



# Hierarchical communities in the larval *Drosophila* connectome: Links to cellular annotations and network topology

Richard Betzel<sup>a,b,c,d,1</sup>, Maria Grazia Puxeddu<sup>a</sup> , and Caio Seguina<sup>a</sup>

Edited by Terrence Sejnowski, Salk Institute for Biological Studies, La Jolla, CA; received November 16, 2023; accepted May 28, 2024

One of the longstanding aims of network neuroscience is to link a connectome's topological properties—i.e., features defined from connectivity alone—with an organism's neurobiology. One approach for doing so is to compare connectome properties with annotational maps. This type of analysis is popular at the meso-/macroscale, but is less common at the nano-scale, owing to a paucity of neuron-level connectome data. However, recent methodological advances have made possible the reconstruction of whole-brain connectomes at single-neuron resolution for a select set of organisms. These include the fruit fly, *Drosophila melanogaster*, and its developing larvae. In addition to fine-scale descriptions of connectivity, these datasets are accompanied by rich annotations. Here, we use a variant of the stochastic blockmodel to detect multilevel communities in the larval *Drosophila* connectome. We find that communities partition neurons based on function and cell type and that most interact assortatively, reflecting the principle of functional segregation. However, a small number of communities interact nonassortatively, forming a “rich-club” of interneurons that receive sensory/ascending inputs and deliver outputs along descending pathways. Next, we investigate the role of community structure in shaping communication patterns. We find that polysynaptic signaling follows specific trajectories across modular hierarchies, with interneurons playing a key role in mediating communication routes between modules and hierarchical scales. Our work suggests a relationship between system-level architecture and the biological function and classification of individual neurons. We envision our study as an important step toward bridging the gap between complex systems and neurobiological lines of investigation in brain sciences.

connectome | community structure | network neuroscience | network science | *Drosophila*

## Introduction

Nervous systems are fundamentally networks (1, 2). They are composed of neural elements—cells, areas, regions—linked to one another via synapses, axonal projections, and myelinated white matter. The complete set of neural elements and their pairwise connections defines a “connectome.” The configuration of the connectome as a network helps shape brain activity and function.

A popular strategy for analyzing connectome data is to represent them as a graph of nodes and edges (3). This simple model generally abstracts away neurobiological detail, but returns the backbone of structural interactions, which can be further analyzed using network science tools. Network science sits at the confluence of statistical physics, engineering, and mathematics, and offers a wide range of tools for summarizing and characterizing the structure and function of connectome data.

In principle, the network model is agnostic to scale; it is equally well-suited for representing large-scale connectivity (regions/areas linked by fiber tracts/projections) (4–6) as it is for representing cellular-level connectivity (synaptically coupled neurons) (7). Network analyses have identified a number of phylogenetically conserved architectural features of connectomes, including efficient processing paths coupled with greater-than-expected clustering (small-worldness 8–11), heterogeneous degree distributions and interlinked hubs (4, 12, 13), cost-efficient spatial embedding (14–16), and neurobiologically meaningful subnetworks or modules (17, 18).

To date, however, most connectome analyses have focused on the macroscale, as data can be acquired cheaply, noninvasively (diffusion MRI + tractography), and for the entire brain at a single-subject level. Indeed, very few connectome datasets exist at both the whole-brain and single-cell levels—the most notable being that of the nematode *Caenorhabditis elegans* (7, 19, 20). Recently, however, methodological advances have made it possible to reconstruct cellular-level connectivity for large volumes (21–28). Importantly, reconstructions of synaptic connectivity are accompanied by rich,

## Significance

Connectomes shape signaling and communication patterns. Understanding the principles by which whole-brain neuronal connectomes are organized is therefore critical to understanding brain function. Here, we focus on community structure—the propensity for networks to be decomposed into meaningful subnetworks. Though only connectivity information is used to detect these communities, we find that they have high levels of cell-type and functional specificity. Moreover, communication pathways through the network traverse communities through a highly stereotyped way; our analysis points to interneurons, specifically, as a key mediator of communication. This work, and others like it, pave the way for future nano-scale network neuroscience studies.

Author affiliations: <sup>a</sup>Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47401; <sup>b</sup>Cognitive Science Program, Indiana University, Bloomington, IN 47401; <sup>c</sup>Program in Neuroscience, Indiana University, Bloomington, IN 47401; and <sup>d</sup>Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455

Author contributions: R.B., M.G.P., and C.S. designed research; R.B. performed research; R.B. and C.S. analyzed data; and R.B., M.G.P., and C.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2024 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

<sup>1</sup>To whom correspondence may be addressed. Email: rbtzel@iu.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2320177121/-DCSupplemental>.

Published September 13, 2024.

high-dimensional sets of neurobiological annotations, which are seldom available for macroscale connectomes. These include details of cells' morphologies, types, and lineages, as well as putative functional assignments, thus making it possible to link high-level architectural features to fine-grained properties of single neurons at the same spatial resolution at which neurobiological processes unfold.

Here, we apply network science tools to the connectome of the larval *Drosophila melanogaster*. Our focus is on its community structure, which we uncover using an extension of the classical stochastic blockmodel (SBM) (29, 30). We find evidence of hierarchical community structure—communities within communities—whose sizes range from tens to hundreds of neurons. Community boundaries sharply delineate different cell types from one another and circumscribe groups of cells with shared functional profiles, e.g., proprioception, nociception, olfaction, and so on. Further, we show that while most communities interact assortatively—i.e. form internally dense and externally sparse subnetworks—a small fraction form nonassortative motifs and that these communities are enriched for interneurons. Next, we show that the larval *Drosophila* brain exhibits a “rich club”—a collection of hubs neurons that are also mutually connected to one another. We show that rich-club neurons, which tend to be interneurons, are present in most coarse-scale modules, but are also concentrated within a select set of “hub” modules, all positioned in the midbrain. Last, we investigate the link between community structure and communication policies, focusing on shortest paths and communicability. We found that neurons that were frequently assigned to the same community across hierarchical levels were also more likely to have greater pairwise communicability and reduced path length. Collectively, these results recapitulate, in a whole-brain nano-scale connectome, a set of architectural features that are largely conserved across phylogeny and scale.

## Results

Here, we analyze a previously published nano-scale connectome for the *D. melanogaster*. The connectome is composed of  $N = 2,952$  cells and approximately  $K = 352,611$  synapses.

**Hierarchical Community Structure.** Connectomes are thought to be modular, meaning that they can be decomposed into meaningful subgraphs referred to as “communities.” Here, we use a hierarchical variant of the SBM to partition the connectome into nested communities (30). Unlike modularity maximization (31) or Infomap (32), which are the most popular community detection methods in network neuroscience but only capable of detecting assortative community structure (internally dense and externally sparse subgraphs), SBMs use an inferential framework to detect generalized classes of communities, including core-periphery and disassortative motifs (29, 30, 33–35) [though note that the SBM can be constrained to detect purely assortative partitions (36)]. In addition, the SBM does not suffer from overfitting issues that permit methods like modularity maximization and Infomap to detect communities in random networks (37).

To obtain an estimate of hierarchical communities, we used the procedure described in ref. 38 to sample a large number of high-quality partitions (10,000 samples) and identify latent modes. We found evidence supporting the hypothesis that there is exactly one mode. Here, we characterize the consensus estimate of that partition. The optimal partition resulted in seven hierarchical levels, dividing the network into 2, 4, 6, 10, 20, 36, and 77 communities (Fig. 1 A–C). For the sake of visualization, we

focus on the fourth hierarchical level. In Fig. 1D, we visualize these communities, coloring neurons, and their morphological trees based on the community to which they were assigned. We note, however, that these ten communities can be both subdivided and aggregated further. We show in Fig. 1E examples of communities in the fourth hierarchical level that fracture into two or three smaller communities in the fifth level. For alternative visualizations of the different hierarchical levels and for a qualitative assessment of the link to synapse type-specific connectomes, see *SI Appendix*, Figs. S1 and S2, respectively.

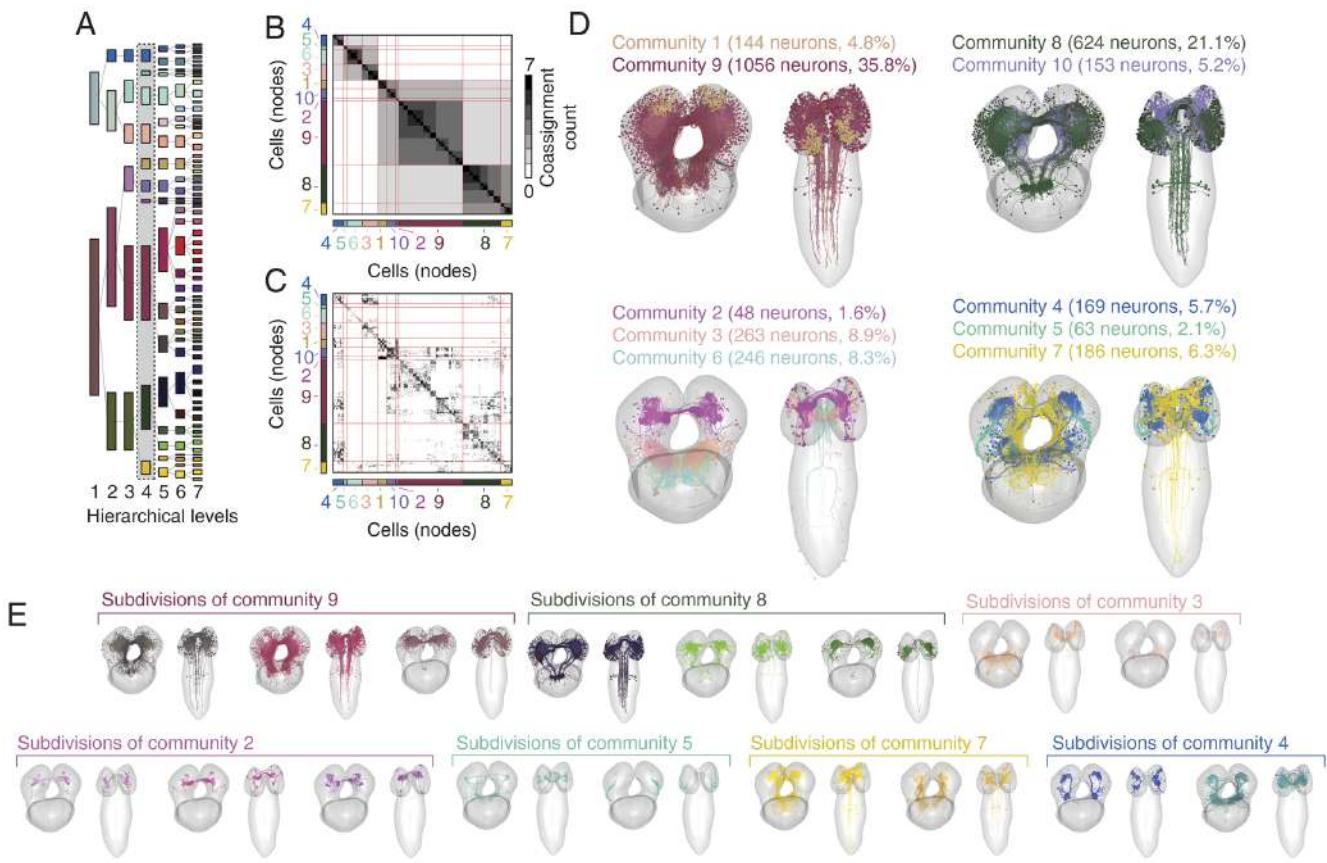
Next, we aimed to characterize the profile of communities across hierarchical levels. As expected, the size of communities—the number of nodes assigned to a given community—decreased monotonically with hierarchical level (Fig. 2A). In parallel, the synaptic and binary densities of communities increased across hierarchical levels (Fig. 2B and C), suggesting that these communities become more internally cohesive.

Next, we examined the spatial properties of communities. We found that the mean and maximum pairwise (Euclidean) distance between soma decreased near-monotonically across hierarchical levels (Fig. 2D and E). We note, however, that there were several outliers—communities whose diameter was far greater than the typical community. These communities were composed largely of sensory neurons in the nerve cord associated with respiratory, gut, and gustatory function (Fig. 2F and G).

In *SI Appendix*, we further investigated spatial properties of the connectome. One of the organizing principles of interareal connectomes is the exponential distance rule (EDR), wherein connection weight decays approximately as an exponential function with interareal Euclidean distance (39, 40). However, whether this rule holds at the nano-scale remains largely untested. To address this question, we examined how the distances between cells were related to their synapse counts (*SI Appendix*, Fig. S3). When the EDR is described in interareal connectome studies, distance is operationalized as the Euclidean distance between areal centers of mass. Accordingly, we considered soma–soma distance as a distance metric. However, synaptic contacts between neurons are often distant from their respective soma, such that two spatially adjacent neurons connect a long distance away from their cell bodies. Therefore, we also defined a second distance metric that takes into account synapse location. For a pre- and postsynaptic neuron, we calculated the shortest path through their arbors from their respective soma to the location of the synapse. We then calculated the wiring cost along the shortest path as a measure of distance. If two neurons were connected via multiple synapses, we identified the union of their shortest paths (the shortest paths backbone) and calculated the Euclidean distance along all arboreal segments comprising said backbone. See *SI Appendix*, Fig. S4 for a direct comparison of these two measures.

In general, we found that although connection probability decayed with soma–soma distance (*SI Appendix*, Fig. S3E), synapse count was not correlated with soma–soma Euclidean distance ( $r = -0.0015$ ; *SI Appendix*, Fig. S3C) but was positively correlated with the shortest path distance ( $r = 0.24$ ;  $P < 10^{-15}$ ; *SI Appendix*, Fig. S3D). We also found that the correlation between synapse count and wiring cost along the shortest path was cell-type specific (*SI Appendix*, Fig. S3G–I).

These results suggest that a simple exponential distance rule may not hold at the level of neuron-to-neuron connectivity. However, if we coarse-grain the network into parcels, we find that connection weights exhibit a negative relationship with distance—both soma–soma distance and wiring cost along the shortest path—that is modulated by the level of granularity. Coarse representations—tens or hundreds of clusters—exhibit



**Fig. 1.** Detected hierarchical modular structure. (A) Hierarchical dendrogram. Each column corresponds to a hierarchical level. Colors correspond to communities at the finest hierarchical level (level five). At coarser levels, they are grouped into larger communities. (B) Community coassignment matrix. Entries correspond to the number of levels in which pairs of neurons were assigned to the same community. Red lines separate level four communities from one another. (C) Connectivity matrix with rows/columns ordered by communities. Panel (D) depicts level-four communities in anatomical space. Communities can, in general, be subdivided or even aggregated into larger clusters. Panel (E) highlights divisions of select communities in the fourth hierarchical level into smaller subcommunities.

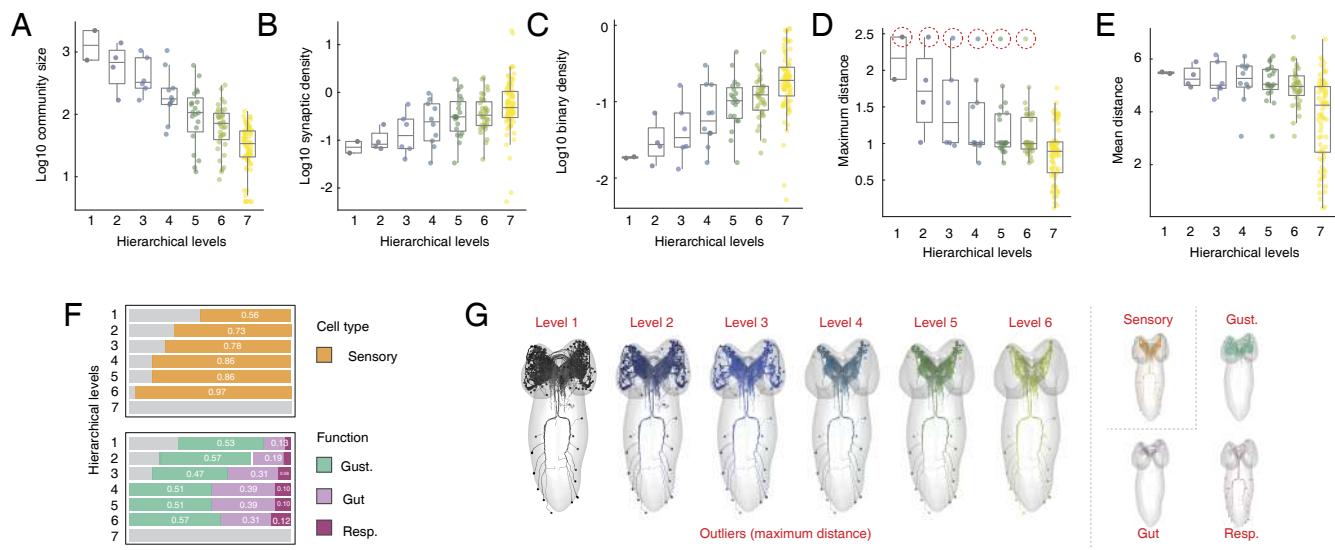
strong anticorrelations (*SI Appendix*, Fig. S3F). However, as  $k \rightarrow N$ , the correlation magnitude increases, approaching 0 in the case of soma–soma distance and  $\approx 0.24$  in the case of distance along the shortest path. These observations suggest that, at single-neuron resolution, the effect of straight-line distance on synapse count is small and possibly inappropriate—i.e. there are better measures of distance that are predictive of synapse count (41).

In addition to the SBM, we also tested a hierarchical variant of modularity maximization, the results of which we report in *SI Appendix*, Fig. S5. Despite differences in objective function and optimization heuristic, the communities detected using both techniques are similar (*SI Appendix*, Fig. S6), yielding correlated enrichment scores (*SI Appendix*, Fig. S7). We also compared the partitions obtained using the SBM with a subset of the partitions reported in ref. 21 (*SI Appendix*, Fig. S10). In general, we found that our solutions were neither perfectly aligned with those clusters, nor were they wildly dissimilar. Last, we also fit the connectomes comprising only axon→axon, axons→dendrite, dendrite→axon, and dendrite→dendrite synapses (*SI Appendix*, Fig. S11). Overall, we found that communities inferred from the axon→axon and axons→dendrite connectomes were most similar to the results described here and that the dendrite→dendrite connectome exhibited marked laterality.

**Linking Communities to Cell Types and Function.** In the previous section, we described a set of hierarchical modules and their properties. How do these modules and their boundaries

relate to cellular annotations? Do they “carve nature at its joints” such that modules circumscribe specific types of cells or functions? Or are annotations intermixed evenly across modules? At the meso-/macroscale, this question can only be approximately addressed by averaging cellular or population-level annotations to generate parcel-based maps (42). Here, however, we take advantage of the fact that communities and annotations are both defined at the single-neuron scale, allowing for direct comparison. Specifically, we asked whether connectivity-defined communities were “enriched” for different types of annotations: neurons associations with specific functions, cell types, and neuron class (e.g., input, output, interneuron) (43). We calculated the overlap of each annotation and community—e.g., the number of nodes labeled as Kenyon cells that were also assigned to community C and compared against a null distribution generated under a null model that preserves the spatial variogram (44) (Fig. 3A).

We found that communities exhibited both a high level of enrichment and also a high degree of specificity in terms of annotations (see Fig. 3B for raw overlap scores and Fig. 3F for normalized overlap, i.e. z-scores). This is interesting, as communities were defined only on the basis of synaptic connectivity using a data-driven algorithm. Virtually every function had a clear correspondence with a community. Community 9 was associated with proprioception; community 8 was dually associated with nociception and chordotonal mechanosensation; community 5 was associated with vision and gustation; communities 3 and 4 were jointly associated with gustation, thermosensation, and olfaction;



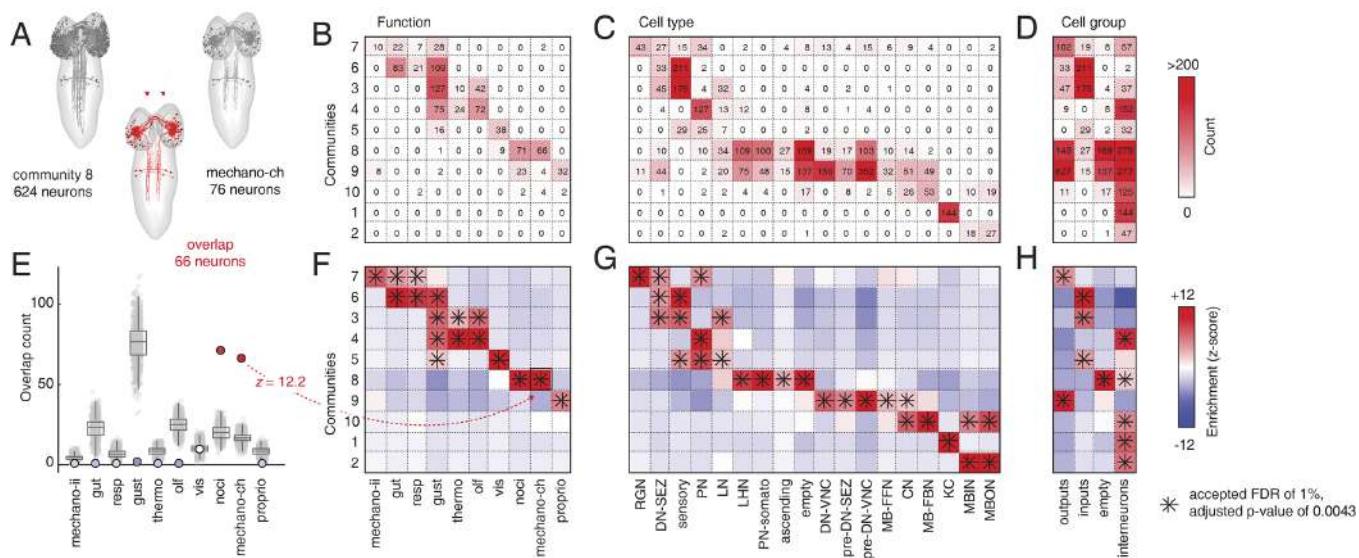
**Fig. 2.** Community statistics across hierarchical levels. (A) Community size (number of nodes). (B) Log synaptic density—number of synapses divided by total number of possible connections within each community. (C) Log binary density—number of connections divided by total number of possible connections. (D) Community diameter—longest distance (Euclidean) between pairs of nodes assigned to the same community. (E) Mean distance between pairs of nodes. At each hierarchical level, there were clear outliers in terms of maximum distance (red circles in panel D). In panel (F), we break down the cell type and functional composition of those communities. In panel (G), we show the anatomical configuration of those communities. In general, they are composed of ascending sensory projections that support gustatory, gut, and respiratory function.

community 6 was associated with gut, respiration, and gustation; community 7 was associated with class II mechanosensation, gut function, and respiration. Interestingly, communities 1, 2, and 10, which were among the smallest and spatially compact, were composed mostly of interneurons and not clearly aligned with any sensorimotor function.

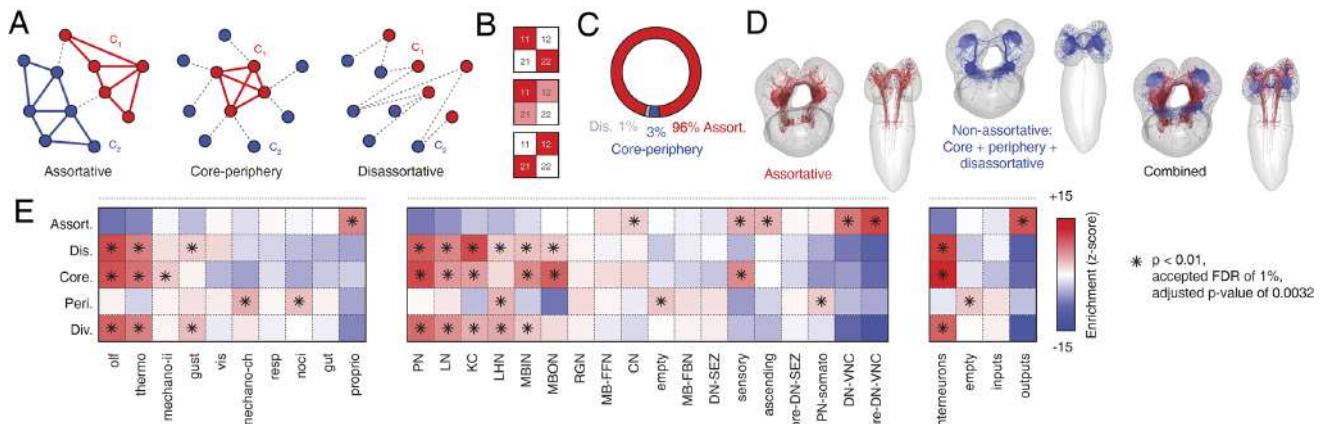
Similarly, cell types (and by extension, the broader cell classes/groups to which cell types were assigned) were also significantly enriched within communities. For example, communities 1 and 8, which were poorly aligned with functional annotations,

were highly enriched for different classes of interneurons, with community 1 aligned with Kenyon cells and mushroom body output neurons and community 8 aligned with mushroom body input, output, and fan-shaped neurons (see Fig. 3C and D for raw overlap; Fig. 3G and H for z-scored enrichment scores).

We also linked connectional properties of communities with function, class, and cell type. Specifically, we calculated “community motifs”, which describe how pairs of communities interact with one another (34). Every pair of communities can be represented as a  $[2 \times 2]$  matrix, whose diagonal elements



**Fig. 3.** Communities are “enriched” for function, cell type, and cell group. (A) Schematic illustrating how we measure overlap. Given two partitions of cells—e.g., one coming from data-driven community labels and another coming from annotation data—we calculate the overlap as the union of the two. Panels (B–D) show the overlap (counts) for each community and for different functional groups, cell types, and macro-cell group labels. (E) The counts are, in general, confounded by community and annotation size—larger maps will, just by chance, tend to have greater overlap with one another. To control for this, we use a space-preserving null model to calculate the expected overlap for each entry in the arrays depicted in panels (B–D), and z-score the observed overlap scores with respect to the null distributions. Panel (F) illustrates this procedure. This allows us to contextualize the scores shown in panels (B–D)—entries with greater overlap than expected signify that a community may be “enriched” for a given function, cell type, or macro-cell group. Panels (F–H) report “enrichment scores” of the raw counts depicted in panels (B–D).



**Fig. 4.** Community motifs distinguish function, cell type, and macro-cell group from one another. (A) Example community interaction motifs. Assortative interactions correspond to “internally dense—externally sparse” communities; “core-periphery” interactions involve a densely connected core that is linked to a sparse periphery; “disassortative” communities form cross-community links, seldom connecting nodes of the same community to one another. (B) Density matrices for the community interaction motifs discussed in (A). (C) Proportion of all two-community motifs classified as either assortative, core-periphery, or disassortative. (D) Anatomical depiction of the most assortative and nonassortative neurons and their arbors. (E) Enrichment (z-score) of assortative, disassortative, core, periphery, and diversity indices within function, cell type, and macro-cell group labels.

correspond to the within-community synaptic densities and the off-diagonal elements correspond to the synaptic density of incoming/outgoing connections. Based on these four elements, the interactions between every pair of communities can be classified as assortative, core-periphery, or disassortative (Fig. 4A and B).

Here, we categorize the interaction between every pair of communities and map these labels back to individual network nodes. This procedure is carried out independently for each hierarchical level, the scores ranked, and then averaged across levels to yield a single assortativity, core, periphery, and disassortativity score per neuron. Note that for core-periphery motifs, we distinguish between the community that acts as the “core” and the community that acts as the “periphery”. Based on these labels, we also calculated an entropy-based “diversity index” whose value is close to 1 if a node participates uniformly in all four classes and is close to 0 if it participates in only a single class. For the sake of visualization, we group the non-assortative labels (core, periphery, and disassortative) together to form a “non-assortative” class (Fig. 4C). We then performed another enrichment analysis, this time testing whether community motif values were significantly concentrated within functional, cell type, and cell class groups.

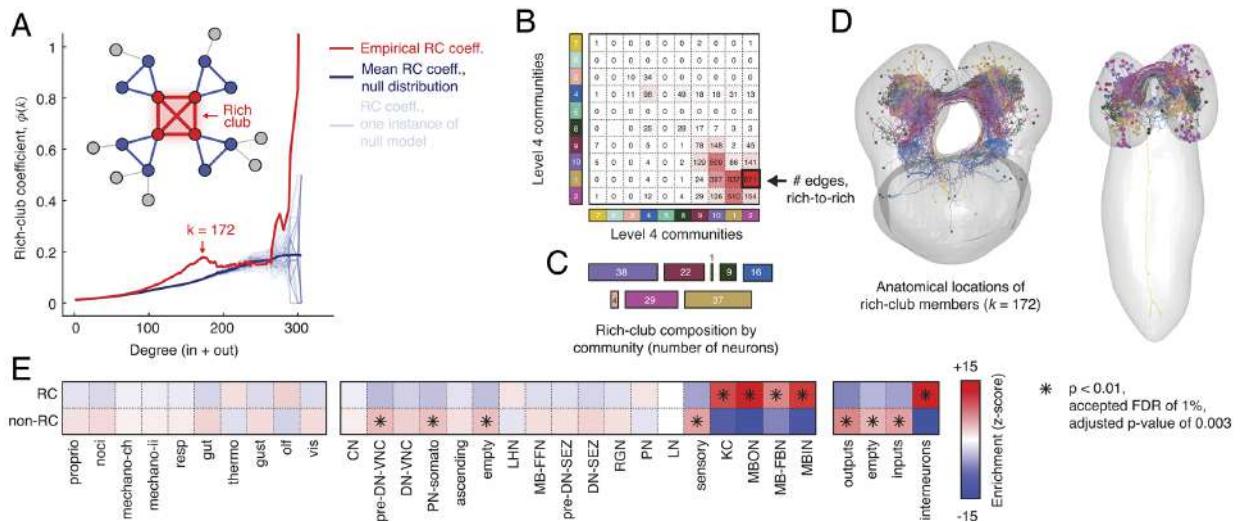
We found, in line with other applications of the stochastic blockmodel to brain network data, that most communities interact assortatively (34, 35). Specifically, 94% of interactions are assortative, with only 4% and 2% classified as core-periphery and disassortative, respectively (Fig. 4C and D). Interestingly, we found that motif classes were enriched within functional groups, cell types, and cell classes. For example, assortative motifs were highly enriched within output neurons, namely descending neurons in the ventral nerve cord associated with proprioception, nociception, and chordotonal mechanosensation (Fig. 4E). Note that this analysis only implies that the level of assortativity within these groups was greater than expected by chance; it does not imply that these groups were the only groups that participated in assortative community motifs. Conversely we found that interneurons specifically were enriched for the non-assortative community motifs—i.e. disassortativity and coreness—and were, in general, among the most diverse in terms of their motif-type participation (Fig. 4D and E).

Collectively, these results support the hypothesis that the *Drosophila* larval connectome exhibits hierarchical modular structure. These communities tend to be assortative and, though data-driven, neatly divide cells into groups based on their type, class, and function. Further, communities occasionally deviate from assortative interactions; the cells that make up these communities tend to be interneurons, supporting the hypothesis that the most assortative and segregated communities support specialized brain function.

**Rich-Club Structure.** Another hallmark feature of brain networks is that their degree distributions tend to be heavy-tailed, signifying that most neural elements maintain few partners, but that a small number make disproportionately many connections. In other brain networks datasets, these highly connected nodes sometimes form a “rich-club” (45–47), wherein hub nodes are more densely connected to one another than expected by chance (12, 13). Rich-clubs are thought to support intermodular communication—human studies have found that rich-club nodes are distributed across cortical modules (48), though other studies using blockmodels to define communities have shown that rich-clubs can form their own, separate community (49). Here, we test whether the larval *Drosophila* connectome exhibits a rich-club and, if so, assess how it interlinks communities to one another and its relationship to known functional classes and cell types.

Specifically, we calculated the directed rich-club coefficient,  $\phi(k)$ , at every degree,  $k$ . We then repeated this procedure for 1,000 randomized networks whose in-/out-degree sequences were identical to that of the original network but where the connections were otherwise formed at random (50). Then, for each value of  $k$ , we calculated the nonparametric  $P$ -value as the fraction of randomized networks whose rich-club coefficient was equal or greater to that of the original network (Fig. 5A). We identified a range of statistically significant rich-clubs, but focused on the local maxima in the normalized rich-club coefficient at  $k = 172$ .

We found that rich-club nodes were distributed across 8/10 communities at the fourth hierarchical level (albeit only barely; community 7 included only 1 rich-club node rich-club connection) (Fig. 5 B–D). However, rich-club nodes were largely concentrated within four communities (1, 2, 9, and 10), whose constituent members were among the most connected



**Fig. 5.** Rich-club structure. (A) Empirical rich-club coefficient (red) compared to randomized null models (blue). We focus on the statistically significant rich-club at  $k = 172$  (combined in-/out-degree). (B) Concentration of rich-club connections (connections that link a rich-club node to another rich-club member). (C) Composition of rich-club by communities. Colors correspond to level two communities; area is proportional to the percentage of the rich-club that comes from each community. (D) Anatomical representation of rich-club nodes and their arbors. (E) Enrichment of function, cell type, and macro-cell group within rich-club labels (part of rich-club and not part of rich-club).

in the network as indexed by degree (*SI Appendix*, Fig. S8). We also tested whether rich-club membership was linked to specific function, cell types, or cell groups using the same enrichment analysis described in the previous section. We found that interneurons, specifically Kenyon cells and a subset of cells in the mushroom body [input, output, and feedback neurons (51)], were preferentially associated with rich-club status ( $P < 10^{-15}$ ; Fig. 5E).

We also used information about synapse type to further decompose and characterize the rich-club (*SI Appendix*, Fig. S12). We found that “rich” connections (links between two rich-club neurons) were most likely to be axon→axon and axon→dendrite synapses, reflecting the fact that these two synapse types were also the most common (*SI Appendix*, Fig. S12B). However, when we controlled for baseline rate, we found that axon→axon synapses were significantly overrepresented while axon→dendrite synapses were significantly underrepresented (permutation test;  $P < 0.01$ ). Interestingly, for more exclusive rich-clubs ( $k > 172$ ) we find evidence that dendrite→axon and axon→dendrite synapses are slightly overrepresented relative to their baseline rate (*SI Appendix*, Fig. S12C), suggesting that these synapses may be important for signaling among highly connected, hub neurons.

Collectively, these results indicate that the *Drosophila* connectome exhibits rich-clubs—groups of highly connected cells that are also connected to one another. Here, the rich-club was detected in a data-driven way but overlaps with known cell types that have been linked to associative learning and memory (52).

**Role of Network Modules in Communication Processes.** Connectomes represent the pathways along which signals propagate. Communication between neurons can be understood as the process by which a “signal” from a source node reaches a prespecified target (53, 54). These types of processes can be modeled using tools from network science. Typically, communication models are situated along a spectrum, ranging from centralized processes like shortest paths routing, in which the message follows the shortest possible path from its source to target, ensuring maximum efficiency, to decentralized processes like diffusion

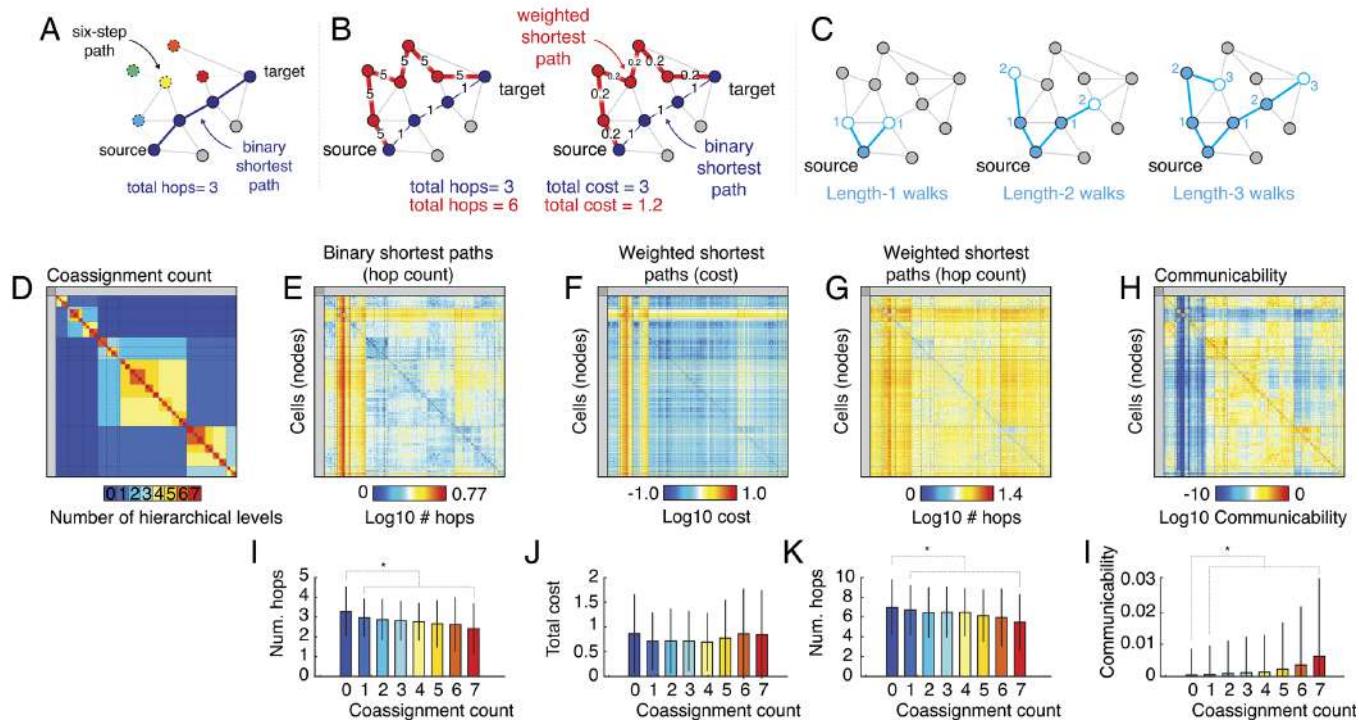
(55, 56), navigation (57), and cascade models (58). Here, we explore communication processes unfolding over the *Drosophila* connectome. Although the set of possible communication policies is vast (59, 60), we restrict our analyses to the following three: “binary shortest paths” (Fig. 6A), “weighted shortest paths” (Fig. 6B), and “weighted and directed communicability” (Fig. 6C). See *SI Appendix* for more details on these measures.

We calculated these measures for every pair of nodes, generating  $N \times N$  network communication matrices. For every pair of nodes, we calculated the number of hierarchical levels in which they were assigned to the same community (Fig. 6D) and linked those values to the three aforementioned measures (Fig. 6E–H).

In general, we found that if nodes were assigned to the same community at any level of the hierarchy, their shortest path length and communicability tended to be less than nodes that were never in the same community ( $t$  test,  $P < 10^{-15}$ ; Fig. 6 I, K, and L). Note, however, that we did not find this effect when we examined the cost matrix (Fig. 6J). We also found that hop distance and communicability decrease and increase monotonically with hierarchical levels, respectively, such that nodes assigned to the same community across all hierarchical levels tend to be connected via fewer hops and their walk density (indexed by communicability) was larger compared to those that appeared in communities less frequently ( $t$  test;  $P < 10^{-15}$ ; Fig. 6 I–L). Taken together these results recapitulate well-established links between network modules and communication in the larval *Drosophila* connectome.

**Linking Shortest-Path Trajectories to Community Hierarchy.** In the previous section, we showed that community hierarchy is, in aggregate, related to measures of communication. Here, we investigate those relationships in greater detail, focusing on shortest path trajectories from source to target nodes, detailing edge usage relative to nodes’ positions in the community hierarchy.

Specifically, we calculated the module coassignment matrix so that every pair of nodes was assigned a value between 0 and 7 depending on the number of hierarchical levels in which they were assigned to the same module. We then “masked” the coas-



**Fig. 6.** Linking processing paths to hierarchical community structure. (*A*) Schematic illustrating example path and shortest path between source and target nodes in a binary network. (*B*) Example of shortest weighted path. Note that weights are typically transformed from measures of affinity to measures of cost in the estimation of shortest weighted paths. Note also that the shortest weighted path may include more hops than the shortest binary path. (*C*) Illustration of walks of different lengths. Communicability counts and sums the number walks of all lengths between pairs of nodes, exponentially discounting the contributions from longer walks. (*D*) Hierarchical community coassignment matrix. Elements range from 0 (pairs of nodes never assigned to the same community at any level of the hierarchy) to 1 (assigned to the same community at every level). Panels (*E–H*) depict the binary shortest path, weighted shortest path (cost), weighted shortest path (hop count), and communicability between all pairs of nodes. For binary and weighted shortest paths, we find mean, mode, and maximum shortest paths of 3.28, 3, and 15 and 5.97, 5, and 23 hops, respectively (note that for the weighted shortest path, the quantity to be minimized is not hop length, but a measure of cost estimated from connections' weights). Panels (*I–L*) compare mean hop count, weighted shortest path (cost and hop count), and communicability between nodes at different levels of the community hierarchy.

signment matrix with the binarized connectivity matrix, setting to zero all entries in the coassignment matrix corresponding to pairs of nodes that were not directly connected (Fig. 7*A*). Using this relabeled matrix, we tracked how often and where edges with labels 0, 1, 2, 3, 4, 5, 6, and 7 appeared in shortest paths. Finally, for paths of all lengths,  $L$ , we calculated the typical trajectory with respect to edge labels, which reveals how shortest paths travel across modular hierarchies on their way from a source to a target cell (Fig. 7*B*). Note that this approach is similar to the model studied in ref. 21, wherein “seed neurons” probabilistically activate their postsynaptic partners, creating “cascades” of activation that propagate across the connectome. Using that model, the authors characterized the timing of activations. Here, we focus on shortest path structure, which represents an extreme case of the cascade model, corresponding to a 100% probability of activating postsynaptic partners. A further distinction between shortest paths and the cascade model is that neurons in the cascade model can enter a “deactivated” state following its own activation, wherein the neuron cannot activate its postsynaptic neighbors.

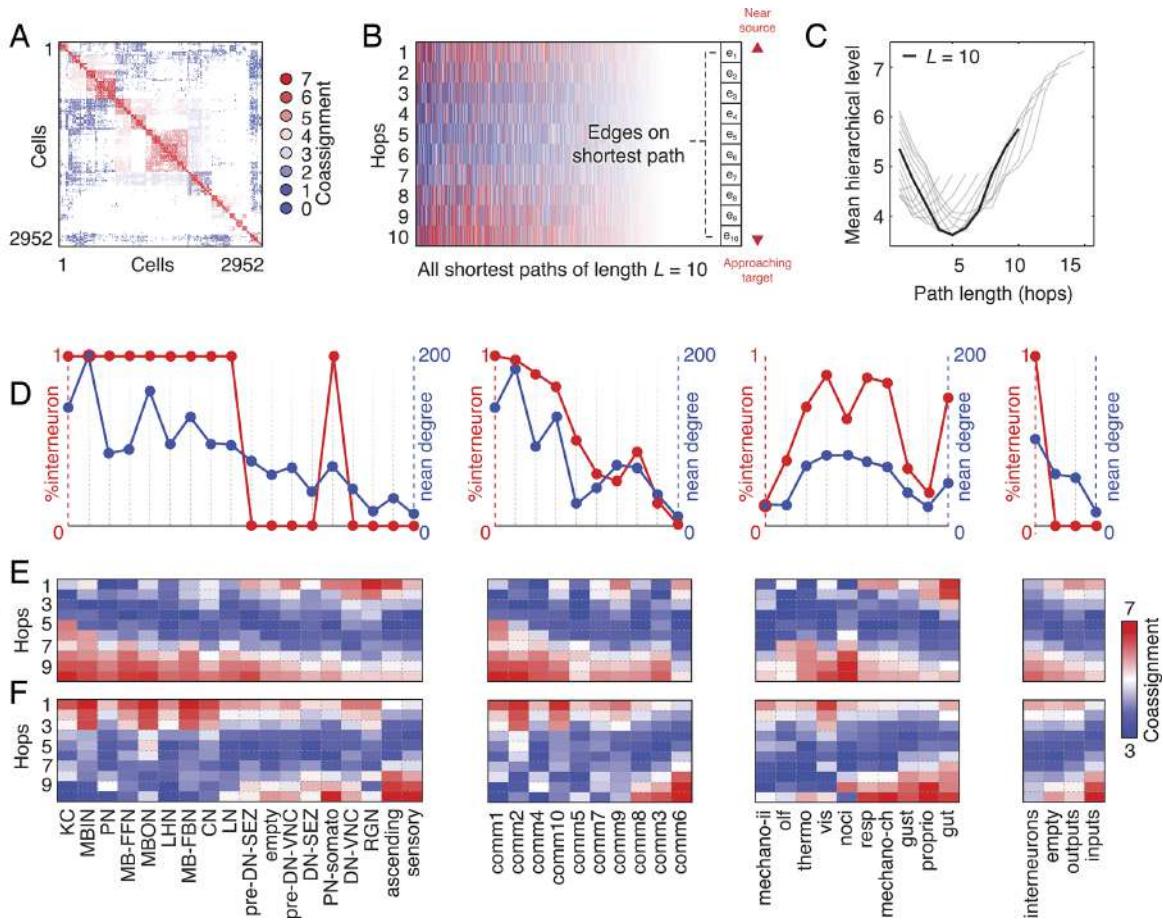
We found that, irrespective of path length, shortest path trajectories tend to be initiated using edges that link to nodes in the same community as source node, advance to cross-community edges, before once again traveling along within-module connections near the vicinity of the target node. These trajectories give rise to a characteristic “u-shape.” However, these characteristic trajectories are estimated by averaging over all pairs of source and target nodes, collapsing across considerable

variation. How much do trajectories vary when the source and target nodes are selected based on specific annotational properties? To address this question, we estimated typical trajectories when the source/target nodes have different cell types, communities, functional profiles, and macro-cell groupings (Fig. 7 *C–E*).

Interestingly, we found heterogeneity across source/target types. For instance, Kenyon cells (KC)—a class of interneuron—deviated considerably from the characteristic u-shaped trajectory. As sources, they followed an increasing and near-linear trajectory, so that the initial step in the shortest path, rather than linking to another node in the same module, immediately takes advantage of cross-community connections. The rest of the shortest path is composed of an “ascent” toward the target, using edges that were increasingly likely to link nodes in the same community (Fig. 7 *D, Left*). We observed a similar, but opposite, trajectory for shortest paths where Kenyon cells were the target (Fig. 7*F*).

In contrast, neurons associated with sensation (sensory neurons in Fig. 7 *D* and *E*) followed a trajectory opposite that of Kenyon cells. These cells, along with descending and ascending neurons in the ventral nerve cord, are enmeshed within segregated and assortative modules. As sources, the first steps on their shortest paths tend to be to other cells in the same group, eventually reaching cells that make cross-module connections and ascending via within-module connections toward the target neuron.

What explains the variation in shortest path trajectories? One hint comes from the trajectories of cell groups. Interneurons, which include KCs, follow monotonically increasing and decreasing pathways when they are grouped by source and target,



**Fig. 7.** Shortest path trajectories. (A) Masked connectivity matrix. Edges are labeled based on coassignment probability: 0, 1, 2, 3, 4, and 5 indicate the number of hierarchical levels that a given pair of connected nodes were assigned to the same community. (B) We can then describe the composition of the shortest path from a given source node to a given target node in terms of the coassignment probabilities of the edges along that path. As an example, we show all shortest paths of length  $L = 10$ . (C) We then calculated the typical trajectory as the mean across all shortest paths, which revealed a characteristic u-shaped trajectory. That is, the first hops away from a source node tend to be to other nodes in the same community as the source node, followed by a series of cross-community hops, before terminating through a series of within-community hops as the shortest path approaches the target node. Panel (C) shows the characteristic trajectory averaged over all source/target pairs for  $L = 2$  to  $L = 15$  (there were very few shortest paths beyond this length and, for brevity, we ignore them). We then asked how these shortest path trajectories varied across cell types, communities, functional annotations, and macro-cell groups. We found that, while most shortest paths follow a general u-shape, a subset deviate from this shape. For instance, shortest paths originating from Kenyon cells almost immediately begin using between-community connections to leave their local community, followed by a quick “ascension” along edges that fall within communities (D). We find an opposite trajectory when we average only shortest paths that end with Kenyon cells (E). These effects can be explained by two observations. Kenyon cells are interneurons and, like other interneurons, tend to be high degree. We find that groupings that deviate from the u-shaped trajectory are both dominated by interneurons and exhibit high average degree (C). Panel (F) shows average trajectory for each annotation, where each annotation is a source. Panel (F) shows the same information, but where each annotation is a target.

deviating from the characteristic u-shaped trajectory (Fig. 7B). Indeed, we find that the percentage of interneurons in each group, with the exception of functional annotations, is largely in agreement with the ordering of trajectories, from those that are similar to the trajectories of Kenyon cells to those that are more u-shaped (Fig. 7C, red curve). This observation suggests that the cell type composition of groups, namely the fraction of cells labeled as interneurons, helps determine the extent to which their respective trajectories deviate from the grand average.

What network property (properties) do interneurons exhibit that might explain this phenomenon? Interneurons tend to be high degree—i.e. they maintain relatively high numbers of incoming/outgoing connections. This aligns with the earlier observations that rich-clubs and core-periphery structures are enriched for interneurons. From the perspective of shortest paths, by virtue of being high-degree and nonassortative, interneurons can use intermodular connections to rapidly hop from their own module to another (unlike input and output neurons, which form largely assortative structures and whose initial steps

along shortest paths tend to be to other neurons in the same module).

As a final analysis, we assessed to what extent the u-shaped trajectories could be attributed to community structure alone, or if their shape and variation revealed something distinct about the organization of the connectome. To address this question, we sampled networks from the fitted SBM (100 repetitions). For each sampled network, we repeated the above analyses. We found that the shortest paths of the sampled networks also exhibited u-shaped trajectories consistent with those previously reported (*SI Appendix, Fig. S14*). Further, we also find that, given these null models, we recover the functional, cell type, group type, and community specificity of shortest paths.

Collectively, our findings suggest the u-shaped trajectories and their specificity are directly related to the nested community structure. That is, our null models preserve topological features of the network-community labels exactly and nodes’ in-/out-degrees approximately but are otherwise not directly informed by neurons’ annotations. However, the results of our study suggest

that many topological features of the larval fly connectome are also correlated with annotations; communities, rich clubs, and local connectivity properties (neurons' degrees) are related to function/cell type/group type. Accordingly, in preserving these topological properties, we are likely inadvertently allowing information about annotations to "leak" into our analysis, further underscoring the need for the development of novel null models for disentangling connectivity data from annotations.

These observations, which are directly in line with previous large-scale connectome studies (53, 57, 61, 62), suggest that the connectional properties of interneurons relative to other cell types may endow them with the ability to rapidly transmit signals to other downstream modules while readily integrating signals from upstream sensory units.

## Discussion

Here, we analyzed the larval *Drosophila* connectome. Our focus was on its community structure—i.e. divisions of the whole network into smaller subnetworks. We presented evidence that the network exhibited hierarchical and largely assortative communities. We also showed that the communities, which were defined based only on connectivity alone, delineated functional groups, cell types, and macro-cell groupings from one another. We reported evidence of a rich-club, comprised largely of high-degree interneurons. Although rich-club members were distributed across all communities (at a coarse scale), they tended to concentrate within a small subset of communities. We showed the community boundaries were related to the ease of communication between neurons, and that in delivering a signal from one neuron to another, shortest paths traversed the hierarchy of communities in a specific way, but deviated when the source/target groups contained high percentages of high-degree interneurons.

**Communities Reflect Cell Types and Function.** Community structure is one of the most studied properties of networks. It has been given proper mathematical treatment for at least half of a century, with early studies in sociology focusing on detecting communities using blockmodels fit to social network data (63–67). The interest in finding latent cluster structure in network data has continued to the present day (31, 32, 68), extending far beyond the social sciences, and, with new approaches and insights, community detection and analysis continues to be a quickly evolving subdiscipline within network science (69, 70).

In network neuroscience, especially, community structure has played a central role in shaping our understanding of brain network organization (17, 19, 71–75). At the large scale—where most network neuroscience applications have occurred—communities are generally thought to reflect functionally related groups of neural elements (76–78).

However, establishing direct links between modules and annotational data—e.g., cytoarchitectonic properties or cell types—has been less successful at the large scale. Only recently has a framework for curating, sharing, and comparing brainmaps become available (42). Even in network science proper, there has been skepticism as to whether detected communities have any basis in "ground truth" community annotations, with studies reporting only modest alignment (70, 79).

Here, and in line with (21), we show that communities estimated based only on connectivity nonetheless are "enriched" for distinct modes of function, cell types, and cell groups. This means that cells carrying these annotations are concentrated within particular communities at a level unexpected by chance.

Critically, we show that at coarse scales the enrichment patterns are nearly one-to-one, so that if a particular functional or cellular annotation is enriched in community "X," it tends to not be enriched in other communities. These findings establish a correspondence between community structure and neurobiological annotation data, reifying the intuition and hypothesis that connectivity-defined communities reflect function and cytoarchitectural properties at the microscale.

One interesting extension would be to incorporate annotation data directly into the generative model as covariates, thereby impacting connection probabilities (80, 81). Indeed, recent extensions of the SBM have made this possible by including annotations as a fixed set of parameters (82, 83). To tease apart which annotations "drive" the formation of communities, one could systematically "lesion" different annotations from the set of parameters used to fit the blockmodel, assessing how the exclusion of different annotation types hinders goodness of fit measures.

**Communities Are Mostly Assortative.** Typically, when the term "modularity" is invoked it is often used to refer to "assortative" community structure—i.e. divisions of a network into segregated subnetworks. This type of structure is thought to support specialized function (84), promote evolvability (85), adaptability (86), robustness to perturbations (87), separation of dynamical timescales (88), and allow for efficient embedding of the network's elements and wiring in three-dimensional space (14).

Indeed, empirical studies of brain networks have consistently revealed precisely this type of organization. On the one hand, these observations could be viewed as evidence that, from the perspective of embodied nervous systems, assortative communities are functionally adaptive features for all (or just some) of the reasons mentioned earlier. On the other hand, the preponderance of assortative modules could also reflect biases in network construction—e.g., correlation-based metrics of functional coupling that might artificially reinforce assortative groupings (89)—or biases in community detection. The de facto methods in network neuroscience—modularity maximization and Infomap—explicitly seek assortative communities. That is, even if other types of community structure were present in a network, these methods are incapable of detecting it.

SBMs, though not without limitations (90), are capable of detecting generic community types, including assortative structure (29), and therefore offer a useful framework for assessing evidence of assortativity in brain networks. If there is a preponderance of statistical evidence supporting core-periphery or disassortative communities, then the SBM will recover said communities.

Here, we use a hierarchical variant of the SBM to detect communities. We find that, overwhelmingly, communities interact assortatively. This observation is in line with other studies that have applied SBMs to connectome data (34, 35, 49). Here, also in line with large-scale imaging studies (33), we find that the most assortative communities are associated with (pre)descending neurons involved in proprioceptive, nociceptive, and mechanosensory function—i.e. sensation and perception.

Although most community interactions are assortative, a small fraction (about 4%) are nonassortative. These motifs represent deviations from the near-uniform assortative interactions observed across the larval *Drosophila* nervous system and cross-link communities to one another. These deviations also demand functional explanations; while assortative communities are thought to support specialized sensory processing, what are the functional roles of disassortativity and core-periphery structures? Future modeling studies should investigate these and related questions in

greater detail. An interesting possibility not explored here is that core-periphery structure emerges from overlapping assortative communities—i.e. if nodes are allowed to participate in multiple assortative communities, a group of nodes with comembership to many of said communities can appear “core-like” (91). In this case, the core-periphery structure observed here may be consistent with the hypothesis that neurons are organized into functionally specialized assortative communities, with the caveat that neurons exhibit mixed memberships. Future studies should investigate this alternative explanation.

**Hubness of Interneurons Gives Them Unique Network Properties.** Heterogeneous and heavy-tailed degree distributions are hallmarks of real-world networks (92). These distributions imply that a small and exclusive subset of nodes make disproportionately many connections, embedding them in influential positions within the network.

Indeed, early network analyses of biological neural networks revealed the presence of putative “hub” nodes (4), an observation that was further refined when it was discovered that hubs link together, forming “rich-club” structures (12). These phenomena are hardly a product of networks reconstructed from magnetic resonance images, and have been observed across spatial scales and phylogeny (47, 93–96).

Here, we find that the neuronal connectome of the larval *Drosophila* exhibits a heavy-tailed degree distribution and that high-degree neurons are comprised largely of interneurons situated in the mid-/central-brain. By virtue of making many connections, these neurons are conferred a number of unique structural (and possibly functional) properties. For instance, high-degree interneurons are the most likely to participate in nonassortative community motifs and rich-clubs. Their status as highly connected units also positions them as key nodes along putative communication pathways (shortest paths), facilitating efficient and increasingly direct routes to downstream target neurons, while also making themselves easily accessible targets for pathways originating in other communities (61).

**Intracommunity Communication Is More Efficient than Intercommunity Communication.** Understanding how the configuration of a connectome’s edges shapes the flow of signaling has been a central aim of network neuroscience (97, 98). Recent work has begun to address this question using network-based models of communication (57, 58)—stylized processes for delivering a “signal” from a source node to prespecified target.

Here, we show that the effect of community structure is imprinted on the efficacy of communication processes—both centralized and decentralized. Specifically, we find that nodes consistently assigned to the same module across hierarchical layers are likely to be connected via shorter paths and exhibit greater “communicability” compared to nodes that are infrequently or never assigned.

Combined with the observation that communities are well aligned with functional annotations and the observations made elsewhere that communication measures are strongly correlated with the magnitude of functional coupling between neural elements (55, 56), our findings position constraints imposed on communication by community structure as a key determinant of a neuron’s functional repertoire.

**Multiscale Network Neuroscience.** For the past two decades, network science has permeated virtually every scientific discipline (99). Part of its success is owed to the generality of network

models; a system’s details are abstracted away, leaving behind a set of circles and lines to represent the system’s elements and their interactions.

This model has proven profoundly useful in neuroscience in applications to interareal connectome data (4–6, 100). These analyses have identified core sets of phylogenetically conserved architectural features, including small-worldness (10), hubs and rich clubs (4, 12), modular structure (101), and wiring cost reduction (39, 102, 103). Though referring to static architectural properties, these features are often interpreted in terms of brain function; modules for specialized and segregated information processing, hubs and small-worlds for integration, and wiring cost as a constraint that limits the total material and metabolic expenses of the brain.

However, due in large part to the paucity of whole-brain, neuron-level connectome data, whether similar organizational principles are evident at the microscale remains unclear (104, 105). As we enter the era of nano-connectomics, it is possible to not only assess whether features described in other scales are evident (106), but to understand altogether new network phenomena (107). Ultimately, this approach holds promise for effectively bridging scales. Starting from MRI data, the smallest characterizable unit is the voxel or surface vertex; probing features at finer spatial scales is impossible. On the other hand, nanoscale data can be coarse-grain to the scale of voxels ( $\approx 1$  mm), presenting an opportunity for truly multiscale network models (108).

**Future Directions.** The focus of this paper was to link network communities derived from synaptic connectivity with neuronal annotations. One of the challenges associated with this type of analysis is adequately addressing, from a statistical perspective, the nested nature of the detected communities as well as the annotations. For instance, we found that community 5 was significantly “enriched” for sensory neurons. That is, the community was composed of more neurons with the “sensory” label than expected under the null model. However, we also found that community 5 was enriched for the labels “gustation” and “vision,” which are nested within the broader “sensory” label. To what extent should we anticipate this second result—enrichment for specific functional annotations—given that the same community was enriched for the “sensory” label? It is straightforward to construct counterexamples where the second outcome does not necessarily follow the first. For instance, a community could be enriched for “sensory” neurons but with each subcategory represented exactly proportional to its baseline rate. While there exist frameworks for dealing with nested hypotheses (109, 110), their application to large datasets with multiple levels of nestedness is not straightforward. With the proliferation of nano-scale connectome data (111) and increasingly rich and nested annotations of both neurons and their connections (112), this statistical issue presents a serious barrier. Future studies should focus on the exploration of frameworks for addressing this challenge.

A second important consideration for future studies concerns the enterprise of community detection and its role in network neuroscience. Community detection algorithms vary across multiple dimensions and, in general, will yield dissimilar estimates of community structure. We highlight an example of this here when we compare SBMs and modularity maximization; the SBM partitions suggest that communities are not strictly assortative, while modularity maximization is restricted to detecting assortative structure. We also compare our SBM partitions, which were derived using connectivity information alone, with partitions from ref. 21, who used spatial information (hemisphere

labels) in their clustering algorithm. In all cases, we found evidence for convergence across algorithms; communities were not identical, but had considerable overlap. Nonetheless, carrying out detailed comparisons of community detection algorithms may not be straightforward in future studies; the increased dimensionality of nano-scale connectomes limits the application of computationally complex community detection algorithms. Relatedly, algorithms that incorporate meta-data and annotations run the risk of limited generalizability, i.e. they can only be applied to connectomes that have the same/similar meta-data. In general, these differences and considerations encourage the exploration and reanalysis of the same dataset using multiple approaches. Indeed, when viewed through the lens of statistical inference, there exists a many-to-one mapping of communities to connectomes (113), such that very different community structure can offer equally good/bad descriptions of a network depending upon the exact function that maps community labels to connectivity.

## Materials and Methods

**Dataset.** We analyzed the larval *Drosophila* connectome as published in ref. 21. The complete connectome included a giant strongly connected component of

1. D. S. Bassett, O. Sporns, Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).
2. O. Sporns, The human connectome: A complex network. *Ann. N. Y. Acad. Sci.* **1224**, 109–125 (2011).
3. M. Rubinov, O. Sporns, Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* **52**, 1059–1069 (2010).
4. P. Hagmann *et al.*, Mapping the structural core of human cerebral cortex. *PLoS Biol.* **6**, e159 (2008).
5. O. Seung Wook *et al.*, A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
6. N. T. Markov *et al.*, A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* **24**, 17–36 (2014).
7. J. G. White, J. Eileen Southgate, N. Thomson, S. Brenner, The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **314**, 1–340 (1986).
8. D. J. Watts, S. H. Strogatz, Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998).
9. D. S. Bassett, E. D. Bullmore, Small-world brain networks. *Neuroscientist* **12**, 512–523 (2006).
10. O. Sporns, J. D. Zwi, The small world of the cerebral cortex. *Neuroinformatics* **2**, 145–162 (2004).
11. M. D. Humphries, K. Gurney, T. J. Prescott, The brainstem reticular formation is a small-world, not scale-free, network. *Proc. R. Soc. B Biol. Sci.* **273**, 503–511 (2006).
12. M. P. Van Den Heuvel, O. Sporns, Rich-club organization of the human connectome. *J. Neurosci.* **31**, 15775–15786 (2011).
13. G. Zamora-López, C. Zhou, J. Kurths, Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. *Front. Neuroinform.* **4**, 1 (2010).
14. D. S. Bassett *et al.*, Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Comput. Biol.* **6**, e1000748 (2010).
15. J. Stiso, D. S. Bassett, Spatial embedding imposes constraints on neuronal network architectures. *Trends Cogn. Sci.* **22**, 1127–1142 (2018).
16. Z. J. Chen, Y. He, P. Rosa-Neto, J. Germann, A. C. Evans, Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cereb. Cortex* **18**, 2374–2381 (2008).
17. O. Sporns, R. F. Betzel, Modular brain networks. *Annu. Rev. Psychol.* **67**, 613–640 (2016).
18. C. Zhou, L. Žemanová, G. Zamora, C. C. Hilgetag, J. Kurths, Hierarchical organization unveiled by functional connectivity in complex brain networks. *Phys. Rev. Lett.* **97**, 238103 (2006).
19. T. A. Jarrell *et al.*, The connectome of a decision-making neural network. *Science* **337**, 437–444 (2012).
20. S. J. Cook *et al.*, Whole-animal connectomes of both *Caenorhabditis elegans* sexes. *Nature* **571**, 63–71 (2019).
21. M. Winding *et al.*, The connectome of an insect brain. *Science* **379**, eadd9330 (2023).
22. A. Lin *et al.*, Network statistics of the whole-brain connectome of *Drosophila*. *bioRxiv* [Preprint] (2024). <https://doi.org/10.1101/2023.07.29.551086>. Accessed 20 November 2023.
23. S. Dorkenwald *et al.*, Neuronal wiring diagram of an adult brain. *bioRxiv* [Preprint] (2023). <https://doi.org/10.1101/2023.06.27.546656> (Accessed 20 November 2023).
24. P. Schlegel *et al.*, Whole-brain annotation and multi-connectome cell typing quantifies circuit stereotypy in *Drosophila*. *bioRxiv* [Preprint] (2023). <https://doi.org/10.1101/2023.06.27.546055> (Accessed 20 November 2023).
25. F. Li *et al.*, The connectome of the adult *Drosophila* mushroom body provides insights into function. *eLife* **9**, e62576 (2020).
26. L. K. Scheffer *et al.*, A connectome and analysis of the adult *Drosophila* central brain. *eLife* **9**, e57443 (2020).
27. K. L. Briggman, M. Helmstaedter, W. Denk, Wiring specificity in the direction-selectivity circuit of the retina. *Nature* **471**, 183–188 (2011).
28. M. Helmstaedter *et al.*, Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature* **500**, 168–174 (2013).
29. B. Karrer, M. E. J. Newman, Stochastic blockmodels and community structure in networks. *Phys. Rev. E* **83**, 016107 (2011).
30. T. P. Peixoto, Hierarchical block structures and high-resolution model selection in large networks. *Phys. Rev. X* **4**, 011047 (2014).
31. M. E. J. Newman, M. Girvan, Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).
32. M. Rosvall, C. T. Bergstrom, Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1118–1123 (2008).
33. R. F. Betzel, M. A. Bertolero, D. S. Bassett, Non-assortative community structure in resting and task-evoked functional brain networks. *bioRxiv* [Preprint] 2018. <https://doi.org/10.1101/355016> (Accessed 20 November 2023).
34. R. F. Betzel, J. D. Medaglia, D. S. Bassett, Diversity of meso-scale architecture in human and non-human connectomes. *Nat. Commun.* **9**, 1–14 (2018).
35. J. Faskowitz, X. Yan, X.-N. Zuo, O. Sporns, Weighted stochastic block models of the human connectome across the life span. *Sci. Rep.* **8**, 1–16 (2018).
36. L. Zhang, T. P. Peixoto, Statistical inference of assortative community structures. *Phys. Rev. Res.* **2**, 043271 (2020).
37. T. P. Peixoto, Descriptive vs. inferential community detection: Pitfalls, myths and half-truths. *arXiv* [Preprint] (2021). <http://arxiv.org/abs/2112.00183> (Accessed 20 November 2023).
38. T. P. Peixoto, Revealing consensus and dissensus between network partitions. *Phys. Rev. X* **11**, 021003 (2021).
39. M. Ercsey-Ravasz *et al.*, A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron* **80**, 184–197 (2013).
40. R. F. Betzel, D. S. Bassett, Specificity and robustness of long-distance connections in weighted, interareal connectomes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4880–E4889 (2018).
41. M. Rivera-Alba *et al.*, Wiring economy and volume exclusion determine neuronal placement in the *Drosophila* brain. *Curr. Biol.* **21**, 2000–2005 (2011).
42. R. D. Markello *et al.*, Neuromaps: Structural and functional interpretation of brain maps. *Nat. Methods* **19**, 1472–1479 (2022).
43. J. Li *et al.*, Network level enrichment provides a framework for biological interpretation of machine learning results. *bioRxiv* [Preprint] (2023). <https://doi.org/10.1101/2023.10.14.562358> (Accessed 20 November 2023).
44. J. B. Burt, M. Helmer, M. Shinn, A. Anticevic, J. D. Murray, Generative modeling of brain maps with spatial autocorrelation. *Neuroimage* **220**, 117038 (2020).
45. E. K. Towson, P. E. Vértes, S. E. Ahnert, W. R. Schafer, E. T. Bullmore, The rich club of the *C. elegans* neuronal connectome. *J. Neurosci.* **33**, 6380–6387 (2013).
46. C.-T. Shih *et al.*, Connectomics-based analysis of information flow in the *Drosophila* brain. *Curr. Biol.* **25**, 1249–1258 (2015).
47. L. Harriger, M. P. Van Den Heuvel, O. Sporns, Rich club organization of macaque cerebral cortex and its role in network communication. *PLoS ONE* **7**, e46497 (2012).
48. M. P. Van den Heuvel, O. Sporns, An anatomical substrate for integration among functional networks in human cortex. *J. Neurosci.* **33**, 14489–14500 (2013).
49. D. M. Pavlovic, P. E. Vértes, E. T. Bullmore, W. R. Schafer, T. E. Nichols, Stochastic blockmodeling of the modules and core of the *Caenorhabditis elegans* connectome. *PLoS One* **9**, e97584 (2014).
50. S. Maslov, K. Sneppen, Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
51. C. Eschbach *et al.*, Recurrent architecture for adaptive regulation of learning in the insect brain. *Nat. Neurosci.* **23**, 544–555 (2020).
52. Y. Aso *et al.*, The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* **3**, e04577 (2014).
53. C. Seguin, O. Sporns, A. Zalesky, Brain network communication: Concepts, models and applications. *Nat. Rev. Neurosci.* **24**, 557–574 (2023).
54. A. Avena-Koenigsberger, B. Misic, O. Sporns, Communication dynamics in complex brain networks. *Nat. Rev. Neurosci.* **19**, 17 (2018).
55. J. Goñi *et al.*, Resting-brain functional connectivity predicted by analytic measures of network communication. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 833–838 (2014).

$N = 2,952$  neurons and  $M = 110,677$  edges. See *SI Appendix* for details related to this dataset.

**Community Detection.** Here, we fit a hierarchical SBM to the larval *Drosophila* connectome following ref. 30 (<https://graph-tool.skewed.de/>). SBMs are generative models of a network in that, given the community assignments of all nodes,  $\sigma = \{\sigma_i\}$ , the model generates the observed network with probability:  $P(W|\theta, \sigma)$ , where  $\theta$  refers to any additional model parameters that specify the link between community labels and the network. The hierarchical variant recursively fits the SBM to network data using an efficient agglomerative algorithm to minimize the posterior probability that the model generated the observed network (114). See *SI Appendix* for details related to community detection and subsequent analyses.

**Data, Materials, and Software Availability.** *Drosophila* connectome data used in the present study is available here: <https://github.com/brain-networks/larval-drosophila-connectome> (115). Code for estimating the hierarchical stochastic blockmodels is available here: [https://figshare.com/articles/dataset/graph\\_tool/1164194](https://figshare.com/articles/dataset/graph_tool/1164194) (116). Code for estimating community motifs is available here: [https://github.com/brain-networks/wsbn\\_sampler](https://github.com/brain-networks/wsbn_sampler) (117). Code for performing spatially constrained permutations of brain maps: <https://github.com/murraylab/brainsmash> (118). All other data are included in the manuscript and/or *SI Appendix*.

56. R. F. Betzel *et al.*, Structural, geometric and genetic factors predict interregional brain connectivity patterns probed by electrocorticography. *Nat. Biomed. Eng.* **3**, 902–916 (2019).
57. C. Seguin, M. P. Van Den Heuvel, A. Zalesky, Navigation of brain networks. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6297–6302 (2018).
58. B. Mišić *et al.*, Cooperative and competitive spreading dynamics on the human connectome. *Neuron* **86**, 1518–1529 (2015).
59. R. F. Betzel, J. Faskowitz, B. Mišić, O. Sporns, C. Seguin, Multi-policy models of interregional communication in the human connectome. *bioRxiv [Preprint]* (2022). <https://doi.org/10.1101/2022.05.08.490752> (Accessed 20 November 2023).
60. F. Z. Esfahlani, J. Faskowitz, J. Slack, B. Mišić, R. F. Betzel, Local structure-function relationships in human brain networks across the lifespan. *Nat. Commun.* **13**, 1–16 (2022).
61. M. P. Van Den Heuvel, R. S. Kahn, J. Gohi, O. Sporns, High-cost, high-capacity backbone for global brain communication. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11372–11377 (2012).
62. M. Boguna, D. Krioukov, K. C. Claffy, Navigability of complex networks. *Nat. Phys.* **5**, 74–80 (2009).
63. S. Wasserman, J. Galaskiewicz, *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences* (Sage, 1994).
64. P. W. Holland, K. B. Laskey, S. Leinhardt, Stochastic blockmodels: First steps. *Soc. Netw.* **5**, 109–137 (1983).
65. Y. J. Wang, G. Y. Wong, Stochastic blockmodels for directed graphs. *J. Am. Stat. Assoc.* **82**, 8–19 (1987).
66. C. J. Anderson, S. Wasserman, K. Faust, Building stochastic blockmodels. *Soc. Netw.* **14**, 137–161 (1992).
67. H. C. White, S. A. Boorman, R. L. Breiger, Social structure from multiple networks. I. Blockmodels of roles and positions. *Am. J. Sociol.* **81**, 730–780 (1976).
68. G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society. *Nature* **435**, 814–818 (2005).
69. S. Fortunato, Community detection in graphs. *Phys. Rep.* **486**, 75–174 (2010).
70. D. Hric, R. K. Darst, S. Fortunato, Community detection in networks: Structural communities versus ground truth. *Phys. Rev. E* **90**, 062805 (2014).
71. W. Ji *et al.*, Modularity in the organization of mouse primary visual cortex. *Neuron* **87**, 632–643 (2015).
72. S. Takemura *et al.*, A visual motion detection circuit suggested by *Drosophila* connectomics. *Nature* **500**, 175–181 (2013).
73. D. Kiyooka *et al.*, Single-cell resolution functional networks during sleep are segregated into spatially intermixed modules. *bioRxiv [Preprint]* (2023). <https://doi.org/10.1101/2023.09.14.557838> (Accessed 20 November 2023).
74. R. F. Betzel, K. C. Wood, C. Angeloni, M. N. Geffen, D. S. Bassett, Stability of spontaneous, correlated activity in mouse auditory cortex. *PLoS Comput. Biol.* **15**, e1007360 (2019).
75. M. G. Puxeddu, J. Faskowitz, O. Sporns, L. Astolfi, R. F. Betzel, Multi-modal and multi-subject modular organization of human brain networks. *Neuroimage* **264**, 119673 (2022).
76. B. T. Thomas Yeo *et al.*, The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
77. S. M. Smith *et al.*, Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13040–13045 (2009).
78. N. A. Crossley *et al.*, Cognitive relevance of the community structure of the human brain functional coactivation network. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11583–11588 (2013).
79. J. Yang, J. Leskovec, Structure and overlaps of ground-truth communities in networks. *ACM Trans. Intell. Syst. Technol.* **5**, 1–35 (2014).
80. B. D. Pedigo *et al.*, Generative network modeling reveals quantitative definitions of bilateral symmetry exhibited by a whole insect brain connectome. *eLife* **12**, e83739 (2023).
81. S. Kerstjens, G. Michel, R. J. Douglas, Constructive connectomics: How neuronal axons get from here to there using gene-expression maps derived from their family trees. *PLoS Comput. Biol.* **18**, e1010382 (2022).
82. M. E. J. Newman, A. Clauset, Structure and inference in annotated networks. *Nat. Commun.* **7**, 11863 (2016).
83. A. C. Murphy *et al.*, Explicitly linking regional activation and function connectivity: community structure of weighted networks with continuous annotation. *arXiv [Preprint]* (2016). <http://arxiv.org/abs/1611.07962> (Accessed 20 November 2023).
84. J. Clune, J.-B. Mouret, H. Lipson, The evolutionary origins of modularity. *Proc. R. Soc. B Biol. Sci.* **280**, 20122863 (2013).
85. N. Kashtan, U. Alon, Spontaneous evolution of modularity and network motifs. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13773–13778 (2005).
86. N. P. Suh, *The Principles of Design* (Oxford University Press, 1990).
87. A. Nematzadeh, E. Ferrara, A. Flammini, Y.-Y. Ahn, Optimal network modularity for information diffusion. *Phys. Rev. Lett.* **113**, 088701 (2014).
88. R. K. Pan, S. Sinha, Modularity produces small-world networks with dynamical time-scale separation. *Europhys. Lett.* **85**, 68006 (2009).
89. A. Zalesky, A. Fornito, E. Bullmore, On the use of correlation as a measure of network connectivity. *Neuroimage* **60**, 2096–2106 (2012).
90. A. Decelle, F. Krzakala, C. Moore, L. Zdeborová, Inference and phase transitions in the detection of modules in sparse networks. *Phys. Rev. Lett.* **107**, 065701 (2011).
91. J. Yang, J. Leskovec, Overlapping communities explain core-periphery organization of networks. *Proc. IEEE* **102**, 1892–1902 (2014).
92. A.-L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
93. K. Uzel, S. Kato, M. Zimmer, A set of hub neurons and non-local connectivity features support global brain dynamics in *C. elegans*. *Curr. Biol.* **32**, 3443–3459 (2022).
94. M. Rubinov, R. J. F. Ypma, C. Watson, E. T. Bullmore, Wiring cost and topological participation of the mouse brain connectome. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10032–10037 (2015).
95. L. Coletta *et al.*, Network structure of the mouse brain connectome with voxel resolution. *Sci. Adv.* **6**, eabb7187 (2020).
96. L. Li *et al.*, Mapping putative hubs in human, chimpanzee and rhesus macaque connectomes via diffusion tractography. *Neuroimage* **80**, 462–474 (2013).
97. C. J. Honey *et al.*, Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2035–2040 (2009).
98. R. F. Galán, On how network architecture determines the dominant patterns of spontaneous neural activity. *PLoS One* **3**, e2148 (2008).
99. A.-L. Barabási *et al.*, *Network Science* (Cambridge University Press, 2016).
100. S. Horváth *et al.*, Spatial embedding and wiring cost constrain the functional layout of the cortical network of rodents and primates. *PLoS Biol.* **14**, e1002512 (2016).
101. R. F. Betzel *et al.*, The modular organization of human anatomical brain networks: Accounting for the cost of wiring. *Netw. Neurosci.* **1**, 42–68 (2017).
102. H. F. Song, H. Kennedy, X.-J. Wang, Spatial embedding of structural similarity in the cerebral cortex. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16580–16585 (2014).
103. R. F. Betzel *et al.*, Generative models of the human connectome. *Neuroimage* **124**, 1054–1064 (2016).
104. M. Schröter, O. Paulsen, E. T. Bullmore, Micro-connectomics: Probing the organization of neuronal networks at the cellular scale. *Nat. Rev. Neurosci.* **18**, 131–146 (2017).
105. D. Kleinfeld *et al.*, Large-scale automated histology in the pursuit of connectomes. *J. Neurosci.* **31**, 16125–16138 (2011).
106. M. P. Van den Heuvel, E. T. Bullmore, O. Sporns, Comparative connectomics. *Trends Cogn. Sci.* **20**, 345–361 (2016).
107. D. L. Barabási *et al.*, Neuroscience needs network science. *J. Neurosci.* **43**, 5989–5995 (2023).
108. R. F. Betzel, D. S. Bassett, Multi-scale brain networks. *Neuroimage* **160**, 73–83 (2017).
109. F. Bretz, W. Maurer, W. Brannath, M. Posch, A graphical approach to sequentially rejective multiple test procedures. *Stat. Med.* **28**, 586–604 (2009).
110. R. F. Barber, A. Ramdas, The p-filter: Multilayer false discovery rate control for grouped hypotheses. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **79**, 1247–1268 (2017).
111. G. Jekely, S. Jasek, M. Guhmann, L. A. Bezerra-Calderon, E. Williams, R. Shahidi, Whole-body connectome of a segmented annelid larva. *bioRxiv [Preprint]* (2024). <https://doi.org/10.1101/2024.03.17.585258> (Accessed 20 November 2023).
112. F. Randi, A. K. Sharma, S. Divali, A. M. Leifer, Neural signal propagation atlas of *Caenorhabditis elegans*. *Nature* **623**, 406–414 (2023).
113. L. Peel, D. B. Larremore, A. Clauset, The ground truth about metadata and community detection in networks. *Sci. Adv.* **3**, e1602548 (2017).
114. T. P. Peixoto, Efficient monte carlo and greedy heuristic for the inference of stochastic block models. *Phys. Rev. E* **89**, 012804 (2014).
115. R. F. Betzel, Larval-drosophila-connectome. *github*. <https://github.com/brain-networks/larval-drosophila-connectome>. Deposited 5 October 2023.
116. T. P. Peixoto, The graph-tool python library. *figshare*. [https://figshare.com/articles/dataset/graph\\_tool/1164194](https://figshare.com/articles/dataset/graph_tool/1164194). Deposited 27 May 2017.
117. R. F. Betzel, J. D. Medaglia, D. S. Bassett, *wsbm\_sampler*. *github*. [https://github.com/brain-networks/wsbm\\_sampler](https://github.com/brain-networks/wsbm_sampler). Deposited 17 July 2019.
118. J. Burt, J. D. Murray, R. Markello, BrainSMASH. *github*. <https://github.com/murraylab/brainsmash>. Deposited 20 January 2020.