Transcriptional complexity in the insect central complex: single nuclei RNA sequencing of adult brain neurons derived from type 2 neuroblasts

Derek Epiney[#], Gonzalo N Morales Chaya[#], Noah R Dillon[#], Sen-Lin Lai[*], and Chris Q. Doe[*]

Institute of Neuroscience, Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403

Neuroblast, adult brain neurons, central complex, type 2 lineage, glia, homeodomain transcription factors, single nuclei RNA-sequencing, transcriptional profiling

# equal contribution
* corresponding authors: *cdoe@uoregon.edu*, *slai@uoregon.edu*

Abstract

In both invertebrates such as *Drosophila* and vertebrates such as mouse or human, the brain contains the most diverse population of cell types of any tissue. It is generally accepted that transcriptional diversity is an early step in generating neuronal and glial diversity, followed by the establishment of a unique gene expression profile that determines neuron morphology, connectivity, and function. In *Drosophila*, there are two types of neural stem cells, called Type 1 (T1) neuroblasts and Type 2 neuroblasts (T2). In contrast to T1 neuroblasts, T2 neuroblasts generate Intermediate Neural Progenitors (INPs) that expand the number and diversity of neurons. The diversity of T2 neuroblast-derived neurons contributes a large portion of the central complex (CX), a conserved brain region that plays a role in sensorimotor integration, including celestial navigation. Recent work has revealed much of the connectome of the CX, but how this connectome is assembled remains unclear. Mapping the transcriptional diversity of T2 neuroblast-derived neurons, including those projecting to the CX, is a necessary step in linking transcriptional profile to the assembly of the connectome. Here we use single nuclei RNA sequencing of T2 neuroblast-derived adult neurons to identify over 150 distinct cell clusters. We map neurotransmitter and neuropeptide expression and identify unique transcription factor combinatorial codes for each cluster (presumptive neuron subtype). This is a necessary step that directs functional studies to determine whether each transcription factor combinatorial code specifies a distinct neuron type within the central complex. We map several well-characterized columnar neuron subtypes to distinct clusters, and two neuronal classes map to a single cluster. Our data support the hypothesis that each cluster represents a one or a few closely related neuron classes.

Introduction

In all organisms, the brain has arguably the most complex cellular diversity, from human (Siletti et al., 2023) to *Drosophila* (Davie et al., 2018). Neuronal diversity is essential for the proper assembly and function of the adult brain, yet the "parts list" of different neuronal and glial cell types remains incomplete. In *Drosophila*, the laterally-positioned optic lobes have been well characterized for transcriptionally distinct neurons (Konstantinides et al., 2018; Konstantinides et al., 2022; Özel et al., 2022), but less is known about the central brain, with a whole brain transcriptome (Davie et al., 2018; Li et al., 2022) and a central brain transcriptome reported (Croset et al., 2018).

The central brain contains many diverse neurons that generate functionally important neuropils, such as the mushroom body (MB) used for learning and memory (Sgammeglia and Sprecher, 2022), or the central complex (CX) used for celestial navigation and sensory-motor integration (Fisher, 2022), among other behaviors. The central brain neurons are all generated from neural stem cells, called neuroblasts (NBs). There are two types of NBs that generate central brain neurons. Type 1 (T1) NBs undergo asymmetric divisions to self-renew the NB and generate a ganglion mother cells (GMCs), which produces two post-mitotic neurons (Pollington et al., 2023; Yu et al., 2010; Yu et al., 2013). There are ~100 T1 NBs in each central brain lobe that each generate 20-100 neurons (Ito et al., 2013; Yu et al., 2013). An example of a T1 neurons are the Kenyon cells of the MB (Ito et al., 1997). The second type of NBs are the type 2 (T2) NBs, which have a more complex division pattern (Bello et al., 2008; Boone and Doe, 2008; Bowman et al., 2008). T2 NBs undergo asymmetric division to self-renew and generate an intermediate neural progenitor (INP); each INP undergoes 4-6 divisions to self-renew and generate a GMC and its two neuronal progeny (Bello et al., 2008; Boone and Doe, 2008; Bowman et al., 2008). Importantly, compared to T1 lineages, T2 NBs generate more neurons per clone (~500) (Yu et al., 2013), have highly complex populations of neurons based on morphology and connectivity (Hulse et al., 2020; Wang et al., 2014; Yang et al., 2013; Yu et al., 2013), and their progeny comprise a large portion of the columnar neurons in the adult CX (Boyan and Williams, 2011; Kandimalla et al., 2023).

Although single cell RNA-seq (scRNA-seq) has been done on adult central brain neurons (Croset et al., 2018), the transcriptomic profile of T1 versus T2 neuronal and glial progeny has not yet been characterized. Furthermore, although the T2 lineages generate diverse neuronal progeny (Hulse et al., 2020; Wang et al., 2014; Yang et al., 2013; Yu et al., 2013), focused transcriptional profiling of T2 NB-derived progeny has not been reported. These data are required for linking neuronal transcriptomes with connectomes in order to gain a molecular understanding of CX development, connectivity, and behavior.

Results

Major classes of neurons are derived from both T1 and T2 neuroblasts

The adult central brain (i.e. brain without optic lobes) is composed of neuronal and glial progeny of both type 1 (T1) and type 2 (T2) lineages. Here we begin by determining the transcriptomes of all central brain neurons derived from T1 and T2 lineages. We dissociated 7-day-old adult fly central brains with optic lobes removed using methods previously described (Li et al., 2022; McLaughlin et al., 2021). The cDNA libraries were then prepared from dissociated nuclei with split-pool-based barcoding for single-nuclei transcription profiling (Rosenberg et al., 2018). We sequenced a total of 30,699 nuclei at 589 median genes per nuclei. The flies were of a genotype that gave permanent lineage tracing of T2 NB-derived progeny in the adult (Figure 1A, A'). In brief, *worniu-gal4,asense-gal80* drives expression of FLP recombinase specifically in T2 NBs to flip out the stop cassette allowing continuous expression of Gal4 under the ubiquitous *actin5C* enhancer. After removal of the stop cassette in T2 NBs, the *actin5c-gal4* can continuously drive expression of the reporter genes (RFP or GFP) and FLP recombinase (FLP) in the T2 neuronal and glial progeny in adult brains (Figure 1A). We assigned 3,125 nuclei which have expression of RFP, GFP, or FLP as being derived from the T2 lineage, while 27,574 triple-negative nuclei as being derived from the T1 lineage. Filtering, integration, and clustering (resolution = 12) were performed by Seurat (Stuart et al., 2019), and we used uniform manifold approximation and projection (UMAP)-based dimension reduction to visualize 152 clusters (Figure 1B,C). The most enriched or least expressed genes (marker genes) for each cluster are shown in Supp. Table 1.

We used known markers (Croset et al., 2018; Davie et al., 2018) to identify distinct cell types in the central brain, including glia, mushroom body neurons, olfactory projection neurons, clock neurons, Poxn+ neurons, serotonergic neurons, dopaminergic neurons, octopaminergic neurons, corazonergic neurons, and hemocytes (Figure 1B). We also found a cluster that represents optic lobe neurons, probably due to contamination during dissection. We observed that 11 clusters were enriched with T2 lineage (>50% of nuclei in the cluster were assigned as T2 lineage) and 141 clusters are enriched with progeny derived from T1 NBs (Figure 1C, Supp. Figure 1). However, we didn't observe any cluster exclusively containing progeny from T1 or T2 lineages. We conclude that both T1 and T2 generate transcriptionally similar neurons and glia, despite their different developmental origins. Alternatively, deeper sequencing may resolve a single cluster into two clusters each derived from T1 or T2 lineages.

We next search the marker genes for the T2 enriched clusters, and we found strong expression of *AstA* (*Allatostatin A*), *SerT* (*Serotonin transporter*), *Tk* (*Tachykinin*), and *Vmat* (*Vesicular monoamine transporter*). These enriched marker genes suggest that *AstA*+ neurons, serotonergic neurons, and *TK*+ neurons may be primarily derived from T2 lineages (Figure 1D).

Major classes of glia are derived from both T1 and T2 neuroblasts

We explored the central brain atlas for glial cell types and their gene expression. We sub-clustered from the whole atlas 3409 glial nuclei from clusters that expressed the pan glial marker *repo* (Figure 1E) (Campbell et al., 1994; Xiong et al., 1994). All clusters contained a

mix of T1 and T2 glial progeny (Figure 1E'). We identified 5 known glial cell types (astrocytes, cortex, ensheathing, and the two surface glial subtypes: perineurial and subperineurial) based on canonical glial makers (Figure 1F; Table 1). We identified a glial cluster that expressed genes associated with extracellular matrix, *Vkg* and *Col4a1* (Figure 1F), which have previously been identified as pan surface glial markers (DeSalvo et al., 2014; Hindle and Bainton, 2014). Interestingly, the two surface glial subtypes, perineurial and subperineurial clusters, do not express these markers at cluster defining levels, and conversely the surface glia cluster 5 does not express perineurial or subperineurial specific makers (Figure 1F; see Discussion). Differential gene expression analysis between Type 1 and Type 2 glial progeny did not show differences across any cell types (Table 1, Supp. Table 2). Similar classes of glia were identified following generation of a T2-derived atlas (Supp. Figure 2). We conclude that the adult central brain contains known glial cell types with no differences in gene expression between T1 and T2 derived glia.

Neuronal classes derived from T1 or T2 neuroblasts

We next explored the diversity of T1 and T2 neurons. We identified T1 neurons by excluding *FLP+/GFP+/RFP+* T2 cells and *repo+* glial clusters from the dataset of the dissociated brains (Figure 1C). We then reran the RunUMAP function (1:50 dimension reduction) to generate a UMAP atlas containing 22,807 T1 neurons that form 114 clusters (with resolution = 12). Marker genes of each cluster are shown in Supp. Table 3. We identified the neurons that are known to be generated by T1 NBs, including MB neurons and olfactory projection neurons (Croset et al., 2018; Davie et al., 2018). We also identified other T1 neurons, including clock neurons, Poxn+ neurons, and neurons that release dopamine, serotonin, octopamine/tyramine, and other neuropeptides (Figure 2A) (Croset et al., 2018; Davie et al., 2018).

We next gathered a list of genes that represented the 10 most significantly enriched genes from each cluster and calculated the scaled averaged expression of each gene from each cluster (Figure 2B). We found that each cluster strongly expressed a unique combination of genes.

In our whole brain atlas described above, T2 neurons represent a minor contribution to the atlas due to their low nuclei numbers relative to the T1 neurons. To generate a comprehensive T2 atlas, we needed to increase the number of nuclei for RNAseq. We used fluorescence-activated cell sorting (FACs) to select the nuclei that are labeled by T2-specific permanent lineage tracing (Figure 1A). We then selected the nuclei *in silico* that showed expression of three UAS-transgenes (FLP, GFP, or RFP, see above) to eliminate contamination of T1 nuclei during FACs. In total, we sequenced 26,692 T2 nuclei. We also included 3,125 UAS-transgene labeled nuclei from the dissociated central brain without sorting (see above) and used the integration function in Seurat to generate a neuronal and glial atlas containing 29,817 nuclei that formed 175 clusters (with resolution = 12). T2 NBs have been estimated to produced ~5000 neurons/glia in the central brain (Ito et al., 2013; Yang et al., 2013), with 3,000 in the central complex (Hulse et al., 2020), giving our atlas ~6x coverage. We next filtered out the *repo+* glial cell clusters and reran the RunUMAP

function (1:50 dimension reduction) in Seurat with the remaining 26,147 non-glial nuclei to generate the T2 neuron cell atlas with 155 clusters (Figure 2C). Marker genes of each cluster are shown in Supp. Table 4.

We identified clusters that are enriched with *doublesex (dsx)*, and those *dsx*+ neurons are essential to regulate courtship behaviors in both males and females (Sato and Yamamoto, 2023). Interestingly, we failed to identify *dsx*+ cluster in the T1 atlas, suggesting that most *dsx*+ neurons may be T2-derived. We also identified the columnar neurons, P-EN, P-FN and PF-R, that regulate fly steering or sun navigation (see details below). We identified clusters that represent neurons secreting serotonin, dopamine, octopamine/tyramine, or other neuropeptides (Figure 2C). Similar to our previous findings in the whole central brain atlas (Figure 1), several neuropeptide producing neurons, such as Allatostatin C, CCAP, Ms and others formed individual clusters in T2 atlas, but not in T1 atlas, suggesting these neuropeptidergic cells are largely T2-derived. In contrast, *Corazonin*+ neurons, essential for fruit fly systemic physiology (Khan et al., 2021), were identified as a major cluster in the T1 atlas, but not in T2 atlas, suggesting *Corazonin*+ neurons were largely T1-derived. We also found 911 mushroom body neurons, which is composed of 3.5% of total sequence nuclei, lower than those identified in T1-derived atlas (2524 nuclei, 11%). We assumed that those mushroom body neurons in the T2 atlas were likely due to the leaky expression of UAS-transgenes during development (Figure. 2C).

We calculated the scaled average expression of top 10 significantly enriched genes gathered from each T2 cluster and found each T2 cluster also strongly expressed a unique suit of genes (Figure 2D). When we compared the lists of significantly enriched genes from T1 and T2 atlas, T1 neurons strongly express 181 unique genes (total 440 genes, ~41%) while T2 strongly express 271 unique genes (total 532 genes, ~51%). These unique genes could be marker genes that may distinguish T1 and T2 neurons. For the remainder of the paper we focus solely on the neuronal diversity of T2 neurons.

Expression of fast-acting neurotransmitters in T2-derived adult central brain neurons

An important aspect of neuronal identity and function are fast-acting neurotransmitter expression. Fast-acting neurotransmitters are a major feature that contributes to neuronal function (Luo, 2020). We determined the neuronal populations that expressed three fast-acting neurotransmitters: glutamatergic, cholinergic and GABAergic (Figure 3A-A''). There are reports of fast-acting neurotransmitters expressed in glia in mammals (see Discussion), and we too observed glial clusters with sparse levels of fast-acting neurotransmitters (Supp. Figure 3). In neurons, we found that cholinergic neurons were most abundant (27%), followed by glutamatergic neurons (15%), and then GABAergic neurons (10%), with small percentages co-expressing 2 or 3 neurotransmitters, and 40% were triple-negative with a log normalized expression <2 (Figure 3B); this is similar to the ratios in the adult midbrain (Croset et al., 2018) and optic lobes (Konstantinides et al., 2022). The nuclei co-expressing two or more fast-acting neurotransmitters may reflect an authentic feature of these neurons, as see in other systems (see Discussion). Alternatively, our double- and triple-

positive neurons may be false positives, respectively (see Discussion). Overall, our data supports Dale's principle of “one neuron, one transmitter” (Dale, 1935).

Expression of neuropeptides and receptors in T2-derived adult central brain neurons

Neuropeptides are often used as markers to distinguish different neuronal populations. We examined whether individual neuropeptide-encoding genes were exclusively expressed within single clusters, representing single neuronal subtypes. Among the 49 neuropeptides examined, we identified 14 that exhibited significant expression in fewer than four clusters, five that displayed a broad and uniform expression pattern across multiple clusters, and 30 that did not show significant expression (Figure 4). We pooled neuropeptides showing any expression and assessed their potential preference for co-expression with specific fast-acting neurotransmitters by comparing them in a dot-plot. For simplicity, we focused on clusters that demonstrated significant expression in any of these selected neuropeptides. Although we observed no significant global correlation, we found some overlap between AstA, NPF, Mip, FMRFa, and AstC in cholinergic neurons, and Proc in GABAergic neurons (Figure 4A).

These data raise several questions. What is the relationship between neuropeptide expression and cluster identity? Is neuropeptide gene expression restricted to one or a few clusters? Can two clusters express the same neuropeptide? Can two neuropeptides be co-expressed by the same cluster? Our data show that both of these patterns exist: in some cases, one cluster expresses two or more neuropeptide genes (e.g. cluster 35 expresses NPF and AstA), whereas we also observe one neuropeptide expressed in multiple clusters (e.g. Ms) (Figure 4A).

To determine the relationship between cluster identity, neuropeptide expression and neuronal morphology required a method for specifically labelling of T2 peptidergic neurons (antibody staining could be contaminated with T1 neurons). We developed a genetic tool to selectively visualize neuropeptide expression exclusively in neurons originating from T2 NB lineages. This technique involves driving FLP recombinase (FLP) under the control of the T2-specific driver *wor-GAL4,ase-GAL80*. The T2-specific FLP catalyzes the excision of a stop codon in the *lexAop-FRT-stop-FRT-myr::gfp* transgene, allowing existing neuropeptide-2A-LexA transgenes (Deng et al., 2019) to drive GFP expression specifically in neuropeptide-positive neurons derived from T2 NB lineages (Figure 4B). This allowed us to elucidate the morphology of peptidergic neurons expressed in a limited number of clusters.

We first assayed neuronal morphology of Ms+ neurons in clusters 127 and 140. We found that Ms-2A-LexA-expressing neurons project to two regions of the central complex: the entire ellipsoid body and the dorsal fan-shaped body (Figure 4C-C'', Supp. Movie 1). Currently we can't tell whether neurons in clusters 127 and 140 project to different CX neuropils, or whether neurons from both clusters share projections into both neuropils.

We next assayed neurons in single clusters that express two neuropeptides. Cluster 35 expressed two neuropeptides, AstA and NPF (Figure 4A). This raised the question: does cluster 35 contain two cell types with similar gene expression, does cluster 35 contain a single cell type that co-expressed both neuropeptides? We used NPF-2A-lexA and AstA-2A-lexA to map the morphology of each class of neurons and found that they generated indistinguishable numbers and morphology: ~20 neurons projecting to a lateral domain of the ellipsoid body and the same

two layers of the fan-shaped body (Figure 4D-E; Supp. Movies 2-3). We conclude that some neurons can be transcriptionally similar (one cluster) and express two different neuropeptides. In addition to cluster 35, this would include clusters 62, 75, 93, 114, 116, 140, 145, and 152. This predicts that each pair of neuropeptide genes expressed in a single cluster have cis-regulatory elements that respond to the same inputs.

Our data provide an important stepping-stone to the analysis of upstream transcription factor codes (see below) that drive neuropeptide expression and may specify neuronal identity, as well as for identifying these neurons in the TEM connectome (Hulse et al., 2020) which would reveal pre- and post-synaptic partners.

Transcription factors expressed in cluster-specific combinatorial codes

A leading hypothesis is that neuronal diversity is generated by transcription factor (TF) combinatorial codes that are uniquely expressed in neuronal subtypes, with an emphasis on homeodomain TFs (Reilly et al., 2020; Sagner et al., 2021). Thus, we assayed each class of TFs to determine how many were expressed in cluster-specific (i.e. neuron-specific) patterns. We began by separating TFs into classes based on their DNA-binding domain: zinc finger, homeodomain (HD), basic, high mobility group (HMG), helix-turn-helix (HTH), and unclassified. We binarized the expression of each gene and determined the combination of TFs from each class expressed in each cluster. We found that that the zinc finger TF class had completely unique expression combinations in each cluster (Figure 5A; quantified in 5C), and HDTFs had unique combinatorial codes for ~70% of the clusters (Figure 5B; quantified in 5C). The other classes of TFs showed fewer cluster-specific combinatorial codes (Figure 5C; Supp. Figure 4). We further investigated HDTFs as they have been shown to have unique expression patterns which regulate molecular identity in each of the 302 neurons of *C. elegans* (Reilly et al., 2020). To determine whether clusters with shared HDTF expression patterns also shared overall molecular identity, we compared the top five marker genes for each cluster and found that non-HDTF gene expression was also highly similar (Figure 5B'). Thus, clusters with shared HDTF expression also share overall molecular identity, raising the testable hypothesis that each HDTF combinatorial code drives neuron-specific gene expression.

Identification of columnar neuron transcriptomes

We sought to link well characterized CX neurons to their transcriptome. We chose columnar neurons due to known cell type markers and genetic access to these neuron subtypes. We first explored clusters expressing the two columnar neuron markers, toy and run (Sullivan et al., 2019) (Figure 6A). We identified PF-R neurons by finding co-expression of the known HDTF marker toy (Sullivan et al., 2017) and rho, the enhancer site of the PF-R transgenic line R37G12 (Figure 6A). We next screened for clusters co-expressing *toy* and *rho* within the P-FR-LexA driver, leading to the identification of *Tdc2*, *cut*, and *Lim1* (Figure 6A). We validated expression of Tdc2 and Cut in the PF-R-LexA neurons (Figure 6B); in this way, we assign cluster 27 in the T2 atlas to PF-R neuron identity.

We next identified P-FN and P-EN neuron clusters by finding co-expression of their known TF marker *runt* (Sullivan et al., 2019) with *wg* and *Ggamma30A*, enhancer sites of the P-FN transgenic line R16D01 and P-EN transgenic line R12D09, respectively (Figure 6A). The clusters 75, 81 and 88 expressed *runt* and *Ggamma30A* and thus are putative P-EN clusters (Figure 6A). We noted that wg had very low expression across the clusters; thus clusters 40, 60, or 158 could be putative P-FN neurons (Figure 6A). We screened for markers that distinguished these *runt* clusters and identified that putative P-EN neurons and the two putative P-FN clusters 40 and 60 both expressed both *Lim1* and *Rx* whereas the HDTF *cut* was differentially expressed in only the putative P-EN clusters (Figure 6A). We validated that P-EN and P-FN neurons expressed both Lim1 and Rx (Figure 6C). Next, we validated Cut as specific to P-EN neurons and not P-FN neurons (Figure 6D); thus, we assigned cluster 40 and 60 as P-FN neurons and clusters 75, 81, and 88 as P-EN neurons. However, not every P-EN or P-FN neuron expresses Lim1 or Rx (Figure 6E). The differential expression of transcription factors probably contributes to the morphological or connectivity diversity found within each class of CX neurons (Hulse et al., 2021; Turner-Evans et al., 2020; Wolff and Rubin, 2018; Wolff et al., 2015). We conclude that our T2 atlas can be used to link known neuron subtypes to their transcriptome. This is a necessary step in determining how cluster-specific gene expression regulates neuron-specific functional properties such as morphology, connectivity, and physiology.

Discussion

We used genetics to express RFP in the progeny of T2 NBs in the adult central brain. We excluded optic lobes by manual dissection, and thus our data are complementary to recent scRNA-seq data from the adult optic lobe (Konstantinides et al., 2022). Similarities include neurons expressing predominantly single fast-acting neurotransmitters (discussed below). Both datasets also show the presence of similar classes of glia (astrocytes, ensheathing, perineurial, subperineurial and cortex); the only exception is the chiasm glia (specific to optic lobe (Lago-Baldaia et al., 2023)), and a small population of Viking+ surface glia (specific to whole atlas and not T2 atlas). Both datasets show major classes of neurons including mushroom body neurons, olfactory projection neurons, clock neurons, Poxn+ neurons, serotonergic neurons, dopaminergic neurons, octopaminergic neurons, corazonergic neurons, and hemocytes. We did not detect neurons expressing motor neuron markers. The mushroom body neurons that appear in the T2-derived progeny may be due to off-target expression of our gal4 driver at stages we have not assayed, or more likely, due to ambient RNA released during the dissociation and sorting steps.

Interestingly, we observed fewer distinct neuronal clusters in the T1-derived population (114 clusters) than in the T2-derived population (155 clusters; Figure 2). This could be due in part to the T1-derived progeny containing a large population of transcriptionally similar Kenyon cells of the mushroom body, whereas T2-derived clusters contain fewer cells that are more transcriptionally diverse than T1-derived progeny. It would be interesting to know whether T1 NB lineages are generally less diverse, or whether the Kenyon cells are exceptional in their lack of diversity. Our results are consistent with lineage analysis of T2

NBs, which make distinct neurons across the temporal axis of T2 NBs, plus unique neurons across the INP temporal axis (Doe, 2017; Ito et al., 2013; Sullivan et al., 2019; Wang et al., 2014; Yang et al., 2013).

We detected all the major types of glia (except the optic lobe-specific chiasm glia) (Konstantinides et al., 2018), with the addition of an unexpected cluster defined by the expression of the collagen IV subunit Viking. Otherwise, these glia are most similar to perineurial and subperineurial glia. They may represent a perineurial/subperineurial glia during a transient physiological state.

We observe that most fast-acting neurotransmitters are uniquely expressed in neuronal populations, but we also observe a minority of neurons that express two or three neurotransmitters. This includes excitatory and inhibitory neurons: e.g. 2.9% of all neurons express genes necessary cholinergic and GABAergic neurotransmitters (Figure 3). Co-expression of neurotransmitters has been reported in other systems (Granger et al., 2020; Granger et al., 2023; Lozovaya et al., 2018; Meye et al., 2016; Saunders et al., 2015; Shabel et al., 2014; Spitzer, 2015; Takács et al., 2018) and it may be biologically relevant in the T2 neurons. Alternatively, some or all of the co-expression may be due to translational control (RNA+ protein-) or background from RNAs released during dissociation.

Neuropeptides are profoundly important for a vast array of behaviors (Schoofs et al., 2017), and thus our mapping of neuropeptide expression to distinct neuronal subtypes may facilitate a functional analysis linking neuron identity, neuropeptide expression, and behavior.

Our data identifies zinc finger TFs and HDTFs as forming combinatorial codes specific for distinct cell types, and raises the possibility that these TF combinations determine neuronal functional properties. This is consistent with data from *C. elegans* (Hobert, 2021; Reilly et al., 2020), mammalian spinal cord (Briscoe et al., 2000; Sagner et al., 2021) and *Drosophila* optic lobe (Holguera and Desplan, 2018; Konstantinides et al., 2018) and leg motor neurons (Baek et al., 2013; Enriquez et al., 2015). Interestingly, we found more zinc finger TFs expressed in unique combinatorial codes compared to HDTFs; this may be a caution to focusing too narrowly on HDTF function. We note that our results are correlative and await functional analysis of the TF combinatorial codes found in our dataset. It is also possible that the current sparse TF expression patterns are due, in part, to false negative results due to low read depths. This can be resolved by increasing read depth or by performing functional assays.

Here we linked the PF-R, P-FN, and P-EN neurons, which have known morphology (Sullivan et al., 2019) and connectivity (Hulse et al., 2021), to their transcriptome. Our pipeline began with a single TF known to be expressed in each target neuron subtype, along with genes whose cis-regulatory DNA was used to generate transgenic Gal4/LexA lines labeling a target neuron subtype (Janelia Flylight project). In this manner, we had two markers that could be used to identify candidate clusters that co-express both markers, followed by validation by immunostaining (Figure 6). In the future, our data can be used in two pipelines for linking neuron subtype to transcriptome. First, start with at least one validated marker for a cluster, then search all labeled clusters for candidate TFs that co-express with the validated marker, followed by validation via RNA in situ (or antibodies if

available). Second, we can collect neurons of interest with FACs or antibody selection (Davis et al., 2020), and perform bulk RNA sequencing, followed by looking for clusters enriched for the TFs found in bulk sequencing data, and then validation via RNA in situ (or antibodies if available). Notably, the second pipeline requires no prior markers beyond the Gal4 or LexA line expressed in single neuronal populations. Both pipelines result in a transcriptome that can be used to identify (a) TFs for their role in neuronal specification and/or maintenance, (b) cell surface molecules that may regulate neuronal morphology and/or connectivity, or (c) functionally relevant genes encoding ion channels, neuropeptides/receptors, signaling pathways.

Materials and Methods

Fly genetics

20xUAS-FLPG5.PEST;worniu-gal4,asense-gal80; Act5c(FRT.CD2)gal4 (Syed et al., 2017); *20xUAS-FLPG5.PEST;worniu-gal4,asense-gal80; lexA(stop.FRT)-mCD8:GFP* (this study); *TI{2A-lexA::GAD}AstA[2A-lexA]/TM3,Sb[1]* (RRID:BDSC_84356); *TI{2A-lexA::GAD}Ms[2A-lexA]/TM3,Sb[1]* (RRID:BDSC_84403); *TI{2A-lexA::GAD}NPF[2A-lexA]/TM3,Sb[1]* (RRID:BDSC_84422); *UAS-RedStinger* (RRID:BDSC_8545); *UAS-unc84-*2xGFP (Henry et al., 2012); *GMR12D09-lexA/CyO* (RRID:BDSC_54419); *GMR16D01-lexA* (RRID:BDSC_52503); *GMR37G12-lexA* (RRID:BDSC_52765); *13xLexAop2-IVS-myr::smGdP-V5* (RRID:BDSC_62215).

Single nuclei isolation, library preparation, and sequencing

The T2 nuclei were labeled by crossing *20xUAS-FLPG5.PEST;worniu-gal4,asense-gal80; Act5c(FRT.CD2)gal4* (Syed et al., 2017) to *UAS-RedStinger* (RFP) or *UAS-unc84-2xGFP* (GFP) flies. The adult flies were aged for 1 week in 25 ˚C before dissection. Equal amounts of male and female central brains were dissected (without optic lobes) in room temperature within 1 hour, flash-frozen in liquid nitrogen, and then stored in -80 ˚C freezer. Nuclei dissociation from the frozen dissected brains were performed by following published protocols (McLaughlin et al., 2022). RFP+ or GFP+ nuclei were collected by sorting dissociated nuclei with SONY-SH800 with 70 µm chip. The total cDNA libraries were composed of 3 samples (1) unsorted: whole brain cells without sorting (50%), (2) RFP: *redstinger*+ nuclei (25%), and (3) GFP: *unc84sfGFP*+ nuclei (25%). The RNAs were barcoded with WT kit (Parse Bioscience, Seattle, WA, USA) by following manufacturer's protocol. The library was then split into 8 sublibraries (1, 2, 3, 4, 5, 6, 7, and 8) as instructed. Sublibrary 1, 3, 4, 6 and sublibrary 2, 5, 7, 8 were grouped for sequencing with pair-ends reads 150bp on Illumina Novaseq 6000 in a single lane (University of Oregon's Genomics and Cell Characterization Core Facility).

snRNA-seq analysis

Our bioinformatic analysis was performed using pipeline from Parse Bioscience and the Seurat R package version 4.4.0. Briefly, pipelines from Parse Biosciences were used to perform demultiplexing, alignment, filtering, counting of barcodes and UMIs with an output being a cell-by-genes matrix of counts. We aligned our sequences to a custom reference genome by adding *flpD5* (RRID:Addgene_32132), *redstinger* (RRID:Addgene_46165), and *unc84sfGFP* (RRID:Addgene_46023) (+SV40 tail) sequences and annotations to the *Drosophila* genome release BDGP6.32.109. To further ensure that only high-quality cells were retained, we removed any cells with fewer than 200 genes or more than 2500 genes expressed and more than 5% mitochondrial RNA. snRNA-seq data were integrated in Seurat to generate an integrated dataset, and the downstream analysis was performed with the same parameters as early described (Seroka et al., 2022). We generated the following datasets for transcriptomic analysis: (1) unsorted.integrated: this dataset was created from

integrating two rounds of sequencing results of unsorted group. This dataset was used for analysis in Figure 1. Glial cells were subsetted from this dataset for analysis in Figure 1E and 1F. (2) unsorted.T1.neurons: we removed *repo+* glial clusters from unsorted.integrated, and reran RunUMAP. This dataset was used for analysis in Figure 2A and 2B. (3) T2_all.integrated.neurons: the nuclei from unsorted, RFP, and GFP samples with raw counts of *redstinger*, *unc84sfGFP*, or *flpD5* > 0 were integrated to generate a dataset that includes all cells derived from T2 NBs. We then removed *repo+* glial clusters and reran RunUMAP. This dataset is used for analysis for Figure 2C, 2D, 3, 4, 5, and 6.

Binarization of Transcription factors

Each gene of interest was given a binary score in each cell using the "Binarize" function in the R package clustifyr (Fu et al., 2020). We then averaged each gene's binary score per cluster. We determined a cutoff point to call a gene "on" in a cluster by comparing different parameters to dotplots. We found that an average binary score of 0.3 or greater was the best representation of a gene that is "on" in a given cluster.

Data and code availability

All data will be placed in a public repository (e.g. github or GEO) upon acceptance for publication.

Immunohistochemistry and imaging

Standard methods were used for adult brain staining (Sullivan et al., 2019). Primary antibodies were chicken anti-V5 at 1 µg/mL (Fortis Life Sciences, Waltham, MA), rabbit anti-Lim1 (1:500) (gift from C. Desplan, New York University), guinea pig anti-Rx (1:500) (gift from C. Desplan, New York University), and mouse anti-Cut at 2 µg/mL (2B10, DSHB, Iowa City, IA) . Secondary antibodies were donkey whole IgG antibody conjugated with DyLight405, Alexa Fluor 488, RR-X, or Alexa Fluor 647 (Jackson ImmunoResearch Laboratories Inc., West Grove, PA) and used at 5 µg/mL. The antibody stained brains were mounted in DPX (https://www.janelia.org/project-team/flylight/protocols) on poly-L-Lysine coated coverslip (Corning, Glendale, AZ) and imaged with Zeiss confocal LSM800 with software Zen.

Figure production

We used Imaris (Bitplane, Oxford Instruments plc, Abingdon, UK) for confocal image processing, and Illustrator (Adobe, San Jose, CA) to assemble figures.

Table 1. Validated glial markers

Cell type	Marker	References
All glia	*repo*	(Campbell et al., 1994; Xiong et al., 1994)
Astrocyte	*Tet*	(Frey et al., 2022)
	AANAT1	(Davla et al., 2020)
	mbc	(Lago-Baldaia et al., 2023)
	alrm	(Doherty et al., 2009)
	wun2	(Lago-Baldaia et al., 2023)
	htl	(Shishido et al., 1997)
Cortex	*wrapper*	(Noordermeer et al., 1998)
	apolpp	(Lago-Baldaia et al., 2023)
	Ork1	(Lago-Baldaia et al., 2023)
	ana	(Ebens et al., 1993)
Ensheathing	*List*	(Lago-Baldaia et al., 2023)
	axo	(Lago-Baldaia et al., 2023)
	Eaat2	(Kremer et al., 2017)
	CG9657	(Lago-Baldaia et al., 2023)
Surface glia	*Vkg*	(DeSalvo et al., 2014; Hindle and Bainton, 2014)
Perineurial	*Tret1-1*	(Volkenhoff et al., 2015)
	Indy	(DeSalvo et al., 2014)
	CG4797	(Konstantinides et al., 2018)
	CG6126	(Brunet Avalos et al., 2019)
	CG5080	(Beckervordersandforth et al., 2008)
Subperineurial	*ltl*	(Lago-Baldaia et al., 2023)
	moody	(Bainton et al., 2005)
	CG10702	(Lago-Baldaia et al., 2023)

Ethical Approval and Consent to participate

n/a

Consent for publication

All authors consent for this to be published

Availability of data and materials

All data and materials will be placed in the appropriate public repository (e.g. github or Bloomington stock center) upon acceptance for publication

Authors' contributions

SLL generated Figures 1, 2, 6, SF1, ST1, ST3, and ST4.
NRD generated Figure 1, 6, SF2, and ST2.
GNMC generated Figures 3, 4, SM1, SM2, and SM3.
DE generated Figures 5, 6, and SF3.
All authors contributed to writing the manuscript.
CQD supervised the project.

Acknowledgements

We thank Jason Carriere and Peter Newstein for comments on the manuscript. We thank the Desplan lab (NYU) for reagents. We thank Jason Carriere (University of Oregon

Genomics and Cell Characterization Core Facility) for technical assistance. Flies were obtained from the Bloomington Stock Center.

Competing Interests

None.

Funding

Funder	Grant reference number	Author
Howard Hughes Medical Institute	None	CQD, SLL, NRD, DE
NIH Training Grant	5-T32-HD07348	NRD
NIH	R01-HD27056	NRD, DE, GNMC

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author ORCIDs

Derek Epiney	https://orcid.org/0000-0002-6056-1817
Gonzalo N Morales Chaya	https://orcid.org/0009-0005-4263-1323
Noah R Dillon	https://orcid.org/0000-0002-6708-4208
Sen-Lin Lai	https://orcid.org/0000-0002-7531-283X
Chris Q Doe	https://orcid.org/0000-0001-5980-8029

References

Baek, M., Enriquez, J. and Mann, R. S. (2013). Dual role for Hox genes and Hox co-factors in conferring leg motoneuron survival and identity in Drosophila. *Development* **140**, 2027–38.

Bainton, R. J., Tsai, L. T.-Y., Schwabe, T., DeSalvo, M., Gaul, U. and Heberlein, U. (2005). moody encodes two GPCRs that regulate cocaine behaviors and blood-brain barrier permeability in Drosophila. *Cell* **123**, 145–156.

Beckervordersandforth, R. M., Rickert, C., Altenhein, B. and Technau, G. M. (2008). Subtypes of glial cells in the Drosophila embryonic ventral nerve cord as related to lineage and gene expression. *Mech. Dev.* **125**, 542–557.

Bello, B. C., Izergina, N., Caussinus, E. and Reichert, H. (2008). Amplification of neural stem cell proliferation by intermediate progenitor cells in Drosophila brain development. *Neural Dev* **3**, 5.

Boone, J. Q. and Doe, C. Q. (2008). Identification of Drosophila type II neuroblast lineages containing transit amplifying ganglion mother cells. *Dev Neurobiol* **68**, 1185–95.

Bowman, S. K., Rolland, V., Betschinger, J., Kinsey, K. A., Emery, G. and Knoblich, J. A. (2008). The tumor suppressors Brat and Numb regulate transit-amplifying neuroblast lineages in Drosophila. *Dev Cell* **14**, 535–46.

Boyan, G. and Williams, L. (2011). Embryonic development of the insect central complex: insights from lineages in the grasshopper and Drosophila. *Arthropod Struct. Dev.* **40**, 334–348.

Briscoe, J., Pierani, A., Jessell, T. M. and Ericson, J. (2000). A Homeodomain Protein Code Specifies Progenitor Cell Identity and Neuronal Fate in the Ventral Neural Tube. *Cell* **101**, 435–445.

Brunet Avalos, C., Maier, G. L., Bruggmann, R. and Sprecher, S. G. (2019). Single cell transcriptome atlas of the Drosophila larval brain. *eLife* **8**, e50354.

Campbell, G., Goring, H., Lin, T., Spana, E., Andersson, S., Doe, C. Q. and Tomlinson, A. (1994). RK2, a glial-specific homeodomain protein required for embryonic nerve cord condensation and viability in Drosophila. *Development* **120**, 2957–66.

Croset, V., Treiber, C. D. and Waddell, S. (2018). Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. *eLife* **7**, e34550.

Dale, H. (1935). Pharmacology and Nerve-Endings. *Proc. R. Soc. Med.* **28**, 319–332.

Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C., et al. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell* **174**, 982-998.e20.

Davis, F. P., Nern, A., Picard, S., Reiser, M. B., Rubin, G. M., Eddy, S. R. and Henry, G. L. (2020). A genetic, genomic, and computational resource for exploring neural circuit function. *eLife* **9**, e50901.

Davla, S., Artiushin, G., Li, Y., Chitsaz, D., Li, S., Sehgal, A. and van Meyel, D. J. (2020). AANAT1 functions in astrocytes to regulate sleep homeostasis. *eLife* **9**, e53994.

Deng, B., Li, Q., Liu, X., Cao, Y., Li, B., Qian, Y., Xu, R., Mao, R., Zhou, E., Zhang, W., et al. (2019). Chemoconnectomics: Mapping Chemical Transmission in Drosophila. *Neuron* **101**, 876-893.e4.

DeSalvo, M. K., Hindle, S. J., Rusan, Z. M., Orng, S., Eddison, M., Halliwill, K. and Bainton, R. J. (2014). The Drosophila surface glia transcriptome: evolutionary conserved blood-brain barrier processes. *Front. Neurosci.* **8**, 346.

Doe, C. Q. (2017). Temporal Patterning in the Drosophila CNS. *Annu Rev Cell Dev Biol* **33**, 219–240.

Doherty, J., Logan, M. A., Tasdemir, O. E. and Freeman, M. R. (2009). Ensheathing glia function as phagocytes in the adult Drosophila brain. *J Neurosci* **29**, 4768–81.

Ebens, A. J., Garren, H., Cheyette, B. N. and Zipursky, S. L. (1993). The Drosophila anachronism locus: a glycoprotein secreted by glia inhibits neuroblast proliferation. *Cell* **74**, 15–27.

Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U. and Mann, R. S. (2015). Specification of individual adult motor neuron morphologies by combinatorial transcription factor codes. *Neuron* **86**, 955–70.

Fisher, Y. E. (2022). Flexible navigational computations in the Drosophila central complex. *Curr. Opin. Neurobiol.* **73**, 102514.

Frey, F., Sandakly, J., Ghannam, M., Doueiry, C., Hugosson, F., Berlandi, J., Ismail, J. N., Gayden, T., Hasselblatt, M., Jabado, N., et al. (2022). Drosophila Tet Is Required for Maintaining Glial Homeostasis in Developing and Adult Fly Brains. *eNeuro* **9**, ENEURO.0418-21.2022.

Fu, R., Gillen, A. E., Sheridan, R. M., Tian, C., Daya, M., Hao, Y., Hesselberth, J. R. and Riemondy, K. A. (2020). clustifyr: an R package for automated single-cell RNA sequencing cluster classification. *F1000Research* **9**, 223.

Granger, A. J., Wang, W., Robertson, K., El-Rifai, M., Zanello, A. F., Bistrong, K., Saunders, A., Chow, B. W., Nuñez, V., Turrero García, M., et al. (2020). Cortical ChAT+ neurons co-transmit acetylcholine and GABA in a target- and brain-region-specific manner. *eLife* **9**, e57749.

Granger, A. J., Mao, K., Saulnier, J. L., Hines, M. E. and Sabatini, B. L. (2023). Developmental regulation of GABAergic gene expression in forebrain cholinergic neurons. *Front. Neural Circuits* **17**, 1125071.

Henry, G. L., Davis, F. P., Picard, S. and Eddy, S. R. (2012). Cell type-specific genomics of Drosophila neurons. *Nucleic Acids Res.* **40**, 9691–9704.

Hindle, S. J. and Bainton, R. J. (2014). Barrier mechanisms in the Drosophila blood-brain barrier. *Front. Neurosci.* **8**, 414.

Hobert, O. (2021). Homeobox genes and the specification of neuronal identity. *Nat. Rev. Neurosci.* **22**, 627–636.

Holguera, I. and Desplan, C. (2018). Neuronal specification in space and time. *Science* **362**, 176–180.

Hulse, B. K., Haberkern, H., Franconville, R., Turner-Evans, D. B., Takemura, S., Wolff, T., Noorman, M., Dreher, M., Dan, C., Parekh, R., et al. (2020). A connectome of the Drosophila central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *bioRxiv* 2020.12.08.413955.

Hulse, B. K., Haberkern, H., Franconville, R., Turner-Evans, D., Takemura, S., Wolff, T., Noorman, M., Dreher, M., Dan, C., Parekh, R., et al. (2021). A connectome of the Drosophila central complex reveals network motifs suitable for flexible navigation and context-dependent action selection. *eLife* **10**, e66039.

Ito, K., Awano, W., Suzuki, K., Hiromi, Y. and Yamamoto, D. (1997). The Drosophila mushroom body is a quadruple structure of clonal units each of which contains a virtually identical set of neurones and glial cells. *Development* **124**, 761–71.

Ito, M., Masuda, N., Shinomiya, K., Endo, K. and Ito, K. (2013). Systematic analysis of neural projections reveals clonal composition of the Drosophila brain. *Curr Biol* **23**, 644–55.

Kandimalla, P., Omoto, J. J., Hong, E. J. and Hartenstein, V. (2023). Lineages to circuits: the developmental and evolutionary architecture of information channels into the central complex. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* **209**, 679–720.

Khan, Z., Tondravi, M., Oliver, R. and Vonhoff, F. J. (2021). Drosophila Corazonin Neurons as a Hub for Regulating Growth, Stress Responses, Ethanol-Related Behaviors, Copulation Persistence and Sexually Dimorphic Reward Pathways. *J. Dev. Biol.* **9**, 26.

Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R. and Desplan, C. (2018). Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. *Cell* **174**, 622-635.e13.

Konstantinides, N., Holguera, I., Rossi, A. M., Escobar, A., Dudragne, L., Chen, Y.-C., Tran, T. N., Martínez Jaimes, A. M., Özel, M. N., Simon, F., et al. (2022). A complete temporal transcription factor series in the fly visual system. *Nature* **604**, 316–322.

Kremer, M. C., Jung, C., Batelli, S., Rubin, G. M. and Gaul, U. (2017). The glia of the adult Drosophila nervous system. *Glia* **65**, 606–638.

Lago-Baldaia, I., Cooper, M., Seroka, A., Trivedi, C., Powell, G. T., Wilson, S. W., Ackerman, S. D. and Fernandes, V. M. (2023). A Drosophila glial cell atlas reveals a mismatch between transcriptional and morphological diversity. *PLoS Biol.* **21**, e3002328.

Li, H., Janssens, J., De Waegeneer, M., Kolluru, S. S., Davie, K., Gardeux, V., Saelens, W., David, F. P. A., Brbić, M., Spanier, K., et al. (2022). Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science* **375**, eabk2432.

Lozovaya, N., Eftekhari, S., Cloarec, R., Gouty-Colomer, L. A., Dufour, A., Riffault, B., Billon-Grand, M., Pons-Bennaceur, A., Oumar, N., Burnashev, N., et al. (2018). GABAergic inhibition in dual-transmission cholinergic and GABAergic striatal interneurons is abolished in Parkinson disease. *Nat. Commun.* **9**, 1422.

Luo, L. (2020). *Principles of Neurobiology*. 2nd edition. Boca Raton, FL: Garland Science.

McLaughlin, C. N., Brbić, M., Xie, Q., Li, T., Horns, F., Kolluru, S. S., Kebschull, J. M., Vacek, D., Xie, A., Li, J., et al. (2021). Single-cell transcriptomes of developing and adult olfactory receptor neurons in Drosophila. *eLife* **10**, e63856.

McLaughlin, C. N., Qi, Y., Quake R., S., Luo, L. and Li, H. (2022). Isolation and RNA sequencing of single nuclei from Drosophila tissues: STAR Protocols. *STAR Procotols* **3**, 101417.

Meye, F. J., Soiza-Reilly, M., Smit, T., Diana, M. A., Schwarz, M. K. and Mameli, M. (2016). Shifted pallidal co-release of GABA and glutamate in habenula drives cocaine withdrawal and relapse. *Nat. Neurosci.* **19**, 1019–1024.

Noordermeer, J. N., Kopczynski, C. C., Fetter, R. D., Bland, K. S., Chen, W. Y. and Goodman, C. S. (1998). Wrapper, a novel member of the Ig superfamily, is expressed by midline glia and is required for them to ensheath commissural axons in Drosophila. *Neuron* **21**, 991–1001.

Özel, M. N., Gibbs, C. S., Holguera, I., Soliman, M., Bonneau, R. and Desplan, C. (2022). Coordinated control of neuronal differentiation and wiring by a sustained code of transcription factors. 2022.05.01.490216.

Pollington, H. Q., Seroka, A. Q. and Doe, C. Q. (2023). From temporal patterning to neuronal connectivity in Drosophila type I neuroblast lineages. *Semin. Cell Dev. Biol.* **142**, 4–12.

Reilly, M. B., Cros, C., Varol, E., Yemini, E. and Hobert, O. (2020). Unique homeobox codes delineate all the neuron classes of C. elegans. *Nature* **584**, 595–601.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L. T., Peeler, D. J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182.

Sagner, A., Zhang, I., Watson, T., Lazaro, J., Melchionda, M. and Briscoe, J. (2021). A shared transcriptional code orchestrates temporal patterning of the central nervous system. *PLoS Biol.* **19**, e3001450.

Sato, K. and Yamamoto, D. (2023). Molecular and cellular origins of behavioral sex differences: a tiny little fly tells a lot. *Front. Mol. Neurosci.* **16**, 1284367.

Saunders, A., Granger, A. J. and Sabatini, B. L. (2015). Corelease of acetylcholine and GABA from cholinergic forebrain neurons. *eLife* **4**, e06412.

Schoofs, L., De Loof, A. and Van Hiel, M. B. (2017). Neuropeptides as Regulators of Behavior in Insects. *Annu. Rev. Entomol.* **62**, 35–52.

Seroka, A., Lai, S.-L. and Doe, C. Q. (2022). Transcriptional profiling from whole embryos to single neuroblast lineages in Drosophila. *Dev. Biol.* **489**, 21–33.

Sgammeglia, N. and Sprecher, S. G. (2022). Interplay between metabolic energy regulation and memory pathways in Drosophila. *Trends Neurosci.* **45**, 539–549.

Shabel, S. J., Proulx, C. D., Piriz, J. and Malinow, R. (2014). Mood regulation. GABA/glutamate co-release controls habenula output and is modified by antidepressant treatment. *Science* **345**, 1494–1498.

Shishido, E., Ono, N., Kojima, T. and Saigo, K. (1997). Requirements of DFR1/Heartless, a mesoderm-specific Drosophila FGF-receptor, for the formation of heart, visceral and somatic muscles, and ensheathing of longitudinal axon tracts in CNS. *Dev. Camb. Engl.* **124**, 2119–2128.

Siletti, K., Hodge, R., Mossi Albiach, A., Lee, K. W., Ding, S.-L., Hu, L., Lönnerberg, P., Bakken, T., Casper, T., Clark, M., et al. (2023). Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046.

Spitzer, N. C. (2015). Neurotransmitter Switching? No Surprise. *Neuron* **86**, 1131–1144.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P. and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21.

Sullivan, L. F., Warren, T. L. and Doe, C. Q. (2019). Temporal identity establishes columnar neuron morphology, connectivity, and function in a Drosophila navigation circuit. *eLife* **8**,.

Syed, M. H., Mark, B. and Doe, C. Q. (2017). Steroid hormone induction of temporal gene expression in Drosophila brain neuroblasts generates neuronal and glial diversity. *eLife* **6**, 6.

Takács, V. T., Cserép, C., Schlingloff, D., Pósfai, B., Szőnyi, A., Sos, K. E., Környei, Z., Dénes, Á., Gulyás, A. I., Freund, T. F., et al. (2018). Co-transmission of acetylcholine and GABA regulates hippocampal states. *Nat. Commun.* **9**, 2848.

Turner-Evans, D. B., Jensen, K. T., Ali, S., Paterson, T., Sheridan, A., Ray, R. P., Wolff, T., Lauritzen, J. S., Rubin, G. M., Bock, D. D., et al. (2020). The Neuroanatomical Ultrastructure and Function of a Biological Ring Attractor. *Neuron* **108**, 145-163.e10.

Volkenhoff, A., Weiler, A., Letzel, M., Stehling, M., Klämbt, C. and Schirmeier, S. (2015). Glial Glycolysis Is Essential for Neuronal Survival in Drosophila. *Cell Metab.* **22**, 437–447.

Wang, Y. C., Yang, J. S., Johnston, R., Ren, Q., Lee, Y. J., Luan, H., Brody, T., Odenwald, W. F. and Lee, T. (2014). Drosophila intermediate neural progenitors produce lineage-dependent related series of diverse neurons. *Development* **141**, 253–8.

Wolff, T. and Rubin, G. M. (2018). Neuroarchitecture of the Drosophila central complex: A catalog of nodulus and asymmetrical body neurons and a revision of the protocerebral bridge catalog. *J. Comp. Neurol.* **526**, 2585–2611.

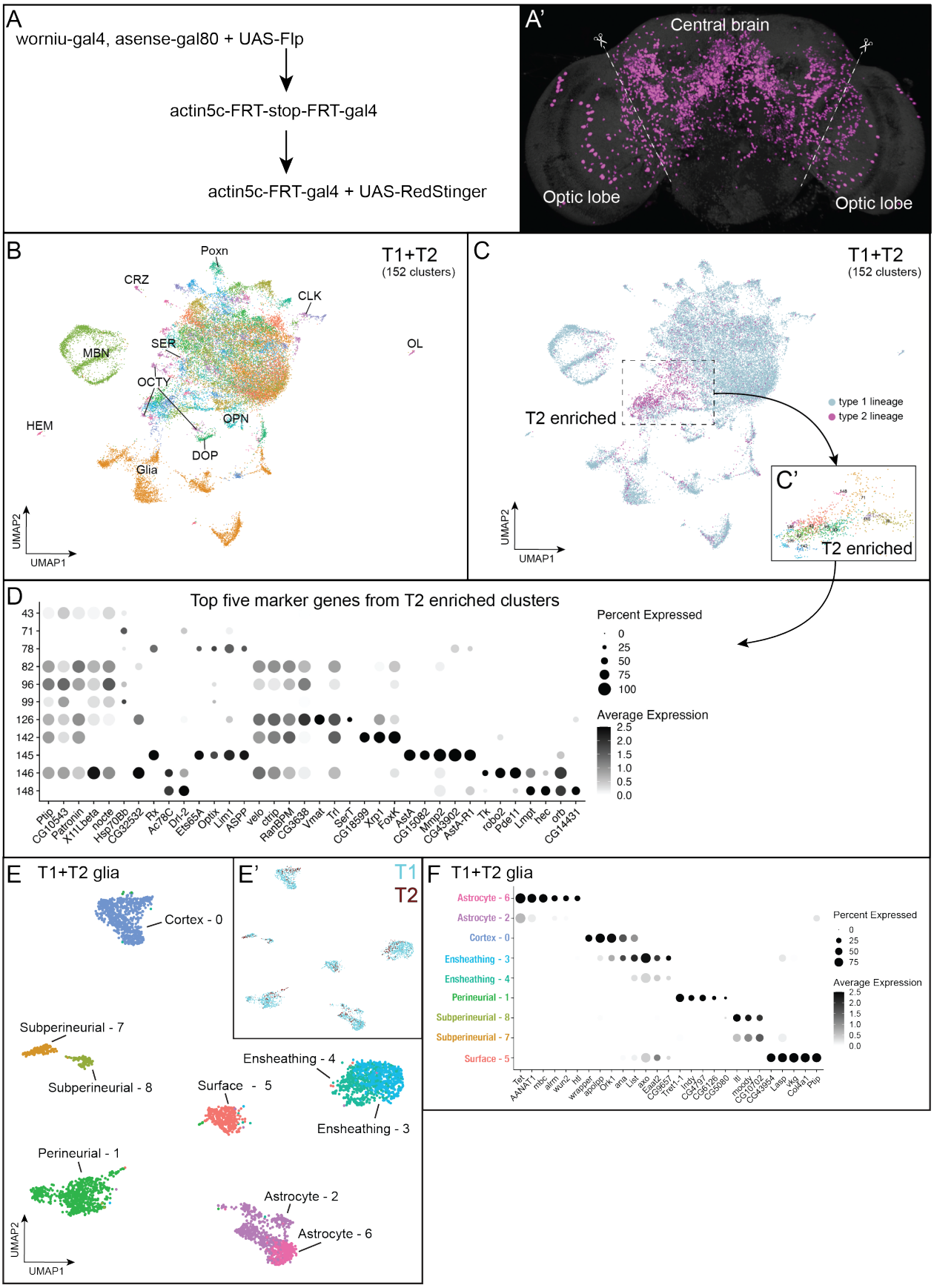
Wolff, T., Iyer, N. A. and Rubin, G. M. (2015). Neuroarchitecture and neuroanatomy of the Drosophila central complex: A GAL4-based dissection of protocerebral bridge neurons and circuits. *J. Comp. Neurol.* **523**, 997–1037.

Xiong, W. C., Okano, H., Patel, N. H., Blendy, J. A. and Montell, C. (1994). repo encodes a glial-specific homeo domain protein required in the Drosophila nervous system. *Genes Dev.* **8**, 981–94.

Yang, J. S., Awasaki, T., Yu, H. H., He, Y., Ding, P., Kao, J. C. and Lee, T. (2013). Diverse neuronal lineages make stereotyped contributions to the Drosophila locomotor control center, the central complex. *J Comp Neurol* **521**, 2645–62.

Yu, H. H., Kao, C. F., He, Y., Ding, P., Kao, J. C. and Lee, T. (2010). A complete developmental sequence of a Drosophila neuronal lineage as revealed by twin-spot MARCM. *PLOS Biol.* **8**,.

Yu, H. H., Awasaki, T., Schroeder, M. D., Long, F., Yang, J. S., He, Y., Ding, P., Kao, J. C., Wu, G. Y., Peng, H., et al. (2013). Clonal development and organization of the adult Drosophila central brain. *Curr Biol* **23**, 633–43.


A
worniu-gal4, asense-gal80 + UAS-Flp
actin5c-FRT-stop-FRT-gal4
actin5c-FRT-gal4 + UAS-RedStinger
A'
Central brain
Optic lobe
Optic lobe
B
T1+T2
(152 clusters)
Poxn
CRZ
CLK
MBN
SER
OL
OCTY
HEM
OPN
DOP
Glia
UMAP2
UMAP1
C
T1+T2
(152 clusters)
T2 enriched
type 1 lineage
type 2 lineage
C'
T2 enriched
UMAP2
UMAP1
D
Top five marker genes from T2 enriched clusters
Percent Expressed
Average Expression
E
T1+T2 glia
Cortex - 0
Subperineurial - 7
Subperineurial - 8
Surface - 5
Ensheathing - 4
Ensheathing - 3
Perineurial - 1
Astrocyte - 2
Astrocyte - 6
UMAP2
UMAP1
E'
T1
T2
F
T1+T2 glia
Astrocyte - 6
Astrocyte - 2
Cortex - 0
Ensheathing - 3
Ensheathing - 4
Perineurial - 1
Subperineurial - 8
Subperineurial - 7
Surface - 5
Percent Expressed
Average Expression

718

Figure 1. Cell atlas of central brain with single nuclei RNAseq.

(A-A') Genetics (A) to label progeny (A') derived from T2 NBs. Dashed lines show the boundary of optic lobes and central brain, and the optic lobes were removed during dissection. **(B)** Central brain atlas labeled with known cell types. Abbreviations: CLK, clock neurons; CRZ, Corazonergic neurons; DOP, dopaminergic neurons; HEM, hemocytes; MBN, mushroom body neurons; OCTY, octopaminergic-tyraminergic neurons; OL, optic lobe; OPN, olfactory projection neurons; SER: serotoninergic neurons. **(C)** Central brain atlas labeled by NB (T1 or T2) lineage. Dash line-outlined box shows the region enriched with the cells derived from T2 NBs, and the identity are shown at the bottom-right box. **(D)** Dot plot of top 5 marker genes of the T2-enriched clusters. **(E)** Atlas of central brain glia labeled with known cell types. **(E')** T1 and T2 derived cells colored cyan and red respectfully. **(F)** Dot plot of known and top marker genes of glial clusters.

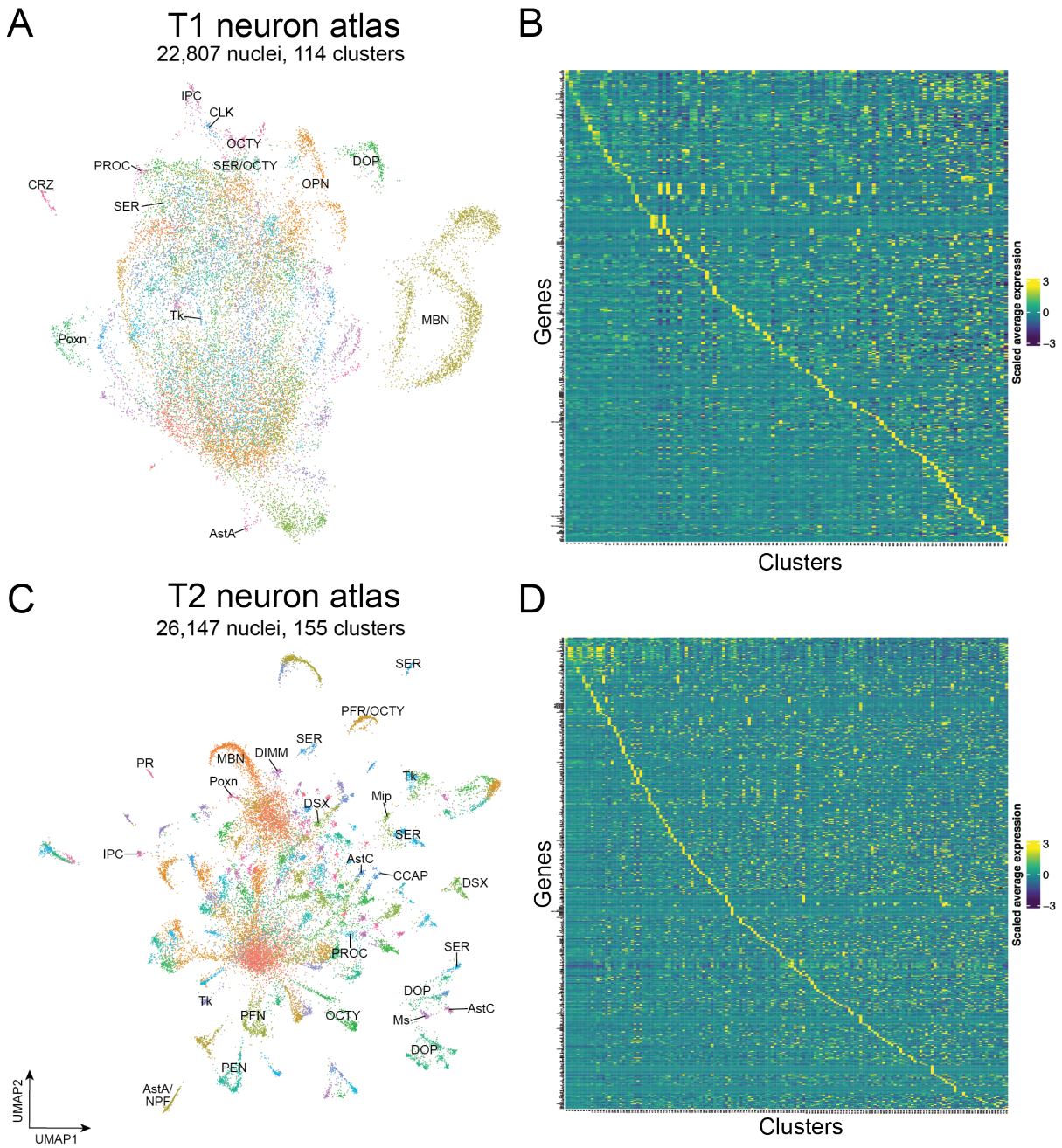


Figure 2. Cell atlas of neurons produced by T1 or T2 NBs.
(A) Cell atlas from T1 NBs. **(B)** Heatmap of scaled average expression of top 10 markers genes from T1 each neuronal cluster. **(C)** Cell atlas from T2 NBs. **(B)** Heatmap of scaled average expression of top 10 markers genes from each T2 neuronal cluster. Annotations: AstA, *Allatostatin A*+ neurons; AstC, *Allatostatin C*+ neurons; CCAP, *Crustacean cardioactive peptide*+ neurons; CLK, clock neurons; CRZ, *Corazonin*+ neurons; DIMM, *dimmed*+ neurons; DOP, dopaminergic neurons; DSX, *doublesex*+ neurons; IPC, insulin-like peptide producing cells; MBN, mushroom body neurons; Mip, *Myoinhibiting peptide precursor*+ neurons, NPF, *neuropeptide F*+ neurons; OCTY, octopaminergic-tyraminergic neurons; OPN olfactory projection neurons; Tk, *Tachykinin*+ neurons; PEN, neurons projects dendrites to protocerebral bridge, and axons to ellipsoid body and

742 nodulus; PFN, neurons project dendrites to protocerebral bridge, and axons to fan-shaped body
743 and nodulus; PFR, neurons project dendrites to protocerebral bridge and fan-shaped body, and
744 axons to round body; Poxn, *Pox neuro*+ neurons; PR, photoreceptors; SER, serotonergic neurons.
745

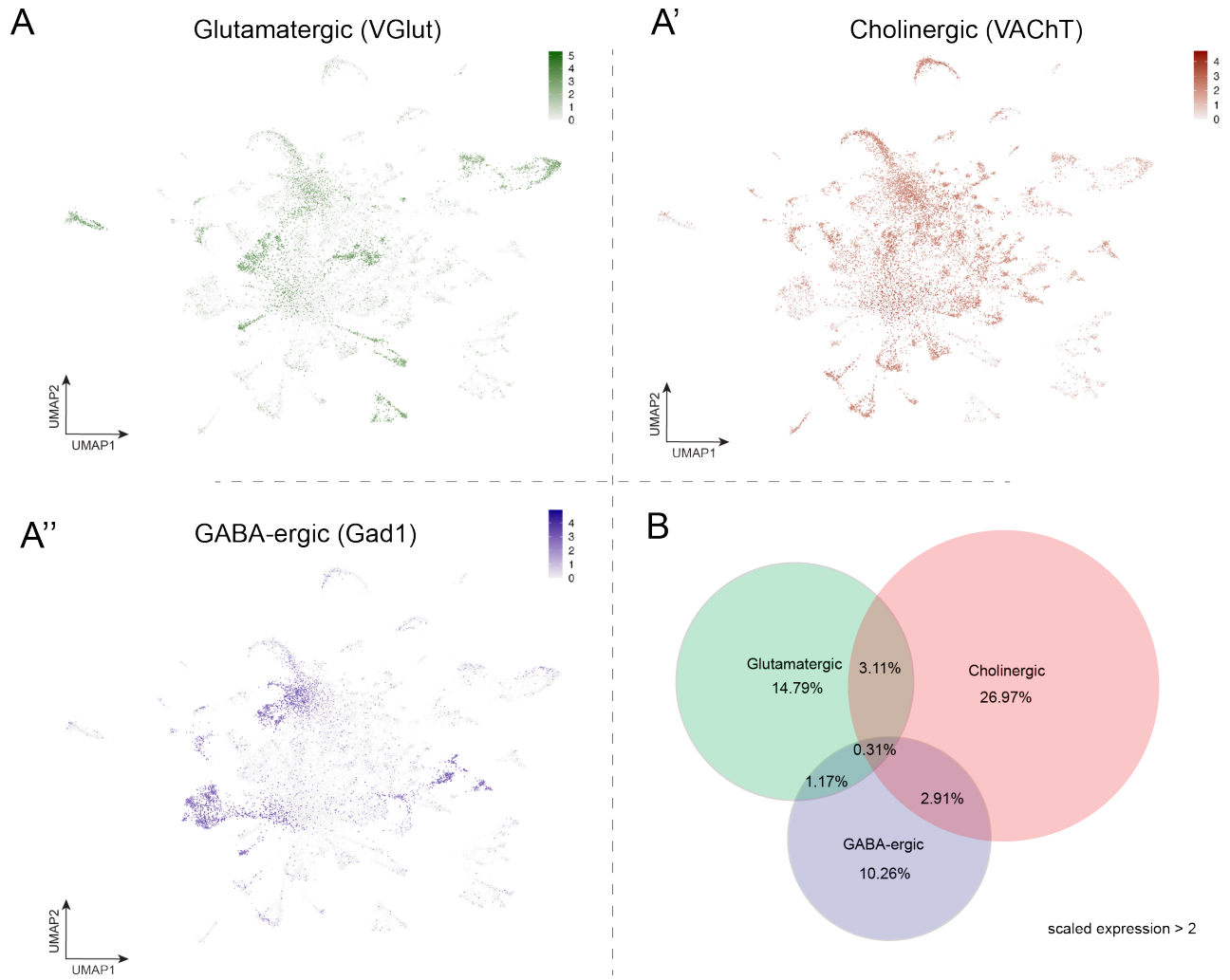


Figure 3. Expression of fast-acting neurotransmitters in neurons derived from T2 neural stem cell (NSC) lineages.

(A-A'') UMAP distribution plots demonstrate the expression of the following neurotransmitters: **(A)** *vesicular glutamate transporter* (VGlut, glutamatergic neurons), **(A')** *vesicular acetylcholine transporter* (VAChT, cholinergic neurons) and **(A'')** *glutamic acid decarboxylase 1* (Gad1, GABAergic neurons). All plots have a minimum cutoff value set at 0. **(B)** Venn diagram quantifies the number of cells in the atlas expressing each neurotransmitter with a scaled expression > 2.

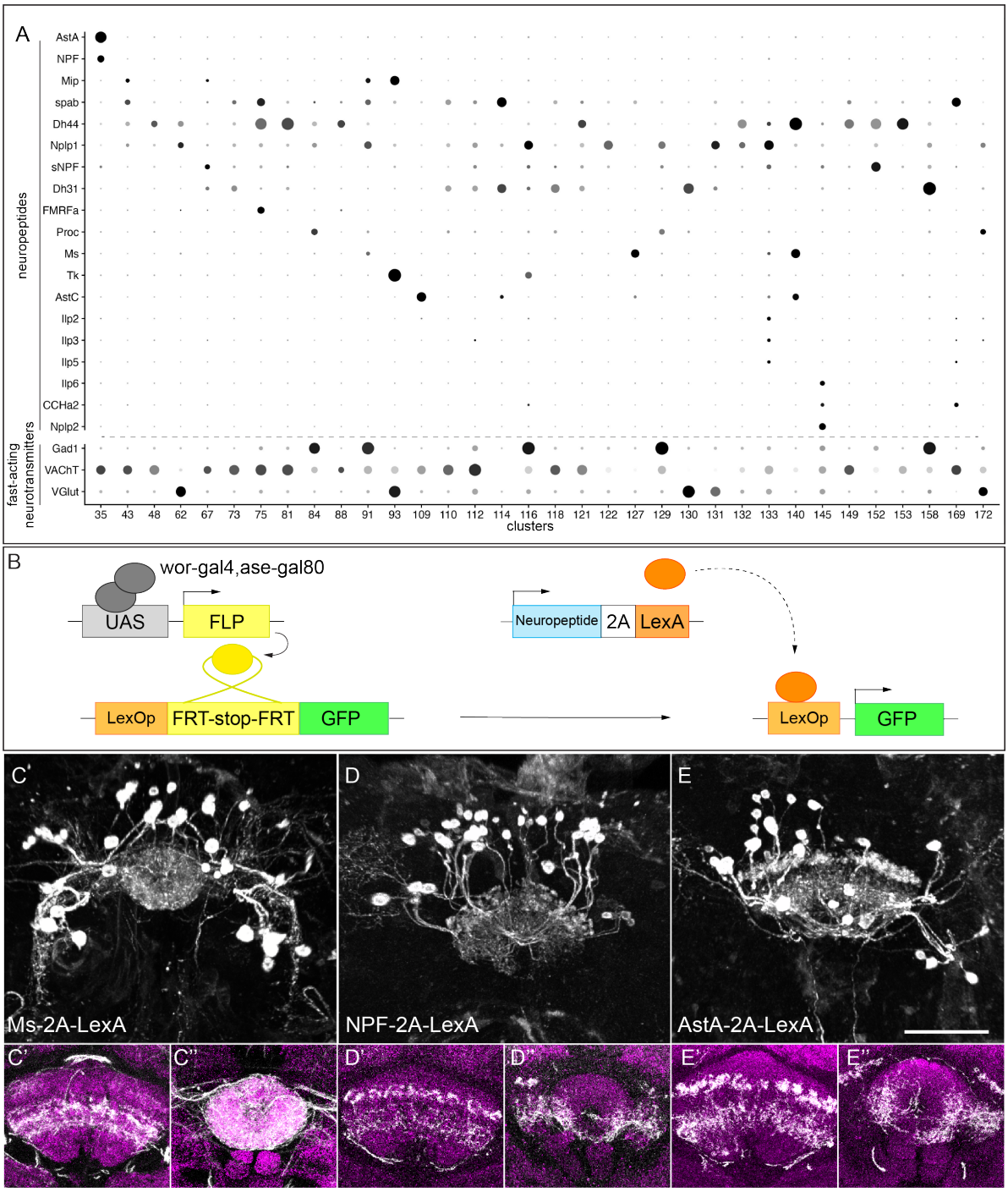
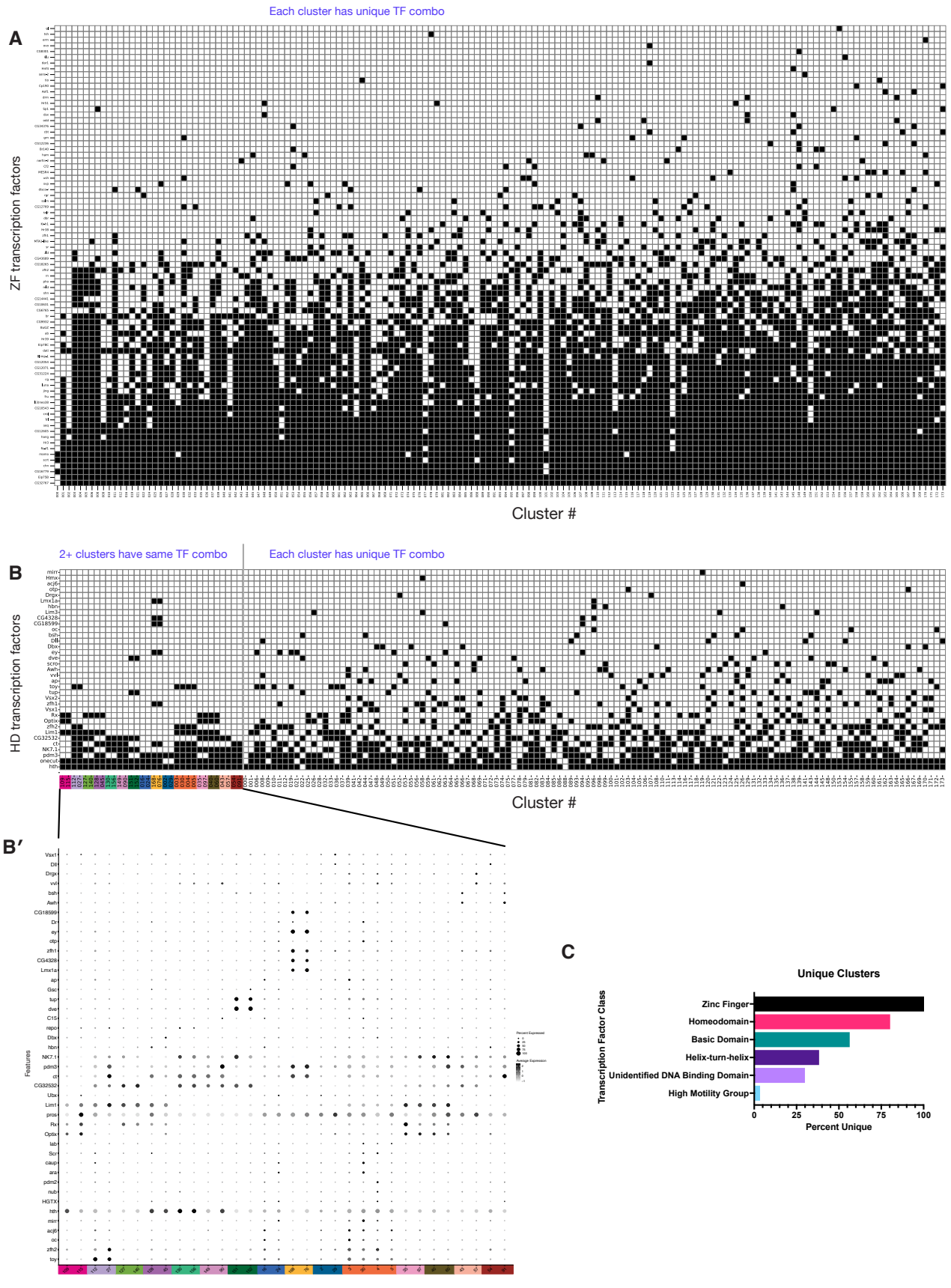


Figure 4. Neuropeptide expression in T2 NSC derived-neurons.
(A) Dot plot showing expression of 12 neuropeptides and their co-expression with the three fast-acting neurotransmitters (dashed line) across selected clusters. **(B)** Genetic scheme to map neuropeptide expression in T2-derived adult neurons. **(C-E)** Neuropeptide-expressing neurons labelled with GFP. **(C-E)** Three-dimensional projections. **(C'-E')** Fan-shaped body projections. **(C''-**

762 **E'')** Ellipsoid body projections. nc82 counterstains (magenta) in the brain for neuropil projections.
763 Scale bar represents 20 μm.


Each cluster has unique TF combo
A
ZF transcription factors
Cluster #
2+ clusters have same TF combo
Each cluster has unique TF combo
B
HD transcription factors
Cluster #
B′
Features
C
Unique Clusters
Transcription Factor Class
Zinc Finger
Homeodomain
Basic Domain
Helix-turn-helix
Unidentified DNA Binding Domain
High Motility Group
0
25
50
75
100
Percent Unique

764

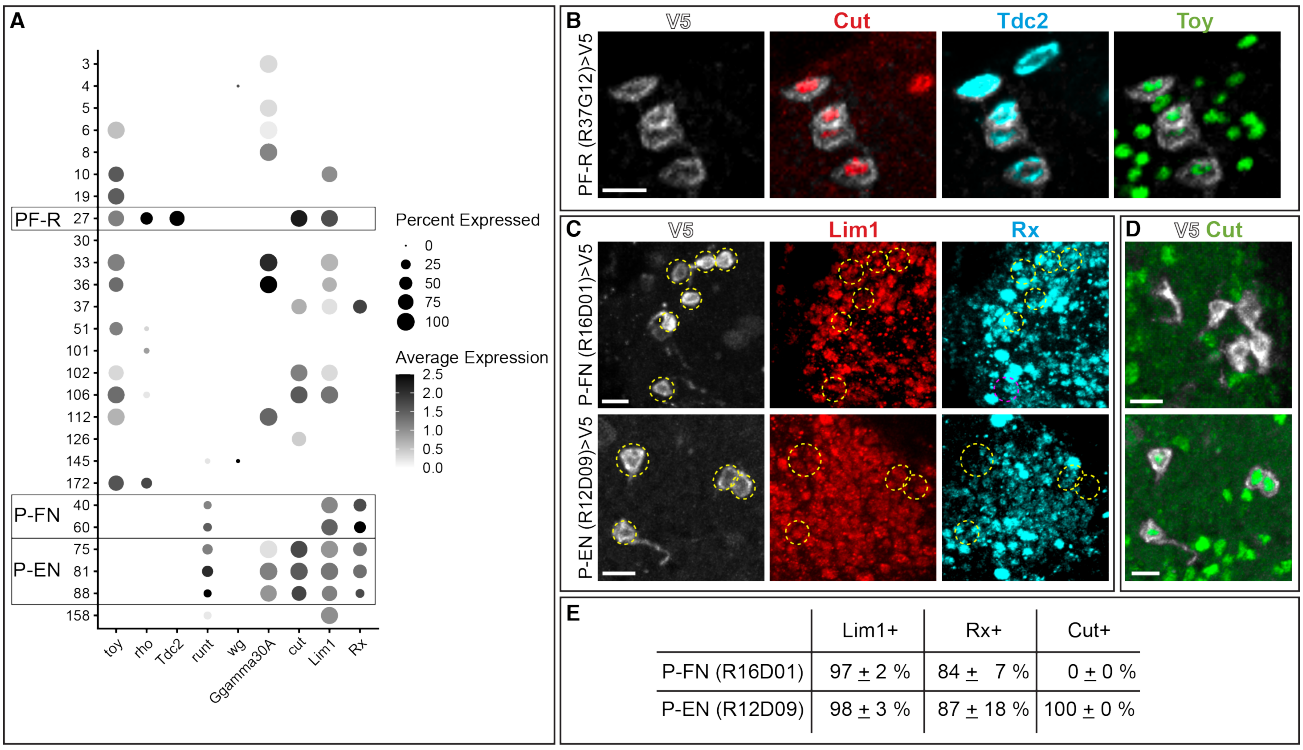
Figure 5. Heatmap of binarized expression of homeodomain and zinc finger transcription factors.

(A) Zinc finger TFs with differential expression among clusters. TFs that are 0 or 1 in all clusters (i.e. not generating specificity in the clusters) not shown and are listed below. Zinc finger TFs with +1 in all clusters: *lola, chinmo, EcR, jim, CG43347, Glut4EF*. Zinc finger TFs with 0 in all clusters: *btd*, *CG17361, CG17359, Phs, CG11456, Neu2, CG11247, CG14655, CG14667, CG10979, Hr83, Zif, CG7963, CG10654, nom, CG8159, ranshi, M1BP, CG8319, CG8478, topi, CG6254, CG4820, CG6689, CG6791, CG14710, CG14711, ouib, CG11560, ssp, CG7368, Lime, CG12744, CG18011, CG12942, CG12391, sug, CG17385, CG8089, CG8388, CG4282, CG15073, CG11906, CG10321, CG9890, CG15710, CG3065, CG4707, CG1233, Kah, Asciz, CG13287, CG10274, CG7386, CG10147, CTCF, ERR, phol, CG6813, CG18476, MBD-R2, CG5245, CG30431, Opbp, CG31365, CG31388, CG31441, CG31612, CG32121, tef, pzg, CG3281, hkb, CG42726, CG42741, Hr4, CG44002, zf30C, prg, ich, eg, CG15269, kmg, Prdm13, CG17801, tll, Sry-beta, CG7691, CG7271, CG30020, CG1663, CG12219, Dlip1, CG7987, CG6654, GATAe, pad, CG17803, CG17806, Odj, CG17186, CG4424, CG4854, trem, CG4360, dmrt93B, lmd, CG4374, Spps, CG10669, CG4730, CG1647, dmrt99B, CG15514, ZIPIC, CG1792, ATbp, MTF-1, CG11398, CG14050, CG18764, CG1602, pita, bowl, cas, CG1603, wdn, noc, sqz, ken, CkIIalpha-i1, Hr78, Hr96, dsf, Dref, D19B, D19A, CG2712, az2, Meics, CG1024, CG11902, CG17328, ci, CG16863, usp, term, disco, dwg, grau, hb, her, kni, knrl, Kr, esg, wor, wek, sens, opa, peb, pnr, sna, srp, Sry-delta, su(Hw), ttk, nerfin-1, sob, CG4496, CG11696, CG10348, CG11695, CG4318, CG17612, Plzf, CG9215, CG8944, CG9609, CG12299, CG7101, CG1529, GATAd, CG13123, CG3407, CG3847, CG13775, CG10431, CG17568, dmrt11E, CG2202, Clamp, CG3032, CG2116, CG2120, CG2129, CG10462, CG18262, CG2889, CG10366, CG10959, CG15435.*

(B) Homeodomain TFs (HDTFs). HD TFs with +1 in all clusters: *pros*. HD TFs with 0 in all clusters: *abd-A, CG11085, CG11294, Six4, OdsH, ind, PHDP, unc-4, Pph13, Ptx1, zen2, caup, ara, unpg, B-H1, repo, CG11617, vis, achi, lms, CG15696, prd, H2.0, eve, bcd, vnd, Antp, lbe, nub, CG34031, pb, exex, HGTX, HHEX, toe, CG9876, CG34367, Gsc, zen, E5, Scr, ro, slou, lab, lbl, gsb-n, ftz, so, eyg, en, ems, Dr, Dfd, cad, al, Abd-B, exd, Ubx, inv, bap, B-H2, pdm2, tin, C15*.

(B') Dot plot of the top five marker genes for each example of two or more clusters that shared the same HDTF combination.

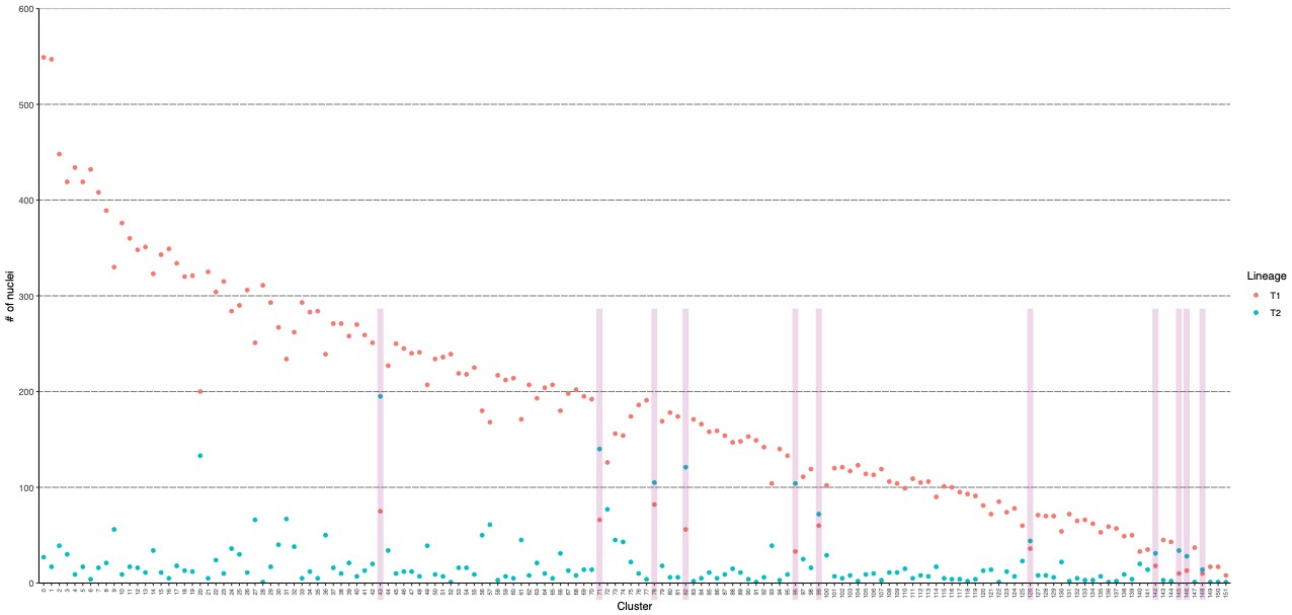
(C) Percentage of neuronal clusters which have unique combinations of each class of TFs: homeodomain TFs = 79.4%, zinc finger TFs = 100%, helix-turn-helix TFs = 37.4%, high motility group TFs = 2.6%, basic domain TFs = 55.5%, unidentified DNA binding domain TFs = 29%).



	Lim1+	Rx+	Cut+
P-FN (R16D01)	97 ± 2 %	84 ± 7 %	0 ± 0 %
P-EN (R12D09)	98 ± 3 %	87 ± 18 %	100 ± 0 %

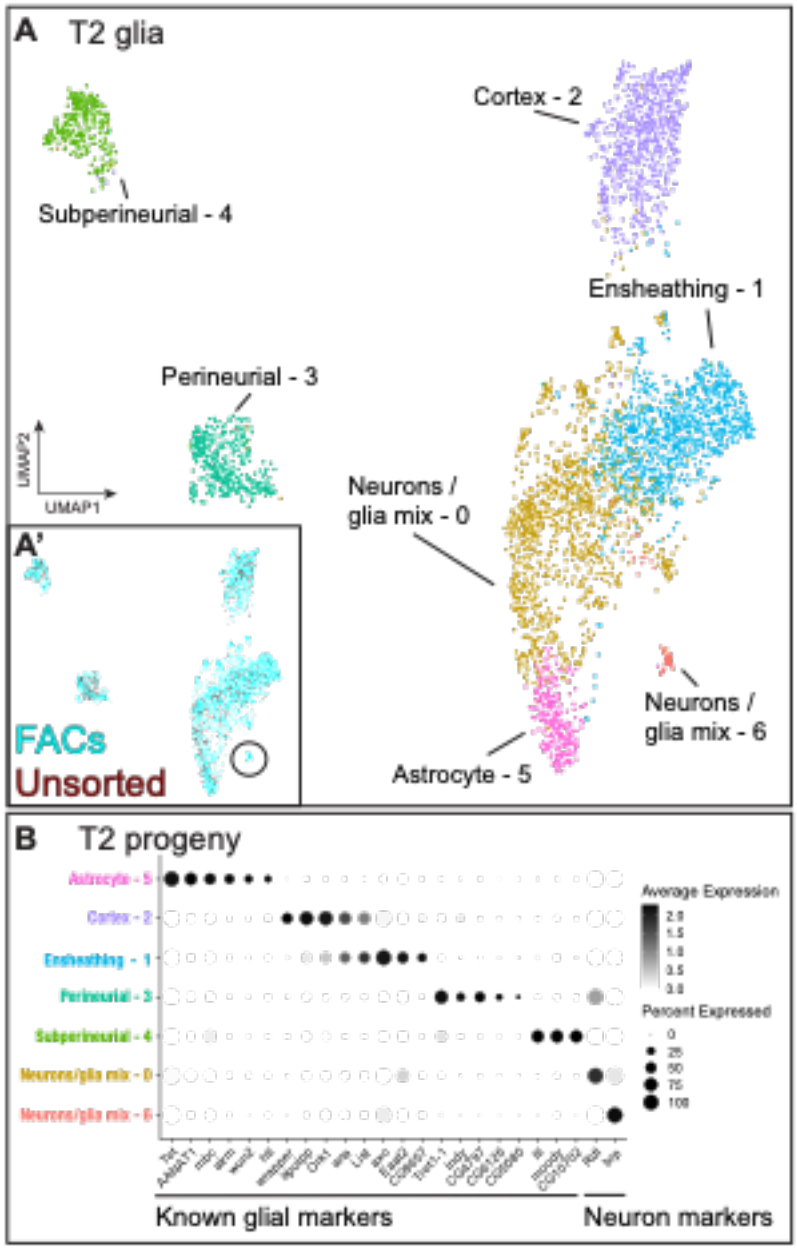
Figure 6. Identification of CX neurons.

(A) Dot plot of the known genes (*runt*, *toy*) expressed in central complex neurons (PF-R, P-FN, and P-EN), lexA-driver enhancer gene (*rho*, *wg*, *Ggamma30A*), and enriched genes found in the central complex neuron clusters (*cut*, *Lim1*, and *Rx*). **(B)** PF-R cell bodies labeled with V5, and co-stained with marker genes boxed in (A). **(C-D)** PF-N and PE-N cell bodies labeled with V5, and co-stained with marker genes boxed in (A). **(E)** Percentage of marker genes in PE-N and PF-N. Scale bar in (B-D): 5 μm.

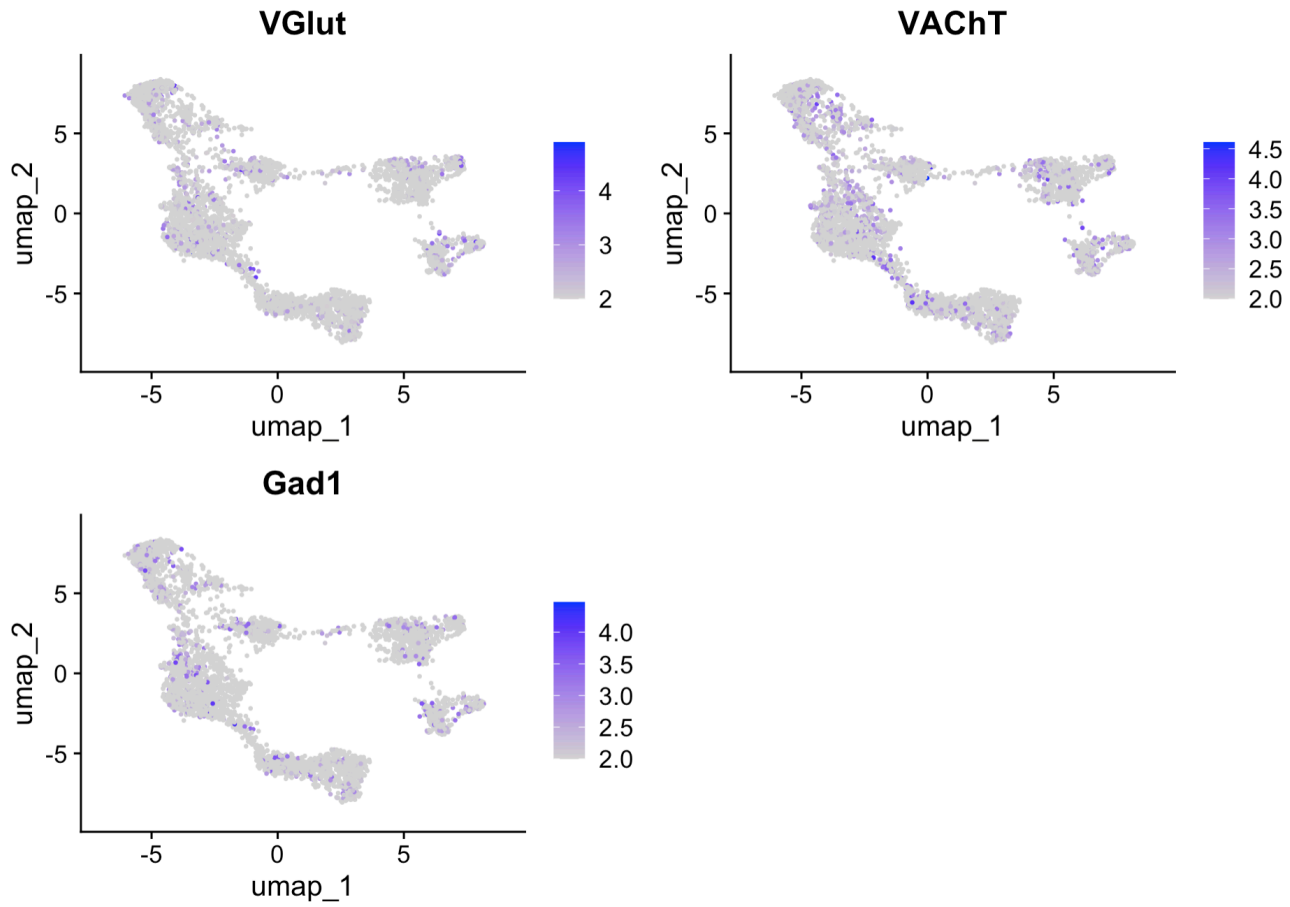


Supplemental Figure 1. Nuclei numbers T1 vs T2

Vertical shades mark the cluster with T2/(T1+T2) > 50%.



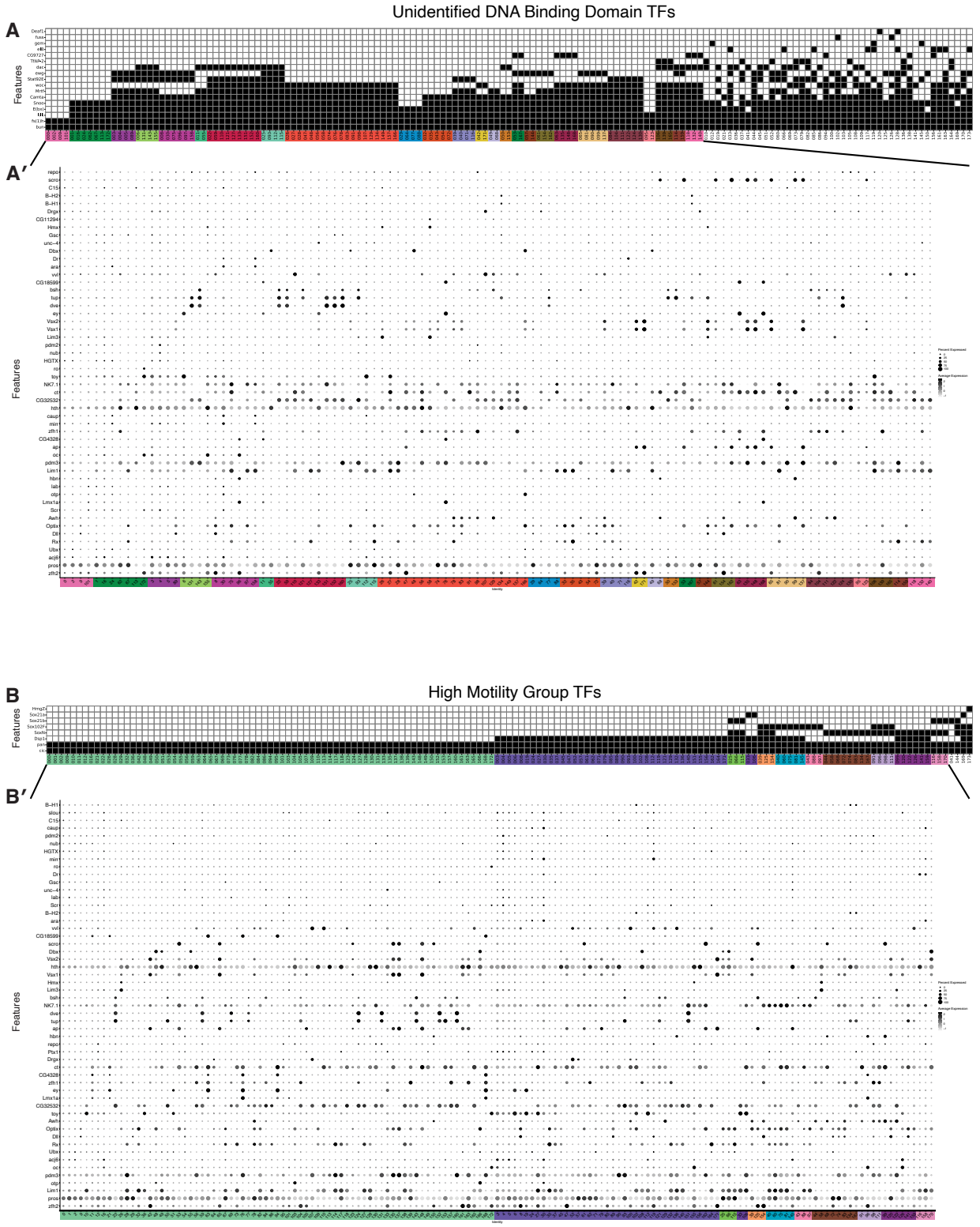
Supplemental Figure 2. Glial cell types present in the T2-derived cell atlas. **(A)** We sub-clustered the T2 atlas 3670 glial nuclei from clusters that expressed the pan glial marker *repo* (Campbell et al., 1994; Xiong et al., 1994). Similar to the whole brain atlas, we assigned T2 glial progeny to 5 glial cell types. We identified two clusters contained a mix of neuronal and glial markers that likely reflect neuronal contamination due to high expression levels of *Rdl* and *brp*, as seen in other glial scRNAseq atlases (Lago-Baldaia et al., 2023). **(B)** The T2 glial progeny subset did not contain the surface glial cell type found in the T1+T2 atlas as we resolved surface glial subtypes perineurial and subperineurial glia. We note that our atlas contains T2 progeny that remains localized to the central brain; therefore, we do not have coverage of glia that may migrate outside of the central brain to the optic lobe.



826

827 **Supplemental Figure 3. Glia show sparse expression of fast neurotransmitters.**

828 Data analyzed with stadard selection criteria of expression >2.



Supplemental Figure 4. Heatmaps of binarized gene expression of different TF classes.

(A) Unidentified DNA binding domain TF heatmap **(A')** Dot plot of top five marker genes for clusters which share Unidentified DNA binding domain TF combination. **(B)** High motility group TF heatmap **(B')** Dot plot of top five marker genes for clusters which share high motility group TF combination. **(C)** Basic domain TF heatmap **(C')** Dot plot of top five marker genes for clusters which share basic domain TF combination. **(D)** Helix-turn-helix TF heatmap **(D')** Dot plot of top five marker genes for clusters which share helix-turn-helix TF combination. **(E)** Zinc finger TF heatmap. Those in each class with zero score are listed below. Unidentified DNA binding domain TFs: *D1, grh, CG32772, CG12659, p53, CG15011, NC2alpha, row, NC2beta, CG11970, brk, Rfx, insv, Mabi, Hsf, stc, z, Su(H).* Basic domain TFs: *ase, Usf, REPTOR-BP, Hand, sisA, Sidpn, maf-S, Irbp18, sage, Fer3, Fer2, dysf, bigmax, Mitf, tx, tgo, Atf-2, da, HLH54F, Hesr, E(spl)m8-HLH, E(spl)m7-HLH, E(spl)mdelta-HLH, nau, slbo, ac, Hey, cyc, Oli, l(1)sc, E(spl)mbeta-HLH, Met, twi, E(spl)m3-HLH, dpn, ato, net, HLH4C, HLH3B, Jra, tap, vri, gt, Max, Xbp1, ss.* Helix-turn-helix TFs: *fd19B, FoxL1, CG15812, Pbp95, danr, dan, jumu, croc, Dlip3, fd96Cb, fd96Ca, dany, CG32006, retn, fd3F, stwl, slp1, Adf1, rib, bin, Poxm, Myb, CG8924, Bdp1.* High motility group TFs: *D, Lpt, polybromo, TFAM, CG12104, CG4617, Hmg-2, Bap111, bbx, Ssrp, Sox15, Sox14, HmgD, Sox100B, tHMG1.*

848

Supplemental Video 1. Rotation of Imaris three-dimensional view of Ms-2A-LexA expression. Genotype: 20xUAS-flp;worniu-gal4,asense-gal80;lexAop-FRT-stop-FRT-myr::gfp x Ms-2A-LexA

Supplemental Video 2 . Rotation of Imaris three-dimensional view of NPF-2A-LexA expression. Genotype: 20xUAS-flp;worniu-gal4,asense-gal80;lexAop-FRT-stop-FRT-myr::gfp x NPF-2A-LexA

Supplemental Video 3. Rotation of Imaris three-dimensional view of AstA-2A-LexA expression. Genotype: 20xUAS-flp;worniu-gal4,asense-gal80;lexAop-FRT-stop-FRT-myr::gfp x AstA-2A-LexA