UNIVERSITY OF CALIFORNIA

Los Angeles

A Molecular Approach to Identify Determinants of Synaptic Specificity

A dissertation submitted in partial satisfaction of the

requirements for the degree of Doctor of Philosophy

in Biological Chemistry

By

Liming Tan

2016

ABSTRACT OF THE DISSERTATION

A Molecular Approach to Identify Determinants of Synaptic Specificity

by

Liming Tan

Doctor of Philosophy in Biological Chemistry

University of California, Los Angeles, 2016

Professor Stephen Lawrence Zipursky, Chair

Information processing in the nervous system relies on precise patterns of synaptic connections between neurons. The level of synaptic specificity within the nervous system is remarkable given its tremendous complexity. How neurons can distinguish their correct synaptic partners from many other neurons during circuit assembly remains a central question in neuroscience. Although important progress has been made on molecular mechanisms regulating neural circuit assembly, the cellular recognition mechanisms mediating synaptic specificity are still poorly understood. The *Drosophila* visual system is well suited to uncovering the molecular mechanisms underlying synaptic specificity, because of the availability of diverse genetic tools, cell type specific markers and electron microscopic reconstruction data. In the medulla neuropil of the *Drosophila* visual system, different neurons form synaptic connections in different layers. Within a layer, neurons form synapses with a unique set of multiple neuronal types, which represents only a subset of

neurons with processes in that layer. In my thesis research, I sought to identify candidate cell recognition molecules underlying this specificity. I did RNA sequencing on closely related neurons with different layer-specific synaptic specificities, lamina neurons L1-L5 and photoreceptor R7 and R8, at the onset of synapse formation. I showed that each of the seven cell types expresses a unique set of hundreds of genes encoding cell surface and secreted proteins. Using these data and additional localization studies on proteins tagged through modification of the endogenous locus, I demonstrated that 21 paralogs of the Dpr family, a subclass of Immunoglobulin (Ig) domain containing cell surface proteins, are expressed in unique combinations in each of the seven cell types during synapse formation. Dpr interacting proteins (DIPs), comprising nine paralogs of another subclass of Ig-containing proteins, are expressed in a complementary layer-specific fashion in subsets of synaptic partners for each of the seven cell types. Thus, I demonstrated that interacting Dprs and DIPs are expressed in synaptic partners during synapse formation in the *Drosophila* visual system. Furthermore, I generated null mutants of Dprs and DIPs via CRISPR-based methods, and showed that mutants of two DIPs and their cognate Dprs had similar phenotypes: neurons normally expressing these DIPs had reduced number of cells in corresponding DIP and cognate Dpr mutants. These data suggested that cognate Dprs and DIPs play important roles in the development of synaptic partners expressing them.

The dissertation of Liming Tan is approved.

Alvaro Sagasti

Leonid Kruglyak

Kelsey C. Martin

Stephen Lawrence Zipursky, Committee Chair

University of California, Los Angeles

2016

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

This dissertation would not have been possible without the support from many people throughout my graduate training.

First and foremost, I am tremendously grateful to my thesis advisor, Larry Zipursky. I am thankful to Larry for his constant guidance, support and encouragement. I have learnt from Larry how to think creatively, how to ask and focus on an important scientific question, how to address scientific problems rigorously, and how to communicate efficiently, which will be my most valuable assets continuously in my scientific career. Moreover, Larry has always encouraged me to follow my passion in science, which will last throughout my future career.

I am also appreciative to the members of my thesis committee, Kelsey Martin, Leonid Kruglyak, Alvaro Sagasti and my previous committee member Jason Ernst for their valuable guidance and support throughout my graduate training.

I am grateful to members of the Zipursky lab for guidance, collaboration, intellectual simulation and friendship. First, I am very thankful to Marta Morey, Jason McEwen, Kelvin Zhang, and Matt Pecot. Marta guided me to initiate the FACS project, and Jason taught me many important techniques in molecular biology, particularly RNA purification, amplification and cDNA library construction, which made my thesis project possible. Kelvin help me analyzed all the RNA sequencing data. Matt provided lamina neuron type specific markers for FACS. I also thank Shuwa Xu, Wael Tadros, Orkun Akin, and Dorian Gunning for their tremendous help in my thesis project. I also appreciate the collaboration on Dpr/DIP project with Shuwa Xu and Qi Xiao. I would also like to thank Punam Patel, Cuong Truong and Western Kramer, excellent undergraduate students

who screened CRISPR null mutants of DIPs and Dprs. Lastly, I would like to thank all the past and present members of the Zipursky lab for valuable comments and suggestions on my projects.

I am also grateful to my collaborators. I'd like to thank Aljoscha Nern from Janelia Research Campus for intellectual discussions and useful reagents. I also want to thank Shin-ya Takemura, Louis Scheffer and Ian Meinertzhagen for sharing results from the EM studies at Janelia Research Campus. I want to thank Sonal Nagarkar-Jaiswal, Pei-Tseng Lee in Hugo Bellen lab at Baylor College of Medicine for sharing and generating MiMIC derived fly lines for DIPs, Dprs, Beats and Sides, and Robert Carrillo, Kaushiki Menon in Kai Zinn lab at Caltech for communication and collaboration on Dpr/DIP project.

My PhD would not be possible without the support of my family. I am thankful to my wife Ran Cheng. I am indebted to her unwavering encouragement, support and friendship throughout the years, especially at difficult times. I am eternally grateful to my parents for their unconditional love, trust and support.

My dissertation work was supported in part by a Graduate Student Fellowship from the China Scholarship Council (CSC), a Graduate Student Fellowship from the University of California, Los Angeles (UCLA) Philip Whitcome Training Program, and the Dissertation Year Fellowship from UCLA Graduate Divisions. I wish to express my appreciation for the financial support.

Chapter 2 is a version of 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, 1756-1769. The authors of this manuscript thank Orkun

Akin, John Carlson, Claude Desplan, Yasushi Hiromi, Frank Laski, Cheng-yu Lee for sharing reagents; members of our laboratory for discussion; Louis K. Scheffer and Ian Meinertzhagen for discussions, sharing results from the EM studies at Janelia Research Campus (HHMI) and comments on the manuscript; Kaushiki Menon and Robert Carrillo (Zinn group) for communicating results prior to publication; Donghui Cheng, Tanya Stoyanova and Owen Witte for assistance in FACS purification of cells; Dorian Gunning for generating Bsh antibody; and Barret Pfeiffer and Gerald Rubin (JRC/HHMI) for reagents to construct the markers for R7 neurons. The work was supported by: a Graduate Student Fellowship from the China Scholarship Council (CSC) and a Graduate Student Fellowship from the University of California, Los Angeles (UCLA) Philip Whitcome Training Program (L.T.); the Canadian Institute of Health Research Fellowship (K.X.Z.); Jane Coffin Childs Memorial Fund for Medical Research (M.P.).; the Robert A. and Renee E. Belfer Family Foundation (S.N.-J.); Target A.L.S. (P.-T.L.); a Long-Term Fellowship from the Human Frontiers Science Program (W.T.); grants from National Institute of Health (R01GM067858 (H.J.B.), R01NS62821 (K.Z.) and R01NS28182 (K.Z.); a Ramon y Cajal contract (RYC-2011-09479) and a Ministerio de Economia y Competitividad grant (BFU2012-32282) (M.M.). H.J.B. and S.L.Z. are investigators of Howard Hughes Medical Institute.

VITA

June, 2010	B.S., Chemistry Zhejiang University Hangzhou, China
2010	Graduate Student ACCESS Graduate Program University of California, Los Angeles, Los Angeles, California
2010	CSC Scholarship China Scholarship Council
2014	Whitcome Fellowship, Philip Whitcome Training Program University of California, Los Angeles, Los Angeles, California
2015	Dissertation Year Fellowship Graduate Division University of California, Los Angeles, Los Angeles, California

PUBLICATIONS AND PRESENTATIONS

Zhang KX, Tan L, Pellegrini M, Zipursky SL, McEwen JM (2016) Rapid Changes in the Translatome during the Conversion of Growth Cones to Synaptic Terminals. *Cell Reports 14*(5), 1258–1271

Tan L, Zhang KX, Pecot MY, Nagarkar-Jaiswal S, Lee PT, Takemura SY, McEwen JM, Nern A, Xu S, Tadros W, Chen Z, Zinn K, Bellen HJ, Morey M, Zipursky SL (2015) Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. *Cell 163*(7), 1756–1769

Tan L, Zhang KX, Pecot MY, Nagarkar-Jaiswal S, Lee PT, Takemura SY, McEwen JM, Nern A, Xu S, Tadros W, Chen Z, Zinn K, Bellen HJ, Morey M, Zipursky SL (2015) Complementary Expression of Immunoglobulin Superfamily Ligands and Receptors in Synaptic Pairs in the *Drosophila* Visual System. Talk presented at the Cold Spring Harbor Laboratory Conference: Neurobiology of *Drosophila*

Tan L, Morey M, Zhang KX, McEwen JM, Pecot MY, Tadros W, Xu S, Chen Y, Liu W, Zipursky SL (2014) Layer-specific expression of Ig-Superfamily Ligands and Receptors in the *Drosophila* Visual System. Poster presented at the Janelia Research Campus: High-throughput Sequencing for Neuroscience

Tan L, Morey M, Zhang KX, McEwen JM, Akin O, Bader M, Chen Y, Zipursky SL (2013) Regulation of Transcription by Growth Cone Receptors during Synapse Formation. Poster presented at 23rd European *Drosophila* Research Conference, Barcelona

Chapter 1. Introduction

Overview

A functional nervous system relies on precise patterns of connections between neurons. Two neurons connect to each other via a synapse, a structure composed of presynaptic site, the synaptic cleft and postsynaptic site. When the presynaptic neuron fires, the presynaptic site releases neurotransmitters, which travel through the synaptic cleft and bind with the neurotransmitter receptors located in the postsynaptic site, resulting in activation or inhibition of postsynaptic neuron. It is estimated that a fruit fly has about 250,000 neurons forming millions of synapses, while about 10^{11} neurons form 10^{14} synapses in the human brain. The integrity of different neural circuits among all the neurons enables an organism to perceive, think, decide and behave. How neurons can distinguish their correct synaptic partners from other neurons during neural circuit assembly remains a central question in neuroscience.

Studies in different model organisms have revealed various mechanisms underlying neural circuit assembly. In a broad sense, these mechanisms can be divided into two major categories, genetically encoded molecular mechanisms and activity or experience-dependent mechanisms. Studies in mammals indicate that activity is critical for layer specificity of retinal ganglion cell (RGC) arbors in the lateral geniculate nucleus (LGN). Axons from the ipsilateral and contralateral eyes initially arborize broadly in the LGN. Activity-dependent processes eliminate branches in inappropriate laminae and refine their arbors in appropriate laminae (Feller, 2009; Penn et al., 1998; Sretavan et al., 1988). However, activity and experience are only important for sharpening retinotopic and laminar projections in the vertebrate visual system, and they are dispensable for formation of

Drosophila visual circuits (Hiesinger et al., 2006; Sanes and Yamagata, 2009). Therefore, genetically encoded molecular mechanisms play more important roles in development of the nervous system. In this Introduction, I will focus on the molecular mechanisms regulating neural circuit assembly.

Diverse cellular recognition strategies regulate circuit assembly

Through regeneration studies in the autonomic nervous system in cat by J. N. Langley (Langley, 1895) and in the optic nerves in amphibia by Roger Sperry (Sperry, 1943, 1944, 1963), it was proposed that differences in molecules between neurons determine the specificity of synaptic connections. Over the past several decades, biochemical and genetic studies have led to the identification of recognition molecules, which are cell surface and secreted proteins regulating the wiring of neural circuits through intercellular signaling pathways. From these studies, three general molecular strategies underlying circuit assembly have emerged.

First, a limited set of conserved cell surface and secreted molecules are used in combination to regulate axon guidance in many different regions of the developing nervous systems in invertebrate and vertebrate. These include ephrins, semaphorins, plexins, DCC, netrins, slits, robos, and various cell adhesion molecules, such as cadherins and immunoglobulin superfamily proteins (Dickson, 2002). These molecules can act as both attractants and repellents mediating short-range (contact mediated) and long-range (diffusible) interactions. Studies in vertebrates and the fruit fly suggest that the same mechanism and, in some cases, the same set of molecules are used in similar guidance processes. One example is the use of attraction and repulsion during midline crossing in both vertebrates and fruit fly. Studies in rodents, chicks, and flies have revealed that commissural growth cones first are attracted by Netrin to the midline through DCC (Kennedy et al., 1994;

Serafini et al., 1994). Before midline crossing, growth cones are insensitive to Slit and, in vertebrates, certain class 3 Semaphorins (Brose et al., 1999; Zou et al., 2000), although these signals are expressed at the midline prior to midline crossing. After crossing, these axons respond to Slit and Semaphorins (Brose et al., 1999) and are repelled from the midline by them (Kidd et al., 1998) and never turn back (Figure 1). Although these molecules are critical for axon guidance, in most cases they are more broadly expressed, and generally not in a cell-type specific fashion. After axons are guided within the target region, they still need to discriminate among neurites from many different neurons in that region and make synapses on appropriate partners. Given the complexity of the nervous system, many different cell recognition molecules expressed in a cell-type specific fashion are likely to be involved in the synaptic target matching after axon guidance.

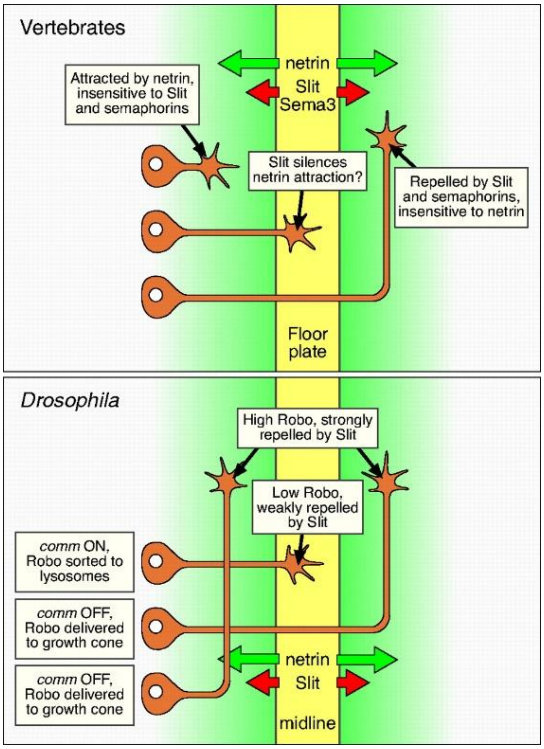


Figure 1-1. Same set of molecules are used by commissural axons to cross the midline in vertebrates and *Drosophila*. As vertebrate commissural axons cross the floor plate, they are attracted by Netrin to the midline through DCC. After crossing the midline they are repelled from midline by Slit and Sema3 in the midline via Robo and Plexin. Drosophila commissural axons are also attracted by netrin to midline through DCC homolog Fra, and also repelled by Slit through Robo after crossing. Adapted from Dickson, 2002.

Second, gradients of cell surface proteins, such as Ephs/Ephrins and Wnts, play a crucial role in establishing topographic maps, (Cang and Feldheim, 2013; Triplett and Feldheim, 2012), a widespread organizational principle in the vertebrate brain. Ephs are a family of receptor tyrosine kinases and are subdivided into A and B classes. Ligands for Ephs, called ephrins, are also divided into A and B classes. In vitro binding assays have demonstrated each EphA receptor can bind with high affinity to all ephrin-As, and each EphB can bind to all ephrin-Bs, with very little cross binding between A and B families (Gale et al., 1996; Himanen et al., 2004). Binding between Eph receptors and ephrin ligands induces repulsive signaling (Egea and Klein, 2007). One simple example for Eph/Ephrin's function in topographic mapping is in the mammalian retinocollicular system. Multiple EphA and ephrin-A family members are expressed in complementary gradients along the nasal-temporal axis of the retina and anterior-posterior axis in the superior colliculus (SC) in mouse (Figure 2A). Both loss and gain of function studies have demonstrated that the EphAs and ephrin-As are required for normal topographic mapping of the temporal-nasal axis of the visual field (Feldheim et al., 2000; Frisén et al., 1998; Pfeiffenberger, Yamada, & Feldheim, 2006; Carreres et al., 2011). Axons expressing different levels of EphAs terminate in the SC based on counterbalancing their tendency to grow toward the posterior SC and the amount of repulsive signal they receive from the posterior-derived ephrin-As. For instance, temporal axons, with high levels of EphA, terminate at anterior SC with low levels of ephrin-A. By contrast, nasal axons, with low EphA receptor levels, terminate at the posterior SC with high levels of ephrin-A (Figure 2B). However, it is not clear that gradients could endow axons with the ability to distinguish among different types of neurons that are physically intermingled rather than spatially arrayed. Within the target region, neurons may require qualitatively distinct molecular tags to form specific microcircuits.

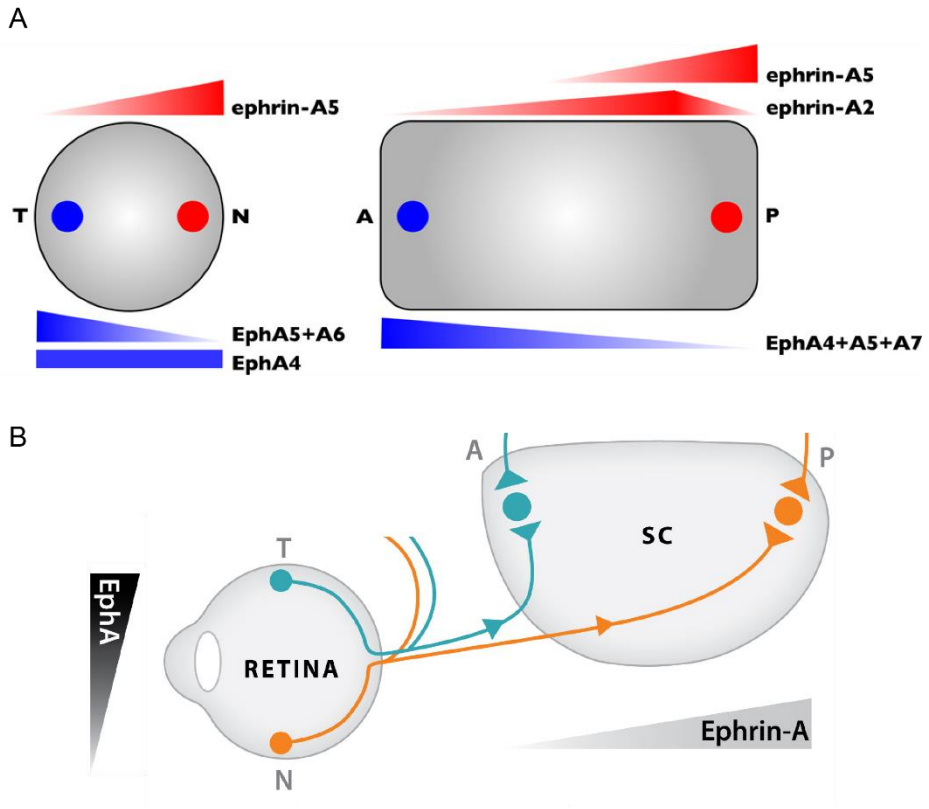


Figure 1-2. Expression patterns of Ephs and ephrins in the mouse retinocollicular system and the topographic projections of retinal axons. The temporal-nasal (T-N) axis of the retina maps along the anterior-posterior (A-P) axis of the SC. (A) Graded expression of EphA and ephrin-A. Expression patterns in each structure are illustrated by blue and red bars, respectively. EphAs are expressed at high levels temporally and anteriorly, while ephrin-As are expressed in a gradient with high levels nasally and posteriorly, which is complementary to EphA gradients. Adapted from Triplett & Feldheim, 2012. (B) Temporal axons, with high levels of EphA, terminate in anterior SC with low levels of ephrin-A, and nasal axons, with low receptor levels, terminate in posterior SC where ephrin-A levels are high. This is a simplified model of the Eph/Ephrin regulation of topographic mapping. Adapted from Cang & Feldheim, 2013.

Third, molecular diversity contributed by large families of related proteins with different recognition specificities provide unique identities to neurons and mediate self-avoidance (Lefebvre et al., 2012; Zipursky & Grueber, 2013). These include Dscam1 proteins in *Drosophila* (Schmucker et al., 2000) and clustered protocadherins in vertebrates (Kohmura et al., 1998; Wu and Maniatis, 1999). For example, the *Drosophila* Dscam1 gene contains three alternative exon

clusters for exon 4, 6 and 9, which have 12, 48 and 33 alternative exons, respectively. And they encode Ig domain 2, 3 and 7 in the extracellular region of the protein. Thus, isoform exclusive alternative splicing within each axon cluster generates 19,008 (12x 48x 33) isoforms with distinct ectodomains (Schmucker et al., 2000) (Figure 3A). Biochemical assays have shown that almost all isoforms exhibit isoform-specific homophilic binding property (Wojtowicz et al., 2007; Wojtowicz et al., 2004) (Figure 3B). Isoform specific homophilic binding mediates contact-dependent repulsion of sister neurites from the same neuron, a process called self-avoidance (Wu et al., 2012). Stochastic expression of different isoforms in different neurons endows each neuron a unique molecular identity (Miura et al., 2013). Thus, Dscam1 proteins mediate self-avoidance of sister neurites; neurites from the same neuron segregate from each other and neurites from different neurons can overlap, while neurites from Dscam1 mutant neurons fail to avoid each other (Matthews et al., 2007; Zhan et al., 2004) (Figure 3C, D). Reducing Dscam1 diversity in all neurons of the same class leads to segregation of their dendrites from each other (Hattori et al., 2007, 2009) (Figure 3D). However, it is unlikely that these protein families can mediate synaptic matching through a lock and key mechanism, as Dscam1 is largely expressed in a probabilistic manner (Miura et al., 2013), and protocadherins also appear to be expressed in this way.

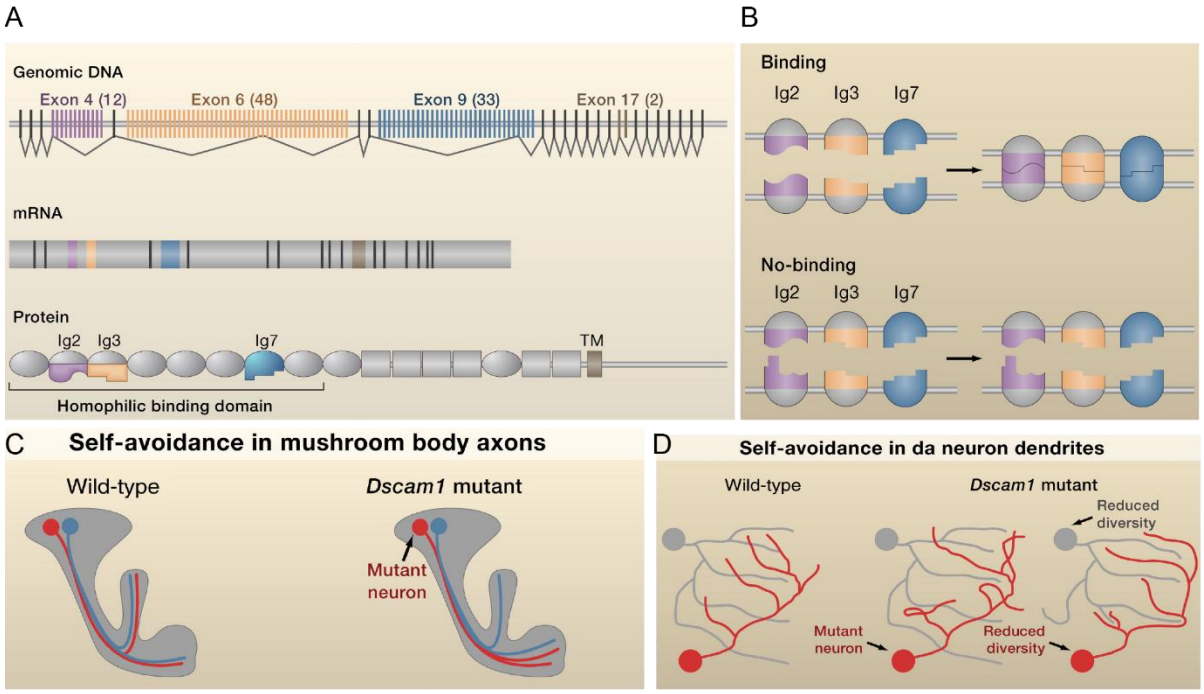


Figure 1-3. Schematics of Dscam1 gene, proteins and function. (A) The *Drosophila* Dscam1 gene contains alternative exon clusters that encode 12, 48 and 33 different variants for Ig2 (purple), Ig3 (orange), and Ig7 (blue), respectively, as well as two different variants for the transmembrane domain (brown). Splicing leads to the incorporation of only one alternative exon from each cluster. Thus, Dscam1 encodes 19,008 isoforms with different ectodomains. (B) Homophilic binding occurs between identical isoforms with the same variable Ig domains through charge and shape complementarity. Isoform pairs that contain two identical variable Ig domains, but differ at the third Ig domain bind poorly or do not bind to one another. This panel shows the mismatch for Ig2; the properties of the other variable domains are analogous to Ig2. (C) Dscam1 mediates self-avoidance in axons of mushroom body (MB) neurons. Each of the MB neurons (2 of 2500 are shown) extends a single axon that bifurcates and projects to two different directions. Each MB neuron expresses a unique combination of isoforms. Therefore, sister branches from the same neuron recognize each other through Dscam1 homophilic binding, which triggers a repulsive signal and subsequent segregation of axons to separate pathways. Axons of single mutant *Dscam1* MB neuron in an otherwise wild type background often fail to segregate. (D) *Dscam1* mediates self-avoidance in dendrites of da sensory neurons. Dendrites of the same neuron are spread out but can cross dendrites from different da neurons. Deletion of *Dscam1* from a single da neuron leads to fasciculation or crossing of its dendrites. Reducing *Dscam1* diversity in all da neurons leads to segregation of their dendrites from each other. Adapted from Zipursky & Sanes, 2010.

Cellular recognition and synaptic specificity

When I started my thesis project, progress has been made in identifying cell surface molecules regulating synaptic specificity in several systems.

In the *Drosophila* olfactory system, axons of ~50 classes of olfactory receptor neurons (ORNs) form one-to-one connections with dendrites of ~50 classes of projection neurons (PNs). Luo and colleagues demonstrate that evolutionarily conserved, epidermal growth factor (EGF)-repeat-containing transmembrane proteins, Teneurins play a role in synaptic matching in the fly olfactory system (Hong et al., 2012). The two Teneurin proteins, Ten-m and Ten-a, are each highly expressed in a unique, but partially overlapping, subset of matching ORNs and PNs in the developing antennal lobe (Figure 4A) and promote homophilic interactions *in vitro*. And Ten-m co-expression in non-partner PNs and ORNs promotes their ectopic connections (Hong et al., 2012). These data suggest that Teneurins match the presynaptic ORNs to postsynaptic PNs through homophilic attraction. However, PN-specific Ten-m knockdown experiments did not show any mismatching phenotype between Ten-m enriched ORNs and PNs. By contrast, a mismatching phenotype between Ten-m enriched ORN and PN was observed from PN-specific Ten-a knockdown, which fails to support that Ten-m is required for homophilic matching. Moreover, many loss of function data on Ten-m and Ten-a were based on RNAi knockdown in all ORNs or PNs. Further experiments to remove specific Teneurins from specific ORNs or PNs are necessary to address the direct matching model. In addition, there was also no developmental analysis of the loss of function phenotype, thus it is not clear how Teneurins function in matching synaptic partners during olfactory circuit assembly. Moreover, since Teneurins are only highly expressed

in subsets of ORNs and PNs, additional cell-surface molecules are also needed to determine connection specificity of the some 50 PN-ORN pairs.

In the vertebrate chick retina, lamina-specific subsets of retinal amacrine and bipolar neurons synapse on dendrites of retinal ganglion cells in specific sublaminae of the inner plexiform layer (IPL). Studies by Yamagata and Sanes show that members from subfamilies of Ig protein superfamily with homophilic binding properties, namely Sidekicks (Sdk), Dscams and Contactins are expressed by subsets of synaptic partners in chick retina. Neurites of these synaptic partners expressing different proteins targeted in a unique or a unique combination of sublaminae within the chick IPL (Yamagata & Sanes, 2008, 2012; Yamagata et al., 2002). Loss and gain of function experiments in the chick suggest that these proteins are required for sublaminae-specific targeting of neurons expressing them (Figure 4B). Importantly, mutually exclusive subsets of retinal cells accounting for approximately 60% of retinal cell population express Sdks and Dscams, and pre- and postsynaptic processes of cells that express the same Sdk or Dscam project to the same sublamina. This raised the possibility that related Ig superfamily proteins regulate layer-specific patterns of synaptic connections between different neurons in the chick retina. However, these studies did not identify the specific cell types expressing these proteins or assess whether these proteins are required for synaptic specificity between synaptic partners expressing them.

By contrast to the chick, Sdks and Dscams are not expressed in a layer-specific fashion in mouse retina, suggesting that they may function differently in mouse. Indeed, recent study by Sanes and colleagues showed that in the mouse retina, a specific retinal ganglion cell (RGC) type which detects moving objects from a visual scene that is also moving expresses Sdk2. The synaptic partner of this RGC type, an excitatory amacrine cell type, also expresses Sdk2. Both loss and gain

of function data are consistent with Sdk2-dependent homophilic interactions as necessary for the selectivity of their connectivity. Sdk2-specified synapses were shown to be essential for the visual responses of the RGC (Krishnaswamy et al., 2015) (Figure 4C). However, it remains unclear how Sdk2 functions during development to control the synaptic specificity between these two neurons. Also, further studies on other synaptic partners co-expressing Sdks and Dscams are necessary to establish a common molecular logic regulating synaptic specificity in the mouse retina. Despite this progress, given the complexity of the nervous system, little is known about the molecular mechanisms underlying synaptic specificity during neural circuit assembly.

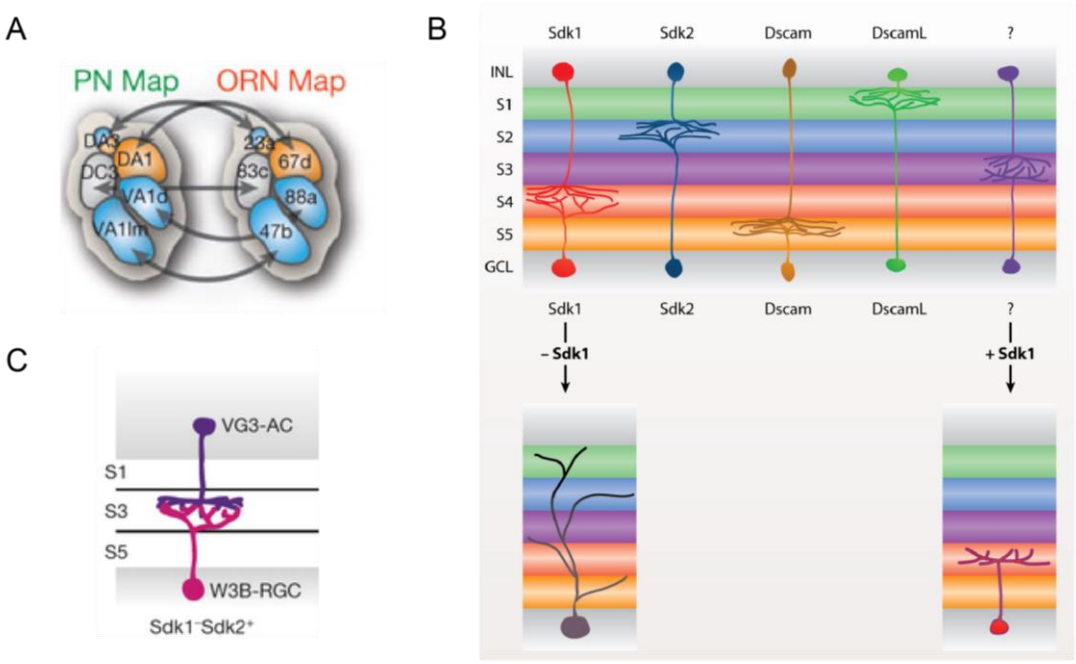


Figure 1-4. Examples of progress on identifying cell surface molecules regulating synaptic specificity. (A) Combined expression patterns of Teneurin proteins in PNs (left) and ORNs (right). Matching Ten-m or Ten-a expression levels between PNs and ORNs mediates the matching between synaptic partners. Blue: Ten-m high; orange: Ten-a high. Adapted from Hong et al., 2012. (B) Mutually exclusive subsets of synaptic partners in chick retina express homophilic adhesion proteins Sidekicks (Sdk) and Dscams. Processes of pre- and postsynaptic cells that express the

same Sdk or Dscam project to the same sublamina. Removal of Sdk or Dscam from cells normally expressing it disrupts restriction of their processes to the appropriate sublamina. Over-expression of Sdk or Dscam in a cell directs its processes to the layer enriched for the same Sdk or Dscam protein. Adapted from Sanes & Yamagata, 2009. (C) Arborization patterns of retinal cell types expressing Sdk2. Sdk2 is expressed in W3B-RGC which detects moving objects from a visual scene that is also moving, and in VG3-AC, an excitatory amacrine cell type. Both loss and gain of function data are consistent with that Sdk2-dependent homophilic interactions are necessary for the selectivity of their connection. Adapted from Krishnaswamy et al., 2015.

The Drosophila visual system as a model for studying synaptic specificity

The *Drosophila* visual system is well suited to uncovering the molecular mechanisms regulating synaptic specificity. The cellular organization and circuitry has been described in detail (Fischbach and Dittrich, 1989; Morante and Desplan, 2008). Remarkably, nearly complete maps of synaptic connections between different neurons in the first two neuropils in the fly visual system, the lamina and medulla, has been generated using serial electron microscopic (EM) reconstruction (Meinertzhagen & O'Neil, 1991; Rivera-Alba et al., 2011; Takemura et al., 2008, 2013, 2015). In addition, molecular markers that specifically label many different cell types both in adult and during development are available (Jenett et al., 2012; Kvon et al., 2014), genetic tools that facilitate gain and loss of function studies at the level of single identified cells in developing and adult tissue have been well established (Lee and Luo, 1999; Venken and Bellen, 2014), and an extensive protein interaction network of extracellular domains of cell surface proteins has been assembled (Özkan et al., 2013).

The *Drosophila* visual system is modular. It comprises the retina and four neuropils, the lamina, medulla, lobula and lobula plate. There are only neuronal processes within each neuropil; the neuronal cell bodies lie outside of neuropils (referred to as cortical regions). Light is detected on the retina, and visual information is processed by the four neuropils. The retina, lamina and

medulla each contains precisely registered sets of ~750 units, called ommatidia, cartridges and columns, respectively. Connections of neurons from a single ommatidium, and lamina neurons from a single cartridge are largely restricted to a single column in medulla, thus retaining strict retinotopic mapping of visual information in higher visual processing regions. Each medulla column is divided into ten layers, designated M1-M10 (Figure 5A-B). And the lobula and lobula plate are also divided into multiple layers. Thus, in a broad sense, columns process information from different points in space and layers process different types of visual information (e.g., ON versus OFF responses).

The fly visual system possesses extraordinary neuronal diversity. There are different classes of neurons and these in turn fall into many subclasses. Each ommatidium in the retina is composed of eight photoreceptor neurons (R cells), which are functionally divided into three subclasses. R1-R6 neurons express the same opsin responding to a broad spectrum of wavelengths and are motion detectors. R7 and R8 express UV and blue/green sensitive opsins, respectively, thus detecting chromatic information. Axons of R1-R6 project to and form synapses within each lamina cartridge (Figure 5B). The lamina contains 12 different subclasses of neurons. These include five subclasses of lamina monopolar neurons (L1-L5), two subclasses of wide-field neurons (Lawf1, Lawf2), three subclasses of putative feedback neurons from the medulla (T1, C2 and C3), lamina intrinsic neuron and lamina tangential neuron (Tuthill et al., 2013).

The medulla is much more complex than the lamina. Axons of R7, R8 and lamina neurons terminate into distinct layers, and often arborize in others in the outer region of medulla (M1-M6 layers), where they synapse on interneurons and transmedullary neurons (Figure 5B) (Takemura et al., 2008, 2013, 2015). Connections of interneurons, such as Mi, Dm and Pm neurons remain

confined in the medulla and promote processing within and between layers (Figure 5B). By contrast, the two classes of transmedullary neurons send axons to the lobula complex: Tm neurons to the lobula and TmY neurons to both the lobula and lobula plate (Figure 5B). The medulla is analogous to the inner plexiform layer in the vertebrate visual system; Mi, Dm and Pm are equivalent to amacrine cells, while Tm and TmY are equivalent to retinal ganglion cells (Sanes and Zipursky, 2010). There are 20~30 morphologically distinct subclasses of medulla interneurons (Mi, Dm and Pm), Tm and TmY neurons, respectively (Fischbach & Dittrich, 1989; Morante & Desplan, 2008; A. Nern, personal communication). Thus, there are over 100 different subclasses of neurons forming synapses in each medulla column, which belong to a few classes based primarily on shared morphological features.

EM reconstruction studies have provided a detailed description of synaptic connections in the lamina (Figure 5D) (Meinertzhagen and O'Neil, 1991; Rivera-Alba et al., 2011). Each lamina cartridge consists of synaptic connections between axon terminals of R1-R6 and the dendrites of lamina monopolar neurons, as well as connections between other neurons. In all cases these connections are multiple contact synapses, with a single presynaptic terminal and multiple postsynaptic elements. The predominant synapses formed by R1-R6 are tetrad synapses, in which single presynaptic R1-R6 terminals juxtapose four postsynaptic elements. Two postsynaptic elements, from L1 and L2, are invariant, with the remaining two derived from amacrine cells, L3 cells, or glial cells (Figure 5C). Behavior studies have shown that L1 and L2 are the key components in ON and OFF pathways in motion detection, respectively (Joesch et al., 2010). Flies with silenced L1 and L2 are blind to motion (Tuthill et al., 2013). L3 and L4 are also important components in OFF pathway (Meier et al., 2014; Serbe et al., 2016). Despite having largely similar synaptic connections with R1-R6, L1 and L2 exhibit distinct synaptic specificity with L4.

Although both L1 and L2 have similar area of contact with L4 within the lamina, L4 only makes synapses with L2 but not with L1 (Rivera-Alba et al., 2011). The tetrad synapses made by R1-R6 and the specific synapses between L2 and L4 provide promising systems to identify specific cell recognition molecules regulating synaptic specificity within lamina.

Some progress has been made in understanding the synaptic specificity in lamina circuitry. For the tetrad synapses formed by R1-R6 and four postsynaptic elements, Millard and colleagues show that *Drosophila Dscam1* and *Dscam2*, genes encoding proteins mediating repulsion upon isoform-specific homophilic binding, are required to ensure the invariable combination of L1 and L2 at postsynaptic sites. In *Dscam1* and *Dscam2* double mutants, postsynaptic elements from the same L1 or L2 were seen within the same tetrad synapse, indicating the strict pairing is lost. Thus, removing these two repulsive proteins allows elements from the same cell (L1 or L2) to incorporate into the same postsynaptic tetrad, altering the specificity of tetrad synapses (Millard et al., 2010). It is still not clear what mechanisms allow L4 to distinguish between L1 and L2 dendrites during synapse formation, although recent data indicate that, Kirre, a member of the evolutionary conserved IRM proteins, is necessary for synapses between L2 and L4 (Lüthy et al., 2014).

The synaptic connectivity in the medulla has been determined using serial section electron microscopic reconstruction (Figure 5D) (Takemura et al., 2008, 2013, 2015). These studies show that the synapses in medulla are also predominantly multiple contact synapses. Axons of R7, R8 and lamina neurons terminate and often arborize into distinct layers in the outer medulla, where they selectively make synapses with subsets of interneurons and transmedullary neurons within these layers. The postsynaptic neurons of R7, R8 and lamina neurons also arborize and form synapses within discrete layers of the medulla, but some of their arbors extend beyond column

boundaries. The EM reconstruction data show that within a layer, neurons form synapses with multiple neuronal types (Takemura et al., 2013, 2015), but these represent only a subset of neurons with processes in the layer. For instance, R7 specifically terminate at M6 layer. More than 20 different cell types elaborate processes in the M6 layer and many of them directly contact R7 axons. However, R7 axons only selectively form synapses with 5 target cell types and the vast majority of the synaptic connections are with one dominant partner Dm8. Remarkably, the most recent reconstruction data of seven adjacent columns reveal that synaptic connections made by different neurons are highly specific and reproducible across columns (Takemura et al., 2015). Although some progress has been made in identifying genes regulating layer-specific targeting in the medulla (Hadjieconomou et al., 2011), genes controlling synaptic specificity within layers have not been identified.

The remarkable reconstruction data set provides a unique opportunity to search for recognition molecules regulating synaptic specificity. The goal in my thesis research was to uncover a common molecular logic underlying synaptic specificity during circuit assembly in the medulla.

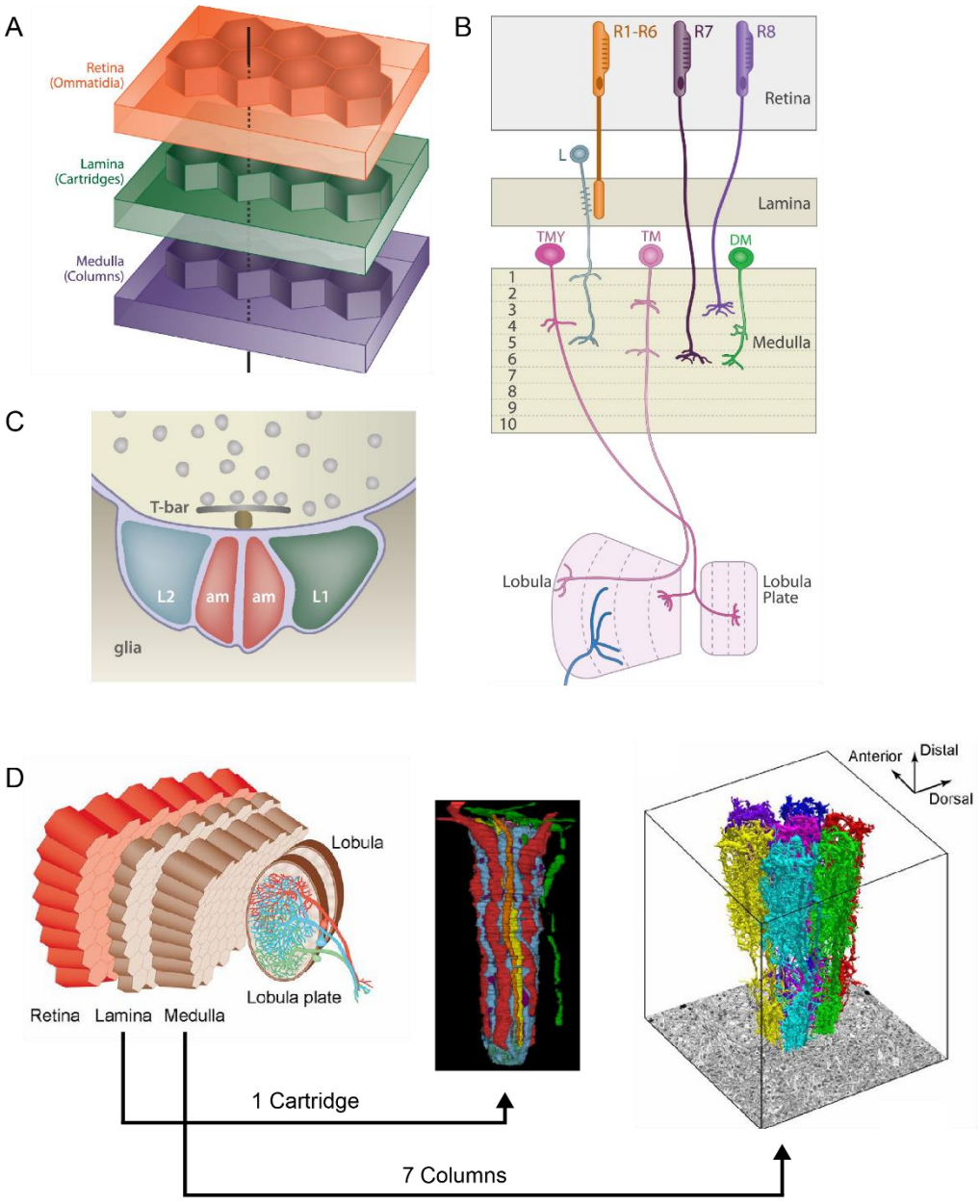


Figure 1-5. The *Drosophila* visual system. (A) In *Drosophila*, retinal ommatidia overlie laminar cartridges, which in turn overlie medullary columns. (B) *Drosophila* visual system, showing retina, lamina, medulla, lobula and lobula plate. A few cell types are shown. (C) Tetrad synapse in *Drosophila* lamina, showing synapses of R cell axon on processes of L1, L2, and amacrine (am) cell dendrites. The presynaptic specialization in *Drosophila* is T bar. Schematic is simplified to show all postsynaptic elements in the same plane. In some cases, one amacrine cell process is replaced with a process from L3. A-C are adapted from Sanes & Zipursky, 2010. (D) Electron microscopic studies have reconstructed one lamina cartridge and seven medulla columns,

revealing synaptic connections within lamina and medulla. Adapted from Rivera-Alba et al., 2011 and Takemura et al., 2015.

References

Brose, K., Bland, K.S., Wang, K.H., Arnott, D., Henzel, W., Goodman, C.S., Tessier-Lavigne, M., and Kidd, T. (1999). Slit Proteins Bind Robo Receptors and Have an Evolutionarily Conserved Role in Repulsive Axon Guidance. Cell *96*, 795–806.

Cang, J., and Feldheim, D. a (2013). Developmental mechanisms of topographic map formation and alignment. Annu. Rev. Neurosci. *36*, 51–77.

Carreres, M.I., Escalante, A., Murillo, B., Chauvin, G., Gaspar, P., Vegar, C., and Herrera, E. (2011). Transcription Factor Foxd1 Is Required for the Specification of the Temporal Retina in Mammals. J. Neurosci. *31*, 5673–5681.

Dickson, B.J. (2002). Molecular mechanisms of axon guidance. Science (80-.). *298*, 1959–1964.

Egea, J., and Klein, R. (2007). Bidirectional Eph–ephrin signaling during axon guidance. Trends Cell Biol. *17*, 230–238.

Feldheim, D.A., Kim, Y.-I., Bergemann, A.D., Frisén, J., Barbacid, M., and Flanagan, J.G. (2000). Genetic Analysis of Ephrin-A2 and Ephrin-A5 Shows Their Requirement in Multiple Aspects of Retinocollicular Mapping. Neuron *25*, 563–574.

Feller, M.B. (2009). Retinal waves are likely to instruct the formation of eye-specific retinogeniculate projections. Neural Dev. *4*, 24.

Fischbach, K.-F., and Dittrich, a P. (1989). The optic lobe of Drosophila melanogaster. I: A.

Golgi analysis of wild-type structure. Cell Tissue Res *258*, 441–475.

Frisén, J., Yates, P.A., McLaughlin, T., Friedman, G.C., O'Leary, D.D.., and Barbacid, M. (1998). Ephrin-A5 (AL-1/RAGS) Is Essential for Proper Retinal Axon Guidance and Topographic Mapping in the Mammalian Visual System. Neuron *20*, 235–243.

Gale, N.W., Holland, S.J., Valenzuela, D.M., Flenniken, A., Pan, L., Ryan, T.E., Henkemeyer, M., Strebhardt, K., Hirai, H., Wilkinson, D.G., et al. (1996). Eph Receptors and Ligands Comprise Two Major Specificity Subclasses and Are Reciprocally Compartmentalized during Embryogenesis. Neuron *17*, 9–19.

Hadjieconomou, D., Timofeev, K., and Salecker, I. (2011). A step-by-step guide to visual circuit assembly in Drosophila. Curr. Opin. Neurobiol. *21*, 76–84.

Hattori, D., Demir, E., Kim, H.W., Viragh, E., Zipursky, S.L., and Dickson, B.J. (2007). Dscam diversity is essential for neuronal wiring and self-recognition. Nature *449*, 223–227.

Hattori, D., Chen, Y., Matthews, B.J., Salwinski, L., Sabatti, C., Grueber, W.B., and Zipursky, S.L. (2009). Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms. Nature *461*, 644–648.

Hiesinger, P.R., Zhai, R.G., Zhou, Y., Koh, T.-W., Mehta, S.Q., Schulze, K.L., Cao, Y., Verstreken, P., Clandinin, T.R., Fischbach, K.-F., et al. (2006). Activity-Independent Prespecification of Synaptic Partners in the Visual Map of Drosophila.

Himanen, J.-P., Chumley, M.J., Lackmann, M., Li, C., Barton, W.A., Jeffrey, P.D., Vearing, C., Geleick, D., Feldheim, D.A., Boyd, A.W., et al. (2004). Repelling class discrimination: ephrin-A5 binds to and activates EphB2 receptor signaling. Nat. Neurosci. *7*, 501–509.

Hong, W., Mosca, T.J., and Luo, L. (2012). Teneurins instruct synaptic partner matching in an olfactory map. Nature *484*, 201–207.

Jenett, A., Rubin, G.M., Ngo, T.T.B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., et al. (2012). A GAL4-Driver Line Resource for Drosophila Neurobiology. Cell Rep. *2*, 991–1001.

Joesch, M., Schnell, B., Raghu, S.V., Reiff, D.F., and Borst, A. (2010). ON and OFF pathways in Drosophila motion vision. Nature *468*, 300–304.

Kennedy, T.E., Serafini, T., de la Torre, J., and Tessier-Lavigne, M. (1994). Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell *78*, 425–435.

Kidd, T., Brose, K., Mitchell, K.J., Fetter, R.D., Tessier-Lavigne, M., Goodman, C.S., and Tear, G. (1998). Roundabout Controls Axon Crossing of the CNS Midline and Defines a Novel Subfamily of Evolutionarily Conserved Guidance Receptors. Cell *92*, 205–215.

Kohmura, N., Senzaki, K., Hamada, S., Kai, N., Yasuda, R., Watanabe, M., Ishii, H., Yasuda, M., Mishina, M., and Yagi, T. (1998). Diversity Revealed by a Novel Family of Cadherins Expressed in Neurons at a Synaptic Complex. Neuron *20*, 1137–1151.

Krishnaswamy, A., Yamagata, M., Duan, X., Hong, Y.K., and Sanes, J.R. (2015). Sidekick 2 directs formation of a retinal circuit that detects differential motion. Nature *524*, 466–470.

Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature.

Langley, J.N. (1895). Note on Regeneration of Prae-Ganglionic Fibres of the Sympathetic. J. Physiol. *18*, 280–284.

Lee, T., and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron *22*, 451–461.

Lefebvre, J.L., Kostadinov, D., Chen, W. V, Maniatis, T., and Sanes, J.R. (2012). Protocadherins mediate dendritic self-avoidance in the mammalian nervous system.

Lüthy, K., Ahrens, B., Rawal, S., Lu, Z., Tarnogorska, D., Meinertzhagen, I.A., and Fischbach, K.-F. (2014). The irre Cell Recognition Module (IRM) Protein Kirre Is Required to Form the Reciprocal Synaptic Network of L4 Neurons in the *Drosophila* Lamina. J. Neurogenet. *28*, 291–301.

Matthews, B.J., Kim, M.E., Flanagan, J.J., Hattori, D., Clemens, J.C., Zipursky, S.L., and Grueber, W.B. (2007). Dendrite Self-Avoidance Is Controlled by Dscam. Cell *129*, 593–604.

Meier, M., Serbe, E., Maisak, M.S., Haag, J., Dickson, B.J., and Borst, A. (2014). Neural Circuit Components of the Drosophila OFF Motion Vision Pathway. Curr. Biol. *24*, 385–392.

Meinertzhagen, I.A., and O'Neil, S.D. (1991). Synaptic organization of columnar elements in the lamina of the wild type inDrosophila melanogaster. J. Comp. Neurol. *305*, 232–263.

Millard, S.S., Lu, Z., Zipursky, S.L., and Meinertzhagen, I.A. (2010). Drosophila Dscam Proteins Regulate Postsynaptic Specificity at Multiple-Contact Synapses.

Miura, S.K., Martins, A., Zhang, K.X., Graveley, B.R., and Zipursky, S.L. (2013). Probabilistic splicing of Dscam1 establishes identity at the level of single neurons. Cell *155*, 1166–1177.

Morante, J., and Desplan, C. (2008). The Color-Vision Circuit in the Medulla of Drosophila.

Curr. Biol. *18*, 553–565.

Özkan, E., Carrillo, R. a., Eastman, C.L., Weiszmann, R., Waghray, D., Johnson, K.G., Zinn, K., Celniker, S.E., and Garcia, K.C. (2013). An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. Cell *154*.

Penn, A.A. (1998). Competition in Retinogeniculate Patterning Driven by Spontaneous Activity. Science (80-.). *279*, 2108–2112.

Pfeiffenberger, C., Yamada, J., and Feldheim, D.A. (2006). Ephrin-As and Patterned Retinal Activity Act Together in the Development of Topographic Maps in the Primary Visual System. J. Neurosci. *26*, 12873–12884.

Rivera-Alba, M., Vitaladevuni, S.N., Mishchenko, Y., Lu, Z., Takemura, S., Scheffer, L., Meinertzhagen, I.A., Chklovskii, D.B., and de Polavieja, G.G. (2011). Wiring Economy and Volume Exclusion Determine Neuronal Placement in the Drosophila Brain.

Sanes, J.R., and Yamagata, M. (2009). Many Paths to Synaptic Specificity. Annu. Rev. Cell Dev. Biol. *25*, 161–195.

Sanes, J.R., and Zipursky, S.L. (2010). Design Principles of Insect and Vertebrate Visual Systems. Neuron *66*, 15–36.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C. a, Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell *101*, 671–684.

Serafini, T., Kennedy, T.E., Galko, M.J., Mirzayan, C., Jessell, T.M., and Tessier-Lavigne, M. (1994). The netrins define a family of axon outgrowth-promoting proteins homologous to C.

elegans UNC-6. Cell *78*, 409–424.

Serbe, E., Meier, M., Leonhardt, A., and Borst, A. (2016). Comprehensive Characterization of the Major Presynaptic Elements to the Drosophila OFF Motion Detector. Neuron *89*, 829–841.

Sperry, R.W. (1943). Visuomotor coordination in the newt (triturus viridescens) after regeneration of the optic nerve. J. Comp. Neurol. *79*, 33–55.

Sretavan, D.W., Shatz, C.J., and Stryker, M.P. (1988). Modification of retinal ganglion cell axon morphology by prenatal infusion of tetrodotoxin. Nature *336*, 468–471.

Takemura, S., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P.K., Katz, W.T., Olbris, D.J., Plaza, S.M., Winston, P., et al. (2013). A visual motion detection circuit suggested by Drosophila connectomics. Nature *500*, 175–181.

Takemura, S.Y., Lu, Z., and Meinertzhagen, I. a. (2008). Synaptic circuits of the Drosophila optic lobe: The input terminals to the medulla. J. Comp. Neurol. *509*, 493–513.

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., et al. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. Proc. Natl. Acad. Sci. U. S. A. *112*, 13711–13716.

Triplett, J.W., and Feldheim, D. a. (2012). Eph and ephrin signaling in the formation of topographic maps. Semin. Cell Dev. Biol. *23*, 7–15.

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

Venken, K.J.T., and Bellen, H.J. (2014). Chemical mutagens, transposons, and transgenes to interrogate gene function in Drosophila melanogaster. Methods *68*, 15–28.

Ward, A., Hong, W., Favaloro, V., and Luo, L. (2015). Toll Receptors Instruct Axon and Dendrite Targeting and Participate in Synaptic Partner Matching in a Drosophila Olfactory Circuit. Neuron *85*, 1013–1028.

Wojtowicz, W.M., Flanagan, J.J., Millard, S.S., Zipursky, S.L., and Clemens, J.C. (2004). Alternative Splicing of Drosophila Dscam Generates Axon Guidance Receptors that Exhibit Isoform-Specific Homophilic Binding. Cell *118*, 619–633.

Wojtowicz, W.M., Wu, W., Andre, I., Qian, B., Baker, D., and Zipursky, S.L. (2007). A Vast Repertoire of Dscam Binding Specificities Arises from Modular Interactions of Variable Ig Domains. Cell *130*, 1134–1145.

Wu, Q., and Maniatis, T. (1999). A Striking Organization of a Large Family of Human Neural Cadherin-like Cell Adhesion Genes. Cell *97*, 779–790.

Wu, W., Ahlsen, G., Baker, D., Shapiro, L., and Zipursky, S.L. (2012). Complementary Chimeric Isoforms Reveal Dscam1 Binding Specificity In Vivo.

Yamagata, M., and Sanes, J.R. (2008). Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. Nature *451*, 465–469.

Yamagata, M., and Sanes, J.R. (2012). Expanding the Ig superfamily code for laminar specificity in retina: expression and role of contactins. J. Neurosci. *32*, 14402–14414.

Yamagata, M., Weiner, J.A., and Sanes, J.R. (2002). Sidekicks: Synaptic adhesion molecules that promote lamina-specific connectivity in the retina. Cell *110*, 649–660.

Zhan, X.-L., Clemens, J.C., Neves, G., Hattori, D., Flanagan, J.J., Hummel, T., Vasconcelos, M.L., Chess, A., and Zipursky, S.L. (2004). Analysis of Dscam Diversity in Regulating Axon

Guidance in Drosophila Mushroom Bodies. Neuron *43*, 673–686.

Zipursky, S.L., and Grueber, W.B. (2013). The molecular basis of self-avoidance. Annu. Rev. Neurosci. *36*, 547–568.

Zipursky, S.L., and Sanes, J.R. (2010). Chemoaffinity revisited: Dscams, protocadherins, and neural circuit assembly. Cell *143*, 343–353.

Zou, Y., Stoeckli, E., Chen, H., and Tessier-Lavigne, M. (2000). Squeezing Axons Out of the Gray Matter: A Role for Slit and Semaphorin Proteins from Midline and Ventral Spinal Cord. Cell *102*, 363–375.

Chapter 2 Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila

This chapter is adapted from the following publication with minor changes.

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, 1756-1769.

Summary

Information processing relies on precise patterns of synapses between neurons. The cellular recognition mechanisms regulating this specificity are poorly understood. In the medulla of the *Drosophila* visual system, different neurons form synaptic connections in different layers. Here, we sought to identify candidate cell recognition molecules underlying this specificity. Using RNA sequencing (RNA seq), we show that neurons with different synaptic specificities express unique combinations of mRNAs encoding hundreds of cell surface and secreted proteins. Using RNA seq and protein tagging, we demonstrate that 21 paralogs of the Dpr family, a subclass of Immunoglobulin (Ig)-domain containing proteins, are expressed in unique combinations in homologous neurons with different layer-specific synaptic connections. Dpr interacting proteins (DIPs), comprising nine paralogs of another subclass of Ig-containing proteins, are expressed in a

complementary layer-specific fashion in a subset of synaptic partners. We propose that pairs of Dpr/DIP paralogs contribute to layer-specific patterns of synaptic connectivity.

Introduction

Neural circuits typically comprise many different neurons linked in precise ways by synaptic connections. How neurites discriminate between one another during circuit assembly remains a central issue in neuroscience. Through regeneration studies in vertebrates, Langley (Langley, 1895) and Sperry (Sperry, 1963) proposed that molecular differences between neurons determine the specificity of synaptic connections. In its simplest formulation Sperry's chemoaffinity hypothesis (Sperry, 1963) suggested that a lock and key mechanism mediates recognition between synaptic partners.

Over the past 30 years, biochemical and genetic approaches have led to the identification of the cell recognition molecules and the intercellular signaling pathways regulating the patterning of axons and dendrites. From these studies three general molecular strategies underlying circuit assembly have emerged. First, combinatorial use of a limited set of conserved cell surface and secreted molecules regulates axon guidance in many different regions of the developing invertebrate and vertebrate nervous systems. These include netrins, slits, semaphorins, and various cell adhesion molecules, such as cadherins and immunoglobulin superfamily proteins (O'Donnell et al., 2009). Second, gradients of cell surface proteins, notably Ephs and Ephrins as well as Wnts, play a crucial role in the establishment of topographic maps (Cang and Feldheim, 2013; Schmitt et al., 2006; Triplett and Feldheim, 2012) a widespread organizational principle in the vertebrate brain (Cang and Feldheim, 2013). And third, molecular diversity contributed by large families of related proteins with different recognition specificities impart unique identities to neurons and

mediate self-avoidance (repulsion between neurites of the same cell (Zipursky and Grueber, 2013; Lefebvre et al., 2012)). These include Dscam1 proteins in *Drosophila* (Schmucker et al., 2000) and clustered protocadherins in vertebrates (Kohmura et al., 1998; Wu and Maniatis, 1999). The molecular diversity of both Dscam1 and protocadherins coupled with their exquisite isoform-specific homophilic binding specificities raised the possibility that they could directly specify patterns of synaptic specificity through a lock and key mechanism. As Dscam1 is largely, if not exclusively expressed in a probabilistic manner (Miura et al., 2013), and protocadherins also appear to be expressed in this way, it is unlikely that these protein families mediate synaptic matching.

Important progress has been made in identifying cell surface molecules regulating synaptic specificity, including Syg1 and Syg2 in the worm (Shen and Bargmann, 2003; Shen et al., 2004), Toll and Teneurin proteins in the fly olfactory system (Hong et al., 2012; Ward et al., 2015) and Sidekick proteins in the mouse retina (Krishnaswamy et al., 2015). Studies by Sanes and colleagues raised the possibility that related Ig superfamily proteins regulate layer-specific patterns of synaptic connections between different neurons in the chick retina (Yamagata and Sanes, 2002, 2008, 2012) (see *Discussion*). As a step towards identifying a common molecular logic underlying synaptic specificity, we sought to identify families of cell surface proteins expressed in a cell-type enriched fashion in closely related neurons with different patterns of synaptic specificity. Here, we set out to do this using RNA seq and molecular genetic approaches in *Drosophila*.

The *Drosophila* visual system is well suited to uncovering the molecular recognition mechanisms regulating synaptic specificity. The cellular organization and circuitry has been described in detail

(Fischbach and Dittrich, 1989; Morante and Desplan, 2008) including serial EM reconstruction to reveal connections between neurons (Takemura et al., 2015, 2013, 2008). In addition, molecular markers for many cell types are readily available (Jenett et al., 2012; Kvon et al., 2014), genetic tools facilitate gain and loss of function studies at the level of single identified cells in developing and adult tissue (Lee and Luo, 1999; Venken and Bellen, 2014), and an extensive protein interaction network of extracellular proteins has been assembled (Özkan et al., 2013).

In this paper, we focus on the medulla region of the fly visual system. It comprises columns and layers (Figure 1A-C). In a broad sense, columns process information from different points in space and layers process different types of visual information (e.g. ON vs OFF responses). The cell bodies of medulla neurons lie outside the neuropil and synaptic specificity is elaborated within a dense meshwork of axonal and dendritic processes. There are over 100 different types of neurons forming synapses in the medulla. These neurons fall into a few general categories based primarily on their morphology and location of their arbors (Fischbach and Dittrich, 1989; Morante and Desplan, 2008; Takemura et al., 2013)(Figure 1A-C). In a landmark study, the synaptic connectivity between neurons in the medulla was determined using serial section electron microscopic reconstruction (Takemura et al., 2013). The shaded electron micrographic sections through the adult column shown in Figure 1D and E are included to emphasize the complexity of the neuropil in one medulla column comprising the processes of on the order of 100 different neuronal cell types (Nern, personal communication) (Figure 1D, E). These patterns of synaptic connections are complex, specific and reproducible (Takemura et al., 2015). In addition, these studies revealed that within a layer, neurons form synapses with multiple neuronal types (Takemura et al., 2013, 2015), but these represent only a subset of neurons with processes in the layer. Although some progress has been made in identifying genes regulating layer-specific

targeting (Hadjieconomou et al., 2011), genes controlling synaptic specificity within layers have not been identified.

In the work described here, we set out to identify proteins regulating synaptic specificity using RNA seq to determine the transcriptome of developing R7 and R8 photoreceptor neurons, and five lamina monopolar neurons, L1-L5 (Figure 1A, B). The axon of each type of neuron elaborates a unique morphology, including layer-specific branches, and forms characteristic patterns of synaptic connections (Fischbach and Dittrich, 1989; Takemura et al., 2013, 2008). In most cases, synapses formed by each neuron type occur in the layer in which the axon terminates or in which interstitial branches arborize. We show that each cell type expresses mRNAs from hundreds of genes encoding cell surface proteins and unique combinations of them. Using a protein interaction database (Özkan et al., 2013), *Minos* mediated integration cassette (MiMIC)-based protein traps to visualize protein expression (Nagarkar-Jaiswal et al., 2015; Venken et al., 2011) and genetic markers for identified medulla cell types, we present evidence that two families of heterophilic recognition molecules of the Ig superfamily, the Dprs and DIPs, are expressed in discrete subsets of synaptic partners within layers. These families are promising candidates for regulating synaptic specificity within the developing *Drosophila* central nervous system.

Results

Purification of developing neurons using Fluorescence Activated Cell Sorting

As a first step towards identifying cell surface and secreted molecules as candidates involved in cellular recognition events regulating synaptic specificity through RNA seq, we developed methods to purify seven neuronal cell types (R7, R8 and L1-L5 neurons) with different layer and synaptic specificities. These neurons were isolated at 40% after puparium formation (APF), just prior to (i.e. R7 and R8) or during early stages of synapse formation (i.e. L1-L5) ((Chen et al., 2014); Pecot and Zipursky, unpublished observations). To purify each cell type at this stage in development, we used a dual labeling approach. We generated transgenes expressing tandem tomato in all cells in the retina or in all cells in the lamina, and combined these with a separate GFP marker expressed selectively in a specific cell type (e.g. R8 in the retina or L3 in the lamina, Figure 1F-I). Each cell type was isolated in a highly purified form as assessed using qPCR for several diagnostic markers (data not shown) and post-hoc analysis of RNA seq data (Figure 2C).

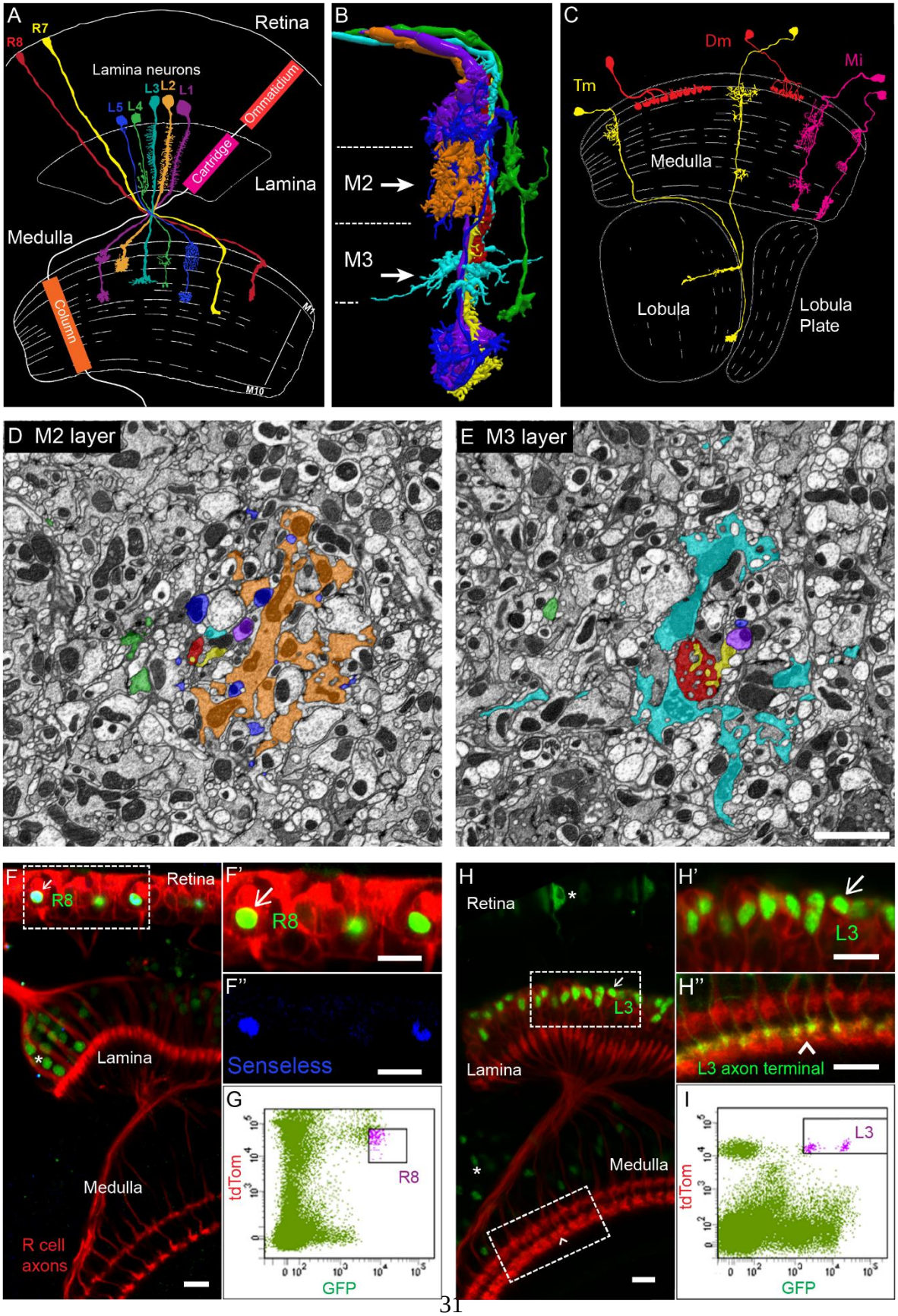

A
R7
R8
Retina
Lamina neurons
L3 L2 L1
L5 L4
Ommatidium
Cartridge
Lamina
Medulla
Column
M1
M10
B
M2
M3
C
Dm
Mi
Tm
Medulla
Lobula
Lobula Plate
D M2 layer
E M3 layer
F
Retina
R8
Lamina
Medulla
R cell axons
F'
R8
F''
Senseless
G
tdTom
R8
GFP
H
Retina
L3
Lamina
Medulla
H'
L3
H''
L3 axon terminal
I
tdTom
L3
GFP

Figure 2-1. FACS isolation of developing neurons with different synaptic specificities.

(A) Schematic of the adult morphologies of R7 and R8 photoreceptors and lamina neurons L1-L5. The visual system comprises topographically matched modules (i.e. ommatidia, cartridges and columns). (B) Axons of R7, R8, and L1-L5 are shown together within a single column as determined from serial EM reconstruction. The color of different neurons is the same as in (A). Dotted lines represent layer boundaries. Arrowheads indicate the plane of section shown in electron micrographs in D and E. (C) Examples of three general classes of medulla neurons that are synaptic targets for R7, R8, and L1-L5 neurons (two examples are shown for each class). Within each class, there are many cell types that display similar morphologies and branch in different layers. (D, E) Cross-sections through medulla columns reconstructed by serial EM within the M2 and M3 layers (see arrows in panel B). Axons are colored as in A and B. Each column contains processes from over 100 different neuronal cell types (Nern, A. personal communication). Scale bar, 2 µ. (F, G) Isolation of R8 neurons at 40 hrs APF using FACS. Only R8 neurons express both retinal-specific TdTom and R8-specific GFP. Senseless is an R8-specific transcription factor. Scale bars, 10 µm. (H, I) Isolation of L3 neurons at 40 hrs APF using FACS. Only L3 neurons express lamina-specific TdTom and L3-specific GFP. Scale bars, 10 µm. In F and H, arrows indicate double-labeled cells in developing tissue and the asterisks indicate single GFP-labeled cells of different cell types (i.e. contaminants). See *Experimental Procedures* for purification protocols for other cells and additional details. Panel A and C are adapted from Fischbach and Dittrich, 1989. Panel B, D, and E courtesy of Takemura, S., Meinertzhagen. I. A, and Scheffer, L. (JRC/HHMI).

Identification of cell type-specific differences in gene expression

To obtain the transcriptomes of purified cell populations, we isolated total RNA, linearly amplified polyA-mRNA using T7 polymerase and generated cDNA libraries that were analyzed in a single lane by 50bp paired-end sequencing on an Illumina HiSeq 2000 platform. At least two independent biological replicate libraries were sequenced for each cell type. We obtained between 221 and 441 million reads from each library, with a percentage of uniquely mapped reads ranging between 33% and 74%. Of these, 19 % to 32% were intergenic and 60% to 75% mapped to exons. A small fraction of reads mapped to intronic regions. (Table S1).

The correlation in the distribution of normalized raw reads between biological replicates for each cell type was high (Figure 2A) and ranged from 0.97-1 for the L4 and R7 libraries, respectively. The correlation coefficients between libraries of different cell types ranged from 0.87 for R7 vs L4, to 0.97 for R7 vs R8. Pairs of neurons from either the retina or the lamina were more closely related to each other than any given retinal to lamina neuron pair. The L1 and L2 pair, both required for the optomotor response (Borst, 2014), are more closely related than other pairs of lamina neurons. These data are consistent with principle component analysis in which R7/R8 are distinct from L1-L5 and that the L1 and L2 pair, as well as L4 and L5 pair, are more closely related to each other than to L3 neurons (Figure 2B).

To assess whether differences revealed through RNA seq reliably reflect differences in expression between cell types, we compared the RPKM (the Reads Per Kilobase of a specific mRNA Per Million reads) values for seven transcripts expressed specifically in each of the seven neuronal cell types as determined by immunohistology. There was an excellent correlation between antibody staining and RPKM values (Figure 2C). Thus, RNA seq provides a reliable method to identify transcripts differentially expressed between these neurons.

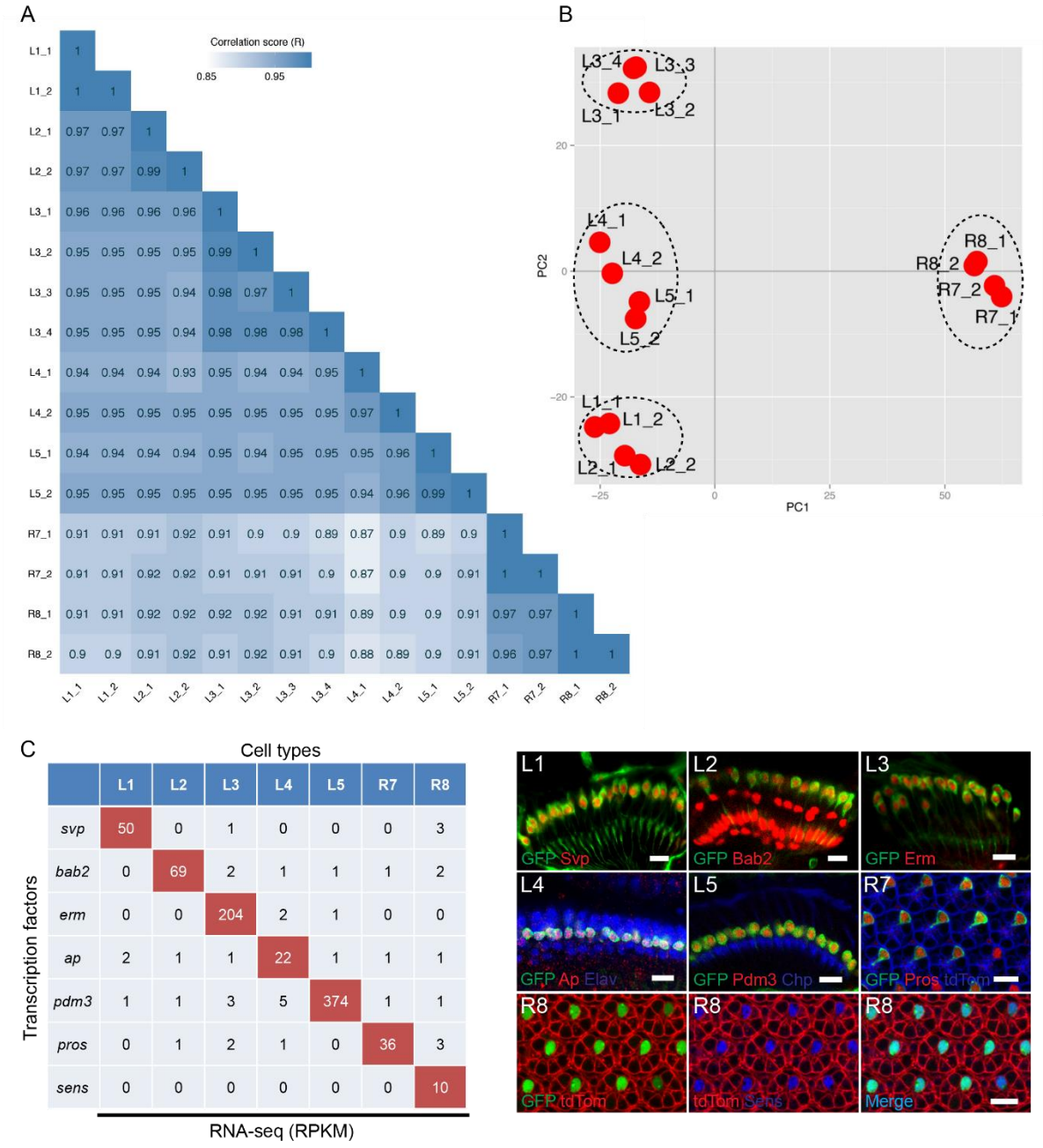


Figure 2-2. RNA seq of visual system neurons.

(A) Correlograms showing the correlation score matrix across all libraries of all seven cell types (R, Pearson correlation coefficient). (B) Principal component analysis plot of the RNA seq data for the indicated cell types. Each red dot represents an RNA seq sample. (C) RPKM values (left) and antibody staining (right) for transcription factors in the lamina (L1-L5) and retina (R7, R8) at 40 hrs APF. Cell-type specific markers are shown in green and antibodies for cell-type specific transcription factors are shown in red (L1-L5 and R7) or blue (R8). Arrows in L2 panels indicate

glial cells also stained with antibody to Bab2. A general retinal marker is shown in blue and red in the R7 and R8 panels, respectively. Scale bars, 10 μm. See also Table S1.

Many genes are differentially expressed between closely related neurons

We set out to gain a global perspective on gene expression differences between R7, R8 and L1-L5 using two different approaches. First, we performed pair-wise comparisons of their transcriptomes and identified differentially expressed genes between different cell types. As we obtained independent verification of cell-type enriched expression for transcripts with a maximum RPKM slightly below 5 using protein traps (Nagarkar-Jaiswal et al., 2015; Venken et al., 2011) (e.g. see DIP-β in Figure S4C), we set conservative criteria for genes differentially expressed (DE) between cell-types. We selected genes exhibiting a difference of >5X between one neuronal cell type and other neurons with expression in at least one cell type exhibiting an RPKM >5 with an adjusted P value of <0.05 (Table S2A). Even with these criteria, the number of DE genes between two cell types was substantial, ranging from 217 to 1,159. In summary, marked differences in gene expression between different neurons were observed at this stage in development.

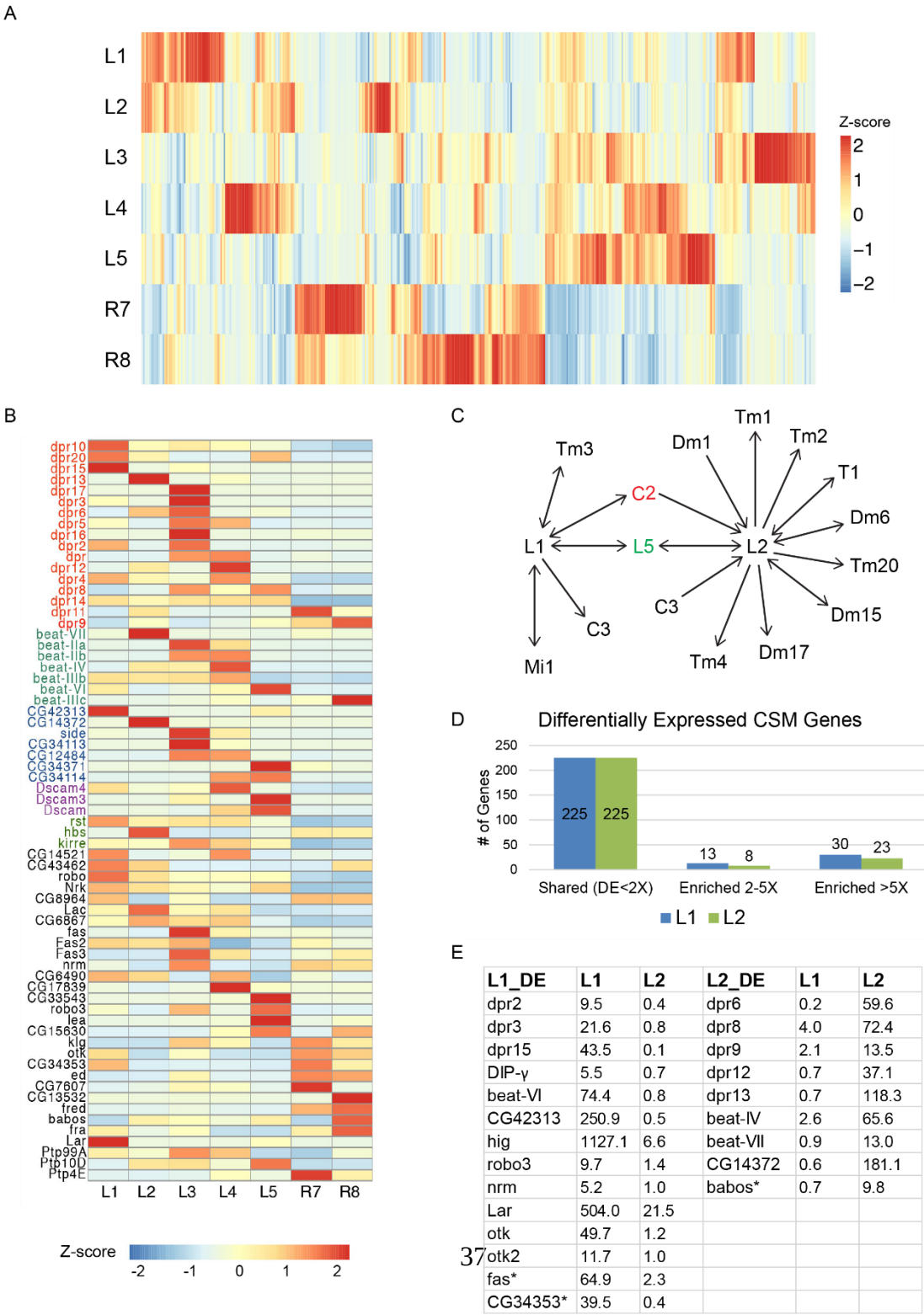
In a second approach, we explored the relationship between patterns of gene expression and specific cell types using a weighted gene co-expression network analysis (WGCNA)(Langfelder and Horvath, 2008). This unsupervised and unbiased analysis identified distinct co-expression modules by clustering transcripts with similar expression patterns across all samples (see *Supplemental Experimental Procedures*). Cell-type specific modules were preferentially enriched in cell surface membrane and secreted molecules (CSMs) (see Figure S1, Table S3). This is consistent with an important role for intercellular interactions as important determinants of patterns of synaptic connectivity.

Each neuronal cell type expresses a unique combination of mRNAs encoding CSMs during synapse formation

We next sought to identify genes encoding CSMs that are expressed in a cell-type enriched fashion, as these are candidates for regulating the cellular interactions underlying synaptic specificity. With a threshold of an RPKM >5 and an adjusted P value <0.05, we observed that each cell type expresses between a quarter to a third (i.e. between 247 (for R7) and 322 (for L3)) of the 976 genes encoding CSMs predicted to be encoded in the fly genome (Figure S2)(Kurusu et al., 2008) (see *Supplemental Experimental Procedures* for the criteria used to establish the list of CSMs) and each cell type exhibits a unique pattern of expression (Figure 3A). To gain an appreciation of the differences in genes encoding CSMs expressed between cell types, we carried out a pair-wise comparison of RPKM values for each pair of cell types. Here, we observed marked differences in expression, as each pair differentially expressed between 49 (between R7 and R8) and 168 (between R7 and L4) CSM genes (RPKM >5 in at least one cell type and >5X difference between the two cell types) (Table S2B). Further analyses revealed that only a small fraction of the CSM transcripts are selectively enriched in only one cell type of the seven profiled. Thus, each cell type expresses many genes encoding CSMs, the majority of which are expressed in multiple cell types, and there are marked differences in expression between cell types.

Several families of genes encoding CSMs known to regulate cellular interactions during circuit assembly were expressed in a cell-type enriched fashion. These included genes encoding immunoglobulin (Ig) (Fischbach et al., 2009; Zipursky et al., 2006), Leucine-Rich Repeat (LRR) (de Wit et al., 2011) and Epidermal Growth Factor (EGF) domain containing proteins

(Kenzelmann et al., 2007; Serafini et al., 1994), as well as many members of the large tetraspanin protein family (Fradkin et al., 2002; Kopczynski et al., 1996) (Figure 3B, Figure S3).



E

L1_DE	L1	L2	L2_DE	L1	L2
dpr2	9.5	0.4	dpr6	0.2	59.6
dpr3	21.6	0.8	dpr8	4.0	72.4
dpr15	43.5	0.1	dpr9	2.1	13.5
DIP-γ	5.5	0.7	dpr12	0.7	37.1
beat-VI	74.4	0.8	dpr13	0.7	118.3
CG42313	250.9	0.5	beat-IV	2.6	65.6
hig	1127.1	6.6	beat-VII	0.9	13.0
robo3	9.7	1.4	CG14372	0.6	181.1
nrm	5.2	1.0	babos*	0.7	9.8
Lar	504.0	21.5			
otk	49.7	1.2			
otk2	11.7	1.0			
fas*	64.9	2.3			
CG34353*	39.5	0.4			

Figure 2-3. Gene expression patterns of cell surface membrane and secreted molecules (CSMs) in each cell type.

(A) Heat map showing expression of all genes encoding CSMs expressed in at least one cell with an RPKM >5 (n =444). See also Figure S2. (B) Heat map representing expression of genes encoding Immunoglobulin (Ig) superfamily of cell surface proteins. Each gene in this list is expressed in at least one cell type with an RPKM >5 and 5X greater in one cell type than at least one of the six other cell types. Genes shown in color are members of gene sub-families. Note all Side family members (with the exception of Side) are shown as CG numbers. See also Figure S3 and Table S4. (C) Synaptic connections of L1 and L2 in medulla neuropil. They are largely different. Arrows indicate directionality of connection from pre-synaptic neuron to post-synaptic neuron. For example, L1 is pre-synaptic to C3. And L1 is both pre- and post-synaptic to L5. C2 in red means that C2 is a shared synaptic partner with both L1 and L2. L5 in green means that L5 is also a shared synaptic partner for both L1 and L2. (D) Numbers of genes exhibiting differences of <2X (shared), 2-5X and >5X in expression between L1 and L2 with RPKM >5 in at least one cell type, with an adjusted P value <0.05. Enriched means the level of a gene in one cell type is higher than the other cell type. Numbers of genes in each category are shown. See also Table S2. (E) Lists of genes encoding Ig superfamily cell surface proteins that are enriched in L1 and L2 by >5X. RPKM values in L1 and L2 are also listed. CG42313 and CG14372 are Side protein family members. Asterisk indicates that the interacting partner of the protein is not known yet. Interacting partners for all other proteins in this table have been identified (Johnson et al., 2006; Linnemannstöns et al., 2014; Özkan et al., 2013; Winberg et al., 2001). See also Figure S1 and Table S3.

Differential expression of Ig superfamily proteins in two closely related lamina neurons

As an additional step towards identifying candidates for regulating synaptic specificity in the medulla, we compared the pattern of expression of CSMs between two closely related neurons in more detail. To do this, we focused on L1 and L2. These neurons have similar patterns of gene expression and morphologies, particularly in the lamina, where their dendrites are postsynaptic to photoreceptor neurons. L1 and L2 are key components of the ON and OFF pathways, respectively, and play overlapping roles in motion detection (Borst, 2014). Their morphologies, layer specificity and patterns of synaptic connections in the medulla sharply diverge. For example, L1 and L2 are presynaptic to five and nine classes of neurons in the medulla, respectively, only one of which is

in common (Figure 3C) (Takemura, Meinertzhagen, and Scheffer, personal communication). Thus, we reasoned that differences in the CSMs expressed between these neurons would be candidates for regulating their synaptic specificity in the medulla.

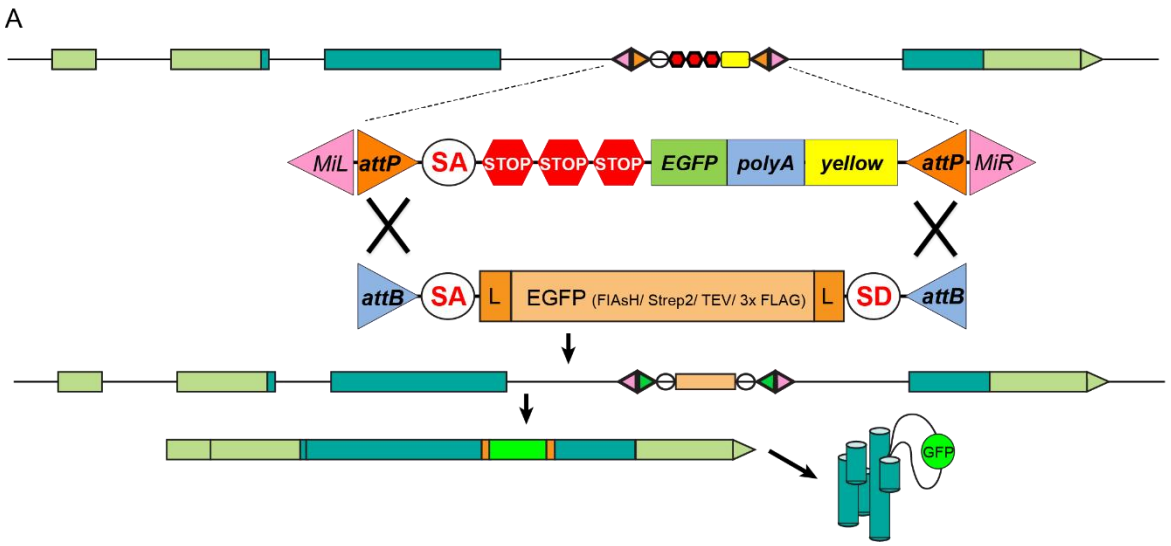
L1 and L2 express a similar number of CSM genes (i.e. ~260, with an RPKM >5, Figure S2). Of these, 225 were expressed at a similar level (i.e. <2X difference) in both neurons. By contrast, 21 and 53 CSM genes were expressed 2-5X and >5X between them, respectively (Figure 3D). Among the 53 CSM genes with >5X difference, 23 encode Ig superfamily cell surface proteins (Figure 3E). This enrichment is highly unlikely to arise by chance (P-values = 6.681e-06). The Dpr sub-family of Ig proteins was also enriched within this category with eight of the 21 family members differentially expressed (i.e. >5X) (P-value = 2.853e-04). Furthermore, as seen in Figure 3B, each lamina neuron expresses a unique combination of *dpr* genes (Figure 3B and Table S4). Based on these findings, we speculated that Dprs were good candidates for regulating cell-type specific patterns of synaptic connectivity in the medulla neuropil.

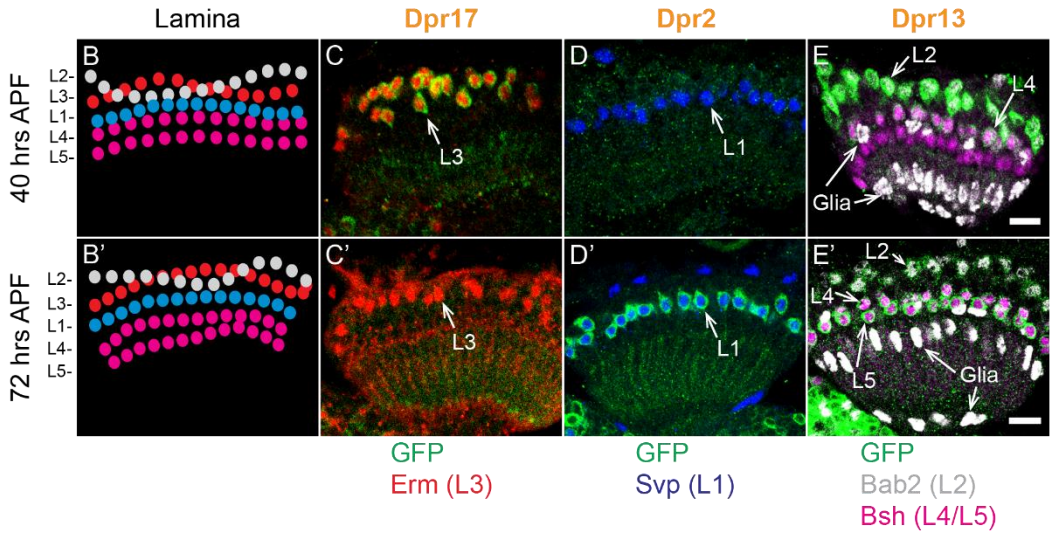
Many Dpr proteins are expressed in a cell-type enriched fashion

The twenty-one Dpr proteins are cell surface proteins comprising two Ig domains (Nakamura et al., 2002). They show a complex pattern of interactions in vitro, with another family, the Dpr Interacting Proteins or DIPs, comprising 3 Ig domains. These interactions were discovered in an Elisa-based assay and the interactions are presumed to occur in trans (Özkan et al., 2013). Their functional significance remains unclear, but they are expressed in the embryonic nervous system (Fisher et al., 2012; Özkan et al., 2013). Each Dpr binds to between one and four DIPs and each DIP binds to between one and seven Dprs (Özkan et al., 2013) (Figure 6F).

To independently assess the pattern of expression of Dprs in R7, R8 and L1-L5, we tagged the proteins produced from the endogenous locus with GFP using recombination mediated cassette exchange of specific MiMIC insertions into Dpr genes (Venken et al., 2011) (Figure 4A). MiMIC insertions into introns separating coding exons for 10 of the 21 Dpr genes were identified and converted into protein traps (Nagarkar-Jaiswal et al., 2015) (see *Supplemental Experimental Procedures*). These contain an open reading frame encoding GFP flanked by splice acceptor and donor sites. Pupal brains were stained at 40 hrs APF just prior to the onset of synapse formation and some 32 hrs later (at 72 hrs APF), a stage at which these neurons continue to add synaptic connections (Chen et al., 2014). In all cases, sufficient protein was detected in the cell body to identify the specific cells expressing the modified Dpr by co-staining the retina and lamina with antibodies to cell-type specific nuclear proteins (Figure 4B-E'). Cell-type enriched expression in lamina neurons was also observed in Beat (Pipes et al., 2001) and Side (Sink et al., 2001) protein families using this method (Figure S4F-H). For each Dpr tested, the protein trap expression pattern correlated well with the RPKM values observed (Figure 4F-F'). For some cell types, the pattern of expression was stable between 40 hrs and 72 hrs APF and for others marked changes were observed. For instance, Dpr15 and Dpr17 are expressed only at 40 hrs APF. By contrast, Dpr2 is selectively expressed at 72 hrs. Thus, from both RNA seq studies and MiMIC expression analysis, all cells express more than one Dpr and they express different combinations of them.

A





F Dpr expression at 40 hrs APF

	L1	L2	L3	L4	L5	R7	R8
Dpr1	4	2	282	298	9	40	10
Dpr2	10	0	12	1	0	0	1
Dpr3	22	1	117	1	0	0	0
Dpr6	0	60	90	11	16	1	3
Dpr10	119	52	66	52	39	12	2
Dpr11	0	5	1	1	1	9	3
Dpr12	1	37	5	95	0	1	4
Dpr13	1	118	3	16	0	2	4
Dpr15	44	0	7	0	0	0	0
Dpr17	1	1	354	2	0	1	1

RNA-seq (RPKM) data for 40 hrs

F' Dpr expression at 72 hrs APF

	L1	L2	L3	L4	L5	R7	R8
Dpr1							
Dpr2							
Dpr3							
Dpr6							
Dpr10*							
Dpr11							
Dpr12							
Dpr13							
Dpr15							
Dpr17							

No RNA-seq data for 72 hrs

GFP++ GFP+ GFP-

Figure 2-4. Dpr proteins are expressed in a neuronal cell-type enriched fashion in the lamina.

(A) Schematic of a MiMIC-based protein trap. MiMIC protein traps contain GFP in frame flanked by splice acceptor and donor sites. They are generated by cassette exchange using ϕC31 recombinase to catalyze recombination with the minos insertion between the AttP and AttB sites. The green inverted arrows after recombination represent recombined recombination sites (i.e. attR sites). Genes modified in this way encode chimeric proteins containing GFP. (B and B') Arrangement of lamina neuron cell bodies at 40 hrs and 72 hrs APF. L2 and L3 are intermingled at the top of lamina cell clusters. L4 and L5 make up the bottom two rows with L5 beneath L4. (C-E') Dpr17, Dpr2 and Dpr13 expression in lamina neurons visualized using MiMIC protein traps. See Figure 2 for lamina neuron markers. Scale bars, 10 μm. (F, F') Summary of Dpr expression using protein trap lines (10 of 21 *dpr* genes). RPKM values from the RNA seq results indicating level of gene expression are included in F. Dpr2, Dpr13 and Dpr17 are orange colored in bold to indicate changes in staining with the preceding panels. Asterisk indicates Dpr10 expression level in L5 is variable at 72hrs APF. See also Figure S4.

DIPs are expressed in a layer-specific fashion

If Dpr proteins regulate interactions with specific neurons, we would anticipate that DIPs would be expressed in neurons with which R7, R8 and L1-L5 interact. To explore the in vivo expression of DIPs, we generated and analyzed GFP protein trap derivatives for six of the nine DIPs (* in Figure S4A, see *Supplemental Experimental Procedures*) and assessed their expression at 24, 40, 72 hrs APF, and in the adult. Consistent with our RNA seq data, DIP-β and DIP-γ are expressed at low levels in a subset of lamina neurons (Figure S4C-D') and the remaining DIPs are not expressed in these cells (Figure S4B, B', E, E').

Each DIP analyzed was expressed in neurons exhibiting unique layer-specific patterns of processes within the medulla neuropil (Figure 5A-Q, Figure S5). Prior to synapse formation (i.e. 24 hrs APF) layered patterns are diffuse and overlap (Figure S5A-G). By 40% APF, the patterns are less diffuse with less overlap between different DIPs (Figure 5D-J). At this stage of development, layers are still forming and most neurons have yet to form synapses. By contrast, some 32 hrs later (at 72hrs

APF) the medulla neuropil has expanded, many neurons have formed synapses (Chen et al., 2014) and the processes expressing DIPs are, in general, more clearly separated (Figure 5K-Q). At this stage, the six DIPs are expressed in one to three layers and all layers are defined by a unique combination of them (Figure 5R). With one exception (DIP-θ) (Figure S5H), layer-specific expression patterns remained the same in the adult as they were at 72 hrs APF. In summary, DIPs are differentially expressed in layers innervated by R7, R8 and L1-L5 neurons.

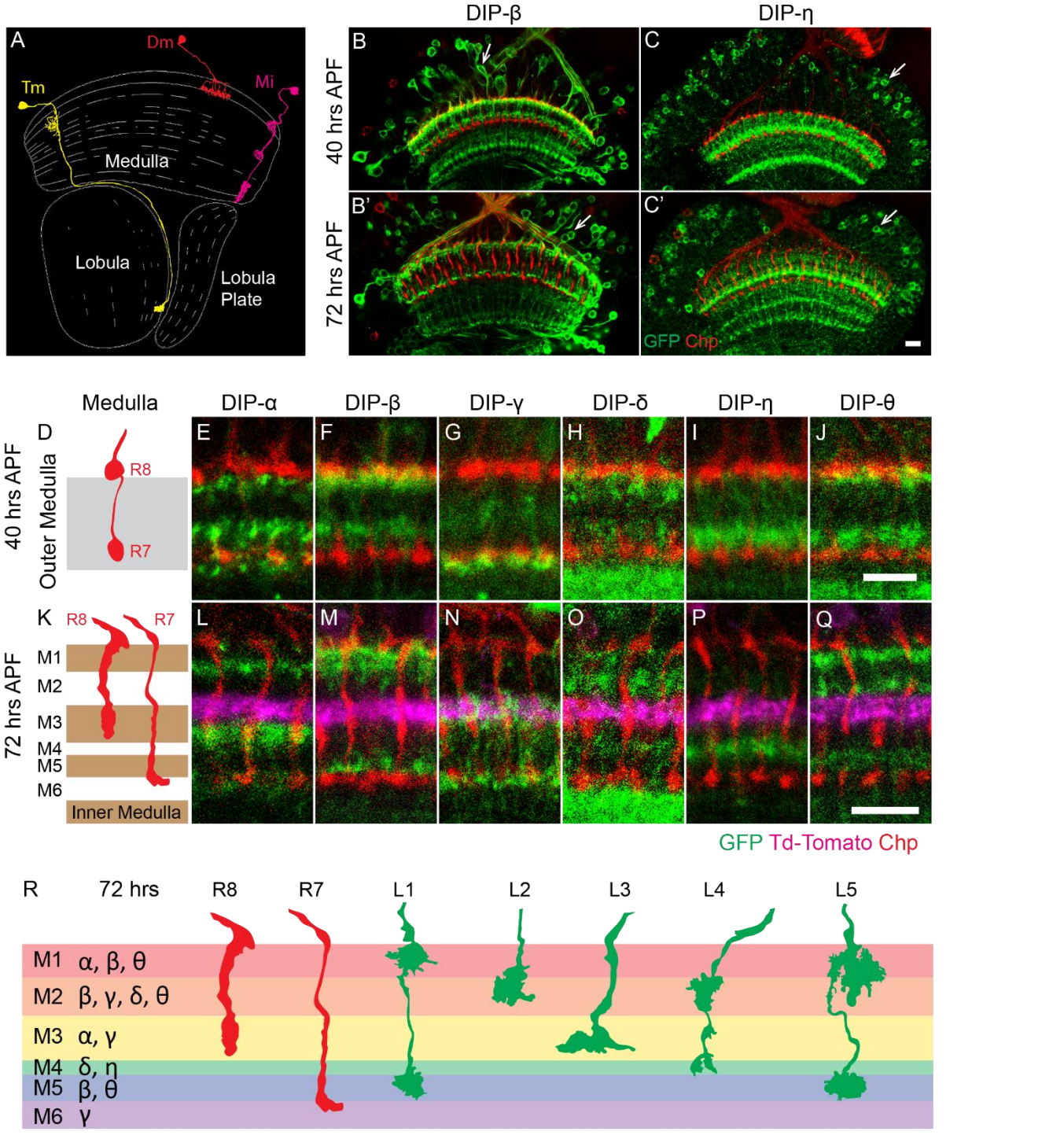


Figure 2-5. DIP proteins are expressed in a layer-specific fashion in the medulla.

(A) Schematic of three classes of medulla neurons. A transmedullary neuron (Tm, in yellow), an amacrine-like distal medulla neuron (Dm, in red) and a medulla intrinsic neuron (Mi, in magenta)

are shown. Each class of medulla neurons can be further divided into specific cell types based on different patterns of layer-specific branching. Adapted from Fischbach and Dittrich, 1989. (B-C') DIP protein traps are expressed in scattered cells in the medulla cortex (arrows) and in layer-specific patterns in the medulla neuropil. DIP-β and DIP-η are shown as examples (green). Photoreceptor axons are visualized by staining for the cell surface protein Chp (red). (D-J) All six DIP genes for which protein trap lines are available were expressed in neurons exhibiting unique layer-specific patterns of processes within the outer medulla neuropil at 40 hrs APF. (D) Schematic of R8 and R7 axon morphology and layer distribution in the outer medulla. (E-J) Protein expression of six DIPs (green) in the outer medulla. The six DIPs are expressed in one to three layers; each layer is defined by a unique combination of DIPs. (K-Q) The DIP expression pattern at 72 hrs is shown. (K) Schematic of R7 and R8 axons at 72 hrs. The medulla expands and R7 and R8 layers change between 40 and 72 hrs. (L-Q) The layered expression of DIPs is largely the same as at 40 hr. Expression in an additional layer, however, appears in DIP-θ. Dm3 axons are labeled with td-tomato (magenta). They run parallel to layers and mark the M2 and M3 border. (R) Summary of expression of DIPs in the medulla and the projection of R8, R7 and L1-L5 terminals at 72 hrs APF. Scale bars, 10 μm. See also Figure S5.

Dprs and DIPs are expressed by synaptic partners

We next sought to assess whether Dpr-expressing lamina neurons and DIP-expressing medulla neurons with processes in the same layers were synaptic partners. To assess the identity of medulla neurons expressing specific DIPs, we crossed flies carrying a DIP-GFP (DIP-α, DIP-δ and DIP-θ) to a panel of GAL4 marker lines for specific medulla neurons and assessed co-localization of the markers (see *Supplemental Experimental Procedures*). For some DIPs (i.e. DIP-β), we used DIP-GAL4 derivatives of MiMICs in combination with FLP-mediated excision to express target UAS reporter constructs active in scattered DIP-expressing cells (Nern et al., 2015). This allowed us to visualize individual DIP-expressing neurons and to identify them by their morphologies. We then correlated the identification of cells expressing specific DIPs and Dprs with the pattern of synaptic connections within layers determined by serial EM reconstruction (Takemura et al., 2013, 2015).

Synaptic Partners	Number of Synapses	
	L5 Pre	L5 Post
Dm1	39	13
Dm10	48	7
Dm18	39	13
C2	35	37
C3	13	24
Mi1	56	11
Mi4	58	0
Tm3	83	0
L1	29	127

Table 2-1. Synaptic partners of L5 expressing Dpr/DIP Pairs

A dense connectome of 7 medulla columns has been completed by serial EM reconstruction (Takemura et al., 2015; Takemura, Meinertzhagen, and Scheffer, personal communication.). This includes a central column and 6 surrounding ones. Here the synaptic partners of L5 with the number of inputs and outputs are listed. Most synapses are made with partners in the same column but processes can also extend into neighboring columns and form synapses. Here we are showing the sum of the synapses made by an L5 neuron in the central column to partners within the same and neighboring columns.

Dense synaptic connectomes for a single column (Takemura et al., 2013) and, more recently, seven adjacent columns, comprising a central one surrounded by six additional ones (Takemura et al., 2015; (Takemura, Meinertzhagen, and Scheffer, personal communication), have been determined. In general, these studies revealed that lamina neurons make synapses with multiple partners and show marked specificity. An example of the synaptic connections made by L5 neurons is shown in Table 1. L1-L5 neurons each express one or more Dpr proteins, which bind to DIPs expressed

in a subset of their synaptic partners (Figure 6A-E). For instance, L1 expresses Dpr2 and Dpr3 and these proteins bind to DIP-θ expressed on one of their synaptic partners, Tm3. Other synaptic partners of L1 do not express DIP-θ (Figure 6A, Table S5). The Dpr/DIP patterns of expression in L5 and its partners provide an example of a more complex relationship between these paralogs than L1 (Figure 6E). L5 expresses Dprs that bind to three different DIPs, DIP-α, DIP-β and DIP-θ, in three synaptic partners: 1. Like L1, L5 makes synapses with Tm3 and these neurons express a matching pair of Dpr1 and DIP-θ, respectively; 2. L5 expresses Dpr6, which binds to DIP-β, which is expressed in post-synaptic C2 neurons; and 3. Dpr6 and Dpr10 bind to DIP-α on Dm1 neurons. By contrast, five other synaptic partners of L5 do not express these DIPs. In addition, we also demonstrated that R7 neurons express Dpr11 and its cognate DIP is expressed in its synaptic partner Dm8 (see *accompanying paper by Carrillo et al.*). Thus, Dprs in R7 and L1-L5 neurons match DIPs expressed in subsets of their synaptic partners.

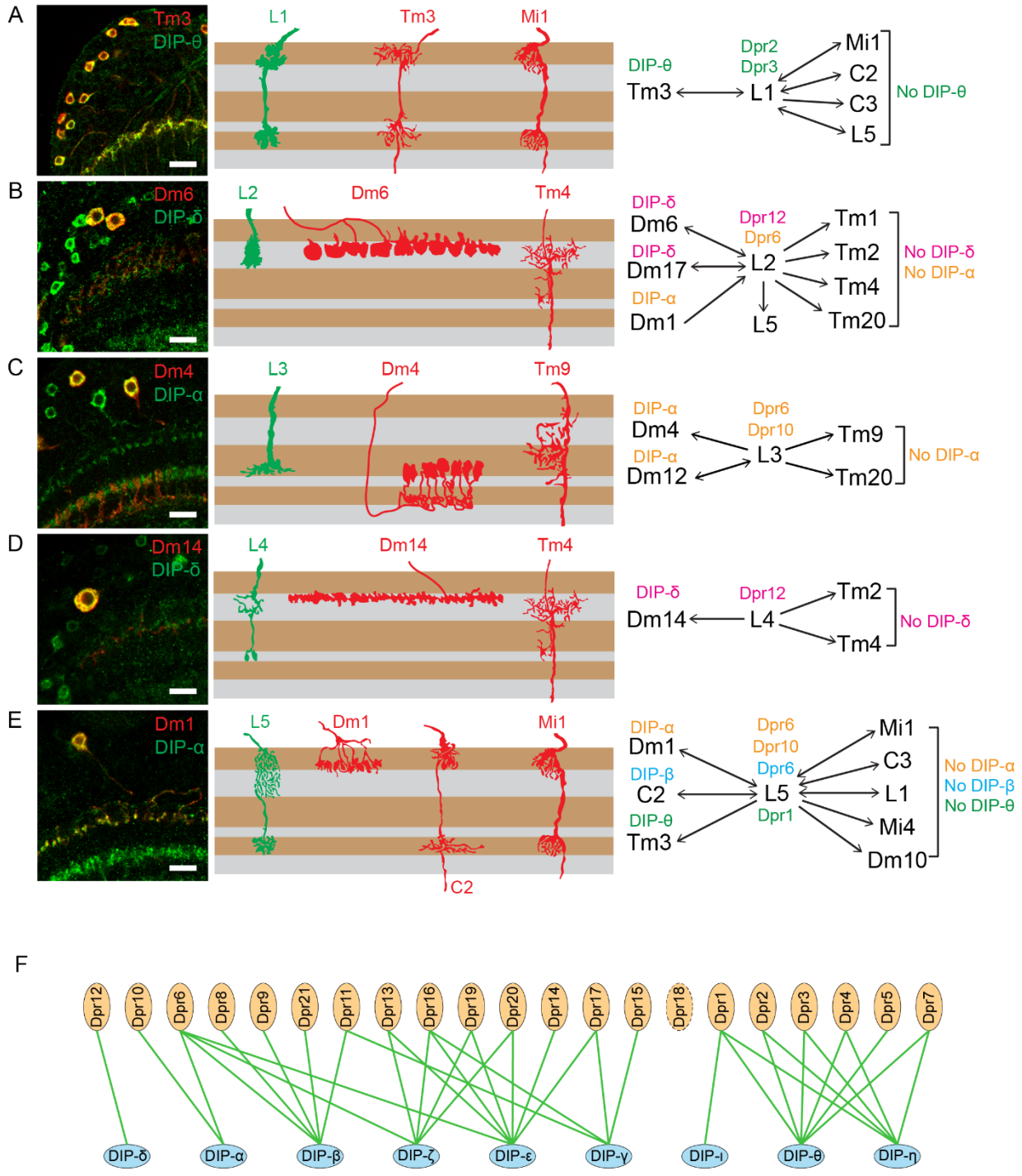


Figure 2-6. Matching of cognate Dpr and DIP expression in synaptic partners.

(A-E) Co-localization of DIPs in synaptic partners of L1-L5. Left panels, co-localization of indicated DIP (green) and cell-type specific marker (red) in the adult; Middle panels, schematic of morphology of lamina neurons (green) and a subset of their synaptic partners (red) within the

medulla neuropil; Right panels, summary of Dpr expression in L1-L5 and DIPs in synaptic partners. Layer patterns for DIPs in the medulla are the same as at 72 hrs. Synaptic partner assignments from Takemura et al., 2015 and Takemura, Meinertzhagen, and Scheffer, personal communication. F. Summary of the Dpr/DIP interactome (Ozkan et al., 2013; see *accompanying paper by Carrillo et al.*). See also Table S5.

Discussion

Here we used RNA sequencing of mRNAs from different, but highly related neuronal cell types, to identify families of cell surface proteins as potential regulators of synaptic specificity. Using a conservative RPKM threshold for expression, we estimate that at the onset of synapse formation neurons express between a quarter and a third of the potential CSMs encoded by the fly genome. Many of these are expressed in a cell-type enriched fashion; they are expressed at least 5X greater in one cell type, than in one or more of the other cell types profiled. Thus, neurons express many different CSMs, and each neuronal cell type expresses a unique combination of them. Many of the differentially expressed proteins have known interacting partners. For instance, of the 23 Ig superfamily proteins differing in expression by more than 5X between L1 and L2, all but three have known interacting partners. Thus, expression studies coupled with the cell surface interactome and genetic analysis provide a multipronged approach to dissecting the cellular interactions leading to neural circuit formation.

Several families of proteins were differentially expressed in different neuronal cell types through RNA seq analysis, and some of these were confirmed using MiMIC protein traps. The most dramatic and complex pattern of expression observed, however, was seen for the Dpr proteins with 17 of the 21 paralogs expressed in a cell-type enriched fashion. This pattern and the extensive

protein interaction network of these proteins defined by Garcia and colleagues (Özkan et al., 2013) prompted us to explore the expression patterns of these proteins and their ligands in further detail.

The cell-type enriched pattern of Dpr expression observed in lamina or photoreceptor neurons was striking. By contrast, most DIPs were not expressed, or were expressed at only very low levels in these cells. Localization of DIP expression using protein traps, however, revealed that each of the outer 6 layers of the medulla neuropil was defined by the expression of one or more DIPs and that DIPs were expressed in only a subset of processes within a layer. And these, in turn, are specific subsets of synaptic targets of cells expressing an interacting Dpr. These observations raise the possibility that Dpr/DIP interactions specify patterns of synaptic connections between neurons within each layer.

These findings, in combination with previous genetic studies, suggest a model for mechanisms regulating the formation of layer-specific connections within the medulla. Previous work demonstrated that at early stages of medulla development, lamina neuron growth cones target to overlapping regions that are established by broadly expressed adhesive (i.e. N-cadherin) and repulsive cell surface molecules (i.e. Plexin/Semaphorin signaling). Growth cones then segregate into discrete domains as the medulla matures through interactions between signals localized to specific layers (Nern et al., 2008; Pecot et al., 2013; Timofeev et al., 2012). We speculate that different combinations of Dpr and DIP proteins specify synaptic connections within a layer.

In support of this notion, Zinn and colleagues (see *accompanying paper by Carrillo et al.*) have shown that loss of Dpr11 and its binding partner DIP-γ show abnormalities in the M6 target layer. Dpr11 is expressed in a discrete subset of R7 neurons and these form synapses with DIP-γ -expressing Dm8 neurons. These include abnormalities in the R7 terminal morphology consistent

with a role in synapse formation and a reduction in Dm8 neurons. Interestingly, we have also recently observed a reduction in the number of DIP-α expressing neurons in DIP-α null mutants suggesting a commonality in the function of Dprs and DIPs in the medulla. The simplest interpretation of the matching of Dprs and DIPs in synaptic partners is that these proteins regulate synaptic specificity. It would not be surprising, however, if these proteins play different roles such as contributing to layer specific targeting, as with N-cadherin (Lee et al., 2001) or Netrin (Timofeev et al., 2012), or cell-type specific trophic support as we described previously for Jeb/Alk signaling (Pecot et al., 2014). Detailed phenotypic analyses of null mutants lacking Dprs and DIPs, and given the redundancy within these families, perhaps genetic analysis of animals lacking combinations of them, will be required to ascertain the precise functions of this family of ligand/receptor pairs in circuit assembly.

The two-step model for synaptic connectivity in the medulla shares intriguing similarities to, and indeed was significantly influenced by, models for layer specificity in the analogous structure in the mouse retina, the inner plexiform layers. Here cadherin and semaphorin/plexin proteins direct processes to layers (Duan et al., 2014; Matsuoka et al., 2011). In a subsequent step, Ig superfamily proteins then promote synaptic matching within them. Important support for this second step comes from recent genetic studies from Sanes and colleagues demonstrating that homophilic interaction between Sdk2 proteins (an Ig superfamily protein) is required for synapses between a specific pair of amacrine and retinal ganglion cell neurons (Krishnaswamy et al., 2015). That this may represent a general strategy for synaptic-specificity in the vertebrate retina is suggested by the layer-specific expression of Sdk2 and related homophilic Ig superfamily proteins Sdk1, Dscam1, Dscam2 and Contactins 1-5 (Yamagata and Sanes, 2002, 2008, 2012). Thus, the studies

in the mouse IPL and the medulla region of the fly allude to a common strategy for achieving synaptic specificity.

Concluding Remarks

Dprs and DIPs are likely to be only a part of the story of synaptic specificity in the medulla. Indeed, a striking feature of the synaptic connectome in the medulla column is its complexity (Takemura et al., 2013, 2015), with synapses between the processes of greater than 100 neuronal cell types (Nern, personal communication). This complexity is mirrored by the unique combination of hundreds of cell surface and secreted molecules expressed by each of the photoreceptor and lamina neurons profiled in this study. How this complexity contributes to specificity remains elusive, but the convergence of improved histological, genetic, physiological, and molecular tools promises to provide important insights into the molecular recognition strategies controlling synaptic specificity.

Experimental procedures

Fly husbandry and stocks

Flies were reared at 25ºC on standard medium. For developmental analysis and sorting experiments white pre-pupae were collected and incubated for the indicated number of hours. See supplemental experimental procedures for the list of stocks used in different experiments.

Sorting cell types and library construction

For tissue dissociation, pupal brain tissue dissected at 40hrs after pupal formation was incubated with a Papain (Worthington) and Liberase TM protease cocktail (Roche) at 25^0C for 15 min in a microfuge shaker at 1000rpm. At 5 and 10 min into this incubation, the tissue was pipetted up and down with a P200. At 15 min the sample was passed through a 25G 5/8-gauge needle until the tissue was completely dissociated. Digestion was inactivated by the addition of rich media with serum, and the cell suspension was passed through a 70μm filter. To concentrate the cells, the suspension was spun down at 1600rpm for 8 min at 4^0C. After decanting the supernatant, cells were re-suspended in approximately 500ul of rich media and sorted in a BD FACSAria II.

RNA was then isolated from sorted cells using the RNA-min elute kit from Qiagen. mRNA was amplified in a linear fashion using Arcuturus RiboAmp HS kit (Life Technologies). cDNA was then generated for quality assessment and paired-end Illumina sequencing libraries were prepared.

Detailed protocols are available upon request.

Microscopy and Image Analysis

Confocal images were acquired on a Zeiss LSM780 confocal microscope. The staining patterns were reproducible between samples. However, some variation on the overall fluorescence signal and noise levels existed between sections and samples. Thus, proper adjustments of laser power, detector gain, and black level settings were made to obtain similar overall fluorescence signals. Single plane or maximum intensity projection confocal images were exported into TIFF files using ImageJ software.

See supplemental experimental procedures for bioinformatics analysis and immunohistochemistry.

References

Borst, A. (2014). Fly visual course control: behaviour, algorithms and circuits. Nat. Rev. Neurosci. *15*, 590–599.

Cang, J., and Feldheim, D.A. (2013). Developmental mechanisms of topographic map formation and alignment. Annu. Rev. Neurosci. *36*, 51–77.

Chen, Y., Akin, O., Nern, A., Tsui, C.Y.K., Pecot, M.Y., and Zipursky, S.L. (2014). Cell-type-specific labeling of synapses in vivo through synaptic tagging with recombination. Neuron *81*, 280–293.

Duan, X., Krishnaswamy, A., De la Huerta, I., and Sanes, J.R. (2014). Type II cadherins guide assembly of a direction-selective retinal circuit. Cell *158*, 793–807.

Fischbach, K.-F., and Dittrich, a P. (1989). The optic lobe of Drosophila melanogaster. I: A. Golgi analysis of wild-type structure. Cell Tissue Res *258*, 441–475.

Fischbach, K.-F., Linneweber, G.A., Andlauer, T.F.M., Hertenstein, A., Bonengel, B., and Chaudhary, K. (2009). The irre cell recognition module (IRM) proteins. J. Neurogenet. *23*, 48–67.

Fisher, B., Weiszmann, R., Frise, E., Hammonds, A., Tomancak, P., Beaton, A., Berman, B., Quan, E., Shu, S., Lewis, S., et al. (2012). BDGP insitu homepage.

Fradkin, L.G., Kamphorst, J.T., DiAntonio, A., Goodman, C.S., and Noordermeer, J.N. (2002). Genomewide analysis of the Drosophila tetraspanins reveals a subset with similar function in the formation of the embryonic synapse. Proc. Natl. Acad. Sci. U. S. A. *99*, 13663–13668.

Hadjieconomou, D., Timofeev, K., and Salecker, I. (2011). A step-by-step guide to visual circuit assembly in Drosophila. Curr. Opin. Neurobiol. *21*, 76–84.

Hong, W., Mosca, T.J., and Luo, L. (2012). Teneurins instruct synaptic partner matching in an olfactory map. Nature *484*, 201–207.

Jenett, A., Rubin, G.M., Ngo, T.T.B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., et al. (2012). A GAL4-Driver Line Resource for Drosophila Neurobiology. Cell Rep. *2*, 991–1001.

Johnson, K.G., Tenney, A.P., Ghose, A., Duckworth, A.M., Higashi, M.E., Parfitt, K., Marcu, O., Heslip, T.R., Marsh, J.L., Schwarz, T.L., et al. (2006). The HSPGs Syndecan and Dallylike bind the receptor phosphatase LAR and exert distinct effects on synaptic development. Neuron *49*, 517–531.

Kenzelmann, D., Chiquet-Ehrismann, R., and Tucker, R.P. (2007). Teneurins, a transmembrane protein family involved in cell communication during neuronal development. Cell. Mol. Life Sci. *64*, 1452–1456.

Kohmura, N., Senzaki, K., Hamada, S., Kai, N., Yasuda, R., Watanabe, M., Ishii, H., Yasuda, M., Mishina, M., and Yagi, T. (1998). Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. Neuron *20*, 1137–1151.

Kopczynski, C.C., Davis, G.W., and Goodman, C.S. (1996). A neural tetraspanin, encoded by late bloomer, that facilitates synapse formation. Science *271*, 1867–1870.

Krishnaswamy, A., Yamagata, M., Duan, X., Hong, Y.K., and Sanes, J.R. (2015). Sidekick 2 directs formation of a retinal circuit that detects differential motion. Nature *524*, 466–470.

Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E., and Zinn, K. (2008). A Screen of Cell-Surface Molecules Identifies Leucine-Rich Repeat Proteins as Key Mediators of Synaptic Target Selection. Neuron *59*, 972–985.

Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics.

Langley, J.N. (1895). Note on Regeneration of Prae-Ganglionic Fibres of the Sympathetic. J. Physiol. *18*, 280–284.

Lee, T., and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron *22*, 451–461.

Lee, C.H., Herman, T., Clandinin, T.R., Lee, R., and Zipursky, S.L. (2001). N-cadherin regulates target specificity in the Drosophila visual system. Neuron *30*, 437–450.

Lefebvre, J.L., Kostadinov, D., Chen, W. V, Maniatis, T., and Sanes, J.R. (2012). Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. Nature *488*, 517–521.

Linnemannstöns, K., Ripp, C., Honemann-Capito, M., Brechtel-Curth, K., Hedderich, M., and Wodarz, A. (2014). The PTK7-related transmembrane proteins off-track and off-track 2 are co-receptors for Drosophila Wnt2 required for male fertility. PLoS Genet. *10*, e1004443.

Matsuoka, R.L., Nguyen-Ba-Charvet, K.T., Parray, A., Badea, T.C., Chédotal, A., and Kolodkin, A.L. (2011). Transmembrane semaphorin signalling controls laminar stratification in the mammalian retina. Nature *470*, 259–263.

Miura, S.K., Martins, A., Zhang, K.X., Graveley, B.R., and Zipursky, S.L. (2013). Probabilistic splicing of Dscam1 establishes identity at the level of single neurons. Cell *155*, 1166–1177.

Morante, J., and Desplan, C. (2008). The Color-Vision Circuit in the Medulla of Drosophila. Curr. Biol. *18*, 553–565.

Nagarkar-Jaiswal, S., Lee, P.-T., Campbell, M.E., Chen, K., Anguiano-Zarate, S., Cantu Gutierrez, M., Busby, T., Lin, W.-W., He, Y., Schulze, K.L., et al. (2015). A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila. Elife *4*, 1–28.

Nakamura, M., Baldwin, D., Hannaford, S., Palka, J., and Montell, C. (2002). Defective proboscis extension response (DPR), a member of the Ig superfamily required for the gustatory response to salt. J. Neurosci. *22*, 3463–3472.

Nern, A., Pfeiffer, B.D., and Rubin, G.M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. Proc. Natl. Acad. Sci. U. S. A. *112*, 2967-2976.

Nern, A., Zhu, Y., and Zipursky, S.L. (2008). Local N-cadherin interactions mediate distinct steps in the targeting of lamina neurons. Neuron *58*, 34–41.

O'Donnell, M., Chance, R.K., and Bashaw, G.J. (2009). Axon growth and guidance: receptor regulation and signal transduction. Annu. Rev. Neurosci. *32*, 383–412.

Özkan, E., Carrillo, R. a., Eastman, C.L., Weiszmann, R., Waghray, D., Johnson, K.G., Zinn, K., Celniker, S.E., and Garcia, K.C. (2013). An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. Cell *154*.

Pecot, M., Chen, Y., Akin, O., Chen, Z., Tsui, C.Y.K., and Zipursky, S.L. (2014). Sequential axon-derived signals couple target survival and layer specificity in the drosophila visual system. Neuron *82*, 320–333.

Pecot, M.Y., Tadros, W., Nern, A., Bader, M., Chen, Y., and Zipursky, S.L. (2013). Multiple interactions control synaptic layer specificity in the Drosophila visual system. Neuron *77*, 299–310.

Pipes, G.C., Lin, Q., Riley, S.E., and Goodman, C.S. (2001). The Beat generation: a multigene family encoding IgSF proteins related to the Beat axon guidance molecule in Drosophila. Development *128*, 4545–4552.

Schmitt, A.M., Shi, J., Wolf, A.M., Lu, C.-C., King, L.A., and Zou, Y. (2006). Wnt-Ryk signalling mediates medial-lateral retinotectal topographic mapping. Nature *439*, 31–37.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C. a, Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell *101*, 671–684.

Serafini, T., Kennedy, T.E., Galko, M.J., Mirzayan, C., Jessell, T.M., and Tessier-Lavigne, M. (1994). The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. Cell *78*, 409–424.

Shen, K., and Bargmann, C.I. (2003). The immunoglobulin superfamily protein SYG-1 determines the location of specific synapses in C. elegans. Cell *112*, 619–630.

Shen, K., Fetter, R.D., and Bargmann, C.I. (2004). Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1. Cell *116*, 869–881.

Sink, H., Rehm, E.J., Lee, R., Bulls, Y.M., and Goodman, C.S. (2001). Sidestep encodes a target-derived attractant essential for motor axon guidance in Drosophila. Cell *105*, 57–67.

Sperry, R.W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. Proc. Natl. Acad. Sci. U. S. A. *50*, 703–710.

Takemura, S., Xu, C., Lu, Z., Rivlin, P., Parag, T., Olbris, D., Plaza, S., Zhao, T., Katz, W., Umayam, L., et al. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. Proc. Natl. Acad. Sci. U. S. A. 1509820112v1-201509820.

Takemura, S., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P.K., Katz, W.T., Olbris, D.J., Plaza, S.M., Winston, P., et al. (2013). A visual motion detection circuit suggested by Drosophila connectomics. Nature *500*, 175–181.

Takemura, S.Y., Lu, Z., and Meinertzhagen, I. a. (2008). Synaptic circuits of the Drosophila optic lobe: The input terminals to the medulla. J. Comp. Neurol. *509*, 493–513.

Timofeev, K., Joly, W., Hadjieconomou, D., and Salecker, I. (2012). Localized netrins act as positional cues to control layer-specific targeting of photoreceptor axons in Drosophila. Neuron *75*, 80–93.

Triplett, J.W., and Feldheim, D.A. (2012). Eph and ephrin signaling in the formation of topographic maps. Semin. Cell Dev. Biol. *23*, 7–15.

Venken, K.J.T., and Bellen, H.J. (2014). Chemical mutagens, transposons, and transgenes to interrogate gene function in Drosophila melanogaster. Methods *68*, 15–28.

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., et al. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nat. Methods *8*, 737–743.

Ward, A., Hong, W., Favaloro, V., and Luo, L. (2015). Toll receptors instruct axon and dendrite targeting and participate in synaptic partner matching in a Drosophila olfactory circuit. Neuron *85*, 1013–1028.

Winberg, M.L., Tamagnone, L., Bai, J., Comoglio, P.M., Montell, D., and Goodman, C.S. (2001). The transmembrane protein Off-track associates with Plexins and functions downstream of Semaphorin signaling during axon guidance. Neuron *32*, 53–62.

de Wit, J., Hong, W., Luo, L., and Ghosh, A. (2011). Role of Leucine-Rich Repeat Proteins in the Development and Function of Neural Circuits. Annu. Rev. Cell Dev. Biol. *27*, 697–729.

Wu, Q., and Maniatis, T. (1999). A striking organization of a large family of human neural cadherin-like cell adhesion genes. Cell *97*, 779–790.

Yamagata, M., Weiner, J.A., and Sanes, J.R. (2002). Sidekicks: synaptic adhesion molecules that promote lamina-specific connectivity in the retina. Cell *110*, 649-660.

Yamagata, M., and Sanes, J.R. (2008). Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. Nature *451*, 465–469.

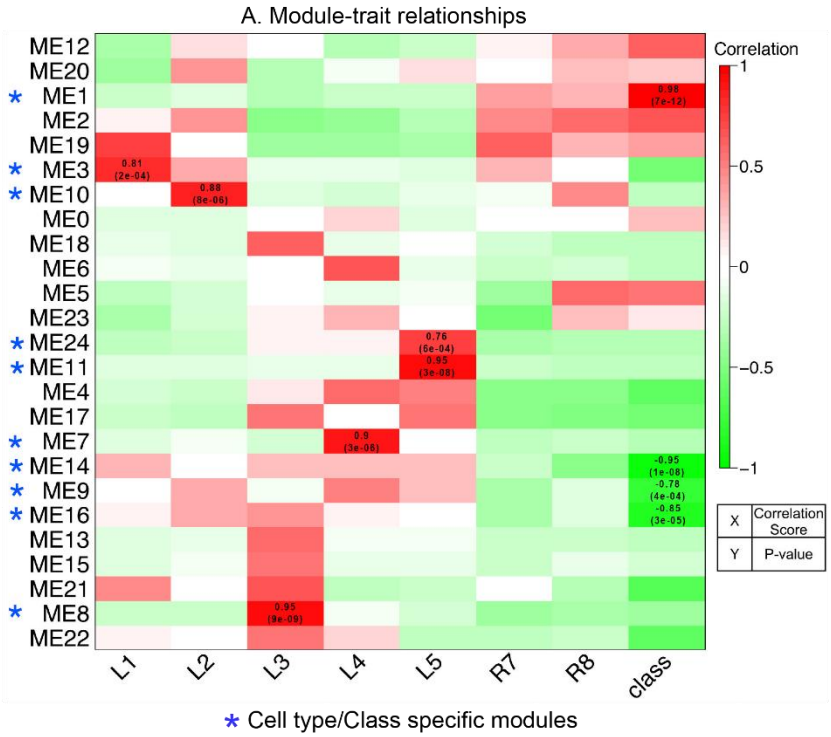
Yamagata, M., and Sanes, J.R. (2012). Expanding the Ig superfamily code for laminar specificity in retina: expression and role of contactins. J. Neurosci. *32*, 14402–14414.

Zipursky, S.L., and Grueber, W.B. (2013). The Molecular Basis of Self-Avoidance. Annu. Rev. Neurosci. *36*, 547–568.

Zipursky, S.L., Wojtowicz, W.M., and Hattori, D. (2006). Got diversity? Wiring the fly brain with Dscam. Trends Biochem. Sci. *31*, 581–588.

Supplemental information

Supplemental Figures and Legends



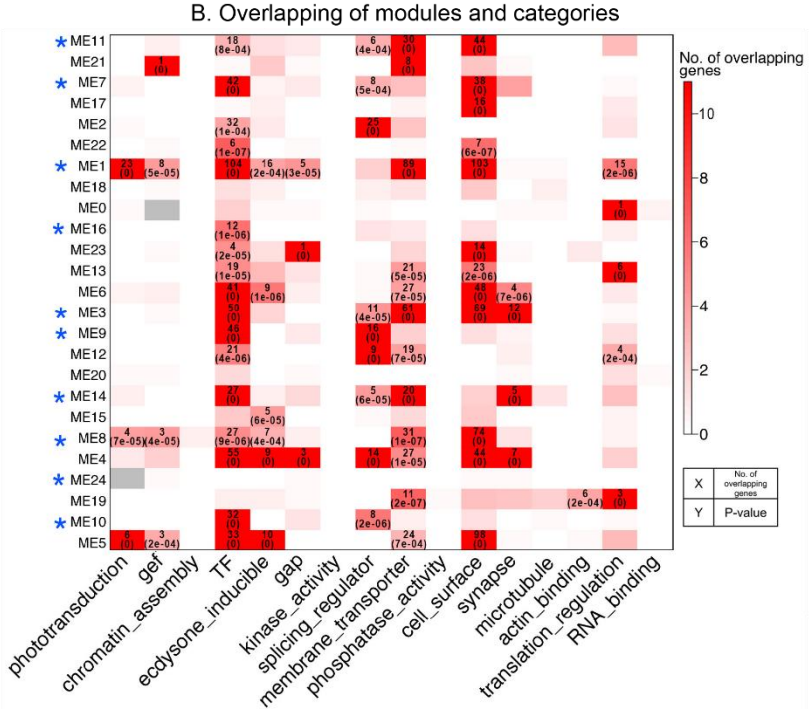


Figure S2-1. Cell-type specific modules and biological functions. Related to Figure 3. (A) Heatmap of the relationship of modules to each cell type (see text). WGCNA was carried out to identify genes expressed in a common pattern. Twenty-five expression modules were identified (see Supplemental Experimental Procedures and Table S3). These were then assessed for their correlation with an idealized cell-type specific pattern (see text). Cells are shown in columns and modules in rows. Here, cell-type specific modules are defined as $r > 0.8$ and P-value < 0.001. Class refers to the comparison of retinal cells (i.e. R7 and R8 together) to lamina neurons (i.e. L1-L5 together). (B) Heatmap showing the significance of gene overlaps between hand-curated biological categories and cell-type specific modules. Each column represents a cell-type specific module and each row represents a different biological category. A P-value of the intersection is based on the permutation test. Number of overlapping genes between the indicated modules and categories are only labeled in cells with P-value < 0.001.

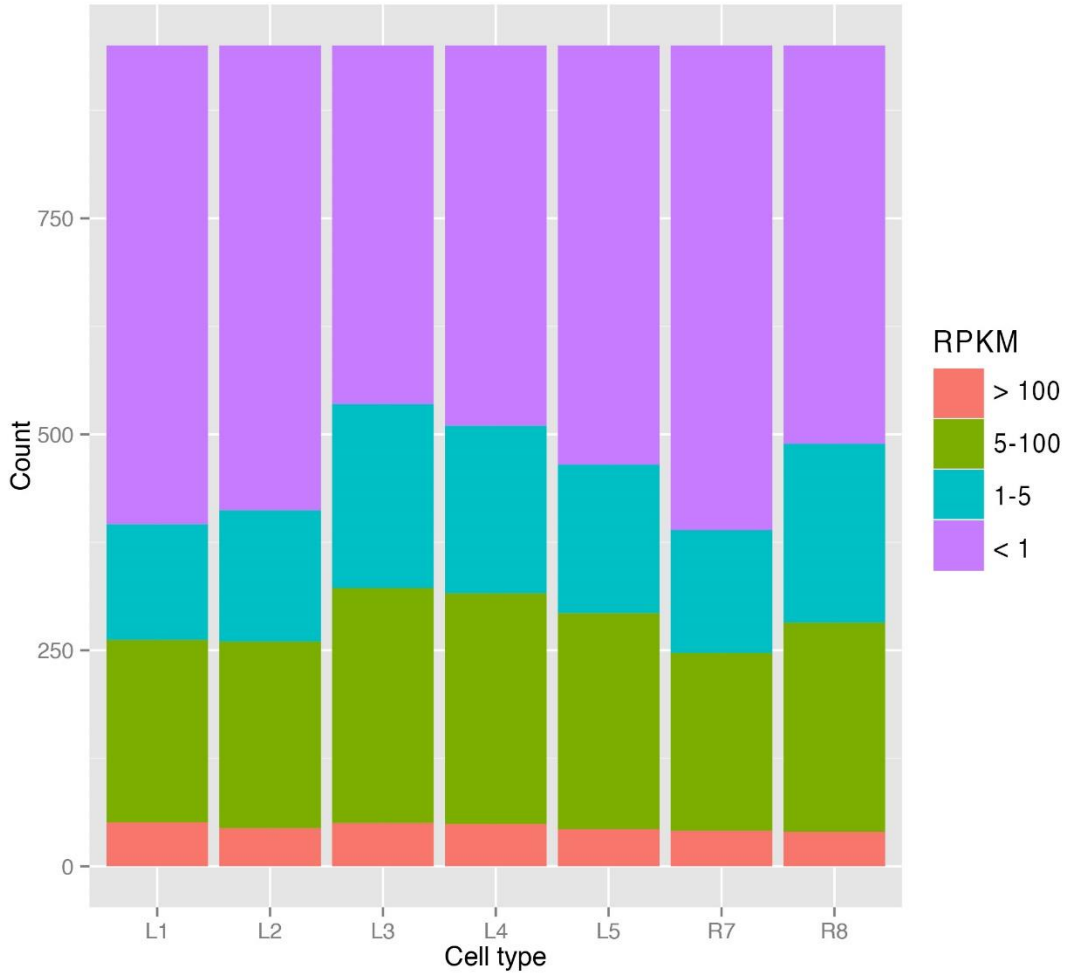


Figure S2-2. The distribution of RPKM values for cell surface and secreted molecules (CSMs) expressed in R7, R8 and L1-L5 neurons. Related to Figure 3.

The number of reads in each range is shown in color. Note, for instance, that for each cell type there are approximately 250 cell surface proteins with levels of expression >5.

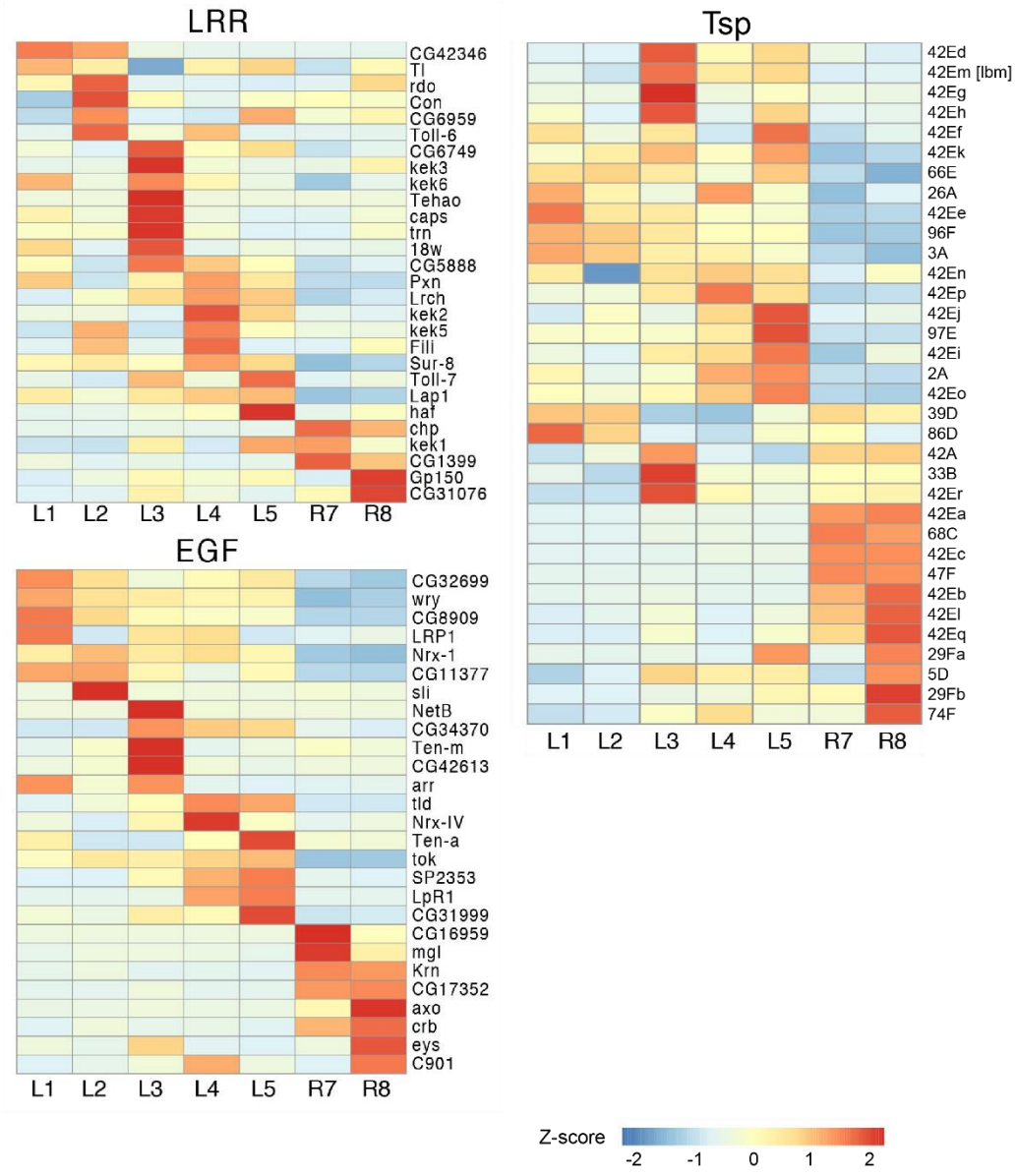
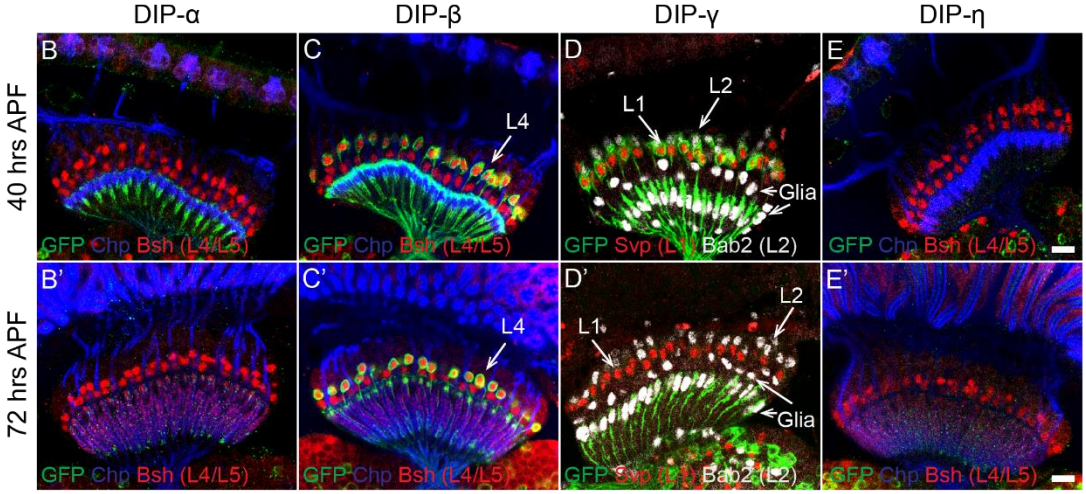


Figure S2-3. Heat maps representing relative expression of different CSM families as indicated. Related to Figure 3. The tetraspanin family of proteins has been implicated in synapse formation. Mutations in *late bloomer* (*lbm*) (Kopczynski et al., 1996) lead to delayed innervation of muscle in the *Drosophila* embryo and mutations in a human tetraspanins are associated with intellectual disability and synapse formation (Bassani et al., 2012). Proteins containing EGF and LRR domains play roles in axon guidance (e.g. netrin) (Timofeev et al. 2012) and synapse formation (e.g. Toll-6,) (Ward et al. 2015) respectively.

A

RPKMs	L1	L2	L3	L4	L5	R7	R8
DIP-α*	0.7	0.8	3.1	0.8	0.6	4.2	3.9
DIP-β*	0.0	0.1	0.1	4.6	0.0	0.0	0.0
DIP-γ*	5.5	0.7	1.3	5.5	0.2	0.9	0.5
DIP-δ*	0.1	0.1	0.1	0.1	0.1	0.1	0.0
DIP-ε	0.1	0.1	0.2	0.1	0.1	0.0	0.0
DIP-ζ	0.1	0.0	0.3	0.2	0.0	0.1	0.1
DIP-η*	0.0	0.1	0.2	0.1	0.1	0.0	0.0
DIP-θ*	0.2	0.2	0.5	0.9	0.1	0.3	0.3
DIP-ι	0.6	0.1	0.2	0.4	0.0	0.1	0.0



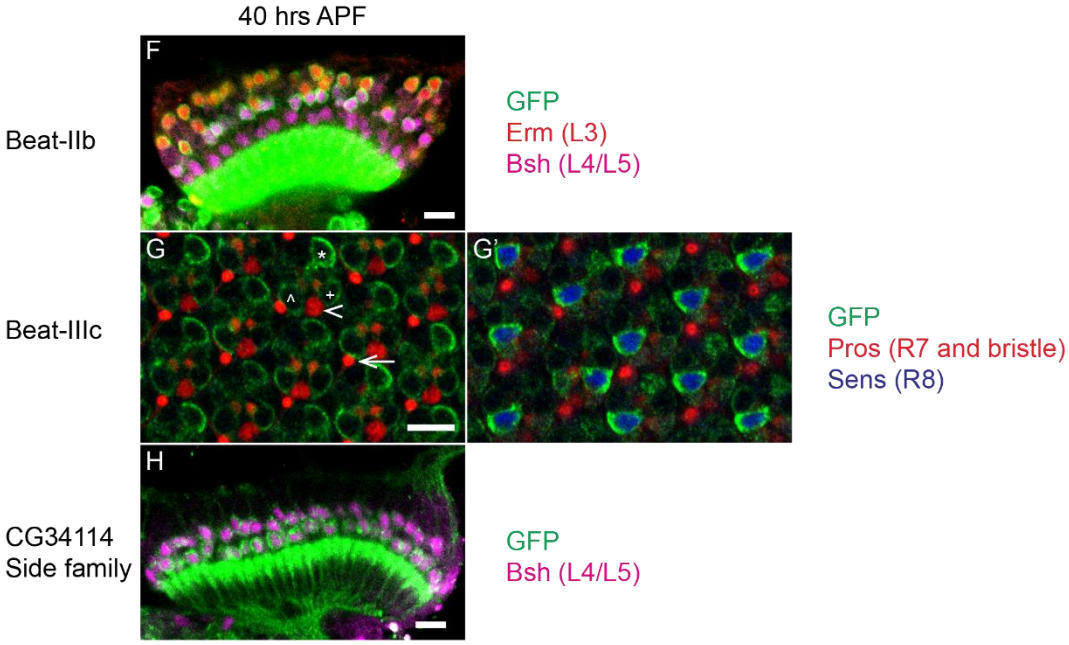


Figure S2-4. Expression of DIP, Beat and Side protein families using MiMIC protein traps. Related to Figure 4. (A) RPKM values of DIP protein families in L1-L5, R7 and R8 at 40 hrs APF. Asterisks indicate DIP genes for which protein trap or Gal4 trap lines are available. (B-E') DIP-α, DIP-β, DIP-γ and DIP-η expression in retina and lamina neurons at 40 hrs and 72 hrs APF using MiMIC derivatives. For DIP-α and DIP-η, MiMIC protein traps were used. For DIP-β and DIP-γ, Gal4 trap derivatives were used in combination with UAS-myr-GFP to visualize DIP expression in lamina neurons. (B, B') DIP-α is not expressed in retina or lamina neurons at 40 and 72 hrs APF. (C, C') DIP-β (green) is specifically expressed in L4 (red) at 40 hrs and 72 hrs APF. (D, D') DIP-γ (green) is expressed strongly in L1 (red) and weakly in L2 (grey) at 40 hrs APF. The expression in L1 and L2 at 72 hrs APF is close to threshold of detection. Although DIP-γ is expressed with an RPKM in L4 similar to the RPKM in L1, we did not detect expression using the MiMIC derivative. DIP-γ is also expressed in glia. (E, E') DIP-η is not expressed in retina or lamina neurons at 40 and 72 hrs APF. Scale bars, 10 µm. (F-H) Beat-IIb expression in lamina neurons and Beat-IIIc expression in retina at 40 hrs APF using MiMIC protein traps. (F) Beat-IIb (green) is specifically expressed in L3 (red) and L4 (upper magenta) in lamina neurons. It is also expressed in glia beneath L5 (lower magenta). The massive green staining is likely to be due to expression in the processes of glial cells within the lamina neuropil. (G, G') Beat-IIIc (green) is specifically expressed in R1 (^), R4 (*), R6 (+) and R8 (blue), but not in R7 (red, arrowhead) in the retina. Bristles (arrow) are also labeled by Pros staining (red). (H) CG34114 (green), a member of the Side protein family is specifically expressed in L4 and L5 (magenta), with higher level in L5 than in L4. The strong staining just beneath these cells is likely to be the proximal region of the L5 axon. Scale bars, 10 µm.

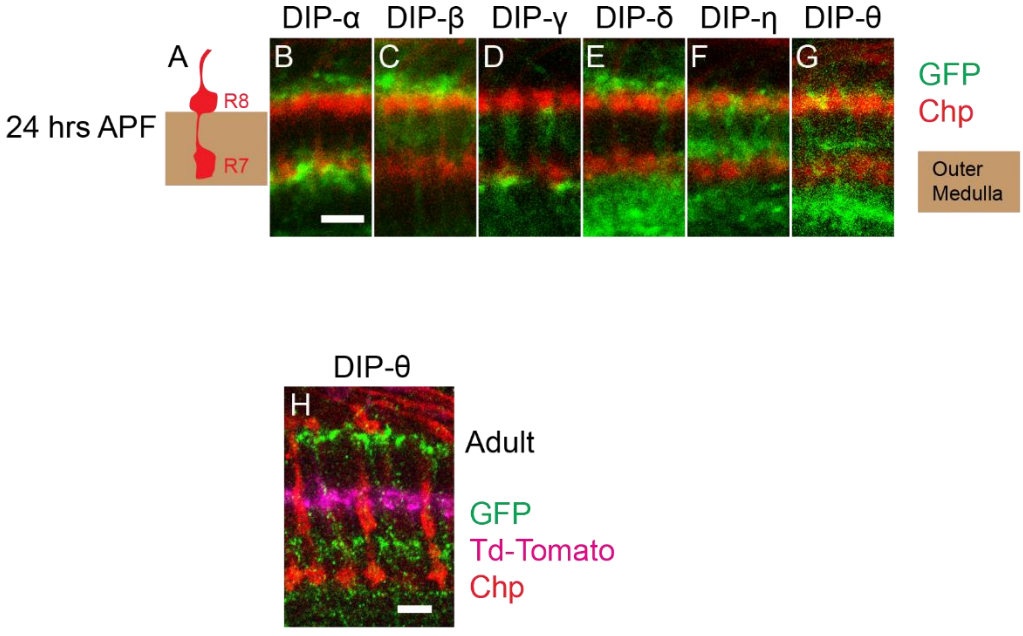


Figure S2-5. DIP protein expression in the medulla. Related to Figure 5. (A-G) Expression in the medulla at 24 hrs APF of the six DIP genes for which protein trap lines are available. They were expressed in neurons exhibiting layer-specific patterns of processes within the outer medulla neuropil. However, prior to synapse formation (i.e. 24 hrs APF), layered patterns are diffuse and overlap. (A) Schematic of R8 and R7 axon morphology in the outer medulla at 24 hrs APF to visualize the boundaries of the outer medulla. (B-G) Protein expression of six DIPs (green) in the outer medulla. The six DIPs are expressed in layer-specific patterns. (H) Expression of DIP-θ in the medulla neuropil in the adult. The axons of Dm3 neurons are labeled with td-tomato (magenta) to mark the M2 and M3 border. Note that the pattern is different from the one at 72 hrs APF (see Figure 5Q); expression at the M2 layer is missing. Layer-specific expression patterns for all other 5 DIPs remained the same in the adult as they were at 72 hrs APF. Scale bars, 5 μm.

Supplemental Tables and Legends

Sample	Total no. of reads	Uniquely mapped (UM) reads		Exonic		Intronic		Intergenic	
		Read No.	(% total)	Read No.	(% UM reads)	Read No.	(% UM reads)	Read No.	(% UM reads)
L1_1	221,006,312	148,723,224	67.3%	369,729,537	63.3%	9,713,039	6.5%	44,867,060	30.2%
L1_2	420,696,072	290,822,341	69.1%	711,518,414	65.1%	21,925,078	7.5%	79,686,317	27.4%
L2_1	365,555,130	226,673,999	62.0%	592,229,130	68.5%	17,757,860	7.8%	53,719,036	23.7%
L2_2	332,657,332	208,233,324	62.6%	540,890,657	74.8%	10,886,378	5.2%	41,553,646	20.0%
L3_1	303,110,842	162,043,016	53.5%	115,714,350	71.4%	13,944,754	8.6%	32,383,912	20.0%
L3_2	251,168,116	130,719,651	52.0%	94,034,900	71.9%	11,855,755	9.1%	24,828,996	19.0%
L3_3	296,900,662	149,077,234	50.2%	88,839,355	59.6%	12,690,643	8.5%	47,547,236	31.9%
L3_4	356,475,762	158,195,099	44.3%	99,036,222	62.6%	16,688,207	10.6%	42,470,670	26.9%
L4_1	308,174,712	112,505,309	36.5%	67,096,676	59.6%	16,358,609	14.5%	29,050,024	25.8%
L4_2	309,755,062	114,510,256	37.0%	81,837,498	71.5%	6,485,738	5.7%	26,187,020	22.9%
L5_1	289,374,230	96,075,572	33.2%	65,161,597	67.8%	10,172,543	10.6%	20,741,432	21.6%
L5_2	366,466,960	200,045,736	54.6%	137,845,618	68.9%	16,182,112	8.1%	46,018,006	23.0%
R7_1	230,988,330	165,914,722	71.8%	110,267,302	66.5%	7,347,038	4.4%	48,300,382	29.1%
R7_2	227,157,836	163,812,834	72.1%	109,919,467	67.1%	7,116,698	4.3%	46,776,669	28.6%
R8_1	441,834,342	327,209,264	74.1%	218,807,624	66.9%	15,306,637	4.7%	93,095,003	28.5%
R8_2	283,842,260	202,625,494	71.4%	142,634,665	70.4%	8,674,083	4.3%	51,316,746	25.3%
Total	3,665,249,114	1,982,734,187		1,331,195,274		142,822,817		508,716,096	

Table S2-1. Alignment statistics of RNA-seq data. Related to Figure 2.

A

	L1	L2	L3	L4	L5	R7	R8
L1	NA						
L2	217	NA					
L3	393	361	NA				
L4	402	352	277	NA			
L5	400	362	337	252	NA		
R7	864	810	916	1159	981	NA	
R8	959	838	803	1087	894	262	NA

B

	L1	L2	L3	L4	L5	R7	R8
L1	NA						
L2	53	NA					
L3	83	96	NA				
L4	80	78	70	NA			
L5	91	81	92	58	NA		
R7	140	136	165	168	157	NA	
R8	146	129	137	139	129	49	NA

Table S2-2. Summary of pair-wise comparisons of transcripts differentially expressed in R7, R8, and L1-L5 neurons. Related to Figure 3. (A) Differentially expressed genes with RPKM value >5 (for at least one neuron in each pair-wise comparison) and a 5X difference between pairs of neurons (see text for results with RPKM values >5) and an adjusted P value of <0.05. (B) Differentially expressed genes encoding CSM genes. The criteria for a differentially expressed gene are an RPKM value >5 for a least one of the pair and at least a 5X difference in RPKM values with an adjusted P value <0.05.

Symbol	ID	L1	L2	L3	L4	L5	R7	R8	Module Labels
FBgn0016070	smg	3.9	2.7	4.7	4.9	3.9	5.3	4.5	0
FBgn0037526	CG10092	2.4	2.5	1.7	2.4	2.7	2.1	2.6	0
FBgn0038058	CG5608	11.8	10.3	10.6	10.5	9.5	9.8	13.0	0
FBgn0038564	CG7785	6.3	6.6	6.4	6.2	6.5	6.2	7.0	0
FBgn0038760	MED25	8.2	10.4	9.9	9.9	8.1	10.6	9.4	0
FBgn0039943	CG17168	63.2	93.7	82.1	93.1	72.2	91.2	79.6	0
FBgn0262481	CG43071	0.3	0.9	0.6	0.0	0.0	0.0	0.6	0
FBgn0263048	CG43343	2.4	4.0	3.7	3.1	3.7	4.0	2.8	0
FBgn0000003	7SLRNA:CR32864	4.0	12.5	40.6	34.0	58.4	119.9	146.7	1
FBgn0000008	a	0.5	0.6	1.6	0.6	3.9	29.6	14.5	1
FBgn0000042	Act5C	1057.5	1140.0	792.2	629.8	703.9	1818.0	1253.1	1
FBgn0000064	Ald	8.9	6.7	27.1	10.5	12.7	83.3	80.3	1
FBgn0000083	AnnIX	13.2	22.4	27.1	20.5	20.3	117.4	64.6	1
FBgn0000109	Aprt	26.8	50.5	62.4	31.9	43.8	111.3	109.3	1
FBgn0000116	Argk	170.8	239.2	258.1	202.3	313.9	352.3	304.2	1
FBgn0000120	Arr1	0.0	0.2	0.0	0.0	0.0	1.0	0.6	1
FBgn0000121	Arr2	2.3	0.6	18.6	3.1	1.2	317.0	540.0	1
FBgn0000139	ash2	17.7	21.2	18.7	17.1	24.4	42.4	33.1	1
FBgn0000150	awd	547.7	1391.6	1306.9	641.2	1029.0	4175.4	3443.6	1
FBgn0000163	baz	0.2	0.2	0.4	0.3	0.3	4.2	3.3	1
FBgn0000173	ben	145.2	154.8	145.8	112.6	114.0	254.4	275.5	1
FBgn0000179	bi	0.5	0.6	2.7	1.9	0.9	5.9	5.1	1
FBgn0000206	boss	1.4	0.6	11.0	2.3	2.1	142.0	166.2	1
FBgn0000228	Bsg25D	5.8	4.3	4.9	4.9	4.9	7.2	7.1	1
FBgn0000239	bur	8.3	8.5	7.6	8.2	9.5	11.4	10.5	1
FBgn0000246	c(3)G	1.0	0.8	2.3	0.9	0.8	1.9	2.7	1
…	…	…	…	…	…	…	…	…	…
FBgn0051602	tRNA:CR31602	0.0	0.0	0.0	0.1	0.6	0.0	0.0	24
FBgn0051712	CG31712	196.3	181.7	234.0	305.1	379.8	103.1	108.6	24
FBgn0051816	CG31816	0.0	0.1	0.3	0.0	1.7	0.5	1.0	24
FBgn0052598	betaNACtes6	0.0	0.0	0.2	0.0	0.7	0.1	0.1	24
FBgn0085757	CR40621	2.8	2.0	10.3	14.0	20.4	3.8	3.9	24
FBgn0085759	CR40640	4.5	3.3	21.1	19.3	62.9	19.8	35.4	24
FBgn0085760	CR40641	0.1	0.1	47.1	47.9	152.0	0.3	0.5	24
FBgn0085761	CR40642	0.0	0.0	10.3	8.7	19.5	0.1	0.1	24
FBgn0085777	CR40959	3.2	2.7	31.1	49.6	71.9	11.4	7.5	24
FBgn0085795	CR41535	1.3	1.5	83.2	69.8	142.2	0.8	1.7	24
FBgn0085796	CR41539	1.2	1.2	13.2	10.0	19.5	0.5	0.7	24
FBgn0085799	CR41544	5.7	4.8	18.3	17.4	56.3	19.8	33.4	24
FBgn0085807	CR41590	2.4	1.4	10.0	14.6	21.3	2.4	4.2	24
FBgn0085815	CR41605	0.3	0.2	20.2	30.6	47.3	0.4	0.6	24
FBgn0085817	CR41607	1.5	0.8	20.0	27.3	44.1	2.2	4.0	24
FBgn0085818	CR41608	0.2	0.3	0.6	0.8	1.4	0.1	0.1	24
FBgn0085822	CR41613	0.0	0.1	22.7	24.1	81.9	0.3	0.6	24
FBgn0085828	CR42195	0.1	0.1	9.5	14.6	23.2	0.3	0.5	24
FBgn0260440	spdo	0.3	0.1	0.3	0.1	0.8	0.3	0.5	24
FBgn0263707	CG43659	0.0	0.1	0.1	0.1	0.5	0.0	0.0	24

Table S2-3. RPKM values for the genes in the 25 distinct co-expression modules identified in the WGCNA analysis. Related to Figure 3 and Figure S1. Table shown here is a shorter version because the original table is too long.

ID	symbol	L1	L2	L3	L4	L5	R7	R8
FBgn0040726	dpr	3.5	1.9	282.0	298.4	9.1	40.2	10.0
FBgn0261871	dpr2	9.5	0.4	12.2	0.6	0.3	0.4	1.0
FBgn0053516	dpr3	21.6	0.8	116.9	1.0	0.4	0.4	0.4
FBgn0053512	dpr4	42.5	24.7	17.0	48.1	13.7	1.6	0.8
FBgn0037908	dpr5	18.3	5.4	56.6	44.6	3.3	5.9	3.8
FBgn0040823	dpr6	0.2	59.6	89.8	11.4	16.4	1.3	2.6
FBgn0053481	dpr7	141.9	154.0	191.1	156.3	178.2	138.4	115.8
FBgn0052600	dpr8	4.0	72.4	426.2	185.3	436.6	38.8	88.0
FBgn0038282	dpr9	2.1	13.5	3.8	8.0	13.4	25.4	39.3
FBgn0052057	dpr10	118.5	51.7	66.0	52.0	39.2	12.0	2.4
FBgn0053202	dpr11	0.2	4.8	0.9	0.8	1.3	9.4	2.7
FBgn0085414	dpr12	0.7	37.1	4.5	94.5	0.4	1.3	4.3
FBgn0034286	dpr13	0.7	118.3	2.7	16.1	0.4	1.7	3.5
FBgn0029974	dpr14	35.3	27.1	34.3	35.9	37.5	1.4	0.9
FBgn0037993	dpr15	43.5	0.1	6.8	0.4	0.0	0.0	0.1
FBgn0037295	dpr16	0.9	0.3	26.1	0.5	0.3	1.3	1.0
FBgn0051361	dpr17	0.6	0.7	354.4	2.0	0.4	1.3	0.7
FBgn0030723	dpr18	2.2	0.7	2.1	1.2	1.7	0.7	2.3
FBgn0032233	dpr19	30.6	22.0	24.3	21.7	22.7	26.0	27.1
FBgn0035170	dpr20	52.7	22.4	3.0	3.8	38.8	1.4	1.3
FBgn0260995	dpr21	82.2	121.2	82.0	47.5	97.6	119.5	45.6

Table S2-4. RPKM values for the 21 dpr transcripts in R7, R8 and L1-L5. Related to Figure 3. Note Dpr18 RPKM values are considerable below 5 RPKM in all cell types and thus the only Dpr we consider not expressed in our data set.

Medulla Neurons Test																
Dip-GFP																
Dip-α	Dm1	Dm4	Dm6	Dm12	Dm14	Dm17	Dm18	Dm19	Mi1	Mi4	Tm1	Tm2	Tm3	Tm4	Tm9	Tm20
Dip-δ	Dm1		Dm6		Dm14	Dm17	Dm18	Dm19	Mi1	Mi4	Tm1	Tm2	Tm3	Tm4		Tm20
Dip-η									Mi1	Mi4		Tm2 ?	Tm3 ?	Tm4		
Dip-θ							Dm18		Mi1	Mi4		Tm2	Tm3	Tm4		

Table S2-5. Summary of Co-localization Experiments. Related to Figure 6. DIP protein traps were tested for co-localization with a set of GAL4 lines driving marker expression in the indicated medulla neurons (see Text). Medulla neuron classes tested for each DIP-GFP are shown. Neuron in red indicates that the cell expresses the DIP and black indicates that it does not. Note question marks for DIP-η to indicate that staining was very weak.

Supplemental Experimental Procedures

Fly Stocks

The following stocks were used:

Genetic labeling of specific cell types for FACS: For experiments involving the isolation of R7 and R8 neurons, GMR-myr-tdtom, a general marker for retinal cells, was utilized in combination with cell-specific GFP markers for R7 and R8 neurons. The sens-F2Bshort-deltaE1-GFP transgene (Pepple et al., 2008) was used to specifically label R8 neurons in the retina. This marker also labels a subset of lamina neurons. To label R7 neurons we generated split GAL4 constructs as previous described (Luan et al., 2012; Pfeiffer et al., 2008, 2010). The R44F08 fragment was used to drive expression of GAL4DBD (Jenett et al., 2012), and the 465bp *pros* core-enhancer (Hayashi et al., 2008) was used to drive the expression of p65AD. Reconstituted GAL4 activity activated the expression of UAS-GFP-tagged RpL10 (McEwen, Zhang and Zipursky, unpublished). To isolate different lamina neurons, L1-L5 neurons were labeled using the pan-lamina driver R27G05-nlsLexAGADfl in su(Hw)attP2 (Janelia Research Campus; GAL4 pattern described in Riddiford et al., 2010) and LexAop-myr::tdTom, and particular subtypes were simultaneously labeled with cell-specific GFP reporters. L3, L4 and L5 neurons were labeled using 9-9-GAL4, *apterous*-GAL4, and 6-60-GAL4 (Nern et al., 2008) to drive expression of UAS-mCD8-GFP, respectively. For isolation of L3 neurons the nuclear marker UAS-H2AGFP (gifts from Barret Pfeiffer and Gerald Rubin (Janelia Research Campus)) was used in combination with UAS-mCD8-GFP. Svp-GAL4 (Kyoto: 103727) and Bab1-GAL4 (BDSC stock #47736. P{GMR73D08-GAL4}attP2) were used to label L1 and L2 neurons, respectively, through expression of UAS-mCD8-GFP.

Expression of Dpr, Beat, and Side proteins using MiMIC-derivatives: Cellular expression patterns were assessed in co-labeling experiments using MiMIC protein trap derivatives (GFP) and antibodies against nuclear proteins expressed specifically in R7, R8 or particular lamina neuron subtypes (L1-L5). The following MiMIC protein trap derivatives were used: MI03102 (*beatIIb*), MI03726 (*beatIIIc*), MI01052 (*CG34114*, *side* family), MI02201 (*dpr1*), MI02530 (*dpr2*), MI05963 (*dpr3*), MI01358 (*dpr6*), MI03557 (*dpr10*), MI01695 (*dpr12*), MI05577 (*dpr13*), MI01408 (*dpr15*), and MI08707 (*dpr17*). GFP expression from the original MiMIC line MI02231 was used to visualize the expression pattern for *dpr11*. The following antibodies were used to assess expression in specific cell types (see below for concentrations): anti-prospero (R7), anti-Senseless (R8), anti-Seven-up (L1), anti-Bab2 (L2), anti-erm (L3), anti-Bsh (L4, L5), anti-apterous (L4), anti-Pdm3 (L5).

Cellular expression of DIPs using MiMIC-derivatives: The following MiMIC protein trap derivatives (GFP) were used: MI02031 (DIP-α, CG32791), MI08287 (DIP-δ, CG34391), MI07948 (DIP-η, CG14010) and MI03191 (DIP-θ, CG31646). To visualize expression of DIPs β and γ MiMIC GAL4 derivatives of MI01971 (DIP-β, CG42343) and MI03222 (DIP-γ, CG14521) were used to drive expression of UAS-myr-GFP. GFP expression from the original MiMIC line MI03222 was also used to assess expression of DIP-γ. In these experiments Dm3 neurons were labeled using R25F07-LexAp65 (Janelia Research Campus) and LexAop-myr::tdTom. The expression of DIPs α, β, δ and θ in subsets of medulla neurons was assessed in co-labeling experiments using MiMIC protein trap lines (see above) and the following cell-type specific GAL4 drivers: Mi1 (R19F01), Mi4 (R72E01), Dm1 (R22D12), Dm4 (R23G11), Dm6 (R38H06), Dm12 (R47G08), Dm13 (R38A07), Dm14 (R47E05), Dm17 (VT43152), Dm18 (VT028450), Dm19 (VT024602), Tm1 (R74G01), Tm2 (R71F05), Tm3 (R59C10), Tm4 (R53C02), Tm9 (R24C08),

and Tm20 (R33H10). All of these driver lines are from the Janelia Research Campus (R) (Jenett et al., 2012) and Vienna Drosophila Resource Center (VT) (Kvon et al., 2014) collections. Lines for individual cell types were identified and characterized as described for Dm neuron markers in (Nern et al., 2015). Each GAL4 line was crossed to pJFRC21-10XUAS-IVS-mCD8::RFP in attP18 (Pfeiffer et al., 2010) to label the cell bodies and processes of a specific medulla neuron type.

Bioinformatics

After the step of quality control, we first filtered out raw reads with low quality and containing sequencing adapters and then mapped reads (pair-end, 50bp in length) to the D. melanogaster reference genome (release FB2013_01) with the gapped aligner Tophat (Trapnell et al., 2009) with the default setting. The fly gene model was downloaded from the Ensembl database (version of Drosophila_melanogaster.BDGP5.73.gtf) and supplied to Tophat as the reference genome annotation. Only reads uniquely aligned were collected. In total for all libraries sequenced, 1,982,734,187 reads were uniquely mapped (corresponding to an overall mappability of 57%) and used for further analysis. The expression levels of genes were quantified using RPKM units (Reads Per Kilobase of exon per Million reads mapped) using customized scripts written in Perl after normalization based on the geometric means as described in DESeq (Anders and Huber, 2010). The accession number for the raw data reported in this paper is GEO: GSE68235.

Differential expression analysis was performed using the packages, DESeq (Anders and Huber, 2010) and edgeR (Robinson et al., 2010) in R (http://www.R-project.org). The original p-values were corrected by the false discovery rate (FDR) for multiple testing errors. In addition to the FDR of <0.05, we considered differentially expressed genes with >5X differences. Thus, in summary, we considered genes as differentially expressed if: 1. the adjusted p-value was less than 0.05; 2.

the expression ratio between two samples was >5X; 3. the maximal RPKM value for at least one group in the comparison was >5; and 4. there was agreement between DESeq and edgeR.

We performed a principal component analysis (PCA) on normalized read counts of all samples to compare gene expression under different developmental stages using the 'prcomp' function in R. In this analysis, we selected 500 top-ranked genes based on their variations across all samples using the function 'rowVars' in R matrixStats package. PCA revealed that our samples were clearly distinguishable by both the first and second principal component (PC1, 45.1% of the total variation; PC2, 28.7% of the total variation).

We performed weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008) to identify neuron-cell type specific modules. This identified co-expression modules by clustering transcripts that exhibit similar expression patterns as revealed through the analysis of all samples. To further understand the cell type specificity of the modules, we correlated the identified module eigengenes with traits/cell types represented as the theoretical expression patterns for all cell types in a binary fashion.

To identify cell surface and secreted membrane molecules (CSMs), we used a gene list established by Kai Zinn and co-workers (Kurusu et al., 2008). In brief, to define genes encoding CSMs that might be relevant to cell recognition during neural development, the fly proteome was searched with sequences of every domain in the "extracellular" portion of the SMART domain database (http://smart.embl-heidelberg.de/browse.shtml). A total of more than 80 domain types had representatives in *Drosophila*. Several hundred proteins from this list were excluded. These included members of large groups containing proteins with almost identical structures, including small chitin-binding proteins, single-domain serine proteases, single-domain C-type lectins,

protease inhibitors, and others thought to be unlikely to play important roles in cell-type specific recognition. In addition, the list did not include ion channels, pumps, transporters, secreted enzymes, and a variety of other classes of CSMs in the database. The final CSM cell-recognition database contains 976 proteins.

Immunohistochemistry

Pupal brains were dissected in PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4) and fixed in PBL (4% paraformaldehide, 75mM lysine, and 37mM sodium phosphate buffer, pH 7.4) for 25 min at room temperature (RT). After several rinses with PBT (PBS 0.5% Triton-X10) at RT, samples were incubated in PBT containing 10% normal goat serum (blocking solution) for at least 1hr at RT. Brains were incubated overnight at 4^{o}C in primary and secondary antibodies for at least one day each with multiple blocking solution rinses at RT in between and afterwards. Incubations were extended for up to 5 days for some antibody combinations to increase the quality of the signal. Brains were mounted in EverBrite mounting medium (Biotium).

The following primary antibodies were used in this study: chicken-anti-GFP (1:1000, Abcam ab13970); rabbit-anti-DsRed (1:200, Clontech 632496); mouse-anti-Seven-up (Kanai et al., 2005) (1:20, a gift from Yasushi Hiromi); rat-anti-Bab2 (1:500, a gift from Frank Laski); rabbit-anti-Erm (Janssens et al., 2014)(1:50, a gift from Cheng-yu Lee); rabbit-anti-Ap (1:5000, a gift from Claude Desplan); guinea pig-anti-Pdm3 (1:500, a gift from John Carlson); guinea pig-anti-Bsh (1:200, generated in the Zipursky lab); mouse-anti-Pros (1:20, MR1A from DSHB); guinea pig-anti-Sens (1:1000, a gift from Hugo Bellen); mouse-anti-24B10 (Zipursky et al., 1984)(1:20, DSHB), rat-anti-Elav (1:500, 7E810 from DSHB).

Secondary antibodies against chicken, rabbit, mouse, guinea pig and rat conjugated to Alexa Fluor -488, -555, -568, -647 or Cy5 with the following references and concentrations were used. From Life Technologies: A11039 (1:1000); A31572 (1:800); A11031 (1:500); A11011 (1:500); A11075 (1:500); A21235 (1:500) and A21450 (1:500). From Jackson ImmunoResearch: 112-175-143 (1:500) and 712-607-003 (1:200).

Supplemental References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Bassani, S., Cingolani, L.A., Valnegri, P., Folci, A., Zapata, J., Gianfelice, A., Sala, C., Goda, Y., and Passafaro, M. (2012). The X-linked intellectual disability protein TSPAN7 regulates excitatory synapse development and AMPAR trafficking. Neuron *73*, 1143–1158.

Hayashi, T., Xu, C., and Carthew, R.W. (2008). Cell-type-specific transcription of prospero is controlled by combinatorial signaling in the Drosophila eye. Development *135*, 2787–2796.

Janssens, D.H., Komori, H., Grbac, D., Chen, K., Koe, C.T., Wang, H., and Lee, C.-Y. (2014). Earmuff restricts progenitor cell potential by attenuating the competence to respond to self-renewal factors. Development *141*, 1036–1046.

Jenett, A., Rubin, G.M., Ngo, T.T.B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., et al. (2012). A GAL4-Driver Line Resource for Drosophila Neurobiology. Cell Rep. *2*, 991–1001.

Kanai, M.I., Okabe, M., and Hiromi, Y. (2005). Seven-up controls switching of transcription factors that specify temporal identities of drosophila neuroblasts. Dev. Cell *8*, 203–213.

Kopczynski, C.C., Davis, G.W., and Goodman, C.S. (1996). A neural tetraspanin, encoded by late bloomer, that facilitates synapse formation. Science *271*, 1867–1870.

Kurusu, M., Cording, A., Taniguchi, M., Menon, K., Suzuki, E., and Zinn, K. (2008). A Screen of Cell-Surface Molecules Identifies Leucine-Rich Repeat Proteins as Key Mediators of Synaptic Target Selection. Neuron *59*, 972–985.

Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics.

Luan, H., Diao, F., Peabody, N.C., and White, B.H. (2012). Command and Compensation in a Neuromodulatory Decision Network. J. Neurosci. *32*, 880–889.

Nern, A., Zhu, Y., and Zipursky, S.L. (2008). Local N-cadherin interactions mediate distinct steps in the targeting of lamina neurons. Neuron *58*, 34–41.

Nern, A., Pfeiffer, B.D., and Rubin, G.M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. Proc. Natl. Acad. Sci. U. S. A. *112*, E2967–E2976.

Pepple, K.L., Atkins, M., Venken, K., Wellnitz, K., Harding, M., Frankfort, B., and Mardon, G. (2008). Two-step selection of a single R8 photoreceptor: a bistable loop between senseless and rough locks in R8 fate. Development *135*, 4071–4079.

Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.-T.B., Misra, S., Murphy, C., Scully, A.,

Carlson, J.W., Wan, K.H., Laverty, T.R., et al. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. Proc. Natl. Acad. Sci. U. S. A. *105*, 9715–9720.

Pfeiffer, B.D., Ngo, T.-T.B., Hibbard, K.L., Murphy, C., Jenett, A., Truman, J.W., and Rubin, G.M. (2010). Refinement of tools for targeted gene expression in Drosophila. Genetics *186*, 735–755.

Riddiford, L.M., Truman, J.W., Mirth, C.K., and Shen, Y.-C. (2010). A role for juvenile hormone in the prepupal development of Drosophila melanogaster. Development *137*, 1117–1126.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Timofeev, K., Joly, W., Hadjieconomou, D., and Salecker, I. (2012). Localized netrins act as positional cues to control layer-specific targeting of photoreceptor axons in Drosophila. Neuron *75*, 80–93.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Ward, A., Hong, W., Favaloro, V., and Luo, L. (2015). Toll receptors instruct axon and dendrite targeting and participate in synaptic partner matching in a Drosophila olfactory circuit. Neuron *85*, 1013–1028.

Zipursky, S.L., Venkatesh, T.R., Teplow, D.B., and Benzer, S. (1984). Neuronal development in the Drosophila retina: monoclonal antibodies as molecular probes. Cell *36*, 15–26.

Chapter 3. Recent result updates

Cognate Dprs and DIPs are expressed by synaptic partners throughout the fly visual system

In the previous section I identified DIP-expressing neurons by crossing flies carrying a DIP-GFP to a panel of GAL4 marker lines for specific medulla neurons and assessed co-localization of the markers. Although this method provides a quick and unambiguous way of identifying neurons, it is not a practical way to generate a complete list of medulla neurons expressing each DIP.

The multi-color flip out (MCFO) method combined with Gal4 derivatives of DIP MiMICs (available for all DIPs, except for DIP-iota (Diao et al., 2015)), is a more efficient way to identify all medulla neurons expressing each DIP (Figure 3-1) (Nern et al., 2015). The DIP-Gal4 in combination with conditional FLP-mediated excision of stop cassettes results in stochastic expression of different combinations of MCFO reporters in scattered DIP-expressing cells. This allowed visualization of individual DIP-expressing neurons in different colors and identification by their morphologies. Through collaboration with a postdoc in our lab, I have done MCFO experiments for DIP-α, DIP-β, DIP-ζ, DIP-ε, DIP-η and DIP-θ, and I'm still in the process of completing the experiment for other DIPs and identifying all the neurons expressing each DIP.

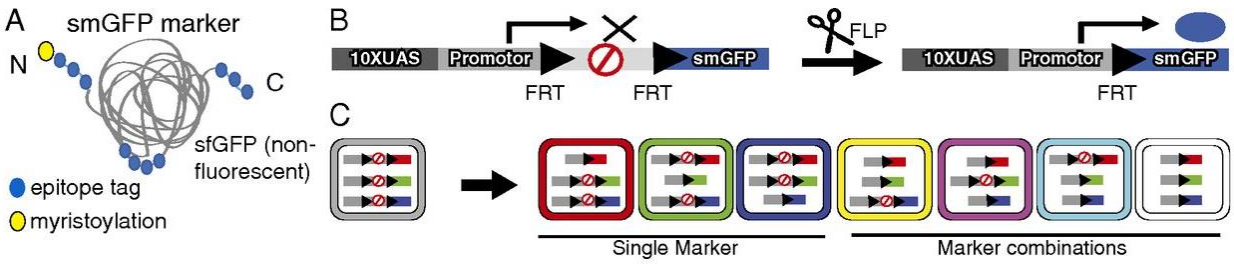


Figure 3-1. Schematic of multi-color flip out (MCFO). (A) Schematic of smGFP markers. The backbone of myristoylated (yellow circle) superfolder GFP (gray) is inserted with groups of multiple copies of a single epitope tag (HA, FLAG, MYC, V5, or OLLAS; blue circles). The insertion of the tags eliminates the endogenous fluorescence; (B) Schematic of an MCFO reporter with 10XUAS and a core promoter for GAL4-activated expression, a transcriptional terminator flanked by FRT sites, and a smGFP marker. The marker is expressed upon the excision of the terminator by Flp-recombinase. (C) Potential color combinations of markers with three MCFO reporters bearing smGFPs inserted with different epitope tags: unlabeled (gray), one marker (red, green, blue), or combinations of two (yellow, magenta, cyan) or three (white) labels. Adapted from Nern et al., 2015.

Results from MCFO experiments demonstrate that each DIP is expressed by approximately five to twenty different subclasses of medulla neurons, as well as different types of neurons in lobula and lobula plate in the fly visual system. The partial list assembled for identified DIP-expressing medulla neurons show that Dm1, Dm4, Dm12 and Lawf1 express DIP-α, consistent with the colocalization data, and nine subclasses of neurons plus about ten other subclasses of unidentified medulla neurons express DIP-ε. Notably, most DIP-expressing medulla neurons only express one particular DIP. Tm3 and several other types of TmY neurons, however, express both DIP-η and DIP-θ. Interestingly, these paralogs are tandemly arranged and interact with many of the same Dprs. Consistent with our previous observations, many of the newly identified types of neurons expressing each DIP are also synaptic partners of lamina neurons or photoreceptors. Furthermore, matching of cognate Dpr and DIP expression was found for most synaptic partners of lamina neurons (data not shown). Other types of identified DIP-expressing medulla neurons only make synapses with medulla neurons, and since Dprs are also expressed by medulla neurons (also lobula and lobula plate neurons), it would be important to assess whether cognate Dprs are also expressed in their synaptic partners.

Genetic analysis suggest a role of DIP-Dpr interaction for survival of neurons

To study the function of cognate Dpr and DIP, I generated null mutants for DIPs and Dprs using CRISPR-Cas9 mediated gene knock-out. To do this, plasmids encoding gRNAs targeting two sites in the gene of interest were injected into flies carrying vasa-Cas9. Typically, I selected target sites close to translation start site and 40-500 bp away from each other within the coding exon of the target gene. A double stranded break at each of the two sites mediated by Cas9 nuclease results in deletion of the regions between them. This usually generates frame shift mutations causing early stop of the protein close to the translational start site and thus results in null mutants of the target gene. I have created such mutants for seven of the nine DIPs and some Dprs, and I am in the process of completing all the DIPs and will then generate nulls in Dprs. In collaboration with other two postdocs in the lab, Shuwa Xu and Qi Xiao, we assessed the phenotype of synaptic pairs expressing three groups of cognate DIP and Dpr in corresponding DIP and Dpr mutants.

Preliminary results on two of the three groups of cognate Dpr and DIP suggest that Dpr-DIP interaction plays a role in the development of DIP-expressing neurons. Dpr11 is expressed in a discrete subset of R7 neurons which form synapses with a subset of Dm8 neurons expressing DIP-γ in the M6 layer. There is a reduction in number of Dm8 neurons in both in DIP-γ and Dpr11 null mutant. However, no phenotype in R7 axon morphology or cell number is seen in these mutants (data not shown). Our data on Dm8 loss and R7 terminal are inconsistent with the study from Zinn and colleagues (Carrillo et al., 2015). In their study, they observed reduction of Dm8 with a similar penetrance to our data using DIP-γ MiMIC line in combination with a deficiency line uncovering DIP-γ genomic region. However, they did not see a reduction of Dm8 using Dpr11 Kyoto line in combination with a deficiency line uncovering Dpr11 genomic region. One possible explanation

for this discrepancy is that the Dpr11 Kyoto line is not a null allele, while the DIP-γ MiMIC line is, at least, very close to a null allele. Furthermore, Zinn and colleagues assessed the phenotype of R7 terminals by over-expressing Brp-short fused with tandem tomato as a presynaptic marker, and observed modest overshoots of R7 terminals in both Dpr11 and DIP-γ loss of function mutants using Dpr11 and DIP-γ MiMICs. In my experiments, I assessed the R7 phenotype using the STaR technique (Chen et al., 2014), in which terminals are visualized using cell-type specific epitope tagging of Brp, at endogenous levels, in R7. No phenotypes in R7 terminal and general Brp distribution along R7 terminal were observed. Given subtle effects in Zinn experiments and our results using a more physiologically appropriate genetic background, I think the previously described R7 terminal phenotype should be considered questionable.

A similar phenomenon is seen for multiple synaptic pairs, with one neuron expressing both Dpr6 and Dpr10 and its partner expressing DIP-α. Dpr6 and Dpr10 are expressed in L3 neurons and DIP-α is expressed in two of its synaptic partners, Dm4 and Dm12, in the M3 layer. In addition, Dm1, which also expresses DIP-α, forms synapses in the M1 layer with L2 and L5 and they, in turn, express Dpr6 and Dpr10. Removal of DIP-α or both Dpr6 and Dpr10, in combination, results in a reduction in the number of all three subclasses of Dm neurons. By contrast to the effect on DIP-expressing cells, no phenotypes in the corresponding lamina neurons were observed (data not shown). We also assessed phenotype on synaptic pairs expressing Dpr12 and DIP-δ. Surprisingly, no phenotypes on cell number or layer specificity were seen for neurons expressing DIP-δ. Therefore, preliminary genetic experiments demonstrate that two groups of cognate Dpr and DIP pairs are required for survival of DIP-expressing neurons. Further developmental analysis is critical to ascertain the mechanism of DIP-Dpr interaction in neuronal survival. Moreover, it is crucial to assess the phenotype on synaptic connectivity between Dpr and DIP expressing synaptic

pairs in corresponding mutants, in order to address their role in synaptic specificity. In addition, phenotypic analysis on synaptic pairs expressing other cognate Dprs and DIPs is important to reveal the general role of DIP-Dpr interactions in neural circuit assembly.

References

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., and Zinn, K. (2015). Control of Synaptic Connectivity by a Network of Drosophila IgSF Cell Surface Proteins. Cell *163*, 1770–1782.

Chen, Y., Akin, O., Nern, A., Tsui, C.Y.K., Pecot, M.Y., and Zipursky, S.L. (2014). Cell-type-specific labeling of synapses in vivo through synaptic tagging with recombination. Neuron *81*, 280–293.

Diao, F., Ironfield, H., Luan, H., Diao, F., Shropshire, W.C., Ewer, J., Marr, E., Potter, C.J., Landgraf, M., and White, B.H. (2015). Plug-and-Play Genetic Access to Drosophila Cell Types using Exchangeable Exon Cassettes. Cell Rep. *10*, 1410–1421.

Nern, A., Pfeiffer, B.D., and Rubin, G.M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. Proc. Natl. Acad. Sci. *112*, E2967–E2976.

Chapter 4. Discussion

Combinatorial use of RNA seq and other methods identifies candidate recognition proteins regulating synaptic specificity

Over the past three decades, several experimental approaches have been taken to identify cell recognition molecules regulating circuit assembly. These include genetic screens (Seeger et al., 1993; Zallen et al., 1998), molecular screens using monoclonal antibodies (Kolodkin et al., 1992), biochemical assays (Cheng et al., 1995; Drescher et al., 1995; Luo et al., 1993; Schmucker et al., 2000; Serafini et al., 1994), differential mRNA analysis (Kopczynski et al., 1996) and candidate gene approaches (Cheng et al., 1995). These studies have provided fundamental insights into diverse molecular mechanisms and, in some cases, the underlying logic regulating circuit formation. However, only limited studies have succeeded in identifying cell surface proteins regulating synaptic specificity.

The newly developed high-throughput RNA sequencing (RNA seq) enables deep profiling of gene expression in population of cells. This provides an effect approach to addressing various questions in neuroscience. For instance, RNA seq has been widely used in assessing neuronal identity and identifying new neuronal cell types and subtypes (Fuzik et al., 2015; Handel et al., 2016; Macosko et al., 2015; Cadwell et al., 2015; Zeisel et al., 2015; Usoskin et al., 2015). It has also been applied to revealing roles of proteins in regulating gene expression (Lee et al., 2016), understanding relationship between functional brain activity and gene expression, and identifying odorant receptors responding to specific odors *in vivo* (Jiang et al., 2015). Despite the widespread application of RNA seq to various questions in neurobiology, there has been little effort directed towards using this technique to identify cell surface proteins regulating synaptic specificity.

The availability of cell type specific markers for purifying defined subclasses of neurons at the appropriate developmental stage in the fly visual system provides an excellent opportunity to identify candidate proteins regulating synaptic specificity using RNA seq. In my thesis research, I performed RNA seq on multiple highly related neuronal cell types with distinct synaptic specificities, lamina neuron L1-L5 and photoreceptor R7 and R8, at the onset of synapse formation. By comparing levels of gene expression between highly related neurons with distinct synaptic specificities, differences in gene expression between cell types are more closely related to synaptic specificity than other general aspects of cell type differentiation. In addition, genes regulating synaptic specificity are more likely to be discovered by doing RNA seq at the onset of synapse formation. From RNA seq data, I showed that at the onset of synapse formation, neurons express hundreds of genes encoding cell surface and secreted proteins. In addition, many of the genes encoding these proteins are expressed in a cell-type enriched fashion, that is, their expression levels are at least 5X higher in one cell type, than in one or more of the other cell types that were profiled. Thus, each of the cell types we profiled expresses a unique signature of hundreds of genes encoding cell surface and secreted proteins at the onset of synapse formation.

Genes encoding several families of cell surface proteins were differentially expressed in different neuronal cell types from RNA seq analysis. These include Ig superfamily proteins (e.g. Beat protein family, Side protein family, Dscams and IRM proteins), LRR proteins (e.g. Toll, Connectin, and Capricious) and EGF-domain containing proteins (e.g. Netrin, Ten-a and Ten-m), which have been shown to play important roles in cellular recognition in the developing nervous system. For instance, Beat-Side interactions guide motor neurons to their muscle targets in the developing embryo (Siebert et al., 2009). Interestingly, the most dramatic and complex pattern of expression was observed for the Dpr genes with 17 of the 21 paralogs expressed in a cell-type enriched fashion.

To confirm the expression pattern of Dpr proteins *in vivo*, I used MiMIC-derived protein trap lines (Nagarkar-Jaiswal et al., 2015) and co-stained the tissue at the stage we did RNA seq with antibodies against GFP and cell type specific transcription factors for each lamina neuron. The staining results largely confirmed Dpr expression patterns *in vivo*.

The striking expression pattern of Dpr proteins and the extensive protein interaction network of them determined by Garcia and colleagues (Özkan et al., 2013) prompted us to explore the expression patterns of the DIPs in further detail. By contrast to the Dprs, most DIPs were not expressed, or were expressed at only very low levels in lamina or photoreceptor neurons. Localization of DIP expression using protein traps revealed that each DIP is expressed by medulla neurons projecting neurites to specific layers in the medulla. Through co-localization with available cell-type specific markers, I demonstrated that each DIP is expressed by only a subset of medulla neurons with processes within a layer or a unique combination of layers. By matching neurons expressing cognate DIP and Dprs using EM reconstruction data (Takemura et al., 2008, 2013, 2015), I found that neurons expressing DIPs are specific subsets of synaptic partners of cells expressing interacting Dprs (Tan et al., 2015). These findings raise the possibility that Dpr/DIP interactions specify synaptic contacts between lamina neurons and subsets of their synaptic partners within discrete layers in the medulla. Therefore, expression studies combined with the cell surface interactome, cell type specific markers, and synaptic connectome in the visual system provide a multipronged approach to identify candidate cell surface and secreted proteins regulating synaptic specificity during neural circuit formation.

Correlation of molecular and cellular complexity of medulla circuits

One surprising observation from the RNA seq data is that even with a very conservative RPKM threshold for gene expression (RPKM>5), neurons that we profiled still express between 247 and 322 genes encoding cell surface and secreted proteins at the onset of synapse formation, corresponding to between a quarter to a third of the potential cell surface proteins encoded by the fly genome. Each neuronal cell type expresses a unique combination of them. Indeed, between 49 and 168 genes encoding cell surface and secreted proteins are differentially expressed by at least 5X in at least one pairwise comparison between neurons we profiled. Interestingly, many of the differentially expressed proteins have known interacting partners. For instance, of the 23 Ig superfamily proteins differing in expression by more than 5X between L1 and L2, all but three have known interacting partners. Furthermore, by matching interacting proteins between two neuron pairs, L1-L4 and L2-L4 (Hong et al., 2012; Özkan et al., 2013; Paré et al., 2014), where L1 and L2 make equal area of contacts with L4 in lamina but only L2 make synapses with L4, I found that each neuron pair has 30-40 pairs of interactions mediated by cell surface proteins expressed on them, and among these interactions each neuron pair exhibits ~10 unique pairs of interactions which could potentially contribute to the synaptic specificity between L2 and L4. Therefore, the level of molecular complexity in developing neurons is remarkably high.

The high level of molecular complexity in developing neurons mirrors the cellular complexity in the medulla circuit. Indeed, a striking feature of the synaptic connectome in a medulla column is its extraordinary complexity (Takemura et al., 2013, 2015), with synapses between the processes of over 100 different neuronal cell types (A. Nern, personal communication). Even once processes from a particular type of neuron get into their target layer in the medulla, they still must select

specific synaptic partners from several dozens of neurons with processes in that layer. For instance, from the most recent seven column reconstruction data (Takemura et al., 2015), 30 different subclasses of neurons are identified that make synapses (> 5 synapses) in M2 layer, yet L2 only make synapses with 11 of them. Furthermore, the number of connections between neurons does not correlated with their relative areas of contact. These studies support the idea that many different cell recognition molecules are likely to be necessary for neurons to discriminate between appropriate and inappropriate synaptic partners.

Interactions mediated by cognate Dprs and DIPs in synaptic partners may provide a general molecular strategy for synaptic specificity

Families of cell recognition molecules could, in principle, provide general molecular strategy regulating synaptic specificity with different family members promoting interactions between different synaptic partners. Two families, the Dscam1 in insects and the clustered protocadherins in vertebrates, have emerged as promising candidates. In both families, complex genomic loci encode large sets of proteins that mediate homophilic binding and are expressed in combinatorial patterns by individual neurons. This raised the possibility that they could directly specify patterns of synaptic specificity through a lock and key mechanism. However, as Dscam1 is largely expressed in a probabilistic manner throughout the fly nervous system (Miura et al., 2013), and protocadherins also appear to be expressed in this way, they are mainly used in neurons to discriminate self vs. non-self, rather than matching synaptic partners.

By contrast, Dprs and DIPs, with complex pattern of heterophilic interactions, are expressed in the fly visual system in a highly deterministic fashion. Each of the DIPs is expressed in a unique subset of medulla neurons, and consistently expressed in these neurons throughout development.

Interestingly, most DIP-expressing neurons identified so far only express one DIP. By contrast, some of the Dprs are expressed in a dynamic fashion during development in lamina neurons and photoreceptors, and each neuronal cell type expresses multiple Dprs. Importantly, matching of cognate Dpr and DIP expression is found in most synaptic partners of lamina neurons and photoreceptors during synapse formation. There are, however, still many other types of medulla neurons expressing Dprs and DIPs that have not been identified yet. As such further studies may expand the matching of Dprs and DIPs in synaptic partners throughout the fly visual system. Therefore, the expression of cognate Dprs and DIPs in synaptic partners raises the possibility that their interaction plays a general role in synaptic specificity during circuit assembly.

Preliminary genetic analysis demonstrated that removal of DIP and cognate Dprs result in loss of DIP-expressing neurons, but not changes in layer-specific targeting of DIP and Dpr expressing neurons. That the same phenotypes were observed in two of the three groups of cognate Dpr and DIP proteins that we have studied so far suggests that these two families of proteins play a common role in the development of DIP-expressing neurons in the medulla. These proteins are not required for axon guidance or layer specific targeting. What that role is remains unclear.

One possible explanation to the phenotypes is that these proteins could play a direct role in cell type specific trophic support, as previously described in Jeb/Alk signaling (Pecot et al., 2014). That is removal of Dpr and DIP interaction results in direct loss of trophic signaling between synaptic partners and subsequent cell death. Over past several decades, studies have demonstrated that neurotrophins (e.g. NGF, BDNF, NT-3 and NT-4), are secreted proteins that play a major role in neuron survival, through interacting with their main tyrosine kinase receptors (TrkA, TrkB, TrkC) and p75NTR (Ichim et al., 2012). For example, in the developing mammalian peripheral

nervous system, sympathetic neurons require nerve growth factor (NGF) for survival. NGF is produced in limiting amounts by the targets innervated by sympathetic neurons, and binds to its specific tyrosine kinase receptor, TrkA, on the surface of the innervating axons (Glebova and Ginty, 2005). The NGF-TrkA signaling in sympathetic neurons promotes growth, hence neurons that fail to bind sufficient NGF undergo naturally occurred cell death. Similar to the interactions between neurotrophins and their receptors, the complex interaction between Dpr and DIP families and their cell type specific expression patterns suggest that DIP-Dpr interaction could regulate circuit assembly by providing cell-type specific trophic support. Different Dpr-DIP interactions could provide trophic support to specific subsets of synaptic partners throughout the fly visual system, such that wild type synaptic partners survive and are able to make synapses, and mutant neurons undergo cell death to avoid making synapses with the wrong partners. This selective cell death would minimize the negative effects on the processing of visual information. By contrast to the DIP-α and DIP-γ cell loss phenotypes, cell loss was not observed for DIP-δ expressing neurons in DIP-δ and Dpr12 null mutants. Here the DIP/Dpr interaction may regulate another process or alternatively parallel pathways may act in these cells to provide trophic support.

Dprs and DIPs could regulate synaptic specificity. Neurons lacking the appropriate DIPs and Dprs may make synapses with the wrong partners or they may not make synapses. Inappropriate connectivity could either lead indirectly to neuronal cell death for some neurons but not others. Consistent with the indirect death model, a previous study in vertebrates demonstrated that incorrectly projecting retinal ganglion cells are preferentially eliminated during the period of naturally occurring cell death (O'Leary et al., 1986). This explanation would be consistent with the phenotypic analysis for some DIP/Dpr pairs in some neurons. Further developmental analysis

of different Dpr/DIP pairs is critical to ascertain the mechanism of DIP-Dpr interaction in regulating the formation of neural circuits.

Based on these and previous studies, we envision a two-step model for regulating the formation of layer-specific connections within the medulla. Previous studies showed that at early stages of medulla development, growth cones of lamina neurons target to overlapping regions with broadly expressed adhesive (i.e. N-cadherin) and repulsive cell surface molecules (i.e. Plexin/Semaphorin signaling). Growth cones then segregate into discrete domains through interactions between signals localized to specific layers as the medulla matures (Pecot et al., 2013). The complex interactome and expression pattern of Dpr and DIP proteins and the preliminary genetic analysis in Dpr and DIP mutants provides a potential general molecular strategy for wiring, that once growth cones get into their target layers, they use these two families of proteins as an IgSF code to specify synaptic connections within a layer in the medulla and, by implication, in other regions of the fly visual system.

The two-step model for synaptic connectivity in the medulla was significantly influenced by models for layer specificity in the analogous structure in the mouse retina, the inner plexiform layers (IPL). Intriguing similarities are shared between these two models. In the mouse IPL, cadherin and semaphorin/plexin proteins direct neuronal processes to layers (Duan et al., 2014; Matsuoka et al., 2011). In the second step, Ig superfamily proteins then promote matching of synaptic partners within layers. Important support for this second step comes from recent elegant study from Sanes and colleagues, demonstrating that homophilic interactions between Sdk2 proteins, an Ig superfamily protein expressed by a specific synaptic pair of amacrine and retinal ganglion cell neurons, is necessary for synapses between them (Krishnaswamy et al., 2015). This

may represent one example of a general strategy for synaptic-specificity in the vertebrate retina, as studies in the chick retina by Yamagata and Sanes (Yamagata et al., 2002, 2008, 2012) revealed that Sidekicks, Dscams and Contactins are widely expressed by neurons forming synapses in specific sub-laminae in chick IPL. Thus, the studies in the mouse IPL and the medulla region of the fly allude to a common strategy for achieving synaptic specificity.

Furthermore, Dprs and DIPs are likely to be only a part of the story of synaptic specificity in the medulla, given the remarkable complexity of synaptic connections in it (Takemura et al., 2013, 2015), with synapses between the processes of >100 neuronal cell types (A. Nern, personal communication). This cellular complexity is reflected by the unique combination of hundreds of cell surface and secreted molecules expressed by each of the photoreceptor and lamina neurons profiled in this study. How the molecular complexity contributes to specificity remains elusive, but the convergence of improved histological, genetic, physiological, behavior and molecular tools promises to provide important insights into the molecular recognition strategies controlling synaptic specificity during formation of neural circuits.

References

Cadwell, C.R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., Reimer, J., Shen, S., Bethge, M., Tolias, K.F., et al. (2015). Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. Nat. Biotechnol. *34*.

Cheng, H.J., Nakamoto, M., Bergemann, a D., and Flanagan, J.G. (1995). Complementary gradients in expression and binding of ELF-1 and Mek4 in development of the topographic retinotectal projection map. Cell *82*, 371–381.

Drescher, U., Kremoser, C., Handwerker, C., Löschinger, J., Noda, M., and Bonhoeffer, F. (1995). In vitro guidance of retinal ganglion cell axons by RAGS, a 25 kDa tectal protein related to ligands for Eph receptor tyrosine kinases. Cell *82*, 359–370.

Duan, X., Krishnaswamy, A., De La Huerta, I., and Sanes, J.R. (2014). Type II cadherins guide assembly of a direction-selective retinal circuit. Cell *158*, 793–807.

Fuzik, nos, Zeisel, A., Calvigioni, D., Yanagawa, Y., Szab, bor, Linnarsson, S., and Harkany, T. (2015). Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. Nat. Biotechnol. *34*.

Glebova, N.O., and Ginty, D.D. (2005). Growth and Survival Signals Controlling Sympathetic Nervous System Development.

Handel, A.E., Chintawar, S., Lalic, T., Whiteley, E., Vowles, J., Giustacchini, A., Argoud, K., Sopp, P., Nakanishi, M., Bowden, R., et al. (2016). Assessing similarity to primary tissue and cortical layer identity in induced pluripotent stem cell-derived cortical neurons through single-cell transcriptomics. Hum. Mol. Genet. *25*, 989–1000.

Hong, W., Mosca, T.J., and Luo, L. (2012). Teneurins instruct synaptic partner matching in an olfactory map. Nature *484*, 201–207.

Ichim, G., Tauszig-Delamasure, S., and Mehlen, P. (2012). Neurotrophins and cell death. Exp. Cell Res. *318*, 1221–1228.

Jiang, Y., Gong, N.N., Hu, X.S., Ni, M.J., Pasi, R., and Matsunami, H. (2015). Molecular profiling of activated olfactory neurons identifies odorant receptors responding to odors in vivo.

Nat. Neurosci. *18*, 1446–1454.

Kolodkin, A.L., Matthes, D.J., O'Connor, T.P., Patel, N.H., Admon, A., Bentley, D., and Goodman, C.S. (1992). Fasciclin IV: sequence, expression, and function during growth cone guidance in the grasshopper embryo. Neuron *9*, 831–845.

Kopczynski, C.C., Davis, G.W., and Goodman, C.S. (1996). A neural tetraspanin, encoded by late bloomer, that facilitates synapse formation. Science *271*, 1867–1870.

Krishnaswamy, A., Yamagata, M., Duan, X., Hong, Y.K., and Sanes, J.R. (2015). Sidekick 2 directs formation of a retinal circuit that detects differential motion. Nature *524*, 466–470.

Lee, J.A., Damianov, A., Lin, C.H., Fontes, M., Parikshak, N.N., Anderson, E.S., Geschwind, D.H., Black, D.L., and Martin, K.C. (2016). Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. Neuron *89*, 113–128.

Luo, Y., Raible, D., and Raper, J.A. (1993). Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones. Cell *75*, 217–227.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

Matsuoka, R.L., Nguyen-Ba-Charvet, K.T., Parray, A., Badea, T.C., Chédotal, A., and Kolodkin, A.L. (2011). Transmembrane semaphorin signalling controls laminar stratification in the mammalian retina. Nature *470*, 259–263.

Miura, S.K., Martins, A., Zhang, K.X., Graveley, B.R., and Zipursky, S.L. (2013). Probabilistic

splicing of Dscam1 establishes identity at the level of single neurons. Cell *155*, 1166–1177.

Nagarkar-Jaiswal, S., Lee, P.-T., Campbell, M.E., Chen, K., Anguiano-Zarate, S., Cantu Gutierrez, M., Busby, T., Lin, W.-W., He, Y., Schulze, K.L., et al. (2015). A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila. Elife *4*, 1–28.

O'Leary, D.D., Fawcett, J.W., and Cowan, W.M. (1986). Topographic targeting errors in the retinocollicular projection and their elimination by selective ganglion cell death. J. Neurosci. *6*, 3692–3705.

Özkan, E., Carrillo, R. a., Eastman, C.L., Weiszmann, R., Waghray, D., Johnson, K.G., Zinn, K., Celniker, S.E., and Garcia, K.C. (2013). An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. Cell *154*.

Paré, A.C., Vichas, A., Fincher, C.T., Mirman, Z., Farrell, D.L., Mainieri, A., and Zallen, J.A. (2014). A positional Toll receptor code directs convergent extension in Drosophila. Nature *515*, 523–527.

Pecot, M., Chen, Y., Akin, O., Chen, Z., Tsui, C.Y.K., and Zipursky, S.L. (2014). Sequential axon-derived signals couple target survival and layer specificity in the drosophila visual system. Neuron *82*, 320–333.

Pecot, M.Y., Tadros, W., Nern, A., Bader, M., Chen, Y., and Zipursky, S.L. (2013). Multiple interactions control synaptic layer specificity in the Drosophila visual system. Neuron *77*, 299–310.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C. a, Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell *101*, 671–684.

Seeger, M., Tear, G., Ferres-Marco, D., and Goodman, C.S. (1993). Mutations affecting growth cone guidance in Drosophila: genes necessary for guidance toward or away from the midline. Neuron *10*, 409–426.

Serafini, T., Kennedy, T.E., Galko, M.J., Mirzayan, C., Jessell, T.M., and Tessier-Lavigne, M. (1994). The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. Cell *78*, 409–424.

Siebert, M., Banovic, D., Goellner, B., and Aberle, H. (2009). Drosophila motor axons recognize and follow a sidestep-labeled substrate pathway to reach their target fields. Genes Dev. *23*, 1052–1062.

Takemura, S., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P.K., Katz, W.T., Olbris, D.J., Plaza, S.M., Winston, P., et al. (2013). A visual motion detection circuit suggested by Drosophila connectomics. Nature *500*, 175–181.

Takemura, S.Y., Lu, Z., and Meinertzhagen, I. a. (2008). Synaptic circuits of the Drosophila optic lobe: The input terminals to the medulla. J. Comp. Neurol. *509*, 493–513.

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., et al. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. Proc. Natl. Acad. Sci. U. S. A. *112*, 13711–13716.

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., et al. (2015). Ig Superfamily Ligand and Receptor Pairs Expressed in Synaptic Partners in Drosophila. Cell *163*, 1756–1769.

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P. V, et al. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing Dmitry. Nat. Neurosci. *18*, 145–153.

Yamagata, M., and Sanes, J.R. (2008). Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. Nature *451*, 465–469.

Yamagata, M., and Sanes, J.R. (2012). Expanding the Ig superfamily code for laminar specificity in retina: expression and role of contactins. J. Neurosci. *32*, 14402–14414.

Yamagata, M., Weiner, J.A., and Sanes, J.R. (2002). Sidekicks: Synaptic adhesion molecules that promote lamina-specific connectivity in the retina. Cell *110*, 649–660.

Zallen, J., Yi, B. a, and Bargmann, C.I. (1998). The conserved immunoglobulin superfamily member SAX-3/Robo direts multiple aspects of axon guidance in C. Elegans. Cell *92*, 217–227.

Zeisel, A., Munoz-Manchado, A.B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science (80-.). *347*, 1138–1142.