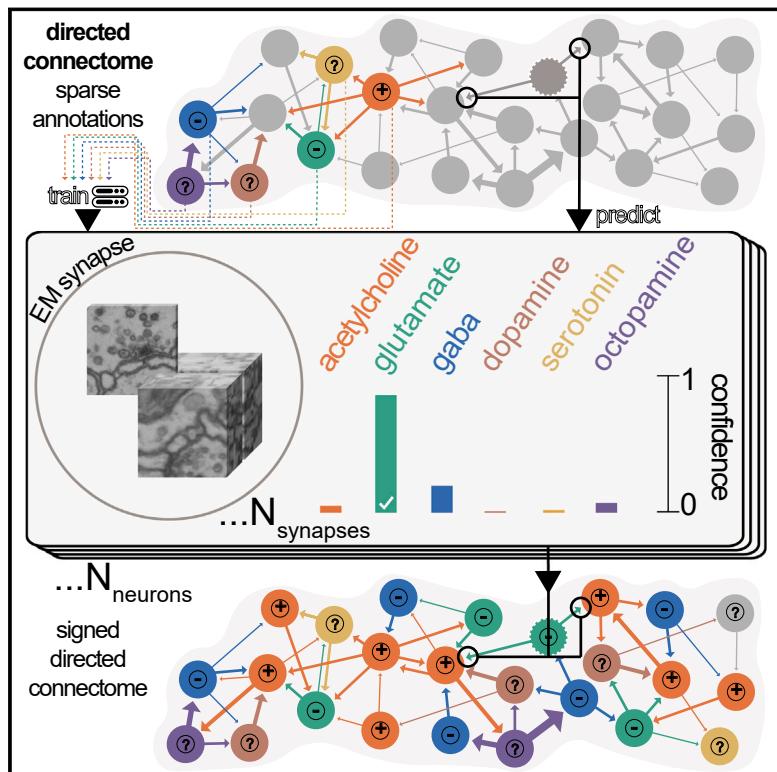


Neurotransmitter classification from electron microscopy images at synaptic sites in *Drosophila melanogaster*

Graphical abstract



Authors

Nils Eckstein, Alexander Shakeel Bates, Andrew Champion, ..., Volker Hartenstein, Gregory S.X.E. Jefferis, Jan Funke

Correspondence

funke@janelia.hhmi.org (J.F.), jefferis@mrc-lmb.cam.ac.uk (G.S.X.E.J.)

In brief

Neurotransmitter identities are predicted directly from electron micrographs with high fidelity across the whole fly brain for acetylcholine, glutamate, GABA, dopamine, serotonin, and octopamine.

Highlights

- Machine learning identifies synaptic transmitters from electron micrographs
- Six transmitters predicted across the whole fly brain connectome
- Explainable AI reveals ultrastructural differences between transmitter identities
- Fly brain hemilineages predominantly express one fast-acting transmitter



Resource

Neurotransmitter classification from electron microscopy images at synaptic sites in *Drosophila melanogaster*

Nils Eckstein,^{1,2,9} Alexander Shakeel Bates,^{3,4,5,9} Andrew Champion,⁶ Michelle Du,¹ Yijie Yin,⁶ Philipp Schlegel,^{3,6} Alicia Kun-Yang Lu,¹ Thomson Rymer,¹ Samantha Finley-May,¹ Tyler Paterson,¹ Ruchi Parekh,¹ Sven Dorkenwald,⁷ Arie Matsliah,⁷ Szi-Chieh Yu,⁷ Claire McKellar,⁷ Amy Sterling,⁷ Katharina Eichler,⁶ Marta Costa,⁶ Sebastian Seung,⁷ Mala Murthy,⁷ Volker Hartenstein,⁸ Gregory S.X.E. Jefferis,^{3,6,*} and Jan Funke^{1,10,*}

¹HHMI Janelia Research Campus, Ashburn, VA, USA

²Institute of Neuroinformatics UZH/ETHZ, Zurich, Switzerland

³Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge, UK

⁴Centre for Neural Circuits and Behaviour, The University of Oxford, Tinsley Building, Mansfield Road, Oxford OX1 3SR, UK

⁵Department of Neurobiology and Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA

⁶Drosophila Connectomics Group, Department of Zoology, University of Cambridge, Cambridge, UK

⁷Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

⁸Department of Molecular Cell and Developmental Biology, University of California, Los Angeles, Los Angeles, CA, USA

⁹These authors contributed equally

¹⁰Lead contact

*Correspondence: funkej@janelia.hhmi.org (J.F.), jefferis@mrc-lmb.cam.ac.uk (G.S.X.E.J.)

<https://doi.org/10.1016/j.cell.2024.03.016>

SUMMARY

High-resolution electron microscopy of nervous systems has enabled the reconstruction of synaptic connectomes. However, we do not know the synaptic sign for each connection (i.e., whether a connection is excitatory or inhibitory), which is implied by the released transmitter. We demonstrate that artificial neural networks can predict transmitter types for presynapses from electron micrographs: a network trained to predict six transmitters (acetylcholine, glutamate, GABA, serotonin, dopamine, octopamine) achieves an accuracy of 87% for individual synapses, 94% for neurons, and 91% for known cell types across a *D. melanogaster* whole brain. We visualize the ultrastructural features used for prediction, discovering subtle but significant differences between transmitter phenotypes. We also analyze transmitter distributions across the brain and find that neurons that develop together largely express only one fast-acting transmitter (acetylcholine, glutamate, or GABA). We hope that our publicly available predictions act as an accelerant for neuroscientific hypothesis generation for the fly.

INTRODUCTION

Generating a synaptic connectome entails identifying all neurons and synapses in a sample. Recent advances in volume electron microscopy (EM) have enabled connectome generation for entire nervous systems.^{1–6} Automated methods for segmenting neurons,^{7–11} detecting synapses,^{12–16} and proofreading¹⁷ have significantly reduced the human effort required. These methods have been applied to create connectomes for the *Drosophila melanogaster* brain^{6,18,19} and its ventral nerve cord.^{5,20} However, EM does not directly tell us about gene expression, most crucially the transmitter pathways active in each neuron. This gap hinders our understanding of key neurobiological processes relevant to circuit function.

The action that a neuron has on its downstream targets depends on the transmitter it releases and the postsynaptic recep-

tors that receive it. The so-called *classical* fast-acting transmitters (i.e., acetylcholine, glutamate, and GABA) are most common.^{21,22} They are contained in small, clear vesicles. Monoamines such as dopamine, serotonin, and octopamine are packaged into clear core or small dense core vesicles.²³ About 53 neuropeptides and peptide hormones^{24–26} are contained in larger dense core vesicles.^{27,28} While co-transmission of a small molecule transmitter and a neuropeptide is common,^{21,22,29} the usage of acetylcholine, glutamate, and GABA is largely mutually exclusive.^{30,31} Therefore, Dale's law, an expectation that each neuron expresses a single small molecule transmitter, often guides our hypotheses about neuronal function.^{32,33} We expect that this transmitter expression will be stereotyped for cell types across individuals, given their shared gene expression profiles.³⁴ We also expect transmitter expression to be organized at higher levels, for example by hemilineage. Hemilineages are basic



developmental units of the insect brain and typically consist of dozens of cell types. In the ventral nerve cord Lacin et al.³¹ comprehensively showed that only one of acetylcholine, glutamate, or GABA is expressed per hemilineage. By analogy with Dale's law, we call this observation Lacin's law. It was unclear whether Lacin's law also holds in the brain.

In larger organisms, experts can distinguish excitatory and inhibitory transmitters based on synaptic vesicle ellipticity^{35–37} or synapse symmetry.³⁸ However, in invertebrates transmitter identity cannot be consistently identified by human annotators from electron micrographs. Instead, investigators use molecular biology and light microscopy pipelines to link RNA expression data or immunoreactivity to proteins involved in transmitter biosynthesis with specific neuronal morphologies.^{22,39–44} Standard high-throughput methods such as single-cell RNA-seq cannot be used since they do not retain morphology information. Therefore, researchers profile sparse genetic driver lines,^{45,46} which target only a few neurons, followed by accurate morphological matching to EM reconstructions.^{47,48} Since these experimental pipelines can take weeks per cell type and are limited by the availability of sparse genetic driver lines, they do not scale to whole nervous system discovery. Consequently, the transmitter identity is known for only ~700 of the ~7,000 cell types of the adult *D. melanogaster* central brain.^{49,50} However, this subset provided ground truth that we can use to learn the features of specific transmission types, by mapping known transmitter identities to previously localized presynaptic sites in two brain datasets (FAFB [the Full Adult Fly Brain]⁶ and the HemiBrain⁵¹).

RESULTS

Assembling ground-truth neurotransmission data

We compiled a list of 356 *D. melanogaster* neuronal cell types from 21 studies (Data S1), selecting those with robust transmitter data. We identified these same cell types within the FAFB-Catmaid and HemiBrain datasets (Figure 1A). We selected cell types with clear transmitter data from RNA expression or immunohistochemistry, complemented by specific cell type targeting through the GAL4/split-GAL4 system.⁴⁶ We did not generally pursue types reported to exhibit co-transmission (see STAR Methods). Given that *D. melanogaster* neurons are highly stereotyped,⁵⁰ these cell types were identifiable in our EM datasets. We chose to proceed with transmitters supported by at least 10 EM reconstructions per dataset, i.e., acetylcholine, glutamate, dopamine, serotonin, and octopamine (Data S1, see STAR Methods). All presynapses for each neuronal reconstruction were assumed to use the single small-molecule transmitter reported in the literature (Figure 1B). In total, we matched 3,025 FAFB-Catmaid neuronal reconstructions (211,564 synapses) and 5,902 HemiBrain reconstructions (840,535) to cell types with a known transmitter (Data S2).

We took a slightly different approach in selecting presynapses from our two datasets. For HemiBrain, we used automatically predicted presynaptic sites across many brain cell types, of which a proportion are false detections.¹³ In FAFB, however, we used manually placed synaptic markers laid by human researchers from a smaller pool of manually reconstructed neurons (see STAR Methods). Therefore, compared with

HemiBrain, our FAFB-Catmaid dataset contained higher-quality synapses from a smaller number of neurons. Note that in later results, to work with the full connectome, we explored transmitter predictions across automatically detected presynapses¹² in the newer FAFB-FlyWire reconstruction project.^{17,18,50} After building our ground truth, we expanded our literature review to find 268 more cell types with the reported transmission of one of our six transmitters (Data S7), which we could use to further validate our results.

Network architecture, training, testing, and validation datasets

For each transmitter $y \in \{\text{GABA, acetylcholine, glutamate, serotonin, octopamine, dopamine}\}$, we partitioned the data into training, testing, and validation sets by randomly assigning entire neurons (neuron split), such that approximately 70% of presynapses were used for training, 10% for validation, and the remaining 20% for testing. This approach mirrors real-world scenarios where we typically know the transmitter of an entire neuron and are interested in predicting the transmitter of a different neuron. We employed a 3D deep convolutional network based on the Visual Geometry Group (VGG) architecture⁵² to predict transmitter identity from cubes of EM image data (edge length 640 nm), each centered on a presynaptic site (Figure 1C). The cube size was chosen to be large enough to provide surrounding context, including synaptic features like vesicles, T-bars, clefts, and postsynaptic densities, while at the same time being small enough to fit into the limited memory of a GPU. The network consisted of four functional blocks, each with two 3D convolution operations, batch normalization, ReLU non-linearities, and subsequent max pooling with a downsampling factor of 2, except for FAFB where we limited downsampling to the x and y dimensions for the first three blocks to account for image voxel anisotropy. The last block was followed by three fully connected layers with dropout ($p = 0.5$) applied to the last one. We trained the network to minimize cross-entropy loss over the six classes (GABA, acetylcholine, glutamate, serotonin, octopamine, and dopamine) using the Adam optimizer.⁵³ We trained for a total of 500,000 iterations in batches containing eight samples and selected the iteration with the highest validation accuracy for testing.

For the FAFB neuron split, the testing set consisted of a total of 40,104 presynapses from 185 neurons that the network was not trained on. The network achieved an average per-transmitter accuracy of 87% on FAFB and 78% on HemiBrain. We assigned each neuron with over 30 presynapses in the testing set a transmitter through a majority vote of its presynapses, yielding an average accuracy of 94% for transmitter prediction per neuron on FAFB-Catmaid and 91% on HemiBrain (Figure 2A). Our goal was to train a network with high prediction accuracy based on ultrastructural features of synapses. Cytological correlates of synapse location or neuron development might have driven transmitter identification by our network. We, therefore, split the data by the neuropil (i.e., brain region) location of synapses⁵⁴ (Figure 2B) and their developmental origin (hemilineage) (Figure 2C). Accuracy in these split datasets remained high (Figure 1D), indicating that the network is unlikely to use related features.

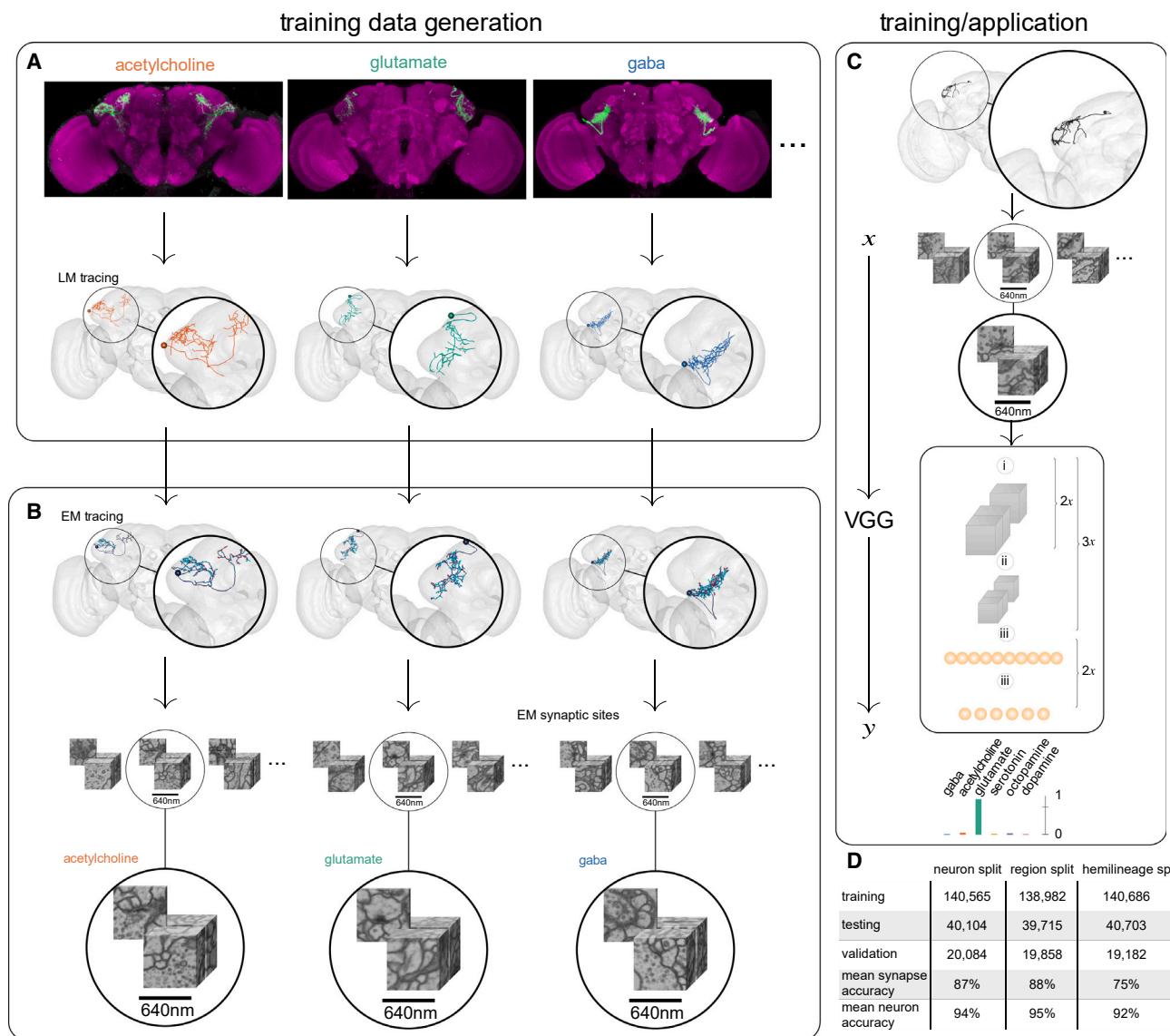


Figure 1. Method overview

We assembled a dataset of neurons with known transmitter expression (see [STAR Methods](#)) in two *D. melanogaster* brain EM datasets (FAFB and HemiBrain) and retrieved corresponding synaptic locations.

(A) Typically, neurons had been genetically tagged to identify their transmitter identity and reconstruct their coarse morphology using light microscopy ([Data S1](#)).

(B) Light microscopy tracings of neurons are then matched to corresponding EM reconstructions with annotated synaptic locations, yielding a dataset of EM volumes of synaptic sites with known transmitter identity.

(C) We used the resulting pair (x, y) where x is a 3D EM volume of a synaptic site and y is the transmitter of that synaptic site (one of GABA, acetylcholine, glutamate, serotonin, octopamine, or dopamine) to train a 3D VGG-style deep neural network to assign a given synaptic site x to one of the six considered transmitters. We used the trained network to predict the transmitter identity of synapses from neurons with so far unknown transmitter identity. Panels i, ii, and iii denote convolution, down-sampling, and fully connected layers, respectively.

(D) Overview of our results on the FAFB dataset. Shown are the number of presynapses for training, testing, and validation as well as average synapse and neuron classification accuracy on the testing set for each data split.

See also [Data S2](#).

Applying Dale's law to the entirety of both datasets, we adopted the most frequent synapse-level transmitter prediction as the neuron-level transmitter prediction and developed a confidence score based on the proportion of presynapses that "voted" for the neuron-level transmitter prediction and our confusion

matrices (see [STAR Methods](#)). The distribution of confidence scores across the two datasets suggested the network was most confident in acetylcholine predictions ([Figure 2D](#)). In general, we found excellent agreement with the literature ([Data S7](#)). We predicted most known cholinergic cell types correctly

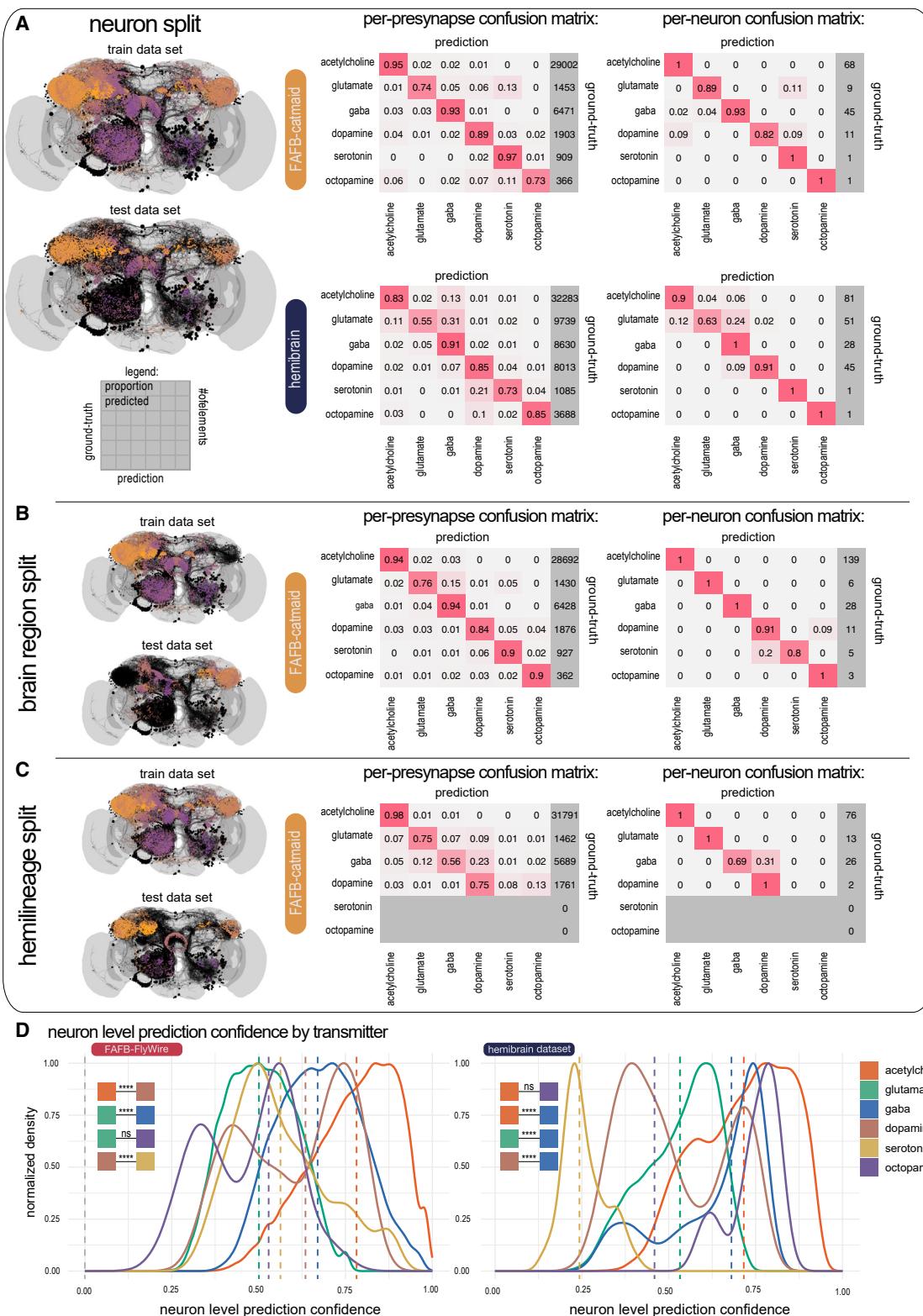


Figure 2. The accuracy of the trained classifier on a per-presynapse and per-neuron basis

(A) Left: visualization of the training (upper) and testing (lower) data (split by entire neurons) that we used for the results in this manuscript. Presynapse locations are color coded according to their z-depth; anterior-posterior shown as purple-orange, neuron skeletons in black. Right: confusion matrices for the trained

(legend continued on next page)

(FAFB-FlyWire, 91%; HemiBrain, 91%), as well as most glutamatergic (FAFB-FlyWire, 91%; HemiBrain, 95%), GABAergic (FAFB-FlyWire, 96%; HemiBrain, 97%), dopaminergic (FAFB-FlyWire, 90%; HemiBrain, 85%), and octopaminergic (FAFB-FlyWire, 85%; HemiBrain, 100%) cell types (Figures S2D and S2F). Notably, the optic lobe, which has the most known transmitter assignment,³⁹ was largely not used in our ground truth due to a relative paucity of FAFB-Catmaid reconstructions and its absence from the HemiBrain dataset. Nevertheless, 96% of ~29,000 cholinergic optic lobe neurons were predicted correctly, as well as 87% of ~3,600 GABAergic neurons and 91% of ~1,600 glutamatergic neurons.

However, we noticed a few clear mispredictions and discrepancies, most commonly in cases of suspected co-transmission (see STAR Methods). For example, Kenyon cells had been mispredicted for dopamine in both HemiBrain and FAFB-FlyWire rather than acetylcholine,⁵⁵ some known serotonergic neurons (Figure S2E) were not predicted for serotonin in either dataset,^{56–60} many first-order sensory neurons and antennal lobe local neurons were mispredicted for serotonin rather than acetylcholine in both datasets^{61–64} and some intrinsic neurons of the fan-shaped body were mispredicted in HemiBrain but not in FAFB-FlyWire and vice versa. Overall, serotonin was our least reliable prediction (FAFB-FlyWire, 33%; HemiBrain, 38%) likely due to its relative paucity in our ground-truth data. Our results for a full ventral nerve cord (MaleVNC, 82% accuracy, limited to acetylcholine, glutamate, and GABA) are reported elsewhere.^{5,65}

Classifier synapse feature analysis

The classifier's high accuracy on test datasets raised questions about how transmitter classes are discriminated, as human annotators cannot reliably determine transmitter identity from EM alone. We reasoned that identifying the visual features that our network used to label transmitters could verify that decisions were not based on class-irrelevant confounders and discover unknown ultrastructural differences between synapse classes. To visualize these features, we explored *post hoc* single-input attribution methods that derive an attribution map for a single input image, highlighting image areas crucial for classification. Existing methods^{66–69} did not provide images we could interpret since the highlighted areas were too large and variable. This may be because single-input attribution methods are subject to highlighting non-class-specific "distractors,"⁷⁰ e.g., features stemming from different orientations of the synapse, section thickness, or intensity variations. We introduced an attribution method to focus on class-relevant features between pairs of classes while disregarding distractors⁷¹ and used the FAFB da-

taset for its superior x,y resolution. Given a real image x_R of a class y_R , we first created a counterfactual image x_C by translating x_R into an image of another class y_C using a CycleGAN,⁷² resulting in paired images of different transmitter types. Crucially, this domain translation of an image from class y_R to class y_C keeps class-irrelevant distractors (e.g., the orientation of the synapse) intact. Class-relevant features, however, are changed due to the adversarially trained discriminator of the CycleGAN. We used our previously trained classifier to confirm that the domain translation was successful: we filtered all images such that x_R was classified as y_R and the counterfactual x_C as y_C . We then identified a small region in x_C , that when swapped with x_R changed the prediction of the trained classifier (Figures 3A and 3B). Specifically, we identified a minimal binary mask m , such that the hybrid image $x_H = m \cdot x_R + (1 - m) \cdot x_C$ was classified as y_R . To find m , we used a modified version of the DeepLift method, where we used the counterfactual x_C as the neutral ref.⁷¹

This method allowed us to manually identify at least one distinguishing feature (Figure 3C) between each pair of transmitters (see Data S8). We only included features consistently observed in both directions, e.g., translating a real GABA image to acetylcholine results in a brighter cleft and translating a real acetylcholine image to GABA produces a darker cleft. Comparing transmitter identities in paired images allowed us to observe features as subtle as sub-pixel changes in vesicle diameters (e.g., between GABA and glutamate), which would be near impossible to pick up in unpaired images. We confirmed the identified features between the classical transmitters GABA, acetylcholine and glutamate on the original synapse images by manually segmenting the synaptic cleft, vesicles, and T-bars of 222 synapse images (75 GABA, 85 acetylcholine, 62 glutamate; annotators were blind to the predicted transmitter class). We found strong support for each of the identified features between those transmitters (Figures 3D–3F): acetylcholine has a brighter cleft than GABA and glutamate ($p \leq 0.0001$), glutamate has larger vesicles than GABA ($p \leq 0.001$) and glutamate has a darker T-bar than acetylcholine ($p \leq 0.001$).

Comparing neurotransmitter predictions for neuron homologs across datasets

The insect brain consists of thousands of isomorphic cell types. Every cell type has a copy on each hemisphere. These cell types contain only a few neurons (median, 2, IQR, 2, neurons per cell type per hemisphere for the HemiBrain dataset, excluding the largest outlier classes of Kenyon cells and sensory receptor neurons), often only a single neuron per hemisphere (41% of cell types in HemiBrain). We refer to these neurons as singletons.

classifier on the testing data, shown per presynapse and as a majority vote per neuron, on datasets FAFB and HemiBrain. We considered only those neurons with more than 30 presynapses.

(B) Classification results on alternative training and testing data (split by brain regions) from FAFB.

(C) Same as (B) but split by hemilineage. It was not possible to generate a fully balanced split and as a result there are no serotonin and octopamine neurons in the testing set, as indicated by the grayed-out rows.

(D) The distribution of neuron-level confidence scores by transmitter, across our pool of central brain neurons in the FlyWire and HemiBrain datasets (FAFB-FlyWire, 136,927; HemiBrain, 24,666). Vertical dashed line, median value. Colored boxes with stars indicate statistical comparisons, Wilcoxon two-sample tests (n.s., not significant; * $p \leq 0.05$; *** $p \leq 0.00001$).

See also Figure S1 and Data S3 and S4.

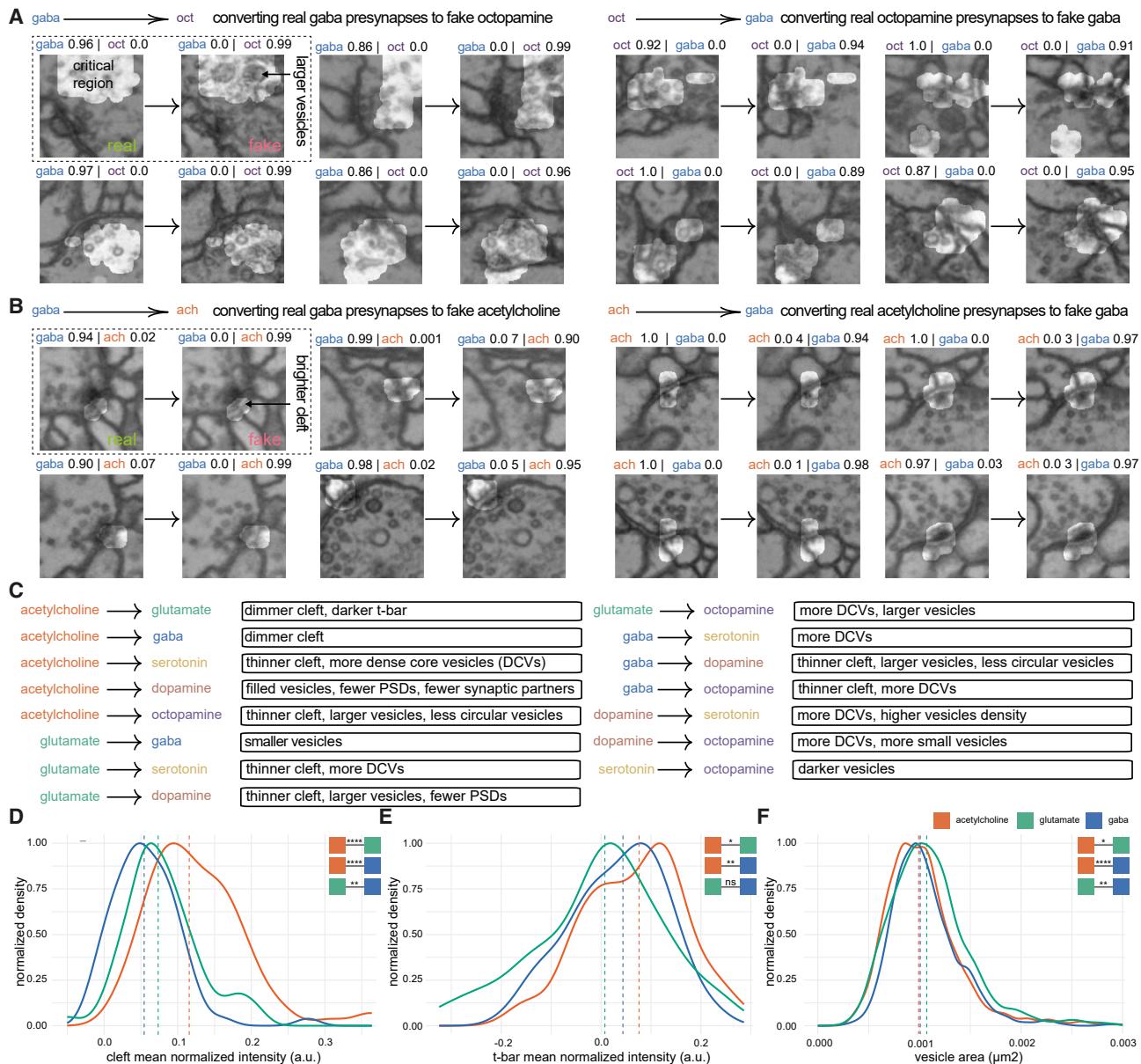


Figure 3. Classifier feature analysis using a discriminative attribution method

- (A) Example translations of real synapse images to fake counterfactual images. Highlights show attribution masks indicating the most important changes between the two classes. Classifier scores are shown above each image. Left, two columns show the translation of real GABA synapses into counterfactual octopamine synapses. Right, same as left but for octopamine to GABA.
- (B) Same as (A) but for GABA-acetylcholine.
- (C) Pairwise differences between transmitters, found through manual inspection of real and counterfactual images. Dense core vesicles, DCVs; postsynaptic densities, PSDs.
- (D) Normalized density plot showing the distribution of cleft intensity among original synapse images. Number of annotated synapses: acetylcholine 84, glutamate 61, GABA 74.
- (E) Same as (C) for T-bar intensities. Number of annotated synapses: acetylcholine 85, glutamate 62, GABA 75.
- (F) Same as (C) for vesicle sizes. Number of annotated vesicles: acetylcholine 1,729, glutamate 1,153, GABA 1,382. Vertical dashed line, median value. Colored boxes with stars indicate statistical comparisons, Wilcoxon tests. Note that the vesicle size comparison assumes that vesicle sizes from the same synapses are conditionally independent given the transmitter (n.s., not significant; $*p \leq 0.05$; $**p \leq 0.001$; $***p \leq 0.00001$).

See also Figure S1 and Data S8.

They provide a natural mechanism by which to test whether our transmitter predictions are consistent as, being unique on each hemisphere, they can be unambiguous matching to their homologs between hemispheres and brains (Figure 4A). We matched 1,586 right hemisphere singletons to their left hemisphere homologs in FAFB-FlyWire and 1,318 to their HemiBrain homologs.⁵⁰

We found good agreement between left-right matched FAFB-FlyWire singletons (Figure 4B) and FAFB-HemiBrain matched singletons (Figure 4C) for cholinergic, glutamatergic, and GABAergic pairs. Inconsistent results between matched neurons were more common with singletons predicted to express dopamine, serotonin, or octopamine. The average confidence score for FAFB-FlyWire right (mean, 0.79, SD, 0.16), FAFB-FlyWire left (mean, 0.79, SD, 0.17), and HemiBrain (mean, 0.67, SD, 0.14) neurons were significantly higher when there was no mismatch between our paired neurons than when there was a conflict (FAFB-FlyWire right: mean, 0.62, SD, 0.22; FAFB-FlyWire left: mean, 0.47, SD, 0.14; hemibrain: mean, 0.43, SD, 0.13) (Figure 4D).

We found that matched singleton pairs and cell types correlated in their neuron-level transmitter prediction scores (Figures 4D–4F and S2A). When the network was less confident it was consistently less confident for both members of a pair, both across hemispheres and datasets, although within-dataset comparisons were more similar (Figure 4E). In cases where changes in the confidence score were correlated, the reason is more likely to be biological than dataset specific. For example, low, correlated scores could indicate that the network has encountered a biological situation that is rare in, or outside of, our training data.

We also examined isomorphic cell types in the HemiBrain dataset for which there was more than one member of the cell type per hemisphere (Figures 4F and S2B). Only 14% of cell types with more than one member per hemisphere did not have the same neuron-level transmitter prediction for each member. Moreover, we matched 2,626 neuronal cell types between FAFB-FlyWire and the HemiBrain datasets and found that 95% agree in their neuron-level transmitter prediction between the two datasets (Figures S2B and S2C). Again, among these conflicted types, the mean neuron-level transmitter prediction score was significantly lower (Figure S2C). Together, this suggests that there may be a biological factor, e.g., the expression of transmitters not in our training data or co-transmission, that leads to lower confidence scores and incorrect or inconsistent predictions with certain cell types, rather than a confound related to the EM data quality.

An overview of neurotransmitter usage in the nervous system

We next wanted to get an overview of transmitter usage in the *D. melanogaster* nervous system (Figures 5A and 5B), including a breakdown by axon and dendrite connections across the brain (Figures 5C, 5D, and 6B). We also explored potential correlations between neuron-level morphological features and transmitter predictions (Figures 5E, 5F, and S3), alongside variations in transmitter usage across sensory systems within the fly brain (Figure 6A).

We calculated neuron-level transmitter predictions for all 24,666 well-reconstructed neurons in the HemiBrain dataset (Data S3), all 49,985 central brain and 86,942 optic lobe FAFB-FlyWire neurons (Data S4), and 23,503 ventral nerve cord neurons from the MaleVNC dataset. In the central brain, the largest fraction of neurons were predicted to be cholinergic with smaller fractions for glutamatergic and GABAergic neurons (Figure 6A). Single-cell RNA sequencing has suggested a breakdown of 44%–45% cholinergic, 14%–15% glutamatergic, and 10%–15% GABAergic neurons^{21,22} in the central brain. Dopaminergic neurons had distinctive features, such as higher mitochondrial density (Figure S3C), a diverse set of input neurons but fewer downstream targets (Figure S3I), primarily in deeper sensory layers (Figure 6C). They are encountered after Kenyon cells in the olfactory system (often thought of as a conditioned stimulus) but before the mushroom body in the gustatory system (typically an unconditioned stimulus). These observations suggest that dopaminergic neurons are highly active and sample widely from more superficial brain layers to provide teaching signals to select target neurons of a deeper sensory layer.

One interesting new insight from large-scale neuron-level transmitter predictions concerns the fan-shaped body. This central brain structure computes navigational variables^{75,76} and is built as a matrix with ~9 rows and ~10 columns.⁷⁷ Our predictions show that row-wise tangential input is overwhelmingly glutamatergic (~84%), with no inhibitory local neurons in the structure. There are some dopaminergic⁷⁸ but very few GABAergic or cholinergic row-wise inputs. In contrast, acetylcholine was predicted for almost all column-wise input types (~87%), intrinsic neurons (~88%), and output types (~96%); the small number of non-cholinergic neurons are likely mispredictions (see STAR Methods). These findings can be compared to the elegant layout of transmitter expression in the mushroom body,⁷⁹ a discovery that has accelerated research in this associative memory structure in recent years.

The identity of the presynaptic and postsynaptic neuronal compartments is another factor, in addition to transmitter usage, that determines the effect of a synaptic connection.⁸⁰ Insect neurons frequently possess arbor with a mix of input and output synapses, but morphological features can resolve axons and dendrites in most cases^{19,73,74,81–83} (Figure 5C). Axons and dendrites can synaptically connect with either being the source or the target (Figures 5C and 5B). We “split”⁷³ thousands of neurons into separate axonal and dendritic compartments (see STAR Methods). Although most presynapses are on the axon (FAFB-FlyWire: median 76%, SD, 21%; hemibrain: median 70%, SD, 21%), neurons had a large proportion of their output sites on their dendrites (FAFB-FlyWire: median 23%, SD, 21%; hemibrain: median 30%, SD, 21%). We found that while the majority of the synaptic budget is spent on axo-dendritic connections (FAFB-FlyWire: 55%; hemibrain: 48%), a large fraction is spent on axo-axonic connections (FAFB-FlyWire: 22%; Hemi-Brain: 20%) and a similar amount on dendro-dendritic connections (FAFB-FlyWire: 19%; hemibrain: 21%). These figures are comparable to those recently reported for the *D. melanogaster* larva¹⁹ (axo-dendritic, 54%; axo-axonic, 36%; dendro-dendritic, 8%; dendro-axonic, 3%). Neurons tended to receive strong input on their axons from just one of acetylcholine, glutamate,

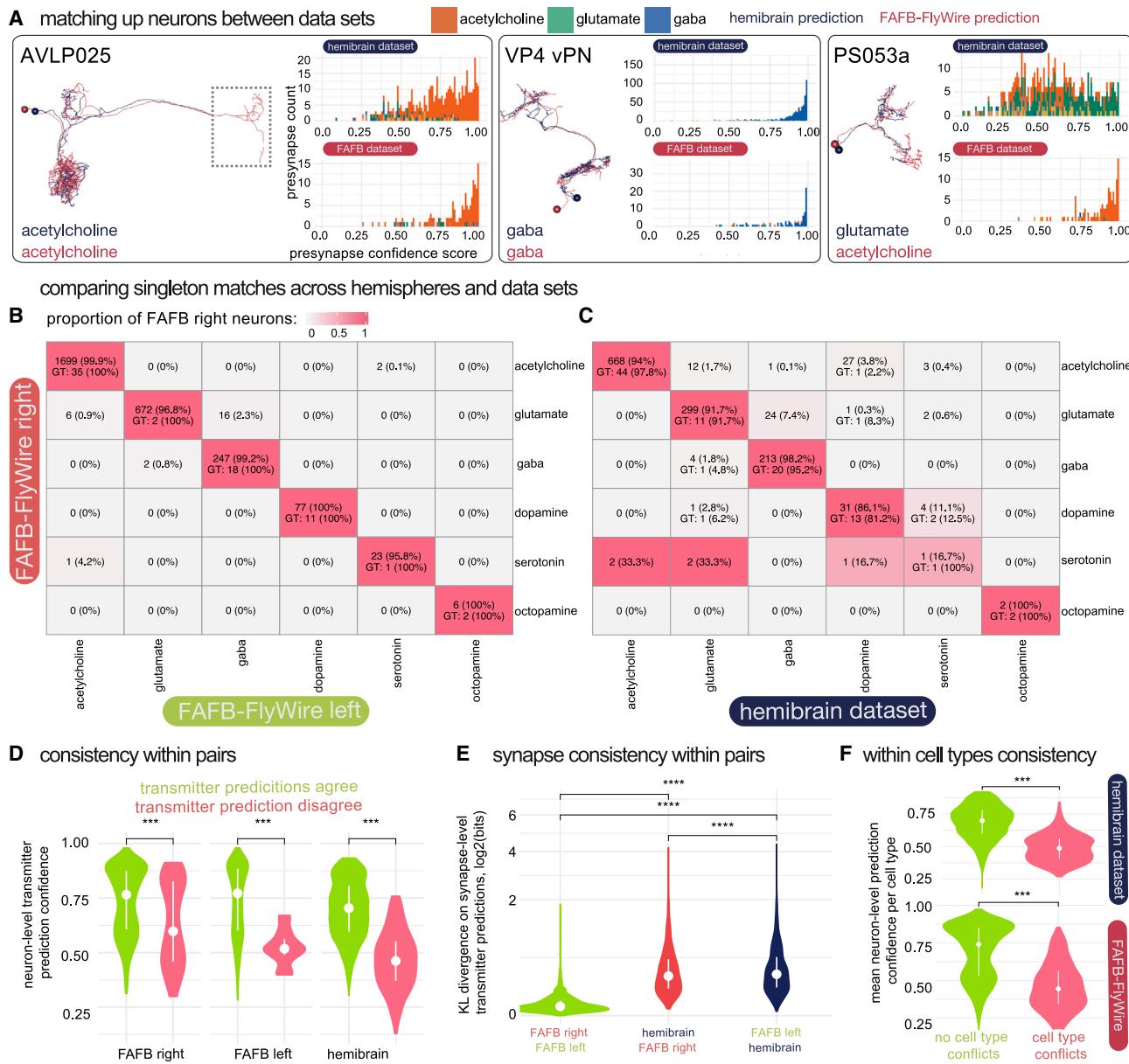


Figure 4. Comparing neuron-level transmitter predictions between connectome datasets from separate animals and between hemispheres

(A) Images of co-registered, matched neurons between the HemiBrain (navy) and the FAFB-FlyWire (red) datasets. Histograms show synapse-level transmitter prediction scores for exemplar pairs. Neurons can be matched despite missing data (left, grey dashed box). PS053a has conflicting neuron-level transmitter predictions.

(B) Confusion matrix compares matched singleton FAFB-FlyWire-right and FAFB-FlyWire-left pairs' neuron-level transmitter predictions (1,586 pairs).

(C) Confusion matrix comparing matched FAFB-FlyWire-right and HemiBrain-right neuron-level transmitter predictions (1,318 pairs). Cells colored by the proportion of FAFB-FlyWire right neurons of each transmitter type (row normalized) that are matched to its homolog-columns give homolog prediction.

(D) Neuron-level transmitter prediction scores between matched singletons that have (red, right) or do not have (green, left) a conflict between their neuron-level transmitter predictions, across all three hemispheres. Matches:mismatches across all comparisons for FAFB-FlyWire right: 2,650:170 FAFB-FlyWire left, 1,562:40, and HemiBrain neurons, 1,088:130.

(E) Comparison of similarity scores for matches (Kullback-Leibler divergence on synapse-level transmitter prediction scores).

(F) The neuron-level transmitter prediction consistency among cell types that have multiple repeats, i.e., not singletons. Green, the mean neuron-level transmitter prediction confidence for cell types where all members of the type are predicted to use the same transmitter. Red, the mean neuron-level transmitter prediction confidence for cell types where not all members of the type are predicted to use the same transmitter. Violin plots show the median value (dot) and the interquartile range (line, 25th to 75th percentiles). Data were compared using Wilcoxon two-sample tests (n.s., not significant; ***p ≤ 0.0001; ****p ≤ 0.00001).

See also Figure S2 and Data S7.

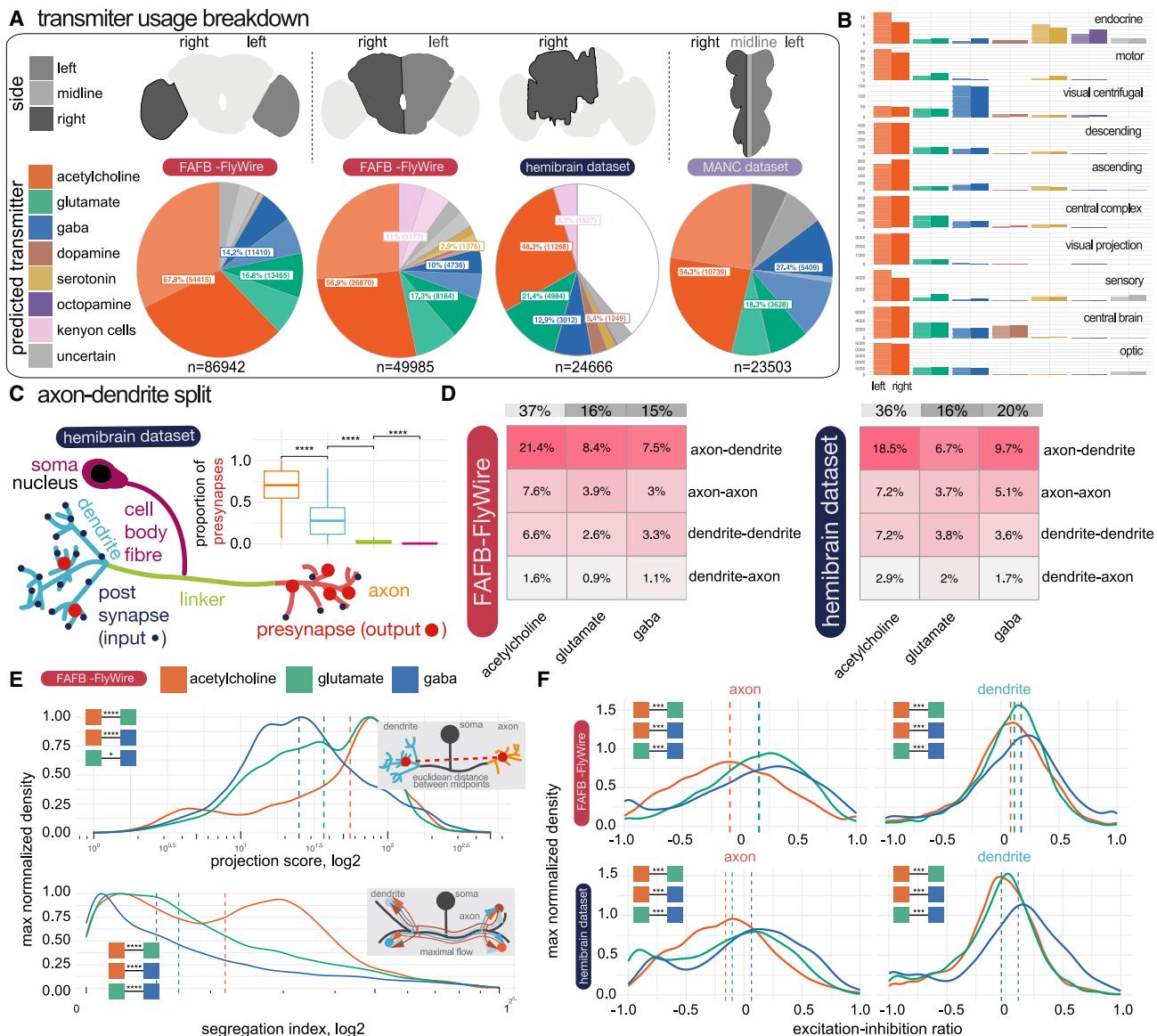


Figure 5. Breakdown of transmitter use across the *D. melanogaster* nervous system

(A) Our neuron-level transmitter predictions across the female optic lobes and central brain and a male ventral nerve cord (see STAR Methods).

(B) Bar plots for the numbers of neurons predicted for different transmitter usages in each super class in the FAFB-FlyWire dataset.⁵⁰

(C) Schematic of a neuron broken into its neuronal compartments. Inset, the proportion of presynapses in each of the four compartment types.

(D) Synaptic budget across different connection types in FAFB-FlyWire (left) and HemiBrain (right). Heatmaps show the proportion of synaptic contacts from neurons of different predicted transmitter types (columns) used in different inter-compartmental connection types (rows). FAFB-FlyWire, 9,123; hemibrain, 10,122 neurons.

(E) Scaled density plots showing neuronal polarity by neuron-level transmitter prediction. Upper, distribution of projection scores, which is the distance in Euclidean space between the dendritic and axonic midpoint. Lower, segregation index: the higher the score, the more polarized the neuron.⁷³

(F) Scaled density plots showing the distribution of excitation-inhibition balance (proportion of excitatory, acetylcholine, input minus the proportion of inhibitory input; GABA, glutamate) across neuron-level transmitter predictions and compartments. Vertical dashed line, median value. Colored boxes with stars indicate statistical comparisons, Wilcoxon two-sample tests (n.s., not significant; * $p \leq 0.05$; ** $p \leq 0.0001$; *** $p \leq 0.00001$).

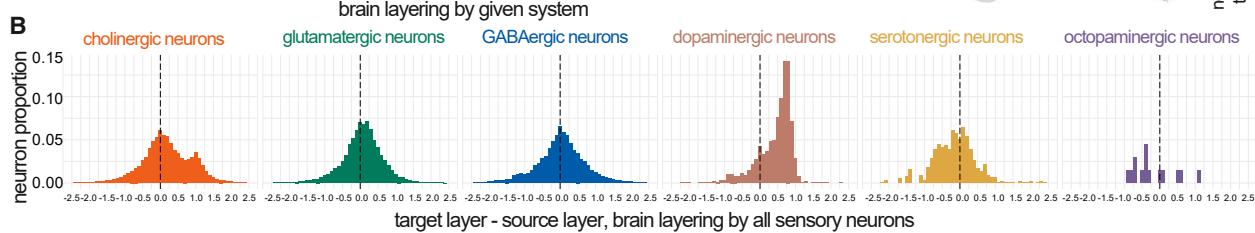
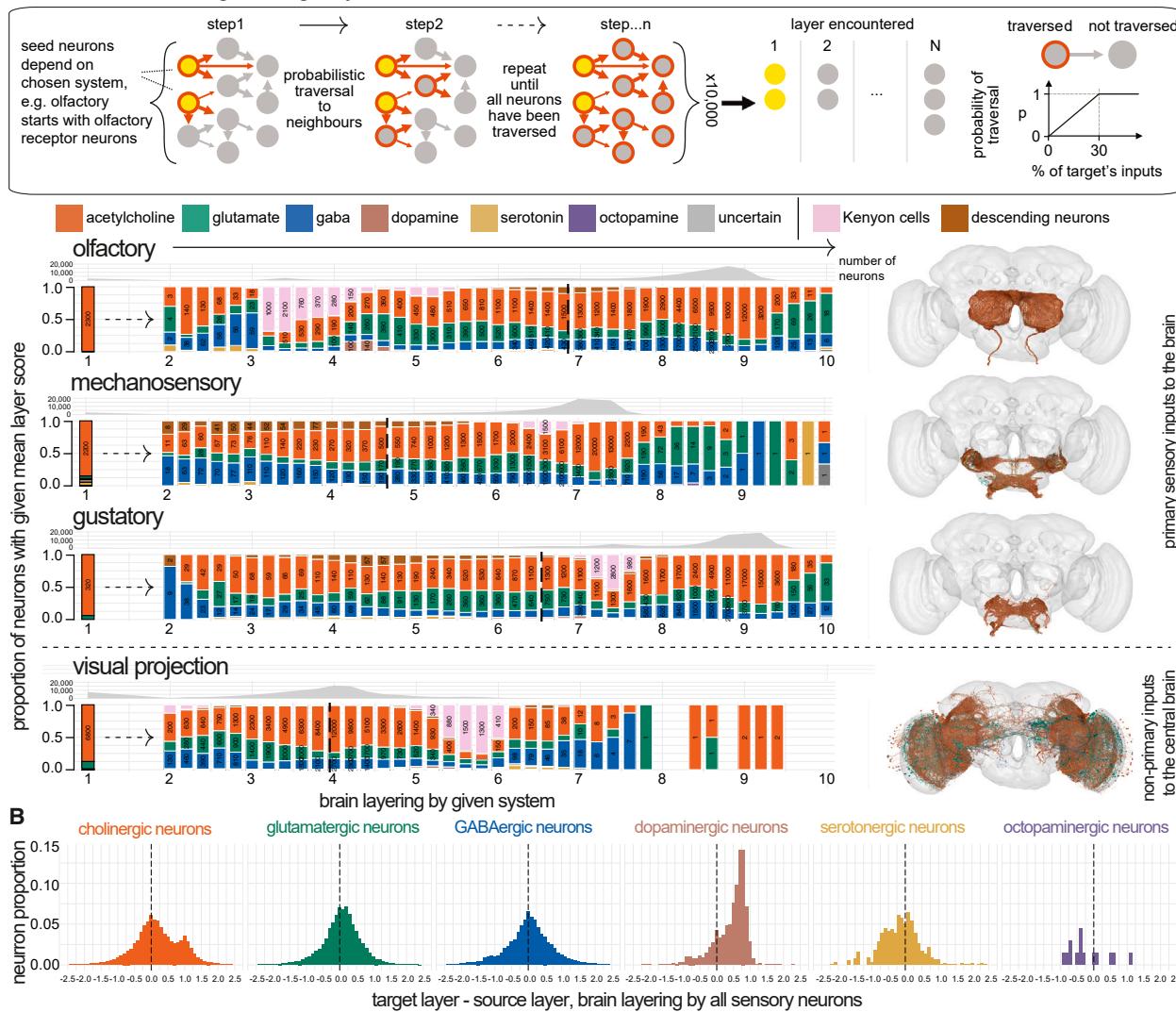
See also Figures S2 and S3 and Data S3 and S4.

or GABA (Figure S4A), in patterns that were consistent between the two datasets (Figure S4B, cosine similarity > 0.9). Axons may therefore be particularly selective in their inputs, particularly inhibitory inputs (Figure S3E).

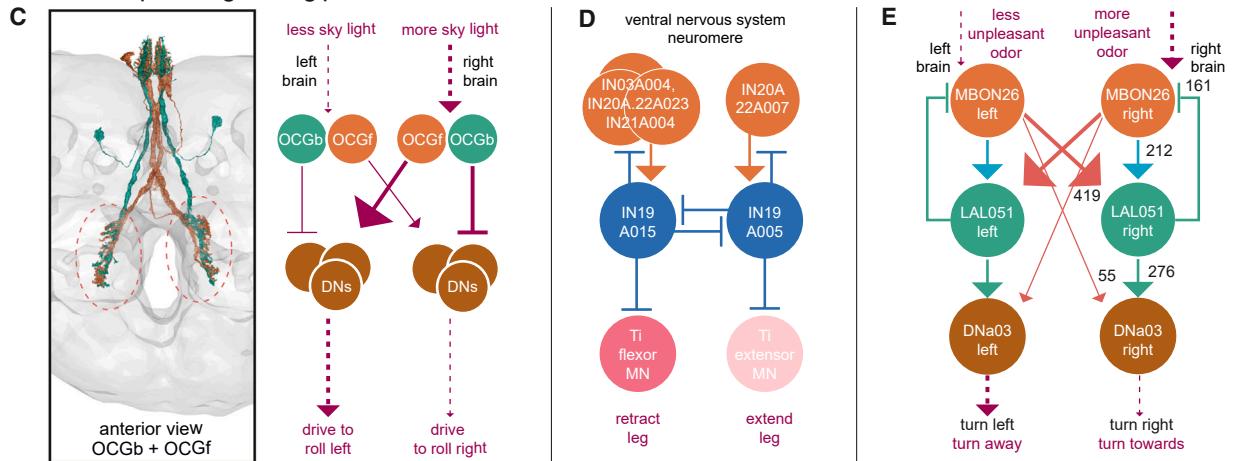
Comparing putative inhibitory and excitatory neurons

Putative excitatory neurons predominate in the brain (Figure 5A). The principal excitatory transmitter is acetylcholine,^{84,85} and the main inhibitory transmitter is GABA.^{86–88} Glutamate can act in

A transmitter usage through layers of the brain FAFB-FlyWire



examples insight using predicted transmitters



(legend on next page)

either an inhibitory^{89–92} or excitatory capacity.^{93–95} Analyzing the FAFB-FlyWire dataset using a probabilistic layer assignment model,^{18,74} we observed shifts in the proportions of transmitter use across sensory systems. For example, there is a greater proportion of GABAergic neurons in the early olfactory system, with glutamatergic neurons rising later. This switch also suggests that neurons in deeper layers are likely to be inhibited by glutamate because GABA is scarce.

Cholinergic (mean 0.21, SD, 0.75) and glutamatergic (mean 0.07, SD, 0.57) neurons tended to target higher-layer neurons (Figure 6B) i.e., predominantly feedforward connectivity; GABAergic neurons had no such bias (mean –0.02, SD, 0.61). On average, GABAergic neurons were more “local” in Euclidean space than cholinergic neurons (Figure 5E upper); they were also smaller by cable length (Figure S3A) and less polarized (Figure 5E lower). However, on average GABAergic neurons had more input neurons and downstream targets (Figure S3I). These connections may have a stronger effect than their cholinergic counterparts for two reasons; first, they show greater synaptic count (Figure S3F), and second, because they are located slightly closer to the target neuron’s primary branch point (Figure S3H), which could enable more powerful inhibition.⁹⁶ Additionally, GABAergic neurons had higher mitochondrial density than cholinergic or glutamatergic neurons (Figure S3C), which could indicate a higher level of energy use and neuronal activity. Lastly, GABAergic neurons received more excitatory than inhibitory drive onto their dendrites compared with cholinergic neurons (Figure S3E). Instead, cholinergic neurons often have large inhibitory inputs onto their axons (median, –0.16). Together, this could mean that GABAergic neurons are more active and integrate a wider array of inputs to inhibit a wider array of downstream neurons than cholinergic neurons; both their axons and dendrites make outputs but these are mainly local, perhaps inhibiting many competing elements in a local circuit.^{73,87,97,98} In particular, they may gate the output of cholinergic axons. On all these metrics, glutamatergic neurons lie between GABA and acetylcholine.

Example transmitter-dependent circuit hypotheses

Our brain-wide transmitter predictions now enable many testable circuit hypotheses. We present three examples that also illustrate how we think about potential confounds.

Ocellar righting circuit

In a vignette on ocellar circuitry by Dorkenwald et al.,¹⁸ our FAFB-FlyWire predictions strongly predicted that 12 OCG01 neurons (with extremely similar axonal morphologies) fall into three glutamatergic and three cholinergic cell types (Figure 6C). A hypothesis emerged: these neurons form pairs, each comprising one inhibitory and one excitatory neuron. For instance, OCG01b (glutamate) and OCG01f (acetylcholine) may collaborate to induce a righting reflex in response to sky-light cues during a roll. The hypothesis was strengthened by the internal control that having two FAFB hemispheres provides; the HemiBrain dataset, in which these neurons are heavily truncated, yielded misleading and likely false predictions (all glutamate).

Leg extension circuit

Cheong et al.⁹⁹ uncovered an inhibitory majority in local circuits controlling leg movement. The GABAergic interneurons IN19A were identified as key regulators, reciprocally inhibiting each other to prevent inappropriate co-contraction of opponent leg muscles. Upstream neurons were predicted cholinergic,⁵ creating a circuit architecture facilitating leg extension; a neuron such as IN21A004 could promote leg extension by inhibiting a flexor and disinhibiting a downstream extensor. In this case, it was initially unclear whether the 6 IN20A projection neurons (one per leg) were cholinergic or glutamatergic since 3/6 were predicted glutamatergic. However, all derive from the same hemilineage, which was majority predicted cholinergic. Since Lacin’s law is strongly followed in the ventral nerve cord,^{31,65} all six IN20A neurons are likely cholinergic.

Simultaneous excitatory and inhibitory control

The cholinergic, bilateral mushroom body output neuron, MBON26, is known to cause turning upon optogenetic activation¹⁰⁰ and innervates mushroom body compartments involved in innate olfactory aversion and appetitive learning.⁷⁹ It directly connects to the downstream turn-control descending neuron, DN03, and indirectly connects via a predicted glutamatergic local neuron, LAL051 (Figure 6E). If glutamate were purely inhibitory, we might expect unilateral MBON26 activation to result in both ipsi- and contralateral DN03 inhibition, with slightly more ipsi-DN03 activity and therefore an ipsilateral turn. If glutamate were purely excitatory, we might expect unilateral MBON26 activation to result in both ipsi- and contralateral DN03 activation, with slightly more contra-DN03 activity and therefore a

Figure 6. Transmitter usage through sensory layers and specific circuits

(A) Schematic depicts the probabilistic graph traversal model used to “layer” different sensory systems, adapted from Schlegel et al.,⁷⁴ underlying data from Dorkenwald et al.¹⁸ Starting from first-order central brain input neurons, we recorded the mean step (“layer”) at which each subsequent FAFB-FlyWire neuron is encountered by the simulation.¹⁸ Bar charts show transmitter input across distinguishable sensory systems.¹⁸ Bars normalized and binned by target neurons’ layer score (width 0.2); text reports neuron count. Kenyon cells are shown in pink and descending neurons, i.e., the last captured point of the sensory-motor transform in the brain, in brown so that the reader can compare layer progression between systems. Vertical line shows the mean descending neuron layer. Olfactory sensory neurons mispredicted for serotonin are corrected to acetylcholine.⁶⁴ “Uncertain” neurons (see STAR Methods) were removed from this analysis.

(B) Feedforward and feedback connectivity across sensory systems by neuron-level transmitter prediction. For each unitary neuron-neuron connection (greater than 100) between a source and target neuron, we calculated a target-source layer difference: the layer value⁷⁴ for the target neuron minus the layer value of the source neuron. Y axis gives the proportion of unitary connections in each bin (width 0.1).

(C) A potential circuit for righting the fly’s body axis relative to celestial cues. Purple arrow weight indicates activity level.

(D) A potential circuit for differential leg extension/retraction control.

(E) A potential circuit for steering away from unpleasant odors. Numbers give synaptic counts from HemiBrain.

See also Figure S4.

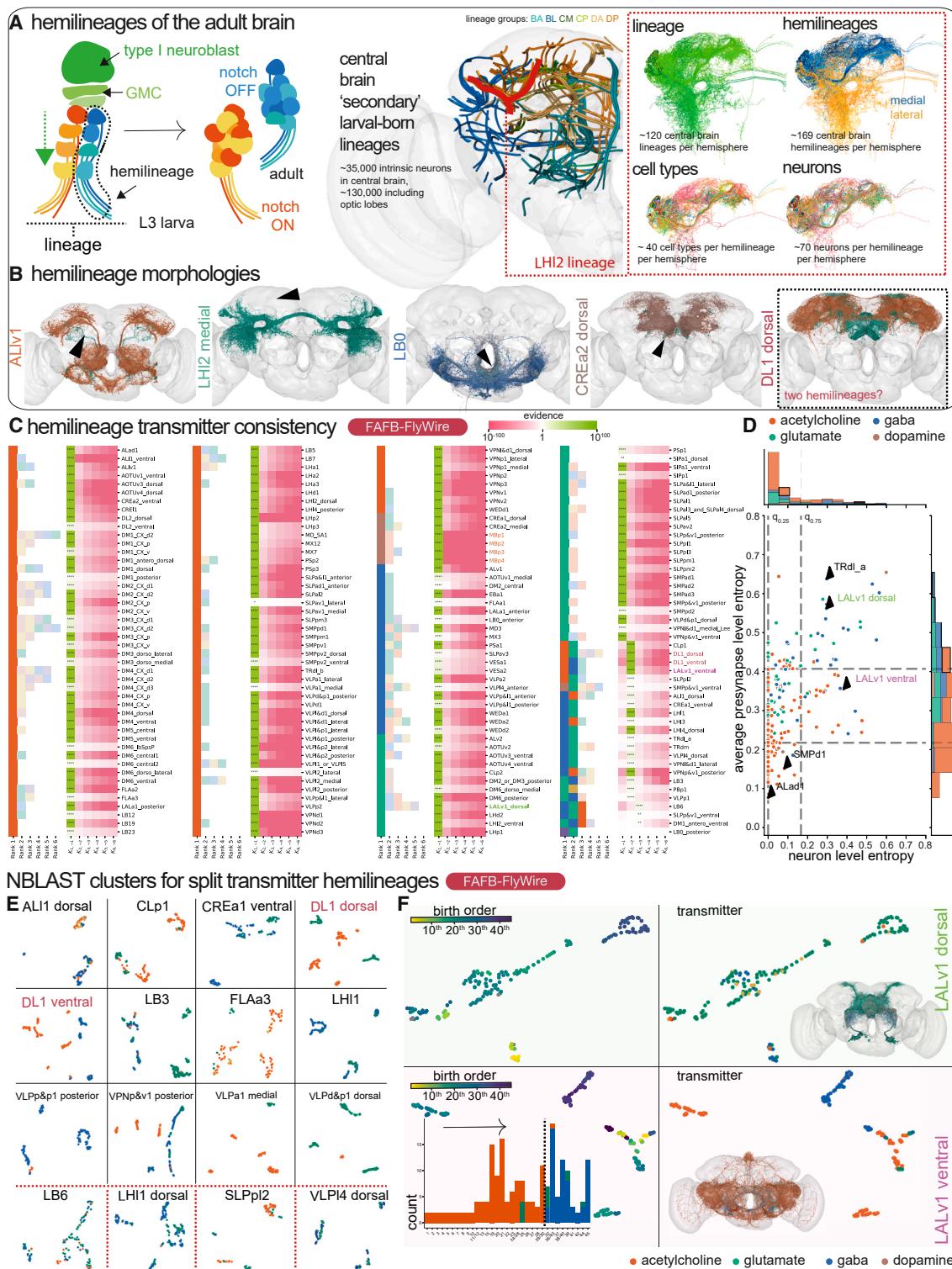


Figure 7. Transmitter usage across all hemilineages in the central fly brain

(A) Left, the progression of a type I neuroblast from third-instar larva (L3) into the adult ganglion mother cell (GMC). Right, a breakdown of a single secondary lineage, "LHI2" into its constituent hemilineages (see STAR Methods).

(legend continued on next page)

contralateral turn. If glutamate excited DNao3 (via AMPA, kainate, or NMDA receptors, e.g., Li et al.¹⁰¹) and inhibited MBON26 (e.g., via *GluCl/Alpha*), we might expect unilateral MBON26 activation to result in ipsilateral DNao3 inhibition and contralateral DNao3 activation, and so a stronger contralateral turn command. Therefore, combining both circuit structure and transmitter predictions, we suspect that glutamate both excites and inhibits in this circuit to steer the fly away from an aversive olfactory stimulus.

The distribution of neurotransmitter predictions within developmental units

The nervous system may already naturally group neurons by their transmitter expression because of their development as hemilineages:³¹ these are groups of ~100 neurons that have developed together in a discrete bundle, the hemilineage tract (see STAR Methods for assignment detail). To assess Lacin's law in the brain, we examined the neuron-level transmitter predictions of all neurons in the 183 secondary (larval-born) hemilineages per brain hemisphere in the FlyWire dataset⁵⁰ (Data S5 and Data S6).

We asked how likely it is to observe a given prediction of transmitters in a hemilineage under some error rate given by the confusion matrix on the test set, if Lacin's law is obeyed. We then compared this likelihood to the alternative hypothesis that a hemilineage consists of neurons with more than one transmitter. We calculated the Bayes factor $K_{2,1} = \frac{p(\hat{y}|m=2)}{p(\hat{y}|m=1)}$ and $K_{3,2} = \frac{p(\hat{y}|m=3)}{p(\hat{y}|m=2)}$ for our selected hemilineages from synapse-level transmitter predictions: i.e., the likelihood ratio of the observed model predictions given that a hemilineage expresses two rather than one transmitter or three rather than two transmitters, respectively (see STAR Methods). Maximal one-versus-rest Bayes factors ($K_{m,-m}$) summarized our predictions of the number and set of transmitters for each hemilineage (Figure 7C). We found only 19 of 183 hemilineages with evidence of expressing two transmitters ($n = 19$ decisive) and 3 hemilineages with evidence of expressing three fast-acting transmitters ($n = 1$ decisive, $n = 2$ good). These hemilineages are flagged in Figure 7C.

However, some of these hemilineages (~12) such as "TRdla" and "LALv1 dorsal" showed high synaptic entropy $H(S_h)$ (see STAR Methods, Figure 7D), indicating that individual neurons within the hemilineage contain substantial multimodal transmitter predictions (Figure 7C). As such, multimodality at the neuron level is at least partially explained by uncertain or inhomogeneous predictions between individual synapses within a neuron. This is in contrast to hemilineage "LALv1 ventral," which has a synaptic entropy within the 75% percentile, and for which a large Bayes factor $K_{2,1}$ value directly stems from neuron-level segregation of the predicted transmitters within the hemilineage. We predicted the remaining 163 hemilineages to express a single transmitter ($n = 161$ decisive, $n = 1$ good, $n = 1$ substantial). Results for 154 HemiBrain hemilineages were very similar (Figure S5A, Data S5).

Therefore, 88% of our hemilineages' predictions were strongly biased toward a singular transmitter identity (Figures 7B, 7C, and S5A). The entropy of neuron-level transmitter predictions for each hemilineage was correlated strongly between left and right hemispheres (Figure S5C), suggesting that the observed variation is biological in origin rather than data quality related. In some cases, only a few neurons deviated from the majority prediction (Figure 7B, black arrows). We suspect that these neurons are the "first-born" neurons of the hemilineage, which often have a divergent morphology from the rest of the hemilineage^{31,102}; they may also be divergent in their transmitter usage.

Other prominently split hemilineages also demonstrated morphology-correlated shifts in transmitter expression (Figures 7C, S5A, and S5D). This suggests discrete switches in expression during development, presumably at stereotyped developmental time points. For instance, LALv1 ventral (Figure 7F) revealed a discernible switch in transmitter expression accompanied by morphological differences. Because a recent analysis delineated the birth order of LALv1 neurons,¹⁰² we were able to map this transmission-morphotype switch into its developmental sequence (Figure 7F, lower). In a few cases (Figure 7E, lower), split expression occurred without overt morphological differences, raising questions about sporadic switches and potential confounds. Often, glutamatergic and GABAergic neurons are mixed—our network's most common confusion

(B) Example of homologous FAFB-FlyWire hemilineages on both sides of the brain, colored by neuron-level transmitter prediction. Black arrows point to one stray member of each hemilineage with a different neuron-level transmitter prediction, which is likely a first-born neuron with distinct morphology^{31,102} (Data S5). The dashed box indicates a hemilineage with potential split transmitter expression.

(C) Bayes factor analysis of hemilineage consistency. For each hemilineage (row), the right set of columns corresponds to the likelihood ratio of the hemilineage expressing that number of transmitters versus the likelihood of any other number ($\lambda = 16$; $c_{\text{exp}} = 0.67$). Evidence strength: substantial * $K \geq 10^{1/2}$; good ** $K \geq 10^1$; strong *** $K \geq 10^{3/2}$; decisive **** $K \geq 10^2$.¹⁰³ The left set of columns indicates the frequency ranked transmitter predictions within the hemilineage, with ranks greater than the maximum likely number of transmitters shaded lighter. "LB0 posterior" did not have substantial evidence for any particular number of transmitters.

(D) Neuron level entropy ($H(N_h)$) versus average synapse level entropy ($H(S_h)$) for all predicted hemilineages with more than 10 neurons and more than 30 pre-synapses per neuron. Dashed lines indicate 25% and 75% percentiles: $q_{25}(H(N_h)) = 0.00$, $q_{25}(H(S_h)) = 0.22$, $q_{75}(H(N_h)) = 0.17$, and $q_{75}(H(S_h)) = 0.41$. Type equation here.

(E) NBLAST UMAP plots of selected hemilineages that exhibit some degree of predicted split transmitter usage. UMAPs are based on NBLAST morphological similarity scores between all possible pairs of neurons in each hemilineage. Points represent neurons, colored by neuron-level transmitter prediction. Black solid lines bound examples with a morphology-transmitter split, red dashed lines bound examples with no such clear divide. Red text label examples for which data annotation issues may explain split usage.

(F) "LALv1" has two hemilineages: dorsal (developmentally defined by Notch-ON) and ventral (Notch-OFF). NBLAST UMAPs for the two "LALv1" hemilineages (rows): dorsal (developmentally defined by Notch-ON) and ventral (Notch-OFF), colored by birth order (left) and neuron-level transmitter prediction (right). Bottom left, histogram of neuron-level transmitter prediction by birth order.¹⁰²

See also Figure S5 and Data S6.

(Figure 2A). Notably, potentially as much as 3% of the brain expresses mRNA for machinery related to both,²¹ and at least a few may transmit both.¹⁰⁴ Comparing ~6 GABA-glutamate mixed hemilineages in FlyWire to HemiBrain indicates that perhaps they more uniformly express GABA (Figure S5A). In two cases (the “DL1” hemilineages), we think that the striking morphology-transmitter correlated splits reveal a hemilineage-based division in a lineage-associated tract that was otherwise hard for human annotators to make, i.e., an issue of data annotation.

Challenges surfaced with the “AL1 dorsal” and “ALv2” hemilineages that produce antennal lobe local neurons, a morphologically variable and diverse class of neuron.¹⁰⁵ They seem to break Dale’s law^{104,106} and Lacin’s law (Figure 7C), with similar morphology types predicted to express different transmitters (Figures S5F and S5G). Despite these challenges, we could align our results with existing literature (see STAR Methods). Surprisingly, 18%–27% of antennal lobe local neurons may be cholinergic, suggesting that lateral excitation is a more prominent feature of antennal lobe processing than previously thought.

DISCUSSION

Using high-level annotations to learn low-level features

In evaluating our predictions as broadly as we could, we have found them to be correct for 91% of 624 FAFB-FlyWire cell types and 91% of 524 HemiBrain cell types (Data S7). This result likely depended on three properties of the data that we selected. (1) For both training and inference we used a specific sub-cellular domain (the synapse), likely to contain image features related to our molecules of interest. (2) We aggregated these features on a per-cell basis; this was crucial for linking ground-truth labels and input data and also for improved prediction accuracy. (3) By using cross-modal matching of cell types between EM and light-level neuronal morphology data, we were able to build an expansive ground-truth dataset from molecular information external to the EM data. We anticipate broad applications for this general approach in the biological sciences. Our methodology could be repurposed to identify differences in sub-cellular structures¹⁰⁷ associated with discrete cell types in a range of neuronal and non-neuronal tissues.

Machine learning reveals key molecular descriptors of neuronal function

D. melanogaster represented a hard case for transmitter prediction because humans cannot tell fast-acting transmitters apart in insects. Our classifier accurately predicts transmitter identity from local 3D EM volumes (Figure 2). We find that the overall accuracy of our synapse-level transmitter predictions were slightly more performant in FAFB than in HemiBrain; in particular, HemiBrain is more likely to confuse glutamate for GABA (Figure 2A). This might be because of different sample preparation and staining protocols used for those datasets, our use of only manually annotated presynapses for FAFB ground truth, and/or the higher lateral resolution of FAFB (4 nm vs. 8 nm).

Given that the relationship between synaptic phenotype and transmitter identity is not understood in *D. melanogaster*, we built an explainable AI method that created counterfactual synthetic

images to investigate class differences⁷¹ (Figures 3A and 3B). This enabled us to manually identify at least one feature difference between each pair of fast-acting transmitters (Figure 3C). However, a simple logistic regression classifier trained on three of those features was not able to discriminate between fast-acting transmitters at the same level as the classifier ($n = 219$, 80%/20% randomized training/testing, balanced accuracy 0.52). This indicates that the features we identified are not exhaustive.

We expect our pipeline to be transferable to other connectomic datasets from diverse species. Analogous work detecting excitatory versus inhibitory synapses in the *Ciona intestinalis* larva has added signs for 49 neurons.¹⁰⁸ In vertebrate EM datasets, human annotators can see differences in the vesicles for excitatory and inhibitory synapses.^{16,35–37,109,110} Symmetric synapses (usually inhibitory) have already been disambiguated from asymmetric ones (usually excitatory) automatically and at scale.³⁸ Using our methodology, performant and more specific transmitter predictions are likely to be achievable in both vertebrate and invertebrate datasets.

Sign labeling across a whole-insect central nervous system

Dale’s law, despite its exceptions, serves as a valuable framework for understanding neural action in the face of incomplete data. We needed to use it to accumulate synapse-level transmitter predictions into singular neuron-level transmitter predictions. Extending our conceptual framework, we found that Lacin’s law³¹ is a useful principle that holds in 88% of hemilineages (Figure 7C). We posit that the major exceptions emerge as correlated, discrete switches in both transmitter use and gross morphotype (Figure 7F). Lacin’s law serves the network anatomist in three main ways: supporting prediction validity (Figures S5A and 2D), flagging potential errors when a neuron-level transmitter prediction deviates from the hemilineage majority, and allowing for preliminary transmitter labels to be applied to whole hemilineages in new insect datasets.^{111,112}

The functional role of a synaptic connection depends on upstream transmitters and downstream receptors. Simulating connectomes directly¹¹³ at a whole nervous system scale may be possible with our method (e.g., Shiu et al.¹¹⁴), but incorporating information on postsynaptic receptor expression can make brain models more accurate.¹¹³ Most importantly, while acetylcholine excites and GABA inhibits, glutamate can perform either function in the fly. In the *D. melanogaster* central brain, most reported examples for glutamate are of inhibition via *GluClAlpha* channels.^{89–92,115} The nature of this inhibition could be different, e.g., possibly subtractive⁹¹ rather than divisive.^{86–88,115} Indeed, glutamate may be used more for specific long-range feedforward inhibition and GABA for local divisive inhibition (Figure 6A). However, central excitatory glutamatergic transmission has also been reported.^{93,95,116,117} At the neuromuscular junction, glutamate primarily excites.⁹⁴ We calculated from single-cell RNA sequencing¹¹⁸ that ~80% of neurons in the central brain express RNA transcripts for *GluClAlpha* as well as at least one excitatory ionotropic glutamate receptor. The sign of a glutamatergic connection may depend on the sub-localization and ratio of receptors at recipient sites.¹¹⁹

Beyond fast-acting transmitters, other factors—such as cellular compartments⁸⁰—influence the strengths and signs of neuron-neuron connections. Notably, a large fraction of the synaptic budget is spent on axo-axonic connections (FAFB-FlyWire, 22%; HemiBrain, 20%). Putative inhibitory connections accounted for two-thirds of the axo-axonic sub-budget (Figure 5D) and axons often receive a skewed excitatory:inhibitory connection ratio (Figure 5F). This suggests that the nervous system may employ different circuit logic around axons compared with dendrites.

Limitations of the study

We can think of this study's limitations in two major ways: misprediction (perhaps largely because of co- or alternative transmission, see STAR Methods) by our classifier and other limitations on sign labeling from transmitter identity alone. Outstanding problems that must be solved before we have highly accurate connection signs minimally include: (1) the correction of transmitter mispredictions, (2) the prediction of other transmitters, notably histamine, tyramine, glycine, nitric oxide, and ~53 different neuropeptides, (3) the annotation of monoaminergic co-transmission, (4) the annotation of peptidergic (co-)transmission, (5) the annotation of postsynaptic receptor expression, and (6) the annotation of gap junctions. For (1), community open science annotation projects such as FlyWire via Codex¹²⁰ and Virtual Fly Brain¹²¹ will be valuable. In particular, our least performant predictions were for serotonin. Notably, ~20 known, large, neuropeptidergic neurons are predicted for serotonin. This may be because we included peptidergic, serotonergic neurons in our ground truth.^{122,123} Answers to (2), (3), and (4) could be achieved by linking morphological and transcriptomic cell types⁴⁹ to build more ground truth. One could extend our method to further transmitters and co-transmission combinations as additional training data becomes available (we have shared our literature review of cell type level co-transmission, Data S7). In addition, the detection of individual vesicles, especially dense core vesicles at synapses, or somata could tell us which neurons express neuropeptides and peptide hormones. For problems (5) and (6), linked transcriptomes alone are unlikely either to solve the issue or to provide suitable ground truth for their detection in EM. One path forward would be to assemble ground-truth data with super-resolution light-level microscopy to observe endogenous protein sub-localization^{119,124,125} or with higher-resolution EM to identify key molecules at identified synapses in dense biological samples.¹²⁶ As we showed here with transmitter usage, the identification of some hundreds of pairs of connected neurons may provide sufficient ground truth for a machine-learning solution to predict the remainder across whole nervous systems.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability

- Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
 - Detail on assembling ground truth data for *D. melanogaster*
 - Chosen neuronal reconstructions
 - Neuron skeletonization and axon-dendrite splits
 - Neuron level transmitter prediction confidence
 - Correlating misprediction with reported unlearned transmission types
 - Working through fan-shaped body neuron expression
 - Working through antennal lobe local neuron expression
- HEMILINEAGE DEFINITION IN *D. MELANOGASTER*
 - Hemilineage assignments in *D. melanogaster*
 - Hemilineage prediction entropy
 - Probability to observe transmitter predictions y
 - Bayes Factor
 - Identification of ultrastructural features
 - Advice for resource users
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2024.03.016>.

ACKNOWLEDGMENTS

This work was supported by a Wellcome Trust Collaborative Award (203261/Z/16/Z), National Institutes of Health (BRAIN Initiative grant 1RF1MH120679-01), an ERC Consolidator grant (649111), and a Neuronex2 award (MC_EX_MR/T046279/1, NSF ref. 2014862), with additional core support from the MRC (MC-U105188491) to G.S.X.E.J.; by the Howard Hughes Medical Institute at Janelia Research Campus (J.F.); and a Boehringer Ingelheim Fonds PhD Fellowship, EMBO fellowship (ALTF 1258-2020), and a Sir Henry Wellcome Postdoctoral Fellowship (222782/Z/21/Z) to A.S.B. Flywire is supported by NIH BRAIN Initiative grant MH117815 to M.M. and S.S. Additional large-scale proofreading and infrastructure was supported by Wellcome awards 203261/Z/16/Z to G.S.X.E.J. and NIH BRAIN Initiative award 1RF1MH120679-01 and NSF NeuroNex award DBI-2014862 to D. Bock and G.S.X.E.J. We thank Laia Serratosa Capdevila (Aelysia Ltd) for copyediting. Please note that we also name individuals from the FAFB community who have assisted with building the data and annotations used in this study in the STAR Methods.

AUTHOR CONTRIBUTIONS

Conceptualization, A.S.B., J.F., and P.S.; methodology A.C., A.S.B., G.S.X.E.J., J.F., M.D., and N.E.; software, A.C., A.S.B., G.S.X.E.J., J.F., M.D., and N.E.; validation, A.C., A.S.B., J.F., M.D., N.E., and P.S.; formal analysis, A.C., A.S.B., J.F., N.E., and Y.Y.; neuroanatomy, A.S.B. and Y.Y.; investigation, A.S.B., J.F., M.D., N.E., and Y.Y.; resources, A.C., A.S.B., G.S.X.E.J., J.F., M.D., N.E., and P.S.; data curation, A.C., A.K.-Y.L., A.S.B., J.F., N.E., P.S., R.P., S.F.-M., T.P., T.R., Y.Y., and V.H.; writing – original draft, A.S.B., G.S.X.E.J., J.F., N.E., and V.H.; writing – review & editing, A.S.B. and J.F.; visualization, A.C., A.S.B., J.F., N.E., and Y.Y.; supervision, A.S.B., G.S.X.E.J., and J.F.; project administration, A.S.B., G.S.X.E.J., and J.F.; funding acquisition, A.S.B., G.S.X.E.J., and J.F. We also credit as authors the key contributors to the FlyWire project, from which we received early data access: FlyWire infrastructure and management, S.D.; FlyWire codex, A.M.; FlyWire proofreader training and community management, A.S., C.M., S.-C.Y., K.E., M.C., and P.S.; FlyWire annotations and cell type matching, P.S., K.E., A.S.B., and G.S.X.E.J.; FlyWire leads, M.M. and S.S.; Cambridge lead, G.S.X.E.J.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 2, 2023
 Revised: October 4, 2023
 Accepted: March 13, 2024
 Published: May 9, 2024

REFERENCES

1. Cook, S.J., Jarrell, T.A., Brittin, C.A., Wang, Y., Bloniarz, A.E., Yakovlev, M.A., Nguyen, K.C.Q., Tang, L.T.-H., Bayer, E.A., Duerr, J.S., et al. (2019). Whole-animal connectomes of both *caenorhabditis elegans* sexes. *Nature* **571**, 63–71.
2. Ohyama, T., Schneider-Mizell, C.M., Fetter, R.D., Aleman, J.V., Franconville, R., Rivera-Alba, M., Mensh, B.D., Branson, K.M., Simpson, J.H., Truman, J.W., et al. (2015). A multilevel multimodal circuit enhances action selection in *drosophila*. *Nature* **520**, 633–639.
3. Phelps, J.S., Hildebrand, D.G.C., Graham, B.J., Kuan, A.T., Thomas, L.A., Nguyen, T.M., Buhmann, J., Azevedo, A.W., Sustar, A., Agrawal, S., et al. (2021). Reconstruction of motor control circuits in adult *drosophila* using automated transmission electron microscopy. *Cell* **184**, 759–774.e18.
4. Ryan, K., Lu, Z., and Meinertzhagen, I.A. (2016). The cns connectome of a tadpole larva of *ciona intestinalis* (l.) highlights sidedness in the brain of a chordate sibling. *Elife* **5**, e16962.
5. Takemura, S.-Y., Hayworth, K.J., Huang, G.B., Januszewski, M., Lu, Z., Marin, E.C., Preibisch, S., Shan Xu, C., Bogovic, J., Champion, A.S., et al. (2023). A connectome of the male *drosophila* ventral nerve cord. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.05.543757>.
6. Zheng, Z., Lauritzen, J.S., Perlman, E., Robinson, C.G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C.B., Sharifi, N., et al. (2018). A complete electron microscopy volume of the brain of adult *drosophila melanogaster*. *Cell* **174**, 730–743.e22.
7. Funke, J., Tschopp, F.D., Grisaitis, W., Sheridan, A., Singh, C., Saalfeld, S., and Turaga, S.C. (2018). Large scale image segmentation with structured loss based deep learning for connectome reconstruction. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1669–1680.
8. Januszewski, M., Kornfeld, J., Li, P.H., Pope, A., Blakely, T., Lindsey, L., Maitin-Shepard, J., Tyka, M., Denk, W., and Jain, V. (2018). High-precision automated reconstruction of neurons with flood-filling networks. *Nat. Methods* **15**, 605–610.
9. Lee, K., Lu, R., Luther, K., and Seung, H.S. (2019). Learning Dense Voxel Embeddings for 3D Neuron Reconstruction. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1909.09872>.
10. Schmidt, M., Motta, A., Sievers, M., and Helmstaedter, M. (2022). RoboEM: automated 3D flight tracing for synaptic-resolution connectomics. Preprint at bioRxiv. <https://doi.org/10.1101/2022.09.08.507122>.
11. Sheridan, A., Nguyen, T.M., Deb, D., Lee, W.-C.A., Saalfeld, S., Turaga, S.C., Manor, U., and Funke, J. (2023). Local shape descriptors for neuron segmentation. *Nat. Methods* **20**, 295–303.
12. Buhmann, J., Sheridan, A., Malin-Mayor, C., Schlegel, P., Gerhard, S., Kazimiers, T., Krause, R., Nguyen, T.M., Heinrich, L., Lee, W.-C.A., et al. (2021). Automatic detection of synaptic partners in a whole-brain *drosophila* electron microscopy data set. *Nat. Methods* **18**, 771–774.
13. Huang, G.B., Scheffer, L.K., and Plaza, S.M. (2018). Fully-Automatic synapse prediction and validation on a large data set. *Front. Neural Circuits* **12**, 87.
14. Kreshuk, A., Funke, J., Cardona, A., and Hamprecht, F.A. (2015). Who is talking to whom: Synaptic partner detection in anisotropic volumes of insect brain. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W.M. Wells, and A. Frangi, eds. (Springer International Publishing), pp. 661–668.
15. Staffler, B., Berning, M., Boergens, K.M., Gour, A., Smagt, P.v.d., and Helmstaedter, M. (2017). Synem, automated synapse detection for connectomics. *Elife* **6**, e26414.
16. Svara, F., Förster, D., Kubo, F., Januszewski, M., Dal Maschio, M., Schubert, P.J., Kornfeld, J., Wanner, A.A., Laurell, E., Denk, W., and Baier, H. (2022). Automated synapse-level reconstruction of neural circuits in the larval zebrafish brain. *Nat. Methods* **19**, 1357–1366.
17. Dorkenwald, S., McKellar, C.E., Macrina, T., Kemnitz, N., Lee, K., Lu, R., Wu, J., Popovych, S., Mitchell, E., Nehoran, B., et al. (2022). FlyWire: online community for whole-brain connectomics. *Nat. Methods* **19**, 119–128.
18. Dorkenwald, S., Matsliah, A., Sterling, A.R., Schlegel, P., Yu, S.-C., McKellar, C.E., Lin, A., Costa, M., Eichler, K., Yin, Y., et al. (2023). Neuronal wiring diagram of an adult brain. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.27.546656>.
19. Winding, M., Pedigo, B.D., Barnes, C.L., Patsolic, H.G., Park, Y., Kazimiers, T., Fushiki, A., Andrade, I.V., Khandelwal, A., Valdes-Aleman, J., et al. (2022). The connectome of an insect brain. *Science* **379**, eadd9330.
20. Azevedo, A., Lesser, E., Mark, B., Phelps, J., Elabbady, L., Kuroda, S., Sustar, A., Moussa, A., Kandelwal, A., Dallmann, C.J., et al. (2022). Tools for comprehensive reconstruction and analysis of *drosophila* motor circuits. Preprint at bioRxiv. <https://doi.org/10.1101/2022.12.15.520299>.
21. Croset, V., Treiber, C.D., and Waddell, S. (2018). Cellular diversity in the *drosophila* midbrain revealed by single-cell transcriptomics. *Elife* **7**, e34550.
22. Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, L., Albar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C., et al. (2018). A single-cell transcriptome atlas of the aging *drosophila* brain. *Cell* **174**, 982–998.e20.
23. Goyal, R.K., and Chaudhury, A. (2013). Structure activity relationship of synaptic and junctional neurotransmission. *Auton. Neurosci.* **176**, 11–31.
24. Abruzzi, K.C., Zadina, A., Luo, W., Wiyanto, E., Rahman, R., Guo, F., Shaffer, O., and Rosbash, M. (2017). RNA-seq analysis of *drosophila* clock and non-clock neurons reveals neuron-specific cycling and novel candidate neuropeptides. *PLoS Genet.* **13**, e1006613.
25. Aso, Y., Ray, R.P., Long, X., Bushey, D., Cichewicz, K., Ngo, T.-T., Sharp, B., Christoforou, C., Hu, A., Lemire, A.L., et al. (2019). Nitric oxide acts as a cotransmitter in a subset of dopaminergic neurons to diversify memory dynamics. *Elife* **8**, e49257.
26. Yew, J.Y., Wang, Y., Barteneva, N., Dikler, S., Kutz-Naber, K.K., Li, L., and Kravitz, E.A. (2009). Analysis of neuropeptide expression and localization in adult *drosophila melanogaster* central nervous system by affinity cell-capture mass spectrometry. *J. Proteome Res.* **8**, 1271–1284.
27. Nässel, D.R., and Winther, A.M.E. (2010). *Drosophila* neuropeptides in regulation of physiology and behavior. *Prog. Neurobiol.* **92**, 42–104.
28. Taghert, P.H., and Veenstra, J.A. (2003). *Drosophila* neuropeptide signaling. *Adv. Genet.* **49**, 1–65.
29. Nusbaum, M.P., Blitz, D.M., and Marder, E. (2017). Functional consequences of neuropeptide and small-molecule co-transmission. *Nat. Rev. Neurosci.* **18**, 389–403.
30. Deng, B., Li, Q., Liu, X., Cao, Y., Li, B., Qian, Y., Xu, R., Mao, R., Zhou, E., Zhang, W., et al. (2019). Chemoconnectomics: Mapping chemical transmission in *drosophila*. *Neuron* **101**, 876–893.e4.
31. Lacin, H., Chen, H.-M., Long, X., Singer, R.H., Lee, T., and Truman, J.W. (2019). Neurotransmitter identity is acquired in a lineage-restricted manner in the *Drosophila* CNS. *Elife* **8**, e43701.
32. Dale, H. (1935). Pharmacology and nerve endings. *Proc. R. Soc. Med.* **28**, 319–332.
33. Eccles, J.C. (1976). From electrical to chemical transmission in the central nervous system: the closing address of the sir henry dale centennial symposium cambridge, 19 september 1975. *Notes and Records of the Royal Society of London* **30**, 219–230.
34. Yang, C.-P., Fu, C.-C., Sugino, K., Liu, Z., Ren, Q., Liu, L.-Y., Yao, X., Lee, L.P., and Lee, T. (2016). Transcriptomes of lineage-specific *drosophila* neuroblasts profiled by genetic targeting and robotic sorting. *Development* **143**, 411–421.

35. Atwood, H.L., Lang, F., and Morin, W.A. (1972). Synaptic vesicles: selective depletion in crayfish excitatory and inhibitory axons. *Science* **176**, 1353–1355.
36. Tao, C.-L., Liu, Y.-T., Sun, R., Zhang, B., Qi, L., Shivakoti, S., Tian, C.-L., Zhang, P., Lau, P.-M., Zhou, Z.H., and Bi, G.-Q. (2018). Differentiation and characterization of excitatory and inhibitory synapses by cryo-electron tomography and correlative microscopy. *J. Neurosci.* **38**, 1493–1510.
37. Uchizono, K. (1965). Characteristics of excitatory and inhibitory synapses in the central nervous system of the cat. *Nature* **207**, 642–643.
38. Dorkenwald, S., Schubert, P.J., Killinger, M.F., Urban, G., Mikula, S., Svara, F., and Kornfeld, J. (2017). Automated synaptic connectivity inference for volume electron microscopy. *Nat. Methods* **14**, 435–442.
39. Davis, F.P., Nern, A., Picard, S., Reiser, M.B., Rubin, G.M., Eddy, S.R., and Henry, G.L. (2020). A genetic, genomic, and computational resource for exploring neural circuit function. *Elife* **9**, e50901.
40. Henry, G.L., Davis, F.P., Picard, S., and Eddy, S.R. (2012). Cell type-specific genomics of drosophila neurons. *Nucleic Acids Res.* **40**, 9691–9704.
41. Hyatt, A.D., and Wise, T.G. (2001). Immunolabeling. In *Immunocytochemistry and In Situ Hybridization in the Biomedical Sciences* (Springer), pp. 73–107.
42. Konstantinides, N., Rossi, A.M., and Desplan, C. (2015). Common temporal identity factors regulate neuronal diversity in fly ventral nerve cord and mouse retina. *Neuron* **85**, 447–449.
43. Long, X., Colonell, J., Wong, A.M., Singer, R.H., and Lionnet, T. (2017). Quantitative mrna imaging throughout the entire drosophila brain. *Nat. Methods* **14**, 703–706.
44. Meissner, G.W., Nern, A., Singer, R.H., Wong, A.M., Malkesman, O., and Long, X. (2019). Mapping neurotransmitter identity in the whole-mount drosophila brain using multiplex high-throughput fluorescence *in situ* hybridization. *Genetics* **211**, 473–482.
45. Jenett, A., Rubin, G.M., Ngo, T.T.B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., et al. (2012). A GAL4-Driver line resource for drosophila neurobiology. *Cell Rep.* **2**, 991–1001.
46. Luan, H., Diao, F., Scott, R.L., and White, B.H. (2020). The drosophila split gal4 system for neural circuit mapping. *Front. Neural Circuits* **14**, 603397.
47. Costa, M., Manton, J.D., Ostrovsky, A.D., Prohaska, S., and Jefferis, G.S.X.E. (2016). NBLAST: Rapid, sensitive comparison of neuronal structure and construction of neuron family databases. *Neuron* **91**, 293–311.
48. Otsuna, H., Ito, M., and Kawase, T. (2018). Color depth MIP mask search: a new tool to expedite Split-GAL4 creation. Preprint at bioRxiv. <https://doi.org/10.1101/318006>.
49. Bates, A.S., Janssens, J., Jefferis, G.S., and Aerts, S. (2019). Neuronal cell types in the fly: single-cell anatomy meets single-cell genomics. *Curr. Opin. Neurobiol.* **56**, 125–134.
50. Schlegel, P., Yin, Y., Bates, A.S., Dorkenwald, S., Eichler, K., Brooks, P., Han, D.S., Gkantia, M., dos Santos, M., Munnelly, E.J., et al. (2023). A consensus cell type atlas from multiple connectomes reveals principles of circuit stereotypy and variation. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.27.546055>.
51. Scheffer, L.K., Xu, C.S., Januszewski, M., Lu, Z., Takemura, S.-Y., Hayworth, K.J., Huang, G.B., Shinomiya, K., Maitlin-Shepard, J., Berg, S., et al. (2020). A connectome and analysis of the adult drosophila central brain. *Elife* **9**, e57443.
52. Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1409.1556>.
53. Kingma, D.P., and Ba, J. (2014). Adam: A method for stochastic optimization. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1412.6980>.
54. Ito, K., Shinomiya, K., Ito, M., Armstrong, J.D., Boyan, G., Hartenstein, V., Harzsch, S., Heisenberg, M., Homberg, U., Jenett, A., et al. (2014). A systematic nomenclature for the insect brain. *Neuron* **81**, 755–765.
55. Barnstedt, O., Owald, D., Felsenberg, J., Brain, R., Moszynski, J.-P., Talbot, C.B., Perrat, P.N., and Waddell, S. (2016). Memory-Relevant mushroom body output synapses are cholinergic. *Neuron* **89**, 1237–1247.
56. Haynes, P.R., Christmann, B.L., and Griffith, L.C. (2015). A single pair of neurons links sleep to memory consolidation in *drosophila melanogaster*. *Elife* **4**, e03868.
57. Omoto, J.J., Nguyen, B.-C.M., Kandimalla, P., Lovick, J.K., Donlea, J.M., and Hartenstein, V. (2018). Putative interactions within the drosophila ellipsoid body neuropil. *Front. Neural Circuits* **12**, 103.
58. Park, S.K., George, R., Cai, Y., Chang, H.Y., Krantz, D.E., Friggi-Grelin, F., Birman, S., and Hirsh, J. (2006). Cell-type-specific limitation on *in vivo* serotonin storage following ectopic expression of the drosophila serotonin transporter, dSERT. *J. Neurobiol.* **66**, 452–462.
59. Shih, H.-W., and Chiang, A.-S. (2011). Anatomical characterization of thermosensory AC neurons in the adult drosophila brain. *J. Neurogenet.* **25**, 1–6.
60. Waddell, S., Armstrong, J.D., Kitamoto, T., Kaiser, K., and Quinn, W.G. (2000). The amnesiac gene product is expressed in two neurons in the drosophila brain that are critical for memory. *Cell* **103**, 805–813.
61. Das, A., Chiang, A., Davla, S., Priya, R., Reichert, H., VijayRaghavan, K., and Rodrigues, V. (2011a). Identification and analysis of a glutamatergic local interneuron lineage in the adult drosophila olfactory system. *Neural Syst. Circuits* **1**, 4.
62. Python, F., and Stocker, R.F. (2002). Immunoreactivity against choline acetyltransferase, gamma-aminobutyric acid, histamine, octopamine, and serotonin in the larval chemosensory system of dosophila melanogaster. *J. Comp. Neurol.* **453**, 157–167.
63. Shang, Y., Claridge-Chang, A., Sjulson, L., Pypaert, M., and Miesenböck, G. (2007). Excitatory local circuits and their implications for olfactory processing in the fly antennal lobe. *Cell* **128**, 601–612.
64. Yasuyama, K., and Salvaterra, P.M. (1999). Localization of choline acetyltransferase-expressing neurons in drosophila nervous system. *Microsc. Res. Tech.* **45**, 65–79.
65. Marin, E.C., Morris, B.J., Stuerner, T., Champion, A.S., Krzeminski, D., Badalamente, G., Gkantia, M., Dunne, C.R., Eichler, K., Takemura, S.-Y., et al. (2023). Systematic annotation of a complete adult male drosophila nerve cord connectome reveals principles of functional organisation. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.05.543407>.
66. Ancona, M., Ceolini, E., Öztireli, C., and Gross, M. (2018). Towards better understanding of gradient-based attribution methods for deep neural networks. In *International Conference on Learning Representations*.
67. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pp. 618–626.
68. Shrikumar, A., Greenside, P., Shcherbina, A., and Kundaje, A. (2016). Not just a black box: Learning important features through propagating activation differences. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1605.01713>.
69. Sundararajan, M., Taly, A., and Yan, Q. (2017). Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning, Volume 70 (JMLR.org)*, pp. 3319–3328.
70. Kindermans, P.-J., Schütt, K.T., Alber, M., Müller, K.-R., Erhan, D., Kim, B., and Dähne, S. (2017). Learning how to explain neural networks: Patternnet and patternattribution. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1705.05598>.
71. Eckstein, N., Bukhari, H., Bates, A.S., Jefferis, G.S., and Funke, J. (2023). Discriminative attribution from paired images. In *Computer Vision-ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part IV (Springer)*, pp. 406–422.

72. Zhu, J.-Y., Park, T., Isola, P., and Efros, A.A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision, pp. 2223–2232.
73. Schneider-Mizell, C.M., Gerhard, S., Longair, M., Kazimiers, T., Li, F., Zwart, M.F., Champion, A., Midgley, F.M., Fetter, R.D., Saalfeld, S., and Cardona, A. (2016). Quantitative neuroanatomy for connectomics in *drosophila*. *Elife* 5, e12059.
74. Schlegel, P., Bates, A.S., Stürner, T., Jagannathan, S.R., Drummond, N., Hsu, J., Serratosa Capdevila, L., Javier, A., Marin, E.C., Barth-Maron, A., et al. (2021). Information flow, cell types and stereotypy in a full olfactory connectome. *Elife* 10, e66018.
75. Pires, P.M., Abbott, L.F., and Maimon, G. (2024). Converting an allocentric goal into an egocentric steering signal. *Nature* 626, 808–818.
76. Westeinde, E.A., Kellogg, E., Dawson, P.M., Lu, J., Hamburg, L., Midler, B., Druckmann, S., and Wilson, R.I. (2024). Transforming a head direction signal into a goal-oriented steering command. *Nature* 626, 819–826.
77. Hulse, B.K., Haberkern, H., Franconville, R., Turner-Evans, D., Take-mura, S.-Y., Wolff, T., Noorman, M., Dreher, M., Dan, C., Parekh, R., et al. (2021). A connectome of the *drosophila* central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *Elife* 10, e66039.
78. Mao, Z., and Davis, R.L. (2009). Eight different types of dopaminergic neurons innervate the *drosophila* mushroom body neuropil: anatomical and physiological heterogeneity. *Front. Neural Circuits* 3, 5.
79. Aso, Y., Hattori, D., Yu, Y., Johnston, R.M., Iyer, N.A., Ngo, T.T.B., Dionne, H., Abbott, L.F., Axel, R., Tanimoto, H., and Rubin, G.M. (2014). The neuronal architecture of the mushroom body provides a logic for associative learning. *Elife* 3, e04577.
80. Liu, T.X., Davoudian, P.A., Lizbinski, K.M., and Jeanne, J.M. (2022). Connectomic features underlying diverse synaptic connection strengths and subcellular computation. *Curr. Biol.* 32, 559–569.e5.
81. Ramon Cajal, S. (1911). Histologie du système nerveux de l'homme et des vertébrés. grand sympathique2 (Paris Maloine), pp. 891–942.
82. Dolan, M.-J., Belliart-Guérin, G., Bates, A.S., Frechter, S., Lampin-Saint-Amaux, A., Aso, Y., Roberts, R.J.V., Schlegel, P., Wong, A., Hammad, A., et al. (2018). Communication from learned to innate olfactory processing centers is required for memory retrieval in *Drosophila*. *Neuron* 100, 651–668.
83. Rolls, M.M. (2011). Neuronal polarity in *drosophila*: sorting out axons and dendrites. *Dev. Neurobiol.* 71, 419–429.
84. Bossy, B., Ballivet, M., and Spierer, P. (1988). Conservation of neural nicotinic acetylcholine receptors from *drosophila* to vertebrate central nervous systems. *EMBO J.* 7, 611–618.
85. Schuster, R., Phannavong, B., Schröder, C., and Gundelfinger, E.D. (1993). Immunohistochemical localization of a ligand-binding and a structural subunit of nicotinic acetylcholine receptors in the central nervous system of *drosophila melanogaster*. *J. Comp. Neurol.* 335, 149–162.
86. Olsen, S.R., Bhandawat, V., and Wilson, R.I. (2010). Divisive normalization in olfactory population codes. *Neuron* 66, 287–299.
87. Olsen, S.R., and Wilson, R.I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature* 452, 956–960.
88. Wilson, R.I., and Laurent, G. (2005). Role of GABAergic inhibition in shaping odor-evoked spatiotemporal patterns in the *Drosophila* antennal lobe. *J. Neurosci.* 25, 9069–9079.
89. Liu, W.W., and Wilson, R.I. (2013). Glutamate is an inhibitory neurotransmitter in the *Drosophila* olfactory system. *Proc. Natl. Acad. Sci. USA* 110, 10294–10299.
90. Lu, J., Behbahani, A.H., Hamburg, L., Westeinde, E.A., Dawson, P.M., Lyu, C., Maimon, G., Dickinson, M.H., Druckmann, S., and Wilson, R.I. (2022). Transforming representations of movement from body- to world-centric space. *Nature* 601, 98–104.
91. McCarthy, E.V., Wu, Y., Decarvalho, T., Brandt, C., Cao, G., and Nitabach, M.N. (2011). Synchronized bilateral synaptic inputs to *drosophila melanogaster* neuropeptidergic rest/arousal neurons. *J. Neurosci.* 31, 8181–8193.
92. Molina-Obando, S., Vargas-Fique, J.F., Henning, M., Gür, B., Schladt, T.M., Akhtar, J., Berger, T.K., and Silies, M. (2019). ON selectivity in the *drosophila* visual system is a multisynaptic process involving both glutamatergic and GABAergic inhibition. *Elife* 8, e49373.
93. Ichinose, T., Aso, Y., Yamagata, N., Abe, A., Rubin, G.M., and Tanimoto, H. (2015). Reward signal in a recurrent circuit drives appetitive long-term memory formation. *Elife* 4, e10719.
94. Jan, L.Y., and Jan, Y.N. (1976). L-glutamate as an excitatory transmitter at the *drosophila* larval neuromuscular junction. *J. Physiol.* 262, 215–236.
95. Zhao, X., Lenek, D., Dag, U., Dickson, B.J., and Keleman, K. (2018). Persistent activity in a recurrent circuit underlies courtship memory in *drosophila*. *Elife* 7.
96. Felsenberg, J., Jacob, P.F., Walker, T., Barnstedt, O., Edmondson-Stait, A.J., Pleijzier, M.W., Otto, N., Schlegel, P., Sharifi, N., Perisse, E., et al. (2018). Integration of parallel opposing memories underlies memory extinction. *Cell* 175, 709–722.e15.
97. Bates, A.S., Schlegel, P., Roberts, R.J.V., Drummond, N., Tamimi, I.F.M., Turnbull, R., Zhao, X., Marin, E.C., Popovici, P.D., Dhawan, S., et al. (2020b). Complete connectomic reconstruction of olfactory projection neurons in the fly brain. *Curr. Biol.* 30, 3183–3199.e6.
98. Wilson, R.I., and Mainen, Z.F. (2006). Early events in olfactory processing. *Annu. Rev. Neurosci.* 29, 163–201.
99. Cheong, H.S.J., Eichler, K., Stuerner, T., Asinof, S.K., Champion, A.S., Marin, E.C., Oram, T.B., Sumathipala, M., Venkatasubramanian, L., Namiki, S., et al. (2023). Transforming descending input into behavior: The organization of premotor circuits in the *drosophila* male adult nerve cord connectome. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.07.543976>.
100. Rubin, G.M., and Aso, Y. (2023). New genetic tools for mushroom body output neurons in *drosophila*.
101. Li, Y., Dharkar, P., Han, T.-H., Serpe, M., Lee, C.-H., and Mayer, M.L. (2016). Novel functional properties of *drosophila* CNS glutamate receptors. *Neuron* 92, 1036–1048.
102. Lee, Y.-J., Yang, C.-P., Miyares, R.L., Huang, Y.-F., He, Y., Ren, Q., Chen, H.-M., Kawase, T., Ito, M., Otsuna, H., et al. (2020). Conservation and divergence of related neuronal lineages in the *drosophila* central brain. *Elife* 9, e53518.
103. Jeffreys, H. (1998). The theory of probability (OUP Oxford).
104. Das, S., Sadanandappa, M.K., Dervan, A., Larkin, A., Lee, J.A., Sudhakaran, I.P., Priya, R., Heidari, R., Holohan, E.E., Pimentel, A., et al. (2011b). Plasticity of local GABAergic interneurons drives olfactory habituation. *Proc. Natl. Acad. Sci. USA* 108, E646–E654.
105. Chou, Y.-H., Spletter, M.L., Yaksi, E., Leong, J.C.S., Wilson, R.I., and Luo, L. (2010). Diversity and wiring variability of olfactory local interneurons in the *drosophila* antennal lobe. *Nat. Neurosci.* 13, 439–449.
106. Carlsson, M.A., Diesner, M., Schachtner, J., and Nässel, D.R. (2010). Multiple neuropeptides in the *drosophila* antennal lobe suggest complex modulatory circuits. *J. Comp. Neurol.* 518, 3359–3380.
107. Xu, C.S., Pang, S., Shtengel, G., Müller, A., Ritter, A.T., Hoffman, H.K., Take-mura, S.-Y., Lu, Z., Pasolli, H.A., Iyer, N., et al. (2021). An open-access volume electron microscopy atlas of whole cells and tissues. *Nature* 599, 147–151.
108. Zhang, A., Shailja, S., Borba, C., Miao, Y., Goebel, M., Ruschel, R., Ryan, K., Smith, W., and Manjunath, B.S. (2022). Automatic classification and neurotransmitter prediction of synapses in electron microscopy. *Biol. Imaging* 2, e6.
109. Colonnier, M. (1968). Synaptic patterns on different cell types in the different laminae of the cat visual cortex. an electron microscope study. *Brain Res.* 9, 268–287.

110. Gray, E.G. (1959). Axo-somatic and axo-dendritic synapses of the cerebral cortex: an electron microscope study. *J. Anat.* 93 (Pt 4), 420–433.
111. Siegler, M.V., Pankhaniya, R.R., and Jia, X.X. (2001). Pattern of expression of engrailed in relation to gamma-aminobutyric acid immunoreactivity in the central nervous system of the adult grasshopper. *J. Comp. Neurol.* 440, 85–96.
112. Witten, J.L., and Truman, J.W. (1998). Distribution of GABA-like immunoreactive neurons in insects suggests lineage homology. *J. Comp. Neurol.* 398, 515–528.
113. Lappalainen, J.K., Tschopp, F.D., Prakhy, S., McGill, M., Nern, A., Shinomiya, K., Takemura, S.-y., Gruntman, E., Macke, J.H., and Turaga, S.C. (2023). Connectome-constrained deep mechanistic networks predict neural responses across the fly visual system at single-neuron resolution. Preprint at bioRxiv. <https://doi.org/10.1101/2023.03.11.532232>.
114. Shiu, P.K., Sterne, G.R., Spiller, N., Franconville, R., Sandoval, A., Zhou, J., Simha, N., Kang, C.H., Yu, S., Kim, J.S., et al. (2023). A leaky integrate-and-fire computational model based on the connectome of the entire adult drosophila brain reveals insights into sensorimotor processing. Preprint at bioRxiv. <https://doi.org/10.1101/2023.05.02.539144>.
115. Groschner, L.N., Malis, J.G., Zuidinga, B., and Borst, A. (2022). A biophysical account of multiplication by a single neuron. *Nature* 603, 119–123.
116. Miyashita, T., Murakami, K., Kikuchi, E., Ofusa, K., Mikami, K., Endo, K., Miyaji, T., Moriyama, S., Konno, K., Muratani, H., et al. (2023). Glia transmit negative valence information during aversive learning in *Drosophila*. *Science* 382, eadf7429.
117. Ni, J.D., Gurav, A.S., Liu, W., Ogunmowo, T.H., Hackbart, H., Elsheikh, A., Verdegaaal, A.A., and Montell, C. (2019). Differential regulation of the drosophila sleep homeostat by circadian and arousal inputs. *Elife* 8, e40487.
118. Park, A., Croset, V., Otto, N., Agarwal, D., Treiber, C.D., Meschi, E., Sims, D., and Waddell, S. (2022). Gliotransmission of d-serine promotes thirst-directed behaviors in drosophila. *Curr. Biol.* 32, 3952–3970.e8.
119. Sanfilippo, P., Kim, A.J., Bhukel, A., Yoo, J., Mirshahidi, P.S., Pandey, V., Bevir, H., Yuen, A., Mirshahidi, P.S., Guo, P., et al. (2023). Mapping of multiple neurotransmitter receptor subtypes and distinct protein complexes to the connectome. Preprint at bioRxiv. <https://doi.org/10.1101/2023.10.02.560011>.
120. Dorkenwald, S., Schneider-Mizell, C.M., Brittain, D., Halageri, A., Jordan, C., Kemnitz, N., Castro, M.A., Silversmith, W., Maitin-Shephard, J., Troidl, J., et al. (2023). CAVE: Connectome annotation versioning engine. Preprint at bioRxiv. <https://doi.org/10.1101/2023.07.26.550598>.
121. Court, R., Costa, M., Pilgrim, C., Millburn, G., Holmes, A., McLachlan, A., Larkin, A., Matentzoglu, N., Kir, H., Parkinson, H., et al. (2023). Virtual fly Brain-An interactive atlas of the drosophila nervous system. *Front. Physiol.* 14, 1076533.
122. Jiang, H., Lkhagva, A., Daubnerová, I., Chae, H.-S., Šimo, L., Jung, S.-H., Yoon, Y.-K., Lee, N.-R., Seong, J.Y., Žitňan, D., et al. (2013). Natalisin, a tachykinin-like signaling system, regulates sexual activity and fecundity in insects. *Proc. Natl. Acad. Sci. USA* 110, E3526–E3534.
123. Scheunemann, L., Plaçais, P.Y., Dromard, Y., Schwärzel, M., and Preat, T. (2018). Dunce phosphodiesterase acts as a checkpoint for drosophila Long-Term memory in a pair of serotonergic neurons. *Neuron* 98, 350–365.e5.
124. Kondo, S., Takahashi, T., Yamagata, N., Imanishi, Y., Katow, H., Hiramatsu, S., Lynn, K., Abe, A., Kumaraswamy, A., and Tanimoto, H. (2020). Neurochemical organization of the drosophila brain visualized by endogenously tagged neurotransmitter receptors. *Cell Rep.* 30, 284–297.e5.
125. Lillvis, J.L., Otsuna, H., Ding, X., Pisarev, I., Kawase, T., Colonell, J., Rokicki, K., Goina, C., Gao, R., Hu, A., et al. (2022). Rapid reconstruction of neural circuits using tissue expansion and light sheet microscopy. *Elife* 11, e81248.
126. Rickgauer, J.P., Grigorieff, N., and Denk, W. (2017). Single-protein detection in crowded molecular environments in cryo-EM images. *Elife* 6, e25648.
127. Matheson, A.M.M., Lanz, A.J., Medina, A.M., Licata, A.M., Currier, T.A., Syed, M.H., and Nagel, K.I. (2022). A neural circuit for wind-guided olfactory navigation. *Nat. Commun.* 13, 4613.
128. Schlegel, P., Gokaslan, A., and Kazimiers, T. (2022). navis-org/skeletor: Version 1.2.3.
129. Sayin, S., De Backer, J.-F., Siju, K.P., Wosniack, M.E., Lewis, L.P., Frisch, L.-M., Gansen, B., Schlegel, P., Edmondson-Stait, A., Sharifi, N., et al. (2019). A neural circuit arbitrates between persistence and withdrawal in hungry *Drosophila*. *Neuron* 104, 544–558.e6.
130. R Core Team (2021). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
131. Xie, Y. (2015). Dynamic Documents with R and knitr (Chapman and Hall).
132. Wickham, H. (2011). Ggplot2. *WIREs Computational Stats.* 3, 180–185.
133. Bates, A.S., Manton, J.D., Jagannathan, S.R., Costa, M., Schlegel, P., Rohlfing, T., and Jefferis, G.S. (2020a). The natverse, a versatile toolbox for combining and analysing neuroanatomical data. *Elife* 9, e53350.
134. Veveryta, L., and Allan, D.W. (2013). Subtype-specific neuronal remodeling during *Drosophila* metamorphosis. *Fly* 7, 78–86.
135. Plaza, S.M., Clements, J., Dolafi, T., Umayam, L., Neubarth, N.N., Scheffer, L.K., and Berg, S. (2022). neuprint: An open access tool for EM connectomics. *Front. Neuroinform.* 16, 896292.
136. Knott, G., Marchman, H., Wall, D., and Lich, B. (2008). Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *J. Neurosci.* 28, 2959–2964.
137. Li, P.H., Lindsey, L.F., Januszewski, M., Zheng, Z., Bates, A.S., Taisz, I., Tyka, M., Nichols, M., Li, F., Perlman, E., et al. (2020b). Automated reconstruction of a Serial-Section EM drosophila brain with Flood-Filling networks and local realignment. Preprint at bioRxiv. <https://doi.org/10.1101/605634>.
138. Busch, S., Selcho, M., Ito, K., and Tanimoto, H. (2009). A map of octopaminergic neurons in the *Drosophila* brain. *J. Comp. Neurol.* 513, 643–667.
139. Ito, M., Masuda, N., Shinomiya, K., Endo, K., and Ito, K. (2013). Systematic analysis of neural projections reveals clonal composition of the *Drosophila* brain. *Curr. Biol.* 23, 644–655.
140. Bräcker, L.B., Siju, K.P., Varela, N., Aso, Y., Zhang, M., Hein, I., Vasconcelos, M.L., and Grunwald Kadow, I.C. (2013). Essential role of the mushroom body in context-dependent CO₂ avoidance in *Drosophila*. *Curr. Biol.* 23, 1228–1234.
141. Dacks, A.M., Christensen, T.A., and Hildebrand, J.G. (2006). Phylogeny of a serotonin-immunoreactive neuron in the primary olfactory center of the insect brain. *J. Comp. Neurol.* 498, 727–746.
142. Dolan, M.-J., Frechter, S., Bates, A.S., Dan, C., Huoviala, P., Roberts, R.J., Schlegel, P., Dhawan, S., Tabano, R., Dionne, H., et al. (2019). Neurogenetic dissection of the drosophila lateral horn reveals major outputs, diverse behavioural functions, and interactions with the mushroom body. *Elife* 8, e43079.
143. Frenkel, L., Muraro, N.I., Beltrán González, A.N., Marcora, M.S., Berribó, G., Hermann-Luibl, C., Romero, J.I., Helfrich-Förster, C., Castaño, E.M., Marino-Busjle, C., et al. (2017). Organization of circadian behavior relies on glycinergic transmission. *Cell Rep.* 19, 72–85.
144. Guo, F., Yu, J., Jung, H.J., Abruzzi, K.C., Luo, W., Griffith, L.C., and Rosbash, M. (2016). Circadian neuron feedback controls the drosophila sleep–activity profile. *Nature* 536, 292–297.
145. Krashes, M.J., DasGupta, S., Vreede, A., White, B., Armstrong, J.D., and Waddell, S. (2009). A neural circuit mechanism integrating motivational state with memory expression in drosophila. *Cell* 139, 416–427.
146. Lai, S.-L., Awasaki, T., Ito, K., and Lee, T. (2008). Clonal analysis of *Drosophila* antennal lobe neurons: diverse neuronal architectures in the lateral neuroblast lineage. *Development* 135, 2883–2893.

147. Liu, C., Meng, Z., Wiggin, T.D., Yu, J., Reed, M.L., Guo, F., Zhang, Y., Rosbash, M., and Griffith, L.C. (2019). A Serotonin-Modulated circuit controls sleep architecture to regulate cognitive function independent of total sleep in *drosophila*. *Curr. Biol.* **29**, 3635–3646.e5.
148. Liu, X., Buchanan, M.E., Han, K.-A., and Davis, R.L. (2009). The GABA_A receptor RDL suppresses the conditioned stimulus pathway for olfactory learning. *J. Neurosci.* **29**, 1573–1579.
149. Nässel, D.R., and Elekes, K. (1992). Aminergic neurons in the brain of blowflies and *drosophila*: dopamine- and tyrosine hydroxylase-immuno-reactive neurons and their relationship with putative histaminergic neurons. *Cell Tissue Res.* **267**, 147–167.
150. Niens, J., Reh, F., Çoban, B., Cichewicz, K., Eckardt, J., Liu, Y.-T., Hirsh, J., and Riemensperger, T.D. (2017). Dopamine modulates serotonin innervation in the *drosophila* brain. *Front. Syst. Neurosci.* **11**, 76.
151. Okada, R., Awasaki, T., and Ito, K. (2009). Gamma-aminobutyric acid (GABA)-mediated neural connections in the *Drosophila* antennal lobe. *J. Comp. Neurol.* **514**, 74–91.
152. Palavicino-Maggio, C.B., Chan, Y.-B., McKellar, C., and Kravitz, E.A. (2019). A small number of cholinergic neurons mediate hyperaggression in female *Drosophila*. *Proc. Natl. Acad. Sci. USA* **116**, 17029–17038.
153. Senapati, B., Tsao, C.-H., Juan, Y.-A., Chiu, T.-H., Wu, C.-L., Waddell, S., and Lin, S. (2019). A neural mechanism for deprivation state-specific expression of relevant memories in *drosophila*. *Nat. Neurosci.* **22**, 2029–2039.
154. Shafer, O.T., Helfrich-Förster, C., Renn, S.C.P., and Taghert, P.H. (2006). Reevaluation of *drosophila melanogaster*'s neuronal circadian pacemakers reveals new neuronal classes. *J. Comp. Neurol.* **498**, 180–193.
155. Shinomiya, K., Takemura, S.-Y., Rivlin, P.K., Plaza, S.M., Scheffer, L.K., and Meinertzhagen, I.A. (2015). A common evolutionary origin for the ON- and OFF-edge motion detection pathways of the *Drosophila* visual system. *Front. Neural Circuits* **9**, 33.
156. Tanaka, N.K., Endo, K., and Ito, K. (2012). Organization of antennal lobe-associated neurons in adult *Drosophila melanogaster* brain. *J. Comp. Neurol.* **520**, 4067–4130.
157. Turner-Evans, D.B., Jensen, K.T., Ali, S., Paterson, T., Sheridan, A., Ray, R.P., Wolff, T., Lauritzen, J.S., Rubin, G.M., Bock, D.D., and Jayaraman, V. (2020). The neuroanatomical ultrastructure and function of a biological ring attractor. *Neuron* **108**, 145–163.e10.
158. Wolff, T., Iyer, N.A., and Rubin, G.M. (2015). Neuroarchitecture and neuroanatomy of the *drosophila* central complex: A GAL4-based dissection of protocerebral bridge neurons and circuits. *J. Comp. Neurol.* **523**, 997–1037.
159. Frechter, S., Bates, A.S., Tootoonian, S., Dolan, M.-J., Manton, J.D., Jamasb, A.R., Kohl, J., Bock, D., and Jefferis, G.S. (2019). Functional and anatomical specificity in a higher olfactory centre. *Elife* **8**, e44590.
160. Huoviala, P., Dolan, M.-J., Love, F.M., Frechter, S., Roberts, R.J.V., Mirevica, Z., Schlegel, P., Bates, A.S., Aso, Y., Rodrigues, T., et al. (2018). Neural circuit basis of aversive odour processing in *Drosophila* from sensory input to descending output. Preprint at bioRxiv. <https://doi.org/10.1101/394403>.
161. Marin, E.C., Roberts, R.J.V., Büld, L., Theiss, M., Pleijzier, M.W., Sarkisian, T., Laursen, W.J., Turnbull, R., Schlegel, P., Bates, A.S., et al. (2020). Connectomics analysis reveals first, second, and third order thermosensory and hygrosensory neurons in the adult *Drosophila* brain. *Curr. Biol.* **30**, 3167–3182.e4.
162. Saalfeld, S., Cardona, A., Hartenstein, V., and Tomančák, P. (2009). Catmaid: collaborative annotation toolkit for massive amounts of image data. *Bioinformatics* **25**, 1984–1986.
163. Sherer, L.M., Catudio Garrett, E., Morgan, H.R., Brewer, E.D., Sirrs, L.A., Shearin, H.K., Williams, J.L., McCabe, B.D., Stowers, R.S., and Cetrel, S.J. (2020). Octopamine neuron dependent aggression requires dVGLUT from dual-transmitting neurons. *PLoS Genet.* **16**, e1008609.
164. Crocker, A., Guan, X.J., Murphy, C.T., and Murthy, M. (2016). Cell-Type-Specific transcriptome analysis in the *drosophila* mushroom body reveals Memory-Related changes in gene expression. *Cell Rep.* **15**, 1580–1596.
165. Prokop, A., and Meinertzhagen, I.A. (2006). Development and structure of synaptic contacts in *Drosophila*. *Semin. Cell Dev. Biol.* **17**, 20–30.
166. Kubrak, O.I., Lushchak, O.V., Zandawala, M., and Nässel, D.R. (2016). Systemic corazonin signalling modulates stress responses and metabolism in *drosophila*. *Open Biol.* **6**, 160152.
167. Nichols, R., Schneuwly, S.A., and Dixon, J.E. (1988). Identification and characterization of a *drosophila* homologue to the vertebrate neuropeptide cholecystokinin. *J. Biol. Chem.* **263**, 12167–12170.
168. Söderberg, J.A.E., Carlsson, M.A., and Nässel, D.R. (2012). Insulin-Producing cells in the *drosophila* brain also express Satiety-Inducing Cholecystokinin-Like peptide, drosulfakinin. *Front. Endocrinol.* **3**, 109.
169. Cao, C., and Brown, M.R. (2001). Localization of an insulin-like peptide in brains of two flies. *Cell Tissue Res.* **304**, 317–321.
170. Nässel, D.R., and Vanden Broeck, J. (2016). Insulin/IGF signaling in *drosophila* and other insects: factors that regulate production, release and post-release action of the insulin-like peptides. *Cell. Mol. Life Sci.* **73**, 271–290.
171. Terhzaz, S., Rosay, P., Goodwin, S.F., and Veenstra, J.A. (2007). The neuropeptide SiFamide modulates sexual behavior in *drosophila*. *Biochem. Biophys. Res. Commun.* **352**, 305–310.
172. Kim, S., Wallace, M.L., El-Rifai, M., Knudsen, A.R., and Sabatini, B.L. (2022). Co-packaging of opposing neurotransmitters in individual synaptic vesicles in the central nervous system. *Neuron* **110**, 1371–1384.e7.
173. Johard, H.A.D., Enell, L.E., Gustafsson, E., Trifilieff, P., Veenstra, J.A., and Nässel, D.R. (2008). Intrinsic neurons of *drosophila* mushroom bodies express short neuropeptide f: relations to extrinsic neurons expressing different neurotransmitters. *J. Comp. Neurol.* **507**, 1479–1496.
174. Sukumar, S.K., Antonydhason, V., Molander, L., Sandakly, J., Umapathy, G., Mendoza-Garcia, P., Masudi, T., Schlosser, A., Nässel, D.R., Wegeiner, C., et al. (2023). The alk receptor tyrosine kinase regulates sparkly, a novel activity regulating neurosecretory protein in the *drosophila* CNS. Preprint at bioRxiv. <https://doi.org/10.1101/2023.06.02.543395>.
175. Kahsai, L., and Winther, A.M.E. (2011). Chemical neuroanatomy of the *drosophila* central complex: distribution of multiple neuropeptides in relation to neurotransmitters. *J. Comp. Neurol.* **519**, 290–315.
176. Hartenstein, V., Cruz, L., Lovick, J.K., and Guo, M. (2017). Developmental analysis of the dopamine-containing neurons of the *drosophila* brain. *J. Comp. Neurol.* **525**, 363–379.
177. Kahsai, L., Carlsson, M.A., Winther, A.M.E., and Nässel, D.R. (2012). Distribution of metabotropic receptors of serotonin, dopamine, GABA, glutamate, and short neuropeptide F in the central complex of *drosophila*. *Neuroscience* **208**, 11–26.
178. Takemura, S.-y., Xu, C.S., Lu, Z., Rivlin, P.K., Parag, T., Olbris, D.J., Plaza, S., Zhao, T., Katz, W.T., Umayam, L., et al. (2015). Synaptic circuits and their variations within different columns in the visual system of *drosophila*. *Proc. Natl. Acad. Sci. USA* **112**, 13711–13716.
179. Raghu, S.V., and Borst, A. (2011). Candidate glutamatergic neurons in the visual system of *drosophila*. *PLoS One* **6**, e19472.
180. Alekseyenko, O.V., Lee, C., and Kravitz, E.A. (2010). Targeted manipulation of serotonergic neurotransmission affects the escalation of aggression in adult male *drosophila melanogaster*. *PLoS One* **5**, e10806.
181. Cetrel, S.J., Leung, A., Lin, C.-Y., Perez, P., Chiang, A.-S., and Kravitz, E.A. (2010). Octopamine neuromodulatory effects on a social behavior decision-making network in *drosophila* males. *PLoS One* **5**, e13248.
182. Yang, J.S., Awasaki, T., Yu, H.-H., He, Y., Ding, P., Kao, J.-C., and Lee, T. (2013). Diverse neuronal lineages make stereotyped contributions to the *drosophila* locomotor control center, the central complex. *J. Comp. Neurol.* **521**, 2645–Spc1.

183. Lyu, C., Abbott, L.F., and Maimon, G. (2022). Building an allocentric travelling direction signal via vector computation. *Nature* **601**, 92–97.
184. Nässel, D.R. (2021). Leucokinin and associated neuropeptides regulate multiple aspects of physiology and behavior in drosophila. *Int. J. Mol. Sci.* **22**, 1940.
185. Troup, M., Yap, M.H., Rohrscheib, C., Grabowska, M.J., Ertekin, D., Randardiya, R., Kottler, B., Larkin, A., Munro, K., Shaw, P.J., and van Swinden, B. (2018). Acute control of the sleep switch in drosophila reveals a role for gap junctions in regulating behavioral responsiveness. *Elife* **7**, e37105.
186. Musso, P.-Y., Junca, P., and Gordon, M.D. (2021). A neural circuit linking two sugar sensors regulates satiety-dependent fructose drive in drosophila. *Sci. Adv.* **7**, eabj0186.
187. Hamid, A., Gattuso, H., Caglar, A.N., Pillai, M., Steele, T., Gonzalez, A., Nagel, K., and Syed, M.H. (2023). The RNA-binding protein, imp specifies olfactory navigation circuitry and behavior in drosophila. Preprint at bioRxiv. <https://doi.org/10.1101/2023.05.26.542522>.
188. De, J., Wu, M., Lambatan, V., Hua, Y., and Joiner, W.J. (2023). Re-examining the role of the dorsal fan-shaped body in promoting sleep in drosophila. *Curr. Biol.* **33**, 3660–3668.e4.
189. Huang, J., Zhang, W., Qiao, W., Hu, A., and Wang, Z. (2010). Functional connectivity and selective odor responses of excitatory local interneurons in drosophila antennal lobe. *Neuron* **67**, 1021–1033.
190. Yaksi, E., and Wilson, R.I. (2010). Electrical coupling between olfactory glomeruli. *Neuron* **67**, 1034–1047.
191. Berck, M.E., Khandelwal, A., Claus, L., Hernandez-Nunez, L., Si, G., Tabone, C.J., Li, F., Truman, J.W., Fetter, R.D., Louis, M., et al. (2016). The wiring diagram of a glomerular olfactory system. *Elife* **5**, e14859.
192. Chou, Y.-H., Yang, C.-J., Huang, H.-W., Liou, N.-F., Panganiban, M.R., Luginbuhl, D., Yin, Y., Taisz, I., Liang, L., Jefferis, G.S.X.E., and Luo, L. (2022). Mating-driven variability in olfactory local interneuron wiring. *Sci. Adv.* **8**, eabm7723.
193. Kumar, A., Bello, B., and Reichert, H. (2009). Lineage-specific cell death in postembryonic brain development of drosophila. *Development* **136**, 3433–3442.
194. Sen, S. (2019). Neurotransmitter identity: A question of lineage. *Elife* **8**, e47162.
195. Truman, J.W., Moats, W., Altman, J., Marin, E.C., and Williams, D.W. (2010). Role of notch signaling in establishing the hemilineages of secondary neurons in drosophila melanogaster. *Development* **137**, 53–61.
196. Jiang, Y., and Reichert, H. (2012). Programmed cell death in type II neuroblast lineages is required for central complex development in the drosophila brain. *Neural Dev.* **7**, 3.
197. Marin, E.C., Watts, R.J., Tanaka, N.K., Ito, K., and Luo, L. (2005). Developmentally programmed remodeling of the *Drosophila* olfactory circuit. *Development* **132**, 725–737.
198. Sullivan, L.F., Warren, T.L., and Doe, C.Q. (2019). Temporal identity establishes columnar neuron morphology, connectivity, and function in a drosophila navigation circuit. *Elife* **8**, e43482.
199. Lovick, J.K., Ngo, K.T., Omoto, J.J., Wong, D.C., Nguyen, J.D., and Hartenstein, V. (2013). Postembryonic lineages of the *Drosophila* brain: I. development of the lineage-associated fiber tracts. *Dev. Biol.* **384**, 228–257.
200. Wong, D.C., Lovick, J.K., Ngo, K.T., Borisuthirattana, W., Omoto, J.J., and Hartenstein, V. (2013). Postembryonic lineages of the *Drosophila* brain: II. identification of lineage projection patterns based on MARCM clones. *Dev. Biol.* **384**, 258–289.
201. Yu, H.-H., Awasaki, T., Schroeder, M.D., Long, F., Yang, J.S., He, Y., Ding, P., Kao, J.-C., Wu, G.Y.-Y., Peng, H., et al. (2013). Clonal development and organization of the adult *Drosophila* central brain. *Curr. Biol.* **23**, 633–643.
202. Kandimalla, P., Omoto, J.J., Hong, E.J., and Hartenstein, V. (2023). Lineages to circuits: the developmental and evolutionary architecture of information channels into the central complex. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **209**, 679–720.
203. Cardona, A., Saalfeld, S., Schindelin, J., Arganda-Carreras, I., Preibisch, S., Longair, M., Tomancak, P., Hartenstein, V., and Douglas, R.J. (2012). TrakEM2 software for neural circuit reconstruction. *PLoS One* **7**, e38011.
204. Rohlfing, T., and Maurer, C.R., Jr. (2003). Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. *IEEE Trans. Inf. Technol. Biomed.* **7**, 16–25.
205. Bello, B.C., Izergina, N., Caussinus, E., and Reichert, H. (2008). Amplification of neural stem cell proliferation by intermediate progenitor cells in drosophila brain development. *Neural Dev.* **3**, 5.
206. Boone, J.Q., and Doe, C.Q. (2008). Identification of drosophila type II neuroblast lineages containing transit amplifying ganglion mother cells. *Dev. Neurobiol.* **68**, 1185–1195.
207. Bowman, S.K., Rolland, V., Betschinger, J., Kinsey, K.A., Emery, G., and Knoblich, J.A. (2008). The tumor suppressors brat and numb regulate transit-amplifying neuroblast lineages in drosophila. *Dev. Cell* **14**, 535–546.
208. Baker, C.A., McKellar, C., Pang, R., Nern, A., Dorkenwald, S., Pacheco, D.A., Eckstein, N., Funke, J., Dickson, B.J., and Murthy, M. (2022). Neural network organization for courtship-song feature detection in drosophila. *Curr. Biol.* **32**, 3317–3333.e7.
209. Eichler, K., Hampel, S., Alejandro-García, A., Calle-Schuler, S.A., Santana-Cruz, A., Kmecova, L., Blagburn, J.M., Hooper, E.D., and Seeds, A.M. (2023). Somatotopic organization among parallel sensory pathways that promote a grooming sequence in drosophila. Preprint at bioRxiv. <https://doi.org/10.1101/2023.02.11.528119>.
210. Engert, S., Sterne, G.R., Bock, D.D., and Scott, K. (2022). Drosophila gustatory projections are segregated by taste modality and connectivity. *Elife* **11**, e78110.
211. Li, F., Lindsey, J.W., Marin, E.C., Otto, N., Dreher, M., Dempsey, G., Stark, I., Bates, A.S., Pleijzier, M.W., Schlegel, P., et al. (2020a). The connectome of the adult drosophila mushroom body provides insights into function. *Elife* **9**, e62576.
212. 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. (2011). Three-dimensional reconstruction of brain-wide wiring networks in drosophila at single-cell resolution. *Curr. Biol.* **21**, 1–11.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
FAFB ssTEM image data	Zheng et al. ⁶	https://temca2data.org/
FAFB-FlyWire	Dorkenwald et al. ¹²⁰	https://ngl.flywire.ai/
FAFB-FlyWire connectome	Dorkenwald et al., ¹⁷ Matsliah et al. ¹²⁷	https://codex.flywire.ai/
FAFB-FlyWire annotations	Schlegel et al. ¹²⁸	https://doi.org/10.5281/zenodo.8077334
HemiBrain FIBSEM image data	Scheffer et al. ¹²⁹	https://neuprint.janelia.org/?dataset=hemibrain:v1.2.1
HemiBrain connectome	Scheffer et al. ¹²⁹	https://neuprint.janelia.org/?dataset=hemibrain:v1.2.1
FAFB-FlyWire synapse-level predictions	This paper	<code>synister fw mat571 t11 synapses.feather</code>
HemiBrain synapse-level predictions	This paper	<code>hemibrain-v1.2-tbar-neurotransmitters.feather.bz2</code>
FAFB-FlyWire neuron-level predictions	This paper	https://codex.flywire.ai
HemiBrain neuron-level predictions	This paper	https://neuprint.janelia.org/?dataset=hemibrain:v1.2.1
All HemiBrain and FlyWire synapse and neuron-level predictions	This paper	https://doi.org/10.5281/zenodo.10593546
Software and algorithms		
R	R Core Team ¹³⁰	RRID SCR 001905
<i>knitr</i> (R)	Xie ¹³¹	RRID SCR 018533
<i>ggplot2</i> (R)	Wickham ¹³²	RRID SCR 014601
<i>tidyverse</i> (R)	Wickham et al. ⁹⁸	RRID SCR 019186
<i>rgl</i> (R)	Wickham et al. ⁹⁸	https://cran.r-project.org/web/packages/rgl/index.html
<i>natverse</i> (R)	Bates et al. ¹³³	https://natverse.org/natverse/
<i>neupintr</i> (R)	Bates et al. ¹³³	https://natverse.org/neupintr/
<i>hemibrainr</i> (R)	Schlegel et al. ¹²³	https://github.com/flyconnectome/hemibrainr
<i>catmaid</i> (R)	Bates et al. ¹³³	https://natverse.org/rcatmaid/
<i>fafbseg</i> (R)	Schlegel et al. ¹²⁸	https://natverse.org/fafbseg
<i>nat.jrcbrains</i> (R)	Bates et al. ¹³³	https://natverse.org/nat.jrcbrains/
<i>nat.nblast</i> (R)	Costa et al. ¹²¹	https://natverse.org/nat.nblast/
python	Van Rossum and Drake ¹³⁴	RRID SCR 008394
<i>navis</i> (python)	Schlegel et al. ⁵¹	https://github.com/navis-org/navis
<i>navis-flybrains</i> (python)	Schlegel et al. ⁵¹	https://github.com/navis-org/navis-flybrains
<i>pymaid</i> (python)	Schlegel et al. ⁵¹	https://github.com/navis-org/pymaid
<i>fafbseg-py</i> (python)	Schlegel et al. ⁵¹	https://github.com/navis-org/fafbseg-py
<i>skeletor</i> (python)	Schlegel et al. ⁷⁴	https://github.com/navis-org/skeletor
<i>neuprint-python</i> (python)	Plaza et al. ¹³⁵	https://github.com/connectome-neuprint/neuprint-python
<i>synister, transmitter prediction</i> (python)	This paper	https://doi.org/10.5281/zenodo.10593546

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jan Funke (funkej@janelia.hhmi.org).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All of our data is also available through Zenodo, <https://doi.org/10.5281/zenodo.10593546>, and will be available through the Virtual Fly Brain project.¹²¹ Our synapse-level transmitter prediction and neuron-level transmitter predictions are also hosted by extant connectome annotation and browsing services for the HemiBrain () and FAFB-FlyWire (<https://codex.flywire.ai/>) datasets.^{120,135} As supplemental data for this paper, we have provided: (1) the studies we have used to generate our ground-truth data (Data S1), (2) identifiers for the neurons we used for our ground truth data (Data S2), (3) our neuron-level transmitter predictions for each complete neuron in the HemiBrain dataset, hemibrain:v1.2.1 (Data S3), and the FAFB-FlyWire dataset, 630 materialization (used in this paper but with links to the newer 783 materialization)(Data S4), (5) images of all discerned brain hemilineages for FlyWire⁵⁰ and HemiBrain, with neurons colored by their predicted transmitter (Data S5), as well as (6) summary data on hemilineage transmitter usage (Data S6), and (7) an transmitter expression summary by brain cell type (Data S7). We have provided a section in these Methods of advice to users on how to use our data.
- Our transmitter classification network and associated training and prediction code are available in our Synister repository (<https://github.com/funkelab/synister>) which also contains instructions on how to access the FAFB-FlyWire and HemiBrain predictions, and through Zenodo, <https://doi.org/10.5281/zenodo.10593546>.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The HemiBrain connectome⁵¹ is a partial female fly brain that has been semi-automatically reconstructed using flood-filling networks⁸ from data acquired by focused ion-beam milling scanning EM (FIBSEM).¹³⁶ Presynapses (i.e., sites with T-bars) and postsynapses were identified completely automatically. The data can be accessed via the NeuPrint connectome analysis service.¹³⁵ Automatically detected mitochondria counts in the HemiBrain dataset were pulled from neuPrint¹³⁵ (<https://connectome-neuprint.github.io/neuprint-python/docs/mitocriteria.html>).

The FAFB ssTEM image data comprises an entire female fly brain. Two auto-segmentations of the data exist,^{17,137} we used the FlyWire segmentation and automatically detected synapses¹² for our biological analyses. However, to build our ground truth data we used high-fidelity manually reconstructed neurons and synapses that were annotated using CATMAID.⁷³ The manually placed FAFB synapses in this study were identified in Catmaid.^{6,73} Catmaid is a collaborative environment in which 27 labs have participated to build connectomes for specific circuits. For these annotations, we thank Ruairi Roberts, Fiona Love, Lisa Marin, Amelia Edmondson-Stait, Xincheng Zhao, Jawaid Ali, Johann Schor, Imaan Tamimi, Arian Jamasb, Marisa Dreher, Markus Pleijzier, Robert Turnbull, Nadiya Sharifi, Steven Calle, Andrew Dacks, Konrad Heinz, Kimberly Meechan, Aidan Smith, Najla Masoodpanah, Serene Dhawan, Peter Gibb, Corey Fisher, Claire Peterson, Jason Polksky, Tansy Yang, Katharina Eichler, Joseph Hsu, Irene Varela, Lucia Kmecova, Istvan Taisz, Jacob Ratliff, Kaylynn Coates, Anna Li, Marta Costa, Tyler Paterson, Claire Managan, Adam Heath, Katie Stevens, Jack McCarty, Nora Forknall, Laurin Buel, Neha Rampally, Zane Mitrevica, Kelli Fairbanks, Stanley Tran, Shada Alghailani, Quinn Vanderbeck, Lauren Warner, Henrique Ludwig, Jeremy Johnson and Levi Helmick, each of whom has contributed over 1,000 synapses. We in particular thank the Wellcome Trust UK and US Drosophila Connectomics, Jefferis, Janelia Connectome Annotation Team, Bock, Prent, Wilson, Dacks, Hampel and Seeds groups for sharing their published and unpublished work in the Catmaid dataset. Development and administration of the Catmaid tracing environment and analysis tools were funded in part by National Institutes of Health BRAIN Initiative grant 1RF1MH120679-01 to Davi Bock and Gregory Jefferis, with software development effort and administrative support provided by Tom Kazimiers (Kazmos GmbH) and Eric Perlman (Yikes LLC). We thank Marissa Sorek for assistance with community management and Ran Lu, Thomas Macrina, Kisuk Lee, J. Alexander Bae, Shang Mu, Barak Nehoran, Eric Mitchell, Sergiy Popovych, Jongpeng Wu, Zhen Jia, Manuel Castro, Nico Kemnitz, Dodam Ih for alignment and segmentation of the FAFB EM volume and registration to the original FAFB EM dataset. Both datasets are from female flies.

METHOD DETAILS

Detail on assembling ground truth data for *D. melanogaster*

We created a list of 356 neuronal cell types from 21 studies (see Data S1) and identified as many of them as we could in the FAFB-Catmaid and HemiBrain datasets (Figure 1A).^{39,40,51,56,57,77–79,88,90,138–158} FAFB neurons were found in extant FAFB-Catmaid data

(<https://neuropil.janelia.org/tracing/fafb>) based on the FAFB-Catmaid literature,^{6,82,96,97,129,142,157,159–161} by using NBLAST⁴⁷ or by manually reconstructing them in Catmaid from scratch using previously reported methodologies to approximate candidate locations for known cell types based on their hemilineage identity.^{6,73,97,162} The transmitter expression evidence for these cell types in the literature came from detecting RNA expression related to transmitter usage (minority) or immunohistochemistry (majority). For these studies, neurons were picked (transcriptomics) or stained (immunohistochemistry) with guidance from green fluorescent protein expression in a GAL4/split-GAL4 lines.⁴⁶ These genetic driver lines target transgene expression to a small constellation of discriminable cell types, or even individual cell types or neurons, which could be linked to specific homologs in the EM datasets. There are only a small number of our identified connectomic cell types for which we found co-transmission of small-molecule transmitters occurred. Mi15 (acetylcholine, serotonin), Mi9 (glutamate, GABA), Mi4 (GABA, nitric oxide), C3 (GABA, nitric oxide), Mi15 (acetylcholine, dopamine, nitric oxide)³⁹ most octopaminergic neurons (octopamine, glutamate)^{21,163} and a subset of PPL1 neurons (dopamine, serotonin).^{78,150} We excluded these cell types from our training data, except for the known octopaminergic neurons because 70% of them are thought to co-express glutamate. Because each neuron in our ground truth data must have a single transmitter label, these were given “octopamine” alone. Co-transmission of a fast-acting transmitter and a neuropeptide was more common, e.g., Kenyon cells (at least acetylcholine, sparkly and short neuropeptide F),^{21,25,55,164} the dorsal paired medial (DPM) neuron (GABA, serotonin, sparkly, nplp1, proctolin, orcokinin, eclosion hormone, CG34136)^{25,56,60} and most central complex, mushroom body output and mushroom body dopaminergic neurons.^{25,157} Of these, we excluded Kenyon cells due to their unusual synapse morphology.

For each neuron reported in the literature, we assumed that all associated presynapses obeyed Dale’s law (Figure 1B). The intention was that our FAFB ground truth data would consist entirely of manually placed presynaptic sites, which has the advantage of later training our network with only high-fidelity presynapses. Synapses were annotated at presynaptic sites, defined by T-bars, vesicles, and a thick dark active zone by a synaptic cleft.¹⁶⁵ We scored each continuous synaptic cleft as a single presynapse regardless of its size or the number of associated T-bars. Note that where we say we used FlyWire data, we looked at transmitter predictions across automatically detected presynapses in FAFB¹² using the newer FlyWire (<https://ngl.flywire.ai/>) reconstruction.^{17,18,50} This enabled us to work with the full fly brain connectome.

Many neuronal cell types with transmitter expression data had already been linked to reconstructions in the HemiBrain dataset by the reporting authors or.⁵¹ We matched the remaining unlinked cell types using the morphological match algorithms NBLAST and/or a color MIP search.^{47,48,50} We used automatically detected presynapses from identified neurons,¹³ which were more numerous than the manually placed ones we had from Catmaid, but potentially contained more low-quality identifications and multiple detections across the same continuous cleft. Manually placed presynapses were not available in this dataset at scale.

In total, we matched 3,025 FAFB-Catmaid neuronal reconstructions to cell types with a known transmitter. The assembled Catmaid ground truth dataset contained 153,593 acetylcholine presynapses (587 neurons), 7,953 glutamate presynapses (50 neurons), 32,836 GABA presynapses (175 neurons), 9,526 dopamine presynapses (83 neurons), 4,732 serotonin presynapses (5 neurons), and 2,924 octopamine presynapses (6 neurons) (see Data S2). We also matched 5,902 HemiBrain reconstructions to cell types with a known transmitter. It contained 451,033 acetylcholine presynapses (3,094 neurons), 75,239 glutamate presynapses (218 neurons), 80,732 GABA presynapses (242 neurons), 117,054 dopamine presynapses (310 neurons), 70,460 serotonin presynapses (38 neurons), and 46,017 octopamine presynapses (21 neurons).

Due to a relative paucity of linked presynapses when we began this study we did not use our annotations for: allatostatin A (2 neurons in the HemiBrain), corazonin (6), drosulfakinin (6), glycine (9), insulin (23), IPNa (2), *Drosophila* NPF (3), and SiFamide (4). Histamine is also used in the *D. melanogaster* brain, primarily by photoreceptor neurons not captured in our original datasets, and a few ascending neurons. Nitric oxide is used as a co-transmitter,²⁵ we did not include it in our dataset. We therefore only carried forward: acetylcholine, glutamate, GABA, dopamine, serotonin and octopamine. At least one of these six transmitters is probably expressed by most neurons in the brain.²¹

We thank Michael Reiser, Vivek Jayaraman, Arthur Zhao, Tatsuo Okubo, Jenny Lu and Kathi Eichler for identifying neuron matches in Catmaid which helped us build our ground truth dataset, Mareike Selcho for helping to confirm our FAFB-FlyWire identifications of octopaminergic neurons and Clare Pilgrim for further cell type to known transmission annotations.

Chosen neuronal reconstructions

We calculated neuron-level transmitter predictions for all 24,666 well reconstructed neurons in the HemiBrain dataset (see Data S3), all 49,985 central brain neurons and 86,942 neurons from FAFB-FlyWire (see Data S4) and 23,503 ventral nerve cord neurons from the MANC dataset. Data was pulled from: HemiBrain: <https://neuprint.janelia.org/?dataset=hemibrain:v1.2.1>, FAFB-FlyWire: <https://codex.flywire.ai>, MANC: <https://neuprint.janelia.org/?dataset=manc:v1.0>. All FAFB-FlyWire data is from the 630 materialization, a specific public version of the FlyWire dataset first released with the flagship FAFB-FlyWire dataset preprint.¹⁸ Our synapse-level transmitter predictions can be associated with these segments and therefore neurons built by the community even if they are further edited (further FAFB-FlyWire dataset releases are planned, i.e., the 783 materialization destined for the publication). Note that the male ventral nerve cord results were generated using our methodology in the MANC dataset (manc:v1.0) but reported on elsewhere.⁵ We included all neurons from the project with the status “Traced” and gave 3638 neurons that had fewer than 100 presynapses the label “uncertain”. Due to the availability of single cell type ground truth, only acetylcholine, glutamate and GABA were predicted in this dataset. Together, these results are part of our data releases and shown in Figure 5A and used in Figure 6.

For other biological analyses (Figures 5D–5F, S3E–S3H, and S4), we analyzed a subset of central brain neurons from HemiBrain and FAFB-FlyWire. We excluded Kenyon cells and neurons because we had very high confidence that these predictions were incorrect. We kept other possible mispredictions. We excluded truncated neurons and neurons with fewer than 100 presynapses. Truncated neurons typically included bilateral neurons in HemiBrain and ascending, descending and first-order sensory neurons in both datasets.

For this purpose, 246,66 neurons in HemiBrain were filtered down to 11,277 by excluding neurons without the HemiBrain project status label “Traced”, neurons with less than 100 presynapses and removing truncated neurons with a large part of their arbor cut from the dataset (label “cropped”), including neurons with the terms: LP, LC, LT, DN and LLP in their cell type label. These neurons were semi-automatically reconstructed by,⁵¹ and come from neuPrint release hemibrain:v1.2.1.

For the FAFB-FlyWire dataset we used 88,115 intrinsic brain neurons. Neurons were selected as well traced if a human annotator could confirm that they appeared to have a full dendrite, axons, cell body fibre tract and cell body (soma). This pool is small compared with the total 136,927 FAFB-FlyWire neurons because 1) they were selected as well reconstructed in FAFB-FlyWire by the start of 2023 and 2) we only selected neurons that we could ‘skeletonize’ and ‘split’ into separable axons and dendrites^{73,133} for our biological analyses and 3) we only chose neurons whose arbors were contained within the central brain.

Neurons are unmatched between datasets in our analyses e.g., (Figures 5 and S3) unless noted otherwise (e.g., Figure 4), i.e., the corpus of neuronal cell types from each dataset is partially overlapping but distinct, cell type content may account for some of the FAFB–HemiBrain differences. The full pool of HemiBrain neurons was not cross-matched to FAFB-FlyWire neurons at the time of writing, only 2626 neuronal cell types.

We thank the flywire.ai community for allowing us to use their semi-automatic FAFB-FlyWire neuronal reconstructions, which took over 1,366,543 edits from human annotators to build from automatically reconstructed segments.¹⁷ We also want to specifically thank those human annotators that contributed to a smaller pool of 27,706 neurons, built between 2019 and 2022, on which our results were initially piloted and preprinted. These neurons were built by 1,366,543 edits of automatically reconstructed segments,¹⁷ from 100 human annotators Those persons contributing more than 1,000 edits were: Doug Bland, Austin T Burke, Yijie Yin, Laia Ser-ratosa Capdevila, Kyle Patrick Willie, Arti Yadav, Ryan Willie, Nash Hadjerol, Zairene Lenizo, Griffin Badalemente, J. Anthony Ocho, Shirleyjoy Serona, Dharini Sapkal, Anjali Pandey, Ben Silverman, Varun Sane, Zeba Vohra, regine salem, Mendell Lopez, J. Dolorosa, Imaan Tamimi, Chitra Nair, Dhwani Patel, Joshua Bañez, Márcia Santos, Katharina Eichler, Shaina Mae Monungolh, Dustin Garner, Jay Gager, Joseph Hsu, Mark Larson, Bhargavi Parmar, Rey Adrian Candilada, Dhara Kakadiya, Alexandre Javier, Itisha Joshi, Michelle Pantuan, Irene Salgarella, James Hebditch, Kaushik Parmar, Darrel Jay Akian, Kendrick Joules Vinson, Marina Gkantia, Ariel Dagohoy, remer tancontian, Chan Hyuk Kang, Hane Two, Markus Pleijzier, Emil Kind, Olivia Sato, Yashvi Patel, Miguel Albero, Eva Munnely, Katie Molloy, Christopher Dunne, Quinn Vanderbeck, Rashmita Rana, Merlin Moore, Lucia Kmecova, Alexis E Santana Cruz, Nadia Seraf, Usb, Claire McKellar, Monika Patel, Mareike Selcho, Greg Jefferis, Steven Calle, Siqi Fang, Arzoo Diwan, Sarah Morejohn, Christa Baker, Brian Reicher, Sangeeta Sisodiya, Tansy Yang, Paul Brooks, Selden, Marlon Blanquart, Hyungjun Choi, Celia D, Sanna Koskela, Joanna Eckhardt, Krzysztof Kruk, Wolf Huettteroth, Alisa Poh, Stefanie Hampel, Wes Murfin, Li Guo, Zhihao Zheng, Szi-chieh Yu, Jones, Farzaan Salman, Amalia Braun, Mark Lloyd Pielago, Nidhi Patel, Ben Gorko, Akanksha Jadia, Fernando J Figueroa Santiago and Urja Verma. We thank Forrest Collman, Casey Schneider-Mizell, Chris Jordan, Derrick Brittain, Akilesh Haligeri for CAVE development and maintenance, and Kai Kuehner, Oluwaseun Ogedengbe, Jay Gager, Will Silversmith, Ryan Morey for Neuroglancer development, tools, and Codex development - these tools made efficient reconstruction possible.

Neuron skeletonization and axon-dendrite splits

Some morphology analyses required “skeletonized” neuronal reconstructions and splitting them into axon and dendrite.⁷³ For HemiBrain we retrieved neuronal skeletons from the HemiBrain neuPrint project,¹³⁵ which were generated by an edge collapse method. For FAFB-FlyWire we calculated skeletons using the “wavefront” (rays = 2) algorithm as implemented in the python library “skeletor”.¹²⁸ We “skeletonized” or retrieved skeletons for a subset of neurons (HemiBrain: 11,277, FAFB-FlyWire: 88,115)

Separate axons and dendrites^{73,81,83} were discerned using a graph theoretic algorithm,⁷³ implemented in the R based natverse toolbox.¹³³ To improve the results of this algorithm, we removed likely erroneous presynaptic detections on the soma, primary neurite or otherwise outside of the synaptic neuropil, a mean of ~ 5% per neuron. In general, the soma, cell body fibre tract and linker cable (primary dendrite) do not contain a large number of synaptic contacts. The raw numbers of synapses after filtering can be quite different between the FAFB-FlyWire and HemiBrain datasets (FAFB-FlyWire: median 202, hemibrain: median 386). This method benefited from the O2 High-Performance Compute Cluster, supported by the Research Computing Group, at Harvard Medical School (<https://it.hms.harvard.edu/our-services/research-computing>).

Neuron level transmitter prediction confidence

Assuming Dale’s law holds for all neurons, we labeled each neuron with a single transmitter and provided a confidence score for this assignment. To do this, we first retrieved all presynaptic sites associated with each neuronal reconstruction. We removed potential erroneous presynapses by employing standard thresholds for synapse detection for each dataset (a cleft score above 50 for FAFB-FlyWire auto-detected synapses,¹⁸ and a confidence score above 0.5 for HemiBrain auto-detected synapses⁵¹). We only considered neurons with at least 100 presynapses. We further validated our cut-off for automatically-detected FAFB presynapses¹² by evenly sampling 4,306 presynapses across all six synapse-level transmitter predictions and neuronal compartments (see Figure 5C) and

determining whether the auto-detected preynapses were valid by cross-checking with a human annotator. 32% of auto-detected presynapses were erroneous. Our chosen threshold eliminates ~13% of valid presynapses and ~60% of erroneous detections (Figure S1A). Because a higher proportion of auto-detected presynapses on the primary dendrite, cell body fiber tract and soma were erroneous, we retained only axonic and dendritic presynapses. (Figure S1B). Presynapses had also to be no less than 15 μm from a neuron's cell body, and 0.1 μm from its primary dendrite, to ensure that erroneous detections on non-synaptic cable were not used, this accounted for ~5% of presynaptic connections per neuron. Filtering presynapses produced a different neuron-level transmitter prediction in 4% of neurons (columns 'top_nt' (unfiltered after the 50 cleft score threshold) and 'conf_nt' (filtered)). Notably, a larger proportion of octopamine predicted presynapses proved erroneous (43%), indicating that the network may be more likely to guess octopamine for dark features of non-synaptic origin (Figure S2F).

To calculate neuron-level transmitter predictions, we determined which synapse-level transmitter prediction was most common for each neuron. If the difference between the top and second highest transmitter was <10%, a neuron's neuron-level transmitter prediction was designated *uncertain*. To determine a neuron-level transmitter prediction confidence score for each neuron (or neuronal compartment), we computed the average confusion matrix value of each presynapse prediction as given in the row of the winning transmitter. More formally, we assigned a confidence value $c(n)$ to a neuron n as:

$$c(n) = \frac{1}{|S_n|} \sum_{s \in S_n} C_{\hat{y}_n \hat{y}_s}, \quad (\text{Equation 1})$$

where S_n is the set of presynapses of neuron n , \hat{y}_n the winning transmitter of the entire neuron, \hat{y}_s the predicted transmitter of synapse s , and C the confusion matrix as computed on the test dataset. For example, if we determined a FAFB-FlyWire neuron was cholinergic because a majority of its presynapses were predicted to transmit acetylcholine, each presynapse predicted as acetylcholine would contribute a value of 0.95—the proportion of cholinergic ground-truth presynapses correctly determined as cholinergic (Figure 2A)—and any presynapse presumably mispredicted as GABA would contribute a value of 0.02 (the proportion of ground-truth cholinergic presynapses mispredicted as GABAergic). A neuron's confidence score is the mean of these values over all presynapses. The distribution in these scores across the FAFB-FlyWire and HemiBrain datasets shows that we have most confident predictions in acetylcholine and least confident in serotonin (Figure 2D). Indeed, our serotonin predictions are less reliable, with several suspected mispredictions including missing known serotonergic neurons (Figure S2E).

It should also be noted that each neuron was given a predicted transmitter identity even if in reality it would express none of the six transmitters with which we trained. When the network is applied to data outside of the training distribution, it is often less confident in its erroneous prediction, but not always.

Correlating misprediction with reported unlearned transmission types

In examining our neuron-level transmitter predictions together with reported expression in the literature, we noticed a few issues and errors, which we have investigated to try to establish the major reasons for misprediction.

For HemiBrain neurons we know to use transmitters outside of the six we learned,⁵¹ we found that glycinergic neurons¹⁴³ were mostly predicted to use acetylcholine (mean confidence 0.66), Allatostatin A¹⁴³ were mostly predicted to use GABA (0.7) (even though it most commonly co-transmits with glutamate²¹), Corazonin¹⁶⁶ use was predicted as acetylcholine (0.3), Drosulfakinin^{167,168} use was mostly predicted as octopamine (0.47), insulin-like peptides^{169,170} use was mostly predicted as acetylcholine or serotonin (0.44), IP-Namide¹⁵⁴ use was mostly predicted as glutamate (0.32), *Drosophila* neuropeptide F¹⁴⁵ use was mostly predicted as octopamine (0.46) and SIFamide¹⁷¹ use was mostly predicted as octopamine (0.43).

Mispredictions in which one of our six learned transmitters is predicted when another should have been, can occur in one dataset but not the other (Figures S2D–S2F). For example, our mispredictions include neurons thought to be dopaminergic being assigned serotonin (FB4L), octopamine (FB4M) or glutamate (PPM1024) in HemiBrain but correctly in FlyWire, and other dopaminergic neurons being mispredicted as glutamatergic (PAM01) in FlyWire but not in HemiBrain, which may happen due to unclear dataset-specific confounds. However, some of the most notable mispredictions occur in both. For example, the largest scale misprediction we caught was our network's incorrect mass prediction of Kenyon cells as dopaminergic. Immunohistochemistry and RNA sequencing data have shown that Kenyon cells express the machinery necessary for cholinergic transmission.^{21,55,164} Despite this, our neuron-level transmitter predictions guessed dopamine across all Kenyon cell types (FlyWire: 99.9%, HemiBrain: 99.9%) with high confidence (FlyWire: mean, 0.627, s.d., 0.99, HemiBrain: mean, 0.54, s.d., 0.046). A possible source of error might be contamination by nearby presynapses belonging to other neurons. For example, dopaminergic PAM neurons were used in our ground truth data, and are synaptic partners of Kenyon cells in the highly synapse-dense region of the mushroom body (Figure S3I). The cube of image data (edge length 640 nm) centered on each predicted presynaptic location was not masked with neuronal identity, and features from proximal presynapses may have skewed the result. However, if we consider presynapses on Kenyon cell axons (in the mushroom body lobes, dense dopaminergic innervation) and dendrites (mushroom body calyx, only sparse dopaminergic innervation⁷⁸) separately, we see that dopamine is still the most common prediction for both (77% of axons and 95% of dendrites in FlyWire). This suggests that the confusion may arise either due to the uncommon biology of the Kenyon cell presynapse or because the network saw Kenyon cell presynaptic features near labelled dopamine presynapses during training – rather than observing dopamine presynaptic features near Kenyon cell presynapses when testing. More generally, across our dataset we observed a small percentage of axon-dendrite

prediction mismatches in their compartment-level transmitter predictions (FAFB-FlyWire: 6.5%, hemibrain: 11.0%) (Figure S3G). In neurons where there was no mismatch, there was a strong correlation between the compartment-level transmitter prediction for the axon and the dendrite (Figure S3A), suggesting that the image features impacting our predictions are the same for both. Because axons are far from dendrites in Euclidean space, these features are most likely features of the neuron and/or its presynapses themselves.

Co-transmission with transmitters outside of our training data may be a major reason why our results can sometimes differ from what we expect. In mammals, cases of co-packaging of fast-acting clear core vesicular transmitters are known, including transmitters of opposing downstream effect.¹⁷² While transmitter-related gene expression and immunoreactivity are largely exclusive between the fast-acting transmitters in flies,^{21,22,30,31} most neuropeptides and monoamines can co-transmit with fast-acting transmitters.^{21,163} Co-transmission of especially fast-acting transmitters alongside neuropeptides and other neuromodulators is expected to be common in the brain, although examples at the resolution of individual cell types are sparse. The $\alpha\beta$ and γ Kenyon cells are known to express the short neuropeptide F precursor, and immunostain for short neuropeptide F,¹⁷³ which is highly indicative of neuropeptide transmission or co-transmission. Kenyon cells may also express the neuropeptide sparkly.^{25,174} In addition, *Ddc*, which encodes a protein responsible for converting L-DOPA to dopamine, is expressed in $\alpha'\beta'$ and γ Kenyon cells, where it could play a role in the biosynthesis of dopamine or another aromatic L-amino acid.²¹ Indeed, we did not include Kenyon cells in our ground truth data, because of this evidence for co-transmission.

The other major mispredictions we noticed include the singleton cell type DPM in the mushroom body, some olfactory sensory neurons of the antennal lobe and some central complex neurons such LPsP, vDelta, hDelta, and a few fan-shaped body tangential neurons, particularly in HemiBrain. The mushroom body, antennal lobe and central complex are especially synapse-dense regions of the brain as revealed by presynaptic antibody staining (see Figure S3I for regional biases in transmitter prediction). They are also known to exhibit diverse neuropeptide expression.^{106,175} DPM has been found to express the fast-acting transmitter GABA the monoamine serotonin and several neuropeptides (sparkly, nplp1, proctolin, orcockinin, eclosion hormone, CG34136).^{25,56,60} DPM were not used in our ground truth data and in both datasets are predicted to be dopaminergic. 20% of olfactory sensory neurons in HemiBrain and 25% in FlyWire are likely mispredicted as serotonergic instead of cholinergic.⁶⁴ There is a glomerular pattern to the olfactory sensory receptor neuron cell types that are predicted serotonergic, and though this does not match up with known glomerular patterns for short neuropeptide F expression in sensory neurons¹⁰⁶ it may match an as-yet unstudied neuropeptide distribution in the antennal lobe. Notably, these presynapses are in a region of the brain (antennal lobe) with a high density of true serotonergic presynapses from the CSD neurons, though the CSD is pan-glomerular. The CSD neurons were represented in our ground truth, but olfactory sensory neurons were not.

Central complex LPsP neurons are potentially mispredicted as glutamatergic instead of dopaminergic^{78,176} in FlyWire and HemiBrain and are in a region of the brain (protocerebral bridge) with a high density of true glutamatergic presynapses from the delta7 neurons. It is a rare case of a known dopaminergic neuron to be mispredicted as something else in both datasets. Interestingly though, known DopR receptors are only weakly expressed in the region of LPsP output, the protocerebral bridge, and so LPsP may have in some way unusual dopaminergic presynapses.¹⁷⁷ In addition, preliminary data suggests LPsP may co-transmit dopamine and glutamate (personal communication, Pablo Reimers). Both delta7 neurons and LPsP were represented in our ground truth data. Other protocerebral bridge neurons, IbSpSP neurons, were predicted for acetylcholine in FlyWire but for glutamate in HemiBrain.

The neurons of the optic lobe, despite being poorly represented in our ground-truth data, are remarkably well predicted. 96% of ~29,000 optic lobe neurons (T4, T5, Tm2-4, Tm9, Tm20, L2-5, Lawf1-2) were correctly predicted acetylcholine. 87% of ~3,600 optic lobe neurons (C2, C2, Pm4, Mi4, Dm10) were correctly predicted GABA. 91% of ~1,600 optic lobe neurons (TmY5a, Dm8) were correctly predicted glutamate. However, some cell types exhibited a glutamate GABA confusion. ~5,700 neurons from the known glutamatergic cell types (L1, Dm1, Dm3, Dm4, Dm11, Dm12) are correctly predicted a putative inhibitory class, but only 61% were predicted glutamate and 30% GABA. Our only unusual result was that just 3% of Dm9 neurons were predicted glutamate as expected. Dm9 neurons have ~20-fold more autapses (self-connections) than usual,¹⁷⁸ which may present a confound. Mi1 neurons have a similar rate of autapses and only 67% correct acetylcholine assignment. We note that one wet-lab study reports a different expression from what we have predicted in some optic lobe cell types, in a study based on mosaic analysis with repressible cell marker (MARCM) clones generated from drivers such as Cha-GAL4.¹⁷⁹ These are TmY5a (acetylcholine from MARCM, glutamate from prediction), Mi2 (acetylcholine from MARCM, glutamate-GABA from prediction) and Pm1a (acetylcholine from MARCM, GABA from prediction). We believe that¹⁷⁹ were misled by MARCM, our experience has been that clones generated from a Cha-GAL4 driver are not faithful to acetylcholine transmission.^{142,159}

Our predictions have also provided a starting place for looking for dopamine, serotonin and octopamine neurons unknown to the literature, but particularly for serotonin we encourage data users to validate predictions in the wet lab. Genetic and histological studies have estimated there to be ~130 dopaminergic neurons, ~80 serotonergic neurons and ~44 octopaminergic neurons¹³⁸ in the brain. Discounting Kenyon cells and sensory receptor neurons, we have predicted 6052 dopaminergic, 2000 neurons and 289 octopaminergic neurons in FlyWire. Many of these cases likely indicate co-transmitting neurons, although especially for dopamine they may also reveal as-yet unknown dopaminergic cell types because the TH-GAL4 line most often used to find dopaminergic neurons is not completely faithful to dopamine transmission.¹⁷⁷ Some known monoaminergic neurons were mispredicted. Two monoaminergic DNs, DNg28 (serotonin¹⁸⁰) and DNg32 (octopamine¹⁸¹) - possibly because most of their presynapses are outside of the

brain, in the ventral nerve cord. 114 photoreceptor neurons, R1-8, which may express histamine and possibly acetylcholine,³⁹ are incorrectly predicted for octopamine in FAFB-FlyWire. We advise users to treat these predictions with caution.

In conclusion, by examining known cases of misprediction we think the main two reasons for misprediction are co-transmission and expression of transmitters outside of our training set. Both issues be remedied by expanding the training data with co-transmission information and by including neurons with other transmitters. The method presented here can be extended to support more than the six transmitters considered so far. The prediction of co-transmission will require an architecture change, as the current architecture can not confidently pick more than one transmitter at the time. One way this could be achieved is to replace the softmax layer in the current architecture with element-wise sigmoid activation functions. In addition to mispredictions stemming from co-transmission or other transmitters, rarer biology (such as autapses, unusual T-bars and unusual postsynapses) may also be a factor some of the time. Generally, we found that when there is conflict, the FlyWire predictions are often superior.

Working through fan-shaped body neuron expression

The fan-shaped body is a central brain neuropil that computes navigational variables, including internal goals for the fly. We focused on it because along with the antennal lobe we considered it one of the worst predicted parts of the brain in the HemiBrain dataset. In brief, it is built as a matrix with 9 rows and 10 columns.⁷⁷ Tangential neurons input projects to specific, whole layers. Other input types (e.g., PFN) target specific columns within a single row. Intrinsic neurons can tile rows (e.g., hDelta types) or columns (e.g., vDelta types). Output neurons sample across several columns (e.g., FC and FS types) or more specific segments (e.g., PFL and FR types). We further found that almost all intrinsic cell types were predicted to transmit acetylcholine (~88%), as well as almost all output cell types (~96%) and column-specific input types (~87%) with some misprediction due to a paucity of automatically detected presynapses in or near an unusual brain region⁵¹ known as the gall (Figure 5G, dashed circles). Synapse detection often fails in the gall,⁷⁷ as presynapses here have elongated T-bars and dense core vesicles, indicative of neuropeptide co-transmission (Figure 5G, circles). We found that almost all tangential neurons are predicted to transmit glutamate (~84%). Of the remainder, we are confident that 4 neurons (cell types: FB5AB, FB1G) are correctly predicted to transmit acetylcholine because they come from a cholinergic hemilineage ('DM4 dorsal') (Figure 7B). Indeed, an intersection with a ChA driver line has suggested that FB5AB and so-called 'vFB' neurons are cholinergic,¹²⁷ though note that vFB morphology is not consistent with FB1G. Interestingly, the only neurons predicted to use GABA in both FlyWire and HemiBrain are FB5A (primary neurons from lineage 'EBa1') and FB3B, FB3C and FB3E (secondary neurons from lineage 'EBa1').^{50,182} Although the FB3 neurons are predicted for glutamate in FlyWire we suspect the HemiBrain prediction for GABA could be correct because 'EBa1' is majority GABAergic (Figure 7B). Indeed, metabotropic GABA receptors have only been found specifically in layers 3 and 5.¹⁷⁷ A few dopaminergic tangential cell types have been reported (FB5H, FB6H, FB7B).⁷⁷ In FlyWire only FB6H is predicted dopaminergic, the others are predicted to transmit glutamate. In HemiBrain, they are all correctly predicted dopamine. In both datasets, predicted dopaminergic neurons include: FB1C, FB1H, FB2A, FB4M and FB4Y. They all come from the known dopaminergic PPM3 cluster.^{77,78,175} FB4L also belongs to the PPM3 cluster but are predicted serotonin in both HemiBrain and FlyWire. Interestingly, TH, DopEcR and DopR1 immunoreactivity can be detected weakly across the fan-shaped body,¹⁷⁷ likely by ExR2 and neurons receiving ExR2 input. However DopR1 is particularly strong in layers 1-2, DopR2 is expressed specifically and strongly in layer 3 (due to discrepancies between HemiBrain annotation and literature, "layer 3" in other literature seems to be related to the projection pattern of many FB4 neurons), and D2R is expressed specifically in dorsal layers 7-8.¹²⁴ This is in agreement with our predicted sites of dopamine transmission in the fan-shaped body.

We have also examined neuropeptides in the fan-shaped body as reported in the literature carefully, and here give our thoughts on how layer-specific expression may have caused some misprediction in this structure. Central complex hDelta cell types^{77,90,127,183} (not in our ground truth) are a set of 190 neurons whose morphologies are very similar, and which segment the fan-shaped body of the fly. However, our HemiBrain prediction results estimate 30% to be cholinergic and 45% to be dopaminergic. While the field lacks authoritative data, we expected them all to express the same transmitter or set of transmitters, most likely acetylcholine. Indeed, in FlyWire they come out as 83% cholinergic and 12% serotonergic, suggesting that all of the class should be cholinergic. The case is even worse with vDelta neurons, another similar class in the same neuropil born from the same set of hemilineages, for which HemiBrain predictions give us neuron-level transmitter predictions that are 40% glutamatergic and 35% dopaminergic, but are predicted 85% cholinergic in FlyWire. Fortunately, this issue is not common, only 14% of cell types in the HemiBrain have a transmitter prediction conflict. The fan-shaped body is known to express leucokinin, allatostatin A, short neuropeptide F, *Drosophila* neuropeptide F, tachykinin, proctolin, dFMRFa, SIFamide and myoinhibitory peptide in a highly specific and layer-wise manner.^{106,175,177,184} Because we expect all intrinsic and output types of the fan-shaped body to be cholinergic based on FlyWire results, we think the glutamate predictions for intrinsic cell types from HemiBrain are erroneous. The most notable example is the cholinergic hDeltaK neurons, which have been likely mispredicted to transmit serotonin in FlyWire and dopamine in HemiBrain because their presynapses lie in layer 6 of the fan-shaped body (Figure 5G, arrows), an unusual layer that is known to express allatostatin A, short neuropeptide F and innixin 6.^{175,184,185} hDeltaK is the only major case of intrinsic neuron misprediction in FlyWire. hDeltaC and hDeltaL are also likely mispredicted to transmit dopamine in HemiBrain and have their axons in layer 6. In addition, short neuropeptide F has been shown to co-localise in 5-6 neurons with cholinergic transmission specifically in layers 5-6,¹⁷⁵ indicative that another hDelta class co-transmits it because vDelta neurons per layer are more numerous. In particular, hDeltaF and hDeltaG are likely mispredicted to transmit dopamine in this region. We suspect one of those types co-transmits short neuropeptide F. hDeltaJ (layer 4) and hDeltaM (layer 5) are predicted to transmit dopamine in HemiBrain but acetylcholine in FlyWire. It is less clear what neuropeptide expression, or not,

may have caused these confounds, but we note that dFMRFa is expressed in layer 4, and SIFamide and proctolin are expressed in layer 5.^{175,186} We suspect tachykinin expression in vDeltaA neurons, which are majority mispredicted to transmit dopamine in HemiBrain and minority mispredicted to transmit serotonin in FlyWire. However, only 16–18 neurons are thought to express tachykinin and there are over 80 vDeltaA neurons. It is also possible that the 16 hDeltaD/E neurons express tachykinin, half of which are mispredicted dopamine in HemiBrain, but likely correctly acetylcholine in FlyWire.

There are also a few mispredictions of glutamatergic tangential neurons, particularly in HemiBrain. Mispredictions are probably lower in these tangential neurons because glutamate has been shown mainly not to co-transmit with the most abundant neuropeptides in the fan-shaped body – short neuropeptide F, tachykinin and myo-inhibitory peptide.¹⁷⁵ However, leucokinin neuropeptide is expressed in layer 6 tangential neurons,¹⁸⁴ e.g., FB6A, which are erroneously predicted octopaminergic in HemiBrain but perhaps correctly for glutamate in FlyWire. We think it is likely that FB6A co-transmits glutamate and leucokinin. Tachykinin is also expressed in layer 2 of the fan-shaped body. Because it does not co-localize with ChAT stains in this layer and because of cell body placement, it might be expressed by specific glutamatergic tangential neurons^{177,187} and may account for why FB2B is mispredicted octopaminergic in HemiBrain, and FB2A and FB2I predicted dopaminergic in FlyWire and HemiBrain. Allatostatin A is expressed in layer 6 in neurons other than “dFB” FB6A neurons.¹⁸⁸ Because in HemiBrain FB6H/L/K/J are potentially mispredicted to transmit dopamine (likely correctly glutamate in FlyWire) we wonder if these could be the allatostatin A neurons. Proctolin is expressed in layer 5 and because FB5S has some dark vesicles and 1/7 in HemiBrain was predicted for octopamine, we wonder whether this is a proctolin co-transmitting type. Curiously, while most 5-HT receptors are not expressed in the fan-shaped body,¹⁷⁷ serotonin immunolabeling revealed that serotonin is used in layers 3, 6 and more ventral layers.^{150,175} One serotonergic source is ExR3, which projects to layer 6, but there may be others innervating more ventral layers. ExR3 is mispredicted to transmit octopamine in HemiBrain and dopamine in FlyWire, and has dense core vesicles indicative of unknown neuropeptide co-transmission. A final notable misprediction is the numerous, expected cholinergic class PFNp, which receives various mispredicted labels in HemiBrain and FlyWire. However, the class appears to have very few detected output presynapses, only a mean of 56 in HemiBrain, indicative of little or unusual neurotransmission, or else automatic synapse detection was problematic in the dorsal nodulus and layer 1 of the fan-shaped body.

Working through antennal lobe local neuron expression

We tackled in detail a difficult case concerning a morphologically variable and diverse class of neurons¹⁰⁵ from two split transmitter hemilineages: local neurons of the antennal lobe. We found we could still make sense of our neuron-level transmitter predictions in relation to the extant literature. We wanted to draw attention to this case because transmitter usage in these local neurons seems to break Dale’s law,^{104,106} Lacin’s law (Figure 7C) and the expectation of uniform transmitter expression within a type (Figures S5F and S5G), and so presented perhaps the biggest challenge for our resource. They are known to express acetylcholine,^{63,189,190} GABA,^{62,89,151} glutamate^{61,89,104} and various neuropeptides transmitters.¹⁰⁶

‘AL1 dorsal’ is a local neuron hemilineage that our Bayesian analysis identified as having a decisive split in its transmitter usage. From the literature, we expected ~ 54 local neurons per hemisphere of which ~ 90% would be GABAergic.¹⁵¹ We also expected perhaps as many as ~ 15 cholinergic local neurons.^{63,189} In our connectomic datasets, we find ~ 126 local neurons per hemisphere for ‘AL1 dorsal’.⁷⁴ We found that 143 (left: 74, right: 69) are predicted GABAergic, 48 (left: 20, right: 28) cholinergic, 13 (left: 9, right: 4) dopaminergic, 14 (left: 7, right: 7) glutamatergic and 37 (left: 18, right: 19) serotonergic. Serotonin and dopamine in the antennal lobe have been well-studied but no reports of either in local neurons exist in the literature. We, therefore, think that these neurons should mostly have been predicted GABAergic or cholinergic, perhaps mostly the latter based on morphological similarity (Figure 7E). However, the other predictions could be valid, albeit surprising. The predicted cholinergic neurons are unilateral and ‘broad’, i.e., pan-glomerular, in accordance with previous descriptions^{63,189} but more numerous. Positioned closely among them are other local ILN1_bc neurons that are predicted to be dopaminergic. We suspect that some cholinergic local neurons have been mispredicted dopaminergic because at least ~ 3 cholinergic neurons co-express allatostatin A and ~ 10–15 co-express myo-inhibitory peptide.¹⁰⁶ The predicted GABAergic local neurons in the hemilineage have cell body fiber tracts slightly offset from the cholinergic/dopaminergic subgroup. Half the local neurons that had previously been categorized as LN2 ‘patchy’ local neurons,^{74,105} i.e., those that innervate discontinuous glomeruli in space, were predicted glutamatergic (cell type: ILN2P_a) and half GABAergic (cell types: ILN2P_b, ILN2P_c). Analogously ‘picky’ local neurons in the larval brain are also known to be glutamatergic, while the grossly morphologically similar ‘choosy’ neurons are GABAergic.¹⁹¹ A subset of ILN1 neurons is known to express both GABA and glutamate to effect synaptic plasticity, but it is not known how large this subset is or whether this phenomenon extends to other local neuron classes.¹⁰⁴ Some of these neurons have been predicted GABAergic and some mispredicted serotonergic. Tachykinin is also known to co-express in ~ 40 GABAergic ‘AL1 dorsal’ neurons per hemisphere¹⁰⁶ and may have been a confound that generated serotonin predictions. Specifically, among them are ~ 13 ILN2T_abc neurons that have a cell body placement and general morphology most similar to reported cholinergic local neurons.^{63,189} The remainder is most likely GABAergic, e.g., ILN2_d gross morphology looks more like related GABAergic neurons.^{63,189}

The remaining, largely bilateral, local neurons come from a second hemilineage ‘ALv2’. This too, appears to be a split hemilineage containing 133 (left: 67, right: 66) glutamatergic neurons alongside a newly identified sub-population of 28 (left: 14, right: 16) potential cholinergic neurons. Previous immunohistochemistry work⁶¹ found no sign of GABA usage at ‘ALv2’ presynapses, and we predict only one ‘ALv2’ GABAergic neuron on each side (v2LN41), perhaps a deviant first-born neuron of the hemilineage. One ‘ALv2’ local neuron was predicted serotonergic (v2LN36), this unusual neuron had previously been determined to be glutamatergic using

immunohistochemistry and able to undergo stark morphological changes dependent on sex and mating state.¹⁹² ‘ALv2’ axons extend to specific glomeruli in the contralateral antennal lobe. Significantly, this means that specific lateral excitation between glomeruli on different sides of the brain should be possible. As with their glutamatergic counterparts, these sparse bilateral local neurons focus on connections mainly between thermosensory glomeruli.⁷⁴ Excluding serotonin and dopamine mispredictions our HemiBrain predictions were similar, but with some misprediction of ‘ALv2’ neurons as using GABA rather than glutamate. We have a little further detail on local neuron misprediction in our Methods section on misprediction correlates. In conclusion, we think there could be far more cholinergic local neurons in the antennal lobe than previously thought which has implications for second-order olfactory information processing. See Figures S5F and S5G. We thank Asa Barth-Maron for discussions and insights into antennal lobe local neurons.

HEMILINEAGE DEFINITION IN *D. MELANOGASTER*

The neurons of the central nervous system are generated by a set of stem cells known as neuroblasts. During division, neuroblasts generate two cells, one additional stem cell and one cell that further divides into two sibling neurons. In only one of these siblings, the so-called Notch pathway is activated, leading to two different *hemilineages* of neurons within each lineage, one Notch positive the other Notch negative^{31,193–195} (or more than two in the case of Type II neuroblasts, which mainly contribute to the central complex, Figure S5A). In some cases, one of a lineage’s hemilineage may apoptose,¹⁹⁶ leaving only a single hemilineage. Neurons can be born in this way during embryogenesis and are known as primary neurons (~ 10% of the adult brain^{134,197}). Many neuromodulatory neurons appear to be primary.^{138,141} Those made during larval development are known as secondary neurons. We found 120 lineages that broke down into 183 hemilineage-associated tracts, per hemisphere. Of these, we think 144 are secondary Type I hemilineages. The remainder are secondary Type II. These eight lineages (Type II; ‘DM1-6’; ‘DL1-2’) produce intermediate progenitors, each in turn delivering two hemilineages. Their relationship with their lineage-associated tracts is often less clear.

Neuronal fibers co-fasciculate with neurons from the same hemilineage and form bundles as they enter the neuropil from the insect brain’s outer layer of cell bodies. Only secondary hemilineages (each ~ 100 neurons) can easily be demarcated in adult *D. melanogaster*. In EM data annotators can make our “lineage-associated tracts” (sometimes called “secondary axon tracts”), many of which contain a single hemilineage. It should be noted that cases have been identified¹⁰² where a single lineage-associated tract is composed of two hemilineages. Type II hemilineages may be less uniform than Type I in their tract organization. Each intermediate progenitor of a Type II lineage behaves like a “small neuroblast” and produces 10-20 neurons, divided into two hemilineages.¹⁹⁸ According to this idea, there should be around 50-100 hemilineages for each type 2 neuroblast. That does not match the morphologically visible tracts in BP106 strings in the larva or adult central brain.

Our lineage-associated tracts have been mapped by light microscopy and assigned to named lineage clones.^{139,199–201,202} We have discovered them in the adult brain FAFB dataset.⁵⁰ Lacin et al.³¹ have shown that Each hemilineage in the adult ventral nerve cord uses just one of the fast-acting transmitters, acetylcholine, glutamate or GABA, even though mRNA transcripts for combinations of these can appear in the nucleus. We have found the same with our predictions in the ventral nerve cord.⁶⁵ We refer to this principle as *Lacin’s law*. Using the FlyWire dataset we show in this work that it largely holds in the adult central brain with some interesting exceptions (Figure 7).

Because HemiBrain dataset is only a partial brain, many neurons have large missing portions or do not exist in this dataset, causing a discrepancy between neuron count and hemilineage content compared with FlyWire. However, we could still identify most central brain hemilineages and see that their predicted transmitter expression and projection envelope are similar to those of FlyWire neurons (Data S6). Comparing GABA-glutamate mixed hemilineages (‘LHI1’, ‘VLPI4 dorsal’, ‘LHI4 dorsal’, ‘SLPav3’, ‘WEDd2’, ‘VLPI4 anterior’, ‘VPNp&v1 posterior’, ‘CREa1 ventral’, ‘VESa1’, ‘EBa1’, ‘DM2 central’) in FlyWire to HemiBrain indicates that perhaps they more uniformly express GABA (Figure S5A). On the other hand, HemiBrain predictions were more likely than FAFB-Catmaid predictions to confuse glutamate for GABA, but not vice versa (Figure 2A). In addition, dopamine misprediction seems be higher in HemiBrain, ‘CLp1’, ‘SMPpv2 dorsal’, ‘SMPpm1’, ‘VLPd&p1 posterior’, ‘SMPpv1’, ‘SMPpd1’, ‘VLPI&p1 posterior’, ‘CREI1’, ‘LHI2 dorsal’, ‘SLPpm3’, ‘VPNp1 medial’, ‘DM6 dorsolateral’, ‘DM3 dorsolateral’ have predicted dopaminergic neurons only in HemiBrain. The largest discrepancy was for ‘DM6 IbSpsP’. IbSpsP neurons were predicted for acetylcholine in FlyWire but for glutamate in HemiBrain.

Hemilineage assignments in *D. melanogaster*

Our work assigning neurons in the FAFB-FlyWire dataset to hemilineages has been reported in Schlegel et al.⁵⁰ We briefly detail the process here. Cell body fiber tracts for identified hemilineages had previously been identified using TrakEM2²⁰³ in a light-level atlas for a *D. melanogaster* brain, stained with an antibody against neurotactin (BP104).¹⁹⁹ We extracted these expertly identified tracts and registered them into a common template brain, JFRC2, using CMTK,²⁰⁴ and then into FAFB space.¹³³ These lineage-associated tracts could then be readily identified in EM stacks based on their point of entry and subsequent trajectory. We generally consider lineage-associated tracts as fiber units belonging to one hemilineage. However, we suspect that the putative hemilineages ‘DL1 dorsal’, ‘DL1 ventral’, ‘DL2 dorsal’ and ‘DL2 ventral’ (derived from the Type II lineages ‘DL1’ and ‘DL2’) may be examples where this has happened, and each of these groupings is a composite of not one but two hemilineages (giving four hemilineages per lineage here, which is possible because these are Type II lineages,^{205–207} a Type II would have two). We identified cell body fiber tracts in the ssTEM

FAFB dataset in the vicinity of the transformed lineage-associated tracts using the flywire.ai Web interface.¹⁷ We compared our candidate neurons to registered images of lineage clones, produced using genetic tools and light microscopy.^{139,201} Hemilineages can be told apart in an image of a lineage clone, by the different placement of their cell body clusters, and the tract their cell body fibers take into the neuropil. From light-level data that labels a whole lineage and all of its hemilineages together,^{139,199–201} it is occasionally difficult to assess whether there is a single hemilineage or multiple, if the cell body clusters happen to not separate sufficiently. While they can be told apart using developmental genetic tools, e.g., in the case of ‘LALv1’,¹⁰² this has not been done at scale.

In Figure 7E we focused on birth order in ‘LALv1’ because ‘LALv1 ventral’ exhibits split transmitter use. ‘LALv1’ is, fortunately, one of few hemilineages whose birth order has been delineated.¹⁰² ‘LALv1’ neurons were matched 1:1 to FAFB-FlyWire reconstructions with a combination of visual inspection and morphological clustering⁴⁷ to the morphological progression by birth order¹⁰².

We have used 183 well-annotated hemilineages of the central brain, which were human-reviewed on both the right and left FAFB-FlyWire hemispheres, in this work. However, assigning neurons to hemilineages is difficult. For example, neurons from the lineage ‘LALv1’ co-bundle between its two hemilineages, making their disambiguation dependent on examining light-level data and single neuron clones.¹⁰² Without prior work¹⁰² on this lineage’s detailed composition, we might have incorrectly assumed that its glutamatergic and cholinergic neurons from its two separate hemilineages intermingled in a single hemilineage. High-quality light-level data is not available for every hemilineage, meaning that we had to make some judgment calls in our hemilineage discrimination.

Indeed, the ‘DL1’ hemilineages at first seemed a striking example of split transmitter use within hemilineages, i.e., disobeying Lakin’s law. However, it is possible that the ‘DL1’ and ‘DL2’ hemilineages we have delineated are actually each not a single hemilineage but two, i.e., there are four hemilineages in total for ‘DL1/2’.^{205–207} The split transmitter use may reveal this hemilineage-based division in a lineage-associated tract that was otherwise hard for human annotators to make. Therefore we suspect that our split-use label in this case reveals an issue of data annotation.

Hemilineage prediction entropy

Neuron level entropy

To quantify multimodality of transmitter predictions on neuron level within a hemilineage, we calculated the entropy H of the transmitter distribution over neurons in the following way: Let $n \in N_h$ be a neuron in hemilineage h and $\hat{y}_n \in Y = \{\text{GABA, acetylcholine, glutamate, serotonin, octopamine, dopamine}\}$ the predicted transmitter of neuron n . Then

$$H(N_h) = - \sum_{y \in Y} p_h(y) \log_6 p_h(y), \text{ with} \quad (\text{Equation 2})$$

$$p_h(y) = \frac{1}{|N_h|} \sum_{n \in N_h} \delta(\hat{y}_n = y). \quad (\text{Equation 3})$$

A value of $H(N_h) = 0$ then means that all neurons within hemilineage h have the same predicted transmitter, while a value of $H(N_h) = 1$ means that within hemilineage h all predicted transmitters are equally common.

Synapse level entropy

Similarly, we quantified the average multimodality over synapses within neurons of a given hemilineage: Let $s \in S_n$ be the synapses in neuron $n \in N_h$ of hemilineage h and \hat{y}_s the predicted transmitter. The entropy of predicted synaptic transmitters $H(S_n)$ in neuron n is then given by:

$$H(S_n) = - \sum_{y \in Y} p_n(y) \log_6 p_n(y), \text{ with} \quad (\text{Equation 4})$$

$$p_n(y) = \frac{1}{|S_n|} \sum_{s \in S_n} \delta(\hat{y}_s = y) \quad (\text{Equation 5})$$

With this the average synaptic entropy over all neurons within hemilineage h is given by:

$$H(S_h) = \frac{1}{|N_h|} \sum_{n \in N_h} H(S_n) \quad (\text{Equation 6})$$

A value of $H(S_h) = 0$ means that all synapses of all neurons in hemilineage h have the same predicted transmitter, while a value of $H(S_h) = 1$ means that in all neurons within hemilineage h all synaptic transmitter predictions are equally common. Figure 7D shows the distribution of $H(N_h)$ and $H(S_h)$ of all predicted hemilineages’ neurons that have more than 30 synapses each.

On the population level, we found relatively higher values of $H(S_h)$ (Synapse level entropy) than $H(N_h)$ (Neuron level entropy). 75% of hemilineages show a synapse level entropy below $q_{75}(H(S_h)) = 0.41$ as compared to $q_{75}(H(N_h)) = 0.20$. This is reassuring as it

suggests less variation of transmitter identity of neurons within a hemilineage compared to variations of transmitter identity predictions within individual neurons, meaning there is improved consensus of predictions when aggregating across populations. However, in cases with a high level of synaptic entropy, such as hemilineage 'TRdl_a', it is less clear whether neuron-level multimodality is an artifact of uncertain, multimodal predictions on the synapse level of individual neurons. In contrast, hemilineages such as 'SMPpd1' show high neuron level entropy $H(N_h) \geq q_{75}$ but low synapse level entropy $H(S_h) \leq q_{25}$, suggesting clear neuron level segregation of predicted transmitters within those hemilineages. Hemilineages such as 'ALad1' with $H(S_h) \geq q_{25}$ and $H(S_n) < q_{25}$ appear homogeneous within each neuron and within the entire hemilineage.

Probability to observe transmitter predictions \hat{y}

Given a neuron has true transmitter $y \in Y$, the probability that we predict transmitter $\hat{y} \in Y$ (assuming that each prediction is independent and identically distributed) is given by the categorical distribution

$$p(\hat{y}|y) = C_{y,\hat{y}}, \quad (\text{Equation 7})$$

where C is the neuron confusion matrix obtained on the test dataset (see Figure 2A).

Let m be the number of different transmitters in hemilineage h . We model the probability $p(\hat{y}|m)$ of observing transmitter predictions $\hat{y} = \{\hat{y}_0, \hat{y}_1, \dots, \hat{y}_n\}$ under the assumption that hemilineage h contains m different transmitters. Here, \hat{y}_j is the predicted transmitter of neuron j in hemilineage h with n neurons total. Let $\mathbb{P}_c(Y)$ be the set of subsets of true transmitters Y with cardinality c , then:

$$p(\hat{y}|m) = \sum_{S \in \mathbb{P}_m(Y)} p(\hat{y}|S) \cdot p(S|m), \quad (\text{Equation 8})$$

where $p(\hat{y}|S)$ is the probability to observe predictions \hat{y} if the hemilineage has true underlying transmitters $y \in S$ and $p(S|m)$ is the probability for the set of true transmitters S given the hemilineage contains m different transmitters. Since we assume i.i.d. predictions \hat{y} , $p(\hat{y}|S)$ factorizes as follows:

$$p(\hat{y}|S) = \prod_j p(\hat{y}_j|S) \quad (\text{Equation 9})$$

and marginalizing over $y \in S$ yields:

$$p(\hat{y}|S) = \prod_j \sum_{y \in S} p(\hat{y}_j|y) \cdot p(y|S) \quad (\text{Equation 10})$$

$$= \prod_j \sum_{y \in S} C_{y,\hat{y}_j} \cdot p(y|S). \quad (\text{Equation 11})$$

Regarding $p(S|m)$ and $p(y|S)$ we assume a flat prior, i.e.,:

$$p(S|m) = \binom{|Y|}{m}^{-1} \quad (\text{Equation 12})$$

$$p(y|S) = \frac{1}{|S|} = \frac{1}{m}. \quad (\text{Equation 13})$$

With this, the probability of observing predictions \hat{y} given m different transmitters becomes:

$$p(\hat{y}|m) = \binom{|Y|}{m}^{-1} \sum_{S \in \mathbb{P}_m(Y)} \left(\prod_j \sum_{y \in S} C_{y,\hat{y}_j} \cdot \frac{1}{|S|} \right) \quad (\text{Equation 14})$$

Bayes Factor

With this formalism in place, we can compare hypotheses about the number of true transmitters m in a given hemilineage by using the Bayes Factor $K = \frac{p(D|M_1)}{p(D|M_2)}$, where D is our observed data (predicted transmitters) and M_1, M_2 are two models about the underlying true transmitters that we wish to compare. The Bayes factor for a model M_1 with m_1 true transmitters per hemilineage and model M_2 with m_2 different transmitters is given by:

$$K_{m_1, m_2} = \frac{p(\hat{y}|m_1)}{p(\hat{y}|m_2)} \quad (\text{Equation 15})$$

$$= \frac{\binom{|Y|}{m_1}^{-1} \sum_{S \in \mathbb{P}_{m_1}(Y)} \left(\prod_j \sum_{y \in S} C_{y,\hat{y}_j} \cdot \frac{1}{m_1} \right)}{\binom{|Y|}{m_2}^{-1} \sum_{S \in \mathbb{P}_{m_2}(Y)} \left(\prod_j \sum_{y \in S} C_{y,\hat{y}_j} \cdot \frac{1}{m_2} \right)}. \quad (\text{Equation 16})$$

So far, we assumed that $p(\hat{y}_j|y) = C_{y,\hat{y}_j}$, i.e., we estimated this distribution on the test dataset. However, because our test set is finite we can not expect that the estimated error rates perfectly transfer to other datasets. To relax our assumptions about this distribution we simulated additional errors, by incorporating additive smoothing on the counts of neurons $N_{y,\hat{y}}$ that have true transmitter y and were predicted as transmitter \hat{y} , i.e.:

$$\tilde{C}_{y,\hat{y}} = \frac{N_{y,\hat{y}} + \beta}{\sum_y N_{y,\hat{y}} + 6\beta}, \quad (\text{Equation 17})$$

where $\beta \in \mathbb{N}_0$ is the smoothing parameter. With $C_{y,\hat{y}} = \frac{N_{y,\hat{y}}}{\sum_y N_{y,\hat{y}}}$ we then have

$$\tilde{C}_{y,\hat{y}} = \frac{C_{y,\hat{y}} + \frac{\beta}{\sum_y N_{y,\hat{y}}}}{1 + 6 \frac{\beta}{\sum_y N_{y,\hat{y}}}} = \frac{C_{y,\hat{y}} + \alpha}{1 + 6\alpha} \quad (\text{Equation 18})$$

and $\alpha \in \mathbb{R}_{\geq 0}$ the count normalized smoothing parameter. In the limit of $\alpha \rightarrow \infty$, $\tilde{C}_{y,\hat{y}}$ approaches the uniform distribution with probability 1/6 for each transmitter, whereas a value of $\alpha = 0$ means we recover the observed confusion matrix C . With this our distributions are now parametrized by α and the Bayes factor becomes:

$$K_{m_1, m_2} = \frac{\int_\alpha p(\hat{y}, \alpha | m_1) p(\alpha) d\alpha}{\int_\alpha p(\hat{y}, \alpha | m_2) p(\alpha) d\alpha} \quad (\text{Equation 19})$$

$$= \frac{\tilde{p}(\hat{y}|m_1)}{\tilde{p}(\hat{y}|m_2)}, \quad (\text{Equation 20})$$

where $\tilde{p}(\hat{y}|m)$ is as defined in 14 but with C_{y,\hat{y}_j} replaced with its expected value $\mathbb{E}_{p(\alpha)}[\tilde{C}_{y,\hat{y}_j}]$.

The prior distribution on α , $p(\alpha)$ allows us to encode our prior knowledge about α and use it to weight the likelihood of the corresponding model. Given the data, a value of $\alpha = \epsilon$ with epsilon small ($0 < \epsilon \ll 1$), should be most probable, while the probability of values $\alpha > \epsilon$ should monotonically decrease as we deviate more from the observed confusion matrix. Values of $\alpha < \epsilon$ should have probability zero, because they correspond to the un-smoothed confusion matrix with zero entries, i.e., a probability of zero for misclassification of certain transmitters. While these probabilities may be small, they are likely greater than zero and an artifact caused by the finite test set. Many distributions fulfill these criteria, in particular the family of exponential distributions with rate parameter λ :

$$p(\alpha) = \begin{cases} \lambda e^{-\lambda(\alpha - \epsilon)} & \alpha \geq \epsilon \\ 0 & \alpha < \epsilon \end{cases}. \quad (\text{Equation 21})$$

Thus λ controls the weight for smoothing parameter α in the integral $\mathbb{E}_{p(\alpha|M)}[\tilde{C}(\alpha)_{y,\hat{y}_j}] = \int_\alpha \tilde{C}_{y,\hat{y}_j} p(\alpha) d\alpha$. For $\lambda \rightarrow 0$, the expected confusion matrix approaches the unweighted average of all $C(\alpha)$ in the integration range. For $\lambda \rightarrow \infty$, the expected confusion matrix approaches the ϵ -smoothed confusion matrix $\tilde{C}_{y,\hat{y}} = \frac{C_{y,\hat{y}} + \epsilon}{1 + 6\epsilon}$.

The rate parameter λ can also be understood via its influence on the expected average accuracy $\tilde{C}_{\text{exp}} = \frac{1}{6} \sum_i \mathbb{E}_{p(\alpha|M)}[\tilde{C}]_{i,i}$. For values of $\lambda \rightarrow 0$, the expected accuracy approaches chance level while for values of $\lambda \rightarrow \infty$, the expected accuracy approaches the ϵ -smoothed, observed accuracy on the test set.

To summarize overall maximum likelihood of number of true transmitters in a given lineage, for a fixed λ we consider a one-versus-rest Bayes factor:

$$K_{m,-m} = \frac{\tilde{p}(\hat{y}|m)}{\sum_{n \neq m} \tilde{p}(\hat{y}|n)} \quad (\text{Equation 22})$$

Identification of ultrastructural features

We identified ultrastructural features by visualizing relevant differences between pairs of transmitters. To this end, we trained CycleGANs⁷² to translate original images of presynapses from one transmitter to counterfactual images of another transmitter. Since

a single CycleGAN can translate images in both directions (e.g., the same CycleGAN can be used to translate GABA → glutamate and glutamate → GABA), this required training of a total of 15 CycleGANs to translate each of the six transmitters into any other transmitter. Each CycleGAN was trained on the same FAFB training dataset we used to train our synapse classifier, but limited to 2D images for efficiency.

For each pair of transmitters (i,j), we then translated each original synapse image of class i into a counterfactual image of class j . We confirmed that the translation was successful by classifying both the original and counterfactual image. To this end, we retrained our synapse classifier on 2D images as well and only kept a pair of images if the original image was correctly classified as i and the counterfactual as j , with a score of at least 0.8 for each. For each image pair, we then found the smallest mask such that swapping the contents of this mask between the original and counterfactual image flips the classification (details of this procedure are given in⁷¹). We then sorted all image pairs by the size of this mask and visually inspected the 40 pairs with the smallest mask between each pair of transmitters (Data S8).

We took note of visual differences between the original and counterfactual images within the masked region to obtain hypotheses about ultrastructural features. Crucially, we only included differences that were symmetric between translations: for each observed difference, we required that the difference is visible when translating from class i to j , as well as the inverse when translating from j to i .

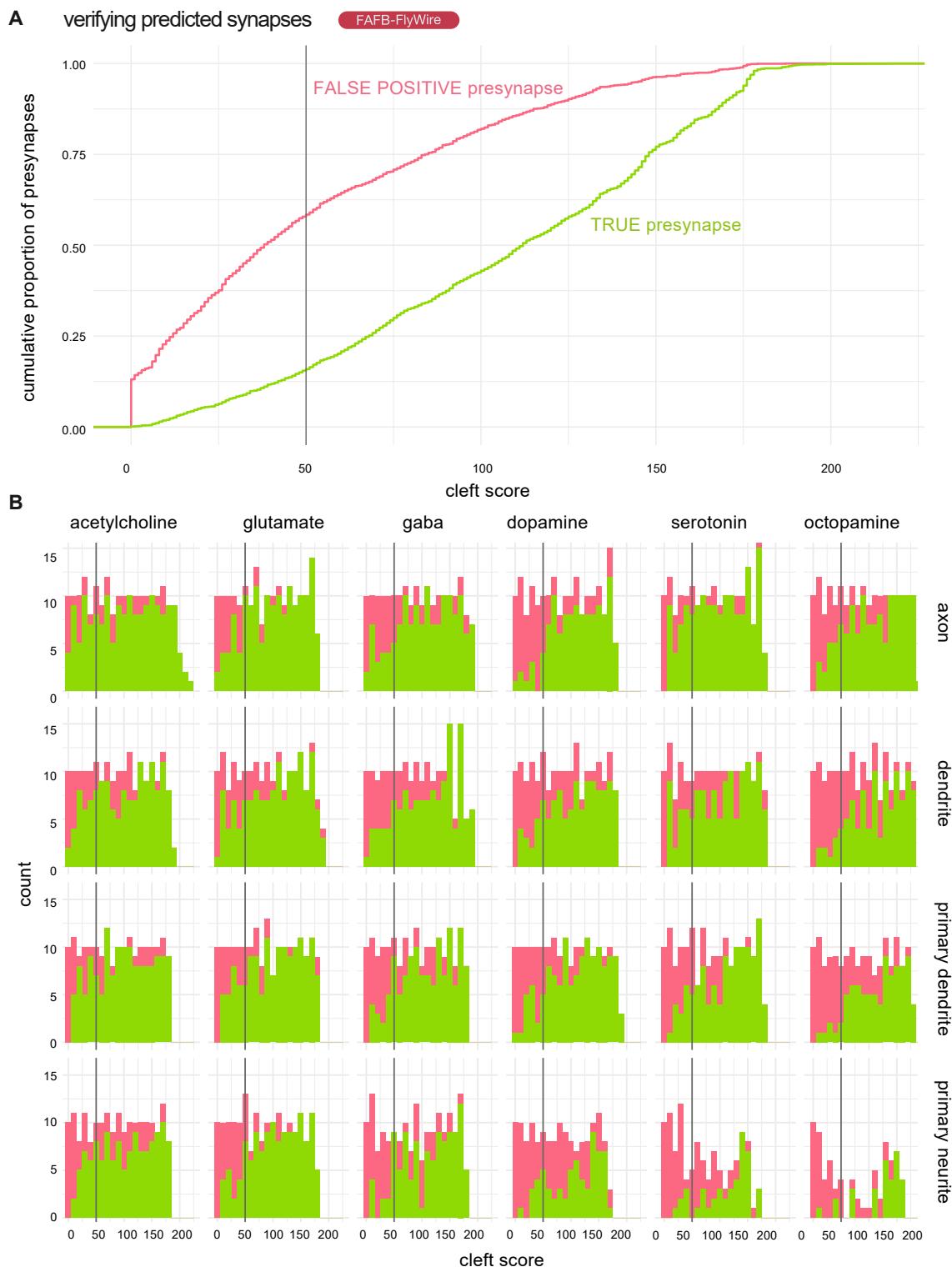
Advice for resource users

We have publicly released (Zenodo, <https://doi.org/10.5281/zenodo.10593546>) our comprehensive results encompassing the entire central nervous system (FAFB-FlyWire, HemiBrain and MaleVNC). We thank Davi Bock, Eric Perlman and Stuart Berg and the flywire.ai project for helping us make our preliminary results available to the community via CAVE. We thank Clare Pilgrim, Alex McLachlan and David Osumi-Sutherland for incorporating our predictions into Virtual Fly Brain. Researchers^{18,50,65,74,76,90,114,202,208–211} have already utilized our preprinted results to formulate neurobiological hypotheses. The most reliable predictions align across hemispheres, exhibit cohesion within a cell type, adhere to Lacin's law and remain consistent across datasets. We have provided two worked examples of less reliable cases, tackling misprediction among major fan-shaped body classes and antennal lobe local neurons. Considerations for users include: (i) Dale's law in the insect should be taken as mutually exclusive between acetylcholine, glutamate and GABA usage but we expect co-transmission with both monoamines and neuropeptides to be common. (ii) Confounds will produce an error rate that may differ between neuron classes. A user can note higher-level features to provide greater clarity, e.g., Lacin's law. Users should scrutinize neuron-level transmitter predictions that do not agree with the majority vote in a hemilineage (Data S6), in particular, GABA versus glutamate confusions (e.g., Figures 5G and S5A). True deviations may occur for distinct morphological subsets or among first-born neurons. (iii) Direct inspection of presynapses in EM data can reveal confounds, such as irregular T-bars or large dense core vesicles for co-transmitted neuropeptides. Unusual-looking presynapses are less likely to yield trustworthy predictions. (iv) Caution is advised for serotonin results, often indicative of uncertainty, while unreported octopamine results may only suggest dense core vesicle use (Figures 2D and S2). (v) One can be more confident in results that are consistent for the same cell type between hemispheres and datasets (Figure 4). Where there is conflict, we often find our FAFB-FlyWire predictions to be more reliable. (vi) Given the neuron-level transmitter prediction confidence scores we see for well-matched neurons, we suggest that if users wish to be conservative they could use a stringent neuron-level transmitter prediction confidence score threshold of ~ 0.62 and ~ 0.53 for FAFB-FlyWire and HemiBrain neurons respectively (1 s.d. lower than the mean score among concurring matches). (vii) One needs also to be wary of the opposite roles glutamate could play in a circuit. (viii) Our model is forced to select one of the six classes, even if there is no evidence for any of them.

QUANTIFICATION AND STATISTICAL ANALYSIS

All violin plot data is presented as the 25th percentile, the median and the 75th percentile, unless otherwise indicated. Unless otherwise indicated, all statistical tests were Wilcoxon two-sample tests, unless otherwise indicated. The Bayes analyses performed and our confidence score calculation are described in our method detail. Statistical comparisons were analyzed R. The N in our analyses are indicated either in the figure legend or in the figure panels themselves. Our Figure 5A provides the total number of neuronal reconstructions by dataset (FAFB-FlyWire: 136,927, HemiBrain: 24,666). For Figures 5D–5F, S3E–S3H, and S4) a subset of total central brain neurons that were skeletonised were used in analysis (FAFB-FlyWire: 88,115, HemiBrain: 11,277). Counts by class, lineage, hemilineage, cell type and matches between datasets (shared cell_type column) can be found in our supplementary data (FAFB-FlyWire: Data S4, HemiBrain: Data S3). Figures were prepared in R¹³⁰ using Ggplot2¹³² and were arranged using Adobe Illustrator.

Supplemental figures



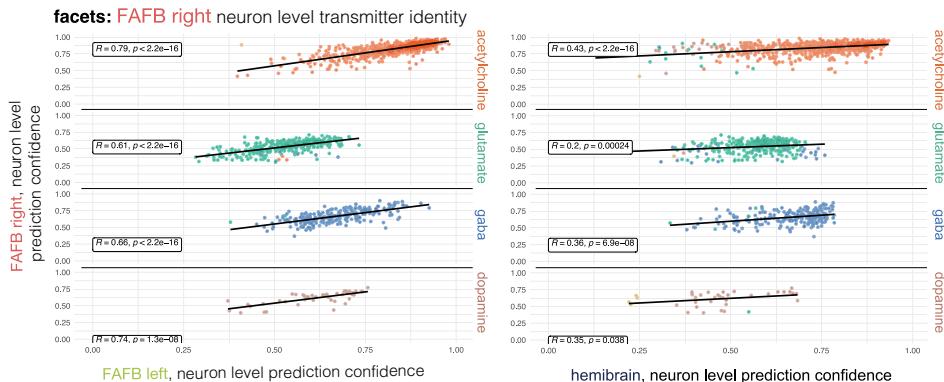
(legend on next page)

Figure S1. Human review of auto-detected presynapses in the FAFB-FlyWire dataset, related to Figures 3 and 2

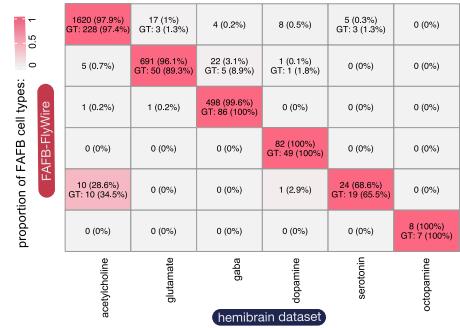
(A) Empirical cumulative density distribution curve for review of 4,306 automatically detected preynapses from the FAFB-FlyWire dataset. A presynapse comprises the synaptic machinery and vesicles on the source neuron's side of the synaptic cleft. Detected preynapses have a "cleft score" that ranges between 0 to over 200, which indicates how discriminable the synaptic cleft at the presynaptic site is for the detection network.¹² Our threshold of 50 is indicated by a vertical grey line. Green, determined to be a true presynapse by a human annotator; pink, determined to not be a true presynapse.

(B) Rates of false presynapse detection across cleft scores, transmitter types, and compartments. We sampled ~180 for each set of conditions. We sampled ~10 presynapses per cleft score bin (width 10), presynaptic transmitter prediction type (columns) and neuronal compartment (rows). Histograms show the number of presynapses determined to be real (green) or not (red) by a human annotator (A.S.B). Presynapses were reviewed using the flywire interface.¹⁷

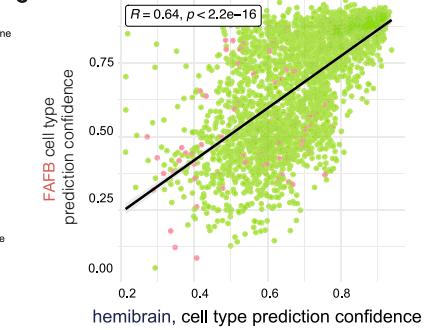
A transmitter prediction confidences correlate between hemispheres and data sets



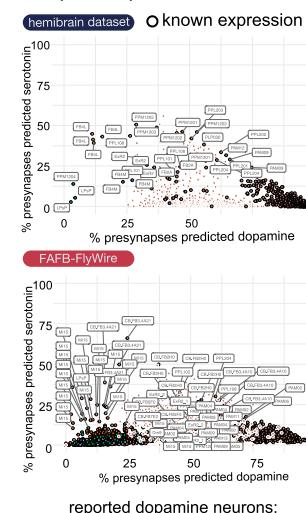
B consistency between matched cell types



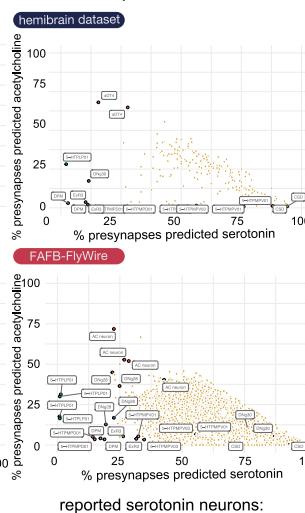
C hemibrain, neuron level prediction confidence



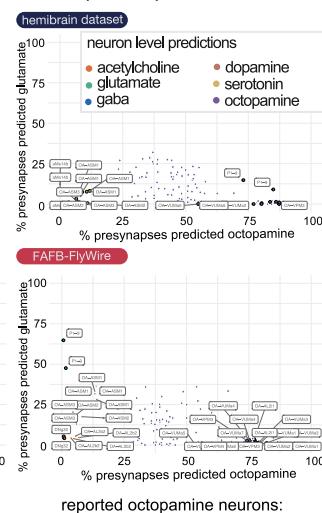
D dopamine predictions



E serotonin predictions



F octopamine predictions

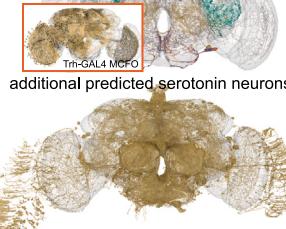


reported dopamine neurons:
TH-GAL4 MCFO

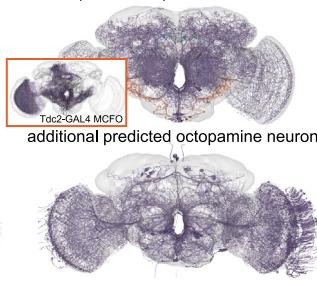
additional predicted dopamine neurons:



reported serotonin neurons:



reported octopamine neurons:



(legend on next page)

Figure S2. Comparing neurons' transmitter predictions between connectome datasets from separate animals (FAFB-FlyWire and HemiBrain) and between two hemispheres in the same dataset (FAFB-FlyWire), related to Figures 4 and 5

(A) Left: Scatterplots comparing neuron-level confidence scores in the transmitter prediction of FAFB-FlyWire-right central brain neurons, faceted by the neuron-level transmitter prediction for the FAFB-FlyWire right side homolog. Individual points colored by their FAFB-FlyWire-left or HemiBrain neuron-level transmitter prediction. Only 40 (2.5%) disagree with the FAFB-FlyWire left side prediction and 94 (7.7%) disagree HemiBrain. Right, scatterplots comparing neuron-level confidence scores in the transmitter prediction of HemiBrain dataset central brain neurons. Individual points colored by their HemiBrain side neuron-level transmitter prediction score. Only 94 (7.7%) disagree with the HemiBrain prediction.

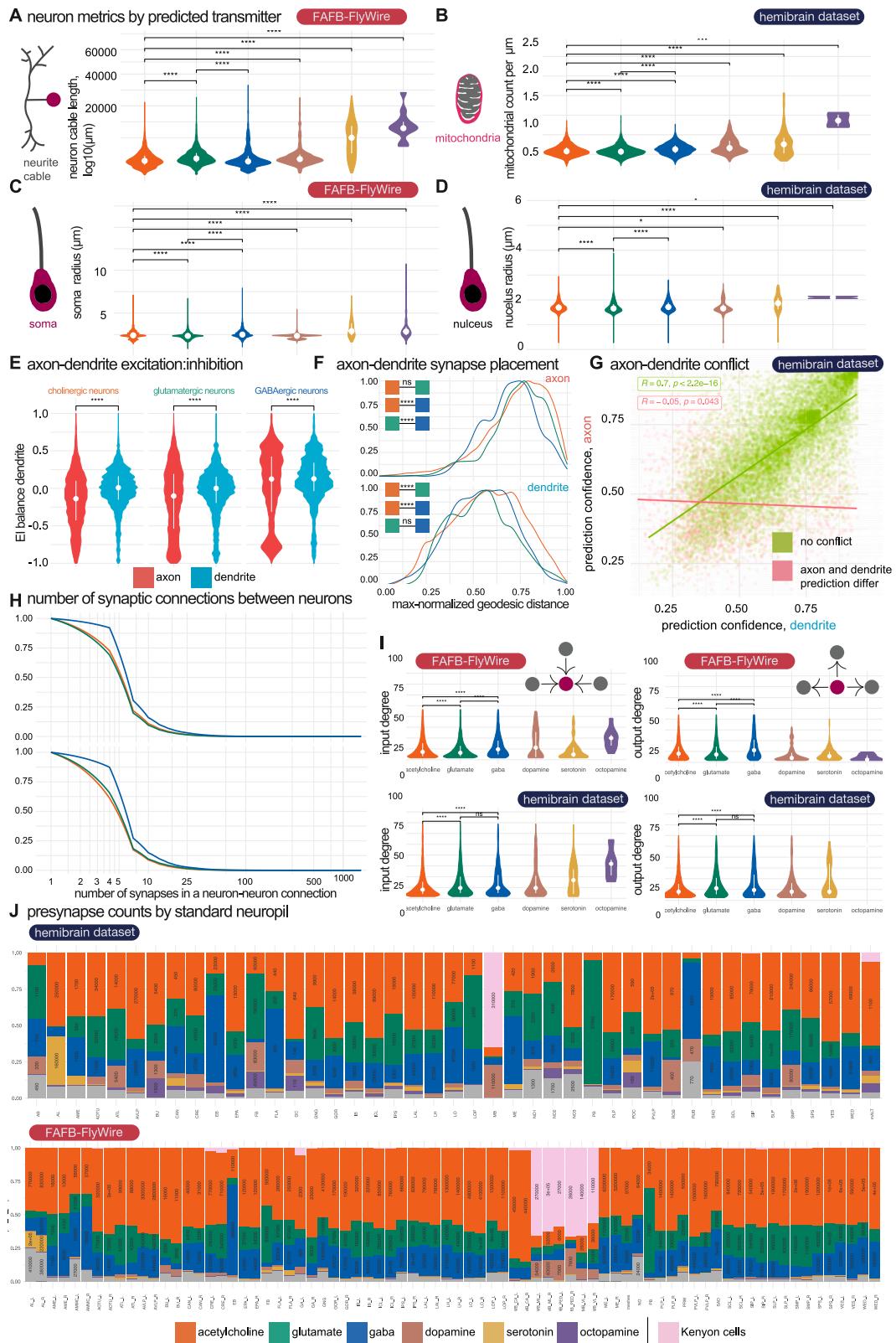
(B) A confusion matrix showing the neuronal cell type level prediction (mode of the neuron-level transmitter predictions per neuronal cell type) for neuronal cell types in the FAFB-FlyWire and HemiBrain datasets. Cells give the number of cross-matched neuronal cell types we examined, and the number of those present in the ground truth data for at least one of the two datasets.

(C) A scatterplot showing the correlation between our mean prediction confidence scores for FAFB-FlyWire and HemiBrain neuronal cell types. Each point is a neuronal cell type identified in both datasets (2626). Green points mean that the transmitter prediction agrees between the two datasets and pink points indicate disagreement. Scatterplots display Pearson's product-moment correlation, giving R, the coefficient and the associated p-value.

(D) A look at dopamine predicted neurons. We show two scatterplots using data and predictions from the HemiBrain (upper) and FAFB-FlyWire (lower) datasets. The proportion of presynapses in each neuron (each point) that are predicted as dopamine (X axis) and serotonin (Y-axis). Neurons that have been predicted as dopaminergic, or known as dopaminergic from the literature (dark circles), are shown. Those neurons from the ground truth data are circled with a black ring. Upper brain plot shows neurons known to be dopaminergic (colored by their neuron-level transmitter prediction). The visual system Mi15 neurons are thought to express dopamine and acetylcholine.³⁹ Lower, brain plot shows neurons strongly predicted to be dopaminergic (>50% of presynapses 'voting' for dopamine), excluding those in the upper plot. Many weakly predicted dopaminergic neurons belong to the central complex and mushroom body, where the density of presynapses from other neurons may have contributed to possible mis-predictions (see STAR Methods). Image inset with orange border shows light-level single neuron skeletons from MultiColor FlipOut experiments from the FlyCircuit project,²¹² from the TH-GAL4 line which labels most putative dopaminergic neurons in the fly brain. All neurons have been transformed onto the right hemisphere of the standard FlyCircuit template brain, FCWB. We have found the FlyCircuit MultiColor FlipOut data (23513 morphologies) to be unfaithful to the expected expression patterns for Cha-GAL4 (cholinergic neurons), vGlut-GAL4 (glutamatergic neurons) and Gad1-GAL4 (GABAergic neurons) and therefore of limited use in assigning transmitters, but provide the data for monoamines here to give the reader some impression of what whole brain expression patterns may look like.

(E) Same as a, but for serotonin predictions. Some PPL101-6 neurons are co-express dopamine and serotonin but are predicted as dopaminergic. Some known serotonergic neurons have low proportions of presynapses predicted as serotonergic. Flycircuit neurons from the Trh-GAL4 driver shown in inset. Trh is involved in serotonin biosynthesis.

(F) Same as a, but for octopamine predictions. Flycircuit neurons from the Tdc2-GAL4 driver shown in inset, which labels putative octopaminergic and tyraminergic neurons. Most octopamine neurons have been identified in prior work. Many of our octopamine predictions (no dark circle) indicate neurons that express some other dense core vesicle transmitter in abundance, for example, PI neurons which express an insulin-like peptide. Interestingly, the putative octopaminergic aMe14b neurons¹³⁸ (also known as OA-AL2b2 neurons) are predicted for acetylcholine. Busch et al. noted that they might not be octopaminergic, as not all neurons in cluster AL2 of NP7088 are OA-immunoreactive, and because OA-AL2b2 (HemiBrain type: aMe14b) was identified in NP7088, but not in tdc2-GAL4. OA-ASM (HemiBrain type: aMe14b) neurons also are not predicted octopaminergic, but serotonergic. On OA-ASM, Busch et al. note: "There are 8 OA-immunoreactive somata localized to the anterior superior medial protocerebrum uniquely labeled by tdc2-GAL4 (the ASM cluster). Yet they are not necessarily octopaminergic, as there are GAL4-positive neurons without OA-immunoreactivity in this cluster."

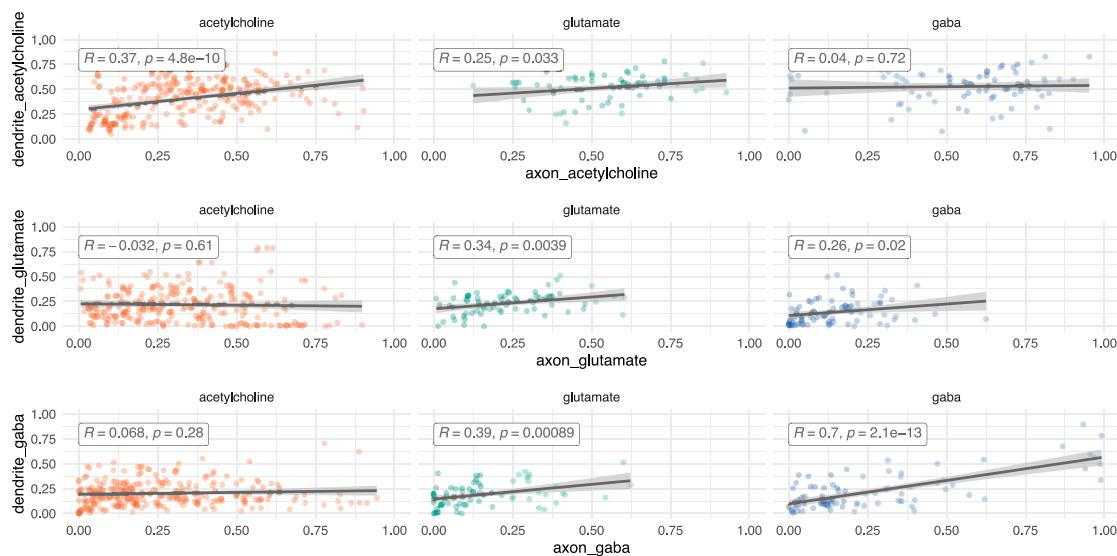


(legend on next page)

Figure S3. Comparing neuron features across transmitter classes, related to Figure 5

- (A) Cable length by neuron-level transmitter prediction.
- (B) Mitochondria density by neuron-level transmitter predictions. Violin plots show the number of automatically detected mitochondria¹³⁵ per micron cubed. Volume measures per neuron originate from the HemiBrain's automatically reconstructed 3D neuron volumes.⁵¹ A mitochondria detection is currently only available in the HemiBrain dataset. The mean number of mitochondria per neuron is 245, s.d. 275.
- (C) Soma, i.e., neuronal cell body, and (D) nucleus size by neuron-level transmitter predictions. The HemiBrain dataset provides a soma segmentation (left), and the FAFB-FlyWire dataset provides a nucleus segmentation (right)^{17,135}.
- (E) Violin plots of excitation:inhibition balance by neuron-level transmitter prediction and compartment.
- (F) Correlation between compartment-level transmitter prediction score for axons and dendrites. Each point is a separate neuron in the HemiBrain dataset, $n = 10,122$. 11.0% disagree on the compartment-level transmitter prediction (red). The scatterplot displays Pearson's product-moment correlation, giving R, the coefficient and the associated p-value.
- (G) Scaled density plot showing the density of input connections onto all FAFB-FlyWire and HemiBrain neurons (facets) at different synaptic weights (X axis, log2).
- (H) Scaled density plots showing the max-normalised geodesic distance (the distance along a neuron's arbour) from input synapses (colored by the source neuron's neuron-level transmitter prediction) to the target neurons' cell body.
- (I) Differences in the number of outgoing and incoming connections by neuron-level transmitter prediction. The input and output degree for a neuron is the number of unitary connections it has incoming and outgoing, respectively (the number of synaptic pairs, regardless of synaptic weight). All source-target connections with a synaptic count ≥ 10 included. Left, boxplots show the distribution of input degrees by the target neurons' neuron-level transmitter prediction. Right, output degrees by the source neurons' neuron-level transmitter prediction. A subset of total central brain neurons that were skeletonized (see STAR Methods) were used for this analysis (FAFB-FlyWire: 88,115, HemiBrain: 11,277).
- (J) Breakdown of neuron-level transmitter predictions by brain region in HemiBrain. Plot shows the proportion of synapses in each HemiBrain neuropil that belong to a neuron of a given neuron-level transmitter prediction (colors). A total of $\sim 4,000,000$ were assigned a neuropil and neuron-level transmitter prediction, which helps buffer erroneous synapse-level transmitter predictions. Number labels give the total number of synapses in each group. Not all the standard neuropils⁵⁴ are shown because the HemiBrain only comprises $\sim 1/3$ of the central brain. Total number of neuronal reconstructions (see STAR Methods) by dataset: FAFB-FlyWire: 136,927, HemiBrain: 24,666. (J) Breakdown of neuron-level transmitter predictions by brain region in FAFB-FlyWire. Neuropils⁵⁴: AB, asymmetric body, AL, antennal lobe, AME, accessory medulla, AOTU, anterior optic tubercle, ATL, antler, AVLP, anterior ventrolateral protocerebrum (incomplete in HemiBrain), BU, bulb, CAN, cantle, CRE, crepine, EB, ellipsoid body, EPA, epaulette, FB, fan-shaped body, FLA, flange, GC, great commissure (incomplete in HemiBrain), GNG, gnathal ganglion (incomplete in HemiBrain), GOR, gorget, IB, inframedial bridge, ICL, inferior clamp, IPS, inferior posterior slope, LAL, lateral accessory lobe, LH, lateral horn, LO, lobula (incomplete in HemiBrain), LOP, lobula plate (incomplete in HemiBrain), ME, medulla (incomplete in HemiBrain), NO1, nodulus compartment 1, NO2, nodulus compartment 2, NO3, nodulus compartment 3, PB, protocerebral bridge, PLP, posterior lateral protocerebrum, POC, posterior optic commissure, PVLP, posterior ventrolateral protocerebrum (incomplete), ROB, round body, RUB, rubus, SAD, saddle, SCL, superior clamp, SIP, superior intermediate protocerebrum, SLP, superior lateral protocerebrum, SMP, superior medial protocerebrum, SPS, superior posterior slope, VES, vest, WED, wedge. Violin plots show the median value (dot) and the inter-quartile range (line, 25th to 75th percentiles). Significance values: ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$.

A matched neurons clustered by transmitter identity of synaptic input



B cross-matched neurons clustered by transmitter identity of synaptic input

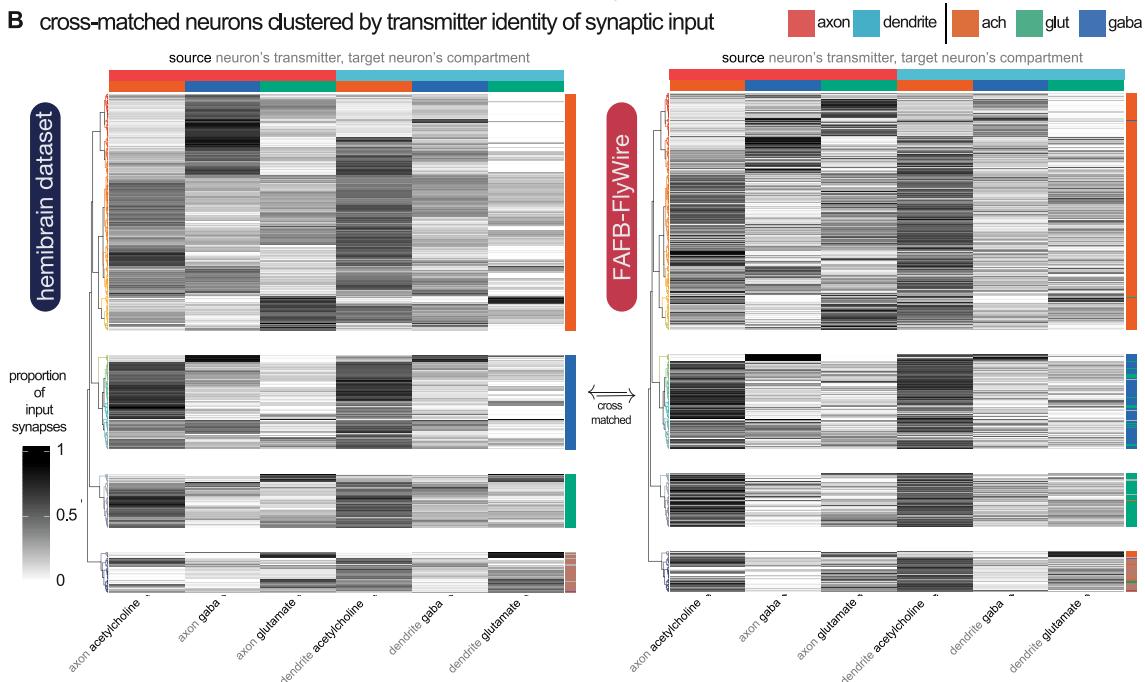


Figure S4. Diversity in targeting by transmitter class and compartment, related to Figure 6

(A) Correlations between opposing input transmitter types by compartment. Plots faceted by the source (upstream) neurons' neuron-level transmitter prediction (axis values). Colored by the target (downstream) neurons' neuron-level transmitter prediction. The X axis shows the proportion of a neuron's input accounted for by the input type on the axis label. Each dot is one neuron. For calculating the R^2 and p -values, neurons for which a proportion on either the X or Y axis fell below 0.1 or above 0.9 were excluded, to remove outlier cases with a very strong input preference.

(B) Equivalent 451 neuronal cell types from the FAFB-FlyWire and HemiBrain datasets clustered by input type. Only neurons for which at least 50% of inputs came from well-reconstructed and predicted neurons in our 88,115 FAFB-FlyWire neurons or 11,277 HemiBrain neurons were used. For each source neuron to target neuron connection, we used the identity (neuron-level transmitter prediction), location (neuronal compartment) and normalized connection weight (number of synaptic contacts made on that compartment / total number of synaptic inputs to the target neuron). We calculated cell type averages, and separated target cell types by their transmitter prediction and then clustered within each grouping. Heatmaps show the proportion of synaptic input onto the axon (upper horizontal color bar, red) and dendrite (blue), separated by the neuron-level transmitter prediction for each input (lower horizontal color bar). Each row is a separate neuronal cell type (see Figure S4C for names). Cell types are grouped by a hierarchical clustering within their neuron-level transmitter prediction class (vertical color

(legend continued on next page)

bar: acetylcholine, glutamate or GABA) employing Ward's clustering criterion. This clustering was performed in the HemiBrain dataset and applied to the FAFB-FlyWire dataset. Dendrogram (left) colors show a split into 30 groups. The same dendrogram is used in both heatmaps. Cosine similarity, $z = 0.892$, p -value < 0.0001, 100,000 row shuffles. A subset of total central brain neurons that had been skeletonized (see [STAR Methods](#)) were used for these analysis (FAFB-FlyWire: 88,115, HemiBrain: 11,277).

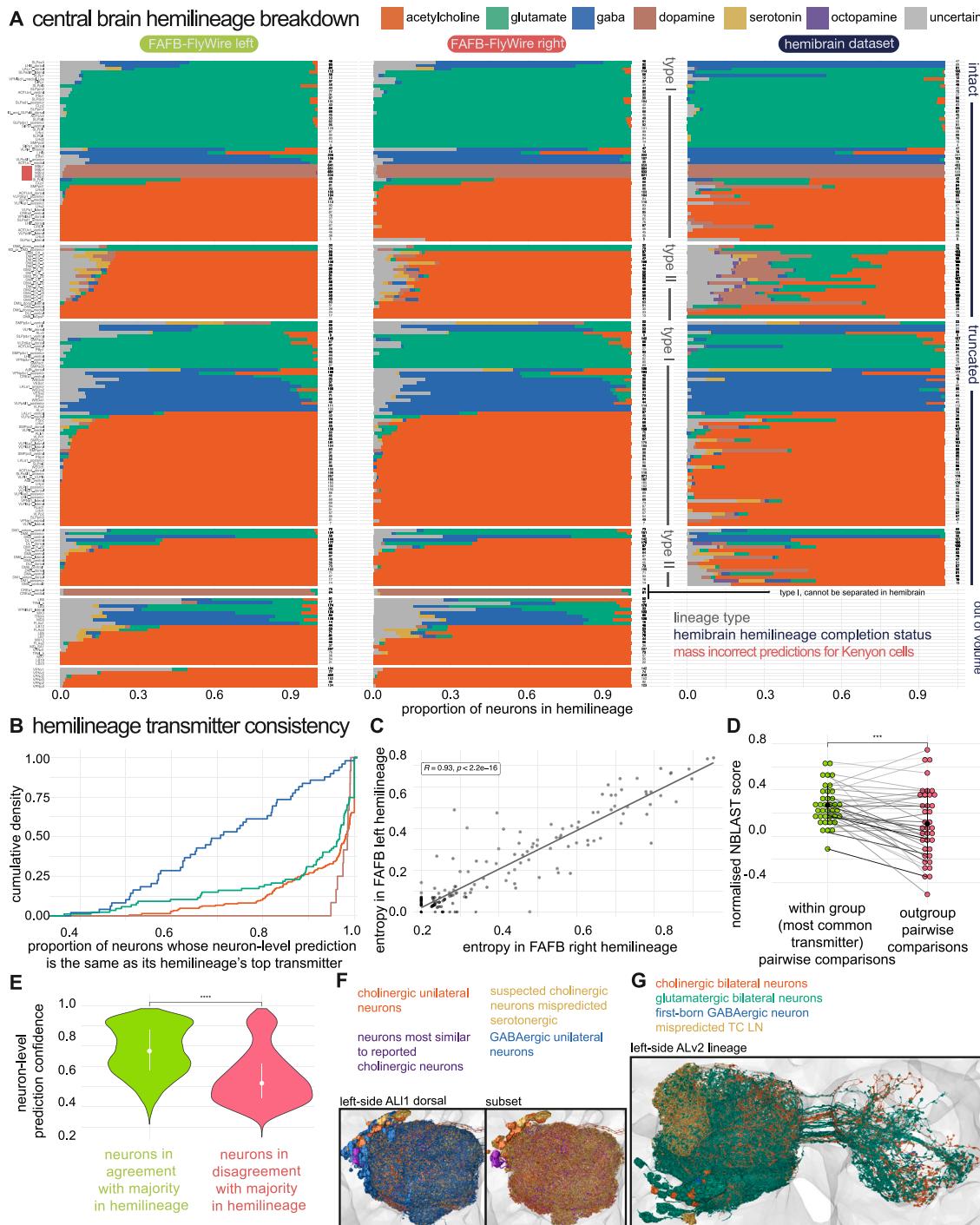


Figure S5. The distribution of neuron-level transmitter predictions within secondary hemilineages, related to Figure 7

(A) Consistency of neuron-level transmitter predictions within selected hemilineages in the central adult *D. melanogaster* brain. Bar plots show the proportion of neurons in each hemilineage predicted to express each of our six transmitters. Data is shown for neurons of the left (left) and right (middle) hemispheres of the FAFB-FlyWire dataset, as well as both hemispheres of the HemiBrain dataset (right). Note that the HemiBrain dataset is only a partial brain, many brain neurons have large missing portions or do not exist in this dataset. Hemilineage names are given on the left of the bar plots, and the numbers of neurons per hemilineage are on the right. The red bar highlights lineages of cholinergic Kenyon cells, MBp1-4, which are mispredicted dopaminergic. The plot is faceted first by presence in the HemiBrain dataset (intact, truncated, missing), then by lineage type (Type I and Type II).

(B) Empirical cumulative density plot shows how consistent a transmitter within each hemilineage is predicted to be. The Y axis gives the proportion of hemilineages, and the X axis gives the proportion of neurons in those hemilineages that “voted” for the top transmitter (color groups).

(legend continued on next page)

(C) How the Shannon entropy (base 6) in the neuron-level transmitter predictions for each hemilineage correlate, between the hemilineage copy on the right (X axis) and left (Y-axis) hemispheres of the FAFB-FlyWire dataset.

(D) Dot plot shows the mean normalized pairwise NBLAST scores⁴⁷ between neurons expressing the majority transmitter within a hemilineage (for each green dot, each member of pair expresses the main transmitter) and between these neurons and those expressing other transmitters (pink, at least 10 neurons express the other transmitter). Dots represent means taken per hemilineage (183).

(E) Violin plot shows the distribution of neuron-level transmitter prediction confidences for neurons that are in agreement with their hemilineage's transmitter use (green, strictly obey Lacin's law) and those that do not (pink, strictly obey Lacin's law).

(F) Majority unilateral, left-side antennal lobe local neurons of 'ALlv1 dorsal'. Neurons colored by their neuron-level transmitter prediction except for a minority of predicted serotonergic neurons with the most ventral cell bodies, given in purple. These are the most similar to described Krasavitz positive cholinergic local neurons.^{63,189} The upper plot shows neurons predicted to transmit acetylcholine with neurons likely mispredicted to transmit serotonin; they have similar primary neurite and soma positions. We suspect they all should be predicted for acetylcholine. Lower, GABAergic predicted neurons have been added in.

(G) Majority bilateral, left-side antennal lobe local neurons of 'ALlv2'. Neuron meshes colored by neuron-level transmitter prediction. Data were compared using Wilcoxon two-sample tests. Significance values: ns: $p > 0.05$; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$; ****: $p \leq 0.0001$.