Neuronal “parts list” and wiring diagram for a visual system

Arie Matsliah[1]*, Szi-chieh Yu[1]*, Krzysztof Kruk[3], Doug Bland[1], Austin Burke[1], Jay Gager[1], James Hebditch[1], Ben Silverman[1], Kyle Willie[1], Ryan Willie[1], Marissa Sorek[1], Amy R. Sterling[1], Emil Kind[4], Dustin Garner[5], Gizem Sancer[4], Mathias Wernet[4], Sung Soo Kim[5], Mala Murthy[1], H. Sebastian Seung[1,2], and the FlyWire Consortium[+]

[1]Neuroscience Institute, Princeton University, USA
[2]Computer Science Department, Princeton University, USA
[3]ul. Grunwaldzka 22/91, 25-736, Kielce, Poland
[4]Institut für Biologie - Neurobiologie, Freie Universität Berlin, Germany
[5]Molecular, Cellular, and Developmental Biology, Univ. California Santa Barbara, USA

*These authors contributed equally.
[+]A list of contributing Consortium authors appears at the end of the paper

A catalog of neuronal cell types has often been called a “parts list” of the brain, and regarded as a prerequisite for understanding brain function. In the optic lobe of *Drosophila*, rules of connectivity between cell types have already proven essential for understanding fly vision. Here we analyze the fly connectome to complete the list of cell types intrinsic to the optic lobe, as well as the rules governing their connectivity. We more than double the list of known types. Most new cell types contain between 10 and 100 cells, and integrate information over medium distances in the visual field. Some existing type families (transmedullary, lobula intrinsic, and lobula plate intrinsic) at least double in number of types, with implications for perception of color, motion, and form. We introduce a new family, serpentine medulla intrinsic, which has more types than any other, and three new families of types that span multiple neuropils. We demonstrate self-consistency of our cell types through automatic assignment of cells by distance in high-dimensional feature space, and provide further validation by selection of small subsets of discriminative features. Our work showcases the advantages of connectomic cell typing: complete and unbiased sampling, a rich array of features based on connectivity, and reduction of the connectome to a drastically simpler wiring diagram of cell types, with immediate relevance for brain function and development.

Introduction

Some of the greatest scientific discoveries of the 20th century concern the neural basis of sensory perception. Visual neuroscience, in particular, garnered Nobels for Hubel and Wiesel. Their discovery of simple and complex cells in the visual cortex not only entered neuroscience textbooks. The hypothetical neuronal wiring diagrams in their 1962 paper (Hubel and Wiesel 1962) inspired convolutional nets (Fukushima 1980; Y. LeCun et al. 1989), which eventually ignited the deep learning revolution in artificial intelligence (Yann LeCun, Bengio, and Hinton 2015). It may come as a surprise that no one

has ever obtained direct evidence supporting such wiring diagrams, however influential they may be. The best existing evidence requires the indirect inference of connectivity from cross-correlations in neural activity (Reid and Alonso 1995; Alonso and Martinez 1998), or the assumption that adjacency implies connectivity (Chapman, Zahs, and Stryker 1991).

Mapping the connections between individual neurons is still highly challenging in mammalian brains. Progress is being made by coupling calcium imaging *in vivo* with synaptic physiology in brain slices (Cossell et al. 2015) and calcium imaging of dendritic spines *in vivo* (Jia et al. 2010; Wilson et al. 2016). The reconstruction of a column of visual cortex from electron microscopic images is also becoming feasible (MICrONS Consortium et al. 2021; Schneider-Mizell et al. 2023). These are tiny slivers of visual systems; scaling up these approaches to tackle the full complexity of mammalian vision is still aspirational.

To imagine the future of mammalian visual neuroscience, it is helpful to extrapolate from a brain of more modest size, that of the fly, for which a complete neuronal wiring diagram is now just a few keystrokes away. Especially over the past 15 years, visual neural circuits have been functionally investigated in *Drosophila* (Currier, Pang, and Clandinin 2023) with great progress in understanding the perception of motion (Borst, Haag, and Mauss 2020; Shinomiya et al. 2022), color (Schnaitmann et al. 2018), and form (Wu et al. 2016), as well as the role of vision in complex behaviors like courtship (Ribeiro et al. 2018). The recent release of a neuronal wiring diagram of a *Drosophila* brain (Zheng et al. 2018; Dorkenwald, Matsliah, et al. 2023; Schlegel et al. 2023) poses an unprecedented opportunity. The first wiring diagram for a whole brain contains as a corollary the first wiring diagram for an entire visual system, as well as all the wiring connecting the visual system with the rest of the brain.

Most visual neurons, and indeed most neurons, are situated in the optic lobes of the *Drosophila* brain (Fig. S1a). In the reconstructed brain (Dorkenwald, Matsliah, et al. 2023), about 37,000 neurons are intrinsic to the right optic lobe, meaning that their synapses are fully contained in this region. A wiring diagram for the *Drosophila* visual system must include this many neurons at least. While minuscule compared to a mammalian visual system, this is still too complex to comprehend or even visualize. It is essential to reduce the complexity by describing the connectivity between types of cells. For example, the roughly 800 ommatidia in the compound eye send photoreceptor axons to roughly 800 L1 cells in the lamina, which in turn connect with roughly 800 Mi1 cells. That is a lot of cells and connections, but they can all be described by the simple rules that photoreceptors connect to L1, and L1 connects to Mi1. Some of these connectivity rules have long been known (Meinertzhagen and O'Neil 1991; Gao et al. 2008), and more have been discovered over the past decade (S.-Y. Takemura et al. 2013, 2015; S. Takemura et al. 2017; Shinomiya et al. 2019, 2022), but this knowledge is fragmentary and incomplete.

In the present work, we exhaustively enumerate all cell types intrinsic to the optic lobe, and find all rules of connection between them. We have effectively collapsed 37,000 intrinsic neurons onto just 224 types, a more than 150× reduction. The wiring diagram is reduced from a 37,000×37,000 matrix to a 224×224 matrix, an even huger compression. The most numerous types contain more than 800 cells, slightly more than the estimated number of ommatidia (Fig. 1a). The least numerous types contain only a single cell. The number of cells per type is multimodally distributed (Fig. 1b), with two sharp modes at 1 and roughly 700. Most of the distribution lies between these modes (Fig. 1b, S1b).

The most numerous (>700 cells) types have long been known (Fischbach and Dittrich 1989). It is the less common types where our knowledge has been incomplete, and arguably they are where much of

the magic of vision happens. As with the photoreceptors, neural activity in the most numerous cell types like L1 and Mi1 mostly encodes information about the image at or near single points in visual space. But perception requires the integration of information from points that can be quite distant from each other, and this is done by the larger neurons that belong to the less common types.

While our work focuses on intrinsic neurons of the optic lobe, we also provide rules of connectivity between intrinsic types and known types of boundary neurons, defined as those that connect the optic lobe with regions in the central brain. Such central brain regions are generally multimodal and/or sensorimotor, mixing information coming from the eyes and other sense organs. Therefore we regard the optic lobe as the fly's visual system.

As well as a study of a visual system, our work is also a contribution to cell typing, which has become a major endeavor of contemporary neuroscience (Zeng and Sanes 2017; BRAIN Initiative Cell Census Network (BICCN) 2021). *Drosophila* serves as an interesting middle ground between small nervous systems and big mammalian brains. In *C. elegans*, the number of cell types is somewhat less than half the number of neurons; most cell types consist of just a pair of mirror symmetric neurons (White et al. 1986). Over 30,000 neurons are intrinsic to the central brain of *Drosophila*, 100× more than *C. elegans*. Many cell types in this brain region consist of just one or a few neurons per hemisphere (Scheffer et al. 2020; Schlegel et al. 2023), meaning that the number of cell types is in the thousands (Fig. S1c). In the optic lobes, the number of cell types is far less than the number of neurons. A large ratio of neurons to cell types is reminiscent of mammalian retina or cortex, where the ratio is even larger. The implications of our work for neuronal cell types in mammals will be described in the Discussion.

The present work is based on version 630 of the FlyWire connectome, which incorporates proofreading up to April 2023. Our work will continue to be updated as new versions of the connectome are released (see Discussion). The names of cell types will be adjusted based on more focused studies by the community (Garnet, Kind, et al., in preparation) and coordinated with existing parallel efforts. In the meantime, we are providing what we call discriminative names for the neuropil-intrinsic neurons, which make up most of our novel cell types. A discriminative name provides enough connectivity information to discriminate a cell type from others. For example, Li⇐TmY4⇒LT79 is the Lobula intrinsic (Li) neuron that receives input from TmY4 and sends output to LT79. These features (with thresholding) are sufficient to distinguish Li⇐TmY4⇒LT79 from all other Li types. We do not expect that they will replace traditional names. However, they could serve as an additional naming system that aids establishing correspondences between the results of parallel cell typing efforts. As the consensus evolves, we will also add names in a more traditional style.

All information can now be browsed, searched, and downloaded at the FlyWire Codex (codex.flywire.ai). This addition to the existing array of FlyWire resources (Dorkenwald, Matsliah, et al. 2023; Schlegel et al. 2023) will accelerate the pace of discovery for many researchers.

Single-cell neuroanatomy

Neurons intrinsic to the optic lobe are those with synapses entirely contained inside the optic lobe, and are the main topic of this study. There are additionally optic lobe neurons that extend outside the optic lobe. We refer to these non-intrinsic neurons as "boundary" neurons, as they straddle the optic lobe and the rest of the brain. Boundary neurons fall into several classes. Visual projection neurons (VPNs) have dendrites in the optic lobe and extend axons to the central brain. Visual centrifugal neurons (VCNs) do

the opposite. Heterolateral neurons are intrinsic to the pair of optic lobes; they extend from one optic lobe to the other while making few or no synapses in the central brain. The lamina and medulla are targeted by photoreceptor axons originating in the roughly 800 ommatidia of the compound eye. Photoreceptors can be regarded as boundary neurons, as they connect the optic lobe with the visual world. Alternatively, photoreceptors can be regarded as their own class of cell, distinct from intrinsic or boundary neurons.

Our work is based on the brain of a single *Drosophila* adult female, which was reconstructed by the FlyWire Consortium (Dorkenwald, Matsliah, et al. 2023). There are roughly 37,500 intrinsic neurons in the right optic lobe, 4,000 VPNs, 250 VCNs, and 200 heterolateral neurons. 80% of the synapses of intrinsic neurons are with other intrinsic neurons, and 20% are with boundary neurons. These statistics suggest that the optic lobe communicates more with itself than with the rest of the brain.

To refine these coarse class distinctions, we rely on a number of classical attributes of single cells. (Fischbach and Dittrich 1989) say a neuron is *columnar* if its axon is oriented parallel to the main axis of the visual columns. They allow the arbor of a columnar neuron to be wider than a single column; what matters is the orientation of the axon, not the aspect ratio of the arbor. A neuron is *tangential* if its axon is oriented perpendicular to the main axis of the visual columns. Strictly speaking, the columnar-tangential distinction only applies to neurons with axons. A neuron without an axon is called an *amacrine* cell.

The optic lobe (Fig. S1d) contains four main neuropils (lamina, medulla, lobula, and lobula plate) and a smaller fifth neuropil (accessory medulla). Cell types are either intrinsic to a single neuropil, or span multiple neuropils. Cell types are grouped into families that share the same neuropils. (Fischbach and Dittrich 1989) further distinguished between distal and proximal medulla (Fig. S1d), effectively regarding them as two neuropils separated by medulla layer 7 (M7). The number of types and number of cells in each family are detailed in Fig. 1c. Some families are wholly or almost wholly new (Fig. 1c,d).

To distinguish between neurons in the same neuropil family, each neuropil has been divided into multiple layers (Fischbach and Dittrich 1989) (Fig. S1d), and a cell type is defined as a set of neurons that stratify across the layers in the same way. However, this approach runs into difficulty when distinguishing between cell types with similar stratification, or when grouping morphological variants into a single type. As we will see below, cell types have distinctive connectivity patterns that can readily be used to surmount such difficulties.

Connectomic approach to cell types

All 224 intrinsic types are available for 3D interactive viewing at the FlyWire Codex (codex.flywire.ai). Table S1 is a master list of all types and their properties. Data S1 contains one “card” for each type, which includes its discriminative logical predicate (see below), basic statistics, diagram in the style of (Fischbach and Dittrich 1989) showing stratification and other single-cell anatomy, and 3D renderings of all the cells in the type. The types look qualitatively correct at first glance. Cells of the same type generally look similar, and cover the entire visual field. Some types have “bald spots” in coverage, but preliminary indications are that most if not all of these areas will become covered in a later data release. Cell counts for the types (Fig. 1a) should be regarded as lower bounds; some counts will increase in later data releases. Quantitative validation of the types will be given below.

Roughly speaking, cell typing went through three stages (Methods). First, neurons were annotated by many people eyeballing single-cell anatomy, sometimes assisted by computational tools that analyzed connectivity. Second, types were merged and split based on connectivity analysis; single-cell anatomy had become insufficient by this time. Third, assignment errors were automatically corrected by assigning each cell to the nearest type. The second and third stages relied on a feature vector for each neuron based on its connectivity with cell types, similar to the computational tool CBLAST (Scheffer et al. 2020; Schlegel et al. 2023).

The final cell types are validated in several ways. We show that our clustering is self-consistent, in the sense that almost all cells end up in the original cluster if we attempt to reassign each cell's feature vector to the nearest cluster. For more interpretable evaluations, we construct compact connectivity-based discriminators that can predict cell type membership. We show that membership can be accurately predicted by a logical conjunction of on average five synaptic partner types. For each intra-neuropil type, we also provide a 2D projection of the connection matrix that can be used to discriminate that type from others in the same neuropil.

Feature vectors from cell-to-type and type-to-cell connectivity

For each cell, we define an output feature vector by the number of output synapses onto neurons of cell type t, which runs from 1 to T. This is a T-dimensional vector, where T is the number of cell types. The output feature vector is a row of the cell-to-type connectivity matrix defined in the Methods. For each cell, we similarly defined an input feature vector by the number of input synapses received from neurons of cell type t. This is a column of the type-to-cell connectivity matrix defined in the Methods. Alternatively, feature vectors can be based on connection number, where a connection is defined as two or more synapses from one neuron to another (Methods), and this gives similar results.

The input and output feature vectors were concatenated to form a $2T$-dimensional feature vector. If we include only intrinsic types (Fig. 2a), then T is 224 and the dimensionality $2T$ is 448. If we include both intrinsic and boundary types, then T is about 500 and the dimensionality $2T$ is about 1000 (not shown). We experimented with both dimensionalities, and as will be explained below, the results do not depend on the boundary types. The lower dimensionality feature vector with only intrinsic types is sufficient for defining types.

In the second and third stages of analysis, a cell type was operationally defined as a set of neurons that share similar feature vectors. We used Jaccard similarity (Methods) to quantify similarity between pairs of feature vectors (Fig. S2a, b). This attains a maximum value of one when the two feature vectors are equal, and is lower bounded by zero. The Jaccard distance is defined as one minus the Jaccard similarity. We used various ways of computing the "center" of a set of feature vectors (Methods). The center of the feature vectors in a type can be interpreted as the typical connectivity of a neuron in that type.

Clustering of high-dimensional feature vectors

For the third and final stage of automatic correction, we estimated the "center" of each type using the element-wise trimmed mean. Then for every cell we computed the nearest type center by Jaccard distance. For 98% of the cells, the nearest type center coincided with the assigned type. We sampled some disagreements and reviewed them manually. In the majority of cases, the algorithm was correct,

and the human annotators had made errors, usually of inattention. The remaining cases were mostly attributable to proofreading errors. There were also cases in which type centers had been contaminated by human-misassigned cells (see Morphological Variation), which in turn led to more misassignment by the algorithm. After addressing these issues, we applied the automatic corrections to all but 0.1% of cells, which were rejected using distance thresholds. Based on this procedure, we estimate that our cell type assignments are between 98% and 99.9% accurate. The above results are essentially the same, whether or not boundary types are included in the feature vector (data not shown).

For another measure of the quality of our cell typing, we computed the "radius" of each type, defined as the average distance from its cells to its center. Here we computed the center by approximately minimizing the sum of Jaccard distances from each cell in the type to the center (Methods). A large type radius can be a sign that the type contains dissimilar cells, and should be split. For our final types, the radii vary, but almost all lie below 0.6 (Fig. S2c, d). Dm9 is the only exception, and Lat and Lai have the next largest radii just below 0.6. Dm9 and Lai both resemble chandeliers with many "candlesticks" that are extremely challenging to proofread. These types are expected to become less noisy in the next release. Lat are serotonergic neurons, and further investigation is needed to determine whether their paucity of synapses is biological reality or an artifact of inaccurate synapse detection. The type radii are essentially the same, whether or not boundary types are included in the feature vector (data not shown).

For another quality indicator, we computed a single linkage hierarchical clustering of the type centers. The resulting dendrogram (Fig. S3) reveals pairs that are close to each other by Jaccard distance. These type pairs contain similar cells, and splitting these cells could potentially be an error. The closest pair is Dm3p and Dm3q. The next closest is Pm⇐Tm1⇒Tm2 and Pm⇒Mi1⇒Tm1, the result of splitting Pm2 into two types. In each of these type pairs, both types have similar stratification, but later on we will see the evidence that they are distinct.

Discrimination with logical predicates

Because the feature vector is rather high dimensional, it would be helpful to have simpler insights into what makes a type. One approach is to find a set of simple logical predicates based on connectivity that predict type membership with high accuracy. For a given cell, we define the attribute "is connected to input type *t*" as meaning that the cell receives at least one connection from some cell of type *t*. Similarly, the attribute "is connected to output type *t*" means that the cell makes at least one connection onto some cell of type *t*. We constrain our search to logical predicates that are conjunctions of at most five input attributes and at most five output attributes (Fig. 2b). The average number of attributes is five.

For example, the logical predicate "is connected to input type Dm13 and output type Pm⇐Tm3⇒Tm2 and output type T5b" predicts Tm2 cells with 99% precision and 97% recall. For all but one of the identified types we found a logical predicate with 5 or fewer input / output attributes that predicts type membership with average f-score 0.93, weighted by the number of cells in type (see Figure S4 and Table S1). Some of the attributes in a predicate are the top most connected partner types, but this is not necessarily the case. The attributes are distinctive partners, which are not always the most connected partners. The predicate for each type is shown on its card in Data S1. For each family, the predicates for all types can be shown together in a single graph containing all relevant attributes (Data S2).

We experimented with searching for predicates after randomly shuffling a small fraction of types (namely, swapping types for 5% of randomly picked pairs of neurons). We found that precision and recall of the best predicates dropped dramatically, suggesting that we are not overfitting. This was expected because the predicates are short.

We also measured the drop in the quality of predicates if excluding boundary types (where the predicates are allowed to contain intrinsic types only). As is the case with the clustering metrics, the impact on predicates is marginal (weighted mean F-score drops from 0.93 to 0.92; Methods).

Discrimination with two-dimensional projections

Another approach to interpretability is to look at low-dimensional projections of the $2T$-dimensional feature vector. For each cell type, we select a small subset of dimensions that suffice to accurately discriminate that type from other types (Fig. 2c). Here we normalize the feature vector so that its elements represent the "fraction of input synapses received from type t" or "fraction of output synapses sent to type t." In these normalized quantities, the denominator is the total number of all input or output synapses, not just the synapses with other neurons intrinsic to the optic lobe.

For example, we can visualize all cells in the Pm family in the two-dimensional space of C3 input fraction and TmY3 output fraction (Fig. 2c). In this space, Pm⇐C3⇒TmY3 cells are well-separated from other Pm cells, and can be discriminated with 100% accuracy by "C3 input fraction greater than 0.01 and TmY3 output fraction greater than 0.01."

More generally, a cell type discriminator is based on thresholding a set of input and output fractions, and taking the conjunction of the result. The search for a discriminator finds a set of dimensions, along with threshold values for the dimensions. To simplify the search, we require that the cell type be discriminated only from other types in the same neuropil family, rather than from all other types. Under these conditions, it almost always suffices to use just two dimensions of the normalized feature vector.

We will apply this approach to intra-neuropil type families, which contain many types that can be difficult to distinguish. The discriminating 2D projections are provided in Data S3. Many though not all discriminations are highly accurate. We have restricted the input and output fractions in the discriminator to cross-neuropil types. This aids biological interpretability, as the functions of some cross-neuropil types have already been studied. We have also included boundary types as discriminative features for the same reason. For example, for some LPi types, we have chosen well-studied LPTCs to be discriminative features. For most of the types with less accurate discriminators in Data S3, accuracy can be improved by allowing intra-neuropil types to also be discriminating partners (data not shown). We have avoided this, however, as it causes composite names (explained below) to become recursive.

Composite names based on distinctive partners

Historically, names for optic lobe cell types have been chosen by combining an alphabetic prefix (such as Tm or Dm) with a numeric suffix that records the order of discovery. For the central complex, composite names have been created by concatenating input and output neuropils (Wolff, Iyer, and Rubin 2015; Hulse et al. 2021). Here, we devise a new composite naming system that incorporates connectivity information into the name itself. For example, we have used the two-dimensional projection

of Fig. 2d to name the type as Pm⇐C3⇒TmY3. The arrows are used to specify whether the distinctive partners are inputs or outputs (Fig. 2d). The threshold values of the input and output fractions are suppressed in the name. To be more explicit, Pm⇐C3⇒TmY3 refers to the Pm cell type that can be accurately discriminated from other Pm types using threshold values for the C3 input fraction and TmY3 output fraction.

Composite names are provided for the intra-neuropil types, which make up the bulk of the new types introduced by this paper. For the cross-neuropil types, we stick to the traditional simple names. As mentioned above, the 2D discriminators are restricted to cross-neuropil types, so that composite names are compositions of simple names, avoiding recursion of composite names. Composite names should be helpful for establishing correspondences with independent efforts to define cell types in the optic lobe.

Type-to-type connectivity

We define a type-to-type connection matrix in which the *st* element is the number of synapses from cell type *s* to cell type *t* (Methods). The type-to-type matrix is drastically smaller than the cell-to-cell matrix. Nevertheless, it is still quite large, presenting challenges to visualization and understanding. Figure 3 attempts to show the entire matrix. The area of each dot encodes the number of synapses in some type-to-type connection. Dot area saturates above 3600 synapses, in order to make weaker connections visible. Figure S4 shows another view of the matrix in which dot area saturates above 10,000 synapses.

A cautionary note is that weaker connections in the matrix could be artifactual, due to false positives of automated synapse detection. There are several ways of guessing whether a connection is artifactual, short of manually inspecting the original EM images. One might distrust weak connections between cells, i.e., those with less than some threshold number of synapses. The type-to-type connection matrix can be recomputed after discarding synapses from weak connections. The threshold depends on the scientific question to be addressed. Codex discards connections with less than five synapses. The predicates of the present work apply a threshold of two synapses rather than five. Another heuristic is to look for extreme asymmetry in the matrix. If the number of synapses from A to B is much larger than from B to A, the latter connection might be spurious. The reason is that the strong connection from A to B means the contact area between A and B is large, which means more opportunity for false positive synapses from B to A.

We can also normalize the type-to-type connection matrix to be the fraction of synapses from cell type *s* to cell type *t*. Depending on the normalization, this could be the fraction of input to type *t* or fraction of output from type *s* (Methods). Input and output fractions are shown in Figs. S5 to S14. Each column of the heatmap contains the input or output fractions of the corresponding cell type listed at the bottom of the heatmap. The rows include all partners that contribute at least 0.02 input or output fraction to any of the reference types at the bottom of the heatmap. The heatmaps include partner boundary types as well as intrinsic types. The colormap has been chosen to saturate at 0.2, in order to make small values visible, so large values (≥0.2) are indistinguishable. The viewer should be alert to artifacts that can arise due to normalization. If a cell type makes very few synapses overall, that can result in large input and output fractions, giving the mistaken impression of connections that are strong when they are actually weak or nonexistent.

Columnar types

Many columnar types have long been known. The lamina monopolar cells (L1-L5) project from lamina to medulla. T1, Lawf1, Lawf2, C2, and C3 project from the medulla to the lamina. T2, T2a, and T3 project from proximal medulla to the lobula. T4 (proximal medulla to lobula plate) and T5 (proximal medulla to both lobula and lobula plate) each come in four types activated by four preferred directions of motion. The wiring diagram of the most numerous of these columnar types is shown in Fig. 4a. The securely known types above amount to 21 in all, not including the photoreceptors. The latter project from eye to lamina (R1-6) and medulla (R7, R8).

While the types in the preceding list are securely known, it is not hard to find new information about their connectivity (Fig. S5). For example, Lawf1 receives input from Mi10 and Dm20, and targets L3 and L2, while Lawf2 receives input from the octopaminergic VCN aMe14b and targets L5 (Fig. S5a). T2, T2a, and T3 partners include many LC, LPLC, and LT types (Fig. S5a).

Four families of columnar types connect the distal medulla, lobula, and lobula plate: Tm, TmY, Y, and Tlp (Fischbach and Dittrich 1989). We have updated all four families by adding and/or eliminating types. Most of the new types fall in the Tm family.

Tm Transmedullary

Tm cells project from the distal medulla to the lobula. Tm1 through Tm26 and Tm28 were defined by (Fischbach and Dittrich 1989), and Tm27/Tm27Y was reported by (Hasegawa et al. 2011). We were able to identify Tm1, 2, 3, 4, 7, 9, 16, 20, 21, 25, and 27. We have merged Tm6 and Tm21 into a single type Tm21. We prefer the latter name because the cells more closely match the Tm21 stratification as drawn by (Fischbach and Dittrich 1989). We have moved Tm23 and Tm24 to the Lobula intrinsic (Li) family. They were originally classified as Tm because their cell bodies are in the distal rind of the medulla, and they send a neurite along the columnar axis of the medulla to reach the lobula (Fischbach and Dittrich 1989). However, they do not form synapses in the medulla, so we regard them as Li neurons in spite of their soma locations.

Perhaps the most significant change in the Tm family concerns Tm5. Tm5 is known to be important for color vision, because it's stratification in the distal medulla near M7 and in the M3 overlaps with the axon terminals of the chromatic photoreceptors R7 and R8 (Gao et al. 2008). We identified six Tm5 types (Fig. 5a), half of which are predicted to be cholinergic and the other half glutamatergic. Each type has a characteristic single-cell anatomy as shown in Fig. 5a. However, there is considerable variability within a type, so assigning cells to types is difficult or impossible based on single-cell anatomy alone. Luckily, each type turns out to have a clear connectivity signature, which we used for typing individual cells (Fig. 5b, S7, Data S1-3).

We established correspondences with three canonical Tm5 types originally defined by single-cell anatomy and Ort expression (Gao et al. 2008; Karuppudurai et al. 2014). Tm5a is cholinergic, the majority of the cells extend one dendrite into the distal medulla, and often has a “hook” at the end of its lobula axon. Tm5b is cholinergic, and many cells extend several dendrites into the distal medulla. Tm5c is glutamatergic and extends its dendrites up to the surface of the distal medulla. Three of our Tm5 types fit these descriptions (Fig. 5a), and receive direct input from R7 or R8.

Our Tm5a and Tm5b receive R7 and Dm8 input (Fig. 4b, 5b, S7), as expected from previous reports (Karuppudurai et al. 2014; Menon et al. 2019; Kind et al. 2021). Tm5a receives more input from Dm8b (Fig. 4b, 5b, S7), which is therefore identified as yellow Dm8 (yDm8) (Menon et al. 2019; Courgeon and Desplan 2019) . Tm5b receives more input from Dm8a, which is therefore identified as pale Dm8 (pDm8). Tm5c receives R8 input (Karuppudurai et al. 2014; Kind et al. 2021), and also strong L3 input (Fig. 4b, 5b, S7). Tm20 receives R8 input (S.-Y. Takemura et al. 2013, 2015; Kind et al. 2021), and also strong L3 input. R7 and R8 do not yet appear in the Figures; quantification is postponed until the next data release, which will incorporate the proofread photoreceptors.

Almost 400 Dm8 cells have been proofread and identified in v630. This is substantially less than the numbers reported by molecular studies (Menon et al. 2019; Courgeon and Desplan 2019). The pale to yellow ratio is roughly even, though the molecular studies reported that Dm8 is mostly yellow. More proofread Dm8 cells will be added in the next release, but it is unlikely that the numbers will change dramatically.

The new Tm5 types (Tm5d, e, and f) are not directly connected with photoreceptors, but they do receive indirect R7 and R8 input. Tm5d and Tm5e are predicted glutamatergic and Tm5f is predicted cholinergic. Tm5d receives indirect R7 input from Tm5b, yDm8, and Tm5a. Tm5e receives indirect R8 input from Tm5c, as well as L5 input (Fig. S7). Tm5f receives indirect R8 input from Tm20.

We divided Tm8 into Tm8a and Tm8b, which are both predicted cholinergic. Tm8 subtypes have no obvious morphological difference. They share some partners, as would be expected from their similar stratifications. For example, Mi9 is the strongest input to both types, and Li⇒LC12⇒Tm8a is another shared input. But overall Tm8a and Tm8b differ markedly in connectivity. Li⇒LC12⇒Tm8a is the strongest output of Tm8a, but weak or nonexistent as an output of Tm8b. Tm5c is a strong input to Tm8b, but weak or nonexistent as an output of Tm8a. The direct connection from Tm5c suggests that Tm8b might play a role in color vision.

Tm1, Tm2, Tm3, Tm4, Tm9, Tm20, and Tm21 are numerous (>600 cells, Fig. 1a), and receive strong input from L cells (Fig. S7). Tm3 drives T4 cells, while Tm1, Tm2, Tm4, and Tm9 drive T5 cells (Fig. S7). Tm27 is driven weakly by T4 cells, as well as T5a.

Most existing Tm types are predicted cholinergic (excepting the three glutamatergic Tm5 types mentioned above), and have some stratification in the distal medulla (Fischbach and Dittrich 1989). We discovered 6 wholly new Tm types, which we named Tm31 through Tm36. Almost all of them contain less than 100 cells, and are not predicted to be cholinergic. (Tm36 is the only exception.) They pass through the distal medulla without making synapses, and stratify in the serpentine and proximal medulla. Many of them are coupled with the chromatic types Tm5 and Tm20 (Fig. S7). For example, the strongest inputs to Tm31 are Tm5d and Tm5c, and a strong output is Tm5a. The strongest input to Tm33 is Tm20, and the strongest outputs are Tm5f and Tm20. Overall, roughly half of the 24 types in the Tm family are novel.

TmY

Transmedullary Y (TmY) cells project from the distal medulla to the lobula and lobula plate. The “Y” refers to the divergence of branches to the lobula and lobula plate. TmY1 to TmY13 were defined by (Fischbach and Dittrich 1989), TmY5a by (Morante and Desplan 2008; Fischbach and Dittrich 1989),

TmY14 by (S.-Y. Takemura et al. 2013), TmY15 by (Takemura et al. 2017), TmY16, TmY18, and TmY20 by (Shinomiya et al. 2019). We identified TmY3, TmY4, TmY5a, TmY10, TmY11, TmY14, TmY15, TmY16, and TmY20. We have divided TmY9 into two types. Their strongest inputs are Tm1 and two types of line amacrine cells Dm3 (Fig. S8). A new type TmY31 receives input from Tm20 and other chromatic Tm types, and projects to numerous LC and LPLC types (Fig. S8).

Y and Tlp Translobula plate

Y cells project from the proximal medulla to the lobula and lobula plate. They are similar to TmY cells, but the latter traverse both the distal and proximal medulla (Fischbach and Dittrich 1989). Y1 and Y3 to Y6 were defined by (Fischbach and Dittrich 1989). Y11 and Y12 were defined by (Shinomiya et al. 2022). We have identified Y1, Y3, Y4, Y11, and Y12 in our reconstruction, and have not found any new Y types. Y1 receives strong T4a and T5a input (Fig. S6b), suggesting it is direction selective, similar to previous reports for Y11 and Y12 (Shinomiya et al. 2022). Y3 receives input from several columnar types (Mi1, Mi4, Mi9, and Tm1) that are highly numerous, and also from TmY5a (Fig. S6b). The strongest input of Y4 is T2a (Fig. S6b).

A translobula plate neuron is defined as projecting from the lobula plate to the lobula. Tlp1 through Tlp5 were defined by (Fischbach and Dittrich 1989). We have identified Tlp1, Tlp4, and Tlp5, and an additional type Tlp6. All four Tlp types receive strong input from T4 and T5 with a single preferred direction, suggesting they are direction selective (Fig. S6b). Based on T4 and T5 input, it should be possible to establish correspondences with Tlp11 through Tlp14 defined by (Shinomiya et al. 2022). Three out of four of the Tlp types also have T4 and T5 outputs. All Y and Tlp types drive TmY14 (Fig. S6b).

Mi medulla intrinsic columnar

The Mi family contains medulla intrinsic types that have an axon and are columnar. Mi is a bit unusual as most columnar types span multiple neuropils, and are not intrinsic to a single neuropil. However, Mi types extend across both distal and proximal medulla, which can be regarded as two different neuropils following (Fischbach and Dittrich 1989). With this convention, the Mi family conforms to the mold of an cross-neuropil columnar type.

We identified five (Mi1, 2, 4, 9, 10) of the dozen Mi types originally defined by (Fischbach and Dittrich 1989). We also identified three (Mi13, 14, 15) types uncovered by EM reconstruction (S.-Y. Takemura et al. 2013). This is a total of eight medulla intrinsic columnar (Mi) types. Mi1, Mi4, and Mi9 are the most numerous (>750). They receive strong L input, and drive the T4 elementary motion detectors (Fig. S5b, S6b). Mi9 also receives T4 input, except not from T4d (Fig. S5b). The less numerous Mi2, Mi13 and Mi14 target a variety of columnar types that are intrinsic to the optic lobe. The main output of Mi10 is Lawf1, which projects back to the lamina.

It was previously noted that Mi15 receives R8 photoreceptor input (Kind et al. 2021). Mi15 also has strong input from Tm5c (Fig. 5b, S5, S7), so it may play a role in color vision. Mi15 is predicted cholinergic, and it is also the only dopaminergic cell intrinsic to the optic lobe (Davis et al. 2020). The main output of Mi15 is medulla columnar (MC) neurons that project to the central brain, specifically the MeTu3 cells (Garner, Kind, et al., in preparation).

Revisiting a definitional fine point, we note that Mi1, Mi4, and Mi9 have dendrite in the distal medulla and axon in the proximal medulla, whereas Mi13 is the other way around. Other Mi types are less polarized, and narrow-field amacrine might be more accurate than columnar. Nevertheless we will adhere to the convention that they are columnar. Narrow-field amacrine cells exist in the mammalian retina (Marc et al. 2014).

Cross-neuropil tangential and amacrine

Most types that span multiple neuropils are columnar. There is only one securely known tangential type that spans multiple neuropils inside the optic lobe: Lat has a tangential axon that projects from the medulla to the lamina (Fischbach and Dittrich 1989). Here we introduce two new families of cross-neuropil types that are tangential (MLt1-8 and LMt1-4), and one that is amacrine (LMa1-4). Along with an additional new tangential type (LLPt) and the known CT1 and Am1 types, that is a total of 19 cross-neuropil types that are non-columnar (Fig. 1c). Each of the new types contains between 10 and 100 cells.

The tangential types connect neuropils within one optic lobe and do not leave the optic lobe. Our usage of the term “tangential” focuses on axonal orientation only. It should not be misunderstood to imply a wide-field neuron that projects out of the optic lobe, which is the case for the well-known lobula plate tangential cells (LPTCs) or lobula tangential cells (LTs). The term “tangential” presupposes that we can identify an axonal arbor for the cell. Roughly speaking, the axon is some portion of the neuron where the ratio of presynapses to postsynapses is high, and there is a preponderance of presynaptic boutons. If there is no axonal arbor, then the cell is amacrine. For many types it is obvious whether there is an axon, but for a few types we have made judgment calls.

MLt Medulla-Lobula tangential

(Kind et al. 2021) described ML1, a tangential neuron projecting from medulla to lobula. We will refer to this type as MLt1, and have discovered more types of the same family, MLt2 through MLt8. Mlt1 and Mlt2 dendrites span both distal and proximal medulla, and Mlt3 dendrites are in the distal medulla, so MLt1 through MLt3 receive L input (Fig. S9a). Mlt4 dendrites are in the proximal medulla, and its strongest inputs are Pm⇒Mi1⇒Tm1 and LMa5 (Fig. S9a). Mlt5 through Mlt8 have substantial arbor overlap with the serpentine layer M7, and are therefore connected with many Serpentine medulla (Sm) types to be discussed later on (Fig. S9a). Interaction between MLt types is fairly weak, with the exception of MLt7 to MLt5 (Fig. S9a). MLt7 and MLt8 are restricted to the dorsal and dorsal rim areas.

LMt Lobula-Medulla tangential

We identified four tangential types (LMt1 through LMt4) that project from lobula to medulla. Their axonal arbors are all in the proximal medulla, thinly stratified near layer M7, so they have many Pm targets (Fig. S9b). Only LMt4 exhibits partial coverage.

LLPt Lobula-Lobula Plate tangential

We discovered one tangential type that projected from lobula to lobula plate, and called it LLPt. This is just a single type, rather than a family.

LMa Lobula Medulla amacrine

We discovered four amacrine types that extend over lobula and medulla, and named them LMa1 through LMa4. We regard these types as amacrine, because the ratio of presynapses to postsynapses is approximately the same in both neuropils, and there is no bouton-heavy part of the arbor. These types are coupled with T2, T2a, and T3, and LMa4 and LMa3 synapse onto T4 and T5 (Fig. S9b).

The LMa family could be said to include CT1, a known amacrine cell that also extends over both the lobula and medulla. The new LMa types, however, consist of smaller cells that each cover a fraction of the visual field, whereas CT1 is a wide-field cell.

(Shinomiya et al. 2022) reported Am1, a wide-field amacrine cell that extends over the medulla, lobula, and lobula plate. We found no other amacrine types like Am1 with such an extended reach. Whether Am1 is truly amacrine is somewhat unclear, as the medulla is definitely its output neuropil by presynapse/postsynapse ratio, but there is no obvious axon defined by morphological criteria.

Medulla intrinsic noncolumnar

We have discussed the types that span multiple neuropils. The remaining types are those that are intrinsic to a single neuropil. These make up the majority of types, but a minority of cells. Lai is the only type intrinsic to the lamina. We start with neurons that are intrinsic to the medulla, but not columnar, i.e., excluding the Mi family discussed earlier. These types were divided into Dm and Pm families by (Fischbach and Dittrich 1989), depending on whether they stratified in the distal or proximal medulla. We have more than doubled the number of Pm types, and slightly increased the number of Dm types. More importantly, we introduce the Serpentine medulla (Sm) family of types, which is almost completely new. For types that are intrinsic to one neuropil, we will usually avoid distinguishing between columnar, tangential, and amacrine, even though it is often possible to make such distinctions.

Dm Distal Medulla

Dm1 through Dm8 were defined by (Fischbach and Dittrich 1989), Dm9 and 10 by (S.-Y. Takemura et al. 2013), and Dm11 through Dm20 by (Nern, Pfeiffer, and Rubin 2015). DmDRA1 and DmDRA2 were defined by (Sancer et al. 2019; Hardcastle et al. 2021). We do not observe Dm5 and Dm7, consistent with a previous study (Nern, Pfeiffer, and Rubin 2015). Most types are predicted to secrete glutamate and GABA, but there are also a few cholinergic types (Table S1).

Based on light microscopic anatomy (Nern, Pfeiffer, and Rubin 2015), transcriptomics (Özel et al. 2021; Kurmangaliyev et al. 2020), and an EM reconstruction of seven medulla columns (S.-Y. Takemura et al. 2015; Özel et al. 2021), it was previously believed that there are two Dm3 types with orthogonally oriented dendrites. In fact there are three types (Dm3v, Dm3p, and Dm3q) with dendrites pointing in three different directions (Fig. 5c, d, e).

We can categorize Dm types by their source of input arriving from the lamina (Fig. S10). For almost all Dm types, one of the strongest inputs is L2-5 or R7-8, and this input dominates all other laminar inputs. L1 synapses onto Mi1, but not onto Dm cells.

Dm types with L2 input include Dm15, Dm6, Dm19, Dm16, and Dm17. Dm types with L3 input include Dm12, Dm9, Dm20, Dm4, and Dm3. Dm14 is the only Dm type for which L4 input dominates. Dm types with L5 input include Dm10, Dm1, Dm18, and Dm13.

Dm types with R7 input include Dm8a and Dm8b, which we think correspond to yDm8 and pDm8, respectively. These are the most numerous and stick close to the serpentine layer. They presumably correspond with two variants of Dm8 identified by molecular studies (Menon et al. 2019; Courgeon and Desplan 2019). Dm11 makes synapses onto MC66 visual projection neurons, and extends chandelier-like dendrites towards the distal surface of the medulla. DmDRA1 cells are only at the dorsal rim. DmDRA1 receives R7 input, but sits squarely in M7. This is a borderline case that could be counted as an Sm type, if not for history. Dm types with R8 input include DmDRA2.

Pm Proximal medulla

Pm1, 1a, and 2 (Fischbach and Dittrich 1989) have each been split into two types. Pm3 and 4 were defined by (Nern, Pfeiffer, and Rubin 2015). We have additionally defined five new Pm types, for a total of 13 Pm types. All are predicted to be GABAergic.

Pm⇒Mi9 and Pm⇐C3⇒TmY3 were formerly known as Pm1. Pm⇐Tm1⇒T3 and Pm⇐Mi1⇒Tm25 were formerly known as Pm1a. Pm⇐Tm1⇒Tm2 and Pm⇒Mi1⇒Tm1 were formerly known as Pm2.

All Pm types receive strong input from one or more of the highly numerous columnar types (Mi1, Mi4, Mi9, Tm1 through Tm4). There is little interaction with chromatic Tm types or with T4 and T5, and a great deal of output onto T2, T2a, and T3 (Fig. S11).

Most types consist of 10 to 100 cells. One exception is Pm⇒C2, a pair of full-field cells stratifying in M8. We refer to them as a “jigsaw pair” because they jointly cover the visual field in an irregular fashion, as if they were cut by a jigsaw. Another exception is Pm⇐TmY5a, a pair of full-field cells stratifying in M9, and sending tendrils up into M8.

Sm Serpentine medulla

We define a large new family of medulla intrinsic neurons consisting of 46 types, almost all new. This class is called “Serpentine medulla” (Sm), because its types have significant stratification in the serpentine layer (M7) of the medulla. Most Sm types are thinly stratified in and around M7, including types recently named “medulla tangential intrinsic” (Kind et al. 2021). We avoid using this term indiscriminately because some Sm types are tangential while others are amacrine. Some types spill into the distal or proximal medulla, and a few arborize thickly across many medulla layers.

In the original classification scheme of (Fischbach and Dittrich 1989), the non-columnar medulla intrinsic cells stratified exclusively in the distal or proximal medulla, and could therefore be divided into Dm and Pm classes. The serpentine layer (M7) is the boundary between distal and proximal medulla. Because there are so many borderline cases, it makes sense to create a new class Sm for them. The Sm class has functional significance, because it consists of the neurons that communicate with the medulla tangential (Mt) cells. These boundary types are an important conduit of information in and out of the optic lobe.

Most of these boundary types are VCNs (Fig. S12). Notable VCN inputs include aMe17a and CB0566. The MeTu types are prominent Sm inputs and outputs. Mi15, Mi10, and especially Dm2 are strong input sources. Connections with chromatic types (Tm5, Dm8) are particularly strong or numerous.

Most types are thinly stratified in or near the serpentine layer. Some types stratify on both sides of the serpentine layer, including Sm⇐Mi4⇐PS129 and Sm⇐CB0042⇒Tm21, which stratify in M1 as well as in the distal medulla. Sm⇒TmY3 stratifies over the distal medulla, serpentine layer, and proximal medulla.

The Sm class more than doubles the number of medulla intrinsic types, relative to the old scheme with only Mi, Pm, and Dm.

Li Lobula intrinsic

Two lobula intrinsic types (Li1 and Li2) were defined by (Fischbach and Dittrich 1989) and 12 more (Li11 through 20 and mALC1 and mALC2) by (Scheffer et al. 2020). Of these we have confirmed Li2, Li12, Li16, mALC1, and mALC2. Li11 has been modified, as explained below. We have split Li18 into three types. We identified 21 additional Li types, but have not been able to make conclusive correspondences with previously identified types. As mentioned earlier, we transfer Tm23 and Tm24 (Fischbach and Dittrich 1989) from the Tm family to the Li family, because they make no synapses in the medulla. This amounts to a total of 31 Li types.

Li2 is a set of eight cells, Li12 is a jigsaw pair, and Li16 is a pair of full-field cells. We postpone our discussion of the Li18 split until the end of this section.

Li outputs include many LC and LPLC types (Fig. S13). The lobula columnar (LC) neurons are the most numerous class of lobula VPN, and converge onto optic glomeruli in the central brain (Otsuna and Ito 2006; Wu et al. 2016). Li outputs also include many LT types, the VPN type aMe20, VCN types, columnar types, and the cross-neuropil tangential type MLt3.

In addition to mALC1 and mALC2, there are two new full-field singletons Li⇒Y3 and Li⇒LC13⇒TmY17.

Li⇐T5a (Tm23) receives strong input from T5a (Shinomiya et al. 2019) and Tm3, and sends strong output to T5 cells with all four preferred directions. This is a rare example of a cell with T5 connections but not T4 connections. Li⇐T5a and Li⇐Tm5_f⇒LT87 were called Tm23 and Tm24 by (Fischbach and Dittrich 1989).

The four cells in Li⇐mALC4⇒LT33 are almost full-field, but their shapes are irregular and interesting. For example, one covers dorsal and ventral regions but not the equator.

Li18 has been split into three types. Li⇐TmY17⇒TmY10 covers the whole visual field, while the other two Li18 types have only partial coverage of the visual field. Li⇐Tm8b⇒aMe20 covers a dorsal region except for the dorsal rim. It is tangentially polarized, with the axon more dorsal than the dendrites. Both axon and dendrite point in the posterior direction, perpendicular to the direction of polarization. The dendrites are more thickly stratified than the axon. Li⇐mALC5⇒Tm16 has ventral coverage only. The

axons are in one layer, and extend over a larger area than the dendrites, which hook around into another layer and are mostly near the ventral rim.

We considered merging Li⇐Tm8b⇒aMe20 and Li⇐mALC5⇒Tm16, but their connectivity is quite different. Furthermore, in a hierarchical agglomerative clustering, Li⇐mALC5⇒Tm16 merges with Li⇐TmY17⇒TmY10 rather than Li⇐Tm8b⇒aMe20.

LPi Lobula plate intrinsic

LPi1-2 and 2-1 were defined by (Shinomiya et al. 2022), generalizing previous definitions of LPi3-4 and 4-3 by (Mauss et al. 2015). Other types reported were LPi2b and LPi34-12 (Shinomiya et al. 2022). (We are not counting fragments for which correspondences are not easy to establish.) We have added nine new types, for a total of 15 LPi types.

LPi names were originally based on stratification in layers 1 through 4 of the lobula plate. Because of collisions, stratification is no longer sufficient for naming. For example, there are now two types that could be called LPi1-2. They can be distinguished by their size, but we will instead apply discriminative names based on connectivity. LPi⇐T5a⇒H2 was called LPi1-2 by (Shinomiya et al. 2022), and is a jigsaw pair of full-field cells (Fig. 5f). H2 is a lobula plate tangential cell (LPTC) that prefers back-to-front motion (Wei et al. 2020). LPi⇐T5a⇒LPLC2 is a more numerous type of smaller cell with similar stratification (Fig. 5g). Here the distinctive output is LPLC2, a VPN activated by looming stimuli (Klapoetke et al. 2017). LPi⇐T5a⇒VS is a third type with T5a input. The distinctive output is VS, which includes multiple LPTCs that prefer downward motion (Joesch et al. 2008).

Four LPi types receive input from T4b and T5b. Two of them, LPi⇐T5b⇒T5b and LPi⇐T5b⇒TmY15, are reciprocally connected with T4b and T5b. LPi⇐T5b⇒T5b is a singleton full-field cell, which likely corresponds with LPi2b of (Shinomiya et al. 2022). LPi⇐T5b⇒CH is another singleton with output onto LPTCs that prefer front-to-back motion (CH and HS), as well as columnar cells with the same preference (T4a and T5a). LPi⇐T5b⇒CH. LPi⇐T5b⇒TmY17, the most numerous type in this group, has little or no output onto T4 and T5. It was called LPi2-1 by (Shinomiya et al. 2022).

Four LPi types receive input from T4c and T5c. The less numerous types LPi⇐T5c⇒LPLC4 and LPi⇐T5c⇒PLP127 each cover only part of the visual field. The strongest outputs of the ventral-only LPi⇐T5c⇒PLP127 are CB0203 and PLP127, VPNs with dendrites only in the ventral portion of the lobula plate. LPi⇐T5c⇒LPLC4 lacks coverage in the medial portion of the lobula plate. This type targets a subset of VS cells, the locations of which are related to the partial coverage. Both LPi⇐T5c⇒LPLC4 (Fig. 5h) and LPi⇐T5c⇒PLP127 are polarized in the tangential plane, showing a clear segregation between dendritic and axonal arbor. This is unlike most LPi types, which are amacrine. LPi⇐T5c⇒Y11 targets a number of columnar types, including columnar VPNs. The most numerous type in this group is LPi⇐T5c⇒VS, which was called LPi3-4 by (Shinomiya et al. 2022).

Three LPi types receive input from T4d and T5d. LPi⇐T5d⇒TmY14 is a full-field singleton that is reciprocally connected to T4d and T5d. LPi⇐T5d⇒LLPC2 and LPi⇐T5d⇒LPLC4 synapse onto columnar cells that target the central brain.

A final cell type is LPi⇐T5c⇐T5d, which receives roughly equal T4 and T5 input from the c and d preferred directions. This breaks the rule of input from one preferred direction, obeyed by all other LPi types. It is reciprocally connected to the same T4 and T5 types.

The full-field singletons (LPi⇐T5b⇒T5b, LPi⇐T5b⇒VCH, and LPi⇐T5d⇒TmY14) and the jigsaw pair (LPi⇐T5a⇒H2) are all predicted to be GABAergic. The remaining types contain tens of cells, and are mostly predicted to be glutamatergic.

The connectivity of LPi types with their top partners is shown in Fig. S14. T4 and T5 are the dominant input to all types, and are also outputs for a few types. LPi types connect with a variety of LPTCs and columnar VPNs. LPi input is also important for most LPi types. One exception is LPi1-H2, which receives only a small fraction of its input from LPi1-LPLC2. Overall, there seems to be more connectivity between LPi types (Fig. S14) than between Li types (Fig. S13).

Morphological variation

Tm1a and Tm4a were defined as morphological variants (Fischbach and Dittrich 1989), but we have found that they do not differ in connectivity and are not common, so we have merged them into Tm1 and Tm4. We merged Tm27Y into Tm27 (Hasegawa et al. 2011), and Tm5Y and TmY5 into TmY5a (Fischbach and Dittrich 1989; Morante and Desplan 2008). These morphological distinctions arose because the projection into the lobula plate, the differentiator between Tm and TmY, can vary across cells in a type.

In general, ignoring some morphological variability has been crucial for correct cell typing; connectivity seems to be more fundamental than morphology. One interesting example is TmY14, a type that was originally identified by (S.-Y. Takemura et al. 2013). Later on, (Shinomiya et al. 2019) recognized that TmY14 does not really belong in the TmY family, because it is a VPN sending a projection to the central brain. In another story twist, our optic lobe turns out to contain a subset of TmY14 that lack the central brain projection (Fig. 6a, b). In these atypical TmY14 cells, the axon remains in the medulla rather than projecting to the central brain. Whether typical or atypical, the TmY14 axon has few synapses and minimal impact on connectivity. And the optic lobe connectivity of the TmY14 cells does not seem to depend on whether or not there is a central brain projection. Therefore we have ignored the morphological difference and grouped all cells into a single type. (In cases like this, we double check the proofreading before concluding that this is true morphological variation rather than an experimental artifact.)

Another interesting example is Tlp4 versus Y11. These two types have similar connectivity patterns (Fig. S6b), and are near each other in the hierarchical clustering of types (Fig. S3). A major difference is that Tlp4 cells, by definition, have no connectivity in the medulla. Except a few of them do, and look like they do not belong in Tlp4 (Fig. 6d). For a long time, we classified these errant cells as Y11. But it turns out that their medullar projections make few synapses, and their connectivity in the lobula and lobula plate matches Tlp4. Therefore we have included such “pseudo-Y11” cells in the Tlp4 type.

It is worth mentioning an unusual example in which ignoring morphological variation is correct in one sense, but happens to be misleading. (Schlegel et al. 2023) annotated three Li11 cells in the present optic lobe, following the hemibrain (Scheffer et al. 2020). Two of these cells are Li⇐Tm21⇒LT61 (Fig. 6e). The third cell can be paired with a fourth to form a “pseudo-Li11” type Li⇐Tm21⇐LT11, which has

a small projection into the central brain (Fig. 6f). This projection is the most striking difference between Li11 and pseudo-Li11 morphologies, but has few synapses and therefore little impact on connectivity. It might be tempting to ignore the projection as a developmental “accident” and merge Li11 and pseudo-Li11 into a single type. But it turns out that Li11 and pseudo-Li11 can be distinguished by connectivity in the lobula. As expressed by the discriminative names, strong Tm21 input distinguishes these two types from other Li types, and they can be distinguished from each other using LT61 output fraction or LT11 input fraction. Pseudo-Li11 has now been identified in the hemibrain (Schlegel et al. personal communication), though there it lacks the small projection. So the central brain projection of pseudo-Li11 exhibits variability across individuals, further evidence that it is a developmental accident.

We also saw a few “weirdo” cells, which looked strange and were usually one-of-a-kind. For example, cell 720575940621331438 resembles an Li full-field cell, but also extends a smaller secondary arbor into the lobula plate and medulla. We decided that this was a developmental accident, and did not include it in our list of types. Whether this odd-looking cell is repeated in the left optic lobe remains to be seen.

Spatial coverage

As a general rule, the cells of a cell type collectively cover the tangential plane of the optic lobe. This property is sometimes called “tiling,” although for some types the arbors overlap so much that the analogy to floor tiles is not so apt. Tiling is also a property of many cell types in mammalian retina.

Tiling means that visual circuits perform the same computation at every point in visual space. This is also a strategy commonly used for the early stages of computer vision algorithms, including convolutional networks. Tiling means that during neural development, cells are capable of avoiding other cells of the same type, without avoiding cells of different types (Grueber and Sagasti 2010).

Although tiling is extremely important for cell typing in the optic lobe, we have said little about it so far, because our feature vector (Fig. 2a) includes no explicit information about the spatial coordinates of a cell. Suppose that we group cells with similar feature vectors into a candidate cell type. If those cells happen to tile, it lends more certainty that the type is correct. In other words, tiling is not an input to typing algorithms; it is an outcome that can be used for validation.

Tiling solves the lumper-splitter dilemma. Suppose we attempt to split one type into two candidate types, based on hierarchical clustering. If both candidate types tile, then we accept them as valid. If the cells of both candidate types seem randomly scattered, that means our split is invalid, because it is presumably discriminating between cells based on noise. (Chromatic types might seem like an exception to this rule, but their locations are only apparently random because they depend on the locations of pale and yellow columns.)

The above are easy cases, but there are also edge cases. Suppose that the two candidate types tile the dorsal field and the ventral field, respectively. In this case, we say that the split is invalid. Instead, we prefer to say that there is a single type that exhibits spatial variation in connectivity. Suppose that one candidate type tiles the dorsal field and the other tiles the full field. This is the only case in which we accept a split that produces a type with incomplete tiling.

Conversely, if hierarchical clustering merges two types with nonoverlapping spatial coverage, the result is likely a correct type. If hierarchical clustering merges two types with overlapping coverage, the result is likely not correct.

We have already discussed quite a few types with partial coverage. One well known example is the Dorsal Rim Area (DRA). It is a different region from the rest of the retina in its organization of the inner photoreceptors. Photoreceptors in non-DRA and DRA differ in their axonal target layers and output cell types (von Philipsborn and Labhart 1990; Sancer et al. 2019). Specifically, DRA-R7 connects with DmDRA1 whereas DRA-R8 connects to DmDRA2. These distinctive connectivity patterns result in DmDRA1 and DmDRA2 types exhibiting an arched coverage primarily in the M6 layer of the dorsal medulla. There are many other cases in which partial coverage can be related to connectivity with a boundary type that only exists in a particular region of the visual field. The Sm family contains the most types with partial coverage. This along with very similar stratification profiles makes Sm the most challenging family for cell typing, and connectivity analysis is indispensable. We discuss the Sm types with partial coverage here because their unusual spatial organization is bound to be related to function. Furthermore cell types with partial coverage deserve extra scrutiny to be accepted as valid.

While most singleton types (e.g. Sm⇐TmY17, Sm⇒aMe5⇐aMe17b, Sm⇒TmY16, and Sm⇒MeTu1⇒Tm35) are full-field, Sm⇐L3⇒Tm5_d is a singleton covering the ventral field only. Sm⇐aMe4⇒Mi15 is a singleton with a dorsal dendritic arbor in M7 and full-field axonal arbor in M1.

Of the types with partial coverage that contain multiple cells, Sm⇐CB0566⇐MeTu2, Sm⇐MC65⇒TmY5a, Sm⇐MeTu3⇒MeTu2, Sm⇐MeTu4⇐MeTu1, Sm⇐MeTu4⇒CB0566, Sm⇐Mi15⇒MLt5, Sm⇐Mi9⇒Tm32, and Sm⇒TmY3 are dorsal-only. Ventral-only types include Sm⇐Mi9⇒CB0165, Sm⇐TmY5a⇒Tm7, and Sm⇐Y12⇒TmY10. Notably, Sm⇐Mi9⇒Tm32 resembles similar stratification to Mi3 in upper and lower boundaries of the M7 layer. However, Sm⇐Mi9⇒Tm32 exhibits an additional stratification within the M4 layer.

Sm⇐Tm31⇒Tm25 and Sm⇐MC66⇐Mi10 are types covering the ventral two thirds of the medulla. Sm⇐CB0081⇐PS123 leaves the dorsal and ventral thirds of the visual field uncovered.

Sm⇐IB029⇐MeTu1 is a trio of H-shaped polarized cells with dendrites at the anterior rim, and axons at the posterior rim. Sm⇐CB0566⇐MeTu2 consists of three sickle-shaped cells at the dorsal rim. Sm⇐IB029⇒Tm34 consists of nine polarized cells with sickle-shaped axons arching in the anterodorsal medulla, and dendrites branching sparsely in the anteroventral region.

The existence of many Sm types with noncanonical shapes and incomplete spatial coverage makes sense, given that Sm cells interact closely with the Mt cells arborizing in the serpentine layer. Some Mt types exist in only one portion of the visual field, presumably to support some visual computation that is specific to that location. Therefore, we expect that some Sm cell types might also specialize to specific regions in the visual field.

Discussion

Early advocates of connectomics predicted that deriving cell types from connectivity would be transformative (H. Sebastian Seung 2012). Now that the connectomic era has finally dawned, we can begin to see how well this prediction holds up. The connectomic approach to cell typing has three

powers. First, it does not suffer from the incomplete and biased sampling that can plague other methods. Second, connectivity turns out to provide a rich set of features for distinguishing between cell types. Third, connectomic cell typing not only yields cell types, but even more importantly tells us how they are wired to each other.

Implications for visual function

Optic lobe cell types and their wiring will be useful for understanding every aspect of fly vision, from color to motion to form. Cell types relevant for color perception (R7, R8, three Tm5 types, Tm20, two Dm8 types, and Dm9) have been known for some time. Tm5a and Tm5b exhibit hue selective responses (Christenson et al. 2023). Tm5c lacks hue selectivity (Christenson et al. 2023), consistent with the fact that L3 is such a strong input and Dm8 is not (Fig. 5b). Tm20 exhibits hue selectivity but it is weak (Christenson et al. 2023), consistent with the fact that Tm20 receives strong input from L3.

We have added three more Tm5 types (Fig. 5a), and their connectivity hints at their possible roles in color vision. It seems likely that Tm5d will inherit hue selectivity from Tm5b, its strongest input. Tm5e might inherit lack of hue selectivity from its strongest input Tm5c. Tm5f might inherit weak hue selectivity from one of its strongest inputs Tm20. Going beyond these speculations to obtain firmer predictions would require a recurrent network model like the one proposed to explain hue selectivity of existing Tm5 types (Christenson et al. 2023).

Downstream from Tm5 are numerous VPN types (Fig. 4b, 5b, S7) that lead to the central brain. Notably, Tm5 types target LC10 (Fig. S7). In male flies, LC10a neurons are known to be important for chasing and orienting to the female during courtship (Ribeiro et al. 2018; Hindmarsh Sten et al. 2021). Color preference in *Drosophila* varies over the day (Schnaitmann et al. 2018), and males court most at dawn and dusk. Fly wings are also extremely colorful (Katayama et al. 2014), and males may use this color information to orient behind females. Other strong outputs of Tm5 include Tm5b to LC6, and Tm5f to LC17, LPLC1, and LPLC2 (Fig. S7). These VPNs are all looming detectors (Klapoetke et al. 2022), and implicated in processing motion cues during social behavior (Cowley et al. 2022). In short, our work indicates that LCs important for male courtship behavior receive information about color as well as motion and form. (Note that our wiring diagram is derived from a female brain, but is likely to be the same or similar for a male brain, as there are few sexual dimorphisms within the optic lobe.)

Motion circuits are the best studied part of the fly visual system, yet it is not hard to make new discoveries about them. Y1 and Pm⇐C3⇒TmY3 are inputs to T4 and T5. Tlp output onto T4 and T5 is now known (Fig. S6a), whereas only Tlp input from T4 and T5 was previously known (Shinomiya et al. 2022). It is also evident that each T4 and T5 type synapses onto other cells of the same type (Fig. S6a)(Shinomiya et al. 2022). This intra-type interaction likely contributes to direction selectivity, as explored in a companion paper. Many LPi partners of T4 and T5 are newly defined types (Fig. S6a). LPi types are all likely inhibitory, and some are thought to mediate opponency between different motion directions (Mauss et al. 2015). But the complex web of interactions between an expanded number of LPi types (Fig. S14) suggests that we may have to go beyond the simple concept of opponency. Furthermore, the connectivity of LPi types with types outside motion circuits will be important for understanding how motion vision and form vision interact. As conjectured in a companion paper, the three Dm3 types, TmY4, TmY9q, and TmY9q$^{\perp}$ likely constitute an optic lobe circuit for form vision (H. Sebastian Seung 2023).

Implications for visual development

Mammalian visual cortex was historically fertile ground for studies of neural circuit development, starting with attempts to explain ocular dominance and orientation columns (Katz and Crowley 2002). These aspects of cortical organization can be seen without single cell resolution, and now seem crude compared to neuronal wiring diagrams. Now that we have a detailed wiring diagram for the fly visual system, we can regard it as the end goal of fly visual system development. Single cell transcriptomics is providing detailed information about the molecules in fly visual neurons (Kurmangaliyev et al. 2020; Özel et al. 2021; Konstantinides et al. 2022). Comparison of transcriptomic and connectomic information is already uncovering molecules important for the development of the fly visual system (Yoo et al. 2023), and this trend is bound to increase in momentum. Our discriminative predicates and type names may be useful in the search for molecules that endow neural connectivity with its exquisite specificity.

Complete and unbiased

Most of our new types are not so common (10 to 100 cells), which may be why they were missed by (Fischbach and Dittrich 1989). These pioneers relied on Golgi staining to sample neurons from multiple individuals, a technique that is well-suited for identifying the most numerous types. We can only speculate about why they missed the Sm family entirely. And we do not know why so many types defined by (Fischbach and Dittrich 1989) cannot be identified in our optic lobe. These authors acknowledge that Golgi staining yields a biased sample of cells, and that morphological variability might have caused them to overestimate the number of types.

Contemporary light microscopic anatomy is more powerful because it can leverage the selective labeling of genetic lines. However, combining light microscopy with genetic control does not evade the limitations of incomplete and biased sampling, as we explain now using the examples of Dm3 and Tm5. Light microscopic anatomy (Nern, Pfeiffer, and Rubin 2015) went beyond (Fischbach and Dittrich 1989) by splitting Dm3 into two types with dendrites at orthogonal orientations. Why did (Nern, Pfeiffer, and Rubin 2015) not find the third Dm3v type, which is so obvious in Fig. 5b? Presumably it is because only Dm3p and Dm3q were labeled in the GAL4 line used by the study. There was no way to see Dm3v because it was invisible. The dendrogram of Fig. S3 shows that Dm3p and Dm3q are more similar to each other than to Dm3v in their connectivity patterns. Presumably this is also true for the gene expression patterns, which could be why the GAL4 line only labeled Dm3p and Dm3q. Subsequent transcriptomic studies (Özel et al. 2021; Kurmangaliyev et al. 2020) did not uncover the third Dm3 type. Neither did analysis of a wiring diagram reconstructed from seven medulla columns via electron microscopy (S.-Y. Takemura et al. 2015; Özel et al. 2021), presumably because this volume is too small to contain more than fragments of Dm3 cells. But the three Dm3 types (Fig. 5b) are unmistakably obvious in our complete and unbiased sample of optic lobe neurons.

A breakthrough in color vision started by genetically labeling neurons that express the histamine receptor Ort (Gao et al. 2008). Researchers reasoned Ort would be expressed by cells postsynaptic to the chromatic photoreceptors R7 and R8, which are histaminergic. Then light microscopic anatomy was used to make fine distinctions between three Tm5 types labeled in the transgenic line (Gao et al. 2008). The present connectomic work has revealed six Tm5 types, a finding that was only foreshadowed by previous work on the same EM dataset (Kind et al. 2021). The three new Tm5 types were presumably missed by previous work because they receive little or no direct photoreceptor input (Fig. 4b), and do

not express Ort. Nevertheless, they may belong to the color vision circuit because they are strongly interconnected with types that receive direct R7 and R8 input. As speculated above, Tm5d might be hue selective, because its strongest input is Tm5b, which is known to be hue selective (Christenson et al. 2023).

The Dm3 and Tm5 examples demonstrate that connectomics can find fresh patches in well-trodden ground. More telling is that connectomics can guide us to entirely new landscapes, such as the 45 Sm types in an entirely new type family. The Introduction alluded to a lower bound of 500 types in the optic lobe, if we include both intrinsic and boundary types. This is considerably greater than transcriptomic estimates of neuronal diversity. For example, one study reported 171 neuronal cell types, a figure that includes visual projection neurons as well as intrinsic neurons (Özel et al. 2021).

Distinguishing cell types using connectivity

The first and second stages of our cell typing relied heavily on single-cell anatomy (Methods). The third stage, however, had to rely more heavily on connectivity. Connectivity-based features (Fig. 2a, Fig. S2a, b) enabled us to discriminate between cell types that stratify in very similar neuropil layers. This was essential, for example, to define Sm types. They are easy to confuse because they are so thinly stratified at or near medulla layer 7.

Stratification constrains connectivity, because neurons cannot connect with each other unless they overlap in the same layers. However, stratification does not completely determine connectivity, because neurons in the same layer may or may not connect with each other. Classical neuroanatomy, whether based on Golgi or genetic staining, relied on stratification because it could be seen with a light microscope. Now that we have electron microscopic data, we can rely on connectivity for cell typing, rather than settle for stratification and other properties from single-cell anatomy.

We have additionally verified self-consistency of the final types in the connectivity feature space. Virtually every cell is assigned to the nearest type in feature space, where nearest is defined by Jaccard distance. Such self-consistency might be viewed as trivial because erroneous type assignments were corrected using Jaccard distance in the final stage of cell typing (see Clustering of high dimensional feature vectors and Methods). It is actually nontrivial because human experts have other means of type assignment beyond the feature vector. They can look at morphology, how the cells tile the visual field, and so on. The nontrivial result is that our type assignments are both consistent with expert opinion and self-consistent according to distances in feature space.

As noted previously (Scheffer et al. 2020), using this feature vector to define types is circular because the feature vector itself depends on the assignment of cells to types. In the initial stage of cell typing, these assignments are based on morphology as well as connectivity. Once the feature vector has become rich enough based on these assignments, it becomes possible to switch to clustering based on connectivity alone. We leave for future work the challenge of evolving this into a connectivity-based clustering algorithm from start to finish.

Role of boundary types

Some might not be surprised by our empirical finding that defining intrinsic types can be done using only synapses between intrinsic neurons, because these make up 80% of all synapses of intrinsic

neurons. Others might be skeptical of our empirical finding because boundary types contribute 20% of the synapses of intrinsic neurons, a fraction that may not sound insignificant. To explain our empirical finding, we speculate that synapses between intrinsic and boundary neurons do not aid typing because the longer feature vector has more redundancy in the information theoretic sense.

While boundary types are not necessary for defining intrinsic types, they are important to our work for two reasons. First, the connectivity between types is arguably the main product of our work, and connectivity between intrinsic and boundary types is of great importance for understanding visual function. Second, using boundary types as discriminative features for intrinsic types can aid the biological interpretation of the latter.

Spatial organization

The wiring diagram for cell types (Fig. 3) is simpler than the connectome from which it was derived. Part of the reduction in complexity comes from ignoring space. If we say that cell types A and B are connected, it is understood implicitly that this rule only applies when the tangential separation between the cells is small enough that their arbors overlap. In some cases connectivity might also depend on the absolute locations of the cells, not only their relative location. Cell types that cover only parts of the visual field (Fig. 7) are the clearest example of dependence on absolute location. The dependence of connectivity on space is essential for vision, and we leave its characterization for future work. We can view both cell type labels and spatial coordinates as discrete and continuous latent variables in a model of connectivity (H. Sebastian Seung 2009).

Artificial intelligence

This paper began by recounting the story (H. S. Seung and Yuste 2010) of how wiring diagrams for visual cortex drawn in the 1960s inspired convolutional nets, which eventually sparked the deep learning revolution in artificial intelligence (AI). Convolutional nets have now been applied to reconstruct the fly brain from electron microscopic images (Dorkenwald, Matsliah, et al. 2023), making the current study possible. Coming full circle, the fly optic lobe turns out to be as literal an implementation of a convolutional net as one could ever expect from a biological system. The columns of the optic lobe form a hexagonal lattice, rather than the square lattice used in computer vision, but it is a highly regular lattice nonetheless. And the activities of the neurons in each cell type are analogous to a feature map in a convolutional net (Lappalainen et al. 2023). One difference is that optic lobe circuits are highly recurrent, which is not the norm for the convolutional nets used in computer vision.

Another difference is that the connections of optic lobe neurons do not appear to be learned in the sense of AI, because they seem to have little dependence on visual experience (Scott, Reuter, and Luo 2003). However, mechanisms based on spontaneous activity (rather than visually evoked activity) might play a role in *Drosophila* visual development (Akin and Zipursky 2020), analogous to mammalian visual development.

Implications for mammalian cell types

For cortical cell types, single cell transcriptomics has been hailed as more advanced than the old-fashioned anatomical approach (Yuste et al. 2020). Proponents are now coming around to a more nuanced view, because neurons of the same transcriptomic type can have highly variable morphological and electrical properties (Scala et al. 2020; Gouwens et al. 2020). It is not yet clear why

this is the case. One possibility has long been suggested, which is that transcriptomic differences might exist during development, and vanish in adulthood. This possibility has been confirmed in the fly visual system. For example, T4 and T5 types with a and b preferred directions can be transcriptionally distinguished from those with c and d preferred directions in adult flies (Özel et al. 2021). But all four preferred directions can be transcriptionally distinguished only at the P50 pupal stage.

The connectomic approach is already being applied to cell types in visual cortex (Schneider-Mizell et al. 2023). There is obvious motivation to scale up the approach, and make it as definitive for the cortex as it now is for the fly visual system. Here the limiting factor is throughput. In mammalian brains, transcriptomics has so far had higher throughput, measured in number of cells per unit of time or money. This is because basic research in transcriptomics can leverage technologies developed by the large DNA sequencing industry. Connectomics is also becoming more economical, and this trend should continue.

Future releases

As emphasized in the Introduction, we are releasing the results of our cell typing as quickly as possible in the FlyWire Codex (codex.flywire.ai), to enhance usability of the *Drosophila* connectome. The next update will occur after the next release of the connectome. One major change will be the addition of the remaining photoreceptors, which by now have been completely proofread and typed, and are already accessible via CAVEclient (Dorkenwald, Matsliah, et al. 2023; Dorkenwald, Schneider-Mizell, et al. 2023). Our naming system is provisional, and may be revised based on community feedback, and comparison with other efforts.

Data availability

Most up to date optic lobe intrinsic cell type annotations can be downloaded directly from the Codex download portal (https://codex.flywire.ai/api/download). They are also versioned and available to download from a dedicated data repository on GitHub:
https://github.com/murthylab/flywire-visual-neuron-types
These annotations will be ported to the next FlyWire Connectome releases when published.

The FlyWire Consortium (optic lobe)

Krzysztof Kruk[2], Ben Silverman[1], Jay Gager[1], Kyle Patrick Willie[1], Doug Bland[1], Austin T Burke[1], James Hebditch[1], Ryan Willie[1], Celia David[1], Gizem Sancer[8], Jenna Joroff[4], Dustin Garner[5], annkri (Anne Kristiansen)[2], Thomas Stocks[2], Amalia Braun[6], Szi-chieh Yu[1], Marion Silies[7], AzureJay (Jaime Skelton)[2], TR77[2], Maria Ioannidou[7], Marissa Sorek[1], Matt Collie[4], Gerit Linneweber[8], Sebastian Mauricio Molina Obando[7], Rey Adrian Candilada[1], Alexander Borst[6], Wei-Chung Lee[4], Philipp Schlegel[9,10], Greg Jefferis[9,10], Arie Matsliah[1], Amy R. Sterling[1], Emil Kind[8], Mathias Wernet[8], Sung Soo Kim[5], Mala Murthy[1], H. Sebastian Seung[1,11]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, USA
[2]Eyewire, Boston, USA
[3]Department of Neuroscience, Yale University, New Haven, USA
[4]Harvard Medical School, Boston, USA
[5]University of California, Santa Barbara, USA

[6]Department Circuits-Computation-Models, Max Planck Institute for Biological Intelligence, Planegg, Germany
[7]Johannes-Gutenberg University Mainz, Mainz, Germany
[8]Institut für Biologie - Neurobiologie, Freie Universität Berlin, Germany
[9]Drosophila Connectomics Group, Department of Zoology, University of Cambridge, Cambridge, UK;
[10]Neurobiology Division, MRC Laboratory of Molecular Biology, Cambridge, UK;
[11]Computer Science Dept, Princeton University, Princeton, USA

Acknowledgements

We are indebted to all those who proofread optic lobe intrinsic neurons (Dorkenwald, Matsliah, et al. 2023). We thank A. Nern and M. Reiser for educating and advising FlyWire community members who engaged in annotation of visual neurons. We thank the FAFB tracing community for supportive and open sharing of methods and data, especially the FAFB optic lobe working group. We are grateful to Y. Kurmangaliyev for his thoughtful comments on the manuscript. We thank R. Behnia and M. Silies for feedback about visual motion detection pathways. We thank J. Wiggins, G. McGrath, and D. Barlieb for computer system administration and M. Husseini for project administration. We are grateful to J. Maitin-Shepard for Neuroglancer. MM and HSS acknowledge support from the National Institutes of Health (NIH) BRAIN Initiative RF1 MH117815, RF1 MH129268 and U24 NS126935, from the Princeton Neuroscience Institute, as well as assistance from Google. DG, GS, and SK were supported by the National Eye Institute of the NIH (DP2EY032737), Searle Scholars Program, Sloan Research Fellowship, and Klingenstein-Simons Fellowship in Neuroscience. EK and MW were supported by Deutsche Forschungsgemeinschaft (DFG) grant WE 5761/4-1, SPP 2205, FOR 5289, and AFOSR grant FA9550-19-1-7005.

Author Contributions

DG and GS annotated cells under the supervision of SSK and MW. MS and ARS recruited, trained, and managed citizen scientists with help from EK. KK annotated cells and created computational cell typing tools for use by the community. SY trained and managed DB, AB, JG, JH, BS, KW, RW to annotate the remaining known cell types and discover and annotate new types. AM and HSS created semiautomated cell typing tools. AM and HSS carried out the final automated stage of typing. AM verified types with predicates. HSS verified types with 2D projections. SY, HSS, MM devised type family names and composite type names. MW and MM identified implications for visual function. AB, JG, JH, BS, KW, RW, SY, AM, HSS created figures. KK, MS, ARS, AM, SY, SSK, MM, HSS wrote the text. MM and HSS supervised the project.

Competing interests

H. S. Seung declares financial interests in Zetta AI.

Main Figures

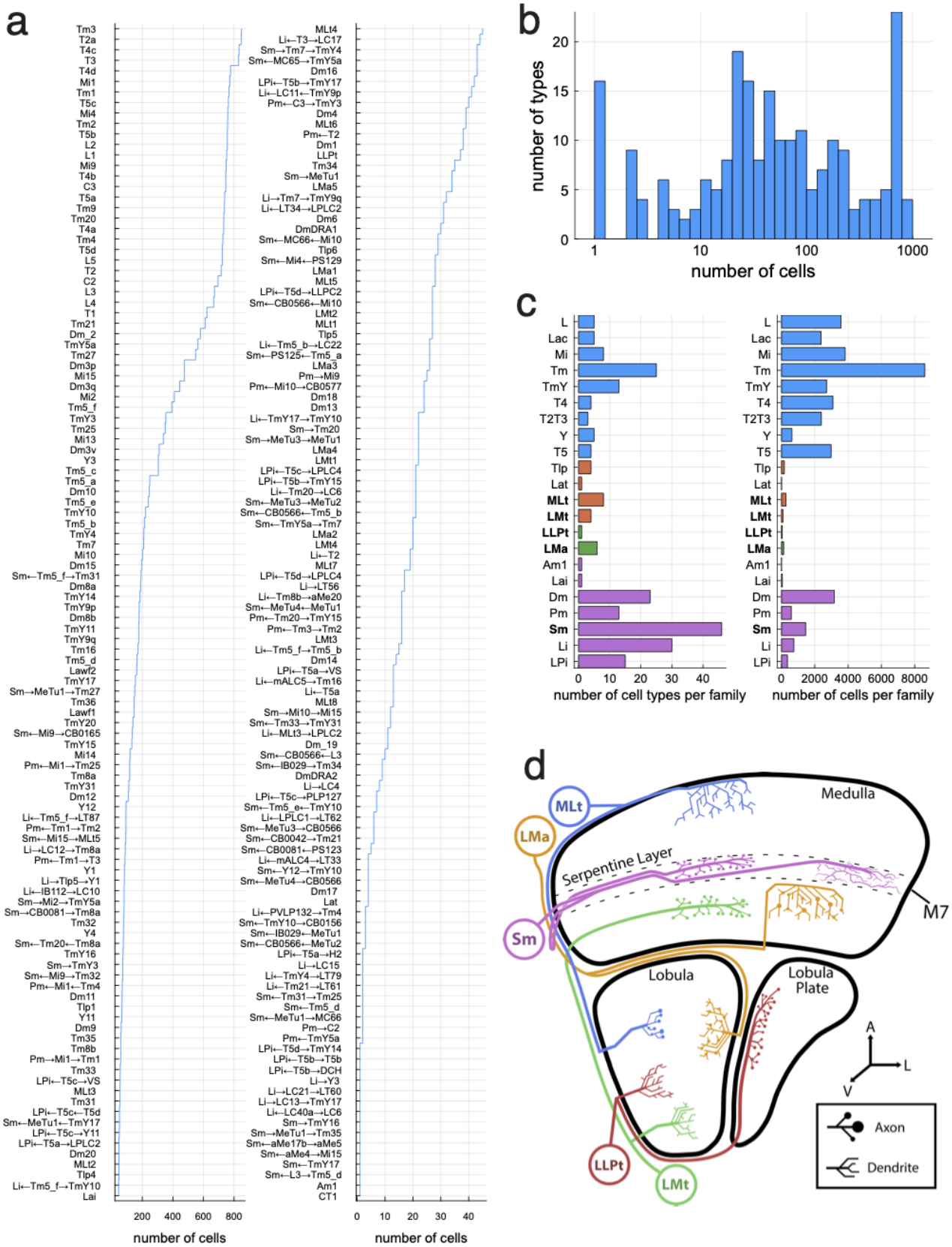

Figure 1. Overview of types and their properties

(a) Number of cells in each type. Y-axis lists cell type names, X-axis denotes their sizes (number of cells). These counts are based on v630, and some will increase in subsequent data releases, due to the addition of more proofread cells.

(b) Distribution of number of types by their size. X-axis denotes type size (log-scale), and Y-axis the number of types with matching size. Around 750 are the columnar cell types with one instance per ommatidia.

(c) Number of types (left) and number of cells (right) in each neuropil-defined family. Bold font indicates families that are entirely new, or almost entirely new.

(d) Illustration of new type families. For tangential families, axon and dendrite are distinguished graphically. We avoid this distinction for Sm, which contains both tangential and amacrine types. (MLt: Medulla Lobula tangential. LMa: Lobula Medulla amacrine. Sm: Serpentine medulla. LLPt: Lobula Lobula Plate tangential. LMt: Lobula Medulla tangential. A: Anterior. L: Lateral. V: Ventral)

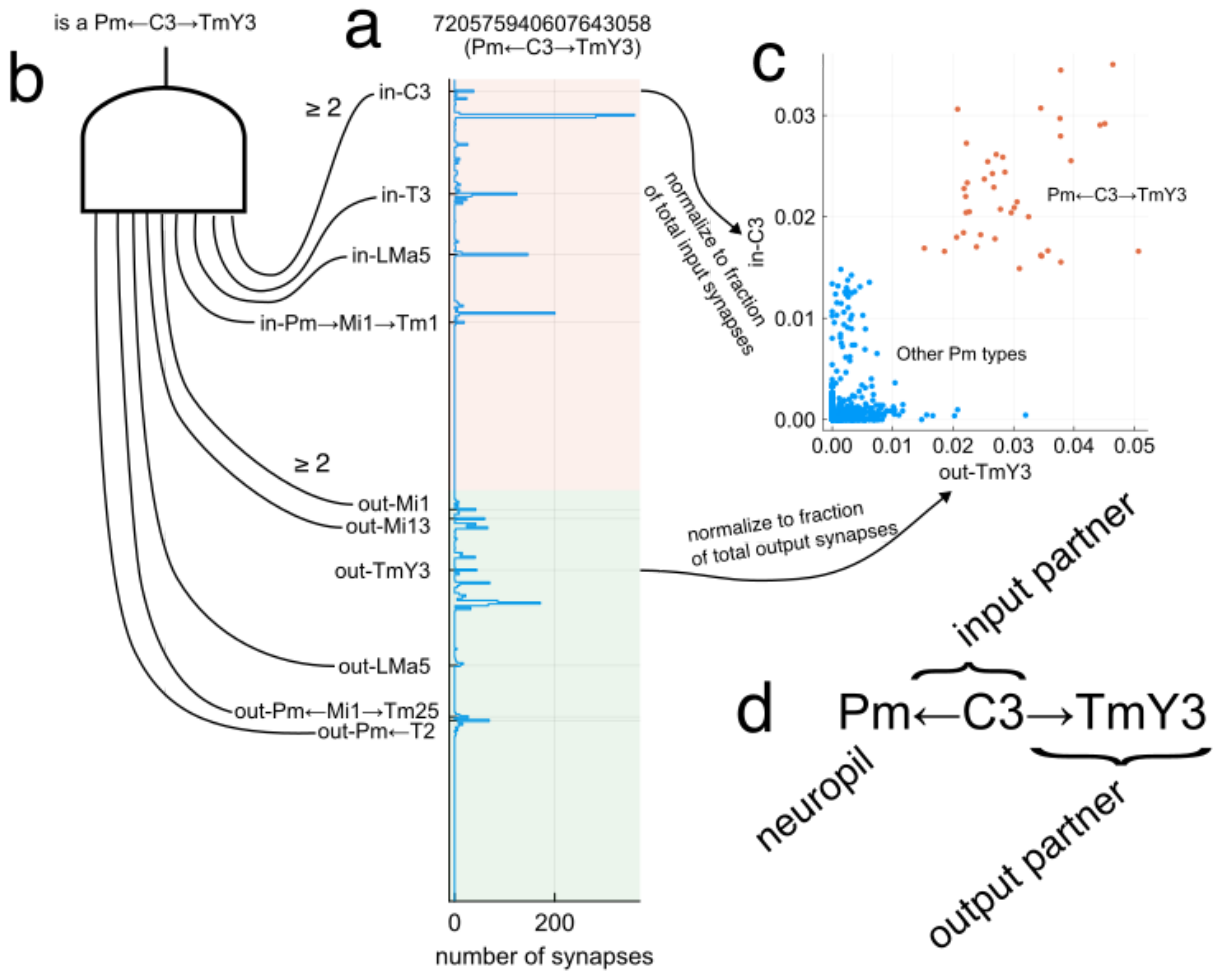

Figure 2. High and low-dimensional representations of cells

(a) High-dimensional feature vector of a single cell. On the Y-axis are the input (red) and output (green) features corresponding to types. On the X-axis is the number of input / output synapses the cell has with those types respectively.

(b) Logical predicate. Predicts if a cell is of type Pm⇐C3⇒TmY3 with a conjunction (logical and) of binary features (corresponding to up to 5 input types and up to 5 output types). Each of the binary features evaluates to True if the corresponding connection has 2 or more synapses.

(c) 2D projection of the discriminating predicates used in the composite name for Pm⇐C3⇒TmY3. Red points indicate Pm⇐C3⇒TmY3 cells, blue points indicate other cells in the same family Pm. On the X and Y axis are the fraction of their inputs / outputs in C3 / TmY3 respectively.

(d) Discriminative type name parts explained.

Figure 3. Wiring diagram of cell types as a matrix.

The number of synapses from one cell type to another is indicated by the area of the corresponding dot. Dot area saturates above 3600 synapses, to make weaker connections visible. For legibility, the type names alternate between left and right edges, and bottom and top edges, and are color coded to match the lines that are guides to the eye.

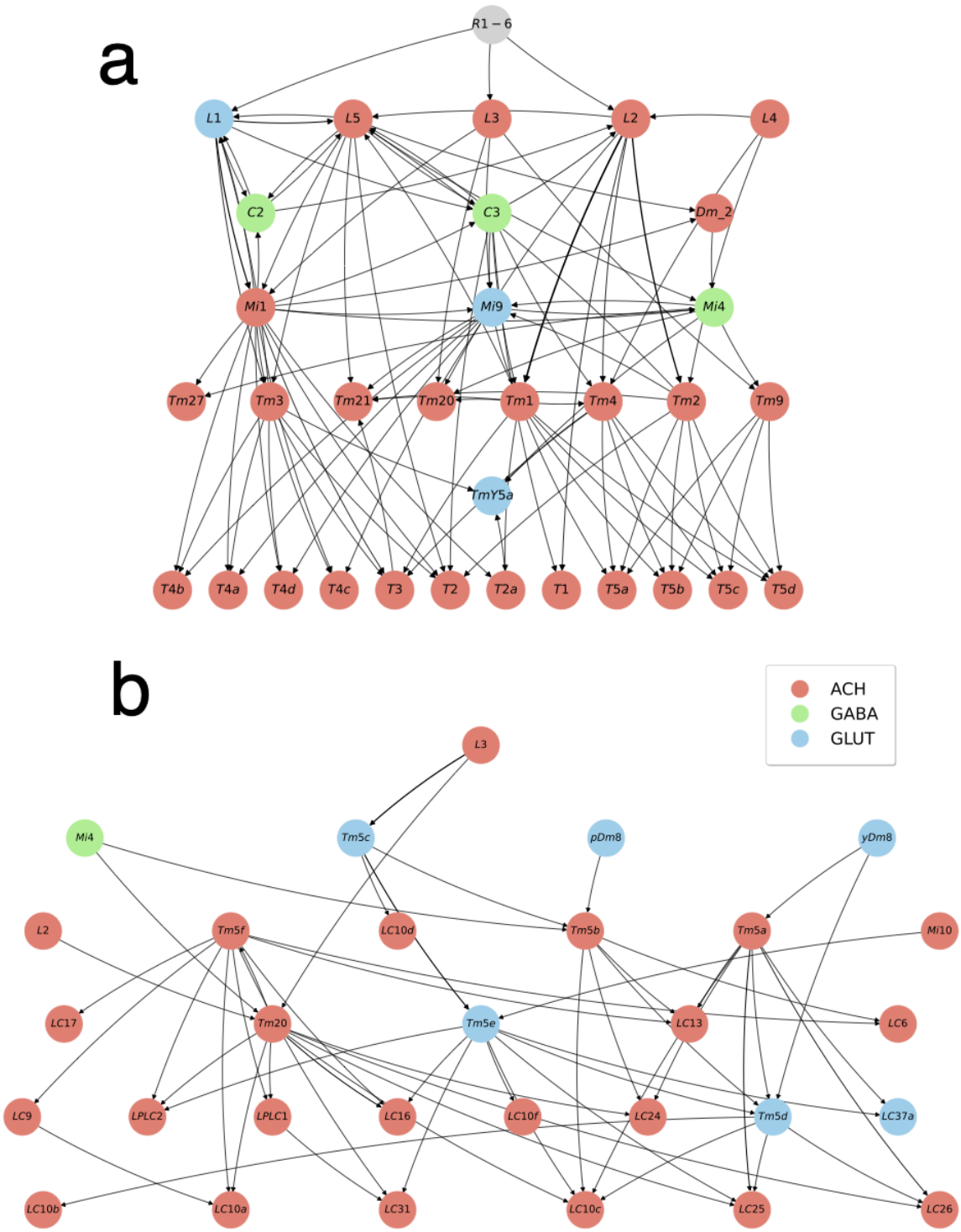

Figure 4. Wiring diagrams of selected cell types as weighted directed graphs.

(a) Synaptic connections between the numerous types, those containing 500+ cells. Types are organized into layers, approximating their distance from the eye. To highlight the strongest pathways, only connections with 5000+ synapses (combined) are shown. Motion pathways are prominent, because the types in motion circuits happen to be numerous types.

(b) Connectivity of Tm5, Tm20, Dm8 types and their prominent partners selected as follows:

(i) Upstream: Types from Medulla and Lamina contributing at least 5% of incoming synapses to any of Tm5, Tm20, Dm8 types

(ii) Downstream: Columnar VPN types (projecting to central brain) that receive at least 5% of their input synapses from Tm5, Tm20, Dm8 types (combined).

(iii) After selecting the partners with these criteria, we include all connections A->B such that the fraction of input synapses coming into B from A is above 2%.

Types are organized into layers algorithmically, minimizing the number of upward and sideway connections. To highlight the strongest pathways, edge widths scale with absolute number of synapses between the respective pair of types.

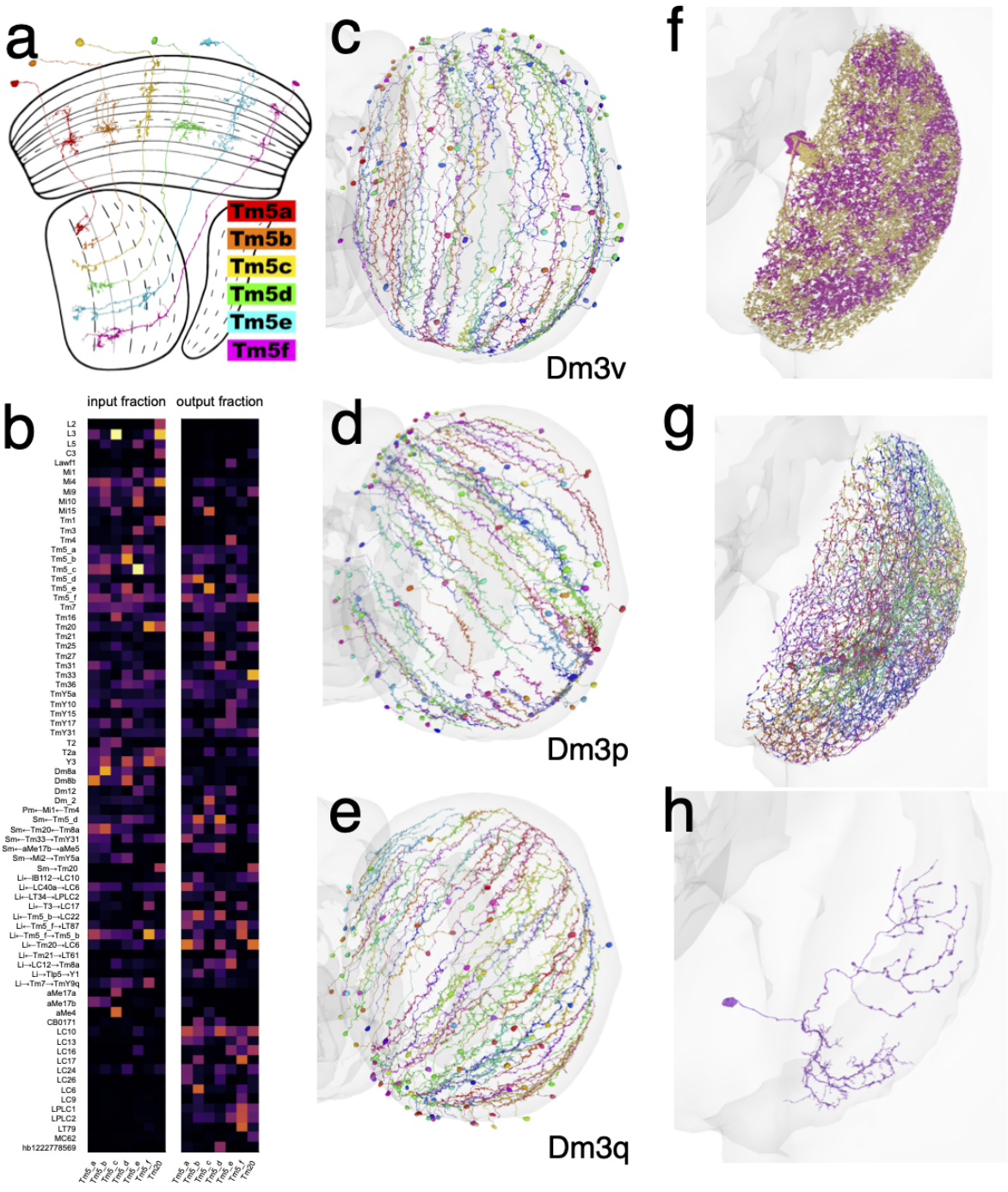

Figure 5. Tm5, Dm3, and LPi

(a) Tm5 types are doubled from three to six, with implications for color vision.
(b) Input and output fractions for chromatic Tm types.
(c) A new type Dm3v is added to known types
(d) Dm3p and
(e) Dm3q, with implications for form vision (H. Sebastian Seung 2023).

(f) LPi⇐T5a⇒H2, called LPi1-2 by (Shinomiya et al. 2022), is a jigsaw pair of full-field cells.
(g) LPi⇐T5a⇒LPLC2 could be called “LPi1-2 junior” based on its stratification and smaller size, but we prefer to use our discriminative names based on connectivity.
(h) LPi⇐T5c⇒LPLC4 is polarized in the tangential plane, and hence not amacrine.

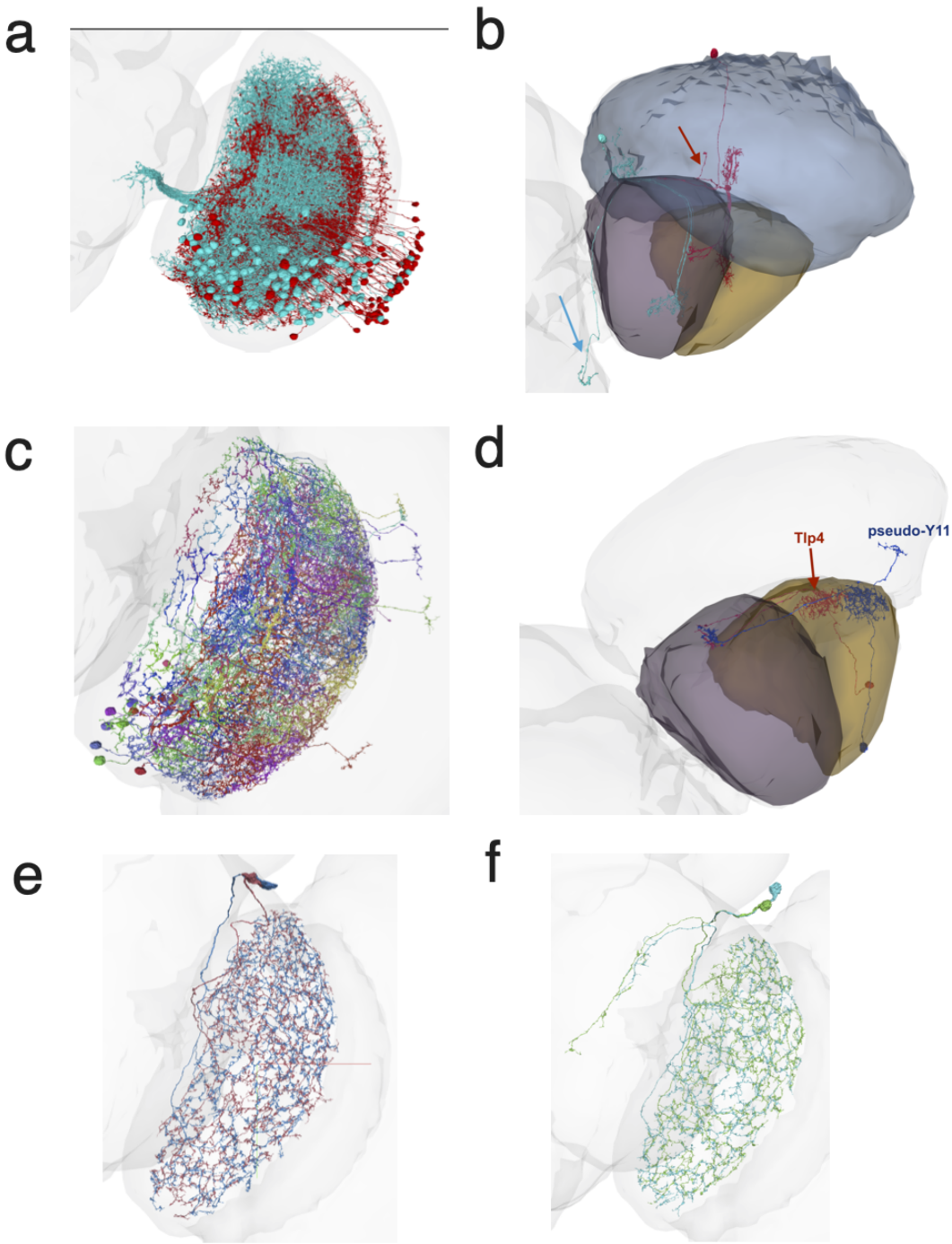

Figure 6. Morphological variation

(a) The TmY14 cell type exhibits a mixture of typical and atypical cell morphologies, with projections extending to the central brain or the medulla. Some of the atypical ones project toward the central brain but retract without reaching it, instead projecting into the medulla. Their connectivity remains consistent regardless of whether the cell projects to the central brain.

(b) Representative typical (cyan) and atypical (red) TmY14 with arbor projecting into the central brain and medulla respectively.
(c) A few Tlp4 cells exhibit Y11 - like morphology, but have the same connectivity as Tlp4. We call these cells pseudo-Y11.
(d) Morphological comparison of Tlp4 and pseudo-Y11. Pseudo-Y11 has an additional branch in the medulla.
(e) Li11 does not project into the central brain.
(f) Pseudo-Li11 has an additional arbor projection into the central brain. This arbor makes a few synapses, and might lead to the conclusion that pseudo-Li11 should be categorized as Li11. However, the connectivity between Li11 and pseudo-Li11 is fundamentally different, making them distinct types.

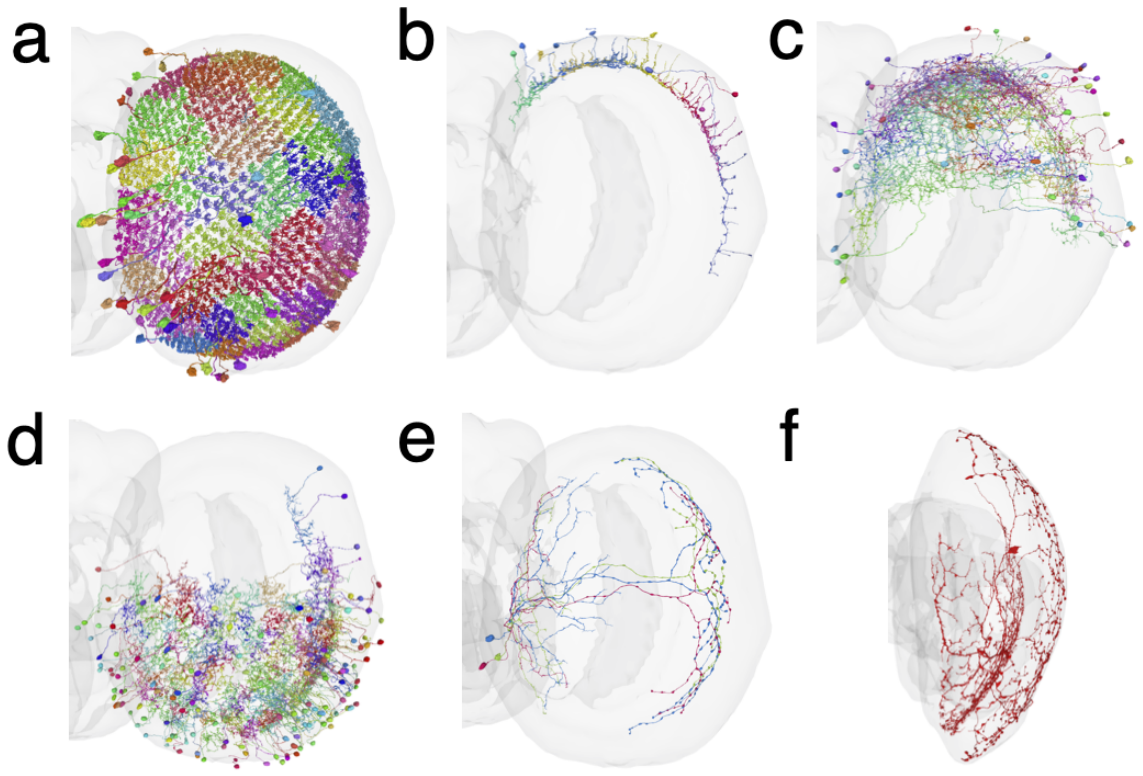

Figure 7. Different kinds of spatial coverage

(a) Example (Dm4) of complete tiling with no overlap
(b) Example (DmDRA2) of dorsal rim coverage
(c) Example (Sm⇐MC65⇒TmY5a) of dorsal hemifield coverage
(d) Example (Sm⇐Mi9⇒CB0165) of ventral hemifield coverage
(e) Example (Sm⇐IB029⇐MeTu1) of H-shaped coverage (anterior and posterior rim)
(f) Singleton (Sm⇐aMe4⇒Mi15) with mixed coverage: dorsal dendritic arbor in M7 and full-field axonal arbor in M1.

Supplementary Figures

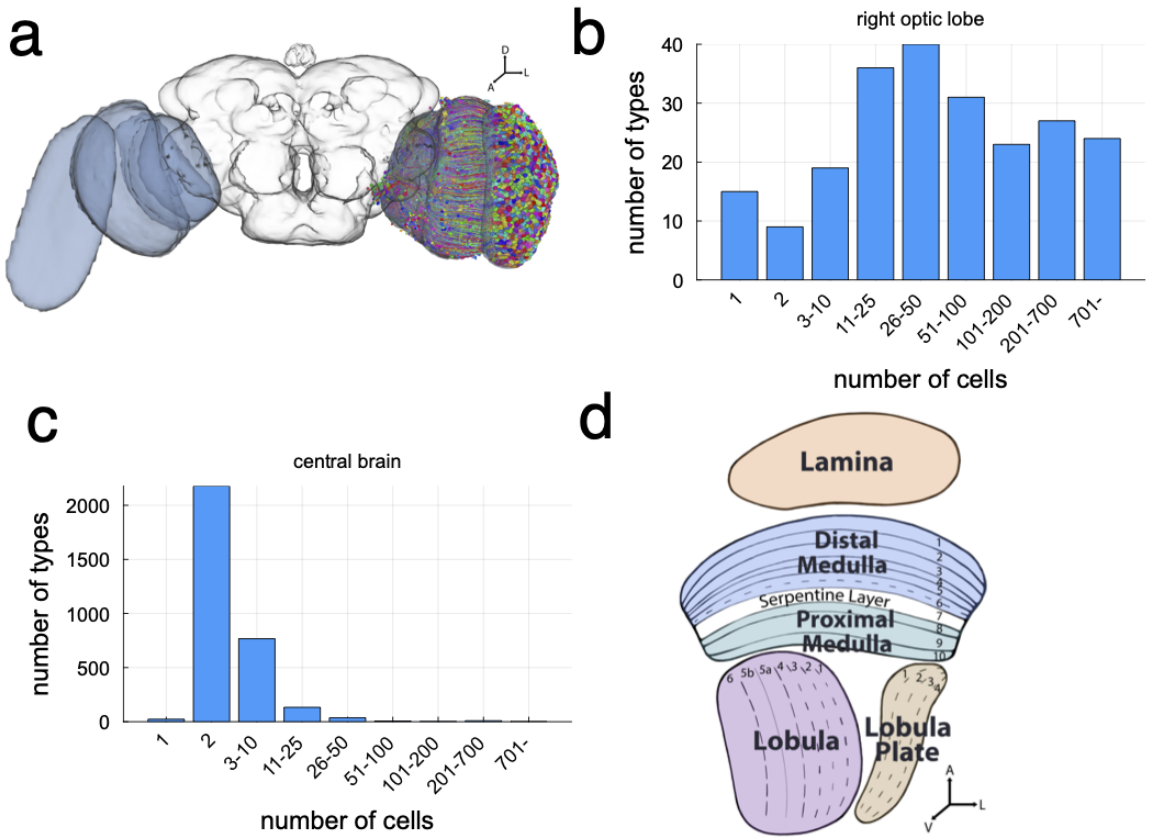

Figure S1. Cell counts of optic lobe types versus central brain types

(a) *Drosophila* central brain and flanking optic lobes. (A: Anterior. L: Lateral. D: Dorsal)

(b) Distribution of number of **optic lobe** types by bucketed unilateral cardinality. Each bar represents types whose cardinality (number of cells) is within the specified range. Most types contain 10+ cells, and a significant portion of types contain hundreds of cells.

(c) Distribution of the number of **central brain** types by bucketed bilateral cardinality. In contrast to the optic lobe, here most types have cardinality 2 (cell and its mirror twin in the opposite hemisphere).

(d) Optic lobe main neuropils (brain regions) and their layering. (A: Anterior. L: Lateral. V: Ventral)

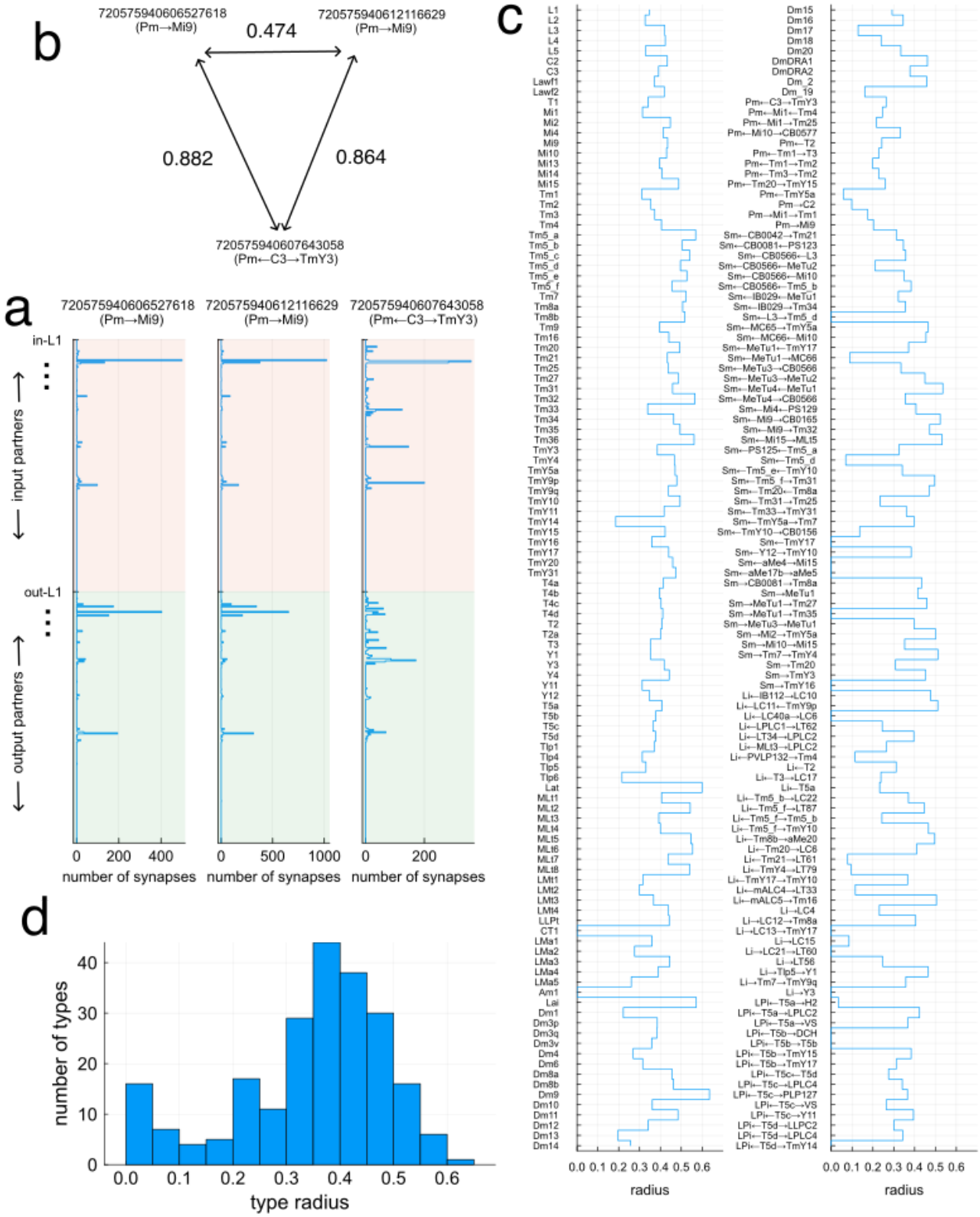

Figure S2. Feature vectors and distances in connectivity space

Panels are labeled in clockwise sequence.

(a) Feature vectors for three example cells. On the Y-axis are the input (red) and output (green) features corresponding to types. On the X-axis is the number of input / output synapses the cell has with those types respectively.
(b) Jaccard distances between feature vectors of example cells.
(c) Radii of types in feature space.
(d) Histogram of type radii.

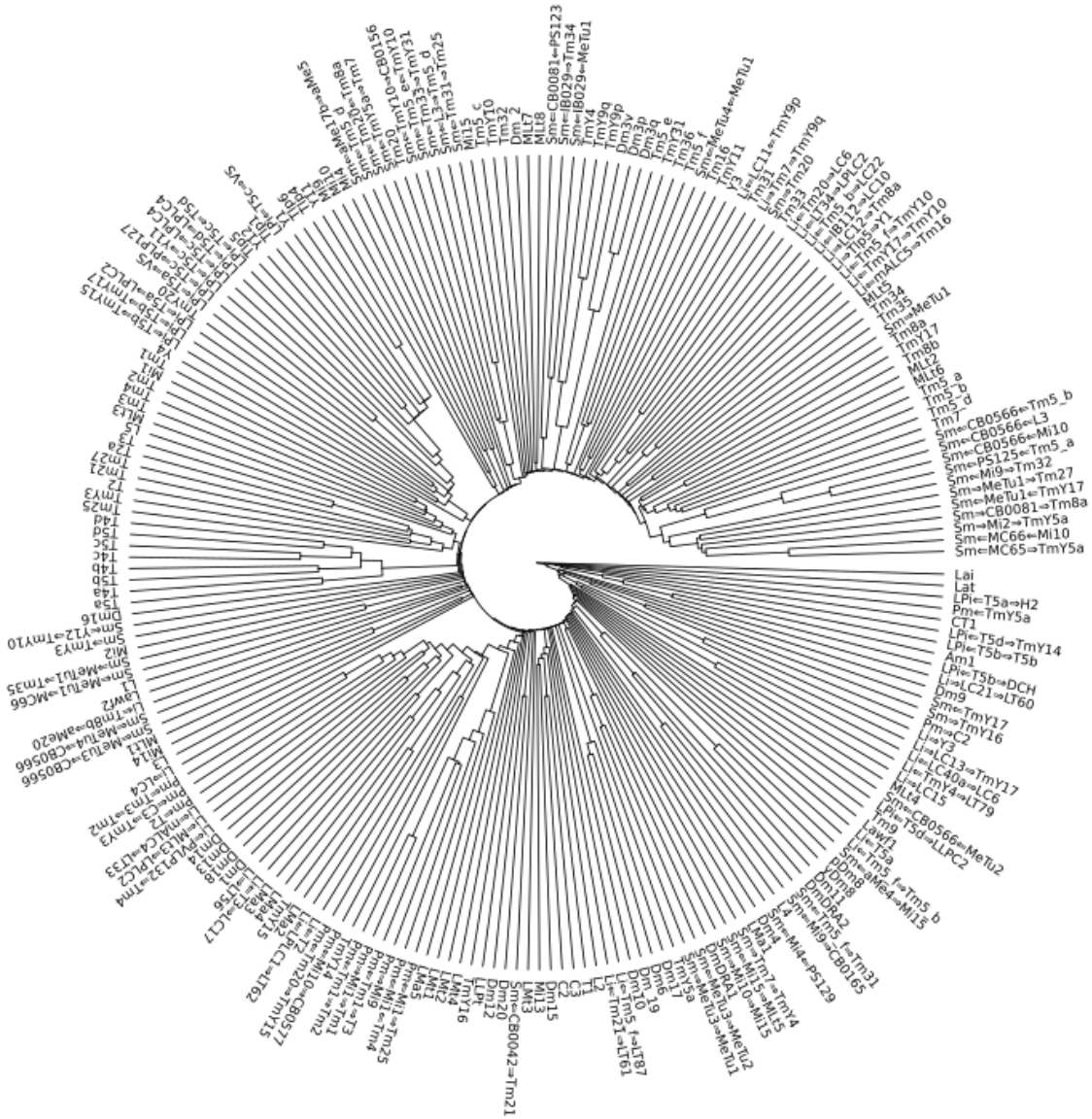

Figure S3. Dendrogram of type centers

Type centers were computed using Eq. (8), and single linkage hierarchical clustering was applied to organize the types into a tree. Radial axis is the square of Jaccard distance for visibility.

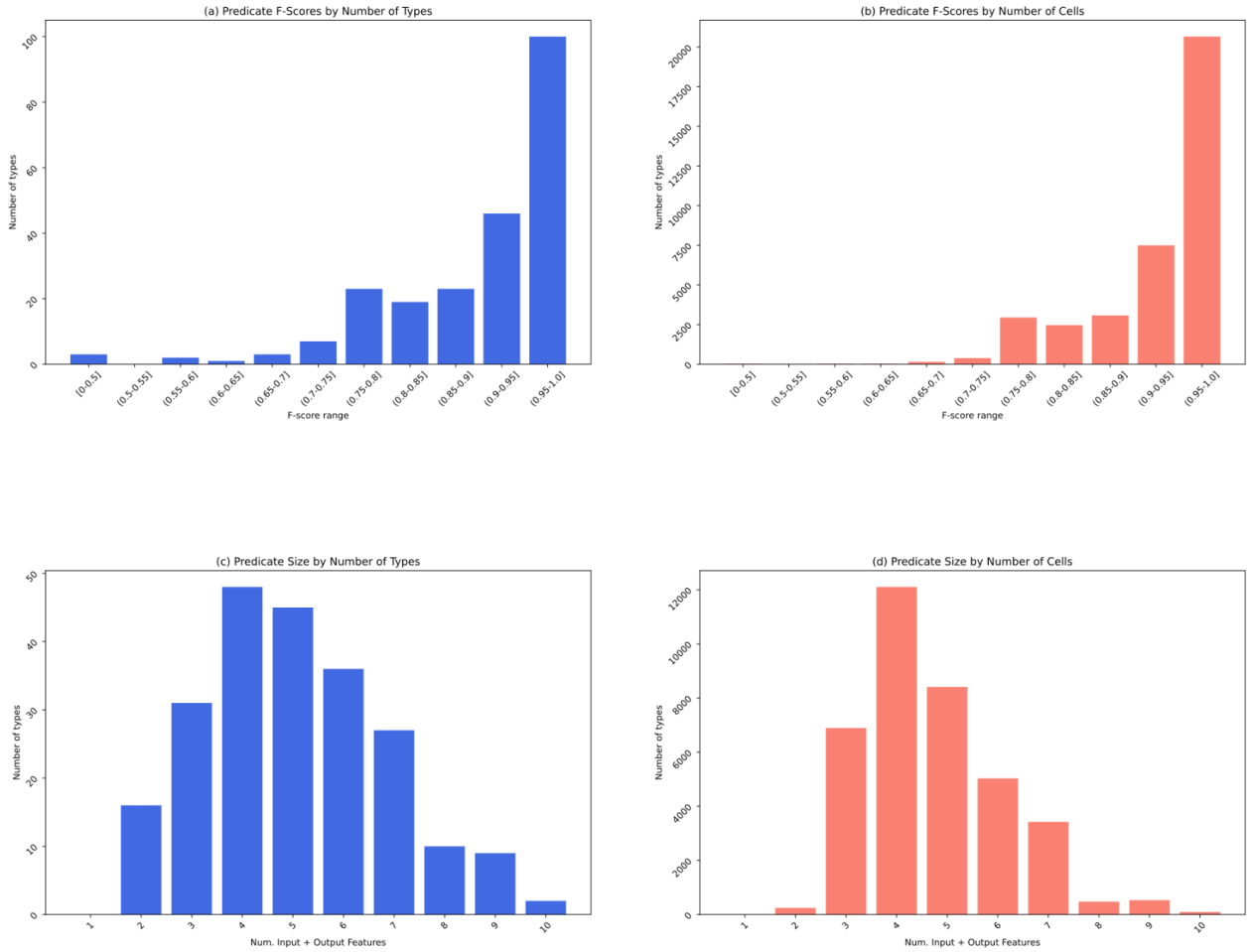

Figure S4. Logical connectivity predicate statistics

(a) Number of types by predicate f-score range.
(b) Number of cells by their types' predicate f-score range.
(c) Number of types by predicate size, that is the sum of the number of input features and output features participating in the binary conjunction.
(d) Number of cells by their types' predicate size.

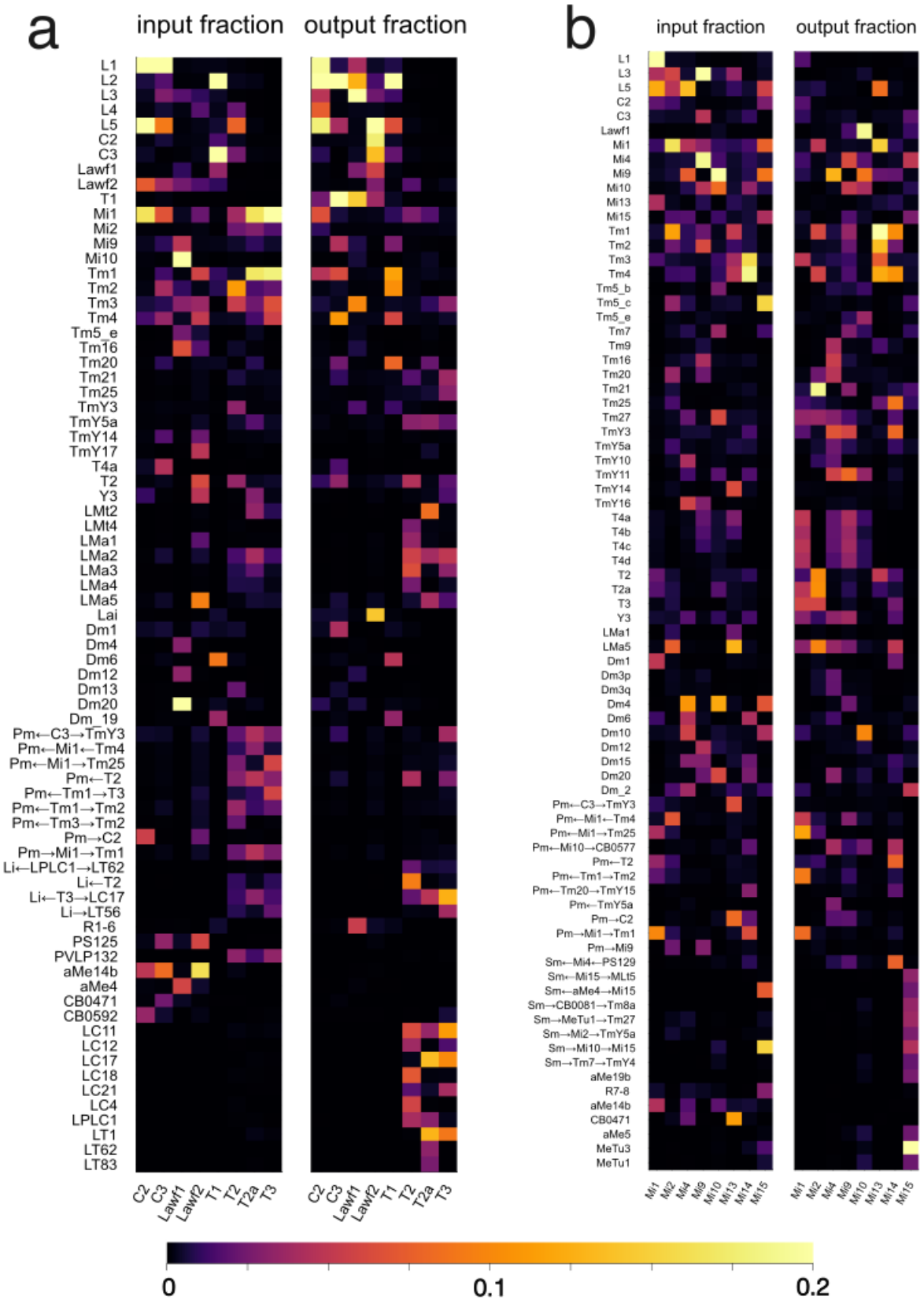

Figure S5. Lamina columnar and Mi input and output fractions.

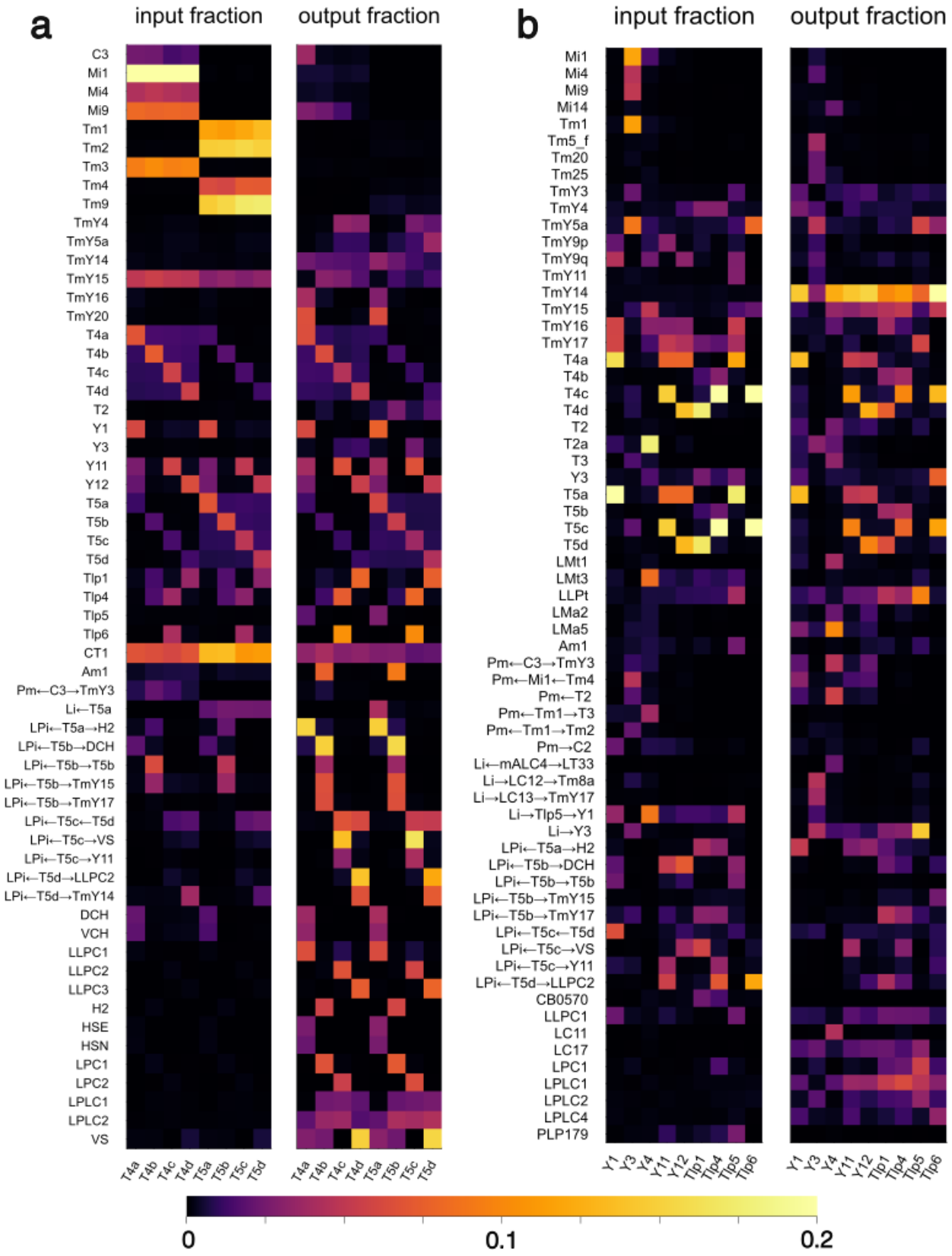

Figure S6. T4, T5, Y and Tlp input and output fractions.

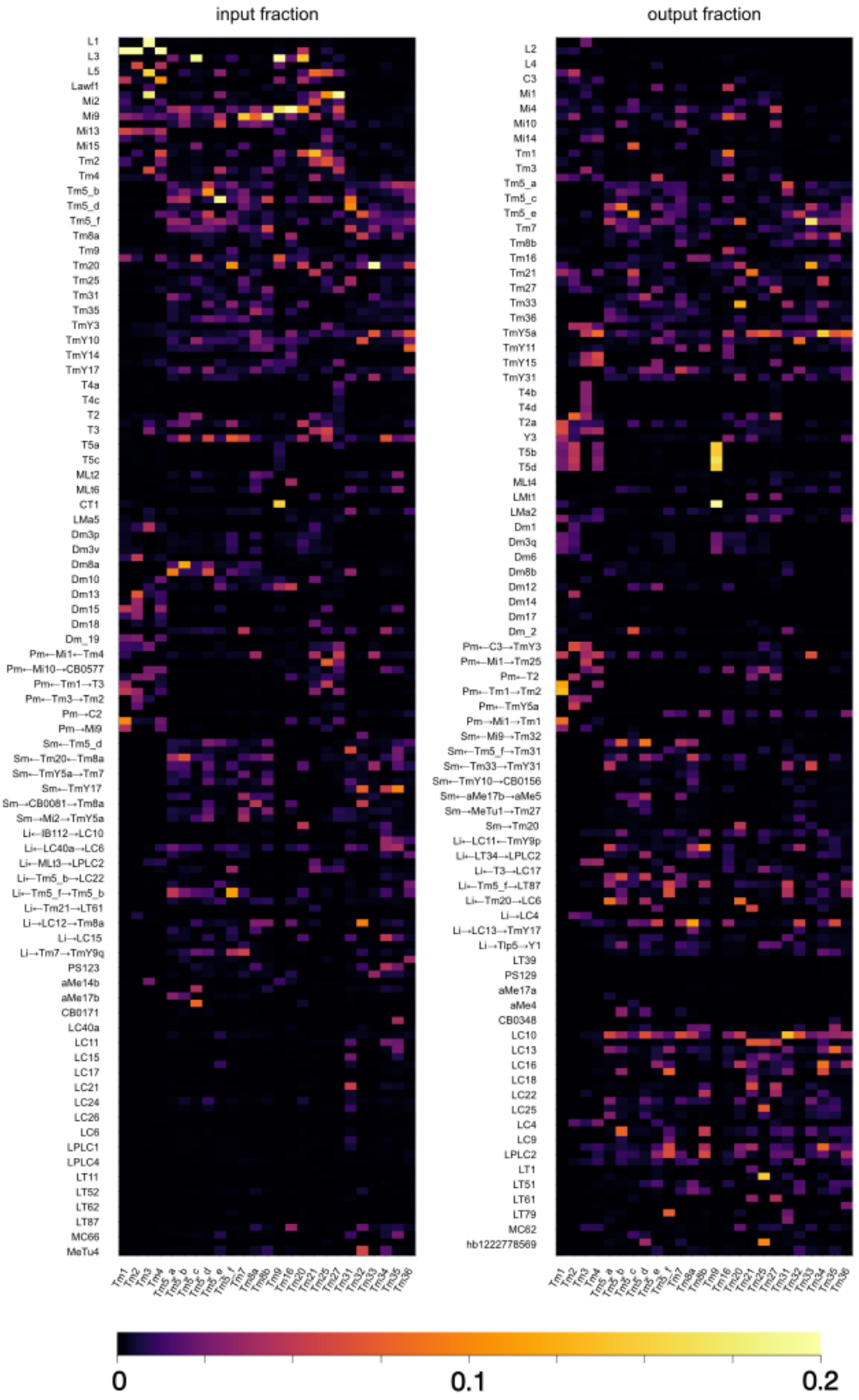

Figure S7. Tm input and output fractions.

For legibility, the type names on the vertical axis alternate between the two columns.

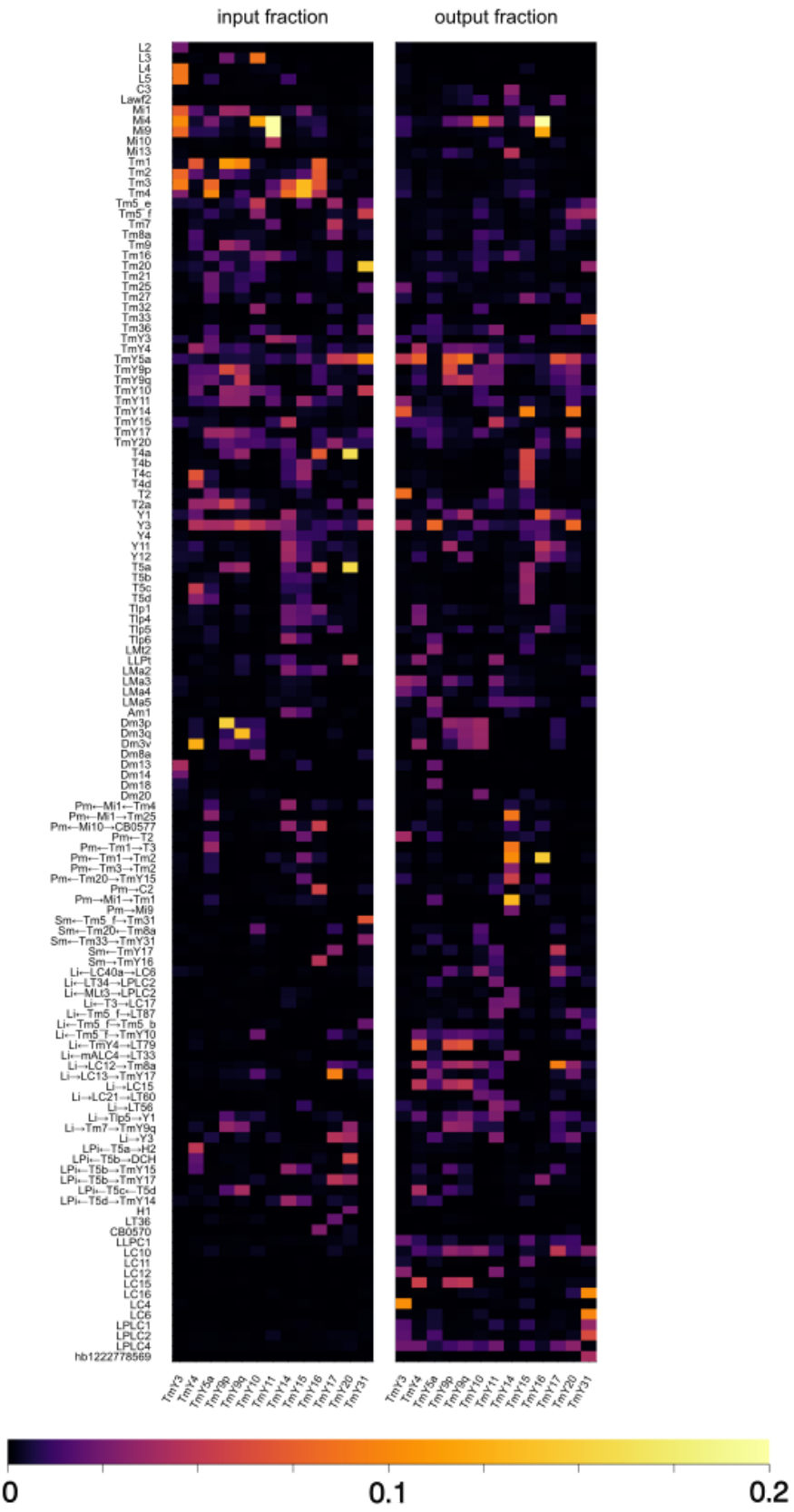

Figure S8. TmY input and output fractions.

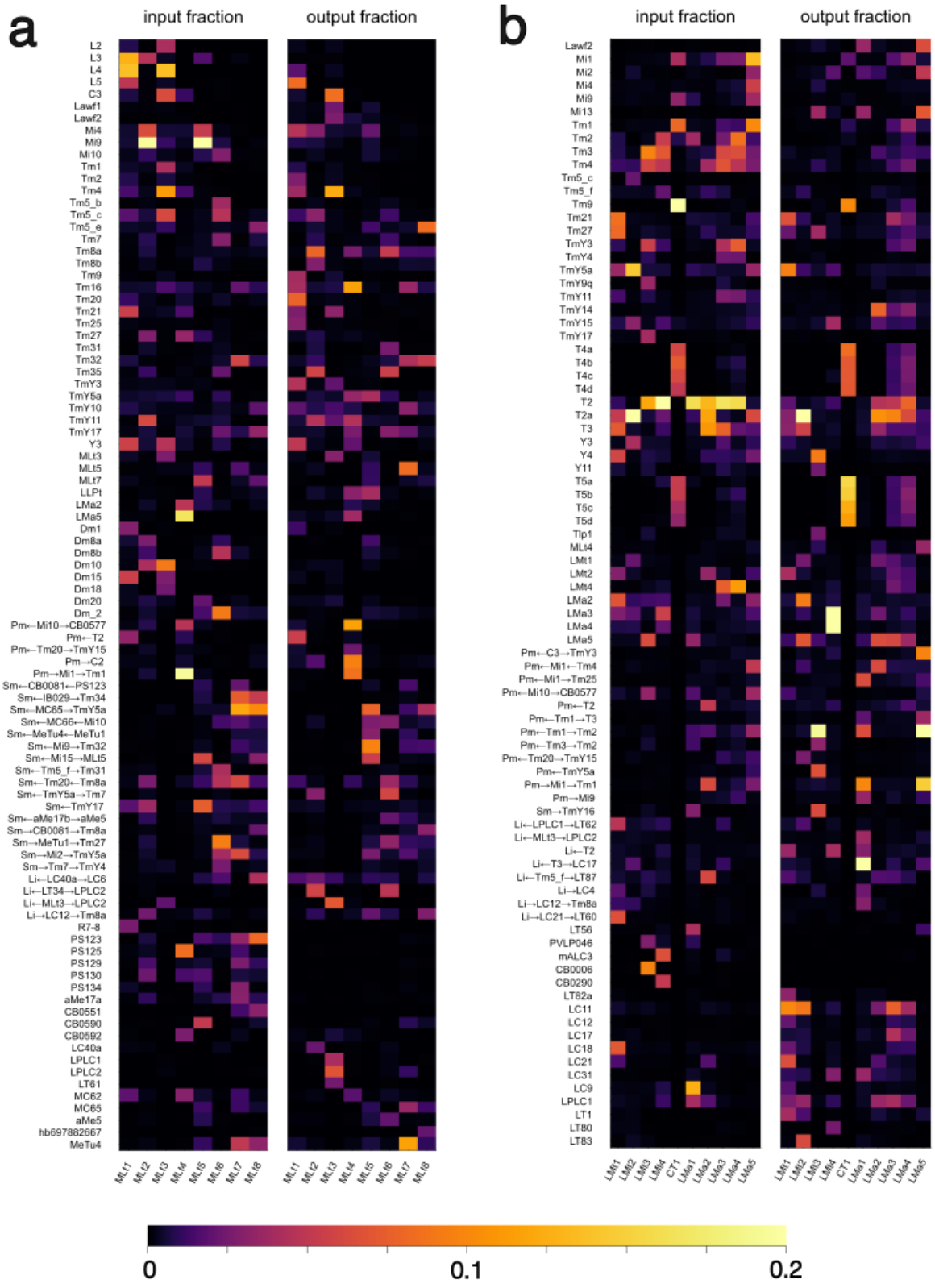

Figure S9. Cross-neuropil tangential and amacrine input and output fractions.

(a) MLt. Medulla Lobula tangential
(b) LMt, LMa. Lobula Medulla tangential and amacrine

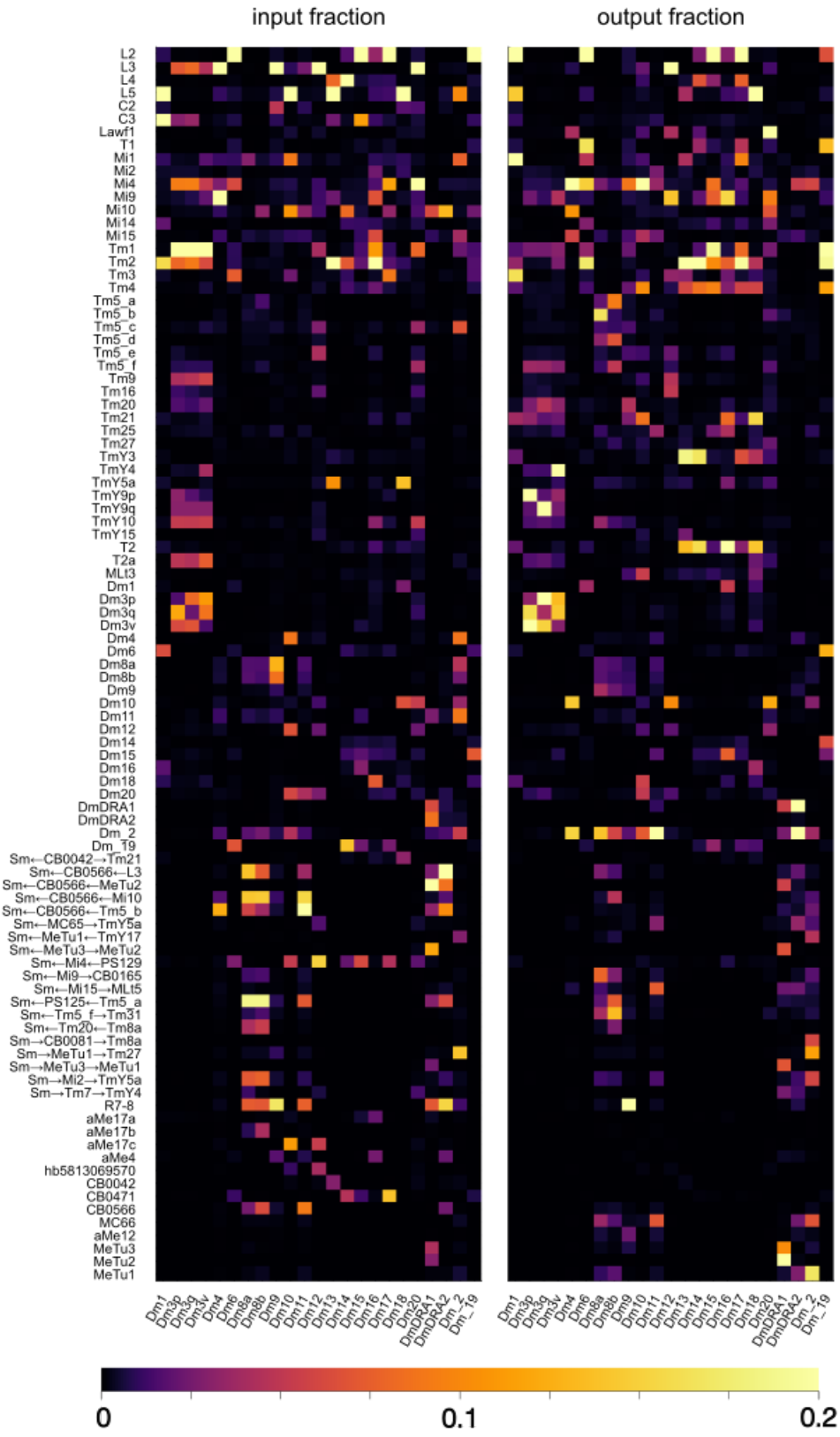

Figure S10. Distal medulla (Dm) input and output fractions.

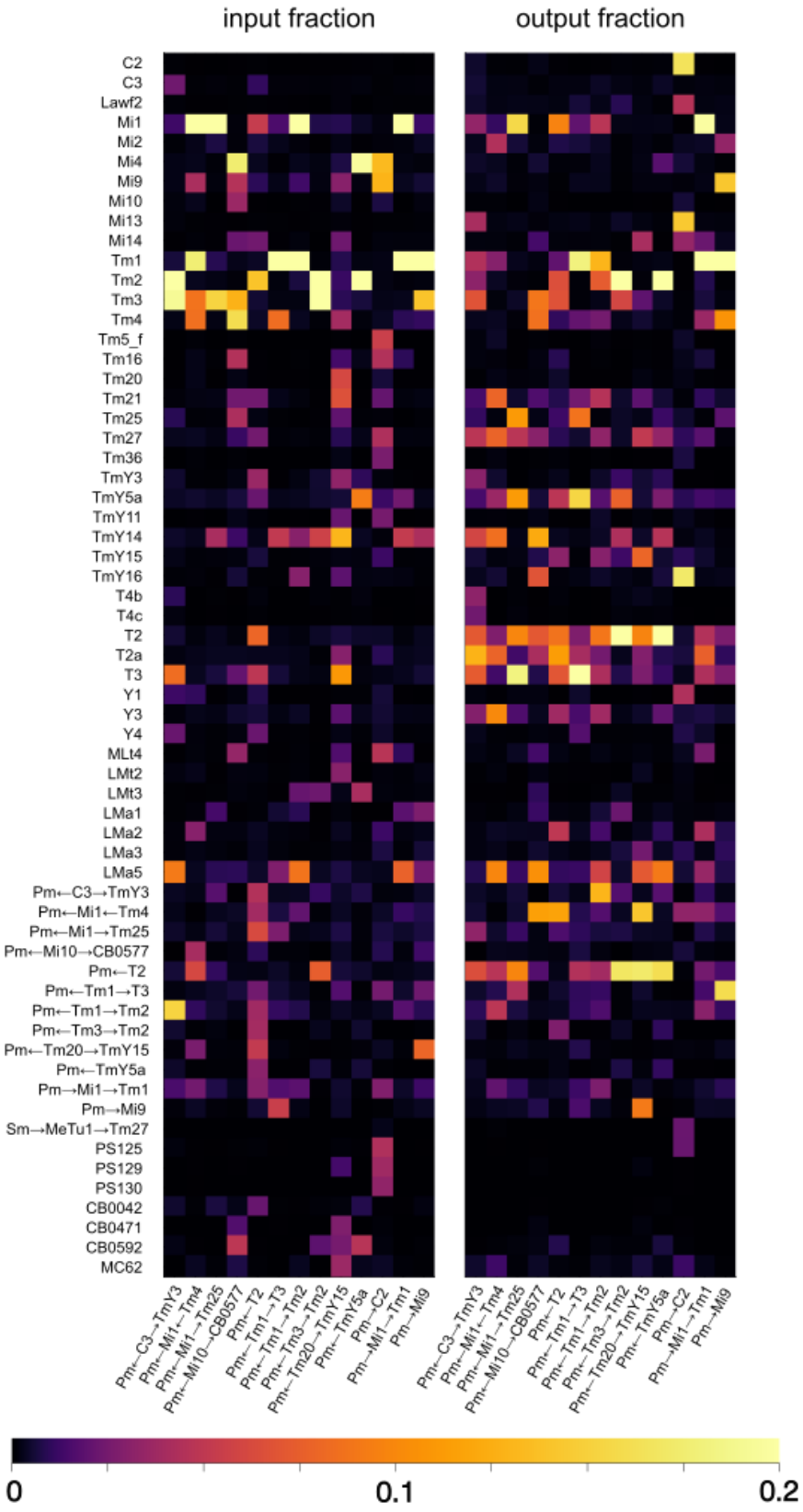

Figure S11. Proximal medulla (Pm) input and output fractions.

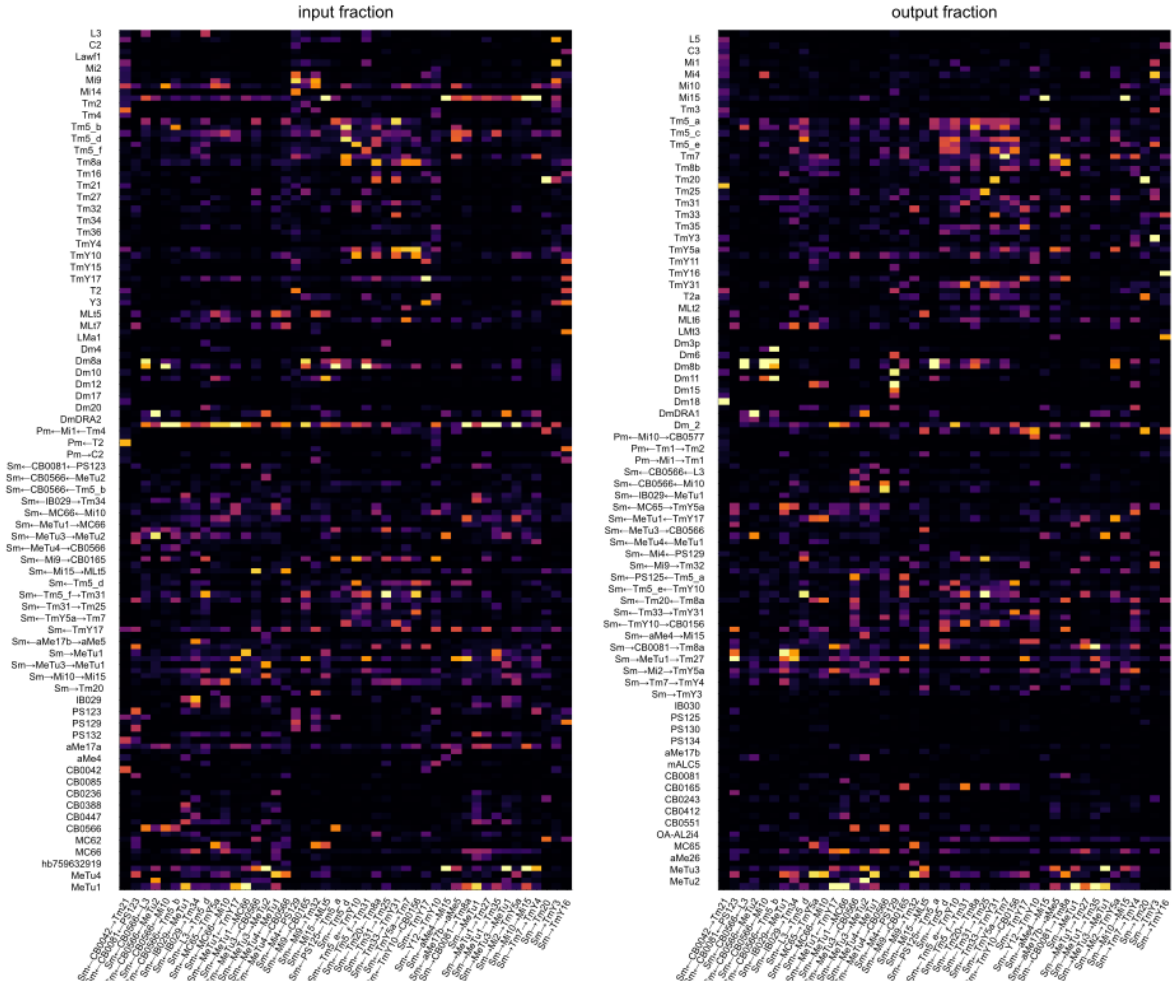

Figure S12. Serpentine medulla (Sm) input and output fractions.

For legibility, the type names on the vertical axis alternate between the two columns.

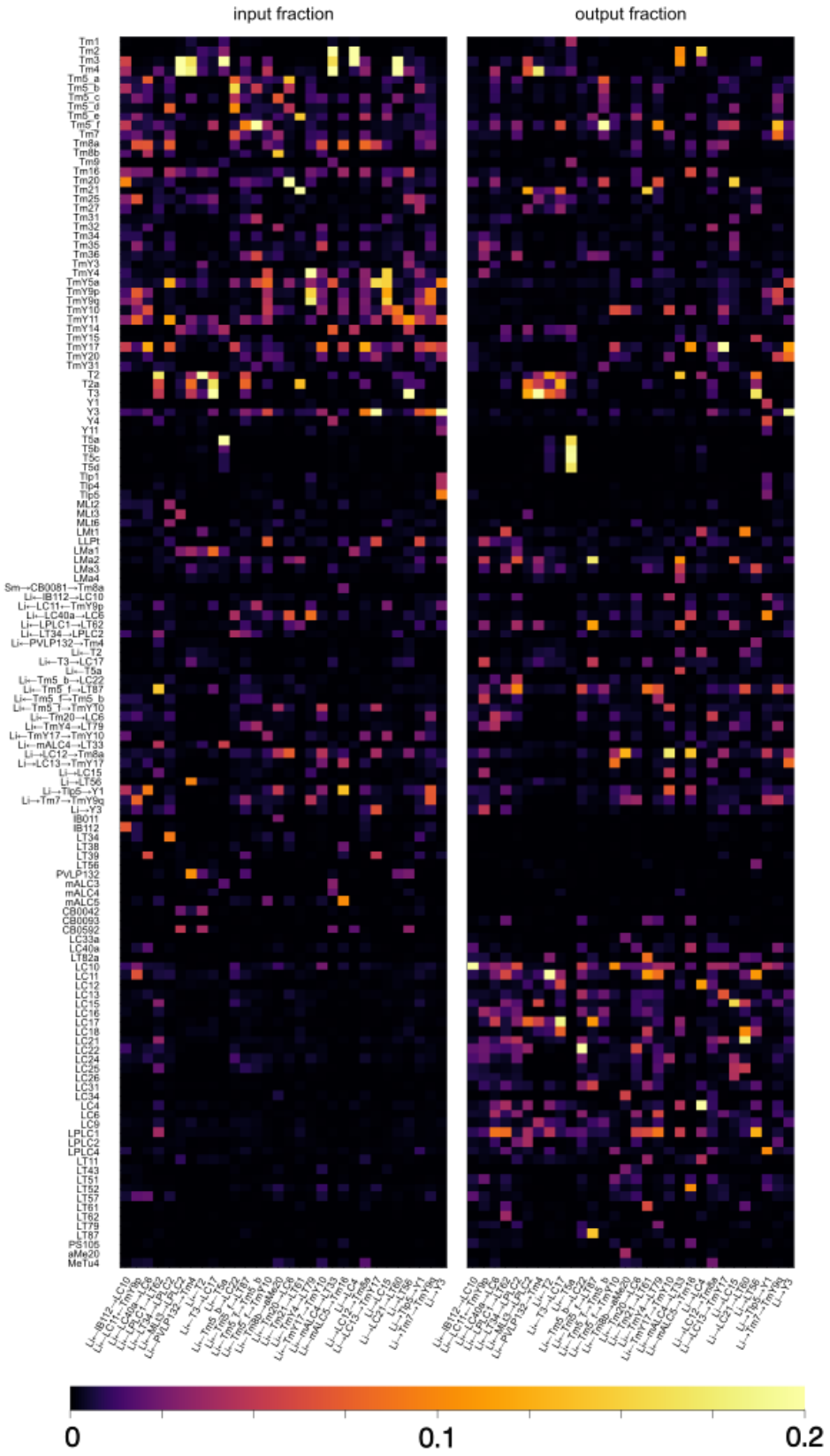

Figure S13. Lobula intrinsic (Li) input and output fractions.

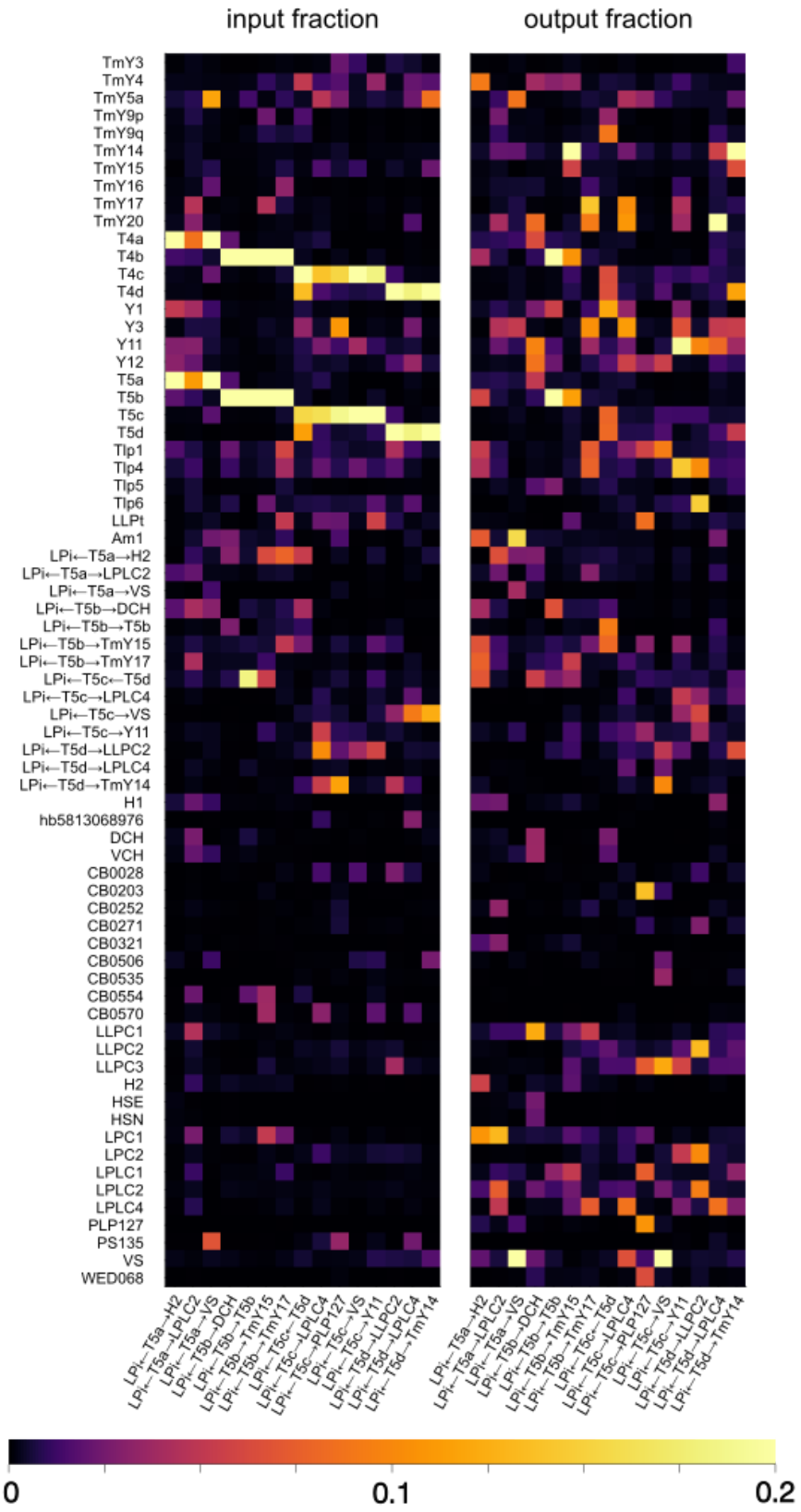

Figure S14. Lobula plate intrinsic (LPi) input and output fractions.

Supplementary Tables

Table S1 (link) Master table of cell types

	AFFINITY	LINKAGE	TYPES	CELLS	TRANS	NEUROPILS
Centrifugal	Cross Neuropil	Axon Bearing / Columnar	2	1441	GABA	ME → ME, LA
Distal Medulla	Neuropil Intrinsic	Non-columnar	21	3137	GLUT	ME → ME
Distal Medulla Dorsal Rim Area	Neuropil Intrinsic	Non-columnar	2	39	GLUT	ME, LOP → ME, LOP
Lamina Intrinsic	Neuropil Intrinsic	Non-columnar	1	45		LA → LA
Lamina Monopolar	Cross Neuropil	Axon Bearing / Columnar	5	3570	ACH	ME, LA → ME
Lamina Tangential	Neuropil Intrinsic	Axon Bearing / Tangential	1	4	SER	ME →
Lamina Wide Field	Cross Neuropil	Axon Bearing / Columnar	2	310	ACH	ME → LA, ME
Lobula Intrinsic	Neuropil Intrinsic	Non-columnar	30	736	GABA	LO → LO
Lobula Lobula Plate Tangential	Cross Neuropil	Axon Bearing / Tangential	1	37	GABA	LO, LOP → LO, LOP
Lobula Medulla Amacrine	Cross Neuropil	Amacrine	6	134	GABA	ME, LO, LOP → ME, LO, LOP
Lobula Medulla Tangential	Cross Neuropil	Axon Bearing / Tangential	4	85	GLUT	LO, ME → LO, ME
Lobula Plate Intrinsic	Neuropil Intrinsic	Non-columnar	15	357	GLUT	LOP, LO → LOP
Medulla Intrinsic	Neuropil Intrinsic	Axon Bearing / Columnar	8	3822	ACH	ME → ME, LOP
Medulla Lobula Lobula Plate Amacrine	Cross Neuropil	Amacrine	1	1	GABA	LOP, LO, ME → LOP, LO, ME
Medulla Lobula Tangential	Cross Neuropil	Axon Bearing / Tangential	8	264	ACH	ME, LO, LOP → ME, LO, LOP
Proximal Medulla	Neuropil Intrinsic	Non-columnar	13	589	GABA	ME, LOP → ME, LOP
Serpentine Medulla	Neuropil Intrinsic	Non-columnar	45	1450	GABA	ME → ME
T Neuron	Cross Neuropil	Axon Bearing / Columnar	12	9085	ACH	ME, LO, LOP → LOP, LO, ME
Translobula Plate	Cross Neuropil	Axon Bearing / Columnar	4	171	GLUT	LOP, LO → LOP, LO
Transmedullary	Cross Neuropil	Axon Bearing / Columnar	25	8643	ACH	ME, LO → ME, LO, LOP
Transmedullary Y	Cross Neuropil	Axon Bearing / Columnar	13	2686	ACH	ME, LOP, LO → LO, ME, LOP
Weirdos	Cross Neuropil	Non-columnar	3	3	GABA	LO, LOP → LO
Y Neuron	Cross Neuropil	Axon Bearing / Columnar	5	625	GLUT	LOP, ME, LO → LO, LOP, ME

Table S2. Type families and their properties.

	⇑ LA	⇑ ME	⇑ LO	⇑ LOP	⇓ LA	⇓ ME	⇓ LO	⇓ LOP
Centrifugal	2,108	103,589	154	1,057	16,643	178,530	22	366
Distal Medulla	15	438,707	47	289	0	414,276	139	492
Distal Medulla Dorsal Rim Area	0	5,018	0	686	0	4,781	0	827
Lamina Intrinsic	6,384	21	0	0	692	5	0	0
Lamina Monopolar	67,820	313,413	0	41	4,418	949,739	0	172
Lamina Tangential	0	108	0	0	0	0	0	0
Lamina Wide Field	200	71,532	59	688	14,180	591	0	0
Lobula Intrinsic	0	430	440,189	128	0	164	292,492	29
Lobula Lobula Plate Tangential	0	6	18,326	1,456	0	13	13,069	6,454
Lobula Medulla Amacrine	0	195,155	116,447	13,119	0	121,027	113,744	9,762
Lobula Medulla Tangential	0	5,466	39,879	465	0	19,898	24,611	1,112
Lobula Plate Intrinsic	0	154	9,768	304,887	0	69	6,232	211,609
Medulla Intrinsic	67	496,704	921	7,295	3	985,632	6,875	48,405
Medulla Lobula Lobula Plate Amacrine	0	2,778	4,126	25,442	0	2,932	4,652	11,363
Medulla Lobula Tangential	0	27,449	2,043	946	0	20,663	10,168	962
Proximal Medulla	0	848,594	7,503	61,893	0	386,648	2,617	27,968
Serpentine Medulla	0	236,536	69	4,546	0	230,297	76	4,094
T Neuron	3,877	573,507	239,977	84,452	100	91,354	297,145	478,375
Translobula Plate	0	37	4,722	88,619	0	100	17,294	49,630
Transmedullary	0	888,620	90,161	13,950	0	949,893	569,737	47,036
Transmedullary Y	0	348,773	118,011	153,162	0	136,047	207,231	67,668
Weirdos	0	22	3,674	189	0	118	3,882	52
Y Neuron	0	75,091	26,682	85,674	0	51,294	80,217	55,958

Table S3. Distribution of synapses over neuropils for each type family.

	PROOFREAD (LEFT)	TYPED (LEFT)	PROOFREAD (RIGHT)	TYPED (RIGHT)
optic	36k	10%	38k	99%
central	16k	55%	16k	55%
sensory	4k	94%	5k	95%
visual_projection	4k	82%	4k	84%
other	2k	38%	2k	36%
all	63k	33%	65k	84%

Table S4. Status of cell typing

FlyWire connectome dataset as of October 2023.

Table S5 (link) [FlyWire Consortium (optic lobe)]

Supplementary Data

Data S1 (link) Cards for all cell types

Each card has 6 panels (clockwise): Logical connectivity predicate diagram, numeric stats, illustration of a typical cell, followed by 3D rendering of all cells from 3 orthogonal viewpoints.

Data S2 (link) Discriminating logical predicates for all types

Each figure contains types from the same family (middle layer) with shared input attributes (left layer) and output attributes (right layer) that are sufficient for discriminating all types in the middle layer. Families with many types are split into multiple figures for clarity of presentation.

Data S3 (link) Discriminating 2D projections for neuropil-intrinsic types

Methods

Stage 1: Crowdsourced annotation of known types

Annotations of optic lobe neurons were initially crowdsourced. The first annotators were volunteers from *Drosophila* labs. They were later joined by citizen scientists. At this stage, the annotation effort was mainly devoted to labeling cells of known types, especially the most numerous types.

Drosophila lab annotators

Labs that researched the *Drosophila* anterior visual pathway reconstructed and annotated medulla neurons that were upstream of the pathway. These included many of the medulla and lamina neurons discussed in this study. The annotated neurons were primarily Dm2, Mi15, R7, and R8, but also comprised various L, Dm, Mi, Tm, C, and Sm cells. Previously known neuron types were identified primarily by morphology and partially by connectivity. Annotators additionally found all Mi1 neurons in both hemispheres in order to find every medulla column. These Mi1 neurons were used to create the map of medulla layers based on Mi1 stratification in (Fischbach and Dittrich 1989). This layer map would later be used by the citizen scientists to identify medulla cell types.

Citizen scientists

The top 100 players from Eyewire (Kim et al. 2014) had been invited to proofread in FlyWire (Dorkenwald, Matsliah, et al. 2023). After three months of proofreading in the right optic lobe, they were encouraged to also label neurons when they felt confident. Most citizen scientists did a mixture of annotation and proofreading. Sometimes they annotated cells after proofreading, and other times searched for cells of a particular type to proofread.

Citizen scientists were provided with a visual guide to optic lobe cells sourced from (Fischbach and Dittrich 1989) and Davis et. al., 2020. FlyWire made available a 3D mesh overlay indicating the four main optic lobe neuropils. Visual identification was primarily based on single-cell anatomy (neuropils, stratification, and morphology). Initially labeling of type families (ie. Dm, Tm, Mi, etc.) was encouraged, especially for novices. Annotation of specific types (ie. Dm3, Tm2, etc.) developed over time. Type names from (Fischbach and Dittrich 1989) were learned by the citizen scientists from the visual guide. Use of canonical names was further enforced by a software tool that enabled easy selection and submission of pre-formatted type names.

Additional community resources (discussion board/forum, blog, shared Google drive, chat, dedicated email, and Twitch livestream) fostered an environment for sharing ideas and information between community members (citizen scientists, community managers, and researchers). Community managers answered questions, provided resources such as the visual guide, shared updates, performed troubleshooting and general organization of community activity. Daily stats including number of annotations submitted per individual were shared on the discussion board/forum to provide project progress. Live interaction, demonstrations and communal problem solving occurred during weekly Twitch video livestreams led by a community manager. The environment created by these resources allowed citizen scientists to self-organize in several ways: community driven information sharing, programmatic tools, and “farms.”

Community-driven information sharing

Citizen scientists created a comprehensive guide with text and screenshots that expanded on the visual guide. They also found and studied any publicly available scientific literature or resources regarding the optic lobe. They shared findings on the Discussion Board, which as of October 10, 2023 has over 2,500 posts. Community managers interacted with citizen scientists by sharing findings from the scientific literature, consulting *Drosophila* specialists on FlyWire, and providing feedback.

Programmatic Tools

Programmatic tools were created to help with searching for cells of the same type. One important script traced partners-of-partners, i.e., source cell → downstream partners → their upstream partners, or source cell → upstream partners → their downstream partners. This was based on the assumption that cells of the same type will probably synapse with the same target cells, which often turned out to be true. The tool could either look for partners-of-all-partners or partners-of-any-partners. The resulting lists of cells could be very long, and were filtered by excluding cells that had already been identified, or excluding segments with small sizes or low ID numbers (which had likely not yet been proofread). Another tool created from LPTCs (e.g. HS, VS, H1) aided definition of layers in the lobula plate. This facilitated identification of various cell types, especially T4 and T5.

Cell “Farms”

Citizen scientists created "farms" in FlyWire or Neuroglancer with all the found cells of a given type visible. Farms showed visually where cells still remained to be found. If they found a bald spot, a popular method to find missing cells was to move the 2D plane in that place and add segments to the farm one after another in search of cells of the correct type. Farms also helped with identifying cells near the edges of neuropils, where neurons are usually deformed. Having a view of all other cells of the same type made it possible to extrapolate to how a cell at the edge should look.

Stage 2: Centralized annotation and discovery of new types

A team of image analysts at Princeton finished the annotation of the remaining cells in known types, and also discovered new types. We began by utilizing community annotations from each type. To ensure accuracy, these annotations were initially cross-referenced with existing literature to confirm their validity. Once validated, these cells were used to query various Codex search tools, including “similar connectivity”, “similar embedding” and “similar type projections”. These tools returned previously unannotated cells exhibiting connectivity similar to that of the cell in the query. The hits from the search query were evaluated by morphology and stratification to confirm match with the target cell type. In some cases where cell type distinctions were uncertain, predicted neurotransmitters (Eckstein et al. 2023) were used for additional guidance. This process allowed us to create a preliminary clustering of all previously known and new types.

Connectivity: cell-to-cell, type-to-cell, cell-to-type, and type-to-type

For further refinement of the preliminary clustering, we switched to techniques based on the feature vectors of Fig. 2a and S2a. The feature vectors are defined formally as follows. Define a (weighted) cell-to-cell connectivity matrix w_{ij}, as the number of synapses from neuron *i* to neuron *j*. The synapse out-degree and in-degree of neuron *i* are:

$$d_i^+ = \sum_j w_{ij} \qquad d_i^- = \sum_j w_{ji} \tag{1}$$

The sums are over all neurons in the brain. If neuron *i* is a cell intrinsic to one optic lobe, the only nonvanishing terms in the sums are due to the intrinsic and boundary neurons for that optic lobe.

Let A_{it} be the 0-1 matrix that assigns neuron *i* to type *t*. The column and row sums of the assignment matrix satisfy

$$n_t = \sum_i A_{it} \qquad 1 = \sum_t A_{it} \tag{2}$$

where n_t is the number of cells assigned to type *t*. The assignment matrix may include only intrinsic types (Fig. 2a and S2a), or both intrinsic and boundary types.

The cell-to-type connectivity matrix O_{it} is the number of output synapses from neuron *i* to neurons of type *t*,

$$O_{it} = \sum_j w_{ij} A_{jt} \tag{3}$$

and the type-to-cell connectivity matrix I_{tj} is the number of input synapses from neurons of type *t* onto neuron *j*,

$$I_{tj} = \sum_j A_{it} w_{ij} \tag{4}$$

The *i*th row and ith column of these matrices are concatenated to form the feature vector (Fig. 2a and S2a) used for computing Jaccard similarities and distances. These matrices are normalized by degree to yield the output and input fractions of cell *i*, O_{it}/d_i^+ and I_{it}/d_i^-. Elements of these matrices are used for the 2D projections (Fig. 2d).

The type-to-type connectivity matrix is the number of synapses from neurons of type *s* to neurons of type *t*,

$$W_{st} = \sum_{ij} A_{is} w_{ij} A_{jt} \tag{5}$$

The synapse degree of type *t* is the sum of the degrees of the cells in type *t*,

$$D_t^+ = \sum_i A_{it} d_i^+ \qquad D_t^- = \sum_i A_{it} d_i^- \tag{6}$$

Normalizing by degree yields the output fractions of type *s*, W_{st}/D_s^+, where *t* runs from 1 to *T*. The input fractions of type *t* are similarly given by W_{st}/D_t^-, where *s* runs from 1 to *T*. Selected output and input fractions of types are shown in Figs. S5 to S14.

Alternatively, the feature vectors can be based on connection number rather than synapse number, where a connection is defined as two or more synapses from one neuron to another. Such a threshold is intended to suppress noise due to false positives in the automated synapse detection. Synapse number and connection number give similar results, and we use both in our analyses.

Similarity and distance measures

The Jaccard similarity between feature vectors **x** and **y** is defined by

$$J(\mathbf{x}, \mathbf{y}) = \frac{\sum_t \min(x_t, y_t)}{\sum_{t'} \max(x_{t'}, y_{t'})} \tag{7}$$

and the Jaccard distance *d*(**x**, **y**) is defined as one minus the Jaccard similarity. These quantities are bounded between zero and one since our feature vectors are nonnegative. In our cell typing efforts, we have found empirically that Jaccard similarity works better than cosine similarity when feature vectors are sparse.

Cluster centers

Given a set of feature vectors $\mathbf{x}^a$, the center **c** can be defined as the vector minimizing

$$\sum_a d(\mathbf{x}^a, \mathbf{c}) \tag{8}$$

This cost function is convex, since *d* is a metric satisfying the triangle inequality. Therefore the cost function has a unique minimum. We used various approximate methods to minimize the cost function.

For auto-correction of type assignments, we used the element-wise trimmed mean. We found empirically that this gave good robustness to noise from false synapse detections. For the type radii, we used a coordinate descent approach, minimizing the cost function with respect to each c_i in turn. The loop included every *i* for which some x_i was nonzero. This converged within a few iterations of the loop.

Logical Connectivity Predicates

Once types are assigned to cells we compute the logical connectivity predicates (see Discrimination with logical predicates). An optimal predicate is constructed for each type that consists of 2 tuples: input types and output types. Both tuples are limited to size 5 at most, and they are optimal with respect to the f-score of their prediction of the subject type, defined as follows:

- Recall of a predicate for type T is the ratio of true positive predictions (cells matching the predicate) to the total number of true positives (cells of type T). It measures the predicate's ability to identify all positive instances of a given type.
- Precision is the ratio of true positive predictions (predictions that are indeed of type T) to the total number of positive predictions made by the logical predicate.
- F-score is the harmonic mean of precision and recall - a single metric that combines both precision and recall into one value.

On a high level, the process for computing the predicates is exhaustive - for each type we look for all possible combinations of input type tuples and output type tuples and compute their precision, recall and f-score. Few optimization techniques are used to speed up this computation, by calculating minimum precision and recall thresholds from the current best candidate predicate and pruning many tuples early.

Stage 3: Connectivity-based splitting and merging of types and auto-correction

In the final stage, we used computational methods to split types that could not be properly split in Stage 2. Some candidates for splitting were suggested by the image analysts of Stage 2. For example, it had been difficult to assign Tm5 to the three known types. Some candidates were suspicious because they contained so many cells. Finally, some candidates were scrutinized because their type radii were large. We applied hierarchical clustering with average linkage, and accepted the splits if they did not violate the tiling principle as described in Spatial Coverage.

We also applied computational methods to merge types that had been improperly split in Stage 2. Here the candidates were types with low spatial coverage of the visual field, or types that were suspiciously close in the dendrogram of type centers (Fig. S4). Merge decisions were made by hierarchical clustering of cells from types that were candidates for merging, and validated if they improved spatial coverage.

Once we arrived at the final list of types, we automatically re-assigned cells to types by going over all cells and checking which cluster center is closest to their feature vector in Jaccard distance. This process was only applied once, and resulted in reassignment of ~680 cell types (~2%). After inspection we believe these corrections are sound. (The auto-reassigned types are labeled in Codex, along with previously human assigned types.)

References

Akin, Orkun, and S. Lawrence Zipursky. 2020. “Activity Regulates Brain Development in the Fly.” *Current Opinion in Genetics & Development* 65 (December): 8–13.

Alonso, J. M., and L. M. Martinez. 1998. “Functional Connectivity between Simple Cells and Complex

Cells in Cat Striate Cortex." *Nature Neuroscience* 1 (5): 395–403.
Borst, Alexander, Jürgen Haag, and Alex S. Mauss. 2020. "How Fly Neurons Compute the Direction of Visual Motion." *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 206 (2): 109–24.
BRAIN Initiative Cell Census Network (BICCN). 2021. "A Multimodal Cell Census and Atlas of the Mammalian Primary Motor Cortex." *Nature* 598 (7879): 86–102.
Chapman, B., K. R. Zahs, and M. P. Stryker. 1991. "Relation of Cortical Cell Orientation Selectivity to Alignment of Receptive Fields of the Geniculocortical Afferents That Arborize within a Single Orientation Column in Ferret Visual Cortex." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 11 (5): 1347–58.
Christenson, Matthias P., Álvaro Sanz Díez, Sarah L. Heath, Maia Saavedra-Weisenhaus, Atsuko Adachi, L. F. Abbott, and Rudy Behnia. 2023. "Hue Selectivity from Recurrent Circuitry in Drosophila." *bioRxiv : The Preprint Server for Biology*, July. https://doi.org/10.1101/2023.07.12.548573.
Cossell, Lee, Maria Florencia Iacaruso, Dylan R. Muir, Rachael Houlton, Elie N. Sader, Ho Ko, Sonja B. Hofer, and Thomas D. Mrsic-Flogel. 2015. "Functional Organization of Excitatory Synaptic Strength in Primary Visual Cortex." *Nature* 518 (7539): 399–403.
Courgeon, Maximilien, and Claude Desplan. 2019. "Coordination between Stochastic and Deterministic Specification in the Drosophila Visual System." *Science* 366 (6463). https://doi.org/10.1126/science.aay6727.
Cowley, Benjamin R., Adam J. Calhoun, Nivedita Rangarajan, Jonathan W. Pillow, and Mala Murthy. 2022. "One-to-One Mapping between Deep Network Units and Real Neurons Uncovers a Visual Population Code for Social Behavior." *bioRxiv*. https://doi.org/10.1101/2022.07.18.500505.
Currier, Timothy A., Michelle M. Pang, and Thomas R. Clandinin. 2023. "Visual Processing in the Fly, from Photoreceptors to Behavior." *Genetics* 224 (2). https://doi.org/10.1093/genetics/iyad064.
Davis, Fred P., Aljoscha Nern, Serge Picard, Michael B. Reiser, Gerald M. Rubin, Sean R. Eddy, and Gilbert L. Henry. 2020. "A Genetic, Genomic, and Computational Resource for Exploring Neural Circuit Function." *eLife* 9 (January). https://doi.org/10.7554/eLife.50901.
Dorkenwald, Sven, Arie Matsliah, Amy R. Sterling, Philipp Schlegel, Szi-Chieh Yu, Claire E. McKellar, Albert Lin, et al. 2023. "Neuronal Wiring Diagram of an Adult Brain." *bioRxiv : The Preprint Server for Biology*, July. https://doi.org/10.1101/2023.06.27.546656.
Dorkenwald, Sven, Casey M. Schneider-Mizell, Derrick Brittain, Akhilesh Halageri, Chris Jordan, Nico Kemnitz, Manual A. Castro, et al. 2023. "CAVE: Connectome Annotation Versioning Engine." *bioRxiv : The Preprint Server for Biology*, July. https://doi.org/10.1101/2023.07.26.550598.
Eckstein, Nils, Alexander Shakeel Bates, Andrew Champion, Michelle Du, Yijie Yin, Philipp Schlegel, Alicia Kun-Yang Lu, et al. 2023. "Neurotransmitter Classification from Electron Microscopy Images at Synaptic Sites in Drosophila Melanogaster." *bioRxiv*. https://doi.org/10.1101/2020.06.12.148775.
Fischbach, K-F, and A. P. M. Dittrich. 1989. "The Optic Lobe of Drosophila Melanogaster. I. A Golgi Analysis of Wild-Type Structure." *Cell and Tissue Research* 258 (3). https://doi.org/10.1007/bf00218858.
Fukushima, K. 1980. "Neocognitron: A Self Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position." *Biological Cybernetics* 36 (4): 193–202.
Gao, Shuying, Shin-Ya Takemura, Chun-Yuan Ting, Songling Huang, Zhiyuan Lu, Haojiang Luan, Jens Rister, et al. 2008. "The Neural Substrate of Spectral Preference in Drosophila." *Neuron* 60 (2): 328–42.
Gouwens, Nathan W., Staci A. Sorensen, Fahimeh Baftizadeh, Agata Budzillo, Brian R. Lee, Tim Jarsky, Lauren Alfiler, et al. 2020. "Integrated Morphoelectric and Transcriptomic Classification of Cortical GABAergic Cells." *Cell* 183 (4): 935–53.e19.
Grueber, W. B., and A. Sagasti. 2010. "Self-Avoidance and Tiling: Mechanisms of Dendrite and Axon Spacing." *Cold Spring Harbor*. https://cshperspectives.cshlp.org/content/2/9/a001750.short.
Hardcastle, Ben J., Jaison J. Omoto, Pratyush Kandimalla, Bao-Chau M. Nguyen, Mehmet F. Keleş, Natalie K. Boyd, Volker Hartenstein, and Mark A. Frye. 2021. "A Visual Pathway for Skylight Polarization Processing in Drosophila." *eLife* 10 (March). https://doi.org/10.7554/eLife.63225.
Hasegawa, Eri, Yusuke Kitada, Masako Kaido, Rie Takayama, Takeshi Awasaki, Tetsuya Tabata, and

Makoto Sato. 2011. “Concentric Zones, Cell Migration and Neuronal Circuits in the Drosophila Visual Center.” *Development* 138 (5): 983–93.
Hindmarsh Sten, Tom, Rufei Li, Adriane Otopalik, and Vanessa Ruta. 2021. “Sexual Arousal Gates Visual Processing during Drosophila Courtship.” *Nature* 595 (7868): 549–53.
Hubel, D. H., and T. N. Wiesel. 1962. “Receptive Fields, Binocular Interaction and Functional Architecture in the Cat’s Visual Cortex.” *The Journal of Physiology* 160 (January): 106–54.
Hulse, Brad K., Hannah Haberkern, Romain Franconville, Daniel Turner-Evans, Shin-Ya Takemura, Tanya Wolff, Marcella Noorman, et al. 2021. “A Connectome of the Drosophila Central Complex Reveals Network Motifs Suitable for Flexible Navigation and Context-Dependent Action Selection.” *eLife* 10 (October). https://doi.org/10.7554/eLife.66039.
Jia, Hongbo, Nathalie L. Rochefort, Xiaowei Chen, and Arthur Konnerth. 2010. “Dendritic Organization of Sensory Input to Cortical Neurons in Vivo.” *Nature* 464 (7293): 1307–12.
Joesch, Maximilian, Johannes Plett, Alexander Borst, and Dierk F. Reiff. 2008. “Response Properties of Motion-Sensitive Visual Interneurons in the Lobula Plate of Drosophila Melanogaster.” *Current Biology: CB* 18 (5): 368–74.
Karuppudurai, Thangavel, Tzu-Yang Lin, Chun-Yuan Ting, Randall Pursley, Krishna V. Melnattur, Fengqiu Diao, Benjamin H. White, et al. 2014. “A Hard-Wired Glutamatergic Circuit Pools and Relays UV Signals to Mediate Spectral Preference in Drosophila.” *Neuron* 81 (3): 603–15.
Katayama, Natsu, Jessica K. Abbott, Jostein Kjærandsen, Yuma Takahashi, and Erik I. Svensson. 2014. “Sexual Selection on Wing Interference Patterns in Drosophila Melanogaster.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (42): 15144–48.
Katz, Lawrence C., and Justin C. Crowley. 2002. “Development of Cortical Circuits: Lessons from Ocular Dominance Columns.” *Nature Reviews. Neuroscience* 3 (1): 34–42.
Kim, Jinseop S., Matthew J. Greene, Aleksandar Zlateski, Kisuk Lee, Mark Richardson, Srinivas C. Turaga, Michael Purcaro, et al. 2014. “Space-Time Wiring Specificity Supports Direction Selectivity in the Retina.” *Nature* 509 (7500): 331–36.
Kind, Emil, Kit D. Longden, Aljoscha Nern, Arthur Zhao, Gizem Sancer, Miriam A. Flynn, Connor W. Laughland, et al. 2021. “Synaptic Targets of Photoreceptors Specialized to Detect Color and Skylight Polarization in Drosophila.” *eLife* 10 (December). https://doi.org/10.7554/eLife.71858.
Klapoetke, Nathan C., Aljoscha Nern, Martin Y. Peek, Edward M. Rogers, Patrick Breads, Gerald M. Rubin, Michael B. Reiser, and Gwyneth M. Card. 2017. “Ultra-Selective Looming Detection from Radial Motion Opponency.” *Nature* 551 (7679): 237–41.
Klapoetke, Nathan C., Aljoscha Nern, Edward M. Rogers, Gerald M. Rubin, Michael B. Reiser, and Gwyneth M. Card. 2022. “A Functionally Ordered Visual Feature Map in the Drosophila Brain.” *Neuron* 110 (10): 1700–1711.e6.
Konstantinides, Nikolaos, Isabel Holguera, Anthony M. Rossi, Aristides Escobar, Liébaut Dudragne, Yen-Chung Chen, Thinh N. Tran, et al. 2022. “A Complete Temporal Transcription Factor Series in the Fly Visual System.” *Nature* 604 (7905): 316–22.
Kurmangaliyev, Yerbol Z., Juyoun Yoo, Javier Valdes-Aleman, Piero Sanfilippo, and S. Lawrence Zipursky. 2020. “Transcriptional Programs of Circuit Assembly in the Drosophila Visual System.” *Neuron* 108 (6): 1045–57.e6.
Lappalainen, Janne K., Fabian D. Tschopp, Sridhama Prakhya, Mason McGill, Aljoscha Nern, Kazunori Shinomiya, Shin-Ya Takemura, Eyal Gruntman, Jakob H. Macke, and Srinivas C. Turaga. 2023. “Connectome-Constrained Deep Mechanistic Networks Predict Neural Responses across the Fly Visual System at Single-Neuron Resolution.” *bioRxiv*. https://doi.org/10.1101/2023.03.11.532232.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep Learning.” *Nature* 521 (7553): 436–44.
LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. 1989. “Backpropagation Applied to Handwritten Zip Code Recognition.” *Neural Computation* 1 (4): 541–51.
Marc, Robert E., James R. Anderson, Bryan W. Jones, Crystal L. Sigulinsky, and James S. Lauritzen. 2014. “The AII Amacrine Cell Connectome: A Dense Network Hub.” *Frontiers in Neural Circuits* 8 (September): 104.
Mauss, Alex S., Katarina Pankova, Alexander Arenz, Aljoscha Nern, Gerald M. Rubin, and Alexander

Borst. 2015. “Neural Circuit to Integrate Opposing Motions in the Visual Field.” *Cell* 162 (2): 351–62.
Meinertzhagen, I. A., and S. D. O’Neil. 1991. “Synaptic Organization of Columnar Elements in the Lamina of the Wild Type in Drosophila Melanogaster.” *The Journal of Comparative Neurology* 305 (2): 232–63.
Menon, Kaushiki P., Vivek Kulkarni, Shin-Ya Takemura, Michael Anaya, and Kai Zinn. 2019. “Interactions between Dpr11 and DIP-γ Control Selection of Amacrine Neurons in Drosophila Color Vision Circuits.” *eLife* 8 (November). https://doi.org/10.7554/eLife.48935.
MICrONS Consortium, J. Alexander Bae, Mahaly Baptiste, Agnes L. Bodor, Derrick Brittain, Joann Buchanan, Daniel J. Bumbarger, et al. 2021. “Functional Connectomics Spanning Multiple Areas of Mouse Visual Cortex.” *bioRxiv*. https://doi.org/10.1101/2021.07.28.454025.
Morante, Javier, and Claude Desplan. 2008. “The Color-Vision Circuit in the Medulla of Drosophila.” *Current Biology: CB* 18 (8): 553–65.
Nern, A., B. D. Pfeiffer, and G. M. Rubin. 2015. “Optimized Tools for Multicolor Stochastic Labeling Reveal Diverse Stereotyped Cell Arrangements in the Fly Visual System.” *Proceedings of the*. https://www.pnas.org/content/112/22/E2967.short.
Otsuna, Hideo, and Kei Ito. 2006. “Systematic Analysis of the Visual Projection Neurons of Drosophila Melanogaster. I. Lobula-Specific Pathways.” *The Journal of Comparative Neurology* 497 (6): 928–58.
Özel, Mehmet Neset, Félix Simon, Shadi Jafari, Isabel Holguera, Yen-Chung Chen, Najate Benhra, Rana Naja El-Danaf, et al. 2021. “Neuronal Diversity and Convergence in a Visual System Developmental Atlas.” *Nature* 589 (7840): 88–95.
Philipsborn, Andreas von, and Thomas Labhart. 1990. “A Behavioural Study of Polarization Vision in the Fly, Musca Domestica.” *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 167 (6): 737–43.
Reid, R. C., and J. M. Alonso. 1995. “Specificity of Monosynaptic Connections from Thalamus to Visual Cortex.” *Nature* 378 (6554): 281–84.
Ribeiro, Inês M. A., Michael Drews, Armin Bahl, Christian Machacek, Alexander Borst, and Barry J. Dickson. 2018. “Visual Projection Neurons Mediating Directed Courtship in Drosophila.” *Cell* 174 (3): 607–21.e18.
Sancer, Gizem, Emil Kind, Haritz Plazaola-Sasieta, Jana Balke, Tuyen Pham, Amr Hasan, Lucas O. Münch, Maximilien Courgeon, Thomas F. Mathejczyk, and Mathias F. Wernet. 2019. “Modality-Specific Circuits for Skylight Orientation in the Fly Visual System.” *Current Biology: CB* 29 (17): 2812–25.e4.
Scala, Federico, Dmitry Kobak, Matteo Bernabucci, Yves Bernaerts, Cathryn René Cadwell, Jesus Ramon Castro, Leonard Hartmanis, et al. 2020. “Phenotypic Variation of Transcriptomic Cell Types in Mouse Motor Cortex.” *Nature*, November. https://doi.org/10.1038/s41586-020-2907-3.
Scheffer, Louis K., C. Shan Xu, Michal Januszewski, Zhiyuan Lu, Shin-Ya Takemura, Kenneth J. Hayworth, Gary B. Huang, et al. 2020. “A Connectome and Analysis of the Adult Drosophila Central Brain.” *eLife* 9 (September). https://doi.org/10.7554/eLife.57443.
Schlegel, Philipp, Yijie Yin, Alexander Shakeel Bates, Sven Dorkenwald, Katharina Eichler, Paul Brooks, Daniel S. Han, et al. 2023. “Whole-Brain Annotation and Multi-Connectome Cell Typing Quantifies Circuit Stereotypy in Drosophila.” *bioRxiv : The Preprint Server for Biology*, July. https://doi.org/10.1101/2023.06.27.546055.
Schnaitmann, Christopher, Väinö Haikala, Eva Abraham, Vitus Oberhauser, Thomas Thestrup, Oliver Griesbeck, and Dierk F. Reiff. 2018. “Color Processing in the Early Visual System of Drosophila.” *Cell* 172 (1-2): 318–30.e18.
Schneider-Mizell, Casey M., Agnes Bodor, Derrick Brittain, Joann Buchanan, Daniel J. Bumbarger, Leila Elabbady, Daniel Kapner, et al. 2023. “Cell-Type-Specific Inhibitory Circuitry from a Connectomic Census of Mouse Visual Cortex.” *bioRxiv : The Preprint Server for Biology*, February. https://doi.org/10.1101/2023.01.23.525290.
Scott, Ethan K., John E. Reuter, and Liqun Luo. 2003. “Dendritic Development of Drosophilahigh Order Visual System Neurons Is Independent of Sensory Experience.” *BMC Neuroscience* 4 (1): 14.
Seung, H. Sebastian. 2009. “Reading the Book of Memory: Sparse Sampling versus Dense Mapping of

Connectomes.” *Neuron* 62 (1): 17–29.
———. 2012. *Connectome: How the Brain’s Wiring Makes Us Who We Are*. HMH.
———. 2023. “Insights into Vision from Interpretation of a Neuronal Wiring Diagram.” *bioRxiv*. https://doi.org/10.1101/2023.11.15.567126.
Seung, H. S., and R. Yuste. 2010. “Neural Networks In Principles of Neural Science, Kandel ER and Jessel TJ, Eds.” New York: Mc Graw-Hill.
Shinomiya, Kazunori, Gary Huang, Zhiyuan Lu, Toufiq Parag, C. Shan Xu, Roxanne Aniceto, Namra Ansari, et al. 2019. “Comparisons between the ON- and OFF-Edge Motion Pathways in the Drosophila Brain.” *eLife* 8 (January). https://doi.org/10.7554/eLife.40025.
Shinomiya, Kazunori, Aljoscha Nern, Ian A. Meinertzhagen, Stephen M. Plaza, and Michael B. Reiser. 2022. “Neuronal Circuits Integrating Visual Motion Information in Drosophila Melanogaster.” *Current Biology: CB* 32 (16): 3529–44.e2.
Takemura, Shin-Ya, Arjun Bharioke, Zhiyuan Lu, Aljoscha Nern, Shiv Vitaladevuni, Patricia K. Rivlin, William T. Katz, et al. 2013. “A Visual Motion Detection Circuit Suggested by Drosophila Connectomics.” *Nature* 500 (7461): 175–81.
Takemura, Shin-Ya, C. Shan Xu, Zhiyuan Lu, Patricia K. Rivlin, Toufiq Parag, Donald J. Olbris, Stephen Plaza, et al. 2015. “Synaptic Circuits and Their Variations within Different Columns in the Visual System of Drosophila.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (44): 13711–16.
Takemura, S., A. Nern, D. B. Chklovskii, and L. K. Scheffer. 2017. “The Comprehensive Connectome of a Neural Substrate for ‘ON’motion Detection in Drosophila.” *eLife*. https://cdn.elifesciences.org/articles/24394/elife-24394-v2.pdf.
Wei, Huayi, Ha Young Kyung, Priscilla J. Kim, and Claude Desplan. 2020. “The Diversity of Lobula Plate Tangential Cells (LPTCs) in the Drosophila Motion Vision System.” *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 206 (2): 139–48.
White, J. G., E. Southgate, J. N. Thomson, and S. Brenner. 1986. “The Structure of the Nervous System of the Nematode Caenorhabditis Elegans.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 314 (1165): 1–340.
Wilson, Daniel E., David E. Whitney, Benjamin Scholl, and David Fitzpatrick. 2016. “Orientation Selectivity and the Functional Clustering of Synaptic Inputs in Primary Visual Cortex.” *Nature Neuroscience* 19 (8): 1003–9.
Wolff, Tanya, Nirmala A. Iyer, and Gerald M. Rubin. 2015. “Neuroarchitecture and Neuroanatomy of the Drosophila Central Complex: A GAL4-Based Dissection of Protocerebral Bridge Neurons and Circuits.” *The Journal of Comparative Neurology* 523 (7): 997–1037.
Wu, Ming, Aljoscha Nern, W. Ryan Williamson, Mai M. Morimoto, Michael B. Reiser, Gwyneth M. Card, and Gerald M. Rubin. 2016. “Visual Projection Neurons in the Drosophila Lobula Link Feature Detection to Distinct Behavioral Programs.” *eLife* 5 (December). https://doi.org/10.7554/eLife.21022.
Yoo, Juyoun, Mark Dombrovski, Parmis Mirshahidi, Aljoscha Nern, Samuel A. LoCascio, S. Lawrence Zipursky, and Yerbol Z. Kurmangaliyev. 2023. “Brain Wiring Determinants Uncovered by Integrating Connectomes and Transcriptomes.” *Current Biology: CB* 33 (18): 3998–4005.e6.
Yuste, Rafael, Michael Hawrylycz, Nadia Aalling, Argel Aguilar-Valles, Detlev Arendt, Ruben Armañanzas, Giorgio A. Ascoli, et al. 2020. “A Community-Based Transcriptomics Classification and Nomenclature of Neocortical Cell Types.” *Nature Neuroscience* 23 (12): 1456–68.
Zeng, Hongkui, and Joshua R. Sanes. 2017. “Neuronal Cell-Type Classification: Challenges, Opportunities and the Path Forward.” *Nature Reviews. Neuroscience* 18 (9): 530–46.
Zheng, Zhihao, J. Scott Lauritzen, Eric Perlman, Camenzind G. Robinson, Matthew Nichols, Daniel Milkie, Omar Torrens, et al. 2018. “A Complete Electron Microscopy Volume of the Brain of Adult Drosophila Melanogaster.” *Cell* 174 (3): 730–43.e22.