent channel expression frequency throughout the brain. We envision that our method can be employed to allow precise quantitation of gene expression characteristics of the constituent cells of the brain and potentially other intact whole organs in a format suitable for sophisticated downstream analysis.

RESULTS

Genetic and imaging tools to facilitate automated whole brain cellular quantitation. To establish a robust quantitative method to measure gene expression frequency and quantify the cell numbers that contribute to *Drosophila* larval brain function, we sought to develop a pipeline with genetic reporters designed to expediate automated neuron and glia quantitation from three-dimensional intact organ images. While membrane associated reporters are generally employed to label *Drosophila* neurons (Pfeiffer et al., 2008; Jenett et al., 2012; Ravenscroft et al., 2020), we generated UAS-driven (Brand and Perrimon, 1993; Wang et al., 2012) fluorescent reporters fused to Histone proteins (Sherer et al., 2020) to target fluorescence only to the nucleus, in order to facilitate subsequent automated segmentation and counting. Through empirical selection of transgene genomic integration sites, we established a set of reporter lines that produced a strong and specific punctate nucleus signal when expression is induced, with little to no unwanted background expression. We then developed a procedure to capture the entire microdissected larval brain volume by light sheet microscopy at multiple angles and with high resolution, imaging only animals within the ~two-hour developmental time window of the wandering third instar larval stage(Ainsley et al., 2008). These multiview datasets were then processed to register, fuse, and deconvolve the entire larval brain volume. The volume was then segmented and cell numbers automatically quantified (Fig 1a-d).

To evaluate the reliability of the procedure, we began by comparing automated counts of distinct neuronal subtypes with manual counting. We separately labelled all dopaminergic neurons (Fig. 1e, MovieS1) [TH>GAL4](Friggi-Grelin et al., 2003a; Mao and Davis, 2009), serotonergic neurons (Fig. 1f, MovieS2)[Trh>GAL4] (Alekseyenko et al., 2010), and neurons that produce both types of neurotransmitter (Fig. 1g, MovieS3)[Ddc>GAL4] (Lundell and Hirsh, 1994) in larval brains. Quantification revealed a high level of concordance (Fig 1h, +/-0.21%, n= 5 for TH>GAL4, +/-1%, n= 5 for Trh>GAL4, +/-0.38%, n= 6 for Ddc>GAL4) between automated and manual measurements of these neuronal subtypes establishing confidence in the procedure.

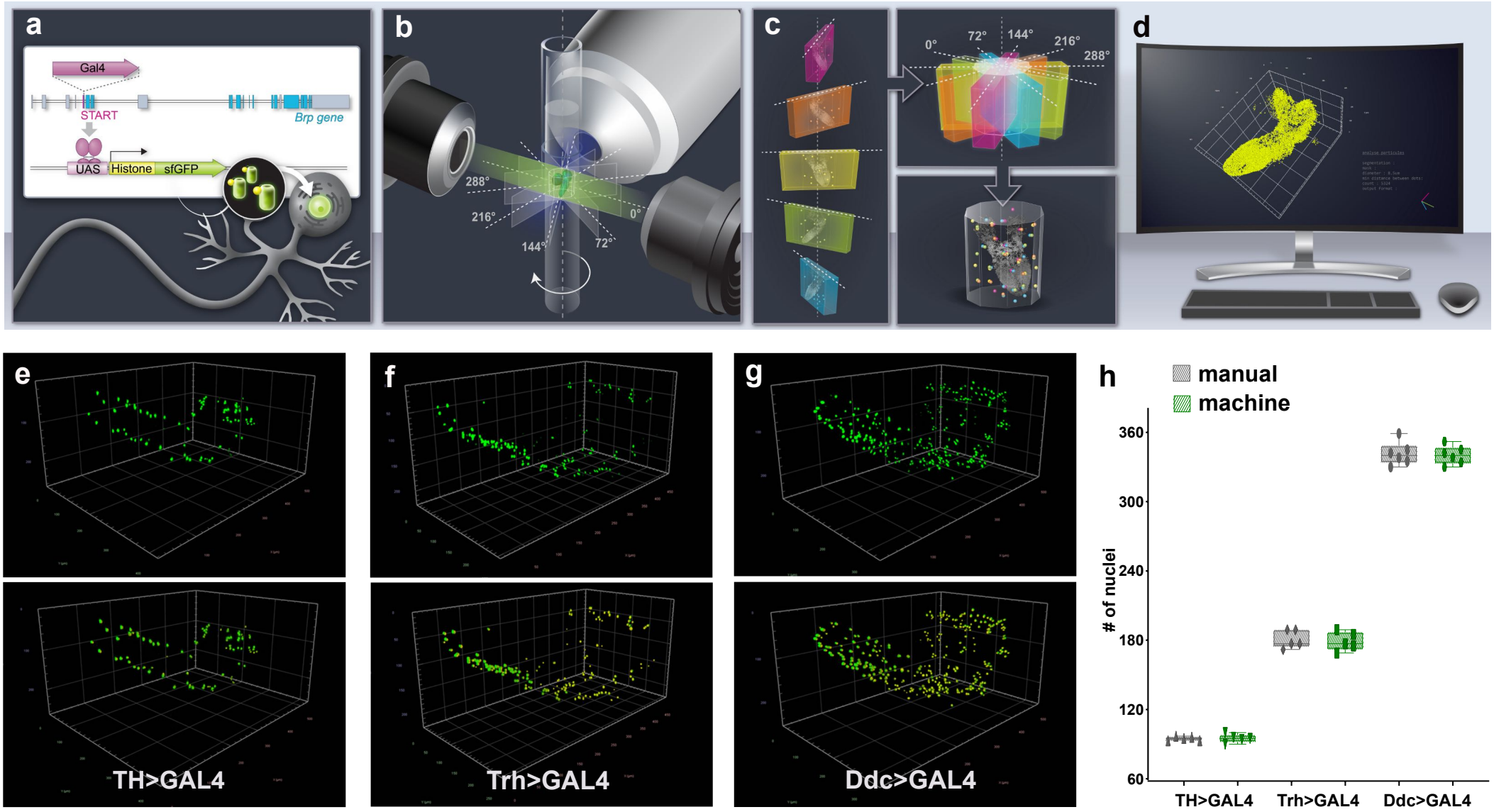


Fig 1. Intact whole brain quantitation pipeline schematic and validation

(**a-d**) The illustration of intact whole brain genetic, imaging and computational pipeline. (**a**) Genetic reagents: GAL4 is introduced to the exonic sequences of genes encoding synaptic proteins (e.g. brp) to capture their expression pattern with high fidelity. GAL4 expression regulates the production of UAS fluorescent-histone reporters which target to the nucleus of cells producing a punctate signal. (**b**) Imaging: intact whole brains were imaged at high resolution using light-sheet microscopy. Images are captured at 5 different angles with 72-degree intervals. (**c**) Assembly: multiview light sheet images are registered, fused and deconvolved. (**d**)Quantitation: volume is segmented, nucleus number and relative position is measured. Three dimensional co-ordinates of the center mass of every nucleus can be calculated to produce a point cloud of nuclei positions.(**e-h**) Pipeline validation. Three dimensional images before (above) and after (below) segmentation panels of (**e**) dopaminergic [TH>GAL4] neurons, (**f**) serotonergic neurons [Trh>GAL4] and (**g**) dopa decarboxylase expressing [Ddc>GAL4] neurons. (**h**) Quantification of manual and automated counting of these volumes produce similar results. Scale squares in **e** and **g** are 100μm and **f** is 50μm.

Number of neurons and glia in the female larval brain. Encouraged by our neuronal subset quantitation results, we next sought to generate GAL4 lines for genes likely to be expressed only in active larval neurons with synaptic connections but not by neuroblasts or by immature neurons. We biased towards generating GAL4 insertions within endogenous loci in order to reproduce endogenous patterns of gene expression with high fidelity.

Bruchpilot (Brp) is a critical presynaptic active zone component widely used to label *Drosophila* synapses, including for large-scale circuit analyses (Wagh et al., 2006). We employed CRISPR/Cas9 genome editing to insert GAL4 within exon 2 of the *brp* gene, utilising a T2A self-cleaving peptide sequence (Diao et al., 2015) to efficiently release GAL4. While this exonic insertion generated a hypomorphic allele of Brp (data not shown) when homozygous, the line was employed in heterozygotes to capture Brp protein expression with high fidelity. To complement this line, we used the Trojan/MiMIC technique (Diao et al., 2015), to generate a GAL4 insert in the *syt1* gene, which encodes Synaptotagmin 1 (Littleton et al., 1994), the fast calcium sensor for synaptic neurotransmitter release (Quiñones-Frías and Littleton, 2021). Lastly, we examined a transgenic line where the promoter of *nsyb* (*neuronal synaptobrevin*) (Deitcher et al., 1998), which encodes an essential presynaptic vSNARE (Südhof and Rothman, 2009), is used to control GAL4 expression (Aso et al., 2014). All three lines were expressed in a similar pattern, labelling a substantial fraction but not all of the total cells in the larval brain (Fig. 2a-c, MovieS4-S6). These lines contrasted with the widely used elav>GAL4(Lin and Goodman, 1994), which appeared to be expressed in larval neurons, but also apparently in immature neurons and potentially in some glia as well(Berger et al., 2007) (Fig. S1a,b).

To ensure that the cells labelled by our lines were exclusively neurons, we compared their expression to that of glial cells labelled by glial specific transcription factor Repo(Xiong et al., 1994; Lin and Potter, 2016) using independent and mutually exclusive QF2 dependent labelling. We found

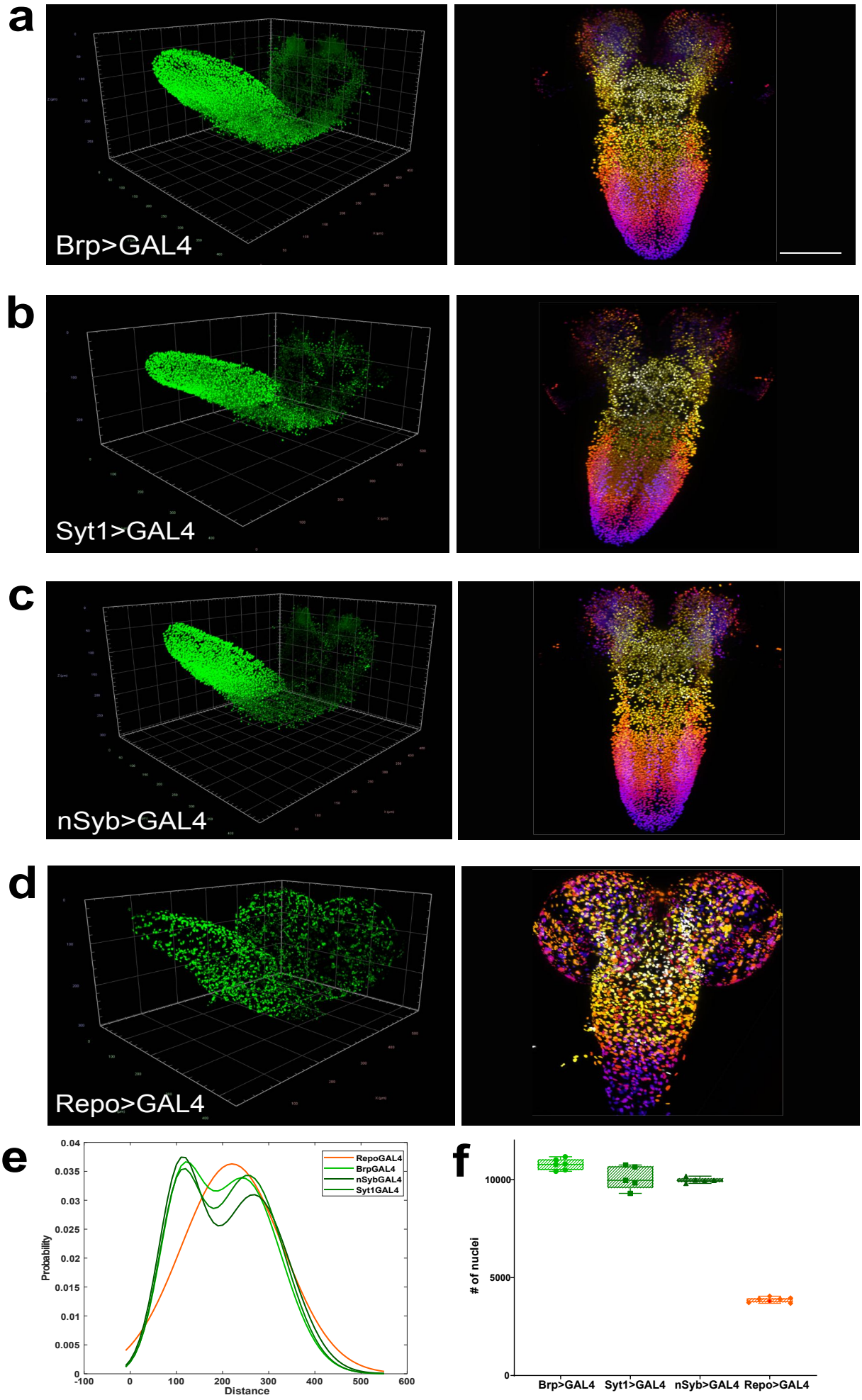


Fig 2. Whole brain quantitation of neurons and glia in the female larval brain.
(**a-d)** Multiview deconvolved images (left) and z-stack projections (right) from brains of (**a**) Brp>GAL4, (**b**) Syt1>GAL4, (**c**) nSyb>GAL4 and (**d**) Repo>GAL4 lines. (**e**) Distribution of inter-nuclei distances for each line. (**f**) Quantification of the number of labelled nuclei in each line. **a-d**: left; scale squares **a** and **c** = 50μm, **b** and **d** = 100μm; right images are identical magnification, scale bar = 100μm. **Jiao et al. Figure 2**

complete exclusion of cells labelled by Brp, Syt and nsyb from cells labelled by Repo (Fig. 2d, Movie S7-S9), consistent with the Brp, Syt1 and nsyb lines labelling neurons but not glial cells.

To further compare these lines, beginning with brains from female animals, we calculated three dimensional coordinates for the centre of mass of all nuclei labelled in the Brp, Syt1 and nSyb GAL4 lines to generate point cloud mathematical objects and compared them to point clouds of glial nuclei labelled by the Repo GAL4 line. We then plotted and compared the distributions of inter-nuclei distances in these lines. Using this measurement, we found that inter-nuclei distance of glial cell nuclei exhibited a unimodal distribution (Fig. 2e). In contrast, all three neuronal lines exhibited a bimodal distribution of inter-nuclei distances (Fig. 2e). We thus observed two patterns of labelled nuclei, one shared among neuronal lines and the other distinct for glia (Fig. 2e).

We next counted the number of nuclei labelled by these neuronal and glial lines, again beginning with females (Fig. 2f). We found that Brp-labelled brains had 10776(+/- 2.65%, n= 6) neurons, Syt1-labelled brains had 10097(+/- 5.96%, n= 5) neurons and nSyb-labelled brains had 9971(+/- 1.35%, n= 5) neurons (Fig. 2f). We tested the statistical difference in the numbers of neurons labelled by these lines and found that while nSyb>GAL4 and Syt1>GAL4 labelled brains were not statistically different from each other, Brp>GAL4 did label significantly more neurons than either Syt1 or nSyb (Brp>GAL4 vs Syt1>GAL4 + 6.72%, p=0.03, Brp>GAL4 vs nSyb>GAL4 + 8.07% p=0.01). Averaging across the lines, we found that female third instar larval brains had 10312 +/- 5.03%, n= 16 neurons. To ensure that our method did not introduce bias in dense data sets, we also manually counted a Brp>GAL4 labelled brain and compared it to the automated count. Similar to brains with sparse labelling, we found good agreement between manual and automated quantification with a difference of just 14 neurons (9430 nuclei manual vs 9444 nuclei automated for this individual brain).

We next counted the number of glia labelled by the Repo GAL4 line (Fig 2d,f, Movie S10). We measured 3860 +/- 3.37%, n= 7 glia in the female brain. This amounted to 37% of the number of neurons, far more than the previously estimated ~10% (Meinertzhagen, 2018; Raji and Potter, 2021).

In sum, we found that female *Drosophila* larval brains had 10312 neurons, between ~15 to 30% fewer than was previously predicted and 3-fold more glia.

Males have fewer neurons and more glia than females. We next carried out a similar analysis on the brains from male larvae (Fig. 3a-c). We found that Brp>GAL4 labelled 9888 (+/-3.15%, n= 5) neurons, Syt1>GAL4 labelled 9012 (+/-3.8%, n= 5) neurons, and nSyb>GAL4 labelled 9286 (+/-5.38%, n= 5) neurons in male larvae (Fig. 3e). In males, Brp>GAL4 did not label significantly more neurons than nSyb>GAL4 but did label more than Syt1>GAL4 (Brp>GAL4 vs Syt1>GAL4 + 876, p=0.01), while the number of neurons labelled by nSyb>GAL4 was not significantly different from Syt1>GAL4, as was found in females. Averaging across the lines we found that male third instar larval brains had 9396 +/- 5.59%, n= 15 neurons, significantly fewer than those of females (-9.75%, P<0.0001). This difference was also consistent within individual genotypes with Brp>GAL4 labelling (-8.98%, P=0.0008), Syt1>GAL4 labelling (-12.04%, P=0.008) and nSyb>GAL4 labelling (-7.38%, P=0.0182) less neurons in males than in females.

We also counted the number of glia labelled by Repo>GAL4 in males (Fig 3d, e). We found that males had 4015 glia, again far more than previous estimates. The number of glia in the male larval brain was also significantly more than in females (+3.86%, P=0.0284). In summary, male *Drosophila* larva have significantly fewer CNS neurons than females but significantly more glia.

Topological analysis detects significant structural differences between males and females. We next wished to determine whether the differences between the point clouds derived from the positions of neuronal nuclei of males and females went beyond simple numerics. To do this we applied the tools of topological data analysis to summarize brains in terms of multi-scale topological structures (Expert et al., 2019). These topological summaries could then be used as the classification features in a support vector machine (SVM). Since the total number of point clouds was relatively small for this type of analysis, we down-sampled each whole brain point cloud randomly to 8000 points 100 times, producing a total of 3100 point clouds, for each of which we then computed a

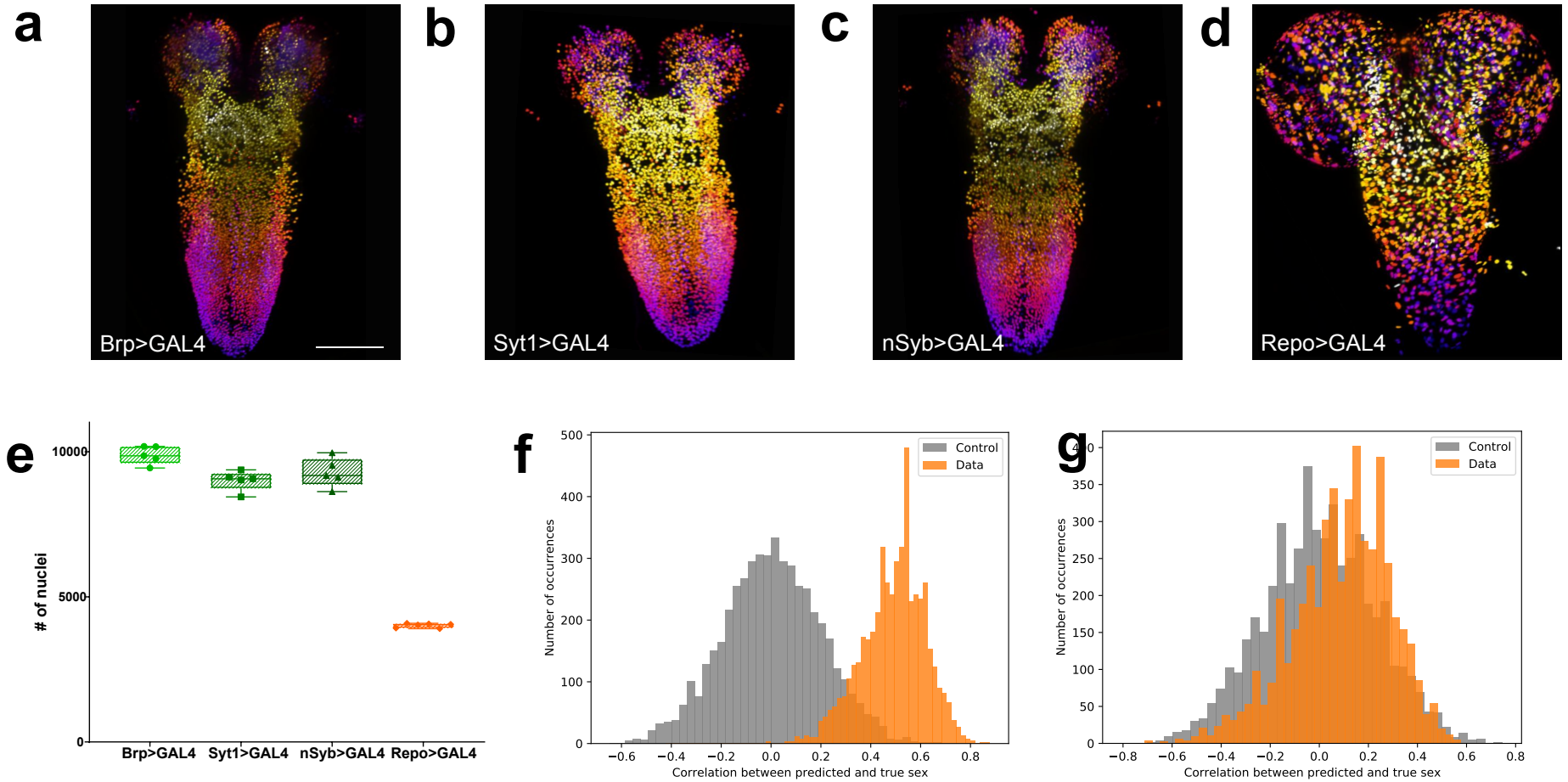


Fig 3. Quantitation of neurons and glia in the male larval brain and topological comparison of sex differences
(**a-d**) Example z-stack projections from male larval brains of (**a**) Brp>GAL4, (**b**) Syt1>GAL4, (**c**) nSyb>GAL4 and (**d**) Repo>GAL4 labeled lines. (**e**) Quantification of the number of labelled nuclei in each line. (**f**) The distribution of correlations between the ground truth and the prediction made by the SVM using topological features is indicative of sexual dimorphism of the higher order structure of neuron point clouds (**g**) Simpler point cloud features such as properties of the distributions of inter-nuclei distances are not indicative of this. a-d: identical magnification, scale bar =100μm

Jiao et al. Figure 3

certain topological summary, called the *degree-1 persistence diagram of its alpha complex* (Edelsbrunner and Mücke, 1994). After fixing the necessary hyperparameters, sex classification experiments were run across 5000 random train/test splits of the topological summaries. In each split, the summaries derived from subsamplings of a single point cloud (brain) were either all in the training set or all in the testing set, to avoid leaking information. Each time, the SVM was trained once with the animals true sex as the target class and once with a randomly assigned sex as target, as a control. We then computed the Pearson correlation between the classifier's output on the testing set and the true (respectively randomized) sex of the animal.

The 5000 splits were used to produce 5000 correlations with the true sex and 5000 correlations with a randomly assigned sex. The distribution of these correlations (Fig. 3f), exhibiting clearly that the SVM is able to extract the sex of the animal reliably: only about 1.9% of the splits result in a higher correlation in the control set than in the true data. Moreover, repeating the procedure with simpler point cloud features, like properties of the distributions of inter-nuclei distances, did not produce a significant signal (Fig. 3g). Thus, the pattern, which seems hard to describe concisely, is not revealed through simpler descriptors of the neuron configurations, leading us to suspect brain sexual dimorphism extends to deeper features of organisation that are both subtle and widely distributed. These results, in addition to the differences in total cell numbers, support sexual dimorphism of male and female brains at the larval stage.

Potassium channel family members have different densities in the brain. Having established a baseline of total numbers of neurons in the larval brain, we next sought to deploy the quantification pipeline to measure the expression frequency of key neuronal function genes throughout the brain. We chose to examine the family of voltage-gated potassium channels, which are essential for many aspects of neuronal function and for which *Drosophila* studies defined the founding members(McCormack, 2003). We generated GAL4 insertions in the Shaker (Sh) [Kv1 family], Shab (Sb) [Kv 2 family], Shaw (Sw) [Kv3 family] and Shal (Sl) [Kv4 family](McCormack, 2003) genes using the Trojan/Mimic technique (Diao et al., 2015). As the Sh gene is x-linked, we carried out our

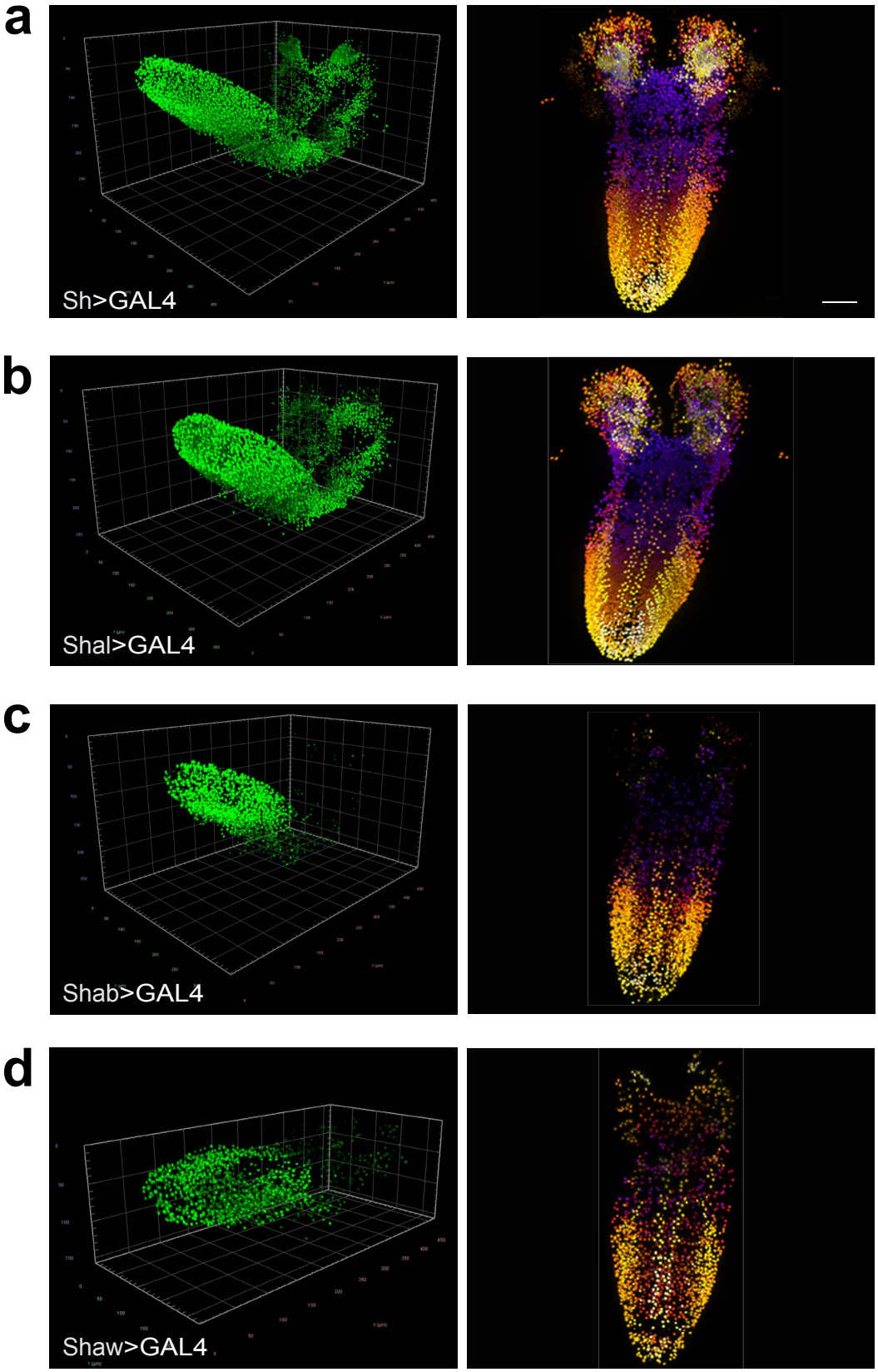


Fig 4. Quantitation of the number of neurons expressing voltage-gated potassium channels.
(**a-d**) Multiview deconvolved images (left) and z-stack projections (right) of potassium channel family members: (**a**) Sh>GAL4, (**b**) Shal>GAL4, (**c**) Shab>GAL4 and (**d**) Shaw>GAL4. (**e**) Quantification of the number of labelled nuclei in each line. a-d: left, scale squares = 50μm, right, identical magnification, scale bar = 50μm.

quantitation analysis in male brains only to avoid potential gene dosage effects. To determine whether our GAL4 reporter lines had patterns of expression consistent with the known properties of these channels, we examined the expression of all four lines in motor neurons, where functional activity for Shaker, Shab, Shaw and Shal has previously been demonstrated by electrophysiological measurements (Covarrubias et al., 1991; Ryglewski and Duch, 2009). We found that the GAL4 reporters for all 4 channels were expressed as expected in motor neurons (Fig. S2), consistent with accurate reproduction of known expression of these proteins.

We next examined the expression frequency of these genes in the entire brain (Fig. 4a-d, Movie S11-S14). We found that Shaker and Shal were expressed in large numbers of neurons 8204 +/- 5.67% and 8261 +/- 3.1%, though significantly less (-12.7 % and -12.1% $p<0.0001$) than the average number of all male neurons (Fig. 4a,b,e). In contrast, Shab (3057+/- 8.21% n=10) and Shaw (1737 +/- 4.3% n=11) were expressed in smaller numbers of neurons (Fig. 4c-e), with expression observed in only 32.5% or 18.5% of total male neurons respectively, suggesting more discrete functions within CNS neurons and contrasting with the collective expression of all four genes within motor neurons (Figure S2). In particular, Shab and Shaw had very reduced expression in the brain lobes of larva (Fig 4c,d) compared with Shaker and Shal (Fig 4a,b). These results establish that our genetic-imaging pipeline can enable quantitation of the expression frequency of families of genes essential for neuronal properties in the entire brain.

DISCUSSION

Establishing the number and identity of cells in the brain is a foundational metric upon which to construct molecular, developmental, connectomic and evolutionary atlases of central nervous systems across species (Lent et al., 2012; Devor et al., 2013). Here, we develop and deploy a methodological pipeline to label discrete cell types in the intact *Drosophila* brain, with genetic reporters designed to facilitate the subsequent segmentation and automated quantification of cell types, in addition to capturing positional coordinates of relative nucleus position throughout the organ. Using this toolset, we find fewer active neurons, as defined by expression of synaptic protein

genes, in the *Drosophila* larval brain than had been previously predicted and substantially more glia. We also discover previously unsuspected differences in both neuron and glial density and brain topology at the larval stage, when external sex organs are absent, with females possessing both more neurons and but fewer glia than males. Topological analysis of the point cloud derived from neuronal nucleus position, which detects potentially subtle and complex geometric structure in the data, also strongly supports the existence differences between males and females. In addition, deploying these tools, we find that while all members of the *Drosophila* voltage-gated potassium channel family are expressed in motor neurons, consistent with prior mutant analyses, the Kv2 channel Shab and Kv3 channel Shaw are expressed in a much smaller number of neurons in the CNS than the Kv1 channel Shaker and the Kv4 channel Shal, suggesting conclusions drawn about the coordinated activity of these channels from studies of motor neurons may not be broadly applicable across the CNS, where these genes are frequently not co-expressed.

A number of semi-quantitative methods have been employed to estimate the number of neurons in the brains of humans and model organisms, including *Drosophila* (Lent et al., 2012; Keller et al., 2018). For example, the number of neurons or other cells in the brain has been estimated using stereological counting of sub-regions. A major limitation of this approach is the assumption of homogenous cell density across the organ or within subregions, which is not supported by the high variability of counts even between samples of similar regions, and thus likely introduces large errors (von Bartheld et al., 2016; Keller et al., 2018). Rough extrapolation of neuronal counts of electron microscope volumes of the regions of the *Drosophila* larval brain had suggested an estimate of ~15,000 neurons (Meinertzhagen, 2018; Eschbach and Zlatic, 2020). An alternate approach is isotropic fractionation, where all cells in large regions or the entire brain are dissociated to produce a homogeneous single-cell suspension. Nuclei in the suspension can then be labelled by immunohistochemistry and cells in a subvolume counted in a Neubauer chamber to estimate the total number of cells present. Limitations of the approach include the necessity to ensure complete dissociation of cells while avoiding tissue loss, the requirement for homogenous antibody labelling,

and highly accurate dilution (Deniz et al., 2018). This approach has recently been used to estimate the total number of neurons and glia in the adult *Drosophila* brain and suggested a number of 199,000 neurons (Raji and Potter, 2021), twice prior estimates (Scheffer and Meinertzhagen, 2019; Allen et al., 2020). In contrast to our results in the larval brain, this study found no significant differences in the number of neurons between sexes and also found that 'non-neuronal' cells, which should include glia, accounted for less than 9% of the total cells counted. In addition to the inherent inaccuracy of the isotropic fractionation technique, which the authors both observed and acknowledge (Raji and Potter, 2021), their use of anti-Elav antibody labelling, which can label some glia in addition to neurons (Berger et al., 2007), or perhaps differences in life stage, could explain some of the discrepancies between our results.

An unpredicted result from our whole brain neuron quantitation was substantial differences in neuron and glial numbers between the sexes in larva. In adult *Drosophila*, sexually dimorphic neural circuitry has been observed in olfactory system (Kimura et al., 2005), and human females have also been reported to have more olfactory bulb neurons and glia than males (Oliveira-Pinto et al., 2014). While sex-specific behavioural differences are obvious in adult *Drosophila* (Jazin and Cahill, 2010), few sexually dimorphic behavioural differences have been reported in larva (Aleman-Meza et al., 2015). However male and female larva do differ in nutritional preference (Rodrigues et al., 2015; Davies et al., 2018), which could potentially account for some aspects of the dimorphism we observe. In addition to differences in total cell numbers, our topological methods, which take into account multi-scale structure, suggests that differences in brain structure between the sexes is both subtle (in the mathematical sense) and non-localised in nature, and indeed are not observable with simpler analysis methods of brain organisation.

In addition to enabling precise counting of genetically labelled cells, our method allows the relative measurement of discrete cell types or gene expression frequencies throughout the brain. For example, the relative frequency of glial cells to neurons in the human brain has been long been debated (von Bartheld et al., 2016) and in the adult *Drosophila* brain it has been suggested there

are 0.1 glial per neuron (Kremer et al., 2017; Scheffer and Meinertzhagen, 2019; Raji and Potter, 2021). In the larval *Drosophila* brain, we found closer to 0.4 glial cells per neuron on average, more similar to the glial-neuron ratios reported for rodents or rabbits (Verkhratsky and Butt, 2018). Our approach may also allow assignment of potential functional classes of neuron types. For example, from our examination of voltage-gated potassium channel family gene expression, all are collectively expressed in motor neurons, however the Shab and Shaw genes have more discrete expression patterns in other CNS neuron classes, potentially imbuing these neurons with unique functional characteristics (Chow and Leung, 2020). Future multiplexing of binary genetic expression systems and reporters (Simpson, 2009; del Valle Rodríguez et al., 2011; Diao et al., 2015) should enable neurons or glia to be further quantitively sub-classified by gene expression features throughout the entire intact brain.

MATERIALS AND METHODS

Drosophila stocks

The following stocks were employed - y[1] w[67c23]; Mi{PT-GFSTF.0}Syt1[MI02197-GFSTF.0]/CyO (BDSC#59788)(Venken et al., 2011), y[1] w[*] Mi{y[+mDint2]=MIC}Sh[MI10885] (BDSC#56260), y[1] w[*];Mi{y[+mDint2]=MIC}Shal[MI10881] (BDSC#56089)(Venken et al., 2011), y[1] w[*]; Mi{y[+mDint2]=MIC} Shab[MI00848] (BDSC#34115)(Venken et al., 2011), nSyb>GAL4(R57C10)(Pfeiffer et al., 2008), repo>GAL4 (BDSC#7415)(Sepp et al., 2001), repo>QF2 (BDSC#66477)(Lin and Potter, 2016), Shaw>TrojanGAL4 (BDSC#60325)(Venken et al., 2011; Li-Kroeger et al., 2018), Ddc>GAL4(BDSC#7009)(Feany and Bender, 2000), TH>GAL4(BDSC#8848)(Friggi-Grelin et al., 2003b), Trh>GAL4(BDSC#38389)(Alekseyenko et al., 2010), UAS>H2A-GFP(Sherer et al., 2020), QUAS>H2B-mCherry(Sherer et al., 2020), Brp>GAL4 (this manuscript), UAS>H2A::GFP-T2A-mKok::Caax (this manuscript). All lines were raised on standard media at 25°C, 50%RH.

Generation of Brp>GAL4 exon 2 insertion line.

A GAL4.2 sequence was inserted in genome, immediately after the start codon of the Brp-RD isoform using CRISPR based gene editing employing the following constructs. *Brp gRNA pCDF3:* Two gRNA sequences targeting each side of the insertion location in exon 2 of *brp*, were selected using the FlyCRISPR algorithm (http://flycrispr.molbio.wisc.edu/), consisting of 20 nucleotides each (PAM excluded), and predicted to have minimal off-targets. Each individual 20-nucleotide gRNA sequence were inserted into pCFD3 plasmid (Addgene #49410) using the KLD enzyme mix (New England Biolabs). *Brp>GAL4 insertion construct*: The 7 following PCR amplified fragments were assembled using HIFI technology - (1) 1198bp Homology arm covering 5'UTR until 5' target site; (2) the region between 5' target site and the start codon were amplified from drosophila nos-cas9 (attp2) genomic DNA (a modified Pam sequence was inserted using overlapping primers); (3) Linker-T2A-GAL4.2 sequence was amplified from pBID-DSCP-G-GAL4 (Wang et al., 2012) (the linker-T2A sequence was added upstream of the forward primer); (4) P10-3'UTR was amplified from pJFRC81-10XUAS-IVS-Syn21-GFP-p10 (Addgene 36432); (5) 3xP3-Hsp70pro-dsRed2-SV40polyA selection cassette, flanked by two LoxP sites, was amplified from pHD-sfGFP Scareless dsRed (Addgene 80811); (6)

The region covering the end of DsRed cassette until 3' target site and (7) the 1079bp Homology arm 2 covering from the 3' target site to exon 2, were amplified from Drosophila nos-cas9 (attp2) genomic DNA. Full length assembly was topo cloned in zero-blunt end pCR4 vector (Invitrogen), all constructs have been verified by sequencing (Microsynth AG, Switzerland) and injections were carried out into a nos-cas9 [attp2] strain (Ren et al., 2013). Correct insertion of GAL4 was verified by genome sequencing. All primer sequences are included in Table S1.

Construction of UAS H2A::GFP-T2A-mKok::Caax

PCR amplifications were performed using Platinium Superfi polymerase (Invitrogen). The three PCR fragments were assembled together using Hifi technology (Invitrogen) - (1) Histone2A (H2A) cDNA was amplified from *pDESTP10 LexO>H2A-GFP* template [Gift from Steve Stowers] with a synthetic 5'UTR sequence (syn21) added upstream to H2A on the forward primer; (2) sfGFP was amplified from template pHD-sfGFP Scareless dsRed (Addgene 80811) and (3) mKok amplified from pCS2+ ChMermaid S188 (Addgene 53617) with the CAAX membrane tag sequence (Sutcliffe et al., 2017) added at the 3' end of the protein using the reverse primer. A *Thosea asigna* virus 2A(T2A) self-cleaving peptide sequence (Diao et al., 2015), was inserted between sfGFP and mKok, using sfGFP reverse and mKok forward overlapping primers. The full length assembly was TOPO cloned into pCR8GW-TOPO vector (Invitrogen) generating pCR8GW-H2A::GFP-T2A-mKok::Caax. The insert, H2A::GFP-T2A-mKok::Caax was , then, transferred to pBID-UASC-G destination vector (Wang et al., 2012) using LR II clonase kit (Invitrogen) to generate pBID-UAS>H2A::GFP-T2A-mKok::Caax. The transgene was generated by injection into the JK66B landing site. All primer sequences are included in Table S1.

Generation of novel Trojan GAL4 lines.

The MiMIC lines generated by the group of Hugo Bellen(Venken et al., 2011) were acquired from the Bloomington Stock Center. Conversion of Mimic lines to Trojan GAL4 lines lines was performed as described previously(Diao et al., 2015).

Larval brain preparation and image acquisition

Wandering 3rd instar larvae were dissected in 1x PBS (Mediatech) and fixed with 4% formaldehyde (Sigma-Aldrich) for 20 mins. 1x PBS were added to remove the fixative, and then brains were

dissected(Hafer and Schedl, 2006) and rinsed with 1xPBS with 4% Triton-X 100 for 2 days at 4°C. After rinses, brains were embedded in 1% low melting temperature agarose (Peq gold) mixed with 200nm red fluorescent beads (1:50000), then introduced into a glass capillary and positioned well separated from each other. After solidification of the agarose, the capillary was mounted to sample holder, transferred to a Zeiss Lightsheet Z.1 microscope and the samples were extruded from the capillary for imaging. Images for brains were acquired with a 20x/1.0 Apochromat immersion detection objective and two 10x/0.2 illumination objectives at 5 different views, with 1μm z-intervals.

Image processing and data analysis

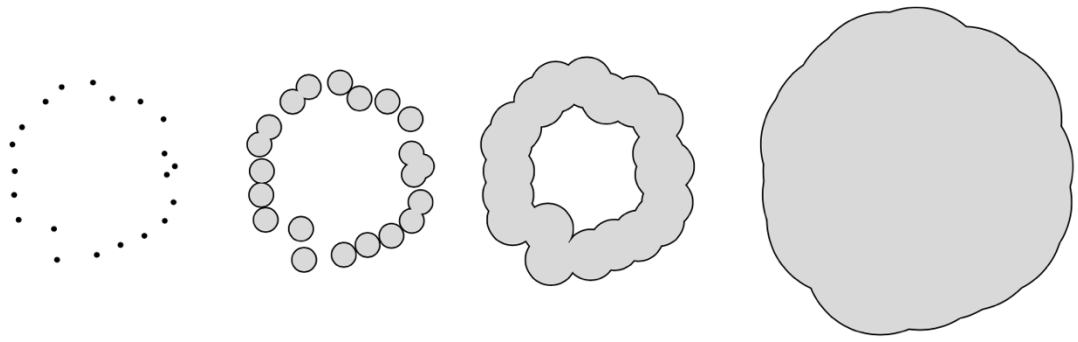
Collected multiview datasets were registered and fused with the Fiji Multiview Reconstruction plugin(Preibisch et al., 2010; Schindelin et al., 2012). Image datasets after Multi-view deconvolution were analyzed with Vision4D 3.0.0 (Arivis AG). A curvature flow filter was first used to denoise the image dataset. Subsequently, a Blob Finder algorithm(Najman and Couprie, 2003) was applied to detect and segment bright rounded 3D sphere-like structures in the images with 4.5μm set as the diameter. Segmented objects with volume less than 15μm^3 were removed from analysis by segmentation filter to avoid unspecific signals. Subsequently, the number of nuclei and the x, y, z coordinates of the center geometry of each nucleus were output from Vision4D. Where manual counting was employed (Fig 1 and a randomly selected Brp>GAL4 labelled brain), Vision4D was used to visualize and iteratively proceed through and manually annotate the dataset. Example whole brain datasets where functional neurons or glia are labelled are available (Jiao and McCabe, 2021a, 2021b). Raw co-ordinates of the center of geometry for the nuclei for whole male and female brains are available in Supplementary Dataset1.

Mathematical analysis

We trained a machine learning classifier, specifically an SVM (support vector machine), on the brain nuclei positions, in order to evaluate its power in determining characteristics of the animal from which it was derived. Correlation significance (classification power) is determined by comparing the performance of the SVM on the actual classification task to one wherein each larva is randomly assigned a class.

Mathematically speaking, the nuclei positions from a single brain form a point cloud, a finite set of points in R^3. A possible, naive approach to SVM feature selection for point clouds would be to consider the mean, variance, or other modes of the distribution of pairwise distances within the cloud. These real-valued features could then be passed through, for example, radial basis function kernels for use in SVMs. We focused on very different kind of features, namely ones obtained from the topology of the point clouds. When the point cloud is of low dimension, such as the three-dimensional point clouds arising from nuclei position data, the following approach is relevant. Let X be a finite point cloud in R^3. For any $r \geq 0$, we let X_r denote the same point cloud, but with each point replaced by a ball of radius r. As r increases, the sequence formed by the X_r expresses different topological features of X. By topological features, we here mean the presence or absence of multiple connected components, unfilled loops, and unfilled cavities.

The figure below illustrates this process in the case of a synthetic 2-dimensional point cloud, but the idea extends to any dimension including whole brain point clouds. When r is small, X_r is topologically very similar to $X = X_0$, and is essentially a collection of disjoint points. When r is very large, X_r is topologically very similar to X_∞, i.e., one giant, featureless blob. As the sequence X_r progresses through the continuum of scales between these two trivial extremes, it undergoes non- trivial topological changes: components merge, and loops form and later get filled. In higher dimensions, cavities of various dimensions likewise form and get filled in.



A small 2-dimensional point cloud X viewed at four different scales $0 < a < b < c$, forming the filtration $X = X_0 \subset X_a \subset X_b \subset X_c$.

In the parlance of topological data analysis (TDA), we refer to this appearance and disappearance of topological structures as the *birth* and *death* of homology classes in various degrees. We capture the whole life cycle with a mathematical object called the *persistent homology* of the point cloud, which can be fully described by its *persistence diagram*, a planar collection of points (labelled by multiplicity), whose coordinates encode the birth and death of homological features. For the filtration in the figure above, the persistence diagram that tracks 1-dimensional features (i.e., unfilled loops) contains only a single point with coordinates (x, y). Here the first coordinate, x, is the radius at which the loop is first formed, and the second coordinate, y, is the radius at which the loop has just been filled in. In the example it is clear that $a < x < b < y < c$.

As multisets of points in the plane, persistence diagrams are not immediately usable as features for SVMs. One way to vectorize persistence diagrams and thus render them digestible by SVMs is to define kernels based on the diagrams, with the heat kernel(Reininghaus et al., 2015) being an oft-used candidate with nice properties. For persistence diagrams P and Q, the heat kernel can informally be defined by the inner product of two solutions of the heat equation — one with an initial condition defined by P, and the other with one defined by Q.

In this analysis in this manuscript, we calculated the persistent homology of the alpha complex(Edelsbrunner and Mücke, 1994) of the point clouds, using GUDHI(The GUDHI Editorial Board, n.d.). The heat kernels were computed using RFPKOG(Spreemann, n.d.). Only the persistence diagrams in degree 1 were used. Since the number of whole brain point clouds was relatively small, we subsampled each one randomly to 8000 points 100 times, producing a total of 3100 point clouds. This was done both in order to test the stability of the method and to ensure that the variability in the number of points in each cloud is not the source of any signal.

The hyperparameters involved, i.e., the SVM regularizer and the heat kernel bandwidth, were determined by a parameter search in the following way. Six point clouds from males and six from females were randomly selected. All 100 subsampled versions of each of these 12 constituted a training set, for a total of 1200 training point clouds. The remaining 1900 subsampled point clouds constituted the testing set. The Pearson correlation between the gender predicted by the SVM on the testing set and the ground truth was computed for each choice of hyperparameters, and a choice in a stable region with high correlation was selected: a regularization parameter C = 10 in the notation of Pedregosa *et al* (Pedregosa et al., n.d.) and a bandwidth of σ = 1/100 in the notation of Reininghaus *et al* (Reininghaus et al., 2015). For the simple distance distribution features, a similar parameter selection process yielded C=10 and a radial kernel bandwidth of 10^5.

Motor neuron preparation and confocal microscopy

Larval fillets from the 3rd instar larvae were dissected and fixed with 4% formaldehyde (Sigma-Aldrich) for 20 mins. After fixation, samples were rinsed with 1x PBS and were washed in PBT overnight at 4°C, and then mounted in VECTASHIELD antifade mounting medium. Z-stack images were taken from Leica SP8 upright confocal microscope.

Statistical Analysis

Column statistics analyses were performed using GraphPad Prism 9 (GraphPad Software). For Fig.1, statistical significance was determined by unpaired t test. For Fig.2-4, statistical significances were determined by Ordinary one-way ANOVA, followed by a Tukey's honestly significant difference test when multiple comparisons were required. The distribution analysis in Fig.2 were performed using matlab (MathWorks). Distances between nuclei coordinates were calculated in matlab and plotted in a histogram of distance distribution.

SUPPLEMENTAL MATERIAL: Supplemental Figures, Movies, Tables and Datasets are available at DOI: 10.5281/zenodo.5643020

ACKNOWLEDGEMENTS: This work was supported in part using the resources and services of the BioImaging & Optics Platform (BIOP) Research Core Facility at the School of Life Sciences of EPFL and we are especially thankful for the assistance of Arne Seitz. We are grateful to Hugo Bellen, Benjamin White, Gerry Rubin, Vanessa Auld, Christopher Potter, Mel Feany, Serge Birman, Ed Kravitz and Pavel Tomancak for generating Drosophila stocks or Software. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study.

FUNDING: This work was supported by the Swiss National Science Foundation grant number: 31003A_179587 to B.M.

AUTHOR CONTRIBUTIONS: W.J, G.S, E.R, S.B, J.A, Y.S., S.S, K.H. and B.M. conceived and designed the study. W.J performed the majority of experiments and analysed the data, assisted by G.S. and Y.S. E.R. generated all novel transgenes, the Brp>GAL4 genome edit and generated illustrations. S.B. generated the Syt1 Trojan line and other useful reagents. S.S. generated the histone fusion transgenes. W.J., G.S, K.H. and B.M. wrote the manuscript.

COMPETING INTERESTS: The authors declare that they have no competing interests**.**

DATA AND MATERIALS AVAILABILITY: Individual whole brain quantitation data and raw co-ordinates of the center of geometry for the nuclei for whole male and female brains and other supporting figures, tables and movies are available as Supplemental Data doi: 10.5281/zenodo.5643020. Example unprocessed whole brain microscopy data is also available for neurons - doi:10.5281/zenodo.5585334 and for glia doi:10.5281/zenodo.5585358.

REFERENCES

Ainsley JA, Kim MJ, Wegman LJ, Pettus JM, Johnson WA. 2008. Sensory mechanisms controlling the timing of larval developmental and behavioral transitions require the Drosophila DEG/ENaC subunit, Pickpocket1. *Developmental Biology* **322**:46–55. doi:10.1016/j.ydbio.2008.07.003

Alekseyenko OV, Lee C, Kravitz EA. 2010. Targeted Manipulation of Serotonergic Neurotransmission Affects the Escalation of Aggression in Adult Male Drosophila melanogaster. *PLOS ONE* **5**:e10806. doi:10.1371/journal.pone.0010806

Aleman-Meza B, Jung S-K, Zhong W. 2015. An automated system for quantitative analysis of Drosophila larval locomotion. *BMC Dev Biol* **15**:11. doi:10.1186/s12861-015-0062-0

Alivisatos AP, Chun M, Church GM, Greenspan RJ, Roukes ML, Yuste R. 2012. The brain activity map project and the challenge of functional connectomics. *Neuron* **74**:970–974. doi:10.1016/j.neuron.2012.06.006

Allen AM, Neville MC, Birtles S, Croset V, Treiber CD, Waddell S, Goodwin SF. 2020. A single-cell transcriptomic atlas of the adult Drosophila ventral nerve cord. *eLife* **9**:e54074. doi:10.7554/eLife.54074

Aso Y, Sitaraman D, Ichinose T, Kaun KR, Vogt K, Belliart-Guérin G, Plaçais P-Y, Robie AA, Yamagata N, Schnaitmann C, Rowell WJ, Johnston RM, Ngo T-TB, Chen N, Korff W, Nitabach MN, Heberlein U, Preat T, Branson KM, Tanimoto H, Rubin GM. 2014. Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. *eLife* **3**:e04580. doi:10.7554/eLife.04580

Bellen HJ, Tong C, Tsuda H. 2010. 100 years of Drosophila research and its impact on vertebrate neuroscience: a history lesson for the future. *Nat Rev Neurosci* **11**:514–522. doi:10.1038/nrn2839

Berger C, Renner S, Lüer K, Technau GM. 2007. The commonly used marker ELAV is transiently expressed in neuroblasts and glial cells in the Drosophila embryonic CNS. *Dev Dyn* **236**:3562–3568. doi:10.1002/dvdy.21372

Brand AH, Perrimon N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**:401–415.

Chiang A-S, Lin Chih-Yung, Chuang C-C, Chang H-M, Hsieh C-H, Yeh C-W, Shih C-T, Wu J-J, Wang G-T, Chen Y-C, Wu Cheng-Chi, Chen G-Y, Ching Y-T, Lee P-C, Lin Chih-Yang, Lin H-H, Wu Chia-Chou, Hsu H-W, Huang Y-A, Chen J-Y, Chiang H-J, Lu C-F, Ni R-F, Yeh C-Y, Hwang J-K. 2011. Three-Dimensional Reconstruction of Brain-wide Wiring Networks in Drosophila at Single-Cell Resolution. *Current Biology* **21**:1–11. doi:10.1016/j.cub.2010.11.056

Chow LWC, Leung Y-M. 2020. The versatile Kv channels in the nervous system: actions beyond action potentials. *Cell Mol Life Sci* **77**:2473–2482. doi:10.1007/s00018-019-03415-8

Covarrubias M, Wei A, Salkoff L. 1991. Shaker, Shal, Shab, and Shaw express independent K+ current systems. *Neuron* **7**:763–773. doi:10.1016/0896-6273(91)90279-9

Davies LR, Schou MF, Kristensen TN, Loeschcke V. 2018. Linking developmental diet to adult foraging choice in Drosophila melanogaster. *J Exp Biol* **221**. doi:10.1242/jeb.175554

Deitcher DL, Ueda A, Stewart BA, Burgess RW, Kidokoro Y, Schwarz TL. 1998. Distinct requirements for evoked and spontaneous release of neurotransmitter are revealed by mutations in the Drosophila gene neuronal-synaptobrevin. *J Neurosci* **18**:2028–2039.

del Valle Rodríguez A, Didiano D, Desplan C. 2011. Power tools for gene expression and clonal analysis in Drosophila. *Nat Methods* **9**:47–55. doi:10.1038/nmeth.1800

Deniz ÖG, Altun G, Kaplan AA, Yurt KK, von Bartheld CS, Kaplan S. 2018. A concise review of optical, physical and isotropic fractionator techniques in neuroscience studies, including recent developments. *J Neurosci Methods* **310**:45–53. doi:10.1016/j.jneumeth.2018.07.012

Devor A, Bandettini PA, Boas DA, Bower JM, Buxton RB, Cohen LB, Dale AM, Einevoll GT, Fox PT, Franceschini MA, Friston KJ, Fujimoto JG, Geyer MA, Greenberg JH, Halgren E, Hämäläinen MS, Helmchen F, Hyman BT, Jasanoff A, Jernigan TL, Judd LL, Kim S-G, Kleinfeld D, Kopell NJ, Kutas M, Kwong KK, Larkum ME, Lo EH, Magistretti PJ, Mandeville JB, Masliah E, Mitra PP, Mobley WC, Moskowitz MA, Nimmerjahn A, Reynolds JH, Rosen BR, Salzberg BM, Schaffer CB, Silva GA, So PTC, Spitzer NC, Tootell RB, Van Essen DC, Vanduffel W, Vinogradov SA, Wald LL, Wang LV, Weber B, Yodh AG. 2013. The Challenge of Connecting the Dots in the B.R.A.I.N. *Neuron* **80**:270–274. doi:10.1016/j.neuron.2013.09.008

Diao Fengqiu, Ironfield H, Luan H, Diao Feici, Shropshire WC, Ewer J, Marr E, Potter CJ, Landgraf M, White BH. 2015. Plug-and-Play Genetic Access to Drosophila Cell Types using Exchangeable Exon Cassettes. *Cell Reports* **10**:1410–1421. doi:10.1016/j.celrep.2015.01.059

Doe CQ. 2017. Temporal Patterning in the Drosophila CNS. *Annu Rev Cell Dev Biol* **33**:219–240. doi:10.1146/annurev-cellbio-111315-125210

Edelsbrunner H, Mücke EP. 1994. Three-dimensional alpha shapes. *ACM Trans Graph* **13**:43–72. doi:10.1145/174462.156635

Eichler K, Li F, Litwin-Kumar A, Park Y, Andrade I, Schneider-Mizell CM, Saumweber T, Huser A, Eschbach C, Gerber B, Fetter RD, Truman JW, Priebe CE, Abbott LF, Thum AS, Zlatic M, Cardona A. 2017. The complete connectome of a learning and memory centre in an insect brain. *Nature* **548**:175–182. doi:10.1038/nature23455

Eschbach C, Zlatic M. 2020. Useful road maps: studying Drosophila larva's central nervous system with the help of connectomics. *Current Opinion in Neurobiology*, Whole-brain interactions between neural circuits **65**:129–137. doi:10.1016/j.conb.2020.09.008

Expert P, Lord L-D, Kringelbach ML, Petri G. 2019. Editorial: Topological Neuroscience. *Network Neuroscience* **3**:653–655. doi:10.1162/netn_e_00096

Feany MB, Bender WW. 2000. A Drosophila model of Parkinson's disease. *Nature* **404**:394–398. doi:10.1038/35006074

Friggi-Grelin F, Coulom H, Meller M, Gomez D, Hirsh J, Birman S. 2003a. Targeted gene expression in Drosophila dopaminergic cells using regulatory sequences from tyrosine hydroxylase. *J Neurobiol* **54**:618–627. doi:10.1002/neu.10185

Friggi-Grelin F, Iché M, Birman S. 2003b. Tissue-specific developmental requirements of Drosophila tyrosine hydroxylase isoforms. *genesis* **35**:175–184. doi:10.1002/gene.10178
Gerber B, Stocker RF, Tanimura T, Thum AS. 2009. Smelling, Tasting, Learning: Drosophila as a Study Case In: Korsching S, Meyerhof W, editors. Chemosensory Systems in Mammals, Fishes, and Insects, Results and Problems in Cell Differentiation. Berlin, Heidelberg: Springer. pp. 187–202. doi:10.1007/400_2008_9

Gowda SBM, Salim S, Mohammad F. 2021. Anatomy and Neural Pathways Modulating Distinct Locomotor Behaviors in Drosophila Larva. *Biology (Basel)* **10**. doi:10.3390/biology10020090

Hafer N, Schedl P. 2006. Dissection of Larval CNS in Drosophila Melanogaster. *J Vis Exp* 85. doi:10.3791/85

Jazin E, Cahill L. 2010. Sex differences in molecular neuroscience: from fruit flies to humans. *Nat Rev Neurosci* **11**:9–17. doi:10.1038/nrn2754

Jenett A, Rubin GM, Ngo T-TB, Shepherd D, Murphy C, Dionne H, Pfeiffer BD, Cavallaro A, Hall D, Jeter J, Iyer N, Fetter D, Hausenfluck JH, Peng H, Trautman ET, Svirskas RR, Myers EW, Iwinski ZR, Aso Y, DePasquale GM, Enos A, Hulamm P, Lam SCB, Li H-H, Laverty TR, Long F, Qu L, Murphy SD, Rokicki K, Safford T, Shaw K, Simpson JH, Sowell A, Tae S, Yu Y, Zugates CT. 2012. A GAL4-driver line resource for Drosophila neurobiology. *Cell Rep* **2**:991–1001. doi:10.1016/j.celrep.2012.09.011

Jiao W, McCabe BD. 2021a. Whole Brain Drosophila Larval Neurons.doi:10.5281/zenodo.5585334

Jiao W, McCabe BD. 2021b. Whole Brain Drosophila Larval Glia. doi:10.5281/zenodo.5585358

Kaiser M. 2015. Neuroanatomy: Connectome Connects Fly and Mammalian Brain Networks. *Current Biology* **25**:R416–R418. doi:10.1016/j.cub.2015.03.039

Keller D, Erö C, Markram H. 2018. Cell Densities in the Mouse Brain: A Systematic Review. *Front Neuroanat* **12**. doi:10.3389/fnana.2018.00083

Kimura K-I, Ote M, Tazawa T, Yamamoto D. 2005. Fruitless specifies sexually dimorphic neural circuitry in the Drosophila brain. *Nature* **438**:229–233. doi:10.1038/nature04229

Kremer MC, Jung C, Batelli S, Rubin GM, Gaul U. 2017. The glia of the adult Drosophila nervous system. *Glia* **65**:606–638. doi:10.1002/glia.23115

Lent R, Azevedo FAC, Andrade-Moraes CH, Pinto AVO. 2012. How many neurons do you have? Some dogmas of quantitative neuroscience under revision. *European Journal of Neuroscience* **35**:1–9. doi:https://doi.org/10.1111/j.1460-9568.2011.07923.x

Li H, Janssens J, Waegeneer MD, Kolluru SS, Davie K, Gardeux V, Saelens W, David F, Brbić M, Leskovec J, McLaughlin CN, Xie Q, Jones RC, Brueckner K, Shim J, Tattikota SG, Schnorrer F, Rust K, Nystul TG, Carvalho-Santos Z, Ribeiro C, Pal S, Przytycka TM, Allen AM, Goodwin SF, Berry CW, Fuller MT, White-Cooper H, Matunis EL, DiNardo S, Galenza A, O'Brien LE, Dow JAT, Consortium FCA, Jasper H, Oliver B, Perrimon N, Deplancke B, Quake SR, Luo L, Aerts S. 2021. Fly Cell Atlas: a single-cell transcriptomic atlas of the adult fruit fly. *bioRxiv* 2021.07.04.451050. doi:10.1101/2021.07.04.451050

Li-Kroeger D, Kanca O, Lee P-T, Cowan S, Lee MT, Jaiswal M, Salazar JL, He Y, Zuo Z, Bellen HJ. 2018. An expanded toolkit for gene tagging based on MiMIC and scarless CRISPR tagging in Drosophila. *Elife* **7**:e38709. doi:10.7554/eLife.38709

Lin C-C, Potter CJ. 2016. Editing Transgenic DNA Components by Inducible Gene Replacement in Drosophila melanogaster. *Genetics* **203**:1613–1628. doi:10.1534/genetics.116.191783

Lin DM, Goodman CS. 1994. Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. *Neuron* **13**:507–523. doi:10.1016/0896-6273(94)90022-1

Littleton JT, Stern M, Perin M, Bellen HJ. 1994. Calcium dependence of neurotransmitter release and rate of spontaneous vesicle fusions are altered in Drosophila synaptotagmin mutants. *PNAS* **91**:10888–10892. doi:10.1073/pnas.91.23.10888

Louis M. 2020. Mini-brain computations converting dynamic olfactory inputs into orientation behavior. *Curr Opin Neurobiol* **64**:1–9. doi:10.1016/j.conb.2019.11.015

Lundell MJ, Hirsh J. 1994. Temporal and Spatial Development of Serotonin and Dopamine Neurons in the Drosophila CNS. *Developmental Biology* **165**:385–396. doi:10.1006/dbio.1994.1261

Mao Z, Davis RL. 2009. Eight different types of dopaminergic neurons innervate the Drosophila mushroom body neuropil: anatomical and physiological heterogeneity. *Front Neural Circuits* **3**. doi:10.3389/neuro.04.005.2009

McCormack TJ. 2003. Comparison of K+-channel genes within the genomes of Anopheles gambiae and Drosophila melanogaster. *Genome Biology* 13.

Meinertzhagen IA. 2018. Of what use is connectomics? A personal perspective on the Drosophila connectome. *Journal of Experimental Biology* **221**. doi:10.1242/jeb.164954

Najman L, Couprie M. 2003. Watershed Algorithms and Contrast Preservation In: Nyström I, Sanniti di Baja G, Svensson S, editors. Discrete Geometry for Computer Imagery. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 62–71. doi:10.1007/978-3-540-39966-7_5
Neckameyer WS, Bhatt P. 2016. Protocols to Study Behavior in Drosophila. *Methods Mol Biol* **1478**:303–320. doi:10.1007/978-1-4939-6371-3_19

Oliveira-Pinto AV, Santos RM, Coutinho RA, Oliveira LM, Santos GB, Alho ATL, Leite REP, Farfel JM, Suemoto CK, Grinberg LT, Pasqualucci CA, Jacob-Filho W, Lent R. 2014. Sexual dimorphism in the human olfactory bulb: females have more neurons and glial cells than males. *PLoS One* **9**:e111733. doi:10.1371/journal.pone.0111733

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D. n.d. Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON* 6.

Pfeiffer BD, Jenett A, Hammonds AS, Ngo T-TB, Misra S, Murphy C, Scully A, Carlson JW, Wan KH, Laverty TR, Mungall C, Svirskas R, Kadonaga JT, Doe CQ, Eisen MB, Celniker SE, Rubin GM. 2008. Tools for neuroanatomy and neurogenetics in Drosophila. *PNAS* **105**:9715–9720. doi:10.1073/pnas.0803697105

Preibisch S, Saalfeld S, Schindelin J, Tomancak P. 2010. Software for bead-based registration of selective plane illumination microscopy data. *Nat Methods* **7**:418–419. doi:10.1038/nmeth0610-418

Quiñones-Frías MC, Littleton JT. 2021. Function of Drosophila Synaptotagmins in membrane trafficking at synapses. *Cell Mol Life Sci* **78**:4335–4364. doi:10.1007/s00018-021-03788-9
Raji JI, Potter CJ. 2021. The number of neurons in Drosophila and mosquito brains. *PLoS One* **16**:e0250381. doi:10.1371/journal.pone.0250381

Ravenscroft TA, Janssens J, Lee P-T, Tepe B, Marcogliese PC, Makhzami S, Holmes TC, Aerts S, Bellen HJ. 2020. Drosophila Voltage-Gated Sodium Channels Are Only Expressed in Active Neurons and Are Localized to Distal Axonal Initial Segment-like Domains. *J Neurosci* **40**:7999–8024. doi:10.1523/JNEUROSCI.0142-20.2020

Reininghaus J, Huber S, Bauer U, Kwitt R. 2015. A stable multi-scale kernel for topological machine learning2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4741–4748. doi:10.1109/CVPR.2015.7299106

Ren X, Sun J, Housden BE, Hu Y, Roesel C, Lin S, Liu L-P, Yang Z, Mao D, Sun L, Wu Q, Ji J-Y, Xi J, Mohr SE, Xu J, Perrimon N, Ni J-Q. 2013. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *PNAS* **110**:19012–19017.

Rodrigues MA, Martins NE, Balancé LF, Broom LN, Dias AJS, Fernandes ASD, Rodrigues F, Sucena É, Mirth CK. 2015. Drosophila melanogaster larvae make nutritional choices that minimize developmental time. *J Insect Physiol* **81**:69–80. doi:10.1016/j.jinsphys.2015.07.002

Ryglewski S, Duch C. 2009. Shaker and Shal Mediate Transient Calcium-Independent Potassium Current in a Drosophila Flight Motoneuron. *J Neurophysiol* **102**:3673–3688. doi:10.1152/jn.00693.2009

Scheffer LK, Meinertzhagen IA. 2019. The Fly Brain Atlas. *Annu Rev Cell Dev Biol* **35**:637–653. doi:10.1146/annurev-cellbio-100818-125444

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682. doi:10.1038/nmeth.2019

Scott K, Brady R, Cravchik A, Morozov P, Rzhetsky A, Zuker C, Axel R. 2001. A chemosensory gene family encoding candidate gustatory and olfactory receptors in Drosophila. *Cell* **104**:661–673. doi:10.1016/s0092-8674(01)00263-x

Sepp KJ, Schulte J, Auld VJ. 2001. Peripheral Glia Direct Axon Guidance across the CNS/PNS Transition Zone. *Developmental Biology* **238**:47–63. doi:10.1006/dbio.2001.0411

Sherer LM, Garrett EC, Morgan HR, Brewer ED, Sirrs LA, Shearin HK, Williams JL, McCabe BD, Stowers RS, Certel SJ. 2020. Octopamine neuron dependent aggression requires dVGLUT from dual-transmitting neurons. *PLOS Genetics* **16**:e1008609. doi:10.1371/journal.pgen.1008609

Silbereis JC, Pochareddy S, Zhu Y, Li M, Sestan N. 2016. The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron* **89**:248–268. doi:10.1016/j.neuron.2015.12.008

Simpson JH. 2009. Chapter 3 Mapping and Manipulating Neural Circuits in the Fly BrainAdvances in Genetics, Genetic Dissection of Neural Circuits and Behavior. Academic Press. pp. 79–143. doi:10.1016/S0065-2660(09)65003-3

Spreemann G. n.d. RFPKOG. *Really Fast Persistence Kernels On GPUs*. https://nonempty.org/software/rfpkog/

Südhof TC, Rothman JE. 2009. Membrane fusion: grappling with SNARE and SM proteins. *Science* **323**:474–477. doi:10.1126/science.1161748

Sutcliffe B, Ng J, Auer TO, Pasche M, Benton R, Jefferis GSXE, Cachero S. 2017. Second-Generation Drosophila Chemical Tags: Sensitivity, Versatility, and Speed. *Genetics* **205**:1399–1408. doi:10.1534/genetics.116.199281

The GUDHI Editorial Board. n.d. GUDHI library. *GUDHI library*. https://gudhi.inria.fr/
Venken KJT, Schulze KL, Haelterman NA, Pan H, He Y, Evans-Holm M, Carlson JW, Levis RW, Spradling AC, Hoskins RA, Bellen HJ. 2011. MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. *Nat Methods* **8**:737–743.

Verkhratsky A, Butt A. 2018. The History of the Decline and Fall of the Glial Numbers Legend. *Neuroglia* **1**:188–192. doi:10.3390/neuroglia1010013

von Bartheld CS, Bahney J, Herculano-Houzel S. 2016. The Search for True Numbers of Neurons and Glial Cells in the Human Brain: A Review of 150 Years of Cell Counting. *J Comp Neurol* **524**:3865–3895. doi:10.1002/cne.24040

Wagh DA, Rasse TM, Asan E, Hofbauer A, Schwenkert I, Dürrbeck H, Buchner S, Dabauvalle M-C, Schmidt M, Qin G, Wichmann C, Kittel R, Sigrist SJ, Buchner E. 2006. Bruchpilot, a Protein with Homology to ELKS/CAST, Is Required for Structural Integrity and Function of Synaptic Active Zones in Drosophila. *Neuron* **49**:833–844. doi:10.1016/j.neuron.2006.02.008

Wang J-W, Beck ES, McCabe BD. 2012. A Modular Toolset for Recombination Transgenesis and Neurogenetic Analysis of Drosophila. *PLOS ONE* **7**:e42102. doi:10.1371/journal.pone.0042102

White JG, Southgate E, Thomson JN, Brenner S. 1986. The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci* **314**:1–340. doi:10.1098/rstb.1986.0056

Xiong WC, Okano H, Patel NH, Blendy JA, Montell C. 1994. repo encodes a glial-specific homeo domain protein required in the Drosophila nervous system. *Genes Dev* **8**:981–994. doi:10.1101/gad.8.8.981