Title

Synaptic specificity determinants identification by converging transcriptome and connectome.

Author

Yoo, Juyoun

Publication Date

2023

UNIVERSITY OF CALIFORNIA

Los Angeles

Synaptic specificity determinants identification by converging transcriptome and connectome.

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Neuroscience

by

Juyoun Yoo

2023

ABSTRACT OF THE DISSERTATION

Synaptic specificity determinants identification by converging transcriptome and connectome.

by

Juyoun Yoo

Doctor of Philosophy in Neuroscience

University of California, Los Angeles, 2023

Professor Stephen Lawrence Zipursky, Chair

Animal brains have as few as a couple hundred to as many as tens of billions of neurons. Each neuron in the developing brain recognizes and makes synapses with appropriate partners to build functional circuits. The principles of wiring have been studied in many scales, from cell-fate specification, cell recognition, local signaling to neural activity. Cell surface molecules have been shown to contribute to wiring in different stages and ways. However, the complexity of the brain hindered a larger and detailed search for the molecular underpinnings of wiring, especially synaptic specificity.

The three studies presented in this dissertation investigate molecular principles of wiring by profiling gene expression programs of the whole visual system and by referencing synaptic connectome as their final map of wiring choices. I'll describe in Chapter 1 an overview of anatomical and genomic methods that are developed up to date and an introduction to the drosophila motion detection circuit to contextualize it as a system where the two lineages of methods can converge and be used for developmental studies.

Chapter 2 and 3 show the transcriptomic approach to profile gene expression programs of individual cell types in the drosophila visual system and the transcriptomic architecture for synaptic wiring. In Chapter 2, I will present how single-cell RNA sequencing of the motion detector revealed the genetic architecture of the motion detector that reflects its wiring patterns. Eight subtypes of motion detectors derived from the same progenitors were profiled during the wiring period. They did not show cell-type specific genetic markers, instead revealed modular architecture of gene expression programs that correlated with dendritic and axonal connections.

In Chapter 3, I'll present an expansion of genetic effort to build a transcriptional atlas that describes developmental genetic programs of all the neurons in the fly visual system. We found that (i) the gene expression programs that are common to all developing neurons which reflect general developmental progress and (ii) the programs that are cell-type specific which endow cell-type specific characteristics and connectivity. This atlas is by far the largest and the most detailed developmental transcriptome that can be linked to the connectome.

In Chapter 4, by converging transcriptome and connectome, I'll show that the genetic underpinnings of the brain wiring can be identified systematically. We show that transcriptionally closely related neurons making alternative synaptic choices do so via pairs of immunoglobulin superfamily molecules. This study paves a way for further studies in larger brains by showing what can be achieved by combining developmental transcriptome and connectome. As the single-cell transcriptome became a household method and the synapse level connectome is being reconstructed for larger mammalian brains, it will elucidate common and unique molecular principles by applying this framework.

The dissertation of Juyoun Yoo is approved.

Mark Arthur Frye

Elissa A. Hallem

Alapakkam P. Sampath

Stephen Lawrence Zipursky, Committee Chair

University of California, Los Angeles

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SUPPLEMENTARY MATERIAL

(Attached separately)

Supplemental Table 4.S1: Genotypes for all figures and supplementary material.

ACKNOWLEDGEMENTS

First and foremost, I want to express my heartfelt appreciation to Larry, my advisor, for his support and patience throughout my graduate studies. As a scientist, he set the bar high for me, and I am committed to upholding those standards throughout my career. The research environment he provided me with a sandbox where I could fearlessly experiment with new ideas without fear of failure, and for that, I am truly grateful.

My time in the Zipursky lab was made delightful and stimulating by both past and current members. I would like to offer a special thanks to Dr. Yerbol Kurmangaliyev, with whom I worked closely for years as a team and who guided me in learning bioinformatics. Additionally, I would like to thank Drs. Liming Tan, Shuwa Xu and Qi Xiao for teaching me the beauty of fly genetics experiments. I am also thankful for Drs. Saumya Jain, Piero Sanfilippo, Sam LoCascio, Mark Dombrovski, Javier Valdes-Aleman, Ying Lin, Bryce Bajar, and Alex Kim, for the invaluable discussions in and out of the lab. Furthermore, I extend my gratitude to Dorian Gunning, Parmis Mirshahidi, Pegah Mirshahidi, and Laura Schreiber for keeping the lab running smoothly.

This achievement would not have been possible or as meaningful without the unwavering support of my family and friends. I thank the friends I have made throughout the years. The adventures we shared, the memories we made are all cherished moments that made me grow as a person. I am deeply grateful to my parents for their enduring love and support. Their unwavering belief in me and encouragement have been the foundation for all I do, and I am forever grateful.

My dissertation work was supported in part by the Dissertation Year Fellowship from UCLA Graduate Divisions. Chapter 2 is modified from published work: Kurmangaliyev, Y. Z., Yoo, J., LoCascio, S. A. & Zipursky, S. L. (2019). Modular transcriptional programs separately define axon and dendrite connectivity. ELife, 8, 505. https://doi.org/10.7554/elife.50822. The authors of this manuscript thank members of the Zipursky lab, Joshua Sanes, Karthik Shekhar, and Jonathan Flint for critical discussion of the manuscript. We thank Gerald Rubin, Hugo J Bellen, Aljosha Nern,

Orkun Akin, and Alain Garces for transgenic fly lines and antibodies. We thank Donghui Cheng and Owen Witte for assistance in FACS purification of cells, and UCLA TCGB and BSCRC Bio Sequencing Core Facility for assistance with single-cell RNA-Sequencing. This work was supported by the NIH National Institute of Neurological Disorders and Stroke (T32NS048004) (SAL), and the G Harold and Leila Y Mathers Foundation (SLZ). SLZ is an Investigator of the Howard Hughes Medical Institute.

Chapter 3 is modified from published work: Kurmangaliyev, Y. Z., Yoo, J., Valdes-Aleman, J., Sanfilippo, P. & Zipursky, S. L. (2020). Transcriptional Programs of Circuit Assembly in the Drosophila Visual System. *Neuron*, *108*(6), 1045-1057.e6. https://doi.org/10.1016/j.neuron.2020.10.006. The authors of this manuscript thank members of the Zipursky lab for critical reading and discussion of the manuscript. We thank the University of California, Los Angeles (UCLA) BSCRC Flow Cytometry core for assistance with fluorescence-activated cell sorting (FACS) purification of cells. We thank Michael Mashock and Xinmin Li from UCLA Technology Center for Genomics and Bioinformatics for assistance with scRNA-seq. We thank Hugo J. Bellen for transgenic fly lines. S.L.Z. is an Investigator of the Howard Hughes Medical Institute

Chapter 4 is modified from a submitted manuscript: Yoo, J., Dombrovski, M., Mirshahidi, P., Nern, A., LoCascio, S.A., Zipursky, S. L., Kurmangaliyev, Y. Z. Brain wiring determinants uncovered by integrating connectomes and transcriptomes. The authors of this manuscript thank Kazunori Shinomiya for sharing connectome data prior to publication. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) and Vienna Drosophila Resource Center were used in this study. We also thank G. M. Rubin, H. J. Bellen, and Qi Xiao for the reagents and the Janelia FlyLight Project Team for some images. We thank members of the Zipursky lab for the critical discussion of the manuscript. S.L.Z is an investigator of the Howard Hughes Medical Institute.

VITA

EDUCATION

Ph.D. Candidate in Neuroscience — 2015 - Present
University of California, Los Angeles, Neuroscience IDP
Thesis advisor: Dr. Larry Zipursky

M.S. in Biological Sciences — 2014
Seoul National University, Department of Biological Sciences

B.S. in Life Sciences — 2012
Ewha Womans University, South Korea

PUBLICATIONS

- Kurmangaliyev Y, **Yoo J**, Valdes-Aleman J, Sanfilippo P, Zipursky SL. (2020) Transcriptional programs of circuit assembly in the Drosophila visual system. *Neuron.* — 2020
- Kurmangaliyev Y*, **Yoo J***, LoCascio SA*, Zipursky SL. (2019) Modular transcriptional program defines neuron subtype-specific connectivity. *eLife* **(***Equal contribution) — 2019
- Xu S, Xiao Q, Cosmanescu F, Sergeeva AP, **Yoo J**, Lin Y, Katsamba PS, Ahlsen G, Kaufman J, Linaval NT, Lee PT, Bellen HJ, Shapiro L, Honig B, Tan L, Zipursky SL. (2018) Interactions between the Ig-Superfamily proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits. *Neuron*. — 2018
- Lim CS, Nam HJ, Lee J, Kim D, Choi JE, Kang SJ, Kim S, Kim H, Kwak C, Shim KW, Kim S, Ko HG, Lee RU, Jang EH, **Yoo J**, Shim J, Islam MA, Lee YS, Lee JH, Baek SH, Kaang BK. *(2017)* PKCα-mediated phosphorylation of LSD1 is required for presynaptic plasticity and hippocampal learning and memory. *Sci Rep*. — 2017
- Ko HG, Kim JI, Sim SE, Kim T, **Yoo J**, Choi SL, Baek SH, Yu WJ, Yoon JB, Sacktor TC, Kaang BK. (2016) The role of nuclear PKMζ in memory maintenance. *Neurobiol Learn Mem* — 2016
- Yu NK, Kim HF, Shim J, Kim S, Kim DW, Kwak C, Sim SE, Choi JH, Ahn S, **Yoo J**, Choi SL, Jang DJ, Lim CS, Lee YS, Kang C, Choi SY, Kaang BK. (2016) A transducible nuclear/nucleolar protein, mLLP, regulates neuronal morphogenesis and synaptic transmission. *Sci Rep.* — 2016
- Kim S, Kim T, Lee HR, Jang EH, Ryu HH, Kang M, Rah SY, **Yoo J**, Lee B, Kim JI, Lim CS, Kim SJ, Kim UH, Lee YS, Kaang BK. (2016) Impaired learning and memory in CD38 null mutant mice. *Mol Brain.* — 2016
- **Yoo J**, Bakes J, Bradley C, Collingridge GL, Kaang BK (2013) Shank Mutant Mice as an Animal Model of Autism. *Philos Trans R Soc Lond B Biol Sci*. — 2013

CONFERENCE PRESENTATIONS

- *Society for Neuroscience (Poster)* 2022
  Title: Minimal molecular code for synaptic specificity in the fly visual system
- *Gordon Research Conference – Neural Development (Poster)* 2022
  Title: Differential expression of matching pairs of interacting IgSF proteins define synaptic specificity in Drosophila motion detection circuit.
- *CSHL Meeting: Neurobiology of Drosophila (Poster)* 2021
  Title: Cell type and subcellular synaptic specificity in the visual system
- *Society for Neuroscience (Talk)* 2019
  Nanosymposium - Molecular Mechanisms of Synaptogenesis and Connectivity
  Title: Modular Transcriptional Programs define neuron subtype-specific connectivity
- *Society for Neuroscience (Poster)* 2013
  Title: Disrupted social behaviors of CD38 Knock-out mice assayed via three-chamber test and pup retrieval test

RESEARCH EXPERIENCE

Janelia Junior Scientist Workshop on Mechanistic Cognitive Neuroscience, 2021
Janelia Research Campus, Ashburn, VA (online)

Gene Regulatory Networks for Development 2018
Marine Biology Laboratory, Woods Hole, MA

HONORS AND AWARDS

- **AKN Outstanding Research Award**, Association of Korean Neuroscientists predoctoral research award 2022
- **Dissertation Year Fellowship**, UCLA Graduate Division 2020
- **Graduate Student Travel Award**, UCLA Brain Research Institute and Semel Institute for Neuroscience & Human Behavior 2019
- **Brain Korea 21 Fellowship**, Seoul National University 2012-2013
- **Best Poster Award**, Undergraduate Research Program, Ewha Womans University 2011
- **Academic Excellence Scholarship**, School of Natural Science, Ewha Womans University 2011
- **Honors Scholarship**, Department of Life Sciences, Ewha Womans University 2018-2019
- **Dean's List**, Department of Life Sciences, Ewha Womans University 2008

CHAPTER ONE

Decoding Brain Connectivity from the Genome

1.1.1 Complexity of the brain connectivity: EM connectome

Ramon y Cajal conducted a meticulous examination of the cellular structures in the nervous system across various species of both vertebrate and invertebrate, which revealed a remarkable diversity of the cells in the brain (Cajal, 1894, 1899; Cajal & Sánchez, 1915). Neuronal shapes play a critical role in determining many of their physiological properties. For example, beginning in the late 50s, the cable model accounting for diameters and length of dendritic branches was devised to mathematically predict experimental results of physiological recordings (Holmes & Rall, 1992; Rall, 1962). Also, the neurite branching patterns provide the opportunities to make synapses with other neurons (Ian A. Meinertzhagen et al., 2009). Complex synaptic input and output patterns are the basis of intricate signal transmission for a neuron.

Over the years, the anatomical taxonomy of the Drosophila visual system has been extensively studied using various techniques (**Figure 1.1A**). Stochastic neuronal staining with the Golgi method was used, and more recently, genetic tools such as GAL4/LexA enhancer-trap strains, combined with light microscopy have been employed (Brand & Perrimon, 1993; Fischbach & Dittrich, 1989). Although Golgi staining based classification of neurons with morphological features have identified many cell types, stochastic methods have limitations in the uncertainty of representation since they may not represent all cell types due to incomplete labeling of neurons. The use of genetic tools, such as fluorescent protein reporters or mosaic methods has advanced our understanding on the diversity of cell types in the fly visual system (Lee & Luo, 1999; Nern et al., 2015) (**Figure 1.1B**). However, to comprehend neuronal function at the system level it is necessary to understand the complex synaptic connectivity between diverse neurons. Without connectivity information, it is difficult to build computational models and determine targets for physiological recordings (Ian A. Meinertzhagen et al., 2009).

In contemporary neuroscience, the depiction of the anatomy, morphology and the connectivity of the brain is best represented by the synapse-level connectome, which can be

obtained through electron microscopic (EM) three-dimensional reconstruction of the brain. Ultra-thin (<40nm sections) electron micrographs are imaged sequentially, and stacked images are segmented to reconstruct 3D volume of the sample (**Figure 1.1C-E**). EM connectome not only reveals the intricate morphology of the neurons but also provides information on the precise synaptic connections among diverse neuronal populations, thereby unveiling the circuitry of the brain (Schneider-Mizell et al., 2023; Shinomiya et al., 2019; S. Takemura et al., 2013; S.-Y. Takemura et al., 2015; Turner et al., 2020).

The compact size of the fly brain has been advantageous in the reconstruction of many parts to date. In the visual system, EM reconstruction has been used to understand the information pathway of the motion detection circuit, from the retina to downstream optic ganglia then to the central brain. The first EM reconstruction of lamina in the early 90s demonstrated the operational rule we can get from connectivity patterns of lamina neurons (I. A. Meinertzhagen & O'Neil, 1991). The Fly EM project from Janelia expanded this effort to medulla (S. Takemura et al., 2013, 2017), and lobula plate (Shinomiya et al., 2022) to provide a near complete synaptic connection map between neurons in the motion detection pathway.

Methods and technologies to generate EM sections, imaging and computational alignment and segmentation are evolving. Also, its utilization has been expanding to the mammalian brain such as mouse, primate, and human, which are much larger in size and thus contain many more cells and connections between them. Precise reconstruction of the connectivity is revealing connectivity pattern differences across species (Loomba et al., 2022; Schneider-Mizell et al., 2023).

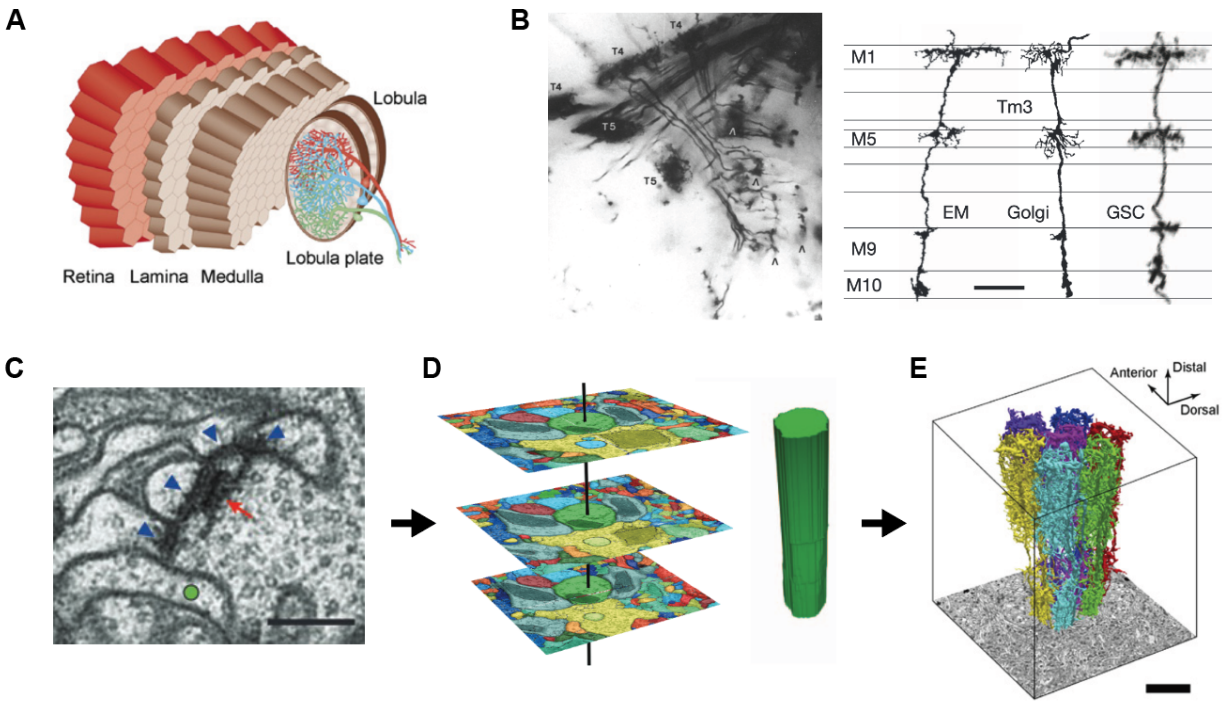


Figure 1.1. The Drosophila visual system and EM reconstruction.
(A) Schematic of fly optic lobe structure (Adapted from Takemura et al., 2015). (B) Comparison of Golgi impregnation of T4/T5 neurons (left). Comparison of EM reconstruction, Golgi impregnation, and genetic single-cell (GSC)-labeled neuron (right). Scale bar, 10um. (Adapted from Fischbach & Dittrich, 1989 and Nern et al., 2015). (C) A representative serial section electron micrograph. T-bar ribbon (red arrow), postsynaptic densities (PSDs) (blue arrowheads). A non-synaptic process (green circle) lacks PSD. scale bar, 250 nm. (D) Neurites reconstructed by linking segmented profiles in consecutive sections. (E) 3D-reconstruction of modular medulla cell types in each of seven reconstructed columns. (Adapted from Takemura et al. 2013; 2015).

1.1.2. Complexity of the brain molecular diversity: High-throughput transcriptomic analysis using single-cell RNA sequencing

Gene expression profiling by enrichment for specific cell types through fluorescence-activated cell sorting (FACS) or microdissection have elucidated changes in gene expression across different cell populations. However, in the developing nervous system, gene expression changes are rapid and dynamic so that it requires longitudinal sampling of a cell population. Some cell types lack genetic access for their enrichment, or if such access exists, often they have temporal limitations. As a result, obtaining longitudinal profiles for each cell type can be a challenging task.

Single-cell RNA sequencing (scRNA-seq) technology emerged in the mid-2010s when I began my thesis project and it revolutionized gene expression analysis without prior cell-type enrichment. Droplet-based method using a microfluidic device captures individual cells in oil droplets along with gel-beads that are covered with unique molecular cell barcodes and unique molecular identifiers (UMI). Each mRNA molecule from a single cell shares the same cell barcode and is marked by a different UMI (Macosko et al., 2015; Zheng et al., 2017) (**Figure 1.2A**). After reverse transcription of the droplet, the resulting cDNA library, which includes unique cell barcode and UMI, is sequenced. Bioinformatic tools then count the number of UMI per gene within each cell barcode, allowing assessment of the gene expression level in each cell. Bioinformatic methods are rapidly improving in parallel to analyze transcriptomic profiles (Butler et al., 2018; Hao et al., 2021; Satija et al., 2015)(**Figure 1.2B**). For example, unsupervised clustering methods can reveal subpopulations within a cell type that was previously considered homogeneous. scRNA-seq is now widely used, including our work, to analyze cell types across many species and different types of tissues, including the brain (Davie et al., 2018; Kurmangaliyev et al., 2020; Özel et al., 2020; Tasic et al., 2018; Yao et al., 2021). .

The high-throughput single-cell transcriptome approach gives unprecedented opportunity to reveal molecular diversity of cell types in the developing nervous system. In Chapters 3 and 4, I'll present how we adapted scRNA-seq to analyze developmental transcriptional programs and transcriptional signatures related to wiring specificity in the Drosophila visual system.

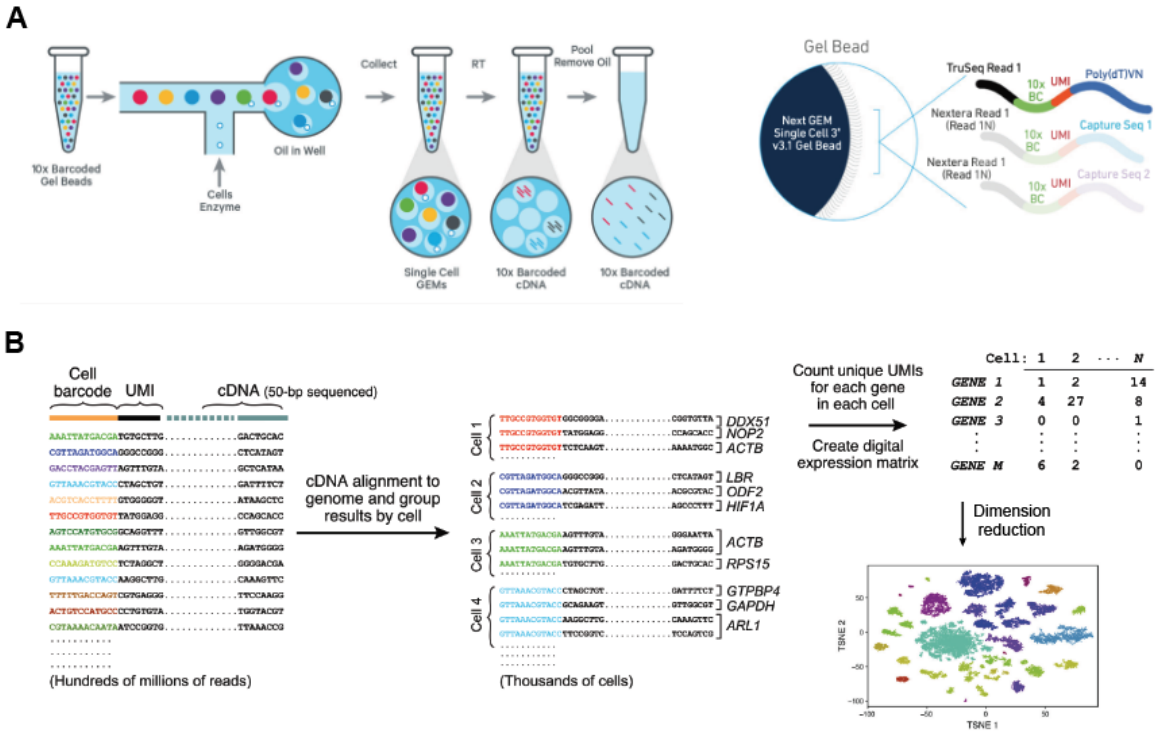


Figure 1.2. High-throughput transcriptomic analysis using single-cell RNA sequencing. (A) Schematics of droplet based single-cell RNA sequencing (http://10XGenomics.com). (B) In silico reconstruction of thousands of single-cell transcriptomes to generate a digital expression matrix. Dimension reduction methods to cluster cells with similar profiles. (Adapted from Macosko et al. 2015).

1.2. Principles of wiring specificity during brain development.

Diversity and complexity of neurons, as well as their synaptic specificity within dense spatial environments are achieved by continuous interactions among developing neurons. Individual brains build stereotypic circuits during development, ensuring reliable functions of adult animals. How do neurons find their right synaptic partners? Over the last couple of decades, significant progress has been made in advancing our understanding of the principles of neural circuit wiring.

1.2.1. Cell adhesion molecules

Sperry proposed the chemoaffinity theory, which suggests that cells and fibers in the brain carry cytochemical “individual identification tags,” enabling them to be distinguished to the level of a single neuron (Sperry, 1963). Biochemical and genetic studies have since identified such cell recognition molecules with heterophilic or homophilic binding affinities that can lead to adhesive or repulsive interactions between contacting neurons for proper wiring in multiple steps during circuit formation. (Sanes & Zipursky, 2020).

Immunoglobulin superfamily (IgSF) proteins encompasses hundreds of cell surface proteins including subfamilies of DIPs/Dprs, Beats/Sides, Sidekicks, Dscams, Syg1/Syg2-related proteins and Contactins (**Figure 1.3A**). In vitro screening showed that 2-Ig domain protein Dprs and 3-Ig domain protein DIP form a network of interactions. With varying binding affinities among the interactors, most Dprs interact with multiple DIPs and vice versa (Carrillo et al., 2015; Özkan et al., 2013) (**Figure 1.3B**).

Bulk RNA-seq of FACS purified cell types in developing Drosophila medulla (R7, R8, and L1-L5 neurons at 40h After Pupal Formation (APF)) showed cell-type specific, unique combination of gene expression patterns of DIPs and Dprs. DIP expression in medulla neurons occupying the same layers as lamina neurons was assessed using DIP-GAL4 and UAS reporter, which showed a correlation between high-affinity interactors of DIPs and Dprs expression in the synaptic

partners (Carrillo et al., 2015; Tan et al., 2015) (**Figure 1.3B**). The DIP-alpha and dpr6/10 binding pairs are shown to have a role in the medulla layer M3 neurons, Dm4 and Dm12. Loss-of-function alleles and affinity mutant alleles of DIP-alpha and dpr6/10 disrupt or change heterophilic interaction, resulting in defects in many wiring features such as arborization within layers, synapse number, layer specificity, cell survival. These results suggest that specific interaction of DIP/dpr families plays a role in wiring (Xu et al., 2018, 2022).

The Beat and Side subfamily include 14 Beats and 8 Side members. Albeit partial, their interaction network has been screened in vitro (Li et al., 2017; Özkan et al., 2013) (**Figure 1.3C**). Founding members of these families (Side and Beat-Ia) were identified in genetic screens as regulators of motor axon guidance in the Drosophila embryo (Fambrough & Goodman, 1996; Siebert et al., 2009; Sink et al., 2001). Functions for other paralogs have not been described. As each neuron type in the developing visual system expresses a unique combination of Beats and Sides proteins during development these proteins may contribute to patterning the connectome more broadly.

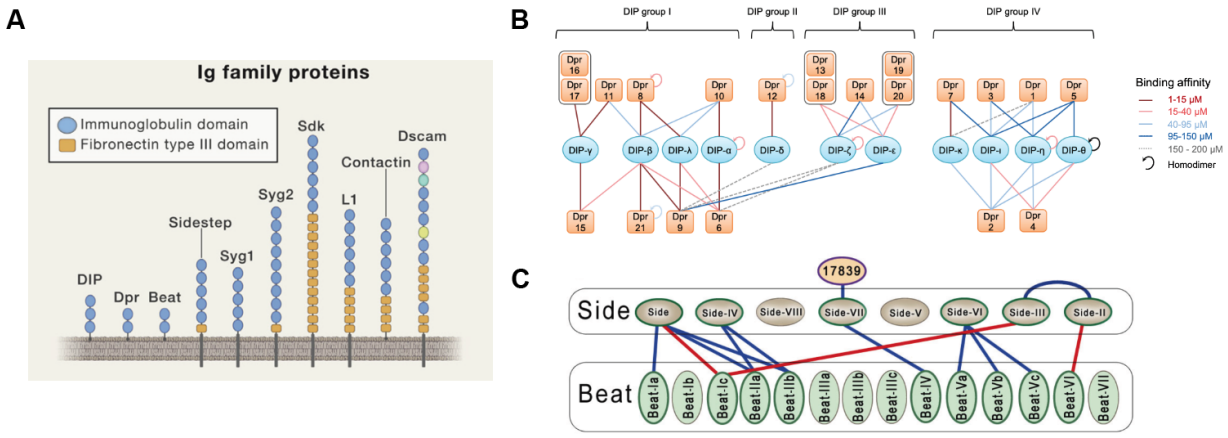


Figure 1.3. Cell adhesion protein families regulating synaptic specificity: Immunoglobulin superfamily proteins.

(A) Ig family protein structures (Sanes and Zipursky, 2020) (B) DIP-dpr protein family heterophilic and homophilic interaction network. Lines are color coded according to the affinity of the binding pairs. Adapted from Cosmanescu et al, 2018 (C) Side-Beat interaction network (Li et al. 2017)

1.2.2. Spatio-temporal regulation of gene expression

Temporal change in the expression patterns of cell adhesion molecules is required to understand the cell surface environment dynamics of the developing brain. Drosophila brain development takes approximately 100 hours during pupal metamorphosis stage in 25 °C (**Figure 1.4A**). Neurogenesis phase ends before or around 24h APF, axon guidance and neurite targeting occurs during earlier pupal development (~48h APF). After most postmitotic neuronal axons and dendrites reach their target, cell-type specific calcium activity recordings using GAL4/LexA drivers and GCaMP with 2-photon microscopy showed that the same cell types tend to show synchronized activity in the early PSINA phase in pupa. The activity pattern changes to an unsynchronized, turbulent phase after ~70h APF. (Akin et al., 2019) (**Figure 1.4B**).

Silencing PSINA resulted in an increased number of presynaptic sites in multiple cell types, showing activity-dependent synapse formation and suggesting the role of activity in circuit development (Bajar et al., 2022). It is not clear yet whether and how the emergence of spontaneous activity and activity pattern changes are related to gene expression changes of neuronal intrinsic property related proteins, such as voltage gated calcium channels, potassium channels.

To capture gene expression changes related to all these developmental events, longitudinal sampling in relevant temporal resolution is necessary. Chapter 4 will delve further into pan-neuronal gene expression and cell-type specific gene expression programs. These studies provide a deeper understanding of the intricate mechanisms that govern Drosophila brain wiring, offering insights into the broader field of neural development.

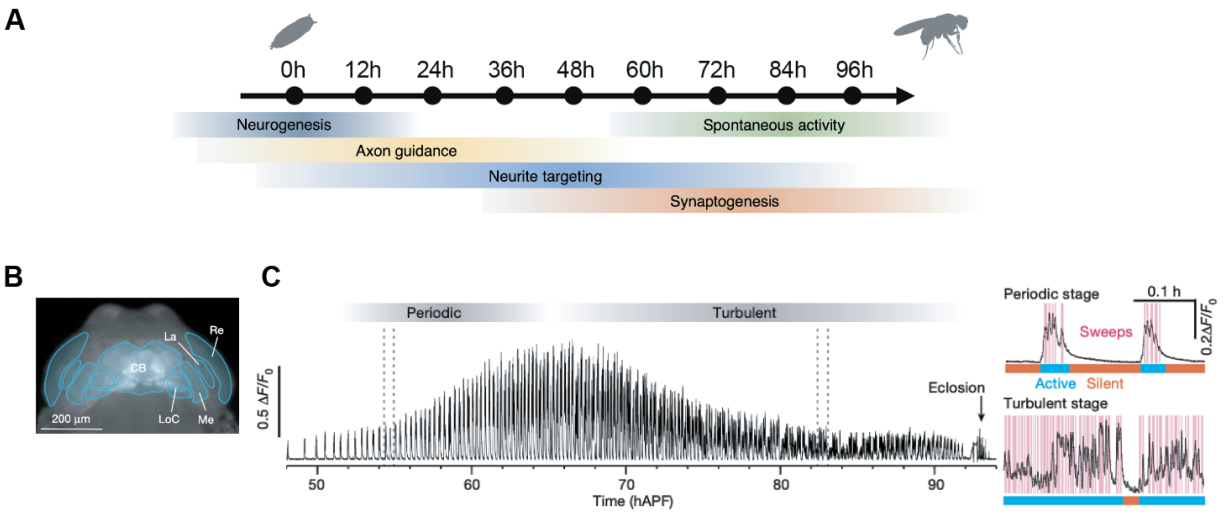


Figure 1.4. Development of Drosophila visual system
(A) Drosophila brain develops during about 100 hours of pupal stage (B) Average intensity projection (AIP) from wide-field fluorescence series of a pupa expressing pan-neuronal GCaMP6s. CB, central brain; La, lamina; Re, retina; LoC, lobula complex; Me, medulla. Scale bar, 200 µm. (C) Representative trace of PSINA, recorded with wide-field fluorescence imaging from a pupa expressing pan-neuronal GCaMP6s. Dotted lines mark limits of inset traces (right) from periodic (top) and turbulent (bottom) stages. Highlights (magenta) mark individual sweeps; bars mark active (cyan) and silent (orange) phases (right) (B-C, Adapted from Bajar et al. 2021).

1.3. Drosophila motion detection pathway

1.3.1. Organization of the Drosophila motion detection pathway

The visual system plays a crucial role in the sensory experience of flies, with half of the brain by volume dedicated to process visual information. Visual navigation, a key aspect of fly behavior, relies on motion computation, which is carried out in a compact nervous system of the optic lobe (**Figure 1.5A**). The drosophila motion detection system has long been a model system for studying the principles of motion sensing and neuronal computation (Borst, 2014). Different cell types that contribute to the motion detection of the drosophila have been identified along the visual pathway. The functional contribution of each cell type has been tested using opto-motor behavioral measurements, and manipulation of each cell type transmission using genetic tools such as UAS-shibire (Tuthill et al., 2013). The T4/T5 subtypes preferentially respond to one of four cardinal directions of visual stimuli (Maisak et al., 2013). The physiological principles of motion detection of T4/T5 neurons have been shown to reflect classical computational models initially proposed in cat and rabbit visual system in the last century (Barlow & Levick, 1965; Groschner et al., 2022; Gruntman et al., 2019; Reichardt et al., 2003; Strother et al., 2017). Motion detector T4/T5 receives spatially and temporally offset inputs, they compute motion through non-linear integration of the signals.

1.3.2. Motion-detector T4/T5

Motion-detector T4/T5 are the first motion-sensitive neurons in the fly visual pathway. There are eight morphological subtypes which are functionally distinct (Borst et al., 2019). T4 neurons receive input from medulla and respond to increment of light stimuli (ON) and T5 neurons receive input from lobula and respond to decrement (OFF) of light. T4 and T5 are each divided into four subtypes depending on their axon terminal target layers in the lobula plate (a-d). Each

lobula plate layer corresponds to one of four cardinal directions of stimuli T4/T5 preferentially respond to (a: front-to-back, b: back-to-front, c: upward, d: downward) (Maisak et al., 2013).

T4/T5 neurons originate from a common progenitor pool in the inner proliferation center (IPC). Two molecularly distinct regions in the IPC give rise to T4/T5 a/b types (Brk+) and c/d types (Dpp+). These progenitors undergo two sequential Notch dependent divisions. The first division distinguishes a from b, and c from d, and then the subsequent division specifies each a-d types into T4 vs T5. Therefore, each progenitor generates two T4 and two T5 neurons with opposite direction selectivity that innervates a single retinotopic column. Thus, the lineage history of T4/T5 are related to their final identity and morphological types (Apitz & Salecker, 2015; Pinto-Teixeira et al., 2018).

During the neurogenesis wave in larval to early pupal stage, post-mitotic T4/T5 neurons begin initial neurite targeting, with dendrites reaching either medulla (T4) or lobula (T5). The axon terminals initially target a thin layer of proto lobula plate, which lacks clear laminar structure. After the neurogenesis window toward the mid-pupal stage, axon terminals form two thick proto-layers, one for the horizontal system (T4/T5ab) and the other for the vertical system (T4/T5cd). From 48h APF to 65h PAF, each proto-layer further separates into two and eventually becomes four discrete layers (**Figure 1.5B**).

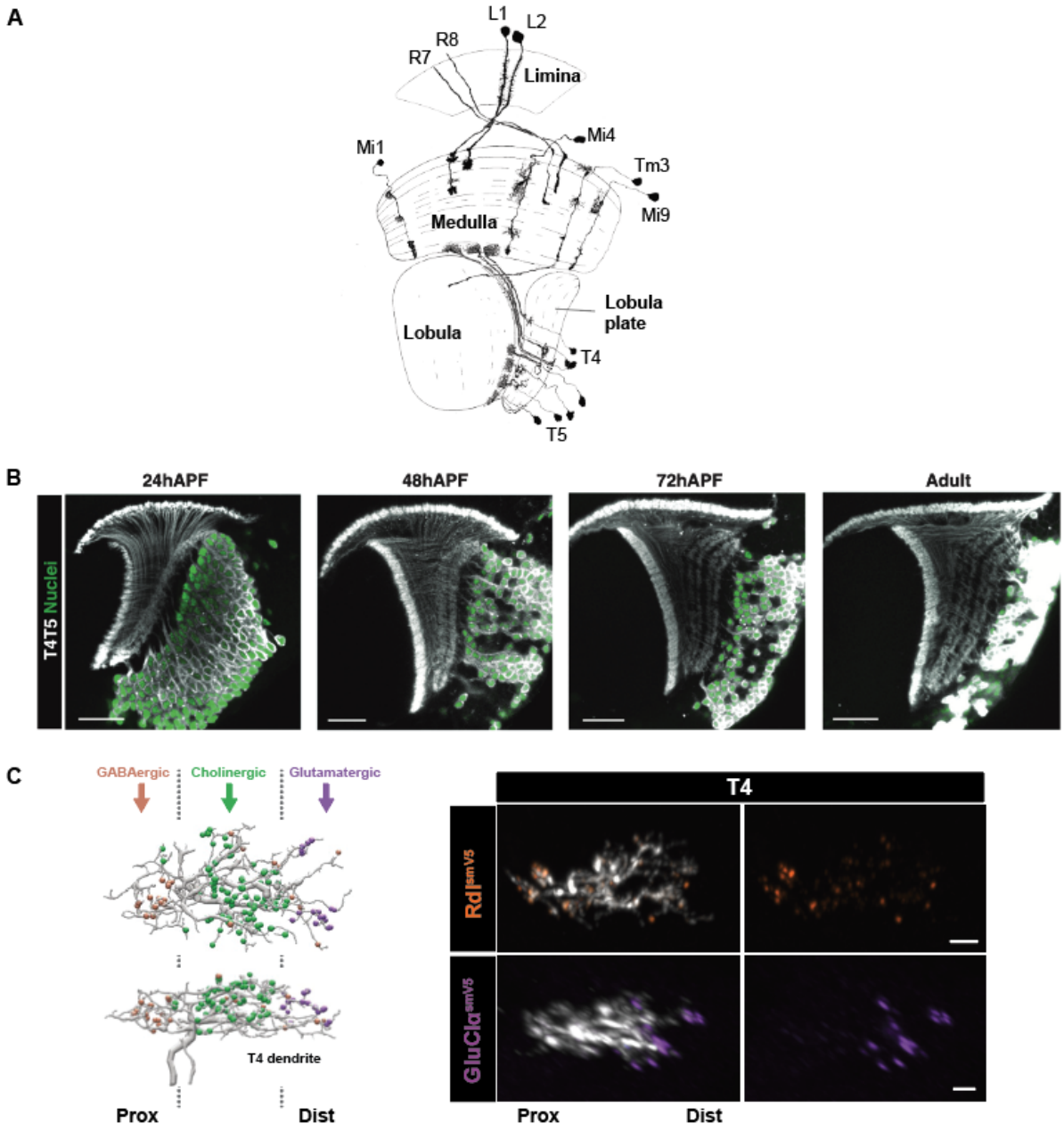


Figure 1.5. Organization of the Drosophila Motion detection system
(A) Schematic of the optic lobe neuropils. ON pathway neurons and motion detector T4/T5 are represented. (Adapted from Fischbach & Dittrich, 1989) (B) Sequential lamination of T4/T5 axons and four LoP layers. 23G12-Gal4 drives membrane localized RFP (grey) and nuclear localized GFP (green) in all T4/T5 neurons throughout pupal development (Adapted from Kurmangaliyev, Yoo, LoCascio & Zipursky. 2019) (C) T4 dendrite receive subcellular domain specific inputs (left) Postsynaptic receptors localizes to subcellular domain of the T4 dendrite: GABA receptor Rdl in proximal (orange), Glutamate receptor GluClα in distal dendrite (purple). Genetic tool that allows cell type specific conditional induction of epitope tag inserted in the endogenous locus of each receptor subunit was used. (Sanfilippo et al. 2023)

1.3.3 Synaptic specificity underlying motion detection: mechanisms of motion detection

The connectome offers a possible molecular explanation for the mechanism of motion detecting computation in the dendrite by showing connectivity distribution of cell type specific inputs onto the dendrite. The proximal region of T4 dendrite receives GABAergic inputs (Mi4, CT1), the middle region receives cholinergic excitatory inputs (Mi1, Tm3), and the distal part receives inhibitory glutamatergic input (Mi9), which supports the proposed models of non-linear dendritic computations (Strother et al., 2017; S. Takemura et al., 2013, 2017). Furthermore, the synaptic specificity identified in the connectome is corroborated by postsynaptic neurotransmitter localization. Receptor localization assays show Rdl localizes to the proximal T4 dendrite, where Mi4 and CT1 neurons make synapses, while GluClα localizes to the tip of the dendrite where Mi9 neurons make synapses (Sanfilippo et al. 2023, unpublished). (**Figure 1.5C**) The biophysical implementation of T4 dendrite computation has been investigated by modeling based on the connectivity of presynaptic neurons on to T4 dendrites, and by silencing or activating the presynaptic partners (Groschner et al., 2022; Strother et al., 2017). Thus, the connectome provides detailed analysis of synaptic connectivity information from EM data and offers insight into the function of neural circuits.

The relationship between presynaptic inputs and T4/T5 neurons has been extensively studied in both connectome and functional studies. The T4/T5 axon terminals connectome in the lobula plate demonstrates the convergence of two streams of information at the synapse level. T4 and T5 neurons targeting the same lobula plate layer share the same postsynaptic partners, thereby integrating the ON and OFF information conveyed by two separate sets of signal relays represented by T4 and T5, respectively. As a result, the information converges on the same downstream pathways that encode directional preference (Shinomiya et al., 2022).

1.3.4. From transcriptome to connectome: How is the brain wiring encoded in the genome?

With advancement of single-cell technologies, it is now possible to access gene expression profiles of individual neuronal types during development. While synapse-level connectivity map provides the final goal that the developing nervous system has to achieve, it remains a challenge to deduce this final connectivity from mRNA expression patterns alone. Reverse engineering the final connectivity requires accounting for many developmental decisions that a single postmitotic neuron makes during development, such as projecting axons and dendrites to the right target, selecting appropriate synaptic partners from a large pool of physical contacts. Thus, the developmental transcriptome not only reflects the final connectivity map, but also the myriad of intracellular and intercellular activities that neurons undergo. The goal of this dissertation is to elucidate the molecular complexity of neurons during development and to identify synaptic specificity rules from gene expression patterns and connectome. In Chapter 4, I'll present a framework we devised to integrate transcriptome and connectome, in order to gain insight into the molecular underpinning of synaptic choice.

1.4. References

Akin, O., Bajar, B. T., Keles, M. F., Frye, M. A. & Zipursky, S. L. (2019). *Cell*-Type-Specific Patterned Stimulus-Independent *Neuron*al Activity in the Drosophila Visual System during Synapse Formation. Neuron, 1–35. https://doi.org/10.1016/j.neuron.2019.01.008

Apitz, H. & Salecker, I. (2015). A region-specific neurogenesis mode requires migratory progenitors in the Drosophila visual system. *Nature Neuroscience*, *18*(1), 46–55. https://doi.org/10.1038/nn.3896

Bajar, B. T., Phi, N. T., Isaacman-Beck, J., Reichl, J., Randhawa, H. & Akin, O. (20*22*). A discrete neuronal population coordinates brain-wide developmental activity. Nature, 1–8. https://doi.org/10.1038/s41586-022-04406-9

Barlow, H. B. & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *The Journal of Physiology*, *178*(3), 477–504. https://doi.org/10.1113/jphysiol.1965.sp007638

Borst, A. (2014). In search of the holy grail of fly motion vision. *European Journal of Neuroscience*, *40*(9), 3285–3293. https://doi.org/10.1111/ejn.12731

Borst, A., Haag, J. & Mauss, A. S. (2019). How fly neurons compute the direction of visual motion. *Journal of Comparative Physiology A*, *206*(2), 109–124. https://doi.org/10.*1007*/s00359-019-01375-9

Brand, A. H. & Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, *118*(2), 401–415. https://doi.org/10.1242/dev.118.2.401

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, *36*(5), 411–420. https://doi.org/10.1038/nbt.4096

Cajal, S. R. y. (18*94*). Neue Darstellung vom histologischen Bau des Central nerven systems. *The American Journal of Psychology*, *6*(3), 450. https://doi.org/10.2307/1411662

Cajal, S. R. y. (1899). Estudios sobre la corteza cerebral humana I: corteza visual. *Rev. Trim. Microgr*, *4*, 1–63.

Cajal, S. R. y & Sánchez, D. (1915). Contribución al conocimiento de los centros nerviosos de los insectos. *Trab Lab Invest Biol (Madrid)*, *13*, 1–164. https://doi.org/10.5962/bhl.title.37839

Carrillo, R. A., Özkan, E., Menon, K. P., Nagarkar-Jaiswal, S., Lee, P.-T., Jeon, M., Birnbaum, M. E., Bellen, H. J., Garcia, K. C. & Zinn, K. (2015). Control of Synaptic Connectivity by a Network of Drosophila IgSF Cell Surface Proteins. Cell, *163*(7), 1770–1782. https://doi.org/10.1016/j.cell.2015.11.022

Davie, K., Janssens, J., Koldere, D., Waegeneer, M. D., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., González-Blas, C. B., Poovathingal, S., Hulselmans, G., Spanier, K. I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. Cell, *174*(4), 982-998.e20. https://doi.org/10.1016/j.cell.2018.05.057

Fambrough, D. & Goodman, C. S. (1996). The Drosophila beaten path Gene Encodes a Novel

Secreted Protein That Regulates Defasciculation at Motor Axon Choice Points. Cell, *87*(6), 1049–*105*8. https://doi.org/10.1016/s0092-8674(00)81*7*99-7

Fischbach, K. F. & Dittrich, A. (1989). The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. *Cell and Tissue Research*. http://link.springer.com/article/10.1007/BF00218858

Groschner, L. N., Malis, J. G., Zuidinga, B. & Borst, A. (2022). A biophysical account of multiplication by a single neuron. Nature, *603*(7899), 119–123. https://doi.org/10.1038/s41586-022-04428-3

Gruntman, E., Romani, S. & Reiser, M. B. (2019). *The computation of directional selectivity in the Drosophila OFF motion pathway*. *8*, 929–926. https://doi.org/10.7554/elife.50706

Hao, Y., Hao, S., Andersen-Nissen, E., III, W. M. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., … Satija, R. (2021). Integrated analysis of multimodal single-cell data. Cell, 1–45. https://doi.org/10.1016/j.cell.2021.04.048

Holmes, W. R. & Rall, W. (1992). Electrotonic length estimates in neurons with dendritic tapering or somatic shunt. *Journal of Neurophysiology*, *68*(4), 1421–1437. https://doi.org/10.1152/jn.1992.68.4.1421

Kurmangaliyev, Y. Z., Yoo, J., Valdes-Aleman, J., Sanfilippo, P. & Zipursky, S. L. (2020). Transcriptional Programs of Circuit Assembly in the Drosophila Visual System. Neuron, *108*(6), 1045-1057.e6. https://doi.org/10.1016/j.neuron.2020.10.006

Lee, T. & Luo, L. (1999). Mosaic Analysis with a Repressible Cell Marker for Studies of Gene Function in Neuronal Morphogenesis. Neuron, 22(3), 451–461. https://doi.org/10.1016/s0896-6273(00)80701-1

Li, H., Watson, A., Olechwier, A., Anaya, M., Sorooshyari, S. K., Harnett, D. P., Lee, H.-K. P., Vielmetter, J., Fares, M. A., Garcia, K. C., Özkan, E., Labrador, J.-P. & Zinn, K. (2017). Deconstruction of the beaten Path-Sidestep interaction network provides insights into neuromuscular system development. *ELife*, *6*, 270. https://doi.org/10.7554/elife.2*8*111

Loomba, S., Straehle, J., Gangadharan, V., Heike, N., Khalifa, A., Motta, A., Ju, N., Sievers, M., Gempt, J., Meyer, H. S. & Helmstaedter, M. (2022). Connectomic comparison of mouse and human cortex. *Science*, *377*(6602), eabo0924. https://doi.org/10.1126/science.abo0924

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A. & McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell, *161*(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., Schilling, T., Bahl, A., Rubin, G. M., Nern, A., Dickson, B. J., Reiff, D. F., Hopp, E. & Borst, A. (2013). A directional tuning map of Drosophila elementary motion detectors. Nature, *500*(7461), 212–216. https://doi.org/10.1038/nature12320

Meinertzhagen, I. A. & O'Neil, S. D. (1991). Synaptic organization of columnar elements in the lamina of the wild type in Drosophila melanogaster. *Journal of Comparative Neurology*, *305*(2),

232–263. https://doi.org/10.1002/cne.903050206

Meinertzhagen, Ian A., Takemura, S., Lu, Z., Huang, S., Gao, S., Ting, C.-Y. & Lee, C.-H. (2009). From Form to Function: the Ways to Know a Neuron. *Journal of Neurogenetics*, *23*(1–2), 68–77. https://doi.org/10.1080/01677060802610604

Nern, A., Pfeiffer, B. D. & Rubin, G. M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proceedings of the National Academy of Sciences*, *112*(22), E2967-76. https://doi.org/10.1073/pnas.1*50*6763112

Özel, M. N., Simon, F., Jafari, S., Holguera, I., Chen, Y.-C., Benhra, N., El-Danaf, R. N., Kapuralin, K., Malin, J. A., Konstantinides, N. & Desplan, C. (2020). Neuronal diversity and convergence in a visual system developmental atlas. Nature, *589*(7840), 88–95. https://doi.org/10.1038/s41586-020-2879-3

Özkan, E., Carrillo, R. A., Eastman, C. L., Weiszmann, R., Waghray, D., Johnson, K. G., Zinn, K., Celniker, S. E. & Garcia, K. C. (2013). An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. Cell, *154*(1), 228–239. https://doi.org/10.1016/j.cell.2013.06.006

Pinto-Teixeira, F., Koo, C., Rossi, A. M., Neriec, N., Bertet, C., Li, X., Rodríguez, A. del V. & Desplan, C. (2018). Development of Concurrent Retinotopic Maps in the Fly Motion Detection Circuit. Cell, 1–26. https://doi.org/10.1016/j.cell.2018.02.053

Rall, W. (1962). Theory of Physiological Properties of Dendrites. *Annals of the New York Academy of Sciences*, *96*(4), 1071–1092. https://doi.org/10.1111/j.1749-6632.1962.tb54120.x

Reichardt, W., Egelhaaf, M. & Guo, A. (2003). Processing of figure and background motion in the visual system of the fly. *Biological Cybernetics*, *61*(5), 327–345. https://doi.org/10.1007/bf00200799

Sanes, J. R. & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell, *181*(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nature Biotechnology, *33*(5), 495–502. https://doi.org/10.1038/nbt.3192

Schneider-Mizell, C. M., Bodor, A., Brittain, D., Buchanan, J., Bumbarger, D. J., Elabbady, L., Kapner, D., Kinn, S., Mahalingam, G., Seshamani, S., Suckow, S., Takeno, M., Torres, R., Yin, W., Dorkenwald, S., Bae, J. A., Castro, M. A., Fahey, P. G., Froudakis, E., … Costa, N. M. da. (2023). Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex. *BioRxiv*, 2023.01.23.525290. https://doi.org/10.1101/2023.01.23.525290

Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C. S., Aniceto, R., Ansari, N., Cheatham, N., Lauchie, S., Neace, E., Ogundeyi, O., Ordish, C., Peel, D., Shinomiya, A., Smith, C., Takemura, S., Talebi, I., Rivlin, P. K., Nern, A., … Meinertzhagen, I. A. (2019). Comparisons between the ON- and OFF-edge motion pathways in the Drosophila brain. ELife, 8, e40025. https://doi.org/10.7554/elife.40025

Shinomiya, K., Nern, A., Meinertzhagen, I. A., Plaza, S. M. & Reiser, M. B. (2022). Neuronal circuits integrating visual motion information in Drosophila melanogaster. *Current Biology*, *32*(16), 3529-3544.e2. https://doi.org/10.1016/j.cub.2022.06.061

Siebert, M., Banovic, D., Goellner, B. & Aberle, H. (2009). Drosophila motor axons recognize and follow a Sidestep-labeled substrate pathway to reach their target fields. *Genes & Development*, *23*(9), 1052–1062. https://doi.org/10.1101/gad.520509

Sink, H., Rehm, E. J., Richstone, L., Bulls, Y. M. & Goodman, C. S. (2001). sidestep encodes a target-derived attractant essential for motor axon guidance in Drosophila. Cell, 105(1), 57–67. https://doi.org/10.1016/s0092-8674(01)00296-3

Sperry, R. W. (1963). CHEMOAFFINITY IN THE ORDERLY GROWTH OF NERVE FIBER PATTERNS AND CONNECTIONS*. Proceedings of the National Academy of Sciences, 50(4), 703–710. https://doi.org/10.1073/pnas.50.4.703

Strother, J. A., Wu, S.-T., Wong, A. M., Nern, A., Rogers, E. M., Le, J. Q., Rubin, G. M. & Reiser, M. B. (2017). The Emergence of Directional Selectivity in the Visual Motion Pathway of Drosophila. Neuron, 94(1), 168-182.e10. https://doi.org/10.1016/j.neuron.2017.03.010

Takemura, S., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P. K., Katz, W. T., Olbris, D. J., Plaza, S. M., Winston, P., Zhao, T., Horne, J. A., Fetter, R. D., Takemura, S., Blazek, K.,

Chang, L.-A., Ogundeyi, O., Saunders, M. A., Shapiro, V., … Chklovskii, D. B. (2013). A visual motion detection circuit suggested by Drosophila connectomics. Nature, 500(7461), 175–181. https://doi.org/10.1038/nature12450

Takemura, S., Nern, A., Chklovskii, D. B., Scheffer, L. K., Rubin, G. M. & Meinertzhagen, I. A. (2017). The comprehensive connectome of a neural substrate for 'ON' motion detection in Drosophila. ELife, 6, e24394. https://doi.org/10.7554/elife.24394

Takemura, S.-Y., Xu, C. S., Lu, Z., Rivlin, P. K., Parag, T., Olbris, D. J., Plaza, S., Zhao, T., Katz, W. T., Umayam, L., Weaver, C., Hess, H. F., Horne, J. A., Nunez-Iglesias, J., Aniceto, R., Chang, L.-A., Lauchie, S., Nasca, A., Ogundeyi, O., … Scheffer, L. K. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(44), 13711–13716. https://doi.org/10.1073/pnas.1509820112

Tan, L., Zhang, K. X., Pecot, M. Y., Nagarkar-Jaiswal, S., Lee, P.-T., Takemura, S.-Y., McEwen, J. M., Nern, A., Xu, S., Tadros, W., Chen, Z., Zinn, K., Bellen, H. J., Morey, M. & Zipursky, S. L. (2015). Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. Cell, 163(7), 1756–1769. https://doi.org/10.1016/j.cell.2015.11.021

Tasic, B., Yao, Z., Graybuck, L. T., Smith, K. A., Nguyen, T. N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M. N., Viswanathan, S., Penn, O., Bakken, T., Menon, V., Miller, J., Fong, O., Hirokawa, K. E., Lathia, K., Rimorin, C., Tieu, M., … Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature, 1–41. https://doi.org/10.1038/s41586-018-0654-5

Turner, N. L., Macrina, T., Bae, J. A., Yang, R., Wilson, A. M., Schneider-Mizell, C., Lee, K., Lu, R., Wu, J., Bodor, A. L., Bleckert, A. A., Brittain, D., Froudarakis, E., Dorkenwald, S., Collman, F., Kemnitz, N., Ih, D., Silversmith, W. M., Zung, J., … Seung, H. S. (2020). *Multiscale and multimodal reconstruction of cortical structure and function*. *72*(5 Pt 2), 056708–056759. https://doi.org/10.1101/2020.10.14.338681

Tuthill, J. C., Nern, A., Holtz, S. L., Rubin, G. M. & Reiser, M. B. (2013). Contributions of the 12 Neuron Classes in the Fly Lamina to Motion Vision. Neuron, 79(1), 128–140.

https://doi.org/10.1016/j.neuron.2013.05.024

Xu, S., Sergeeva, A. P., Katsamba, P. S., Mannepalli, S., Bahna, F., Bimela, J., Zipursky, S. L., Shapiro, L., Honig, B. & Zinn, K. (2022). Affinity requirements for control of synaptic targeting and neuronal cell survival by heterophilic IgSF cell adhesion molecules. *Cell Reports*, *39*(1), 110618. https://doi.org/10.1016/j.celrep.2022.110618

Xu, S., Xiao, Q., Cosmanescu, F., Sergeeva, A. P., Yoo, J., Lin, Y., Katsamba, P. S., Ahlsen, G., Kaufman, J., Linaval, N. T., Lee, P.-T., Bellen, H. J., Shapiro, L., Honig, B., Tan, L. & Zipursky, S. L. (2018). Interactions between the Ig-Superfamily Proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits. Neuron, 100(6), 1369-1384.e6. https://doi.org/10.1016/j.neuron.2018.11.001

Yao, Z., Velthoven, C. T. J. van, Nguyen, T. N., Goldy, J., Sedeno-Cortes, A. E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., Ding, S.-L., Fong, O., Garren, E., Glandon, A., Gouwens, N. W., Gray, J., Graybuck, L. T., Hawrylycz, M. J., Hirschstein, D., … Zeng, H. (2021). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell, *184*(12), 3222-3241.e26. https://doi.org/10.1016/j.cell.2021.04.021

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*(1), 14049. https://doi.org/10.1038/ncomms14049

CHAPTER TWO

Modular transcriptional programs separately define axon and dendrite connectivity.

2.1. Abstract

Patterns of synaptic connectivity are remarkably precise and complex. Single-cell RNA sequencing has revealed a vast transcriptional diversity of neurons. Nevertheless, a clear logic underlying the transcriptional control of neuronal connectivity has yet to emerge. Here, we focused on Drosophila T4/T5 neurons, a class of closely related neuronal subtypes with different wiring patterns. Eight subtypes of T4/T5 neurons are defined by combinations of two patterns of dendritic inputs and four patterns of axonal outputs. Single-cell profiling during development revealed distinct transcriptional programs defining each dendrite and axon wiring pattern. These programs were defined by the expression of a few transcription factors and different combinations of cell surface proteins. Gain and loss of function studies provide evidence for independent control of different wiring features. We propose that modular transcriptional programs for distinct wiring features are assembled in different combinations to generate diverse patterns of neuronal connectivity.

2.2. Introduction

Brain function relies on precise patterns of synaptic connections between neurons. At the cellular level, this entails each neuron adopting a specific wiring pattern, the combination of specific synaptic inputs and outputs. In invertebrates, stereotypical wiring patterns are genetically encoded in the programs regulating the development of neurons. Much of the specificity of inputs and outputs of neurons in the mammalian CNS is also genetically determined (Sanes & Zipursky, 2020).

Vast numbers of neurites from a diversity of neurons are intermingled within the developing central nervous system, and they form highly specific synaptic connections with a discrete subset of the neurons they contact. Studies in both vertebrates and invertebrates have

led to the identification of cell surface proteins (CSPs) that mediate selective association between neurites (Allan & Thor, 2015; Enriquez et al., 2015; Hobert, 2016). Gain and loss of function genetic studies have shown that combinations of different CSP families regulate this specificity (Zarin et al., 2014). Indeed, neuronal subtypes express highly diverse repertoires of CSPs during circuit assembly (Hanqing Li et al., 2017; Sarin et al., 2018; Tan et al., 2015). Conserved regulatory strategies involving combinations of transcription factors (TFs) establish unique neuronal identities (Allan & Thor, 2015; Enriquez et al., 2015; Hobert, 2016). However, the programs regulating expression of CSPs for specific neuronal wiring features are still poorly understood.

Single-cell RNA sequencing (RNA-Seq) provides an unsupervised approach to uncover the genetic programs underlying specific wiring features by exploring subtype-specific transcriptomes during development. As neuronal subtypes exhibit differences in characteristics other than wiring patterns, the relationship between genes and wiring specificity may be obscured by genes contributing to other aspects of neuronal diversity. Therefore, sets of closely related neurons with highly specific differences in wiring patterns are ideally suited to uncover the genetic programs specific to wiring. Here, we explore the genetic logic underlying synaptic specificity in one such set of neurons: T4/T5 neurons of the Drosophila visual motion detection pathway. We envision that our findings in this system will provide insights into the genetic logic of wiring specificity more broadly in both vertebrate and invertebrate systems.

T4/T5 neurons share a common developmental origin, physiological function, and general morphology, but differ in their precise wiring patterns and preferred stimulus (Apitz & Salecker, 2018; Fischbach & Dittrich, 1989; Maisak et al., 2013; Pinto-Teixeira et al., 2018; Shinomiya et al., 2019). There are eight morphological subtypes of T4/T5 neurons in each column of the lobula plate (LoP) neuropil (see below), comprising the most abundant cell type in the fly visual system.

These subtypes can be classified into two quartets of subtypes based on dendritic inputs: the four T4 subtypes share a common set of dendritic inputs in the medulla, and the four T5 subtypes share a different set of dendritic inputs in the lobula (**Figure 2.1A-C**). T4 neurons respond to ON stimuli (i.e. bright edges moving against a dark background) and T5 to OFF stimuli (i.e. dark edges moving across a bright background). T4/T5 neurons can also be classified into four pairs of subtypes (a-d) based on the location of their axon terminals within a given column in layers a-d of the LoP. Each pair responds selectively to visual motion in one of four cardinal directions: posterior, anterior, upwards, and downwards, respectively (**Figure 2.1A-C**). Although transcriptional profiling of the adult Drosophila brain revealed a common transcriptional signature for all T4/T5 neurons, genetic programs for individual subtypes have not been identified (Davie et al., 2018; Konstantinides et al., 2018). We hypothesized that identification of gene expression programs for individual T4/T5 subtypes during circuit assembly would provide insight into the genetic programs regulating discrete wiring features.

Here, we report that independent transcriptional programs define the dendritic inputs and axonal outputs of T4/T5 neurons. We present gain and loss of function studies indicating that these programs control their corresponding morphological features. Our findings suggest that the modular assembly of separate dendritic and axonal transcriptional programs contributes to the diversity of wiring patterns in complex nervous systems.

2.3. Materials and methods

Animal husbandry

Flies (Drosophila melanogaster) were reared at 25℃ on standard medium. For developmental analysis by immunohistochemistry, sorting, and sequencing, white pre-pupae (0 hr APF) were collected and incubated for an indicated number of hours.

Fly stocks

Multiple transgene genotypes are enclosed in brackets. The following transgenic lines were used in this study: MCFO-1 {pBPhsFLP2::PEST;+; UAS-FSF-smGdP::HA_V5_FLAG} (gift from Aljoscha Nern and Gerald Rubin), 10XUAS-IVS-myr::tdTomato (Bloomington Drosophila Stock Center (BDSC#32222)), 23G12-Gal4 (BDSC #49044), {R59E08-p65ADZp (attP40); R42F06-ZpGdbd (attP2)}(JRC_SS00324, Aljoscha Nern and Gerald Rubin), UAS-H2A::GFP (Barret Pfeiffer and Gerald Rubin),UAS-CD4-tdGFP (BDSC #35839), 23G12-LexA (BDSC #65044), LexAop-myr::tdTomato (Zipursky laboratory), 10XUAS-myr::GFP (Zipursky laboratory), 10XUAS-IVS-mCD8::RFP (BDSC #32219), Mi{PT-GFSTF.1}klg[MI02135-GFSTF.1] (BDSC #59787), Mi{PT-GFSTF.1}beat-IV[MI05715-GFSTF.1] (BDSC#66506), dpr2-Gal4 (Zipursky laboratory), P{w[+mW.hs]=GawB}grn[05930-GAL4] (BDSC #42224), Mi{y[+mDint2]=MIC}beat-VI[MI13252] (BDSC #58680), P{y[+t7.7] v[+t1.8]=TRiP.HMS01085}attP2 (UAS-grn-RNAi) BDSC #33746), UAS-grn.ORF.3xHA (FlyORF #F001916), 42F06-Gal4 (BDSC #41253), UAS-p35 (BDSC #5072).

For visualization of T4a clone, virgin females {pBPhsFlp2::PEST; 10XUAS-IVS-myr::tdTomato;UAS-FSF-smGdP::HA_V5_FLAG/CyO::TM6B} were crossed to males with the T4/T5-specific Split-Gal4 driver {R59E08-p65ADZp (attP40); R42F06-ZpGdbd (attP2)}

(JRC_SS00324). White pre-pupae were heat shocked at 37℃ for 3 min. For FACS sorting of GFP+ T4/T5 neurons, 23G12-Gal4 was used to drive UAS-H2A::GFP. For T4/T5 developmental time course, 23G12-Gal4 was used to driveUAS-H2A::GFP and 10XUAS-IVS-mCD8:RFP. For visualization of all T4/T5 neurons with subtype-specific markers, female virgins of the genotypes: {23G12-LexA; LexAop-myr::tdTomato; 10XUAS-myr::GFP/CyO::TM6B} or {w; 10xUAS-IVS-mCD8::RFP; 23G12-Gal4} were crossed to males with MiMICs or their derivatives for klg, beat-IV, dpr2, beat-VI, or to w[1] males for Fas2 immunolabeling. For visualization of grn-expressing neurons, grn-Gal4 was used to drive 10XUAS-myr::GFP. For grn phenotypes, virgin females {w[1];UAS-CD4-tdGFP;23G12-Gal4} were crossed to males with UAS-grn-RNAi or UAS-grn.ORF.3xHA transgenes. For p35 rescue experiments, virgin females with 23G12-Gal4,UAS-CD4-tdGFP and with or without UAS-grn.ORF.3xHA were crossed to males with 42F06-Gal4,and with or without UAS-p35 transgene, as indicated. All RNAi and overexpression crosses were raised at 29℃.

Immunohistochemistry/Immunofluorescence

Brains were dissected in ice-cold Schneider's Drosophila Medium (Gibco #21720–024), and fixed in PBS (Bioland Scientific LLC #PBS01-03) containing 4% paraformaldehyde (Electron Microscopy Sciences, Cat#15710) for 25 min at room temperature (RT). Brains were rinsed repeatedly with PBST (PBS containing 0.5% Triton-X100 (Sigma #T9284)), and incubated in blocking solution (PBST containing 10% Normal Goat Serum (Sigma #G6767)) for at least 1 hr at RT prior to incubation with antibody. Brains were incubated sequentially with primary and secondary antibodies diluted in blocking solution overnight at 4C, with at least 2 PBST rinses followed by 2 hr incubations at RT in between and afterwards. Brains were transferred to 50% (for 30 min), then 100% EverBrite mounting medium (Biotium #23001) and mounted on slides for

confocal microscopy. Primary antibodies and dilutions used in this study were as follows: chicken anti-GFP (abcam #13970, 1:1000), rabbit anti-dsRed (Clontech #632496, 1:200), mouse anti-Brp (nc82 from Developmental Studies Hybridoma Bank (DSHB), 1:30), mouse anti-Fasciclin II (1D4 from DSHB, 1:20), mouse anti-V5 (abcam #ab27671, 1:300), rabbit anti-Dcp-1 (Cell Signaling Technology #9578, 1:50). Secondary antibodies and dilutions used in this study were as follows: goat anti-chicken Alexa Fluor 488 (AF488) (Invitrogen #A11039, 1:200), goat anti-mouse AF488 (Invitrogen #A11029, 1:500), goat antirabbit AF568 (Invitrogen #A11011, 1:200), goat anti-rat 568 (Invitrogen #A11077, 1:500), goat antirabbit AF647 (Invitrogen #A27040, 1:200), and donkey anti-mouse Cy5 (Jackson ImmunoResearch #715-175-150, 1:200).

Confocal microscopy and image analysis

Immunofluorescence images were acquired using a Zeiss LSM 880 confocal microscope with Zen digital imaging software. Optical sections or maximum intensity projections were level-adjusted, cropped, and exported for presentation using Image J software (Fiji). Reported expression patterns were reproducible across three or more biological samples. For cell number quantifications, optic lobes were mounted with ventral side facing objective (as in Figure 2.1), and a single optical section per lobe was acquired at 3/8 total depth in z-dimension through M10. For quantification of apoptosis, optic lobes were mounted with posterior side facing objective, and a superficial optical section with approximately 300 T4/T5 cell bodies was acquired per lobe. The section depth was determined with Dcp-1 immunofluorescence channel turned off. Files were randomized, and cell numbers and proportion of apoptotic cells were quantified blind to condition using Fiji.

Single-cell transcriptome profiling

Purification of genetically labeled T4/T5 neurons Males with 23G12-Gal4 driver were crossed to virgin females with UAS-H2A::GFP reporter. F1-generation female white pre-pupae were collected at 0 hr APF and reared at 25℃. Optic lobes were dissected out at 24 hr and 48 hr APF from 27 and 18 pupae, respectively. Brain tissue was incubated in papain (Worthington #LK003178) and Liberase protease (Sigma-Aldrich #5401119001) cocktail at 25℃ for 15 min. Next, tissue was gently washed twice with 1X PBS and dissociated mechanically by pipetting. Cell suspension was filtered through 20mm cell-strainer (Corning #352235). Single-cell suspension was FACS sorted (BD FACSAria II) to isolate GFP-positive cells. Single-cell library preparation and sequencing FACS-sorted single-cells were captured from a cell suspension using the 10X Chromium platform(~6500–7000 cells loaded). Single-cell RNA-Seq libraries were

generated using Chromium Single Cell Reagent Kit V2 according to the manufacturer's protocol, with 12 cycles of PCR for cDNA amplification. RNA-Seq libraries were sequenced using Illumina Hiseq 4000 platform (paired-end 100 bp reads). Each sample was captured and sequenced using one lane of 10X Chromium and one lane of HiSeq 4000.

Raw data processing

Raw Illumina base call files (*.bcl files) were converted into fastq files using bcl2fastq (–use-basesmask=Y26 n*,I8n*,Y100n*). Fastq files were processed using Cell Ranger (2.2.0) pipeline with default parameters. Reference transcriptome package for Cell Ranger was generated using Drosophila melanogaster genome sequence and gene annotations from FlyBase (release 6.22). Both samples were sequenced at mean depth of 92,000 reads per cell (92% saturation). Average fractions of reads uniquely (confidently) mapped to genome and transcriptome were 93% and 83%, respectively.

Single-cell data analysis

All steps of single-cell data analysis were performed using functions and methods implemented in Seurat package (2.3.4) (Butler et al., 2018). Analysis for 24 hr and 48 hr datasets were performed separately.

Quality control and data pre-processing For 24 hr dataset, we recovered 3833 cells (median of 1447 genes and 3353 transcripts per cell). For 48hr dataset, we recovered 3894 cells (median of 1633 genes and 4389 transcripts per cell). Initial set of cells was pre-filtered based on total number of detected genes (min. 1000; max. 2000), and percentage of mitochondrial transcripts (max. 5%). After pre-filtering, 3312 and 3620 cells remained for 24hr and 48hr datasets,

respectively. Raw transcript counts were log-normalized using NormalizeData function. Next, we regressed out total number of transcripts per cell (nUMI) and scaled expression values to Z-scores using ScaleData function.

Preliminary dimensionality reduction and detection of outlier cells Sets of highly variable genes were selected using FindVariableGenes function (x.low.cutoff: 0.1, x. high.cutoff: 5, y.cutoff: 0.5). Highly variable genes were used to perform independent component analysis (ICA) using RunICA function. Independent components (ICs) were manually inspected to identify and flag outlier cells. In total, 241 and 60 cells were flagged as outliers in 24 hr and 48 hr datasets, respectively. Outlier cells were removed from subsequent steps of the analysis. After quality control and filtering, final datasets included 3071 cells for 24 hr, and 3557 cells for 48 hr datasets.

Dimensionality reduction and clustering (48hr APF) We repeated selection of highly variable genes on final datasets using the same parameters (2290 genes), and used them to perform ICA. Inspection of results of ICA revealed that the three ICs separated cells into two discrete populations of approximate halves. Final clustering was performed based on these 3 ICs using the graph-based clustering approach implemented in FindClusters function with default parameters. In addition to ICA, we performed principal component analysis (PCA) using the same set of highly variable genes. Comparison of ICA and PCA results revealed robustness of clusters identified by both methods (Figure 2.1—Figure 2.S1). t-distributed stochastic neighbor embedding (tSNE) was used to visualize cellular heterogeneity based on ICA and PCA results using RunTSNE function (perplexity: 100). Clusters were validated and matched to eight morphological T4/T5 subtypes using in vivo expression patterns of marker genes (Figure 2.3).

Dimensionality reduction and clustering (24hr APF) Similar to 48hr dataset, we selected highly variable genes (2198 genes), and used them to perform ICA. We selected three ICs that were driven by similar sets of genes as ICs used for clustering of 48hr dataset. Selected ICs were used to perform clustering and tSNE using same parameters as 48hr dataset. Cluster identities were matched with T4/T5 subtypes using expression patterns of the same sets of marker genes (Figure 2.S4). In comparison to 48hr dataset, differences between subtypes at 24hr were less pronounced. This may reflect a lower degree of transcriptional divergence among distinct subtypes at earlier stages of development.

Differential gene expression analysis Differentially expressed genes (DEGs) were identified using Wilcoxon rank-sum test implemented in FindMarkers function (min.pct: 0.25, min.diff.pct: 0.25; fold-change: 1.5). We performed this analysis for each cluster against all other cells in dataset (‘one versus all’, Figure 2.1), and between pairs of individual clusters (‘one versus one’, Figure 2.4 and Figure 2.S3.

2.4. Results

2.4.1. Single-cell RNA-Seq reveals eight transcriptionally distinct populations of T4/T5 cells.

As a step towards uncovering genetic programs that control neuronal wiring patterns, we performed single-cell RNA-Seq on developing T4/T5 neurons. Sequencing was performed at 48hr after puparium formation (APF). This developmental time point precedes a period of widespread synaptogenesis in the visual system, and coincides with the appearance of four discrete synaptic layers (a, b, c, d) in the LoP neuropil. Neurons were purified from dissected optic lobes by FACS using a transgenic line with nuclear GFP selectively expressed in all T4/T5 neurons (**Figure 2.1D**). RNA-Seq libraries were generated using 10X Chromium technology (Zheng et al., 2017) and sequenced to a mean depth of 92,000 reads per cell. In total, we profiled 3894 cells with a median of 1633 genes and 4389 transcripts captured per cell. After quality control and removal of outlier cells, our final dataset consisted of 3557 cells with 1000–2000 genes per cell.

We applied independent component analysis (ICA) followed by a graph-based clustering method to separate transcriptionally distinct cell populations (Butler et al., 2018; Saunders et al., 2018). Unsupervised analysis revealed eight clusters of approximately equal numbers of cells (**Figure 2.1E** and **Figure 2.S1**), suggesting that each cluster corresponded to a single T4/T5 subtype.

To identify genes preferentially expressed in T4/T5 subtypes, we performed differential gene expression analysis between each of the eight individual clusters and all other cells in the dataset (i.e. ‘one versus all,’ see Materials and methods). This revealed 69 genes which were strongly expressed in some clusters and not in others. Cluster-enriched genes, however, were not specific to single clusters. By contrast, for instance, each of the five subtypes of lamina neurons is defined by at least one subtype-type specific transcription factor (Tan et al., 2015).

Thus, while T4/T5 subtypes separated into eight transcriptionally distinct clusters, they were not defined by unique molecular markers (**Figure 2.1F**).

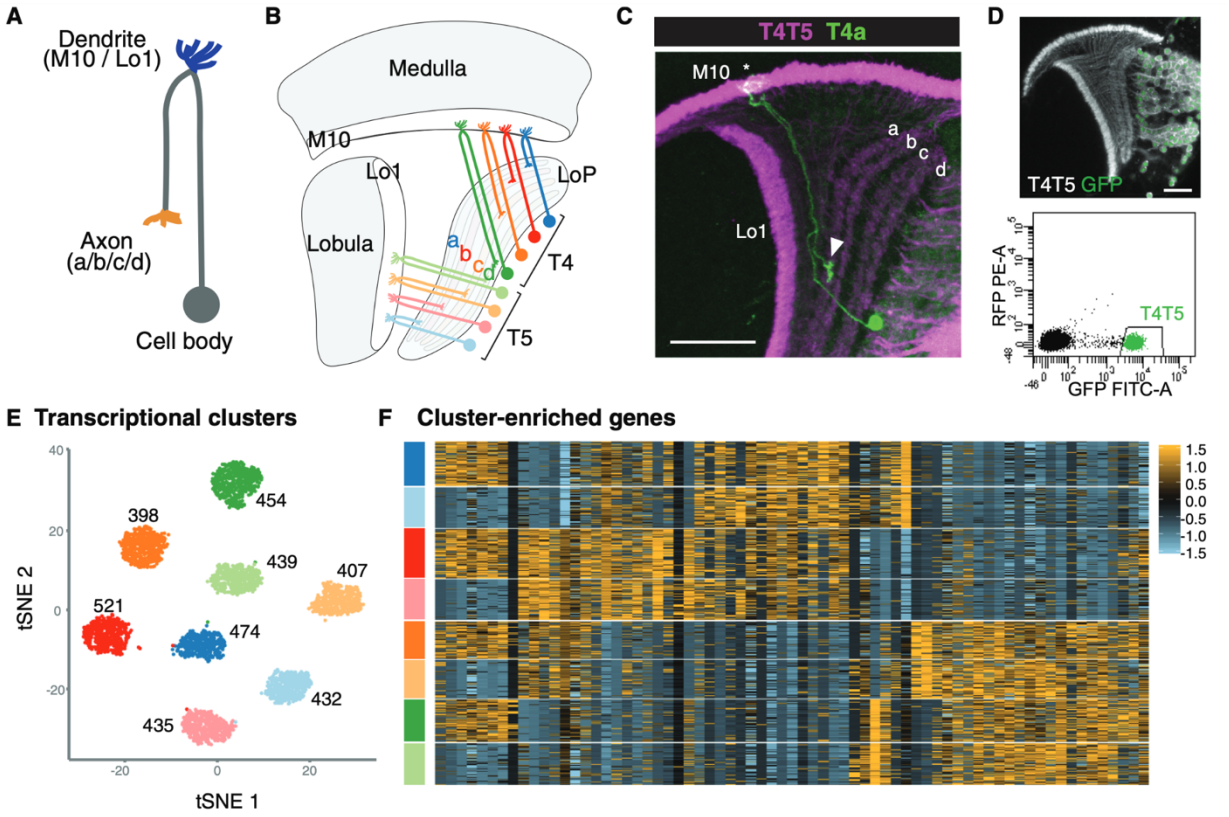


Figure 2.1. Single-cell sequencing reveals eight transcriptionally distinct populations of T4/T5 neurons. (A) Common morphology of a T4/T5 neuron, with axon and dendrite wiring pattern variations in parentheses. (B) Arrangement of the eight T4/T5 subtypes in the optic lobe. Each subtype is defined by a combination of one dendrite (M10 or Lo1) and one axon (LoP a, b, c, or d) wiring pattern. (C) A single T4a neuron (green) with dendrites in M10 (asterisk) and axon terminal in LoP layer a (arrowhead). All T4/T5 neurons labeled in magenta. Scale bar, 20 mm. (D–F) Single-cell sequencing of T4/T5 neurons at 48 hr APF. Unsupervised analysis revealed eight distinct transcriptional clusters. (D) T4/T5 neurons were labeled with nuclear GFP, purified by FACS and used for single-cell RNA-Seq. (E) t-distributed stochastic neighbor embedding (tSNE) plot of 3557 single-cell transcriptomes. Clusters are color-coded according to subtype identity based on following results. Cell numbers are displayed for each cluster. See also Figure 2.S1. (F) Heatmap of expression patterns of cluster-enriched genes ('one versus all', see Materials and methods). Cells (rows) grouped by cluster identities as in (E). Genes (columns) are ordered by similarity of their expression patterns. Scaled expression levels are indicated, as in scale.

2.4.2. Eight T4/T5 transcriptional clusters are separated by three primary axes of transcriptional diversity.

The absence of unique markers for individual subtypes suggested that they were instead defined by unique combinations of genes. ICA has been shown to capture groups of genes corresponding to discrete biological phenomena (Saunders et al., 2018). Intriguingly, three independent components (ICs) each split the eight T4/T5 clusters into two groups of four, each in a different way (**Figure 2.2A**). Together, these three ICs were sufficient to define all eight clusters (**Figure 2.2B**).

Each of the three ICs defined an axis of transcriptional diversity (hereafter referred to as Axis 1, 2, 3) driven by a group of genes differentially expressed along each axis. Many of these genes were expressed in a binary (ON/OFF) pattern in one of the two groups of clusters separated along each axis (**Figure 2.2C**). A small set of TFs were among the genes with the highest contributions to each axis and illustrate this pattern. Binary expression of bifid (bi) defined the two groups of clusters separated by Axis 1, grain (grn) defined the clusters separated by Axis 2, and TfAP-2 defined the clusters separated by Axis 3 (**Figure 2.2D**). Thus, while no individual cluster is uniquely defined by the expression of a single gene, each cluster expresses a unique combination of genes. In this way, three axes of diversity with orthogonal ON/OFF expression patterns of TFs define the eight T4/T5 clusters.

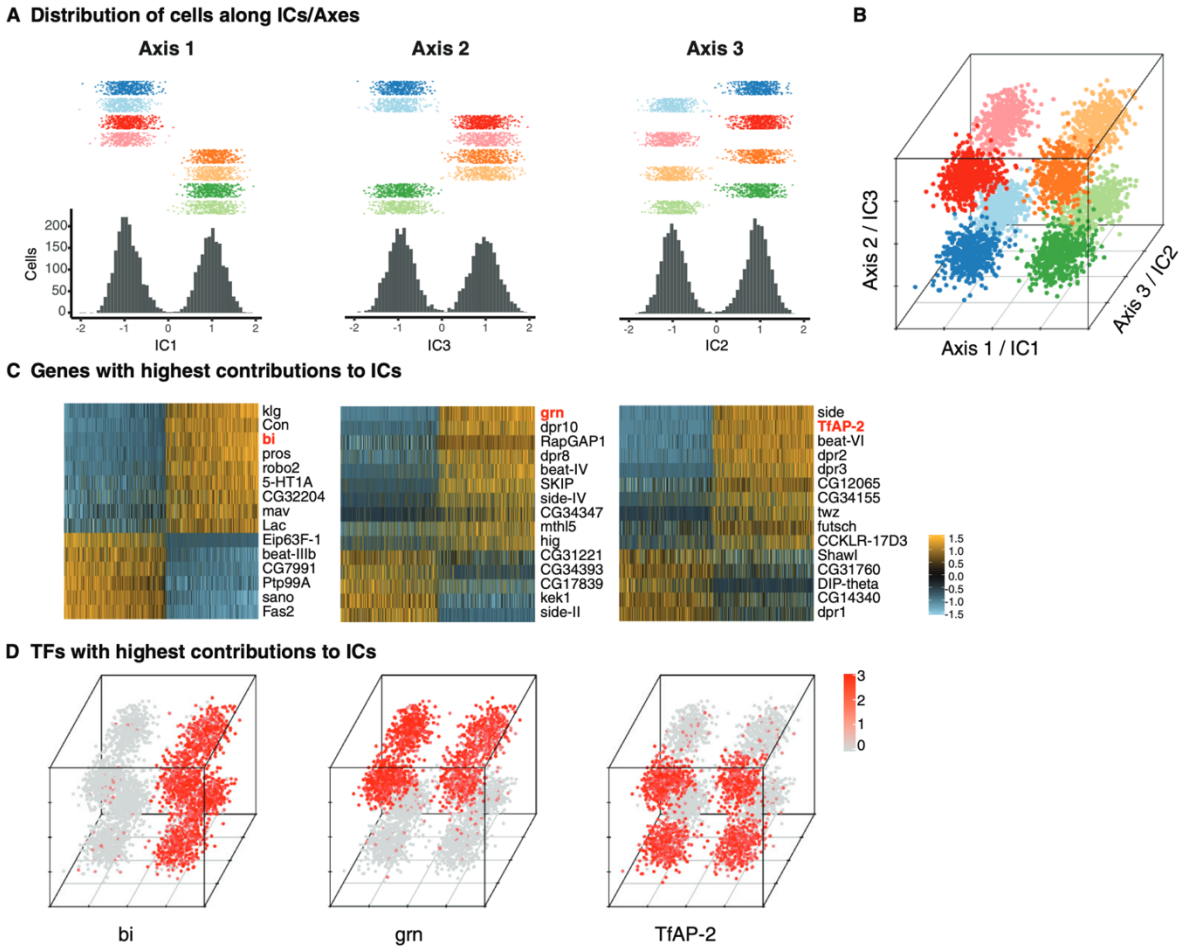


Figure 2.2. Three primary axes of transcriptional diversity define eight T4/T5 populations. (A) Three independent components (ICs, henceforth Axis 1, 2, 3) separate cells into approximate halves. Histograms (bottom) and 1-D scatterplots (top) show the distributions of cells along each axis. Cells are grouped into rows based on cluster identities. ICs/Axes are ordered according to following results. Clusters are color-coded as in Figure 2.1E. See also Figure 2.S1. (B) 3-D scatterplot of the distributions of cells along the three ICs/Axes. (C) Heatmaps of expression patterns of the top 15 genes with highest contribution (loading) to each IC/Axis. Cells (columns) are ordered according to a score for each IC/Axis. Genes (rows) are ordered according to the contribution to each IC/Axis. Scaled expression levels are indicated, as in scale. Axes ordered as in (A). (D) 3-D scatterplots with expression patterns of transcription factors (TFs) with highest contribution to each IC/Axis. Normalized expression levels are indicated by color, as in scale. Axes are arranged as in (B).

2.4.3. Primary axes of transcriptional diversity correspond to axon and dendrite wiring patterns.

We next sought to map transcriptional clusters to T4/T5 subtypes, and to determine the biological significance of the observed axes of transcriptional diversity. We inspected in vivo expression patterns of genes associated with the three primary axes of diversity using transgenic reporters inserted into the endogenous loci (Venken et al., 2011).

Axis 1 separated clusters into two groups of four that were defined by mutually exclusive binary expression of two genes, Fasciclin 2 (Fas2) and klingon (klg), respectively (**Figure 2.3A**), each encoding immunoglobulin (Ig) superfamily proteins. Fas2 was expressed in LoP layers a/b, whereas klg was expressed in LoP layers c/d (**Figure 2.3B**). Clusters expressing Fas2 and klg also expressed previously described markers for T4/T5 subtypes a/b (dachshund (dac)) and c/d (bi, Connectin (Con)) (Apitz & Salecker, 2018)(**Figure 2.S2**). Thus, Axis 1 separated LoP layer a/b and c/d subtypes, defining specificity of axonal outputs between two broad domains of the LoP (**Figure 3C**). This corresponds to separation of horizontal (posterior/anterior) and vertical (upwards/ downwards) motion detection circuits, respectively.

Axis 2 separated clusters into two groups of four defined by binary expression of beat-IV (an Ig superfamily protein) and the TF grn. Both genes were expressed in LoP layers b/c, but not a/d. Thus, Axis 2 further separated subtypes into inner (b and c) and outer (a and d) LoP layer subtypes in a symmetrical fashion, defining specificity of axonal outputs between adjacent layers within the two broad domains of the LoP (**Figure 2.3A'–2.3C'**). This corresponds to further separation of each of the motion detection circuits into two subcircuits detecting motion in two opposing directions (i.e. horizontal into posterior and anterior, and vertical into upwards and downwards motion).

Axis 3 separated clusters into two groups of four defined by binary expression of dpr2 and beat-VI, each encoding an Ig superfamily protein. In vivo, both genes were expressed in all LoP layers, and M10 but not Lo1. Thus, Axis 3 separated all T4 from all T5 subtypes, defining specificity of dendritic inputs (**Figure 2.3A''–2.3C''**). This corresponds to separation into two parallel circuits for ON and OFF motion detection, respectively.

Taken together, three primary axes of diversity defined distinct wiring features of T4/T5 subtypes and in combination defined wiring patterns of each T4/T5 subtype (**Figure 2.3A–2.3A''**). A combination of Axis 1 and Axis 2 defined four types of axonal outputs (a, b, c, d), and Axis 3 defined two types of dendritic inputs (T4 and T5).

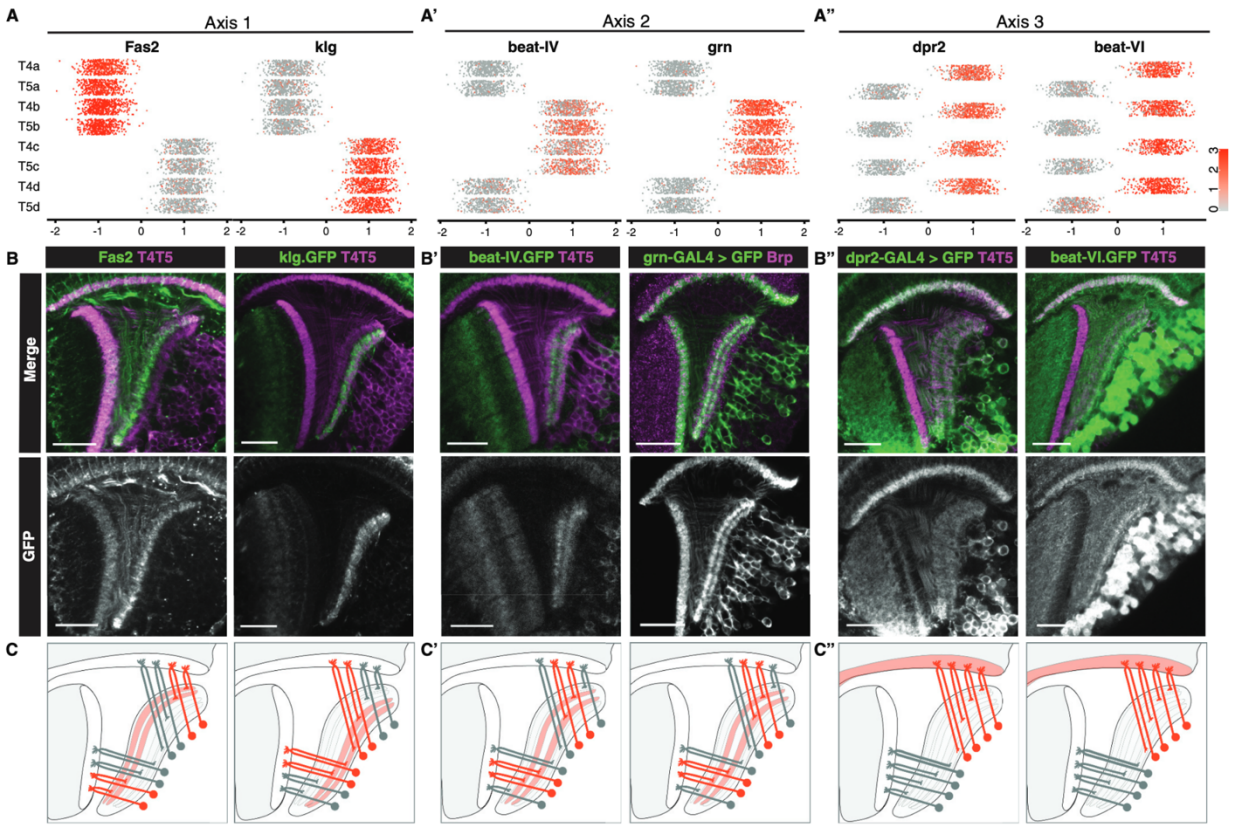


Figure 2.3. Primary axes of transcriptional diversity define groups of T4/T5 subtypes with shared wiring patterns.

(A–A") 1-D scatterplots show distribution of cells along Axis 1, 2, and 3 for each cluster. Normalized expression levels are indicated by color, as in scale. (B–B") In vivo expression of marker genes for each axis at 48 hr APF. Fas2 labels LoP layers a/b, klg labels LoP layers c/d, beat-IV and grn label LoP layers b/c, dpr2 and beat-VI label M10 but not Lo1. Scale bars, 20 mm. Sets of positive clusters in (A) are matched to specific sets of T4/T5 subtypes based on in vivo expression patterns in (B). Individual cluster identities are deduced based on combination of expression patterns. For example, T4a is Fas2+ (a/b), beat-IV- (not b/ c), dpr2+ (M10). (C–C") Schematic of wiring patterns of T4/T5 subtypes corresponding to the expression patterns of marker genes (red). See also Figure 2.S2.

2.4.4. Transcriptional program of a single T4/T5 subtype

In addition to the transcriptional differences between groups of T4/T5 subtypes described above, further variation might exist at the individual subtype level. To examine this possibility, we focused on a single subtype (T4a) and performed comprehensive pairwise comparisons with each of the other subtypes (i.e. 'one versus one').

First, we compared T4a and each subtype that differed by a single wiring feature: either axonal outputs (T4b, T4c, T4d), or dendritic inputs (T5a) (**Figure 2.4**, upper dot plots). Comparison of T4a to T4c or T4d, which have axonal outputs in non-adjacent LoP layers, yielded the largest number of differentially expressed genes (DEGs, 58 and 55, respectively). T4a and T4b have axonal outputs in adjacent LoP layers, and were separated by an intermediate number of DEGs (16). Finally, only a small number of DEGs (9) separated T4a and T5a, which share axonal output but receive different dendritic inputs.

Expression patterns of the DEGs from pairwise comparisons across all T4/T5 subtypes revealed a general pattern (**Figure 2.4**, lower dot plots): virtually all DEGs were co-regulated across all subtypes according to either specificity of axonal outputs or dendritic inputs. In other words, distinct sets of DEGs were expressed in each pair of subtypes with shared axonal outputs, but different dendritic inputs (e.g. in T4a and T5a, **Figures 2.4A–C**). Similarly, a distinct set of DEGs was expressed in groups of subtypes with shared dendritic inputs, but different axonal outputs (i.e. all T4 subtypes, **Figure 2.4D**).

The co-regulation of DEGs according to wiring patterns was not limited to comparisons between subtypes that differed by a single wiring feature. DEGs between T4a and subtypes that differed by both axon and dendrite wiring patterns (e.g. T5c), were also expressed in groups of subtypes sharing either axonal outputs or dendritic inputs (**Figure 2.S3**).

In addition to three primary axes of transcriptional diversity, this analysis shows that a number of DEGs exhibited more distinct LoP layer-specific patterns. For example, many DEGs were specifically expressed or suppressed in T4/T5 subtypes from a single LoP layer (**Figure 2.4C**).

Many of the DEGs have been implicated in neuronal wiring specificity (**Figure 2.4**). Approximately half of the DEGs encoded CSPs with cell adhesion domains (Ig/LRR), including multiple members of the dpr/DIP and beat/side families of interacting proteins (Apitz & Salecker, 2018; Zinn & Özkan, 2017); specific members of these families have been shown to regulate axon guidance and synaptic specificity in the developing fly nervous system.

Taken together, these results reveal that the transcriptional organization of T4/T5 neurons mirrored their wiring patterns. Discrete groups of co-regulated genes reiteratively defined either shared axon or shared dendrite wiring patterns among different subtypes. These groups of genes were assembled in different combinations to uniquely define the eight T4/T5 subtypes.

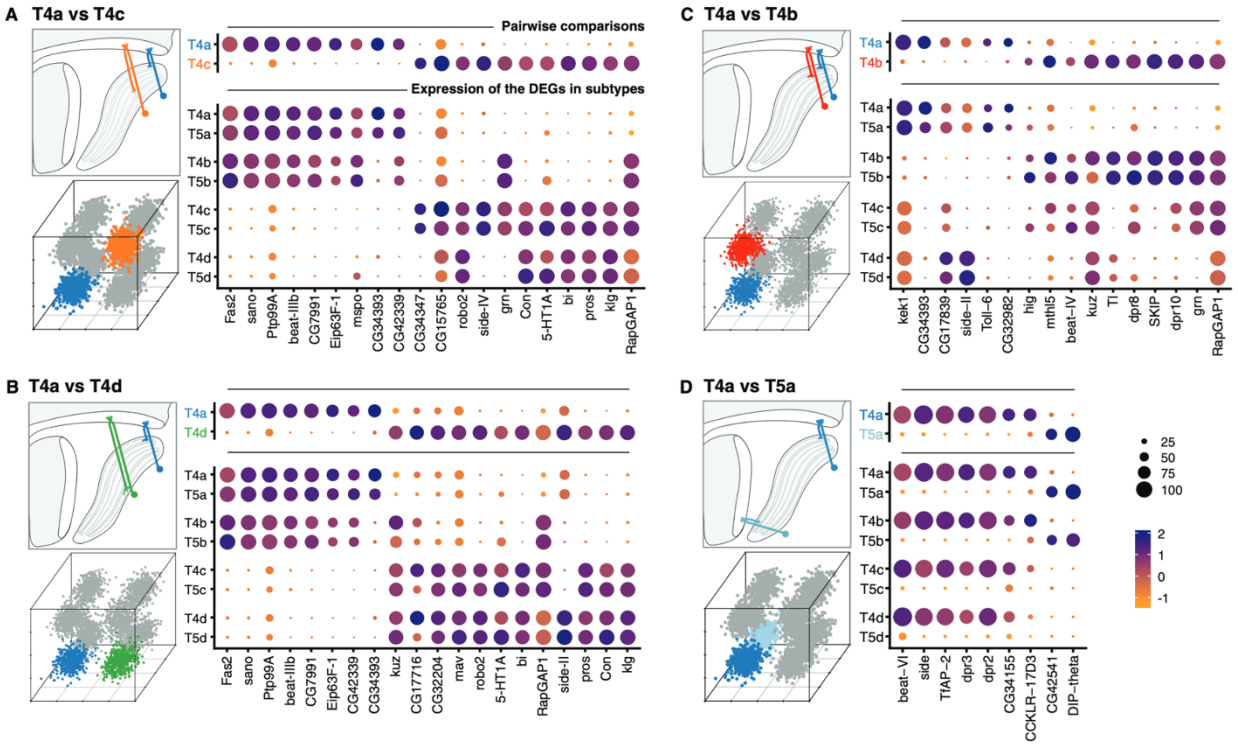


Figure 2.4. Transcriptional program of a single T4/T5 subtype.
Pairwise comparisons between T4a and other subtypes ('one versus one', see Materials and methods) that differ by either axonal outputs (A–C), or dendritic inputs (D). For each comparison, insets indicate morphologies (upper left) and cluster distributions along axes of transcriptional diversity (lower left). Expression patterns of differentially expressed genes (DEGs) for each pairwise comparison are shown in upper right. Dot size indicates the percentage of cells in which the DEG was detected, color represents average scaled expression, as in scale. Genes are ordered by fold-change values. Top 20 DEGs are shown for (A) and (B). Expression patterns of DEGs among all eight subtypes are shown in lower right. See also Figure 2.S3.

2.4.5. Stable and dynamic features of T4/T5 transcriptional programs during development

To evaluate how gene expression in T4/T5 subtypes changes during development, we profiled T4/ T5s at an earlier time point, 24hr APF. Similar to our dataset at 48hr APF, we identified eight distinct populations separated by three equivalent axes of transcriptional diversity. Many of the same genes were associated with these axes at both time points, allowing us to match subtypes between 24hr and 48hr APF (**Figure 2.S7**).

Comparison of 24hr and 48hr datasets revealed stable and dynamic features of T4/T5 transcriptional programs. TFs defining the primary axes of diversity (bi, grn, TfAP-2) were expressed in the same sets of subtypes at both time points, suggesting they may contribute to stable subtype identities during development (**Figure 2.5A**). Some CSPs also exhibited stable expression, marking subtypes with shared axon or dendrite wiring patterns at both time points. Other CSPs were dynamically regulated and were specific to subtypes only at a particular stage of development (**Figures 2.5B-D** and **Figure 2.S8**).

Interestingly, dynamic changes in gene expression were also coordinated among subtypes with shared wiring features (**Figures 2.5B–D**). For example, dpr3 and a few other CSPs were synchronously upregulated in all T4 subtypes from 24 hr to 48 hr APF. Similarly, Toll-6 was synchronously upregulated in both LoP layer ‘a’ subtypes (T4a/T5a). These data indicate that similar transcriptional programs unfold in parallel among T4/T5 subtypes with shared wiring features.

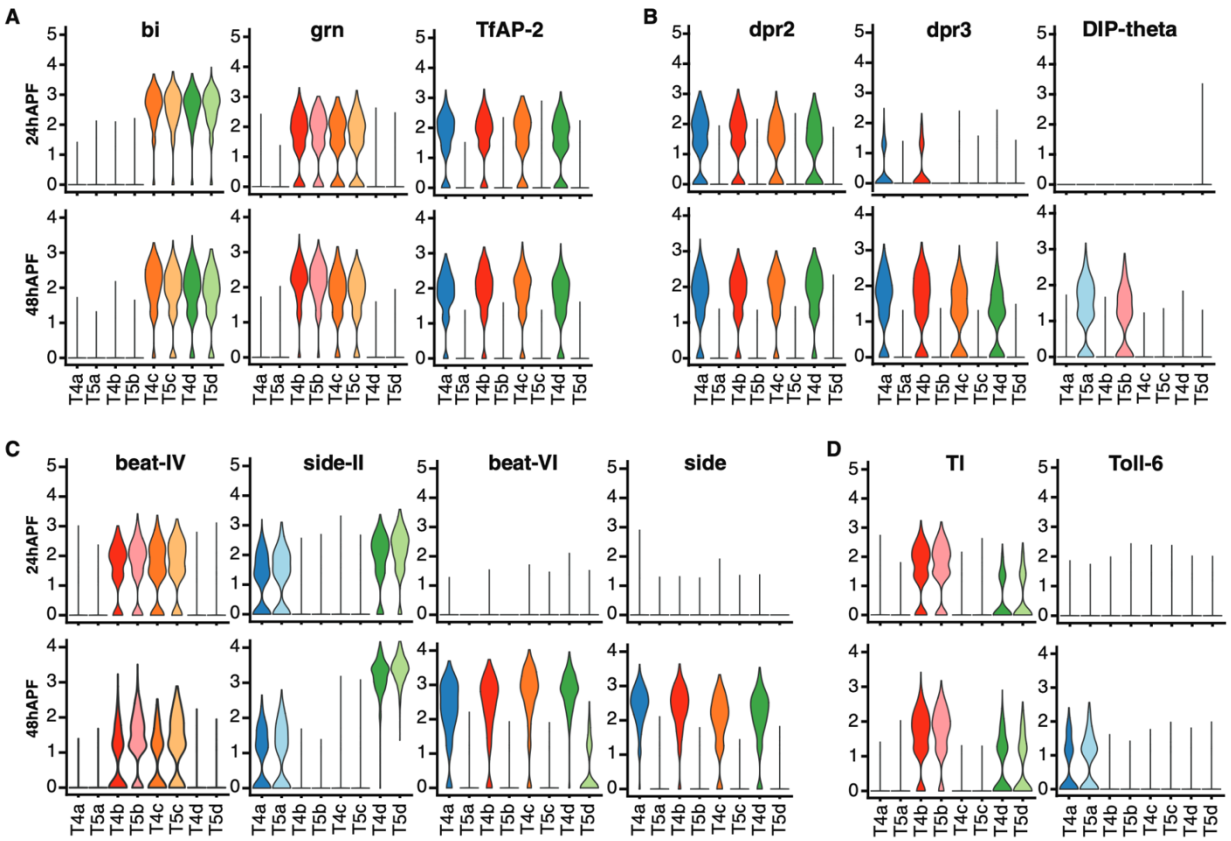


Figure 2.5. Dynamics of T4/T5 transcriptional programs during development. Distributions of normalized expression levels of TFs (A) and selected families of CSPs (B-D) at 24 hr and 48 hr APF. Distributions for each subtype are color-coded as in Figure 2.S4.

2.4.6. Axon-specific transcriptional programs of T4/T5 neurons control lamination of LoP layers.

A remarkable correspondence between transcriptional programs and wiring patterns suggested that these programs control development of corresponding features of T4/T5 neurons. During the period covered in our study (24–48 hr APF) four LoP layers form in two discrete lamination steps (**Figure 2.S5**). The T4/T5 axon terminals first laminate into two broad domains corresponding to layers a/b and c/d. These domains then further sublaminate into two pairs of adjacent layers to form the four discrete LoP layers, a, b, c, and d. The two primary axes of transcriptional diversity mirrored these two stages of LoP layer formation, suggesting a regulatory code for axon wiring (**Figure 2.3**). Mutually exclusive expression of the TFs dac and bi/pros separated a/b (dac+) from c/d (bi+/ pros+) subtypes (i.e. Axis 1, Apitz & Salecker, 2018). These subtypes were further separated by expression of the TF grn into inner (b and c, grn+) and outer (a and d, grn-) layer subtypes in a symmetrical fashion (i.e. Axis 2). This suggested that two levels of transcriptional regulation, acting either sequentially or in a temporally overlapping way, control development of four types of T4/T5 axonal outputs.

We sought to experimentally address this issue. Previous studies indicate that bi specifies c/d subtypes and formation of corresponding LoP layers. RNAi of bi in all T4/T5 neurons results in loss of the c/d domain of the LoP, whereas overexpression results in loss of the a/b domain. In both cases, further sublamination of remaining inner and outer LoP layer pairs still occurs (Apitz & Salecker, 2018). We performed RNAi of grn, which resulted in a different phenotype; whereas distinct a/ b and c/d LoP domains were still separated by a pronounced gap and differential expression of Con (a marker for LoP layers c/d), both domains failed to sublaminate into inner and outer layers, instead forming a single layer each (**Figure 2.6A–B**). The overall morphological organization of T4/T5 neurons was otherwise unaffected. Overexpression of grn in all T4/T5 neurons also resulted in a specific failure of a/b and c/d LoP domains to sublaminate.

RNAi and overexpression of grn resulted in significant loss of T4/T5 neuron numbers between 24 hr and 48 hr APF (**Figure 2.S6**), associated with an increase in apoptosis (**Figure 2.6C–D**). Failure of LoP layer sublamination could result from death of specific subtypes during development. Alternatively, differential expression of grn might be required to direct T4/T5 axons to discrete layers. Expression of baculovirus caspase inhibitor p35 in developing T4/T5s rescued cell death associated with grn overexpression. Nevertheless, T4/T5 axons still failed to sublaminate into four discrete LoP layers (**Figure 2.6E–F**). Thus, differential expression of grn is specifically required for sublamination of T4/T5 axons between pairs of adjacent LoP layers.

We conclude that axon-specific transcriptional programs defined by binary (ON/OFF) expression patterns of two TFs, bi and grn, control the formation of four LoP layers and corresponding T4/T5 axonal wiring patterns.

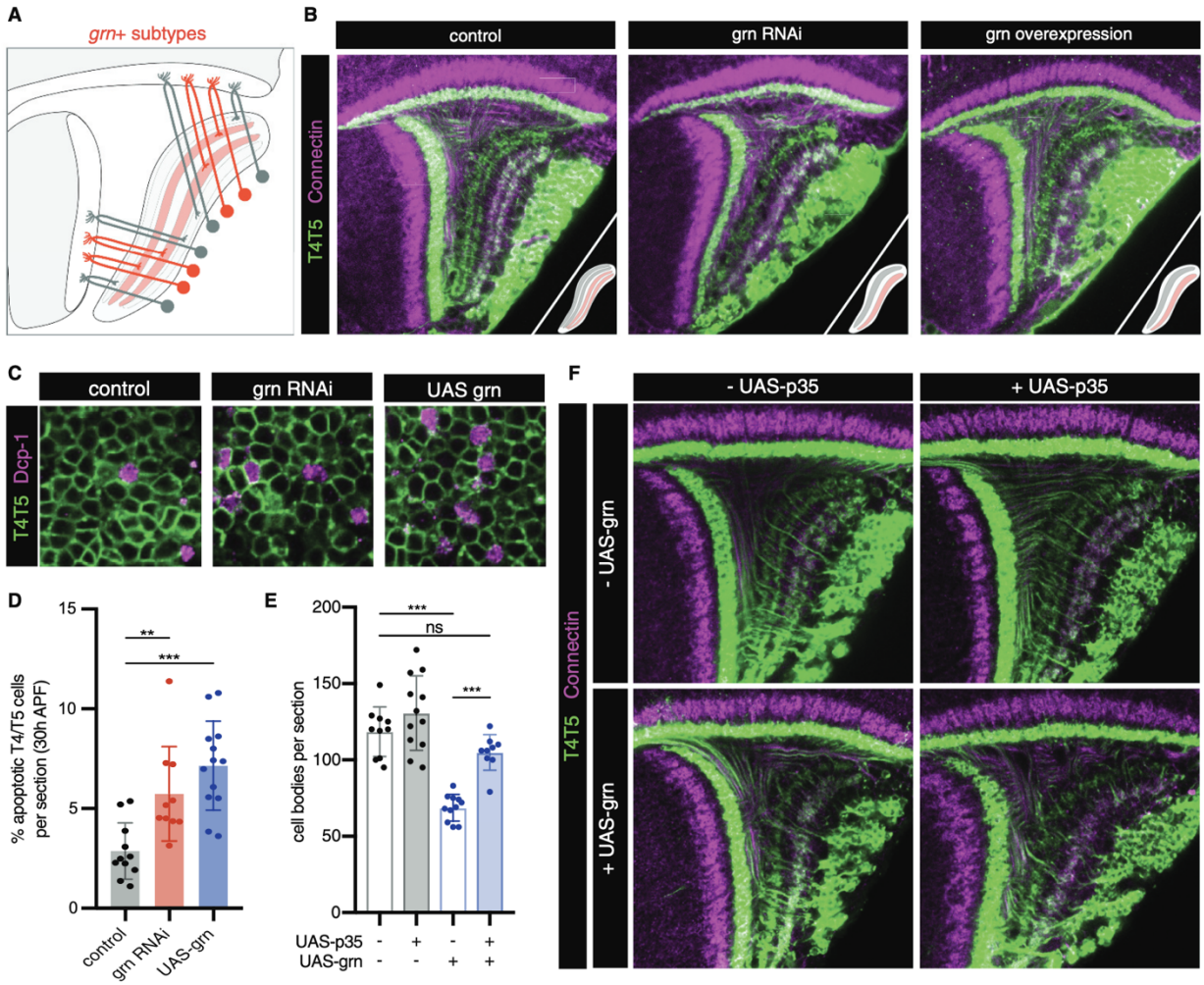


Figure 2.6. grn controls sublamination of T4/T5 axons into inner and outer LoP layers.
(A) Schematic of grn+ (red) T4/T5 subtypes in wild-type optic lobe. grn expression defines inner LoP layer subtypes. See also Figure 6—figure supplement 1. (B) grn RNAi and grn overexpression in all T4/T5 neurons specifically disrupts sublamination of a/b (Con-) and c/d (Con+) LoP subdomains into inner and outer layers. Insets depict LoP phenotypes. (C–D) Immunostaining for Death caspase-1 (Dcp-1) reveals increased apoptotic T4/T5 neurons under grn RNAi and overexpression (UAS-grn) conditions at 30 hr APF. See also Figure 2.S6, and Figure 6—source data 1. (E) Ectopic expression of p35 in T4/T5 neurons (UAS-p35) rescues apoptotic cell death associated with overexpression of grn. (F) grn overexpression specifically disrupts axon sublamination when apoptosis is blocked. Statistical significance assessed by one-way ANOVA with Tukey's multiple comparison test (**p<0.01, ***p<0.001). Bars and whiskers represent mean and standard deviation. Dots represent values for individual optic lobes.

2.5. Discussion

Single-cell transcriptional profiling has the potential to transform our understanding of the genetic programs controlling wiring in complex nervous systems (Klingler et al., 2018; Hongjie Li et al., 2017; Tasic et al., 2018). However, neurons exhibit a vast diversity of wiring patterns, morphologies, and molecular identities, making it difficult to extract the transcriptional logic underlying specific wiring features. Here, we turned to the closely related T4/T5 subtypes of the Drosophila visual system, which differ by specific variations in wiring, with the expectation that transcriptional differences among them would reflect the specificity of dendritic inputs and axonal outputs. A unique attribute of T4/T5 neurons is that the same dendritic and axonal wiring patterns are reiteratively used among different subtypes; each neuron can be described by a unique combination of one of four types of axonal outputs and one of two types of dendritic inputs. We anticipated that this property of T4/T5 neurons would provide an opportunity to assess the relationship between specific genetic programs and fundamental features of neuronal architecture.

Unsupervised analysis revealed that separable transcriptional programs correlate with these specific wiring features. We demonstrate through gain and loss of function experiments that these programs control specific axonal targeting features, which are separable from other features (e.g. dendrite targeting). These programs can be re-assembled in a modular fashion to generate neuronal subtypes with different combinations of wiring features. A modular transcriptional architecture may provide a general strategy for discrete modifications to neuronal connectivity in development and evolution.

A common T4/T5 neuronal identity is defined by a unique combination of TFs expressed in all subtypes (e.g. Lim1, Drgx, acj6) (Davie et al., 2018; Konstantinides et al., 2018). Perturbation of TFs expressed in all subtypes disrupts overall organization of T4/T5 neurons, including both dendritic and axonal morphologies(Contreras et al., 2018; Schilling et al., 2019). We find that this common T4/T5 transcriptional program is further diversified by separable feature-specific

transcriptional programs. These programs are defined by three binary (ON/OFF) TF expression patterns, with two TF patterns defining the four axonal outputs and one TF pattern defining the two dendritic inputs. In this way, modular TF codes defining common and feature-specific transcriptional programs give rise to eight T4/T5 subtypes (**Figure 2.7**).

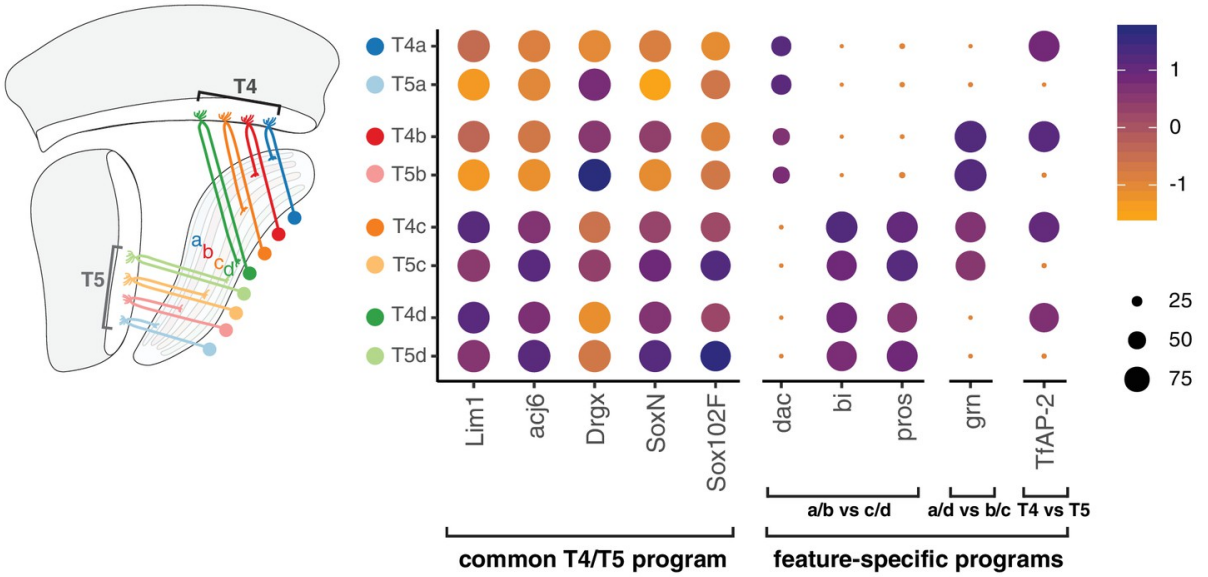


Figure 2.7. Modular transcription factor codes define eight T4/T5 subtypes.
A common T4/T5 regulatory program is defined by TFs expressed in all subtypes (Davie et al., 2018; Konstantinides et al., 2018; Contreras et al., 2018; Schilling et al., 2019). This program is diversified by modular combinations of feature-specific TFs defining unique wiring patterns of eight T4/T5 subtypes. Dot size indicates the percentage of cells in which the TF was detected, color represents average scaled expression, as in scale. Data shown for 48 hr APF. See also Figure 2.S7 and 2.S8

Four pairs of T4/T5 subtypes with shared axonal outputs (and different dendritic inputs) each target one of four LoP layers, a-d. The ultimate layered architecture of neuropils develops through sequential lamination into increasing numbers of layers (Millard & Pecot, 2018; Sanes & Zipursky, 2010). Together with previous results, our findings suggest that the lamination of T4/T5 axonal outputs occurs via two distinct processes, each controlled by a separate TF. Binary expression of bi is required for lamination of the broad a/b from c/d LoP domains (Apitz & Salecker, 2018), whereas binary expression of grn is required for sublamination of each of these two domains into separate LoP layers. Importantly, perturbation of each TF exclusively disrupts the corresponding lamination step, while not affecting other morphological features of T4/T5 neurons. Similarly, two quartets of subtypes with shared dendritic inputs (and different axonal outputs) were defined by binary expression of TfAP-2. Arborization of dendrites in M10 (T4) or Lo1 (T5) occurs during initial neurite guidance steps, preceding the developmental stages covered in this study (Pinto-Teixeira et al., 2018). We hypothesize that DEGs between T4 and T5 subtypes identified in our analysis contribute to the connections with two distinct sets of presynaptic partners (Shinomiya et al., 2019).

The binary expression patterns of TFs also mirror the developmental lineages of T4/T5 neurons. a/b and c/d subtypes arise from bi- and bi+ progenitor populations. Neuroblasts from each population undergo two terminal Notch-dependent asymmetric divisions to give rise to the eight subtypes (Pinto-Teixeira et al., 2018). These divisions correspond to binary expression patterns of grn and TfAP-2, respectively, which act with Notch signaling to regulate wiring. Remarkably, despite divergent developmental trajectories separated by multiple divisions and distinct progenitor pools, all T4 and all T5 subtypes converge onto the same transcriptional programs associated with two types of dendritic inputs. Three regulatory dichotomies could also reflect the evolutionary origin of T4/T5 subtypes and correspond to consecutive duplications of ancestral cell types and circuits (Arendt et al., 2016; Shinomiya et al., 2015).

Each axonal and dendritic transcriptional program is characterized by a specific pattern of TFs, as well as a set of CSPs, many of which are implicated in regulating wiring in other developmental contexts. These include Ig superfamily proteins in which different paralogs exhibit discrete heterophilic binding specificities, including the beat/side and the dpr/DIP interacting protein families (Zinn & Özkan, 2017). Interestingly, dynamic expression of these proteins in neurons with shared wiring features was developmentally coordinated. We envision that the synaptic specificity of T4/T5 dendrites and axons are determined by the combined activity of these recognition molecules through interactions with synaptic partners. Future experiments utilizing gain and loss of function analysis, either alone or different combinations, will provide insights into the cellular recognition mechanisms by which synaptic specificity is established.

The composite morphological properties of T4/T5 subtypes allowed us to identify, and thus decouple transcriptional programs for dendrite and axon wiring. Combining separate dendritic and axonal programs, and variations on them, may contribute to the diversification of synaptic specificity in different neuronal subtypes across complex nervous systems.

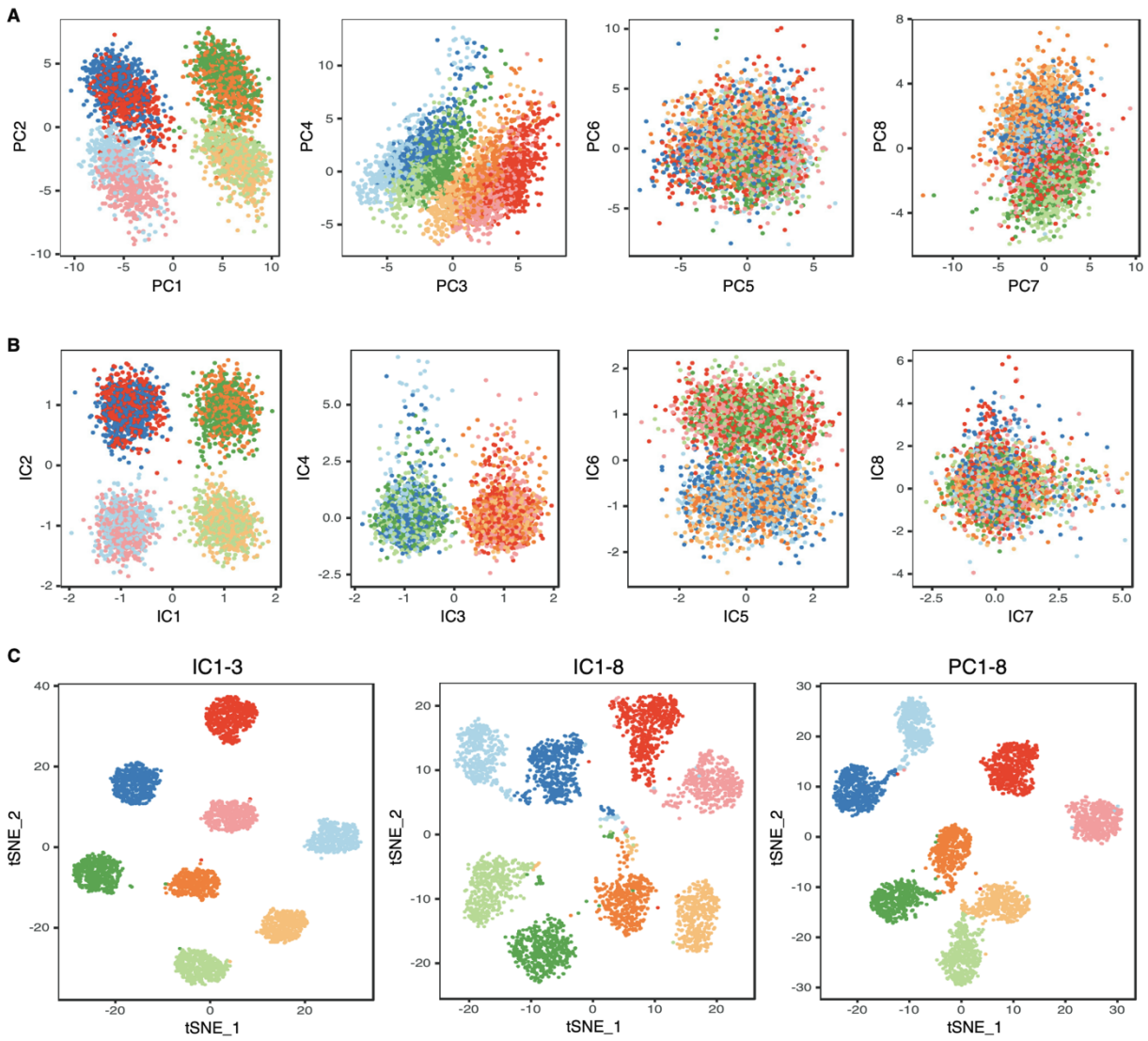


Figure 2.S1. T4/T5 neurons robustly cluster into eight transcriptionally distinct populations (48 hr APF).

(A) Principal component analysis (PCA). (B) Independent component analysis (ICA). Distributions of cells along eight principal components (PCs) and eight independent components (ICs). (C) tSNE plots based on IC 1–3 (left), IC 1–8 (middle), PC 1–8 (right). Cells are color coded according to the final clustering results based on IC 1–3, as in Figure 2.1

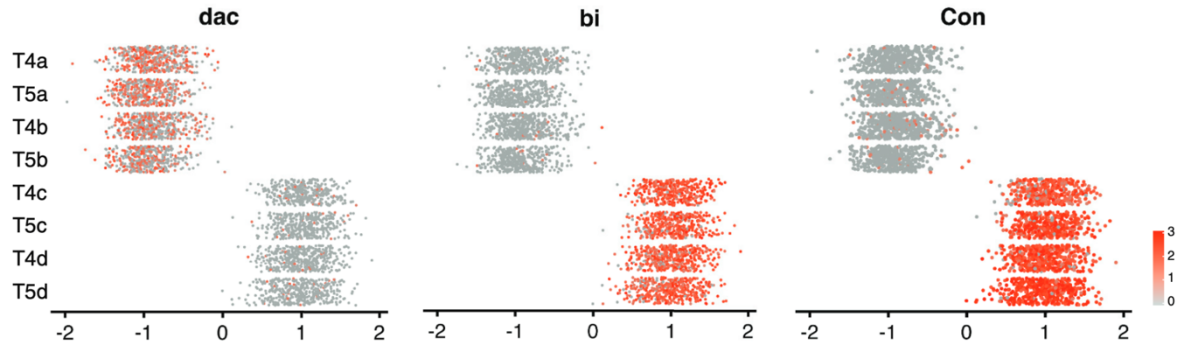


Figure 2.S2. Expression patterns of known marker genes for a/b and c/d subtypes along Axis 1 at 48 hr APF. See legend of Figure 2.3 for details.

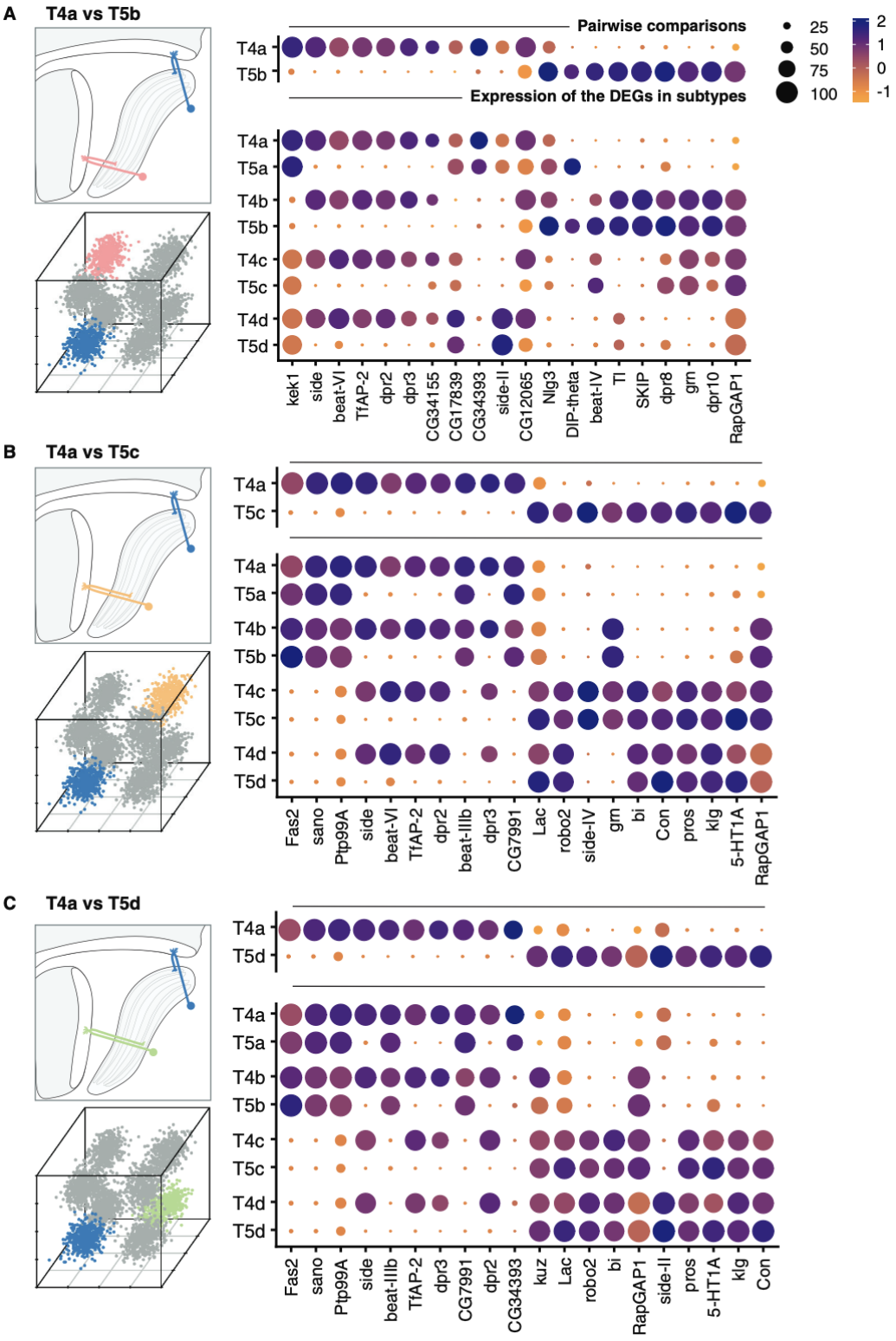


Figure 2.S3. Pairwise comparisons between T4a and subtypes that differ by both axonal outputs and dendritic inputs. Top 20 DEGs are shown for each comparison. See legend of Figure 2.4 for details.

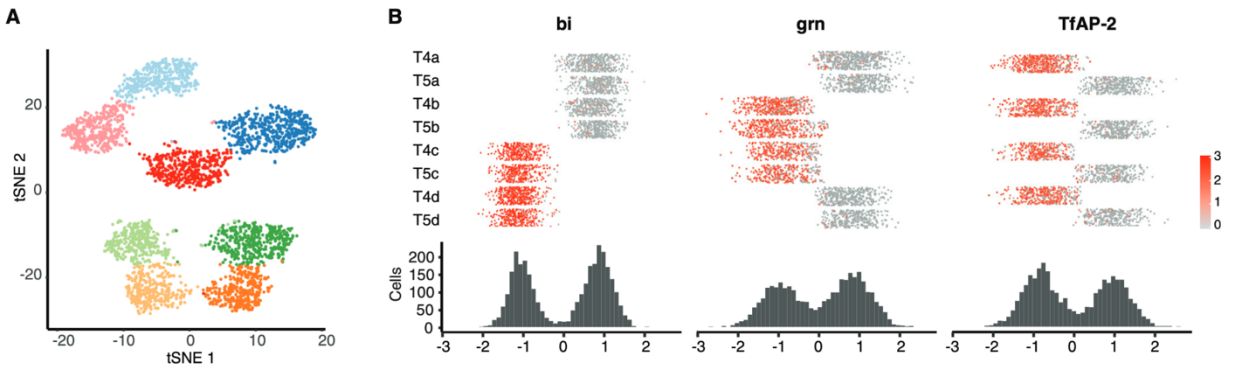


Figure 2.S4. Single-cell profiling of T4/T5 neurons at 24 hr APF. Unsupervised analysis revealed eight transcriptionally distinct populations. (A) tSNE plot of 3833 single-cell transcriptomes. (B). Distribution of cells along three primary axes of transcriptional diversity, and expression patterns of TF with highest contribution to each axis. See legends of Figures 2.1–2.3 for details.

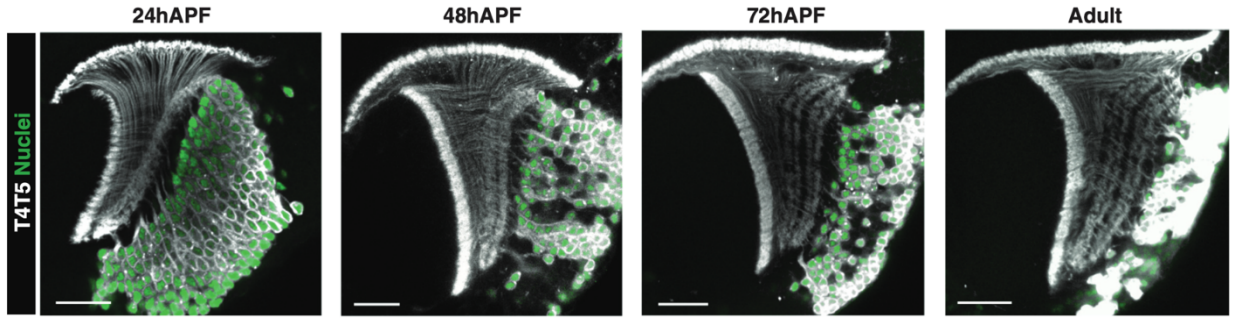


Figure 2.S5. Sequential lamination of T4/T5 axons and four LoP layers. 23G12-Gal4 drives membrane localized RFP (grey) and nuclear localized GFP (green) in all T4/T5 neurons throughout pupal development.

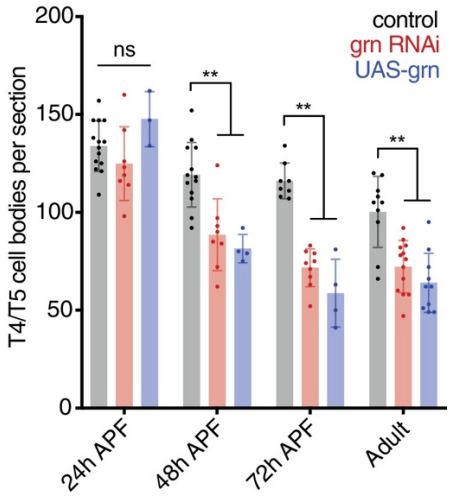


Figure 2.S6. grn RNAi and grn overexpression cause significant loss of T4/T5 neurons between 24 and 48 hr APF.

Statistical significance assessed by one-way ANOVA with Dunnett's multiple comparison test (**p<0.01).

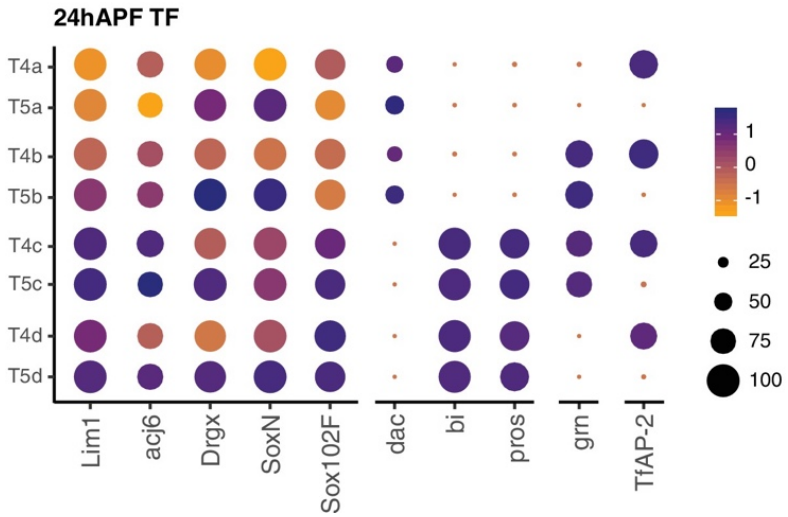


Figure 2.S7. Expression patterns of TFs at 24hr APF.

See Figure 2.7 legend for details.

Figure 2.S8. Expression patterns of subtype-enriched CSPs with cell adhesion domains. Expression patterns of subtype-enriched CSPs with cell adhesion domains (e.g. Ig and LRR). Expression patterns are shown for 24hr and 48hr APF. CSPs are ordered by similarity of their expression patterns. See Figure 2.7 legend for details.

2.6. References

Allan, D. W. & Thor, S. (2015). Transcriptional selectors, masters, and combinatorial codes: regulatory principles of neural subtype specification. Wiley Interdisciplinary Reviews: Developmental Biology, 4(5), 505–528. https://doi.org/10.1002/wdev.191

Apitz, H. & Salecker, I. (2018). Spatio-temporal relays control layer identity of direction-selective neuron subtypes in Drosophila. Nature Communications, 1–16. https://doi.org/10.1038/s41467-018-04592-z

Arendt, D., Musser, J. M., Baker, C. V. H., Bergman, A., Cepko, C., Erwin, D. H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M. D. & Wagner, G. P. (2016). The origin and evolution of cell types. Nature Publishing Group, 17(12), 744–757. https://doi.org/10.1038/nrg.2016.127

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature Biotechnology, 36(5), 411–420. https://doi.org/10.1038/nbt.4096

Contreras, E. G., Palominos, T., Glavic, Á., Brand, A. H., Sierralta, J. & Oliva, C. (2018). The transcription factor SoxD controls neuronal guidance in the Drosophila visual system. Scientific Reports, 8(1), 13332. https://doi.org/10.1038/s41598-018-31654-5

Davie, K., Janssens, J., Koldere, D., Waegeneer, M. D., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., González-Blas, C. B., Poovathingal, S., Hulselmans, G., Spanier, K. I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. Cell, 174(4), 982-998.e20. https://doi.org/10.1016/j.cell.2018.05.057

Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U. & Mann, R. S. (2015). Specification of Individual Adult Motor Neuron Morphologies by Combinatorial Transcription Factor Codes. Neuron, 86(4), 955–970. https://doi.org/10.1016/j.neuron.2015.04.011

Fischbach, K. F. & Dittrich, A. (1989). The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. Cell and Tissue Research. http://link.springer.com/article/10.1007/BF00218858

Hobert, O. (2016). Terminal Selectors of Neuronal Identity (1st ed., Vol. 116). https://doi.org/10.1016/bs.ctdb.2015.12.007

Klingler, E., Prados, J., Kebschull, J. M., Dayer, A., Zador, A. M. & Jabaudon, D. (2018). Single-cell molecular connectomics of intracortically-projecting neurons. BioRxiv, 378760. https://doi.org/10.1101/378760

Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R. & Desplan, C. (2018). Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. 174(3), 622-635.e13. https://doi.org/10.1016/j.cell.2018.05.021

Li, Hanqing, Watson, A., Olechwier, A., Anaya, M., Sorooshyari, S. K., Harnett, D. P., Lee, H.-K. P., Vielmetter, J., Fares, M. A., Garcia, K. C., Özkan, E., Labrador, J.-P. & Zinn, K. (2017).

Deconstruction of the beaten Path-Sidestep interaction network provides insights into neuromuscular system development. ELife, 6, 270. https://doi.org/10.7554/elife.28111

Li, Hongjie, Horns, F., Wu, B., Xie, Q., Li, J., Li, T., Luginbuhl, D. J., Quake, S. R. & Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. Cell, 171(5), 1206-1207.e22. https://doi.org/10.1016/j.cell.2017.10.019

Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., Schilling, T., Bahl, A., Rubin, G. M., Nern, A., Dickson, B. J., Reiff, D. F., Hopp, E. & Borst, A. (2013). A directional tuning map of Drosophila elementary motion detectors. Nature, 500(7461), 212–216. https://doi.org/10.1038/nature12320

Millard, S. S. & Pecot, M. Y. (2018). Strategies for assembling columns and layers in the Drosophila visual system. 1–17. https://doi.org/10.1186/s13064-018-0106-9

Pinto-Teixeira, F., Koo, C., Rossi, A. M., Neriec, N., Bertet, C., Li, X., Rodríguez, A. del V. & Desplan, C. (2018). Development of Concurrent Retinotopic Maps in the Fly Motion Detection Circuit. Cell, 1–26. https://doi.org/10.1016/j.cell.2018.02.053

Sanes, J. R. & Zipursky, S. L. (2010). Design principles of insect and vertebrate visual systems. Neuron, 66(1), 15–36. https://doi.org/10.1016/j.neuron.2010.01.018

Sanes, J. R. & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell, 181(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Sarin, S., Zuniga-Sanchez, E., Kurmangaliyev, Y. Z., Cousins, H., Patel, M., Hernandez, J., Zhang, K. X., Samuel, M. A., Morey, M., Sanes, J. R. & Zipursky, S. L. (2018). Role for Wnt Signaling in Retinal Neuropil Development: Analysis via RNA-Seq and In Vivo Somatic CRISPR Mutagenesis. Neuron, 98(1), 109-126.e8. https://doi.org/10.1016/j.neuron.2018.03.004

Saunders, A., Macosko, E. Z., Wysoker, A., Goldman, M., Krienen, F. M., Rivera, H. de, Bien, E., Baum, M., Bortolin, L., Wang, S., Goeva, A., Nemesh, J., Kamitaki, N., Brumbaugh, S., Kulp, D. & McCarroll, S. A. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. Cell, 174(4), 1015-1030.e16. https://doi.org/10.1016/j.cell.2018.07.028

Schilling, T., Ali, A. H., Leonhardt, A., Borst, A. & Pujol-Marti, J. (2019). Transcriptional control of morphological properties of direction-selective T4/T5 neurons in Drosophila. Development (Cambridge, England), 146(2), dev169763-15. https://doi.org/10.1242/dev.169763

Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C. S., Aniceto, R., Ansari, N., Cheatham, N., Lauchie, S., Neace, E., Ogundeyi, O., Ordish, C., Peel, D., Shinomiya, A., Smith, C., Takemura, S., Talebi, I., Rivlin, P. K., Nern, A., … Meinertzhagen, I. A. (2019). Comparisons between the ON- and OFF-edge motion pathways in the Drosophila brain. ELife, 8, e40025. https://doi.org/10.7554/elife.40025

Shinomiya, K., Takemura, S.-Y., Rivlin, P. K., Plaza, S. M., Scheffer, L. K. & Meinertzhagen, I. A. (2015). A common evolutionary origin for the ON- and OFF-edge motion detection pathways of the Drosophila visual system. Frontiers in Neural Circuits, 9, 33. https://doi.org/10.3389/fncir.2015.00033

Tan, L., Zhang, K. X., Pecot, M. Y., Nagarkar-Jaiswal, S., Lee, P.-T., Takemura, S.-Y., McEwen, J. M., Nern, A., Xu, S., Tadros, W., Chen, Z., Zinn, K., Bellen, H. J., Morey, M. & Zipursky, S. L. (2015). Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. Cell, 163(7), 1756–1769. https://doi.org/10.1016/j.cell.2015.11.021

Tasic, B., Yao, Z., Graybuck, L. T., Smith, K. A., Nguyen, T. N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M. N., Viswanathan, S., Penn, O., Bakken, T., Menon, V., Miller, J., Fong, O., Hirokawa, K. E., Lathia, K., Rimorin, C., Tieu, M., … Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature, 1–41. https://doi.org/10.1038/s41586-018-0654-5

Venken, K. J. T., Schulze, K. L., Haelterman, N. A., Pan, H., He, Y., Evans-Holm, M., Carlson, J. W., Levis, R. W., Spradling, A. C., Hoskins, R. A. & Bellen, H. J. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nature Methods, 8(9), 737–743. https://doi.org/10.1038/nmeth.1662

Zarin, A. A., Asadzadeh, J., Hokamp, K., McCartney, D., Yang, L., Bashaw, G. J. & Labrador, J.-P. (2014). A transcription factor network coordinates attraction, repulsion, and adhesion combinatorially to control motor axon pathway selection. Neuron, 81(6), 1297–1311. https://doi.org/10.1016/j.neuron.2014.01.038

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. Nature Communications, 8(1), 14049. https://doi.org/10.1038/ncomms14049

Zinn, K. & Özkan, E. (2017). Neural immunoglobulin superfamily interaction networks. Current Opinion in Neurobiology, 45, 99–105. https://doi.org/10.1016/j.conb.2017.05.010

CHAPTER THREE

Transcriptional Programs of Circuit Assembly in the *Drosophila* Visual System

3.1 Abstract

Precise patterns of synaptic connections between neurons are encoded in their genetic programs. Here, we use single-cell RNA sequencing to profile neuronal transcriptomes at multiple stages in the developing Drosophila visual system. We devise an efficient strategy for profiling neurons at multiple time points in a single pool, thereby minimizing batch effects and maximizing the reliability of time-course data. A transcriptional atlas spanning multiple stages is generated, including more than 150 distinct neuronal populations; of these, 88 are followed through synaptogenesis. This analysis reveals a common (pan-neuronal) program unfolding in highly coordinated fashion in all neurons, including genes encoding proteins comprising the core synaptic machinery and membrane excitability. This program is overlaid by cell-type-specific programs with diverse cell recognition molecules expressed in different combinations and at different times. We propose that a pan-neuronal program endows neurons with the competence to form synapses and that cell-type-specific programs control synaptic specificity.

3.2. Introduction

Neural circuits in brains from worms to mammals are largely hard wired. Genetic and molecular studies have provided mechanistic insights into how circuits develop (Sanes & Zipursky, 2020). A small, largely evolutionarily conserved set of intercellular signaling molecules and a more diverse set of transcription factors (TFs) act early in development to pattern neural tissue and regulate neuronal cell fate (Holguera & Desplan, 2018). TFs called terminal selectors have been identified that regulate selective features of neurons, such as the specific neurotransmitter a neuron uses (Hobert & Kratsios, 2019). Combinatorial transcriptional programs have been identified in Drosophila that specify unique dendritic and axonal morphologies (Enriquez et al., 2015). At later stages of development, transcriptional programs also specify patterns of synaptic

connectivity, with examples of specific TFs controlling the expression of genes encoding proteins that regulate wiring (Liu et al., 2018; Morey et al., 2008). Recent single-cell sequencing studies in the Drosophila olfactory and visual systems suggest that a relatively small number of TFs drive neuron diversity and that a much larger set of cell surface proteins regulated by them specify patterns of synaptic connectivity (Kurmangaliyev et al., 2019; Li et al., 2017).

The formation of connections between neurons unfolds during development in a stepwise fashion. During axon guidance, growth cones at the leading edge of axons extend along stereotyped paths to their target regions. Here, other neurons elaborate dendrites. Within the target region, axons intermingle with dendritic processes of tens to hundreds of different neuron types. It is during this period in development that neurites form synaptic connections and acquire their mature signaling properties. Whereas genetic and biochemical studies have led to a detailed molecular description of axon guidance (Dickson, 2002; Tessier-Lavigne & Goodman, 1996), our understanding of how neurites discriminate between one another to select their appropriate synaptic partners remains fragmentary (Sanes & Zipursky, 2020). Several recent advances have brightened the prospects of uncovering molecules, mechanisms, and principles underlying synaptic specificity. Among these, single-cell RNA sequencing (scRNA-seq) provides a unique opportunity to characterize expression of cell recognition molecules expressed in synaptic partners during synapse formation.

The Drosophila visual system is well suited to studying the genetic programs of synaptic specificity due to the availability of synapse-level connectivity maps of more than a hundred well-defined cell types (Fischbach & Dittrich, 1989; S. Takemura et al., 2013; S.-Y. Takemura et al., 2015) and the availability of genetic and molecular tools to manipulate discrete classes of neurons. A number of recent studies used single-cell sequencing to characterize neurons in the adult fly brain, including the visual system (Allen et al., 2020; Croset et al., 2018; Davie et al., 2018; Konstantinides et al., 2018). However, these datasets are of limited use for studying

developmental processes, as transcriptomes during development are dynamic and quite different from their adult counterparts (Kurmangaliyev et al., 2019; Li et al., 2017; Sarin et al., 2018).

Here, we set out to generate a transcriptional atlas of the developing Drosophila visual system during circuit assembly. We devised a strategy for single-cell sequencing of multiple time points in a pooled fashion to minimize the confound of batch effects. We identified 162 distinct neuronal populations and followed 88 of them at seven time points just prior to, during, and following synapse formation. This enabled us to uncover a common (pan-neuronal) transcriptional program in all neuron types that proceeds in a synchronous fashion. This program includes genes encoding core protein components of synapses and membrane excitability. The diversity of cell-type-specific programs was driven by dynamically expressed sets of cell recognition molecules and other genes involved in intercellular interactions. We propose that the pan-neuronal program endows all neurons with a competence to form synapses, whereas synaptic specificity is determined by cell-type-specific repertoires of cell surface proteins with distinct temporal expression patterns and recognition specificities.

3.3. Materials and Methods

Drosophila DGRP heterozygote animals

Flies were reared at 25°C on standard medium. For experiments with conventional design (W1118), w[1118] (BDSC #5905) female white pre-pupae (0h APF) were collected and reared until dissection at target time points (5-6 animals per sample). For experiments with pooled design(DGRP), virgin females w[1118] (BDSC #5905) were crossed to males of 35 isogenic wild-type strains from the Drosophila Reference Genetic Panel (Mackay et al. 2012; Huang et al. 2014). F1-generation female white pre-pupae were collected and reared until dissection at target time points. Animals for pooled samples with mixtures of target time points were staged with 12h intervals and were dissected and processed simultaneously. For every target time point, we used three individual animals with unique genotypes (total 27 animals per sample). For two replicates of DGRP experiments, genotypes used for each time point were reshuffled. All genotypes used in the study are shown in Supplementary Figure1A-B.

Drosophila transgenic lines

The following transgenic lines were used for validation of regional subtypes of Tm9 and T4/T5 neurons: 24C08-LexA (Tm9, BDSC #62012), 23G12-LexA (T4/T5, BDSC #65044), Wnt10-TG4 (gift from Hugo Bellen, CR01661), CG15537-TG4 (gift from Hugo Bellen, CR02107), LexAop-myr::tdTomato (Zipursky laboratory), 10XUAS-myr::GFP (Zipursky laboratory).

Tissue dissociation and single-cell suspensions

Optic lobes were collected in a single Eppendorf tube per sample. Brain tissue was incubated in papain (Worthington #LK003178) and Liberase protease (Sigma-Aldrich #5401119001) cocktail at 25°C for 15 min. Tissue was gently washed twice with PBS, then washed with 0.04% BSA in PBS and dissociated mechanically by pipetting. Cell suspension was filtered through a 20 µm

cell-strainer (Corning #352235), stained with DRAQ5 (abcam #ab108410), and sorted by FACS (BD FACSAria II) to isolate single-cells and measure cell concentrations.

Immunohistochemistry / Immunofluorescence

Brain dissections and immunostaining were performed as described in (Kurmangaliyev et al. 2019). Brains were dissected in ice-cold Schneider's Drosophila Medium (Gibco #21720–024), and fixed in PBS (Bioland Scientific LLC #PBS01-03) containing 4% paraformaldehyde (Electron Microscopy Sciences, Cat#15710) for 25 min at room temperature (RT). Brains were rinsed repeatedly with PBST (PBS containing 0.5% Triton-X100 (Sigma #T9284)), and incubated in blocking solution (PBST containing 10% Normal Goat Serum (Sigma #G6767)) for at least 1 hr at RT prior to incubation with anti- body. Brains were incubated sequentially with primary and secondary antibodies diluted in blocking solution overnight at 4C, with at least 2 PBST rinses followed by 2 hr incubations at RT in between and afterwards. Brains were transferred to 50% (for 30 min), then 100% EverBrite mounting medium (Biotium #23001) and mounted on slides for confocal microscopy.

Primary antibodies and dilutions used in this study: chicken anti-GFP (Abcam #13970, 1:1000), rabbit anti-dsRed (Clontech #632496, 1:200), rat anti-NCad (Developmental Studies Hybridoma Bank (DSHB) DN-Ex#8, 1:40). Secondary antibodies and dilutions used in this study were as follows: goat anti-chicken Alexa Fluor 488 (AF488) (Invitrogen #A11039, 1:200), goat anti-rabbit AF568 (Invitrogen #A11011, 1:200), goat anti-rat AF647 (Invitrogen #A21247, 1:500).
Immunofluorescence images were acquired using a Zeiss LSM 880 confocal microscope with Zen digital imaging software. Optical sections or maximum intensity projections were level-adjusted, cropped and exported for presentation using Image J software (Fiji). Reported expression patterns were reproducible across three or more biological samples.

Confocal microscopy and image analysis

Immunofluorescence images were acquired using a Zeiss LSM 880 confocal microscope with Zen digital imaging software. Optical sections or maximum intensity projections were level-adjusted, cropped and exported for presentation using Image J software (Fiji). Reported expression patterns were reproducible across three or more biological samples.

Droplet-based single-cell RNA-Seq

Single-cell suspensions were processed using the 10XGenomics Chromium 3' v3 platform. For W1118 experiments, 4 lanes of a Chromium Chip were loaded per each sample; for DGRP experiments, 8 lanes of a Chromium Chip were loaded per each sample (i.e. replicates A and B). Loading volumes were estimated based on cell concentrations to capture around 8,000 single cells per one lane. Single-cell RNA-Seq libraries were generated using the manufacturer's protocol, with 12 cycles of PCR for cDNA amplification. All RNA-Seq libraries were sequenced using 4 lanes of Illumina NovaSeq 6000 S4 platform (28bp + 150bp).

QUANTIFICATION AND STATISTICAL ANALYSIS

Single-cell RNA-seq data analysis

Raw data processing

Fastq files with raw reads were processed using Cell Ranger (3.1.0) with default parameters. Reference genome and transcriptome were based on FlyBase (release 6.29).

Demultiplexing of DGRP samples

Single-cell transcriptomes from DGRP samples were demultiplexed based on parental genotypes using demuxlet with default parameters (version 2, https://github.com/statgen/popscle, Kang et al. 2018). The genotypes of wild-type strains were

downloaded from the DGRP web-site (http://dgrp2.gnets.ncsu.edu/, Mackay et al. 2012; Huang et al. 2014). Demultiplexing was based on genotypes of 35 DGRP strains that were used in experiments (each sample included 27 of them and the remaining genotypes were used as negative controls). The coordinates of genomic variants for DGRP strains were updated from dm3 to dm6 version of the Drosophila melanogaster reference genome using CrossMap (Zhao et al. 2014). The exonic variants for w[1118] (BDSC #5905) were called based on BAM files produced by Cell Ranger for W1118 samples. Variants were called using bcftools pipeline with default parameters (version 1.8, Heng, 2011). The genotypes of F1 heterozygotes were assembled based on parental genotypes. The genomic variants used for demultiplexing were filtered using following criteria: (1) only biallelic single-nucleotide polymorphisms (SNPs) with maximum minor allele count of 8 among analyzed DGRP strains; (2) SNPs had to be represented by a reference genome allele in w[1118] strain (min.depth>10). In total, 176,636

SNPs were used for demultiplexing. Only 37 of 246,308 cell barcodes (before filtering) were erroneously assigned to the genotypes that were not used in the given samples (negative controls), indicating high accuracy of recovered sample identities of single cells.

Quality control and filtering of single cells

The initial set of cell barcodes called by Cell Ranger were filtered based on following criteria: (1) number of transcripts from 2,000 to 20,000; (2) maximum 10% of mitochondrial transcripts; (3) for DGRP samples, we removed predicted doublets and cells assigned to wrong genotypes; (4) we also removed cells with more than 3 transcripts aligned to male-specific lncRNA:roX1 and lncRNA:roX2 genes. Only one of the samples had a considerable number of potential male cells suggesting contamination by a male pupa (W1118, 48h APF, replicate A, 15% of cells). In total, 208,976 cells passed all criteria and were used for further analysis (Figure S1E).

Integrative analysis of the main dataset (24-96h APF)

All steps of single-cell data analysis were performed using methods implemented in Seurat V3 (version 3.1.2, Butler et al., 2018, PMID:29608179; Stuart et al. 2019). Samples from 24h to 96h APF were analyzed together. Raw gene counts were normalized by total number of transcripts per cell and log-transformed (function NormalizeData). Integration was performed on 3000 highly variable genes selected across all samples (function SelectIntegrationFeatures). We excluded from this set mitochondrial genes, ribosomal protein genes, genes encoding heat-shock proteins and genes from oxidative phosphorylation complexes (FlyBase gene groups). We performed integration of samples using "reference-based" implementation of CCA (canonical correlation analysis) and MNN (mutual nearest neighbors) based workflow. W1118 samples from 48h (replicate A) and 72h APF were used as reference datasets, and dimensionality of the dataset was set to 200 (functions FindIntegrationAnchors and IntegrateData). Integrated dataset was scaled and used for principal component analysis (PCA). The first 200 PCs were used for graph-based clustering (functions FindNeighbors and FindClusters, k.param = 50, resolution = 10). The same set of PCs was used to generate t-distributed stochastic neighbor embedding (tSNE) plots.

Integrative analysis of the early dataset (0h-24h APF)

DGRP samples from early time points were integrated separately. The analysis was performed similar to the main dataset. We used DGRP 24h samples as reference datasets. All parameters of the analysis were the same as those used for the main dataset. Matching transcriptional clusters to major classes of cell types (main dataset): The first three PCs of integrated dataset separated the major classes of cell types: PC1 separated neurons (elav+) from other populations; PC2 separated two groups of non-neuronal clusters from each other (repo+ glia from repo- cells); and PC3 separated photoreceptor cells (chp+) from the remaining cells. We computed centroids of each cluster in the space of these 3 PCs and applied K-means clustering. This allowed us to group transcriptional clusters into four major classes of cell types (Figure S3).

Matching transcriptional clusters to cell types (main dataset)

Neuronal clusters were matched to known morphological cell types using two bulk RNA-Seq datasets from adult animals (Davis et al. 2020; Konstantinides et al. 2018). For Davis et al. (2020) normalized expression profiles per cell type were downloaded from NCBI GEO (GSE116969, dataTable4). For Konstantinides et al. (2018), expression matrix with raw gene counts was downloaded from NCBI GEO (GSE103772). Flybase gene IDs were matched to gene symbols, and gene counts were normalized using edgeR (Robinson et al. 2010). For both datasets we focused on reference profiles for individual optic lobe neurons. Figure 3.S4A-B).

For single-cell transcriptomes, we used cells from the W1118/96h sample. The analysis performed on original expression values (i.e. uncorrected expression matrix). We computed the average expression profile for each cell type. Next, we identified genes enriched in each neuronal cluster in the same sample. Cluster-enriched genes were identified using Wilcoxon rank-sum test (function FindMarkers, fold-change > 2, adjusted $p < 0.01$). Each cluster was compared to the background set of neurons with equal proportions of cells from each neuronal cluster (max. 30 cells from each cluster). The union of top 20 enriched genes for each cluster was used for correlation analysis (top marker genes, total 826 genes).

Pearson's correlation coefficients were computed between log-transformed reference expression profiles and the average expression profile of each transcriptional cluster for top marker genes (Figure 3.S4A-B). Almost all reference datasets had best mutual match with a single transcriptional cluster (only exclusion was LC16 cells from Davis et al. 2020). Matching of cell types that were present in both reference datasets were also concordant. In a few cases multiple clusters matched with a single reference profile with similar correlation values. In these cases, reference cell types included multiple distinct transcriptional subtypes (both known and newly identified, see main text for details). The final set of matched cell types was manually curated based on expression of known marker genes (Figure 3.S4C). In particular, we curated known

subtypes of T4/T5 neurons, newly identified subtypes for Dm3 and Tm9 neurons, and we also inferred identity of T2a cluster based on similarity to the related and matched T2 and T3 clusters. We used two levels of cluster identities for few cell types: cell types and cell subtypes. Subtypes were defined for T4/T5 neurons, Tm9 and Dm3 neurons. All further analyses were performed on the level of cell types (e.g. all T4/T5 neurons are considered as a single cell type). We also annotated two large clusters of photoreceptors that corresponded to R1-6 and R7/R8 cell types. The remaining cluster of R cells represented only 0.3% of photoreceptors and likely represents cells undergoing apoptosis (defined by expression of grim and rpr). In total, 58 neuronal and 2 photoreceptor clusters were matched to known cell types and subtypes in the main dataset.

Matching transcriptional clusters to cell types (early dataset)

The early dataset included cells from DGRP/24h samples. These cells were also analyzed in the main dataset and have assigned cluster and cell type identities. We used these shared cells to directly transfer cluster identities between two datasets (on the level of cell types, see above). Clusters from the early and the main datasets were considered matching if the majority of 24h cells in both clusters were the same (minimum 90% cells in each direction). In this way, we were able to match and transfer identities for 58 of 196 transcriptional clusters from the main to the early dataset (Figure 3.S5). In few cases, matching of closely related clusters were curated manually (i.e. T4/T5 neurons and photoreceptors). Also note that some of the matched clusters in the early datasets largely consisted of cells from 24h APF (i.e. was not detected at 0h and 12h APF). 22 neuronal clusters that were matched to the main dataset were represented by at least 10 cells at both 0h and 12h APF. Some other clusters included cells only from 0h and 12h APF. These clusters may represent cells from other parts of the brain captured due to differences in tissue dissections at different pupal stages, or represent transient progenitor populations.

Cell type-specific transcriptional profiles

The average transcriptional profiles were computed for each cluster at given time point in given sample/replicate (e.g. Mi1 in DGRP, 24h APF, replicate A, see Figure 2D). Averaging was performed in non-log space for the original normalized expression values (i.e. uncorrected expression matrix). The downstream analyses of cell type-specific profiles were focused on DGRP samples (i.e. analyses of pan-neuronally coordinated and highly variable genes). In particular, we focused on 88 neuronal and 6 glial clusters that could be followed at every time point from 24h to 96h APF (minimum 10 cells in one of the replicates). Expression values for two DGRP replicates were further averaged to obtain summarized cell type-specific transcriptional profiles. For visualizations of cell type-specific expression patterns we used log1p-transformed expression values. In heatmaps, we capped the maximum expression values to 20.

Pan-neuronally coordinated genes

The analysis was based on 88 neuronal and 6 glial cell types in DGRP samples from 24h to 96h APF (see above). First, we defined neuron-specific genes by comparing average expression in neurons and glia (fold-change > 8). For each gene, enrichment was estimated at a time point in which it was expressed in the largest number of neurons, or (in case of ties) with highest average expression. Next, we computed the average correlation between expression patterns of neuron-specific genes across different neurons. Pearson's correlation coefficients were averaged using Fisher's Z transformation (Corey, Dunlap and Burke, 2010). Genes expressed in at least half of the neurons (n > 44) and with average Pearson's correlation coefficient higher than 0.75 were defined as pan-neuronally coordinated genes. The expression levels at earlier time points (for cell types that could be followed in the early dataset) were used only for visualizations.

Highly variable genes

The analysis was focused on 8 neuronal cell types from 24h to 96h APF (Figure 3.5). We determined genes with highly variable expression patterns across these cell types. Coefficients of variation (CV) were calculated for all genes that were expressed in any of these cell types (min.expression 1). For each gene, CV was calculated at a time point in which it was expressed in the largest number of neurons, or (in case of ties) with highest average expression. To take into account the relationship between variability and average expression, CV values were grouped into 10 bins based on similarity of mean expression and transformed into Z-scores within each of these bins (scaled CV). We used a cut-off of 1.5 for scaled CV values to define the set of top highly variable genes.

Functional categories of genes

We defined several functional categories of genes. Some genes could be classified into more than one category. Thus, classification was performed sequentially, and every gene was assigned to only one category (e.g. gene classified as “ion channel” cannot be classified as “synaptic genes”, see below). Gene classifications were based on gene groups (GG) and gene ontology (GO) annotations obtained from FlyBase (release 6.29). Annotations for cell adhesion molecules were obtained from FlyXCDB (http://prodata.swmed.edu/FlyXCDB, Pei et al. 2018). Functional categories were defined in following order: (1) ribosomal proteins (from GG); (2) transcription factors (from GG); (3) RNA binding proteins (from GO, term GO:0003723; excluding ribosomal proteins (GG), translation factors (GG) and tRNA genes (GG)); (4) non-coding RNA genes (based on gene annotations); (5) cell adhesion molecules (from FlyXCDB, protein domains: Ig, EGF, LRR, fn3, Cadherin); (6) receptor and ligands (from GG, groups: “transmembrane receptors”, “receptor ligands”). (7) ion channels (from GG); (8) synaptic genes (from GO, GO:0007268).

3.4 Results

3.4.1 Profiling of Different Stages of Neuronal Development in Parallel in a Single Experiment

In this study, we generated gene expression maps of postmitotic neurons in the fly visual system throughout pupal development (100 h) using high-throughput scRNA-seq (Zheng et al., 2017). During this time, neurons complete axon guidance, elaborate distinct axonal and dendritic morphologies, select appropriate synaptic partners, form synapses, and acquire distinct biophysical properties.

Two different experimental designs were used. First, we profiled optic lobes at 24, 48, 72, and 96 h after puparium formation (APF) (**Figures 3.1A** and **3.S1**). Each time point was profiled as a separate experiment, and for the 48h time point, we profiled two independent samples. We refer to this dataset as W1118 based on the strain used for these experiments. This provided a high-resolution dataset for unsupervised identification of cell types at different stages of development. The largest number of cells (51,000) was profiled at 48 h APF, when the transcriptional differences between neuronal cell types were maximal (Li et al., 2017).

A second set of experiments was undertaken to follow changes in gene expression at 12-h intervals to more precisely define the developmental trajectory of each cell type (**Figures 3.1B** and **3.S1**). We developed a strategy to profile neurons from multiple time points in a single experiment. This strategy exploits natural genetic variation in wild-type strains from the Drosophila Genetic Reference Panel (DGRP) to mark cells isolated at different developmental stages (Huang et al., 2014; Kang et al., 2017; Mackay et al., 2012). Males from individual DGRP strains were crossed to females from a common reference strain(W1118) to reduce differences in genetic background (Fear et al., 2016). Pupae from the resulting progeny were collected every 12 h from 0 to 96 h APF. For every time point, we staged three individual animals, each tagged by a unique wild-type genotype. Optic lobes for all time points were dissected, pooled, and processed as a

single sample from tissue dissociation to sequencing. The resulting single-cell transcriptomes for each time point and individual animals were separated based on SNPs (single-nucleotide polymorphisms) captured in mRNA sequences and matched with the genotypes of parental strains (**Figure 3.1C**). In this way, we profiled cells at nine stages of development with three biological replicates for each time point (total 27 samples) in a single experiment. We performed two rounds of this experiment with reshuffled wild-type parental genotypes (**Figures 3.1B** and **3.S1**).

Profiling pools of cells together from all time points in a single experiment has several advantages. First, it is highly efficient and cost effective. The entire experiment, from genetic crosses to single-cell capture, took 2 weeks. Second, it largely eliminates batch effects. An inherent limitation in the conventional design of scRNA-seq is the technical variation (batch effects) be-tween different samples (**Figure 3.S2**). In droplet-based scRNA-seq. for instance, mRNA released from tissue dissociation prior to cell capture is a source of contaminants confounding comparison between samples. In the conventional design, for instance, a contaminant (e.g., a transcript expressed at high levels at a specific time in a subset of cells) may appear as ubiquitously upregulated at this time in all neurons. In the pooled design, however, background contaminants that change with time are distributed among all cells and times and thus are not erroneously attributed to biological differences. Similarly, in pooled samples, other sources of technical variability of unknown etiology arising in the conventional design are eliminated. Finally, genetic multiplexing enables identification of cell doublets, further improving the quality of datasets and reducing cost through maximizing single-cell capture rates in scRNA-seq platforms (Kang et al., 2018).

In summary, we generated a comprehensive dataset of gene expression in visual system neurons covering every 12h of pupal development. After quality control, the dataset comprises 208,976 single-cell transcriptomes with a median of 1,650 genes and 5,148 transcripts detected per cell (minimum 2,000 transcripts). Each of the W1118 and DGRP approaches represented

approximately half of the dataset. The complete experimental design and distribution of cells across time points are shown in Figure 3.S1.

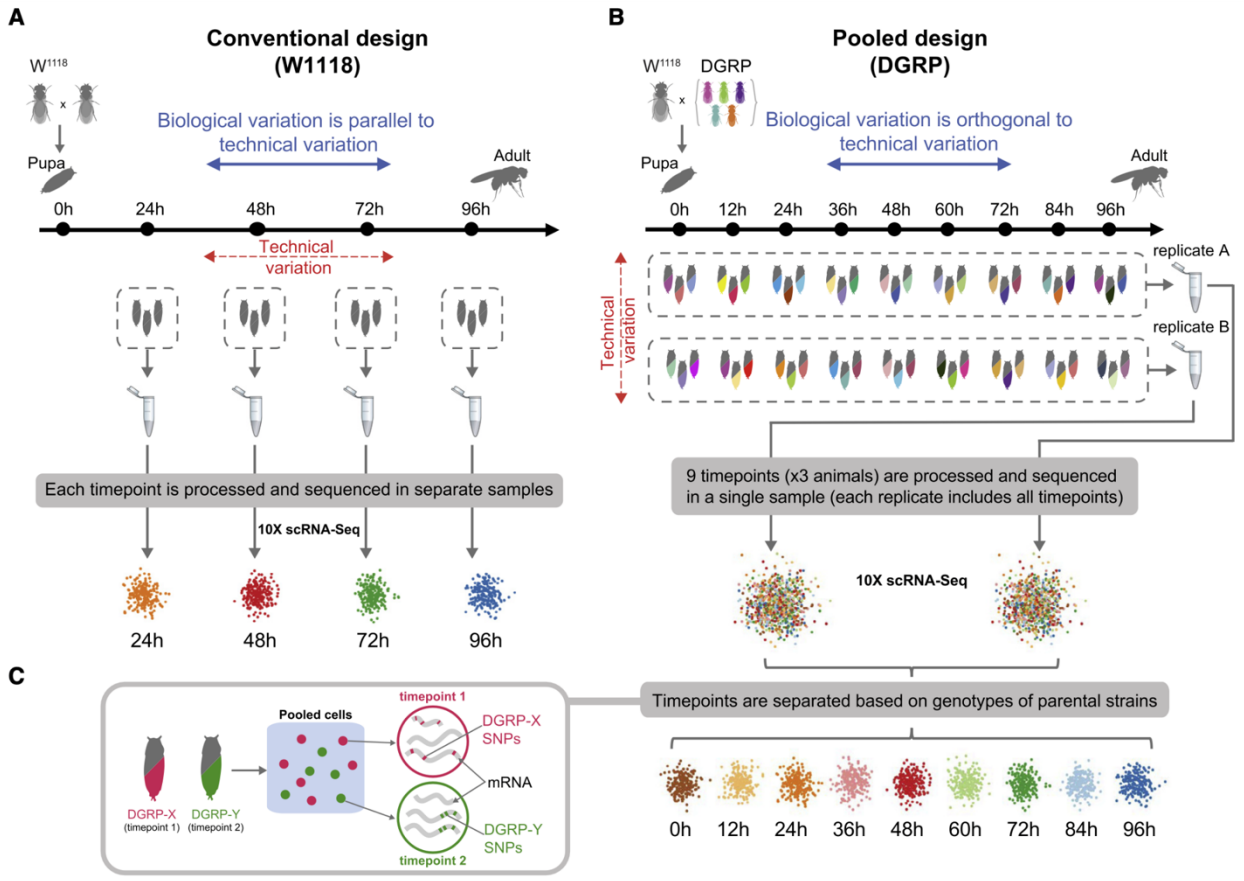


Figure 3.1. Transcriptional Profiling of Multiple Stages of Neuronal Development in a Single Experiment.

Schematics of experimental designs. (A) Conventional design (W1118 dataset). Four stages of pupal development were profiled in separate experiments. (B) Pooled design (Drosophila Genetic Reference Panel [DGRP] dataset). Nine stages of pupal development were pooled and profiled in the same experiment. Cells from each time point were tagged using three unique wild-type strains from the DGRP. Wild-type strains were reshuffled between time points in two independent replicates. (C) Single-cell transcriptomes for each time point were separated based on SNPs (single-nucleotide polymorphisms) captured in mRNA molecules and matching with unique genotypes of pupae. In the conventional design, biological differences between time points cannot be separated from technical variation between different samples. In the pooled design, differences between time points are independent of technical variation between samples. This minimizes the confound of batch effects on constructing an accurate time series of gene expression between time points (see effects on background contributed by ambient RNA in Figure 3.S2). See also Figure 3.S1 for details on experimental design and workflow of the analysis.

3.4.2. A transcriptional atlas of the developing Drosophila visual system

Transcriptional differences between time points for a single neuronal cell type can be comparable to the differences between cell types (Li et al., 2017; Sarin et al., 2018). Therefore, to follow individual cell types over time, we first integrated single-cell datasets across different stages of development. Both W1118 and DGRP samples were analyzed together to achieve maximal resolution in cell type diversity and to generate a unified atlas of transcriptional cell types. Samples from 0 and 12 h APF differed substantially from other time points both in tissue composition and coverage and were analyzed separately (the early dataset). The main analysis was focused on samples from 24 to 96 h APF (the main dataset).

The integrative analysis of the main dataset was performed using CCA (canonical correlation analysis)- and MNN (mutual nearest neighbors)-based workflow implemented in Seurat 3 (Stuart et al., 2019). The integration was performed on the levels of individual time points and replicates. An integrated dataset was used for dimensionality reduction followed by a graph-based clustering method that revealed 196 transcriptionally distinct cell populations (see Methods for details). The results of dataset integration and clustering analysis were visualized using t-distributed stochastic neighbor embedding (tSNE) plots (**Figures 3.2** and **3.S3**). Importantly, the integrative analysis was used strictly for joint clustering of cell types across different samples (i.e., time points and replicates). Once we identified clusters, however, we returned to the original (pre-integration) normalized gene expression values for each time point and sample (**Figure 3.2D**)

Most of the clusters were detectable in all samples and comprise a comparable fraction of cells across different time points, while some were distributed less evenly (**Figure 3.S3**). The variation in cluster proportions could be due to technical reasons (e.g., sampling variation or differences in tissue dissections) or biological differences in cell composition at different stages. Unsupervised analysis grouped clusters into four major classes (**Figure 3.S4**), which were annotated based on expression of known marker genes (see STAR Methods for details): 162

clusters of neurons (elav+), 3 clusters of photoreceptor cells (chp+), 19 clusters of glia (repo+), and 11 clusters of non-neuronal cells of unknown origin (elav-/repo-).

Next, we matched neuronal clusters to previously identified cell types. As a reference, we used two recent datasets from bulk sequencing of several dozen morphologically distinct cell types in the adult (Davie et al., 2018; Davis et al., 2020). We identified a set of cluster-specific marker genes and used them for correlation analysis (**Figures 3.S5A** and 3.**S5B**). In addition, we used expression patterns of TFs to manually curate matched clusters (**Figure 3.S5C**). Most of the reference cell types had a one to-one match with a single cluster. In a few cases, a reference dataset had a high correlation with multiple clusters representing further heterogeneity. For example, T4/T5 neurons matched eight clusters with similar correlation values. Expression of previously identified marker genes allowed us to match these clusters to the eight known morphological subtypes of T4/T5 neurons (Kurmangaliyev et al., 2019). Moreover, we identified new transcriptional subtypes (e.g., two Dm3 and two Tm9 neurons; see below). Overall, cell-type-specific expression profiles of known morphological cell types were in good agreement with previously validated marker genes (e.g., refs.(Cosmanescu et al., 2018; Tan et al., 2015)). For example, the TF bsh has been shown to be expressed in three cell types (Hasegawa et al., 2013), and it was detected only in the corresponding clusters (**Figure 3.S5C**). The transcriptional similarities between clusters were also consistent with known anatomical and functional relationships between cell types (**Figure 3.S6**). In total, we were able to match identities of 58 neuronal clusters to known morphological cell types (**Figure 3.2A**).

The optic lobe neuropils consist of repeated columns with stereotyped cell-type composition and connectivity (S.-Y. Takemura et al., 2015). Many cell types are present in a single copy in each column (unicolumnar neurons). In agreement with this, all matched unicolumnar neurons were among the largest 40–50 transcriptional clusters in our data, and each represented on average 0.7%–1.7% of the cells in each sample. We estimate that these 40–50 largest clusters represent cell types with roughly one copy per column (**Figure 3.2B**).

In parallel to the main atlas, we performed similar integrative analysis for early time points (i.e., 0 and 12 h APF; **Figure 3.S7**). These samples were integrated together with samples from 24 h APF to generate a separate atlas covering the first day of pupal development (from 0 to 24 h APF). Since cells from 24 h APF were present in both parts of the atlas, we were able to match clusters between the two datasets. The differences between cell types were less prominent at earlier stages, and fewer cells were profiled. As a result, many related cell types were not resolved into separate clusters. For example, Tm1, Tm2, and Tm4 neurons and one unidentified cluster (N19) grouped together, suggesting a close developmental relationship between these cell types.

Transcriptional heterogeneity can reflect neuronal diversity beyond morphological cell types. For example, Dm8 neurons comprise two transcriptionally distinct subtypes with the same morphology but different synaptic specificities (Courgeon & Desplan, 2019; Menon et al., 2019). We found new subtypes for Dm3, Tm9, and T4 neurons. Tm9 neurons, for instance, formed two clusters defined by mutually exclusive expression of Wnt4 and Wnt10 genes. Wnt10 expression in vivo was restricted to dorsal Tm9 neurons (**Figure 3.3A**). We conclude that identified clusters represent ventral (Wnt4+) and dorsal (Wnt10+) subtypes of Tm9 neurons. We also noticed a stable subpopulation of T4 neurons expressing an uncharacterized gene (CG15537) throughout development that was enriched in T4 a/b subtypes. This group of cells did not form a separate cluster at the resolution of the analysis. The expression of CG15537, however, was restricted to T4 neurons targeting the dorsal third of the medulla and lobula plate layers a/b (**Figure 3.3B**). The product of CG15537 is predicted to be a cell surface protein with a hormone binding domain. Among the other genes specific to the same subpopulation of T4 neurons are a neuropeptide receptor (TrissinR) and Tbh, a key component of the octopamine biosynthesis pathway. We hypothesize that these neurons are a regional T4 subtype with distinct neuromodulatory characteristics. In this way, we found and validated two cases of transcriptionally defined regional subtypes of neurons with unique spatial distributions. The discovery of new subtypes, even for

such well-studied cell types as T4/T5 neurons, underscores the power of single-cell sequencing to reveal subtle aspects of cellular diversity.

Taken together, we generated a comprehensive transcriptional atlas of the developing Drosophila visual system. This atlas covers more than 150 transcriptionally distinct neuronal populations, including most of the known abundant cell types. The one to-one correspondence between transcriptional clusters and known morphological cell types, the expression of known markers, and the match with the expected proportions of cells indicate that this is a highly reliable resource for following individual cell types at multiple stages of development (**Figure 3.2C**). We tracked 88 neuronal cell types every 12 h of pupal development from 24 to 96 h APF (minimum of 10 cells in either DGRP replicate). This largely encompasses developmental times preceding synapse formation through the formation of the mature connectome. We also tracked 22 of these cell types in the early dataset (both at 0 and 12 h APF) covering all 100 h of pupal development. The cell-type-specific gene expression profiles were highly reproducible across replicates (**Figure 3.2D**). For an additional transcriptional analysis of the developing Drosophila visual system, see the contemporaneous study from the Desplan lab (Özel et al., 2020).

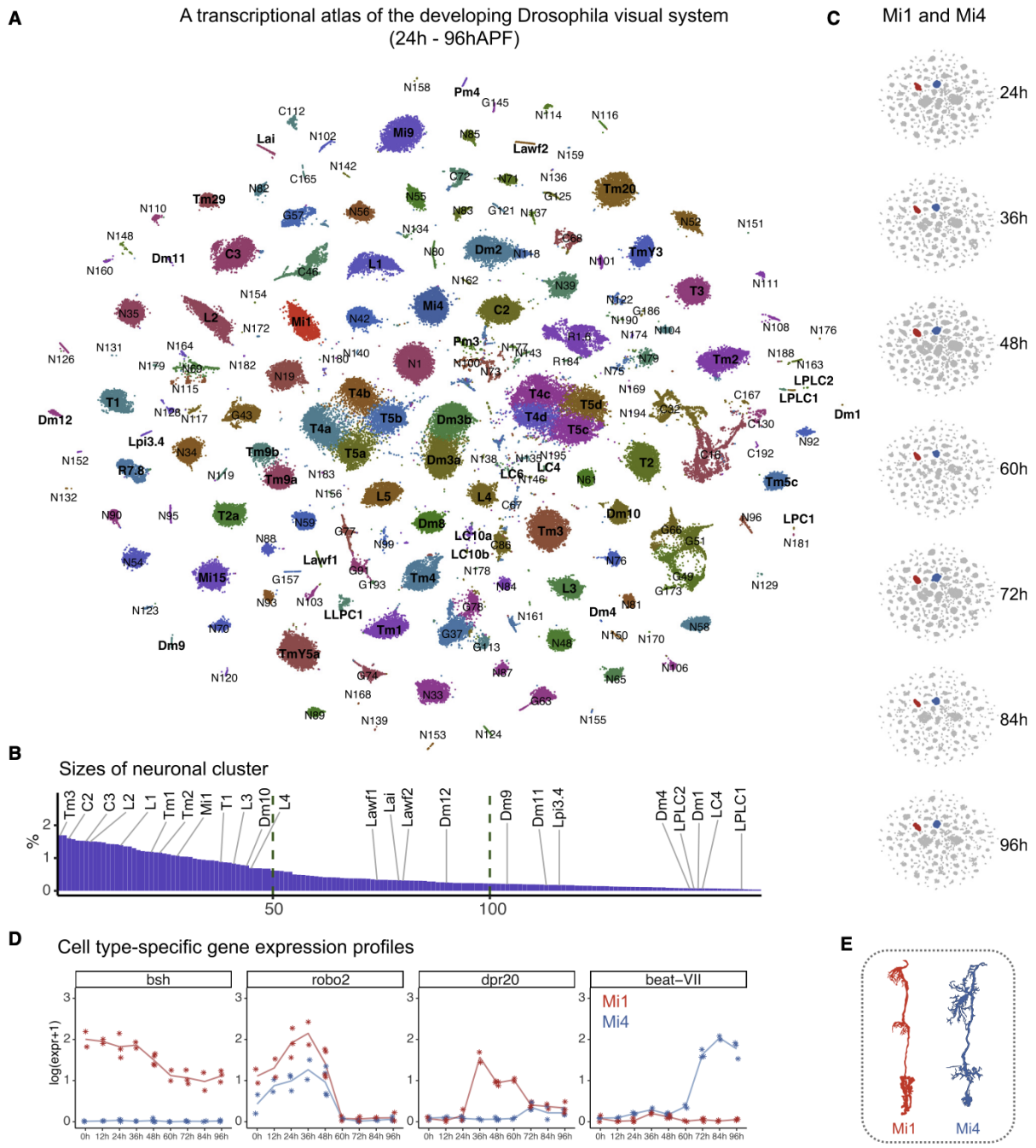


Figure 3.2. A Transcriptional Atlas of the Developing Drosophila Visual System (A) tSNE plot of the main dataset. This includes samples from 24 to 96 h APF (DGRP and W1118). All samples were integrated and clustered together. Distinct transcriptional clusters are color coded and labeled by identities of clusters (i.e., names of neurons for known morphological cell types). (B) The proportion of each neuronal cluster relative to the total number of neurons (median across all samples). Vertical bars correspond to the neuronal clusters ordered by average proportion (selected clusters of known morphological cell types are highlighted). All of the identified unicolumnar neurons were among the largest 50 clusters.

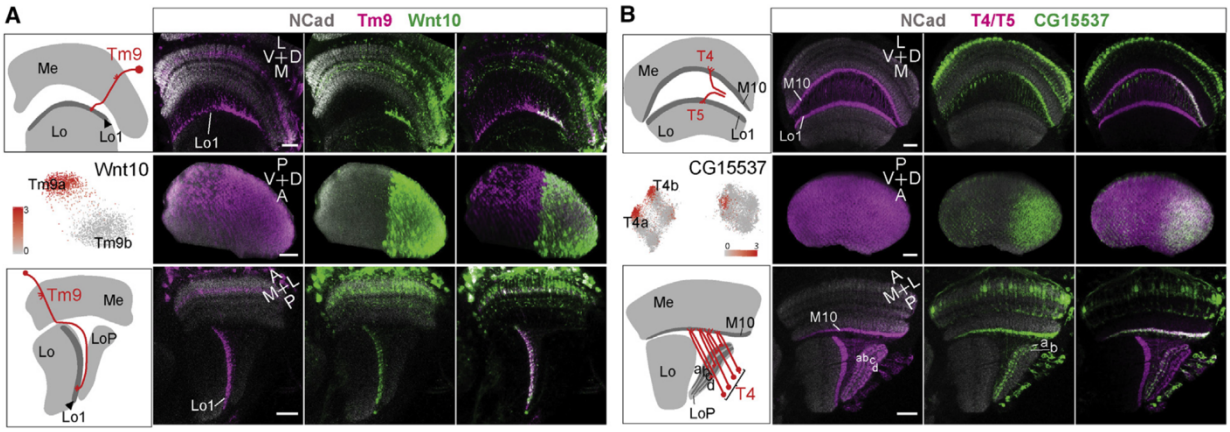


Figure 3.3. Regional Specialization of Neurons (A and B) Transcriptionally distinct variants of Tm9 (A) and T4/T5 (B) neurons. The arrangement of cell types in the optic lobe and tSNE plots with expression patterns of marker genes for newly identified subtypes are shown on the left in each panel. All Tm9 (A) and T4/T5 (B) neurons are labeled in magenta. Expression patterns of marker genes in adult are shown in green. (A) There are two transcriptionally distinct populations of Tm9 neurons. They are defined by mutually exclusive expression of Wnt10 and Wnt4 (not shown) genes. Wnt10 is restricted to Tm9 neurons targeting the dorsal half of the lobula. The middle row shows the cropped lobula 1 layer (Lo1). (B) A subpopulation of T4 neurons is defined by expression of CG15537 restricted to T4 neurons targeting the dorsal third of the medulla and lobula plate layers a/b. The middle row shows the cropped medulla 10 layer (M10). Scale bar, 20 mm

3.4.3. Common and Cell-type-Specific Components of Neuronal Genetic Programs A Common (Pan-neuronal) Program of Neuronal Development

All neurons share similar properties (e.g., synaptic transmission and membrane excitability). We sought to assess whether there was a common neuronal differentiation pathway giving rise to these features or whether they emerge in different ways in different neurons.

To do so, we followed 88 neuronal and 6 glial cell types from 24 to 96 h APF (see STAR Methods for details). First, we defined neuron-specific genes that were highly enriched in neurons compared to glia. Genes known to be involved in neuronal development and function were enriched among neuron-specific genes (**Figure 3.4A**).

Neuron-specific genes showed a wide range of temporal dynamics during development. We calculated Pearson's correlation coefficients between expression patterns of each gene across different cell types and averaged them using Fisher's Z transformation (Corey et al., 1998). Many neuron-specific genes were expressed in a highly coordinated fashion across different cell types. We hypothesize that they are part of a common genetic program unfolding in all neurons in parallel (a ''pan-neuronal program''). Pan-neuronal genes were defined using the following criteria: (1) a gene was detected in at least half the neuronal cell types, and (2) the average Pearson's correlation coefficient was >0.75 (see STAR Methods for details). In total, 200 genes met these criteria (**Figure 3.4A**). Expression patterns of these genes are shown in Figures 4, S8, and S9. The list of reported genes represents a subset of the top pan-neuronally genes with different levels of coordination. For instance, increasing the requirement for number of cell types in which gene was detected to 80% results in a slightly more conservative list of the genes (182 of 200 reported genes still pass this criterion).

The core components of synapses and membrane excitability were among the most tightly coordinated groups of genes (**Figures 3.4, 3.S8**, and 3.**S9**; Harris and Littleton, 2015). These include components of the synaptic release machinery (e.g., Snap25 and nSyb), the presynaptic

active zone (e.g., Brp and Rbp), and many ion channels (e.g., Cac and Para). Some of the neurotransmitter receptor subunits were also expressed in a highly coordinated fashion in most neurons (e.g., nAChRbeta1 and Rdl), and others were cell-type specific, with diverse temporal dynamics (e.g., nAChRalpha7). A number of genes implicated in synaptic function in more specific contexts were also regulated in a pan-neuronal fashion (e.g., Drep2; Andlauer et al., 2014). These included cell adhesion molecules (CAMs) that could contribute to synaptic adhesion complexes (Nrx-1, Hasp, and Lrp4; Li et al., 2007, Nakayama et al., 2016, Mosca et al., 2017). Other cell recognition molecules that contribute to connectivity in many contexts (e.g., Dscam1 and CadN) were also expressed pan-neuronally but were not specific to neurons or did not change appreciably during development. Other genes with similar expression patterns may represent uncharacterized constitutive components of synapses.

Most of the genes encoding presynaptic proteins were gradu-ally upregulated during mid-pupal stages (36–48 h APF) andpeaked60 h APF. The timing and synchrony of this expression parallels the synchronous onset of patterned stimulus-indepen-dent neuronal activity (PSINA) in all neurons at48 h APF (Akinet al., 2019). The gene encoding the gap junction protein ShakB is expressed pan-neuronally and peaks at mid-pupal stages (36–60 h APF). This would be consistent with a role for gap junctions at early stages of PSINA; indeed, attenuating ShakB activity de-presses PSINA throughout the developing visual system (B. Ba-jar, S.L.Z., and O. Akin, unpublished data). Many other pan-neuronal genes were expressed transiently at particular stages of development. These patterns may pinpoint undescribed discrete steps in neuronal differentiation in all neurons (**Figures 3.4,S8**, and 3.**S9**).

The synchrony of pan-neuronal genes suggests a common regulatory mechanism underlying their developmental timing. Several TFs exhibited temporally specific and pan-neuronally coordinated expression patterns (**Figures 3.4, 3.S8**, and **3.S9**). Among these, Blimp-1 and Hr3 are controlled by EcR as part of the ecdysone regulated transcriptional cascade. Other genes in this cascade were also expressed in a tightly coordinated fashion in all neuron types,

but they were not specific to neurons (data not shown). Another steroid hormone receptor, Hr38, was also expressed largely in a pan-neuronal fashion and was restricted to the second half of pupal development. Hr38 is an activity regulated immediate early gene (Chen et al., 2016). It has also been shown to be indirectly activated by ecdysteroids, but independently of the EcR pathway (Baker et al., 2003). Other genes expressed in a pan-neuronal fashion also suggest a role for intercellular signaling in regulating the pan-neuronal program. These include two G-protein coupled receptors: DopEcR, a receptor for both ecdysone and dopamine (Srivastava et al., 2005), and SIFaR, a receptor for the conserved neuropeptide SIFamide (Jørgensen et al., 2006).

Pan-neuronal genes can also be regulated at post-transcriptional levels. A number of RNA-binding proteins (RBPs) and long non-coding RNA (lncRNA) genes were pan-neuronally coordinated (**Figures 3.4, 3.S8**, and **3.S9**). For instance, Elav has been shown to function in neuron-specific splicing and lengthening of 30 untranslated regions (30 UTRs) of transcripts (Lisbin et al., 2001; Hilgers et al., 2012). Changes in 3' UTR length have been implicated in the regulation of RNA localization and translation, two prominent features of gene regulation in neurons (Tian and Manley, 2017). Two other Elav paralogs, Fne and Rbp9, implicated in synaptic growth at the larval neuromuscular junction (Zaharieva et al., 2015), are also expressed pan-neuronally, peaking at mid-pupal stages. RBPs involved in translational regulation, such as Bol, are also expressed panneuronally. Bol is an inhibitor of axon pruning in the mushroom body (Hoopfer et al., 2008) but may act more broadly to modulate expression of a core set of neuronal genes.

Taken together, these data suggest that the expression of pan-neuronal genes may be regulated at both the transcriptional and post-transcriptional levels and may, in addition, be coordinated by intercellular signaling.

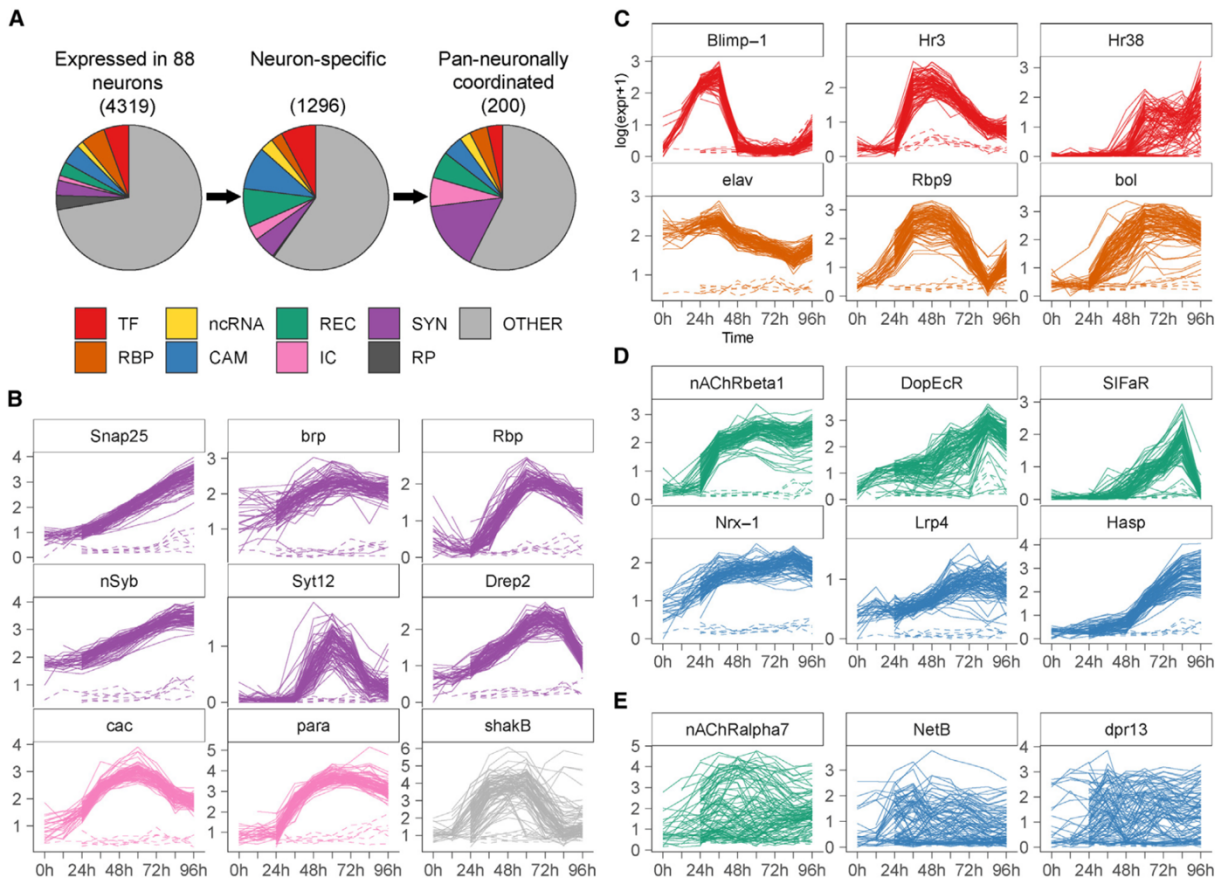


Figure 3.4. A Common (Pan-neuronal) Program of Neuronal Development

(A) Distribution of genes in different categories expressed in neurons (left), enriched in neurons compared to glia (neuron-specific) (middle), and coordinated in a pan-neuronal fashion (right). Functional categories are color coded, including transcriptions factors (TFs), RNA-binding proteins (RBPs), non-coding RNA(ncRNA), cell adhesion molecules (CAMs), receptors and ligands (RECs), ion channels (ICs), synaptic proteins (SYNs), and ribosomal genes (RP).(B–D) Expression patterns of selected pan-neuronally coordinated genes for 88 neurons (solid) and 6 glial (dashed) cell types from 24 to 96 h APF. For a subset of cell types, the expression is also shown at 0 and 12h APF (i.e., clusters from the early dataset). Colors indicate functional categories of genes as in (A). Plots Shown in log-scaled normalized expression levels. (B) Synaptic proteins and ICs. (C) Transcription factors (TFs) and RNA-binding proteins (RBPs). (D) Receptors And CAMs.(E) For comparison, examples of broadly expressed genes with cell-type-specific expression patterns that are not coordinated are shown. Expression Patterns of all 200 pan-neuronally coordinated genes ordered by average correlation values are shown in Figures 3.S8 and 3.S9

3.4.4. Cell-type-Specific Programs of Neuronal Development

Neurons exhibit great diversity in their terminal morphologies, connectivity patterns, and physiological properties. We sought to characterize differences in transcriptomes that could underlie the development of these cell-type-specific features of neurons.

To explore these differences, we focused on eight cell types representing different morphological and functional classes of neurons (**Figure 3.5A**). Some of them represent pairs of anatomically related neurons with different connectivity patterns (e.g., C2/C3 and Tm1/Tm2). We defined a set of 362 highly variable genes that were differentially expressed among these cell types (see STAR Methods for details). They represented 12% of all expressed genes. Three classes, TFs, CAMs, and receptors with ligands, accounted for 41% of highly variable genes (**Figure 3.5B**). TFs and CAMs were previously shown to be primary drivers of transcriptional diversity of neurons in different parts of the nervous system, both during development and in adults (Li et al., 2017a; Allen et al., 2020).

The expression patterns of TFs and CAMs were qualitatively different (**Figures 3.5C** and **3.5D**). Most TFs were expressed in a binary fashion across cell types (i.e., on/off manner) and were expressed throughout development. Closely related cell types often expressed similar combinations of TFs (e.g., C2/C3 cells). Each neuron expresses discrete combinations of TFs that define their cell type. CAMs were more volatile in their expression patterns (Figure 3.5D). Many were expressed transiently during development, and their specificity could be switched between cell types at different time points. Moreover, quantitative differences in expression levels between neurons or in the same neuron at different developmental times were common. For example, dpr17 was expressed at higher levels but with different temporal patterns by Tm1 (36–48 h APF) and C3 cells (48–96 h APF) and at lower levels and with similar temporal patterns by Tm2 and Mi4 cells (72–84 h APF). Thus, neurons express repertoires of CAMs with continuous variation in their expression levels. These diverse and rapidly changing cell recognition

landscapes mirror the changes in morphology and patterns of synaptic connections elaborated by different neurons.

Genes encoding neurotransmitter-specific functions exhibited interesting temporal patterns of expression. For example, the vesicular acetylcholine transporter (VAChT) is expressed exclusively by cholinergic neurons. However, the temporal dynamics and expression levels of VAChT are tightly coordinated between different types of cholinergic neurons (**Figure 3.5G**). Thus, the choice of neurotransmitter is strictly cell-type specific, but the temporal patterns of expression are shared between different cell types. Neurons also expressed genes involved in neuronal circuit functions in cell-type specific ways. These included genes encoding neurotransmitter biosynthetic enzymes, transporters, receptors, and ion channel subunits (**Figure 3.5E**). These genes can define functional diversity of neurons in adult circuits. For example, different combinations of neurotransmitter receptor subunits could result in distinct properties of synaptic connections. Interestingly, some of the genes that control synaptic communication and membrane excitability in the adult were also expressed during development. They may regulate intercellular communication underlying circuit assembly, most notably during the second half of pupal development in which each neuron type exhibits unique patterns of neuronal activity (i.e., PSINA).

Neuronal diversity is further shaped by differential expression of genes involved in post-transcriptional and post-translational mechanisms of gene regulation. We identified a number of cell-type specific RBPs and lncRNAs (**Figure 3.5F**). These genes could contribute to cell identity at post-transcriptional levels. For example, expression of the cell-type-specific RBP muscleblind (Mbl) regulates alternative splicing of pan-neuronally expressed Dscam2 to generate isoforms with different recognition specificities (Li and Millard, 2019). We also identified several genes involved in post-translational modification of proteins, including glycosylation and proteolytic enzymes, which are expressed in a cell-type-specific fashion (**Figure 3.5F**).

In summary, genetic programs of neuronal development comprise a common program of neuronal maturation regulating synapse formation and membrane excitability properties, overlaid by cell-type-specific programs encoding dramatically changing repertoires of cell surface proteins. These proteins regulate interactions between different neuron types controlling the assembly and specificity of connections between neurons.

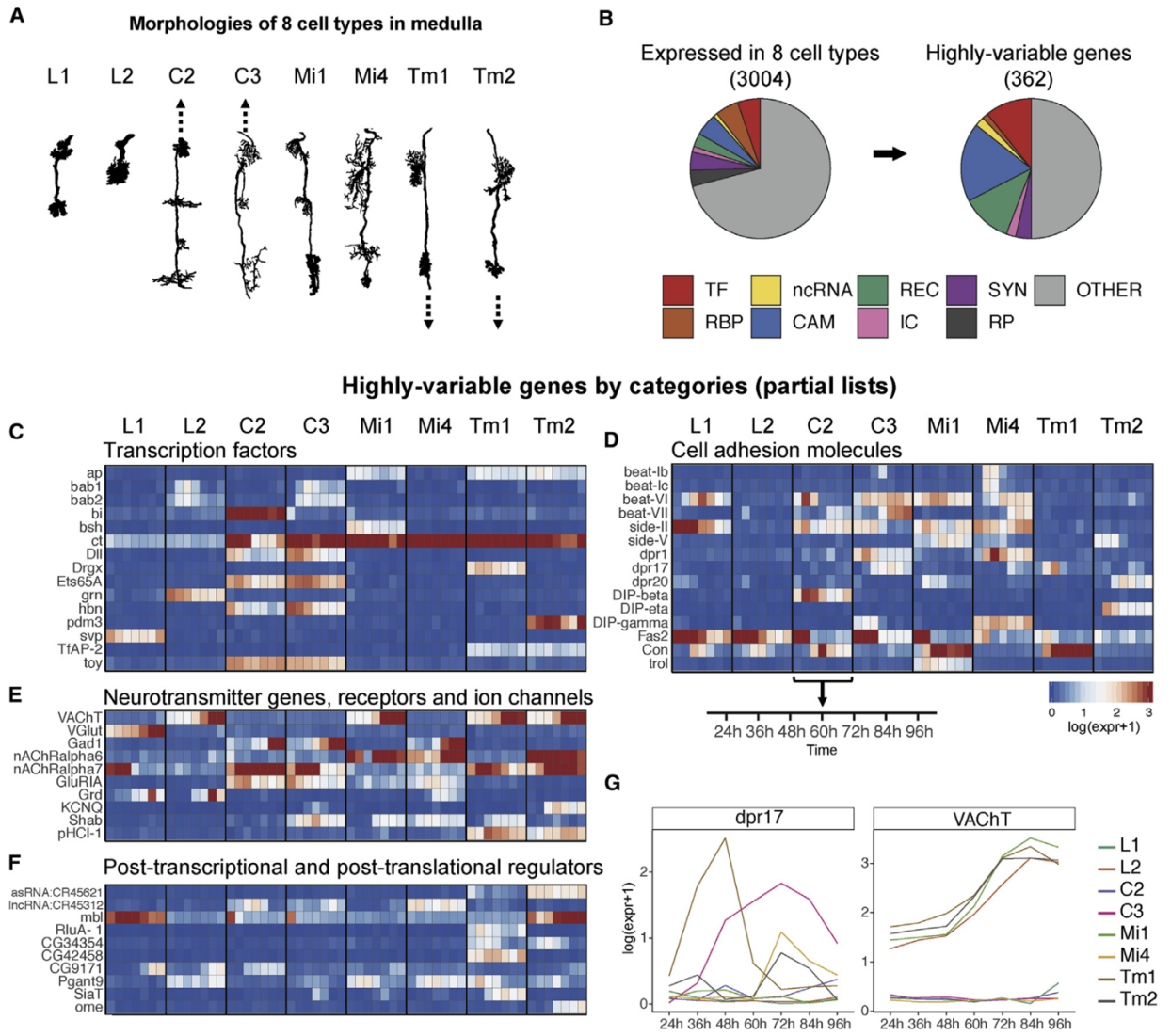


Figure 3.5. Cell-type-Specific Programs of Neuronal Development (A) Morphologies of eight neuronal cell types in the medulla (adapted from Fischbach and Dittrich, 1989). Arrows indicate direction of further extension of axons away from the cell body. (B) Distribution of functional categories of 362 highly variable genes (HVGs; see STAR Methods) expressed in any one of these eight cell types. Functional categories are color coded as in Figure 3.4. (C–F) Heatmaps with expression patterns of selected HVGs from the indicated functional categories. Expression patterns of other HVGs are shown in Figure 3.S10. (G) Examples of temporal patterns of cell-type-specific genes. dpr17 is expressed with different dynamics and expression levels in different cell types. VAChT is also expressed in cell-type-specific fashion, but temporal patterns and expression levels are coordinated across cell types. Expression patterns from seven time points at 12-h intervals between 24 and 96 h APF are shown. Values are shown in log-scaled normalized expression levels.

3.4.5. Approaching Synaptic Specificity through Transcriptomics of Synaptic Partners

Each neuron type in the fly visual system has a unique set of synaptic inputs and outputs and each of the different neuron types profiled expresses cell recognition proteins with different combinations, levels, and temporal profiles. This transcriptional atlas, in combination with the connectome, provides a resource to explore synaptic specificity mechanisms by combining it with developmental, biochemical, and genetic studies (**Figure 3.S11**).Here, as an example of how this approach may be applied, we consider a simpler part of the connectome in the lamina and use a retrospective analysis to correlate expression data with previously described biochemical and genetic studies (**Figure 3.6**).

During early pupal development, L4 neurons extend dendrites that contact L1 and L2 axons. Later in pupal development, L4 Selectively forms synapses with L2 (Rivera-Alba et al., 2011). Previous genetic studies have shown that three different families of cell recognition molecules of the immunoglobulin superfamily regulate this process. Dscam2 and Dscam4 are required for targeting of L4 dendrites to a bundled pair of L1 and L2 axons (Tadros et al., 2016). Kirre, a synaptic adhesion protein, is required for synapses between L4 and L2 (Lüthy et al., 2014), and DIP-β promotes L2 specificity (Xu et al., 2019). Previous biochemical studies have reported the binding specificities of Dscam2 and Dscam4 (Millard et al., 2007; Özkan et al., 2013), the affinities of different Dprs for binding to DIP- β (Cosmanescu et al.,2018), and the interaction of Kirre with its ligands (Özkan et al., 2013).

The temporal expression patterns of these genes in L4 are predictive of their function (**Figures 3.6A-C**). Dscam2 and Dscam4 are expressed during targeting (see **Figure 3.S11**). As their levels decrease, there is a concomitant increase in the expression of Kirre and DIP-β just prior to the onset of synaptogenesis (**Figures 3.6A** and **3.6B**). As the Kirre ligands (Kirre itself, Rst, and Hbs) (Özkan et al., 2013) are expressed in a similar way in L1 and L2, it is unlikely that Kirre alone determines specificity (**Figure 3.6A**). By contrast, the expression patterns of DIP-β ligands (i.e., Dprs) and their affinities (Cosmanescu et al., 2018), however, would favor

interactions between L4 and L2, consistent with the decrease in the fidelity of synaptic specificity seen in DIP-β mutants (Xu et al., 2019) (**Figures 3.6B** and **3.6C**). As both the Kirre and DIP-β synaptic phenotypes are partial, other recognition) molecules must act in a redundant fashion during synapse for-mation (Lüthy et al., 2014; Xu et al., 2019). We sought to address whether the atlas provides candidates for these redundant functions. Our dataset is a rich resource for doing so, allowing us to follow the expression of cell surface proteins expressed by these neurons during synapse formation, of course, in an analogous way for many other synaptic partners in the visual system (**Figure 3.S11**). At the onset of synapse formation, eight pairs of recognition proteins (both heterophilic and homophilic) are expressed in both the synaptic partner L2/L4 pair and in the L1/L4 pair, which do not form synapses (**Figure 3.6D**). Several of these proteins are associated with synapse formation in flies and mice (e.g., N-cadherin; Schwabe et al.,2014), the L1 family protein Nrg (Ango et al., 2004;Tai et al.,2019), and the mammalian Igsf9b ortholog turtle (Woo et al.,2013), and these may act in a redundant way with Kirre to regulate synapse formation. As different Beat/Side protein family pairs (Siebert et al., 2009; Özkan et al., 2013;Li et al., 2017b) are differentially expressed in L1 and L2, they may work in parallel to DIP/Dpr interactions to determine specificity. These inter-actions may promote synapses between L2/L4, inhibit inappropriate L1/L4 synapses, or both. These data support the idea that multiple protein pairs on opposing membranes may contribute to synapse formation between these neurons, and it is the differences in the cell-type-specific expression of highly diversified protein families (i.e., DIPs/Dprs and Beats/Sides) that bias the specificity of synapse formation toward L2 (Xu et al., 2019).

We have observed a similar level of complexity in the potential interactions between cell surface proteins expressed by other neurons. For example, there are over a hundred neuron types that extend axons and dendrites into the medulla neuropil and they must discriminate between one another to select appropriate synaptic partners. Each neuron expresses many dozens of cell recognition molecules in a cell-type-specific and tempo-rally dynamic fashion (**Figure 3.S11**). These findings provide a molecular correlate for the modest phenotypes seen in

many wiring mutants, alluding to marked redundancy. It raises the possibility that specificity may largely reflect biases in connectivity rather than all or none selectivity. That is, neurons may select synaptic partners not by a mechanism demanding an ‘‘exact match’’ between cell surface molecules but rather by selecting the ‘‘best match’’ among potential partners. This may confer robustness to the establishment of a complex stereotyped connectome. Unravelling the precise mechanisms by which this complexity in cellular recognition landscapes translates into highly specific patterns of synaptic connectivity will rely on future studies incorporating biochemical, genetic, and computational analyses.

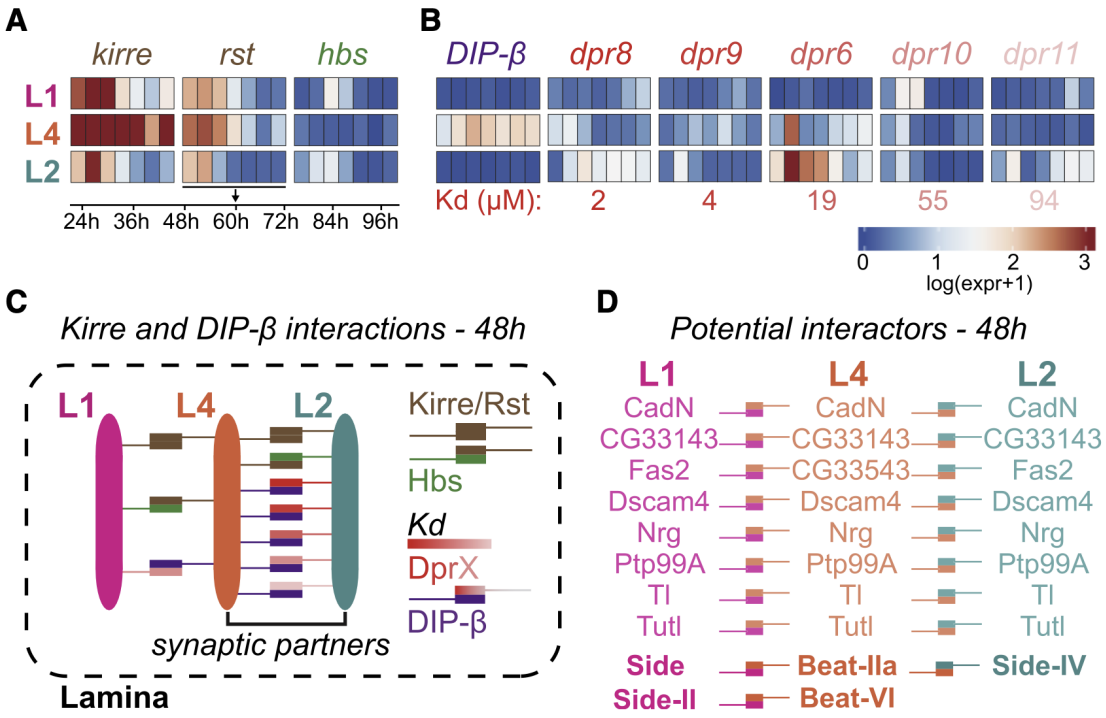


Figure 3.6. Cell Surface Interactions betweenL4 and L1 or L4 and L2 (A) Expression patterns of kirre, rst, andhbs.(B) Expression patterns of DIP-band its Dpr binding partners from scRNA-seq data in this paper. Different Dpr proteins bind to DIP-b with different affinities, as shown in red (Cosmanescu et al., 2018).(C) Protein interactions regulating synapses between L4 and L2 are summarized. Kirre is required for synapses, but its ligands are not differentially expressed between L1 and L2. DIP-b biases syn-apses to L2. This may be the result of differential expression of Dpr paralogs with different affinities between L1 and L2 (Xu et al., 2019). Interactions shown are based on genes expressed at 48 h (i.e.,onset of synapse formation). (D) Many other cell adhesion proteins of the immunoglobulin, leucine-rich repeat, and cadherin superfamilies are also expressed inL1, L2, and L4 at 48 h. Based On Biochemical Studies (Özkan et al., 2013), these proteins may bind to each other on the surface of opposing membranes as indicated. Values are shown in log-scaled normalized expression levels. For expression patterns of these proteins, see Figure 3.S11. See text for further discussion.

3.5. Concluding Remarks

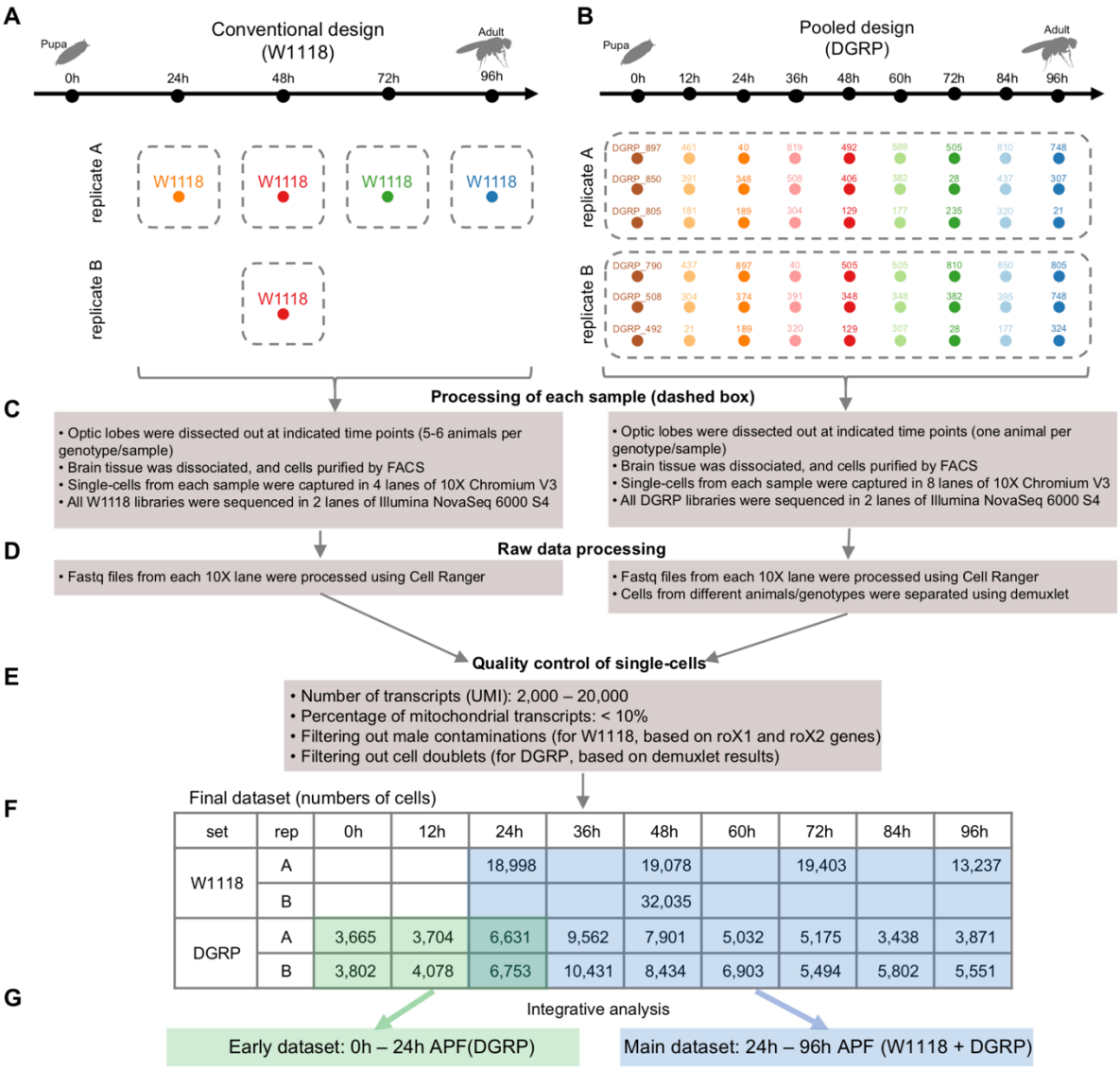
Understanding how patterns of synaptic connections are determined during development remains a central question in neural development. In this study, we followed the development of 88 different neuron types at many time points spanning synaptogenesis. In so doing, we uncovered a pan-neuronal program running synchronously in all neurons regardless of cell type. This program was enriched, in particular, for synaptic proteins and proteins regulating electrically excitable membranes. The neuron-specific expression of pan-neuronal genes may be controlled by a common gene regulatory program. Alternatively, each cell type may regulate expression of these genes independently through a myriad of cell-type specific regulatory programs. The coordinated expression of these genes in all neurons may be controlled by a separate timing mechanism. Whatever the underlying mechanisms are for regulating this process in a similar way in all neurons, we speculate that this program endows neurons with a broad potential to form synapses.

By contrast to the pan-neuronal program of development, the expression of cell recognition molecules is profoundly asynchronous and highly cell-type specific. This diversity mirrors cell type-specific morphological changes of axons and dendrites in the target region and the complexity and specificity of synaptic connectivity each neuron elaborates. These rapid changes may be necessary to choreograph the specific interactions between the processes of many different neuron types as they seek to discriminate between appropriate and inappropriate synaptic partners and determine the number and distribution of synapses between them. Whether the dynamic expression pattern we observe is largely cell autonomous in nature or whether it reflects a dialog between neurons during circuit assembly is unknown. Many signaling pathway components, including those regulated by neural activity, growth factors, neuropeptides, and ecdysteroids, all of which can influence gene expression, are expressed in these neurons.

Traditionally, circuit assembly has been studied genetically one gene at a time. Molecules to study were selected based upon their patterns of expression or phenotypes of mutants. In these

approaches, investigators were largely ignorant of other recognition molecules expressed in these neurons. Phenotypes in almost all cases were incomplete and often embarrassingly mild, supporting a general consensus in the field that this reflects substantial redundancy, which is advantageous for constructing a robust circuit but a bane to researchers. Transcriptome profiling revealed that each neuron expresses a unique combination of a vast number of cell recognition molecules, many of which have the potential to promote interactions between them. These findings and many genetic studies underscore the confound of redundancy in studying the logic and molecular mechanisms underlying specificity. In a broad sense we speculate that these proteins act together through common mechanisms and variations on them to sculpt circuitry. Uncovering the molecular logic of synaptic specificity will rely on combining transcriptomics with detailed studies of specific synapses using an increasingly sophisticated tool kit of genetic, biochemical, physiological, and imaging approaches.

The development of high-throughput single-cell technologies has led to an exponential increase in transcriptional studies in many different systems (Zheng et al., 2017). Profiling of multiple experimental conditions with internal controls for technical variation in a single experiment contributes to these advances (Kang et al., 2018). The pooled strategy, as we describe here, can be easily adapted for a variety of experimental designs, including comparative studies in different mutant backgrounds or in altered environments. In principle, similar approaches could be used in many other systems, where the genotype information could be obtained in parallel to transcriptional profiling. The rapid and economically feasible generation of temporally resolved transcriptional datasets can be used to tackle a variety of longstanding questions in developmental biology.



set	rep	0h	12h	24h	36h	48h	60h	72h	84h	96h
W1118	A			18,998		19,078		19,403		13,237
	B					32,035				
DGRP	A	3,665	3,704	6,631	9,562	7,901	5,032	5,175	3,438	3,871
	B	3,802	4,078	6,753	10,431	8,434	6,903	5,494	5,802	5,551

Figure 3.S1. The experimental design and workflow of the analysis. Top panels indicate genotype combinations used for each time point and replicate.

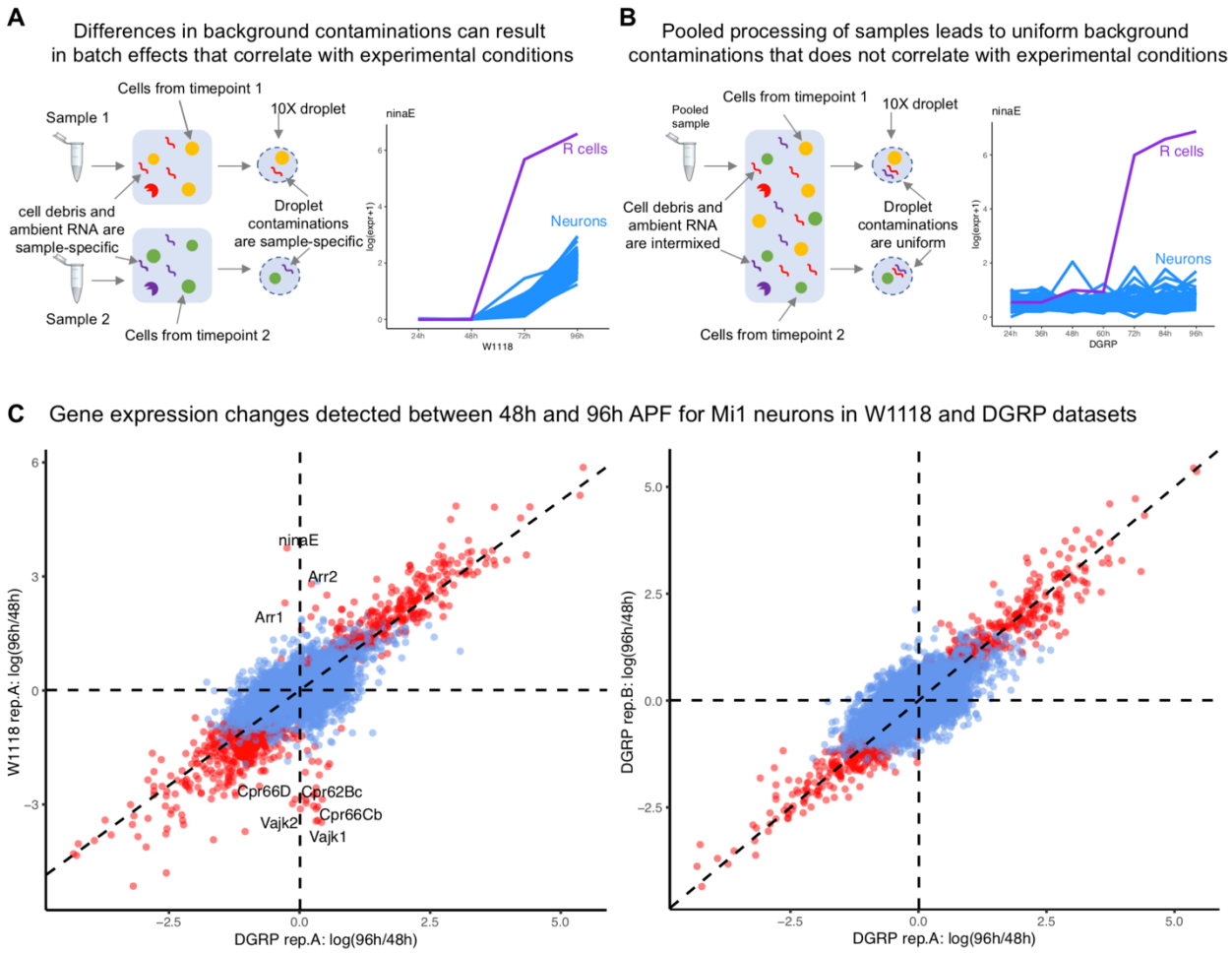


Figure 3.S2. Background contamination improved from pooled processing. (A-B) Cell free (ambient) RNA contaminations in conventional (W1118), and pooled (DGRP) designs. In droplet-based scRNA-Seq, cell-free mRNA molecules present in cell suspensions are a source of contaminants confounding time course studies. For example, the main rhodopsin gene (ninaE) is expressed exclusively by photoreceptors (R cells). It is highly abundant at later stages of development, and, thus, contributes to background ambient RNA in a time-dependent manner. As a consequence, in W1118 dataset, ninaE expression erroneously appears to increase in all cell types during development. In DGRP dataset, the background levels of ninaE is still detectable in single-cell transcriptomes, but the level does not correlate with time. (C) Gene expression changes between 48h and 96h APF for Mi1 neurons. Log2-fold change values are shown separately for DGRP replicate A (X-axis), DGRP replicate B (Y axis, right), and W1118 replicate A samples (Y axis, left). Genes with statistically significant changes in either comparison are shown in red. Detected changes were highly concordant between two DGRP replicates (right). Some genes were significantly upregulated or downregulated exclusively in W1118 dataset. Among those, genes abundantly expressed in adult photoreceptors (e.g. ninaE, Arr1, Arr2), and a group of cuticular proteins of unknown origin specifically enriched in 48h samples (e.g. Vajk1, Cpr66Cb and others). We assume that these genes represent time point-dependent contaminations of single-cell transcriptomes.

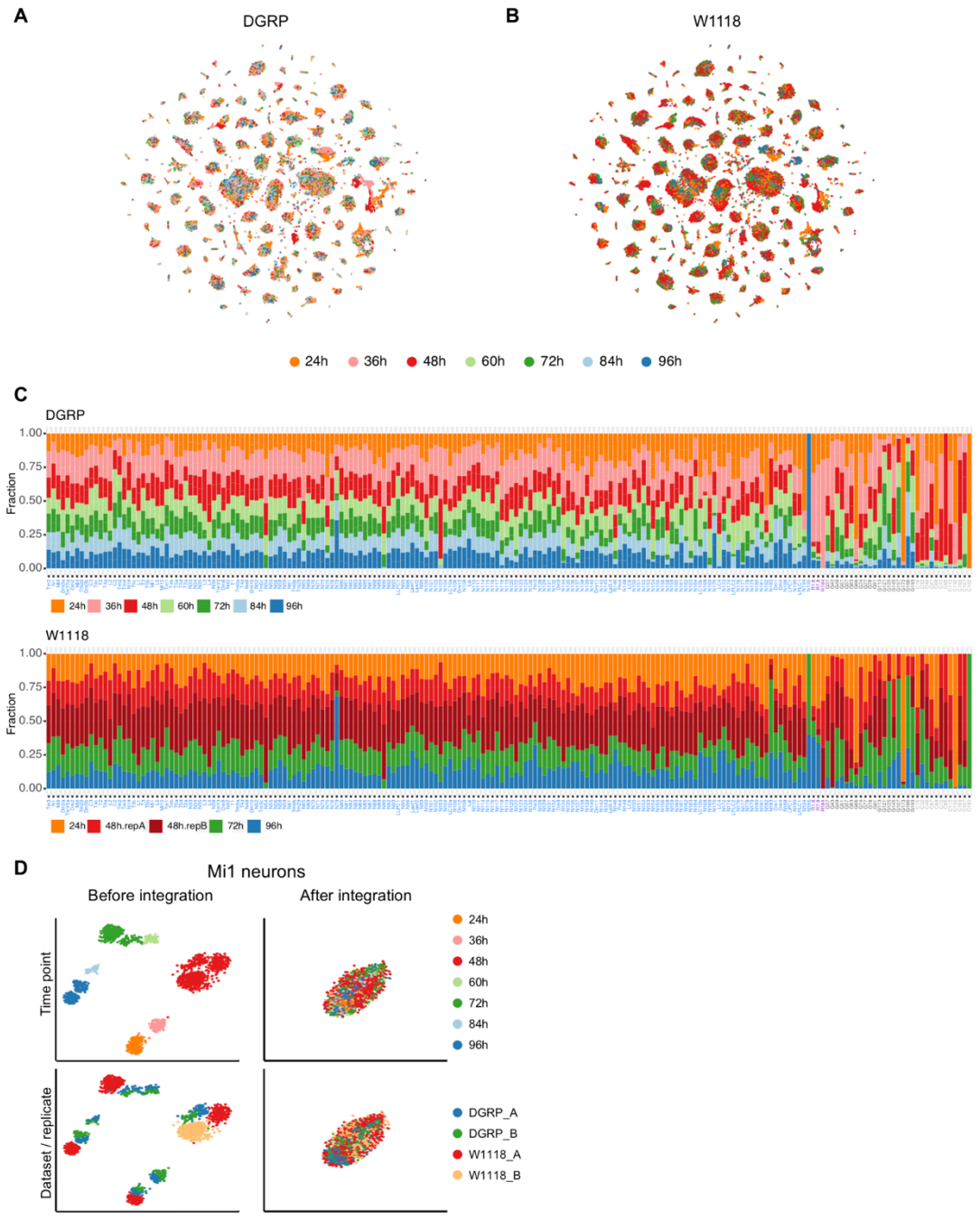


Figure 3.S3. Data integration (A-B) tSNE plots for the main dataset (24-96h APF). Cells are color coded by time points, as indicated. tSNE plots are shown separately, for DGRP (A) and W1118 (B) datasets. (C) Proportion of cells in each cluster by datasets and time points. Cluster labels are color coded and ordered by cell classes (as in Figure S4). For each major class, clusters are ordered by sizes from the largest (left) to the smallest (right). (D) tSNE plots for Mi1 neurons before and after dataset integration. Before integration cells form separate clusters for each time points and sample (i.e. dataset/replicate). Dimensionality reduction before and after integration is performed using same set of genes and parameters.

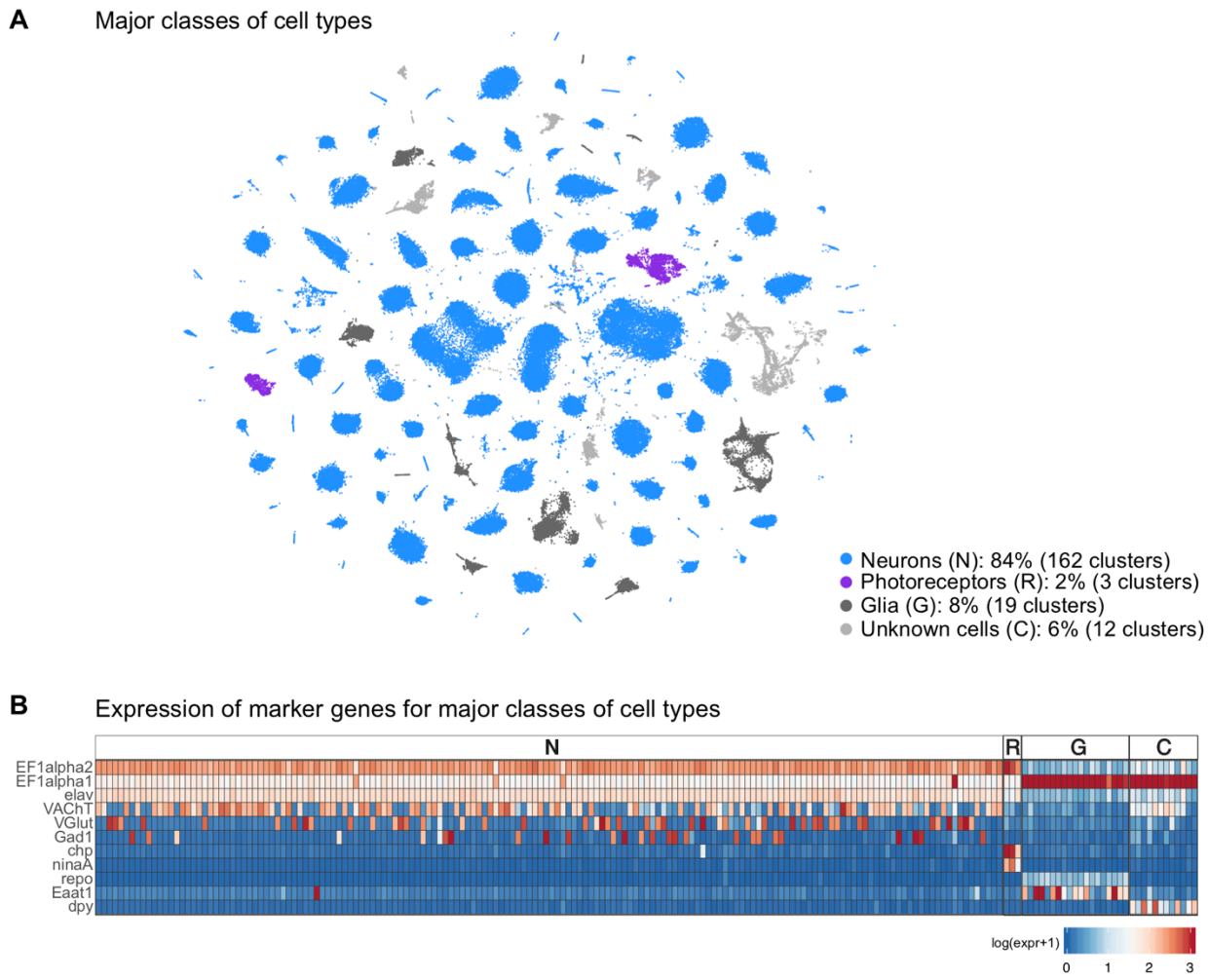


Figure 3.S4. Cell type class composition from dataset

(A) tSNE plots for the main dataset (24-96h APF). Cells are color-coded by major classes, as indicated. (B) Heatmap with expression patterns of known and novel marker genes for cell classes. Each column represents a single transcriptional cluster. Clusters are grouped by cell classes. Expression levels are averaged across all samples and timepoints (from 24h to 96h APF). Values are shown in log-scaled normalized expression levels, as indicated.

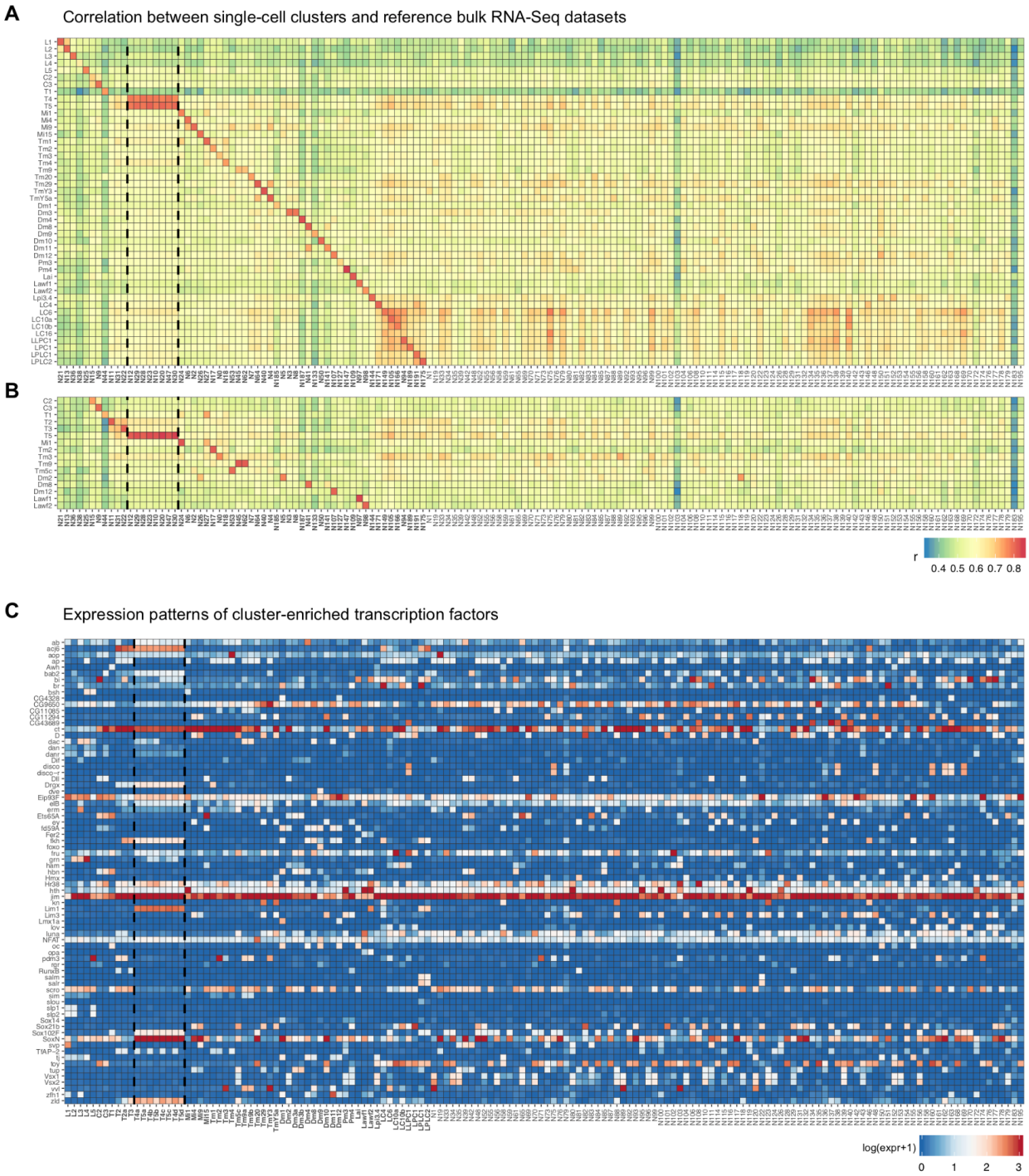

A
Correlation between single-cell clusters and reference bulk RNA-Seq datasets
B
r
0.4 0.5 0.6 0.7 0.8
C
Expression patterns of cluster-enriched transcription factors
log(expr+1)
0 1 2 3

Figure 3.S5. (A-B) Heatmaps with correlation between reference bulk RNA-Seq datasets from previous studies (rows) and single-cell transcriptional clusters from the data in this study (columns). Comparisons are shown for datasets from Davis et al., 2020 (A) and Konstantinides et al., 2018 (B). (C) Expression patterns of cluster-enriched TF were used for additional curation of matches between clusters and cell types. For example, T4/T5 reference datasets matched with eight transcriptional clusters with similar correlation values (bracketed by dashed lines). These eight clusters correspond to known subtypes of T4/T5 identified genes (Kurmangaliyev et al., 2019). Clusters that were matched with known cell types were renamed as indicated on X-axis of (C). Values are shown in log-scaled normalized expression levels, as indicated.

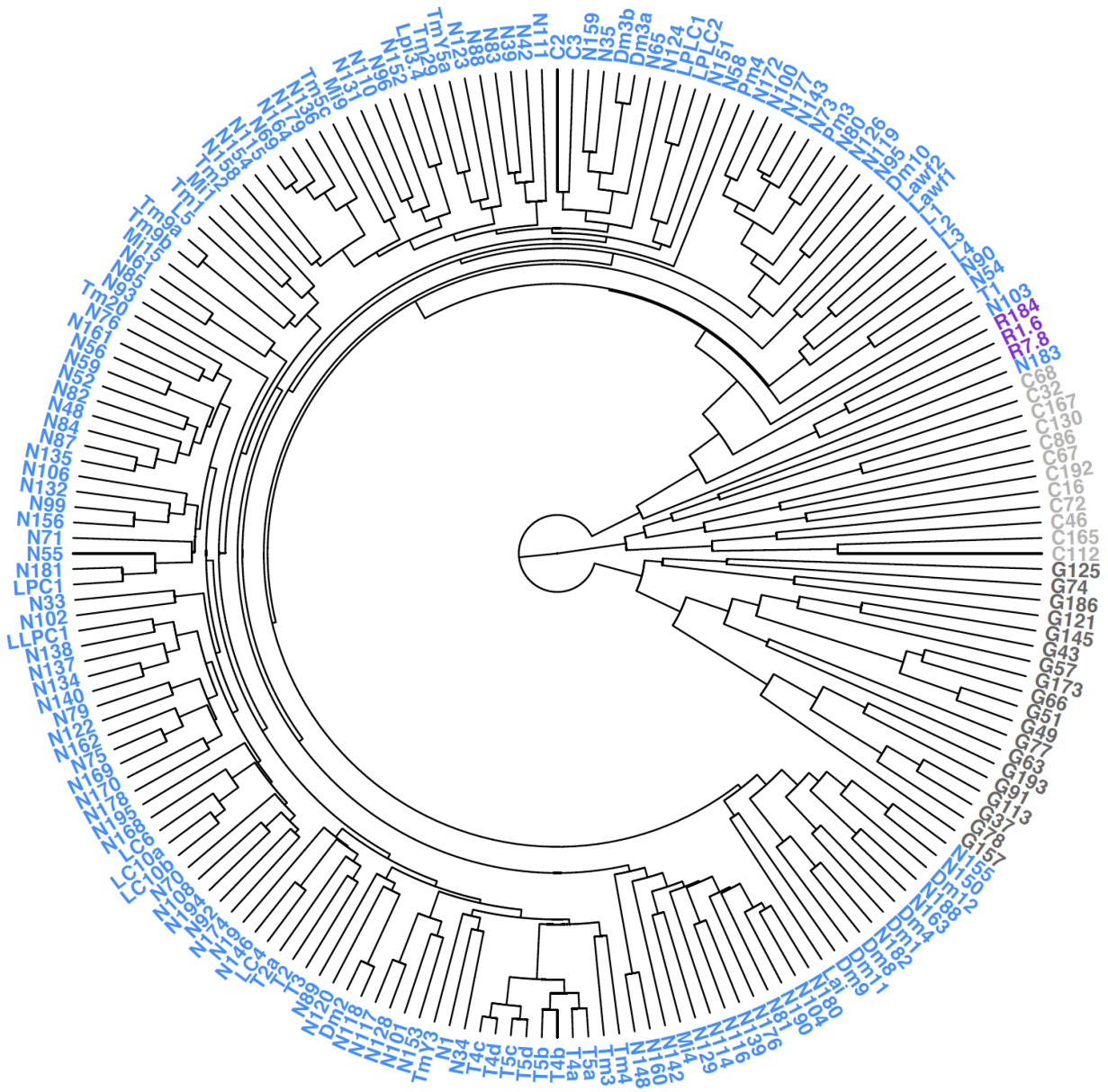


Figure 3.S6. Dendrogram of similarities between transcriptional clusters. Clusters are color coded by cell classes, as in Figure 3.S4. Note that similarities were estimated for the integrated dataset and represent the average similarity across entire dataset. The relationships between clusters can vary across different time points.

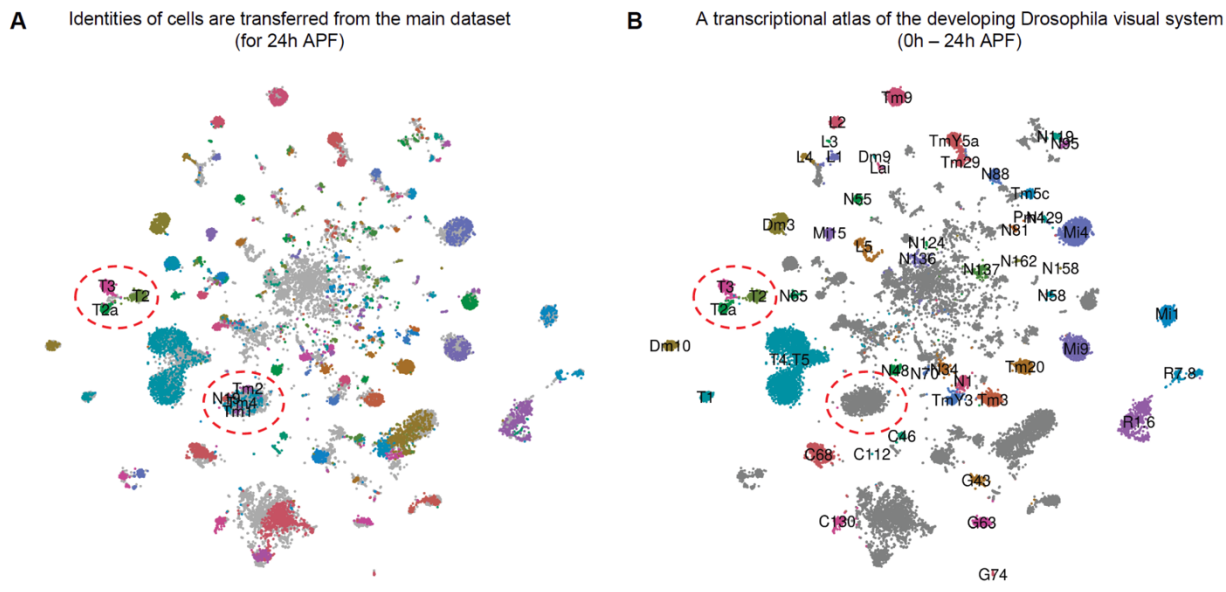


Figure 3.S7. **tSNE plot of the early dataset.** The early dataset includes samples from 0h to 24h APF (DGRP). All samples were integrated and clustered together. Cells from 24h samples were also analyzed as a part of main dataset. Identities of these cells in the main dataset are color-coded as in Figure 3.2A. Clusters with one-to-one match between the early and the main datasets were named as in Figure 2B. Examples of clusters that could be followed between the early and the main datasets (e.g. T2/T2a/T3), and examples of clusters that were unresolvable in the early dataset (e.g. Tm1, Tm2, Tm4, and cluster N19) are encircled by the red dashed lines.

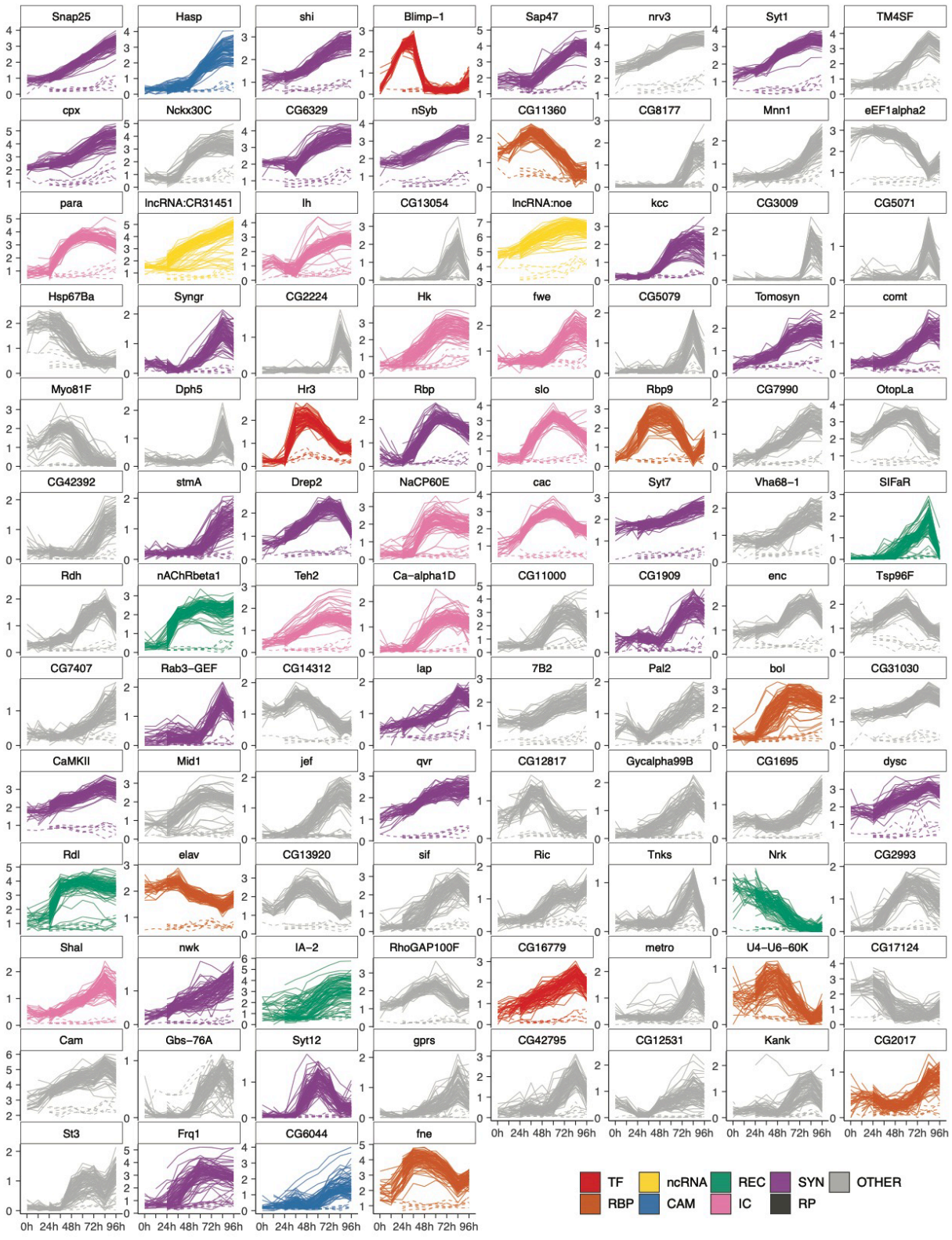

Snap25
Hasp
shi
Blimp−1
Sap47
nrv3
Syt1
TM4SF
cpx
Nckx30C
CG6329
nSyb
CG11360
CG8177
Mnn1
eEF1alpha2
para
lncRNA:CR31451
Ih
CG13054
lncRNA:noe
kcc
CG3009
CG5071
Hsp67Ba
Syngr
CG2224
Hk
fwe
CG5079
Tomosyn
comt
Myo81F
Dph5
Hr3
Rbp
slo
Rbp9
CG7990
OtopLa
CG42392
stmA
Drep2
NaCP60E
cac
Syt7
Vha68−1
SIFaR
Rdh
nAChRbeta1
Teh2
Ca−alpha1D
CG11000
CG1909
enc
Tsp96F
CG7407
Rab3−GEF
CG14312
lap
7B2
Pal2
bol
CG31030
CaMKII
Mid1
jef
qvr
CG12817
Gycalpha99B
CG1695
dysc
Rdl
elav
CG13920
sif
Ric
Tnks
Nrk
CG2993
Shal
nwk
IA−2
RhoGAP100F
CG16779
metro
U4−U6−60K
CG17124
Cam
Gbs−76A
Syt12
gprs
CG42795
CG12531
Kank
CG2017
St3
Frq1
CG6044
fne
0h 24h 48h 72h 96h
TF
RBP
ncRNA
CAM
REC
IC
SYN
RP
OTHER

Figure 3.S8. **Expression patterns of pan-neuronally coordinated genes ordered by average correlation.** Second half is shown in Figure 3.S7. Transcription factors (TF); RNA-binding proteins (RBP); non-coding RNA (ncRNA); cell adhesion molecules (CAM); receptors and ligands (REC); ion channels (IC); synaptic proteins (SYN); ribosomal genes (RP). See legend of Figure 3.4 for details.

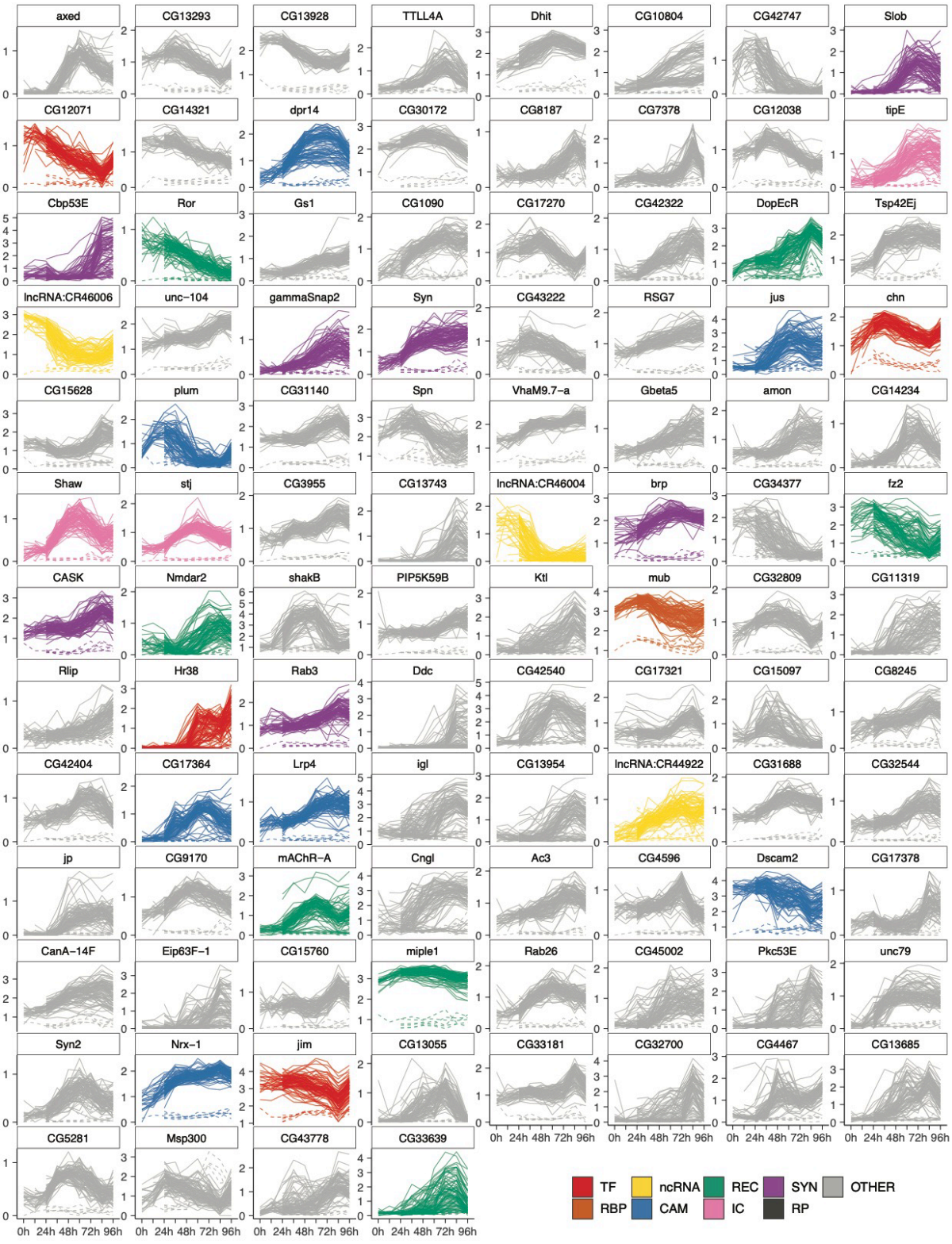

axed
CG13293
CG13928
TTLL4A
Dhit
CG10804
CG42747
Slob
CG12071
CG14321
dpr14
CG30172
CG8187
CG7378
CG12038
tipE
Cbp53E
Ror
Gs1
CG1090
CG17270
CG42322
DopEcR
Tsp42Ej
lncRNA:CR46006
unc−104
gammaSnap2
Syn
CG43222
RSG7
jus
chn
CG15628
plum
CG31140
Spn
VhaM9.7−a
Gbeta5
amon
CG14234
Shaw
stj
CG3955
CG13743
lncRNA:CR46004
brp
CG34377
fz2
CASK
Nmdar2
shakB
PIP5K59B
Ktl
mub
CG32809
CG11319
Rlip
Hr38
Rab3
Ddc
CG42540
CG17321
CG15097
CG8245
CG42404
CG17364
Lrp4
igl
CG13954
lncRNA:CR44922
CG31688
CG32544
jp
CG9170
mAChR−A
Cngl
Ac3
CG4596
Dscam2
CG17378
CanA−14F
Eip63F−1
CG15760
miple1
Rab26
CG45002
Pkc53E
unc79
Syn2
Nrx−1
jim
CG13055
CG33181
CG32700
CG4467
CG13685
CG5281
Msp300
CG43778
CG33639
0h 24h 48h 72h 96h
TF
RBP
ncRNA
CAM
REC
IC
SYN
RP
OTHER

Figure 3.S9. Expression patterns of pan-neuronally coordinated genes ordered by average correlation. First half is shown in Figure 3.S6. Transcriptions factors (TF); RNA-binding proteins (RBP); non-coding RNA (ncRNA); cell adhesion molecules (CAM); receptors and ligands (REC); ion channels (IC); synaptic proteins(SYN); ribosomal genes (RP). See legend of Figure 3.4 for details.

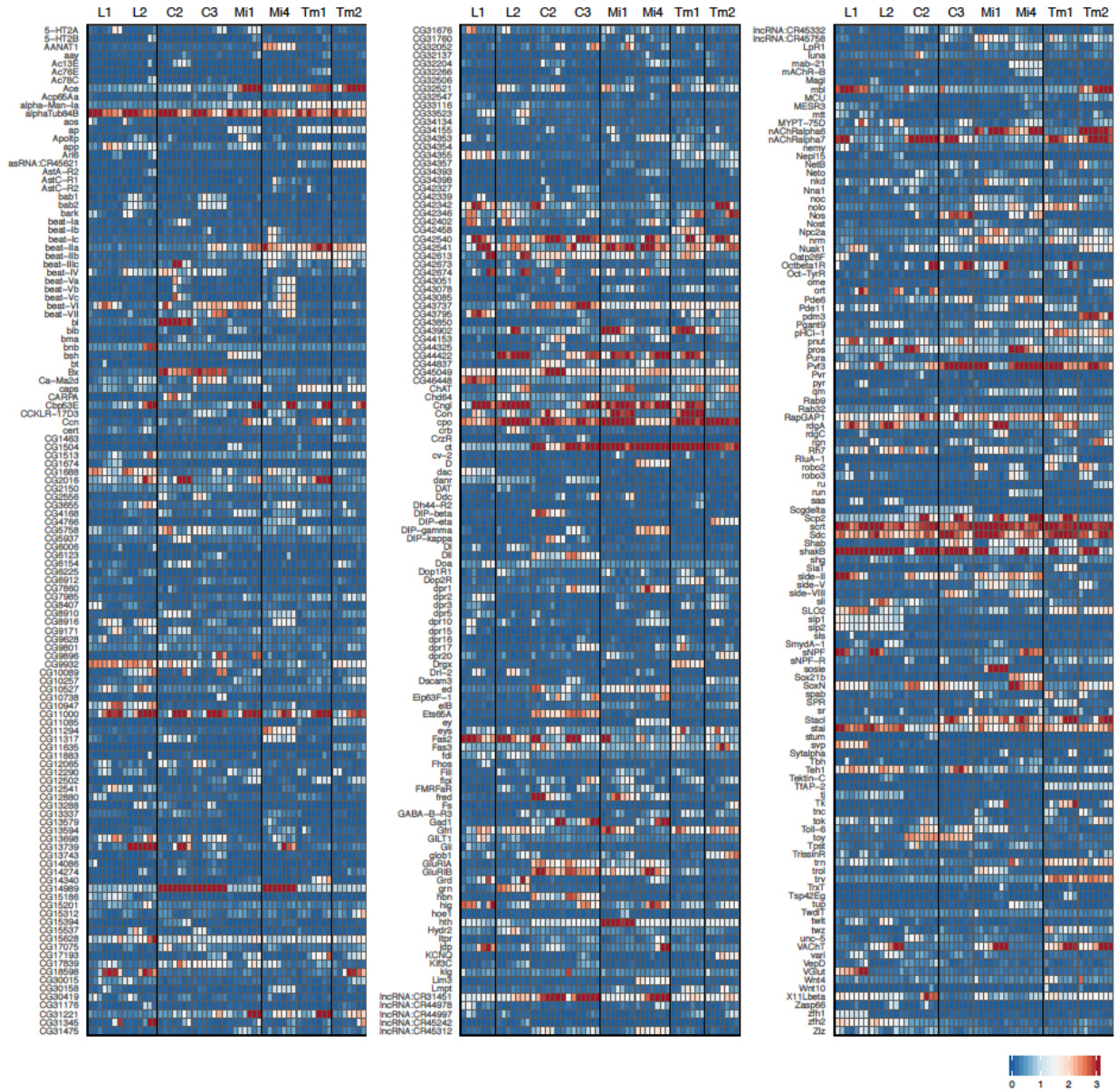


Figure 3.S10. Expression patterns of highly variable genes. See legend of Figure 3,4 for details.

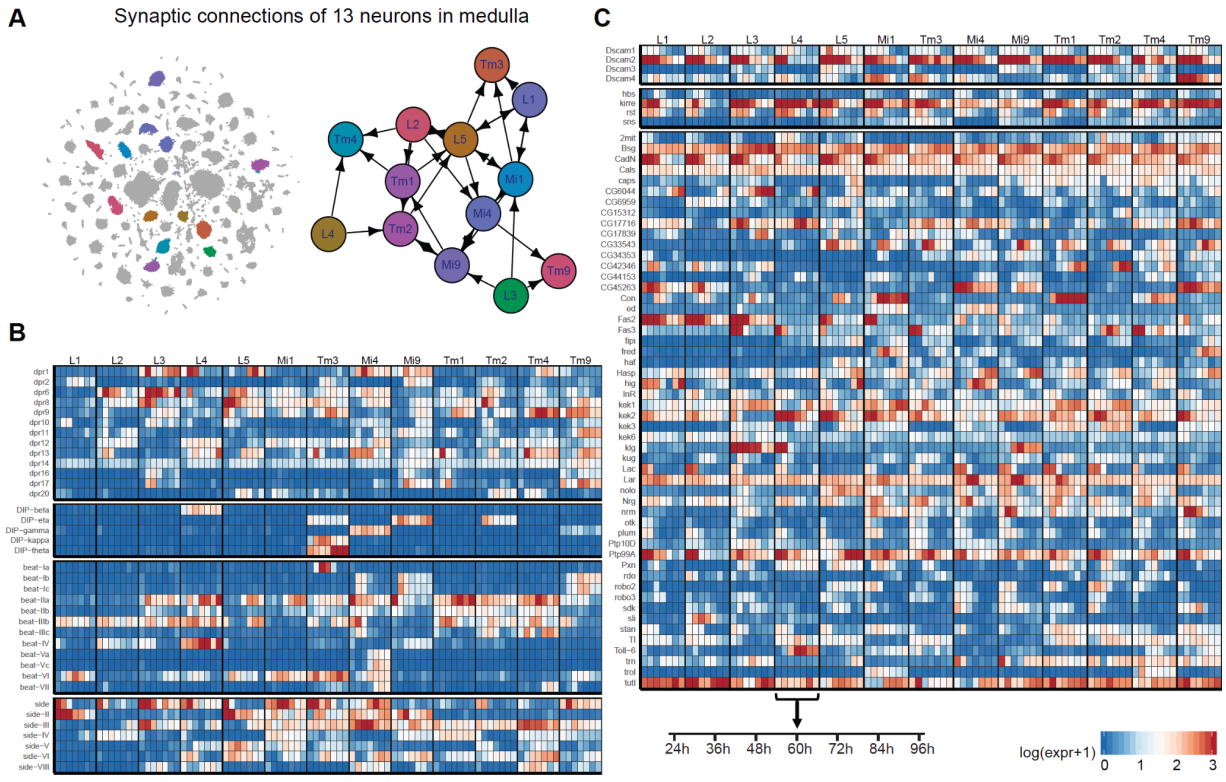


Figure 3.S11. (A) Transcriptomes and connectomes of 13 neuronal cell types in medulla. Connections are shown for synaptic partners with minimum 10 synapses (adapted from Takemura et al., 2015). (BC) Expression patterns of Ig, LRR and Cadherin superfamily proteins expressed in these cell types (min. Expression > 5). Several families of proteins are grouped together. Values are shown in log-scaled normalized expression levels.

3.6. References

Allen, A. M., Neville, M. C., Birtles, S., Croset, V., Treiber, C. D., Waddell, S. & Goodwin, S. F. (2020). A single-cell transcriptomic atlas of the adult Drosophila ventral nerve cord. *ELife*, *9*, e54074. https://doi.org/10.7554/elife.54074

Cosmanescu, F., Katsamba, P. S., Sergeeva, A. P., Ahlsen, G., Patel, S. D., Brewer, J. J., Tan, L., Xu, S., Xiao, Q., Nagarkar-Jaiswal, S., Nern, A., Bellen, H. J., Zipursky, S. L., Honig, B. & Shapiro, L. (2018). *Neuron*-Subtype-Specific Expression, Interaction Affinities, and Specificity Determinants of DIP/Dpr *Cell* Recognition Proteins. Neuron, *100*(6), 1385-1400.e6. https://doi.org/10.1016/j.neuron.2018.10.046

Courgeon, M. & Desplan, C. (2019). Coordination between stochastic and deterministic specification in the Drosophila visual system. *Science*, *366*(6463), eaay6727-13. https://doi.org/10.1126/science.aay6727

Croset, V., Treiber, C. D. & Waddell, S. (2018). Cellular diversity in the Drosophila midbrain revealed by single-cell transcriptomics. ELife, 7, e34550. https://doi.org/10.7554/elife.34550

Davie, K., Janssens, J., Koldere, D., Waegeneer, M. D., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., González-Blas, C. B., Poovathingal, S., Hulselmans, G., Spanier, K. I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. Cell, *174*(4), *982*-998.e20. https://doi.org/10.1016/j.cell.2018.05.057

Davis, F. P., Nern, A., Picard, S., Reiser, M. B., Rubin, G. M., Eddy, S. R. & Henry, G. L. (2020). A genetic, genomic, and computational resource for exploring neural circuit function. ELife, 9, e50901. https://doi.org/10.7554/elife.50901

Dickson, B. J. (2002). Molecular Mechanisms of Axon Guidance. Science, *298*(5600), 1959–1964. https://doi.org/10.1126/science.1072165

Enriquez, J., Venkatasubramanian, L., Baek, M., Peterson, M., Aghayeva, U. & Mann, R. S. (2015). Specification of Individual Adult Motor Neuron Morphologies by Combinatorial Transcription Factor Codes. Neuron, *86*(4), 955–970. https://doi.org/10.1016/j.neuron.2015.04.011

Fear, J. M., León-Novelo, L. G., Morse, A. M., Gerken, A. R., Lehmann, K. V., Tower, J., Nuzhdin, S. V. & McIntyre, L. M. (2016). Buffering of Genetic Regulatory Networks in Drosophila melanogaster. *Genetics*, *203*(3), 1177–1190. https://doi.org/10.1534/genetics.116.188797

Fischbach, K. F. & Dittrich, A. (1989). The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. *Cell and Tissue Research*. http://link.springer.com/article/10.1007/BF00218858

Hasegawa, E., Kaido, M., Takayama, R. & Sato, M. (2013). Brain-specific-homeobox is required for the specification of neuronal types in the Drosophila optic lobe. *Developmental Biology*, *377*(1), 90–99. https://doi.org/10.1016/j.ydbio.2013.02.012

Hobert, O. & Kratsios, P. (2019). ScienceDirect Neuronal identity control by terminal selectors in worms, flies, and chordates. *Current Opinion in Neurobiology*, *56*, 97–105. https://doi.org/10.1016/j.conb.2018.12.006

Holguera, I. & Desplan, C. (2018). Neuronal specification in space and time. Science, *362*(6411), 176–180. https://doi.org/10.1126/science.aas9435

Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ràmia, M., Tarone, A. M., Turlapati, L., Zichner, T., Zhu, D., Lyman, R. F., Magwire, M. M., Blankenburg, K., Carbone, M. A., Chang, K., Ellis, L. L., Fernandez, S., Han, Y., Highnam, G., Hjelmen, C. E., … Mackay, T. F. C. (2014). Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. *Genome Research*, *24*(7), 1193–1208. https://doi.org/10.1101/gr.*171*546.113

Kang, H. M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., Wan, E., Wong, S., Byrnes, L., Lanata, C. M., Gate, R. E., Mostafavi, S., Marson, A., Zaitlen, N., Criswell, L. A. & Ye, C. J. (2017). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nature Biotechnology*, *36*(1), 89–94. https://doi.org/10.1038/nbt.4042

Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R. & Desplan, C. (2018). *Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features*. *174*(3), 622-635.e13. https://doi.org/10.1016/j.cell.2018.05.021

Kurmangaliyev, Y. Z., Yoo, J., LoCascio, S. A. & Zipursky, S. L. (2019). Modular transcriptional programs separately define axon and dendrite connectivity. ELife, 8, 505. https://doi.org/10.7554/elife.50822

Li, H., Horns, F., Wu, B., Xie, Q., Li, J., Li, T., Luginbuhl, D. J., Quake, S. R. & Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-Cell RNA Sequencing. Cell, 171(5), 1206-1207.e22. https://doi.org/10.1016/j.cell.2017.10.019

Liu, J., Reggiani, J. D. S., Laboulaye, M. A., Pandey, S., Chen, B., Rubenstein, J. L. R., Krishnaswamy, A. & Sanes, J. R. (2018). Tbr1 instructs laminar patterning of retinal ganglion cell dendrites. *Nature Neuroscience*, *21*(5), 659–670. https://doi.org/10.1038/s41593-018-0127-z

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., Casillas, S., Han, Y., Magwire, M. M., Cridland, J. M., Richardson, M. F., Anholt, R. R. H., Barrón, M., Bess, C., Blankenburg, K. P., Carbone, M. A., Castellano, D., Chaboub, L., Duncan, L., … Gibbs, R.

A. (2012). The Drosophila melanogaster Genetic Reference Panel. Nature, *482*(7384), 173–178. https://doi.org/10.1038/nature10811

Menon, K. P., Kulkarni, V., Takemura, S.-Y., Anaya, M. & Zinn, K. (2019). *Interactions between Dpr11 and DIP-γ control selection of amacrine neurons in Drosophila color vision circuits*. *8*, 13901–13932. https://doi.org/10.7554/elife.48935

Morey, M., Yee, S. K., Herman, T., Nern, A., Blanco, E. & Zipursky, S. L. (2008). Coordinate control of synaptic-layer specificity and rhodopsins in photoreceptor neurons. Nature, *456*(7223), 795–799. https://doi.org/10.1038/nature07419

Özel, M. N., Simon, F., Jafari, S., Holguera, I., Chen, Y.-C., Benhra, N., El-Danaf, R. N., Kapuralin, K., Malin, J. A., Konstantinides, N. & Desplan, C. (2020). Neuronal diversity and convergence in a visual system developmental atlas. Nature, *589*(7840), 88–95. https://doi.org/10.1038/s41586-020-2879-3

Sanes, J. R. & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell, *181*(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Sarin, S., Zuniga-Sanchez, E., Kurmangaliyev, Y. Z., Cousins, H., Patel, M., Hernandez, J., Zhang, K. X., Samuel, M. A., Morey, M., Sanes, J. R. & Zipursky, S. L. (2018). Role for Wnt Signaling in Retinal Neuropil Development: Analysis via RNA-Seq and In Vivo Somatic CRISPR Mutagenesis. Neuron, 98(1), 109-126.e8. https://doi.org/10.1016/j.neuron.2018.03.004

Takemura, S., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P. K., Katz, W. T., Olbris, D. J., Plaza, S. M., Winston, P., Zhao, T., Horne, J. A., Fetter, R. D., Takemura, S., Blazek, K., Chang, L.-A., Ogundeyi, O., Saunders, M. A., Shapiro, V., … Chklovskii, D. B. (2013). A visual motion detection circuit suggested by Drosophila connectomics. Nature, *500*(7461), 175–181. https://doi.org/10.1038/nature12450

Takemura, S.-Y., Xu, C. S., Lu, Z., Rivlin, P. K., Parag, T., Olbris, D. J., Plaza, S., Zhao, T., Katz, W. T., Umayam, L., Weaver, C., Hess, H. F., Horne, J. A., Nunez-Iglesias, J., Aniceto, R., Chang, L.-A., Lauchie, S., Nasca, A., Ogundeyi, O., … Scheffer, L. K. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(44), 13711–13716. https://doi.org/10.1073/pnas.1509820112

Tan, L., Zhang, K. X., Pecot, M. Y., Nagarkar-Jaiswal, S., Lee, P.-T., Takemura, S.-Y., McEwen, J. M., Nern, A., Xu, S., Tadros, W., Chen, Z., Zinn, K., Bellen, H. J., Morey, M. & Zipursky, S. L. (2015). Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. Cell, *163*(7), 1756–1769. https://doi.org/10.1016/j.cell.2015.11.021

Tessier-Lavigne, M. & Goodman, C. S. (1996). The Molecular Biology of Axon Guidance. Science, *274*(5290), 1123–1133. https://doi.org/10.1126/science.274.5290.1123

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*(1), 14049. https://doi.org/10.1038/ncomms14049

CHAPTER FOUR

Brain wiring determinants uncovered by integrating connectomes and transcriptomes

4.1 Abstract

How do developing neurons select their synaptic partners? To identify molecules matching synaptic partners, we integrated the synapse-level connectome of neural circuits with the developmental expression patterns and binding specificities of cell adhesion molecules (CAMs) on pre- and postsynaptic neurons. We focused on parallel synaptic pathways in the Drosophila visual system, in which closely related neurons choose between two alternative synaptic targets. We show that the choice of synaptic partners correlates with the matching expression of receptor-ligand pairs of Beat and Side proteins of the immunoglobulin superfamily (IgSF) CAMs. Genetic analysis demonstrates that these proteins determine the choice between alternative synaptic targets. Combining transcriptomes, connectomes, and protein interactome maps provides a framework to uncover the molecular logic of synaptic connectivity.

4.2 Introduction

Advances in electron microscopy (EM) level connectomics have demonstrated the extraordinary complexity and specificity of synaptic connectivity (Dorkenwald et al., 2021; Motta et al., 2019; Scheffer et al., 2020; Scheffer & Meinertzhagen, 2019; Shinomiya et al., 2019, 2022; Zheng et al., 2018). Developing neurons encounter the axons and dendrites of many different neuron types and form synapses with only a subset of them. Considerable progress has been made in understanding the molecular mechanisms by which neurons reach their target regions (axon guidance (Kolodkin & Tessier-Lavigne, 2011)) and form synapses (synaptogenesis (Südhof, 2021)). Here we focus on the less well-understood step of synaptic specificity, how neurons choose their synaptic partners (Sanes & Zipursky, 2020).

Sperry proposed that molecular labels allow neurites to discriminate between one another (Sperry, 1963). Identifying these labels and understanding how they work is key to uncovering the molecular basis of brain wiring. Neurons express complex repertoires of many different cell

surface molecules that mediate heterophilic and homophilic binding (Davie et al., 2018; Kurmangaliyev et al., 2020; Li et al., 2017; Özel et al., 2020; Shekhar et al., 2016). These molecules contribute to synaptic specificity in different ways (Canzio & Maniatis, 2019; Duan et al., 2018; Millard et al., 2010; Sanes & Zipursky, 2020; Shen & Bargmann, 2003; S. Xu et al., 2018). Due to the extraordinary complexity of synaptic circuitry in the central nervous system, it has been challenging to identify matching receptor-ligand pairs that specify the choice of synaptic partners (Südhof, 2017).

4.3 Materials and Methods

Connectome analysis

T4/T5 connectomes (Shinomiya et al., 2019, 2022) were downloaded from the neuPrint database (https://neuprint.janelia.org/, dataset: fib19:V1.0, (Clements et al., 2020)) using natverse (0.2.4) package (Bates et al., 2020). Synaptic connectivity data (inputs and outputs) were downloaded for 40 representative T4/T5 neurons from published data (Shinomiya et al., 2022) (five instances of each T4/T5 subtype). For each representative neuron, we summed the total number of synapses between a given instance of T4/T5 and all synaptic partners of the same cell type (e.g. total number of synapses between a single T4a and any Mi1 neurons). This data was plotted as a heatmap of synaptic weights between partner neuron types and individual instances of T4/T5 neurons (Figure 1B). Synaptic weights were averaged across all instances of each T4/T5 subtype to generate a cell type-level connectome graph. The connectome graph was visualized using Cytoscape (Shannon et al., 2003) and igraph (Csardi & Nepusz, 2006) (Figure 1A). Synaptic partners were restricted to cell types that make more than 10 synapses with a single T4/T5 neuron. Only connections with more than 7 synapses have been plotted. Presynaptic inputs were restricted to connections in the medulla and lobula; postsynaptic outputs were restricted to connections in the lobula plate. Reconstructions of representative LLPC2 and LLPC3 neurons in Fig. 4D were visualized in neuPrint.

Transcriptome analysis

Analysis of the transcriptional atlas of the Drosophila visual system Single-cell analysis was performed using Seurat (V4.1.1, (Butler et al., 2018)). All functions were used with default parameters, unless otherwise indicated. In this study, we use single-cell RNA-Seq data from a previously generated comprehensive transcriptional atlas of the developing visual system (Kurmangaliyev et al., 2020). We focus on the main dataset including samples from seven

developmental time points taken every 12 hours from 24 to 96h APF. In the initial version of the visual system atlas (V1.0), 58 transcriptional clusters were matched to known morphological cell types. Two of these clusters were annotated as LLPC1 and LPC1 neurons based on correlation analysis with available bulk reference datasets (Davis et al., 2020). A more detailed evaluation of these clusters revealed further heterogeneity in the LLPC1 cluster. We subsetted and reclustered LLPC1 and LPC1 clusters separately from the rest of the dataset. This analysis was performed after the integration of the main dataset as previously described (Kurmangaliyev et al., 2020). After subsetting, a new set of 1000 highly variable genes was selected, scaled, and used for principal component analysis (functions: FindVariableFeatures, ScaleData, RunPCA). The first nine principal components were used for the generation of tSNE plots and clustering (functions: RunTSNE, FindNeighbors, FindClusters, resolution = 0.1). The analysis revealed five transcriptionally distinct clusters of LPC/LLPC neurons. Clusters were annotated based on the expression of cell-type-specific marker genes (Figure 4.S2). Cell types of LPC/LLPC neurons were renamed in the main dataset of the visual system atlas (V1.1).

Visualization of gene expression patterns: Expression patterns of genes were visualized using average expression levels for each cell type and time point. Averaging was performed for each replicate in non-log space for the original normalized expression values (TP10K, transcripts-per-10,000). For the heatmaps, we used log1p-transformed expression values averaged across replicates (capped at the maximum expression value of 20). For the line plots, we used expression values in linear scale. The list of all IgSF proteins was obtained from FlyXCDB (http://prodata.swmed.edu/FlyXCDB; (Pei et al., 2018)).

Hierarchical clustering of neuronal cell types: Clustering was performed for neuronal cell types from one sample at 48h APF (W1118 sample, replicate B). This analysis was performed based on the original normalized expression values (pre-integration). We selected 1000 highly variable genes and computed average expression levels for each cell type (functions: FindVariableFeatures, AverageExpression). Hierarchical clustering was performed based on

Pearson's correlation coefficients between log-transformed expression profiles of each cell type (distance metric: 1 - Pearson's r; clustering method: ward.D2) (Murtagh & Legendre, 2014). Clustering results were visualized as a dendrogram (function: ape::plot.phylo) (Paradis & Schliep, 2018).

Differential gene expression analysis: We identified differentially expressed genes using Wilcoxon rank-sum test (function: FindMarkers, min.pct = 0.35, pseudocount.use = 0.01, max.cells.per.ident = 1000, fold-change > 2, adjusted p < 0.01). Marker genes common to all LPC/LLPC neurons were identified by comparison of all LPC/LLPC clusters to all other neurons in the atlas; cell-type-specific markers were identified by comparison of individual LPC/LLPC clusters to other LPC/LLPC neurons. Marker genes were identified for all time points and replicates together. Expression patterns of select markers are shown in Figure 4.S2. Differential analysis of T4/T5 and LLPC neurons in Lop3 and Lop4 for Figure. 4.2 was performed at 48h APF (pseudocount.use = 0.01, fold-change > 3, adjusted p < 0.01).

The analysis of connectome and transcriptome data was performed in R (4.1.0). Single-cell RNA-Seq data used in this study and new cell type annotations for the visual system atlas (V1.1) are available through NCBI GEO: GSE156455. The code used in this work is available on request.

Fly Genetics

Flies were kept on standard cornmeal medium at 25°C, 12-hour light/dark cycles. All RNAi experiments were run with UAS-Dcr2.D. Detailed genotypes for all figures and supplementary material are included in a separate spreadsheet (Table 4.S1)

Generation of null alleles using CRISPR For side-II and beat-VI, two protospacer sequences targeting the first coding exon were chosen to create a short deletion leading to a frameshift mutation of the protein sequence. For side-IV, two protospacer sequences spanning the whole gene were chosen to create a total 12.9kb deletion. High score protospacer sequence was chosen

from UCSC Genome Browser crisprTarget table. Oligos were made from selected gRNA sequence and inserted into pU6-2 vector (Ren et al., 2013). gRNA plasmid was injected into vas-Cas9 line (BDSC 51323) via Bestgene Inc. Injected larvae were crossed with balancer lines and F1 progeny was screened for mutation. A mutant allele line was established from this single F1 progeny. sgRNA sequences are listed. Detailed protocols are available upon request.

side-II[13] (side-II null) deleted sequence(44bp):
TCCGGCGGAGGCAGCAGCATGGGTCCTGGCGGAGGAGGATCCGG

side-IV[4-5] deleted sequence(12.9kb):
AACGCGTATTCGCACCCACACACAAGTGAAGTCGGCTCT……..
GGAACTCTCCGGCACTCCGGTATTCCGGAATTCCGTTGCTCCGGTGGTC

beat-VI[4] (beat-VI null) deleted sequence(19bp): AAGGATACGGAGCCGGCCA

MARCM-STaR experiment MARCM-STaR labels cell morphology and presynaptic machinery (Brp) in single homozygous null mutant neurons in otherwise heterozygous backgrounds (Chen et al., 2014). The *side-II*[13] allele was recombined with FRT40 for MARCM (Lee & Luo, 2001). Mitotic recombination was induced at the third instar larva stage with 37°C heat shock for 2-3 min. FLP activated the FRT-flanked stop signal resulting in expression of R recombinase under GAL4 control. R recombinase then excised sequences encoding the stop codon flanked by R-specific-recombination sites (RSRT). This resulted in the insertion of the V5 tag into Brp and due to the 2A site expression of the linked LexA coding sequence. LexA then induces LexAop-myr-tdTOM to label the cell membranes to high cell morphology. (Figure 4.S6, detailed genotype in Table 4.S1). Flies were dissected within two days after eclosion. The brains were visualized by immunofluorescence staining as described below.

The T4/T5 GAL4 driver marks all subtypes. In order to classify the identities of T4/T5 MARCM clones, we used the unique dendritic orientation of T4/T5s (i.e. T4c/T5c, dorsal to ventral; T4d/T5d, ventral to dorsal). Brains were mounted to have confocal image stacks along the dorsal to ventral axis, so that the Z-axis in the final volume corresponds to the D-V axis of the compound eye (i.e. the visual field). Images were analyzed in IMARIS to enable 3D visualization. For each T4/T5 dendrite, orientation was determined by the angle of the primary dendritic branch extending away from the axon shaft (i.e. extension away from the axon) and the position of distal tips of the dendrite. T4c/T5c and T4d/T5d dendrites were oriented in opposite directions.

RNAi-MCFO experiment RNAi was expressed by a GAL4 driver expressed in both LLPC2 and LLPC3. To visual single LLPC2 and LLPC3 neurons we combined MultiColor FlipOut (MCFO) with beat VI knock down ((Nern et al., 2015) Figure 4.4E and 4.4H). Heat shock (8-10 min) was induced in the mid-pupal stage for sparse labeling.

Immunohistochemstry / Immunofluorescence and confocal microscopy

Brains were dissected in ice-cold Schneider's Drosophila Medium (GIBCO #21720-024), and fixed in PBS containing 4% paraformaldehyde (PFA) for 25 min at room temperature (RT). Brains were washed three times with PBST (PBS containing 0.5% Triton X-100), and incubated in blocking solution (PBST containing 10% Normal Goat Serum) for at least 2 hr at RT prior to incubation with antibody. Brains were incubated sequentially with primary and secondary antibodies diluted in blocking solution for 2 days at 4°C, with 3 PBST washes followed by 2 hr incubation at RT in between and afterward. Then brains were mounted with Everbrite mounting media (Biotium #23001) or processed for DPX mounting (see below).

Antibody information Primary antibodies and dilutions used in this study: chicken anti-GFP (1:1000, Abcam #13970), rabbit anti-dsRed (1:200, Clontech#632496), mouse anti-Nc82 (1:40,

Developmental Studies Hybridoma Bank (DSHB) Nc82), chicken anti-V5 (1:300, Fortis Life Sciences #A190-118A), rabbit anti-HA (1:200, Cell Signaling Technology #3724), guinea pig anti-Pdm3 (1:20, a gift from John Carlson), mouse anti-Br (1:20, Developmental Studies Hybridoma Bank (DSHB) 25E9.D7). Secondary antibodies and dilutions used in this study: goat anti-chicken Alexa Fluor 488 (AF488) (1:1000, Invitrogen #A11039), goat anti-rabbit AF568 (Invitrogen #A11011, 1:200), goat anti-guinea pig 568 (1:500, ThermoFisher #A11075), goat anti-mouse AF568 (1:500, ThermoFisher #A11031) goat anti-mouse AF647 (1:500, ThermoFisher #A21235).

Tissue clearing and DPX mounting Antibody stained brains were mounted with DPX following Janelia FlyLight protocol (https://www.janelia.org/project-team/flylight/protocols). Briefly, after secondary antibody wash as described above, brains were post-fixed with 4% PFA for at least 3 hr in RT. Brains were washed with PBS and mounted on polylysine-L coated coverslip and sequentially dehydrated in increasing concentration of ethanol (50%, 75%, 90%, 100%, 100%, 100%, 10-min each). Dehydrated brains on coverslip were incubated in Xylene (Fisher Scientific, X5-500) (5min x 3) for tissue clearing. Then the coverslip was embedded in DPX (Electron Microscopy Sciences, #13510) and cured in the chemical hood for more than 2 days before imaging. DPX mounting images are noted in the genotype table (Table 4.S1)

Confocal microscopy Immunofluorescence images were acquired using Zeiss LSM 880 confocal microscope with Zen digital imaging software. Optical sections or maximum intensity projections were level-adjusted, cropped and exported for presentation using ImageJ software (Fiji) or IMARIS 9 (Oxford Instruments). Reported expression patterns were reproducible across three or more biological samples.

Data and materials availability

Connectome and transcriptome datasets used in this study are available through Neuprint (neuprint.janelia.org) and NCBI GEO (GSE156455). Newly generated mutant alleles and split-Gal4 lines are available upon request.

4.4 Results

4.4.1. Matching the connectome to the developmental transcriptome

To overcome the complexity of synaptic circuitry, we focused on the motion detection circuit and, in particular, closely related neurons which choose between similar synaptic targets (Borst et al., 2019; Shinomiya et al., 2022). Dense EM reconstructions identified most of the synaptic connections in this circuit (**Figures 4.1A-B**). A set of eight closely related subtypes of T4/T5 neurons lie at its center. Each subtype is defined by a combination of one of two patterns of dendritic inputs and one of four patterns of axonal outputs (**Figures 4.1A** and **4.1F**). T4 and T5 neurons arborize their dendrites in the medulla and lobula, respectively. Each of these groups is further subdivided into four subtypes (a/b/c/d) based on their axon terminals in four synaptic layers of the lobula plate (Lop1/2/3/4). Each pair of T4 and T5 neurons which project axons to the same layer respond optimally to motion in one cardinal direction; T4 neurons respond to the movement of bright edges (ON pathway) and T5 neurons respond to dark edges (OFF pathway) (Maisak et al., 2013). T4 and T5 axons terminating in the same layer converge onto the same postsynaptic partners (Shinomiya et al., 2022). Furthermore, some of the postsynaptic neurons in different layers are also closely related cell types (see below). In this way, information from the ON (T4) and OFF (T5) pathways corresponding to each cardinal direction converge onto four parallel synaptic pathways (**Figure 4.1B**). We hypothesize that T4 and T5 subtypes, which form synapses with the same set of postsynaptic neurons in each layer of the lobula plate, do so through the same molecular mechanisms (Kurmangaliyev et al., 2019).

To identify these molecules, we integrated the synaptic connectome and the transcriptome of developing neurons in the Drosophila visual system (**Figures 4.1A-C**). We previously generated a comprehensive transcriptional atlas of the developing visual system using single-cell RNA sequencing (Kurmangaliyev et al., 2020). This atlas covers more than 160 neuronal populations at seven developmental time points; ~100 of them were matched to cell types in the

connectome (Davis et al.,2020; A.N., Y.Z.K, S.L.Z., *in preparation*). This includes transcriptional profiles of all T4/T5 subtypes and 17 of their synaptic partners (**Figure 4.1C**). We focus on five types of morphologically similar postsynatpic partners (Davis et al., 2020; Shinomiya et al., 2022): two LPC (Lobula Plate Columnar) and three LLPC (Lobula-Lobula Plate Columnar) neurons (**Figures 4.1D** and 4.**S1**). Each of these neuron types receives its major input from one pair of T4 and T5 subtypes (in one layer of the lobula plate) (**Figures 4.1B** and 4.**1G**). In our initial version of the transcriptional atlas these five cell types were not resolved. A more detailed analysis revealed distinct transcriptional clusters for each of them, which were validated by *in vivo* expression patterns of marker genes (**Figures 4.S1** and **4.S2**) and mRNA profiling of purified cell types (A.N., Y.Z.K, S.L.Z., *in preparation*). Hierarchical clustering of transcriptomes of all neuronal populations in the visual system confirmed that the eight T4/T5 subtypes were closely related. Similarly, four LPC/LLPC types (except LPC1) were also closely related to each other (**Figure 4.1E**). Taken together, T4/T5 subtypes and LPC/LLPC types assemble into parallel synaptic pathways comprising homologous pairs of pre- and postsynaptic partners (**Figures 4.1H and 4.1I**).

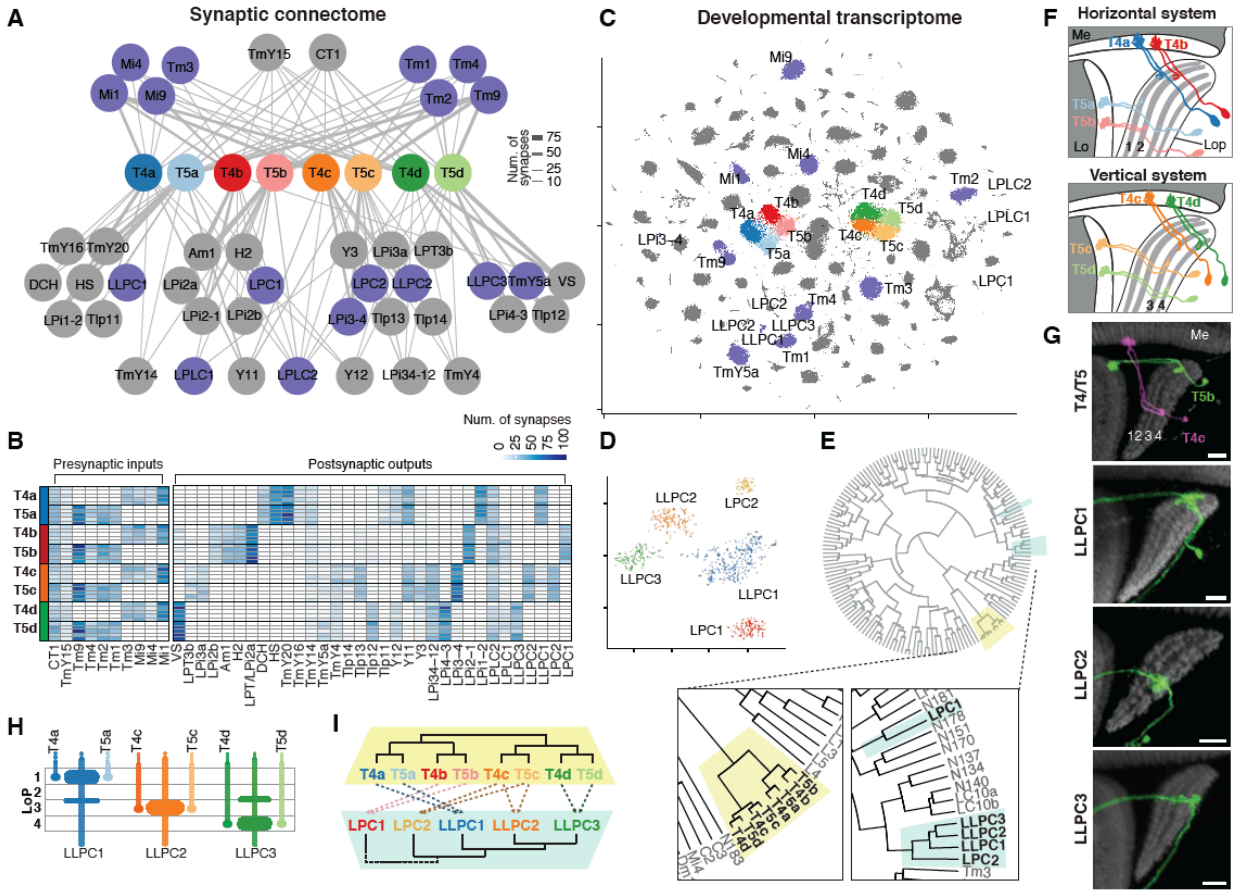


Figure 4.1. Coupled transcriptome-connectome map of the Drosophila motion detection circuit. (A-B) The connectome of T4/T5 subtypes from (Shinomiya et al., 2019, 2022). (A) Cell-type level connectome graph. Presynaptic inputs (top) and postsynaptic outputs (bottom) of T4/T5s are grouped and averaged by cell types. Purple, cell types with known transcriptomes. (B) Connectomes of individual T4/T5s, as an adjacency matrix. Five instances for each T4/T5 subtype (rows). Synaptic partners are grouped by cell type (columns). Each pair of T4 and T5 subtypes converge onto the same set of postsynaptic targets. (C) tSNE of the Drosophila visual system atlas (Kurmangaliyev et al., 2020). T4/T5s and their synaptic partners are labeled as in A. (D) LPC/LLPC clusters. See also Figure 4.S1. (E) Hierarchical clustering of transcriptomes in the visual system atlas; T4/T5 and LPC/LLPC neurons are shown in the zoom-in. (F) Morphology of eight T4/T5 subtypes. Me, medulla; Lo, lobula; LoP, lobula plate. (G) Sparsely labeled T4/T5s and LLPC neurons. Neuropil marker (grey), brp. Scale bars, 10 μm. (H) Layer targeting of T4/T5 and LLPC neurons in the LoP. For LPCs see Figure 4.S1 (I) Closely related pairs of T4 and T5 subtypes converge onto the same postsynaptic targets.

4.4.2. Beat/Side IgSF CAM expression defines synaptic specificity

To understand how neurons choose their synaptic partners, we focused on synapses between the most closely related T4/T5 pairs and their closely related postsynaptic targets. T4c and T5c form synapses with LLPC2, and T4d and T5d form synapses with LLPC3 (**Figure 4.2A**). Each of these pairs of T4/T5 subtypes converge onto transcriptional programs that correlate with the specificity of their axonal outputs (Kurmangaliyev et al., 2019). Virtually all differentially expressed genes (DEGs) between T4c and T4d are the same as between T5c and T5d (**Figure 4.2B**). There were only nine such DEGs, and of these, seven were cell surface proteins (**Figure 4.2C**). Similarly, a small number of cell surface proteins were differentially expressed between LLPC2 and LLPC3 neurons (**Figure 4.2D**). We hypothesize that matched pairs of cell recognition molecules specific to each pair of synaptic partners regulates this binary choice in synaptic specificity.

Although each T4/T5 subtype and LLPC types express many CAMs (**Figure 4.2E**), only two pairs of interacting CAMs correlated with synaptic specificities of these two sets of synaptic partners (**Figure 4.2F**). Both pairs belong to the Side and Beat families of IgSF proteins, which form a heterophilic protein interaction network (Li et al., 2017; Özkan et al., 2013) (**Figure 4.2I**). Founding members of these families (Side and Beat-Ia) were identified in genetic screens as regulators of motor axon guidance in the Drosophila embryo (Fambrough & Goodman, 1996; Siebert et al., 2009; Sink et al., 2001). Functions for other paralogs have not been described. As each neuron type in the developing visual system expresses a unique combination of 14 Beat and eight Side proteins during development, these proteins may contribute to synaptic specificity more broadly (**Figure 4.S3**).

The top differentially expressed genes between T4/T5 subtypes are side-IV (specific to T4c/T5c) and side-II (specific to T4d/T5d). LLPC2 and LLPC3 neurons express interacting Beats in a matching fashion; beat-IIa and beat-IIb are specific to LLPC2, and beat-VI is specific to LLPC3

(**Figure 4.2G**). At least two other major synaptic targets of T4c/T5c (LPC2 and LPi3-4) express beat-IIa/IIb but not beat-VI, and the main synaptic target of T4d/T5d (VS) expresses Beat-VI but not Beat-IIb (**Figure 4.2G and 4.2H**). In this way, the matching expression of two pairs of interacting IgSF CAMs correlates with synaptic specificity in this circuitry (**Figure 4.2J**).

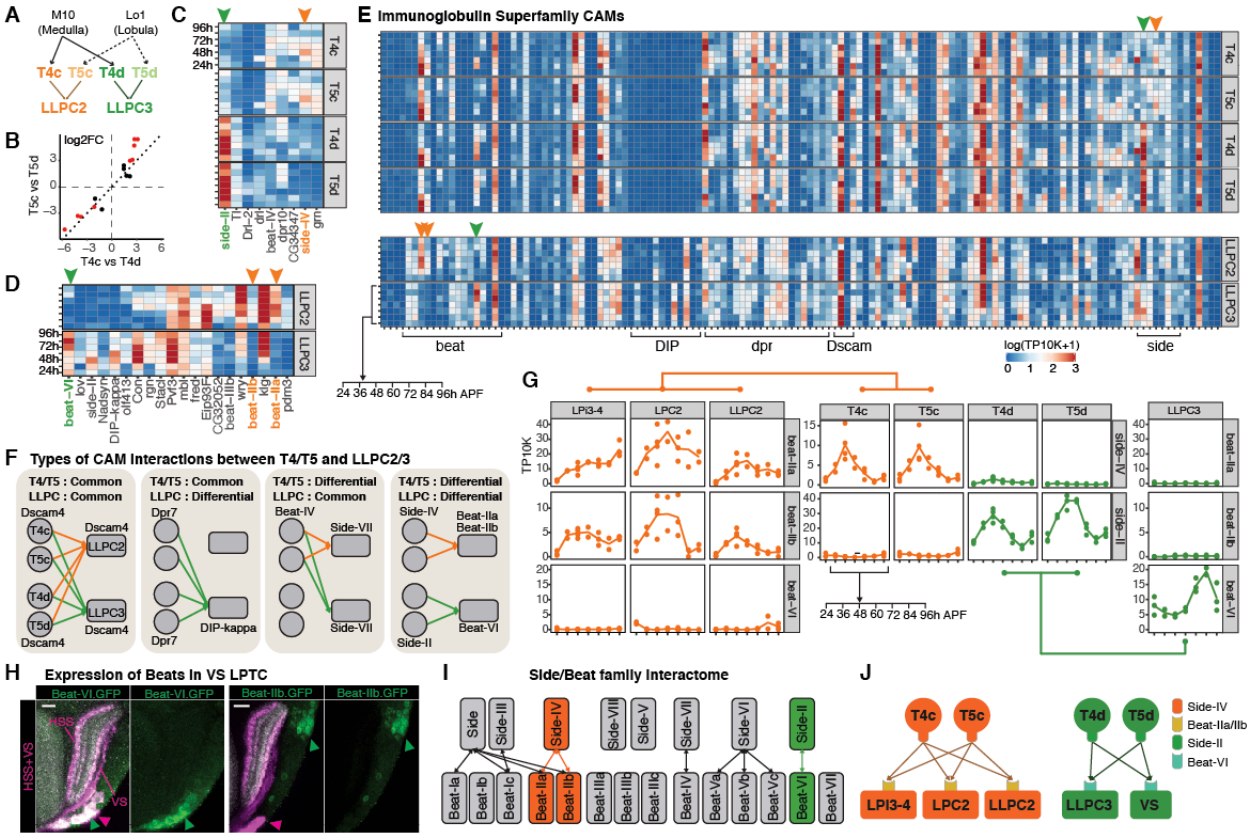


Figure 4.2. Matching receptor-ligand pairs correlate with synaptic specificity.
(A) Connectome of T4/T5-LLPC2/3 circuit. (B to D) Differentially expressed genes (DEGs) between T4/T5 and LLPC subtypes. See Methods for details and thresholds. (B) Log-fold-change for DEGs between T4c and T4d (x-axis), and T5c and T5d (y-axis). (C) Heatmap of expression patterns of DEGs significant in both comparisons in B (red). (D) DEGs between LLPC2 and LLPC3 (E) All annotated IgSF CAMs. Data is shown for 7 time points after pupal formation (APF). (F) Examples of matching CAM interactions between T4/T5 and LLPC. Only two pairs of interactors correlate with connectome: Side-IV::Beat-IIa/b and Side-II::Beat-VI (orange and green arrows in C to E). (G) Line plots with expression patterns of these Beat/Side proteins in the T4/T5-LLPC2/3 circuit, and other major targets with known transcriptomes. Dots are replicates. (H) *In vivo* expression of Beat-IIb and Beat-VI in VS neurons (adult), the main target of T4d/T5d (Figure 1D). Neuropil marker (grey), brp. Scale bars, 10 μm. (I) Interactome of Beat/Side families from (Li et al., 2017). (J) Matching receptor-ligand pairs between T4/T5s and their targets in Lop3 (orange) and Lop4 (green).

4.4.3. Side-II/Beat-VI is required for synaptic layer separation

We sought to assess the roles of Beat/Side interactions in the wiring of T4/T5 axons. In wild-type, these axon terminals form four layers in the lobula plate (**Figure 4.3A**). In homozygous *side-II*^{null} mutant animals, T4/T5 axon terminals formed a single fused Lop3/4 layer (**Figure 4.3B**). Lop1 and Lop2 were normal. Beat-VI is a high-affinity binding partner of Side-II. Homozygous *beat-VI*^{null} mutants phenocopy *side-II*^{null} mutants. These results were confirmed with an insertion and a deletion disrupting these genes (**Figure 4.S4A** and 4.**S4B**). The layer fusion phenotype resembles the earlier stages of development (**Figure 4.3A**).

Next, we tested whether Side-II was required in T4/T5 neurons using RNA interference (RNAi) (**Figure 4.3C-D**). Removing side-II specifically from T4/T5s phenocopied *side-II*^{null} mutants. These results were confirmed using an independent side-II RNAi line (**Figure 4.S4A**). Removing beat-VI from T4/T5s did not result in the fusion of Lop3/4, whereas a pan-neuronal RNAi of beat-VI did. Removal of side-II from T4/T5s also disrupted dendritic morphologies of their main postsynaptic partners (**Figure 4.3E**). Thus, Side-II and Beat-VI is a receptor-ligand pair that regulates separation of neuronal processes into adjacent synaptic layers of the lobula plate.

Expression patterns of Side-IV and Beat-IIa/IIb in Lop3 suggested a similar role for this receptor-ligand pair in lobula plate development. However, in homozygous *side-IV*^{null} mutant animals, we did not observe defects in the lamination of T4/T5 axon terminals (**Figure 4.S4C**). This suggests that these proteins have a different function, or other recognition molecules may act to compensate for their loss.

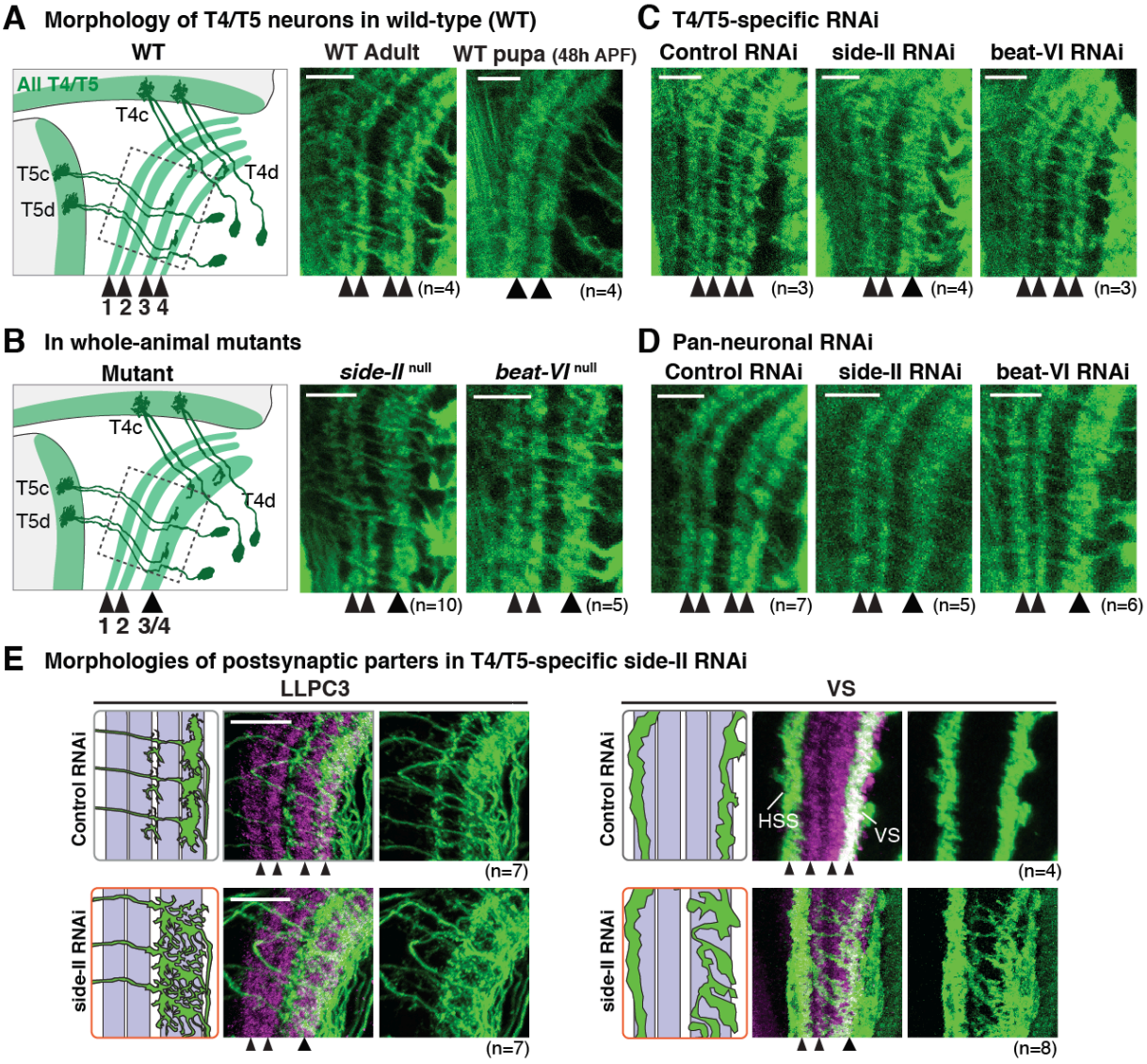


Figure 4.3. Side-II::Beat-VI segregates synaptic partners into layers.
(A to D) Morphology of T4/T5s in the wild-type and mutant backgrounds. In immunofluorescence images, all T4/T5 neurons are labeled. (A) During development (*right*), axon terminals first laminate into two broad domains (Lop1/2 and Lop3/4). In adults (*middle*), these layers further separate into four layers (Lop1/2/3/4). (B) In *side-II*^{null} and *beat-VI*^{null} mutants, axons of T4c/T5c and T4d/T5d form a single fused Lop3/4 layer. (C) T4/T5-specific RNAi of Side-II (but not Beat-VI) also results in fused Lop3/4. (D) Pan-neuronal RNAi of both Side-II and Beat-VI results in fused Lop3/4. (E) Morphology of the main postsynaptic partners of T4/T5s (green) upon removal of Side-II from T4/T5s (via RNAi). In controls, LLPC3 and VS arborize in Lop4. In side-II RNAi, they span fused Lop3/4 layer. Neuropil marker (magenta), brp. Scale bars, 10 μm.

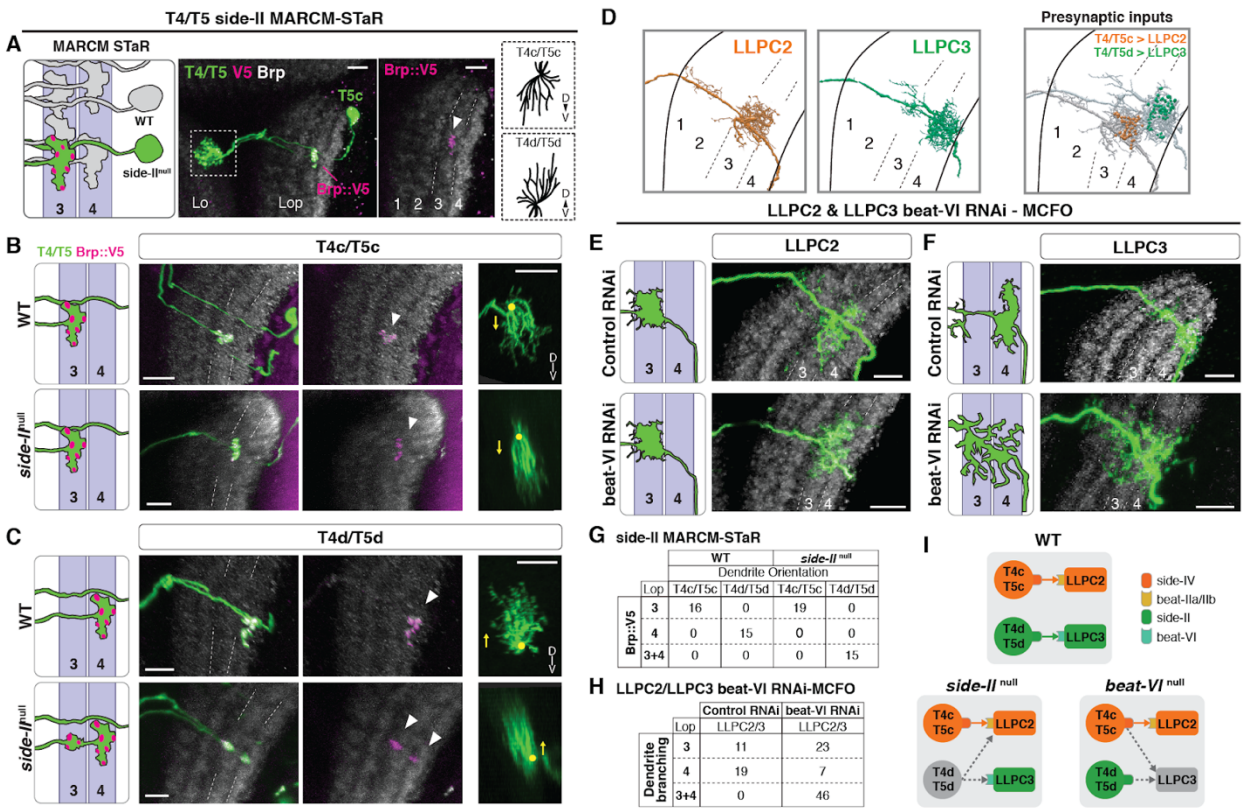
4.4.4. Side-II and Beat-VI control synaptic specificity.

Next, we assessed the role of Side-II/Beat-VI interactions at the level of single cells. For T4/T5 neurons, we visualized the morphologies and presynaptic sites of sparsely distributed homozygous *side-II*^{null} mutant neurons where most, if not all, other neurons are wild-type (heterozygous; see Methods, **Figure 4.4A-C**). Wild-type T4c/T5c and T4d/T5d have presynaptic sites in Lop3 and Lop4, respectively. Mutant T4d/T5d axons terminate in Lop4, however, their presynaptic sites span both layers. The penetrance of this phenotype was complete (15 of 15 single mutant neurons). Mutant T4c/T5c were similar to wild-type. The simplest interpretation is that mutant T4d/T5d neurons form ectopic synapses with inappropriate synaptic partners in Lop3.

We next removed Beat-VI from postsynaptic partners, LLPC2 and LLPC3, using RNAi and visualized individual mutant neurons (see Methods, **Figure 4.4E-G**). In controls, we detected wild-type LLPC2 (dendrites in Lop3) and LLPC3 (dendrites in Lop4). In beat-VI RNAi, we detected wild-type LLPC2 and many abnormal neurons spanning both Lop3 and Lop4. Fewer wild-type LLPC3s than expected were observed (**Figure 4.4H**, Fisher's exact test, $p = 0.004$). As beat-VI is specific to LLPC3, we conclude that abnormal neurons are mutant LLPC3s. This phenotype was confirmed using an independent beat-VI RNAi line (**Figure 4.S5**). The single-cell phenotypes are consistent with each other. That is, presynaptic sites of Side-II mutant T4d/T5d and postsynaptic dendrites of Beat-VI-deficient LLPC3 are no longer restricted to Lop4 and accumulate in Lop3.

In summary, Side-II is expressed in T4d/T5d axons and Beat-VI in dendrites of their postsynaptic partners. Side-II and Beat-VI bind to each other. Removal of this receptor-ligand pair leads to defects in synaptic circuitry in Lop4. These proteins may control several steps of circuit assembly. At an early stage, they promote adhesion between axons and dendrites leading to their segregation into layers. At a later time, they may directly specify connections between synaptic

partners. Regardless of mechanistic details, Side-II and Beat-VI interactions play a critical role in matching synaptic partners.



G	WT		side-II null	
Dendrite Orientation				
Lop	T4c/T5c	T4d/T5d	T4c/T5c	T4d/T5d
Brp::V5 3	16	0	19	0
Brp::V5 4	0	15	0	0
Brp::V5 3+4	0	0	0	15

H	Control RNAi	beat-VI RNAi
Lop	LLPC2/3	LLPC2/3
Dendrite branching 3	11	23
Dendrite branching 4	19	7
Dendrite branching 3+4	0	46

Figure 4.4. Side-II::Beat-VI determines the choice between alternative synaptic targets
(A to C) MARCM-STaR. Mosaics with homozygous mutant (*side-II*^{null}) T4/T5 neurons (green) with presynaptic marker Brp-V5 (magenta), and wild-type controls (WT). (left) Schematics of T4/T5 axon terminals in Lop3 and Lop4. (right) Dendrite orientation discriminates between T4c/T5c and T4d/T5d (D: dorsal, V: ventral, based on the visual field coordinates). (B) WT and *side-II*^{null} T4c/T5c form synapses in Lop3. (C) WT T4d/T5d form synapses in Lop4; *side-II*^{null} T4d/T5d form synapses in Lop3 and Lop4. (D) LLPC2 and LLPC3 morphologies and T4/T5 inputs from EM reconstruction (Shinomiya et al., 2022). (E and F) beat-VI RNAi in LLPC2 and LLPC3. Single neurons were visualized using MCFO. In controls, both LLPC2 and LLPC3 dendrites were wild-type (as in D). In beat-VI RNAi, we detected wild-type LLPC2 and fewer LLPC3 than expected. We observed a large number of abnormal neurons spanning both Lop3 and Lop4. As beat-VI is specific to LLPC3, we conclude that abnormal neurons are LLPC3. (G) Quantification of the labeled neurons from B and C. (H) Quantifications of labeled neurons from E and F. (I) Model for circuit rewiring upon removal of Side-II and Beat-VI. Neuropil marker (gray), brp. Scale bars, 10 μm.

4.5. Discussion

How do neurons choose their synaptic partners? We have shown that any pair of neurons typically express dozens of CAMs that can promote molecular interactions between them (Kurmangaliyev et al., 2020). The choice between alternative targets can be determined by a differential match in a single receptor-ligand pair. If we remove these interactions alternative targets become molecularly less distinct. For instance, in the absence of Side-II, LLPC2 and LLPC3 appear more similar to T4d/T5d than in wild-type (**Figure 4.4I**). As a consequence, T4d/T5d forms synapses in Lop3 in addition to Lop4. Our findings support the idea that a hierarchy of wiring decisions sequentially restricts the pool of possible targets (Duan et al., 2018; Kolodkin & Hiesinger, 2016; Zhang et al., 2022) and synaptic specificity is determined by the relative preference between alternative partners (C. Xu et al., 2019).

This work expands the diverse repertoire of families of IgSF proteins which contribute to synaptic specificity in the *Drosophila* brain (Millard et al., 2010; S. Xu et al., 2018). Each family forms complex receptor-ligand networks including homophilic (e.g. Dscams (Wojtowicz et al., 2007) thousands of isoforms) and heterophilic interactions (e.g. DIP/Dpr (Cosmanescu et al., 2018; Özkan et al., 2013) and Side/Beat (Li et al., 2017; Özkan et al., 2013; Siebert et al., 2009) comprising 50+ interacting pairs). These proteins are expressed in highly dynamic and cell-type-specific ways, and with other cell surface proteins endow each neuron with a unique cell surface protein composition (Kurmangaliyev et al., 2020; Sperry, 1963). The logic in wiring the mammalian brain may be similar, with an expanded cadherin superfamily largely taking the place of IgSF diversity (Sanes & Zipursky, 2020).

Coupled transcriptome-connectome maps provide a description of gene expression patterns for both sides of synaptic connections. These maps can be correlated with binding specificities of cell surface proteins to chart possible molecular interactions between neurons (Verschueren et al., 2020; Wojtowicz et al., 2020). As more connectomes (Abbott et al., 2020;

Motta et al., 2019) and developmental transcriptomes (Zeng, 2022) become available, a similar approach may prove fruitful in uncovering matched receptor-ligand pairs regulating synaptic specificity in the mammalian brain.

Acknowledgments

We thank Kazunori Shinomiya for sharing connectome data prior to publication. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) and Vienna Drosophila Resource Center were used in this study. We also thank G. M. Rubin, H. J. Bellen, and Qi Xiao for the reagents and the Janelia FlyLight Project Team for some images. We thank members of the Zipursky lab for the critical discussion of the manuscript. Funding: S.L.Z is an investigator of the Howard Hughes Medical Institute. Authors contributions: J.Y., A.N., S.L.Z and Y.Z.K., conceptualization; J.Y., M.D., P.M, A.N., S.A.L, Y.Z.K., investigation and formal analysis; J.Y., S.L.Z and Y.Z.K., writing. Competing interests: No competing interests declared. Data and materials availability: Connectome and transcriptome datasets used in this study are available through Neuprint (neuprint.janelia.org) and NCBI GEO (GSE156455). Newly generated mutant alleles and split-Gal4 lines are available upon request.

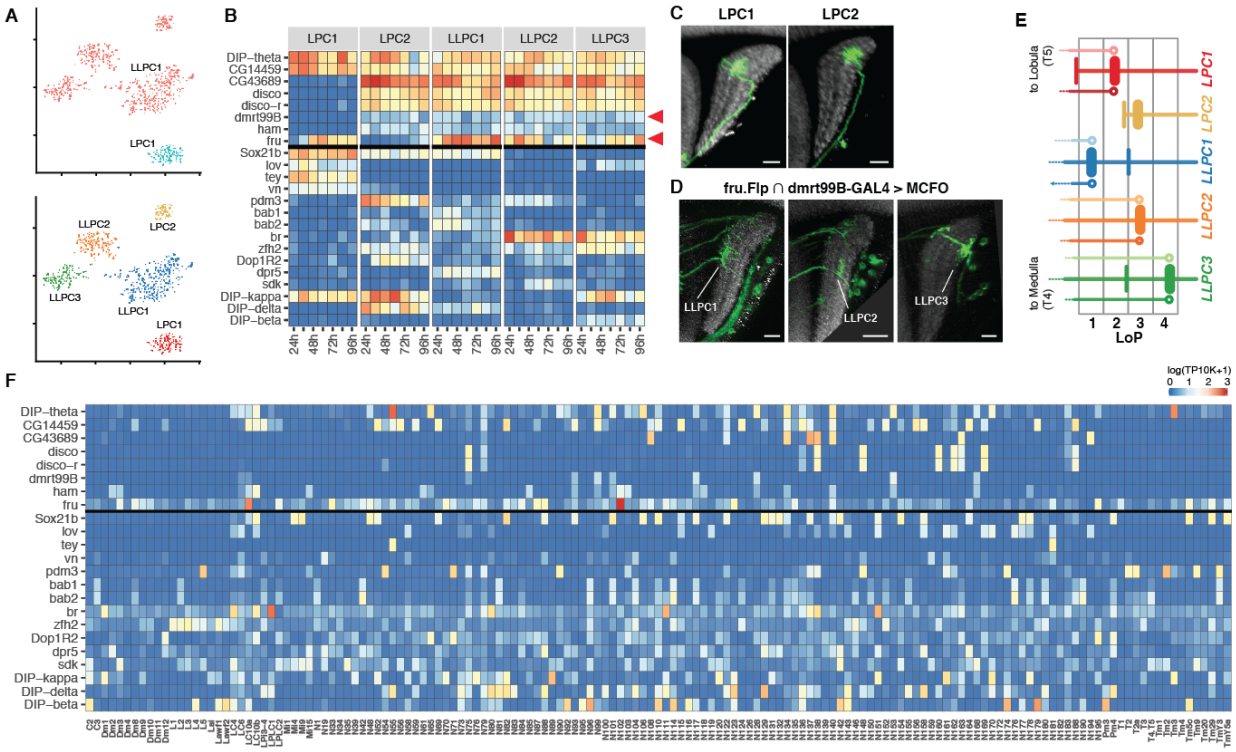


Figure 4.S1. Transcriptomes of LPC/LLPC neurons.

(A) tSNE plots of LPC/LLPC neurons. Cells are color-coded according to annotations in the atlas V1.0 (top) and V1.1 (bottom). In the first version of the atlas, all LLPC and LPC2 neurons were clustered together (as LLPC1). (B) Expression patterns of common and cell-type-specific markers of LPC/LLPC neurons. (C) Sparsely visualized LPC1 and LPC2 neurons. (D) The intersection of transcription factors *fru* and *dmrt99B* visualized by MCFO. Expression patterns of these genes are shown by arrows in B. As predicted, this intersection captures LLPC neurons. (E) Layer targeting of T4/T5 and LPC/LLPC neurons in the LoP. (F) Expression patterns of marker genes from C in all other neuronal clusters in the visual system atlas [14]. Neuropil marker (gray), brp. Scale bars, 10 μm.

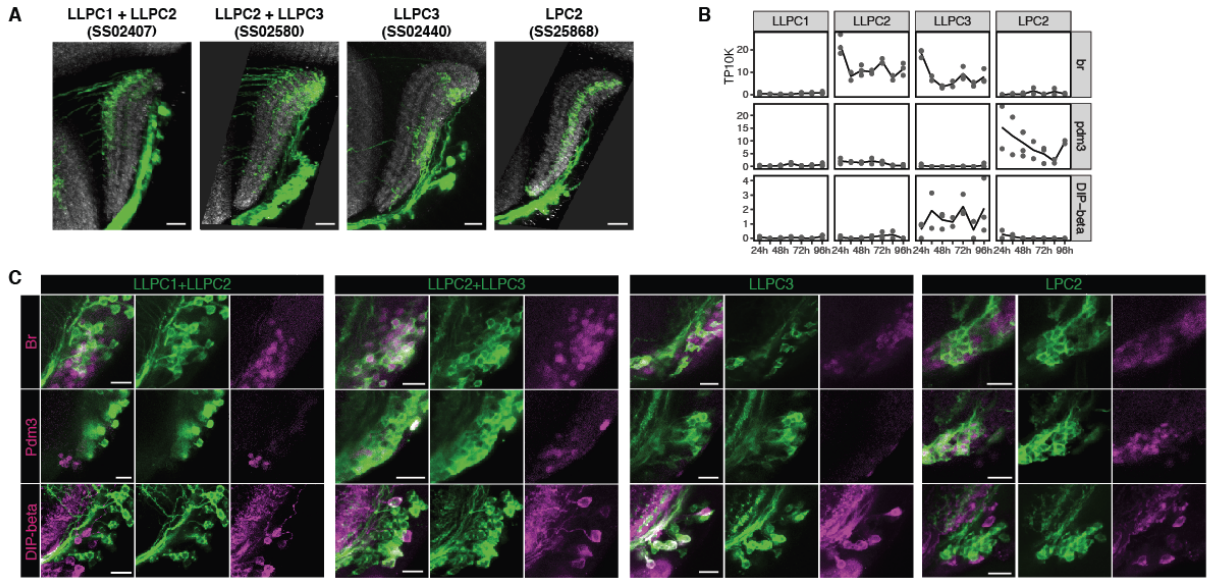


Figure 4.S2. *In vivo* expression patterns of cell-type-specific marker genes of LPC/LLPC neurons.

(A) Expression patterns of Gal4 drivers targeting subsets of LPC/LLPC types (green). (B) Expression patterns of select cell-type-specific marker genes in transcriptomes. (C) *In vivo* expression of marker genes from B (magenta), colocalized with LPC/LLPC cell bodies visualized using Gal4 drivers from A (green). Expression of *br* and *pdm3* is visualized using immunostaining for these proteins. Expression of DIP-beta is visualized using DIP-beta-lexA. Neuropil marker (gray), brp. Scale bars, 10 μm.

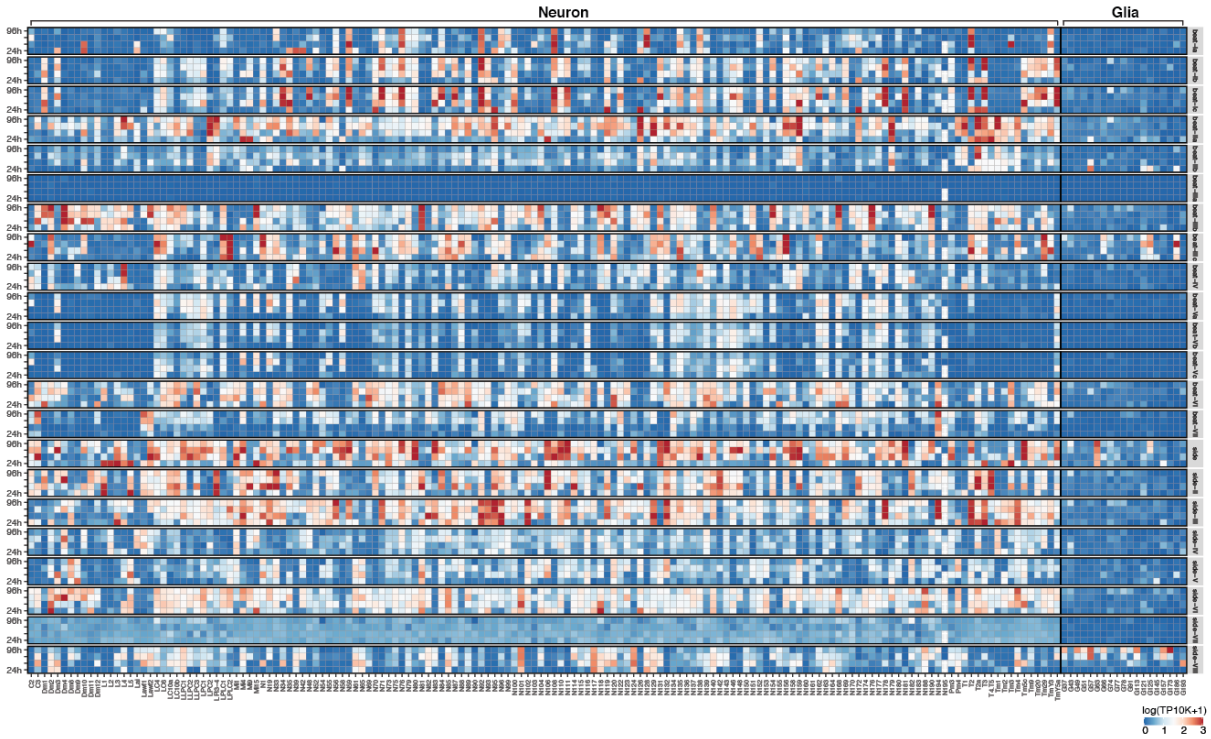


Figure 4.S3. Expression patterns of Beat/Side families of proteins in the Drosophila visual system.

Expression patterns are shown for all neuronal (left) and glial (right) clusters from the Drosophila visual system atlas (Kurmangaliyev et al., 2020). Data is shown for 4 time points (24h, 48h, 72h, and 96h APF). Lack of beat-IIIa expression could be due to technical reasons (e.g. misannotation of transcripts) or it is not expressed in the brain.

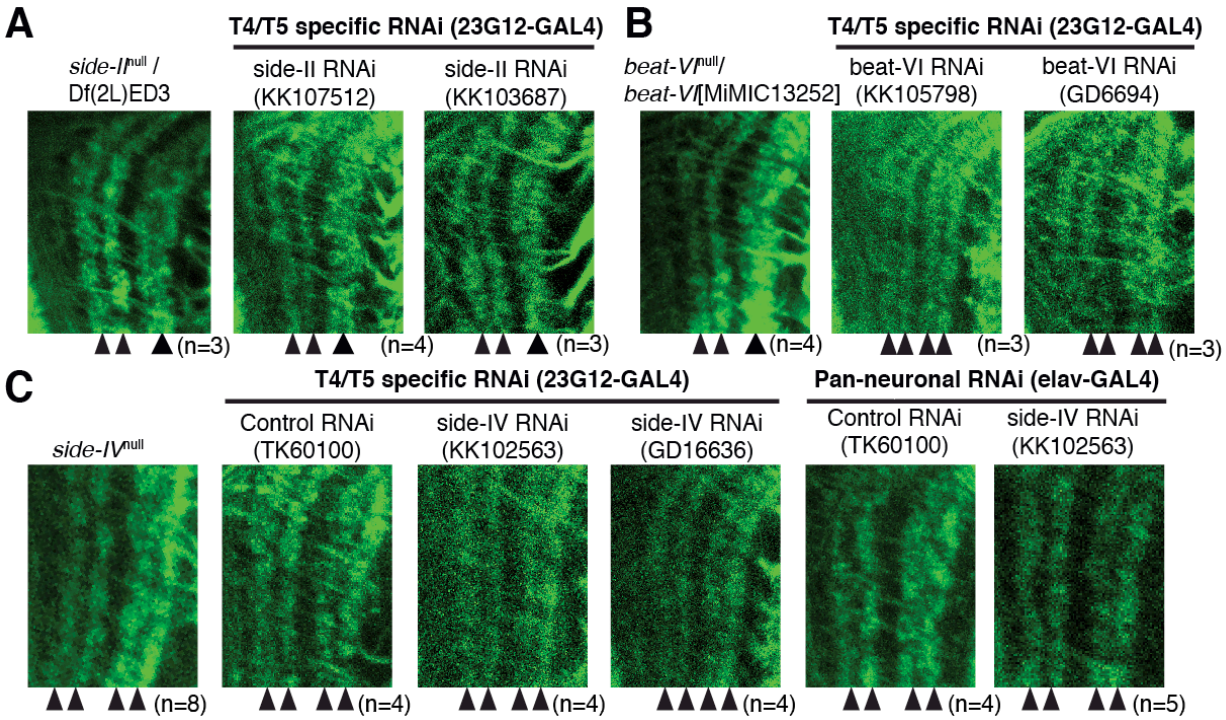


Figure 4.S4. Side/Beat mutants.

(A to C) Morphology of T4/T5 neurons in various mutant backgrounds, visualized as in Figure. 4.3. (A) *side-II*^{null} allele over deficiency and two independent side-II RNAi are indistinguishable from homozygous *side-II*^{null} animals (KK107512 is also shown in Figure 4.3). (B) *beat-VI*^{null} allele over a MiMIC insertion into a beat-VI gene phenocopies homozygous *beat-VI*^{null} animals. Two independent beat-VI RNAi lines expressed in T4/T5s are indistinguishable from controls. By contrast, the same RNAi lines expressed in all neurons phenocopies homozygous *beat-VI*^{null} animals (KK105798 is also shown in Figure 4.3). (C) In *side-IV*^{null} and side-IV RNAi, T4/T5 axons form four layers in the LoP as in wild-type.

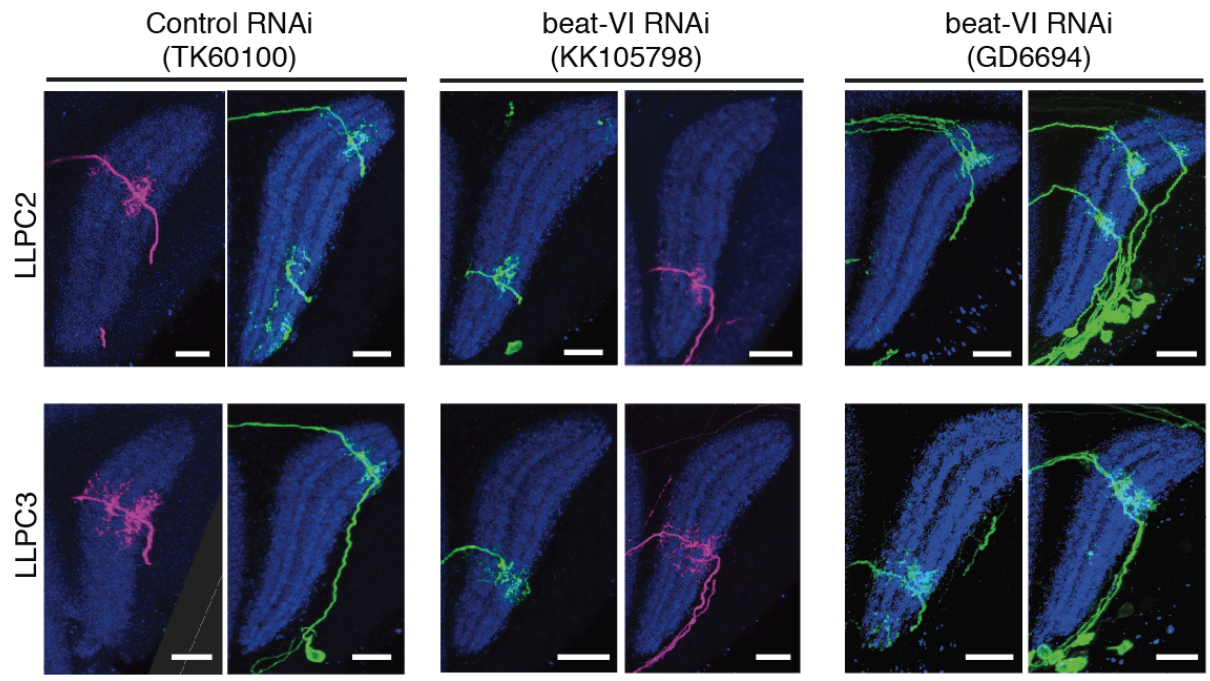


Figure 4.S5. LLPC2 and LLPC3 beat-VI MCFO-RNAi
Additional examples of sparsely labeled LLPC2 and LLPC3 neurons coupled with beat-VI RNAi (same experiment as in Figure 4.4E-G). Data is shown for two independent beat-VI RNAi (KK105798 is used in Figure 4.4). Neuropil marker (blue), brp. Scale bars, 10 μm.

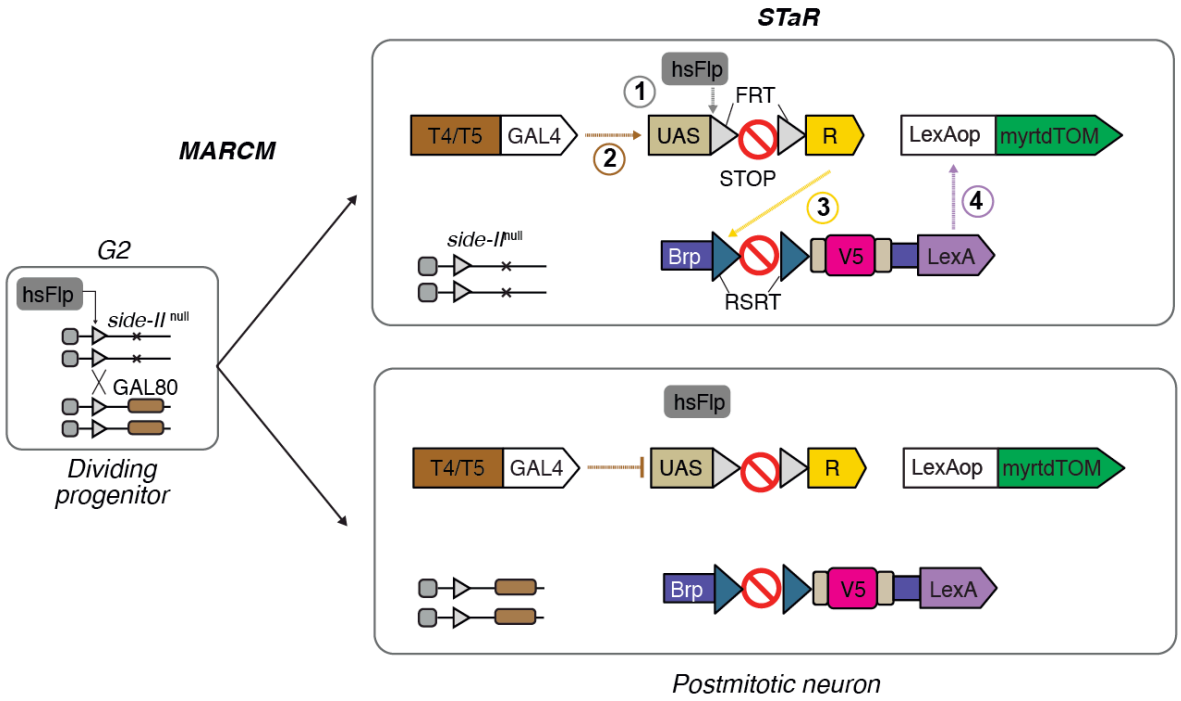


Figure 4.S6. The MARCM-STaR methods.
MARCM-STaR labels cell morphology and the presynaptic active zone protein Brp in single sparsely distributed null mutant neurons in otherwise wild-type backgrounds. Heatshock Flp induces mitotic recombination (MARCM) to generate homozygous null mutant clones. In the post mitotic mutant neuron where T4/T5-GAL4 is active, heatshock Flp excises the stop signal by mediating recombination between the flanking FRT recombination sites. This allows for expression of R recombinase. R recombinase, in turn, excises the stop cassette in the Brp gene by mediating recombination between RSRT sites. This recombination event inserts the V5 epitope tag into Brp and through T2A sequence (not shown) allows for co-expression of LexA; LexA then drives expression of myrtdTOM.

4.6 References

Abbott, L. F., Bock, D. D., Callaway, E. M., Denk, W., Dulac, C., Fairhall, A. L., Fiete, I., Harris, K. M., Helmstaedter, M., Jain, V., Kasthuri, N., LeCun, Y., Lichtman, J. W., Littlewood, P. B., Luo, L., Maunsell, J. H. R., Reid, R. C., Rosen, B. R., Rubin, G. M., … Essen, D. C. V. (2020). The Mind of a Mouse. *Cell*, *182*(6), 1372–1376. https://doi.org/10.1016/j.cell.2020.08.010

Bates, A. S., Manton, J. D., Jagannathan, S. R., Costa, M., Schlegel, P., Rohlfing, T. & Jefferis, G. S. (2020). The natverse, a versatile toolbox for combining and analysing neuroanatomical data. *ELife*, *9*, e53350. https://doi.org/10.7554/elife.53350

Borst, A., Haag, J. & Mauss, A. S. (2019). How fly neurons compute the direction of visual motion. *Journal of Comparative Physiology A*, *206*(2), 109–124. https://doi.org/10.1007/s00359-019-01375-9

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, *36*(5), 411–420. https://doi.org/10.1038/nbt.4096

Canzio, D. & Maniatis, T. (2019). The generation of a protocadherin cell-surface recognition code for neural circuit assembly. *Current Opinion in Neurobiology*, *59*, 213–220. https://doi.org/10.1016/j.conb.2019.10.001

Chen, Y., Akin, O., Nern, A., Tsui, C. Y. K., Pecot, M. Y. & Zipursky, S. L. (2014). Cell-type-specific labeling of synapses in vivo through synaptic tagging with recombination. *Neuron*, *81*(2), 280–293. https://doi.org/10.1016/j.neuron.2013.12.021

Clements, J., Dolafi, T., Umayam, L., Neubarth, N. L., Berg, S., Scheffer, L. K. & Plaza, S. M. (2020). neuPrint: Analysis Tools for EM Connectomics. *BioRxiv*, 2020.01.16.909465. https://doi.org/10.1101/2020.01.16.909465

Cosmanescu, F., Katsamba, P. S., Sergeeva, A. P., Ahlsen, G., Patel, S. D., Brewer, J. J., Tan, L., Xu, S., Xiao, Q., Nagarkar-Jaiswal, S., Nern, A., Bellen, H. J., Zipursky, S. L., Honig, B. & Shapiro, L. (2018). Neuron-Subtype-Specific Expression, Interaction Affinities, and Specificity Determinants of DIP/Dpr Cell Recognition Proteins. Neuron, *100*(6), 1385-1400.e6. https://doi.org/10.1016/j.neuron.2018.10.046

Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695. https://igraph.org

Davie, K., Janssens, J., Koldere, D., Waegeneer, M. D., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., González-Blas, C. B., Poovathingal, S., Hulselmans, G., Spanier, K. I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., … Aerts, S. (2018). A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. Cell, *174*(4), 982-

*99*8.e20. https://doi.org/10.1016/j.cell.2018.05.057

Davis, F. P., Nern, A., Picard, S., Reiser, M. B., Rubin, G. M., Eddy, S. R. & Henry, G. L. (2020). A genetic, genomic, and computational resource for exploring neural circuit function. ELife, 9, e50901. https://doi.org/10.7554/elife.50901

Dorkenwald, S., Turner, N. L., Macrina, T., Lee, K., Lu, R., Wu, J., Bodor, A. L., Bleckert, A. A., Brittain, D., Kemnitz, N., Silversmith, W. M., Ih, D., Zung, J., Zlateski, A., Tartavull, I., Yu, S.-C., Popovych, S., Wong, W., Castro, M., … Seung, H. S. (2021). Binary and analog variation of synapses between cortical pyramidal neurons. ELife, 11. https://doi.org/10.7554/elife.76120

Duan, X., Krishnaswamy, A., Laboulaye, M. A., Liu, J., Peng, Y.-R., Yamagata, M., Toma, K. & Sanes, J. R. (2018). Cadherin Combinations Recruit Dendrites of Distinct Retinal Neurons to a Shared Interneuronal Scaffold. Neuron, 99(6), 1145-1*154*.e6. https://doi.org/10.1016/j.neuron.2018.08.019

Fambrough, D. & Goodman, C. S. (1996). The Drosophila beaten path Gene Encodes a Novel Secreted Protein That Regulates Defasciculation at Motor Axon Choice Points. Cell, *87*(6), 1049–*105*8. https://doi.org/10.1016/s0092-8*67*4(00)81799-7

Kolodkin, A. L. & Hiesinger, P. R. (2016). Wiring visual systems: common and divergent mechanisms and principles. Current Opinion in Neurobiology, *42*, 128–135. https://doi.org/10.1016/j.conb.2016.12.006

Kolodkin, A. L. & Tessier-Lavigne, M. (2011). Mechanisms and Molecules of Neuronal Wiring: A Primer. *Cold Spring Harbor Perspectives in Biology*, *3*(6), a001727. https://doi.org/10.1101/cshperspect.a001727

Kurmangaliyev, Y. Z., Yoo, J., LoCascio, S. A. & Zipursky, S. L. (2019). Modular transcriptional programs separately define axon and dendrite connectivity. ELife, 8, 505. https://doi.org/10.7554/elife.50822

Kurmangaliyev, Y. Z., Yoo, J., Valdes-Aleman, J., Sanfilippo, P. & Zipursky, S. L. (2020). Transcriptional Programs of Circuit Assembly in the Drosophila Visual System. Neuron, *108*(6), 1045-1057.e6. https://doi.org/10.1016/j.neuron.2020.10.006

Lee, T. & Luo, L. (2001). Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. *Trends in Neurosciences*, *24*(5), 251–254. http://www.sciencedirect.com/science/article/pii/S0*166*223600017914

Li, H., Watson, A., Olechwier, A., Anaya, M., Sorooshyari, S. K., Harnett, D. P., Lee, H.-K. P., Vielmetter, J., Fares, M. A., Garcia, K. C., Özkan, E., Labrador, J.-P. & Zinn, K. (2017). Deconstruction of the beaten Path-Sidestep interaction network provides insights into

neuromuscular system development. ELife, 6, 270. https://doi.org/10.7554/elife.28111

Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., Schilling, T., Bahl, A., Rubin, G. M., Nern, A., Dickson, B. J., Reiff, D. F., Hopp, E. & Borst, A. (2013). A directional tuning map of Drosophila elementary motion detectors. Nature, *500*(7461), 212–216. https://doi.org/10.1038/nature12320

Millard, S. S., Lu, Z., Zipursky, S. L. & Meinertzhagen, I. A. (2010). Drosophila dscam proteins regulate postsynaptic specificity at multiple-contact synapses. Neuron, 67(5), 761–768. https://doi.org/10.1016/j.neuron.2010.08.030

Motta, A., Berning, M., Boergens, K. M., Staffler, B., Beining, M., Loomba, S., Hennig, P., Wissler, H. & Helmstaedter, M. (2019). Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science*, *366*(6469). https://doi.org/10.*112*6/science.aay3134

Murtagh, F. & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, *31*(3), 274–295. https://doi.org/10.1007/s00357-014-9161-z

Nern, A., Pfeiffer, B. D. & Rubin, G. M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proceedings of the National Academy of Sciences*, *112*(22), E2967-76. https://doi.org/10.1073/pnas.1*50*6763112

Özel, M. N., Simon, F., Jafari, S., Holguera, I., Chen, Y.-C., Benhra, N., El-Danaf, R. N., Kapuralin, K., Malin, J. A., Konstantinides, N. & Desplan, C. (2020). Neuronal diversity and convergence in a visual system developmental atlas. Nature, *589*(7840), 88–95. https://doi.org/10.1038/s41586-020-2879-3

Özkan, E., Carrillo, R. A., Eastman, C. L., Weiszmann, R., Waghray, D., Johnson, K. G., Zinn, K., Celniker, S. E. & Garcia, K. C. (2013). An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. Cell, 154(1), 228–239. https://doi.org/10.1016/j.cell.2013.06.006

Paradis, E. & Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pei, J., Kinch, L. N. & Grishin, N. V. (2018). FlyXCDB—A Resource for Drosophila Cell Surface and Secreted Proteins and Their Extracellular Domains. *Journal of Molecular Biology*, *430*(Part B), 3353–3411. https://doi.org/10.1016/j.jmb.2018.06.002

Ren, X., Sun, J., Housden, B. E., Hu, Y., Roesel, C., Lin, S., Liu, L.-P., Yang, Z., Mao, D., Sun, L., Wu, Q., Ji, J.-Y., Xi, J., Mohr, S. E., Xu, J., Perrimon, N. & Ni, J.-Q. (2013). Optimized gene

editing technology for Drosophila melanogaster using germ line-specific Cas9. Proceedings of the National Academy of Sciences, *110*(47), 19012–19017. https://doi.org/10.1073/pnas.1318481110

Sanes, J. R. & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell, *181*(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Scheffer, L. K. & Meinertzhagen, I. A. (2019). The Fly Brain Atlas. *Annual Review of Cell and Developmental Biology*, *35*(1), 1–17. https://doi.org/10.1146/annurev-cellbio-100818-125444

Scheffer, L. K., Xu, C. S., Januszewski, M., Lu, Z., Takemura, S., Hayworth, K. J., Huang, G. B., Shinomiya, K., Maitlin-Shepard, J., Berg, S., Clements, J., Hubbard, P. M., Katz, W. T., Umayam, L., Zhao, T., Ackerman, D., Blakely, T., Bogovic, J., Dolafi, T., … Plaza, S. M. (2020). A connectome and analysis of the adult Drosophila central brain. ELife, 9, e57443. https://doi.org/10.7554/elife.57443

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Shekhar, K., Lapan, S. W., Whitney, I. E., Tran, N. M., Macosko, E. Z., Kowalczyk, M., Adiconis, X., Levin, J. Z., Nemesh, J., Goldman, M., McCarroll, S. A., Cepko, C. L., Regev, A. & Sanes, J. R. (2016). Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. Cell, 166(5), *1308*-1323.e30. https://doi.org/10.1016/j.cell.2016.07.054

Shen, K. & Bargmann, C. I. (2003). The Immunoglobulin Superfamily Protein SYG-1 Determines the Location of Specific Synapses in C. elegans. Cell, 112(5), 619–630. https://doi.org/10.1016/s0092-8674(03)00113-2

Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C. S., Aniceto, R., Ansari, N., Cheatham, N., Lauchie, S., Neace, E., Ogundeyi, O., Ordish, C., Peel, D., Shinomiya, A., Smith, C., Takemura, S., Talebi, I., Rivlin, P. K., Nern, A., … Meinertzhagen, I. A. (2019). Comparisons between the ON- and OFF-edge motion pathways in the Drosophila brain. ELife, 8, e40025. https://doi.org/10.7554/elife.40025

Shinomiya, K., Nern, A., Meinertzhagen, I. A., Plaza, S. M. & Reiser, M. B. (2022). Neuronal circuits integrating visual motion information in Drosophila melanogaster. *Current Biology*, *32*(16), 3529-3544.e2. https://doi.org/10.1016/j.cub.2022.06.061

Siebert, M., Banovic, D., Goellner, B. & Aberle, H. (2009). Drosophila motor axons recognize and follow a Sidestep-labeled substrate pathway to reach their target fields. *Genes & Development*, *23*(9), 1052–1062. https://doi.org/10.1101/gad.520509

Sink, H., Rehm, E. J., Richstone, L., Bulls, Y. M. & Goodman, C. S. (2001). sidestep encodes a target-derived attractant essential for motor axon guidance in Drosophila. Cell, 105(1), 57–67. https://doi.org/10.1016/s0092-8674(01)00296-3

Sperry, R. W. (1963). CHEMOAFFINITY IN THE ORDERLY GROWTH OF NERVE FIBER PATTERNS AND CONNECTIONS*. Proceedings of the National Academy of Sciences, 50(4), 703–710. https://doi.org/10.1073/pnas.50.4.703

Südhof, T. C. (2017). Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. Cell, *171*(4), 745–769. https://doi.org/10.1016/j.cell.2017.10.024

Südhof, T. C. (2021). The cell biology of synapse formation. *The Journal of Cell Biology*, *220*(7). https://doi.org/10.1083/jcb.202103052

Verschueren, E., Husain, B., Yuen, K., Sun, Y., Paduchuri, S., Senbabaoglu, Y., Lehoux, I., Arena, T. A., Wilson, B., Lianoglou, S., Bakalarski, C., Franke, Y., Chan, P., Wong, A. W., Gonzalez, L. C., Mariathasan, S., Turley, S. J., Lill, J. R. & Martinez-Martin, N. (2020). The Immunoglobulin Superfamily Receptome Defines Cancer-Relevant Networks Associated with Clinical Outcome. Cell, 182(2), 329-344.e19. https://doi.org/10.1016/j.cell.2020.06.007

Wojtowicz, W. M., Vielmetter, J., Fernandes, R. A., Siepe, D. H., Eastman, C. L., Chisholm, G. B., Cox, S., Klock, H., Anderson, P. W., Rue, S. M., Miller, J. J., Glaser, S. M., Bragstad, M. L., Vance, J., Lam, A. W., Lesley, S. A., Zinn, K. & Garcia, K. C. (2020). A Human IgSF Cell-Surface Interactome Reveals a Complex Network of Protein-Protein Interactions. Cell, 182(4), 1027-1043.e17. https://doi.org/10.1016/j.cell.2020.07.025

Wojtowicz, W. M., Wu, W., Andre, I., Qian, B., Baker, D. & Zipursky, S. L. (2007). A vast repertoire of Dscam binding specificities arises from modular interactions of variable Ig domains. Cell, 130(6), 1134–1145. https://doi.org/10.1016/j.cell.2007.08.026

Xu, C., Theisen, E., Maloney, R., Peng, J., Santiago, I., Yapp, C., Werkhoven, Z., Rumbaut, E., Shum, B., Tarnogorska, D., Borycz, J., Tan, L., Courgeon, M., Meinertzhagen, I. A., Bivort, B. de, Drugowitsch, J. & Pecot, M. Y. (2019). Control of Synaptic Specificity by Establishing a Relative Preference for Synaptic Partners. Neuron, 1–21. https://doi.org/10.1016/j.neuron.2019.06.006

Xu, S., Xiao, Q., Cosmanescu, F., Sergeeva, A. P., Yoo, J., Lin, Y., Katsamba, P. S., Ahlsen, G., Kaufman, J., Linaval, N. T., Lee, P.-T., Bellen, H. J., Shapiro, L., Honig, B., Tan, L. & Zipursky, S. L. (2018). Interactions between the Ig-Superfamily Proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits. Neuron, 100(6), 1369-1384.e6. https://doi.org/10.1016/j.neuron.2018.11.001

Zeng, H. (2022). What is a cell type and how to define it? Cell, *185*(15), 2739–2755.

https://doi.org/10.1016/j.cell.2022.06.031

Zhang, C., Hellevik, A., Takeuchi, S. & Wong, R. O. (2022). Hierarchical partner selection shapes rod-cone pathway specificity in the inner retina. *IScience*, *25*(9), 105032. https://doi.org/10.1016/j.isci.2022.105032

Zheng, Z., Lauritzen, J. S., Perlman, E., Robinson, C. G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C. B., Sharifi, N., Calle-Schuler, S. A., Kmecova, L., Ali, I. J., Karsh, B., Trautman, E. T., Bogovic, J. A., Hanslovsky, P., Jefferis, G. S. X. E., Kazhdan, M., … Bock, D. D. (2018). A Complete Electron Microscopy Volume of the Brain of Adult Drosophila melanogaster. Cell, 174(3), 730-743.e22. https://doi.org/10.1016/j.cell.2018.06.019

CHAPTER FIVE

Conclusion and future directions

5.1. Summary of results

In this dissertation, I presented three studies focused on identifying genetic programs of postmitotic neurons during wiring, utilizing single-cell transcriptome and synapse-level connectome. In Chapter 2, we discovered a modular architecture of gene expression programs in motion detector T4/T5 neurons. Strikingly, none of the eight subtypes of T4/T5 neurons had a unique cell-type specific marker, but instead, we observed three orthogonal transcriptional programs, or heterogeneity axes. These programs were expressed in a binary fashion for each subtype and were related to dendritic and axonal wiring of T4/T5 subtypes. This finding shows that specific gene expression programs associated with wiring can be identified by analyzing the transcriptional profiles of neurons during their wiring phase.

In Chapter 3, we expanded the scope of transcriptional profiling to the whole visual system during total pupal development, achieving more than 150 distinct neuronal populations annotated along nine time points during pupal development. Through this comprehensive longitudinal analysis, we were able to discover pan-neuronal and cell-type specific programs, as well as the dynamic nature of gene expression within these programs. An important contribution of this paper is the ability to obtain gene expression profiles of synaptic partners at the exact same developmental time point from the same brain, as all cell types were sampled simultaneously. To my knowledge, this atlas is the largest and the most detailed developmental transcriptome that can be mapped to an EM level connectome, to date. Additionally, the genetic multiplexing methods we employed to differentiate between different timepoints have proven useful for other studies where multiplexing is required, such as perturb-seq experiments utilizing different DGRP lines for control and RNAi knock-down conditions (Jain et al., 2022).

In Chapter 4, we built upon the developmental transcriptome atlas from Chapter 3 and combined it with synapse-level connectome to address questions regarding synaptic specificity.

The EM connectome allowed us to map the connectivity of eight T4/T5 subtypes and their postsynaptic partners. Through transcriptomic analysis, we discovered a group of postsynaptic neurons that were transcriptionally similar. By investigating the factors that determine why T4/T5 subtypes choose certain transcriptionally related postsynaptic neurons, we identified the differential expression of Beat/Side IgSF family genes. Further genetic experiments demonstrated that side-II is required for final synaptic partner choice of T4d/T5d within the lobula plate layer 4. Overall these findings provide insights into the molecular mechanisms underlying synaptic specificity in developing visual systems.

5.2. Big data for comprehensive understanding of neuronal wiring and beyond.

Single-cell RNA sequencing is now offering unprecedented access to gene expression profiles of any parts of an animal in any species. In parallel to the Human Cell Atlas, the Drosophila community around the world is making a collaborative effort to build a single-cell transcriptome atlas. Launched in 2020, the Fly Cell Atlas is building a database of single-cell transcriptomes of entire adult Drosophila, including not only the brain, but also many parts of the Drosophila body: antenna, gut and legs (flycellatlas.org). In addition to studying wiring and synaptic specificity, the optic lobe developmental atlas from our work and similar datasets from others add valuable resources to address various biological questions (Kurmangaliyev et al., 2020; Özel et al., 2020).

On the other hand, recent studies of EM connectome in both Drosophila brain and preliminary studies in mouse retina and cortex have revealed an unprecedented level of complexity and organizational rules that are yet to be fully understood (Kim et al., 2014; Shinomiya et al., 2022; Takemura et al., 2013; Turner et al., 2020). These studies have not only uncovered the connectivity between different cell types, but also the detailed subcellular locations of connections, providing insights into the elaborate and controlled wiring achieved during

development. Such findings spur a deeper exploration of the molecular mechanisms underlying this specificity.

The accumulation of big data in single-cell research will continue to provide increasingly complex information on individual cells. With rapid progress of methods to gain various modalities, such as single-cell ATACseq for chromatin state changes (Janssens et al., 2022), emerging long-read sequencing for alternative splicing measurement (Traunmüller et al., 2023), along with cell-type specific proteome (J. Li et al., 2020; Shuster et al., 2022) will enable a more comprehensive description of single neurons and their changes during development. With the challenge shifting from data scarcity to data overload, it is crucial to use these data effectively to answer biological questions one has.

The T4/T5 motion detector system is an excellent model to study synaptic specificity as we have prior knowledge of their development from progenitors to their morphological maturation trajectory of eight subtypes (Apitz & Salecker, 2018; Pinto-Teixeira et al., 2018). The synaptic connectivity of both pre- and postsynaptic partners is well mapped at the EM level, revealing shared and distinct wiring patterns between each subtype. This information allows us to determine where the eight subtypes make diverging choices to build differential connectivity (Shinomiya et al., 2022; Takemura et al., 2017). Among eight subtypes, T4c and T4d subtypes have extremely similar developmental trajectories. Their axon terminals initially target their axon terminals within the same domain of the lobula plate (proto-layer c/d), and as they mature, they make subtypes specific synapses by selecting different partners via small differences in their gene expression (**Figure 1.5B**). T4c and T4d undergo extremely similar morphological development and similar batteries of gene expression programs in their postmitotic developmental period. While any two different neuronal types exhibit many differences in characteristics including diverse morphologies, molecular identities, and wiring patterns, the similarity between T4c and T4d allows to link the small differences in gene expressions to the diverging choices made by T4c and T4d

to connect with two different postsynaptic partners, LLPC2 and LLPC3. As such, prior knowledge on the properties of neuron provides an opportunity to assess the relationship between specific genetic programs and specific neuronal features.

5.3. Disentangling cell surface protein's role in synaptic specificity

The process of proper wiring and synapse formation involves multiple stages where interactions between cell adhesion molecules are crucial. While it may not always be easy to distinguish between axon guidance and synaptic partner choice process, the appropriate use of adhesion molecules and secreted molecules at each stage is necessary for progression to the next step of decision making. For example, canonical guidance cues such as Netrin, Slits, Semaphorins, and Ephrins play an important role in driving axon terminals to the domain where they locate their synaptic partners and establish synapses (Sanes & Zipursky, 2020).

Cell adhesion molecules participate in extremely complex networks of protein-protein interactions in both vertebrate and invertebrate systems (Kolodkin & Hiesinger, 2016; Sanes & Zipursky, 2020). IgSF proteins such as DIP/dpr and Side/Beat families, have a wide range of affinities (Cosmanescu et al., 2018; H. Li et al., 2017), and display cell-type specific dynamic expression patterns (Kurmangaliyev et al., 2020; Tan et al., 2015). Given the diverse array of adhesion molecules expressed by neurons and the dynamic changes in their composition during development, it is extremely challenging to identify the molecules responsible for the decisions neurons make among the complex proteome environment of neighboring contacts. Despite this complexity, there are some examples that highlight the importance of these interactions, such as the DIP-α and dpr6/10 pairs, which have been implicated in dendrite arborization, synapse numbers, and cell survival in medulla neurons (Xu et al., 2018, 2022). These limited examples where a few cell surface molecules are disrupted by gene deletion or protein avidity mutations shows that it may be challenging to understand the role of each cell adhesion molecule.

One approach for disentangling the complexity of cell adhesion molecules would be assessing the expression level and dynamic composition of the cell surface proteome of each cell type at any given time. However, this method is not practically feasible. An alternative, more feasible approach is presented in Chapter 4, which focuses on closely related neurons that share developmental changes both in gene expression patterns and the external environment factors. This approach allows for the testing of the role of genes or molecules of interest. It is important to note that the same molecules may have different effects depending on the context in which they are present, and thus minimizing variability is crucial. As described in the previous section, the confined developmental context of time, space of T4/T5 subtypes enabled us to find the role of Side/Beat pairs as the factor that tips the balance for one synaptic choice over the other.

5.4. Future directions

This dissertation can be developed in three ways to deepen our understanding of the molecular mechanisms underlying brain development. First, further investigation into gene regulation is needed. While the transcriptome provides valuable information on cell-type specific gene expression patterns during development, a thorough understanding of the gene regulatory mechanisms of developing nervous system requires examination of chromatin state changes during development (Janssens et al., 2022). This can help identify the regulatory mechanisms that impact synaptic choices and wiring during development. T4/T5 progenitors undergo two Notch dependent asymmetric divisions, which correspond to binary expression patterns of grn and TfAP-2, respectively. Although these two divisions diverge developmental trajectories for T4/T5, all T4 and all T5 converge onto the same transcriptional programs that regulate their dendritic targeting and wiring. The binary but converging regulatory programs and resulting cell surface molecule expression regulation may be related to T4/T5 neurons chromatin state characteristics.

Moreover, capturing spatial information is crucial as the current transcriptome is mainly generated from dissociated cells, which results in the loss of spatial information. For instance, our work in Chapter 3 showed that some genes are expressed in dorso-ventral gradients along the spatial axes of the whole optic lobe. A recent study showed that the whole brain scale spatial axis can affect synaptic gradients (Dombrovski et al., 2023). Local translation events from localized mRNA may occur in axons or dendrites, which cannot be captured from cell bodies. While it is unclear whether this actively occurs in the Drosophila optic lobe, evidence exists of mRNA transport to cell structures beyond cell bodies, indicating the possibility of local translation.

Second, understanding protein dynamics is necessary to better comprehend the molecular environment of developing neurons, as transcriptomes may not always correlate with protein expression. Localization of protein also needs to be understood. For example, the temporal gene expression pattern and expression levels of GluClα in all T4/T5 subtypes are the same, but the protein localization during development differs between T4 and T5. Before 72h APF, GluClα localizes to both T4 and T5 dendrites, but after 72hAPF, GluClα in T5 dendrites clear out. This disparity between the transcriptome and proteome at the local level indicates that relying solely on the transcriptome to predict the cell surface environment can be limiting. Additionally, the role of alternative splicing that produces multiple protein isoforms as a product adds complexity especially for cell surface proteomes (Furlanis et al., 2019; Traunmüller et al., 2023).

Finally, large-scale transcriptome studies are advancing our understanding of molecular cell types (Berg et al., 2021; (BICCN) et al., 2021; Yao et al., 2021). Also, EM-level connectome data are beginning to be generated at a large scale in vertebrate brains including mouse, primate, and human. (Loomba et al., 2022; Motta et al., 2019; Schneider-Mizell et al., 2023). The studies I presented in this dissertation provide a framework to utilize developmental transcriptome and synapse-level connectome to understand molecular mechanisms for brain development.

5.5 References

Apitz, H. & Salecker, I. (2018). Spatio-temporal relays control layer identity of direction-selective neuron subtypes in Drosophila. Nature Communications, 1–16. https://doi.org/10.1038/s41467-018-04592-z

Berg, J., Sorensen, S. A., Ting, J. T., Miller, J. A., Chartrand, T., Buchin, A., Bakken, T. E., Budzillo, A., Dee, N., Ding, S.-L., Gouwens, N. W., Hodge, R. D., Kalmbach, B., Lee, C., Lee, B. R., Alfiler, L., Baker, K., Barkan, E., Beller, A., … Lein, E. S. (2021). Human neocortical expansion involves glutamatergic neuron diversification. Nature, 1–29. https://doi.org/10.1038/s41586-021-03813-8

(BICCN), B. I. C. C. N., authors, B. I. C. C. N. (BICCN) C., Callaway, E. M., Dong, H.-W., Ecker, J. R., Hawrylycz, M. J., Huang, Z. J., Lein, E. S., Ngai, J., Osten, P., Ren, B., Tolias, A. S., White, O., Zeng, H., Zhuang, X., investigators, B. contributing principal, Ascoli, G. A., Behrens, M. M., Chun, J., … Sunkin, S. (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. Nature, 598(7879), 86–102. https://doi.org/10.1038/s41586-021-03950-0

Cosmanescu, F., Katsamba, P. S., Sergeeva, A. P., Ahlsen, G., Patel, S. D., Brewer, J. J., Tan, L., Xu, S., Xiao, Q., Nagarkar-Jaiswal, S., Nern, A., Bellen, H. J., Zipursky, S. L., Honig, B. & Shapiro, L. (2018). Neuron-Subtype-Specific Expression, Interaction Affinities, and Specificity Determinants of DIP/Dpr Cell Recognition Proteins. Neuron, 100(6), 1385-1400.e6. https://doi.org/10.1016/j.neuron.2018.10.046

Dombrovski, M., Peek, M. Y., Park, J.-Y., Vaccari, A., Sumathipala, M., Morrow, C., Breads, P., Zhao, A., Kurmangaliyev, Y. Z., Sanfilippo, P., Rehan, A., Polsky, J., Alghailani, S., Tenshaw, E., Namiki, S., Zipursky, S. L. & Card, G. M. (2023). Synaptic gradients transform object location to action. Nature, 613(7944), 534–542. https://doi.org/10.1038/s41586-022-05562-8

Furlanis, E., Traunmüller, L., Fucile, G. & Scheiffele, P. (2019). Landscape of ribosome-engaged transcript isoforms reveals extensive neuronal-cell-class-specific alternative splicing programs. Nature Neuroscience, 22(10), 1709–1717. https://doi.org/10.1038/s41593-019-0465-5

Jain, S., Lin, Y., Kurmangaliyev, Y. Z., Valdes-Aleman, J., LoCascio, S. A., Mirshahidi, P., Parrington, B. & Zipursky, S. L. (2022). A global timing mechanism regulates cell-type-specific wiring programmes. Nature, 603(7899), 112–118. https://doi.org/10.1038/s41586-022-04418-5

Janssens, J., Aibar, S., Taskiran, I. I., Ismail, J. N., Gomez, A. E., Aughey, G., Spanier, K. I., Rop, F. V. D., González-Blas, C. B., Dionne, M., Grimes, K., Quan, X. J., Papasokrati, D., Hulselmans, G., Makhzami, S., Waegeneer, M. D., Christiaens, V., Southall, T. & Aerts, S. (2022). Decoding gene regulation in the fly brain. Nature, 1–7. https://doi.org/10.1038/s41586-021-04262-z

Kim, J. S., Greene, M. J., Zlateski, A., Lee, K., Richardson, M., Turaga, S. C., Purcaro, M., Balkam, M., Robinson, A., Behabadi, B. F., Campos, M., Denk, W., Seung, H. S. & EyeWirers. (2014). Space-time wiring specificity supports direction selectivity in the retina. Nature, 509(7500), 331–336. https://doi.org/10.1038/nature13240

Kolodkin, A. L. & Hiesinger, P. R. (2016). Wiring visual systems: common and divergent mechanisms and principles. Current Opinion in Neurobiology, 42, 128–135. https://doi.org/10.1016/j.conb.2016.12.006

Kurmangaliyev, Y. Z., Yoo, J., Valdes-Aleman, J., Sanfilippo, P. & Zipursky, S. L. (2020). Transcriptional Programs of Circuit Assembly in the Drosophila Visual System. Neuron, 108(6), 1045-1057.e6. https://doi.org/10.1016/j.neuron.2020.10.006

Li, H., Watson, A., Olechwier, A., Anaya, M., Sorooshyari, S. K., Harnett, D. P., Lee, H.-K. P., Vielmetter, J., Fares, M. A., Garcia, K. C., Özkan, E., Labrador, J.-P. & Zinn, K. (2017). Deconstruction of the beaten Path-Sidestep interaction network provides insights into neuromuscular system development. ELife, 6, 270. https://doi.org/10.7554/elife.28111

Li, J., Han, S., Li, H., Udeshi, N. D., Svinkina, T., Mani, D. R., Xu, C., Guajardo, R., Xie, Q., Li, T., Luginbuhl, D. J., Wu, B., McLaughlin, C. N., Xie, A., Kaewsapsak, P., Quake, S. R., Carr, S. A., Ting, A. Y. & Luo, L. (2020). Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators. Cell, 180(2), 373-386.e15. https://doi.org/10.1016/j.cell.2019.12.029

Loomba, S., Straehle, J., Gangadharan, V., Heike, N., Khalifa, A., Motta, A., Ju, N., Sievers, M., Gempt, J., Meyer, H. S. & Helmstaedter, M. (2022). Connectomic comparison of mouse and human cortex. Science, 377(6602), eabo0924. https://doi.org/10.1126/science.abo0924

Motta, A., Berning, M., Boergens, K. M., Staffler, B., Beining, M., Loomba, S., Hennig, P., Wissler, H. & Helmstaedter, M. (2019). Dense connectomic reconstruction in layer 4 of the somatosensory cortex. Science, 366(6469). https://doi.org/10.1126/science.aay3134

Özel, M. N., Simon, F., Jafari, S., Holguera, I., Chen, Y.-C., Benhra, N., El-Danaf, R. N., Kapuralin, K., Malin, J. A., Konstantinides, N. & Desplan, C. (2020). Neuronal diversity and convergence in a visual system developmental atlas. Nature, 589(7840), 88–95. https://doi.org/10.1038/s41586-020-2879-3

Pinto-Teixeira, F., Koo, C., Rossi, A. M., Neriec, N., Bertet, C., Li, X., Rodríguez, A. del V. & Desplan, C. (2018). Development of Concurrent Retinotopic Maps in the Fly Motion Detection Circuit. Cell, 1–26. https://doi.org/10.1016/j.cell.2018.02.053

Sanes, J. R. & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. Cell, 181(3), 536–556. https://doi.org/10.1016/j.cell.2020.04.008

Schneider-Mizell, C. M., Bodor, A., Brittain, D., Buchanan, J., Bumbarger, D. J., Elabbady, L., Kapner, D., Kinn, S., Mahalingam, G., Seshamani, S., Suckow, S., Takeno, M., Torres, R., Yin, W., Dorkenwald, S., Bae, J. A., Castro, M. A., Fahey, P. G., Froudakis, E., … Costa, N. M. da. (2023). Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex. BioRxiv, 2023.01.23.525290. https://doi.org/10.1101/2023.01.23.525290

Shinomiya, K., Nern, A., Meinertzhagen, I. A., Plaza, S. M. & Reiser, M. B. (2022). Neuronal circuits integrating visual motion information in Drosophila melanogaster. Current Biology, 32(16), 3529-3544.e2. https://doi.org/10.1016/j.cub.2022.06.061

Shuster, S. A., Li, J., Chon, Ur., Sinantha-Hu, M. C., Luginbuhl, D. J., Udeshi, N. D., Carey, D. K., Takeo, Y. H., Xie, Q., Xu, C., Mani, D. R., Han, S., Ting, A. Y., Carr, S. A. & Luo, L. (2022).

In situ cell-type-specific cell-surface proteomic profiling in mice. Neuron. https://doi.org/10.1016/j.neuron.2022.09.025

Takemura, S., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P. K., Katz, W. T., Olbris, D. J., Plaza, S. M., Winston, P., Zhao, T., Horne, J. A., Fetter, R. D., Takemura, S., Blazek, K., Chang, L.-A., Ogundeyi, O., Saunders, M. A., Shapiro, V., … Chklovskii, D. B. (2013). A visual motion detection circuit suggested by Drosophila connectomics. Nature, 500(7461), 175–181. https://doi.org/10.1038/nature12450

Takemura, S., Nern, A., Chklovskii, D. B., Scheffer, L. K., Rubin, G. M. & Meinertzhagen, I. A. (2017). The comprehensive connectome of a neural substrate for 'ON' motion detection in Drosophila. ELife, 6, e24394. https://doi.org/10.7554/elife.24394

Tan, L., Zhang, K. X., Pecot, M. Y., Nagarkar-Jaiswal, S., Lee, P.-T., Takemura, S.-Y., McEwen, J. M., Nern, A., Xu, S., Tadros, W., Chen, Z., Zinn, K., Bellen, H. J., Morey, M. & Zipursky, S. L. (2015). Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. Cell, 163(7), 1756–1769. https://doi.org/10.1016/j.cell.2015.11.021

Traunmüller, L., Schulz, J., Ortiz, R., Feng, H., Furlanis, E., Gomez, A. M., Schreiner, D., Bischofberger, J., Zhang, C. & Scheiffele, P. (2023). A cell-type-specific alternative splicing regulator shapes synapse properties in a trans-synaptic manner. Cell Reports, 42(3), 112173. https://doi.org/10.1016/j.celrep.2023.112173

Turner, N. L., Macrina, T., Bae, J. A., Yang, R., Wilson, A. M., Schneider-Mizell, C., Lee, K., Lu, R., Wu, J., Bodor, A. L., Bleckert, A. A., Brittain, D., Froudarakis, E., Dorkenwald, S., Collman, F., Kemnitz, N., Ih, D., Silversmith, W. M., Zung, J., … Seung, H. S. (2020). Multiscale and multimodal reconstruction of cortical structure and function. 72(5 Pt 2), 056708–056759. https://doi.org/10.1101/2020.10.14.338681

Xu, S., Sergeeva, A. P., Katsamba, P. S., Mannepalli, S., Bahna, F., Bimela, J., Zipursky, S. L., Shapiro, L., Honig, B. & Zinn, K. (2022). Affinity requirements for control of synaptic targeting and neuronal cell survival by heterophilic IgSF cell adhesion molecules. Cell Reports, 39(1), 110618. https://doi.org/10.1016/j.celrep.2022.110618

Xu, S., Xiao, Q., Cosmanescu, F., Sergeeva, A. P., Yoo, J., Lin, Y., Katsamba, P. S., Ahlsen, G., Kaufman, J., Linaval, N. T., Lee, P.-T., Bellen, H. J., Shapiro, L., Honig, B., Tan, L. & Zipursky, S. L. (2018). Interactions between the Ig-Superfamily Proteins DIP-α and Dpr6/10 Regulate Assembly of Neural Circuits. Neuron, 100(6), 1369-1384.e6. https://doi.org/10.1016/j.neuron.2018.11.001

Yao, Z., Velthoven, C. T. J. van, Nguyen, T. N., Goldy, J., Sedeno-Cortes, A. E., Baftizadeh, F., Bertagnolli, D., Casper, T., Chiang, M., Crichton, K., Ding, S.-L., Fong, O., Garren, E., Glandon, A., Gouwens, N. W., Gray, J., Graybuck, L. T., Hawrylycz, M. J., Hirschstein, D., … Zeng, H. (2021). A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell, 184(12), 3222-3241.e26. https://doi.org/10.1016/j.cell.2021.04.021