A systems analysis of neural connectivity and functionality in *Caenorhabditis elegans*

Christopher Aaron Brittin

School of Computing

University of Leeds

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

December 2017

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement

The right of Christopher Aaron Brittin to be identified as Author of this work has been asserted by Christopher Aaron Brittin in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgements

I would like to first thank my supervisor Netta Cohen for providing guidance and thoughtful discussion during the course of this PhD. Also many thanks to Ian Hope for allowing me to work in his lab. I would like to thank the Leeds International Research Scholarship for funding my postgraduate studies.

I would also like to thank my lab-mates Elpiniki Kalogeropoulou, David Pertab and Robert Holbrook who graciously offered their expertise in molecular biology and helped me pick up the necessary bench skills for this research.

The first part of this thesis has been carried out by a team which has included Steven Cook (SC) and myself. My own contributions, fully and explicitly indicated in the thesis, have been conception of the project, data collection, and data analysis. The other members of the group and their contributions have been as follows: SC helped with data collection.

The second part of this thesis has been carried out by a team which has included Netta Cohen (NC), Jung-uk Shim (JS), Jinyang Chung (JC) and myself. My own contributions, fully and explicitly indicated in the thesis, have been conception of the project, maintenance and engineering of worm strains, construction of the fluid control system, software for image acquisition and analysis and data analysis. The other members of the group and their contributions have been as follows: NC conceived the project and helped with data analysis, JS and JC designed and fabricated the microfluidics device and JC collected and analyzed the data.

I would like to thank my parents, Melanie and Dennis, for their support through the year. Finally, I would like to thank my partner Skye Beare who provided endless support during this PhD.

Abstract

Part I considers the stereotypic patterns of synaptic connectivity in neural circuits, referred to as wiring specificity. Two aspects of wiring specificity are lamina specificity - placing synaptic partners in close spatial proximity - and synaptic specificity - making the appropriate synaptic connections among many physically adjacent neurons. Combinatorial expression patterns of surface proteins could be used to uniquely label neurons for the purposes of synaptic specificity. To test this model in the worm, the C. elegans nerve ring was volumetrically reconstructed from serial sectioned legacy electron micrographs, which provides detailed spatial and morphological information of neural processes in the nerve ring. Comparing the spatial data with the synaptic wiring diagram shows that the C. elegans nerve ring exhibits both lamina-like specificity and synaptic specificity. Combinatorial expression patterns of CAM genes in nerve ring neurons were analyzed. If a few key conditions regarding gene expression are satisfied, then the number of known combinatorial CAM expression patterns is sufficient to uniquely label neurons and can account for more than 90% of the observed synaptic specificity.

Part II develops a new rotatable microfluidic device for simultaneously imaging calcium activity in bilateral neuron pairs within the same animal. Typically when imaging calcium activity in *C. elegans*, only the left or right side of the animal is imaged. This is due to the natural orientation of the worm which places neurons on the left and right side in different focal planes. Our new device allows the animal to be rotated, placing cells on the right and left side of the animal in the same focal plane.

Contents

List of Figures

Chapter 1

Introduction

This thesis is motivated by the notion that a biological organism is more than the sum of its parts. As scientists, we are trained to break a system apart into its component pieces, study each piece individually, and then aggregate that information to better understand how the system works. The problem with this, particularly in a biological system, is that the behavior of an individual component likely depends on the behavior of other components in the system. In a perspective article, Marder & Taylor (2011) argued that biological insights may not come from considering system components independently but rather by looking at the interaction of the individual components. With this in mind, this works tries to apply a systems approach to two different problems related to the nervous system of the nematode *Caenorhabditis elegans* (*C. elegans*). Part I performs a top-down analysis of wiring specificity in the *C. elegans* nerve ring, specifically addressing if different gene expression patterns across neurons can explain how neurons choose synaptic partners. Part II develops a novel tool for simultaneously imaging calcium activity in left/right bilateral neurons, a step forward for understanding how neural activity of individual neurons is correlated. While Part I and II are treated individually, in keeping with the systems approach, they should also be viewed as being complementary. Part I considers the development of a biological neural network while Part II considers the activity of that network.

Part I

Wiring specificity in the *C. elegans* nerve ring

Chapter 2

Wiring specificity in nervous system development

The brain's 'wiring' allows us to coordinate movement, store memories and learn. The human brain has on the order of 86 billion neurons with an individual neuron making and receiving hundreds to thousands of synaptic contacts. Understanding how these contacts are negotiated at the network level in order to create functioning neural circuitry is of fundamental importance. What are the general underlying mechanisms implemented by an individual neuron to discriminate synaptic partners from the many physically adjacent neighboring neurons? How do the general mechanisms vary across different neurons in order to arrive at a cohesive neural network? Finally, how robust and with what degree of precision are these mechanisms implemented during the development of a neural network?

Wiring specificity is the stereotypic patterns of synaptic connectivity in a neural circuit. Understanding wiring specificity requires solutions to three main questions. First, how do neurons physically make contact with potential synaptic partners? A neuron can only synapse onto a cell at points of physical contact. Therefore, neurons must be placed in close spatial proximity to synaptic partners. This requires guiding axon and dendrite growth towards the appropriate anatomical region where neurons can come into physical contact with target cells. Directing neurons to the correct anatomical region is referred to as lamina specificity. Second, how do neurons identify synaptic partners from the myriad of neighboring cells? Neurons only make synapses with a fraction of the neighboring

cells with which they make physical contact (Hamos *et al.*, 1987). This suggests that neurons are able to ascertain the identities of their neighboring cells and correctly identify synaptic partners. Choosing synaptic partners from neighboring cells is referred to as synaptic specificity. Finally, once synapses are created, how are the synapses maintained or if need be eliminated? Construction of a neural network is an ongoing process and does not terminate after initial synaptic connectivity has been established. During network activity, frequently used synapses are strengthened and maintained while infrequently used synapses are eliminated (Yogev & Shen, 2014). Some organisms may even undergo complete nervous system rewiring during certain developmental phases (Walthall *et al.*, 1993; White *et al.*, 1978). In this way, the network is constantly being resculpted in order to adapt to changing environmental conditions and to minimize the overall metabolic cost of maintaining neural network activity. This study will focus on lamina and synaptic specificity with no further discussion on synapse maintenance and elimination. Readers interested in synapse maintenance and elimination are directed to following excellent reviews (Shen & Scheiffele, 2010; Yogev & Shen, 2014).

2.1 Cell adhesion molecules mediate wiring specificity

It has long been postulated that the molecular diversification of cell adhesion molecules provides specific surface identities to neurons (de Wit & Ghosh, 2015). Classic work by Langley (Langley, 1895) and Sperry (Sperry, 1963) showed that nerve fibers regenerate with striking specificity. They proposed the "chemoaffinity hypothesis which states that neurons possess unique cytochemical labels that allow neurons to selectively navigate to their target cells (Meyer, 1998). In order to regulate wiring specificity and synaptic diversity, such surface labels would need to be expressed in distinct neuronal populations, act in *trans* with membrane binding partners and be sufficiently diverse to confer unique identities to both cells and synapse. Genomic and proteomic analysis has identified several superfamilies of cell adhesion protein that meet these requirements. Of these,

the most extensively documented are the immunoglobulin superfamily (IgSF), the Leucine-Rich Repeat (LRR) proteins, cadherin family members and neurexins. These superfamilies are very briefly discussed here with references to more extensive reviews provided.

Immunoglobulin superfamily

The IgSF proteins, characterized by the presence of an Ig homology domain, are the largest and most diverse superfamily found in both vertebrates and invertebrates (de Wit & Ghosh, 2015; Shimono *et al.*, 2012). The Ig domains possess a characteristic Ig-fold, formed by two anti-parallel beta-sheets (Barclay, 2003). IgSF proteins are able to bind specifically to other proteins making them ideal components for cell-surface receptors and cell adhesion molecules. Members of IgSF have been shown to play a role in axon pathfinding, synapse formation, neuronal axon and soma adhesion, axonal maintenance and neurotransmitter receptor clustering (Carrillo *et al.*, 2015; Yogev & Shen, 2014).

Leucine-Rich Repeat (LRR) proteins

The LRR is a 20-30 amino acid long structural motif and is one of the most common domain repeats across organisms. LRRs have a flexible structure that allows them to interact with a large number of diverse ligands, making it a versatile protein interaction motif (Bella *et al.*, 2008). Many extracellular LRR (eLRR) genes are expressed in the nervous system and exhibit specific expression patterns (de Wit *et al.*, 2011). LRR proteins have been shown to regulate axon guidance, synapse specificity, axon myelination and neural circuit stability.

Cadherin superfamily

The cadherin superfamily consists of over 100 transmembrane glycoproteins that can be grouped in several subfamilies, of which the classic cadherins and the protocadherins have been the most extensively studied in relation to synaptic connectivity (Angst *et al.*, 2001; de Wit & Ghosh, 2015; Takeichi, 2007). Their extracellular domains contain repeated domains called cadherin repeats that contain sequences for calcium binding. Cadherins have been shown to play a role in

cell adhesion, cell-cell recognition, cytoskeletal organization, signal transduction and growth control.

Neurexins

Neurexins are presynaptic cell adhesion molecules that play a key role in neural circuit assembly and restructuring by interacting with various pre- and postsynaptic ligands (Südhof, 2017). The vertebrate neurexin family only consists of three genes (*Nrxn1-Nrxn3*), each of which has two promoters for generating long (α) and short (β) domains. Neurexins exhibit extensive alternative splicing and the theoretical number of variants ranges in the thousands (Ullrich *et al.*, 1995). However, only hundreds have been observed experimentally (Schreiner *et al.*, 2014). Neurexins are primarily expressed by neurons and localized to synapses (Ushkaryov *et al.*, 1992) and are abundantly produced by astrocytes (Zhang *et al.*, 2014). Neurexins are expressed early in development before synapse formation and are thought to play a role in synthesizing synaptic proteins in preparation of synaptic assembly (Daly & Ziff, 1997).

2.2 Lamina specificity

Lamina specificity requires spatially organizing neurons so that synaptic partners come into physical contact. Organization of neuropile into distinct anatomical regions appears to be a common feature of the nervous system. The vertebrate neocortex, olfactory bulb and visual system all exhibit a stereotyped multi-layered structure (Baier, 2013; Gilmore & Herrup, 1997; Nagayama *et al.*, 2014; Sanes & Zipursky, 2010). The layered structures can take the form of continuous planar lamina (the neocortex and visual system) or as spherical glomeruli (the olfactory system). The layered structures aggregate synapses with similar functional properties to restricted anatomical regions. The vertebrate retina is a good illustration of synaptic aggregation.

The vertebrate retina has six main cell types: photoreceptors (rods and cones), projection neurons (retina ganglion cells [RGCs]), three types of interneurons (horizontal cells [HC], amacrine cells [AC] and bipolar cells [BC]) and glial cells

(Muller glia). The retina is organized into layers with either cell bodies (outer and inner nuclear layer and the ganglion cell layer) or synapses (the inner and outer plexiform layers). In the outer plexiform layer (OPL), BCs and ACs receive synapses from the photoreceptors. In the inner plexiform layer (IPL) the RGCs receive synapses from the BCs and ACs. Connectivity in the IPL determines the ON/OFF response properties RGCs have to light, which is crucial for correct vision. At a few micrometers thick, the IPL has a high synaptic density (approximately one synapse per μm^3, Helmstaedter *et al.* (2013)), making locating and identifying the correct synaptic partners a challenge. This problem is partially addressed by further organizing the IPL into sublamina where specific synapses are formed.

The sublamina specificity of the IPL is in part mediated by a combination of both homophilic and repulsive interactions of cell adhesion molecules (Baier, 2013; Shen & Scheiffele, 2010). In the chick retina, three families of immunoglobulin superfamily (IgSf) adhesion molecules - Dscams, Sidekicks (Sdk) and Contactins (Cnts) (Yamagata & Sanes, 2008, 2012)- are expressed in non-overlapping sets of the BCs, ACs and RGCs and concentrated at synaptic sites. Pre- and postsynaptic neurons expressing the same molecule arborize to the same sublamina. Reducing expression levels in the RGCs leads to dendrites losing sublamina confinement. Ectopic expression of new adhesion proteins cause RGC dendrite and afferent processes to mistarget new sublamina (Yamagata & Sanes, 2008). Collectively, these observations suggest the sublamina specificity is achieved through homophilic interactions. In the mouse retina, the transmembrane protein semaphorin 6A (Sema6A) and its receptor plexin A2 (PlexA2) are required for the correct radial arborization and stratification of the starburst amacrine cells (SAC) (Sun *et al.*, 2013). PlexA2 is expressed in both ON and OFF SACs, while the Sema6A is only expressed on ON SACs. Mutant mice lacking Sema6A do not exhibit the ON/OFF SAC stratification (ON SAC grow in both the ON and OFF layers) nor the ON SAC radial arborization (ON SAC dendrite branches cross paths) observed in wild type. These results suggest that sublamina specificity is also achieved through repulsive interactions.

Lamina specificity appears to be an evolutionary conserved principle for the structural organization of neural circuits. The *Drosophila* visual system also ex-

hibits the sublamina specificity discussed above (Baier, 2013; Sanes & Zipursky, 2010). The first three components of the fly visual system are the retina, lamina and medulla, which contain approximately 750 column-like units called ommatidia, cartridge and column, respectively. Each ommatidia is composed of eight photoreceptors called retinula (R) cells. There are three classes of R cells – R1-R6, R7 and R8 – each of which respond to different wavelengths of light (think of the R cells as the rods and cones of the fly visual system). Unlike the vertebrate retina, there are no synapses in the fly retina. Instead, R1-R6 neurons project axons to the lamina, the structure immediately beneath the retina. Lamina neurons (L1-L5) along with R7 and R8 project axons onto distinct layers within the medulla where they synapse onto interneurons and transmedullary neurons which project onto the lobula complex. The medulla is divided into 10 layers, with R7, R8 and L1-L5 arborizing and making synaptic contacts in one or a few of the outer six layers, M1-M6. R7 and R8 synapses from a single ommatidia and L1-L5 synapses from a single cartridge are generally restricted to a single medulla column.

As in the vertebrate retina, CAMs are used to mediate sublamina specificity in the *Drosophila* visual system. The classical cadherin N-cadherin mediates the medulla specificity of R7 and select lamina cells Lee *et al.* (2001); Nern *et al.* (2008). R7 cells lacking N-cadherin terminate in the M3 rather than the M6 layer. Conversely, loss of function of N-cadherin in L3 causes the the cell to target M6 rather than M3. N-cadherin mediated homophilic interactions also promote the growth of L5 branches along the growth cone of L2 from the M1 to M2 layer. The leucine-rich repeat (LRR) protein Capricious is sufficient to promote specificity of R8 (Shinza-Kameda *et al.*, 2006). Capricious is expressed in R8 and M3, the target of the R8 growth cone. R7 does not express Capricious, but misexpression in R7 results in targetting of M3. This suggests that Capricious acts via homophilic interactions.

Understanding lamina specificity could provide insights into the developmental mechanisms that drive wiring specificity, which could ultimately help illuminate the fundamentals of connectivity patterns and the nature of neural computation. Lamina specificity has been observed across organisms suggesting that it is a fundamental mechanism for spatially organizing neuropile (Sanes & Zipursky,

2010). This shows that there is value in understanding lamina specificity in simpler organisms, because any developmental mechanism identified in the simple organism may be applicable to more complex organisms.

This section has focused on the lamina organization of the retina as a model for lamina specificity, but it is also worth noting that nervous systems also use other mechanisms to direct neuron projections. For example, transient nonpartner cells, called guidepost cells, can act as intermediate targets for axon guidance events. Such cells may be neuronal cells, Cajal Retzius cells and the GABAergic interneurons in the hippocampus (Sanes & Yamagata, 1999), or non-neuronal cells, glia cells (Barres, 2008). Further discussion of guidepost cells can be found in Section 2.5, where the wiring specificity of the model organism *C. elegans* is discussed.

2.3 Synaptic specificity

Neurons form synapses with only a fraction of the neurons that they physically touch. An anatomical analysis of a single X-cell axon in the lateral geniculate nucleus showed that only 4 of 43 neighboring neurons received synapses from the X-cell (Hamos *et al.*, 1987). This raises the question of how do neurons identify synaptic partners among neighboring cells? One particularly attractive hypothesis is that multigene families of adhesion molecules with distinct binding properties are either combinatorially or differentially expressed within a neuron population thereby giving each individual neuron a distinct molecular identity (Südhof, 2017; Zipursky & Sanes, 2010). If true, then a large number of neurons could be uniquely identified either by a single gene or a relatively small gene cluster. Two gene families, Dscams in insects and protocadherins in vertebrates, have emerged as promising molecular candidates for mediating synaptic specificity.

In the Dscam1 gene, alternative exons at three positions in the dscam RNA have the possibility to generates 19,000 different ectodomain isoforms (Wojtowicz *et al.*, 2007). Expression of 17,000 of the isoforms has been confirmed by high-throughput sequencing (Zipursky & Sanes, 2010) each with isoform-homophilic binding specificity (Sawaya *et al.*, 2008). Moreover, it has been found that single cells express 10-50 distinct isoforms and that splice isoforms being stochastically

expressed, with the same cell types expressing different Dscam1 isoforms (Neves *et al.*, 2004). Given the large number of variants, the stochastic expression would leave every cell with a unique Dscam1 signature.

The clustered protocadherins in vertebrates also encode a large diversity of proteins and has been proposed as the molecular code for neuron individuality in the brain (Morishita & Yagi, 2007; Yagi, 2012). In mice, exons that encode the extracellular and transmembrane domain are arranged in three groups (α, β and γ) that generate 58 protocadherins variants (Wu & Maniatis, 1999). Protocadherins variants are generated by the alternative use of separate promoters upstream of each ectoderm. Alternative splicing joins an α or γ ectoderm/transmembrane exon to a set of three constant exons in the group (Zipursky & Sanes, 2010). While expression levels and patterns vary among isoforms, the overall impression is that isoforms are broadly expressed throughout the nervous system and that the expression overlaps across regions. Single-cell RT-PCR analysis of Purkinje cells has provided evidence for stochastic, combinatorial expression in individual cells (Esumi *et al.*, 2005). The parallels with Dscam1 suggest that protocadherins could define the molecular identify of neurons in the vertebrate nervous system.

2.4 The potential role of timing in wiring specificity

Timing plays a critical role in nervous system development. Timing of progenitor cell division is tightly regulated, where neurons from a common pool of progenitors will adopt distinct cell fates according to their birth order (Kao & Lee, 2010). Birth-dating and cell lineage studies have shown that the *Drosophilla* nerve cord, retina and cerebral cortex are populated by distinct cell types in a precise temporal sequence (Toma *et al.*, 2016). Studies have shown that this temporal sequencing is regulated by both internal and external neural stem cell cues. Heterochronic transplant studies of neural progenitor cells of the ferret cortex showed that the sensing of enviromental cues was dependent on the age of the cell (Desai & McConnell, 2000; McConnell & Kaznowski, 1991). Early born

neural progenitor cells can sense and acquire the cell identities of the environment to which they are transplanted, while older progenitor cells maintain their normal cell identities. Thus, it is the combination of intrinsic (gene regulation) and environmental cues that contributes to the developmental competence of progenitor cells. One canonical example of how these internal and external cues are regulated is the Notch signaling pathway.

The evolutionarily conserved Notch signaling pathway regulates cell differentiation via intercell communication between adjacent cells (Artavanis-Tsakonas *et al.*, 1999; Kopan & Ilagan, 2009). Notch influences the decision between alternative cell fates during cell differentiation (Gaiano & Fishell, 2002) and is important for the maintenance of neural progenitors and the correct timing of differentiation (Imayoshi *et al.*, 2010). In the mammalian nervous system, proneural genes *Mash1* and *Ngn2* induce expression of Notch ligands such as Deltalike 1 (Dll1) (Castro *et al.*, 2006). Dll1 activates the transmembrane Notch receptor in neighboring cells causing the Notch intercellular domain (NICD) to be released. The NICD moves from the transmembrane region to the nucleus where it dimerizes with the DNA binding protein RBPj (Ohtsuka *et al.*, 1999). The NICD-RBPj complex induces expression of the basic helix loop helix factors Hes1 and Hes5, which repress expression of proneural genes (including *Mash1* and *Ngn2*), thereby creating a feedback loop between the two cells. Because high Hes1/Hes5 signaling in one cell reduces levels of Hes1/Hes5 expression in neighboring cells, Notch signaling creates a mutually repressive relationship referred to as lateral inhibition (Kageyama *et al.*, 2008).

Not surprisingly, CAM genes expression is also spatiotemporally regulated. N-CAM is a glycoprotien of the immunoglobulin super family that is expressed on the membrane surface of neurons, glia and muscle (Moore *et al.*, 1987; Weledji & Assob, 2014). N-CAM exhibits a spatiotemporal expression pattern during development (Crossin *et al.*, 1985; Moscoso & Sanes, 1995). At the early stage of neural tube formation, N-CAM is expressed in the neuroepithelium. At later stages, N-CAM is expressed by postmitotic neurons with a positional dependence on the anteriorposterior axis (Bally-Cuif *et al.*, 1993). Disruptions of either N-CAM binding or expression leads to altered morphogenesis (Cremer *et al.*, 1994;

Fraser *et al.*, 1988; Tomasiewicz *et al.*, 1993) and synapse formation (Stoenica *et al.*, 2006).

CAMs are regulated by internal gene products. The N-CAM gene promoter responds to cues from both homeobox (Hox) and paired box (Pax) gene products. Hox genes encode transcription factors that contain the conserved DNA binding domain called the homeodomain (Mallo & Alonso, 2013). Pax genes encode transcription factors that contain the conserved DNA binding domain called the paired domain (Blake & Ziman, 2014). Both Hox and Pax genes are required for proper tissue segmentation during development. In a cotransfection study of 3T3 cells, N-CAM promoter activity was controlled by *Hox-2.5* and *Hox-2.4* in a concentration-dependent manner (Jones *et al.*, 1992). Cotransfection of *Hox-2.5* with N-CAM increase N-CAM reporter expression while cotransfection with *Hox-2.4* mitigated reporter expression. Two homeodomain binding sites (HBS) on the N-CAM promoter were identified, which mediated responsiveness to homeobox genes. Later experiments showed that the N-CAM promoter was capable of responding to a variety of different homeodomain proteins (Jones *et al.*, 1993). Additional cotransfection studies using Cos cells showed that Pax-8 proteins increased N-CAM expression using binding sites on the N-CAM promoter using paired domain binding sites (PDS) distinct from the HBS above (Holst *et al.*, 1994), while Pax-6 proteins reduced N-CAM promoter activity (Holst *et al.*, 1997). Finally, the same group also showed that mutations of HBS and PDS sequences on the N-CAM promoter leads to defects in N-CAM patterning in the spinal cord (Holst *et al.*, 1997; Wang *et al.*, 1996).

Neurons likely use CAM proteins to mediate responses to external cues when selecting synaptic partners. A study from Lohmann and Bonhoeffer suggests that differences in intracellular signaling could contribute to synaptic specificity (Lohmann & Bonhoeffer, 2008). They used time-lapse microscopy to examine changes in intracellular calcium transients in dendritic filopodia after contact with axons in the CA3 region of the hippocampus. Filopodia frequently make contact with axons, but only some contacts are selectively stabilized. Filopodia never made stabilized contacts with inhibitory neurons. Local dendritic calcium transients were observed shortly after contact formation. Stabilized contacts exhibited a higher frequency of calcium transients compared to temporary contacts.

These results suggest that there is a relationship between local intracellular calcium signals and target recognition. Moreover, because calcium transients were observed after contact formation, but not before; CAMs rather than diffusible factors are the likely candidates to mediate contact-induced calcium transients.

CAM molecules have been implicated in regulating calcium signals during neuron growth. Studies have shown that CAMs directly promote axonal growth via homophilic binding mechanisms (Doherty *et al.*, 1990; Lemmon *et al.*, 1989). It is believed that CAMs activiate intracellular signaling in order to promote axonal growth because soluble versions of L1 (Doherty *et al.*, 1995) and N-CAM (Meiri *et al.*, 1998) can promote axonal growth as effectively as CAMs expressed at the surface of transfected cells. Calcium is a key second messenger that mediate the rate and direction of growth cone extension (Mattson & Kater, 1987; McCaig, 1989; Williams *et al.*, 1992). Studies have shown that growth cone elongation is due to spatial and temporal changes in calcium concentration (Gomez *et al.*, 1995; Kater *et al.*, 1994). Experiments using barium as the charge carrier have shown that CAM stimulated neurite outgrowth is modulated by localized submembrane increases in calcium changes rather than cell-wide changes (Archer *et al.*, 1999). Furthermore, the CAM stimulated localized calcium fluxes is dependent on the kinase activity of fibroblast growth factor receptor (FGFR). The proposed signaling pathway is that activation of the FGFR is followed by activation of phospholipase C (PLC), which generates diacylclycerol (DAG), which is converted to arachidonic acid, which then activates voltage dependent calcium channels (VDCC) giving rise to an induced calcium influx (Sheng *et al.*, 2013). This is supported by the observation that inhibitors of FGFR and PLC reduce the increase in intracellular calcium levels in response to N-CAM in cultured hippocampal neurons (Kiryushko *et al.*, 2006). Thus, CAM induced localized calcium signaling could provide a mechanism for directing neuronal growth.

Interestingly, some aspects of CAM signaling may not require cell adhesion. When FGFR binds to its ligand FGF, it undergoes dimerization and autophosphorylation, triggering signalling cascades (Knights & Cook, 2010). It has been proposed that N-cadherin positively regulates FGFR dimerization, initiating growth-factor independent signaling (Williams *et al.*, 2001). The FGFR dimerization is due to the *cis*-dimerization of N-cadherin, which is distinct from the adhesive

activities (Utton *et al.*, 2001). Moreover, it has been demonstrated that downstream inhibitors of FGF signaling also inhibit N-cadherin mediated cell motility (Nieman *et al.*, 1999), suggesting that N-cadherin is able to stimulate FGF-independent signaling. N-cadherin also affects ligand-dependent FGFR signalling by preventing receptor internalization, thereby increasing the number of receptors at the membrane surface which leads to sustained activation of FGFR by FGF (Suyama *et al.*, 2002). This results in the increased migration and motility of single cells, indicating that cell-cell adhesion is not required. In contrast to N-cadherins, N-CAM negatively regulates FGF activity, repressing FGF-signalling, cell proliferation and matrix adhesion (Cavallaro *et al.*, 2001; Francavilla *et al.*, 2007). There is evidence to suggest that N-CAM can act as a noncanonical ligand for FGFR and notably both the signaling cascade and the intracellular fate of N-CAM stimulated FGFR is distinct from those induced by FGF (Francavilla *et al.*, 2009). A number of studies suggest the interplay between N-CAM and FGFR may be independent of cell-adhesion. For example, N-CAM-FGFR complexes form on the surface of single cells resulting in FGFR activation (Cavallaro *et al.*, 2001). Additionally, the FGFR-binding motif and the modules involved in homophilic interactions are located at significantly distant regions on the N-CAM ectodomain (Kiselyov *et al.*, 2003; Soroka *et al.*, 2003). Taken together, these results suggest that cell-adhesion and FGFR-mediated signaling may be distinct and independent activities of N-CAM. For a review of other CAM transduction mechanisms which are independent of cell-adhesion, see (Cavallaro & Dejana, 2011).

Finally, we cannot conclude this sections without briefly addressing the role of microRNAs (miRNA) in the temporal development of the nervous system. MiRNAs are a class of non-coding RNAs that induce translational repression or degradation of a target mRNA by imperfect base pairing to its 3' untranslated region. The biogenesis and mechanisms of miRNA are discussed elsewhere (Bartel, 2004; Bushati & Cohen, 2007) and will not be discussed here. MiRNAs play a critical role in nervous system development (Cao *et al.*, 2016; Fiore *et al.*, 2008), having been implicated in neural stem cell proliferation/differentiation (Bian *et al.*, 2013), neuronal migration (Pedersen *et al.*, 2013), axon outgrowth/guidance (Vo

et al., 2005) and synaptogenesis and synaptic plasticity (Schratt, 2009). In addition, a number of miRNAs have been found capable of regulating molecules that mediate the cell-adhesion molecules (Valastyan & Weinberg, 2011). One such miRNA is miR-8, which regulates CAMs FasIII and Nrg during synapse formation in *Drosophilla*(Lu *et al.*, 2014). It is believed that FasIII and Nrg interact downstream of miR-8 in order to promote accurate target recognition. Moreover, miR-8 appears to control key efforts on both sides of synapse formation during embryogenesis. While research on the interaction between miRNAs and CAMs is still an emerging field, current evidence seems to suggest that miRNA regulated CAM signaling pathways will be crucial for our understanding wiring specificity in the nervous system.

In summary, while I do not consider the temporal dynamics of CAM expression in the present analysis, it is nonetheless a crucial factor in determining the wiring specificity of the nervous system. Given what is known, it seems plausible that different CAM genes could be expressed at different developmental time points and operate on different time scales. Initial expression of some subset of CAM genes may be used to modulate tissue segmentation and therefore control initial ganglia placement. Expression of these CAM genes may be internally regulated by Hox or Pax proteins. Another set of CAM molecules (not necessarily distinct from the early CAM set) may be expressed later and mediate process guidance and target recognition in the nerve ring. The later set of CAM molecules may respond to proper target identification and thereby modulate FGF and other intracellular signaling in order to guide processes into their proper neighborhoods. This guidance process may not require actual adhesion between cells thereby allowing cells to more quickly find their proper neighborhoods. Eventual adhesions and synapse formation with proper synaptic targets may then be regulated by miRNA or other regulatory mechanisms.

2.5 *C. elegans* as a model of wiring specificity

The nematode *C. elegans* offers unique advantages for understanding wiring specificity. The worm has a small nervous system – 302 and 383 neurons in the hermaphrodite and male, respectively. The synaptic connectivity for both sexes

has been well defined using serial sectioned electron micrographs (Cook *et al.*, 2017; Jarrell *et al.*, 2012; Varshney *et al.*, 2011; White *et al.*, 1986). *C. elegans* is very amenable to genomic studies, with the neurobiology-related gene families being particularly well documented (Bargmann, 1998; Hobert, 2005). Moreover, *C. elegans* has a number of representatives from each of the CAM proteins classes that have been extensively documented in the development of wiring specificity of various species (IgSf, LRR, cadherins and neurexins). A number of examples of wiring specificity in the *C. elegans* nervous system have already been observed.

Glia-like sheath cells coordinate the synapse formation between interneurons in the *C. elegans* thermotaxis circuit (Colón-Ramos *et al.*, 2007). The sheath cell processes express UNC-6/netrin, a class of proteins involved in axon guidance, and converge at the point where interneuron AIY synapses onto interneuron RIA. Both AIY and RIA express the netrin receptor UNC-40/DCC. UNC-40 elicits normal axon guidance behavior in RIA while in AIY it cell-autonomously promotes assembly of presynaptic terminals at the point of the sheath cell convergence. This suggests that the glia-like sheath cells function as guidepost cells during neural circuit development.

Guidepost epithelial cells act as a placeholder for presynaptic specializations in the HSN egg-laying motor neurons (Shen & Bargmann, 2003; Shen *et al.*, 2004). The subcellular synapse localization is mediated by the heterologous interaction between the two IgSf proteins, SYG-1 and SYG-2. During early HSN synaptogenesis, the SYG-2 ligand is expressed transiently in the epithelial cells while the SYG-1 receptor is expressed in HSN and localizes to points of synapse formation. Loss-of-function mutants *syg-1* and *syg-2* form ectopic synapses onto inappropriate targets. This suggests that guidepost cells can be used to determine the subcellular specificity of synapses.

Subcellular specificity is also achieved through inhibitory Wnt signaling in the neuromuscular connectivity of motor neuron DA9 (Klassen & Shen, 2007). Neuromuscular junctions are restricted to a specific part of the DA9 axon. DA9 synapses do not form at the most posterior end of its axon, which has the highest concentration of the Wnts LIN-44 and EGL-20 expressed in both DA9 and surrounding tail cells. Loss-of-function mutants lin-44 and egl-20 form ectopic

synapses in this DA9 subdomain. Conversely, ectopic expression of LIN-44 inhibits synapses formation in adjacent axon segments. This shows that local Wnt signaling can be used to shape the synaptic domains of neurons.

More recently, it was shown that multiple CAM proteins interact to mediate wiring specificity in the *C. elegans* male mating circuit (Kim & Emmons, 2017). The male-specific sensory neuron HOA synaptically targets the sex-shared interneuron AVG and the sex-shared sensory neurons PHC. The cadherin CASY-1 and IgSf RIG-6 proteins are both expressed in AVG. The neurexin BAM-2 in HOA binds to CASY-1 and the IgSf SAX-7 in PHC binds to RIG-6. Loss-of-function mutants casy-1, rig-6 and bam-2 exhibit an altered pattern of presynaptic punta in the male-specific sensory neuron HOA. The altered synaptic pattern is due to local fasciculation defects along the HOA-AVG-PHC bundle. Interestingly, the synaptic defect in HOA is also observed both when PHC is ablated and in *sax-7* mutants. This illustrates how multiplexed CAM expression from multiple cells affects synapse formation in an individual cell.

2.6 Contribution of this work

In 1986, White *et al.* published their seminal work on the reconstructed wiring diagram of synaptic connectivity in *C. elegans*, where they meticulously scored the synapses between all 302 hermaphrodite neurons from serial sectioned electron micrographs (EMs). While small additions have been made to the wiring diagram in the intervening years [Hall & Russell (1991); Varshney *et al.* (2011)], this valuable data set has stood the test of time. To date, this remains the only 'complete' whole-animal wiring diagram and has been an incredibly useful tool for experimentally probing the development and function of the *C. elegans* nervous system. However, the data set in some sense still remains incomplete because it does not provide a comprehensive view of the spatial proximity of neurons, referred to as neuron adjacency. White and colleagues did provide an initial analysis of neuron adjacency in the nerve ring (White, 1985; White *et al.*, 1983). They measured the adjacency for roughly half of the NR neurons by counting the number of EM sections in which neurons made physical contact with neighboring neurons. While impressive, the data had several limitations. The adjacency data

was collected from the larval stage 4 data set (JSH), which cannot be directly compared to the canonical connectivity data collected from a young adult data set (N2U). Adjacency was only scored for one neuron from each neuron class, which makes it impossible to asses adjacency variability of neurons within a class. Finally, only adjacency data for a few selected neurons (AIAR, AIBR, AQR and PQR) was published (White *et al.*, 1983).

This work presents the first volumetric reconstruction of an organism's main neuropile at two different developmental stages. The datasets provide adjacency data for all neurons with processes in the NR for both the larval stage 4 (L4) and young adult. Adjacency is quantified by measuring the surface area of membrane contact between neurons and not simply by counting EM sections. Upon publication, the data will be made publicly available on wormwiring.org and can be easily correlated with connectivity and gene expression data.

Knowing the spatial proximity of neurons provides a useful framework for exploring nervous system development. During nervous system development, cell surface molecules modulate attraction and repulsion between neurons in order to ensure that neurons make physical contact with the appropriate partners (Yogev & Shen, 2014). Once appropriate neurons make contact, they must form a synapse which could require subcellular specificity (Yogev & Shen, 2014). Thus, it's not only important to know which neurons make contact and form synapses but also where along the neurite synapses are formed; all of these processes are seemingly regulated at the molecular level. Knowing the spatial organization of neurons has been useful for experimentally probing these mechanisms in other model organisms. For example, knowing the lamina organization of the vertebrate retina allowed researchers to identify immunoglobulin superfamily (IgSF) adhesion molecules which leads to a mismatches in sublamina specificity when inappropriately expressed (Yamagata & Sanes, 2008). Researchers effectively exploited the spatial organization of neurons in order to experimentally uncover the molecular mechanisms that give rise to the lamina structure.

Comparing the spatial data with the synaptic wiring diagram shows that the C. elegans nerve ring exhibits both lamina-like specificity and synaptic specificity. A statistical model is developed that captures the variability in synaptic

connectivity among neurons. The results of the model suggest that synaptic connectivity does not strongly depend on the amount of physical contact between neurons, but does strongly depend on the cell autonomous characteristics of the pre- and postsynaptic neuron. With this in mind, wormbase data was used to analyze the combinatorial expression patterns of CAM genes in nerve ring neurons. If isoforms of alternatively spliced genes are included in the expression, then the number of known combinatorial CAM expression patterns is sufficient to uniquely label neurons and can account for more than 90% of the observed synaptic specificity. Finally, comparing CAM expression patterns of postsynaptic neurons suggests that multiple gene combinations are used to elicit synapses from the presynaptic neuron.

2.7 Acknowledgments

This work was a collaboration between Steven Cook (SC) and myself (CB). CB conceived the project. SC reconstructed the synaptic wiring for the JSH and N2U data series. SC volumetrically reconstructed the JSH data series. CB volumetrically reconstructed the N2U data series. CB conducted all of the data analysis.

Chapter 3

Volumetric reconstruction of the *C. elegans* nerve ring

The nematode *C. elegans* has a compact well-described nervous system[1]. The adult hermaphrodite nervous system consists of 302 neurons interconnected by more than 6400 chemical synapses and 900 gap junctions (White *et al.*, 1986). The nervous system is comprised of two mostly independent nervous systems: a large somatic (282 neurons) and a small pharyngeal (20 neurons) nervous system[2]. The somatic nervous system consists of a head ganglia, a complex head sensory system, a ventral cord, dorsal cord and a tail ganglia (Thomas & Lockery, 1999). Among wild type worms, neuron structure and cell body positions are nearly invariant from animal to animal (Durbin, 1987), and it is possible to identify each neuron *in vivo* (Sulston *et al.*, 1983), though some variation has been characterized (Bargmann & Avery, 1995). The 302 neurons are divided into 118 classes based on morphology and synaptic connectivity. Neuron names follow a standardized nomenclature. In sensory neurons, interneurons and head motor neurons, the first three letters describe the class of the neuron. The one or two letter suffix indicates the member of the neuron class and indicates anatomical placement of the cell body. Suffix letters use L, R, D, and V to distinguish left,

[1] Unless otherwise stated, all discussion refers to the adult hermaphrodite somatic nervous system (White *et al.*, 1986), which is considerably less complex than the adult male nervous system (Jarrell *et al.*, 2012).

[2] Details of the pharyngeal nervous system are discussed in (Albertson & Thomson, 1976).

right, dorsal and ventral, respectively. For example, RMDVL is the neuron in the RMD class whose cell body is located dorsally and to the left. For ventral cord and dorsal cord motor neurons, the first two letters describe the neuron class and the suffix is a number which describes the neuron's sequential placement along the cord proceeding posteriorly from the nerve ring. For example, VA1 is the first neuron in the VA class. No suffix is used when a neuron class has only one member, e.g. DVA. When referring to a neuron class rather than to a specific neuron, the two or three letter class name will be used.

The largest collection of neurons is around the nerve ring (NR), so called because of the shape it takes during development as the pharynx grows through the surrounding neurons (Altun & Hall, 2011). The NR is essentially an enlarged commissure[1] encircling the pharyngeal isthmus. The NR sits between five ganglia (anterior, dorsal, lateral, ventral and retrovesicular ganglion; Figure 3.1) and is often referred to as the 'brain' of the worm because it is the most densely innervated part of the nervous system and contains the majority of sensory and interneurons. There are 180 head and tail neurons that project axons/processes into the NR and approximately 80% and 60% of all known chemical synapses and gap junctions, respectively, are contained in the NR. During embryonic development, anterior, dorsal and some lateral ganglion neurons (e.g. AVA and RIV) innervate the NR directly. The remaining neurons first innervate the ventral nerve cord (VC) via the commissural and longitudinal nerve bundles that join the VC and the VC in turn innervates the NR. During NR innervation, neurons must be able to recognize the NR region, make the left/right side choices to enter the NR and then make appropriate synaptic contacts to physically adjacent neurons. Also at this time, development of head/neck muscles and other support cells (e.g. glia-like sheath cells) must be coordinated with the longitudinal and commissure tract development.

Surprisingly, given the NR's prominence in the *C. elegans* nervous system, little is known about the process and control of NR development and in particular

[1] A commissure is a circumferential tract created by neurons growing through a dorsoventral route. The NR commisure serves as a junction where left/right bilaterally segregated neurons can meet and synaptically couple – many left/right neurons from the same class are coupled by gap junctions in the commisure.

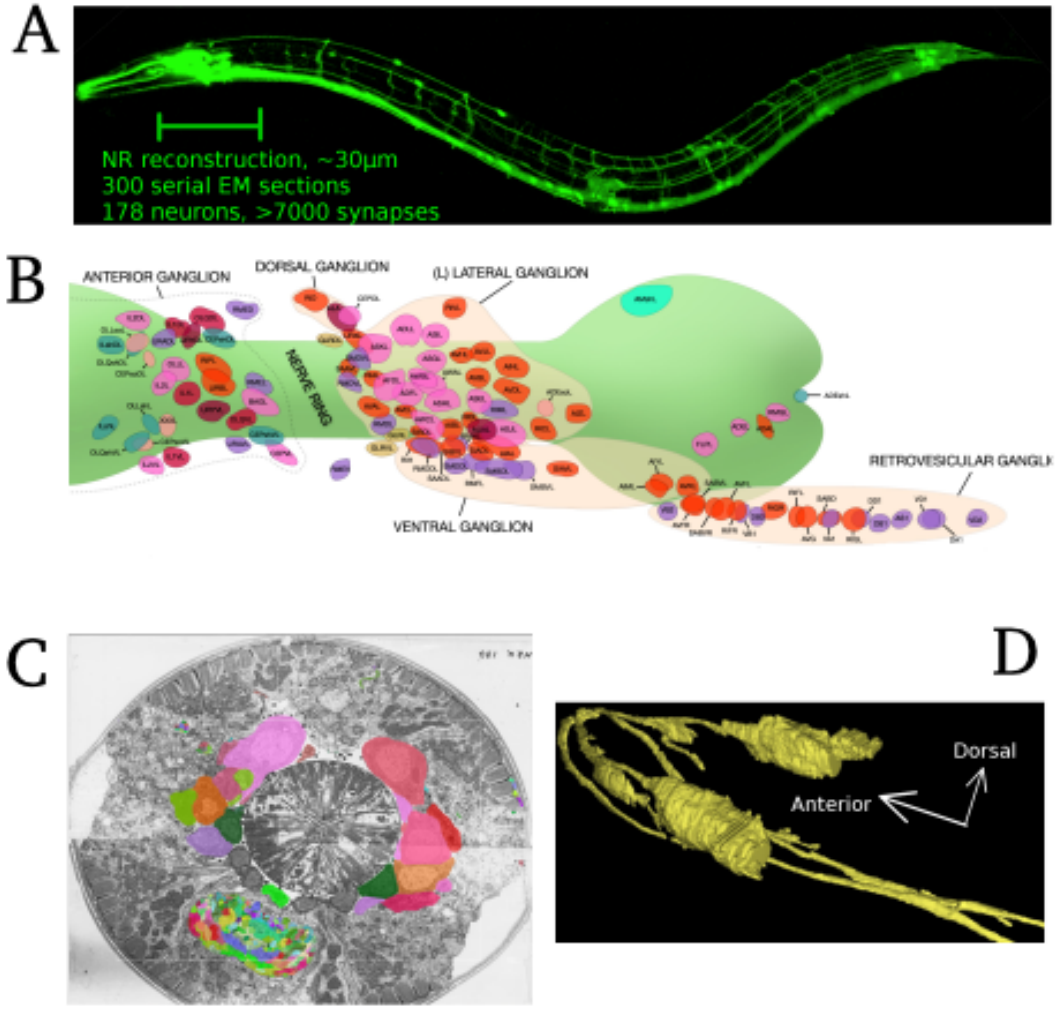


Figure 3.1: **Overview of anatomy and volumetric reconstruction.** (A) A worm expressing fluorescent GFP in its entire nervous system. The bar outlines the segment of the worm (~30μm) used for the NR volumetric reconstruction from legacy serial EMs (Modified image from Hang Ung, Jean-Louis Bessereau laboratory, France.) (B) The five head ganglia (anterior, dorsal, lateral, ventral and retrovesicular) with axons/processes that project into the NR. The NR encircles the pharyngeal isthmus between the anterior and posterior bulbs.(Modified image from wormatlas.org.) (C) A segmented EM taken from where the VC enters the NR. Neurons are manually segmented using TrakEM2 where neurons are assigned different colors. The segmentation was performed for each EM. (D) A 3D reconstruction of neurons AVAL and AVAR generated from the segmentation data.

the wiring specificity of the NR. This is most certainly due to the high neural and synaptic density of the *C. elegans* NR which has 180 processes packed into 300 μm^3 and makes studies using fluorescent markers challenging to interpret. A key piece of missing information is a complete and comprehensive volumetric map of the spatial organization of axons/processes within the NR commissure. Recall how characterization of the sublamina of the IPL in the vertebrate retina aided many developmental insights concerning wiring specificity (Baier, 2013). While the synaptic connections in the NR have been extensively characterized the spatial relations between neurons has not. Just as the synaptic wiring diagram has been a valuable tool for experimentally probing the function of the *C. elegans* neural circuits (Emmons, 2015), a volumetric map combined with a map of gene expression could provide insights into the wiring specificity of neural circuits.

3.1 Reconstructing from legacy serial EMs

The NR was volumetrically reconstructed from legacy serial section electron micrographs (EM) originally published by White *et al.* (1983). Two NR reconstructions were performed, one from an adult and the second from a larval stage 4 (L4) worm (approximately 10 hours younter than the adult). Both datasets start in the anterior ganglia and finish in the ventral ganglia, covering approximately 36 μm (Figure 3.1a). The adult and L4 data series consist of 300 and 400 EM sections, respectively, each section is ~90 nm thick. The adult has fewer EM sections because every other ventral ganglia section was skipped during imaging (the adult sections may also be slightly thicker). While all 178 axon/processes in the nerve ring were reconstructed, only somas in the anterior and ventral ganglia were reconstructed. Also, the dendritic processes of the amphid and labial sensory neurons which extend towards the nose from the lateral and anterior ganglia, respectively, were not reconstructed. Because these regions of the neuron processes do not have many synapses, they were not of immediate interest.

The EMs from White *et al.* (1986) were previously donated from the MRC/LMB archives to the Hall laboratory. The EMs have since been digitized and are available at wormimage.org. This study uses the 'N2U' and 'JSH' data sets which

were taken from an adult and L4 hermaphrodite, respectively. The synaptic connectivity of these data sets was previously reconstructed by White *et al.* (1986) and Varshney *et al.* (2011), but we used the most recent reconstruction reported by Cook *et al.* (2017) and available at www.wormwiring.org. Unlike the data of Varshney *et al.* (2011); White *et al.* (1986), the Cook *et al.* (2017) data also contains the spatial locations of synapses and is conveniently stored in a MySQL database (Xu *et al.*, 2013). The volumetric reconstruction was manually done using TrakEM2 software (Cardona *et al.*, 2012). The software provides tools to allow the user to segment neurons, track the segments and stores the data in XML format. Volumetric reconstructions are generated by combining the tracked segments. Measurements of the physical contact between neurons was taken directly from the segmented XML data.

The volumetric data can be visualized using a custom built web app. At the time of this writing, the web app, MeshApp is available at wormwiring.org and the source code is available at GitHub. MeshApp is written in javascript using the three.js library. Wavefront files (.obj) are saved for each neuron using the TrakEM software, which are then used as the data source for MeshApp. Both the wavefront files and app are hosted on wormwiring.org, so the user does not need to download or install anything. To use MeshApp, the user simply goes wormwiring.org/beta/meshapp, clicks the Select neurons button and inputs the neuron names into the dialog. The MeshApp can be used to view any number of nerve ring neurons. Visualization of each neuron can be turned on/off using the side menu. The color of each neuron can also be adjusted using the side menu.

I constructed an algorithm to measure the physical connectivity between neurons. Two neurons that touch are said to be physically connected. The algorithm measures physical connectivity directly from the segmented XML data, which consists of the (x, y)-coordinates of the segment boundaries. For every segmented EM, the algorithm classifies two neuron segments as physically connected if the distance between their respective boundaries is less than 10 pixels ($\sim$50 nm). To measure the amount of physical contact between neuron segments in a given image, the algorithm counts the number of boundary points that are physically connected. Each boundary point is estimated to be 5 nm long, which

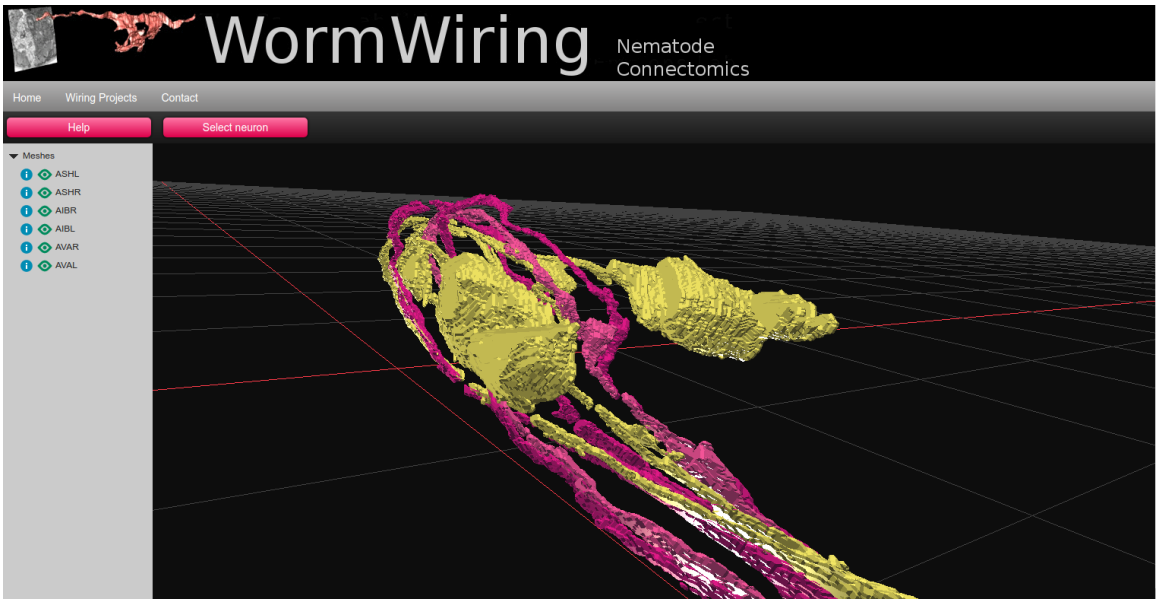


Figure 3.2: **App for viewing volumetric data.** Screenshot of the web app used to view the volumetric data at wormwiring.org/beta/meshapp.

is used to convert the number of boundary points to a physical length. The total physical contact between two neurons is the sum of physical contact over all images.

To check the accuracy of the algorithm, two EM sections were manually scored for physically connected partners and compared to the algorithm results (data not shown). In both cases the algorithm outperformed the manual scorer, recognizing connected partners not identified manually. Any failure of the algorithm to identify adjacent partners (false negatives) was mostly due to poor manual segmentation. For example, the person doing the segmentation may not have properly traced around the cell boundary. There were a small number of cases where the algorithm incorrectly labeled two neurons as connected (false positives). In these cases, the length of contact was very small and could be screened out in later analysis by requiring connected partners have a minimum contact length. Finally, connected partners were previously reported for a small subset of neurons based on a sparse analysis of physical connectivity in the L4 (White *et al.*, 1983). Our physically connected partners match those previously reported. Thus, we concluded that, for appropriately segmented neurons, the algorithm adequately

identifies all physically connected partners.

3.2 Metrics for quantifying physical and synaptic connectivity

Network analysis provides a convenient framework for describing how neurons are both physically and synaptically connected (Newman *et al.*, 2006). A network is a set of points called vertices which are connected by lines called edges. In our representation, vertices are neurons and edges are either the physical or synaptic connections between neurons. A network can be represented in matrix form. I define two matrices: the adjacency matrix and the synaptic matrix.

The adjacency matrix A describes the physical connectivity of the NR[1]. Two neurons that make physical contact are said to be adjacent. The adjacency matrix A contains elements a_{ij} where $a_{ij} = 1$ if neuron i and j are adjacent and $a_{ij} = 0$ if i and j do not make physical contact. The adjacency network is undirected, so A is a symmetric matrix, i.e. $a_{ij} = a_{ji}$. The synaptic connectivity is represented by the synaptic matrix S where the element $s_{ij} = 1$ if there is a chemical synapse from the presynaptic neuron i to the postsynaptic neuron j and $s_{ij} = 0$ otherwise. Unlike physical contacts, synaptic contacts are directed because the presynaptic neuron signals the postsynaptic neurons. Therefore, in general, $s_{ij} \neq s_{ji}$ and S is asymmetric.

The *degree* of a neuron is the number of connections made by a neuron. Neuron j is a neighbor of neuron i if there is an edge between i and j. The neighborhood N_i is the set of neighbors of neuron i. The degree d_i is the size of N_i. The degree can be computed from the adjacency matrix A as

$$d_i = \sum_{j=1}^{n} a_{ij}, \tag{3.1}$$

[1]The nomenclature here is slightly unfortunate. In standard network and graph theory, the adjacency matrix is any matrix used to represent a finite graph. However, White *et al.* (1983) and Durbin (1987) use adjacency to refer to neurons that make physical contact. For the sake of biological continuity, I use the term adjacency matrix to specifically refer to the graph of physical connectivity.

where n is the number of neurons. Because S is directed, we distinguish between presynaptic and postsynaptic edges. A chemical synapse from i to j is a presynaptic edge for i and a postsynaptic edge for j because i is the presynaptic neuron and j is the postsynaptic neuron. The presynaptic degree d_i^{pre} counts the number of presynaptic edges extending from neuron i to its neighbors,

$$d_i^{pre} = \sum_{j=1}^{n} s_{ij}. \tag{3.2}$$

The postsynaptic degree d_i^{post} counts the number of postsynaptic edges received by i from its neighbors,

$$d_i^{post} = \sum_{j=1}^{n} s_{ji}. \tag{3.3}$$

Degree measures are typically reported as the integer values above, but we normalize the degree by the number of neurons n. The normalized degree is defined as

$$\widetilde{d_i} = \frac{1}{n} \sum_{j=1}^{n} a_{ij}. \tag{3.4}$$

The $\sim$ signifies that the degree is normalized. Both $\widetilde{d_i^{pre}}$ and $\widetilde{d_i^{post}}$ are defined similarly.

The above treats edges as binary, either the edge exists or it does not exist, but in many instances we want to associate values to each edge. For the adjacency network, we want to quantify the amount of contact between neurons. For each adjacency edge, we associate two attributes: section contact and touch density. The section contact w_{ij} for edge (i, j) is the number of EM sections where the physical contact was scored. The touch density t_{ij} for edge (i, j) is the fraction of neuron i's membrane that makes contact with neuron j. Formally, this is computed by dividing the number of pixels where i and j touch by the total surface area of neuron i within our dataset. In general, $t_{ij} \neq t_{ji}$ because neuron i and j make the same amount of physical contact with each other but the neurons will have different surface membrane areas. Section contact is useful when comparing the amount of physical contact with the amount of synaptic contact because they are in the same units (number of sections). Touch density is useful for comparing the amount of adjacency between neuron pairs, both within and

between datasets, because it is normalized by the cell surface area, thus controlling for differences in neuron sizes. For the synaptic network, we also quantify the amount of synaptic contact between neurons. For each synaptic edge, we associate the section contact. As with the adjacency section contact, the synaptic section contact for edge (i, j) is the number of EM sections where a synapse from i to j was scored.

Finally, we define the connectivity fraction which measures the fraction of physical contacts that are also synaptic contacts. The connectivity fraction is defined as the ratio of a neuron's synaptic degree to its adjacency degree. Formally, the pre and post connectivity fractions are respectively defined as:

$$C_i^{pre} = \frac{d_i^{pre}}{d_i} \text{ and } C_i^{post} = \frac{d_i^{post}}{d_i}. \tag{3.5}$$

3.3 NR exhibits conserved spatial structure

The above metrics already reveal some basic spatial structure in the NR. The adjacency degree distribution is approximately normally distributed (Figure 3.3a). Thus, the adjacency degree distribution does not exhibit properties found in other commonly studied networks. For example, the single independent edge probability found in Poisson distributed networks (Newman *et al.*, 2001) or the over representation of hubs (nodes with high degree) found in scale-free networks (Barabasi & Albert, 1999). Instead, it appears that neurons make physical contacts with varying probability with no probability particularly overrepresented. Neurons typically make physical contact with 40-80 other NR neurons. This only represents a small fraction ($\sim 1/3$) of the entire NR population indicating that NR neurons are spatially segregated.

Touch densities are exponentially distributed (Figure 3.3c). That is, most touch densities are small – 80% of touch densities take up less than 5% of the cell membrane – and a few touch densities are relatively large – 10% of touch densities take up more that 30% of the cell membrane. This suggests that neurons make a few "preferential" longer contacts with neighboring neurons in addition to making many more smaller sporadic physical contacts. This is consistent with previous observations that neuron axons/processes fasciculate the NR in process bundles

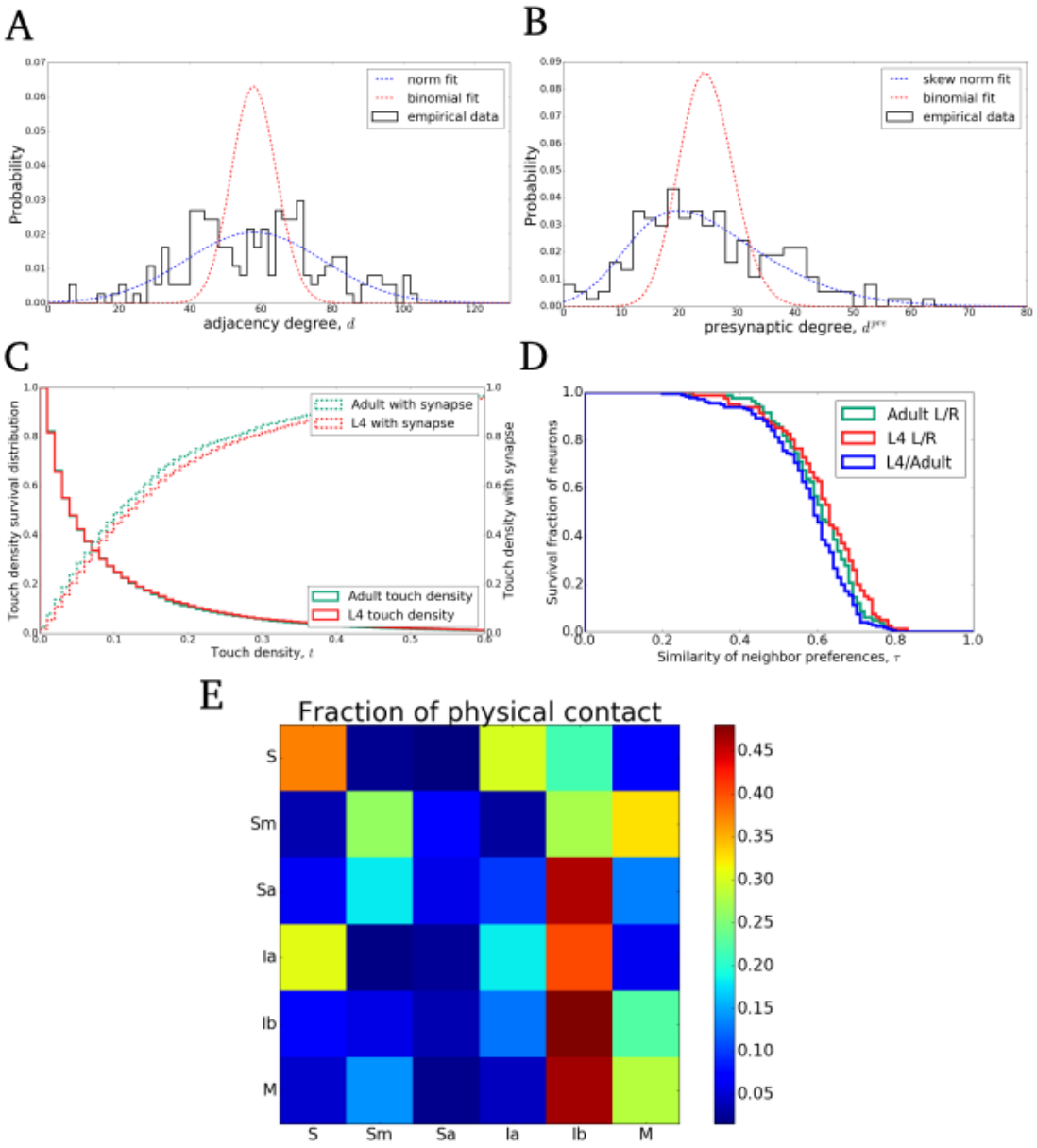


Figure 3.3: **Distribution of physical contacts.** (A) Degree distribution of the adult adjacency network (black). Fit with a binomial distribution (red dashed) and a normal distribution (blue dash). (B) Presynaptic degree distribution of the adult synaptic network (black). Fit with a binomial distribution (red dash) and a skew normal distribution (blue dash, $skew = 0.53$). The L4 networks (not shown) are qualitatively similar to the adult. (C) Survival distribution of touch densities. Touch density is given as the fraction of surface membrane that makes physical contact with the adjacent neuron. Also shown is the conditional distribution function of touch densities for physical contacts that result in at least one synapse. (D) Comparison of inter- and intraworm variability of neighbor preferences. Survival distribution of similarity of neighbor preferences (τ) between adult (green) and L4 (red) bilaterally symmetric left/right neurons and between homologous adult and L4 neurons (blue). Most neurons ($> 90\%$), neighbor preferences are more similar ($\tau > 0.5$) than dissimilar. (E) Distribution matrix of surface area contact (pixels) across neurons. Neurons are placed in one of six categories based on their anatomical location and function: sensory (S), mechanosensory (Sm), anterior sensory (Sa), first-layer interneuron (Ia), second layer interneuron (Ib) and motorneurons. Elements of the matrix are normalized by row sums such that the sum of each row is 1. Red and blue indicate large and small fractions of physical contact, respectively.

White *et al.* (1983, 1986), referred to as "neighborhoods". White *et al.* (1986) pointed out that certain neighbors make more persistent contact than others, while other neighbors move in and out of direct physical contact along the length of the process. My measurements show that roughly 60% of neighbors will change over 1-2μm of process length and 10% of neighbors will persist for longer than 10μm of process length (data not shown). So, given that neighbor preferences do exist, how consistent are these preferences across worms?

The neighbor preferences are both developmentally and bilaterally conserved between the L4 and the adult. For each neuron, density scores were used to rank neighbor preferences. Neuron i prefers neuron j over neuron k, if $t_{ij} > t_{ik}$. For neuron i, let O_i be the set of neighbors ordered by increasing preference. The set O_i can be decomposed into $\frac{1}{2}M(1-M)$ ordered pairs. Let P_i be the set of ordered pairs from O_i. Now consider neuron j with ordered set O_j which can be decomposed into the set of ordered pairs P_j. For convenience, we require that set O_i and O_j have the same set of M neighbors. If this is not the case, we simply consider the intersection of O_i and O_j. The Kendall rank coefficient is used to compare the ordered preferences O_i and O_j, given by

$$\tau_{ij} = 1 - \frac{2\Delta(P_i, P_j)}{M(M-1)}, \tag{3.6}$$

where $\Delta(P_i, P_j)$ is the symmetric difference operation which counts the number of ordered pairs that belong to either P_i or P_j, but not both. Because τ_{ij} is determined from counting the number of different pairs between sets, it can be interpreted in a probabilistic context (Abdi, 2007). Specifically, for two ordered sets O_i and O_j, τ_{ij} can be interpreted as the probability that the sets are in the same order minus the probability that the sets are in different orders. Thus, τ_{ij} provides a measure of conserved neighbor preferences between neurons. The significance of τ_{ij} can theoretically always be computed because the ranked sets O_i and O_j are finite. For details on how to compute the significance, the reader is referred to Abdi (2007). Here, we will simply note that for sufficiently large M the sampling distribution of τ converges towards a normal distribution with mean 0 and variance

$$\sigma_\tau^2 = \frac{2(2M+5)}{9M(M-1)}. \tag{3.7}$$

The associated p-values can then be computed using standard procedures.

For each neuron, we only ranked neighbors that are shared by both the L4 and adult. Correlation coefficients were positive for all neurons, with 80% of neurons having a $\tau \geq 0.5$ with $p < 0.01$ for each neuron (Fig. 3.3d). Therefore, we can reject the null hypothesis and conclude there is significant agreement of touch density rankings between the L4 and adult neurons. For comparison, we computed ranking coefficients for bilaterally symmetric left/right homologous neurons in both the L4 and adult and arrived at similar scores. This indicates that the differences in t rankings are not larger than what would be expected from inter-worm variability.[1]

We have thus far only focused on the adjacency of individual neurons, but the NR does exhibit higher level spatial organization. Neurons were classified into one of six categories based on anatomy and function: sensory (S), amphid sensory (Sa), mechanosensory (Sm), first layer interneurons (Ia), second layer interneurons (Ib) and motor neurons (M). These categories are discussed in more detail in Section 4.1. Here, we merely consider the breakdown of synaptic connectivity between these classes. Approximately 90% of amphid sensory surface area is dedicated to other amphid sensory neurons, first- and second-layer interneurons. Roughly 90% of mechanosensory surface area is dedicated to second-layer interneurons, motor neurons and other mechanosensory neurons. The anterior sensory neurons mostly mix with second-layer interneurons. First-layer interneurons mostly mix with amphid sensory and second-layer interneurons. Second-layer interneurons and motor neurons mostly mix with motor neurons and other second-layer interneurons. This suggests that the NR exhibits a lamina-like structure.

[1]Some outliers are neuron classes ADE/AVD/BAG/FLP/RID/RME/SDQ which consistently had a $\tau < 0.5$; indicating that neighborhood preferences for these neurons are less likely to be conserved.

3.4 Synaptic contacts are not correlated with adjacency

Previous studies have suggested that synapse probability is correlated to the spatial proximity between neurons This has previously been formalized as Peters' rule, with one simple interpretation being that axons make synapses in direct proportion to the number of proximal synaptic targets (Binzegger *et al.*, 2004; Braitenberg & Schüz, 1998). However, there has never been a clear consensus on how this rule should be applied and evaluated (Rees *et al.*, 2017). Recent studies have shown that Peters' rule is not a good predictor of synaptic connectivity (Kasthuri *et al.*, 2015; Mishchenko *et al.*, 2010; Shepherd *et al.*, 2005), on the other hand algorithms using variations of Peters' rule have been able to simulate and reconstruct synaptic connectivity (Markram *et al.*, 2015; Reimann *et al.*, 2015). Consequently, there is no consensus on if Peters' rule can predict connectivity.

Comparison of the adjacency and synaptic networks shows that there is no direct correlation between physical and synaptic contacts in the *C. elegans* NR. The adjacency degree distribution is normally distributed (Figure 3.3a) while the presynaptic degree distribution has a skewed normal distribution (Figure 3.3b). A linear map from physical to synaptic contacts would shift or scale but not skew the degree distribution, suggesting any possible relation is likely nonlinear. For each neuron, there is minimal correlation between $\widetilde{d}$ and $\widetilde{d}^{pre}$ ($r = 0.25$, Figure 3.4a) while there is some correlation between $\widetilde{d}$ and $\widetilde{d}^{post}$ ($r = 0.68$, Figure 3.4b), possibly suggesting that there are pre- and postsynaptic mechanisms that operate independently. Ultrastuctural analysis of the rat hippocampal neuropile suggested that axo-dendritic touch density is a good predictor of synapse density (Mishchenko *et al.*, 2010). However, in the NR dataset there is no significant correlation between the amount of physical contact and synapse size ($r = 0.25$, Figure 3.4c) nor between the amount of physical contact and the number of synapses ($r = 0.42$, Figure 3.4d). Finally, both C^{pre} and C^{post} exhibit large variation across neurons (Figure 3.4e,f) and there is essentially no correlation between C^{pre} and C^{post} (r=0.02, data not shown). This suggests that synaptic properties are likely different for each neuron. Collectively, these results suggest that there is no simple linear relationship between physical and synaptic contacts.

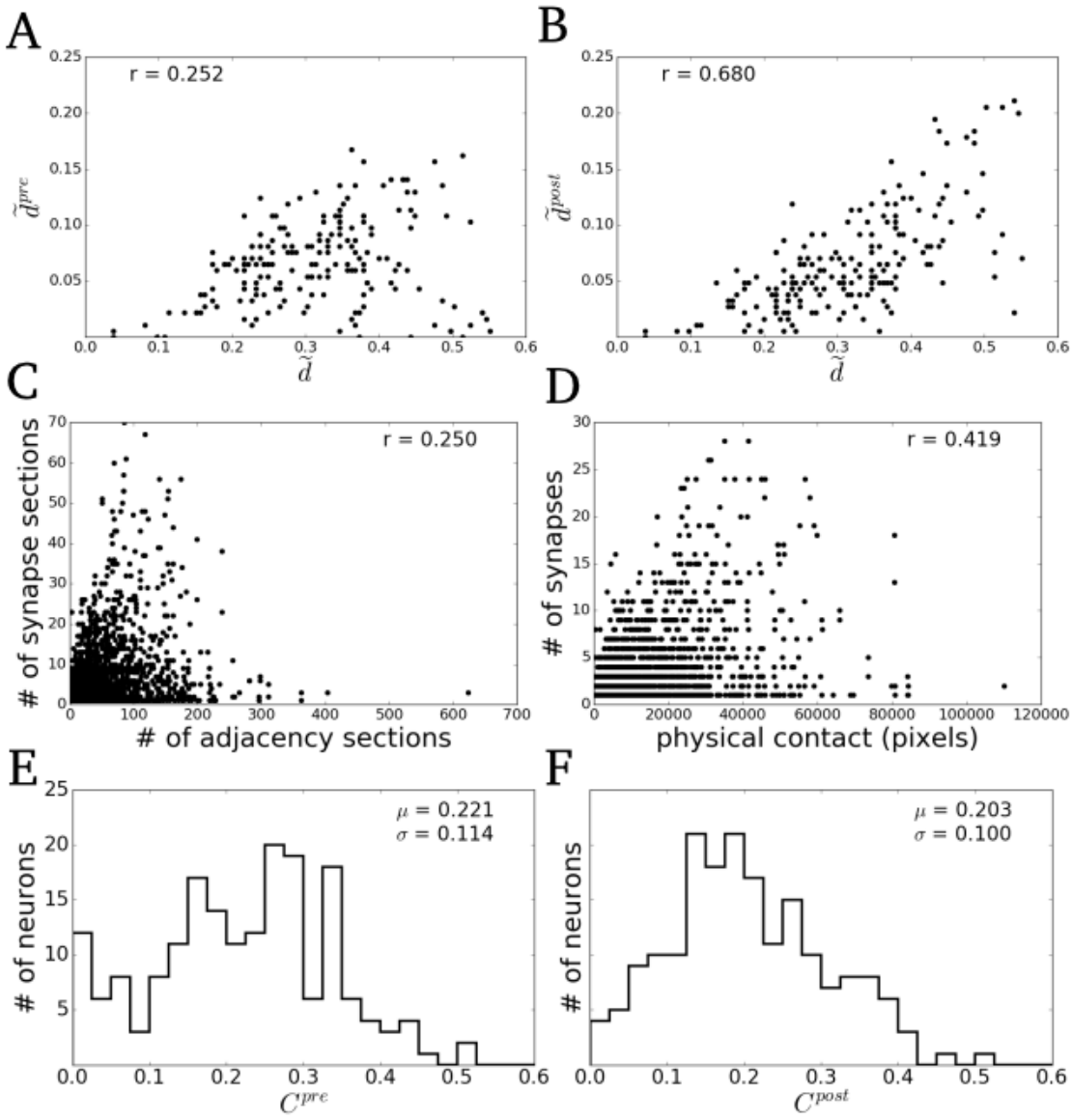


Figure 3.4: **Correlations between adjacency and synaptic contacts.** (A-B) The normalized presynaptic degree ($\widetilde{d}^{pre}$) versus the normalized adjacency degree ($\widetilde{d}$). (B) The normalized postsynaptic degree ($\widetilde{d}^{post}$) versus $\widetilde{d}$. (C) For every pair of synaptic partners, the number of synaptic sections versus the number of adjacency sections. (D) For every pair of synaptic partners, the number of synapses versus the amount of physical contact (pixels). (D-E) The distribution of pre (C^{pre}) and post (C^{post}) connectivity fractions. The correlation coefficients (r), mean (μ) and standard deviation (σ) are given.

Chapter 4

Measuring wiring specificity

4.1 NR exhibits lamina-like structure

In order to assess how the nerve ring is physically organized, the spatial positions of neural processes and synapses within the NR was mapped. Because of its torus shape, spatial positions in the NR are conveniently characterized in terms of cylindrical coordinates (r,ϕ,z) (Fig. 4.1a). The radius (r) is measured as the distance from the outer edges of the pharynx to the neuron or synapse. The azimuth angle (ϕ) is measured with respect to the ventral axis, with positive ϕ moving in the clockwise direction. The z coordinate gives the position along the anterior-posterior axis of the worm. For each EM section, both r and ϕ are plotted as a function of z for both right and left side neurons (Fig. 4.1c-d). Neurons within the same category typically exhibit similar r distances along z. Most neurons complete a half cycle around the nerve ring, starting at $\phi = 0$ or $\phi = \pm\pi$ and stopping at $\phi = \pm\pi$ or $\phi = 0$, respectively. A few neurons will cycle around more than half of the NR (e.g. AIB) and still some wrap completely around the NR (e.g. AQR). Organization patterns in the NR are more easily observed by grouping neurons into one of six categories: mechanosensory, anterior sensory, amphid sensory, first-layer interneurons, second-layer interneurons and motor neurons.

With few exceptions, the first two categories consist of sensory neurons in the anterior ganglion that send processes into the NR. The first category, mechanosensory neurons respond to touch stimuli and primarily innervate motor neurons and

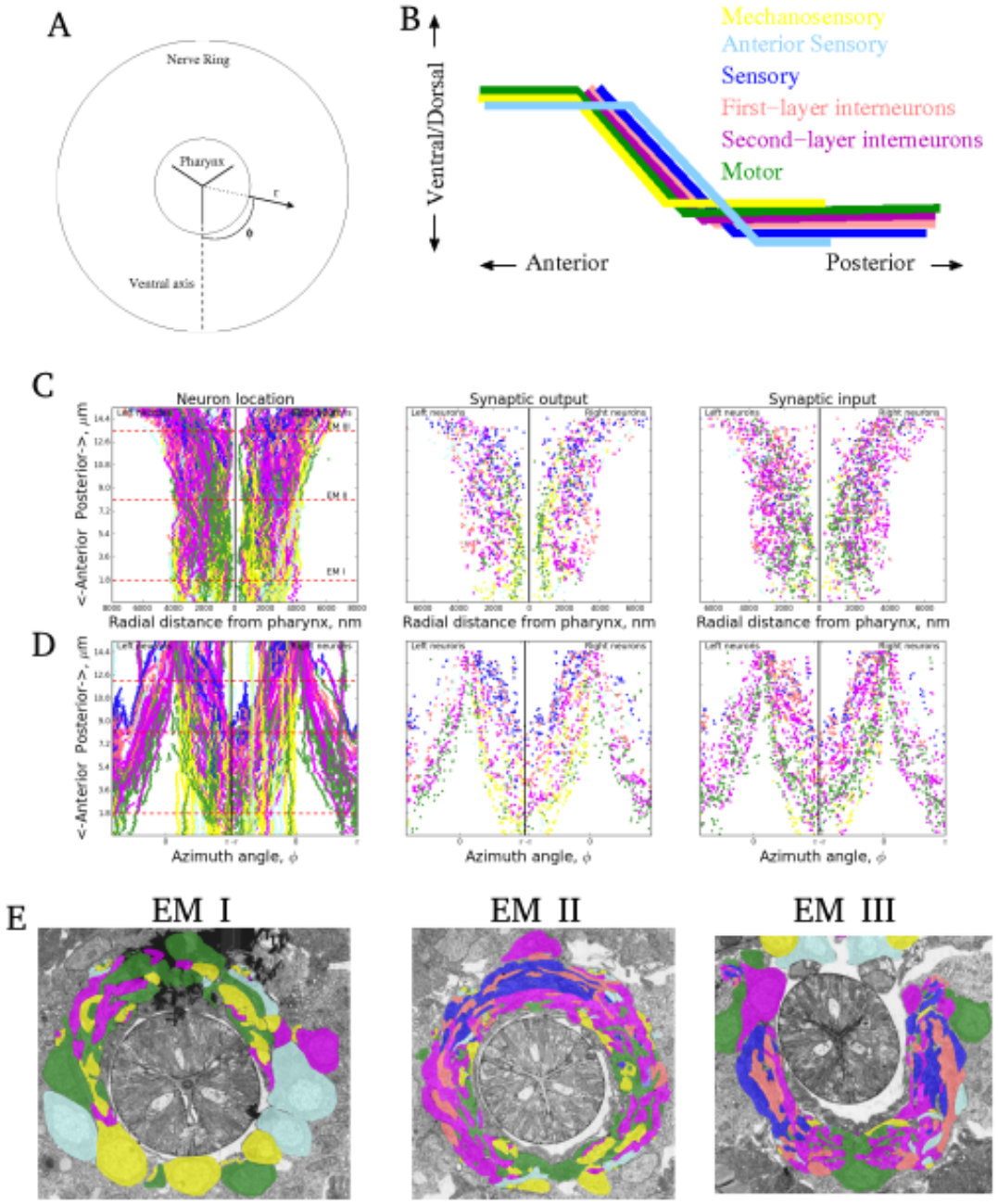


Figure 4.1: **Spatial organization of the NR.** (A) Positions of neurons/synapses are given in terms of cylindrical coordinates (r, ϕ, z). r is the distance to the outer edge of pharynx. ϕ is azimuth angle with respect to the ventral axis; $+\phi$ is clockwise; $-\phi$ is anti-clockwise. z is the position along the anterior-posterior axis. (B) Heuristic model for the NR organization. Neurons are grouped into 6 categories: anterior sensory (cyan), mechanosensory (yellow), amphid sensory (blue), first-layer interneurons (pink), second-layer interneurons (magenta), motor neurons (green). See main text for explanation. (C) Radial distance (r) of neuron processes (left), synaptic outputs (middle) and synaptic inputs (right) as a function of z. Red dashed lines show z positions of representative EM sections, shown in E. (D) ϕ as a function of z for neuron processes, synaptic outputs and synaptic inputs. (E) Representative EM sections taken from z positions given in (C).

muscles within the inner segments of the NR. This placement of mechanosensory neurons has obvious implications for synaptic connectivity and ultimately behavior. The majority of mechanosensory synaptic output is onto motor neurons and muscle. By placing mechanosensory neurons next to motor neurons and muscle, the nervous system reduces the delay between touch stimuli and resulting behavior. The benefits of this are straightforward. If a predator is close enough to touch the animal, then the animal needs to be able to respond and escape with minimal delay. The second category, anterior sensory neurons consists of the remaining sensory neurons in the anterior ganglia. The majority of these neurons are primarily involved in O_2 sensation and CO_2 avoidance. Other functions of these neurons include lifespan regulation and pheromone sensing. These neurons primarily innervate the outer segments of the NR along with interneurons. Thus, sensory information from these neurons incurs more processing steps.

The third category consists of the amphid sensory neurons which have cell bodies in the lateral ganglion. Amphid sensory cell bodies send processes that wrap circumferentially around the ventral body wall muscles and then fasciculate together into a large nerve bundle under the ventral ganglion where they are surrounded by interneurons and motor neurons. The processes begin in the lower quadrants of the nerve bundle (Fig. 4.1e), but then migrate to the outside and eventually to the upper quadrants of the nerve bundle. Once the amphid sensory processes enter the nerve ring, they appear to mix more freely with first-layer and to a lesser extent second-layer interneurons. Unlike mechanosensory neurons, the amphid sensory neurons have large r distances, placing these neurons in the outer segments of the NR. The amphid sensory processes extend the shortest distance into the nerve ring. They are among the first processes to enter the NR from the posterior side and are among the first posterior processes to terminate at the commissure.

Interneurons have cell bodies in posterior ganglia and send processes anteriorly into the NR. The interneurons are separated into two categories first- and second-layer, based on synaptic connectivity. The first-layer interneurons receive significant synaptic input from the amphid sensory neurons. These neurons follow and mix with amphid sensory neurons within the NR and proceed only slightly farther than amphid sensory neurons into the NR. First-layer interneurons are

followed by second-layer interneurons. Second-layer interneurons mix with first-layer interneurons and motor neurons and bridge the connectivity between the two neuron categories. Among interneurons, second-layer interneurons proceed the farthest into the NR, allowing more opportunity for contact with motor neurons.

The final category consists of motor neurons which form neuromuscular junctions onto head muscle arms. Motor neuron cell bodies are located in both the anterior and posterior ganglia and thus innervate the NR from both sides. Coming from either side, motor neurons primarily mix with mechanosensory and second-layer interneurons. Motor neurons primarily innervate the inner segments of the NR, conveniently placing them next to head muscle arms. Unlike most sensory and interneurons, motor neurons span the entire length of the NR along the z-axis.

In summary, the NR exhibits clear macro-level spatial organization. It is useful to think of the NR in terms of overlapping layers (Fig. 4.1b). Mechanosensory and motor processes span the entire length of the NR within the inner cylindrical segments. The anterior sensory processes extend posteriorly within the outer cylindrical segments along the entire length of the NR. From the posterior side, amphid sensory neurons are the first to enter and terminate within the outer cylindrical segments. The first- and second-layer interneurons then fill the gap between amphid sensory and motor neurons. This macro-level organization is both bilaterally conserved between left/right neurons and between the L4 and adult. To be clear, this abstract spatial model of the NR is a simplification because it attempts to place neuron categories that are continuously mixing into discreet layers. However, the model provides a useful heuristic for thinking about the spatial organization within the NR.

4.2 NR exhibits synaptic specificity

Synapse specificity refers to the precision with which neurons choose synaptic partners. A number of studies have shown that neurons only make synapses with a small fraction of their neighbors (Hamos *et al.*, 1987; White *et al.*, 1986). Some have taken this as support that neurons actively choose synaptic partners

from surrounding cells (Shen & Scheiffele, 2010). However, a small connectivity fraction only indicates that a small number of synaptic partners are chosen, it does not rule out that the synaptic partners were chosen randomly.

Before proceeding, it is useful to develop some terminology. Left/right neuron pairs are referred to as homologous neurons. For example, (ASHL,ASHR) and (AVAL,AVAR) are both homologous neuron pairs. Because AVAL and AVAR are physically adjacent neighbors of ASHL and ASHR,respectively, we say that AVA is a homologous neighbor of ASH. Bilaterally conserved synaptic connections are synaptic connections that occur on both the left and right side of the animal. For example, the synaptic connections ASHL→AVAL and ASHR→AVAR are bilaterally conserved connections. We also say that ASH→AVA is a symmetric connection. A synaptic connection that is not bilaterally conserved, i.e. a synaptic connection that occurs on either the left or right side, is said to be an asymmetric connection. An asymmetric connection on the left side is said to be left asymmetric while an asymmetric connection on the right side is said to be right asymmetric.

One way to test if synaptic partners are chosen randomly is to assess if the number of symmetric connections could have occurred randomly. I tested the null-hypothesis that the observed number of symmetric connections is random. I compared the actual number, c_a, of symmetric connections with the expected number, c_e, of symmetric connections had synaptic partners been chosen randomly. If $c_a > c_e$ and the probability of randomly observing c_a connections is small, then the null-hypothesis can be rejected.

The expected number of random symmetric connections is computed as follows. Let M be the number of homologous neighbors, s the number of symmetric connections, a_l the number of left asymmetric connections and a_r the number of right asymmetric connections. The left and right connectivity fraction are given by $\frac{s+a_l}{M}$ and $\frac{s+a_r}{M}$, respectively. The left/right connectivity fractions are assumed to be constant while the choice of synaptic partners is random. The number of ways of randomly choosing $s+a_l$ synaptic partners from M neighbors is given by the binomial coefficient $\binom{M}{s+a_l}$. The number of ways of choosing $s+a_r$ synaptic partners from M neighbors is $\binom{M}{s+a_r}$. The number of possible combinations between the left and right homologous neuron is given by $\binom{M}{s+a_l}\binom{M}{s+a_r}$. The number

of ways of having s symmetric connections is given by the multinomial coefficient

$$\binom{M}{s, a_l, a_r} = \binom{M}{s}\binom{M-s}{a_l}\binom{M-s-a_l}{a_r}. \tag{4.1}$$

Therefore, the probability of randomly having s symmetric connections is given by

$$p(s) = \frac{\binom{M}{s,a_l,a_r}}{\binom{M}{s+a_l}\binom{M}{s+a_r}}. \tag{4.2}$$

To test the null hypothesis, we need to compute the probability of having s or greater symmetric connections. Without loss of generality, assume that $a_l \leq a_r$. Then the maximum possible number of symmetric connections is $s + a_l$. Let k be a dummy variable such that $0 \leq k \leq a_l$. Note that if the number of symmetric connections is increased to $s + k$, then the number of left and right asymmetric connection must be reduced to $a_l - k$ and $a_r - k$, respectively. The number of possible ways of having $s + k$ symmetric connections is given by

$$\binom{M}{s+k, a_l-k, a_r-k} = \binom{M}{s+k}\binom{M-s-k}{a_l-k}\binom{M-s-a_l}{a_r-k}. \tag{4.3}$$

Then the probability of having $s + k$ symmetric connections is given by

$$p(k) = \frac{\binom{M}{s-k,a_l-k,a_r-k}}{\binom{M}{s+a_l}\binom{M}{s+a_r}}, \tag{4.4}$$

where the probability is a function of k and not $s + k$ because s is held constant while k is allowed to vary. Finally, the probability of observing at least s symmetric connections is given by

$$\Pr(k \geq 0)) = \sum_{k=0}^{a_l} p(k). \tag{4.5}$$

Using the standard Type I error rate $\alpha = 0.05$, we say that a given pair of homologous neurons exhibit specificity if $\Pr(k \geq 0) \leq 0.05$. Here, we have computed the probability of bilaterally conserved presynaptic connections, but the probability of bilaterally conserved postsynaptic connections is computed in a similar way.

I computed the probability of pre- and postsynaptic bilaterally conserved connections for each pair of homologous neurons in both the L4 and adult. Most

neurons (>60%) have probabilities less than 0.05, indicating that for most neurons the conserved connections are statistically not random and the null-hypothesis can be rejected (Figure 4.2a-b). Hence, the connectivity appears to be specified. I find that 45% of neurons have both pre- and postsynaptic connections specified, 41% of neurons have either pre- or postsynaptic connections specified and for 14% of neurons we cannot rule out that both pre- and postsynaptic connections are random (Figure 4.2c). It should be noted that the second group contains neurons which either have relatively few presynaptic connections (e.g. locomotion interneurons AVA/AVB/AVE/AVD which have significant synaptic output in the body but not the NR) or relatively few postsynaptic connections (e.g. many sensory neurons), which could explain why probability of conserved matches is so high. For the third group of neurons, I find that less than half of the synaptic connections are bilaterally conserved. For the SDQ neurons, this is due to morphological differences. For the other neurons, this could suggest that these neurons are developmentally or functionally different.

In addition to testing if conserved connections are random, I also tested if developmentally conserved connections are random. Developmentally conserved connections are synaptic connections that occur in both the L4 and the adult. For example, if the synaptic connection ASHL→AVAL is present in both the L4 and the adult, the connection is said to be developmentally conserved. If the appropriate substitutions are made, equation (4.5) can also be used to compute the probability of observing at least s developmentally conserved connections. Specifically for a given neuron, let M the number of shared neighbors in both the L4 and the adult, let a_l be the number of synaptic connection in the L4 but not the adult and let a_r be the number of synaptic connections in the adult but not the L4. The probability of developmentally conserved connections is less that 0.05 for $\sim$ 60% of neurons (Figure 4.2a-b). In general, I also find that many neurons either have pre- or postsynaptic developmentally conserved connections. Thus, in addition to be bilaterally conserved, synaptic connections are also developmentally conserved.

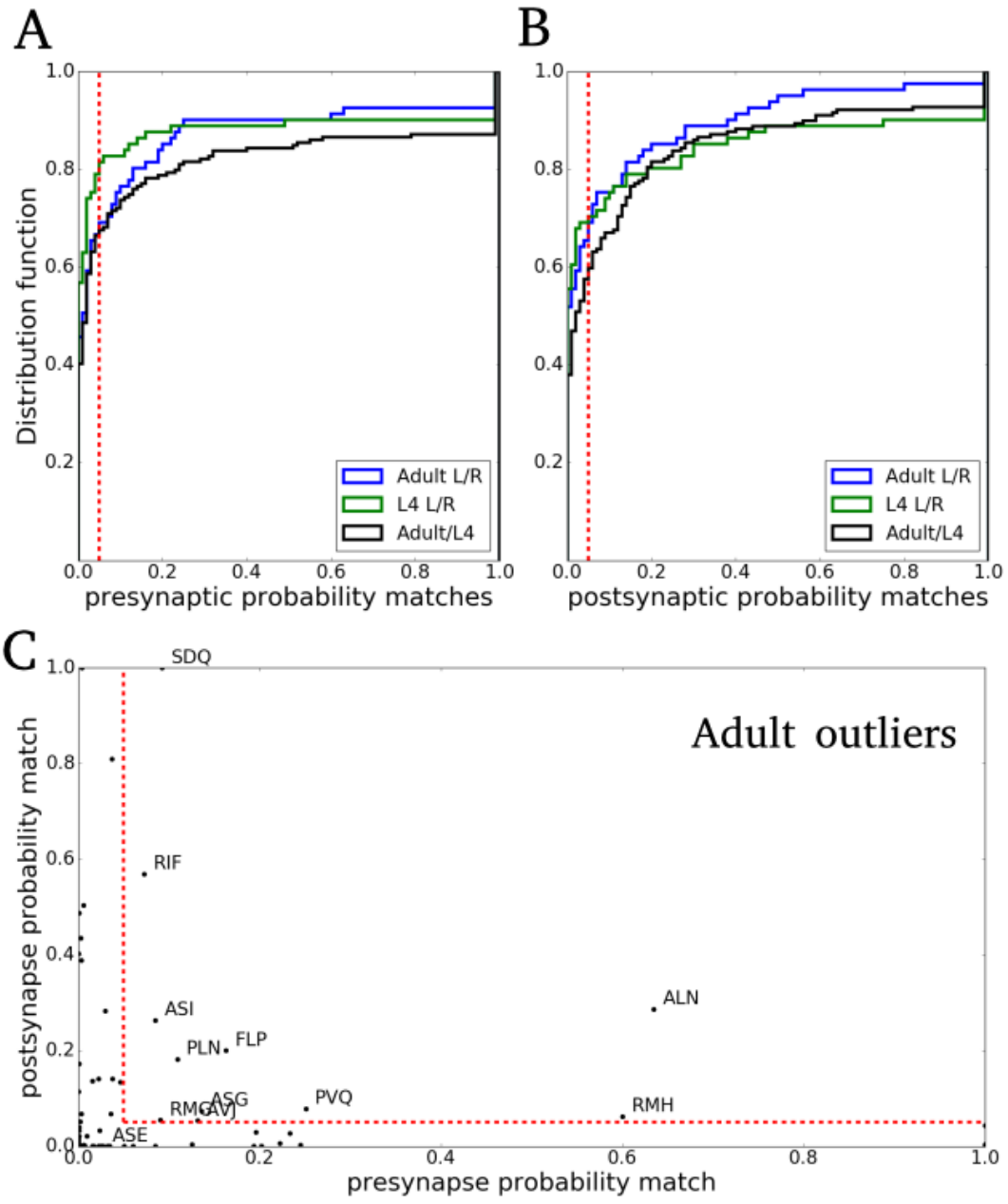


Figure 4.2: **Probability of conserved connections.** (A) Distribution function of the probability of presynaptic connections the are bilaterally conserved in the adult (blue), L4 (green) and developmentally conserved between the L4 and adult (black). (B) Distribution function of the probability of bilaterally and developmentally conserved postsynaptic connections. (C) Plot of bilaterally conserved presynaptic vs. postsynaptic connections in the adult. Outlier homologous neurons are labeled. Homologous neurons are considered outliers if both the pre- and postsynaptic probabilities are greater than $\alpha = 0.05$. In all plots, red dashed line marks where the probability is 0.05.

4.3 NR exhibits subcellular specificity

4.4 Modeling synaptic connectivity

Specified synaptic connections occur with higher probability than would be expected by chance, but what factors increase or decrease the probability of synapse formation? In this section, I outline the framework of a statistical model that will be used to address this question. The model will be used to predict the number of synaptic connections for each neuron and will be evaluated based on how closely the predicted value matches the measured number of synaptic connections.

The model starts by defining the probability of a synaptic connection between neurons i and j. Let X_{ij} and Y_{ij} be two random variables that relate i to j. The random variable X_{ij} describes some measured quantity, e.g. the physical contact, and $Y_{ij} = 1$ if there is a synaptic connection from i to j and $Y_{ij} = 0$ otherwise. Then the joint probability function (p.f.) $f_{XY}(x, y)$ describes the sample space of possible connections and $f_X(x)$ is the marginal p.f. from summing over all possible $Y_{ij} \in \{0, 1\}$. The probability of a synaptic connection from i to j given $X_{ij} = x$ is

$$\Pr(Y_{ij} = 1 | X_{ij} = x) = \frac{\Pr(Y_{ij} = 1 \cap X_{ij} = x)}{\Pr(X_{ij})} = \frac{f_{XY}(x, 1)}{f_X(x)}, \tag{4.6}$$

where the middle expression is the definition of a conditional probability. Thus, in order to compute the synapse probability, we only need to compute $f_{XY}(x, 1)$ and $f_x(x)$ which can be derived from the associated distribution functions (d.f.) $F_{XY}(x, 1)$ and $F_X(x)$, respectively. The final expression of synapse probability is given by

$$\Pr(x) = \frac{F'_{XY}(x, 1)}{F'_X(x)}. \tag{4.7}$$

where F'_{XY} and F'_X are the derivatives of the distribution functions and on the left I have dropped the more formal notation of Equation (4.6) for clarity.

In practice, deriving the distribution functions from the data is straightforward. $F_X(x)$ is the cumulative distribution of possible X values and $F_{XY}(x, 1)$

is the cumulative distribution of all X values where there is also a synaptic connection. The distributions are then fit with a modified logistic equation,

$$g(x) = A\left(\frac{1}{D + e^{-Bx}} - C\right) \tag{4.8}$$

where A, B, C and D are constants determined using the Levenberg-Marquardt algorithm for least squares curve fitting. For the random values of X analyzed in this study (touch density and p-score, see below), equation (4.8) provided a good fit for the distribution functions. Equation (4.7) states that the derivatives of the distribution functions are needed, which are given by the derivative of (4.8),

$$g'(x) = A\left(\frac{Be^{-Bx}}{(D + e^{-Bx})^2}\right). \tag{4.9}$$

In summary, synapse probability is computed by (i) deriving the distribution functions from the data, (ii) computing the constants needed to fit the distribution functions with (4.8), (iii) plugging the constants into (4.9) to determine the derivatives of the distribution functions and (iv) plugging the derivatives into (4.7) to compute the probability.

In practice, I found that (4.7) can be ill-conditioned depending on the behavior of the distribution functions. As $F_X \to 1$ for $x \to \infty$, the derivative $F'_X \to 0$ which can cause the expression in (4.7) to go to infinity. This behavior was mitigated by deriving an equivalent expression for (4.7). Let $h = \frac{F_{XY}}{F_X}$, then the chain rule for derivatives gives

$$h' = \frac{F'_{XY}}{F_X} - \frac{F'_X F_{XY}}{(F_X)^2} \tag{4.10}$$

Rearranging terms gives the expression

$$\frac{F'_{XY}}{F'_X} = \frac{F_X}{F'_X}h' + h \tag{4.11}$$

So alternatively, the synapse probability can be computed as

$$\Pr(x) = \frac{F_X(x)}{F'_X(x)}h'(x) + h(x). \tag{4.12}$$

I assessed the model's ability to capture variation in synaptic connectivity among neurons by comparing the actual number of synaptic connections for each

neuron with the value predicted by the model. Let the random variable $Z_i = Y_{i1} + Y_{i2} + \cdots + Y_{iM}$ be the total number of synaptic connections that neuron i makes with its M neighbors. Y_{ij} is binomially distributed and I assume that each synaptic connection is made independently of all the other connections. Therefore, the expected number of synaptic connections is given by

$$E(Z_i) = \sum_{j}^{M} \Pr(Y_{ij} = 1 | X_{ij} = x) \tag{4.13}$$

with variance

$$Var(Z_i) = \sum_{j}^{M} \Pr(Y_{ij} = 1 | X_{ij} = x)(1 - Pr(Y_{ij} = 1 | X_{ij} = x)) \tag{4.14}$$

I next compute the p-value, the probability of observing a discrepancy as great or greater by chance between the actual and expected number of synaptic connections. A representative p-value for all the neurons is computed using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001), which corrects for the increased chance of observing a Type I error (i.e. falsely rejecting the null hypothesis) and has greater statistical power than the more commonly used Bonferroni correction (Perneger, 1998). For m neurons, the p values are arranged in ascending order, $p_1 \leq p_2 \leq \ldots \leq p_m$, and each p value is adjusted to $p_i^a = \min(mp_i/i, 1)$. The multiple hypothesis adjusted p value is defined as $p_{adj} = \min(\{p_i^a\})$, which is then compared to the false discovery rate $\alpha = 0.05$. When $p_{adj} < 0.05$, we reject the null-hypothesis that the model captures the variation in synaptic connectivity.

4.5 Touch density does not predict connectivity

Synaptic connections require a minimal amount of mutual contact between neurons, but does the likelihood of a synaptic connection increase with the amount of contact? Two observations suggest that synaptic connectivity does not depend strongly on the amount of physical contact. First, the distribution function of synaptic connections as a function of touch density shows that 40% of synaptic connections require less that 1% of membrane contact (Figure 4.3). Second, there

is no correlation between touch density and the number of synapses formed between neurons (Figure 3.4c-d). These observations show that (i) synapses require very little contact between neurons and (ii) the number and size of synapses does not strongly depend on the amount of physical contact. But does more contact increase the probability of forming at least one synapse between neurons?

Durbin (1987) noted that over as little as 10 EM sections of mutual contact, there is a sharp rise in synapse frequency. I also observe a similar correlation between touch density and synapse likelihood. Let the random variable X be the touch densities. Figure 4.3 shows the marginal distribution function $F_X(x)$ of touch densities and the distribution function $F_{XY}(x, 1)$ of touch densities that result in a synaptic connection. The probability of forming a synaptic connection as a function of touch density x is computed using (4.12). Most of the increase in synapse probability occurs over 10% of membrane contact, after which the probability asymptotes to 0.7. This shows that more contact between neurons may increases the chances of a synapse but does not guarantee a synapse. But does touch density predict synaptic connectivity?

For each neuron, I loop through each of its neighbors, look up the touch density and the subsequent probability of a synaptic connection from Figure 4.3b. The predicted number of synaptic connections is then the sum of these probabilities. I then compute the probability of observing a discrepancy as large or larger between the predicted and actual number of synaptic connections. For both pre- and post synaptic connections, probability of discrepancy is $p < 0.05$ (Figure 4.3c-d). Thus, the null-hypothesis that touch density is a good predictor of synaptic connectivity must be rejected.

4.6 Connectivity fraction product predicts connectivity

I now consider a model where the probability of a synaptic connection depends on the pre and post connectivity fraction. In this model, I assume that every neuron has independent probabilities of being pre- or postsynaptic and that the pre- and postsynaptic probabilities of any two neurons are independent. Define

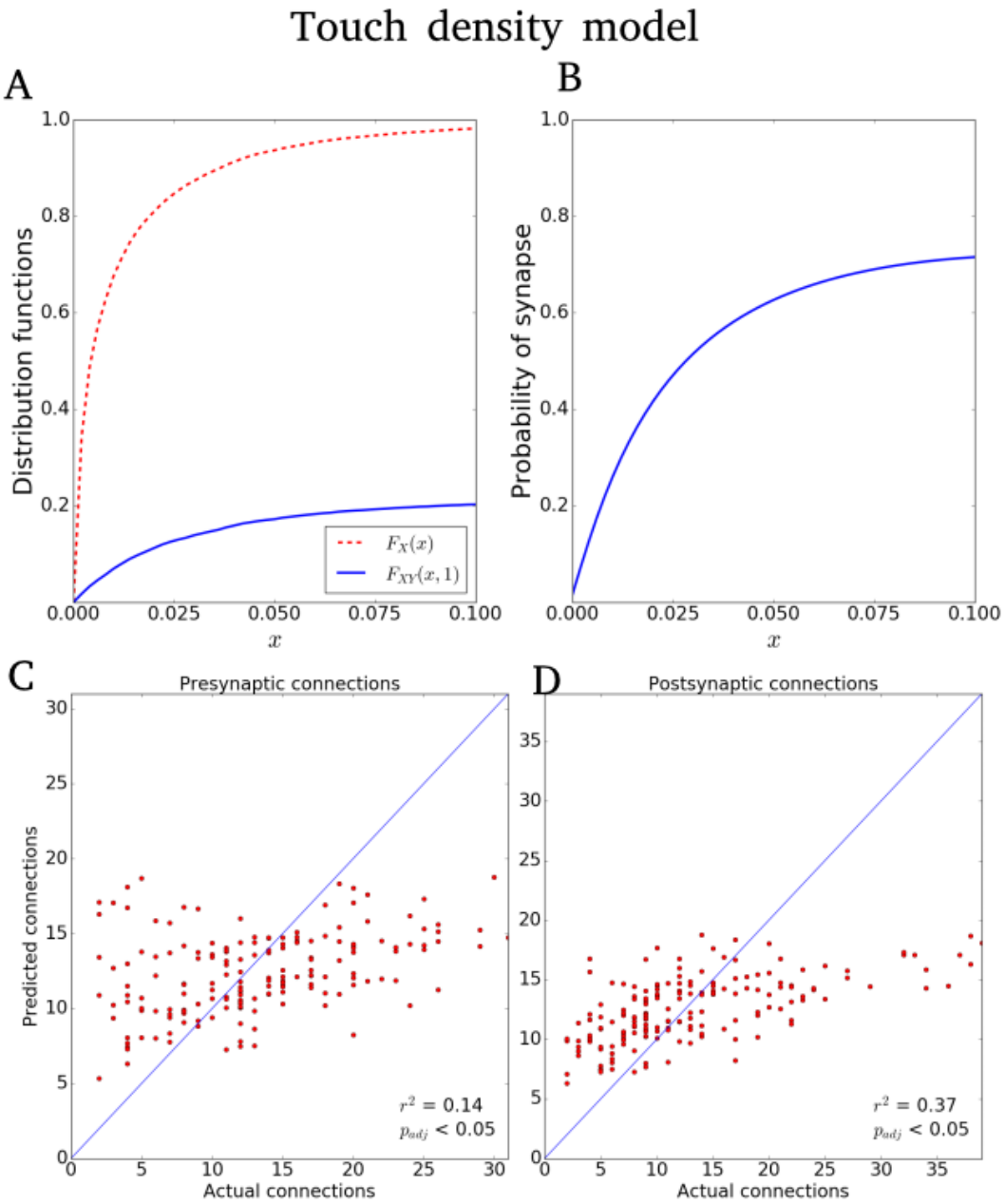


Figure 4.3: **Touch density does not predict connectivity** (A) The dashed line is the marginal distribution of touch densities. Solid line is the joint distributions of touch densities with an associated synaptic contact. (B) Probability of forming a synaptic connection as a function of touch density. (C-D) Touch density does not predict the number of synaptic connections, $p < 0.05$.

the random variable

$$X_{ij} = C_i^{pre} C_j^{post} \tag{4.15}$$

where C_i^{pre} is the pre connectivity fraction of neuron i and C_j^{post} is the post connectivity fraction of neuron j (see Equation 3.5). In words, X_{ij} reflects the product of probabilities that neuron i is presynaptic and neuron j is postsynaptic. I call X_{ij} the product score from neuron i to neuron j. For brevity, I will refer to the product score as simply the p-score in the remainder of this study. The reader should be careful not to confuse p-score and p-value.

The marginal distribution $F_X(x)$ of all p-scores and the joint distribution $F_{XY}(x, 1)$ of p-scores with a synaptic connection are given Figure 4.4(a). The probability of a synaptic connection, computed using (4.12), is given in Figure 4.4b. The number of synaptic connection predicted by the p-score model is in close agreement with the actual number of connections (Figure 4.4c-d). The probability of as large or larger discrepancy is $p > 0.5$. Thus, the null hypothesis that the probability of synaptic connection depends on the product of pre and post connectivity fractions cannot be rejected.

I tested the robustness of this model in two ways: by comparing left/right connectivity and by comparing L4/adult connectivity. For left/right connectivity, I determined distributions of p-scores for neurons on the left side of the animal and then used the derived probability to predict the number of connections on the right side of the animal. For both the L4 and adult, the derived left probability function predicts the actual number of connections on the right with $p > 0.05$ (data not shown). Thus, the model is robust to intra-worm variability. For L4/adult connectivity, I determined the distribution of p-scores for L4 neurons and then used the derived probability to predict the number of synaptic connections in the adult. The derived L4 probability function is not a good predictor of synaptic connections in the adult ($p < 0.05$, data not shown). Thus, the model is not robust to developmental differences in synaptic connectivity between the L4 and adult. This could be an indicator that there are continuing developmental changes in the pre- and postsynaptic properties of neurons from the L4 to the adult. These continued developmental changes could account for the increased connectivity observed in the adult relative to the L4. At this point, it is worth addressing a few points concerning the p-score model.

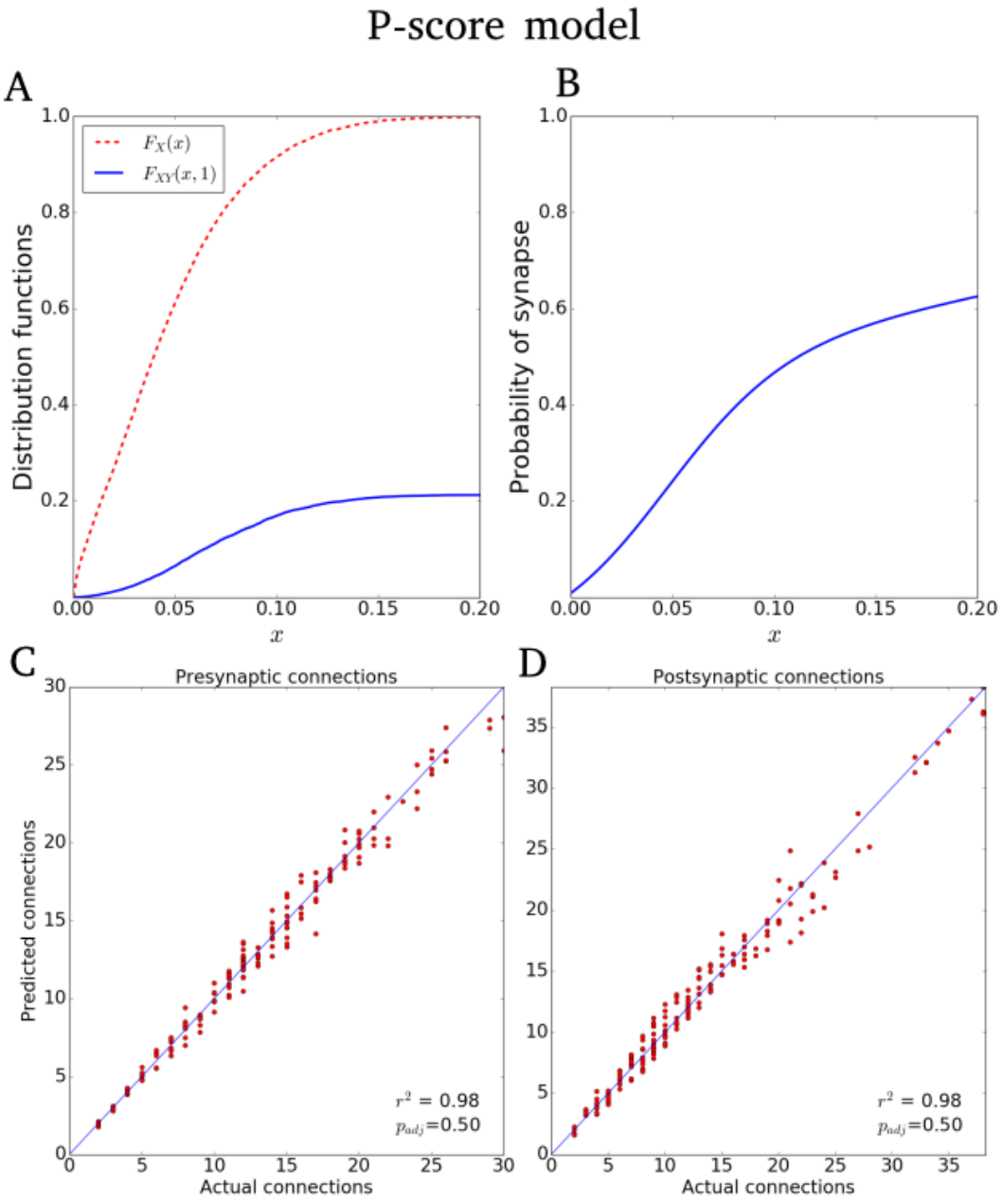


Figure 4.4: **P-score model predicts connectivity.** (A) The dashed line is the marginal distribution of p-scores. Solid line is the joint distributions of p-scores with an associated synaptic contact. (B) Probability of forming a synaptic connection as a function of p-score. (C-D) P-scores predict the number of synaptic connections, $p > 0.5$.

It is not obvious that the p-score model should predict synaptic connectivity as well as it does. One may be tempted to interpret the connectivity fractions as synaptic probabilities. In which case, the synapse probability could be the product of the pre- and postsynaptic probability. However, there are a few reasons why this is faulty reasoning. First, the connectivity fractions are not probabilities. The connectivity fraction is a single measurement of the likelihood that a neuron is pre- or postsynaptic. In order to properly estimate the true pre- and postsynaptic probabilities of a neuron, one would have to make repeated measurements of the connectivity fractions for each neuron. As it stands, I have only one adult and one L4 connectivity fraction for each neuron, which is insufficient to characterize the pre- and postsynaptic probabilities. Second, I assessed the p-score model's ability to capture variation in synapse probability not the mean. The p-score model is essentially an average synapse probability approximated from the connectivity fractions. However, there is no obvious reason to expect the synapse probabilities to be tightly distributed around the mean. Indeed, for the touch density model the distribution around the mean was large, which caused the model to be rejected. That the p-score model is able to predict the actual synaptic connectivity in the data, indicates the spread of synapse probabilities around the mean is quite small. This suggests that the p-scores are a good description of the data.

Unlike touch density, it is difficult to provide a precise physical interpretation of the p-score. The p-score is the product of connectivity fractions which in turn are simplified representations of the connectivity properties of neurons. There are are likely many underlying molecular mechanisms that work in combination and have evolved over time to give rise to the observed connectivity fractions. The p-score is a very simple way of scoring how these mechanisms link neighboring neurons. Higher p-scores are correlated with a higher synapse probability (Figure 4.4b).

Finally, what are the biological implications of this result? The two assumptions of the model are (i) neurons have a pre and post connectivity fraction and (ii) the synapse probability is proportional to the product of the pre and post connectivity fractions of neighboring neurons. The first assumption indicates that the mechanisms for giving and receiving a synapse are different, which is consistent with what has been observed experimentally. For example, the cell adhesion

proteins neurexin and neuroligin are known to help bind cells at the synaptic cleft and are expressed in the pre- and postsynaptic neuron, respectively. The second assumption indicates that the synapse probability depends on the connectivity properties of both the pre- and postsynaptic neurons and that these properties are largely independent. One could imagine a scenario where the synapse probability is only determined by connectivity traits of either the pre- or postsynaptic neuron, but not both. However, the p-score model suggests that the pre- and postsynaptic neurons have cell autonomous pre and post connectivity traits, respectively, and it is the interaction of these traits that determines the probability of a synapse. Furthermore, the pre and post connectivity properties interact in a nonlinear way to give rise to the synapse probability (Figure 4.4b). Finally, the p-score model is probabilistic and its ability to predict connectivity suggests that synapses are formed in a probabilistic manner. Previously, I showed that there are more bilaterally and developmentally conserved synaptic connections than would be expected by chance. In light of the p-score model, these results suggest that conserved synaptic connections are synaptic connections that occur with very high probability. Assumption (ii) would then suggest that the conserved synaptic connections involve pre and post connectivity traits that increase synapse probability. Conversely, the nonconserved synaptic connections can still occur, but they occur with a decreased probability. This could suggest that it is the differential expression of pre and post connectivity traits among neurons that mediates the synaptic specificity in the NR.

Chapter 5

Combinatorial CAM expression can support synaptic specificity

Synaptic specificity requires that neurons be able to identify synaptic partners from the pool of physically adjacent cells. In the vertebrate central nervous, there is on the order of 10^4 cell types, making the task of uniquely labeling cells challenging. What molecular-genetic mechanisms exist that could uniquely label such a vast number of cells? One proposal is that unique cell labels could be created by the combinatorial expression of a small-number of cell surface molecules (Baier, 2013). Informal calculations indicate that 3×10^{10} variations of clustered protocadherins could be expressed in each neuron in the brain (Yagi, 2012) and experimental finding have shown that surface proteins can exist in tens of thousands of isoforms each with their own binding affinities (Wojtowicz *et al.*, 2007). The information-carrying capacity of combinatorial codes is very large and would grow exponentially with each additional molecule. A combinatorial code with two molecules x and y could encode 4 unique labels (not x and not y, x and not y, y and not x, x and y). For N molecules, the number of possible labels is 2^N.

To encode unique labels for each of the 180 *C. elegans* NR neurons would only require 8 molecules ($2^8 = 256$) and there are at least 35 CAM genes expressed in the NR (Table 5.1). So theoretically, the combinatorial expression model could be a plausible explanation for the observed synaptic specificity in the NR. However, this calculation is overly simplistic, because it does not take into account the actual expression patterns of the *C. elegans* NR nor how expression patterns are

coordinated between synaptic partners. A critical assessment of the feasibility of the combinatorial expression model should address whether or not the known expression patterns in the NR are sufficiently unique to allow neurons to discriminate between synaptic partners and other nonsynaptic physically adjacent cells.

5.1 CAM expression in the *C. elegans* NR

In a review of the *C. elegans* neuronal genome, Hobert (2005) identified 106 CAM genes. Wormbase (WS259) identifies 3950 genes that are expressed in the head neurons (Wormbase, 2017). Comparing the two lists, there are 55 CAM genes that are expressed in NR neurons. These CAM genes were placed into one of two categories, CAM I and CAM II, based on how precisely the expression patterns have been characterized. CAM I consists of 35 genes whose expression in the NR has been clearly identified and linked to specific neurons (Table 5.1). CAM II consists of 17 genes that are said to be expressed in all NR neurons or for whom subsets of NR neurons are not clearly identified (Table 5.2). For example, expression of the CAM II gene *egl-15* is observed in hypodermal cells, sex myoblasts, the type I vulva muscles and some "unidentified" head neurons (Huang & Stern, 2004). Because the head neurons were not clearly identified, *egl-15* was placed in CAM II. In order to keep the results as conservative as possible, CAM II neurons were removed from the analysis and only CAM I genes were considered. Henceforth, when CAM genes are mentioned, it should be understood that the CAM I genes are being referenced.

The expression of CAM genes in the NR is sparse with neurons typically expressing a relatively small number of genes and single genes being expressed across multiple neurons. There are 28 NR neurons that have no known CAM expression and were subsequently removed from the analysis. However, most of the remaining neurons express up to 5 CAM genes (Figure 5.1d) and over 60% of CAM genes are expressed in at least 5 neurons (Figure 5.1e). Neuron PVT expresses the most CAM genes (11) and gene *cam-1* is expressed in the most neurons (70). There is no discernible structure to the expression matrix (Figure 5.1a). A number of algorithms were applied to the matrix (e.g. diagonalization

and bipartite graph clustering), but none yielded any meaningful oranizational insights.

Relatively few neurons have unique expression patterns. Expression patterns were determined from the expression matrix E_{ij}, where $E_{ij} = 1$ if neuron i expresses gene j and $E_{ij} = 0$ otherwise (Figure 5.1a). Hence, the ith row of the expression matrix can be used to generate a binary string that labels the expression pattern of the ith neuron. Two neurons are said to have equivalent expression patterns if they have equivalent expression labels. The NR hase 64 unique gene expression labels. Neurons with equivalent expression labels can conveneiently be displayed in a network graph, where edges exist between pairs of neurons with equivalent expression labels. The resulting graph shows 10 isolated neurons that do not share an expression label with any other neuron (Figure 5.2). The remaining neurons are linked to at least one other neuron, forming expression clusters. Much of this clustering can be attributed to homologous neurons in the same class (e.g. ASHL and ASHR) having the same expression patterns, but there are 15 clusters that consist of neurons from more than one class. Thus, despite 35 CAM genes being expressed in the NR, there are not enough expression patterns to uniquely label each neuron.

The relatively few expression labels and the large number of expression clusters would seem to suggest that there are not enough combinatorial expression patterns to satisfy the combinatorial expression model. However, given the spatial segregation of neurons, the expression label of neurons only needs to be locally unique, i.e. expression labels only need to be unique in the adjacency neighborhood of a given neuron. Furthermore, if a neuron only needs to distinguish between synaptic and nonsynaptic neighbors, then the uniqueness requirement is even weaker. The synaptic partners of a neuron may share expression labels and nonsynaptic neighbors of a neuron may share expression labels, but the overlap of expression labels between synaptic partners and nonsynaptic neighbors should be minimized.

To check for local uniqueness, I computed the fraction of posynaptic expression labels that are not shared by nonsynaptic neighbors. For each neuron, the expression label of each postsynaptic neuron was determined. The expression label of each postsynaptic neuron was then compared to the expression labels of

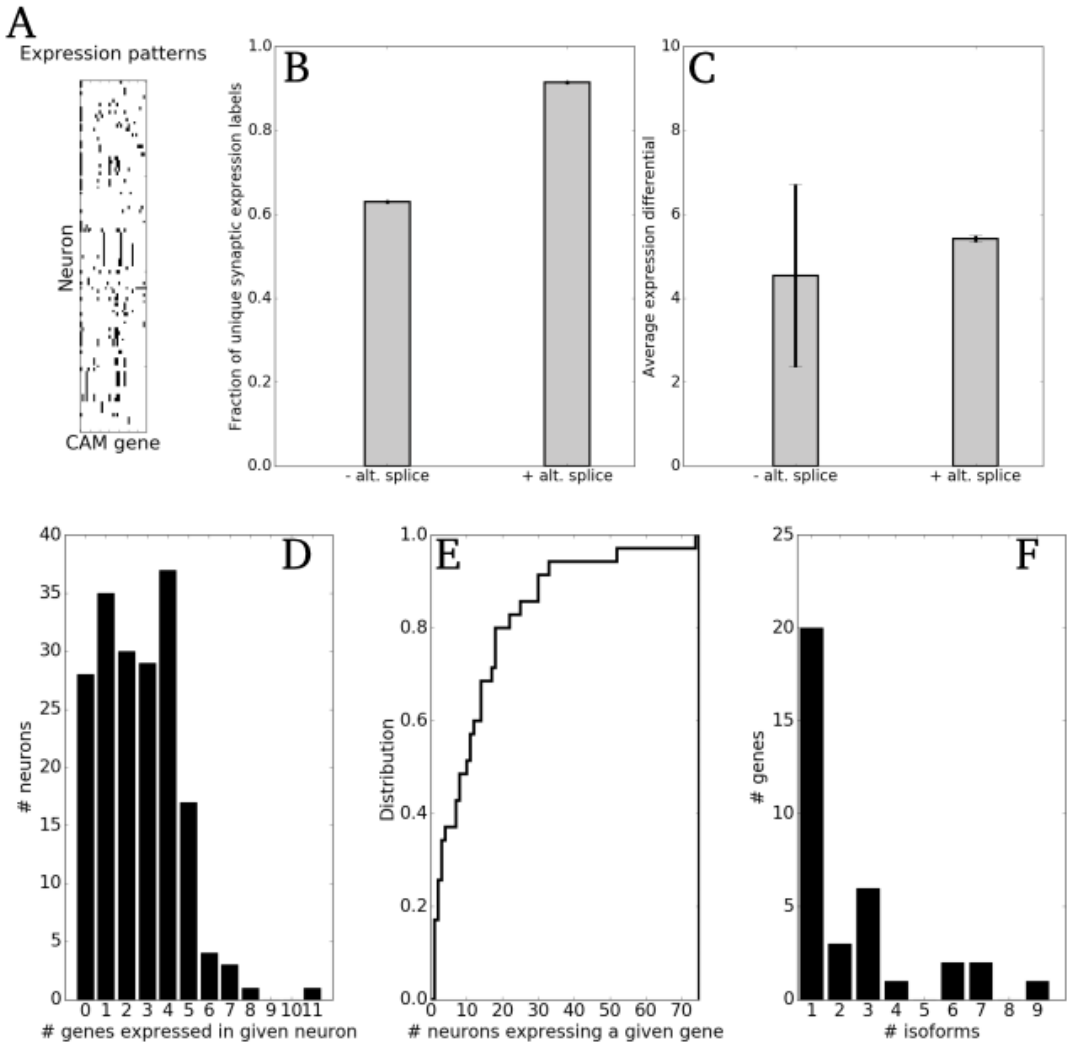


Figure 5.1: **CAM expression in the NR.** (A) Matrix of CAM genes (rows) expressed in the NR neurons (columns). (B) Fraction of postsynaptic connections that have locally unique cam expression labels without (-) and with (+) alternative splicing. (C) Average expression differential between postsynaptic and nonsynaptic neighboring neurons without and with alternative splicing. (D) Histogram of the number of genes expressed in an individual NR neuron. (E)Cumulative distribution of the number of NR neurons that express a given gene. (F) Histogram of the number of isoforms of CAM genes expressed in the NR. Error bars represent standard deviations.

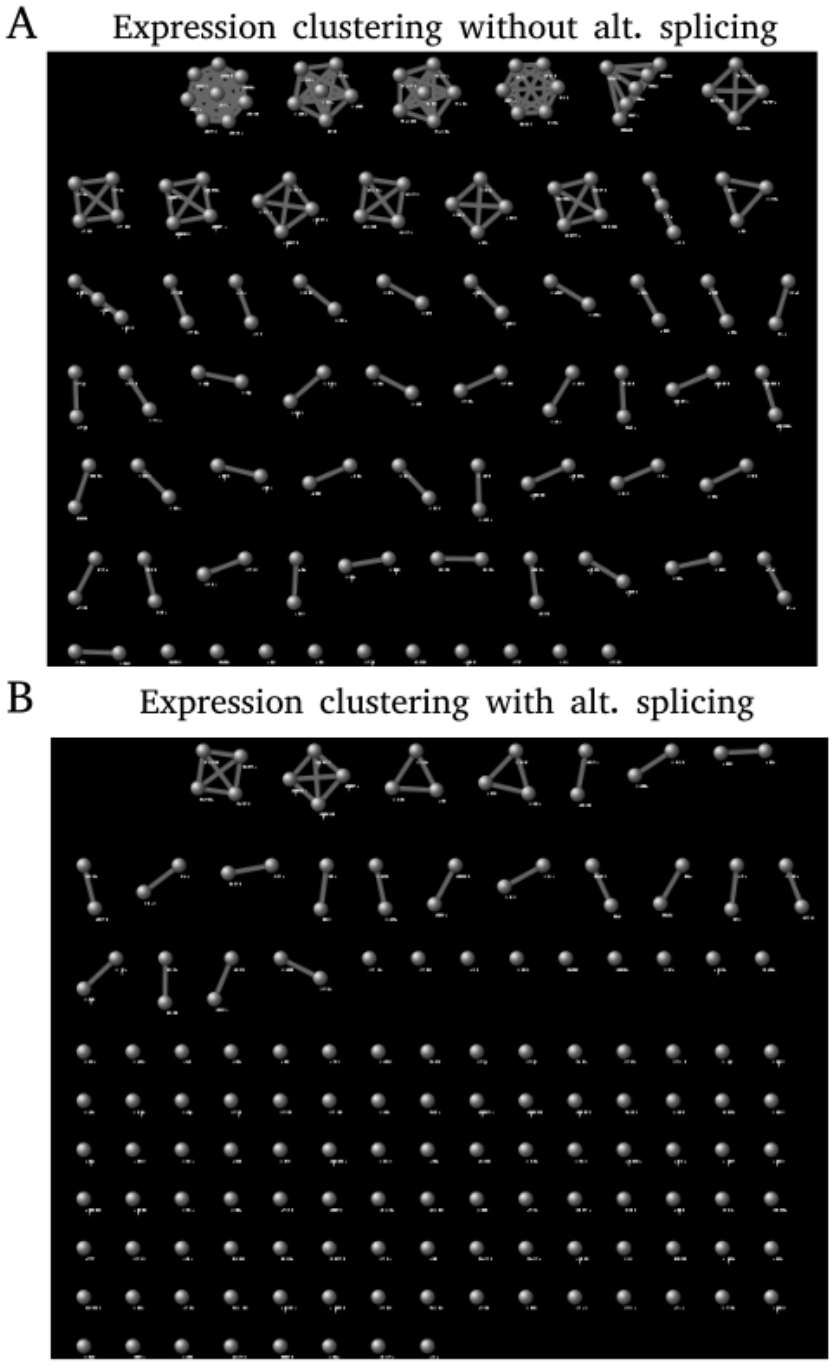


Figure 5.2: **CAM expression clustering.** Graphs showing which neurons have the same CAM expression patterns without (A) and with (B) alternative splicing. Neurons that have the same CAM expression pattern are connected by an edge. Alternative splicing gives less clusters and thus more uniquely labeled neurons.

the neuron's nonsynaptic physically adjacent neighbors. I counted the number of times a postsynaptic expression label matched at least one of the nonsynaptic neighbors expression labels. The number of matches were counted and divided by the total number of comparisons made. The local uniqueness score was computed as 1 minus this fraction, i.e.

$$\text{local uniqueness} = 1 - \frac{\#\text{ of expression label matches}}{\#\text{ of expression label comparisons}}. \tag{5.1}$$

The local uniqueness fraction is 0.65 indicating that only 2/3's of postsynaptic neurons are distinguishable from nonsynaptic neighbors based on CAM expression patterns (Figure 5.1b).

The low local uniqueness fraction can in part be attributed to how matches were determined. Matches are assessed by comparing a single postsynaptic expression label to the expression labels of all nonsynaptic neighbors. If the postsynaptic expression label is equivalent to at least one nonysnaptic expression label, then a match is scored. However, most bilaterally symmetric neurons have the same expression patterns. Thus, for example, if a neuron synapses onto AIBL and is also physically adjacent to AIBR, then a match will be scored because AIBL and AIBR have equivalent expression labels. If comparisons between bilaterally symmetric neurons are excluded, then the local uniqueness fraction increases to about 0.85. However, there is no compelling biological justification for making such an exclusion. Under the combinatorial expression model, if a neuron synapses onto AIBL then why would the same neuron not also synapse onto AIBR if both AIBL and AIBR have equivalent expression labels? Thus, we cannot make exceptions for bilaterally symmetric neurons without violating the combinatorial expression model.

These results suggest that the currently defined CAM genes expression patterns cannot sufficently provide unique labels that differentiate between synaptic partners and nonsynaptice neighbors in the NR. More unique expression labels are required for the combinatorial expression model to be feasible. To increase the number of expression labels, we must either increase the number of CAM genes or consider isoform protein expression of the current CAM genes. Evidence from other organisms suggest that the diversity in isoforms expression of surface

proteins could be used to generate unique cell labels (Zipursky & Sanes, 2010) and it has been shown that alternative splicing of a Lamellipodin homolog (MIG-10) is required for the synapse response to netrin in some *C. elegans* neurons (Stavoe *et al.*, 2012). With this in mind, I assessed if alternative splicing of CAM genes could be used to uniquely label neurons.

5.2 Alternative splicing is a necessary condition for the combinatorial expression model

Alternative splicing allows for a gene to code for multiple isoforms of proteins. Alternatively spliced CAM genes yield slightly different surface proteins which could sufficiently differentiate cells and allow neurons to identify synaptic partners. Compared to other oraganisms, *C. elegans* has relatively little isoform diversity. It is estimated that up to 25% of protein-coding genes in *C. elegans* exhibit alternative splicing (Wani & Kuroyanagi, 2017) compared to 95% in humans (Pan *et al.*, 2008). Moreover, there are rarely more than 10 isoforms expressed by an alternatively spliced CAM gene, compared to the thousands to tens-of-thousands expressed in other organisms (Zipursky & Sanes, 2010). However, the *C. elegans* nervous system is also significantly smaller and may not require such isoform diversity.

There are 15 CAM genes with known alternative splicing. These genes can code up to 9 isoforms (Figure 5.1f). Unfortunately, precise isoform expression of CAM genes in NR neurons is not generally known. Instead, I simulated alternative splicing by randomly assigning splice variants. For example, neuron ADAL expresses CAM genes *cam-1* and *sax-7* which have 3 and 6 splice variants, respectively. In the simulation, neuron ADAL is randomly assigned one of the 3 splice variants for *cam-1* and one of the 6 splice variants for *sax-7*, each with equal probability. This is repeated for each neuron. Because splice variants are simulated, each neuron has many possible expression patterns instead of a single expression pattern. For example, ADAL has $6 \times 3 = 18$ possible expression patterns due to the splice variants of *cam-1* and *sax-7*. The number of possible ways of assigning

expression patterns to all neurons is then the product of these possible individual arrangements for each neuron. There are $\sim 10^{190}$ possible ways of assigning expression patterns to all of the NR neurons. It is computationally prohibitive to test all 10^{190} possible combinations. Instead, 1000 assignments were randomly sampled and the mean results taken. On average, there are 140 unique expression patterns when alternative splices are randomly assigned. There are now only 22 clusters and 107 isolated nodes in the expression graph (Figure 5.2b). The average local uniqueness fraction is 0.9, indicating 90% of postsynaptic expression labels are distinguishable from nonsynaptic neighbors.

These results suggest that alternative splicing is a necessary conditon for the combinatorial expression model in the *C. elegans* NR. Making the additional assumption that alternatively spliced genes will encode for isoform proteins increases both the number of unique expression labels and the local uniqueness fraction. Including alternative splicing solves the major problem encountered above, namely, discriminating between bilaterally symmetric neurons. The vast majority of neuron classes express at least one CAM gene with a splice variant (Table 5.3). The results suggest that alternative gene splicing could be used to distinguish between neurons within the same class. Interestingly, a large number of neuron classes express *cam-1* (Figure 5.3) which is required for asymmetric cell division (Forrester *et al.*, 1999), indicating *cam-1* is required for differentiating cells. However, no single alternatively spliced gene is expressed across all neuron classes, indicating that no single alternatively spliced gene is used to differentiate intra-class neurons.

5.3 Neurons make use of multiple expression patterns for synaptic connectivity

Are there specific gene patterns that induce synapse formation? On average, there is a difference of 4 to 5 genes between postsynaptic neurons and nonsynaptic neighbors (Figure 5.1c). This indicates that, on average, there is a cluster of genes that differentiate postsynaptic and nonsynaptic neurons. These gene clusters are good candidates for gene combinations that elicit synapse formation because the

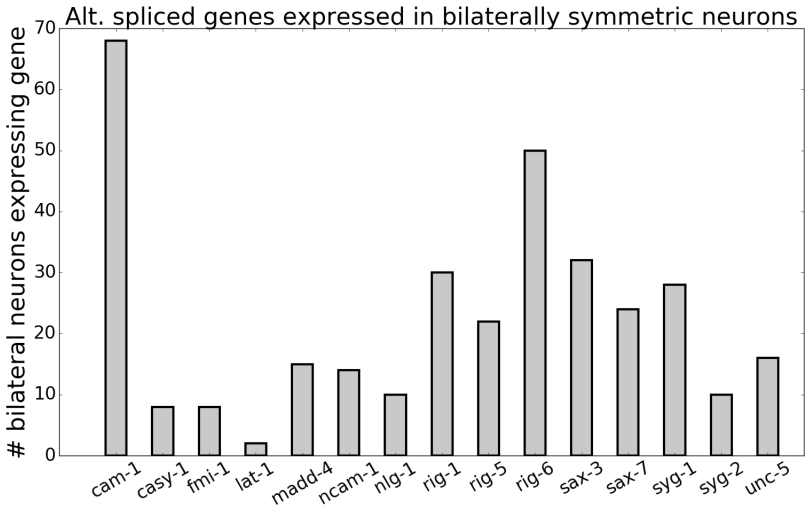


Figure 5.3: **Expression of alternatively spliced genes.** The number of neurons expressing alternatively spliced genes.

gene clusters are found in the postsynaptic partner and not in the nonsynaptic neighbor. Assuming this is the case, the next logical question is whether there is a single gene cluster or multiple gene clusters that elicit a synapse in a given presynaptic neuron. To address this question, gene expression patterns of postsynaptic neurons were compared and the unique gene clusters were determined. For this part of the analysis, alternative splicing was not considered.

Some care must be taken when comparing postsynaptic expression because most synapses are polyadic. A synapse is polyadic when there is more than one neuron that is identified as postsynaptic partners (Figures 5.4a). This is due to the nature of chemical synapses between *C. elegans* neurons. In *C. elegans* electron micrographs, chemical synapses are identified by dark electron-dense regions near the membrane in the presynaptic neuron. The electron-dense regions are referred to as the *presynaptic density*. The neurons directly apposed to the presynaptic density are scored as the postsynaptic partners. If only one neuron is directly apposed to the presynaptic density, the synapse is monadic. If more than one neuron is directly apposed to the presynaptic density, the synapse is polyadic. Roughly 60% of synapses in the NR are polyadic.

Special attention needs to be given to polyadic synapses because it is unclear

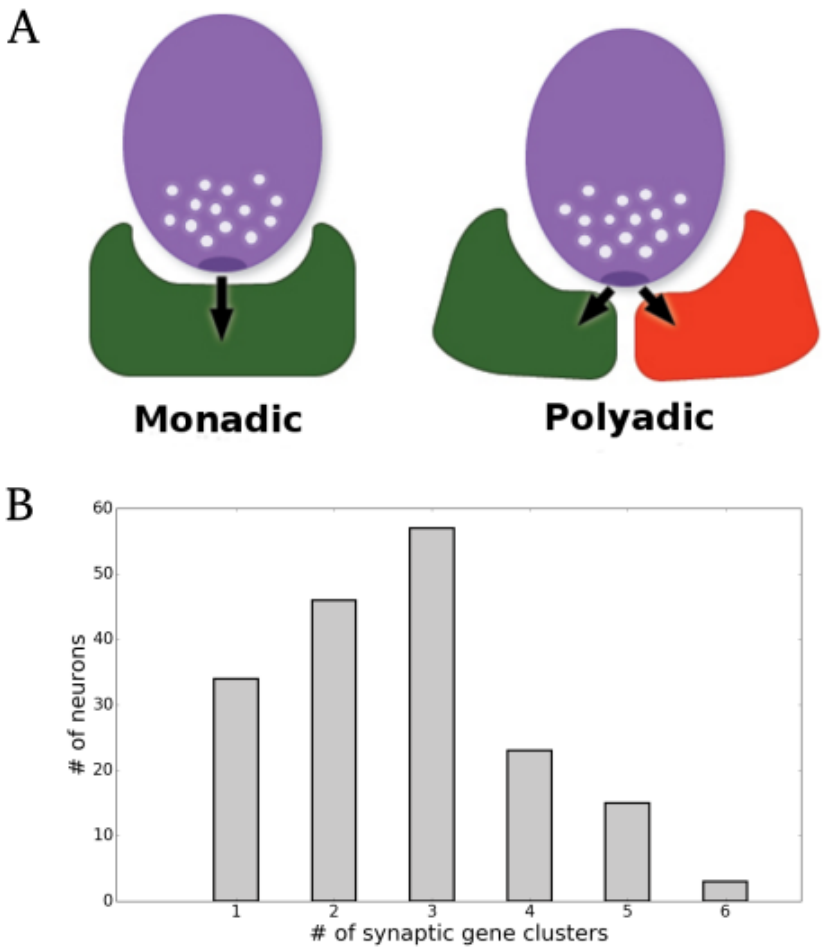


Figure 5.4: **Distinct expression clusters among postsynaptic neurons.** (A) Presynaptic neurons can have multiple postsynaptic partners. For monadic synapses, the presynaptic density is apposed to a single neuron. For polyadic synapses, the presynaptic density is apposed to multiple postsynaptic neurons. Image taken from wormatlas.org. (B)Distribution of the number of unique gene clusters. Gene clusters are from the combined expression of all postsynaptic partners at a synapse.

how polyadic synapses in the NR behave. It has been shown that all postsynaptic cells at polyadic neuromuscular junctions in the ventral cord are stimulated by the presynaptic neuron (Liu *et al.*, 2007), but similar studies have not been done for neurons in the NR. Hence, it is unclear whether one or all of the neurons apposed to the presynaptic density are actually active postsynaptic partners. However, many of the polyadic postsynaptic partners are bilaterally (left/right) and developmentally (L4/adult) conserved suggesting that they could be functional. What is particularly relevant here is that some of the neurons that have been erroneously labeled as postsynaptic may exhibit different expression patterns from the "correct" postsynpatic neurons. Therefore, by including the erroneous neurons which may have different gene expression patterns, we may overcount the number of unique gene clusters expressed by postsynaptic neurons.[1] To avoid this, the gene expression patterns for all scored postsynaptic neurons at a synapse were combined into a single expression pattern. The single gene expression pattern is the union of all the genes expressed by the postsynaptic neurons. Therefore, any differences in gene expression from a "wrong" postsynaptice neuron will be offset by the "correct" postsynaptic neurons.

The algorithm used to determine the number of unique gene clusters among postsynaptic neurons is as follows. For each neuron, the synapses are identified and the combined gene exression of the postsynaptic neurons is determined. Let L_i be the list of unique clusters for presynaptic neuron i and let $e_1, e_2, \ldots, e_m$ be the lists of genes expressed by the postsynaptic neurons at each of the m synapses for neuron i. Unique gene clusters were identified in the following way. Start by placing e_1 into L_i and give it the new label l_1. Next compare e_2 with l_1. If there is a subset of genes common to both e_2 and l_1, then replace l_1 with this common gene subset. Otherwise, place e_2 into L_i and give it the new label l_2. Next, repeat this process with e_3, comparing it to all elements in L_i. Do the same for the remaining expression patterns. After all comparisons have been made, the final list L_i will have distinct elements that are the smallest nonoverlapping gene

[1] I did also consider polyads in the previous analysis comparing gene expression patterns between postsynaptic neurons and nonsynaptic neighbors. Controlling for polyadic synapses did not have any noticeable effect on the results. Therefore, for brevity and clarity, these results were not included in the previous analysis.

clusters between the postsynaptic neurons. Note that it may be the case that a gene cluster only consists of one gene. Pseudocode for the above algorithm is given in Appendix 5.B.

The majority of neurons have multiple distinct expression clusters among their postsynaptic neurons (Figure 5.4b). Only ∼30% of neurons have a single distinct postsynaptic expression cluster while the remaining neurons mostly have 2 to 3 distinct expression clusters. These results suggest that for a given neuron multiple postsynaptic genes clusters can elicit a synapse. Such a set-up offers a number of benefits. By not rigidly enforcing a single gene cluster for each neuron, the network has more flexibility in how it is wired together. Furthermore, using mulitple gene clusters provides a mechanism for neurons to coordinate subcellular specificy of synapse formation. These points are discussed further in the next chapter.

5.A CAM gene lists

Table 5.1 CAM genes with well characterized expression in NR neurons. Pre and post columns indicate whether genes are expressed in the pre- and/or postsynaptic neuron.

CAM I genes				
Protein family	**Gene name**	**Pre**	**Post**	**Isoforms**
Ig domain	cam-1	+	+	3
	ver-3	+	+	1
	igcm-1	+	+	1
	igcm-2	+	+	1
	oig-1	+	+	1
	oig-3	+	+	1
	rig-1	+	+	2
	rig-3	+	+	1
	rig-4	+	+	1
	rig-5	+	+	7
	rig-6	+	+	4
	ncam-1	+	+	3
	sax-3	+	+	2
	sax-7	+	+	6
	lad-2	+	+	1
	syg-1	+	+	2
	syg-2	+	+	7
	unc-40	+	+	1
	unc-5	+	+	6
	zig-1	+	+	1
	zig-2	+	+	1
	zig-3	+	+	1
	zig-4	+	+	1
	zig-5	+	+	1
	zig-8	+	+	1
	madd-4	+	+	3
Ig + LRR	pxn-2	+	+	1
eLRR	slt-1	+	+	1
	tol-1	+	+	1
	dma-1	+	+	1
cadherins	cdh-3	+	+	1
	fmi-1	+	+	3
	casy-1	+	+	3
neurexin superfamily	nlr-1	+	-	1
neurexin ligands	lat-1	-	+	3
	nlg-1	-	+	9

Table 5.2 CAM genes that do not have well characterized expression in NR neurons.

CAM II genes				
Protein family	**Gene name**	**Pre**	**Post**	**Isoforms**
Ig domain	egl-15	+	+	16
	mig-6	+	+	3
	oig-2	+	+	1
	oig-4	+	+	1
	oig-5	+	+	1
	zig-6	+	+	1
	zig-7	+	+	1
	zig-10	+	+	1
	igeg-1	+	+	2
Ig + LRR	pxn-1	+	+	1
	iglr-1	+	+	1
	iglr-3	+	+	2
eLRR	fshr-1	+	+	2
	+	+	1	
	lron-9	+	+	4
	lron-14	+	+	2
cadherins	cdh-4	+	+	1
neurexin superfamily	nrx-1	+	-	13
	bam-2	+	-	1

Table 5.3 Alternatively spliced CAM genes expressed in bilaterally symmetric neurons.

Bilaterally expressed alt. spliced genes	
Neuron class	**Genes**
ADA	cam-1,sax-7
ADE	cam-1
ADF	syg-1
ADL	cam-1,syg-1
AIB	ncam-1,rig-1,rig-6
AIM	cam-1
AIN	cam-1,ncam-1,rig-1,rig-5,syg-1
AIY	cam-1,lat-1,nlg-1
AIZ	cam-1,madd-4,syg-2
ALM	cam-1,rig-6,sax-7
ALN	cam-1,syg-2
ASE	unc-5
ASG	madd-4
ASH	cam-1
ASI	cam-1,ncam-1
ASJ	ncam-1
ASK	cam-1
AUA	cam-1,rig-1,rig-5,rig-6
AVA	cam-1,rig-1,rig-6,sax-3
AVB	cam-1,ncam-1,rig-1,rig-6,sax-3
AVD	cam-1,rig-1,rig-5,sax-3
AVE	cam-1,ncam-1,rig-1,rig-6
AVH	cam-1,madd-4,rig-1,syg-1
AVJ	cam-1,rig-1
AVK	cam-1,madd-4
BDU	cam-1
FLP	cam-1
HSN	cam-1,fmi-1,nlg-1,rig-6,sax-3,syg-1
IL1	sax-7
IL2	sax-7,unc-5
OLL	casy-1,madd-4,sax-7,unc-5
OLQ	sax-3,sax-7,unc-5
PLN	syg-2
PVC	cam-1,ncam-1,rig-1,rig-6,sax-3
PVP	fmi-1
PVQ	cam-1,fmi-1,sax-3
RIA	madd-4,unc-5
RIB	rig-6
RIC	cam-1,madd-4,rig-1,rig-5,rig-6,syg-2
RIF	rig-5,rig-6,syg-1
RIG	syg-1
RIM	cam-1,rig-6,syg-1,syg-2
RIV	cam-1
RMD	cam-1,casy-1,rig-1,rig-5,rig-6,sax-3
RME	cam-1,madd-4,rig-6
RMG	cam-1,sax-3,sax-7
SAA	rig-5,rig-6,syg-1
SDQ	cam-1,fmi-1
SIA	rig-6,sax-3,syg-1
SIB	rig-6,sax-3,syg-1
SMD	casy-1,rig-1,rig-5,rig-6,sax-3
URA	nlg-1
URB	nlg-1
URX	cam-1,rig-6

5.B Algorithm to find synaptic gene clusters

Algorithm 1 Find distinct postsynaptic expression clusters

Input: A presynaptic neuron i.

Output: List L of distinct gene expression clusters.

$S \leftarrow$ synapse list of i
$L :=$ empty list
while S **do**
    $s \leftarrow$ pop synapse from S
    $e \leftarrow$ union of gene expression for postsynaptic neurons in s
    if L is empty **then** push e
    for l in L **do**
        $temp := e \cap l$
        if $temp \neq \emptyset$ **then**
            $L \leftarrow$ push $temp$
        else
            $l := temp$
return L

Chapter 6

Discussion and future directions

How neural circuits achieve wiring specificity remains an open question in developmental neurobiology. Two important aspects of wiring specificity are lamina specificity, placing synaptic partners in close spatial proximity, and synaptic specificity, choosing synaptic partners from the myriad of physically adjacent neighboring cells. The combinatorial CAM expression model has been proposed as the mechanism that neurons use to discriminate between synaptic partners and nonsynaptic neighbors. In essence, the model states that cells use combinations of cell surface proteins to uniquely label and thereby identify synaptic partners. CAM gene expression data combined with the wiring diagram of which neurons are synaptically coupled is critical for assessing the feasibility of the model, but equally important is knowing which neurons make physical contact without being synaptically coupled. The former provides a list of possible CAM expression patterns linked to synapse formation, but the latter can be used to test if expression labels of synaptic partners are indeed distinguishable from the expression labels of nonsynaptic neighbors.

This work set out to test the feasibility of the combinatorial CAM expression model in the *C. elegans* NR, the "brain" of the worm. The wiring diagram of the worm has been known for over 30 years (White *et al.*, 1986), but detailed spatial information of neural processes within the NR has unfortunately been absent until now. For the first time, this work presents a complete volumetric reconstruction of an animal's major neuropile at two different developmental stages, the larval L4 and young adult. The new volumetric data proved to be critical for rigorously

establishing and characterizing the wiring specificity in the *C. elegans* NR, which in turn provided biological insights as to what key gene expression conditions must be satisfied in order for the combinatorial CAM expression model to be feasible.

Wiring specificity specificity has never been rigorously characterized in the *C. elegans* NR. Studies from the Shen lab have provided empirical evidence demonstrating wiring specificity in the *C. elegans* ventral cord (Klassen & Shen, 2007; Shen & Bargmann, 2003; Shen *et al.*, 2004) and one of their studies did demonstrate that glial cells act as guidepost cells which are critical for synapse formation between the NR neuron AIY and RIA (Colón-Ramos *et al.*, 2007). Studies from the Hobert lab have have shed some light on the transcriptional factors regulating neural development and synapse formation (Howell *et al.*, 2015). White and colleagues generated the electron micrographs used in this study and provided ample anecdotal evidence of conserved neighborhoods in the NR (White *et al.*, 1983, 1986), illustrating that the NR exhibits some type of spatial specificity in neural process placement. But at what level is synaptic wiring specified in the NR?

There are at least two models that could support the reproducible synaptic connectivity observed in the NR. In the first model, neural process placement is strictly controlled such that neurons only make physical contact with potential synaptic partners. Because process placement is strictly controlled, neurons can randomly synapse onto their neighboring neurons and the "correct" synapses will be made with high probability. In the second model, neural process placement is slightly more relaxed which requires neurons to be more discriminative when identifying synaptic partners. In this model, neurons randomly making synaptic connections would lead to a lower probability of the "correct" synapses being made. In order to make the "correct" synapses, neurons must somehow distinguish synaptic partners from neighboring cells. Thus, neurons must exercise synaptic specificity.

This analysis shows that the second model is correct. Process placement within the NR does appear to be regulated, forming a lamina-like structure that is both bilaterally and developmentally conserved. However, process placement is not so tightly regulated that it could account for the number of bilaterally

and developmentally conserved synaptic connections in the NR. Indeed, 86% of NR neurons exhibit either pre- or postsynaptic connection that are statistically unlikely to be random. Thus, in addition to lamina-like specificity, NR neurons also exhibit synaptic specificity.

Given that there is synaptic specificity, what variables affect synaptic specificity? This study focused on two possible variables: the amount of physical contact between neurons and the independent connectivity characteristics of the pre- and postsynaptic neuron.

If we assume a model where synapse probability depends on the amount of physical contact between neurons, then synapse specificity could be determined by the amount of mutual contact. Synaptic partners would be expected to have more physical contact than nonsynaptic neighbors. Synapses could still be formed probabilistically, but synapses would be biased toward neurons pairs that make more physical contact. This study shows that this model can be rejected. There is little correlation between synapse formation and the amount of physical contact between neurons. Furthermore, a model where synapse probability only depends on the amount of physical contact between neurons cannot capture the variability in synaptic connectivity across neurons.

However, a model where synapse probability depends on the cell-autonomous connectivity properties of the pre- and postsynaptic neuron does capture the variation in synaptic connectivity. In this model, the presynaptic neuron makes a synapse with some fraction of its neighbors (C^{pre}), the postsynaptic neuron receives a synapse from some fraction of its neighbors (C^{post}) and the synapse probability depends on the product of both. The key assumption of this model is that the pre and post connectivity fractions are independent, suggesting that synapse probability is largely determined by the cell-autonomous properties of the pre- and postsynaptic neuron. This is in agreement with the combinatorial CAM expression model.

Using CAM expression data obtained from Wormbase, this study shows that combinatorial CAM expression could support synaptic specificity provided that alternatively spliced CAM genes are differentially expressed among neurons. The combinatorial expression of just 8 surface molecules is theoretically sufficient to generate unique labels for the 180 NR neurons. However, in practice this is not

the case. Expression of CAM genes is sparse and overlapping, but many neighboring neurons share the same expression labels. Thus, using only the known CAM expression, there would be no way for neurons to reliably distinguish between synaptic partners and nonsynaptic neighbors. However, stochastically expressing isoforms of known alternatively spliced genes substantially increases the number of unique expression labels. Stochastic isoform expression has been observed in other organisms and has been proposed as a mechanism for uniquely labeling cells (Zipursky & Sanes, 2010). Typically, transcription reporter studies have shown that left/right homologous neurons express the same CAM genes. However, we observe that some neurons will make physical contact with both the left and right homologous neurons while only synapsing onto either the left or right neuron. This cannot be explained by the combinatorial CAM expression model if alternative splicing is not considered. However, if alternative splicing is stochastically assigned in the model, many of the left/right homologous are assigned unique expression labels. Moreover, when alternative splicing is included in the model, 90% of synaptic partners are distinguishable from nonsynaptic neighbors.

Finally, this study shows that multiple gene clusters elicit synapses in a presynaptic neuron. If the combinatorial CAM expression model is true, then there are two possible variations. In the first variation, for a given neuron only one gene combination will elicit a synapse. I refer to this as the *single key* model. In the second variation, for a given neuron multiple gene combinations can elicit a synapse. I refer to this as the *key ring* model. If the single key model is true, then all of the post synaptic neurons for a given presynaptic neuron should share some common gene expression pattern. This is rarely the case, indicating that the key ring model is correct.

There are a number of interesting implications for the key ring model. First, utilizing multiple gene combination 'keys' for eliciting a synapse affords some level of robustness during synaptic wiring of a neural circuit. If only a single gene combination key were used then every postsynaptic neuron would have to have the correct key for each presynaptic neuron from which it receives a synapse. If at any point the neuron lost the key expression, then that synaptic connection would be lost. Allowing for multiple gene combinations in the key ring model means that if one key is 'lost' then the neuron could fall back on another key to

elicit the synapse. Second, the key ring model offers a mechanism for controlling synaptic connectivity with subcellular specificity. The ordering of synapses along neural processes in the NR appears to be mostly conserved between left/right homologous neurons and between equivalent L4/adult neurons (White *et al.*, 1986). This suggests that synapses are made with subcellular specificity. This could be explained by the key ring model where different molecular 'locks' are distributed along the length of the neural process. Then at each lock only postsynaptic neurons with the appropriate molecular key can elicit a synapse. In this way, the presynaptic neuron controls where along its process the specific postsynaptic neurons can elicit synapses.

A very targeted prediction of this study is that bilaterally symmetric neurons need to differentially express alternatively spliced genes. A good test case for this would be the AIB interneurons whose left and right processes extend around the NR, making them one of the few aphid neurons which innervate both the right and left side of the NR. As a result, neurons AIBL and AIBR make physical contact with many pairs of bilaterally symmetric neurons. For example, AIBL and AIBR make physical contact with both the nociceptive sensory neurons ASHL and ASHR. However, AIBL is only postsynaptic to ASHL and AIBR is only postsynaptic to ASHR. Thus, ASHL and ASHR are able to differentiate between AIBL and AIBR, despite both AIB neurons exhibiting equivalent CAM expression labels. How is this possible? If the combinatorial expression hypothesis is true, then according to the results of this analysis AIBL and AIBR must differentially express at least one of their shared alternatively spliced genes (*ncam-1*,*rig-1*,*rig-6* in Table 5.3). This could be confirmed using fluorescence splicing reporters which enable visualization of alternative splicing patterns in cells (Wani & Kuroyanagi, 2017). Under an AIB specific promoter (e.g. *odr-2*b), a genomic fragment of interest is cloned upstream of the fluorescent protein marker. The genomic fragment would be cloned from one of the alternatively spliced genes and artificial termination codons and/or frameshifts would be strategically introduced into the exons at splice points. Expression of the marker would indicate a specific splicing pattern. Expression of distinct splice patterns in AIBL and AIBR would provide supporting evidence for the combinatorial expression model.

This study is obviously not without its limitations. The most frequently cited limitation is the low sample size. The connectivity and volumetric data were both taken from just two worms, each at a different developmental stage. Low sample sizes raise legitimate concerns as to the robustness of the results. I have tried mitigate these concerns by framing questions in a way that allows for maximal information to be derived from the data. I tried to avoid questions concerning specific neurons (e.g. ASH, AVA, etc.) which can typically only be addressed with 2-4 data points in this data set. I also tried to avoid questions concerning the overall NR network because there are only two networks to compare. Instead, I asked more basic questions relating to the general nature of neurons and synapses where the neurons could be treated more abstractly. For example, I would not ask "What is the connectivity fraction of ASH?" Instead, I asked "What is the distribution of connectivity fractions of NR neurons?" While the first question limits me to 4 data points to work with (ASHL and ASHR in the adult and L4), the second question gives me 180 NR neurons from which to derive a distribution. Also, instead of simply generating statistics about NR connectivity, I tried to take a model driven approach. Statistically characterizing NR connectivity would requiring analyzing many worms, while the model approach allowed me to use data from a few worms to test specific hypotheses. For example, I assumed models where physical contacts are made at random, synaptic contacts are made a random and synaptic connectivity depends strongly on touch density. I then asked if these models are supported by the data and they were not. This allowed me to conclude that the NR exhibits both lamina-like and synaptic specificity.

A second legitimate concern is noisy and incomplete data. CAM expression was determined from data curated from Wormbase. I had to remove 17 CAM II genes from the analysis because the genes could not be linked to specific neurons. There may be additional CAM genes with NR expression that have not been reported by Wormbase or that have simply not yet been discovered. However, additional genes are unlikely to change the major conclusions of this study. While additional CAM genes would decrease the overall need for isoform expression, it would not provide the needed isoform expression to distinguish left/right homologous neurons. Expression studies are typically done with transcription reporters and aside from a few notable exceptions (expression of *gcy-7* in ASEL and *gcy-5*

in ASER, for example) left/right homologous neurons typically express the same promoters. Therefore, unless it was found that the additional CAM genes were strictly expressed in either the right or left neuron, the additional genes would not be able to convey left/right identification without considering isoform expression. Finally, the key ring model was inferred from the observation that postsynaptic neurons have distinct gene expression labels. Additional CAM genes would not change this fact.

Ultimately, this study shows that synaptic wiring likely makes use of a complex combinatorial gene expression code. To validate or disprove this model will likely require more sophisticated and perhaps more creative molecular approaches. Simple gene knockouts may be used to link genes to specific synapses, but to unravel the molecular code will require altering genes in well thought out combinations. The problem becomes even more challenging when isoforms of alternatively spliced genes are considered. One positive takeaway from this study is that *C. elegans* exhibits many of the wiring specificity characteristics found in other organisms, e.g. lamina-like specificity, synaptic specificity and alternative splicing. This simple model organism could potentially be a powerful tool for unraveling the complex mechanisms of wiring specificity. Thus, it would appear that 30 years on, this humble organism still has something to teach us.

Part II

A rotatable microfluidic device for simultaneous calcium imaging of bilateral chemosensory neurons

Chapter 7

Introduction

Calcium imaging has become a widely used tool for studying the cellular activity of neurons and muscles. However, imaging multiple cellular structures within a small animal is challenging because different cellular structures within an organism are generally not visible within a single focal plane. Thus, imaging multiple cellular structures requires either optical sectioning techniques using expensive confocal imaging systems (Nguyen *et al.*, 2016; Schrödel *et al.*, 2013) or the capability to orient the animal such that the desired cellular structures lie within the same focal plane. To our knowledge, only two microchip devices have demonstrated the capability to rotate small animals (specifically *C. elegans*). The first device uses a glass capillary tube to pneumatically grab and rotate the worm within a microchannel (Ardeshiri *et al.*, 2016), while the second device rotates worms using microvorticies controlled by an acoustic field (Ahmed *et al.*, 2016). While both devices may be adequate for imaging the structural properties of cells, the invasive methods used to rotate the worm may interfere with measurements of neural activity. For example, a number of *C. elegans*' sensory neurons respond to both light (Hart *et al.*, 1999) and harsh touch(Li *et al.*, 2011) and sufficiently high ultra-sound pulses can effect *C. elegans* behavior (Ibsen *et al.*, 2015). Furthermore, these devices have not yet implemented methods of controlled stimulus delivery needed to measure neuronal responses (Chronis *et al.*, 2007a).

The nematode *C. elegans* is a popular model organism because of its well characterized nervous system (White *et al.*, 1986) and its amenability to genetic study (de Bono & Maricq, 2005). The hermaphrodite *C. elegans* has 302 neurons

that have been assigned to 118 distinct classes according to their topology and synaptic connections. Of these, 98 neuron classes consist of left/right bilaterally symmetric neuron pairs which, except for neurons ASE and AWC, have seemingly identical morphology, function and gene expression. Because of how the worm naturally orients itself, the left/right neuron pairs cannot be viewed simultaneously because they lie on different focal planes. Typically, with the above exceptions, when measuring neural activity using fluorescence imaging, only one neuron is measured (either left or right) and the other neuron is assumed to behave similarly. However, this has never been rigorously validated by simultaneously viewing left/right neuron pair activity in response to the same stimulus. Here, we present a novel device for simultaneously imaging neural activity in left/right bilateral neurons.

7.1 Microfluidics as a tool for biological study

Microfluidic systems are used to manipulate minute fluid volumes using channels with dimension tens of micrometers wide. At this micron-scale, fluid flow is easily controlled due to the low Reynolds number (Whitesides, 2006). This ease of fluid flow offers several advantages for experiments where accurate flow rates, concentration gradients and shear rates are required. Furthermore, microfluidic devices are typically made using soft lithography replica molding of a flexible silicone elastomere, PDMS (polydimethylsiloxane) (Duffy *et al.*, 1998). PDMS is compatible with aqueous solutions, non-toxic, gas permeable and optically transparent making it suitable for biological experiments that require microscopic observation. Microfluidics has been used in studies of cells (Andersson & van den Berg, 2003), *Drosophilla* (Lucchetta *et al.*, 2005) and *C. elegans*.

Microfluidics is having a growing impact on *C. elegans* research. Microfluidic devices provide a platform for easily moving and manipulating the worm (Hulme *et al.*, 2010; Stirman *et al.*, 2010). The devices provide a tool for immobilizing the worm during imaging studies without the need for glues or anesthetics (Chronis *et al.*, 2007b; Hulme *et al.*, 2007; Lockery, 2007). The devices allow for more sophisticated behavioral studies (Zhang *et al.*, 2005; Zimmer *et al.*, 2009) and longitudinal studies (Hulme *et al.*, 2010; Pincus *et al.*, 2011). High-throughput

imaging (Krajniak & Lu, 2010; Lockery *et al.*, 2012; Stirman *et al.*, 2010), phenotype screening (Chung *et al.*, 2008; Crane *et al.*, 2009; Lee *et al.*, 2013), and laser ablation and nerve regeneration (Allen *et al.*, 2008; Ben-Yakar *et al.*, 2009) studies have also been performed using microfluidic devices. In short, microfluidics is making it easier to perform more rapid and more precise experiments on the worm.

7.2 The ASH neurons mediate nociceptive

C. elegans contains a number of nociceptive neurons that respond to both mechanical and chemical stimuli (Tobin & Bargmann, 2004). Of these neurons, the bilateral pair of ASH neurons have probably been the most exhaustively studied. The ASH neurons are associated with the amphid sensory organ. Amphid sensory neurons have cilia that are directly exposed to the environment through small pores, enabling them to detect chemical and osmotic repellents (Perkins *et al.*, 1986). ASH has been shown to detect aversive nose touch (Kaplan & Horvitz, 1993) and a number of other aversive cues such as: high osmotic strength (Troemel *et al.*, 1995), SDS (Hilliard *et al.*, 2002), and heavy metals (Sambongi *et al.*, 1999, 2000). Because of their broad range sensory functionality, ASH neurons are considered to be polymodal.

Response to nose touch and hyperosmotic solutions is believed to be segregated by different levels of glutamate release from ASH. This is supported by analysis of glutamate receptors glr-1, glr-2, and nmr-1 which are present in the command interneurons AVA and AVB. Mutants deficient in glr-1 and glr-2 exhibit no response to nose touch and only partial osmotic avoidance (Maricq *et al.*, 1995; Mellem *et al.*, 2002). Mutants deficient in nmr-1 exhibit no compromised nose touch response and full osmotic avoidance. Combined nmr-1,glr-1,glr-2 mutants exhibit a larger defect in osmotic avoidance than either nmr-1 or glr-1,glr-2 mutants (Mellem *et al.*, 2002). This suggests a model where ASH directs avoidance via (excitatory) glutamatergic synapses onto forward (AVB) and backward (AVA) command interneurons.

Interestingly, many aversive responses mediated by ASH also require other amphid sensory neurons. The ADL neurons are required for avoidance of high

osmolarity (Bargmann *et al.*, 1990), octanol (Troemel *et al.*, 1995), copper and cadmium (Sambongi *et al.*, 1999). The neurons ASK and ASE are required for avoidance of water-soluble repellents (Sambongi *et al.*, 1999, 2000). Clearly, in these instances, sensory inputs from ASH are being combined with input from other amhpid sensory neurons in order to elicit the aversive response (Hilliard *et al.*, 2002). Furthermore, ASH response to noxious stimuli can vary depending on the condition of the worm. For example, octanol avoidance is almost entirely mediated by ASH in well-fed animals, but after an hour of starvation, octanol avoidance is distributed between ASH, ADL and AWB (Chao *et al.*, 2004). This switch between the two circuits appears to be governed by a combination of neuromodulators: serotonin, dopamine, tyramine, octopamine and numerous neuropeptides (Komuniecki *et al.*, 2012; Mills *et al.*, 2012; Wragg *et al.*, 2007). Thus, it is possible for one behavior to be governed by multiple sub-circuits.

7.3 ASE neurons are the primary NaCl chemosensors

Chemoattraction to salts is distributed among four pairs of amphid sensory neurons (ADF, ASE, ASG, ASI, ASK and ASJ), of which ASE is the most important (Bargmann & Horvitz, 1991). In addition, ASE senses other water soluble attractants, including anions, cations, cAMP, biotin and lysine (Bargmann & Horvitz, 1991). ASE also plays minor roles in avoidance of Cd^{2+} and Cu^{2+}, which is only evident when the sensory neurons ASH are missing (Sambongi *et al.*, 1999). ASE is also a primary CO_2 sensor (Bretscher *et al.*, 2011).

Despite being morphologically similar, the neurons ASE left (ASEL) and ASE right (ASER) exhibit different gene expressions, developmental programs and physiological properties (Chang *et al.*, 2003; Pierce-Shimomura *et al.*, 2001; Suzuki *et al.*, 2008). The principal postsynaptic partners of ASEL/R are AIY and AIB, which in turn interconnect with each other and AIZ (White *et al.*, 1986). The sensory neurons ASE respond to step changes in NaCl. Calcium imaging experiments show that ASEL responds with transient depolarization to

NaCl upsteps while ASER responds with transient depolarization to NaCl downsteps(Suzuki *et al.*, 2008). This suggests that ASEL and ASER behave like ON and OFF cells, respectively. The ASE neurons express different ion sensitivities, where ASEL is primarily sensitive to Na^+ ions and ASER is primarily sensitive to Cl^- ions (Pierce-Shimomura *et al.*, 2001; Suzuki *et al.*, 2008). Furthermore, activation of ASEL with NaCl upsteps promotes bouts of forward locomotion and activation of ASER with NaCl downsteps promotes turns (Suzuki *et al.*, 2008). Ablating ASER and applying a NaCl upstep yields a deficit in forward locomotion, suggesting that the positive contribution of ASER to forward locomotion is the result is desuppression of forward probability (Suzuki *et al.*, 2008).

7.4 Contributions of this work

This work presents a rotatable microfluidic chip that can be used to image neural activity in *C. elegans*' left/right bilateral neurons. The chip is minimally invasive and implements current established methods of olfactory stimulus delivery. Rather than rotate the worm within the microchannel, the microchip itself is physically rotated. Thus, left/right bilateral neurons can be simultaneously imaged without manipulating the worm and without using expensive optics. To demonstrate the functionality of the chip, we imaged the calcium activity of ASHL/R neurons which respond synchronously to noxious stimuli. We find that ASHL and ASHR can exhibit uncorrelated rise times in their transient response and uncorrelated response probabilities to mid-range NaCl stimulus. The ASEL/R neurons, which are known to respond asymmetrically to NaCl, can also exhibit independent response probabilities within the same worm. We believe that this new device will provide a useful tool to explore how organisms coordinate bilateral neural activity.

7.5 Acknowledgments

This work was done in collaboration with Netta Cohen (NC), Jung-uk Shim (JS), Jinyang Chung (JC) and myself (CB). NC and CB conceived the project. JS and JC designed and fabricated the microfluidic device. CB procured, engineered and

maintained worm strains. CB engineered the fluid control system and the image acquisition and image analysis software. JC collected the calcium imaging data. JC and CB analyzed the data.

Chapter 8

Methods and Results

8.1 Fabrication of a rotatable chip

The cell bodies of left/right symmetric neuron pairs are generally located on the same coronal plane but lie in different sagittal planes. Because the worm's body muscles are located along its ventral and dorsal sides, the worm naturally orients itself so that it lies on either its left or right side. This places the sagittal planes parallel to the focal plane, meaning left/right neuron pairs cannot be viewed simultaneously. A 90° rotation about the anterior-posterior axis roughly places the coronal plane parallel to the focal plane, making simultaneous viewing possible. Therefore, a simple solution to the left/right imaging problem is to load the worm into the microfluidic chip, let the worm naturally orient itself and then rotate the chip by 90°.

Our rotatable chip design is a modified version of the so-called olfactory chip (Chronis *et al.*, 2007a) which is designed to examine the neural activity of chemosensory neurons in response to stimuli. The olfactory chip traps the worms in a narrow channel in such a way that the nose protrudes into a perpendicular channel where chemical stimuli flows across the nose of the worm (Figure 8.1b). The olfactory chip implements a four-channel system where the flow of either a stimulus (channel 2) or a control buffer (channel 3) is directed across the nose of the worm and switching between the stimulus and buffer is controlled by an external three-way valve that directs flow to one of two side channels (channels 4-1 and 4-2) (Figure 8.1a). When flow is directed to side channel 1, buffer

is directed towards the worm. When flow is directed to side channel 4, stimulus is directed towards the worm (Figure 8.1c). The chip is bonded to a coverslip through which the worm is imaged and the worm trap is located between two outlet channels, both of which limit how the chip can be oriented.

To make the olfactory chip rotatable, we made the following modifications. We moved the outlet channel from behind the worm trap to behind the the four-channel system (Figure 8.1a). This allowed the worm trap to be moved closer to the side of the chip. We cut the side of the chip so that the distance between the side of the chip and the worm trap is less that 1mm. Instead of bonding the chip to a cover slip, we bonded the chip to a thin layer of PDMS which could be cut to the dimensions of the chip (Figure 8.1d). This resulted in a self-contained cuboid chip that could easily be rotated by 90° (Figure 8.1e). The advantage of our design is that neurons can either be viewed in the sagittal plane or the coronal planes, which we refer to as the top- and side-views, respectively. Figure 8.1f shows the same worm in top-view where only one neuron is visible and from from the side-view after a 90° rotation of the chip where now both the left and right neurons are visible. In general, we found that when only the left or right neuron is visible in top-view, then both the left and right neurons are visible from side-view, which demonstrates that the 90° rotation is sufficient to place left/right neuron pairs in the same focal plane.

8.2 Fluid control system

The fluid control system is a low-cost version of the design presented in Chronis *et al.* (2007a). Briefly, a compressed air supply ($\sim$ 10 psi) is coupled to a three-way air manifold block which serves as the pneumatic inlet for the microfluidic system. The three outlets of the manifold are each coupled to a 15 ml falcon tube reservoir. Pneumatic inlets and fluid outlets are created for the reservoirs by piercing the falcon tube lid with two 19 gauge needles and then fixing the needles in place with epoxy. One buffer filled tube is coupled directly to channel 3 and the stimulus filled tube is coupled to channel 2 on the chip. The second buffer tube is coupled to a 3-way solenoid valve (Lee company/LFAA2403410H) that is controlled by a custom circuit (see below) which in turn is coupled to channels

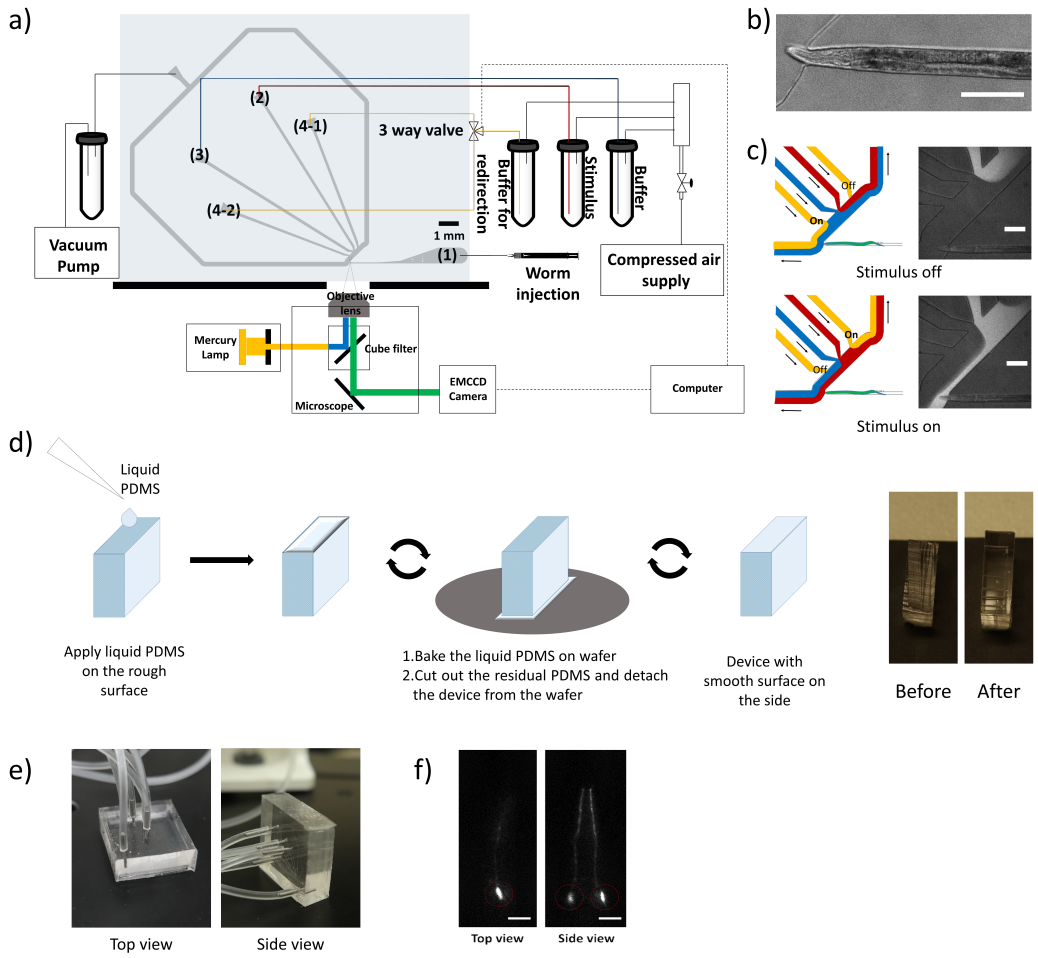
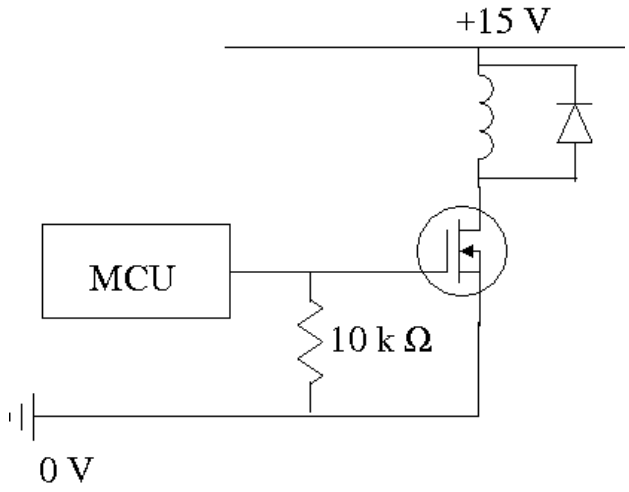


Figure 8.1: **Rotatable microfluidic chip.** Image created by Jinyang Chung. Used with permission. (a) Fluid control and imaging setup for the rotatable chip. (b) Image of worm loaded into the trap. (c) Switching between buffer (blue) and stimulus (red) via flow redirection of the control channels (yellow). (d) Bonding of the chip to PDMS to create a rotatable cuboid shape. (e) The worm can either be imaged from the top-view or the rotated side-view. (f) Image of the same worm in top- and side-view.

4-1 and 4-2 on the chip. Falcon tubes were coupled to the chip using polyethylene tubing which was inserted into the chip channels using custom metal tubing. The chip outlet was coupled to a waste reservoir which in turn was connected to a vacuum pump ($\sim$ 5 psi).

Solenoid valves were driven by a MOSFET transistor whose gate is controlled by a microcontroller unit (MCU). Serial communication over USB from the PC operates the MCU (Arduino board). The MCU outputs to the MOSFET gate which switches the transistor and completes the circuit for the solenoid valve (Figure 8.2, Table 8.1). A 10 kΩ pull-down resistor ensures that the gate is held low when the MCU does not send a high signal. A flyback diode protects the transistor from the reverse voltage-spike that results from the abrupt change in the magnetic field across the coils of the solenoid valve when power to the coils is turned off.



(a)

Figure 8.2: **Robotic solenoid valve switch.** The MCU modulates the gate of a MOSFET transistor which connects the circuit for the solenoid valve. The 10 kΩ pull-down resistor keeps the gate low when the MCU does not output the 3.3 V high signal. The flyback diode protects the transistor from voltage spikes that result when power to the solenoid valve is turned off.

Table 8.1 Components used in robotic switch.

Electronic Components	
Component	**Type/Model**
MCU	Genuino 101 Intel® Curie™ Microcontroller
Flyback diode	1N4002S 1A Silicon Rectifier
MOSFET transistor	2N7000G small signal MOSFET
3/2 Solenoid valve	Lee Co. part#: LFAA1201610H

8.3 Image acquisition

For each experiment, a single worm was manually loaded into channel 1 of the microfluidic chip. Worms were loaded in CTX buffer (4 mM KH_2PO_4/K_2HPO_4 pH 6, 1 mM $CaCl_2$ and 1 mM $MgSO_4$) using a 1 mL syringe by eluting a single worm in a 5 mL droplet of CTX buffer, drawing the worm into the syringe and then injecting the worm into the chip. Once injected, the worm is slowly pushed into the trap by controlling the plunger of the syringe (Figure 8.1b). After the worm is in the trap, the worm is allowed to settle for 6 minutes before commencing imaging. Prior to imaging, the chip was rotated 90° to place the left and right neurons in the same focal plane (Figure 8.1c). During imaging, worms were subjected to a salt solution pulse (CTX plus NaCl at the desired concentration) and calcium responses were recorded.

Imaging was performed on a fluorescence microscope (BIM800F, Bioimager) equipped with a 40x long distance working lens (BIM5-40xF, numerical aperture = 0.6, working distance = 2.2 mm, Bioimager) and mercury lamp (Bum-HBOB, 100W, Bioimager), 10% Neutral density filter (Model, company) and FITC filter (ex: 480 15 nm / em: 535 20 nm, Chroma Technology Corporation). Images were captured with EMCCD camera (Ixion, Andor Technology Ltd) with an exposure time of 31 - 100 ms. Image acquisition and stimulus delivery was coordinated with a custom MicroManager plugin (Edelstein *et al.*, 2015).

8.4 Adaptive neuron tracking

While the worm body is mostly fixed in the microfluidic chip, the worm can still exert small translational movements within the trap making long term calcium imaging challenging. If the worm remained perfectly fixed for the duration of the experiment, then a fixed ROI around the soma could be defined for fluorescence analysis. However, head movements by the worm and settle changes of pressure in the chip will cause translational movements of the neurons. Manually tracking the neurons through each frame would be laborious and time consuming, therefore automated visual tracking methods are preferred. To this end, we implemented a method of adaptive cross-correlation filters know as minimum output sum of squared errors (MOSSE) (Bolme *et al.*, 2010).

The cross-correlation is a measure of the similarity of two signals. Suppose we have two signals f and h. The cross-correlation of these two signals is defined as

$$(h \star f)[n] = \sum_{-\infty}^{\infty} h^*[m] f[m+n].$$

If f and h have a "similar" shape, then the cross-correlation is maximal when the two functions are aligned. Now let f be the part of the image with the soma. For tracking purposes, we want to pick a filter h that maximizes the cross-correlation response g, which we assume to be Gaussian. In other words, we need to compute h which solves the equation

$$g = h \star f.$$

Computational speed is enhanced by computing the cross-correlation in the Fourier domain with the Fast Fourier Transform (FFT). The 2D transforms of the input image $F = \mathcal{F}(f)$ and filter $H = \mathcal{F}(h)$ are computed. The Convolution Theorem states that the Fourier transform of a convolution is equal to the product of the pointwise product of the Fourier transforms. Hence, the Fourier transform of the response $G = \mathcal{F}(g)$ is

$$G = H^* \odot F \tag{8.1}$$

where $\odot$ is the Hadamard product denoting element-wise multiplication. The response function g is obtained from G by taking the inverse FFT. Hence, for an

image with p pixels, the upper bound of the computational time is determined by the cost of computing the forward and inverse FFT which is $O(p \log p)$.

The MOSSE algorithm produces an optimal filter from a set of training images. The algorithm starts with an initial set of training images f_i with corresponding outputs g_i. In practice, the initial training images are generated by taking the first frame and applying small random affine transformations. The outputs are assumed to be a 2D Gaussian shape with a peak centered on the target of the tracking frame f_i. Training is done in the Fourier domain. Rearranging terms in Equation 8.1 shows that the filter is computed as

$$H_i^* = \frac{G_i}{G_i} \tag{8.2}$$

where the division is performed element-wise. The optimal filter that maps the training images to the desired training outputs is the filter H the minimizes the difference between the actual and desired output of the convolution. This sets up the following minimization problem

$$\min_{H^*} \sum_i |H^* \odot F_i - G_i|^2. \tag{8.3}$$

Solving the optimization problem is not difficult because the computation is elementwise in the frequency domain. That is, each element of H (indexed by ω and ν) can be written in terms of its real and imaginary component, i.e. $H_{\omega,\nu}$ and $H^*_{\omega,\nu}$, respectively. We differentiate each element of Equation 8.3 with respect to each $H^*_{\omega\nu}$ and set equal to 0,

$$0 = \frac{\partial}{\partial H^*_{\omega\nu}} \sum_i |H^* \odot F_i - G_i|^2. \tag{8.4}$$

It is straightforward to solve Equation 8.4 to get a closed form expression of the MOSSE filter:

$$H^* = \frac{\sum_i G_i \odot F_i^*}{\sum_i F_i \odot F_i^*}. \tag{8.5}$$

This can be interpreted as the correlation between the input image and desired output divided by the energy spectrum of the input image.

We find that the MOSSE algorithm provides excellent performance for tracking neurons in our imaging data. The tracker is able to track neurons of varying

degrees of expression levels in real time. Moreover, the MOSSE algorithm can be used to independently track multiple neurons in the same image.

Once the neurons have been tracked it is straightforward to compute the fluorescence, because within the tracking window the neurons remain fixed. The tracking window is extracted from the movie and the neurons can be segmented within the tracking window. We found it easiest to manually segment the neurons using predefined geometries such as rectangles or circles that encapsulated the region of interest (ROI) around the neuron or soma, respectively. The background was then defined as all pixels within the tracking window not in the ROI. The fluorescence F_m of the ROI was taken to be the mean pixel value within the ROI. Similarly, the background fluorescence F_b was defined as the mean pixel value within the background. The fluorescence was computed as $F = F_m - F_b$. The baseline fluorescence F_0 was defined as the average change in fluorescence during 3 seconds prior to stimulus delivery. The change in fluorescence was computed as $\Delta F = F - F_0$. The calcium activity of the neuron was computed as the normalized relative change in fluorescence, $\Delta F/F_0$.

8.5 ASHL and ASHR can exhibit uncorrelated responses

The ASH left and right neurons (ASHL and ASHR, respectively) are considered to be the major nociceptive sensory neurons in *C. elegans* (Tobin & Bargmann, 2004); responding to mechanical, osmotic and chemical stimuli (Kaplan & Horvitz, 1993). Previously, ASH activity has been measured in response to these stimuli using the GCaMP calcium indicator (Chronis *et al.*, 2007a; Larsch *et al.*, 2013). However, it was assumed that ASHL and ASHR responded synchronously while only imaging ASHL or ASHR. To confirm that ASHL and ASHR do indeed respond synchronously, we simultaneously imaged ASHL and ASHR in response to a range of NaCl concentrations.

ASH responds to high concentrations of NaCl, possibly due to the high osmolarity, but not to low concentrations of NaCl, which is known to be an attractant

for the worm (Bargmann *et al.*, 1990). We tested if ASHL and ASHR exhibit synchronous calcium responses within the same worm in response to 3 second pulses of low (100 mM), medium (200-300mM) and high (500mM) concentrations of NaCl using the calcium indicator GCaMP. The GCaMP was expressed under the ASH specific promoter *sra-6* (strain CX10979, a kind gift from the Bargmann lab). We hypothesized that ASHL and ASHR would give synchronous responses to low and high NaCl concentrations, but left open the possibility of a middle range where the left/right responses could be less synchronized.

To test ASHL and ASHR responses, worms were imaged in side-view in the rotatable chip. ASH neurons respond to blue light, which is required to stimulate the GCaMP. In order to inure ASH neurons to blue light stimulation, worms were preexposed to blue light for 60 seconds immediately before stimulus delivery. Worms were then given a 3 second NaCl pulse and calcium responses were recorded.

As expected, neither ASHL nor ASHR responded to 100 mM in the 8 animals tested (Figure 8.3a). Conversely, ASHL and ASHR always responded to 500 mM in the 16 animals tested. For 300 mM, both ASHL and ASHR responded in 4 of 6 animals and neither ASHL nor ASHR responded in the remaining 2 animals. Interestingly, for 200 mM, we observed animals where either ASHL or ASHR responded (2 of 12 animals, Figure 8.3b). For the remaining 10 animals, either both ASH neurons responded (5 of 10) or none of the ASH neurons responded (5 of 10).

We next computed the correlation between the ASHL and ASHR calcium responses. Analyzing the correlations between left and right neurons does require some care. We could take a statistical approach and compute a correlation coefficient (e.g. a Pearson correlation), but this assumes that the data is independently drawn from some distribution. However, our data points are clearly not independent because it is a time series, where the data point at time t_i is somehow linked to data points at t_{i-1} and t_{i+1}. In signal processing, the standard practice for measuring the similarity of two series is to take the cross-correlation. However, cross-correlations are more appropriate for analyzing patterns in time series where there is some repeated stimuli or noise rather than transient signals in response to a single stimulus pulse, as in the case of our data. Therefore,

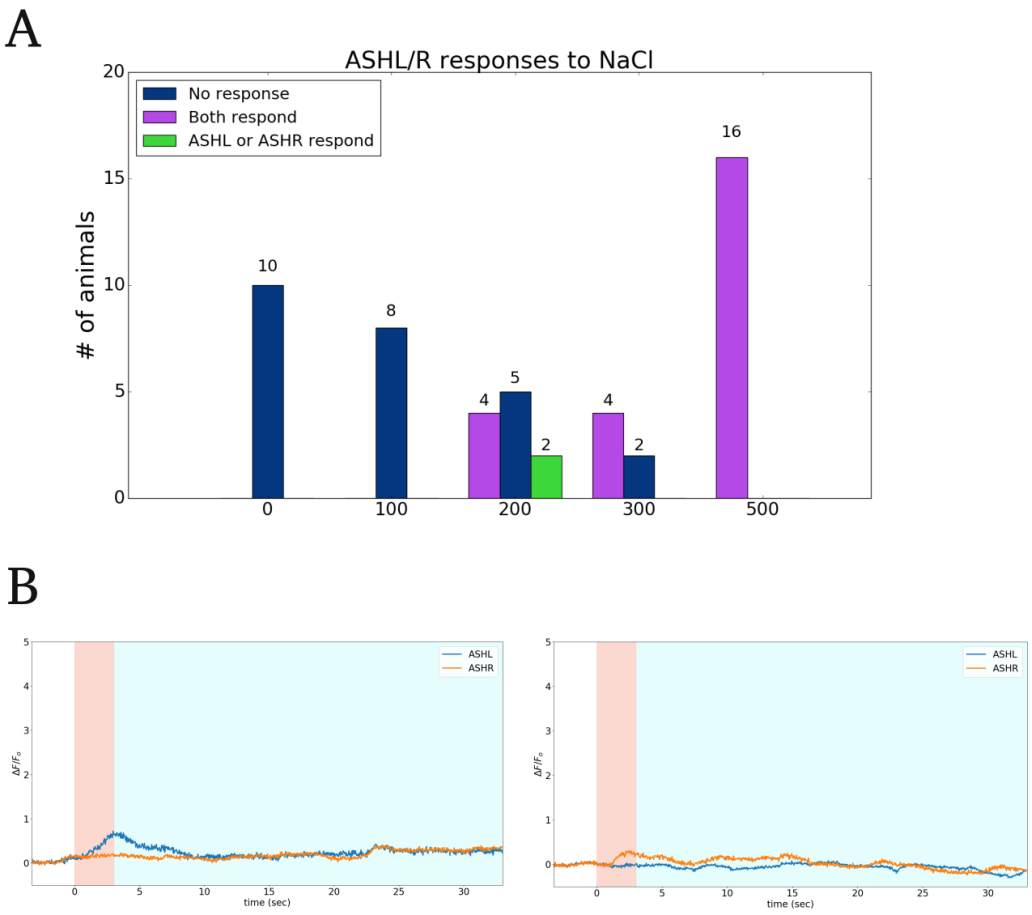


Figure 8.3: **ASH NaCl responses.** (A) Breakdown of ASHL and ASHR responses for 0, 100, 200, 300 and 500 mM NaCl. There were three cases: neither ASHL nor ASHR responded (blue), both ASHL and ASHR responded (purple) and either ASHL or ASHR responded (green). (B) Individual traces where ASHL exhibits a response but ASHR does not (left) and where ASHR exhibits a response but ASHL does not (right). White background shows the basal line activity before 3 sec 100 mM NaCl pulse. Red background indicates the duration of the pulse.

we need to construct a more appropriate metric for measuring correlation in our data. Fortunately, we can take advantage of the structure exhibited in the ASH transient signals.

We compared the rise and fall time of ASHL and ASHR neurons in response to 500 mM NaCl. Both ASHL and ASHR consistently respond to 500 mM, making it the ideal condition for comparing the signals of the two neurons. Both ASH neurons exhibit increasing calcium activity during the 3 second NaCl pulse (Figure 8.4a). The calcium signals peak shortly after the end of the NaCl pulse, at which point the calcium activity slowly returns to baseline levels over the course of tens of seconds. Hence, we defined the rise time as the portion of the signal during the 3 second pulse and the fall time as the portion of the signal during the 30 seconds after the pulse.

Both the rise and fall times were analyzed in the same way. First, the ASHL and ASHR signals were scaled to be between 0 and 1, and the trajectory of ASHL vs. ASHR was plotted. For rise times, the trajectory moves forward in the positive x and y direction (Figure 8.4b, color of points goes to red moving up the diagonal). For fall times, the trajectory moves in the negative x and y direction (Figure 8.4c, color of points goes to red moving down the diagonal). If the two signals were perfectly correlated, then their trajectory would fall along the diagonal of the plot. If ASHL rises (falls) before ASHR, then the trajectory moves below (above) the diagonal. If ASHR rises (falls) before ASHL, then the trajectory moves above (below) the diagonal. Therefore, an intuitive measure of correlation is the distance of the trajectory from the diagonal.

Let $\vec{F}(t)$ be the vector that describes the fluorescence trajectory of ASHL and ASHR at time t. At time t, $\vec{F}(t) = \binom{x(t)}{y(t)}$ where $x(t)$ and $y(t)$ are the scaled fluorescence of ASHL and ASHR, respectively, at time t (Figure 8.5). Let $\vec{c}$ be the vector that points along the perfectly correlated diagonal. A proper correlation measure should give both the distance between $\vec{F}$ and $\vec{c}$ and should also indicate whether $\vec{F}$ is above or below the diagonal. Both pieces of information can be derived from the cross-product $\vec{F}(t) \times \vec{c}$. The distance is given by $\frac{|\vec{F}(t) \times \vec{c}|}{|\vec{c}|}$. Let $\vec{k}$ be the unit vector normal to the (x, y) plane. If the product $(\vec{F}(t) \times \vec{c}) \cdot \vec{k}$ is positive, then $\vec{F}$ is below the diagonal. Conversely, if the product is negative,

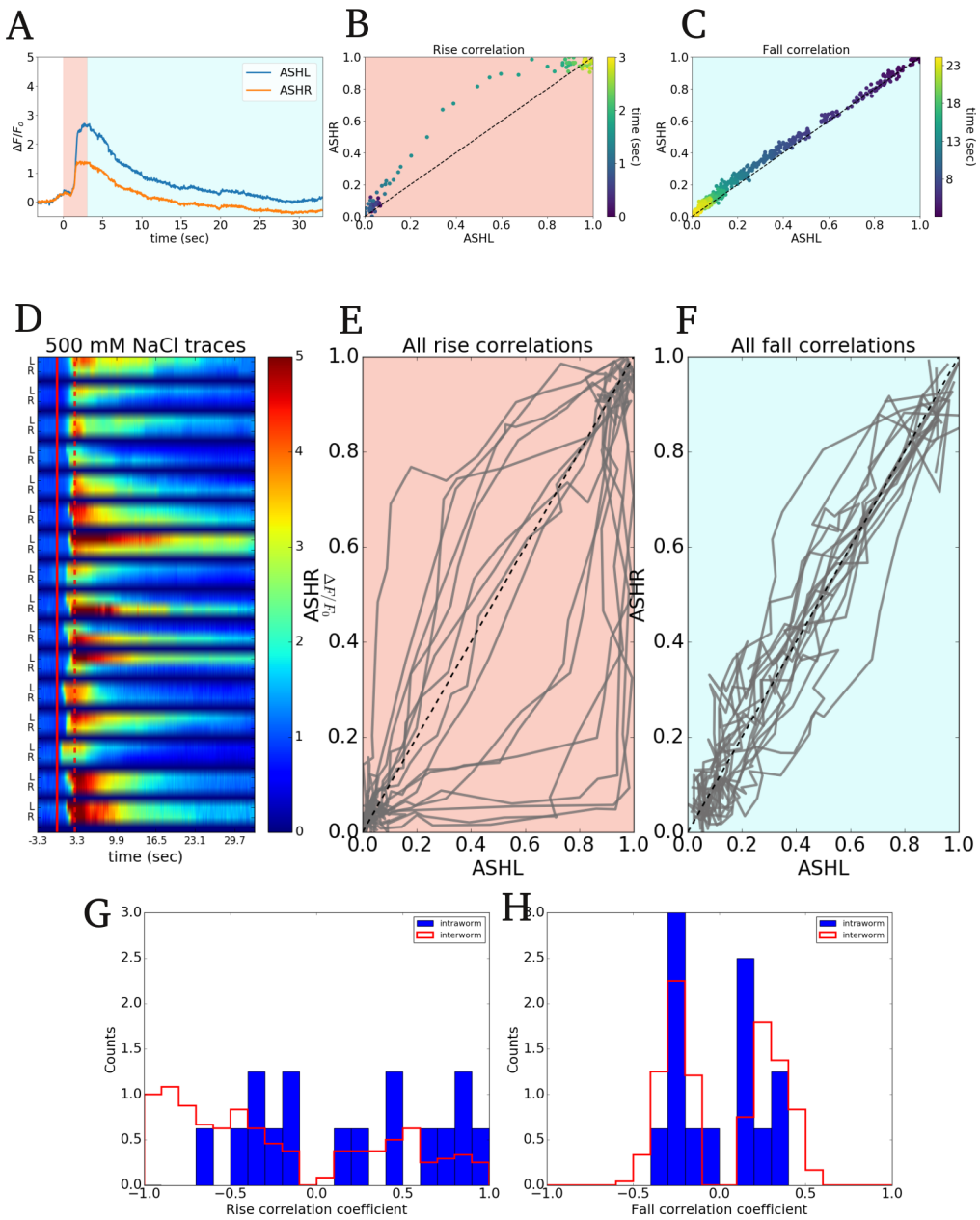


Figure 8.4: **ASH 500 mM NaCl response.** Top: Data from representative individual worm. (A) ASHL (blue) and ASHR (green) trace. (B) ASHL vs. ASHR during 3 sec pulse. (C) Response following the pulse. Colors indicate time after pulse. Dashed line represents a perfectly correlated response. Middle: Data from 16 worms. (D) Heat map of fluorescence response for ASHL (L) and ASHR (R) neurons. Neurons from same worm are in consecutive rows. Solid red line and dashed red line are pulse on and off times, respectively. (E) Correlated responses during 3 sec pulse. (F) Correlated responses after pulse. (G) Distribution max correlation coefficients when pulse is on. (H) Distribution of max correlation coefficients after pulse.

then $\vec{F}$ is above the diagonal. Putting the above expression together, we define the time dependent correlation as

$$\rho(t) = \sqrt{2}\frac{\vec{F}(t) \times \vec{c}}{|\vec{c}|} \cdot \vec{k}, \tag{8.6}$$

where the $\sqrt{2}$ factor ensures the coefficient is scaled to $[-1, 1]$. In practice, we find that trajectories $\vec{F}$ are strictly above or below the diagonal, i.e. $\vec{F}$ does not oscillate across the diagonal. Therefore, for simplicity, we can reduce the time dependent coefficient to a single value by taking the coefficient with the largest magnitude, ρ_{max}. To find ρ_{max}, we find the t_i that produces the largest coefficient in Eq (8.6), then $\rho_{max} = \rho(t_i)$. Using this metric, $\rho_{max} = 0$ indicates ASHL and ASHR are perfectly correlated, $\rho_{max} > 0$ indicates ASHL rises before ASHR and $\rho_{max} < 0$ indicates that ASHR rises before ASHL.

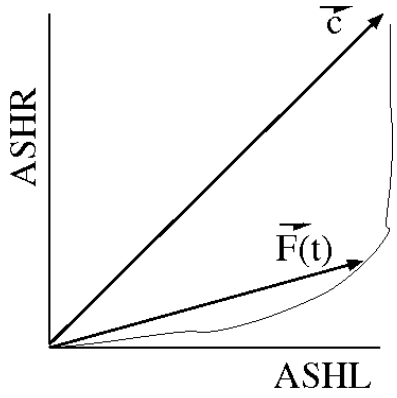


Figure 8.5: **Fluorescence correlation between ASHL and ASHR.** The vector $\vec{F}(t)$ describes the fluorescence trajectory (curve) and $\vec{c}$ is the diagonal vector. See text for explanation.

In general, we find that the rise time is less correlated than the fall time. The spread of the ρ_{max} distribution for rise times is larger than that for the fall times (Figure 8.4g-h). This discrepancy in rise and fall time correlations is in part likely due to the different time constants for the rise and fall times. Typically, the rise of the ASH transient signal occurs over hundreds of milliseconds, while the return to baseline occurs over tens of seconds. The fast rise times means that once one ASH neuron is triggered, it quickly outpaces the other (Figure 8.4e). However, once the stimulus is removed, the slow decay times mean that the ASH neurons can keep pace with each other.

These results clearly show that ASHL and ASHR do not respond simultaneously to the NaCl pulse. This response differential is likely not due to one neuron being exposed to the stimulus before the other. Relative to the worms, the stimulus typically comes from either the ventral or dorsal direction. Hence, both ASHL and ASHR should be exposed to the stimulus at approximately the same time. At least, the differential in exposure time should be much less than the response time differential. Indeed, we find no correlation between the worm's orientation in the chip and whether ASHL or ASHR responds to the stimulus first (data not shown). Taken together, this result could suggest that stimulus responses of ASHL and ASHR are independent.

8.6 ASEL and ASER exhibit independent responses

The ASE neurons are the main neurons that modulate chemoattraction to salts (Bargmann & Horvitz, 1991). Despite being morphologically similar, the neurons ASE left (ASEL) and ASE right (ASER) exhibit different gene expressions, developmental programs and physiological properties (Chang *et al.*, 2003; Pierce-Shimomura *et al.*, 2001; Suzuki *et al.*, 2008). Calcium imaging experiments show that ASEL responds to upsteps in NaCl concentration, while ASER responds to downsteps in NaCl concentration (Suzuki *et al.*, 2008). Typically, either ASEL or ASER is imaged in response to NaCl upshift or downshifts, but the neurons are not imaged together. We imaged ASEL and ASER within the same worm in response to a range of NaCl upshift and downshifts.

We expressed GCaMP under the *flp-6* promoter, which is expressed in both ASEL and ASER. A 3kb genomic fragment immediately upstream of the *flp-6* gene was PCR amplified from wormbank fosmid WRM068aB09 using primers 5'–ACAGGCCGGCCGAAGACTAAGGTGTTCGATCG and 3'– AAACCCGGGC-CACGAGAGTTCATATTCTGG. The amplicon was inserted into a pSM:GCaMP3 vector (also a kind gift of the Bargmann lab) using restriction sites Fse1 and Sma1. Germline transformation were carried out by standard microinjection techniques

(Mello *et al.*, 1991). *flp-6*::GCaMP was co-injected with a *unc-122*:RFP coelemocyte marker at concentrations of 60 ng/μL and 20 ng/μL, respectively, into a wild type (N2 Bristol) worm obtained from the *Caenorhabditis* Genetics Center (Minneapolis, MN, USA).

We simultaneously tested the ASEL and ASER response to upsteps and downsteps of 100 NaCl. To test both neurons, worms were given a 60 second NaCl pulse. We expected ASEL to respond to the beginning of the pulse and ASER to the end of the pulse. We tested 22 worms (Figure 8.6). For 12 worms both ASEL and ASER responded, for 7 worms only ASEL responded, for 2 worms only ASER responded and for 1 worm neither ASEL nor ASER responded. Because there was an appreciable number (9/22) of worms where either ASEL or ASER responded but not both, it suggests that the two neurons respond independently to NaCl within the same worm.

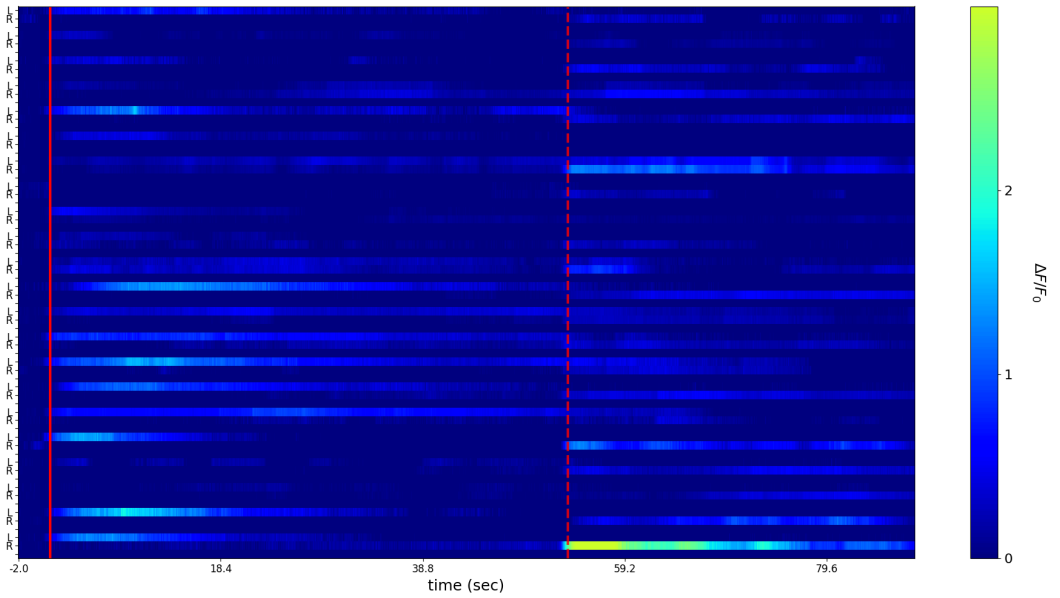


Figure 8.6: **ASE 100 mM NaCl response.** ASEL (L) and ASER (R) responses to 100 mM NaCl upstep (solid red line) followed by a downstep (dashed red line) 60 sec later. Data from 22 worms.

Chapter 9

Discussion

This work presents a novel rotatable microfluidic chip that can be used to image left/right bilateral neurons in *C. elegans*. Unlike other chips that have previously demonstrated the ability to rotate the worm (Ahmed *et al.*, 2016; Ardeshiri *et al.*, 2016), our device does not require physical manipulation of the worm. Instead, the chip itself is rotated, placing left/right bilateral neurons in the same focal plane. Furthermore, our chip has adapted currently established methods for stimulus delivery (Chronis *et al.*, 2007a), meaning left/right neural responses to olfactory stimuli can be reliably measured in an individual animal.

We demonstrated the functionality of the rotatable chip by imaging calcium responses of ASE and ASH to various concentrations of NaCl. ASEL and ASER are known to have asymmetric responses to NaCl, which was easily observed in our chip. However, it was not previously known that ASEL and ASER respond independently to stimuli. We differentiate between independent response, where response of one neuron does not inform the other, and asynchronous responses, where the two neurons respond to different stimuli. That ASEL and ASER do not always respond to NaCl in the same worms suggesting the responses are independent.

As expected ASHL and ASHR responded synchronously to both low ($\leq$ 100 mM) and high (500 mM) concentrations of NaCl. However, while ASHL and ASHR did typically respond synchronously to mid-range concentrations (200-400 mM), we did observe a small number of instances where either ASHL responded or ASHR responded but not both. Moreover, we found the rise times of ASHL

and ASHR are not always correlated. Taken together, these results could suggest that the responses of ASHL and ASHR are independent.

References

Abdi, H. (2007). Kendall rand correlation. In N. Salkind, ed., *Encyclopedia of Measurement and Statistics*, Sage Publications, Thousand Oaks (CA). 30

Ahmed, D., Ozcelik, A., Bojanala, N., Nama, N., Upadhyay, A., Chen, Y., Hanna-Rose, W. & Huang, T.J. (2016). Rotational manipulation of single cells and organisms using acoustic waves. *Nature Communications*, **7**, 11085. 76, 97

Albertson, D.G. & Thomson, J.N. (1976). The Pharynx of Caenorhabditis elegans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **275**, 299–325. 20

Allen, P.B., Sgro, A.E., Chao, D.L., Doepker, B.E., Edgar, J.S., Shen, K. & Chiu, D.T. (2008). Single-synapse ablation and long-term imaging in live C. elegans. *Journal of neuroscience methods*, **173**, 20. 78

Altun, Z.F. & Hall, D.H. (2011). Nervous system, general description. In *WormAtlas*. 21

Andersson, H. & van den Berg, A. (2003). Microfluidic devices for cellomics: a review. *Sensors and Actuators B: Chemical*, **92**, 315–325. 77

Angst, B.D., Marcozzi, C. & Magee, A.I. (2001). The cadherin superfamily: diversity in form and function. *Journal of cell science*, **114**, 629–41. 5

Archer, F.R., Doherty, P., Collins, D. & Bolsover, S.R. (1999). CAMs and FGF cause a local submembrane calcium signal promoting axon outgrowth

without a rise in bulk calcium concentration. *The European journal of neuroscience*, **11**, 3565–73. 13

Ardeshiri, R., Mulcahy, B., Zhen, M. & Rezai, P. (2016). A hybrid microfluidic device for on-demand orientation and multidirectional imaging of C. elegans organs and neurons. *Biomicrofluidics*, **10**, 064111. 76, 97

Artavanis-Tsakonas, S., Rand, M.D. & Lake, R.J. (1999). Notch signaling: cell fate control and signal integration in development. *Science (New York, N.Y.)*, **284**, 770–6. 11

Baier, H. (2013). Synaptic Laminae in the Visual System: Molecular Mechanisms Forming Layers of Perception. *Annual Review of Cell and Developmental Biology*, **29**, 385–416. 6, 7, 8, 23, 51

Bally-Cuif, L., Goridis, C. & Santoni, M.J. (1993). The mouse NCAM gene displays a biphasic expression pattern during neural tube development. *Development (Cambridge, England)*, **117**, 543–52. 11

Barabasi & Albert (1999). Emergence of scaling in random networks. *Science (New York, N.Y.)*, **286**, 509–12. 28

Barclay, A.N. (2003). Membrane proteins with immunoglobulin-like domains–a master superfamily of interaction molecules. *Seminars in immunology*, **15**, 215–23. 5

Bargmann, C.I. (1998). Neurobiology of the Caenorhabditis elegans Genome. *Science*, **282**, 2028–2033. 16

Bargmann, C.I. & Avery, L. (1995). Laser killing of cells in Caenorhabditis elegans. *Methods in cell biology*, **48**, 225–50. 20

Bargmann, C.I. & Horvitz, H.R. (1991). Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in C. elegans. *Neuron*, **7**, 729–42. 79, 95

Bargmann, C.I., Thomas, J.H. & Horvitz, H.R. (1990). Chemosensory cell function in the behavior and development of Caenorhabditis elegans. *Cold Spring Harbor Symposia on Quantitative Biology*, **55**, 529–38. 79, 90

Barres, B.A. (2008). The Mystery and Magic of Glia: A Perspective on Their Roles in Health and Disease. *Neuron*, **60**, 430–440. 9

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–97. 14

Bella, J., Hindle, K.L., McEwan, P.A. & Lovell, S.C. (2008). The leucine-rich repeat structure. *Cellular and Molecular Life Sciences*, **65**, 2307–2333. 5

Ben-Yakar, A., Chronis, N. & Lu, H. (2009). Microfluidics for the analysis of behavior, nerve regeneration, and neural cell biology in C. elegans. *Current Opinion in Neurobiology*, **19**, 561–567. 78

Benjamini, Y. & Hochberg, Y. (1995). Controlling The False Discovery Rate - A Practical And Powerful Approach To Multiple Testing. *Journal of the Royal Statistical Society. Series B: Methodological*, **57**, 289–300. 44

Benjamini, Y. & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, **29**, 1165–1188. 44

Bian, S., Xu, T.L. & Sun, T. (2013). Tuning the cell fate of neurons and glia by microRNAs. *Current Opinion in Neurobiology*, **23**, 928–934. 14

Binzegger, T., Douglas, R.J. & Martin, K.A.C. (2004). A Quantitative Map of the Circuit of Cat Primary Visual Cortex. *Journal of Neuroscience*, **24**, 8441–8453. 32

Blake, J.A. & Ziman, M.R. (2014). Pax genes: regulators of lineage specification and progenitor cell maintenance. *Development (Cambridge, England)*, **141**, 737–51. 12

Bolme, D., Beveridge, J.R., Draper, B.A. & Lui, Y.M. (2010). Visual object tracking using adaptive correlation filters. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2544–2550, IEEE. 87

Braitenberg, V. & Schüz, A. (1998). Peters' Rule and White's Exceptions. In *Cortex: Statistics and Geometry of Neuronal Connectivity*, 99–101, Springer Berlin Heidelberg, Berlin, Heidelberg. 32

Bretscher, A.J., Kodama-Namba, E., Busch, K.E., Murphy, R.J., Soltesz, Z., Laurent, P. & de Bono, M. (2011). Temperature, oxygen, and salt-sensing neurons in C. elegans are carbon dioxide sensors that control avoidance behavior. *Neuron*, **69**, 1099–113. 79

Bushati, N. & Cohen, S.M. (2007). microRNA Functions. *Annual Review of Cell and Developmental Biology*, **23**, 175–205. 14

Cao, D.D., Li, L. & Chan, W.Y. (2016). MicroRNAs: Key Regulators in the Central Nervous System and Their Implication in Neurological Diseases. *International journal of molecular sciences*, **17**. 14

Cardona, A., Saalfeld, S., Schindelin, J., Arganda-Carreras, I., Preibisch, S., Longair, M., Tomancak, P., Hartenstein, V. & Douglas, R.J. (2012). TrakEM2 Software for Neural Circuit Reconstruction. *PLoS ONE*, **7**, e38011. 24

Carrillo, R.A., Özkan, E., Menon, K.P., Nagarkar-Jaiswal, S., Lee, P.T., Jeon, M., Birnbaum, M.E., Bellen, H.J., Garcia, K.C. & Zinn, K. (2015). Control of Synaptic Connectivity by a Network of Drosophila IgSF Cell Surface Proteins. *Cell*, **163**, 1770–1782. 5

Castro, D.S., Skowronska-Krawczyk, D., Armant, O., Donaldson, I.J., Parras, C., Hunt, C., Critchley, J.A., Nguyen, L., Gossler, A., Göttgens, B., Matter, J.M. & Guillemot, F. (2006). Proneural bHLH and Brn Proteins Coregulate a Neurogenic Program through Cooperative Binding to a Conserved DNA Motif. *Developmental Cell*, **11**, 831–844. 11

Cavallaro, U. & Dejana, E. (2011). Adhesion molecule signalling: not always a sticky business. *Nature Reviews Molecular Cell Biology*, **12**, 189–197. 14

Cavallaro, U., Niedermeyer, J., Fuxa, M. & Christofori, G. (2001). N-CAM modulates tumour-cell adhesion to matrix by inducing FGF-receptor signalling. *Nature Cell Biology*, **3**, 650–657. 14

Chang, S., Johnston, R.J. & Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans. *Genes & development*, **17**, 2123–37. 79, 95

Chao, M.Y., Komatsu, H., Fukuto, H.S., Dionne, H.M. & Hart, A.C. (2004). Feeding status and serotonin rapidly and reversibly modulate a Caenorhabditis elegans chemosensory circuit. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15512–7. 79

Chronis, N., Zimmer, M. & Bargmann, C.I. (2007a). Microfluidics for in vivo imaging of neuronal and behavioral activity in Caenorhabditis elegans. *Nature methods*, **4**, 727–31. 76, 82, 83, 89, 97

Chronis, N., Zimmer, M. & Bargmann, C.I. (2007b). Microfluidics for in vivo imaging of neuronal and behavioral activity in Caenorhabditis elegans. *Nature methods*, **4**, 727–31. 77

Chung, K., Crane, M.M. & Lu, H. (2008). Automated on-chip rapid microscopy, phenotyping and sorting of C. elegans. *Nature Methods*, **5**, 637–643. 78

Colón-Ramos, D.A., Margeta, M.A. & Shen, K. (2007). Glia promote local synaptogenesis through UNC-6 (netrin) signaling in C. elegans. *Science (New York, N.Y.)*, **318**, 103–6. 16, 69

Cook, S., Brittin, C., Jarrell, T., Wang, Y., Bloniarz, A., Yakovlev, M., Nguyen, K., Tang, L., Bayer, E., Buelow, H., Hobert, O., Hall, D. & Emmons, S. (2017). Whole-animal connectomes of the two adult sexes of Caenorhabditis elegans. *In preparation*. 16, 24

Crane, M.M., Chung, K. & Lu, H. (2009). Computer-enhanced high-throughput genetic screens of C. elegans in a microfluidic system. *Lab on a chip*, **9**, 38–40. 78

Cremer, H., Lange, R., Christoph, A., Plomann, M., Vopper, G., Roes, J., Brown, R., Baldwin, S., Kraemer, P., Scheff, S., Barthels, D., Rajewsky, K. & Wille, W. (1994). Inactivation of the N-CAM gene in mice results in size reduction of the olfactory bulb and deficits in spatial learning. *Nature*, **367**, 455–459. 11

Crossin, K.L., Chuong, C.M. & Edelman, G.M. (1985). Expression sequences of cell adhesion molecules. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 6942–6. 11

Daly, C. & Ziff, E.B. (1997). Post-transcriptional regulation of synaptic vesicle protein expression and the developmental control of synaptic vesicle formation. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **17**, 2365–75. 6

de Bono, M. & Maricq, A.V. (2005). Neuronal substrates of complex behaviors in C. elegans. *Annual review of neuroscience*, **28**, 451–501. 76

de Wit, J. & Ghosh, A. (2015). Specification of synaptic connectivity by cell surface interactions. *Nature Reviews Neuroscience*, **17**, 4–4. 4, 5

de Wit, J., Hong, W., Luo, L. & Ghosh, A. (2011). Role of Leucine-Rich Repeat Proteins in the Development and Function of Neural Circuits. *Annual Review of Cell and Developmental Biology*, **27**, 697–729. 5

Desai, A.R. & McConnell, S.K. (2000). Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development (Cambridge, England)*, **127**, 2863–72. 10

Doherty, P., Fruns, M., Seaton, P., Dickson, G., Barton, C.H., Sears, T.A. & Walsh, F.S. (1990). A threshold effect of the major isoforms of NCAM on neurite outgrowth. *Nature*, **343**, 464–466. 13

Doherty, P., Williams, E. & Walsh, F.S. (1995). A soluble chimeric form of the L1 glycoprotein stimulates neurite outgrowth. *Neuron*, **14**, 57–66. 13

Duffy, D.C., McDonald, J.C., Schueller, O.J.A. & Whitesides, G.M. (1998). Rapid Prototyping of Microfluidic Systems in Poly(dimethylsiloxane). *Analytical Chemistry*, **70**, 4974–4984. 77

Durbin, R.M. (1987). *Studies on the development and organisation of the nervous system of Caernorhabditis elegans*. Ph.D. thesis, University of Cambridge. 20, 26, 45

Edelstein, A.D., Tsuchida, M.A., Amodaj, N., Pinkard, H., Vale, R.D. & Stuurman, N. (2015). Advanced methods of microscope control using μManager software. *Journal of biological methods*, **1**. 86

Emmons, S.W. (2015). The beginning of connectomics: a commentary on White et al. (1986) 'The structure of the nervous system of the nematode Caenorhabditis elegans'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20140309–20140309. 23

Esumi, S., Kakazu, N., Taguchi, Y., Hirayama, T., Sasaki, A., Hirabayashi, T., Koide, T., Kitsukawa, T., Hamada, S. & Yagi, T. (2005). Monoallelic yet combinatorial expression of variable exons of the protocadherin-α gene cluster in single neurons. *Nature Genetics*, **37**, 171–176. 10

Fiore, R., Siegel, G. & Schratt, G. (2008). MicroRNA function in neuronal development, plasticity and disease. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1779**, 471–478. 14

Forrester, W.C., Dell, M., Perens, E. & Garriga, G. (1999). A C. elegans Ror receptor tyrosine kinase regulates cell motility and asymmetric cell division. *Nature*, **400**, 881–885. 58

Francavilla, C., Loeffler, S., Piccini, D., Kren, A., Christofori, G. & Cavallaro, U. (2007). Neural cell adhesion molecule regulates the cellular

response to fibroblast growth factor. *Journal of Cell Science*, **120**, 4388–4394. 14

Francavilla, C., Cattaneo, P., Berezin, V., Bock, E., Ami, D., de Marco, A., Christofori, G. & Cavallaro, U. (2009). The binding of NCAM to FGFR1 induces a specific cellular response mediated by receptor trafficking. *The Journal of Cell Biology*, **187**, 1101–1116. 14

Fraser, S.E., Carhart, M.S., Murray, B.A., Chuong, C.M. & Edelman, G.M. (1988). Alterations in the Xenopus retinotectal projection by antibodies to Xenopus N-CAM. *Developmental biology*, **129**, 217–30. 12

Gaiano, N. & Fishell, G. (2002). The Role of Notch in Promoting Glial and Neural Stem Cell Fates. *Annual Review of Neuroscience*, **25**, 471–490. 11

Gilmore, E.C. & Herrup, K. (1997). Cortical development: layers of complexity. *Current biology : CB*, **7**, R231–4. 6

Gomez, T.M., Snow, D.M. & Letourneau, P.C. (1995). Characterization of spontaneous calcium transients in nerve growth cones and their effect on growth cone migration. *Neuron*, **14**, 1233–46. 13

Hall, D.H. & Russell, R.L. (1991). The posterior nervous system of the nematode Caenorhabditis elegans: serial reconstruction of identified neurons and complete pattern of synaptic interactions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **11**, 1–22. 17

Hamos, J.E., Van Horn, S.C., Raczkowski, D. & Sherman, S.M. (1987). Synaptic circuits involving an individual retinogeniculate axon in the cat. *The Journal of Comparative Neurology*, **259**, 165–192. 4, 9, 37

Hart, A.C., Kass, J., Shapiro, J.E. & Kaplan, J.M. (1999). Distinct signaling pathways mediate touch and osmosensory responses in a polymodal sensory neuron. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **19**, 1952–8. 76

Helmstaedter, M., Briggman, K.L., Turaga, S.C., Jain, V., Seung, H.S. & Denk, W. (2013). Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, **500**, 168–174. 7

Hilliard, M.A., Bargmann, C.I. & Bazzicalupo, P. (2002). C. elegans Responds to Chemical Repellents by Integrating Sensory Inputs from the Head and the Tail. *Current Biology*, **12**, 730–734. 78, 79

Hobert, O. (2005). The neuronal genome of Caenorhabditis elegans. In *WormBook: The online review of C. elegans Biology.*. 16, 52

Holst, B.D., Goomer, R.S., Wood, I.C., Edelman, G.M. & Jones, F.S. (1994). Binding and activation of the promoter for the neural cell adhesion molecule by Pax-8. *The Journal of biological chemistry*, **269**, 22245–52. 12

Holst, B.D., Wang, Y., Jones, F.S. & Edelman, G.M. (1997). A binding site for Pax proteins regulates expression of the gene for the neural cell adhesion molecule in the embryonic spinal cord. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 1465–70. 12

Howell, K., White, J.G. & Hobert, O. (2015). Spatiotemporal control of a novel synaptic organizer molecule. *Nature*, **523**, 83–87. 69

Huang, P. & Stern, M.J. (2004). FGF signaling functions in the hypodermis to regulate fluid balance in C. elegans. *Development (Cambridge, England)*, **131**, 2595–604. 52

Hulme, S.E., Shevkoplyas, S.S., Apfeld, J., Fontana, W. & Whitesides, G.M. (2007). A microfabricated array of clamps for immobilizing and imaging C. elegans. *Lab on a Chip*, **7**, 1515. 77

Hulme, S.E., Shevkoplyas, S.S., McGuigan, A.P., Apfeld, J., Fontana, W. & Whitesides, G.M. (2010). Lifespan-on-a-chip: microfluidic chambers for performing lifelong observation of C. elegans. *Lab on a chip*, **10**, 589–97. 77

Ibsen, S., Tong, A., Schutt, C., Esener, S. & Chalasani, S.H. (2015). Sonogenetics is a non-invasive approach to activating neurons in Caenorhabditis elegans. *Nature Communications*, **6**, 8264. 76

Imayoshi, I., Sakamoto, M., Yamaguchi, M., Mori, K. & Kageyama, R. (2010). Essential Roles of Notch Signaling in Maintenance of Neural Stem Cells in Developing and Adult Brains. *Journal of Neuroscience*, **30**, 3489–3498. 11

Jarrell, T.A., Wang, Y., Bloniarz, A.E., Brittin, C.A., Xu, M., Thomson, J.N., Albertson, D.G., Hall, D.H. & Emmons, S.W. (2012). The Connectome of a Decision-Making Neural Network. *Science*, **337**, 437–444. 16, 20

Jones, F.S., Prediger, E.A., Bittner, D.A., De Robertis, E.M. & Edelman, G.M. (1992). Cell adhesion molecules as targets for Hox genes: neural cell adhesion molecule promoter activity is modulated by cotransfection with Hox-2.5 and -2.4. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 2086–90. 12

Jones, F.S., Holst, B.D., Minowa, O., De Robertis, E.M. & Edelman, G.M. (1993). Binding and transcriptional activation of the promoter for the neural cell adhesion molecule by HoxC6 (Hox-3.3). *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 6557–61. 12

Kageyama, R., Ohtsuka, T., Shimojo, H. & Imayoshi, I. (2008). Dynamic Notch signaling in neural progenitor cells and a revised view of lateral inhibition. *Nature neuroscience*, **11**, 1247–51. 11

Kao, C.F. & Lee, T. (2010). Birth time/order-dependent neuron type specification. *Current opinion in neurobiology*, **20**, 14–21. 10

Kaplan, J.M. & Horvitz, H.R. (1993). A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 2227–31. 78, 89

Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., Roberts, M., Morgan, J.L., Tapia, J.C., Seung, H.S., Roncal, W.G., Vogelstein, J.T., Burns, R., Sussman, D.L., Priebe, C.E., Pfister, H. & Lichtman, J.W. (2015). Saturated Reconstruction of a Volume of Neocortex. *Cell*, **162**, 648–61. 32

Kater, S., Davenport, R. & Guthrie, P. (1994). Filopodia as detectors of environmental cues: signal integration through changes in growth cone calcium levels. In *Progress in brain research*, vol. 102, 49–60. 13

Kim, B. & Emmons, S.W. (2017). Multiple conserved cell adhesion protein interactions mediate neural wiring of a sensory circuit in C. elegans. *eLife*, **6**, e29257. 17

Kiryushko, D., Korshunova, I., Berezin, V. & Bock, E. (2006). Neural Cell Adhesion Molecule Induces Intracellular Signaling via Multiple Mechanisms of Ca2+ Homeostasis. *Molecular Biology of the Cell*, **17**, 2278–2286. 13

Kiselyov, V.V., Skladchikova, G., Hinsby, A.M., Jensen, P.H., Kulahin, N., Soroka, V., Pedersen, N., Tsetlin, V., Poulsen, F.M., Berezin, V. & Bock, E. (2003). Structural basis for a direct interaction between FGFR1 and NCAM and evidence for a regulatory role of ATP. *Structure (London, England : 1993)*, **11**, 691–701. 14

Klassen, M.P. & Shen, K. (2007). Wnt Signaling Positions Neuromuscular Connectivity by Inhibiting Synapse Formation in C. elegans. *Cell*, **130**, 704–716. 16, 69

Knights, V. & Cook, S.J. (2010). De-regulated FGF receptors as therapeutic targets in cancer. *Pharmacology & Therapeutics*, **125**, 105–117. 13

Komuniecki, R., Harris, G., Hapiak, V., Wragg, R. & Bamber, B. (2012). Monoamines activate neuropeptide signaling cascades to modulate nociception in C. elegans: a useful model for the modulation of chronic pain? *Invertebrate neuroscience : IN*, **12**, 53–61. 79

Kopan, R. & Ilagan, M.X.G. (2009). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell*, **137**, 216–233. 11

Krajniak, J. & Lu, H. (2010). Long-term high-resolution imaging and culture of C. elegans in chip-gel hybrid microfluidic device for developmental studies. *Lab on a Chip*, **10**, 1862. 78

Langley, J.N. (1895). Note on Regeneration of Prae-Ganglionic Fibres of the Sympathetic. *The Journal of physiology*, **18**, 280–4. 4

Larsch, J., Ventimiglia, D., Bargmann, C.I. & Albrecht, D.R. (2013). High-throughput imaging of neuronal activity in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E4266–73. 89

Lee, C.H., Herman, T., Clandinin, T.R., Lee, R. & Zipursky, S.L. (2001). N-cadherin regulates target specificity in the Drosophila visual system. *Neuron*, **30**, 437–50. 8

Lee, H., Crane, M.M., Zhang, Y. & Lu, H. (2013). Quantitative screening of genes regulating tryptophan hydroxylase transcription in Caenorhabditis elegans using microfluidics and an adaptive algorithm. *Integrative biology : quantitative biosciences from nano to macro*, **5**, 372–80. 78

Lemmon, V., Farr, K.L. & Lagenaur, C. (1989). L1-mediated axon outgrowth occurs via a homophilic binding mechanism. *Neuron*, **2**, 1597–603. 13

Li, W., Kang, L., Piggott, B.J., Feng, Z. & Xu, X.S. (2011). The neural circuits and sensory channels mediating harsh touch sensation in Caenorhabditis elegans. *Nature Communications*, **2**, 315. 76

Liu, Q., Chen, B., Hall, D.H. & Wang, Z.W. (2007). A quantum of neurotransmitter causes minis in multiple postsynaptic cells at the Caenorhabditis elegans neuromuscular junction. *Developmental neurobiology*, **67**, 123–128. 61

Lockery, S. (2007). Channeling the worm: microfluidic devices for nematode neurobiology. *Nature Methods*, **4**, 691–692. 77

Lockery, S.R., Hulme, S.E., Roberts, W.M., Robinson, K.J., Laromaine, A., Lindsay, T.H., Whitesides, G.M. & Weeks, J.C. (2012). A microfluidic device for whole-animal drug screening using electrophysiological measures in the nematode C. elegans. *Lab on a chip*, **12**, 2211–20. 78

Lohmann, C. & Bonhoeffer, T. (2008). A role for local calcium signaling in rapid synaptic partner selection by dendritic filopodia. *Neuron*, **59**, 253–60. 12

Lu, C.S., Zhai, B., Mauss, A., Landgraf, M., Gygi, S. & Van Vactor, D. (2014). MicroRNA-8 promotes robust motor axon targeting by coordinate regulation of cell adhesion molecules during synapse development. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **369**. 15

Lucchetta, E.M., Lee, J.H., Fu, L.A., Patel, N.H. & Ismagilov, R.F. (2005). Dynamics of Drosophila embryonic patterning network perturbed in space and time using microfluidics. *Nature*, **434**, 1134–1138. 77

Mallo, M. & Alonso, C.R. (2013). The regulation of Hox gene expression during animal development. *Development*, **140**, 3951–3963. 12

Marder, E. & Taylor, A.L. (2011). Multiple models to capture the variability in biological neurons and networks. *Nature Neuroscience*, **14**, 133–138. 1

Maricq, A.V., Peckol, E., Driscoll, M. & Bargmann, C.I. (1995). Mechanosensory signalling in C. elegans mediated by the GLR-1 glutamate receptor. *Nature*, **378**, 78–81. 78

Markram, H., Muller, E., Ramaswamy, S., Reimann, M.W., Abdellah, M., Sanchez, C.A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., Kahou, G.A.A., Berger, T.K., Bilgili, A., Buncic, N., Chalimourda, A., Chindemi, G., Courcol, J.D., Delalondre, F., Delattre, V., Druckmann, S., Dumusc, R., Dynes, J., Eilemann, S., Gal, E., Gevaert, M.E., Ghobril, J.P., Gidon, A., Graham, J.W., Gupta, A., Haenel, V., Hay, E., Heinis, T., Hernando, J.B., Hines, M., Kanari, L., Keller, D., Kenyon, J.,

Khazen, G., Kim, Y., King, J.G., Kisvarday, Z., Kumbhar, P., Lasserre, S., Le Bé, J.V., Magalhães, B.R.C., Merchán-Pérez, A., Meystre, J., Morrice, B.R., Muller, J., Muñoz-Céspedes, A., Muralidhar, S., Muthurasa, K., Nachbaur, D., Newton, T.H., Nolte, M., Ovcharenko, A., Palacios, J., Pastor, L., Perin, R., Ranjan, R., Riachi, I., Rodríguez, J.R., Riquelme, J.L., Rössert, C., Sfyrakis, K., Shi, Y., Shillcock, J.C., Silberberg, G., Silva, R., Tauheed, F., Telefont, M., Toledo-Rodriguez, M., Tränkler, T., Van Geit, W., Díaz, J.V., Walker, R., Wang, Y., Zaninetta, S.M., DeFelipe, J., Hill, S.L., Segev, I. & Schürmann, F. (2015). Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, **163**, 456–92. 32

Mattson, M.P. & Kater, S.B. (1987). Calcium regulation of neurite elongation and growth cone motility. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **7**, 4034–43. 13

McCaig, C.D. (1989). Studies on the mechanism of embryonic frog nerve orientation in a small applied electric field. *Journal of cell science*, **93 (Pt 4)**, 723–30. 13

McConnell, S.K. & Kaznowski, C.E. (1991). Cell cycle dependence of laminar determination in developing neocortex. *Science (New York, N.Y.)*, **254**, 282–5. 10

Meiri, K.F., Saffell, J.L., Walsh, F.S. & Doherty, P. (1998). Neurite outgrowth stimulated by neural cell adhesion molecules requires growth-associated protein-43 (GAP-43) function and is associated with GAP-43 phosphorylation in growth cones. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **18**, 10429–37. 13

Mellem, J.E., Brockie, P.J., Zheng, Y., Madsen, D.M. & Maricq, A.V. (2002). Decoding of Polymodal Sensory Stimuli by Postsynaptic Glutamate Receptors in C. elegans. *Neuron*, **36**, 933–944. 78

Mello, C.C., Kramer, J.M., Stinchcomb, D. & Ambros, V. (1991). Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO Journal*, **10**, 3959–70. 96

Meyer, R.L. (1998). Roger Sperry and his chemoaffinity hypothesis. *Neuropsychologia*, **36**, 957–80. 4

Mills, H., Wragg, R., Hapiak, V., Castelletto, M., Zahratka, J., Harris, G., Summers, P., Korchnak, A., Law, W., Bamber, B. & Komuniecki, R. (2012). Monoamines and neuropeptides interact to inhibit aversive behaviour in Caenorhabditis elegans. *The EMBO journal*, **31**, 667–78. 79

Mishchenko, Y., Hu, T., Spacek, J., Mendenhall, J., Harris, K.M. & Chklovskii, D.B. (2010). Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron*, **67**, 1009–20. 32

Moore, S.E., Thompson, J., Kirkness, V., Dickson, J.G. & Walsh, F.S. (1987). Skeletal muscle neural cell adhesion molecule (N-CAM): changes in protein and mRNA species during myogenesis of muscle cell lines. *The Journal of cell biology*, **105**, 1377–86. 11

Morishita, H. & Yagi, T. (2007). Protocadherin family: diversity, structure, and function. *Current Opinion in Cell Biology*, **19**, 584–592. 10

Moscoso, L.M. & Sanes, J.R. (1995). Expression of four immunoglobulin superfamily adhesion molecules (L1, Nr-CAM/Bravo, neurofascin/ABGP, and N-CAM)in the developing mouse spinal cord. *The Journal of Comparative Neurology*, **352**, 321–334. 11

Nagayama, S., Homma, R. & Imamura, F. (2014). Neuronal organization of olfactory bulb circuits. *Frontiers in neural circuits*, **8**, 98. 6

Nern, A., Zhu, Y. & Zipursky, S.L. (2008). Local N-Cadherin Interactions Mediate Distinct Steps in the Targeting of Lamina Neurons. *Neuron*, **58**, 34–41. 8

Neves, G., Zucker, J., Daly, M. & Chess, A. (2004). Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nature Genetics*, **36**, 240–246. 10

Newman, M.E.J., Strogatz, S.H. & Watts, D.J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, **64**, 026118. 28

Newman, M.E.J.M.E.J., Barabasi, A.L. & Watts, D.J. (2006). *The structure and dynamics of networks*. Princeton University Press. 26

Nguyen, J.P., Shipley, F.B., Linder, A.N., Plummer, G.S., Liu, M., Setru, S.U., Shaevitz, J.W. & Leifer, A.M. (2016). Whole-brain calcium imaging with cellular resolution in freely behaving Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E1074–81. 76

Nieman, M.T., Prudoff, R.S., Johnson, K.R. & Wheelock, M.J. (1999). N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. *The Journal of cell biology*, **147**, 631–44. 14

Ohtsuka, T., Ishibashi, M., Gradwohl, G., Nakanishi, S., Guillemot, F. & Kageyama, R. (1999). Hes1 and Hes5 as Notch effectors in mammalian neuronal differentiation. *The EMBO Journal*, **18**, 2196–2207. 11

Pan, Q., Shai, O., Lee, L.J., Frey, B.J. & Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, **40**, 1413–1415. 57

Pedersen, M.E., Snieckute, G., Kagias, K., Nehammer, C., Multhaupt, H.A.B., Couchman, J.R. & Pocock, R. (2013). An Epidermal MicroRNA Regulates Neuronal Migration Through Control of the Cellular Glycosylation State. *Science*, **341**, 1404–1408. 14

Perkins, L.A., Hedgecock, E.M., Thomson, J.N. & Culotti, J.G. (1986). Mutant sensory cilia in the nematode Caenorhabditis elegans. *Developmental biology*, **117**, 456–87. 78

Perneger, T.V. (1998). What's wrong with Bonferroni adjustments. *BMJ (Clinical research ed.)*, **316**, 1236–8. 44

Pierce-Shimomura, J.T., Faumont, S., Gaston, M.R., Pearson, B.J. & Lockery, S.R. (2001). The homeobox gene lim-6 is required for distinct chemosensory representations in C. elegans. *Nature*, **410**, 694–8. 79, 80, 95

Pincus, Z., Smith-Vikos, T. & Slack, F.J. (2011). MicroRNA Predictors of Longevity in Caenorhabditis elegans. *PLoS Genetics*, **7**, e1002306. 77

Rees, C.L., Moradi, K. & Ascoli, G.A. (2017). Weighing the Evidence in Peters' Rule: Does Neuronal Morphology Predict Connectivity? *Trends in neurosciences*, **40**, 63–71. 32

Reimann, M.W., King, J.G., Muller, E.B., Ramaswamy, S. & Markram, H. (2015). An algorithm to predict the connectome of neural microcircuits. *Frontiers in computational neuroscience*, **9**, 120. 32

Sambongi, Y., Nagae, T., Liu, Y., Yoshimizu, T., Takeda, K., Wada, Y. & Futai, M. (1999). Sensing of cadmium and copper ions by externally exposed ADL, ASE, and ASH neurons elicits avoidance response in Caenorhabditis elegans. *Neuroreport*, **10**, 753–7. 78, 79

Sambongi, Y., Takeda, K., Wakabayashi, T., Ueda, I., Wada, Y. & Futai, M. (2000). Caenorhabditis elegans senses protons through amphid chemosensory neurons: proton signals elicit avoidance behavior. *Neuroreport*, **11**, 2229–32. 78, 79

Sanes, J.R. & Yamagata, M. (1999). Formation of lamina-specific synaptic connections. *Current opinion in neurobiology*, **9**, 79–87. 9

Sanes, J.R. & Zipursky, S.L. (2010). Design Principles of Insect and Vertebrate Visual Systems. *Neuron*, **66**, 15–36. 6, 8

Sawaya, M.R., Wojtowicz, W.M., Andre, I., Qian, B., Wu, W., Baker, D., Eisenberg, D. & Zipursky, S.L. (2008). A Double S Shape Provides the Structural Basis for the Extraordinary Binding Specificity of Dscam Isoforms. *Cell*, **134**, 1007–1018. 9

Schratt, G. (2009). microRNAs at the synapse. *Nature Reviews Neuroscience*, **10**, 842–849. 15

Schreiner, D., Nguyen, T.M., Russo, G., Heber, S., Patrignani, A., Ahrné, E. & Scheiffele, P. (2014). Targeted Combinatorial Alternative Splicing Generates Brain Region-Specific Repertoires of Neurexins. *Neuron*, **84**, 386–398. 6

Schrödel, T., Prevedel, R., Aumayr, K., Zimmer, M. & Vaziri, A. (2013). Brain-wide 3D imaging of neuronal activity in Caenorhabditis elegans with sculpted light. *Nature Methods*, **10**, 1013–1020. 76

Shen, K. & Bargmann, C.I. (2003). The immunoglobulin superfamily protein SYG-1 determines the location of specific synapses in C. elegans. *Cell*, **112**, 619–30. 16, 69

Shen, K. & Scheiffele, P. (2010). Genetics and cell biology of building specific synaptic connectivity. *Annual review of neuroscience*, **33**, 473–507. 4, 7, 38

Shen, K., Fetter, R.D. & Bargmann, C.I. (2004). Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1. *Cell*, **116**, 869–81. 16, 69

Sheng, L., Leshchyns'ka, I. & Sytnyk, V. (2013). Cell adhesion and intracellular calcium signaling in neurons. *Cell Communication and Signaling*, **11**, 94. 13

Shepherd, G.M.G., Stepanyants, A., Bureau, I., Chklovskii, D. & Svoboda, K. (2005). Geometric and functional organization of cortical circuits. *Nature neuroscience*, **8**, 782–90. 32

Shimono, Y., Rikitake, Y., Mandai, K., Mori, M. & Takai, Y. (2012). Immunoglobulin superfamily receptors and adherens junctions. *Sub-Cellular Biochemistry*, **60**, 137–170. 5

Shinza-Kameda, M., Takasu, E., Sakurai, K., Hayashi, S. & Nose, A. (2006). Regulation of Layer-Specific Targeting by Reciprocal Expression of a Cell Adhesion Molecule, Capricious. *Neuron*, **49**, 205–213. 8

Soroka, V., Kolkova, K., Kastrup, J.S., Diederichs, K., Breed, J., Kiselyov, V.V., Poulsen, F.M., Larsen, I.K., Welte, W., Berezin, V., Bock, E. & Kasper, C. (2003). Structure and interactions of NCAM Ig1-2-3 suggest a novel zipper mechanism for homophilic adhesion. *Structure (London, England : 1993)*, **11**, 1291–301. 14

Sperry, R.W. (1963). CHEMOAFFINITY IN THE ORDERLY GROWTH OF NERVE FIBER PATTERNS AND CONNECTIONS. *Proceedings of the National Academy of Sciences of the United States of America*, **50**, 703–10. 4

Stavoe, A.K.H., Nelson, J.C., Martínez-Velázquez, L.A., Klein, M., Samuel, A.D.T. & Colón-Ramos, D.A. (2012). Synaptic vesicle clustering requires a distinct MIG-10/Lamellipodin isoform and ABI-1 downstream from Netrin. *Genes & development*, **26**, 2206–21. 57

Stirman, J.N., Brauner, M., Gottschalk, A. & Lu, H. (2010). High-throughput study of synaptic transmission at the neuromuscular junction enabled by optogenetics and microfluidics. *Journal of Neuroscience Methods*, **191**, 90–93. 77, 78

Stoenica, L., Senkov, O., Gerardy-Schahn, R., Weinhold, B., Schachner, M. & Dityatev, A. (2006). In vivo synaptic plasticity in the dentate gyrus of mice deficient in the neural cell adhesion molecule NCAM or its polysialic acid. *European Journal of Neuroscience*, **23**, 2255–2264. 12

Südhof, T.C. (2017). Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. *Cell*, **171**, 745–769. 6, 9

Sulston, J.E., Schierenberg, E., White, J.G. & Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental biology*, **100**, 64–119. 20

Sun, L.O., Jiang, Z., Rivlin-Etzion, M., Hand, R., Brady, C.M., Matsuoka, R.L., Yau, K.W., Feller, M.B. & Kolodkin, A.L. (2013). On and off retinal circuit assembly by divergent molecular mechanisms. *Science (New York, N.Y.)*, **342**, 1241974. 7

Suyama, K., Shapiro, I., Guttman, M. & Hazan, R.B. (2002). A signaling pathway leading to metastasis is controlled by N-cadherin and the FGF receptor. *Cancer cell*, **2**, 301–14. 14

Suzuki, H., Thiele, T.R., Faumont, S., Ezcurra, M., Lockery, S.R. & Schafer, W.R. (2008). Functional asymmetry in Caenorhabditis elegans taste neurons and its computational role in chemotaxis. *Nature*, **454**, 114–7. 79, 80, 95

Takeichi, M. (2007). The cadherin superfamily in neuronal connections and interactions. *Nature Reviews Neuroscience*, **8**, 11–20. 5

Thomas, J.H. & Lockery, S. (1999). Neurobiology. In I. Hope, ed., *C. Elegans: a Practical. Approach*, 143–176., Oxford University Press, Oxford. 20

Tobin, D.M. & Bargmann, C.I. (2004). Invertebrate nociception: behaviors, neurons and molecules. *Journal of neurobiology*, **61**, 161–74. 78, 89

Toma, K., Wang, T.C. & Hanashima, C. (2016). Encoding and decoding time in neural development. *Development, Growth & Differentiation*, **58**, 59–72. 10

Tomasiewicz, H., Ono, K., Yee, D., Thompson, C., Goridis, C., Rutishauser, U. & Magnuson, T. (1993). Genetic deletion of a neural cell adhesion molecule variant (N-CAM-180) produces distinct defects in the central nervous system. *Neuron*, **11**, 1163–74. 12

Troemel, E.R., Chou, J.H., Dwyer, N.D., Colbert, H.A. & Bargmann, C.I. (1995). Divergent seven transmembrane receptors are candidate chemosensory receptors in C. elegans. *Cell*, **83**, 207–18. 78, 79

Ullrich, B., Ushkaryov, Y.A. & Südhof, T.C. (1995). Cartography of neurexins: more than 1000 isoforms generated by alternative splicing and expressed in distinct subsets of neurons. *Neuron*, **14**, 497–507. 6

Ushkaryov, Y.A., Petrenko, A.G., Geppert, M. & Südhof, T.C. (1992). Neurexins: synaptic cell surface proteins related to the alpha-latrotoxin receptor and laminin. *Science (New York, N.Y.)*, **257**, 50–6. 6

Utton, M.A., Eickholt, B., Howell, F.V., Wallis, J. & Doherty, P. (2001). Soluble N-cadherin stimulates fibroblast growth factor receptor dependent neurite outgrowth and N-cadherin and the fibroblast growth factor receptor co-cluster in cells. *Journal of neurochemistry*, **76**, 1421–30. 14

Valastyan, S. & Weinberg, R.A. (2011). Roles for microRNAs in the regulation of cell adhesion molecules. *Journal of cell science*, **124**, 999–1006. 15

Varshney, L.R., Chen, B.L., Paniagua, E., Hall, D.H. & Chklovskii, D.B. (2011). Structural Properties of the Caenorhabditis elegans Neuronal Network. *PLoS Computational Biology*, **7**, 21. 16, 17, 24

Vo, N., Klein, M.E., Varlamova, O., Keller, D.M., Yamamoto, T., Goodman, R.H. & Impey, S. (2005). From The Cover: A cAMP-response element binding protein-induced microRNA regulates neuronal morphogenesis. *Proceedings of the National Academy of Sciences*, **102**, 16426–16431. 14

Walthall, W.W., Li, L., Plunkett, J.A. & Hsu, C.Y. (1993). Changing synaptic specificities in the nervous system ofCaenorhabditis elegans: Differentiation of the DD motoneurons. *Journal of Neurobiology*, **24**, 1589–1599. 4

Wang, Y., Jones, F.S., Krushel, L.A. & Edelman, G.M. (1996). Embryonic expression patterns of the neural cell adhesion molecule gene are regulated by homeodomain binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 1892–6. 12

Wani, S. & Kuroyanagi, H. (2017). An emerging model organism Caenorhabditis elegans for alternative pre-mRNA processing in vivo. *Wiley Interdisciplinary Reviews: RNA*, **8**, e1428. 57, 72

Weledji, E.P. & Assob, J.C. (2014). The ubiquitous neural cell adhesion molecule (N-CAM). *Annals of medicine and surgery (2012)*, **3**, 77–81. 11

White, J. (1985). Neuronal connectivity in Caenorhabditis elegans. *Trends in Neurosciences*, **8**, 277–283. 17

White, J., Southgate, E., Thomson, J. & Brenner, S. (1983). Factors That Determine Connectivity in the Nervous System of Caenorhabditis elegans. *Cold Spring Harbor Symposia on Quantitative Biology*, **48**, 633–640. 17, 18, 23, 25, 26, 30, 69

White, J.G., Albertson, D.G. & Anness, M.A.R. (1978). Connectivity changes in a class of motoneurone during the development of a nematode. *Nature*, **271**, 764–766. 4

White, J.G., Southgate, E., Thomson, J.N. & Brenner, S. (1986). The Structure of the Nervous System of the Nematode Caenorhabditis elegans. *Philosophical Transactions of the Royal Society B Biological Sciences*, **314**, 1–340. 16, 17, 20, 23, 24, 30, 37, 68, 69, 72, 76, 79

Whitesides, G.M. (2006). The origins and the future of microfluidics. *Nature*, **442**, 368–373. 77

Williams, E.J., Doherty, P., Turner, G., Reid, R.A., Hemperly, J.J. & Walsh, F.S. (1992). Calcium influx into neurons can solely account for cell contact-dependent neurite outgrowth stimulated by transfected L1. *The Journal of cell biology*, **119**, 883–92. 13

Williams, E.J., Williams, G., Howell, F.V., Skaper, S.D., Walsh, F.S. & Doherty, P. (2001). Identification of an N-cadherin Motif That Can Interact with the Fibroblast Growth Factor Receptor and Is Required for Axonal Growth. *Journal of Biological Chemistry*, **276**, 43879–43886. 13

Wojtowicz, W.M., Wu, W., Andre, I., Qian, B., Baker, D. & Zipursky, S.L. (2007). A Vast Repertoire of Dscam Binding Specificities Arises from Modular Interactions of Variable Ig Domains. *Cell*, **130**, 1134–1145. 9, 51

Wormbase (2017). web site. 52

Wragg, R.T., Hapiak, V., Miller, S.B., Harris, G.P., Gray, J., Komuniecki, P.R. & Komuniecki, R.W. (2007). Tyramine and octopamine

independently inhibit serotonin-stimulated aversive behaviors in Caenorhabditis elegans through two novel amine receptors. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **27**, 13402–12. 79

Wu, Q. & Maniatis, T. (1999). A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell*, **97**, 779–90. 10

Xu, M., Jarrell, T.A., Wang, Y., Cook, S.J., Hall, D.H. & Emmons, S.W. (2013). Computer Assisted Assembly of Connectomes from Electron Micrographs: Application to Caenorhabditis elegans. *PLoS ONE*, **8**, e54050. 24

Yagi, T. (2012). Molecular codes for neuronal individuality and cell assembly in the brain. *Frontiers in Molecular Neuroscience*, **5**, 45. 10, 51

Yamagata, M. & Sanes, J.R. (2008). Dscam and Sidekick proteins direct lamina-specific synaptic connections in vertebrate retina. *Nature*, **451**, 465–469. 7, 18

Yamagata, M. & Sanes, J.R. (2012). Expanding the Ig Superfamily Code for Laminar Specificity in Retina: Expression and Role of Contactins. *Journal of Neuroscience*, **32**, 14402–14414. 7

Yogev, S. & Shen, K. (2014). Cellular and molecular mechanisms of synaptic specificity. *Annu Rev Cell Dev Biol*, **30**, 417–437. 4, 5, 18

Zhang, Y., Lu, H. & Bargmann, C.I. (2005). Pathogenic bacteria induce aversive olfactory learning in Caenorhabditis elegans. *Nature*, **438**, 179–184. 77

Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keeffe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., Deng, S., Liddelow, S.A., Zhang, C., Daneman, R., Maniatis, T., Barres, B.A. & Wu, J.Q. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **34**, 11929–47. 6

Zimmer, M., Gray, J.M., Pokala, N., Chang, A.J., Karow, D.S., Marletta, M.A., Hudson, M.L., Morton, D.B., Chronis, N. & Bargmann, C.I. (2009). Neurons Detect Increases and Decreases in Oxygen Levels Using Distinct Guanylate Cyclases. *Neuron*, **61**, 865–879. 77

Zipursky, S.L. & Sanes, J.R. (2010). Chemoaffinity Revisited: Dscams, Protocadherins, and Neural Circuit Assembly. *Cell*, **143**, 343–353. 9, 10, 57, 71