eLife

Reviewed Preprint
v1 • October 9, 2025
Not revised

Reviewed Preprint
v2 • May 18, 2026
Revised by authors

For correspondence:
okuno-takuto@tmu.ac.jp

Competing interests: No competing interests declared

Funding: See page 27

Reviewing editor: Casey M Schneider-Mizell, Allen Institute for Brain Science, United States

Functional connectivity, structural connectivity, and inter-individual variability in *Drosophila melanogaster*

Takuto Okuno[a,b], Alexander Woodward[c], Hideyuki Okano[b,d], Junichi Hata[a,b]

[a]Graduate School of Human Health Sciences, Tokyo Metropolitan University, Tokyo, Japan • [b]Laboratory for Marmoset Models of Brain Diseases, RIKEN Center for Brain Science, Wako, Japan • [c]Intelligent Vision Systems Lab, The University of Auckland, Auckland, New Zealand • [d]Keio University Regenerative Medicine Research Center, Tokyo, Japan

eLife Assessment

This paper presents a collection of analyses relating structure and function in the whole-brain Drosophila EM connectome and whole-brain calcium imaging data. The linkage of detailed anatomical structure with population activity is of broad interest in circuit neuroscience in light of increasingly detailed brain maps, but the methods used made the evidence **inadequate** due to a lack of consideration of neurotransmitter identity and technical issues with the network analysis. The conclusions are **useful** for specific network observations, but a more thorough analysis of the anatomical and functional data is needed to support the overall claims.

https://doi.org/10.7554/eLife.107990.2.sa3

Abstract

Clarifying the relationship between structure and function is important for understanding the brain. In *Drosophila melanogaster*, FlyEM and FlyWire electron microscopy-based connectome data and whole-brain calcium imaging data are available. We applied pre-processing methods from fMRI to whole-brain calcium imaging data and comprehensively investigated the optimal parameters. Then, we found that the FC-SC (functional and structural connectivity) correlation decreased linearly with region of interest count, and this trend was the same in flies and humans. We also developed a new, more robust method to quantify the degree of pre– and post-synaptic segregation and investigated this in the fly whole-brain. This revealed that many neurons have unsegregated synapses. We extracted highly unsegregated synapses and compared them with random-extracted null SC matrices. Their FC-SC correlation was significantly higher, indicating that these synapses contribute to FC well. Conversely, highly segregated-synapses showed significantly lower FC-SC correlation and contribute less to FC. Neurons with unsegregated synapses like non-spiking neurons are spread throughout the whole-brain, and they are thought to have a significant influence on FC.

Introduction

Functional connectivity (FC) describes the relationship between functional activity in one brain region and that in another. In previous approaches, two spike trains were analyzed by a cross-correlation histogram [1, 2]; then, this method was applied to positron emission tomography and functional magnetic resonance imaging (fMRI) [3]. Subsequently, the time-series correlation between two regions has become widely used as a measure of undirected FC [4]. Underlying such functional connectivity, a structural connectivity (SC) substrate exists, and its relationship with FC has been investigated using various modalities [5, 6, 7, 8, 9, 10, 11]. In humans, FC-SC correlations between resting-state (rs-) fMRI and diffusion tensor imaging have been explored, and studies have shown high relationships of r=0.66 for 66 regions and r=0.54 for 998 regions [6, 7]. Other

studies have examined the relationship between rs-fMRI and SC based on retrograde tracing from 145 injection sites in the left cortex of marmosets (r=0.379 for 9862 voxels, Ext.Data.Fig.2-1c) [10] and the relationship between calcium imaging and SC based on FlyEM connectome data of *Drosophila melanogaster (D. melanogaster)* [12] (r=0.74 for 37 regions) [9]. Because of the disparity in species, region of interest (ROI) count, and modality, a direct comparison across them becomes difficult. Therefore, in this study, we used electron microscopy (EM)-based FlyEM [12] and FlyWire [13] connectome data, which are currently considered to be the most detailed, and whole-brain calcium imaging data [14] of *D. melanogaster* to comprehensively investigate the relationship between FC and SC across various ROI shapes and counts.

High-quality whole-brain calcium imaging data have recently been published in [14] and, to our knowledge, is the only available open dataset. Therefore, an exact registration process between this type of functional data and compatible structural data has not been sufficiently investigated. Therefore, we applied registration procedures used in fMRI, quantified by FC-SC correlation and FC-SC detection, and investigated their optimality. Specifically, we investigated the application of motion correction and slice timing correction, as used in SPM [15], the application of nuisance factor removal methods [10], and the application of spatial smoothing to improve the accuracy of 2nd-level (group) analysis [16], aiming to calculate FC more accurately. This analysis revealed a relationship between ROI count and FC-SC correlation in the fly brain, and we further explored interspecies differences in FC-SC correlation between humans and marmosets.

Because of the strong relationship between FC and SC, they can be used to investigate inter-individual variability in *D. melanogaster*. FC varied across the eight flies in our dataset, and spatial smoothing can improve group analysis. Furthermore, SC also varied between FlyWire and FlyEM connectome data, and existing comparative studies [17, 18] have observed large differences in cell type classification [18]. Since these two EM connectomes have not been quantitatively compared to FC, we compared them based on the averaged FC of eight flies. FlyEM connectome data has a “confidence” threshold that indicates the certainty of an annotated synapse. Conversely, FlyWire has a “CleftScore” [19] threshold that shows the degree of synaptic detection. These initial thresholds adopted by FlyWire and FlyEM resulted in substantial differences in post-synaptic counts. We comprehensively investigated the impact of these thresholds by using FC-SC correlation and FC-SC detection in the EM connectome data of each fly.

In calcium imaging, the pre-synaptic activity in each region is detected as a calcium signal [20] representing functional activity. In textbook terms, post-synaptic sites on dendrites sense neurotransmitters and graded potentials if they are excitatory; and pre-synapses on axons release neurotransmitters [21, 22]. Then the FC is detected as the synchronization of functional activities. However, the synaptic arrangement on the neurons revealed in the EM connectome shows pre– and post-synapses that are highly intermingled in many neurons (i.e. input and output locations may be intermingled across axon and dendritic arbors); thus, signal transmission may be complex [22, 17, 13, 23]. Nerve cells include non-spiking neurons such as visual neurons, CT1 amacrine cells [24, 25, 26, 27, 28], interneurons in the antennal lobe [29], and APL cells in the mushroom body [30, 31, 32]. In *D. melanogaster*. these cells show localized functional activity taking place on a portion of the dendrite. In CT1 and APL cells, the pre– and post-synapses are highly intermingled (Fig.4a, Ext.Data.Fig.4-2). Therefore, we wanted to investigate whether these unsegregated synaptic neurons exist in the brain and we defined the pre– and post-synapse segregation index (PPSSI) to measure the degree of synaptic segregation for neurons in the whole brain. We also investigated the effect of such unsegregated synapses on FC. Unsegregated synapses are found not only in flies but also in mammals [33, 34, 21, 35, 36, 37], including humans (Ext.Data.Fig.1-2 from [38]). Of note, retinal amacrine cells have extensive reciprocal synapses [33, 34, 39, 40, 41], which provide local feedback inhibition of burst inputs [34]. This microscopic function in the retina, specifically compartmentalization via unsegregated synapses, may be generally applicable to the entire brain. The relationship between unsegregated synapses and functional connectivity could support this idea.

Results

Investigation of the optimal nuisance factor removal method for calcium imaging

First, the optimal pre-processing of the calcium imaging data of *D. melanogaster* was investigated. Whole-brain calcium imaging data of *D. melanogaster* were acquired by Brezovec et al. [14] and are publicly available as an open dataset. We downloaded and used it for our study. The registration method used in fMRI was applied to the calcium imaging data (details in Methods, Ext.Data.Fig.1-1) to register to the FDA template [14] (Fig. 1b). Then, various nuisance factor removal methods (details in Methods) were applied to the ROI time-series to generate the FC matrices of eight flies (Fig. 1a). Three types of ROIs were used in the analysis: ROIs based on the anatomical neuropil, commonly used in *Drosophila* (hemibrain ROI) [42], ROIs based on k-means clustering of the SC matrix of 146,604 voxels in the hemibrain (SCKm), and ROIs based on distance-based clustering of voxels in the hemibrain (DistKm). SCKm can provide ROIs close to the shape of the anatomical neuropil across various ROI counts (Ext.Data.Fig.2-1a), while DistKm can provide ROIs that are completely unrelated to the anatomical neuropil. We investigated whether nuisance factor removal methods remained robust across these various ROI types. Next, a correlation analysis and ROC curve analysis were performed using the SC matrices based on FlyEM and the FC matrices with these ROI types (Fig. 1a). Fig. 1d-f shows the FC-SC correlation results, FC-SC detection results, and the normalized total score of FC-SC detection and correlation (details in Methods), respectively. Since signals are assumed to be transmitted between regions based on SC, when SC is treated as the ground truth, we considered a pipeline with higher similarity of FC-SC and higher detection to be better. After the comparison of the raw (leftmost) and removal methods for various ROI types, the combination of polynomial detrending [43] & tCompCor [44] was found to yield the optimal results. The optimal removal methods were found to differ between fMRI [10] and calcium imaging data, and these results are further described in the Discussion section.

Comparison of the FC-SC correlation between humans and flies

Spatial smoothing is useful for absorbing inter-individual variability and conducting second-level group analysis [10, 16]. Here, we explored the effect of spatial smoothing (of the calcium imaging data) on various ROI counts in DistKm (Fig. 2a) and SCKm (Ext.Data.Fig.2-1a) ROIs. In DistKm, smoothing was highly effective, and the optimal kernel radius of the Gaussian filter was around 23 voxels (52.44 μm) for various ROI counts (Fig. 2b). In SCKm, smoothing was ineffective for ROI count below 100, while the kernel radius above 16 voxels (36.48 μm) was effective for ROI count above 100 (Ext.Data.Fig.2-1b). Increasing the smoothing size improved the correlation and AUC between group-averaged FC and SC, indicating the presence of inter-individual variability in FC. SCKm has a larger variation in volume for each ROI count than DistKm (Fig. 1c, Fig. 2e); as such, the effect of spatial smoothing is likely unstable. Next, we investigated whether the combination of spatial smoothing and Poly-tCompCor was effective. We found that this combination was significantly more effective than smoothing alone for SCKm and DistKm (Fig. 2c, Mann–Whitney U test p<1.67e-4 Bonferroni corrected). Conversely, the combination of a high-pass filter and nuisance factor removal overlapped for the purpose of noise removal, and the effect was degraded (Ext.Data.Fig.2-1d), as found in fMRI [10].

The optimal FC-SC correlation for various ROI counts was determined. Results showed that the correlation decreased proportionally to the logarithm of the ROI count (Fig. 2d). This was also compared between flies and mammals. At approximately 50 ROIs, the difference was nearly 1.08-fold to 1.22-fold (r=0.66 for humans and r=0.71-0.81 for flies [SCKm, DistKm]); and at approximately 1,000 ROIs, the difference was roughly 1.12-fold to 1.24-fold (r=0.54 for humans and r=0.6-0.67 for flies), which was not a considerable difference (not significant by Mann-Whitney U test, p=0.267). The difference was slightly larger (1.21 to 1.5-fold: r=0.378 for marmosets and

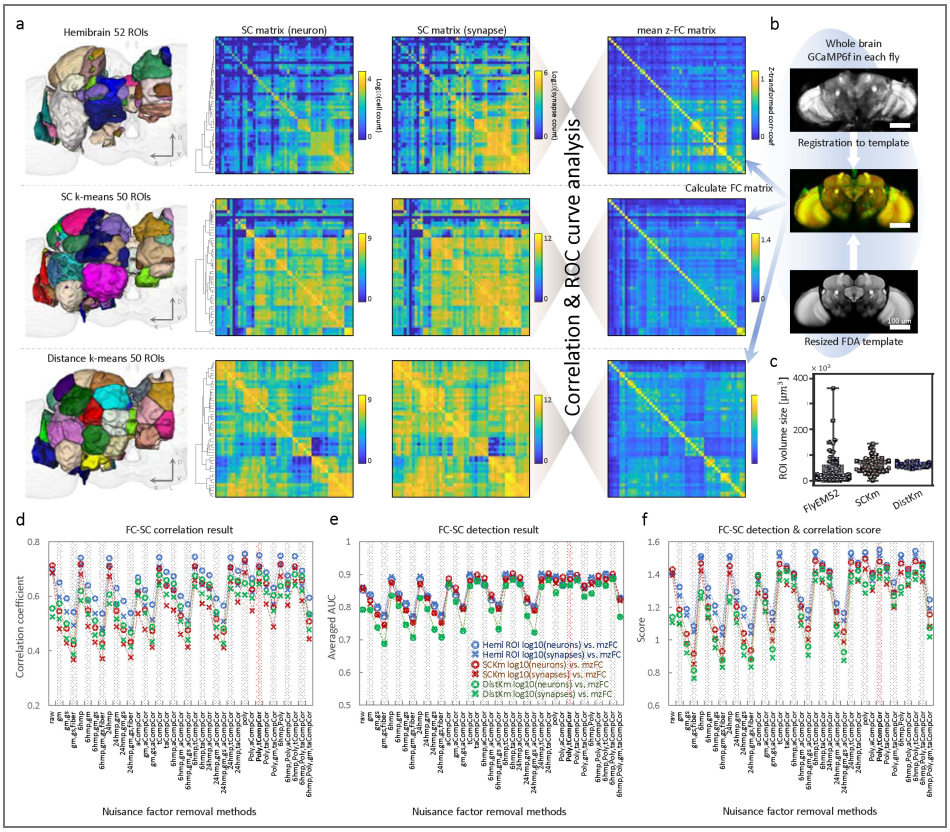


Fig. 1. Comparison between SC and FC, and optimal nuisance factor removal methods.

a, Schematic image of the comparison between SC matrices and FC matrices. (left) Three ROI types: Hemibrain anatomical 52 ROIs, SCKm 50 ROIs, and DistKm 50 ROIs. (center) SC matrices (neuron) and SC matrices (synapse) for each ROI type. (right) FC matrices for an average of 8 flies for each ROI type. **b,** Schematic image of calcium imaging data registration. **c,** Swarm plot of ROI volume for each ROI type. **d,** Results of nuisance factor removal method investigation. The vertical axis shows the correlation coefficient, and the horizontal axis shows the FC-SC correlation for each method. Results are plotted for the three ROI types. **e,** The vertical axis shows the averaged AUC of 100 thresholds, and the horizontal axis is the same as **d**. **f,** The vertical axis shows the FC-SC Detection & Correlation score, and the horizontal axis is the same as **d**.

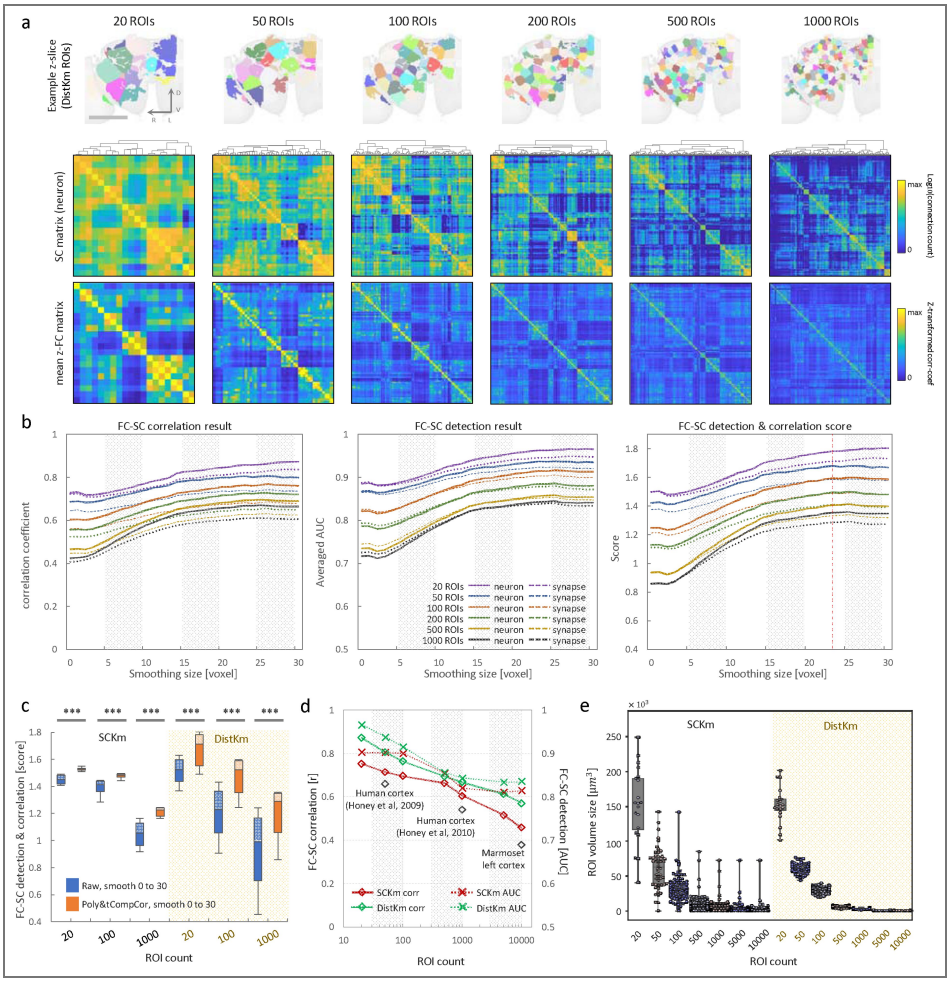


Fig. 2. Investigating spatial smoothing size, and comparison of FC-SC correlation between mammals and flies.

a, (top) Example z-slice of the reduced FDA template with DistKm ROIs from 20 to 1000 ROIs. (middle) SC matrices (neuron) of DistKm ROIs from 20 to 1000 ROIs. (bottom) FC matrices of DistKm ROIs from 20 to 1000 ROIs. **b,** From left to right, FC-SC correlation, FC-SC detection and FC-SC Detection & Correlation score, respectively. The solid line shows the FC vs. SC matrix (neuron), and the dashed line shows FC vs. SC matrix (synapse) at each ROI count. The vertical axis shows the correlation coefficient, averaged AUC and FC-SC Detection & Correlation score, respectively. The horizontal axis shows spatial smoothing size (voxels). **c,** Comparison between raw & smoothing (0 to 30 voxels) vs. Poly-tCompCor & smoothing (0 to 30 voxels) at each ROI count. Non-parametric Mann-Whitney U test was performed and Bonferroni correction was applied to correct for the familywise error rate (*** p<1.67e-4). **d,** Relation between FC-SC correlation and ROI count. (left vertical axis) Solid line shows FC-SC correlation (neuron) of SCKm and DistKm ROI type at each ROI count. (right vertical axis) The dashed line shows FC-SC detection (neuron) of SCKm and DistKm at each ROI count. **e,** Swarm plot of ROI volume of SCKm and DistKm at each ROI count.

r=0.46-0.57 for flies) than the marmoset results at around 10,000 ROIs (Ext.Data.Fig.2-1c). Although SCs for flies and humans were obtained from single individuals, the tracer data is a composite from 52 marmosets [45], which might contribute to the difference with the fly's FC-SC correlation.

Comparison between FlyEM and FlyWire connectome data

We used the FlyEM's SC to compare the averaged FC of eight flies. In this section, we also used the FlyWire connectome data [46] to compare the FC and each SC, and possibly reveal the inter-individual variability of the two EM connectome data. As mentioned earlier, FlyEM has a synaptic "confidence" threshold and contains 21.4 million post-synapses in the hemibrain region at its initial value of 0. Conversely, FlyWire has a 'CleftScore' [19] threshold and 18.8 million post-synapses at its initial value of 50 in hemibrain primary ROIs. Since the number of post-synapses differs at each threshold, we set the confidence threshold of 0.8 as high as possible, thereby increasing the FC-SC correlation (Fig. 3d, ExtFig.3-1b). Then we adjusted the CleftScore to 140, which has a post-synapse number of approximately 11.4 million (Fig. 3a). These thresholds were used for all analyses in this study. When we displayed the maximum Z-projection of the synapse point cloud of post-synapse counts in each voxel, we found that the trends in synapse detection completely varied (Fig. 3b). In each region, FlyEM had more synapses in the mushroom bodies (aL, bL, etc.) and in the major neuropils of the central nervous system (CNS; EB, FB, NO, and PB). In comparison, FlyWire had more synapses in the visual system (LO and ME) and in the periesophageal neuropils (CAN and FLA; Fig. 3c). In the FC-SC correlation of each region, FlyEM had higher values in most regions at 58:5 (Fig. 3e, ExtFig.3-1a). As a result, FlyEM r=0.73 differed from FlyWire r=0.68 at the matrix level (Fig. 3d). Both datasets were from genotype Canton-S-G1×w1118 adult females; FlyEM was five-day-old [12], and FlyWire was seven-day-old [13]. Although brain development may differ depending on the rearing environment, the tendency to detect synapses substantially varied. This may be caused by differences in detection accuracy resulting from the resolution of EM scanning, but not to inter-individual variability.

We next used the FlyWire connectome data to investigate the ratios of neurotransmitter input (ExtFig.3-2a) and output (Fig. 3f) in each region. The highest percentage was acetylcholine (62.8%±14.2%), followed by glutamine (15.5%±9.2%), then GABA (11.4%±8.2%) in neurotransmitter output. The analysis of the relationship between this ratio and FC-SC correlation in each neurotransmitter revealed significant correlations for acetylcholine (r=0.39, p=0.0013) and GABA (r=-0.25, p=0.046) (Fig. 3g). That is, the higher the percentage of excitatory connections, the higher the FC-SC correlation; conversely, the higher the percentage of inhibitory connections, the lower the FC-SC correlation. In humans, FC-SC correlations are lower in the evolutionarily expanded Brodmann area 39,45 and the temporal lobe [8]; one possibility is that a high number of long-range inhibitory connections [47, 48, 49] may result in a low FC-SC correlation.

Unsegregated synapses contribute considerably to functional connectivity

So far, we have compared FC and SC at the macroscopic region level. We next investigated the relationship between the synapses, the source of functional activity (pre-synaptic population dynamics), and FC. We examined the synaptic segregation in FlyWire neurons because many neurons showed pre-and post-synaptic intermingling. Then, we determined how unsegregated and segregated synapses affected the FC. To quantify synaptic segregation, we developed the PPSSI (Fig. 4d). Using DBSCAN [50] with a morphology-based distance threshold of 3 μm (details on Methods), we clustered synapses in each neuron and calculated the degree of segregation for each cluster (cPPSSI). This cPPSSI is also assigned to synapses in each cluster. PPSSI is then the average of the cPPSSI values of all synapses in a neuron. A similar index to our PPSSI, namely, the segregation index, is defined in the literature [22], but we have simplified from this. Since the cPPSSI of each synaptic cluster and the PPSSI of a cell are linearly related, they can be displayed as a common metric. Fig. 5h illustrates this, where the cPPSSI and PPSSI values within each synaptic cluster are uniformly displayed within the [0, 1] range. For clustering, we simply applied DBSCAN by using a morphology-based distance between synaptic nearest arbors. The segregation

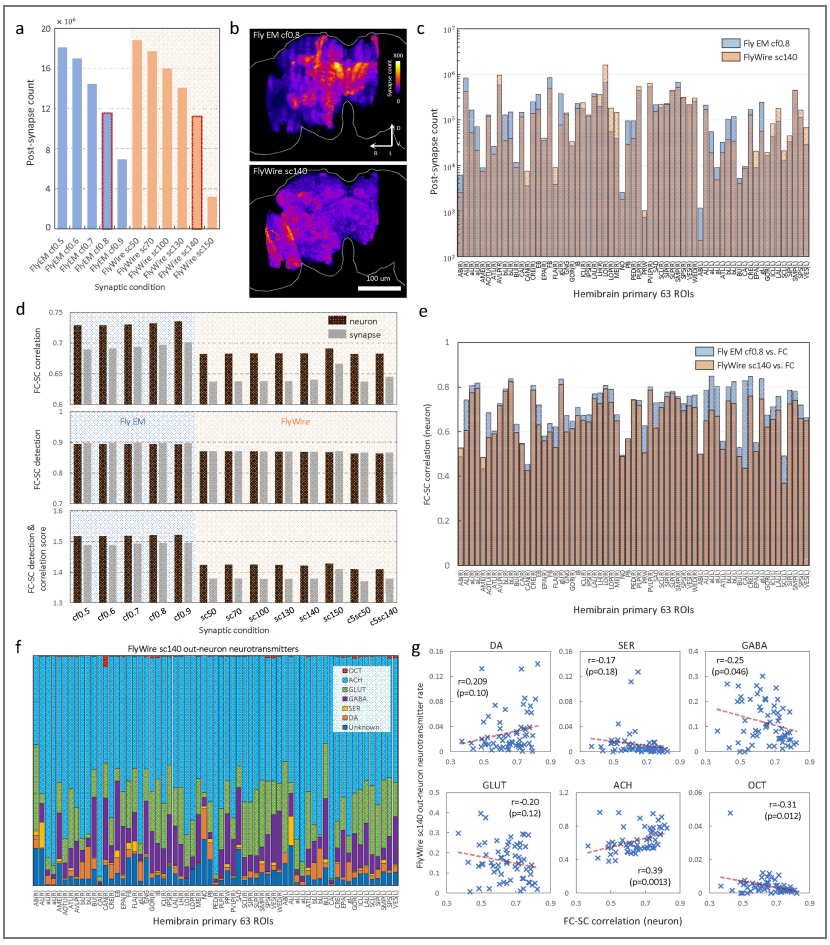


Fig. 3. Comparison between FlyEM and FlyWire connectome data.

a, Bar graph of post-synapse count in hemibrain region from FlyEM and FlyWire connectome data. The vertical axis shows post-synapse count, and the horizontal axis shows confidence thresholds for FlyEM and Cleft score thresholds for FlyWire. **b,** (top) Maximum Z-projection of synapse point cloud of FlyEM (confidence thresholds=0.8 [hereafter, cf0.8]) (bottom) Maximum Z-projection of synapse point cloud of FlyWire (Cleft score threshold=140 [hereafter, sc140]) **c,** Bar graph of post-synapse count in hemibrain primary ROIs. Blue bar shows FlyEM (cf0.8), red bar shows FlyWire (sc140). The vertical axis shows post-synapse count, and the horizontal axis shows hemibrain primary ROIs. **d,** Bar graphs of FC-SC correlation, FC-SC detection, and FC-SC Detection & Correlation score, respectively. The horizontal axis shows confidence thresholds for FlyEM and Cleft score thresholds for FlyWire. **e,** Bar graph of FC-SC correlation in hemibrain primary ROIs. Blue bars show FlyEM (cf0.8) vs FC (neuron), red bars show FlyWire (sc140) vs FC (neuron). The vertical axis shows FC-SC correlation, and the horizontal axis shows hemibrain primary ROIs. **f,** Bar graph of neurotransmitter rate of output neurons in hemibrain primary ROIs based on FlyWire (sc140) connectome data [46]. The horizontal axis shows hemibrain primary ROIs. (DA: dopamine, SER: serotonin, GABA, GLUT: glutamine, ACH: acetylcholine, OCT: octopamine) **g,** Scatter plots of neurotransmitter rate of output neurons vs. FC-SC correlation (neuron) in hemibrain primary ROIs (FlyWire sc140). Each of the six neurotransmitters was compared.

index is impossible to calculate for neurons such as AN and DN, which have only input or output synapses, because it results in calculations of 0 × ∞ or division by 0. Our coefficient simplifies the calculation and addresses these issues. Fig. 4b shows the result of applying PPSSI to 139,255 whole-brain neurons in the FlyWire dataset. This showed a histogram with the most frequent values near the center. When we created an unsegregated ranking (details in Methods) by considering the number of synapses, APL and CT1 cells occupied the top 1–4 positions, indicating that we efficiently extracted unsegregated neurons (Fig. 4a, Ext.Data.Fig.4-3a). APL was also at the top for FlyEM (Ext.Data.Fig.4-1a). Similarly, we established a segregated ranking and successfully extracted neurons with a high degree of segregation (Fig. 4c, Ext.Data.Fig.4-3; Ext.Data.Fig.4-1c and Ext.Data.Fig.4-4 for FlyEM).

Next, synapses with cPPSSI values between 0 and 0.1 were extracted, and the SC matrices (unsegregated synapses) of neurons and synapses were created. Synapses were then randomly sampled to generate 99 null SC matrices (random synapses), centered around the total number of connections (1,122,072) in the SC matrix (unsegregated synapses), forming a normal distribution (Fig. 4e, p=0.82, details in Methods). The FC-SC correlations of the SC matrices (unsegregated synapses) of neurons and synapses were significantly higher than those of the randomly extracted samples (Fig. 4e, p=2.89e-29, 1.36e-25; Ext.Data.Fig.4-1d for FlyEM). Therefore, unsegregated synapses might strongly contribute to FC. A similar analysis was performed on synapses with a cPPSSI of 0.9–1, which surprisingly showed a significantly lower FC-SC correlation (Fig. 4f, p=4.97e-6, 0; Ext.Data.Fig.4-1e for FlyEM). Therefore, segregated synapses may not be able to contribute to strongly to the FC, and these findings are further explored in the Discussion section.

Among unsegregated synapses, reciprocal synapses (Fig. 4g, Ext.Data.Fig.4-1f) have been well studied [33, 34, 39, 40, 41] and their function is considered to be clearer than unsegregated synapses. The shortest distance of such reciprocal synapses was extracted from whole-brain neurons, shown in Fig. 4h (Ext.Data.Fig.4-1g for FlyEM). They are particularly abundant within 2 μm (Ext.Data.Fig.4-2a,c for CT1 and APL cells), and acetylcholine and GABA constitute the majority (Fig. 4h). The comparison of the SC matrix based on reciprocal synapses (<2 μm) with the null SC matrices showed that the FC-SC correlation was significantly higher (Fig. 4I, p=0, 2.7e-32; Ext.Data.Fig.4-1h for FlyEM). Therefore, unsegregated synapses and reciprocal synapses (<2 μm) strongly contributed to FC. Furthermore, since the FC was the average of eight flies, the contribution of unsegregated synapses to FC may be a common phenomenon across individuals.

Region suppression analysis during fly walking

The whole-brain calcium imaging dataset from Brezovec et al. [14] provides fly walking data. Since fMRI pre-processing methods can be applied to calcium imaging data, we investigated whether voxel-level analysis using a general linear model (GLM) [51] could be applied to the fly walking task (Fig. 5a, details in Methods). Fig. 5b shows the GLM analysis result. Similar to the results of Brezovec et al., our findings revealed that LAL, FB, and IPS neuropils showed strong activity; the GNG neuropil, which is highly connected to leg neuropils in the ventral nerve cord [52], also exhibited strong activity. Although Brezovec et al. detected voxels that are strongly correlated with velocity, the GLM extracted voxels associated with walking behavior. Moreover, we found inactivity in the Wedge and AVLP neuropils. We searched for a GABA neuron that would fit with this phenomenon and managed to find one that extended its dendrites widely from the Wedge to AVLP neuropils (Fig. 5e, h, Root ID: 720575940644632087). The synapse point cloud could also fit in this region (Fig. 5c). We called this neuron the Wedge-AVLP GABA neuron (WAGN). The Wedge neuropil receives mechanosensory (wind, sound) input and connects to WPNb neurons (Ext.Data.Fig.5-1a,b), which then connect to ER1 ring neurons that act as the fly's compass [53, 54]. Interestingly, the left WAGN was strongly interconnected with WPNb tier 2/3 (left) [55] and received input from 22 ascending neurons (ANs; Fig. 5d). Therefore, it may be a pathway that provides walking behavior feedback to the compass; however, wind caused by walking was absent

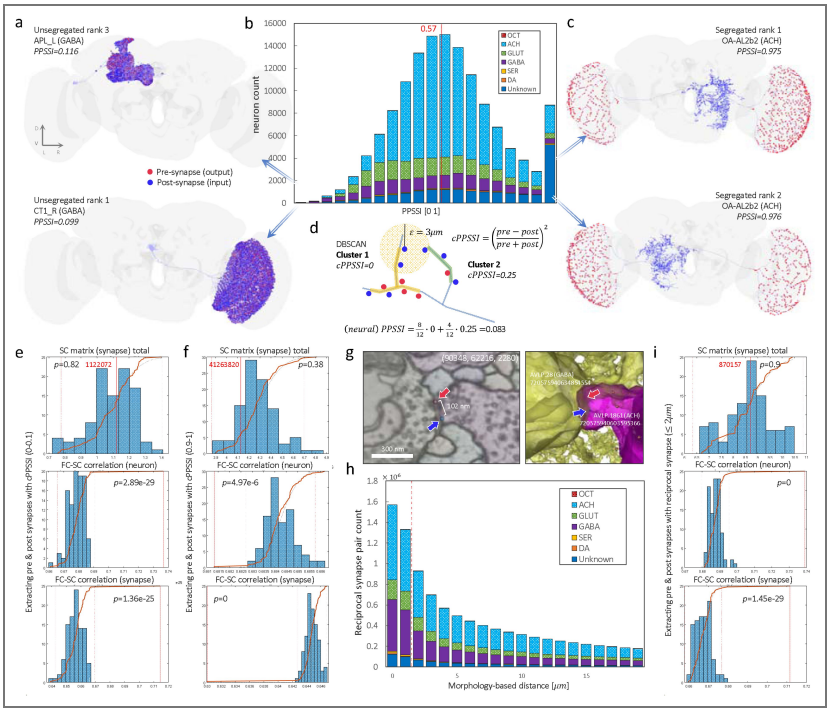


Fig. 4. Quantification of pre– and post-synapse segregation, and the relationship between FC and synapses.

a, Unsegregated first (CT1_R) and third (APL_L) ranked cells in the FlyWire (sc140) connectome data. A red dot shows pre-synapse, a blue dot shows post-synapse. **b,** The PPSSI histogram of all neurons (139,255) in FlyWire (sc140) connectome data. The vertical axis shows neuron count, and the horizontal axis shows PPSSI [0 1]. **c,** Segregated first (OA-AL2b2 cell) and second (OA-AL2b2 cell) ranking in FlyWire (sc140) connectome data. **d,** Schematic image of pre– and post-synapse clustering based on DBSCAN with a morphology-based distance threshold of 3 μm, and calculation of cPPSSI and PPSSI. **e,** Histogram of null SC matrices (blue bar) and extracted SC matrix with cPPSSI (0-0.1) (red solid line). Black dotted line shows cumulative distribution function of the normal distribution, brown solid line shows cumulative distribution function of null & extracted SC matrices, and red dotted line shows Bonferroni-corrected p<0.05 threshold. Top shows SC matrix (synapse) total, middle shows FC-SC correlation (neuron), and bottom shows FC-SC correlation (synapse). **f,** Histogram of null SC matrices and extracted SC matrix with cPPSSI (0.9-1). **g,** Example of a reciprocal-synapse in the EM image (left) and in 3D (right). **h,** Histogram of reciprocal-synaptic minimum distances in FlyWire (sc140) connectome data. The vertical axis shows reciprocal-synapse pair count, and the horizontal axis shows distance [μm]. **i,** Histogram of null SC matrices and extracted SC matrix for reciprocal-synapses (< 2μm).

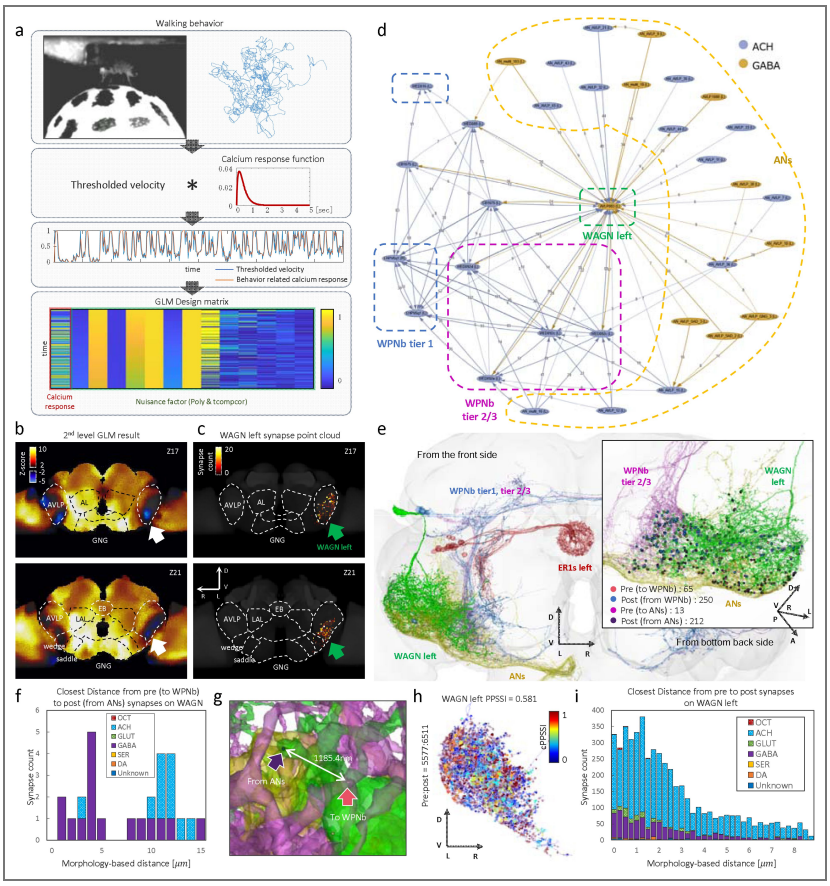


Fig. 5. General Linear Model analysis result and Wedge-AVLP GABA neuron.

a, Schematic image of GLM analysis. From the top, experimental image and walking trajectory, calcium response function, thresholded velocity, and GLM design matrix, respectively. **b,** 2nd level (group) GLM analysis result from 9 flies. Top shows Z=17 slice of the resized FDA template, and the bottom shows the Z=21 slice. The white arrow indicates the suppressed region during the fly walking task. **c,** Synapse point cloud of the left WAGN in FlyWire (sc140) connectome data. Top shows Z=17 slice of the resized FDA template, and bottom shows the Z=21 slice. **d,** Network graph of the left WAGN, ANs, WPNb tier 1 and tier 2/3 neurons from FlyWire codex (sc50, connection threshold>5). Blue is acetylcholine, ocher is GABA neuron. The small number on the edges shows synapse count. **e,** 3D image of the left WAGN, ANs, WPNb tier 1, tier 2/3 and left ER1 neurons, frontal view. Inside the black square shows the left WAGN, ANs and WPNb tier 2/3 from the back side. **f,** Histogram of closest distance from pre-synapse (to WPNb) to post-synapse (from ANs) in the left WAGN. Horizontal axis shows morphology-based distance [μm]. Neurotransmitter indicates input ANs to post-synapses. **g,** Example 3D image of pre-synapse (to WPNb) to post-synapse (from ANs). **h,** All pre– and post-synapses of the left WAGN from the front. Color indicates cPPSSI in each synapse. **i,** Histogram of closest distance from pre-synapse to post-synapse in the left WAGN. The horizontal axis shows morphology-based distance [μm]. Neurotransmitter indicates input neurons to post-synapses.

in the experimental setup (the fly was fixed and walked on a treadmill), and may relate to the strong regional suppression that was observed. As such, further experimentation is needed to pursue this topic.

The left WAGN resembles a hub neuron; however, it contains many unsegregated synapses (PPSSI=0.581). Segregated and unsegregated synapses are mixed throughout the cell, which is reflected in the intermediate PPSSI values (Fig. 5h). It received input from ANs (212 synapses) mainly on the ventral side, and outputs to WPNb (65 synapses), widely from the medial to the lateral side (Fig. 5e). A histogram of the shortest morphology-based distances between pre-synapses to WPNb and post-synapses from ANs is shown in Fig. 5f, and the closest distance was 1185.4 nm (Fig. 5g). Many acetylcholine inputs (disynaptic inhibition [28]) were closer than this value (Fig. 5i), and many GABA inputs from ANs (Fig. 5f) further suppress such disynaptic inhibition. AVLP ventral suppression (Fig. 5b white arrows) may be the suppression from these ANs. Thus, a highly delicate and localized control is possibly exercised from WAGN to WPNb tier 2/3 (left).

Discussion

Although nuisance factor removal methods such as global mean (GM) and aCompCor [44] are optimal for fMRI [10], calcium imaging showed different results. GM is the average of all voxels, and the regression of this signal is useful in fMRI because noise from the equipment and coils affects the entire voxel image. Conversely, in two-photon microscopy, noise does not affect all voxels because of the excitation of specific molecules by laser absorption, and GM does not show a positive effect. aCompCor extracts the component time-series from the voxel time-series of white matter and cerebrospinal fluid (CSF) via PCA and then performs a regression. In flies, the CSF had no ROI mask, but fibers had ROIs; therefore, fibers were used instead of the white matter. However, many synapses may exist around these fibers, and aCompCor was not as effective as in humans. Conversely, polynomial detrending [43] was highly effective (Fig. 1d-f), likely because of linear thermal noise and other factors caused by the continuous excitation of molecules (for 30 min). tCompCor [44] is the component time-series extracted by PCA from voxels with the top 1000 standard deviations among all voxels. It potentially captures minute fluctuations caused by fly movements and effectively removes them.

The SC matrix from a cPPSSI of 0.9–1 synapses resulted in a significantly lower FC-SC correlation than the SC matrix from randomly selected synapses (Fig. 4f for FlyWire, Ext.Data.Fig.4-1e for FlyEM). This result was surprising, but we propose several possible reasons. Neurons with highly segregated synapses include ascending neurons (ANs) and descending neurons (DNs) (Ext.Data.Fig.4-3b). They also have a small number of unsegregated synapses (i.e. Sep rank11: PPSSI=0.963). They are thought to be specialized for input or output in the CNS regions and may show large differences between SC and FC. Another reason is that these neurons may have widespread logical OR inputs and outputs. In Fig. 4c, an action potential at one point in the input synapse (blue) is likely transmitted to all red synapses. Such an OR circuit-like disproportionate input–output possibly causes a mismatch with the FC, because the functional activity in the output region originates from activity at a single location. These neurons are likely those whose FC are difficult to capture. Conversely, neurons with highly unsegregated synapses may capture the inputs of neighboring synapses, such as forming a logical AND circuit, correlating well with the output regions.

The FC-SC correlation was highly dependent on ROI count in both flies and humans (Fig. 2d). Through optimal nuisance factor removal and smoothing, we achieved r=0.87 at 20 ROIs. Furthermore, by applying Spearman correlation instead of Pearson, we obtained r=0.928. Thus, for a reasonably small ROI count the correlation was very high; however, for 10,000 ROIs, the correlation decreased to r=0.57. This phenomenon could be attributed to the presence of highly segregated synapses. Such neurons are widespread in the CNS (Fig. 4b, Ext.Data.Fig.4-2f) and are thought to provide logical OR circuits. When ROI volume was large (when the ROI count was small), a single ROI could cover the distributed input synapses. However, when a ROI was divided, input synapses were also divided across many ROIs, and the mismatch between SC and FC would

increase. Other possibilities include the presence of gap junctions and FC due to neuropeptides [56], but they are outside of the scope of the present study and their relationship with ROI count is unclear.

In this study, we applied fMRI techniques directly to calcium imaging in flies and found many similarities between humans and flies (Fig. 2d). The calcium signal in the fly brain represents pre-synaptic population dynamics and the BOLD (Blood Oxygenation Level Dependent) signal in the mammalian brain is also thought to contain the same type of population dynamics. It has been estimated that 43% of the gray matter's energy consumption is due to synaptic activity [57]. Local oscillations of synaptic activity are detected in the BOLD signal in fMRI through neurovascular coupling [58]. However, the BOLD signal contains biological noise such as cardiac or respiratory noise, etc. We used nuisance factor removal methods to remove these noise sources to approximate pre-synaptic population dynamics and calculate inter-regional FC. Despite the differences in modalities, this commonality in FC is likely responsible for the observed trends in high FC-SC correlation. Studies on FC-SC correlation have also been conducted in C. elegans [59, 56], but results showed a low correlation. In C. elegans, the FC of calcium dynamics in the soma was compared with the SC of the whole cell. Neurons in C. elegans are non-spiking, and calcium dynamics have been shown to differ between compartments in a neuron [60]. While fly and human studies compared the FC and SC of pre-synaptic population dynamics in brain regions, C. elegans studies compared the SC of cells (not regions) with the FC of dynamics in the cell body (in terms of compartments). This discrepancy is thought to explain the differing results.

Our study revealed that unsegregated synapses [28, 33, 34, 39, 40] are widespread in neurons throughout the brain (Fig. 4b, Ext.Data.Fig.4-2f). They are evenly distributed across cells of various neurotransmitters, and some may be non-spiking [24, 25, 26, 27, 28, 29]. Retinal amacrine cells have extensive reciprocal synapses [33, 34, 39, 40, 41], which provide local feedback inhibition of burst inputs to improve signal dynamic range [34]. Grimes et al. [34] noted "The retina is a beautiful example of a neural network that optimizes signal processing capacity while minimizing cellular cost." To maintain signal dynamic range, A17 amacrine cells must optimize processing units and wiring costs. If one unit equaled one cell, an enormous number of cell bodies would be required, reducing the number of processing units per volume and increasing the energy cost during development. To optimize this, Grimes et al. proposed arranging units capable of parallel processing within a single cell, thereby maximizing processing unit count and wiring cost per volume. Signal bursts might also occur in the CNS, in which case CNS neurons would also require dynamic range adjustment. The idea of optimizing processing unit count per volume is highly compelling and is thought to apply not only to the retina but throughout the entire brain. If we elevate such a "design concept," the fly can be a powerful model organism that may also be useful for mammalian studies.

Materials and Methods

Preprocessing of *D. melanogaster* calcium imaging data Whole-brain calcium imaging data of *D. melanogaster* were acquired by Brezovec et al. [14] and downloaded from DANDI (*https://dandiarchive.org/dandiset/000727/0.240106.0043*). 4D NIfTI image data (256×128×49 voxels, 3384 frames) were extracted from a NWB file by an in-house MATLAB script. The FDA template, made from in-vivo calcium imaging, was resized (256×128×49 voxels, 2.45×2.28×3.72μm for a voxel) and used as a template because the original thresholded FDA template [61] was extremely large (1652×768×479 voxels), and the calcium imaging file was large (more than 40 GB); therefore, a smaller template with the imaging data was preferably used. Preprocessing and image registration were performed using Statistical Parametric Mapping (SPM12) [15] and ANTs [62] (Ext.Data.Fig.1-1a). Motion and slice timing correction were performed using SPM12, and the averaged NIfTI image was registered by ANTs to the resized FDA template. Lastly, 4D NIfTI images were transformed using the transform information. The data contained n=9 flies, but 8 flies were used for the ROI-based analysis because of missing time-series in some ROIs in one fly. Nine flies were used in the GLM analysis.

Combinations of the following nuisance factor removal methods were investigated: global mean (average signal across all voxels), global signal (average signal across all brain voxels), mean white matter (average signal across fiber voxels), 6HMP (head motion parameters), 24HMP, aCompCor, tCompCor [44], and polynomial detrending (Poly) [43]. These methods were implemented by an in-house MATLAB script [63]. A full width at half maximum (FWHM) from 1 voxel (2.28 μm) to 30 voxels (68.4 μm), at a 1 voxel step, was investigated to determine the optimal spatial smoothing size for calcium imaging data. Temporal high-pass filters were also examined from 0.1 Hz to 0.001 Hz.

Transformation of EM connectome data into the resized FDA template

In the FlyWire case, connectome data (v783), such as neurons and synapses, were downloaded from FlyWire Codex (*https://codex.flywire.ai/*) and Zenodo (*https://zenodo.org/records/10676866*) [13, 64, 46]. Detailed synapse data were saved in an Apache feather format, which was then converted to a CSV file. This dataset contains 139,255 neurons and 54.5M (pair of pre– and post-) synapses as described in Dorkenwald et al. [13], with 18.8M post-synapses in the regions corresponding to the hemibrain primary ROIs. Synapse points were transformed from the FlyWire space to a JRC2018 female template [65] by the Python navis-flybrains package [65]. Since FlyWire coordinates are left and right reversed [13], they were flipped on the JRC2018 female template, and these points were transformed to the resized FDA template by ANTs. The use of these two templates facilitated precise data alignment across different modalities.

In the FlyEM case, the latest connectome data (v1.2) were downloaded from Janelia's DVID website (*https://dvid.io/blog/release-v1.2/*). This data is provided in the format defined in *https://neuprint.janelia.org/public/neuprintuserguide.pdf*, and we extracted the neurons and synapses from it. The entire segmentation body is 28M segmentations, containing 99,644 Traced neurons. In addition, there were 73M (pre– or post-alone) synapses, 87M *synapseSets* records and 128M *synapseSet-to-Synapse* records. When we extracted post-synapses between Traced neurons, the total number was 21.4M (i.e., connections from Traced neurons to other body fragments, like Orphans, were excluded). Like those in the FlyWire case, synapse points were transformed from the FlyEM space to a JRC2018 female template by the Python navis-flybrains package. These points were transformed to a resized FDA template by ANTs.

Generation of regions of interest for the hemibrain

The hemibrain neuropil atlas (114 ROIs, including 63 primary and 51 sub-regions, aligned with the FlyEM space) was downloaded from Virtual Fly Brain (*https://www.virtualflybrain.org/*). ROI files were transformed from the FlyEM space to a JRC2018 female template by Java template-building scripts (*https://github.com/saalfeldlab/template-building*) [65]. They were transformed to a resized FDA template by ANTs. In Fig. 3, 63 primary ROIs mentioned above were used. The 52 ROIs in Fig. 1 were selected from the primary ROIs, and approximately 50 ROIs (those used by Turner et al. [9]) were selected. ROIs were segmented by SCKm (Fig. 1 and Fig. 2) via the k-means clustering of the SC matrix of 146,604 voxels in the hemibrain. They were also segmented by DistKm (Fig. 1 and Fig. 2) via the k-means clustering of the distance of 146,604 voxels in the hemibrain.

The SC matrix was calculated according to synapse locations and ROIs. Since synapse data formats differ between FlyEM and FlyWire, they were standardized by using post-synapses within ROIs. The input and output neurons at each ROI were extracted from the paired pre– and post-synapse data. Then the input and output neurons and their synapses that were common between ROI pairs were counted to create SC matrices for neurons and synapses. The FC matrix was calculated using the pre-processed voxel time-series and ROIs. The intensity of the voxels in the ROI was averaged to obtain the average intensity time-series of the ROI. Then, Pearson correlation was calculated between ROI pairs to generate the FC matrix. The correlation coefficients were z-transformed, calculated for eight flies, and then averaged to obtain the final FC matrix.

FC-SC detection and correlation

FC-SC correlation was calculated by taking the Pearson correlation between the logarithm of the SC matrix and the FC matrix, as done in other studies [6, 7, 10]. In the SC matrix, elements with zero connections become *-Inf* when logarithmically converted, so we substituted zero for them. Also, the diagonal elements of the FC matrix become *Inf* when z-transformed, so we replaced them with the maximum value of the matrix. For one FC matrix, we calculated its correlation with both the neuron SC matrix and the synapse SC matrix. If there is a neuron from post-synapse to pre-synapse between two regions, that neuron is counted as one, and the number of synapses in the connected region is used as the synapse SC matrix. FC-SC detection was calculated by ROC (Receiver Operating Characteristic) curve analysis. This gives the detection accuracy of the FC matrix when the SC matrix is used as ground truth. Thresholds were taken from 0% to 99% by 1% steps of all elements of the SC matrix, and 100 AUCs (Area Under the Curve) were calculated. The average of these AUCs is called the FC-SC detection. FC-SC correlation indicates the similarity of the two matrices, while FC-SC detection indicates the detection accuracy. Although the digit number of connections differs between neurons and synapses, FC-SC detection can absorb such differences. The FC-SC Detection & Correlation score was calculated as follows:

$$FC-SC_{\text{det\&corr score}} = |FC-SC_{\text{correlation}}| + 2 \times |FC-SC_{\text{detection}} - 0.5| \tag{1}$$

The FC-SC Detection & Correlation score is in the range of [0 2], where the larger the score, the higher the correlation and detection accuracy.

Pre– and Post-synapse Segregation Index (PPSSI)

The PPSSI (Fig. 4d) was used in DBSCAN [50] with a distance threshold of 3 μm to generate pre– and post-synaptic clusters. This distance is not a straight-line distance, but rather a morphology-based distance between synapses along the nearest arbors (Ext.Data.Fig.4-1i). The threshold of 3 μm was empirically defined by several criteria. Firstly, CT1 and APL cells have the shortest distance of reciprocal synapses, within 2 μm (Ext.Data.Fig.4-2a,c). In zebrafish, reciprocal synapses were defined within 2.5 μm [41]. Schneider-Mizell et al. [22] defined a twig as a branch within 3 μm from the backbone of the cytoskeleton, and 80% of synaptic inputs in motor neurons are in the twig. In addition, 97% of pre-synaptic sites are located at less than 3 μm from the mitochondria. Therefore, 3 μm was deemed appropriate for the synapse cluster distance threshold. Once the synapse clusters were determined by DBSCAN, the degree of synaptic segregation for each cluster was calculated using the following formula:

$$c\text{PPSSI}_i = \left(\frac{\text{pre}_i - \text{post}_i}{\text{pre}_i + \text{post}_i}\right)^2, \tag{2}$$

where pre_i and $post_i$ are the pre– and post-synapse counts in cluster *i*, respectively. The cPPSSI of each cluster (each synapse) is in the range of [0 1]. The PPSSI of each neuron was calculated as follows:

$$PPSSI = \sum_{i=1}^{N} \frac{(\text{pre}_i + \text{post}_i)}{W} \cdot \text{cPPSSI}_i \tag{3}$$

where *N* is the cluster number, and *w* is the total synapse count for a neuron. This cPPSSI is also assigned to synapses in each cluster. The PPSSI is then the average of the cPPSSI values of all synapses in a neuron.

The unsegregated ranking was calculated based on the PPSSI and total synapse count of neuron k:

$$\text{unsegregated score}_k = \log_{10}(W_k) \cdot (1 - \text{PPSSI}_k) \tag{4}$$

where $PPssI_k$ is the PPSSI of neuron k and w_k is the total synapse count of neuron k. This allows us to extract neurons with many inputs and outputs and highly unsegregated synapses. The segregated ranking was also calculated in the same way, using the following score:

$$\text{segregated score}_k = \log_{10}(W_k) \cdot \text{PPSSI}_k \quad (5)$$

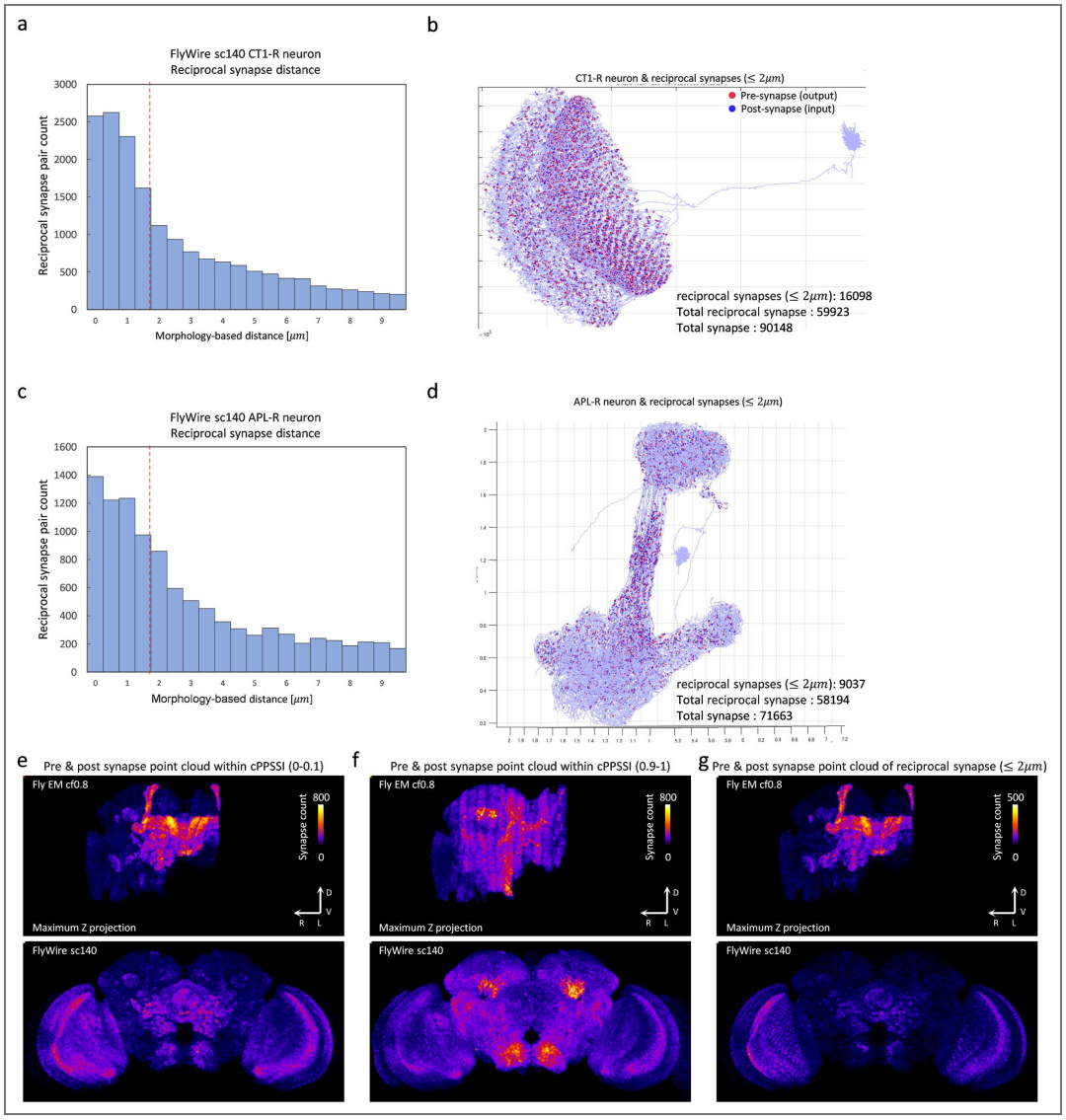
GLM Analysis for the fly walking task

GLM analysis (Fig. 5a) was performed to investigate the whole-brain functional activity during the fly walking task. Movement velocity data were extracted from an NWB file by an in-house MATLAB script. Walking was defined as 1σ (34%) or more of the movement velocity and divided into walking and other. Less than 10% of movement velocity was zero cut. The canonical calcium response function (CRF) was created by the gamma probability density function (a=1.7, b=0.25) in MATLAB based on the GCaMP6f response [66]. The walking-task time-series was transformed into a calcium response time-series by convolution with the CRF. GLM design matrices were created using the calcium response time-series and nuisance factors (polynomial detrending and tCompCor). In whole-brain voxels, smoothing was applied with a kernel radius of 8 voxels (18.24 μm). Then, a Tukey taper (taper size = 8) was used for GLM pre-whitening [67]. A mixed-effects model [16] was used for group analysis, and the t-value of each voxel was calculated via the second-level analysis of OLS regression with a Tukey taper. These analyses were implemented using in-house MATLAB scripts [68].

Statistical Analysis

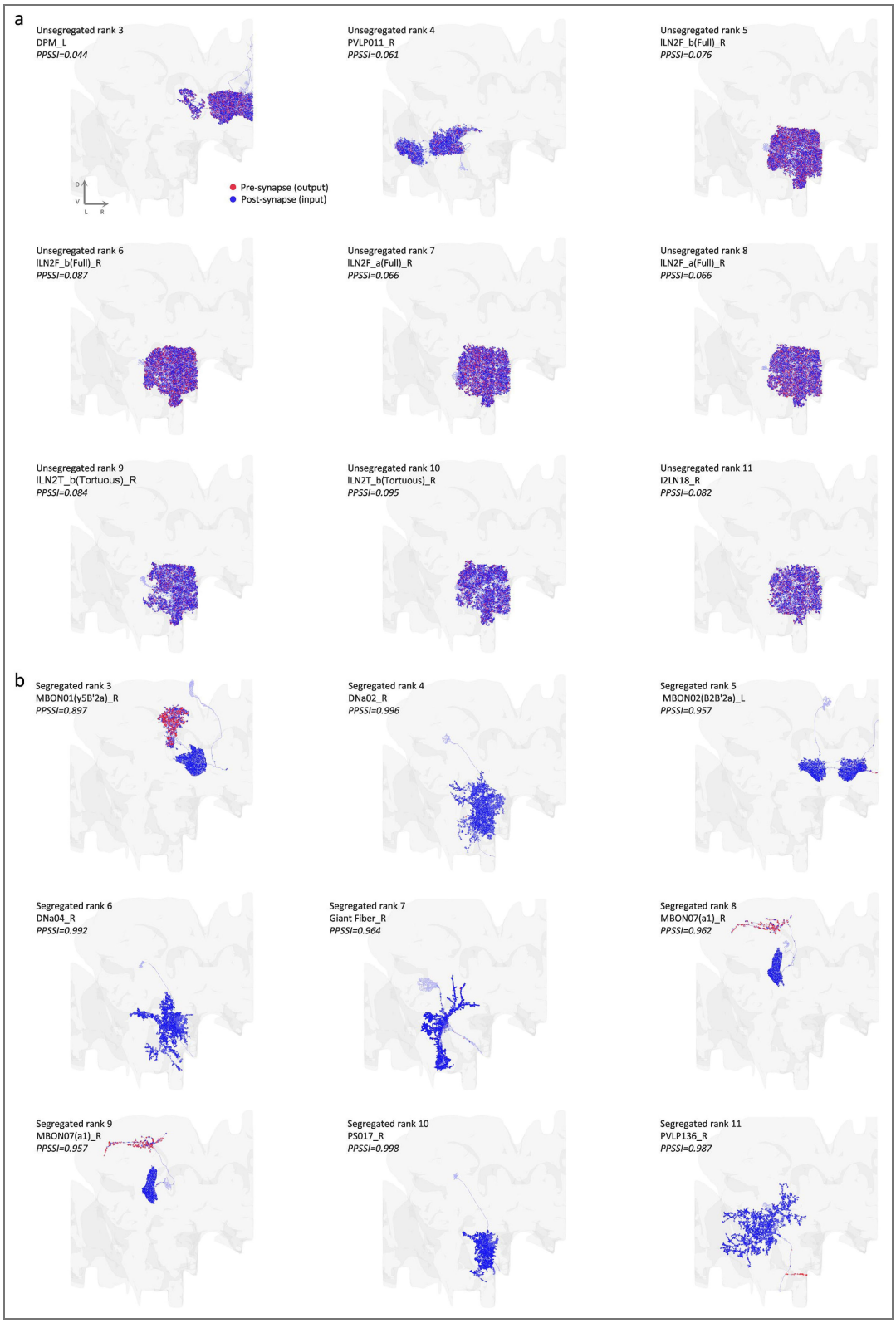
A Mann–Whitney U test was used to examine whether the combination of spatial smoothing and nuisance factor removal was effective. Statistical significance was set at $p<0.05$, and Bonferroni correction was applied to correct the familywise error (FWE) rate. A permutation test was performed to analyze whether specific synapses (segregated, unsegregated, or reciprocal synapses) affect SC-FC correlation. The SC matrix (containing specific synapses) in the 63 ROIs in the hemibrain was generated, and its SC-FC correlation was calculated. Then, 99 null SC matrices (containing random synapses) were generated so that they would be normally distributed around the total number of connections in the SC matrix. Because the sparsity rate of the matrix also affects SC-FC correlation, we generated null SC matrices to match the sparsity rate of the target SC matrix (Ext. Data Fig.4-5). Since the SC-FC correlations of the null SC matrices also roughly followed a normal distribution, the T-value of the SC-FC correlation of the SC matrix was calculated using the cumulative distribution function of the normal distribution. For GLM analysis, a mixed-effects model was used for group analysis.

Extended Data Figures



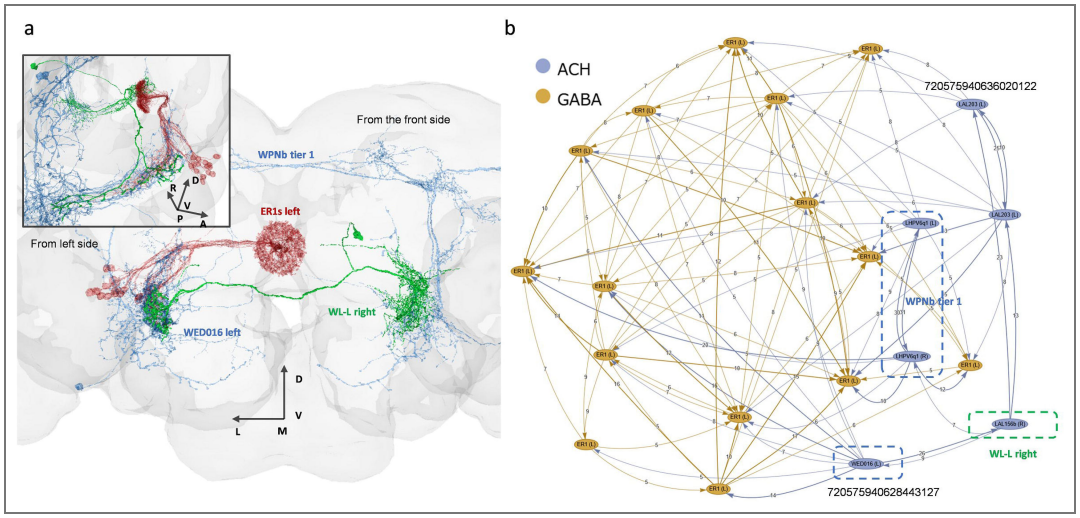
Extended Data Fig.1-1. Schematic image of registration process and connectome data transformation in *Drosophila melanogaster*. **a**, Schematic image of calcium image registration to the resized FDA template. Motion and slice timing correction were conducted using SPM12 [15], and the averaged NIfTI image was registered by ANTs [62] to the resized FDA template. Lastly, 4D NIfTI images were transformed using the transformation information. **b,** Schematic image of synapse points transformation from FlyWire space to the resized FDA template. **c,** Schematic image of synapse points transformation from FlyEM space to resized FDA template.

Extended Data Fig.1-2. Pre– and Post-synapses of pyramidal cell in human temporal cortex layer 3-4. a, (left) Input to (post-synapses) a pyramidal cell in the human temporal cortex, from EM connectome data [38]. (center insert) A slice of human temporal cortex and target pyramidal cell. (right) Output from (pre-synapses) a pyramidal cell in the human temporal cortex, from EM connectome data. Blue dots show post-synaptic sites, Red dots show pre-synaptic sites.

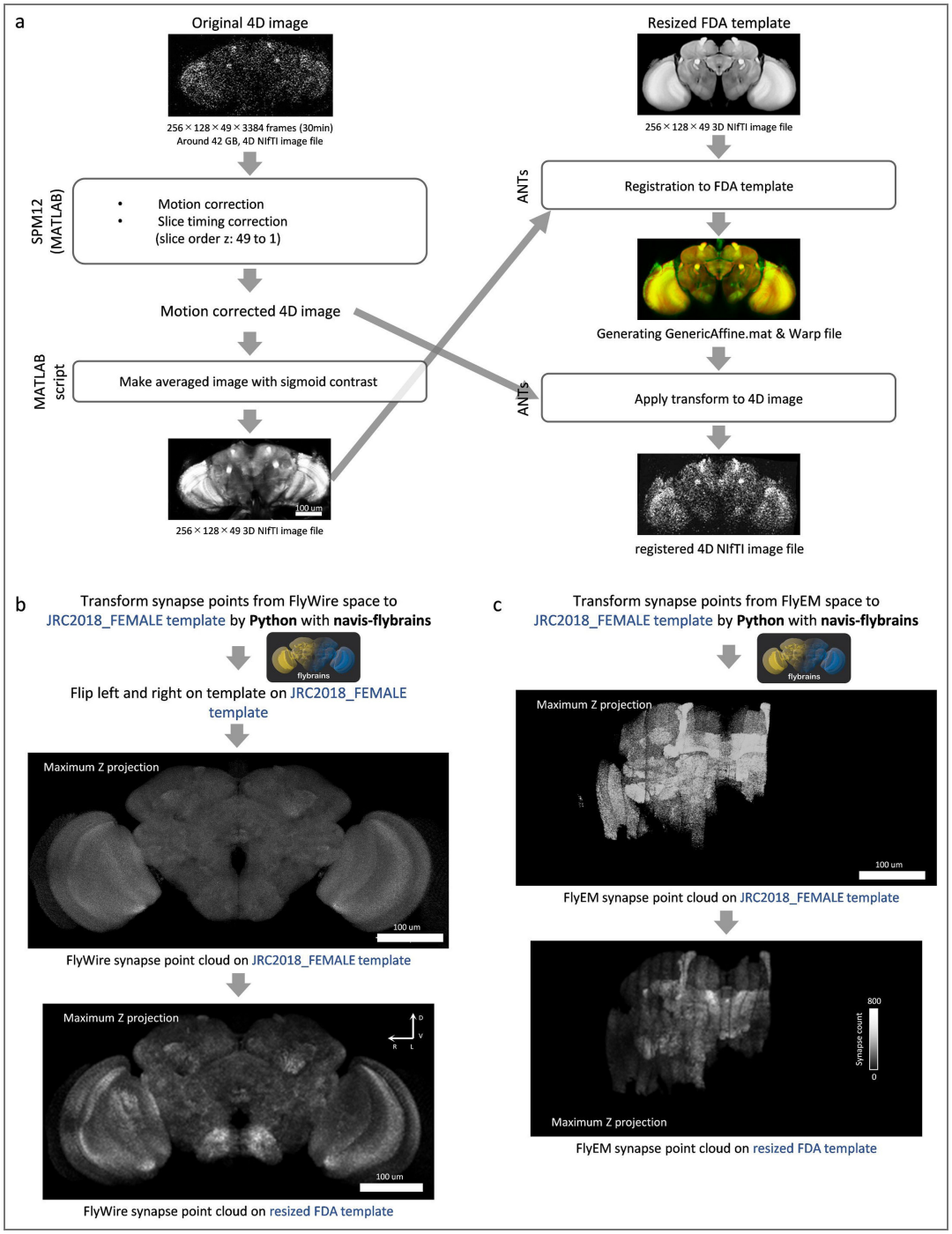


Extended Data Fig.2-1. Investigating spatial smoothing size, and high-pass filtering. a, (top) Example z-slice of reduced FDA template with SCKm ROIs from 20 to 1000 ROIs. (middle) SC matrices (neuron) of SCKm ROIs from 20 to 1000 ROIs. (bottom) FC matrices of SCKm ROIs from 20 to 1000 ROIs. **b,** From left to right, FC-SC correlation, FC-SC detection, and FC-SC Detection & Correlation score, respectively. Solid lines shows FC vs. SC matrix (neuron), and dashed lines shows FC vs. SC matrix (synapse) for each ROI count. The vertical axis shows the correlation coefficient, averaged AUC, and FC-SC Detection & Correlation score, respectively. The horizontal axis shows spatial smoothing size (voxels). **c,** Scatter plot of FC-SC correlation in the marmoset left cortex [10] (145 injection×9862

voxels, r=0.379). **d,** Result of high-pass filtering and nuisance factor removal combination of DistKm ROIs. From left to right, FC-SC correlation, FC-SC detection, and FC-SC Detection & Correlation score, respectively. The high-pass filter degraded the FC-SC relationship.

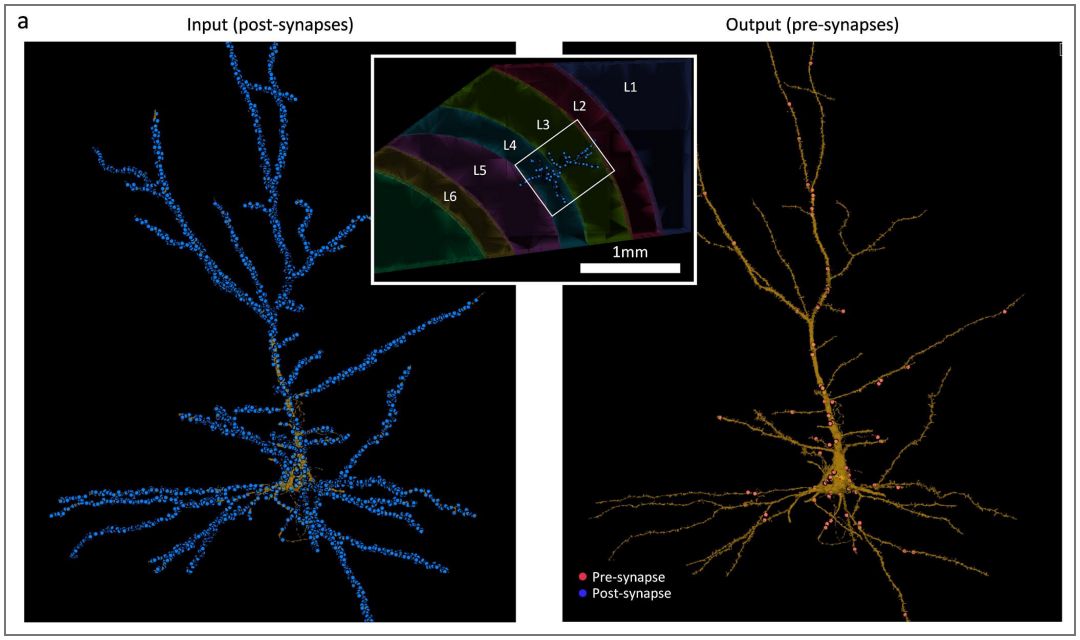


Extended Data Fig.2-2. Investigating gaussian resampled SC vs. log-scaled SC. a, Relation between FC-SC correlation and ROI count. Solid line shows FC-SC correlation (neuron) of SCKm and DistKm ROI type with gaussian resampled SC. Dashed line shows FC-SC correlation with log-scaled SC. **b**, Sparsity rate by ROI count and ROI type. **c**, Example scatter plot of FC-SC correlation for a CmKm ROI count of 10000 with gaussian resampled SC. **d**, Scatter plot of FC-SC correlation in the marmoset left cortex with gaussian resampled SC.

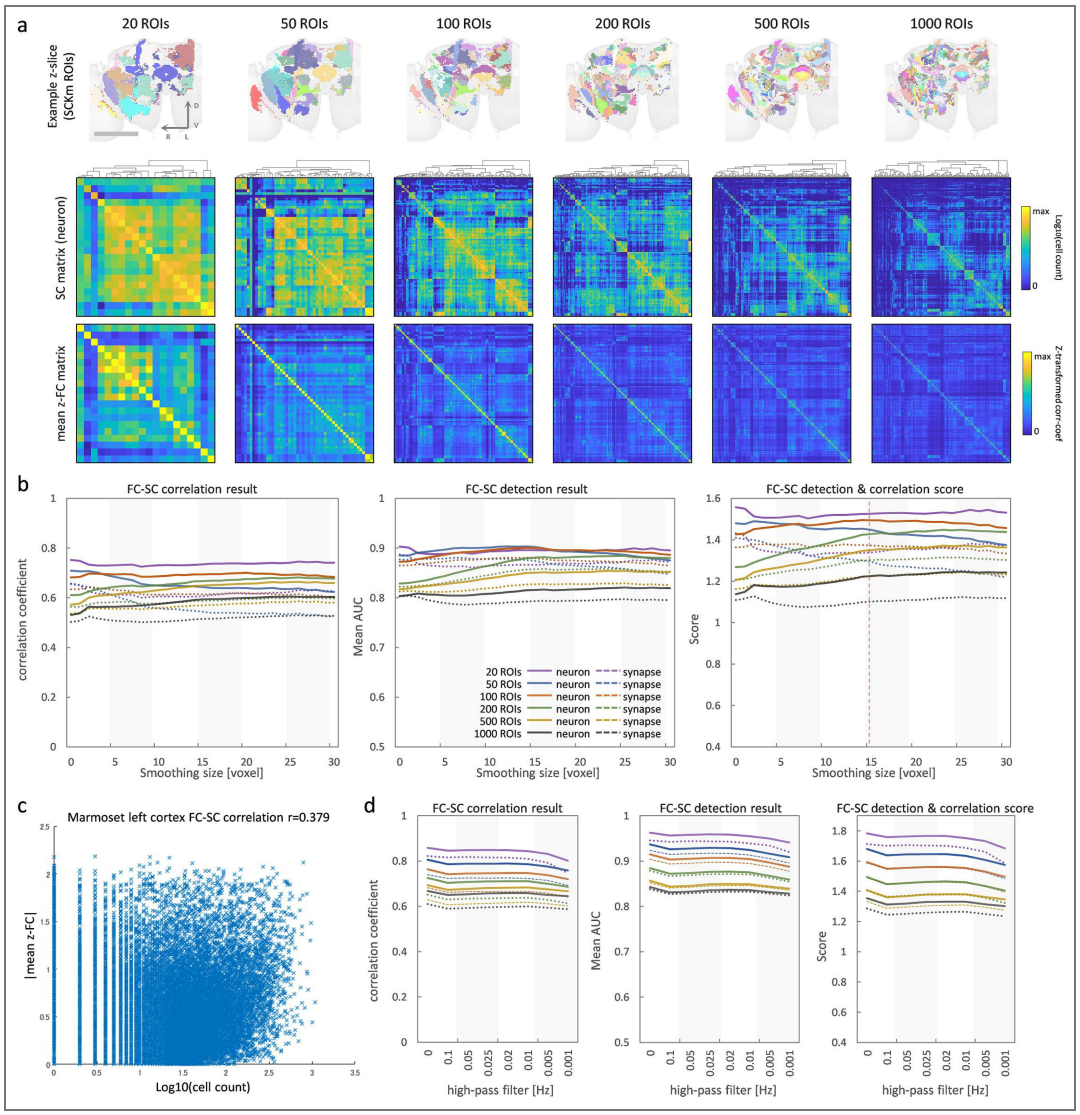


Extended Data Fig.3-1. Comparison between FlyEM and FlyWire connectome data. a, (left) A scatter plot of regional post-synapse count between FlyEM (cf0.8) and FlyWire (sc140) in the hemibrain primary 63 ROIs. The Pearson correlation is not so similar, r=0.778, but the Spearman correlation is much more similar, r=0.9. (right) A scatter plot of log10 of the regional post-synapse count between FlyEM (cf0.8) and FlyWire (sc140) in the hemibrain primary ROIs. Pearson correlation shows r=0.919, a higher similarity. **b,** From left to right, FC-SC correlation, FC-SC detection, and FC-SC Detection & Correlation score, respectively. Solid lines shows FC vs. SC matrix (neuron), and dashed lines shows FC vs. SC matrix (synapse) of FlyEM (cf0.5-0.9) and FlyWire (sc50-140) corresponding Fig.3d. The vertical axis shows the correlation coefficient, averaged AUC, and FC-SC Detection & Correlation score, respectively. The horizontal axis shows spatial smoothing & nuisance factor removal combinations. **c,** From left to right, SC matrix (neuron) (FlyEM cf0.8), SC matrix (neuron) (FlyWire sc140) and FC

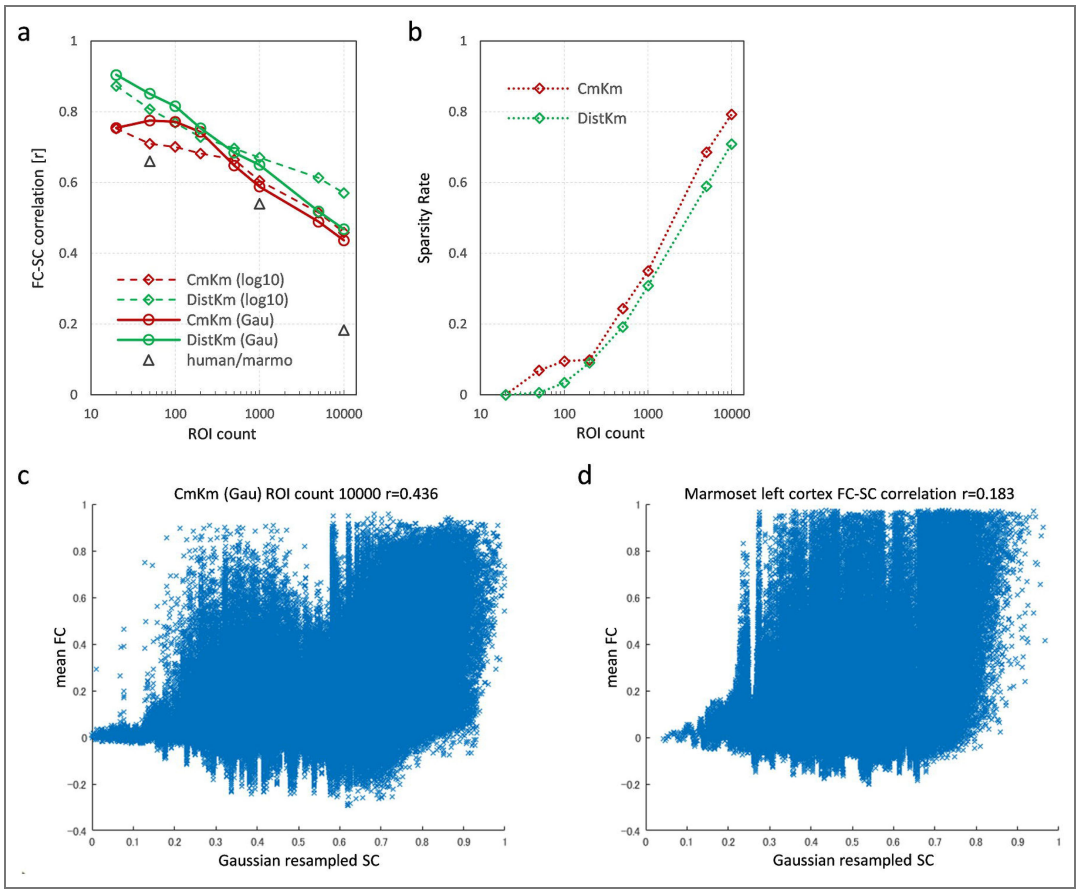
matrix (Poly-tCompCor, no smoothing) in the hemibrain primary 63 ROIs. **d,** (right) SC matrix (neuron) result of FlyEM (cf0.8) minus FlyWire (sc140) in hemibrain primary ROIs. (left) SC matrix (synapse) result of FlyEM (cf0.8) minus FlyWire (sc140). FlyEM shows a higher number of synaptic connections.



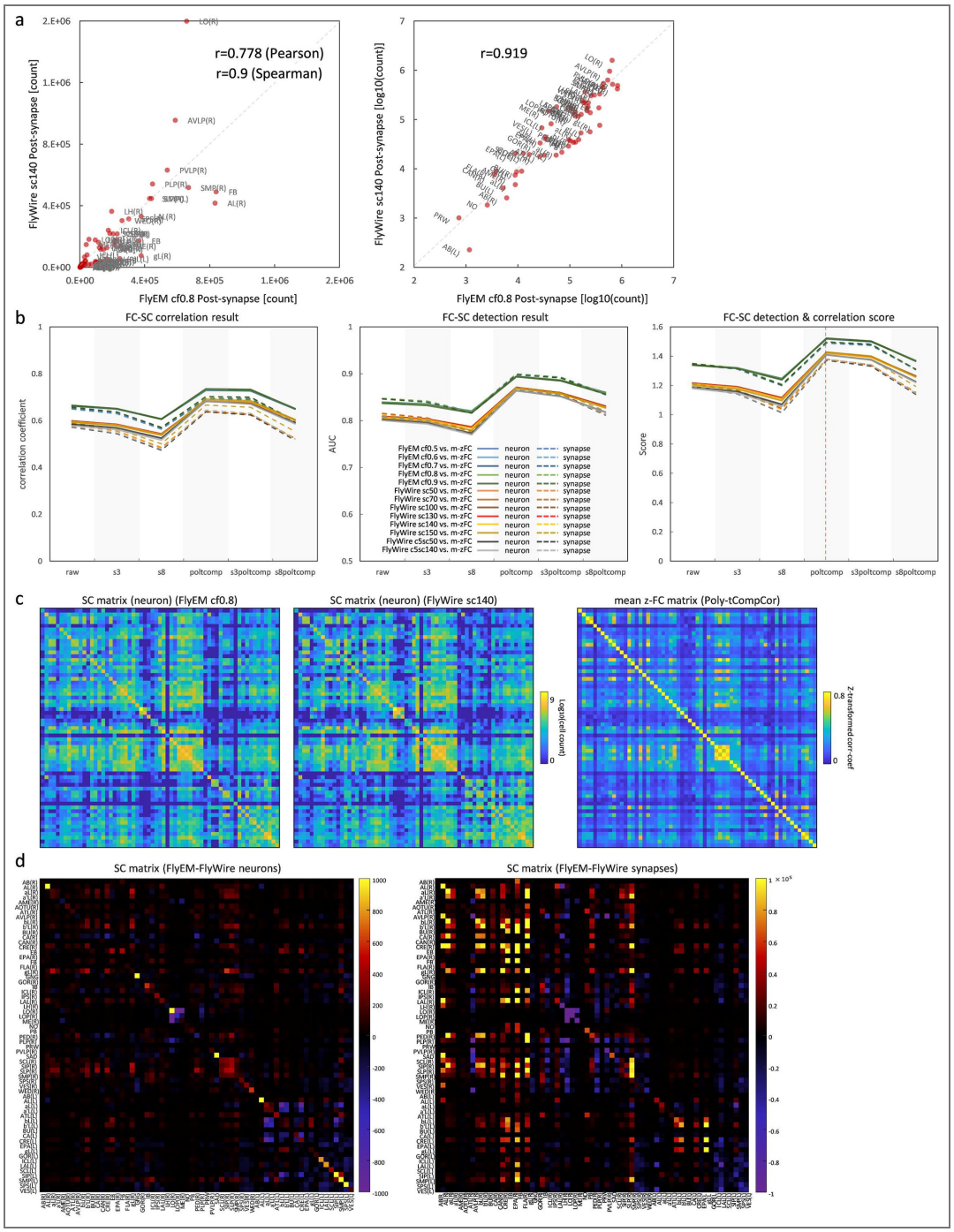
Extended Data Fig.3-2. In-neuron neurotransmitters and comparison between FlyEM and FlyWire connectome data. a, Bar graph of neurotransmitter rate of input neurons in hemibrain primary 63 ROIs based on FlyWire (sc140) connectome data [46]. The horizontal axis shows hemibrain primary ROIs. (DA: dopamine, SER: serotonin, GABA, GLUT: glutamine, ACH: acetylcholine, OCT: octopamine) **b,** Scatter plots of neurotransmitter rate of input neurons vs. FC-SC correlation (neuron) in hemibrain primary ROIs (FlyWire sc140). Each of the six neurotransmitters was compared. **c,** Comparison between FlyEM and FlyWire APL right neuron (synapse point cloud mask). Red voxels are FlyEM, green voxels are FlyWire, and yellow voxels are both. Sørensen-Dice index (0.716) shows similarity, so neuropil shape is similar between two. **d,** (left) A scatter plot of post-synapse count between FlyEM and FlyWire APL right neuron. Pearson correlation shows r=0.104, so not similar at all. (center) Same as left, but synapse point cloud was spatially smoothed by gaussian kernel (2 voxels). Similarity was increased (r=0.379). (right) Same as center, but smoothed by gaussian kernel (8 voxels). Similarity was increased (r=0.66). These results support that spatial smoothing can absorb inter-individual variability.



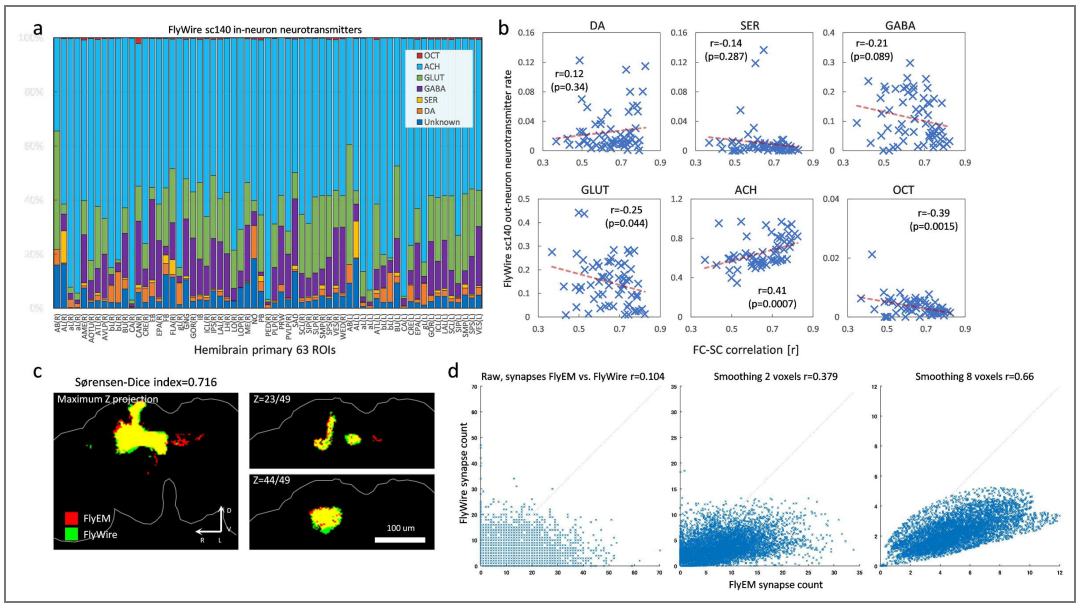
Extended Data Fig.4-1. Quantification of pre– and post-synapse segregation and relationship between FC and synapses (FlyEM). a, Unsegregated first (APL_L cell) and second (DPM_R cell) ranked cells in the FlyEM (cf0.8) connectome data. Red dots shows pre-synapses, blue dot shows post-synapses. **b,** The PPSSI histogram of hemibrain neurons (99,644) in FlyEM (cf0.8) connectome data. The vertical axis shows neuron count, and the horizontal axis shows PPSSI [0 1]. **c,** Segregated first (AOTU019_R cell) and second (DNa02_R cell) ranked cells in the FlyEM (cf0.8) connectome data. **d,** Histogram of null SC matrices (blue bar) and extracted SC matrix with cPPSSI (0-0.1) (red solid line). The black dotted line shows the cumulative distribution function of the normal distribution, the brown solid line shows the cumulative distribution function of null & extracted SC matrices, and the red dotted line shows the Bonferroni-corrected p<0.05 threshold. Top shows the SC matrix (synapse) total, middle shows FC-SC correlation (neuron), and bottom shows FC-SC correlation (synapse). **e,** Histogram of null SC matrices and extracted SC matrix with cPPSSI (0.9-1). **f,** Example of a reciprocal-synapse in the EM image (left) and 3D image (right). g, Histogram of reciprocal-synaptic minimum distances in the FlyEM (cf0.8) connectome data. The vertical axis shows reciprocal-synapse pair count, and the horizontal axis shows distance [μm]. **h,** Histogram of null SC matrices and extracted SC matrix with reciprocal-synapses (≤2μm). **i,** The segregation index histogram of all neurons (139,255) in FlyWire (sc140) connectome data. The horizontal axis shows segregation index [0 1]. **j,** The segregation index histogram of hemibrain neurons (99,644) in the FlyEM (cf0.8) connectome data. The horizontal axis shows the segregation index [0 1].



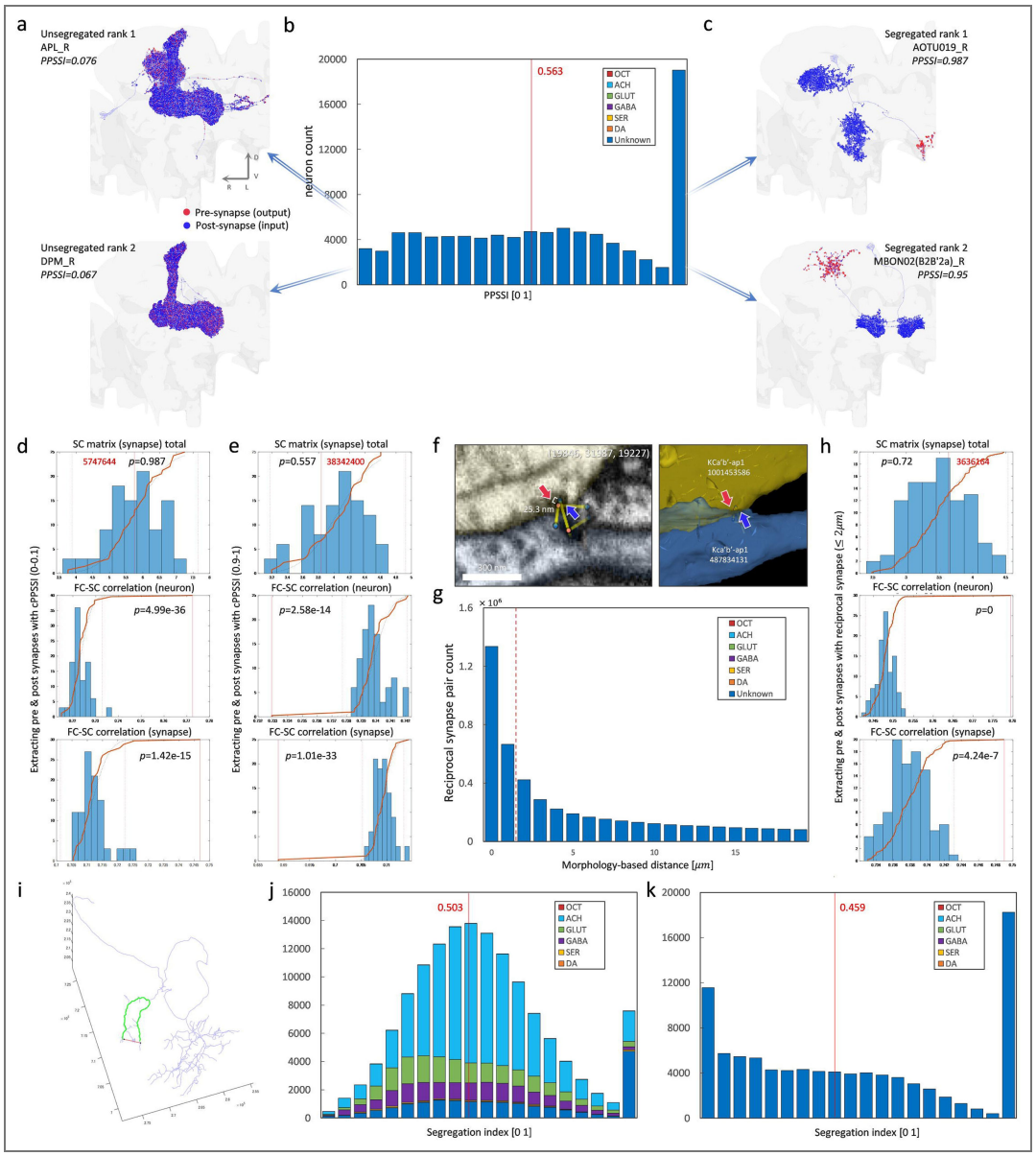
Extended Data Fig.4-2. Reciprocal synapses of CT1 and APL neurons, and synapse point cloud for several conditions. a, Histogram of reciprocal-synaptic minimum distances of the CT1-R neuron in the FlyWire (sc140) connectome data. The vertical axis shows reciprocal-synapse pair count, and the horizontal axis shows distance [μm]. **b,** Skeleton view of the CT1-R neuron. Red dots are pre-synapses and blue dots are post-synapses of reciprocal synapses (< 2μm). **c,** Histogram of reciprocal-synaptic minimum distances of the APL-R neuron in the FlyWire (sc140) connectome data. **d,** Skeleton view of the APL-R neuron, the same way as b. **e,** (top) Maximum Z-projection of pre & post synapse point cloud within cPPSSI (0-0.1) in the FlyEM data (cf0.8). (bottom) Maximum Z-projection of pre & post synapse point cloud within cPPSSI (0-0.1) in the FlyWire data (sc140). **f,** The same as in e, but with the condition cPPSSI in 0.9-1. **g,** The same as in e, but with the condition of only reciprocal synapses (< 2μm).



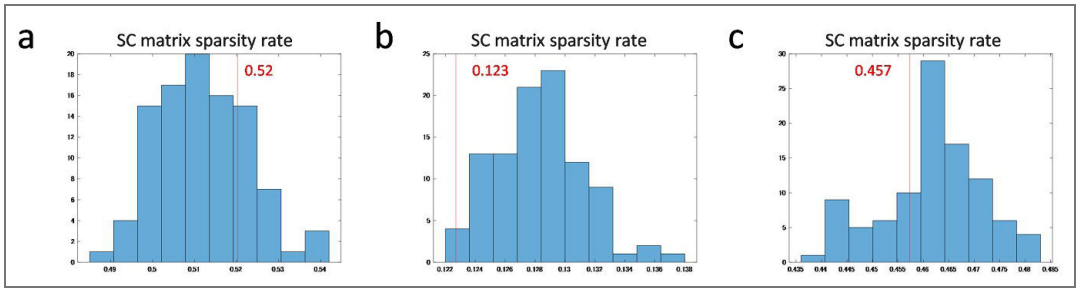
Extendrded Data Fig.4-3. Unsegregated and segregated cell ranking examples (FlyWire sc140). a, Unsegregated cell ranking from 3 to 11th in the FlyWire$_{rd}$(sc140) connectome data. Red dots shows pre-synapses, blue dots shows post-synapses. **b,** Segregated cell ranking from 3 to 11th.



Extended Data Fig.4-4. Unsegregated and segregated cell ranking (FlyEM cf0.8). a, Unsegregated cell ranking from 3 to 11th in the $FlyE_{rd}M$ (cf0.8) connectome data. Red dots shows pre-synapse, blue dots show post-synapses. **b,** Segregated cell ranking from 3 to 11th.



Extended Data Fig.4-5. Sparsity rate of extracted SC matrix. a, Sparsity rate histogram of SC matrix with cPPSSI (0-0.1) and subsampled null SC matrices corresponding to Fig.4e. The red line indicates sparsity rate of the SC matrix with cPPSSI (0-0.1). **b**, Sparsity rate histogram of SC matrix with cPPSSI (0.9-1) and subsampled null SC matrices, corresponding to Fig.4f. **c**, Sparsity rate histogram of SC matrix with reciprocal synapse (≤2μm) and subsampled null SC matrices, corresponding to Fig.4i.



Extended Data Fig.5-1. Mechanosensory pathway from Wedge to Ellipsoid Body. a, 3D image of WL-L right, WED016 left, WPNb tier 1, and ER1 left neurons, frontal view. Black square insert shows WL-L right, WED016 left, WPNb tier 1 and ER1 left neurons from the left lateral view. **b,** Network graph of WL-L right, WED016 left, WPNb

tier 1 and ER1 left neurons from the FlyWire codex (sc50, connection threshold≥5). Blue is for acetylcholine, ocher for GABA neurons. The small number on the edges shows synapse count.

Data availability

Whole-brain calcium imaging data and walking behavior data in Drosophila Melanogaster can be downloaded from the DANDI web site (*https://dandiarchive.org/dandiset/000727/0.240106.0043*). FlyWire connectome data can be downloaded from flywire codex (*https://codex.flywire.ai/*) and the Zenodo web site (*https://zenodo.org/records/10676866*). FlyEM connectome data can be downloaded from Janelia's DVID web site (*https://dvid.io/blog/release-v1.2/*). Software and code used in this study is open source and available from *https://github.com/takuto-okuno-riken/flywalk*.

Acknowledgements

The authors wish to thank Dr. Clandinin for providing valuable calcium imaging data and answering our question.

This research was supported by the program for Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS 2.0) from the Japan Agency for Medical Research and Development, AMED. Grant number: JP24wm0625408.

Additional information

Authors' Contributions

T.O. conceived of the presented idea. T.O. developed the theory, performed the computations. T.O. downloaded and processed fly data. T.O., A.W., H.O. and J.H. discussed the results and contributed to the final manuscript.

Funding

Funder	Grant reference number	Author
Japan Agency for Medical Research and Development (AMED)	JP24wm0625408	Takuto Okuno

Author ORCID iDs

Takuto Okuno: https://orcid.org/0000-0002-1373-8448
Hideyuki Okano: https://orcid.org/0000-0001-7482-5935

References

[1] **Gerstein G. L.,** Perkel D. H. (1969) Simultaneously recorded trains of action potentials: Analysis and functional interpretation. *Science* **164**:828-830 https://doi.org/10.1126/science.164.3881.828 | PubMed

[2] **Gerstein G. L.,** Bedenbaugh P., Aertsen A. M. (1989) Neuronal assemblies. *IEEE Trans. Biomed. Eng* **36**:4-14 https://doi.org/10.1109/10.16444 | PubMed

[3] **Friston K. J.,** Frith C. D., Liddle P. F., Franckowiak R. S. (1993) Functional Connectivity: The Principal-Component Analysis of Large (PET) Data Sets. *J. Cereb. Blood Flow Metab* **13**:5-14 https://doi.org/10.1038/jcbfm.1993.4 | PubMed

[4] **Friston K.,** Moran R., Seth A. K. (2013) Analysing connectivity with Granger causality and dynamic causal modelling. *Curr. Op. in Neuro* **23**:172-178 https://doi.org/10.1016/j.conb.2012.11.010 | PubMed

[5] **Hagmann P.,** Cammoun L., Gigandet X., Meuli R., Honey C. J., Wedeen V. J., Sporns O. (2008) Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biol* **6**:e159 https://doi.org/10.1371/journal.pbio.0060159 | PubMed

[6] **Honey C. J.,** Sporns O., Cammoun L., Gigandet X., Thiran J. P., Meuli R., Hagmann P. (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. USA* **106**:2035-2040 https://doi.org/10.1073/pnas.0811168106 | PubMed

[7] **Honey C. J.,** Thivierge J.-P., Sporns O. (2010) Can structure predict function in the human brain?. *NeuroImage* **52**:766-776 https://doi.org/10.1016/j.neuroimage.2010.01.071 | PubMed

[8] **Baum G. L.,** Cui Z., Roalf D. R., Ciric R., Betzel R. F., Larsen B., Cieslak M., Cook P. A., Xia C. H., Moore T. M., *et al.* (2020) Development of structure–function coupling in human brain networks during youth. *PNAS* **117**:771-778 https://doi.org/10.1073/pnas.1912034117 | PubMed

[9] **Turner M. H.,** Mann K., Clandinin T. R. (2021) The connectome predicts resting-state functional connectivity across the Drosophila brain. *Current Biology* **31** https://doi.org/10.1016/j.cub.2021.03.004 | PubMed

[10] **Okuno T.,** Hata J., Kawai C., Okano H., Woodward A. (2024) A Novel Directed Seed-Based Connectivity Analysis Toolbox Applied to Human and Marmoset Resting-State fMRI. *J. Neurosci* **44**:e0389242024 https://doi.org/10.1523/jneurosci.0389-24.2024 | PubMed

[11] **Hira R.,** Townsend L. B., Smith I. T., Yu C.-H., Stirman J. N., Yu Y., Smith S. L. (2025) Mesoscale functional architecture in medial posterior parietal cortex. *eLife* 105213v1 https://doi.org/10.7554/elife.105213.1

[12] **Scheffer L. K.,** Xu C. S., Januszewski M., Lu Z., Takemura S., 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 https://doi.org/10.7554/eLife.57443 | PubMed

[13] **Dorkenwald S.,** Matsliah A., Sterling A. R., Schlegel P., Yu S., McKellar C. E., Lin A., Costa M., Eichler K., Yin Y., *et al.* (2024) Neuronal wiring diagram of an adult brain. *Nature* **634** https://doi.org/10.1038/s41586-024-07558-y | PubMed

[14] **Brezovec B. E.,** Berger A. B., Hao Y. A., Chen F., Druckmann S., Clandinin T. R. (2024) Mapping the neural dynamics of locomotion across the Drosophila brain. *Current Biology* **34**:710-726 https://doi.org/10.1016/j.cub.2023.12.063 | PubMed

[15] **Penny W.,** Friston K., Ashburner J., Kiebel S., Nichols T. (2007) *Statistical Parametric Mapping: The Analysis of Functional Brain Images* London: Elsevier.

[16] **Holmes A. P.,** Friston K. J. (1998) Generalisability, Random Effects & Population Inference. *NeuroImage* **7**:S754 https://doi.org/10.1016/s1053-8119(18)31587-8

[17] **Schlegel P.,** Bates A. S., Stürner T., Jagannathan S. R., Drummond N., Hsu J., Capdevila L. S., 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 https://doi.org/10.7554/eLife.66018 | PubMed

[18] **Schlegel P.,** Yin Y., Bates A. S., Dorkenwald S., Eichler K., Brooks P., Han D. S., Gkantia M., Santos M. D., Munnelly E. J., *et al.* (2024) Whole-brain annotation and multi-connectome cell typing of Drosophila. *Nature* **634**:139-152 https://doi.org/10.1038/s41586-024-07686-5 | PubMed

[19] **Heinrich L.,** Funke J., Pape C., Nunez-Iglesias J., Saalfeld S. (2018) Synaptic Cleft Segmentation in Non-Isotropic Volume Electron Microscopy of the Complete Drosophila Brain. *arXiv* 1805.02718 https://doi.org/10.1007/978-3-030-00934-2_36

[20] **Helmchen F.,** Waters J. (2002) Ca2+ imaging in the mammalian brain in vivo. *Eur J Pharmacol* **447**:119-129 https://doi.org/10.1016/s0014-2999(02)01836-8 | PubMed

[21] **Branco T.,** Häusser M. (2010) The single dendritic branch as a fundamental functional unit in the nervous system. *Current Opinion in Neurobiology* **20**:494-502 https://doi.org/10.1016/j.conb.2010.07.009 | PubMed

[22] **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., *et al.* (2016) Quantitative neuroanatomy for connectomics in Drosophila. *eLife* **5**:e12059 https://doi.org/10.7554/eLife.12059 | PubMed

[23] **Shafer O. T.,** Gutierrez G. J., Li K., Mildenhall A., Spire D., Marty J., Lazar A. A., Fernandez M. D. (2022) Connectomic analysis of the Drosophila lateral neuron clock cells reveals the synaptic basis of functional pacemaker classes. *eLife* **11**:e79139 https://doi.org/10.7554/eLife.79139 | PubMed

[24] **Meier M.,** Borst A. (2019) Extreme Compartmentalization in a Drosophila Amacrine Cell. *Current Biology* **29**:1545-1550 https://doi.org/10.1016/j.cub.2019.03.070 | PubMed

[25] **Behnia R.,** Clark D. A., Carter A. G., Clandinin T. R., Desplan C. (2014) Processing properties of ON and OFF pathways for Drosophila motion detection. *Nature* **512**:427-430 https://doi.org/10.1038/nature13427 | PubMed

[26] **Gruntman E.,** Reimers P., Romani S., Reiser M. B. (2021) Non-preferred contrast responses in the Drosophila motion pathways reveal a receptive field structure that explains a common visual illusion. *Current Biology* **31**:5286-5298 https://doi.org/10.1016/j.cub.2021.09.072 | PubMed

[27] **Groschner L. N.,** Malis J. G., Zuidinga B., Borst A. (2022) A biophysical account of multiplication by a single neuron. *Nature* **603**:119-123 https://doi.org/10.1038/s41586-022-04428-3 | PubMed

[28] **Braun A.,** Borst A., Meier M. (2023) Disynaptic inhibition shapes tuning of OFF-motion detectors in Drosophila. *Current Biology* **33**:2260-2269 https://doi.org/10.1016/j.cub.2023.05.007 | PubMed

[29] **Schenk J. E.,** Gaudry Q. (2023) Nonspiking Interneurons in the Drosophila Antennal Lobe Exhibit Spatially Restricted Activity. *eNeuro* **10**:ENEURO.0109-22.2022 https://doi.org/10.1523/eneuro.0109-22.2022 | PubMed

[30] **Inada K.,** Tsuchimoto Y., Kazama H. (2017) Origins of Cell-Type-Specific Olfactory Processing in the Drosophila Mushroom Body Circuit. *Neuron* **95**:357-367 https://doi.org/10.1016/j.neuron.2017.06.039 | PubMed

[31] **Amin H.,** Apostolopoulou A. A., Suárez-Grimalt R., Vrontou E., Lin A. C. (2020) Localized inhibition in the Drosophila mushroom body. *eLife* **9**:e56954 https://doi.org/10.7554/eLife.56954 | PubMed

[32] **Liu X.,** Davis R. L. (2009) The GABAergic anterior paired lateral neuron suppresses and is suppressed by olfactory learning. *Nat. Neurosci* **12**:53-59 https://doi.org/10.1038/nn.2235 | PubMed

[33] **Grimes W. N.,** Li W., Chávez A. E., Diamond J. S. (2009) BK channels modulate pre– and postsynaptic signaling at reciprocal synapses in retina. *Nat. Neurosci* **12**:585-592 https://doi.org/10.1038/nn.2302 | PubMed

[34] **Grimes W. N.,** Zhang J., Graydon C. W., Kachar B., Diamond J. S. (2010) Retinal Parallel Processors: More than 100 Independent Microcircuits Operate within a Single Interneuron. *Neuron* **65**:873-885 https://doi.org/10.1016/j.neuron.2010.02.028 | PubMed

[35] **Euler T.,** Detwiler P. B., Denk W. (2002) Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature* **418**:845-852 https://doi.org/10.1038/nature00931 | PubMed

[36] **Patel J. C.,** Witkovsky P., Avshalumov M. V., Rice M. E. (2009) Mobilization of Calcium from Intracellular Stores Facilitates Somatodendritic Dopamine Release. *J. Neurosci* **29**:6568-6579 https://doi.org/10.1523/jneurosci.0181-09.2009 | PubMed

[37] **Grienberger C.,** Chen X., Konnerth A. (2015) Dendritic function in vivo. *Trends Neurosci* **38**:45-54 https://doi.org/10.1016/j.tins.2014.11.002 | PubMed

[38] **Shapson-Coe A.,** Januszewski M., Berger D. R., Pope A., Wu Y., Blakely T., Schalek R. L., Li P. H., Wang S., Maitin-Shepard J., *et al.* (2024) A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science* **384**:eadk4858 https://doi.org/10.1126/science.adk4858 | PubMed

[39] **Cervantes-Sandoval I.,** Phan A., Chakraborty M., Davis R. L. (2017) Reciprocal synapses between mushroom body and dopamine neurons form a positive feedback loop required for learning. *eLife* **6**:e23789 https://doi.org/10.7554/eLife.23789 | PubMed

[40] **Takumura S.,** Aso Y., Hige T., Wong A., Lu Z., Xu C. S., Rivlin P. K., Hess H., Zhao T., Parag T., *et al.* (2017) A connectome of a learning and memory center in the adult Drosophila brain. *eLife* **6**:e23789 https://doi.org/10.7554/eLife.23789

[41] **Wanner A. A.,** Genoud C., Masudi T., Siksou L., Friedrich R. W. (2016) Dense EM-based reconstruction of the interglomerular projectome in the zebrafish olfactory bulb. *Nat. Neurosci* **19**:816-825 https://doi.org/10.1038/nn.4290 | PubMed

[42] **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 https://doi.org/10.1016/j.neuron.2013.12.017 | PubMed

[43] **Friman O.,** Borga M., Lundberg P., Knutsson H. (2004) Detection and detrending in fMRI data analysis. *NeuroImage* **22**:645-655 https://doi.org/10.1016/j.neuroimage.2004.01.033 | PubMed

[44] **Behzadi Y.,** Restom K., Liau J., Liu T. T. (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**:90-101 https://doi.org/10.1016/j.neuroimage.2007.04.042 | PubMed

[45] **Majka P.,** Bai S., Bakola S., Bednarek S., Chen J. M., Jermakow N., Passarelli L., Reser D. H., Theodoni P., Worthy K. H., *et al.* (2020) Open access resource for cellular-resolution analyses of corticocortical connectivity in the marmoset monkey. *Nat. Comm* **11**:1133 https://doi.org/10.1038/s41467-020-14858-0 | PubMed

[46] **Eckstein N.,** Bates A. S., Champion A., Du M., Yin Y., Schlegel P., Lu A. K.-Y., Rymer T., Finley-May S., Paterson T., *et al.* (2024) Neurotransmitter classification from electron microscopy images at synaptic sites in Drosophila melanogaster. *Cell* **187**:2574-2594.e23, https://doi.org/10.1016/j.cell.2024.03.016 | PubMed

[47] **Tomioka R.,** Okamoto K., Furuta T., Fujiyama F., Iwasato T., Yanagawa Y., Obata K., Kaneko T., Tamamaki N. (2005) Demonstration of long-range GABAergic connections distributed throughout the mouse neocortex. *Eur. J. Neurosci* **21**:1587-1600 https://doi.org/10.1111/j.1460-9568.2005.03989.x | PubMed

[48] **Higo S.,** Akashi K., Sakimura K., Tamamaki N. (2009) Subtypes of GABAergic neurons project axons in the neocortex. *Front. Neuroanat* **3** https://doi.org/10.3389/neuro.05.025.2009 | PubMed

[49] **Mazo C.,** Nissant A., Saha S., Peroni E., Lledo P.-M., Lepousez G. (2022) Long-range GABAergic projections contribute to cortical feedback control of sensory processing. *Nat. Comm* **13**:6879 https://doi.org/10.1038/s41467-022-34513-0 | PubMed

[50] **Ester M.,** Kriegel H.-P., Sander J., Xu X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mini. pp. 226-231

[51] **Buckner R. L.,** Andrews-Hanna J. R., Schacter D. L. (2008) The Brain's Default Network Anatomy, Function, and Relevance to Disease. *Ann. N.Y. Acad. Sci* **1124**:1-38

[52] **Namiki S.,** Dickinson M. H., Wong A. M., Korff W., Card G. M. (2018) The functional organization of descending sensory-motor pathways in Drosophila. *eLife* **7**:e34272 https://doi.org/10.7554/eLife.34272 | PubMed

[53] **Okubo T. S.,** Patella P., D'Alessandro I., Wilson R. I. (2020) A Neural Network for Wind-Guided Compass Navigation. *Neuron* **107**:924-940 https://doi.org/10.1016/j.neuron.2020.06.022 | PubMed

[54] **El Jundi B.,** Dacke M. (2021) Insect Orientation: The Drosophila Wind Compass Pathway. *Current Biology* **31**:PR83-R85 https://doi.org/10.1016/j.cub.2020.11.033 | PubMed

[55] **Coates K. E.,** Calle-Schuler S. A., Helmick L. M., Knotts V. L., Martik B. N., Salman F., Warner L. T., Valla S. V., Bock D. D., Dacks A. M. (2020) The Wiring Logic of an Identified Serotonergic Neuron That Spans Sensory Networks. *J. Neurosci* **40**:6309-6327 https://doi.org/10.1523/jneurosci.0552-20.2020 | PubMed

[56] **Randi F.,** Sharma A. K., Dvali S., Leifer A. M. (2023) Neural signal propagation atlas of Caenorhabditis elegans. *Nature* **623**:406-414 https://doi.org/10.1038/s41586-023-06683-4 | PubMed

[57] **Harris J. J.,** Attwell D. (2012) The Energetics of CNS White Matter. *J. Neurosci* **32**:356-371 https://doi.org/10.1523/jneurosci.3430-11.2012 | PubMed

[58] **Iadecola C.** (2017) The Neurovascular Unit Coming of Age: A Journey through Neurovascular Coupling in Health and Disease. *Neuron* **1**:17-42 https://doi.org/10.1016/j.neuron.2017.07.030 | PubMed

[59] **Yemini E.,** Lin A., Nejatbakhsh A., Varol E., Sun R., Mena G. E., Samuel A. D., Paninski L., Venkatachalam V., Hobert O. (2021) NeuroPAL: A Multicolor Atlas for Whole-Brain Neuronal Identification in C. elegans. *Cell* **184**:272-288 https://doi.org/10.1016/j.cell.2020.12.012 | PubMed

[60] **Hendricks M.,** Ha H., Maffey N., Zhang Y. (2012) Compartmentalized calcium dynamics in a C. elegans interneuron encode head movement. *Nature* **487**:99-103 https://doi.org/10.1038/nature11081 | PubMed

[61] **Brezovec B. E.,** Berger A. B., Hao Y. A., Lin A., Ahmed O. M., Pacheco D. A., Thiberge S. Y., Murthy M., Clandinin T. R. (2024) BIFROST: A method for registering diverse imaging datasets of the Drosophila brain. *Proc Natl Acad Sci USA* **121**:e2322687121 https://doi.org/10.1073/pnas.2322687121 | PubMed

[62] **Avants B. B.,** Tustison N. J., Song G., Cook P. A., Klein A., Gee J. C. (2011) A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* **54**:2033-2044 https://doi.org/10.1016/j.neuroimage.2010.09.025 | PubMed

[63] **Okuno T.,** Woodward A. (2021) Vector Auto-Regressive Deep Neural Network: A Data-Driven Deep Learning-Based Directed Functional Connectivity Estimation Toolbox. *Front. Neurosci* **15**:764796 https://doi.org/10.3389/fnins.2021.764796 | PubMed

[64] **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. *Nature Methods* **18**:771-774 https://doi.org/10.1038/s41592-021-01183-7 | PubMed

[65] **Bogovic J. A.,** Otsuna H., Heinrich L., Ito M., Jeter J., Meissner G., Nern A., Colonell J., Malkesman O., Ito K., *et al.* (2020) An unbiased template of the Drosophila brain and ventral nerve cord. *PLoS ONE* **15**:e0236495 https://doi.org/10.1371/journal.pone.0236495 | PubMed

[66] **Chen T.-W.,** Wardill T. J., Sun Y., Pulver S. R., Renninger S. L., Baohan A., Schreiter E. R., Kerr R. A., Orger M. B., Jayaraman V., *et al.* (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**:295-300 https://doi.org/10.1038/nature12354 | PubMed

[67] **Woolrich M. W.,** Ripley B. D., Brady M., Smith S. M. (2001) Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *NeuroImage* **14**:1370-1386 https://doi.org/10.1006/nimg.2001.0931 | PubMed

[68] **Okuno T.,** Ichinohe N., Woodward A. (2024) A reappraisal of the default mode and frontoparietal networks in the common marmoset brain. *Front. Neuroimaging* **2**:1345643 https://doi.org/10.3389/fnimg.2023.1345643 | PubMed

Luke Brezovec (2024) Mapping the Neural Dynamics of Locomotion across the Drosophila Brain. DANDI. ID 000727/0.240106.0043 https://dandiarchive.org/dandiset/000727/0.240106.0043

FlyWire Consortium (2024) FlyWire Whole-brain Connectome Connectivity Data. Zenodo. https://doi.org/10.5281/zenodo.10676866

Bill Katz (2020) FlyEM Hemibrain V1.2 Release. FlyEM Hemibrain V1.2. ID release-v1.2 https://dvid.io/blog/release-v1.2/

Peer reviews

Reviewer #1 (Public review):

In this revision the authors address some of the points, but they also make some technical errors. My overall view of the manuscript hasn't changed since the original evaluation.

Previously I noted that SC sparsity presents an issue when comparing to full FC matrices. They authors misinterpreted the Honey et al paper. They resampled ALL entries of the SC matrix (including zeros) from a Gaussian distribution. In effect, this assigns zeros small (but uniform) weights. In Honey et al, the authors resampled only existing edge weights from a gaussian distribution (the rationale at the time was that there might be pushback against the extremely heavy-tailed edge weight distribution). In other words, the zeros are still zeros following this resampling procedure.

That said, I agree that the log transform is likely useful or necessary given edge weight distributions.

In short, I still think that the approach is interesting and meritorious, I just don't think the execution is correct.

https://doi.org/10.7554/eLife.107990.2.sa2

Reviewer #3 (Public review):

Summary:

In this manuscript, Okuno et al. re-analyze whole-brain imaging data collected in another paper (Brezovec et al., 2024) in the context of the two currently available Drosophila connectome datasets: the partial "FlyEM" (hemibrain) dataset (Scheffer et al., 2020) and the whole-brain "FlyWire" dataset (Dorkenwald et al., 2024). They apply existing fMRI signal processing algorithms to the fly imaging data and compute function-structure correlations across a variety of post-processing parameters (noise reduction methods, ROI size), demonstrating an inverse relationship between ROI size and FC-SC correlation. The authors go on to look at structural connectivity amongst more polarized or less polarized neurons, and suggest that stronger FC-SC correlations are driven by more polarized neurons.

Strengths:

(1) The result that larger mesoscale ROIs have higher correlation with structural data is interesting. This has been previously discussed in Drosophila in Turner et al., 2021, but here it is quantified more extensively.

(2) The quantification of neuron polarization (PPSSI) as applied to these structural data is a promising approach for quantifying differences in spatial synapse distribution. The revision now uses morphological cable length for some analyses rather than straight-line distance, which improves the realism and interpretability of these results.

Weaknesses:

One should not score noise/nuisance removal methods solely by their impact on FC-SC correlation values, because we do not know a priori that direct structural connections correspond with strong functional correlations. In fact, work in C. elegans, where we have access to both a connectome and neuron-resolution functional data, suggests that this relationship is weak (Yemini et al., 2021; Randi et al., 2023). Similarly, I don't think it's appropriate to tune the confidence scores on the EM datasets using FC-SC correlations as an output metric. While it is likely that some FC-SC relationship does exist at large scales, it does not in my view justify use of this metric for evaluating noise removal methods, since such methods may inadvertently remove real neural correlates. This concern remains unaddressed in the revision.

Any discussion of FC-SC comparisons should include an analysis of excitatory/inhibitory neurotransmitters, which are available in the fly connectome dataset. The authors examine the ratios of input and output neurotransmitters in different defined regions. However, I

think it would be more useful to integrate the neurotransmitter information more fully into the assessment of SC, for instance: examining the signed weight (excitatory - inhibitory), or by examining the excitatory and inhibitory networks separately.

Comparisons between fly and human MRI data are also premature here. Firstly, the fly connectomes, which are derived from neuron-scale EM reconstructions, are a qualitatively different kind of data from human connectomes, which are derived from DSI imaging of large-scale tracts. Likewise, calcium data and fMRI data are very different functional data acquisition methods-the fact that similar processing steps can be used on time-series data does not make them surprisingly similar, and does not in my view constitute evidence of "similar design concepts."

The comparison of FlyEM/FlyWire connectomes concludes that differences are more likely a result of data processing than of inter-individual variability. If this is the case, the title should not claim that the manuscript covers individual variability.

The analysis of the wedge-AVLP neuron strikes me as highly speculative, given that the alignment precision between the connectome and the functional data is around 5 microns (Brezovec* et al, PNAS 2024).

https://doi.org/10.7554/eLife.107990.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *In this paper, the authors analyze connectome data from Drosophila and compare the physical wiring with functional connectivity estimated from calcium imaging data. They quantify structure-function relationships as a correlation of the two connectivity modalities. They report correlations roughly comparable to what has been described in the literature on sc/fc relationships in mammalian connectome data at the meso-scale. They then repeat their analysis, focusing on segregated versus unsegregated synapses. They derive separate connectomes using one or the other class of synapse. They show differential contributions to the sc/fc relationships by segregated versus unsegregated synapses.*
>
> *Strengths:*
>
> *There is nice synthesis of multimodal imaging data (Ca and EM data from flies and meso-scale data from human and marmoset).*

Thank you very much for your comments.

> *Weaknesses:*
>
> *(1) The paper is written in an unusual way. The introduction intermingles results with background, making it hard to figure out what precisely is being tested.*

Thank you for pointing this out. We have revised the introduction to make it more concise.

> *(2) There are also major methodological gaps. Though the mammalian connectomes are used as a point of reference, no descriptions of their origins or processing are included.*

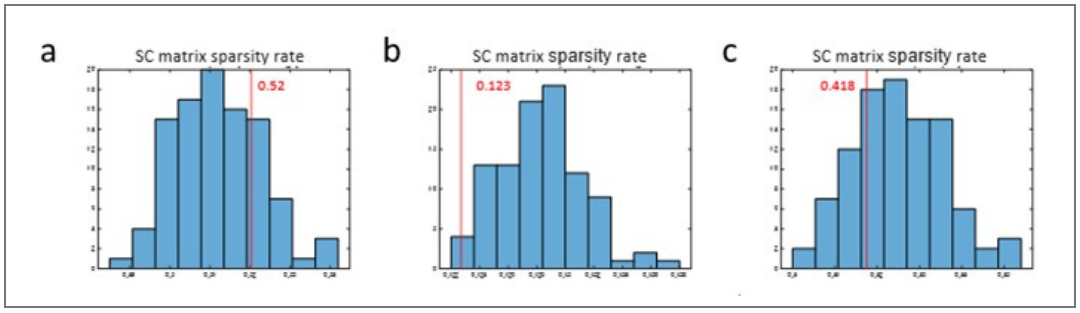
The reanalysis of marmoset data is presented in Ext. Data Figure. However, as pointed out by other reviewers, the data was obtained in [10], and the processing is also described in [10]. Therefore, we have revised the caption and removed the Ethics Declaration.

> *(3) A major weakness stems from the actual calculation of the sc/fc correlation. In general, SC is sparse. In the case of the EM connectomes, it is *exceptionally* sparse (most neural elements are not connected to one another). The authors calculated sc/fc coupling by correlating the off-diagonal elements of sc (the logarithm of its edge weights) and fc matrices with one another. The logarithmic transformation yields a value of infinity for all zero entries. The authors simply impute these elements with 0. This makes no sense and, depending on whether these zero elements are distributed systematically versus uniformly random, could either inflate or deflate the sc/fc correlations. Care must be taken here.*

Thank you for pointing this out. As you mentioned, the SC matrix becomes increasingly sparse as the number of ROIs increases (Ext. Data Fig.2-2b). In contrast, the FC matrix may contain values even when there are no direct connections between ROIs (indirect connections). We conducted an investigation into this issue. To deal with this issue, Honey et al. (2009) [6] resampled the elements of the SC matrix in rank order using a Gaussian distribution and calculated the FC-SC correlation between this resampled SC and FC.

Ext. Data Fig.2-2a shows a comparison between resampled SC (Honey et al.'s method) and log-scaled SC (our method). Up to 200 ROIs, the proportion of SC matrix elements that are zero is less than 10% (Ext. Data Fig.2-2b), and there is little zero replacement of logarithmic elements. In this situation, replacing with Gaussian arithmetic tends to increase the correlation coefficient (Ext. Data Fig.2-2a). On the other hand, with 10,000 ROIs, where sparsity is extremely high, the proportion of SC matrix elements that are zero exceeds 70%. In this situation, 70-80% of the zeros are randomly assigned from the smaller end of the Gaussian distribution, which causes a lowering of the correlation coefficient (Ext. Data Fig.2-2a, c, d). For these reasons, we believe that log-scaled SC has less bias than resampling with a Gaussian distribution, and conclude that using log-scaled SC as is in this paper is reasonable. Log-scaled SC has also been used in previous studies [9, 68] and is considered a simple method for showing the relationship (correlation) between FC and SC. To show that we have considered this issue, Ext. Data Fig.2-2 has been added to the manuscript.

> *(4) Further, in constructing the segregated versus unsegregated connectomes, they use absolute thresholds for collecting synapses. It is unclear, however, whether similar numbers of synapses were included in both matrices. If the number is different, that might explain the differential relationship with fc; one matrix has more non-zero entries (and as noted earlier, those zero entries are problematic).*



Author response image 1. a, Sparsity rate histogram of SC matrix with cPPSSI (0-0. 1) and subsampled null SC matrices corresponding Fig.4e. Red line indicates sparsity rate of SC matrix with cPPSSI (0-0.1). b, Sparsity rate

histogram of SC matrix with cPPSSI (0.9-1) and subsampled null SC matrices corresponding Fig.4f. c, Sparsity rate histogram of SC matrix with reciprocal synapse ($\leq 2\mu m$) and subsampled null SC matrices corresponding Fig.4i.

Thank you for pointing this out. The number of synaptic connections in the SC matrix shows a large difference between those extracted from cPPSSI (0-0.1) and cPPSSI (0.9-1) (Fig. 4e, f). However, when null SC matrices (99) were generated for each and compared with the cPPSSI-extracted matrices, the FC-SC correlation was significantly higher or lower. At this point, since the sparsity rates of the null SC matrices differed a lot from that of the SC matrices extracted by cPPSSI, we regenerated the null SC matrices in Fig. 4e and 4i. As shown in Author response image 1, we ensured that the extracted SCs (red lines) fit within the null-generated matrices. This figure was added to Ext. Data Fig.4-5, and the main text was also revised. The sparsity rates are 0.52 for cPPSSI (0-0.1) and 0.123 for cPPSSI (0.9-1). Since both cases involve comparisons with null SC matrices that have closely similar sparsity rates, we believe comparison using log-scaled SC is appropriate.

> *(5) There was also considerable text (in the results) describing the processing of the Ca data. In this section, the authors frequently refer to some pipelines as "better" or "worse" (more or less effective). But it is not clear what measures they adopted to assess the effectiveness of a pipeline.*

Detailed registration flow of Ca data is described in “Preprocessing of *D. melanogaster* calcium imaging data” in Materials and Methods section (Ext. Data Fig. 1-1a). Then, optimal nuisance factor removal methods and smoothing size were investigated. We used both correlation analysis (FC-SC correlation) and ROC curve analysis (FC-SC detection). Since signals are assumed to be transmitted between regions based on SC, when SC is treated as the ground truth, we considered a pipeline with a FC-SC higher similarity and higher detection to be better. We updated the Results section to include this point.

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *Okuno et al. investigate the structure-function relationship in the fruit fly Drosophila melanogaster. To do so, they combine published data from two recent synapse-level connectomes ("hemibrain" and "FlyWire") with a dataset comprising functional whole-brain calcium imaging and behavioural data. First, they investigate the applicability of fMRI pre-processing techniques on data from calcium imaging. They then cross-correlate this pre-processed functional data with structural data extracted from the connectomes, including a comparison to humans. The authors proceed to compare the two connectomes and find significant differences, which they attribute to differences in the accuracy of the synapse detections. Next, they present a novel algorithm to quantify whether neurons are segregated (pre- and postsynapses are spatially separate) or unsegregated (pre- and postsynapses are mixed). Using this approach, they find that unsegregated neurons may contribute more to function than segregated neurons. Applying a general linear model to the functional dataset suggests that activity in two brain areas (Wedge and AVLP) is suppressed during walking. The authors identify a GABAergic neuron in the connectome that could be responsible for this effect and suggest it may provide feedback to the fly's "compass" in the central complex.*
>
> *Strengths:*
>
> *The study tackles a relevant question in connectomics by exploring the relationship between structural and functional connectivity in the Drosophila brain. The authors apply a range of established and adapted analytical methods, including fMRI-style preprocessing and a novel synaptic segregation index. The effort to integrate multiple datasets and to compare across species reflects a broad and methodical approach.*

Thank you very much for your comments.

> *Weaknesses:*
>
> *The manuscript would benefit from a clearer overarching narrative to unify the various analyses, which currently appear somewhat disjointed. While the technical methods are extensive, the writing is often convoluted and lacks crucial details, making it difficult to follow the logic and interpret key findings. Additionally, the conclusions are relatively incremental and lack a compelling conceptual advance, limiting the overall impact of the work.*
>
> *(1) The introduction currently contains a number of findings and conclusions that would be better placed in the results and discussion to clearly delineate past findings from new results and speculations.*

Thank you for pointing this out. We have revised the introduction to make it more concise.

> *(2) The narrative would benefit greatly from some clear statements along the lines of "we wanted to find out X, therefore we did Y".*

Thank you for pointing this out. In many biology papers, the problem is clear, but as you say, this paper starts by comparing the very fine SC and FC of flies, which makes the problem unclear and the results sporadic. We have revised the structure of the introduction.

> *(3) More concise terminology would be helpful. For example, the connectomes are currently referred to as either "hemibrain", "FlyEM", "whole-brain", or "FlyWire".*

Thank you for pointing this out. We revised the manuscript to separate "hemibrain" and "whole-brain" from "connectome." "hemibrain" and "whole-brain" retain their original meanings.

> *(4) The abstract claims "a new, more robust method to quantify the degree of pre- and post-synaptic segregation". However, the study fails to provide evidence that this method is indeed more robust than existing methods.*

We apologize, but this information was not included in the main figures or the Results section. It is presented in the Methods section and Ext. Data Fig. 4-1i, j. We moved related texts from the Methods to the Results section.

> *(5) The authors define unsegregated neurons as having mixed pre- and postsynapses in the same space. However, this ignores the neurons' topology: a neuron can exhibit a clearly defined dendrite with (mostly) postsynapses and a clearly defined axon with (mostly) presynapses, which then occupy the same space. This is different from genuinely unsegregated neurons with no distinct dendritic and axonal compartments, such as CT1.*

Thank you for pointing this out. Regarding this point, we think it is difficult to discuss the neuron's topology in this paper. We defined PPSSI and demonstrated only that unsegregated neurons with mixed pre- and post-synapses are scattered throughout the brain (Ext. Data Fig. 4-2e). Further research is needed to determine the relationship with morphology in individual neurons.

One possibility is that inhibitory, non-spiking unsegregated neurons, such as CT1 amacrine cell [24, 27, 28] or interneurons in Antennal Lobe [29], may be widely used throughout the brain (WAGN is also a candidate for this). Grimes et al. [34] mentioned "The retina is a beautiful example of a neural network that optimizes signal processing capacity while minimizing cellular cost." To maintain the signal dynamic range, A17 amacrine cells must optimize the processing units and wiring costs. If one unit equaled one cell, an enormous

number of cell bodies would be required, reducing the number of processing units per volume and increasing the energy cost during development. To optimize this, they proposed arranging units capable of parallel processing within a single cell, thereby maximizing the processing units and wiring costs per volume.

Signal bursts might also occur in the central nervous system (CNS), in which case CNS neurons also require dynamic range adjustment. The concept of optimizing processing units per volume is highly compelling and is thought to apply not only to the retina but throughout the entire brain.

> *(6) It is not entirely clear where the marmoset dataset originates from. Was it generated for this study? If not, why is there a note in the Ethics Declaration?*

Marmoset data were reported in [10] and it was not generated for this study. We therefore removed the Ethics Declaration.

> *(7) On the differences between hemibrain and FlyWire: What is the "18.8 million post-synapses" for FlyWire referring to? The (thresholded) FlyWire synapse table has 130M connections (=postsynapses). Subsetting that synapse cloud to the hemibrain volume still gives ~47M synapses. Further subsetting to only connections between proofread neurons inside the hemibrain volume gives 19.4M - perhaps the authors did something like that? Similarly, the hemibrain synapse table contains 64M postsynapses. Do the 21M "FlyEM" post-synapses refer to proofread neurons only? If the authors indeed used only (post-)synapses from proofread neurons, they need to make that explicit in results and methods, and account for differences in reconstruction status when making any comparisons. For example, the mushroom body in the hemibrain got a lot more attention than in FlyWire, which would explain the differences reported here. For that reason, connection weights are often expressed as, e.g., a fraction of the target's inputs instead of the total number of synapses when comparing connectivity across connectomic datasets. Furthermore, in Figure 3b, it looks like the FlyWire synapse cloud was not trimmed to the exact hemibrain boundaries: for example, the trimmed FlyWire synapse cloud seems to extend further into the optic lobes than the hemibrain volume does.*

Thank you for pointing this out. FlyEM connectome data version 1.2 was downloaded and used as described in Data Availability. This data is provided in the format defined by *https://neuprint.janelia.org/public/neuprintuserguide.pdf*, and we extracted neurons and synapses from it.

The entire segmentation body is 28M segmentations, and there were 99,644 Traced proofread neurons. In addition, there were 73M (pre- or post- alone) synapses, 87M records in synapseSets and 128M records in synapseSet-to-synapse. When we extracted post-synapses between Traced neurons, the total number was 21.4M (i.e., connections from Traced neurons to other body fragments like Orphans were excluded).

The FlyWire dataset (v783) was downloaded from the flywire codex and Zenodo. This dataset contained 139,255 proofread neurons and 54.5M (pair of pre- and post-) synapses, as described in Dorkenwald et al. [13], with 18.8M post-synapses in the regions corresponding to the hemibrain primary ROIs. We have updated the Results and Methods sections by taking into account your comment.

In Fig. 3b, these images were created using a mask that extended the boundaries of the hemibrain primary ROIs, making the boundaries unclear. Therefore, we corrected the images in Fig. 3b by adjusting the mask so that the boundaries were properly aligned.

> ***Reviewer #3 (Public review):***

> *Summary:*
>
> *In this manuscript, Okuno et al. re-analyze whole-brain imaging data collected in another paper (Brezovec et al., 2024) in the context of the two currently available Drosophila connectome datasets: the partial "FlyEM" (hemibrain) dataset (Scheffer et al., 2020) and the whole-brain "FlyWire" dataset (Dorkenwald et al., 2024). They apply existing fMRI signal processing algorithms to the fly imaging data and compute function-structure correlations across a variety of post-processing parameters (noise reduction methods, ROI size), demonstrating an inverse relationship between ROI size and FC-SC correlation. The authors go on to look at structural connectivity amongst more polarized or less polarized neurons, and suggest that stronger FC-SC correlations are driven by more polarized neurons.*
>
> *Strengths:*
>
> *(1) The result that larger mesoscale ROIs have a higher correlation with structural data is interesting. This has been previously discussed in Drosophila in Turner et al., 2021, but here it is quantified more extensively.*
>
> *(2) The quantification of neuron polarization (PPSSI) as applied to these structural data is a promising approach for quantifying differences in spatial synapse distribution.*

Thank you very much for your comments.

> *Weaknesses:*
>
> *One should not score noise/nuisance removal methods solely by their impact on FC-SC correlation values, because we do not know a priori that direct structural connections correspond with strong functional correlations. In fact, work in C. elegans, where we have access to both a connectome and neuron-resolution functional data, suggests that this relationship is weak (Yemini et al., 2021; Randi et al., 2023). Similarly, I don't think it's appropriate to tune the confidence scores on the EM datasets using FC-SC correlations as an output metric.*

Thank you for pointing this out. We believe that the FC in *C. elegans* uses cell body dynamics, which is different from the synaptic population dynamics in a region of fly calcium imaging or fMRI data (BOLD [Blood Oxygenation Level Dependent] signal). The BOLD signal in a region is thought to correspond to the neurovascular coupling of synaptic population dynamics. Furthermore, compartmentalization of a neuron has been observed in *C. elegans* (Hendricks et al., 2012)*, showing different dynamics across neuron compartments. Thus, the dynamics of the cell body and the dynamics of the synaptic population in other regions are different in *C. elegans*. We speculate that there is some relationship between FC-SC between regions, because the FC-SC correlation in the fly brain reached r=0.87 with 20 ROIs (Fig. 2d). We believe that this result is different from the cell body dynamics in *C. elegans*.

*Hendricks et al., “Compartmentalized calcium dynamics in a *C. elegans* interneuron encode head movement,” Nature 487, 99-103 (2012)

> *Any discussion of FC-SC comparisons should include an analysis of excitatory/inhibitory neurotransmitters, which are available in the fly connectome dataset. However, here the authors do not perform any analyses with neurotransmitter information.*

A comparison between FC-SC and neurotransmitter has been written in the Results section. We investigated the ratios of neurotransmitter input (ExtFig.3-2a) and output (Fig. 3f) in each region, and investigated the relationship between this ratio and FC-SC correlation in each neurotransmitter. This revealed significant correlations for acetylcholine (r=0.39, p=0.0013) and GABA (r=-0.25, p=0.046) (Fig. 3g). That is, the higher the percentage of excitatory

connections, the higher the FC-SC correlation; conversely, the higher the percentage of inhibitory connections, the lower the FC-SC correlation.

> *Comparisons between fly and human MRI data are also premature here. Firstly, the fly connectomes, which are derived from neuron-scale EM reconstructions, are a qualitatively different kind of data from human connectomes, which are derived from DSI imaging of large-scale tracts. Likewise, calcium data and fMRI data are very different functional data acquisition methods-the fact that similar processing steps can be used on time-series data does not make them surprisingly similar, and does not in my view, constitute evidence of "similar design concepts."*

Thank you for pointing this out. As you say, fiber bundles of DTI and EM connectome are completely different. Nevertheless, the fact remains that the FC-SC correlation is high in both the fly and human brains. As mentioned above, both regional signal from calcium imaging and BOLD signal from fMRI are based on synaptic population dynamics. It was estimated that 43% of the energy consumption in the gray matter is due to synaptic activity of neurons (Harris et al., 2012), and the BOLD signal fluctuates greatly due to this activity. Furthermore, synaptic activity is thought to be much faster than the activity of microglia and astrocytes, so the FC of fMRI is thought to mainly capture the regional correlation of synaptic activity. In other words, in both flies and humans, although the size is different, the pre-synaptic activity in one region and the pre-synaptic activity in another region via neural fibers are being compared in a common manner in the form of FC-SC.

In addition, non-spiking unsegregated neuron exists in mammals as well, such as the amacrine cell of the retina [34], and even pyramidal cells in the neocortex show local mixtures of pre- and post-synapses (Ext. Data Fig.1-2). If a functional unit is realized by local compartment in a neuron as mentioned in [34], the fly will be a powerful model organism for investigating them, and its functional “design concept” may also be useful for mammals.

Harris et al., “The Energetics of CNS White Matter,” J. Neurosci., 2012, 32 (1) 356-371

> *The comparison of FlyEM/FlyWire connectomes concludes that differences are more likely a result of data processing than of inter-individual variability. If this is the case, the title should not claim that the manuscript covers individual variability.*

Thank you for pointing this out. Inter-individual variability is relevant to both SC and FC. Regarding SC, we think the difference in the number of synapses between the two individuals is due to the difference in detection power caused by differences in the resolution of the electron microscope. Regarding FC, as stated in the Results section, “Spatial smoothing is useful for absorbing inter-individual variability and conducting second-level group analysis.” Increasing the smoothing size improves the correlation and AUC between group-averaged FC and SC, indicating the presence of inter-individual variability in FC (Fig. 2b, Ext. Data Fig. 2-1b, especially when the number of ROIs is high). We added this text in the Introduction and Results sections to address your comment.

> *The analysis of the wedge-AVLP neuron strikes me as highly speculative, given that the alignment precision between the connectome and the functional data is around 5 microns (Brezovec* et al, PNAS 2024).*

As you mentioned, functional analysis has limitations in spatial resolution. In particular, the resolution in the Z axis is 4 μm, which is 1,000 times lower than the resolution of electron microscopy data. This makes it difficult to perfectly match synaptic activity to a synapse in the structural data. Furthermore, spatial smoothing is applied to functional images to absorb inter-individual variability, which can only provide blurred results for group analyses. These are considered limitations of the methods used in fMRI analysis. Despite these limitations, we applied GLM analysis to walking behavior and observed clear inactivity region. This region

roughly corresponds to the synaptic cloud of a neuron named WAGN (Fig.5b and c). This neuron also connects to WPNb and ANs in the connectome data, suggesting a possibility that it is related to walking behavior. This is merely a screening reference; therefore, further biological experimentation is needed to pursue this topic.

> ***Recommendations for the authors:***
>
> ***Reviewing Editor Comments:***
>
> *We should emphasize that the reviewers encouraged revision and resubmission. If the reviewers' comments were to be addressed in full in a revision to strengthen the evidence, this would significantly increase the impact of the findings and the relevance of the work to the fly neuroscience community and to the connectomics field more broadly.*

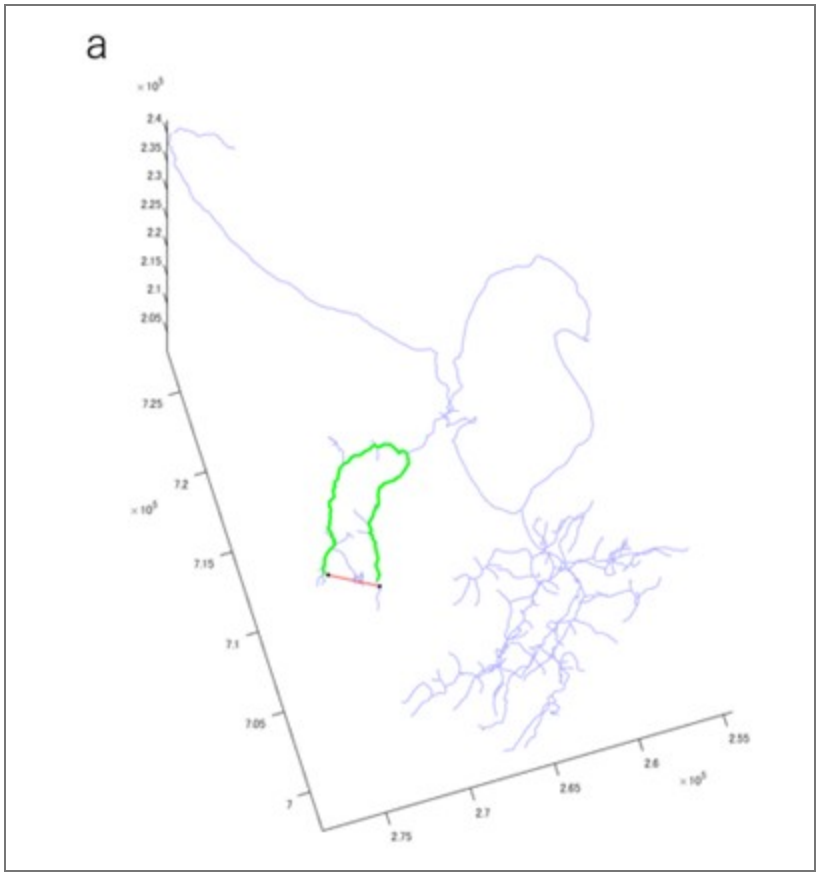
Thank you very much for your comments.

> *Major Issues:*
>
> *(1) Structural correlation and functional correlation measure very different aspects of network data, yet a simple correlation between the off-diagonal elements of the two is used. It would be expected that this would not be directly proportional, and it's not clear why this would be a sensible measure. The authors need a better solution for dealing with the zero entries in the SC matrix. Replacing the infinities with zeros and then running the linear regression to get an SC/FC relationship is not appropriate. Even with a better metric, given that both intuition and other studies have shown a weak correlation between FC and SC, using FC-SC correlation as a quality descriptor for other properties is not proper. Furthermore, the authors don't account for neurotransmitter identity in the structural data, which would have strong implications for the relationships between FC and SC.*

Thank you for pointing this out. To investigate this issue we compared the FC-SC correlation between the Gaussian resampled SC approach used in Honey et al. (2009) [6] and the log-scaled SC used in this study (Ext. Data Fig.2-2a). With a small number of ROIs, the sparsity rate is low (Ext. Data Fig.2-2b), resulting in less zero replacement. Therefore, log-scaled SC is likely to more accurately represent the FC-SC relationship. Furthermore, with a large number of ROIs, the sparsity rate exceeds 70%, and Gaussian resampled SC randomly assigns a large number of zero elements from the smaller end of the distribution. This tends to lower the correlation (Ext. Data Fig.2-2c, d), suggesting that log-scaled SC provides fairer results. Log-scaled SC has been used in previous studies [9, 68] and is considered a simple method for showing the relationship (correlation) between FC and SC. When zero replacement is undesirable, using connection weights (the proportion of connections originating from the target region among all connections) can yield results similar to log-scaled SC (data not shown). It may be possible to compare various methods, but this is outside the scope of this study and requires further research.

The *C. elegans* studies presented by Reviewer #3 showed a weak correlation between FC and SC. However, *C. elegans* neurons do not fire and exhibited different calcium fluctuations depending on the region (Hendricks et al., 2012). This suggested that the cell body and various synaptic terminal regions have different FCs, which is consistent with the objective of our study (neuronal compartmentalization). If a functional unit is locally composed of multiple neurons and synapses, it is expected that SC and FC from that region will show a strong relationship. Larger regions would include multiple functional units, and a relationship between SC and FC would also be found, which is consistent with the results of our study. The *C. elegans* study compared FC of the cell body (a region) with SC of whole cell (not a same region), which would be inconsistent.

> *(2) Synaptic segregation on neurons can be topologically present even if pre- and post-synaptic synapses are present in similar regions of space, as an axon branch and dendrite branch can overlap in space but remain distinct along the arbor. The authors emphasize a region-based definition that does not reflect cellular anatomy. Moreover, the authors do not make an argument for their claim of better robustness of their new synaptic segregation measures.*



Author response image 2. Distance calculation for DBSCAN. a, Example synapse pair (black dot) of distance calculation. Red line shows the straight-line distance, and green line shows the morphology-based distance. DBSCAN will places two synapses in the same cluster based on straight-line distance, but they will be in different clusters based on the morphology-based distance.

Thank you for pointing this out. We changed from using DBSCAN based on the straight-line distance between synapses to DBSCAN based on the morphology-based distance via the branch nearest to the synapse (Author response image 2a). This resulted in a synaptic segregation measure that incorporates cellular anatomy. We updated all related figures, such as Figure.4, Ext. Data Figure.4-1, 4-2, 4-3, 4-4, Figure.5h. Also, we updated related text in the Results and Methods sections.

> *(3) Reviewers found the overall structure of the paper is difficult to follow, with sections appearing disjoint and the aims of different sections not well described. This extended to the paper organization as well, with the introduction not clearly setting up the questions and being distinct from the results. The manuscript would benefit from a clearer overarching narrative to unify the various analyses.*

Thank you for pointing this out. We have revised the introduction to make it more concise.

> *(4) Similarly, there are several descriptions of data and analysis that are unclear or lacking, including the source of the marmoset data and how the FlyWire synapse was subsampled.*

As pointed out by other reviewers, the marmoset data was obtained in [10], and the processing is also described in [10]. Therefore, we have revised the caption and removed the Ethics Declaration.

We have updated the Results and Methods sections regarding the extraction of "traced" neurons and synapses in FlyEM connectome data, and the extraction of post-synapses in hemibrain primary ROIs in FlyWire connectome data.

> *(5) Comparisons between FlyWire and Hemibrain have shown many similarities and some clear examples of inter-individual variability. There was concern that technical decisions with handling FlyWire synapse sampling were responsible for some of the differences observed between the datasets.*

In response to Reviewer #2's question, we answered that both FlyEM and FlyWire use proofread neurons and their connecting synapses. We also updated Fig. 3b and the Results and Methods sections.

> ***Reviewer #1 (Recommendations for the authors):***
>
> *The paper is written in an unusual way. It would be helpful if the introduction read more like a standard introduction. Describe the relevant background that the reader needs to understand the results that come later. Frame the experiments in terms of a question or hypothesis. Results should be relegated to the results section (or, if you like, a final paragraph that summarizes the findings). They should not be intermingled throughout the introduction.*

Thank you for pointing this out. We have revised the introduction to make it more concise.

> *The authors must be more attentive in terms of how they construct the segregated/unsegregated connectomes. I suggest exploring various thresholds/bins, but also considering proportionality thresholds that match the number of synapses.*

Thank you for pointing this out. As pointed out by other reviewers, we changed from using DBSCAN based on the straight-line distance between synapses to DBSCAN based on the morphology-based distance via the branch nearest to the synapse (Author response image 2a). This resulted in a synaptic segregation measure that incorporates cellular anatomy.

We also considered about the sparsity rates of the SC matrices. Since the sparsity rates of the null SC matrices differed a lot from that of the SC matrices extracted by cPPSSI, we regenerated the null SC matrices, shown in Fig. 4e and 4i. As shown in Author response image 1, we ensured that the extracted SCs fit within the null-generated matrices. This figure was added to Ext. Data Fig.4-5, and the main text was also revised.

> *The authors need a better solution for dealing with the zero entries in the sc matrix. Replacing the infinities with zeros and then running the linear regression to get an sc/fc relationship is not appropriate.*

Thank you for pointing this out. To investigate this issue, as pointed out by other reviewers, we compared the FC-SC correlation between the Gaussian resampled SC approach used in Honey et al. (2009) [6] and the log-scaled SC used in this study (Ext. Data Fig.2-2a). With a small number of ROIs, the sparsity rate was low (Ext. Data Fig.2-2b), resulting in less zero replacement. Therefore, log-scaled SC is likely to more accurately represent the relationship. Furthermore, with a large number of ROIs, the sparsity rate exceeds 70%, and resampled SC

randomly assigns a large number of zero elements from the smaller end of the distribution. This tends to lower the correlation (Ext. Data Fig.2-2c, d), suggesting that log-scaled SC provides fairer results. Using connection weights (the proportion of connections originating from the target region among all connections) can yield results similar to log-scaled SC (data not shown), because this matrix can also be very sparse. It may be possible to compare various methods, but this is outside the scope of this study and requires further research.

> *It would be useful to include a description of where the human/marmoset datasets came from. It would be useful to describe the processing of those datasets and whether they're comparable to how the fly data was processed.*

As pointed out by other reviewers, the marmoset data was obtained in [10], and the processing is also described in [10]. Therefore, we have revised the caption and removed the Ethics Declaration.

The pre-processing of fly calcium imaging data is described in the Methods section. Unfortunately, this processing method is not comparable to that used in humans/marmosets as it was highly customized.

> *The authors report sc/fc correlations for the human/marmoset datasets based on single papers. However, in the human case, especially, the strength of sc/fc correlations is highly variable. Not just based on number/size of parcels, but based on amount of data, processing pipeline, single-subject versus group averaged (incidentally, single-subject sc/fc is 'much'* lower than group-averaged, which has big implications for this study, where the fly datasets are, in essence, N=1 studies).*

Yes, there are numerous FC-SC correlation studies. We think Honey et al. (2009) [6] to be a highly representative study. It showed r = 0.39 to 0.48 for individual participants in 998 ROIs, and r = 0.36 for averaged one, but it increased r = 0.53 excluding absent or inconsistent structural connections. So, single-subject may not be much lower than group-averaged. Since the SC for a fly is an N=1 study, the FC-SC correlation for the same individual cannot be calculated. We think further research will be necessary.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *Abstract:*
>
> *Please introduce the term "ROI"*

Thank you for pointing this out. We have revised the Abstract.

> *Introduction:*
>
> *(1) On a general note: the introduction reads like an extended abstract (i.e., a mix of results and discussion).*

Thank you for pointing this out. We have revised the introduction to make it more concise.

> *(2) Line 43: Does this mean FC-SC correlation is higher in flies but not significantly so? Please clarify.*

We performed Mann-Whitney U test and it was not significant (p= 0.2667).

> *(3) Line 51: The "confidence" score does not indicate the degree of synaptic detection.*

In the NeuPrint user guide, *https://neuprint.janelia.org/public/neuprintuserguide.pdf* it states "confidence - The certainty that an annotated synapse is correct and valid." Since "degree of synaptic detection" may be difficult to understand, we changed it to "certainty of an annotated synapse."

> *(4) Line 59-61: This statement needs refining: post-synapses do not "receive" neurotransmitters, action potentials aren't conducted along nerve fibres.*

We changed “receive” to “sense.” About “action potentials,” we changed “conduct an action potential” to “graded potentials”, and removed “along nerve fibers.”

> *(5) Line 61: calcium activity as detected via GCaMP correlates with (electric) neuronal activity - please cite relevant GCaMP literature here.*

We added F. Helmchen and J. Waters, "Ca2+ imaging in the mammalian brain *in vivo*," Eur J Pharmacol., vol. 447, pp. 119-129, 2002.

> *(6) Line 76: "interconnected" is rather vague; just say "many Drosophila neurons are reciprocally connected".*

Thank you for pointing this out. Lin et al., (2024) showed motif analysis and there are many reciprocal, three-node and rich-club connections. However, introduction was updated and this sentence was removed.

> *(7) Line 77: comparing unsegregated vs reciprocal synapses is overly simplistic; these are separate features of the same object - i.e., a synapse can be reciprocal and at the same time be segregated in the presynaptic neuron but unsegregated in the postsynaptic neuron.*

Thank you for pointing this out. As you say, the relationship is complicated. In this paper, we are concerned with the degree of segregation of pre- and post-synapses, and we are looking at the segregation within a neuron. In this case, nearby reciprocal synapses (<=2 μm) are included in unsegregated synapses. We have made a correction to the sentence.

> *(8) Line 79: I don't understand how we get from unsegregated synapses to local activity.*

Retinal amacrine cells have extensive unsegregated synapses, which provide local feedback inhibition of burst inputs [34]. We changed the text around these descriptions.

> *(9) Line 80: What does "more essential function" mean?*

We removed this sentence.

> *(10) Line 85: "as shown earlier": Is this based on results in this study or prior work? See also the general above note on mixing results/discussion into the introduction.*

Thank you for pointing this out. We have revised the introduction to make it more concise.

> *(11) Line 85-87: I don't understand how the applicability of certain fMRI analysis methods in turn means that functional activity is locally compartmentalized. Did you mean to say something along the lines of "we applied common fMRI methods which showed functional activity is locally compartmentalized"?*

These sentences discuss the commonality between fMRI (BOLD signal) and calcium signal, which both represent presynaptic population dynamics within a local region (voxel). Furthermore, unsegregated synapses are widespread throughout the fly brain (Ext. Data Fig.4-2) and can also be observed in human pyramidal cells (Ext. Data Fig.1-2). Unsegregated synapses suggest local compartment activity [33, 34, 39, 40] and contribute more to functional activity (Fig.4b). Therefore, the similar trend in FC-SC correlation (Fig.2d) between humans and flies suggest that both species exhibit localized compartmental activity via unsegregated synapses throughout the entire brain.

Because these sentences contain many conclusions, they have been moved from the Introduction to the Discussion section.

> *(12) Line 87: Please provide a reference for "common among various species".*

Thank you for pointing this out. Because these sentences contain many conclusions, they have been moved from the Introduction to the Discussion section.

> *Results:*
>
> *(1) Line 91-92:*
>
> *(a) Please explain where the calcium data came from, how it was generated, etc.*

We added the data source and a reference (Brezovec et al. [14]).

> *(b) Please clarify: what registration method?*

This is not simple. Please see the Methods section and Ext. Data Fig.1-1. This is also indicated in the text.

> *(c) "calcium image" → "calcium image data"?*

We changed “calcium image” to “calcium imaging data”.

> *(d) What is the "FDA template"?*

This is a brain template created by Brezovec et al. [14]. JRC2018 is a well-known brain template, but it was created by immunostaining postmortem brains and did not fit well with calcium imaging data from living flies. Therefore, we used the FDA template.

> *(2) Line 93: Please introduce the term "ROI".*

We added “(Region of Interest)” in Line 38.

> *(3) Line 94: Ito et al., Neuron (2014) "A systematic nomenclature for the insect brain" is a better reference for Drosophila neuropils; for the hemibrain, the ROIs were generated to match that original atlas*

Thank you for pointing this out. We added a reference.

> *(4) Line 95/96: It is unclear what was used as the basis for the k-means/distance-based clustering*

This was because we wanted to investigate whether nuisance factor removal methods are robust, even for such diverse types of ROI. We added this point to the text.

> *(5) Line 120ff: I'm not sure how the total number of ROIs is relevant for comparing flies and humans, given (a) the huge difference in brain size and (b) the difference in resolution of the functional data.*

Indeed, the fly brain and the human neocortex are completely different. We are investigating whether there are commonalities between them using a metric called FC-SC correlation. As described in our answer for (11), both the fMRI (BOLD signal) and calcium signal represent presynaptic population dynamics within a local region (voxel). FC represents the synchronization of synaptic activity between regions, and SC represents the structural connectivity of neurons. Both flies and humans showed high SC-FC correlation and showed similar trends (Fig. 2d), so we believe it would be interesting to investigate this phenomenon.

> *(6) Line 123: "by contrast" is misleading here since, as you say, there isn't really a difference.*

We changed “by contrast” to “and.”

> *(7) Line 141: I'm somewhat worried that the differences between FlyWire and hemibrain synapse counts are due to the issues mentioned above.*

Thank you for the comment but we are not sure about “the issues mentioned above” is referring to.

> *(8) Line 148: There is no evidence that any differences in synapse are due to the resolution or anisotropy (as suggested in the introduction).*

We apologize that we don’t have direct evidence for it. We changed this to the sentence “This may be caused by differences in detection accuracy resulting from the resolution of EM scanning, but not to inter-individual variability.”

> *(9) Line 155: References "39,45" have no brackets.*

These are not referencing numbers, but brain regions of Brodmann area 39 and 45.

> *(10) Line 155-157: I don't think we can infer the composition of brain areas in humans based on a tenuous correlation in flies; this is highly speculative and really should be in the discussion.*

In humans, there are areas with strong and weak FC-SC correlations [8], which may be due to the E-I (Excitatory-Inhibitory) balance of connections. We investigated this possibility by comparing the correlation between neurotransmitters and FC-SC correlations in the fly brain. We slightly changed this sentence.

> *(11) Line 159: I find the first 2-3 sentences in this paragraph confusing. Are you saying that you did all these things in the prior results sections, or that you wanted to look at X and therefore you did Y? Maybe there is an issue with the tense here?*

We changed the sentences around this description.

> *(12) Line 161: "whole-brain" = FlyWire?*

We changed “whole-brain” to “FlyWire”.

> *(13) Line 163: Please explain the "PPSSI" acronym.*

This is now explained on Line 75.

> *(14) Line 165: The description of how the cPPSSI was calculated is hard to follow. For example, what's the "fraction of synapse number".*

We changed our sentences around this description to be clearer. The cPPSSI is the degree of segregation within a cluster and is also assigned to each synapse. The PPSSI is then the average of the cPPSSI values of all synapses in a neuron.

> *(15) Line 166: Is there a difference between "cPPSSI" and "PPSSI"?*

Yes, there is. Please see our answer for (14).

> *(16) Line 167: "The result showed a histogram resembling a normal distribution" → I suggest running a normality test.*

Thank you for pointing this out. We tested it by Lilliefors test and the result was p=0.001 (significantly not a normal distribution). Since there are numerous values with PPSSI=1, it is not judged to be a normal distribution. We therefore changed this description.

> *(17) Line 173: I am somewhat worried about a selection bias in your correlation of segregated vs unsegregated synapses. First, it seems like only a small fraction of neurons are in the 0-0.1 and 0.9-1 PPSSI range. I would suggest running a proper correlation between PPSSI and FC-SC correlation instead of looking at just the two extremes. Second, your examples for segregated neurons (APL + CT1) are large neurons that densely innervate spatially close and functionally very similar neuropils. If the sample of unsegregated neurons consists mainly of these large interneurons, I'm not at all surprised that they contributed strongly to FC-SC correlation.*

Thank you for pointing this out. For this work we investigated synapses (not neurons), extracting those with cPPSSI of 0-0.1 and 0.9-1, and performed a rank text with the FC-SC correlation of random sub-sampled synapses. We aimed to demonstrate that unsegregated synapses in particular, strongly contribute to FC-SC, and we hope to investigate overall trends in a future study.

> *(18) Line 185: I don't think the function of reciprocal synapses is "considered to be clear". There are examples of feedback inhibition through reciprocal synapses, in particular in the visual system, but that does not mean that this is true across the board.*

We changed “considered to be clear” to “considered to be clearer than unsegregated synapses.” Of course, the function of reciprocal synapses is unknown for the whole brain, but we think it is more well-studied than unsegregated synapses.

> *(19) Line 188 / Figure 4h: that figure panel does not appear to show transmitter pairs.*

Figure 4h (FlyWire) showed transmitter pairs. Ext. Data Fig.4-1g did not, because FlyEM does not have transmitter information.

> *(20) Line 192: Please clarify "functionally common".*

We changed our sentences to clarify this.

> *(21) Line 199: "ventral nerve code" → "ventral nerve cord".*

We fixed this typo.

> *(22) Line 201: I don't think you can use "conversely" here.*

We changed “Conversely” to “Moreover.”

> *(23) Line 201: How certain are you that the WAGN neuron is the only candidate? Also, it would be nice to provide the neuron IDs so that people can identify them in the connectome.*

Thank you for pointing this out. We added Root ID: 720575940644632087 in the text. Actually, we found several GABA neuron candidates, such as 720575940637611365, 720575940644632087, 720575940613552947, 720575940640333109 and 720575940612264817. We investigated whether ER1(L) was present in these downstream connections and found that 720575940644632087 had the strongest connection with the largest number of synapses, so we adopted this.

> *(24) Line 207: When you say "the left WAGN was strongly connected", are those connections not also present for the right WAGN?*

There is a right WAGN (Root ID: 720575940624377224), but it does not have strong interconnections with WPNb tier 2/3 (left) neurons. For the right WAGN, there are few inputs from WPNb tier 2/3 (left). We added “(left)” in the text.

> *(25) Line 212: I don't think you can use "however" here.*

We removed “however.”

> *(26) Line 214: "well unsegregated" → "very unsegregated"?*

This sentence was removed, because we recalculated Fig. 5h.

> *Ethics Declaration:*
>
> *It seems the marmoset data were reported on in [10], so why is there a reference to the generation of the dataset?*

Yes, marmoset data were reported in [10], so we removed the Ethics Declaration.

> ***Reviewer #3 (Recommendations for the authors):***
>
> *(1) In my opinion, the title and framing of this manuscript dramatically overstate the results presented here. Also, the results presented in the different figures in this manuscript seem disjointed and are not very related to each other.*

Thank you for pointing this out. We have rewritten our manuscript slightly to address this. Inter-individual variability is relevant to both SC and FC. Regarding SC, we think the difference in the number of synapses between the two individuals is due to the difference in detection power caused by differences in the resolution of the electron microscope. Regarding FC, as stated in the Results section, “Spatial smoothing is useful for absorbing inter-individual variability and conducting second-level group analysis.” Increasing the smoothing size improves the correlation and AUC between group-averaged FC and SC, indicating the presence of inter-individual variability in FC (Fig. 2b, Ext. Data Fig. 2-1b, especially when the number of ROIs is high). We added this text in the Introduction and Results sections.

> *(2) There are multiple ways to compute structural correlation matrices-the methods the authors implemented should be discussed in greater detail in the manuscript.*

Thank you for pointing this out. To investigate this issue, as pointed out by other reviewers, we compared the FC-SC correlation between the Gaussian resampled SC approach, used in Honey et al. (2009) [6] and the log-scaled SC approach, used in this study (Ext. Data Fig.2-2a). With a small number of ROIs, the sparsity rate was low (Ext. Data Fig.2-2b), resulting in fewer zero replacement. Therefore, log-scaled SC is likely to more accurately represent the relationship in our study. Furthermore, with a large number of ROIs, the sparsity rate exceeds 70%, and resampled SC randomly assigns a large number of zero elements from the smaller end of the Gaussian distribution. This tends to lower the correlation (Ext. Data Fig.2-2c, d), suggesting that log-scaled SC provides fairer results. Using connection weights (the proportion of connections originating from the target region among all connections) can yield results similar to log-scaled SC (data not shown), because this matrix can be also very sparse. The log-scaled SC aprroach has been used in previous studies [9, 68] and is considered a simple method for showing the relationship (correlation) between FC and SC. It may be possible to compare various methods in-depth, but this is outside the scope of this study and requires further research.

> *(3) The use of the FC-SC detection score defined by the authors should be discussed and justified more extensively in the text.*

Thank you for pointing this out. This has already been discussed in [10]. We defined our own "FC-SC detection score," but we consider the overall approach to be well established in the literature. For example, Stafford et al. (2014) carried out FC-SC detection for 168 mouse cortical regions, and obtained 78.26% sensitivity and 81.69% specificity for the top 1% of SC. Hori et al. (2020) also investigated FC-SC detection for 55 cortical regions of the marmoset brain left hemisphere, achieving an AUC of 0.72. We think FC-SC detection is an index that evaluates the relationship between FC and SC from a different angle than FC-SC correlation and is worthwhile.

Hori et al., (2020). Comparison of resting-state functional connectivity in marmosets with tracer-based cellular connectivity. NeuroImage, 204, 116241.

Stafford et al., (2014). Large-scale topology and the default mode network in the mouse connectome. Proc. Natl. Acad. Sci. U.S.A., 111(52), 18745-18750.

https://doi.org/10.7554/eLife.107990.2.sa0