HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2019 July 26.

Published in final edited form as:

Cell. 2018 July 26; 174(3): 622–635.e13. doi:10.1016/j.cell.2018.05.021.

Phenotypic convergence: distinct transcription factors regulate common terminal features

Nikolaos Konstantinides[1,*], **Katarina Kapuralin**[2,*], **Chaimaa Fadil**[1,2], **Luendreo Barboza**[1,3,4], **Rahul Satija**[1,4], and **Claude Desplan**[1,2,+]

[1]Department of Biology, New York University, New York, NY 10003, USA

[2]New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, UAE

[3]Neuroscience Institute, NYU Langone Medical Center, New York, NY 10016, USA

[4]New York Genome Center, New York, NY 10013, USA

Summary

Transcription factors regulate the molecular, morphological, and physiological characters of neurons and generate their impressive cell-type diversity. To gain insight into general principles that govern how transcription factors regulate cell-type diversity, we used large-scale single-cell RNA-sequencing to characterize the extensive cellular diversity in the *Drosophila* optic lobes. We sequenced 55,000 single cells and assigned them to 52 clusters. We validated and annotated many clusters using RNA-sequencing of FACS-sorted single cell-types and cluster-specific genes. To identify transcription factors responsible for inducing specific terminal differentiation features, we generated a 'random forest' model and we showed that the transcription factors Apterous and Traffic-jam are required in many, but not all cholinergic or glutamatergic neurons. In fact, the same terminal characters can often be regulated by different transcription factors in different cell types, arguing for extensive phenotypic convergence. Our data provide deep understanding of the developmental and functional specification of a complex brain structure.

ETOC

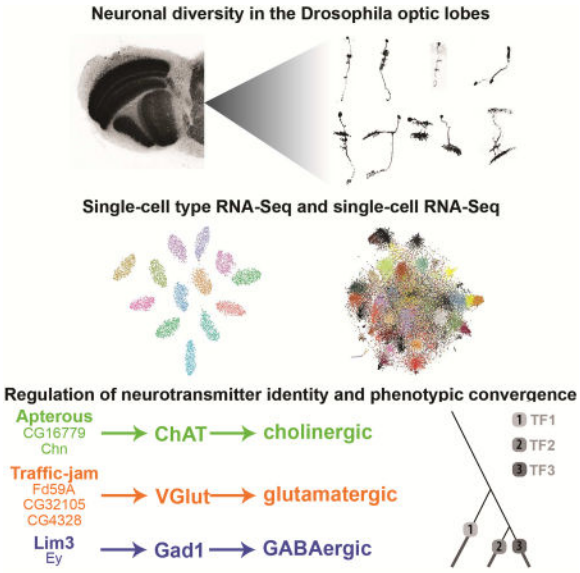
Correspondence: Nikolaos Konstantinides nk1845@nyu.edu, Claude Desplan cd38@nyu.edu.
[+]Lead contact, cd38@nyu.edu
[*]These authors contributed equally

Author Contributions

Conceptualization, N.K, K.K., and C.D.; Methodology, N.K, K.K., R.S., and C.D.; Investigation, N.K, K.K., C.F., L.B.; Writing – Original Draft, N.K, K.K., and C.D.; Writing – Review & Editing, all authors; Supervision, R.S. and C.D.

Declaration of Interests

The authors declare no competing interests.



A single cell analysis of the fly optic lobe reveals extensive phenotypic convergence, with different sets of transcription factors promoting similar outcomes in different cell types.

Keywords

Neuronal diversity; neuronal development; single cell sequencing; scRNA-seq; transcription factors; neurotransmitters; gene regulation; modeling; cell type evolution; *Drosophila* optic lobe

Introduction

Different cellular characteristics define unique cell types. Before the molecular revolution, cell types were distinguished by their morphology (Cajal, 1915; Fischbach and Dittrich, 1989; Morante and Desplan, 2008), and whenever possible, by their function and physiology (Kepecs and Fishell, 2014). The advent of RNA sequencing technologies has allowed us to revisit cell type classification, identify new cell types and corroborate pre-existing ones (Macosko et al., 2015; Shekhar et al., 2016). Transcription factors are activated by other transcription factors or by signaling pathways and activate downstream effector genes, thereby shaping these cell types. Therefore, transcription factors represent the core of cell type identity by controlling structural (Santiago and Bashaw, 2014), molecular (Hobert, 2016), and physiological (Kratsios et al., 2015) properties of cells. Thus, a fundamental question in developmental biology is: How do transcription factors generate cell types with specific characteristics?

In *Drosophila,* temporal, spatial, and morphology transcription factors are expressed at different developmental stages (Bayraktar and Doe, 2013; Brody and Odenwald, 2000; Enriquez et al., 2015; Erclik et al., 2017; Li et al., 2013) to activate terminal selectors that regulate the genes that define the morphology and physiology of adult neurons (Hobert, 2011). Terminal selectors represent the ‘core regulatory complex’ of transcription factors, which *(i)* establishes cell types and *(ii)* can lead to the evolution of new cell types, when

their expression is altered (Arendt et al., 2016). This core regulatory complex shapes neuronal diversity in two timescales: first, by implementing their identity over the course of development and, second, by diversifying cell types over millions of years of evolution. Therefore, the transcription factor underpinnings of neuronal morphology, physiology, and molecular identity are at the intersection of the evolutionary and developmental history of a neuron.

Only recently have we started to comprehend molecularly the huge diversity of mammalian neurons (Macosko et al., 2015; Poulin et al., 2016; Wichterle et al., 2013). In parallel, systematic work in the *Drosophila* optic lobe, which consists of ~70,000 cells, has provided an in-depth and nearly exhaustive description of more than 100 cell types (Cajal, 1915; Fischbach and Dittrich, 1989; Nern et al., 2015). Therefore, the *Drosophila* optic lobe represents an ideal system to seek a mechanistic understanding of the way transcription factors regulate neuronal identity.

We currently have a deep understanding of how different neuronal types are specified from a pool of neuroblasts in the developing optic lobe. The intersection of temporal and spatial transcription factors in the neuroblasts, as well as Notch-driven binary cell fate decisions of the ganglion mother cells and apoptotic cell death are responsible for most, if not all, of the observed neuronal diversity (Erclik et al., 2017; Li et al., 2013; Pinto-Teixeira et al., 2016; Suzuki et al., 2013). However, it is still not understood how a neuron differentiates once specified, *i.e.* how it acquires the characters that endow its identity.

Although a fair number of Gal4 drivers that label a single cell type in the optic lobe are now available (Jenett et al., 2012; Nern et al., 2015), most of the neuronal types in the optic lobes remain molecularly inaccessible. Single cell RNA sequencing techniques, such as Drop-seq, have revolutionized our access to different cell types by offering an unbiased way of sampling single cells from the tissue of interest and determining their transcriptomes (Karaiskos et al., 2017; Macosko et al., 2015).

We combined Drop-seq of single neurons and RNA sequencing of FACS-sorted single cell types to obtain the transcriptome of all neuronal and glial cells in the adult optic lobes. We used the FACS-sorted cell type-specific sequencing data to identify biologically meaningful clusters in the single cell clustering. We annotated additional clusters using Gal4 lines for marker genes specific to each cluster. A ‘random forest’ machine learning model was trained from the data and was used to predict the expression of terminal genes based on the transcription factor profile of each cell, allowing us to establish causal relationships between transcription factors and the genes that participate in the generation and release of neurotransmitters. Notably, we find that each cell type uses distinct combinations of transcription factors to regulate the expression of effector genes. Altogether our study provides a detailed understanding of the developmental and functional specification of nearly all neurons in a complex brain structure and sets the groundwork to address the evolutionary and developmental origin of this diversity.

Results

Drop-seq of single neurons and glia in the Drosophila adult optic lobe

To obtain an unbiased characterization of all cell types and their transcriptomes, we used Drop-seq (Macosko et al., 2015) to sequence a large number of single cells from the *Drosophila* adult optic lobes (Figure 1A). *Drosophila* neurons' diameter is ~3um, while it is ~10um in mice (Shekhar et al., 2016), which results in an approximately 30-fold volume difference. For this reason, cDNA recovery after a Drop-seq run and reverse transcription was too low to allow library preparation with current protocols (Figure S1A) (Macosko et al., 2015). We adjusted the protocol (STAR Methods) and increased cDNA yield >10-fold (Figure S1A).

We sequenced the transcriptome of ~57,601 adult optic lobe cells with a minimum gene-per-cell cutoff at 200 genes (STAR Methods and Figure S1B) and used Seurat (Satija et al., 2015) for the downstream analysis. 1,303 variable genes were used to perform principal component analysis (PCA). Glial genes contributed significantly to the first, second and third PCs, indicating a clear separation of glia and neurons (Figure 1B). The next PCs separated neurons from each other based on their terminal identity (expression of genes involved in neurotransmitter synthesis and delivery, neurotransmitter receptors, cell adhesion molecules etc). For example, PC6 separated glutamatergic and GABAergic neurons from the other neurons, while PC11 split glutamatergic from GABAergic neurons (Figure 1B). We identified 61 transcriptionally distinct clusters (STAR Methods, Figure S1C). We used t-SNE to visualize the clusters (Figure 1C). To avoid over-clustering, we assessed all terminal nodes of a hierarchical clustering tree (STAR Methods and Figure S1C–L). This allowed us to eliminate clusters 5 and 36 (enriched in heat-shock proteins indicating stressed cells) and to merge clusters 23 and 40, 29 and 35, 11 and 50, 20 and 41, and 18 and 34. After eliminating the two smallest clusters (59 and 60), we ended up with 52 clusters for further analyses.

We then performed hierarchical clustering using transcription factor expression as a metric (Figure 1D). As expected from the PCA, the first split differentiated glia and neurons (Figure 3). We also identified specific markers for each of the Drop-seq clusters using Receiver Operating Characteristic (ROC) curves (Table S1 and Figure 1E). Most of the markers were not unique to a specific cluster, which was expected as different neuronal types share gene batteries for performing specific functions (Achim and Arendt, 2014; Hartwell et al., 1999). Nonetheless, most of the clusters could be distinguished by the expression of a combination of genes.

In summary, this unbiased technique identified 52 reliable clusters comprising a total of 54,974 cells that represent a large proportion of the optic lobe cellular diversity.

Alignment of FACS-sorted and Drop-sequenced neuronal transcriptomes and annotation of Drop-seq clusters

A concern that arises from clustering cells using Drop-seq data is whether these clusters actually represent cell types, or if the clustering is based on other parameters that could introduce biases. Differentially tuning the parameters of the k-nearest neighbor algorithm

can over- or under-cluster the single cells. To address this concern, we sequenced the bulk transcriptome of single neuronal cell types from the adult optic lobes after FACS-sorting. We selected 17 cell types that are a representative fraction of optic lobe neurons. They comprise unicolumnar and multicolumnar neurons, local and projection neurons, and cholinergic, GABAergic, and glutamatergic neurons (Figure S2A–B).

To compare the two datasets, we used genes differentially expressed amongst the Drop-seq clusters. A caveat for this comparison is that the two datasets were obtained using vastly different methods. To account for these differences, we simulated single cell data originating from the FACS-sorted cell type-specific transcriptomes. We then plotted the expression of the cluster-specific markers on a heatmap and compared the “simulated single cell” heatmap from each FACS-sorted cell type (Figure S2C) to the real single cell Drop-seq clusters (Figure 1E). Figures 2A and S2D show that 15 out of the 17 genetically labeled neuronal types mapped to a unique cluster, verifying the accuracy of our single cell clustering. The remaining 2 cell types (Tm5c and Dm12 neurons) mapped to two or more clusters (Figure S2E). Pearson correlations between the simulated single cell transcriptomes of FACS-sorted cells and the transcriptomes of Drop-seq clusters supported the visual matching (Figure 2B and S2E). In general, the correlation between cell types and clusters was higher for cells types that are represented by more neurons in each brain (unicolumnar neurons), as opposed to less abundant cell types (multicolumnar neurons) and cell types that are heterogeneous (e.g. Tm5ab consists of 2 cell types).

The optic lobe consists of more than 100 different cell types. Since we recovered 52 clusters, some clusters must contain cells of more than one cell type. Attesting to the quality of the clustering, we observed that very similar cell types clustered into the same cluster, *i.e.* C2/C3 (cluster 8), T2/T3 (cluster 6), T4/T5 (cluster 2), and Lawf1/Lawf2 (cluster 25) (Figure 2A, S2D, and 3B).

To further demonstrate the accuracy of our clustering and its potential to discriminate between different cell types, we wanted to identify instances where a seemingly homogeneous population actually consisted of more than one cell type. For this purpose, we used a Gal4 line that marks two different cell types - T1 and Tm1 (Figure 2C). We obtained the cell type-specific transcriptome, generated “simulated single cell” transcriptomes, and compared it to the single cell cluster heatmap. The “simulated single cells” mapped to two different clusters (12 and 23), which correspond to neuronal types Tm1 and T1 (Figures 2D and 2E).

The mapping of the transcriptomes of these single cell types to unique clusters, as well as the splitting of a heterogeneous cell population into its constituent cell types, indicate the robustness of our clustering technique and the reliability of the clusters.

Identification and annotation of different glial cell types

One of the drawbacks of Drop-seq is that it does not provide the identity of the sequenced single cells and requires other approaches for their annotation. Matching cell type-specific transcriptomes to unique clusters is a perfect way to annotate the clusters for which FACS-sorted cell transcriptomes were available (Figure 2 and Figure 3B). Although this validates

the approach, it does not generate new information about the remaining cell types. We thus aimed to annotate the remaining clusters that should correspond to cells that were not in the collection of FACS-sorted cells.

As a simplified starting point, we focused on annotating glial clusters, which account for ~15% of the total number of clusters (7 out of 52) and represent ~10% of the cells in the adult *Drosophila* central nervous system. Several glial cell types have been described in the *Drosophila* optic lobes: surface glia (comprising perineurial and subperineurial glia), cortex/satellite glia, ensheathing glia, neuropile glia, astrocyte-like glia, and chiasm glia (Kremer et al., 2017).

To annotate the glial clusters, we used available transcriptomic information and reporters (Figure 3A). *AdamTS-A* is more highly expressed in surface glia (perineurial and subperineurial glia) as compared to other glia (DeSalvo et al., 2014) and was expressed in clusters 49 and 58. *CG4797* and *gemini* are expressed in perineurial glia that correspond to cluster 49 in which they are also abundant (Figure 3A–Figure S3A). Three clusters (26, 30, and 38) correspond to ensheathing glia: two of them contained *Glutamate synthase 2* (*Gs2*) and *Excitatory aminoacid transporter 1 (Eaat1)* and correspond to neuropile glia and astrocyte-like glia (clusters 26 and 38) (Figure 3A, 3C and Figure S3B). The two clusters were further distinguished by the expression of *ebony*, which is only expressed in neuropile glia (cluster 26) (Richardt et al., 2002). Cluster 30 expressed *Draper* and corresponds to ensheathing glia with phagocytic activity (Doherty et al., 2009). *Draper* was also expressed in cortex glia and was found in higher levels in cluster 56, which was annotated as cortex glia. Furthermore, *wrapper* was only found in cluster 56. We used a wrapper-Gal4 line to show that *wrapper* is indeed only expressed in cortex glia (Figure 3D). Finally, *hoepel1* was most highly found in cluster 57, and *hoepel1* is expressed in chiasm glia (Figure S3C). Therefore, we were able to identify all known glial types in our clusters and assign identity to all seven glial clusters based on specific markers found in each cluster (Figure 3A).

Molecular markers for the characterization and annotation of neuronal Drop-seq clusters

To annotate more unidentified neuronal clusters, we used known and newly identified cell type-specific markers. Neuroblasts that express *hth* are responsible for generating Mi1, Pm1, Pm2, and Pm3 neurons (Erclik et al., 2017). Mi1 expresses *bsh* and *hth* while Pm1, Pm2, and Pm3 express *Lim3* and *hth*. Moreover, Pm1 and Pm2 express *svp*, Pm1 expresses *tsh*, and Pm3 expresses *Vsx1*. We looked for clusters expressing these combinations of genes and identified cluster 17 as Mi1 (which was also identified by the Drop-seq-FACS comparison – Figure 2A), cluster 52 as Pm3, and cluster 13 as Pm1/Pm2 (Figure 3B). Pm1 and Pm2 are very similar cell types, which explains why they map to the same Drop-seq cluster.

To annotate some of the unidentified clusters, we used Gal4 lines for cluster-specific markers. *kn* (*collier*) was only expressed in cluster 15. We used Multi-Color Flip-Out (MCFO) (Nern et al., 2015) with a *kn*-Gal4 line to generate single cell clones in the adult brain to identify the corresponding neuron. The line marked only one cell type, which was identified as TmY14 (Figure 3E). Similarly, *CG42458* was highly expressed in cluster 12. A *CG42458*-Gal4 line was expressed in TmY8 (Figure 3F). *CG42458* was also present in

cluster 2, which has already been identified as T4/T5 cells, and in cluster 45, which was annotated as Mt1 neurons, as indicated in Figures S3D–D'.

In conclusion, we were able to assign 32 cell types to 23 clusters among the 52. This provides an in-depth description of the molecular properties of cell types that were previously only known from their morphology or a few molecular markers.

Transcription factor-based hierarchical clustering is different from whole transcriptome-based clustering

Single cell transcriptomes are too shallow to recover the expression of every transcription factor in every single cell. To study which transcription factors regulate downstream effectors and cell type identity, we treated the data as cluster transcriptomes rather than as single cells. The reads of all single cells that belong to one cluster were merged and normalized (STAR Methods). We first looked at the expression of transcription factors in the clusters. There are two categories of transcription factors: 72 transcription factors were found at similar levels in most clusters (11 ubiquitous and 61 pan-neuronal), while 598 cell type-specific transcription factors were expressed at significantly higher levels in only one or few clusters (Figure 4A). We also performed weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008) and observed that the transcription factors of each co-expression module were mostly expressed in one cluster (Figure 4A' and S4A). There was no extensive transcription factor fingerprint overlap between cell types that are closely related. This suggests that apparently similar cell types have different compositions of transcription factors, which we will address in more detail below.

Having assigned transcription factor fingerprints to each cluster, we sought to determine how they regulate terminal characters. One of the main terminal characters of a neuron is its neurotransmitter identity. We thus analyzed the neurotransmitter composition in each of the clusters. The majority of clusters expressed *Choline acetyltransferase* (*ChAT)* and were thus cholinergic. Most of the others were either glutamatergic (*Vesicular glutamate transporter* - *VGlut*) or GABAergic (*Glutamic acid decarboxylase 1* - *Gad1*) (Figure 4B). Indeed, antibody staining for *ChAT* (Figure 4C) and *VGlut* (Figure 4C') showed broad expression patterns in the optic lobes. Two clusters corresponded to neuronal types that appear to be monoaminergic as they expressed the *Vesicular monoamine transporter Vmat* (Figure 4C'') (Nassel and Elekes, 1992). Tyrosine hydroxylase (*ple*) was also expressed in these neurons, indicating that they are likely dopaminergic (Daubner et al., 2011). We did not detect expression of *Tryptophan hydroxylase* (*Trh*) or *Tyrosine decarboxylase 2* (*Tdc2*) in any of the clusters suggesting that serotonergic, octopaminergic, and tyraminergic neurons (Cole et al., 2005) do not have cell bodies in the optic lobes. However, some modulatory neurons innervate the medulla as we detected expression of the octopamine receptors in different clusters. There are indeed octopaminergic neurons that reside in the central brain and send their axons to the optic lobe (Busch et al., 2009).

It is interesting to note that the neuronal cell types did not cluster according to their neurotransmitter profile in the transcription factor-based hierarchical clustering. To address this, we generated a hierarchical clustering tree based on the whole transcriptome rather than only transcription factors. The two trees were different (Figure S4B): we calculated the AU

(Approximately Unbiased) p-values using pvclust (Suzuki and Shimodaira, 2006); the strongly supported groups of clusters (red boxes in Figure S4B) were very different from each other in the composition of clusters. Moreover, we noticed that the whole transcriptome tree correlated better with the neurotransmitter composition as glutamatergic and GABAergic neurons clustered closer to each other (Figure S4C and C'). This result suggests that although two cell types may have a similar whole transcriptome composition, this similarity may be achieved by convergence of different combinations of transcription factors.

A 'random forest' model to predict transcription factors responsible for the regulation of downstream effectors

To directly address the regulation of different genes by transcription factors and establish a causal relationship between transcription factors and downstream targets, we generated a 'random forest' model that could predict the gene expression of a cell type based on its transcription factor profile potentially identify the transcription factors responsible for specific traits (Figure 5A and S5A). The predicted gene expression for the neuronal clusters was 98% accurate with a Pearson correlation ranging from 84.9% to 97.5% between actual and predicted transcriptomes. The performance of the model was lower for glia, with a Pearson correlation of 70.2% (Figure 5A–B). This is probably due to the lower number of glial clusters and, therefore, the incomplete training of the model. To address whether housekeeping genes were mainly driving the high accuracy, we compared the actual and predicted expression of the 1,303 variable genes we had identified earlier (Figure S5B). While the accuracy was still very high for neurons (88%), it was much lower for glia (27%), indicating that more clusters are necessary for training. Although most transcription factors appear to be expressed highly in few clusters each (Figure 4A), we were able to predict gene expression with high accuracy. This is likely because they are expressed more broadly at lower levels (Figure 4A and S5C).

Using this model, we identified the top transcription factors predicted to be responsible for cholinergic (*apterous*), glutamatergic (*traffic-jam*), GABAergic (*Lim3*), and monoaminergic (*CG33695*) fate. To rigorously test the model and verify that we can infer a causal relationship between the identified transcription factors and the neurotransmitter identity, we knocked-down in adult flies the transcription factor that was predicted to regulate a specific gene (to avoid developmental defects or apoptosis during neuronal development) and assessed the effect of the knock-down on the expression levels and pattern of that gene. A heat-inducible flip-out actin-Gal4 was activated in adult flies to drive expression of RNAi against *ap.* We performed antibody staining against *ChAT* in the presence and absence of the RNAi. *ChAT* staining was severely reduced in several medulla and lobula layers suggesting that *ap* is necessary for the expression of *ChAT* in many, although not all, cell types (Figure 5D). We also tested the effect of *tj* on *VGlut* expression. Synaptic boutons in medulla layers M1 and M6 that clearly expressed *VGlut* in wild-type optic lobes no longer expressed *VGlut* upon *tj* knock-down (Figure 5E). However, as with *ap* and *ChAT*, some glutamatergic neurons still expressed *VGlut*, suggesting that not all glutamatergic neurons rely on *tj* to regulate *VGlut* expression.

To assess the sufficiency of these transcription factors to control neurotransmitter identity, we mis-expressed them in cell types that normally do not express them. Mi1 is an ap^+ cholinergic cell type. We ectopically expressed *tj* in Mi1 in adults, using a temperature-sensitive Gal80 to avoid earlier expression. We did not observe any detectable upregulation of VGlut in *tj+* Mi1 cells; however, Gal4 expression was stochastically downregulated in Mi1 cells, which indicates a partial fate change (Figure S5D). We performed the opposite experiment by mis-expressing *ap* in the tj^+, glutamatergic adult Dm12 cells. The ectopic expression of *ap* changed the fate of the Dm12 cells, which no longer expressed Gal4; therefore, we were not able to assess the neurotransmitter expression in these cells. These results highlight the importance of *ap* and *tj* for the identity of different cell types and the expression of terminal effector genes. However, *tj* does not appear to be sufficient to drive *VGlut* expression.

We also tested whether there was cross-repression between cholinergic and glutamatergic identity, *i.e.* between *ap* and *VGlut*, and *tj* and *ChAT*. We knocked down *ap* in adult Mi1 cells, using a temperature-sensitive Gal80 to repress the RNAi expression during development, and assessed whether *VGlut* was upregulated. Similarly, we knocked down *tj* in adult Dm12 cells and assessed whether *ChAT* was upregulated. In both cases, we observed partial fate changes, but no upregulation of *VGlut* and *ChAT* (S5E–F).

Finally, in the absence of a functional antibody against Gad1, we performed qPCR to assess the effect of Lim3 knock-down on the expression of Gad1. Gad1 was significantly reduced upon activation of the Lim3 RNAi in adult cells (Figure 5F). We performed the same experiment knocking down *ap* and *tj* and assessing the mRNA of levels of *ChAT* and *VGlut.* We saw downregulation of the terminal genes upon knocking down their predicted regulators (Figure 5F), which verified our previous observations.

These results show that the ‘random forest’ model can predict transcription factors responsible for regulating downstream effectors. The relationships that we tested were causal, as knocking down the predicted transcription factor eliminated the expression of neurotransmitter genes in specific cell types. Interestingly, tampering with the expression of any of these key factors in different cells affected the fate of these cell types, which prevented us from rigorously addressing their sufficiency. Importantly, we found that a single transcription factor does not regulate neurotransmitter identity in all cell types; rather, cell type-specific transcription factors are employed in different cell types.

How do transcription factors generate cell types with specific characteristics?

We then set to identify the transcription factors that regulate the expression pattern of a given neurotransmitter gene in cell types where the highest-scoring transcription factor is not expressed. *CG16779* and *charlatan* were predicted by the ‘random forest’ model to regulate *ChAT*’s expression pattern in neurons where *ap* was not expressed. Similarly, the model predicted three other genes for *VGlut* besides *tj* (*forkhead domain 59A*, *CG32105*, and *CG4328*), and one more gene for *Gad1* besides *Lim3* (*eyeless*). We only found one gene for *Vmat* (*CG33695*) (Figure 6A–A’’’).

We wondered whether the regulation of the same gene by different transcription factors in different cell types is a general phenomenon in the fly optic lobes. Two cell types that share common traits (*e.g.* produce the same neurotransmitter) may either have inherited this trait from a common ancestor, or they may have acquired it independently (Figure 6B). In the first case, the trait would be regulated by the same transcription factor(s) in the two cell types. In the latter case, the trait could be regulated by different (convergence) or by the same transcription factors (deep homology (Tschopp and Tabin, 2017)) (Figure 6B'). We addressed this question by correlating the expression of transcription factors in different clusters with all genes of the same cluster. If a single transcription factor was responsible for the expression of a specific gene, their expression distribution across all clusters should be correlated.

To distinguish between the two models, we selected a gene and asked which transcription factor or combination of transcription factors could best recapitulate its expression pattern. We repeated this for every gene expressed in the optic lobe (9738 genes excluding transcription factors), and asked which model (Figure 6B') was better supported by each of these genes. 3085 genes were better explained by a single activator model, while 6653 genes were better explained by a model where different transcription factors are required to generate its expression pattern (Figure 6C). Since many genes are expressed in one or a few cell types, we wondered whether the 3085 genes of the single activator model corresponded to such genes. Indeed, these genes were expressed on average in 2.2 clusters, while the 6653 genes were expressed in a larger number of clusters (22.2) (Figure 6C'). We also performed Gene Ontology analysis and found that most of the genes related to neural development and differentiation were in the gene pool better explained by the cell type-specific activator model (Table S2).

These results strongly suggest that most neural traits are regulated independently by different combinations of transcription factors in different cell types. They also corroborate our hypothesis regarding the extensive degree of convergence that explains the difference between transcription factor-based and whole transcriptome-based hierarchical clustering trees.

Discussion

In this study, we generated a 'random forest' model that can successfully predict the expression of a gene based on the transcription factor profile of a cell. In parallel, using this model, we identified transcription factors that are responsible for the expression of different effector genes in different cell types. We used RNAi to establish a causal relationship between the expression of the transcription factor and its downstream target. Two of the transcription factors that regulate GABAergic (*Lim3*) and cholinergic (*ap*) identity are LIM homeobox-containing transcription factors, whose interplay has been shown in both vertebrates and invertebrates to regulate neurotransmitter identity (Hobert and Westphal, 2000; Pfaff et al., 1996; Thor et al., 1999; Wenick and Hobert, 2004; Zhang et al., 2014). Interestingly, Lim3 has been associated with either glutamatergic or GABAergic neurons in different systems (Bretzner and Brownstone, 2013; Joshi et al., 2009; Ladewig et al., 2014; Serrano-Saiz et al., 2013; Thor et al., 1999), indicating plasticity, which agrees with our

model of extensive phenotypic convergence (Figure 6B). Our data represent an invaluable resource for studying gene regulatory networks and identifying potential terminal selectors and the effectors they regulate.

We noted that the transcription factors that were found to regulate neurotransmitter identity in adult brains (i.e. Apterous, Traffic-jam, and Lim3) are expressed at the time these neurons were born and maintained throughout their development. More importantly, their expression in newly-born neurons is regulated by temporal factors and by Notch signaling activity. Specifically, *ap* is expressed in $Notch^{ON}$ neurons, while *Lim3* and *tj* are expressed in $Notch^{OFF}$ neurons in four of the five temporal windows, *Hth*, *Ey*, *Slp*, and *D* (Li et al., 2013 and unpublished data). This means that, for the cell types whose neurotransmitter identity relies on these transcription factors, their identity is decided at their birth in a temporal sequence and Notch-dependent manner. We propose that this is the ground plan for the establishment of neurotransmitter identity in different cell types, upon which evolution has acted to generate the phenotypic convergence that we observe today.

Our study also provides evidence regarding the mechanisms by which different transcription factors regulate effector genes in diverse cell types. We show that distinct transcription factors (or combinations) are used in different cell types to drive the expression of effector genes, such as neurotransmitters (Figure 6B'). In *C. elegans,* distinct transcription factor combinations control *VGluT* expression in distinct glutamatergic neurons (Serrano-Saiz et al., 2013). Similarly, cholinergic traits are controlled by distinct transcription factor combinations (Pereira et al., 2015; Wenick and Hobert, 2004; Zhang et al., 2014), and so are GABAergic traits (Gendrel et al., 2016). Here, we present multiple lines of evidence that phenotypic convergence is a more general phenomenon than has been described in worms. The utilization of the same regulatory mode in nematodes and arthropods hints towards a universal strategy for the generation of neuronal diversity that likely also applies to vertebrates.

This is corroborated by the comparison of the hierarchical clustering trees generated using either the whole transcriptome, or transcription factors only. We therefore consider the transcription factor-based tree a better indicator of the developmental or evolutionary history of a neuronal type, although we cannot distinguish between ontogeny and phylogeny (Arendt et al., 2016). In contrast, the whole transcriptome hierarchical clustering is influenced by convergence and delineates functional similarities between adult neurons.

While exploring the data, we made two interesting observations that it will be important to analyze in more detail:

i. Two of our clusters express more than one neurotransmitter – cluster 20 expresses acetylcholine and glutamate, while cluster 54 expresses acetylcholine and GABA. Although there have been indications of neurons co-releasing GABA or glutamate with acetylcholine (Gendrel et al., 2016; Raghu and Borst, 2011; Serrano-Saiz et al., 2017), it is unknown if this is the case in the *Drosophila* optic lobe. An alternative possibility is that these clusters contain more than one cell type that are highly related but differ in their neurotransmitters.

ii. There is a neuronal type (cluster 14) whose whole transcriptome resembles glial cells more than neurons (Figure 1E). Like T1, it expresses *Eaat1*, the glial glutamate transporter. Despite its transcription factor identity, the absence of *repo* expression and the expression of *elav* suggest that this cell type is a neuron. It will be interesting to determine its identity and role in the optic lobe. Given that generation of neurons from glia has been reported several times (Bernardos et al., 2007; Doetsch, 2003; Sammut et al., 2015), one cannot exclude the presence of such a mixed neuronal type in the adult *Drosophila* optic lobe. Moreover, the astrocyte-like glia cluster shows markedly increased expression of *deadpan* (a neuroblast marker), which agrees with the capacity of glial cells to play the role of neural stem cells.

Our data, in combination with single-cell data from the central brain (Croset et al., 2018) or the entire brain (Davie et al., 2018), represent the first fly brain cell atlas, which will be an invaluable resource for future developmental neurobiology studies.

Finally, we generated cell type-specific transcription factor fingerprints. Given that transcription factors may be considered the drivers of cell type evolution (Achim and Arendt, 2014; Arendt, 2008; Arendt et al., 2016), such data will provide a framework for future comparative studies aimed at determining how neuronal diversity has evolved in the optic lobe across the invertebrates (Perry et al., 2017).

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Claude Desplan (cd38@nyu.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All flies used in this study were maintained in fly room at 18–25°C using standard fly husbandry methods. Female flies aged between 3 and 5 days after eclosion were used for bulk RNA sequencing, single-cell sequencing, and antibody stainings after all genetic manipulations. Fly lines used in this study are listed in the Key Resources Table.

METHOD DETAILS

Drop-seq experimental procedure—Drop-seq was performed as previously described and following the protocol of the McCarroll lab (http://mccarrolllab.com/dropseq/).

Drosophila optic lobes were dissected from Canton-S 3-day old females in PBS and were dissociated into single cell suspension by incubating in 2mg/mL collagenase and 2mg/mL dispase in PBS for 1,5 hours at 25°C. The enzymes were then carefully removed and replaced with PBS + 0.1% BSA. The brains are soft but remain intact if pipetted slowly. The brains were pipetted up and down many times (>200–300) until most large chunks of tissue are dissociated. The cells/tissue were kept cold by putting the tubes in ice. The cells were then filtered using 40um cell strainers. The cells were counted immediately before

generating the droplets. The concentration of cells loaded on the microfluidic device was 250 cells/ul.

The lysis buffer was slightly modified to contain 0.2mg/ml Proteinase K. The lysis buffer contained: 0.2M Tris pH 7.5, 20mM EDTA, 50mM DTT, 0.2mg/mL Proteinase K, 6% Ficoll PM-400, and 0.2% Sarkosyl.

The droplets (diameter=125um) were then generated as specified in the Drop-seq protocol and collected in a 50ml falcon tube for a run time of 30 minutes. The emulsions were visualized microscopically in a hemocytometer. 5–10% of the droplets contained a bead and we could observe less than 5% bead-occupied droplets with two beads. After droplet formation, the droplets were incubated in a metal bead bath system at 55°C for 10 min. After incubation, the droplets were broken and the samples were processed as described in the Drop-seq protocol.

The samples from each run (three runs in total) were sequenced in 4 lanes using the Rapid Run mode of HiSeq 2500.

Drop-seq data analysis: filtering, clustering, elimination of over-clustering, and marker selection

Filtering: The Drop-seq data were processed with the standard pipeline available from the McCarroll lab (http://mccarrolllab.com/wp-content/uploads/2016/03/Drop-seqAlignmentCookbookv1.2Jan2016.pdf).

For the downstream analyses, we used Seurat (Satija et al., 2015). To select the minimum gene-per-cell cutoff, we arranged the cell barcodes in decreasing order of size and plotted the cumulative fraction of reads. As we did not observe an inflection point, we used the FACS simulated single cell data to decide the cutoff. We used different cutoffs ranging from 100 to 400 genes/cell and assigned each single cell to clusters. We then measured how many of the single cells were correctly assigned. The highest accuracy (~88%) was achieved when selecting cut-offs of 200 and 300 genes/cell. We selected to keep all cells that had more than 200 genes/cell to maintain the greatest number of single cells. We also filtered cells enriched for mitochondrial gene expression (>20%), which is indicative of stressed cells.

Clustering: We identified 26 significant PCs following the same jackstraw-inspired procedure used in Macosko et al., (2015). These PCs were used as independent “metagenes” to cluster the single cells. For the clustering, we used a k-nearest neighbor algorithm. To decide on the ‘granularity’ of the clustering (resolution parameter), we used the FACS simulated single cell data again and clustered them using different ‘granularities’. We used resolution values ranging from 0.6 to 8, and we observed that when the resolution was between 4 and 8, we were able to recover unique cluster for every cell type. The accuracy of the clustering was higher when using resolution values 4 and 6. For this purpose, we decided to continue with a resolution value of 4.

Elimination of over-clustering: After clustering the single cells, we assessed whether some of the produced clusters were the result of overclustering. We used Seurat to train ‘random

forest' classifiers for each of the terminal nodes and calculated the specificity of the classifier. Nodes with a classifier error higher than 15% were assessed separately for differentially expressed genes, using a likelihood-ratio test for single cell gene expression (McDavid et al., 2013) (Figure S1). In all cases, we were able to recover many highly significantly differentially expressed genes, but we only kept the nodes where the differentially expressed genes were transcription factors, neurotransmitters, cell adhesion molecules, all three of which are indicative of different neuronal types (Figure S1), or well known markers.

Marker selection: To select markers for each of the clusters, we used Seurat to perform a ROC test and calculate the 'classification power' for each individual marker. We selected the ten markers with the highest classification power for each of the clusters. To compare the Drop-seq single cell data and the FACS-sorted cell type RNA sequencing data, we selected the markers that were expressed in both datasets. We constructed the heatmap of Figure 1 and all downstream applications with this shared marker set.

Cluster transcriptome analysis—The single cluster transcriptome was generated by adding the counts of all single cells that constituted this cluster and calculating reads per million for each cluster.

For illustrations of Figures 4 and 6, we used binned data. For this purpose, we binned the expression of each gene according to the distribution of its expression in the 52 clusters and scaled it from 0 to 1. For the binning, we generated a histogram (n=40) for each gene's expression in the 52 clusters and merged the bins that formed "islands" (i.e. they were separated from other "islands" by bins of zero size). We ended up with 2 to 17 bins for each gene that were scaled from 0 to 1 for illustration purposes.

FACS experimental procedure—We identified lines for individual cell types (Key Resources Table) and crossed each Gal4 line to UAS-Red Stinger to label the nuclei of the specific neuronal type. Dissected brains were dissociated and cells expressing the transgene were sorted on the basis of their fluorescent signal using FACS (FACS Aria III). Cells were sorted directly into extraction buffer and we extracted total RNA using the Arcturus PicoPure RNA Isolation Kit (Applied Biosystems). We assessed RNA quality by Bioanalyzer using RNA 6000 Pico chips (Agilent). Smart-Seq v4 Ultra Low Input RNA Kit was used to generate full-length double stranded cDNA with 300 to 500 pg of total RNA input.

Libraries were prepared using Illumina Nextera XT DNA Library Prep Kit and sequenced on the Illumina HiSeq 2500 System. Three barcoded libraries were pooled per sequencing lane and paired-end 100 bp reads were generated.

Three biological replicates were obtained for each cell type.

Simulated single cell generation—To compare shallow Drop-seq generated single cell transcriptomes with deep FACS-sorted cell type bulk RNA sequencing, we generated simulated single cells from the FACS-sorted cell types. We generated 900 single cells for

each cell type (300 from each of the triplicates). To calculate the number of reads for each simulated cell, we picked a random integer from a normal distribution with the same mean and standard deviation as the number of reads of all Drop-seq generated single cells (i.e. mean = 520.8779, sd = 316.0117). When this number was larger than 200, we sampled this amount of reads randomly from all the genes, with the probability of picking each gene defined by its expression level in this cell type in the bulk RNA sequencing. The simulated single cells were then treated exactly like the Drop-seq sequenced single cells, using Seurat to only keep all cells with reads for more than 200 genes.

Antibody stainings—*Drosophila optic* lobes were fixed in 4% formaldehyde for 15 minutes at room temperature. After washing, they were incubated for 2 days with primary antibodies at 4°C. After washing the primary antibody, the brains were incubated with the secondary antibodies overnight at 4°C. The secondary antibodies were washed and the brains were mounted in Slowfade and imaged at a confocal microscope (Leica SP5).

Multi-color flip out clone induction—To generate the multi-color flip out clones, flies carrying the MCFO-1 construct were crossed to ones carrying the specific Gal4 driver. The embryos were raised at 18°C. Once the pupae hatched, the flies were transferred at 37°C for 5 minutes and were again transferred to 18°C. The brains were dissected and stained 2 days after the clone induction.

WGCNA—Using a weighted-gene co-expression network analysis (WGCNA), we performed a single block network construction and module detection from which we obtained 40 gene-network modules. Given the manageable size of the dataset, we were able to analyze all genes in a single block as opposed to a block-wise network construction method, by setting the maxBlockSize parameter to exceed the number of genes.

The first step for building the weighted network consisted in computing an adjacency matrix reporting the strength of connection between each pair of genes. The adjacency a_{ij} of two genes i and j with expression levels x_i and x_j was computed using a power adjacency function given by $a_{ij} = cor(x_i, x_j)$. Upon analysis of network topology for various soft-thresholding powers, we chose β to be the lowest power for which approximate scale free topology is attained (i.e.: the power at which the scale free topology fit index curve flattens out upon reaching a high value).

Following that, genes were clustered by average linkage hierarchical clustering, and modules were identified based on topological overlap measure with a minimum module size of 30 genes.

Among the detected modules, we merged modules with highly correlated module eigengenes (a module eigengene can be considered as an average expression profile of the module). For that, we defined a mergeCutHeight value of 0.25, to merge any two modules whose correlation was greater than 0.75.

RNAi—Flies carrying the hs-flp, actin-flip-out-Gal4, UAS-GFP constructs were crossed to UAS-RNAi lines. The embryos were reared at 18°C. Once the pupae hatched, the flies were

transferred at 37°C for 1 hour and 30 minutes to induce the flip-out and activate Gal4 and were then kept at 25°C for one week. The brains were then dissected and stained.

qPCR—Whole brains were dissected from wild-type or RNAi flies. RNA was isolated using Trizol and first strand cDNA synthesis was performed with the SuperScript III RT kit. The iTaq Universal SYBR Green Supermix was used for the qPCR, which was performed on a Biorad CFX96 Real Time PCR Detection System in biological and technical triplicates.

Antibody generation—A polyclonal antibody against *ChAT* was generated in guinea pigs. The antibody was generated by Genscript (http://www.genscript.com). The epitope used to immunize the guinea pigs were aminoacids 1–480 of the full length protein: MASNEASTSAAGSGPESAALFSKLRSFSIGSGPNSPQRVVSNLRGFLTHRLSNITPSDT GWKDSILSIPKKWLSTAESVDEFGFPDTLPKVPVPALDETMADYIRALEPITTPAQLE RTKELIRQFSAPQGIGARLHQYLLDKREAEDNWAYYYWLNEMYMDIRIPLPINSNPG MVFPPRRFKTVHDVAHFAARLLDGILSHREMLDSGELPLERAASREKNQPLCMAQY YRLLGSCRRPGVKQDSQFLPSRERLNDEDRHVVVICRNQMYCVVLQASDRGKLSES EIASQILYVLSDAPCLPAKPVPVGLLTAEPRSTWARDREMLQEDERNQRNLELIETAQ VVLCLDEPLAGNFNARGFTGATPTVHRAGDRDETNMAHEMIHGGGSEYNSGNRWF DKTMQLIICTDGTWGLCYEHSCSEGIAVVQLLEKIYKKIEEHPDEDNGLPQHHLPPPE RLEWHVGPQLQLRFAQASKSVDK

QUANTIFICATION AND STATISTICAL ANALYSIS

Images were analyzed with FIJI (https://fiji.sc/). All data are expressed as mean ± the standard error of the mean (SEM). ROC tests were used to identify differentially expressed genes between clusters. Pearson correlation on log data was used to test the correlation between FACS-sorted simulated data and Drop-seq clusters, as well as to evaluate the 'random forest' model predictions.

Bioanalyzer—Data was obtained following the standard published protocol: http://www.agilent.com/library/usermanuals/Public/G2938-90014_KitGuideDNA1000Assay_ebook.pdf (Agilent Technologies).

DATA AND SOFTWARE AVAILABILITY

RNA sequencing data has been deposited in NCBI with the following IDs: GEO: GSE103771 and GEO: GSE103772.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are indebted to the fly community for gifts of antibodies and fly stocks, to H. Bellen, Michael Reiser, CH Lee, T.Cook for fly lines, to H. Aberle for the VGlut antibody, and to three anonymous reviewers for constructive feedback on our paper. We are very grateful to A. Butler for help with the Seurat package and A. Powers for help with the Drop-seq setup. We also want to thank the NYUAD sequencing and bioinformatics core. Finally, we thank the Desplan and Satija lab members for critical discussion and comments on the manuscript. This work was supported by grants from the NIH (R01 EY017916) and from the NYUAD institute (G-1205C) to CD, and by NIH

DP2-HG-009623 (New Innovator award) to RS. NK was supported by postdoctoral EMBO (365-2014) and HFSP (LT000122/2015-L) fellowships. This study is dedicated to the memory of our friend and colleague, Jean-Philippe Grossier.

References

Achim K, Arendt D. Structural evolution of cell types by step-wise assembly of cellular modules. Curr Opin Genet Dev. 2014; 27:102–108. [PubMed: 24998387]

Arendt D. The evolution of cell types in animals: emerging principles from molecular studies. Nat Rev Genet. 2008; 9:868–882. [PubMed: 18927580]

Arendt D, Musser JM, Baker CV, Bergman A, Cepko C, Erwin DH, Pavlicev M, Schlosser G, Widder S, Laubichler MD, et al. The origin and evolution of cell types. Nat Rev Genet. 2016; 17:744–757. [PubMed: 27818507]

Bayraktar OA, Doe CQ. Combinatorial temporal patterning in progenitors expands neural diversity. Nature. 2013; 498:449–455. [PubMed: 23783519]

Bernardos RL, Barthel LK, Meyers JR, Raymond PA. Late-stage neuronal progenitors in the retina are radial Muller glia that function as retinal stem cells. J Neurosci. 2007; 27:7028–7040. [PubMed: 17596452]

Bretzner F, Brownstone RM. Lhx3-Chx10 reticulospinal neurons in locomotor circuits. J Neurosci. 2013; 33:14681–14692. [PubMed: 24027269]

Brody T, Odenwald WF. Programmed transformations in neuroblast gene expression during Drosophila CNS lineage development. Dev Biol. 2000; 226:34–44. [PubMed: 10993672]

Busch S, Selcho M, Ito K, Tanimoto H. A map of octopaminergic neurons in the Drosophila brain. J Comp Neurol. 2009; 513:643–667. [PubMed: 19235225]

Cajal SR, Sanchez D. Contribucion al conocimiento de los centros nerviosos de los insectos. Trab Lab Invest Biol. 1915; XIII:1–167.

Cole SH, Carney GE, McClung CA, Willard SS, Taylor BJ, Hirsh J. Two functional but noncomplementing Drosophila tyrosine decarboxylase genes: distinct roles for neural tyramine and octopamine in female fertility. J Biol Chem. 2005; 280:14948–14955. [PubMed: 15691831]

Croset V, Treiber CD, Waddell S. Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. Elife. 2018; 7:e34550. [PubMed: 29671739]

Daubner SC, Le T, Wang S. Tyrosine hydroxylase and regulation of dopamine synthesis. Arch Biochem Biophys. 2011; 508:1–12. [PubMed: 21176768]

Davie K, Janssens J, Koldere D, De Waegeneer M, Pech U, Kreft L, Aibar S, Makhzami S, Christiaens V, Bravo González-Blas C, et al. A single-cell transcriptome atlas of the ageing Drosophila brain. Cell. 2018 (same issue).

DeSalvo MK, Hindle SJ, Rusan ZM, Orng S, Eddison M, Halliwill K, Bainton RJ. The Drosophila surface glia transcriptome: evolutionary conserved blood-brain barrier processes. Front Neurosci. 2014; 8:346. [PubMed: 25426014]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21. [PubMed: 23104886]

Doetsch F. The glial identity of neural stem cells. Nat Neurosci. 2003; 6:1127–1134. [PubMed: 14583753]

Doherty J, Logan MA, Tasdemir OE, Freeman MR. Ensheathing glia function as phagocytes in the adult Drosophila brain. J Neurosci. 2009; 29:4768–4781. [PubMed: 19369546]

Enriquez J, Venkatasubramanian L, Baek M, Peterson M, Aghayeva U, Mann RS. Specification of individual adult motor neuron morphologies by combinatorial transcription factor codes. Neuron. 2015; 86:955–970. [PubMed: 25959734]

Erclik T, Li X, Courgeon M, Bertet C, Chen Z, Baumert R, Ng J, Koo C, Arain U, Behnia R, et al. Integration of temporal and spatial patterning generates neural diversity. Nature. 2017; 541:365–370. [PubMed: 28077877]

Fischbach KF, Dittrich AP. The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. Cell Tissue Res. 1989; 258:441–445.

Gendrel M, Atlas EG, Hobert O. A cellular and regulatory map of the GABAergic nervous system of C. elegans. Elife. 2016; 5

Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. Nature. 1999; 402:C47–52. [PubMed: 10591225]

Hobert O. Regulation of terminal differentiation programs in the nervous system. Annu Rev Cell Dev Biol. 2011; 27:681–696. [PubMed: 21985672]

Hobert O. A map of terminal regulators of neuronal identity in Caenorhabditis elegans. Wiley Interdiscip Rev Dev Biol. 2016; 5:474–498. [PubMed: 27136279]

Hobert O, Westphal H. Functions of LIM-homeobox genes. Trends Genet. 2000; 16:75–83. [PubMed: 10652534]

Jenett A, Rubin GM, Ngo TT, Shepherd D, Murphy C, Dionne H, Pfeiffer BD, Cavallaro A, Hall D, Jeter J, et al. A GAL4-driver line resource for Drosophila neurobiology. Cell Rep. 2012; 2:991–1001. [PubMed: 23063364]

Joshi K, Lee S, Lee B, Lee JW, Lee SK. LMO4 controls the balance between excitatory and inhibitory spinal V2 interneurons. Neuron. 2009; 61:839–851. [PubMed: 19323994]

Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, Kocks C, Rajewsky N, Zinzen RP. The Drosophila embryo at single cell transcriptome resolution. Science. 2017

Kepecs A, Fishell G. Interneuron cell types are fit to function. Nature. 2014; 505:318–326. [PubMed: 24429630]

Kratsios P, Pinan-Lucarre B, Kerk SY, Weinreb A, Bessereau JL, Hobert O. Transcriptional coordination of synaptogenesis and neurotransmitter signaling. Curr Biol. 2015; 25:1282–1295. [PubMed: 25913400]

Kratsios P, Stolfi A, Levine M, Hobert O. Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. Nat Neurosci. 2011; 15:205–214. [PubMed: 22119902]

Kremer MC, Jung C, Batelli S, Rubin GM, Gaul U. The glia of the adult Drosophila nervous system. Glia. 2017; 65:606–638. [PubMed: 28133822]

Ladewig J, Koch P, Brustle O. Auto-attraction of neural precursors and their neuronal progeny impairs neuronal migration. Nat Neurosci. 2014; 17:24–26. [PubMed: 24241396]

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. [PubMed: 19114008]

Li X, Erclik T, Bertet C, Chen Z, Voutev R, Venkatesh S, Morante J, Celik A, Desplan C. Temporal patterning of Drosophila medulla neuroblasts controls neural fates. Nature. 2013; 498:456–462. [PubMed: 23783517]

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015; 161:1202–1214. [PubMed: 26000488]

McDavid A, Finak G, Chattopadyay PK, Dominguez M, Lamoreaux L, Ma SS, Roederer M, Gottardo R. Data exploration, quality control and testing in single cell qPCR-based gene expression experiments. Bioinformatics. 2013; 29:461–467. [PubMed: 23267174]

Morante J, Desplan C. The color-vision circuit in the medulla of Drosophila. Curr Biol. 2008; 18:553–565. [PubMed: 18403201]

Nassel DR, Elekes K. Aminergic neurons in the brain of blowflies and Drosophila: dopamine- and tyrosine hydroxylase-immunoreactive neurons and their relationship with putative histaminergic neurons. Cell Tissue Res. 1992; 267:147–167. [PubMed: 1346506]

Nern A, Pfeiffer BD, Rubin GM. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. Proc Natl Acad Sci U S A. 2015; 112:E2967–2976. [PubMed: 25964354]

Pereira L, Kratsios P, Serrano-Saiz E, Sheftel H, Mayo AE, Hall DH, White JG, LeBoeuf B, Garcia LR, Alon U, et al. A cellular and regulatory map of the cholinergic nervous system of C. elegans. Elife. 2015; 4

Perry M, Konstantinides N, Pinto-Teixeira F, Desplan C. Generation and Evolution of Neural Cell Types and Circuits: Insights from the Drosophila Visual System. Annu Rev Genet. 2017

Pfaff SL, Mendelsohn M, Stewart CL, Edlund T, Jessell TM. Requirement for LIM homeobox gene Isl1 in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. Cell. 1996; 84:309–320. [PubMed: 8565076]

Pinto-Teixeira F, Konstantinides N, Desplan C. Programmed cell death acts at different stages of Drosophila neurodevelopment to shape the central nervous system. FEBS Lett. 2016; 590:2435–2453. [PubMed: 27404003]

Poulin JF, Tasic B, Hjerling-Leffler J, Trimarchi JM, Awatramani R. Disentangling neural cell diversity using single cell transcriptomics. Nat Neurosci. 2016; 19:1131–1141. [PubMed: 27571192]

Raghu SV, Borst A. Candidate glutamatergic neurons in the visual system of Drosophila. PLoS One. 2011; 6:e19472. [PubMed: 21573163]

Richardt A, Rybak J, Stortkuhl KF, Meinertzhagen IA, Hovemann BT. Ebony protein in the Drosophila nervous system: optic neuropile expression in glial cells. J Comp Neurol. 2002; 452:93–102. [PubMed: 12205712]

Sammut M, Cook SJ, Nguyen KC, Felton T, Hall DH, Emmons SW, Poole RJ, Barrios A. Glia-derived neurons are required for sex-specific learning in C. elegans. Nature. 2015; 526:385–390. [PubMed: 26469050]

Santiago C, Bashaw GJ. Transcription factors and effectors that regulate neuronal morphology. Development. 2014; 141:4667–4680. [PubMed: 25468936]

Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single cell gene expression data. Nat Biotechnol. 2015; 33:495–502. [PubMed: 25867923]

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. Nature methods. 2012; 9:676–682. [PubMed: 22743772]

Serrano-Saiz E, Pereira L, Gendrel M, Aghayeva U, Battacharya A, Howell K, Garcia LR, Hobert O. A Neurotransmitter Atlas of the Caenorhabditis elegans Male Nervous System Reveals Sexually Dimorphic Neurotransmitter Usage. Genetics. 2017; 206:1251–1269. [PubMed: 28684604]

Serrano-Saiz E, Poole RJ, Felton T, Zhang F, De La Cruz ED, Hobert O. Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins. Cell. 2013; 155:659–673. [PubMed: 24243022]

Shekhar K, Lapan SW, Whitney IE, Tran NM, Macosko EZ, Kowalczyk M, Adiconis X, Levin JZ, Nemesh J, Goldman M, et al. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. Cell. 2016; 166:1308–1323 e1330. [PubMed: 27565351]

Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 2006; 22:1540–1542. [PubMed: 16595560]

Suzuki T, Kaido M, Takayama R, Sato M. A temporal mechanism that produces neuronal diversity in the Drosophila visual center. Dev Biol. 2013; 380:12–24. [PubMed: 23665475]

Thor S, Andersson SG, Tomlinson A, Thomas JB. A LIM-homeodomain combinatorial code for motor-neuron pathway selection. Nature. 1999; 397:76–80. [PubMed: 9892357]

Tschopp P, Tabin CJ. Deep homology in the age of next-generation sequencing. Philos Trans R Soc Lond B Biol Sci. 2017; 372

Wenick AS, Hobert O. Genomic cis-regulatory architecture and transacting regulators of a single interneuron-specific gene battery in C. elegans. Dev Cell. 2004; 6:757–770. [PubMed: 15177025]

Wichterle H, Gifford D, Mazzoni E. Neuroscience. Mapping neuronal diversity one cell at a time. Science. 2013; 341:726–727. [PubMed: 23950522]

Zhang F, Bhattacharya A, Nelson JC, Abe N, Gordon P, Lloret-Fernandez C, Maicas M, Flames N, Mann RS, Colon-Ramos DA, et al. The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types. Development. 2014; 141:422–435. [PubMed: 24353061]

Highlights

- *Drosophila* glia and neurons were profiled with Drop-seq and pooled in 52 clusters
- A random forest model identified transcription factors that regulate terminal genes
- We established causal relations between transcription factors and neurotransmitters
- Distinct transcription factors regulate the same effector genes in different cells

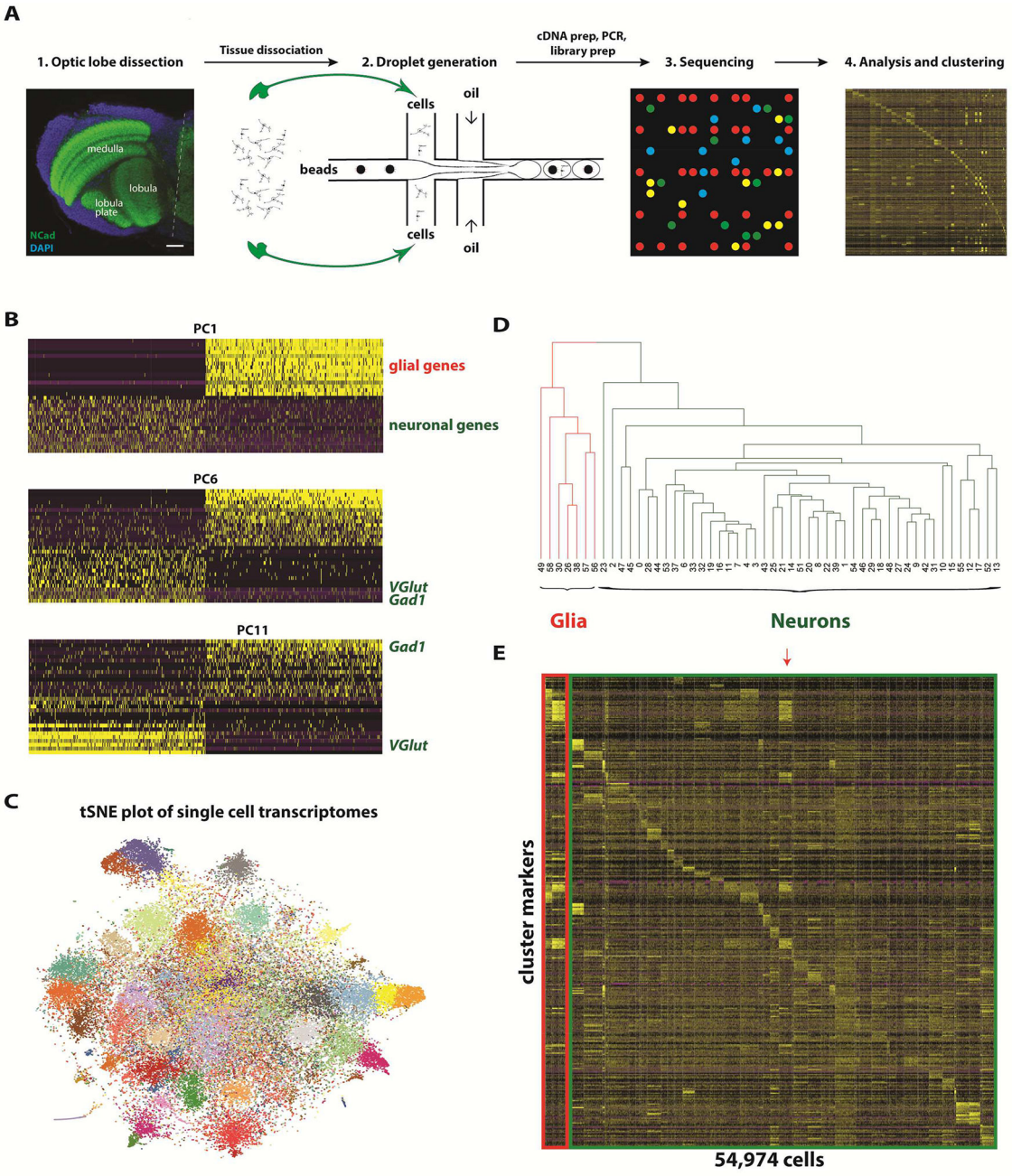


Figure 1. Drop-seq experimental procedure, analysis, and clustering, see also Figure S1 and Table S1

(A) We dissected the optic lobes of the *Drosophila* central nervous system and dissociated them into single cells. The cell bodies can be seen by the DAPI staining in the cortex and rim of the three neuropils (visualized using an antibody against *NCad*). The single cells were then fed into the microfluidic device, alongside the beads (which were in lysis buffer) and the oil, in a setup that resulted in the generation of aqueous droplets in an oil background. Each droplet may be empty, carrying a bead and a single cell, or carrying one of the two. After lysis, transcript annealing, droplet breakage, cDNA preparation and PCR

amplification, we sequenced the pooled single cell transcriptomes, analyzed the results using the Seurat package in R and clustered the single cells in 52 clusters. Scale bar, 20um.
(B) We performed PCA to reduce the dimensions of the data for further analysis. Genes (rows) and cells (columns) are ordered according to their PCA scores and the 500 most extreme cells and 30 most extreme genes on both sides of the distribution are shown in the heatmap. The first PCs (as indicated here by PC1) were responsible for the separation of neurons from glia, as indicated by the positive contribution of glial genes (such as *nrv2*, *Inx2*, *alrm*, and *ogre*) in PC1 and the opposite one for genes enriched in neurons (*VGlut* and nicotinic acetylcholine receptors). Later PCs divide the neurons based on their neurotransmitter identity, as can be seen for PC6 (glutamatergic and GABAergic neurons are separated from the rest, mainly cholinergic ones) and PC11 (glutamatergic and GABAergic neurons are separated from each other).
(C) The tSNE plot of all single cells included in our analysis shows the separation of different clusters. We used a k-nearest neighbor algorithm to call 61 clusters, which are shown in different colors on the tSNE plot.
(D) Transcription factor-based hierarchical clustering of the Drop-seq cluster transcriptomes. Clusters are numbered from 0 to 58. The first split of the tree represents the separation of 7 glial clusters (red) from 45 neuronal ones (blue), as expected from the PCA. Numbers at the bottom of the tree indicate clusters.
(E) The expression of 401 selected Drop-seq cluster markers (rows) is shown in all Drop-seq single cells (columns) (see Table S1). Clusters are separated by white lines and are arranged according to the tree in Figure 1D. Glial clusters are highlighted in red, while neuronal clusters are in blue. Interestingly, a single neuronal cluster that expresses elav but not repo (cluster 14 – red arrow) shares many common markers with glia (see Discussion).

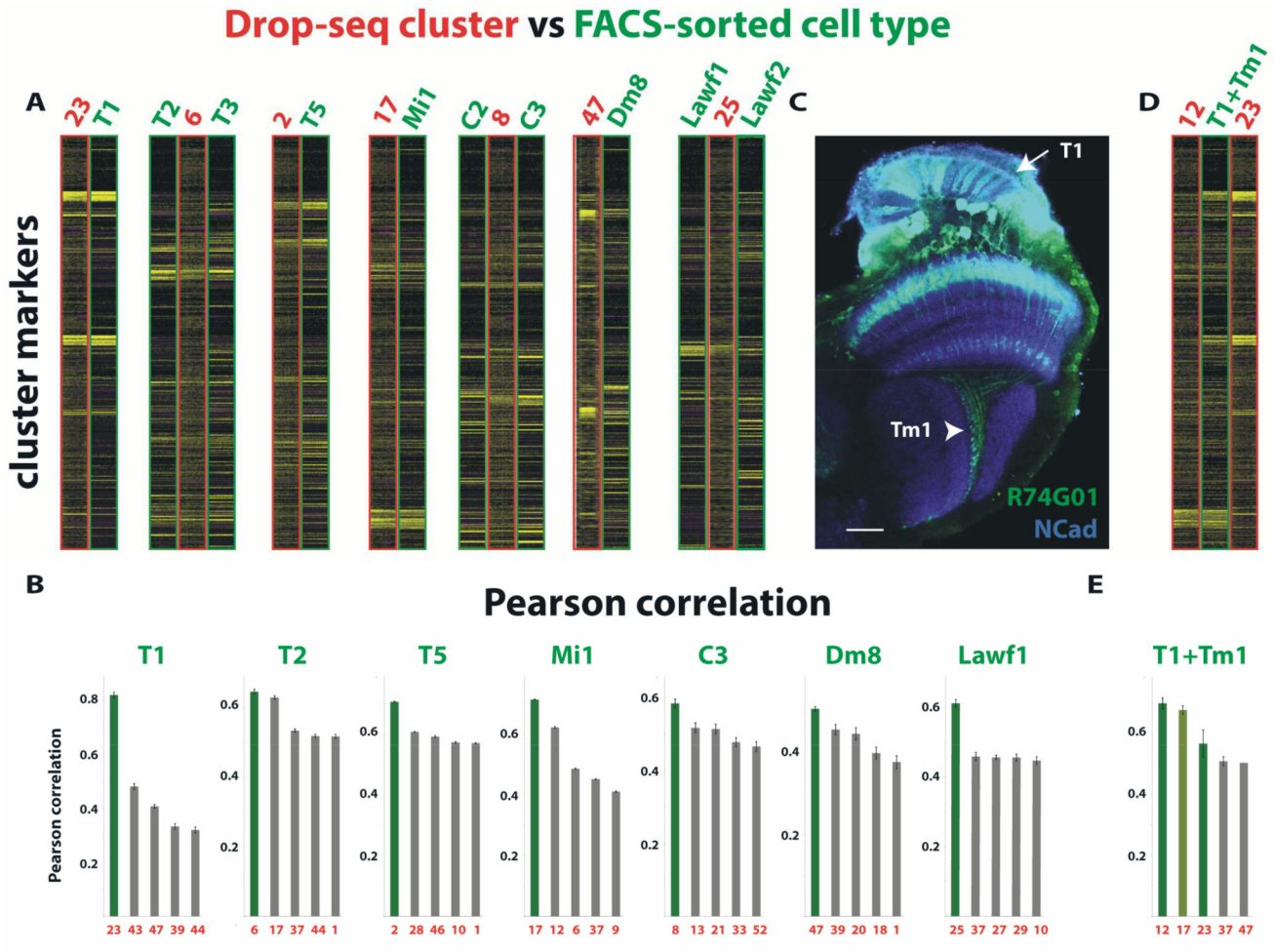


Figure 2. Comparison of Drop-seq cluster transcriptomes and FACS-sorted cell type transcriptomes shows striking similarities between certain clusters and cell types and is used to annotate the Drop-seq clusters, see also Figure S2

(A) The expression levels of 401 selected Drop-seq cluster markers (rows) is shown for "simulated single cells" (columns) representative of FACS-sorted cell types (green – cell type name indicated on top) and for single cells of the respective Drop-seq clusters (red – cluster number indicated on top) (see Table S1). Each FACS-sorted cell type corresponds clearly to one Drop-seq cluster.

(B) Histograms showing the Pearson correlation of the transcriptome of each FACS-sorted cell type with the transcriptome of the more correlated clusters. Most of the cell types map to one cluster. Error bars represent standard error of the mean of the triplicates' Pearson correlation with the more related clusters.

(C) *R74G01-Gal4>UAS-myrGFP* is expressed in two cell types: Tm1, whose projections in the lobula are indicated by the arrowhead, and T1, whose projections in the lamina are marked by the arrow. *NCad* is used to visualize the neuropils. Scale bar, 20um.

(D) The heatmap shows the expression levels of the 401 selected markers (rows) in "simulated single cells" (columns) that represent R74G01-Gal4 and in single cells of the respective Drop-seq clusters, 12 and 23. Since R74G01-Gal4 is expressed in two different cell types, T1 and Tm1, its transcriptome matches two Drop-seq clusters.

(E) Histogram showing the Pearson correlation of the T1 and Tm1 mixed population transcriptome with the more correlated Drop-seq clusters. It maps to two clusters, 12 and 23.

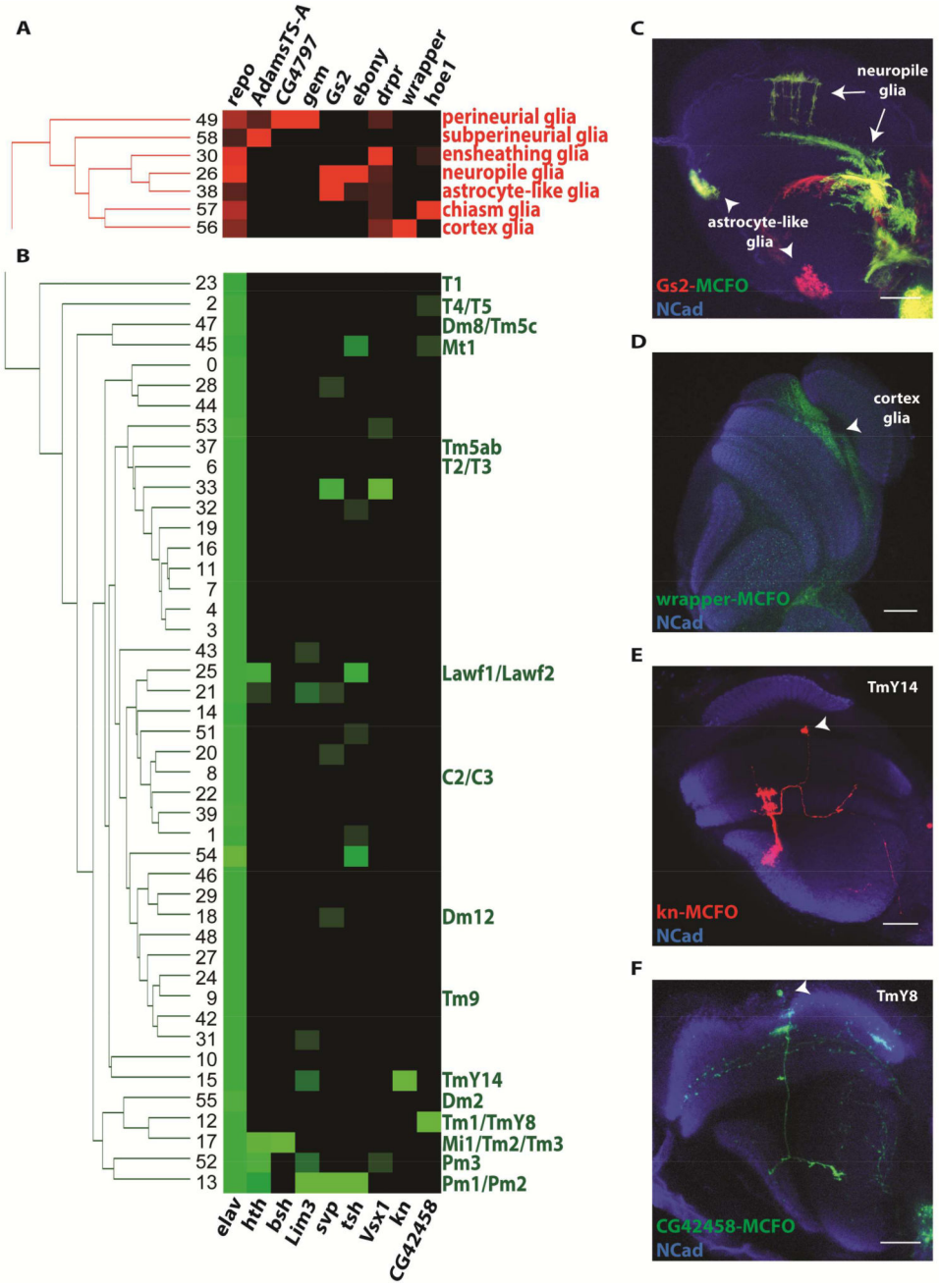


Figure 3. Annotation of glial (red) and neuronal (green) clusters, see also Figure S3

(A) Annotation of all glial clusters using glial markers. *Repo* is expressed in all glial clusters. *AdamsTS-A* is expressed in perineurial and subperineurial glia, while *CG4797* and *gemini* are only expressed in perineurial glia. *Gs2* is expressed in astrocyte-like glia and neuropile glia (see also Figure 3C). *Drpr* is mainly expressed in phagocytic ensheathing glia. *Wrapper* is only expressed in cortex glia (see also Figure 3D) and *hoe1* mainly in chiasm glia (see also Figure S3C)

(B) Annotation of neuronal clusters using three different techniques: 1) based on their correspondence to the FACS-sorted cell type transcriptomes (see also Figure 2), 2) based on

known markers (Mi1 expresses *bsh* and Pm1, Pm2, and Pm3 express *Lim3*, Pm1 and Pm2 express *svp*, Pm1 expresses *tsh*, and Pm3 expresses *Vsx1*), 3) based on newly identified markers (*kn* is expressed in cluster 15 and corresponds to TmY14, and *CG42458* is mainly expressed in cluster 12 and corresponds to TmY8).

(C) A swapped MIMIC line expressing Gal4 in the pattern of *Gs2* was used to drive MCFO (Nern et al., 2015). Single cell clones were generated in the adult brain and are shown in red and green. *Gs2* is expressed in neuropile glia (arrow) and astrocyte-like glia (arrowhead).

(D) A *wrapper*-Gal4 line was used to drive MCFO (Nern et al., 2015). Single cell clones were generated in the adult brain and are shown in green. *Wrapper* is expressed in cortex glia.

(E) A swapped MIMIC line expressing Gal4 in the pattern of *kn* was used to drive MCFO (Nern et al., 2015). Single cell clones were generated in the adult brain and are shown in red. *Kn* is expressed in TmY14.

(F) A swapped MIMIC line expressing Gal4 in the pattern of *CG42458* was used to drive MCFO (Nern et al., 2015). Single cell clones were generated in the adult brain and are shown in green. *CG42458* is expressed in TmY8 (arrowhead).

NCad labels the neuropils in C–F. Scale bar, 20um.

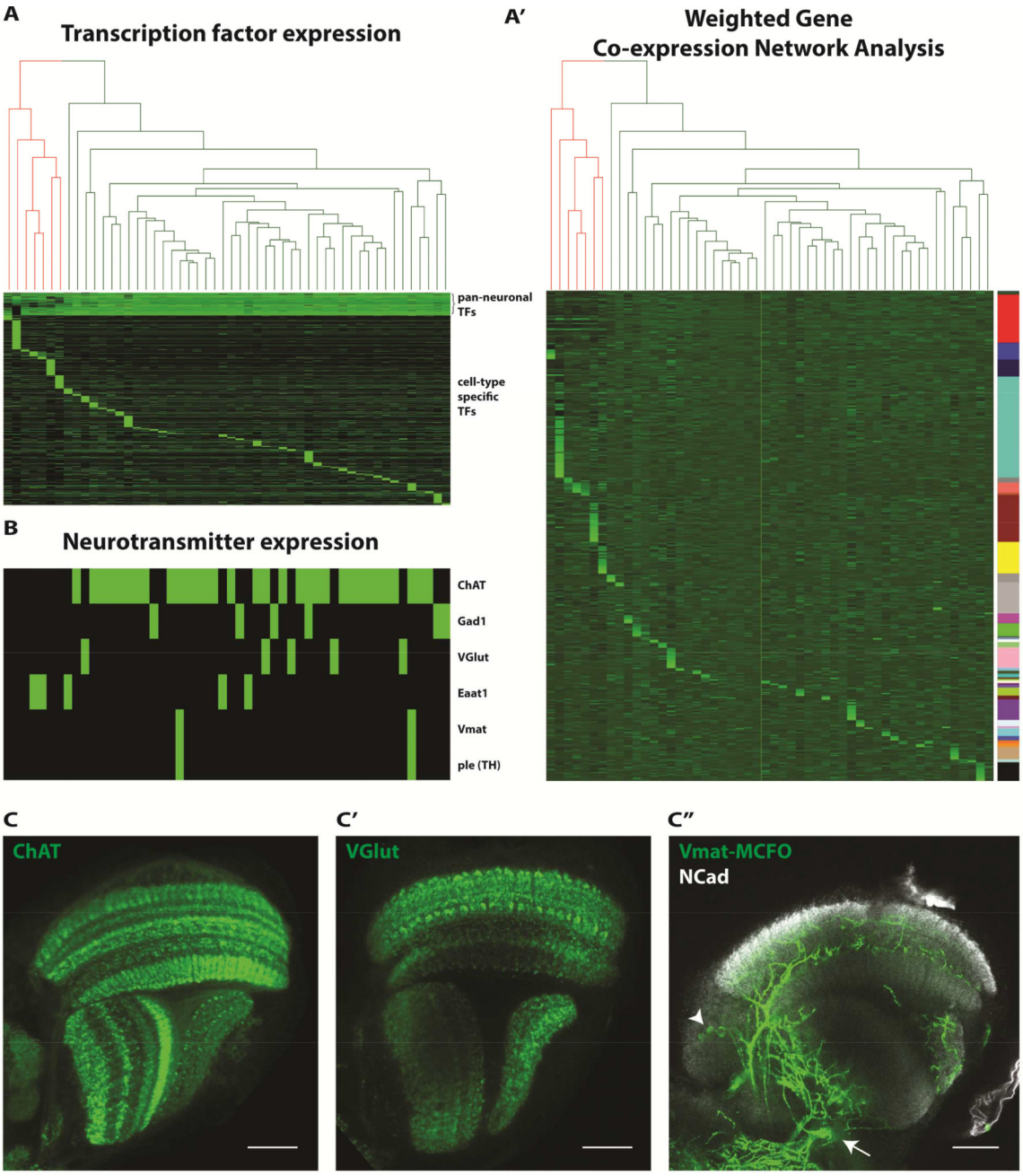
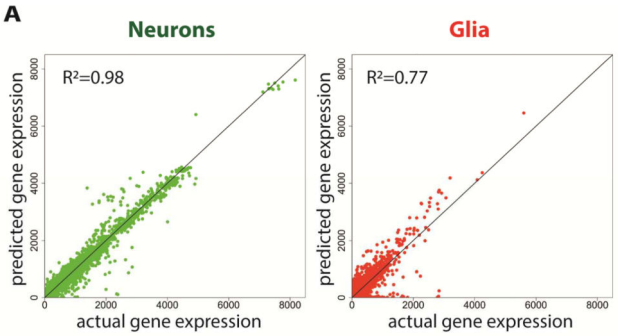


Figure 4. Transcription factor and neurotransmitter expression in the Drop-seq clusters, see also Figure S4

(A) Heatmap showing the expression of transcription factors (rows) in all Drop-seq clusters (columns) – green in different intensities indicates expression in different levels, black corresponds to no expression. The transcription factors that are expressed in the adult optic lobe neurons and glia can be separated into two categories: 72 ubiquitous/pan-neuronal transcription factors are found at similar levels in most cell types, while 598 cell type-specific transcription factors are expressed in significantly higher levels in only one or few cell types.

(A') Heatmap showing the expression of transcription factors (rows) in all Drop-seq clusters (columns). Transcription factors are organized in modules (color-coded on the right), which were defined by weighted-gene co-expression network analysis (WGCNA). We observe that each module of transcription factors is mainly expressed in a single cluster, indicating that similar cell types have different compositions of transcription factors.
(B) Heatmap showing the expression of neurotransmitter related genes (rows) in all Drop-seq clusters (columns) – green indicates expression, black corresponds to no expression. *ChAT* is expressed in cholinergic neurons, *Gad1* in GABAergic, *VGlut* in glutamatergic, *Eaat1* is an excitatory aminoacid transporter that is used to uptake glutamate, *Vmat* marks aminergic neurons and *ple* (tyrosine hydroxylase) is expressed in dopaminergic neurons. Most of the neurons in the optic lobe are cholinergic.
(C–C') Antibody staining against *ChAT* and *VGlut* showing the presence of cholinergic and glutamatergic neurons in the *Drosophila* optic lobe.
(C'') A *Vmat*-Gal4 line was used to drive MCFO (Nern et al., 2015). Single cell clones were generated in the adult brain and are shown in green. Two aminergic neuronal cell types can be seen with their cell bodies in the medulla rim (arrowhead) and the lobula cortex (arrow). Scale bar, 20um.

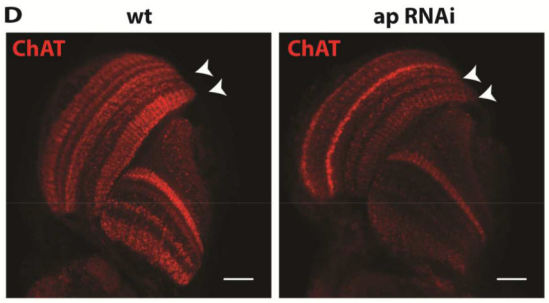


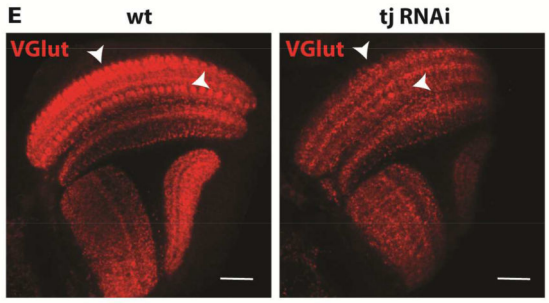
B

Cluster	Pearson correlation between predicted and actual transcriptome
2	0.975
3	0.975
18	0.972
6	0.970
8	0.969
7	0.968
16	0.965
0	0.965
45	0.931
42	0.928
51	0.878
53	0.849
58	0.702

C

Phenotype	Gene tested	Top hit transcription factors
cholinergic	ChAT	Apterous
glutamatergic	VGlut	Traffic-jam
GABAergic	Gad1	Lim3
dopaminergic	Vmat	CG33695





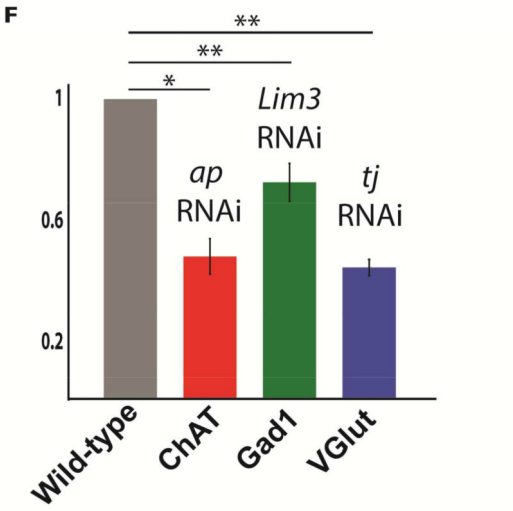


Figure 5. A 'random forest' model identifies transcription factors that regulate terminal genes involved in neurotransmitter expression, see also Figure S5

We trained a 'random forest' model using 39 clusters as a training set and 13 clusters as a test set.

(A) The generated model can faithfully predict the expression of all genes in the test clusters given the transcription factor expression. The accuracy of the prediction was 98% for the neuronal clusters and 77% for the glial clusters (mainly due to the fewer clusters that led to incomplete training of the model).

(B) The Pearson correlation between the predicted and the actual transcriptome ranged from 84.9% to 97.5% in the neuronal clusters and was 70.2% in the glial cluster.

(C) Using the 'random forest' model, we identified the transcription factors that are mainly responsible for the generation of each of the four neurotransmitter identities: cholinergic (*apterous*), glutamatergic (*traffic-jam*), GABAergic (*Lim3*), and monoaminergic (*CG33695*).
(D) The expression of *ChAT* was predicted to rely on the expression of *ap* in a subset of the clusters. Knock-down of *ap* in the adult optic lobe led to the downregulation of *ChAT* in specific medulla layers, M6 and M10, as well as in the lobula.
(E) Effect of *tj* knock-down in the expression of *VGlut*. The expression of *VGlut* in synaptic boutons in medulla layers M1 and M6 is reduced upon downregulation of *tj.*
(F) The expression of *ChAT, Gad1,* and *VGlut* is predicted to be regulated by Ap, Lim3, and Tj, respectively. Quantitative-PCR for the mRNA of *ChAT, Gad1,* and *VGlut* shows that the genes encoding these transcription factors are significantly downregulated upon activation of RNAi against their respective predicted regulators. Data are represented as mean ± SEM. Scale bar, 20um.

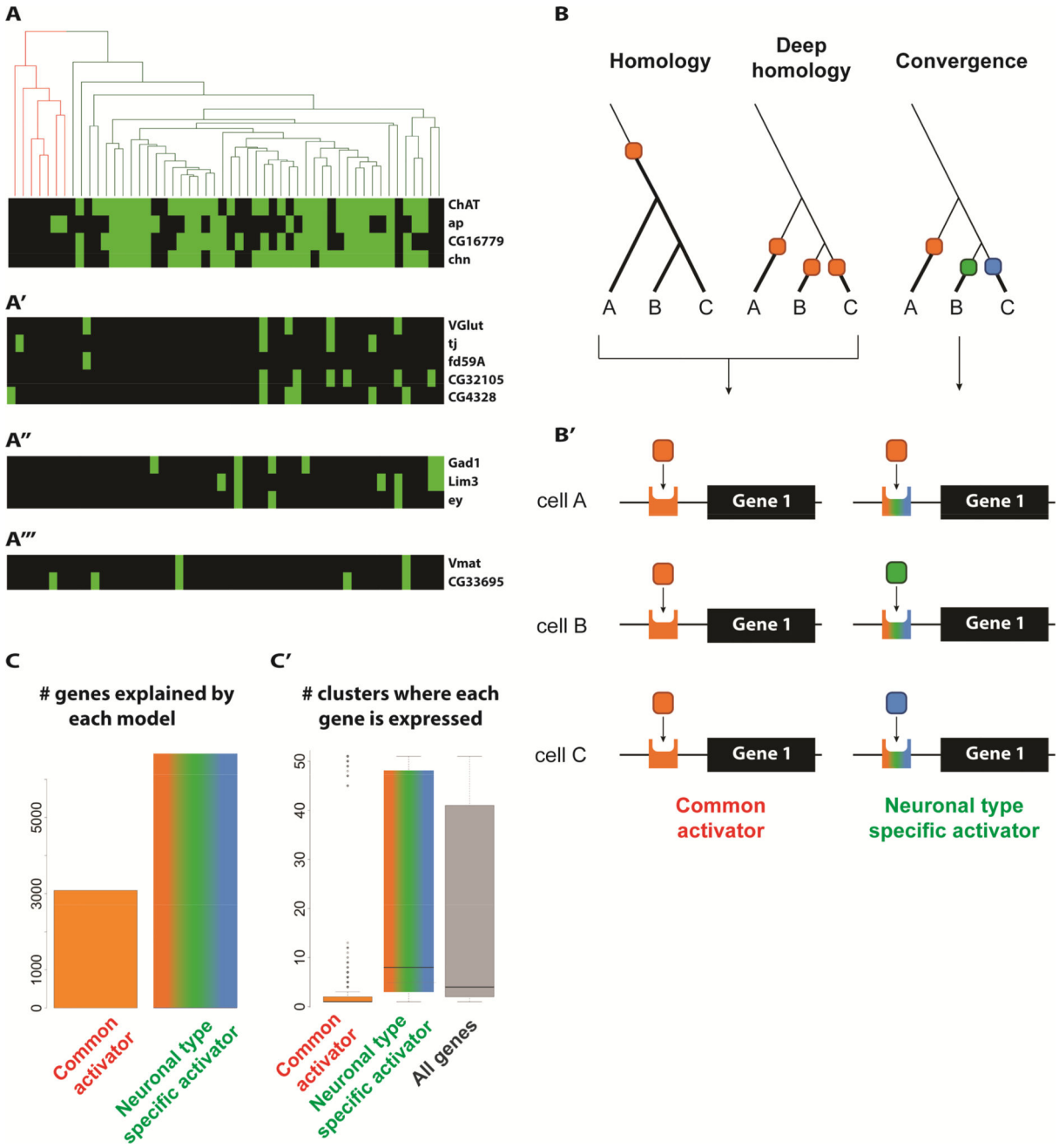


Figure 6. Cell type-specific transcription factors regulate terminal genes in different cells, see also Table S2

(A) We used the 'random forest' model to identify the top transcription factors that better correlated with the expression of effector genes. *ChAT* expression correlated best with the expression of three transcription factors, *ap, chn,* and *CG16779*.

(A') Similarly, the expression pattern of *VGlut* may be generated by the combination of four different transcription factors (*traffic-jam*, *fd59A*, *CG32105*, and *CG4328*).

(A'') The candidates for generating *Gad1* expression pattern are *Lim3* and *ey*.

(A''') One transcription factor, *CG33695*, was needed to explain the expression of *Vmat*.

(B) Two extant cell types that express an effector gene (indicated by the bold line) may either share a common ancestor that was expressing this gene or they may have independently evolved the capacity to express it. In the latter case, the expression of this gene may rely on the same or different transcription factor.
(B') As a consequence of the evolutionary history of the effector gene, its expression in different cell types of the optic lobe may either rely on the same transcription factor (single activator) or different transcription factors may regulate its expression in different cell types (neuronal type-specific activator).
(C) Out of the 9738 genes that are expressed in the adult *Drosophila* optic lobe (excluding transcription factors), 3085 genes support a single activator model, while 6653 genes are better explained by a cell type-specific activator model.
(C') The genes that supported the single activator model are expressed in few clusters (2.2 clusters on average), while the 6653 genes explained by the cell type-specific activator model covered a larger range of clusters (22.2 on average).

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rat monoclonal anti-DN-Cad	DSHB	DN-Ex #8
sheep polyclonal anti-GFP	Bio-Rad	Cat. No 4745-1051
rabbit polyclonal anti-HA	Cell Signaling Techonologies	Cat. No 3724S
rat monoclonal anti-FLAG	Novus Bio	Cat. No NBP1-06712SS
mouse polyclonal anti-V5:DyLight 550	AbD Serotec	Cat. No MCA1360D550GA
guinea pig polyclonal anti-ChAT	this study	N/A
rabbit polyclonal anti-Vglut	(Mahr and Aberle, 2006)	N/A
donkey anti-sheep Alexa 488	Jackson ImmunoResearch	Cat. No 713-545-147
donkey anti-rabbit Alexa 488	Jackson ImmunoResearch	Cat. No 711-545-152
donkey anti-rabbit Alexa 594	Jackson ImmunoResearch	Cat. No 711-585-152
donkey anti-guinea pig Alexa 488	Jackson ImmunoResearch	Cat. No 706-545-148
donkey anti-guinea pig Alexa 594	Jackson ImmunoResearch	Cat. No 706-585-148
donkey anti-rat Alexa 647	Jackson ImmunoResearch	Cat. No 712-605-153
Chemicals, Peptides, and Recombinant Proteins		
DAPI	Sigma	Cat. No 10236276001
Droplet generation oil	Bio-Rad	Cat. No 186-4006
Ficoll PM-400	GE Healthcare	Cat. No 17-0300-10
Sarkosyl	Sigma	Cat. No L7414
Perfluorooctanol	Sigma	Cat. No 370533
TRIzol reagent	Thermo Fisher Scientific	Cat. No 15596026
Critical Commercial Assays		
Maxima H Minus Reverse Transcriptase	Thermo Fisher Scientific	Cat No. EP0753
ExoI buffer	New England Biolabs	Cat No. B0293S
ExoI enzyme	New England Biolabs	Cat No. M0293L
Agencourt AMPure XP - PCR Purification	Beckman Coulter	Cat. No. A63880
Agilent High Sensitivity DNA Kit	Agilent Technologies	Cat. No 5067-4626
Nextera kit	Illumina	Cat. No. FC-131-1096
iTaq Universal SYBR Green Supermix	Biorad	Cat No. 1725121
cDNA Synthesis with SuperScript® III RT	Thermo Fisher Scientific	Cat. No 11752050
Deposited Data		
Single-cell RNA sequencing of Drosophila melanogaster optic lobe cells (Raw and analyzed data)	This paper	GEO: GSE103771
RNA sequencing of Drosophila melanogaster optic lobe cell types (Raw and analyzed data)	This paper	GEO: GSE103772
Experimental Models: Organisms/Strains		
Fly: D. melanogaster: Canton S	Bloomington	stock no. 64349
Fly: D. melanogaster: R74G01-Gal4	Bloomington	stock no. 39868
Fly: D. melanogaster: MCFO-1	Bloomington	stock no. 64085

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Fly: D. melanogaster: Gs2-Gal4	gift by Hugo Bellen	N/A
Fly: D. melanogaster: wrapper-Gal4	Bloomington	stock no. 45784
Fly: D. melanogaster: kn-Gal4	gift by Hugo Bellen	N/A
Fly: D. melanogaster: CG42458-Gal4	gift by Hugo Bellen	N/A
Fly: D. melanogaster: gem-Gal4	gift by Hugo Bellen	N/A
Fly: D. melanogaster: Eaat1-Gal4	Bloomington	stock no. 8849
Fly: D. melanogaster: hoe1-Gal4	this study	N/A
Fly: D. melanogaster: Vmat-Gal4	gift by Hugo Bellen	N/A
Fly: D. melanogaster: UAS-Ap RNAi	Bloomington	stock no. 26748
Fly: D. melanogaster: UAS-Tj RNAi	Bloomington	stock no. 51506
Fly: D. melanogaster: Mi1-Gal4	Bloomington	stock no. 48852
Fly: D. melanogaster: Tm2-Gal4	gift by T. Cook	otd-Gal4
Fly: D. melanogaster: Tm3-Gal4	Bloomington	stock no. 48569
Fly: D. melanogaster: Tm5ab-Gal4	(Melnattur et al, 2014)	OrtC1aDBD#3; dvP16AD-24g
Fly: D. melanogaster: Tm5c-Gal4	(Melnattur et al, 2014)	OrtC1a-Gal4DBD#3 OK371AD
Fly: D. melanogaster: Tm9-Gal4	Bloomington	stock no. 48050
Fly: D. melanogaster: T1-Gal4	Bloomington	stock no. 49685
Fly: D. melanogaster: T2-Gal4	Bloomington	stock no. 40519
Fly: D. melanogaster: T3-Gal4	Bloomington	stock no. 49484
Fly: D. melanogaster: T5-Gal4	Bloomington	stock no. 50172
Fly: D. melanogaster: C2-Gal4	gift by M. Reiser	25B02AD; 48D11DBD
Fly: D. melanogaster: C3-Gal4	gift by M. Reiser	26H02AD ; 29G11DBD
Fly: D. melanogaster: Lawf1-Gal4	gift by M. Reiser	R52H01AD; R17C11DBD
Fly: D. melanogaster: Lawf2-Gal4	gift by M. Reiser	R11D03AD; R61H02DBD
Fly: D. melanogaster: Dm2-Gal4	Bloomington	stock no. 49204
Fly: D. melanogaster: Dm8-Gal4	gift by CH Lee	ort C2b-GAL4
Fly: D. melanogaster: Dm12-Gal4	Bloomington	stock no. 50328
Fly: D. melanogaster: UAS-RedStinger	Bloomington	stock no. 8547
Fly: D. melanogaster: LexA-2xhrGFP.nls	Bloomington	stock no. 29955
Fly: D. melanogaster: UAS-tj	FlyORF	stock no. F000221
Fly: D. melanogaster: UAS-ap	Bloomington	stock no. 42222
Fly: D. melanogaster: UAS-Lim3 RNAi	Bloomington	stock no. 26227
Oligonucleotides		
Act42A_qFor: AAGTGTGTGCAGCGGATAACT	this study	N/A
Act42A_qRev: AAAGCTGCAACCTCTTCGTC	this study	N/A
EF1a100E_qFor: GCAGCGTTGCCGAGTAATA	this study	N/A
EF1a100E_qRev: ATCTTCTCCTTGCCCATCCT	this study	N/A
ChATqFor: CCGAGTCTGTGGACGAGTTT	this study	N/A
ChATqRev: ATAGTCGGCCATCGTTTCAT	this study	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Gad1qFor: AAATGTCGCTGAATCCCAAC	this study	N/A
Gad1qRev: GTCACTGTTGTGGGCATGAG	this study	N/A
VGlutqFor: CATGTGGTGATTTGCGTGA	this study	N/A
VGlutqRev: CCAGAAACGCCAGATACCAT	this study	N/A
Software and Algorithms		
R Statistical Computing Software version 3.3.2	N/A	http://r-project.org
Seurat version 1.4	(Satija et al., 2015)	http://satijalab.org/seurat/
Picard tools	N/A	http://broadinstitute.github.io/picard/
STAR	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
FIJI version 2.0.0	(Schindelin et al., 2012)	https://fiji.sc
PANTHER Overrepresentation Test (release 20170413)	Gene Ontology Consortium	http://geneontology.org