GENE EXPRESSION PROFILES OF THE *C. ELEGANS* NERVOUS SYSTEM REVEAL TARGETS OF THE SYNAPTIC PROTEIN RPM-1

By

Joseph Daniel Watson

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Neuroscience

August, 2007

Nashville, Tennessee

Approved:

Bruce Appel

Kendal Broadie

Randy Blakely

David Miller

To Betsy,

my strength

ACKNOWLEDGEMENTS

I barely remember who I was when I started the Vanderbilt Interdisciplinary Graduate Program. That was seven years ago, and I am still amazed at the amount of effort needed to complete a Ph.D in Neuroscience and Developmental Biology. I learned a considerable amount of theory through my undergraduate work, but the practical, hands-on application necessary to drive scientific discovery was very limited at the University of Tennessee, Martin. The breadth of research available at Vanderbilt provided me with the chance to explore science in ways I had never considered, and I am grateful to the institution itself for giving me the chance to develop my skills in such an ideal setting.

After IGP, I joined the Miller lab, which studies neural development in the nematode C. elegans. David Miller, my mentor, had his work cut out for him when I joined the lab. While I view myself as a logical thinker and a hard worker, I have always been petrified of presenting to large groups. David provided me the opportunity to present my work multiple times, in front of a variety of groups. This treatment had two effects: first, the elimination of any pride I had in my ability to communicate clearly with others and second, it provided me with a new perspective on how to present an interesting talk to a neutral crowd. Ultimately, David's teaching has given me the opportunity to grow as a researcher, in my thinking, my writing, and in my public appearances. He is a brilliant person and his 'massaging' of my manuscripts and abstracts has only made them more accessible.

The Miller lab ranged in size from 5 members, up to around 15 (during the Summer months), but we were never really large, and the membership of the lab changed rather frequently. Only two people were with me in the lab throughout the majority of graduate school. Steve, thanks so much for listening to my ideas, no matter how crazy or off-the-wall. While I didn't always agree with your advice, I always listened to it. Rebecca, I have always admired your persistence and willingness to take on any task, no matter how large or complex. We made it through the entire graduate school process together, and I still would really like to see Oswego! Thanks to both of you. We made it through our qualifying exam, NRSA, and even graduate school as a group. Six years of togetherness is a long time, and both of you have added so many new dimensions to my career and my personal life (especially Trogdor).

I would also like to thank other members of the lab, both current and former. Kathie, thank you for injecting for me when I needed it and for listening to me complain about my family. Clay, thanks for putting up with all the crap I gave you over the years. You were a great running partner and also a good person to slap around. Laurie, thanks for being an outlet to channel my rage and frustration. You are a fun person to work with and one of the most knowledgeable scientists I know. Julie, thanks for showing me how to be strong during difficult times. Jud, Sarah, and Rachel, while I don't know you as well as the others, I appreciate your comments on my work and your critiques of my presentations.

Thanks to my mom and dad, who provided me a solid base to begin my career in science. Thanks to my sister Angie, who shares my feelings like no one else can. Thanks

to Jim and Bettye Griffin for providing me with a wonderful wife and loving me from day one.

To my wife Betsy, you have supported me through this entire journey. You have been there through every ounce of sadness. You have never wavered, even through some of the scariest moments of my life. Your presence alone inspires me, and your compassion for others is my source of hope. I pray I can give you half of the encouragement you have shown me as you finish your doctorate.

I would like to thank the Program in Neuroscience for funding my first two years of research. The National Institute of Neurological Disorders and Stroke (NINDS) provided funding for three years of my graduate work. Finally, I would like to thank the Cell and Developmental Biology, Program in Development, and Program in Neuroscience, each of which enhanced my training by providing Journal Clubs and outlets to practice presenting data in front of large groups.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

DEFINING A LINK BETWEEN SYNAPTIC ASSEMBLY AND GENE TRANSCRIPTION

Introduction

Several steps are necessary in order to establish a functional nervous system. In the developing spinal cord, signaling molecules are released from the notochord and floorplate and act, in a concentration dependent manner, to activate signaling cascades that regulate transcription factors (Lee and Pfaff 2001; Shirasaki and Pfaff 2002). In turn, the transcription factors initiate a genetic program that defines the neural domains, and provides each individual neuron with a functional identity. The neuroblast integrates these transcriptional pathways with external cues to target axon outgrowth to a particular destination, and then selectively synapse with postsynaptic cells (Arber, Han *et al*. 1999; Thaler, Harrison *et al*. 1999). The mechanistic pathways that define axon outgrowth and neural differentiation have been studied extensively (Lee and Pfaff 2001; Chilton 2006). In the past few years, more definition has been given to the mechanisms that control synaptic development. However, many questions remain. This introduction will review recent insights into the mechanisms that define synaptic structure. In addition, it will discuss genomic methods that can be used to identify the transcriptional targets of synaptic development. The final chapter will focus on enhancing gene expression studies using novel RNA amplification techniques.

Synaptic structure

Neuronal communication depends on synaptic structure. EM reconstruction experiments across species have revealed two basic types of chemical synapses (Jin). In *terminaux* synapses, the presynaptic region forms at the end of an axon. In contrast, in *en passant* synapses, connections form on closely apposed axon tracts. These synapses contain common features. Presynaptically, chemical synapses contain an active zone, an electron dense region directly apposed to a postsynaptic target. The active zone contains the machinery to initiate synaptic vesicle fusion. Surrounding the active zone are synaptic vesicles, the communication packets of the synapse. Vesicles are arranged into several compartments. The first set of vesicles form the ready-releasable pool. These vesicles interact with proteins contained within the electron dense region of the synapse and can rapidly fuse in response to action potentials. The second set of vesicles form a reserve pool, which are mobilized after an action potential to replace the ready releasable pool. Outside of this region lies the periactive zone, which is poorly understood but is linked to synaptic formation and the maintenance of synaptic structure (Jin).

The structural development of the neuromuscular junction has been studied extensively in both *Drosophila Melanogaster* and *Caernohabiditis elegans*. Both species are genetically pliable. *Drosophila* have large, glutamatergic synapses at the NMJ that are readily amenable to electrophysiological recordings of synaptic strength (Ataman, Budnik *et al*. 2006; Ruiz-Canada and Budnik 2006; Collins and DiAntonio 2007). In addition, the nervous system of *Drosophila* is somewhat plastic, and has yielded new insights into models of synaptic remodeling (Sanyal and Ramaswami 2006). *C. elegans* offers the unique additional advantage of a simple, well-defined nervous system (White,

Southgate *et al*. 1976; White, Albertson *et al*. 1978). This fact, when combined with powerful hemaphrodite genetics, provides a powerful reason to use *C. elegans* for forward genetic screens. Together, these two model organisms have the potential of providing new insights into the mechanisms of synaptic development.

Synaptic Assembly through RPM-1/Highwire/Esrom/PAM

Recent work in both *Drosophila* and *C. elegans* has elucidated new mechanisms of synaptic assembly. In *C. elegans*, a forward genetic screen revealed mutations in Regulator of Presynaptic Morphology 1, RPM-1, an E3 ubiquitin ligase (Schaefer, Hadwiger *et al*. 2000; Zhen, Huang *et al*. 2000) (Figure 1.1 A,B). In this screen, GABAergic synapses were marked with GFP-tagged synaptobrevin, an integral synaptic vesicle marker (SNB-1::GFP). Normally, SNB-1::GFP localizes to discrete, regularly spaced puncta within the developing nerve cord (Jin, Jorgensen *et al*. 1999). In contrast, SNB-1::GFP puncta in *rpm-1* mutants are abnormally shaped and sporadically placed along the nerve cord (Zhen, Huang *et al*. 2000). In wildtype animals, GABAergic neurons have a single active zone at each *en passant* synapse. In constrast, EM reconstruction of the *rpm-1* motor neuron circuit revealed multiple active zones per synapse (Zhen, Huang *et al*. 2000). In some cases, no active zones were observed for large stretches of the VNC. Instead, electron-dense synaptic debris filled the terminals and, while synaptic vesicles were present, their number was greatly reduced. These data suggest that, in GABA neurons, RPM-1 functions to define key structural components of the presynaptic density. In a parallel study, Schafer, *et al*, identified *rpm-1* as a mutation

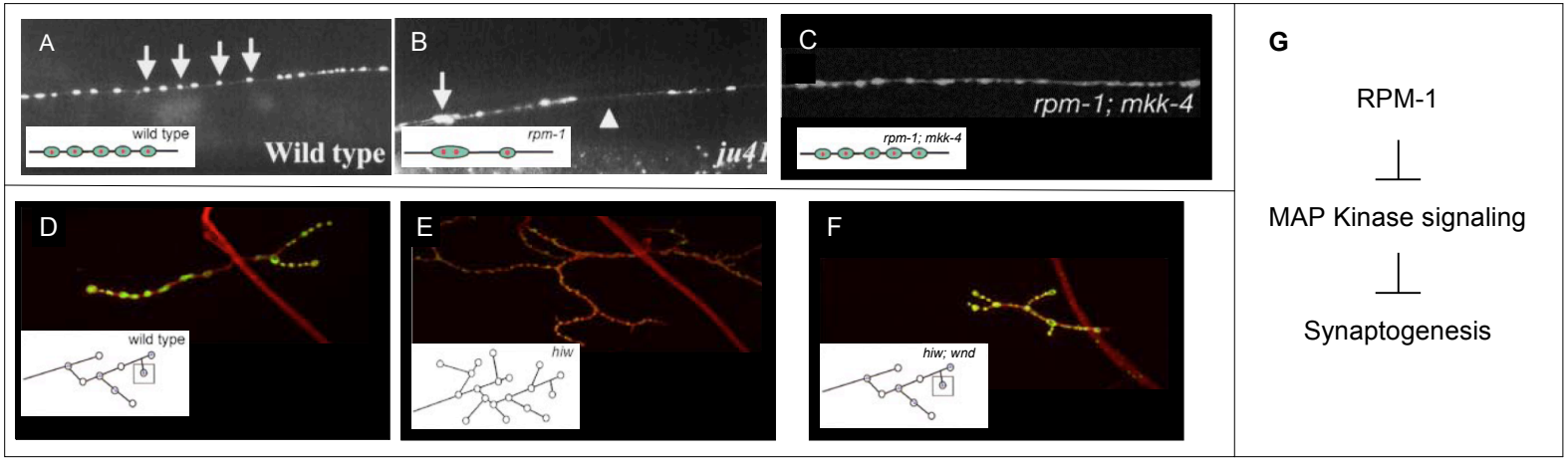


Figure 1.1 RPM-1 regulates synaptic morphology through the MAP Kinase cascade.

A. C. elegans GABAergic synapses (marked by SNB-1::GFP) are normally evenly spaced and punctate.

B. SNB-1::GFP in *rpm-1* mutants is irregularly spaced and forms large aggregates. EM reconstruction revealed multiple active zones per single synapse (inset).

C. Mutations in *mkk-4* (MAP Kinase Kinase) components rescue *rpm-1* synaptic defects.

D. Drosophila glutamatergic NMJ synapses are visualized with the neuronal membrane marker HRP (red) and the synaptic vesicle marker DVGLUT (green).

E. Mutations in the *rpm-1* homolog *highwire* disrupt the NMJ. Synaptic arbors are overgrown and synaptic boutons are decreased in size.

F. Mutations in the MAP Kinase Kinase Wallenda rescues the *highwire* synaptic defects.

(Figure modified from Zhen, et al 2000, Chang, et al 2000 Nakata, et al 2005, and Collins et al 2006)

that disrupts touch neuron synaptic morphology (Schaefer, Hadwiger *et al*. 2000). These findings suggest that RPM-1 functions in either the initial stages of synaptic development or in synaptic maintenance. Schafer, *et al*, also identified an additional axon outgrowth defect in touch neuron processes. These axon outgrowth defects are not observed in GABAergic processes. This finding suggests that RPM-1 fulfills different roles in distinct neuron classes.

A role for RPM-1 proteins in synaptic assembly appears to be conserved in *Drosophila*. Mutations in the RPM-1 homologue *highwire* also disrupt synaptic structure at the NMJ (Chang and Balice-Gordon 2000; Wan, DiAntonio *et al*. 2000) (Figure 1.1 D, E). In wildtype *Drosophila*, presynaptic neurons innervate muscle at many points. The axons that innervate muscle form synaptic arbors, complex structures that enable a single neuron to innervate a large muscle (Collins and DiAntonio 2007). Many synaptic boutons are located within each synaptic arbor, and each of these contains multiple active zones. Mutations in *highwire* increase the number of arbors that are found within muscle (Wan, DiAntonio *et al*. 2000). Correspondingly, the number of synaptic boutons is also increased. Interestingly, the size of each synaptic bouton is reduced, along with the strength of the synaptic signal (Wan, DiAntonio *et al*. 2000).

Additional work in vertebrate model systems has confirmed the evolutionary conservation of RPM-1/Highwire and its role in synaptic development (Burgess, Peterson *et al*. 2004; D'Souza, Hendricks *et al*. 2005). Mutations in esrom, the zebrafish homologue of *rpm-1*, disrupt the structure of the optic tectum (D'Souza, Hendricks *et al*. 2005). In wildtype fish, retinal axons from the anterior eye grow as a single fasicle before branching at the posterior tectum. In contrast, mutations in *esrom* result in early

branching of retinal axons in the anterior tectum. Therefore, esrom is required for correct fasciculation and targeting of retinal axons to the tectum. Similar targeting defects are observed in *Phr1* (PAM/Highwire/RPM-1) mouse mutants (Burgess, Peterson *et al*. 2004). These animals die of respiratory failure at birth. Numerous morphological defects are observed in the neurons that innervate the diaphram. For instance, the number of axons observed in the phrenic nerve bundle is substantially reduced, which limits the innervation of the diaphram. Synaptic fields are also disrupted, with nerve terminal sprouting beyond the postsynaptic receptor plaque. These data suggest that the function of *rpm-1*/*highwire* is highly conserved.

Antibody staining of RPM-1 and Highwire revealed that both proteins are localized at the synapse (Wan, DiAntonio *et al*. 2000; Zhen, Huang *et al*. 2000). This finding suggests that RPM-1/Highwire function locally to regulate synaptic morphology. Furthermore, *rpm-1/highwire* encodes an E3 ubiquitin ligase, which selectively target proteins for degradation. Therefore, RPM-1 could regulate levels of a target protein to specify synaptic inputs. The following experiments provide evidence for this model.

Targets of *rpm-1* signaling

Recent experiments have shown that RPM-1 can ubiquitinate target proteins *in vitro* (Nakata, Abrams et al. 2005). In addition, point mutations in the E3 domain disrupt synaptic structure (Zhen, Huang et al. 2000). These studies suggest that protein degradation may play a role in synaptic assembly. Nakata, *et al*, developed a clever mutant screen to identify potential targets of *rpm-1* (Nakata, Abrams *et al*. 2005). To conduct this screen, they made use of a synergistic movement phenotype created from the

combination of two mutations that affect synaptic assembly. Despite defects in synaptic structure, *rpm-1* mutants show normal movement. Similarly, mutations in *syd-1*, a PDZ domain containing protein (see below) that also affects synaptic structure, movement is only mildly affected (Hallam, Goncharov *et al*. 2002). *rpm-1 syd-1* double mutants, however, are severely uncoordinated. Nakata, *et al*, reasoned that a forward genetic screen might detect mutations in RPM-1 targets as rare animals that show improved movement. For example, if the *rpm-1* mutation results in the ectopic activation of a target that disrupts synaptic function, then a mutation that removes this downstream gene should also eliminate the synaptic defects. This screen revealed three MAP Kinase components – *dlk-1* (MAP KKK), *mkk-4* (MAPKK), and *pmk-3* (p38 MAPK) (Nakata, Abrams *et al*. 2005). MAP Kinase signaling has been implicated in many neural pathways, including axon guidance, neural fate decisions, and synaptic development (Sagasti, Hisamoto *et al*. 2001; Guan, Kim *et al*. 2003; Panicker, Buhusi *et al*. 2003; Sharma and Carew 2004; Nakata, Abrams *et al*. 2005; Collins, Wairkar *et al*. 2006).

Analysis of *rpm-1* double mutants combined with these MAP Kinase genes revealed virtually complete rescue of the Rpm-1 synaptic defects (Nakata, Abrams *et al*. 2005) (Figure 1.1C). Several lines of evidence indicate that the MAP Kinase pathway acts cell autonomously. GFP reporters revealed that all three MAPK components (*dlk-1*, *mkk-4*, and pmk-3) are expressed in neurons and functional DLK-1::GFP localizes to the synapse. Additionally, the rescued synapses of *rpm-1 dlk-1* double mutants can be re-disrupted by DLK-1 expression in neurons. Overexpression of DLK-1 using its endogenous promoter can actually enhance the synaptic defects of *rpm-1* (Nakata,

Abrams *et al*. 2005). These data indicate that DLK-1 functions cell autonomously to regulate synaptic morphology.

Why do mutations in MAP Kinase rescue *rpm-1* synaptic defects? One study reveals that *rpm-1* can ubiquitinate *dlk-1 in vitro* (Nakata, Abrams *et al*. 2005). These data, combined with the synaptic localization of *dlk-1*, suggest that *rpm-1* modulates the levels of MAP Kinase signaling by targeting excess DLK-1 for degradation. Since *dlk-1* activates downstream MAP Kinase components, disrupting any component downstream of DLK-1 (*pmk-3* and *mkk-4*) rescues the synaptic defects of *rpm-1*.

Additional work in *Drosophila* has shown that the MAP Kinase pathway is also conserved downstream of Highwire (Collins, Wairkar *et al*. 2006). Mutations in *highwire* can be fully rescued by corresponding mutations in *wallenda*, a MAPKKK and homologue of *C. elegans dlk-1* (Figure 1.1F). Similar to results from *C. elegans*, Wallenda overexpression is sufficient to cause synaptic overgrowth (Collins, Wairkar *et al*. 2006). Furthermore, mutations in the MAP kinase *jnk* rescue *highwire* synaptic defects. These data suggest that the downstream mechanisms regulated by *rpm-1* are conserved from worm to fly. One interesting difference between these two systems is the regulation of JNK by Highwire, rather than p38 MAP Kinase (Collins, Wairkar *et al*. 2006). This finding could mean that different MAPK pathways are regulated by RPM-1/Highwire in flies and worms. Alternatively, it is also possible that these alternative MAPK pathways could be selectively regulated by RPM-1/Highwire in different classes of neurons. Both p38 and JNK kinases have established roles in synaptic plasticity, indicating that both may be necessary for the proper development of a functional synapse (Guan, Kim *et al*. 2003; Kandel 2004; Sharma and Carew 2004).

RPM-1 regulates gene expression

In our lab, a genetic screen for regulators of motor neuron fate revealed mutations in *rpm-1*. In this screen, I searched for genes that regulate the expression of the a motor neuron specific marker, *acr-5*::YFP (nACh receptor) (Winnier, Meir *et al*. 1999). I identified three mutants, *wd72, wd55, wd67*, that substantially reduced the expression of *acr-5*::GFP in the ventral nerve cord. Additionally, these alleles disrupted the expression of TGF-β (*unc-129*) (Colavita, Krishna *et al*. 1998). Subsequent mapping experiments revealed that these mutations reside in the *rpm-1* locus. We discovered that *rpm-1* regulates gene expression around the time that the Jin lab discovered that MAP Kinase signaling acts downstream of *rpm-1*. I therefore hypothesized that *rpm-1* controls synaptic development by utilizing a MAP Kinase signaling cascade to regulate transcription of specific target genes. Previously published work in *Aplysia* (for a more detailed description, see synaptic plasticity below) provides a well characterized example of the role of transcription in the regulation of synaptic structure (Guan, Kim *et al*. 2003; Kandel 2004; Sharma and Carew 2004). In *Aplysia*, several kinases, including MAP Kinase, act upstream of the transcription factor CREB to mediate long term facilitation. In long term facilitation (LTF), synaptic effectiveness is dramatically increased over a extended period of time. Changes in synaptic structure are a hallmark of LTF and CREB activity is required for these changes (Guan, Kim *et al*. 2003; Kandel 2004; Sharma and Carew 2004).

Direct support for the hypothesis that *rpm-1* controls gene expression comes from work in the DiAntonio lab where the *highwire* synaptic defect is rescued by mutations in the transcription factor Fos (Collins, Wairkar *et al*. 2006). In *Drosophila*, Highwire

inhibits MAP Kinase, which prevents Fos activation. In *highwire* mutants, MAPK activates Fos, altering transcription and disrupting synaptic connections. While the downstream transcription factor driving *rpm-1* changes in synaptic morphology has yet to be isolated, the strong conservation of the *rpm-1* and highwire pathways suggests a similar mechanism of action in *C. elegans* (Figure 1.1G).

Synaptic Development

If transcription factors are indeed acting downstream of *rpm-1*, then an in depth analysis of genes that are transcriptionally regulated by RPM-1 could reveal the downstream mechanism of RPM-1 function. I will now discuss the signaling pathways that are known to regulate synaptic development by controlling transcription. I will also discuss the role of tubulin in synaptic formation since our data suggest that it is regulated by RPM-1.

TGF-β Signaling

The TGF-β/BMP signaling pathway plays an established role in neural differentiation and axon guidance. Recent evidence suggests that BMP signaling also affects synapse formation (Aberle, Haghighi *et al*. 2002; Haghighi, McCabe *et al*. 2003; McCabe, Marques *et al*. 2003; McCabe, Hom *et al*. 2004). In *Drosophila*, mutations in the BMP receptor *wishful thinking* (wit) result in a drastic reduction in synaptic arbor size at the NMJ (Aberle, Haghighi *et al*. 2002). A similar defect is observed in flies that are mutant for the BMP signaling molecule *glass bottom boat* (*gbb*) (McCabe, Marques *et al*. 2003). Rescue experiments revealed that these synaptic defects are exclusively rescued

by postsynaptic expression (*i.e.* in muscle) of GBB (McCabe, Marques *et al*. 2003). This observation suggests that BMP signaling molecules released from post-synaptic targets bind to receptors on the presynaptic membrane and guide development of the presynaptic apparatus. Further work indicates that Wit acts through the SMAD transcription factors Mad and Medea (McCabe, Hom *et al*. 2004). Several lines of evidence corroborate this finding. For example, mutations in either *mad* or *medea* phenocopy *wit* synaptic defects whereas loss-of-function mutations in *wit* disrupt the accumulation of phosphorylated-Mad in the nucleus (McCabe, Hom *et al*. 2004). These data suggest a transcriptional requirement downstream of the BMP signaling pathway to drive synaptogenesis (Figure 1.2).

BMP signaling interacts with Highwire to specify synaptic development

Further work has connected BMP signaling with Highwire. Mutating both *gbb* and *highwire* suppresses the synaptic arbor overgrowth phenotype of *highwire* mutants, while also rescuing arbor undergrowth phenotype associated with *gbb* (McCabe, Hom *et al*. 2004). These data suggest that both BMP signaling and Highwire function are necessary for appropriate synaptic development and may be indicative of shared pathway components. However, the *gbb* mutant only partially rescues the *highwire* defect; in *highwire gbb* double mutants, synaptic boutons are still abnormally small, a characteristic of *highwire* mutants (McCabe, Hom *et al*. 2004). In contrast, mutations in *wallenda* completely rescue Highwire synaptic defects. These data suggest that BMP signaling

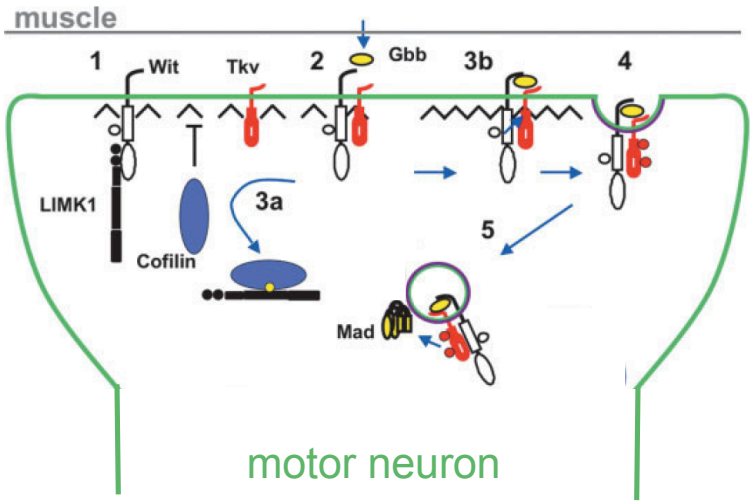


Figure 1.2 BMP signaling regulates synaptic development.
At the Drosophila NMJ, BMP signals released from muscle bind to the pre-synaptic TGF-Beta receptors Wishful thinking (Wit) and Thick-veined (Tkv) (2). This model speculates that the activation of these receptors would relieve the inhibition of Wit on the kinase LIMK1 (1, 3a), which would then inactivate Cofilin and allow the entire receptor complex to be internalized (3b, 4). Tkv could then phosphorylate the transcription factor Mad (5) and activate downstream targets.
(Figure modified from Marques 2005)

may function downstream of the Highwire/MAP Kinase pathway and may be necessary for specifying one aspect of Highwire-regulated synaptic development. It is also possible that BMP signaling and Highwire function independently to affect synaptic development, and the rescued synaptic arbors are merely due to synergistic effects (Collins, Wairkar *et al*. 2006). Further dissection of the downstream transcriptional pathways regulated by both Highwire and Medea may reveal the co-regulation of shared components that define synaptic structure.

Wnt signaling

As shown above, retrograde signals help sculpt the developing synapse. Anterograde signals are also necessary for synaptogenesis. Here I describe the role of wnt, which functions both as a retrograde and anterograde signal to regulate synaptic assembly. In the developing cerebellum, WNT-7a is released from granule cells where it stimulates synaptic formation in mossy fiber axons (Packard, Koo *et al*. 2002; Speese and Budnik 2007). In the mouse, mutations in *wnt-7a* disrupt the onset of synapse maturation (Packard, Koo *et al*. 2002). Wnt can signal through a variety of pathways, but the canonical wnt pathway is the best characterized and appears to function downstream of *wnt-7a* (Speese and Budnik 2007). In this pathway, wnt binds the presynaptic receptor frizzled, which triggers the degradation of GSK-3β, relieving the inhibition of the transcription factor β-Catenin. This link suggests that wnt regulates synaptic development by activating a presynaptic cascade that controls transcription.

Further work in *Drosophila* suggests a postsynaptic role for wnt in synaptic formation. Mutations that disrupt wingless or its receptor frizzled drastically change

synaptic morphology at the NMJ (Packard, Koo *et al*. 2002). For instance, the number of synaptic boutons is reduced by ~50%. Some synapses lack active zones, while in others the active zone structure is grossly abnormal. Post-synaptically, glutamate receptor localization is disrupted (Packard, Koo *et al*. 2002). Therefore, wnt controls development of both pre and post-synaptic structures. How does wnt do this? All of the defects produced by wnt receptor mutants can be rescued by the expression of the wnt receptor frizzled in muscle (Griffith and Budnik 2006; Collins and DiAntonio 2007; Speese and Budnik 2007), whereas in the cerbellum, wnt appears to signal presynaptically. However, none of the canonical wnt pathway components are expressed in muscle. Recent evidence suggests that wnt functions at the *Drosophila* NMJ by activating the frizzled nuclear import pathway (Mathew, Ataman *et al*. 2005; Ataman, Ashley *et al*. 2006). In this pathway, wnt binding to frizzled triggers receptor complex endocytosis. Frizzled is then clipped at the C-terminus and imported into the nucleus, where it may act to modulate transcription. This finding suggests that transcription may also play a role postsynaptically in active zone formation.

One other interesting fact about wnt signaling is that disruption of this signal only alters certain components of presynaptic physical structure. Vesicle pools are still recruited to disrupted synaptic regions (Ghost boutons) (Speese and Budnik 2007). This finding suggests independent signals for active zone formation and vesicle clustering, a finding with parallels in *C. elegans* (see *syd-1* and *syd-2*, below).

The role of Agrin in NMJ formation

Studies of the vertebrate NMJ have also implicated transcription in synapse formation. Agrin is one of the earliest studied synaptic proteins with a demonstrated role in the regulation of transcription. Agrin encodes a heparan sulphate proteoglycan (Ngo, Noakes *et al*. 2007). Agrin is secreted at the developing NMJ, binds to its receptor MuSK (muscle specific kinase), and maintains the presence of AChR clusters at the postsynaptic density (Kummer, Misgeld *et al*. 2006). Additional work has now demonstrated that Agrin induces AChR clustering indirectly by acting in opposition to a dispersal signal (acetylcholine signaling) (Misgeld, Kummer *et al*. 2005; Kummer, Misgeld *et al*. 2006). In this model, acetylcholine release stimulates postsynaptic receptors, which are downregulated through endocytosis. This stimulation also blocks the transcription of nAChR genes in specialized myonuclei that lie beneath the postsynaptic membrane. Agrin blocks both the removal of AChR from the postsynaptic membrane and the transcriptional repression provided by ACh signaling. Mechanisms that control the initial transcriptional activation of AChRs are unknown, but the interaction of Agrin and MuSK is required for the transcriptional maintenance of the densely packed AChRs (Kummer, Misgeld *et al*. 2006). The mechanisms of presynaptically targeting a neuron to the NMJ remains controversial. Originally, it was thought that a presynaptic signal was required to induce AChR clustering, but recent analysis suggests that clusters can also act instructively to induce an adjacent neurite to establish a synapse (Kummer, Misgeld *et al*. 2006). Further work is needed to reveal proteins required for initial clustering and targeting of the axon to the postsynaptic muscle.

Tubulin dynamics

The formation of a functioning synapse depends on the precise regulation of cytoskeletal dynamics. Recent research highlights the key roles of tubulin and its interacting partners in synaptic assembly. Microtubules have the capacity to grow or destabilize quickly and are localized at the central core of the growth cone (Ruiz-Canada and Budnik 2006). At the synapse, static microtubule images suggest one of two conformations for tubulin. Boutons with looping microtubules are correlated with stable presynaptic densities whereas boutons with disassembled microtubule loops are usually associated with an actively growing synapse (Ruiz-Canada and Budnik 2006). Whereas mutations in tubulin usually lead to early lethality in *Drosophila* (Llamazares, Tavosanis *et al*. 1999), mutations in genes that interact with or control cytoskeletal dynamics have provided valuable insight into the role tubulin plays in synaptic development. Mutations in *futsch*, a conserved microtubule binding protein (MAP1B homologue) (Hummel, Krukkert *et al*. 2000; Roos, Hummel *et al*. 2000) result in decreased NMJ bouton number and size. Antibody staining of a weak loss-of-function allele of *futsch* revealed that microtubule structure is disrupted at the NMJ (Roos, Hummel *et al*. 2000). The normal circular tubulin staining seen in synaptic boutons of wildtype animals is replaced with fragmented tubulin throughout the bouton. While the exact function of Futsch is unknown, incubation of the human homologue of Futsch, MAP1B, with microtubules *in vitro* promotes changes in microtubule structure and produces "wavy" microtubules, suggesting that Futsch could be responsible for bending microtubules in vivo (Roos, Hummel *et al*. 2000). Furthermore, Futsch is more likely to intermingle with microtubules at terminal boutons that may be in the process of active division. (Roos,

Hummel *et al*. 2000). These data argue that Futsch binds to microtubules and affects their configuration at the synapse. Additional work suggests translational control of *futsch* by the RNA-binding protein Fragile X provides the neuron with an additional method to adjust tubulin dynamics (Zhang, Bailey *et al*. 2001).

Recently, the activation state of MAP1B/Futsch has been associated with wnt signaling (Marques 2005; Gogel, Wakefield *et al*. 2006). WNTS are known to induce axonal remodeling in developing neurons (Speese and Budnik 2007). In one study, overexpression of disheveled (DVL) in neuroblastoma cell lines induced growth cone enlargement and increased the size of tubulin bundles in the axon (Speese and Budnik 2007). These changes are characteristic of axonal remodeling. Using this same model system, the authors discovered that changes in tubulin dynamics are correlated with decreased GSK-3 activity(Speese and Budnik 2007). Transcriptional changes, however, do not mediate these changes in tubulin. Work in *Drosophila* suggests that active GSK-3 phosphorylates Futsch and modulates its activity (Gogel, Wakefield *et al*. 2006). Decreasing the levels of GSK-3 may alter tubulin dynamics by modulating the MAP1B phosphorylation state. These data lead to the speculation that MAP1B acts downstream of DVL to control tubulin dynamics at the developing synapse (Figure 1.3) (Marques 2005).

Note that mutations in Futsch decrease the number of synaptic arbors but increase bouton size at the NMJ (Roos, Hummel *et al*. 2000). This phenotype is the opposite of that observed in *highwire* mutants (Wan, DiAntonio *et al*. 2000). Highwire seems to restrain synaptic growth whereas futsch promotes it. The pathways utilized by these two

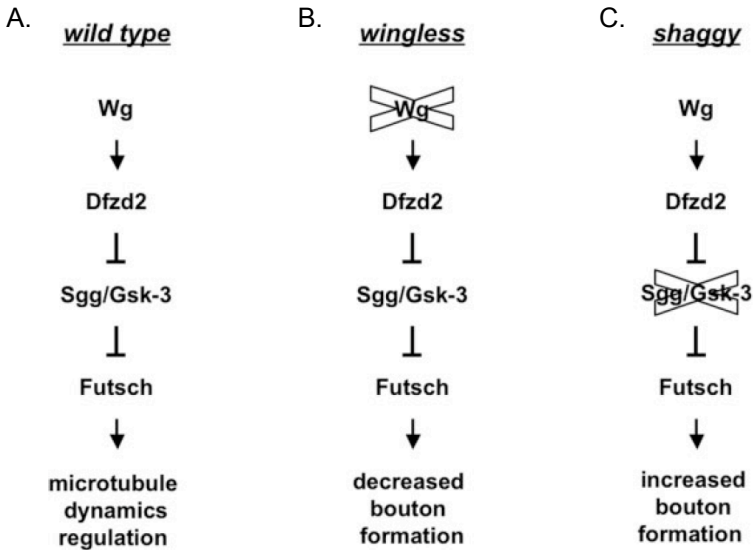


Figure 1.3 Model for the role of Wingless signaling in synaptic bouton formation and tubulin dynamics

A. In wild type animals, wingless (Wg) binds to the receptor Frizzled (Dfzd2), and regulates the activity of the kinase Gsk-3. The microtubule binding protein Futsch is dephosphorylated and regulates tubulin dynamics.

B. The loss of Wg relieves the inhibition of Gsk-3, and increasing Futsch inhibition. As a result, the number of synaptic boutons are decreased.

C. Mutations in Gsk-3 have the opposite effect. Futsch is now hyperactive, tubulin is more stable, and the number of synaptic boutons are increased.

(Figure taken from Marques 2005)

proteins may be shared. Hence, our discovery that RPM-1 regulates beta tubulin levels (Chapter 3) could provide a connection between these pathways. Further work is necessary, however, to develop a plausible model.

Regulation of Tubulin dynamics at the growth cone

PDZ domain containing proteins have also been implicated in regulating microtuble dynamics (Ruiz-Canada, Ashley *et al*. 2004). The Par-3/Par-6/aPKC complex was first characterized in *C. elegans* as a key regulator of embryonic polarity (Kemphues 2000). This finding led to the hypothesis that the Par genes could also affect the establishment of neural polarity. In mice hippocampal cultures, neurons extend one, and only one axon. Antibody staining of these axons revealed that mPar-6 and mPar-3 are selectively enriched in the growth cone (Shi, Jan *et al*. 2003). Ectopic expression of mPar-3 disrupts axon specification. Now, rather than elaborating one long process, the neuron extends multiple long processes from the cell body. Antibody staining of the axon specific microtubule Tau1 revealed the formation of multiple axon branches. In addition, disruption of aPKC activity in these cultured neurons eliminated axon outgrowth. These data suggest that the Par protein complex defines neural polarity.

The Par complex also plays a role at the *Drosophila* NMJ. aPKC is expressed in both pre- and post-synaptic cells at the NMJ and co-localizes with the presynaptic cytoskeleton (Ruiz-Canada, Ashley *et al*. 2004). Altering the activity of aPKC leads to a reduced number of synaptic boutons, similar to *futsch* mutants. Further characterization of an *apkc* loss-of-function mutant revealed that tubulin morphology is disrupted at actively growing terminal boutons. During NMJ expansion, tubulin within terminal

boutons generally adopts a splayed appearance. In *apkc* mutants, however, tubulin appears punctate and diffuse throughout the terminal. Additional work revealed that mutations in *par-3* (*bazooka*) or *par-6* also reduce the number of synaptic boutons and all three of these proteins partially colocalize at the synapse (Ruiz-Canada, Ashley *et al*. 2004). These data demonstrate a role for the Par complex in synaptic assembly.

The phenotype of *apkc* mutants resembles that of *futsch*. Co-immunoprecipitation experiments showed that antibodies to aPKC can immunoprecipitate both Futsch and Tubulin, suggesting that these proteins act in a complex (Ruiz-Canada, Ashley *et al*. 2004). Therefore, aPKC may act to modulate the interaction of futsch with tubulin to control synaptic structure as well as provide an additional link to tubulin dynamics and synapse formation.

aPKC also seems to play a critical role in long-term potentiation (LTP) in mammals and in persistent memory in flies (Drier, Tello *et al*. 2002; Ling, Benardo *et al*. 2002). In an odor avoidance paradigm, specific odors are associated with electrical shocks. Researchers discovered that activating an inducible form of aPKC in neurons after training enhances learning (Drier, Tello *et al*. 2002). These data suggest that synaptic plasticity depends on the Par signaling complex.

syg-1/syg-2

As shown above, both pre- and post-synaptic signals are necessary for synaptic formation (*e.g.* agrin, wnt, and BMP signaling). But how are synapses targeted to a specific cell type? Recent studies suggest that the mechanism of synapse formation may use guidepost molecules to select synaptic targets (Shen and Bargmann 2003; Shen,

Fetter *et al*. 2004). Work in *C. elegans* has identified two members of the immunoglobulin (IG) superfamily, *syg-1* and *syg-2*, which selectively target a synapse to its appropriate position. Mutations in *syg-1* and *syg-2* disrupt the placement of synapses from HSN (hermaphrodite specific neuron) in the developing egg-laying circuit (Shen and Bargmann 2003; Shen, Fetter *et al*. 2004) (Figure 1.4 A-H). Normally, HSN synapses are located in a region dorsal to the vulva. In *syg-1* and *syg-2* mutants, however, synapses are displaced anteriorly to the ventral nerve cord. Rescue experiments have shown that *syg-1* and *syg-2* are expressed exclusively in two different cell types. SYG-1 is expressed in the HSN neuron, whereas SYG-2 is expressed in the primary vulval epithelial cells. SYG-2 appears to act as a molecular guidepost to target synapses to adjacent cells. In other words, epithelial SYG-2 positions the HSN presynaptic apparatus in the correct location adjacent to motor neuron and muscle targets. As described below, recent experiments have linked the SYG-1/SYG-2 interaction to the localization of key proteins, SYD-1 and SYD-2, that nucleate the assembly of the presynaptic apparatus (Dai, Taru *et al*. 2006; Patel, Lehrman *et al*. 2006). These data provide a mechanistic link between a synaptic targeting molecule and synaptic assembly and suggest that a limited set of molecules (*syg-1* and *syg-2*) may be sufficient to induce the initial localization of the pre-synaptic apparatus.

syd-1/syd-2

A genetic screen for defects in GABAergic motor neuron synapses revealed mutations in *syd-1* and *syd-2* (Zhen and Jin 1999; Hallam, Goncharov *et al*. 2002). SYD-1 is a PDZ domain-containing protein and may function in scaffolding at the synapse;

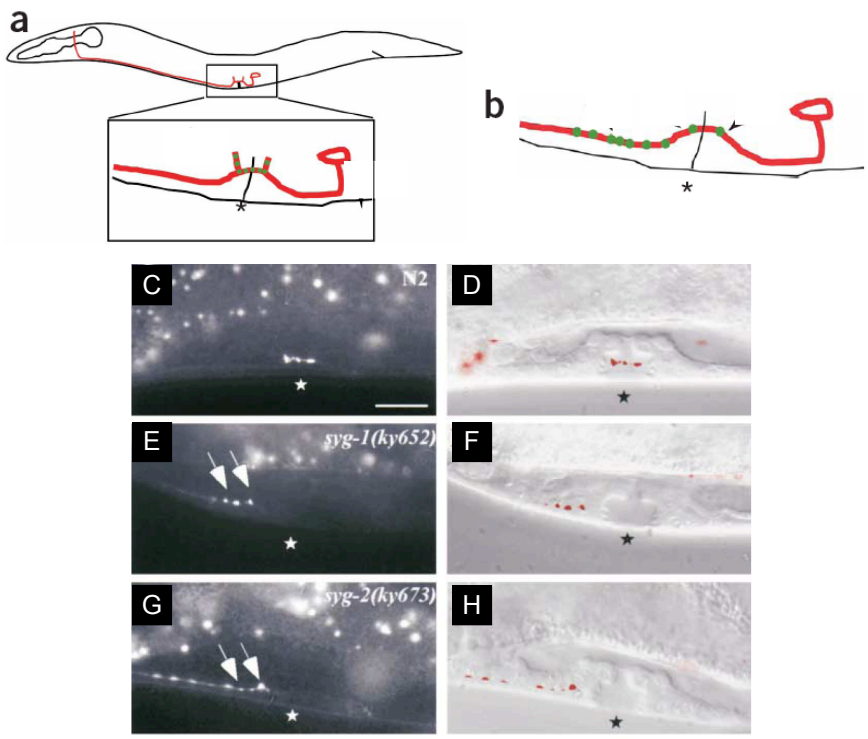


Figure 1.4 SYG-2 expression in vulval epithelium targets HSN synapses to adjacent cells.

A. Schematic of the worm HSN circuit. The HSN (red circle) sends a process to the anterior of the worm. Synapses are formed (green spots) with vulval muscle and VC motor neurons just dorsal to the developing vulva (*).

B. In syg-1 and syg-2 mutants, synapses shift anterior.

C and D. SNB-1 driven in HSN neurons of a wildtype animal.

E and F. SNB-1 is displaced anteriorly in syg-1 (ky652) mutant animals and in (G and H) syg-2 animals. This phenotype can only be rescued by expressing syg-2 in vulval epithelial cells, which do not form synapses with HSN.

(Figure modified from Shen, et al 2004 and Patel, et al 2006)

SYD-2 is a member of the liprin family which are known to regulate RPTPs (receptor proteins with tyrosine phosphatase activity) (Kaufmann, DeProto *et al*. 2002). Mutations in either *syd-1* or *syd-2* increase the size of the presynaptic specialization of GABAergic motor neurons. Work in *Drosophila* suggests that Dliprin-α, the *syd-2* homologue, also restrains active zone growth (Dunah, Hueske et al. 2005). In *C. elegans*, *syd-1*; *syd-2* double mutants phenocopy either single mutant alone, suggesting function in a common pathway (Hallam, Goncharov *et al*. 2002). Recent work has focused on the role of *syd-1* and *syd-2* in HSN synaptic development (Dai, Taru *et al*. 2006; Patel, Lehrman *et al*. 2006). In *syd* mutants, GFP-labeled synaptic proteins (e.g., G-protein RAB-3::GFP) are dispersed throughout the HSN process rather than localized to discrete puncta near the vulva as in the wildtype (Patel, Lehrman *et al*. 2006). Synaptic proteins that are not associated with synaptic vesicles, such as the active zone protein ELKS-1, are also mislocalized. These data suggest that both *syd-1* and *syd-2* are necessary for the recruitment and localization of most synaptic components to the developing synapse. How are *syd-1* and *syd-2* localized to the appropriate position in HSN? *syg-1* and *syg-2* are likely candidates for this upstream component, as mutants in either IG protein disrupt *syd-1* and *syd-2* localization (Patel, Lehrman *et al*. 2006). The upstream role of the SYG-1/SYG-2 interaction is solidified by the finding that *syg-1* and *syg-2* are localized normally in syd mutants (Patel, Lehrman *et al*. 2006). These results have led to the model that *syg-1* and *syg-2* restrain synaptic formation to a particular area within the developing HSN.

Genetic experiments have shown that gain-of-function mutations in *syd-2* can rescue *syd-1* synaptic defects (Dai, Taru *et al*. 2006). This result suggests that *syd-1* may

function as a positive regulator of *syd-2*. *syd-2* gain-of-function mutants can also induce the formation of synapses dorsal to the vulva in *syg-1* and *syg-2* mutants, similar to those seen in wildtype animals (Dai, Taru *et al*. 2006). It is hypothesized that *syd-2* gain-of-function mutations promote synapse formation by expanding "residual synapses" in these regions (Dai, Taru *et al*. 2006). However, many questions still remain. For example, how are synaptic components mislocalized in *syd* mutants? A potential clue is derived from the observation that the *syd-2* orthologue liprin-α binds to the kinesin protein UNC-104 (Miller, DeProto *et al*. 2005). Synaptic vesicle proteins normally bind to kinesin and travel down the axon before being deposited at synaptic sites. In *syd* mutants, synaptic vesicle proteins are diffusely localized (Figure 1.5 C, D). One models suggests that liprin at the site of the developing synapse unloads synaptic vesicles from kinesin (Patel, Lehrman *et al*. 2006).

Interestingly, synaptic formation in HSN is unaltered in *rpm-1* mutants (Patel, Lehrman *et al*. 2006). This result suggests that not all synapses are "created equal" and may indicate that the assembly of distinct synapses may be specified by mechanisms involving different sets of molecules.

ELKS/RIM/Bruchpilot

Further work has focused on components downstream of *syd-1* and *syd-2*. ELKS and RIM are active zone proteins that have been shown to affect mouse synapses (Ohtsuka, Takao-Rikitsu *et al*. 2002; Powell, Schoch *et al*. 2004). ELKS encodes a glutamate, leucine, lysine, and serine-rich protein, and is homologous to Bruchpilot.

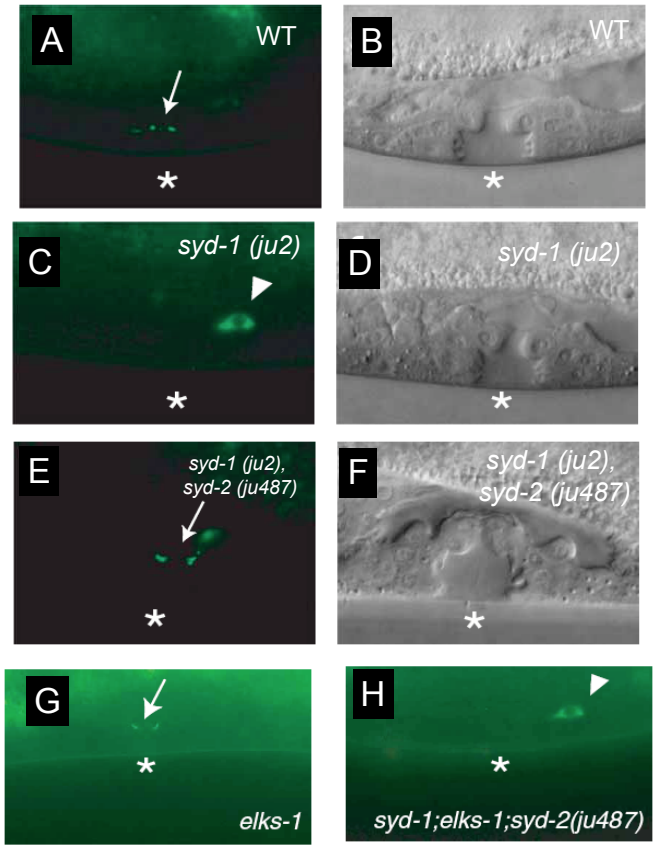


Figure 1.5 Gain of function mutations in *syd-2 (ju487)* rescue *syd-1* HSN synaptic defects

A, B *unc-86*::SNB-1::YFP is expressed in HSN and localizes dorsal to the vulva.

C, D Mutations in the PDZ domain containing protein syd-1 disrupt synaptic localization in HSN.

E, F A gain of function mutation in *syd-2 (ju487*) rescues *syd-1* HSN synaptic defects.

G. Mutations in elks-1 do not disrupt HSN synaptic placement. However, mutations in *elks-1* prevent the *syd-2 (ju487)* dependent rescue of *syd-1* synaptic defects.

Arrow = synapses Arrowhead = HSN cell body *=vulva

Modified from Dai, et al 2006

RIM encodes a Rab3a-interacting molecule. Mutations in *C. elegans* RIM (*unc-10*) disrupt synaptic physiology and behavior, and share some of the defects seen in mouse RIM knockouts (Koushika, Richmond *et al*. 2001). RIM is believed to mediate synaptic vesicle fusion. ELKS interacts with RIM *in vitro* and may mediate . However, in *C. elegans*, null alleles of ELKS fail to show any synaptic defects (Koushika, Richmond *et al*. 2001). Recent evidence suggests that a redundant pathway masks an ELKS synaptic phenotype. As mentioned above, in a *syd-2* gain-of-function (gf) mutant rescues *syd-1*(0) such that synaptic components are properly localized near the vulval epithelium in HSN (Dai, Taru *et al*. 2006) (Figure 1.5 E,F). In contrast, in a *syd-1*; *syd-2* (gf); *elks-1* triple mutant, synaptic components are again mislocalized (Dai, Taru *et al*. 2006) (Figure 1.5 G,H). SYD-2 may therefore be necessary to localize ELKS-1 to the developing synapse. SYD-2 and ELKS-1 could then function, in conjunction with SYD-1 to, recruit synaptic vesicles and other synaptic components (Dai, Taru *et al*. 2006). Dai *et al* observed that ELKS shares homology with the N terminus of Bruchpilot, which regulates the formation of the electron-dense T-bar at the *Drosophila* synapse (Kittel, Wichmann *et al*. 2006; Wagh, Rasse *et al*. 2006). As mentioned above, a single mutation in *elks-1* does not produce a synaptic defect. Rather, *elks-1* seems to act in conjunction with *syd-2*. This result suggests that Bruchpilot may function in the place of multiple proteins during synaptic development.

Synaptic Plasticity

Many of the synaptic genes we have examined (e.g. *rpm-1*/Highwire, *syd-1* and *syd-2*, wnt, TGF-β, etc.) have been shown to function early in the life of the organism,

during the initial stages of synaptic development. However, in most organisms, new synaptic connections are generated throughout life. During learning and memory, synaptic connections can be either created or modified. Extensive work on this aspect of synaptic development has been performed in the Kandel lab on the giant marine snail *Aplysia* (Kandel 2004). *Aplysia* exhibit three types of behavioral learning responses: habituation, sensitization, and classical conditioning. The Kandel lab was able to describe, at the molecular level, the basis for many of these responses. In response to a tactile stimulus, *Aplysia* can withdraw the respiratory organ into the body cavity. Tail shocks given prior to the tactile stimulus can sensitize the animal and increase the strength of the withdrawal response. Multiple shocks to the tail in conjunction with tactile stimulus can sensitize the animal and produce a larger withdrawal even several days after the stimulus. At the synapse, calcium influx sensitizes the neuron to further stimuli, increasing the chances for an action potential. Multiple shocks lead to the induction of new protein synthesis and eventually to the formation of new synaptic connections. Further work has shown that this process requires a variety of molecular components, including the transcription factor CREB. Notably, the synthesis of new synapses during learning and memory is also dependent on the activation of the MAP Kinase cascade and proteasomal degradation (Kandel 2004). While the homologue of *rpm-1* in *Aplysia* has yet to be characterized, these data, combined with the conservation of *rpm-1* from worms to humans, suggest a common mechanism for the generation of synapses that is likely to be conserved. The downstream transcriptional targets that drive synaptic remodeling remain a mystery. The identification of these targets may elucidate the mechanisms of long term potentiation.

Transcriptional control of Synaptic Choice

Transcription can also influence synaptic choice. The motor neuron circuit of *C. elegans* provides one such example. VA and VB motor neurons arise from a common precursor but are connected to different presynaptic partners: AVA interneurons synapse with VA motor neurons to mediate backward movement, whereas AVB interneurons synapse with VB motor neurons to mediate forward locomotion (White, Southgate *et al*. 1986). The homeodomain transcription factor UNC-4 regulates the formation of this circuit (Miller, Shen *et al*. 1992; White, Southgate *et al*. 1992). In *unc-4* mutants, AVB interneurons, which normally synapse selectively with VB motor neurons, now also make gap junctions with VA motor neurons; AVA inputs to VA motor neurons are lost. As a result of this wiring defect, *unc-4* mutants are unable to execute backward locomotion. UNC-4 function depends on interaction with the transcriptional co-repressor UNC-37 (Groucho) (Miller, Niemeyer *et al*. 1993). These findings suggest a model in which UNC-4 functions in VA motor neurons to preserve VA type inputs by inhibiting expression of VB-specific genes.

Cell-specific microarray profiling experiments (see below) identified a list of candidate *unc-4* regulated transcripts and recent work has established that at least one gene, the homeodomain transcription factor, CEH-12, functions downstream of *unc-4* to regulate synaptic choice (Von Stetina, Fox *et al*. 2007). *ceh-12* encodes the nematode homolog of HB9, a well-established determinant of motor neuron fate in flies, birds and mammals (Lee and Pfaff 2001; Odden, Holbrook et al. 2002). Consistent with the proposed model of *unc-4* action, *ceh-12* is normally restricted to VB motor neurons in the wildtype but is also expressed in VAs in *unc-4* and *unc-37* mutants. The results of

additional genetic experiments (*i.e*. *ceh-12* mutants partially suppress Unc-4) suggest that *ceh-12* is de-repressed in parallel to at least one additional pathway to induce the Unc-4 miswiring phenotype. An RNAi screen of the microarray data set of candidate *unc-4* target genes is currently underway in the Miller lab to detect these downstream *unc-4* pathway genes. Ultimately, this work should identify a comprehensive set of effector genes that control the specificity and placement of synapses in this circuit.

A microarray strategy to identify RPM-1-regulated genes

This introduction emphasizes that transcription plays a key role in synaptogenesis, and that several different pathways influence synaptogenesis through a transcriptional mechanism. For instance, Wnt signaling stimulates beta-catenin dependent synapse generation in mossy fiber neurons (Speese and Budnik 2007). Likewise, BMP released from post-synaptic targets activate the transcription factors Mad and Medea to sculpt the presynaptic membrane (McCabe, Marques *et al*. 2003; McCabe, Hom *et al*. 2004). An Agrin/MuSK interaction mediates the transcriptional activation of acetylcholine receptors in myonuclei (Kummer, Misgeld *et al*. 2006). Long term potentiation is dependent on the activation of the transcription factor CREB (Kandel 2004). Furthermore, the *unc-4* pathway provides a signal that directs a presynaptic interneuron to the appropriate synaptic region (White, Southgate *et al*. 1992; Miller and Niemeyer 1995). Transcription plays a vital role in all of these processes. Together, my findings and recent work by the DiAntonio lab suggest that RPM-1/Highwire controls transcription and that this regulation is necessary for the development of normal synaptic structures (Collins, Wairkar *et al*. 2006). In order to understand this mechanism, it will be necessary to

identify the downstream genes that RPM-1 controls. To achieve this goal, I collaborated with Stephen Von Stetina in the Miller lab to optimize a powerful microarray-based strategy, the mRNA-tagging method, to profile gene expression in the *C. elegans* nervous system. This approach has now established a data set of ~700 candidate *rpm-1*-regulated genes (Chapter 3). In addition to confirming the hypothesis that *rpm-1* controls transcript levels in *C. elegans* neurons, the mRNA-tagging method has also provided a comprehensive description of native gene expression throughout the *C. elegans* nervous system (Chapter 4). Moreover, our successful use of mRNA-tagging to profile a specific subset of neurons (i.e. larval A-type motor neurons) demonstrates the utility of this approach for defining gene expression in *C. elegans* at the resolution of a single neuron type. In this section, I set the stage for this work by discussing the results of earlier microarray profiling experiments in *C. elegans*. In addition, I review methods of RNA amplification, a key technical requirement necessary for these experiments. Using these methods, I demonstrate in Chapter 5 the applicability of a new method of RNA amplification to microarray profiling studies of the *C. elegans* nervous system.

Microarray analysis of global gene expression in *C. elegans*.

Microarray profiling offers a powerful strategy for monitoring changes in transcript levels across the entire genome. Two types of platforms have been used for microarray studies in *C. elegans*; spotted arrays and Affymetrix arrays (Dalma-Weiszhausz, Warrington *et al*. 2006). The spotted arrays used to study *C. elegans* are generated by a single lab (Wixon, Blaxter *et al*. 2000; Kim, Lund *et al*. 2001; Stuart, Segal *et al*. 2003). These arrays contain fragmented exon rich DNA representing a large

number of genes (~12,000 to ~20,000, depending on the experiment). Because these arrays are generated in a lab, quality can vary greatly. However, the hybridization of two differentially labeled cDNA samples to a single array reduces much of this variation. Affymetrix arrays, on the other hand, are one-color platforms that depend on high quality standards to minimize variability between arrays (Dalma-Weiszhausz, Warrington *et al*. 2006). *C. elegans* Affymetrix GeneChips are composed of 25 nucleotide oligos, that represent ~20,000 unique transcripts. One-color arrays provide a signficant advanatage over spotted arrays since data generated for a multitude of studies can be compared across chips. Using either chip, sample RNA is converted into a labeled single stranded target (either cRNA or cDNA) for hybridization (Dalma-Weiszhausz, Warrington *et al*. 2006). Initially, *C. elegans* researches exclusively employed spotted arrays. In these early studies, microarray profiling experiments were performed using RNA extracted from the entire animal. For example, RNA was isolated from a variety of different backgrounds in order to assign functions to previously uncharacterized genes. These studies included a large number of transcripts isolated from various mutants (mutations in dauer formation, aging, germline, neural development, etc), as well as profiles generated for specific time points during *C. elegans* development. In all, gene expression data was compiled for 553 experimental conditions. A retrospective study derived from these data sought to establish patterns of co-regulated gene expression (Kim, Lund *et al*. 2001; Stuart, Segal *et al*. 2003). These experiments yielded a vast network of data that was subsequently analyzed on a topographic map. This topographic map was based on a correlation matrix – for example, genes that changed expression in relation to mutations that affected the germline would be grouped into certain mountains within the

topographic map. These experiments resulted in 43 independent "mountains," representing a variety of functions within the worm, and helped to categorize genes with no previous functions (Kim, Lund *et al*. 2001; Stuart, Segal *et al*. 2003).

Analysis of the individual data sets produced from these profiles resulted in a number of significant advances. The germline studies mentioned above produced some of the most relevant findings. Reinke, *et al*, used whole worm microarray experiments to profile wildtype animals versus glp-4 mutants, in which the germline cells fail to proliferate (Reinke, Smith *et al*. 2000). This dataset contained many genes with unknown functions in germline development. Miller, *et al*, used this dataset to identify one of the signaling components that modulates oocyte maturation, the Ephrin receptor VAB-1 (Miller, Ruest *et al*. 2003). The success of this strategy suggests that novel, biologically relevant pathway components can be identified in microarray profiles.

Stephen Von Stetina in the Miller lab initially attempted to identify the transcriptional targets of UNC-4 by comparing microarray profiles of wildtype vs *unc-4* or *unc-37* mutant animals. Total RNA was extracted from a synchronized population of L2 animals, reverse transcribed, and hybridized to spotted arrays. This approach, however, failed to identify *unc-4*-regulated target genes, because *unc-4* regulates gene expression in only a small subset of larval cells (~2%) whereas RNA for these experiments was extracted from the entire animal. Von Stetina concluded that "whole worm" microarray profiles are insufficiently sensitive to detect *unc-4* regulated transcripts (SEV and DMM, unpublished data). A new approach was needed that could identify transcripts that function in a small number of cells within the worm. A cell specific profile could be sensitive enough to detect transcipts with unique roles in specific

cell types, and can even be used to identify targets of transcription factors that control key developmental decisions. Recent work from a number of labs has resulted in technology that enhances the resolution of microarray profiling (Roy, Stuart *et al*. 2002; Zhang, Ma *et al*. 2002; Colosimo, Brown *et al*. 2004; Cinar, Keles *et al*. 2005; Fox, Von Stetina *et al*. 2005; Kunitomo, Uesugi *et al*. 2005; Pauli, Liu *et al*. 2005; Touroutine, Fox *et al*. 2005; Yang, Edenberg *et al*. 2005; Von Stetina, Fox *et al*. 2007). Some of these experiments have used the Affymetrix system and provide data verifying the utility of that platform for *C. elegans* researchers.

Methods for microarray profiling of embryonic *C. elegans* cells

The development of *C. elegans* cell culture technique facilitated the acquisition of cell-specific microarray profiles (Christensen, Estevez et al. 2002). Embryos can be readily dissociated and the resultant blastomeres maintained in primary culture where they differentiate *in vitro*. The expression of GFP reporters in these cultures has allowed the use of FACS to isolate specific cell types for microarray profiling. The Miller lab used this method, for example, to generate a comprehensize gene expression profile of embryonic A-type motor neurons and coined the term MAPCeL (Micro-Array Profiling of *C. elegans* Cells) to describe this approach(Fox, Von Stetina *et al*. 2005). The MAPCeL strategy has now been used to profile gene expression in various classes of neurons and muscle cells (Zhang, Ma *et al*. 2002; Colosimo, Brown *et al*. 2004; Cinar, Keles *et al*. 2005; Fox, Von Stetina *et al*. 2005; Touroutine, Fox *et al*. 2005; Von Stetina, Fox *et al*. 2007).

While profiles of sorted cells have provided robust characterizations of embryonic transcriptomes, these profiles are limited to embryonic cell types since Post-embryonic cells fail to diffentiate in culture. In addition, this method likely disrupts cell signaling pathways necessary for some aspects of differentiation. Roy, *et al*, developed mRNA tagging to circumvent this limitation (Roy, Stuart *et al*. 2002). mRNA-tagging exploits specific binding of the Poly-A binding protein (PAB) to the 3' polyadenylated tail of mRNA. In this method, an epitope-tagged PAB is expressed using a cell specific promoter. Formaldehyde fixation cross-links PAB to poly-A mRNA. Worms are lysed, and the epitope-tagged PAB bound mRNA is isolated by immuno-precipitation wth anti-FLAG beads. This approach was initially utilized to profile gene expression in larval muscle cells. Profiles of ciliated neurons and intestine have also been obtained by mRNA-tagging (Kunitomo, Uesugi *et al*. 2005; Pauli, Liu *et al*. 2005). Additional work in the fly has identified the transcriptomes of the *Drosophila* photoreceptor cells (Yang, Edenberg *et al*. 2005). We have used this approach to identify cell specific profiles of larval A-class motor neurons and the entire *C. elegans* nervous system (Von Stetina, Watson *et al*. 2007) (see chapter 4). These profiles have been particularly robust, as we have identified ~1700 neuron-enriched transcripts. We have confirmed these results using GFP reporters. We have used these data to identify a novel conserved gene expressed exclusively in the nervous system (F29G6.2). We have also shown the enrichment of a broad range of neuropeptides in the larval nervous system. In contrast, using the larval A-class profiles, we have confirmed enrichment of select neuropeptides. Additionally, we have shown that these data can be used to identify genes expressed exclusively in a single neural type.

Our lab has employed this strategy to identify bona fide targets of the transcription factor *unc-4* (Von Stetina, Fox *et al*. 2007). In these experiments, wildtype larval A-class neurons were compared to *unc-4* mutant A-class neurons. This comparison resulted in the idenification of the transcription factor *ceh-12*, which was then confirmed, *in vivo*, to be an UNC-4-regulated transcript and a part of the pathway that controls synaptic choice (Von Stetina, Fox et al. 2007). I have used a similar strategy to identify transcriptional targets of the E3 Ubiquitin Ligase RPM-1 (see chapter 3).

RNA amplification methods to identify neural transcripts

mRNA quantity is not a limiting factor in whole worm profiling experiments. The ten ug of poly A RNA necessary for hybridization can be obtained from large batches of worms. RNA amplification is needed, however, for profiles of single cell types. Three techniques have been used in *C. elegans* to overcome this limitation. The first, and most laborious method, is the use of brute force to generate enough RNA from an mRNA-tagging line to hybridize to an array. This method eliminates any bias that could be generated through amplification of isolated RNA. However, the amount of time spent harvesting the samples necessary for amplification makes this method impractical. The second method, which has been used in most profiling experiments in *C. elegans*, has used a modified version of Eberwine, a two-round *in vitro* transcription (IVT) method (Nygaard and Hovig 2006). In this method, mRNA is reverse transcribed into sscDNA using a poly-A T7 primer, a second strand is transcribed using DNA polymerase, and then an RNA polymerase generates aRNA from the dscDNA *in vitro*. A second round of

amplification is performed, but, rather than T7 primers, random primers are used for this round. Afterwards, the resulting material is biotinylated, fragmented, and then hybridized to Affymetrix arrays. This method has generated robust results, (see microarray experiments described above) and is much more affordable than SAGE analysis. However, the Eberwine method is also time consuming and requires a considerable amount of starting material (25 ng).

A newer method, called the WT-Pico Ovation method, uses a different strategy to amplify RNA (Singh, Maganti *et al*. 2005). In this method, RNA is reverse transcribed into cDNA using reverse transcriptase and a DNA/RNA chimeric primer. RNA in the resulting dimer is heat fragmented. DNA polymerase then uses these fragments to generate a second strand of cDNA. In the next step, called Ribo-SPIA, RNase H, chimeric DNA/RNA primers, and an additional DNA polymerase are added to the mixture. RNase H specifically degrades duplex RNA, so the original DNA/RNA primer that is still attached to the dscDNA is degraded, opening up the cDNA and allowing the binding of a new DNA/RNA primer. DNA polymerase then synthesizes a new second strand, and the process cycles. This method results in robust amplifications, and scatter plots from amplified Mouse kidney and Mouse universal reference RNA show significant correlation between the Eberwine method and the WT-Pico (Singh, Maganti *et al*. 2005). This method is especially promising for *C. elegans* researchers, since WT-Pico amplifications take a quarter of the time of traditional Eberwine amplifications (Singh, Maganti *et al*. 2005). In addition, in our hands, WT-PICO amplification is more robust than IVT (i.e. fewer failed amplification reactions) (Watson *et al*, unpublished data) (see Chapter 5). We have used Eberwine and WT-Pico to generate an additional profile of the

entire nervous system. Our data indicate that, while both methods identify many of the same transcripts, a subset of genes are identified by only one or the other method. We have confirmed neural expression of these transcript using GFP reporters. Using microarray technology, we have generated a robust profile of the *C. elegans* nervous system. In addition, we have discovered genes regulated transcriptionally throughout the nervous system by the E3 ubiquitin ligase RPM-1. We believe that novel targets necessary for synaptic function are contained within the RPM-1-regulated gene list. Further characterization of the Rpm-1 phenotype and its transcriptional targets should enhance our understanding of the *rpm-1* pathway and synaptic development.

CHAPTER II

THE PRESYNAPTIC PROTEIN RPM-1 REGULATES NEURONAL GENE EXPRESSION

Introduction

My work originally focused on identifying genes that regulate motor neuron fate. In *C. elegans*, motor neurons are divided into distinct subclasses, based on morphology and function. For example, B-type motor neurons adopt posteriorly directed axonal processes and are necessary for coordinated forward movement (White, Southgate *et al*. 1976). The nicotinic acetylcholine receptor *acr-5* is expressed exclusively in B-type motor neurons in the ventral nerve cord (Figure 2.1A), but not in A-type motor neurons, which extend anterior axons and control backward locomotion (Winnier, Meir *et al*. 1999). Mutations in the homeodomain transcription factor *unc-4* lead to ectopic expression of *acr-5* in A-type motor neurons (Figure 2.1B) (Winnier, Meir *et al*. 1999). *unc-4* mutant animals are unable to move backward (Brenner 1974). Although this movement defect does not depend on *acr-5*, ectopic expression of *acr-5*::GFP in A-class motor neurons is a reliable indicator of defects in *unc-4* pathway genes (Winnier, Meir *et al*. 1999). Mutations in the COE transcription factor *unc-3* also affect *acr-5* expression but, in contrast to *unc-4*, *unc-3* mutant animals lose expression of *acr-5* in specific ventral nerve cord motor neurons (Figure 2.1C) (K. Lickteig, J. Meir, D. Miller, personal communication). We speculated that screening for mutants that disrupt expression of *acr-5* could identify additional members of the *unc-3* or *unc-4* pathways, and could also

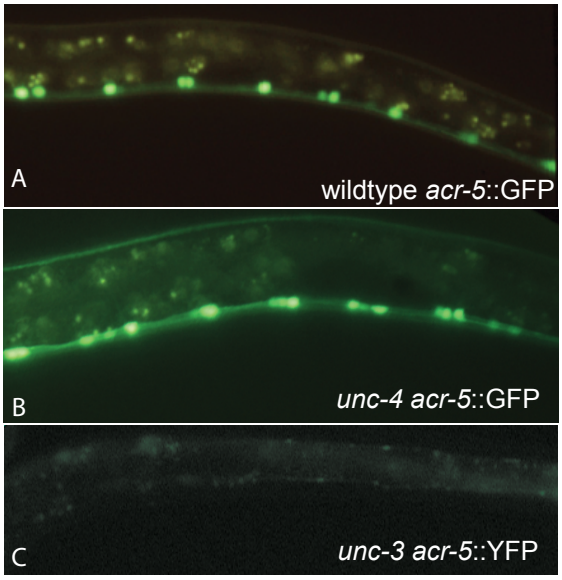


Figure 2.1 *acr-5*::GFP is negatively regulated by UNC-4 and positively regulated by UNC-3.
acr-5::GFP is expressed in VB and DB motor neurons in wildtype animals (A). Mutations in the transcription factor *unc-4* result in the ectopic expression of *acr-5*::GFP in VA and DA motor neurons (B). In contrast, mutations in the transcription factor *unc-3* eliminate expression of *acr-5*::YFP in motor neurons (C).

potentially reveal other uncharacterized pathways governing motor neuron differentiation. To that end, I performed a forward genetic screen to identify mutations that perturb expression of *acr-5*::YFP. This work revealed several mutations that **reduce** *acr-5*::YFP expression in the ventral nerve cord. Three of these strains include alleles of the synaptic gene, *rpm-1*.

rpm-1 encodes an E3 ubiquitin ligase that is required for normal presynaptic assembly (Schaefer, Hadwiger *et al*. 2000; Zhen, Huang *et al*. 2000). *rpm-1* was originally identified in a genetic screen in which GFP-tagged synaptobrevin-1 (SNB-1::GFP) marked pre-synaptic densities in GABAergic motor neurons (also independently discovered by in Mike Nonet's lab using SNB-1::GFP in touch neurons) (Schaefer, Hadwiger *et al*. 2000; Zhen, Huang *et al*. 2000). In wildtype animals, evenly spaced, punctate SNB-1::GFP clusters are clearly visible along the dorsal and ventral nerve cords (Fig 2.7A) (Jin, Jorgensen *et al*. 1999). In *rpm-1* mutants, the pattern and frequency of these clusters are altered (Fig 2.7B) (Zhen, Huang *et al*. 2000). The number of GFP clusters is significantly reduced, GFP puncta are unevenly distributed and frequently coalesce to produce large 'globs' of GFP in both the ventral and dorsal cord. Ultrastructural analysis in the electron microscope has revealed the nature of these 'globs.' Normally, in wildtype animals, GABAergic synapses at neuromuscular junctions contain one electron-dense presynaptic active zone. In contrast, *rpm-1* mutant synapses typically show either partially formed active zones, or "overdeveloped" active zones that contain more than one presynaptic density. Antibody staining revealed that RPM-1 is localized near the synapse, but not within the active zone. Rather, RPM-1 resides in an adjacent region named the "periactive" region. These data have been

interpreted to mean that that RPM-1 acts locally to control synaptic structure (Chang and Balice-Gordon 2000; Schaefer, Hadwiger *et al*. 2000; Zhen, Huang *et al*. 2000; Liao, Hung *et al*. 2004). RPM-1 is a very large protein (3766 a.a.) with several highly conserved domains, including an RCC1 GEF domain, two Pam/Highwire/RPM-1 domains, and a RING finger domain. Recent experiments have shown that RPM-1 ubiquitinates substrates *in vitro*, suggesting a role in protein degradation (Nakata, Abrams *et al*. 2005). The suggested targets are members of the MAP Kinase signaling cascade, which are ubiquitinated by RPM-1 *in vitro*. My work now suggests that RPM-1 may also regulate synaptic morphology by controlling gene expression downstream of this pathway (see chapter 3).

Materials and Methods

Nematode strains

Nematodes were grown as described (Brenner 1974). An integrated strain (NC574) expressing *acr-5*::YFP was generated by Christina Gleason in the Miller lab. NC574 was selected for the motor neuron mutant screen, due to the intensity and the high penetrance (approaching 100%) of YFP expression in B-class motor neurons. These animals were backcrossed 5x against the canonical laboratory strain, N2 (Bristol) to remove potential background mutations. All experiments were performed at 20° C (NC574 grows poorly at higher temperatures). Strains isolated from the mutagenesis are listed in Table 2.2.

Mutagenesis

NC574 was grown to confluence on four 100 mm NGM + OP50 plates. The worms were harvested and the embryos were collected using a standard bleach/NaOH protocol. Embryos were placed at 20° in M9 overnight (~16 hours) on a nutator and allowed to hatch. Synchronized L1 stage larvae were spread over four 100 mm NGM + OP50 plates, then grown to the L4 stage (approximately 50-60 hours for this strain). The worms were washed and then mutagenized with ethane methyl sulfonate (EMS) (.1M) for 4 hours. Mutagenized worms were washed in M9 buffer and then distributed to two 100 mm NGM plates. After an hour, viable adult hermaphrodites (~200) were picked to 60 mm plates (5 worms/plate) and grown at 20°C. Subsequently, 500 F1 progeny were transferred to individual 60 mm plates over the next 2-3 days in order to distribute screening of F2 progeny over a period of several days. A minimum of 20 F2 animals were screened on each plate for altered *acr-5*::YFP expression. This "F2 clonal" screen was performed four times with a total of ~2300 F1s.

Automated isolation of mutants affecting acr-5::YFP expression in the Union Biometrica Biosort

Worms were synchronized and mutagenized as described above. After the mutagenesis, P0 were allowed to self, and F1s were collected in M9. F1 hermaphrodite progeny (~100/plate) were placed onto 66 100mM plates and allowed to self. The F2 progeny of these F1s (between 5500 and 6000) were examined for changes in *acr-5*::YFP using a Union Biometrica COPAS Biosort following the manufacturer's instructions. Mutants with changes in gene expression were picked to individual plates and allowed to self.

Total RNA extraction from mixed stage larvae

Worms were grown to confluence at 20°C on four 100 mm NGM + OP50 plates, washed 3x with M9, and then floated on sucrose to remove bacteria. A Pasteur pipette with the tip removed (diamond tipped pen) was used to collect purified worms. After 3 washes in M9, worms were resuspended in 1.5 mL of M9 and flash-frozen in liquid nitrogen. The size of the worm pellet was estimated. The frozen worm pellet was homogenized by grinding with a baked mortar and pestle under liquid nitrogen. 1 mL of Trizol was added for every 100 μL of frozen worms. mRNA was purified using a Trizol/phenol chloroform extraction and precipitated with isopropanol. The pellet was resuspended in DEPC-treated water and stored at -80C.

mRNA extraction

A Fasttrack RNA purification kit from Invitrogen was used to isolate mRNA from a mixed population of frozen worms following the manufacturer's instructions. 5 μL of the resulting solution was run on a formaldehyde gel to check RNA integrity and mRNA enrichment.

RT-PCR

1.6 μL of purified mRNA was reverse transcribed using Promega enzymes (4 μL of 5x M-MLV reaction buffer, 2 μL of 10μM dNTP, 1 μL 40 units/uL RNasin, 2 μL of 50 μM random primers, 8 μL of depc H20, 2μL of 0.5-1 μg/μL mRNA, and 200 units/μL Promega RT enzyme, program RT-PCR, 42° for 30 minutes, 94° for 4 minutes).

Sequencing reactions

unc-4 cDNA was PCR-amplified for sequencing using the forward primer U4P9 (gatcggtgcactgcatgc) and the reverse primer U4P10 (gtcgacggatcctcagcaaccgtagtcaatgc), program A79 (94° 30 s, 55° 45 s, 72° 1 min, 35 cycles). Amplified cDNA was gel purified using the Biorad Freeze 'n' Squeeze kit protocol. An ethanol precipitation was performed to clean the final product. Big-dye reagents were used to label 70 ng of purified DNA using the U4P9 and U4P10 primers, and labeled using the Seq program at 96° for 30 sec, 50° for 15 sec, and 60° for 4 min, 30 cycles. Labeled DNA was precipitated (74 μL of 70% EtOH + 0.5 mM MgCl2 at -20° C), incubated at room temp in the dark for 10 min, and then spun at 12000 X g for 30 min. The pellet was washed in 80% EtOH and spun for 5 minutes at 4°C. The 80% Ethanol was removed. The pellet was then dried with a Speed-Vac and submitted for sequencing.

Mapping and strains

Mapping experiments were performed using a combination of Snip-SNP, 2- and 3-factor, and deficiency mapping (See Below). NC682 (*wd54*) was originally isolated as a double mutant. Outcrossing with N2 led to the isolation of NC742 (*wd72*) (see Results). Snip-SNP mapping was performed using CB4856, the Hawaiian strain of *C. elegans*. Males isolated from this population were mated into NC742 (*wd72*). F1s from these matings were picked to individual plates and allowed to self. Animals with the characteristics of NC742 (*wd72*) (i.e., reduced *acr-5*::YFP expression in VB motor neurons) were selected and genotyped. Linkage to a Bristol SNP indicated chromosomal linkage. Two sets of

primers covering SNPs on chromosome V were used for Snip-SNP mapping. One SNP, found in cosmid F32D8, is positioned at +3 map units on the right arm of Chromosome V, 5' primer GTCGCACCTTTTGCTCAATC, 3' primer TCGAAAATTGCCCTCCCTAC. The other SNP, found in Y51A2D, is positioned at +17 on the right arm of Chromosome V, 5' primer CAGGCATATTACATGGGATAGG, 3' primer CAATCTCACCTCCATTCTGTG. 2-factor mapping was performed with *unc-42* and *dpy-11*. Heterozygote animals of either *unc-42* (e270) or *dpy-11* (e224) were mated into NC742 (*wd72*). Unc-42 and Dpy-11 animals were scored for the NC742 (*wd72*) phenotype. A double mutant in *dpy-11* and *unc-42* (DR108) was used for 3-factor mapping. Heterozygous *dpy-11 unc-42* males were mated into homozygous NC742 (*wd72*), *acr-5*::YFP doubles. F1s were selected and allowed to self. Plates that contained Dpy Uncs were scored for recombination frequency against NC742 (*wd72*). Deficiency mapping was performed with several lines spanning the region between *dpy-11* and *unc-42*. Strains utilized were BC2617 *dpy-18*(e364)/eT1 III; mDf1/eT1 V, BC3954 (sDf71) *dpy-18*(e364)/eT1 III; *unc-4*6(e177) sDf71/eT1[let-500(s2165)]V, BW219 nDf31/*unc-23*(e25) *sma-1*(e30)V, and CB2619 eDf1 eDp21/*dpy-11*(e224)V. *rpm-1* (*ok364*) was obtained from the *Caenorhabditis* Genetics Center (CGC, from Theresa Stiernegle**.**). Bulk Segregant Analysis (BSA) (Wicks, Yeh *et al*. 2001) and the analysis of SNPs from individual recombinant larvae placed *wd72* near the center of chromosome V. Animals were scored for the presence of ectopic *unc-129*::GFP. These results were consistent with linkage data factor with visible markers on chromosome V, *dpy-11* (0.00), *unc-42* (+2.16), and *unc-60* (-18.79) (Fig 2.6C). *wd72* showed tight linkage to *dpy-11* (0 out of 57 F3s, $\leq$2 map units) and *unc-42* (1 out of 37 F3s or ~3 map

units), and was weakly linked to *unc-60* (7 out of 24 F3s or ~30 map units). Three-factor mapping with *dpy-11* and *unc-42* placed NC742 (*wd72*) to the right of *dpy-11* (5 of 5 *dpy-11* recombinants were *wd72*); the failure to detect recombinants with *unc-42* (0/6) placed *wd72* either closely linked (~ 1/7 of map distance between *dpy-11* and *unc-42* = 0.3 map units) and on the left of *unc-42* or to the right of *unc-42*.

Non-complementation tests vs deficiencies in the *unc-42* region were used to refine the genetic interval containing the *wd72* mutation (Fig 2.6C). One large deficiency, mDf1, spanned +1.64 to +2.44, covering both the left and right side of *unc-42*. *wd72* failed to complement this region (Fig 2.6B). Deletions spanning +1.41 to +1.88 and +1.41 to +2.20 also failed to complement *wd72*. Conversely, deletions to the right of *unc-42* (nDf31 and eDf1 spanning +2.04 to +2.29 and +2.13 to +2.29, respectively) complemented the *wd72* mutant phenotype (i.e. the expression of *unc-129*::GFP was altered). These deficiency mapping data indicated that *wd72* is located between +1.64 and +1.88 on chromosome V.

Single-worm PCR and Bulk Segregate Analysis

Single-worm PCR was used to prepare animals for SNP genotyping per standard amplification protocol. 1μL of each single worm solution was combined into a single solution for bulk segregate analysis. 2 μL of bulk solution was added to 23 μL of PCR mix (1X, 16 μL dH20, 2.5 μL Promega Buffer B, 1.5 μL Promega 25mM $MgCl_2$, 0.5 μL 10 mM dNTP mix, 0.5 μL Promega Taq, 2 μL primer mix 12.5 μM each) in 0.5 μL tubes. Primers used for SNP mapping on chromosome IV are described above. Samples were amplified with program A79 (94° 30 s, 55° 45 s, 72° 1 min, 35 cycles). After bulk

segregate analysis, individual frozen single worm lysates were genotyped. 0.2 μL of single worm lysate was added to 24.8 μL of PCR mix (1X, 17.8 μL dH20, 2.5 μL Promega Buffer B, 1.5 μL Promega 25mM $MgCl_2$, 0.5 μL 10 mM dNTP mix, 0.5 μL Promega Taq, 2 μL primer mix 12.5 μM each) and amplified with program A79. Single worm PCR was also used to genotype *rpm-1* (*ok364*). Amplification with the 5' rpm1O1 primer (cgaatctcctccacggaata) and the 3' rpm1O2 primer (atcgatttgatggtacggga), running the program JOSEPH (94° 5 min, 94° 30 s (1 cycle), 48° 45 s, 72° 4 min 30 s, (35 cycles), 72° C (1 cycle) produced a ~1600 bp fragment in animals containing the *ok364* deletion. The predicted wildtype *rpm-1* PCR product is much larger (~3150 bp) and only weakly amplified under these conditions. Therefore, to confirm *rpm-1* homozygosity, ≥ 15 progeny were analyzed for each experiment.

Snip-SNP mapping restriction digests

For the SNPs in both F32D8 and Y51A2D (Wicks, Yeh *et al*. 2001), 1 μL of PCR material was added to a restriction enzyme solution. (F32D8 = 1 μL αtaq1, 2 μL 10x buffer 3, 13 μL of water, and 2 μL of 10x BSA, digested for an hour and a half at 65°, Y51A2D = 1 μL of Dra1, 2 μL 10x buffer 4, 13 μL of water, and 2 μL of 10x BSA, digested for an hour at 37°C) (NEB buffer 3: 100mM NaCl, 50 mM Tris-HCl, 10 mM MgCl2, 1 mM dithiothreitol (pH7.9 at 25°C, NEB buffer 4: 20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium Acetate, 1 mM Dithiothreitol, pH 7.9 @ 25°C).

Complementation tests

Heterozygous *unc-4* (*e120*) males were mated with NC596 (*wd47*). Male progeny were scored for the Unc-4 phenotype. To identify NC742 (*wd72*), NC669, and NC742, complementation tests were performed with *rpm-1* (*ok364*) expressing *unc-129*::GFP. F1 cross progeny were scored for ectopic *unc-129*::GFP expression in VB motor neurons.

Results

Genetic screens isolate mutants with defective expression of the motor neuron marker, *acr-5*::YFP

The B-class specific motor neuron marker, *acr-5*::YFP, was used to screen for mutants with potential defects in motor neuron differentiation. An initial direct, visual screen in a stereodissecting microscope revealed > 200 mutant lines with altered *acr-5*::YFP expression (Table 2.1). In order to increase the throughput of this screen, we used the Union Biometric COPAS Biosort to automate identification of *acr-5*::YFP expression mutants. The Biosort is similar to FACs but, instead of sorting cells based on their fluorescent intensities, this device sorts worms (Fig 2.2A). A reconstruction experiment was initially conducted to confirm that the Biosort could isolate animals with altered *acr-5*::YFP expression. First, we confirmed that the Biosort could detect *acr-5*::YFP over base-line by comparing the scatter plots of both N2 and *acr-5*::YFP positive animals (Figure 2.2B). Next, we used the strain NC682 (*wd54*), isolated in the visual screen and in which *acr-5*::YFP expression is decreased in VNC B-class motor neurons, to set the

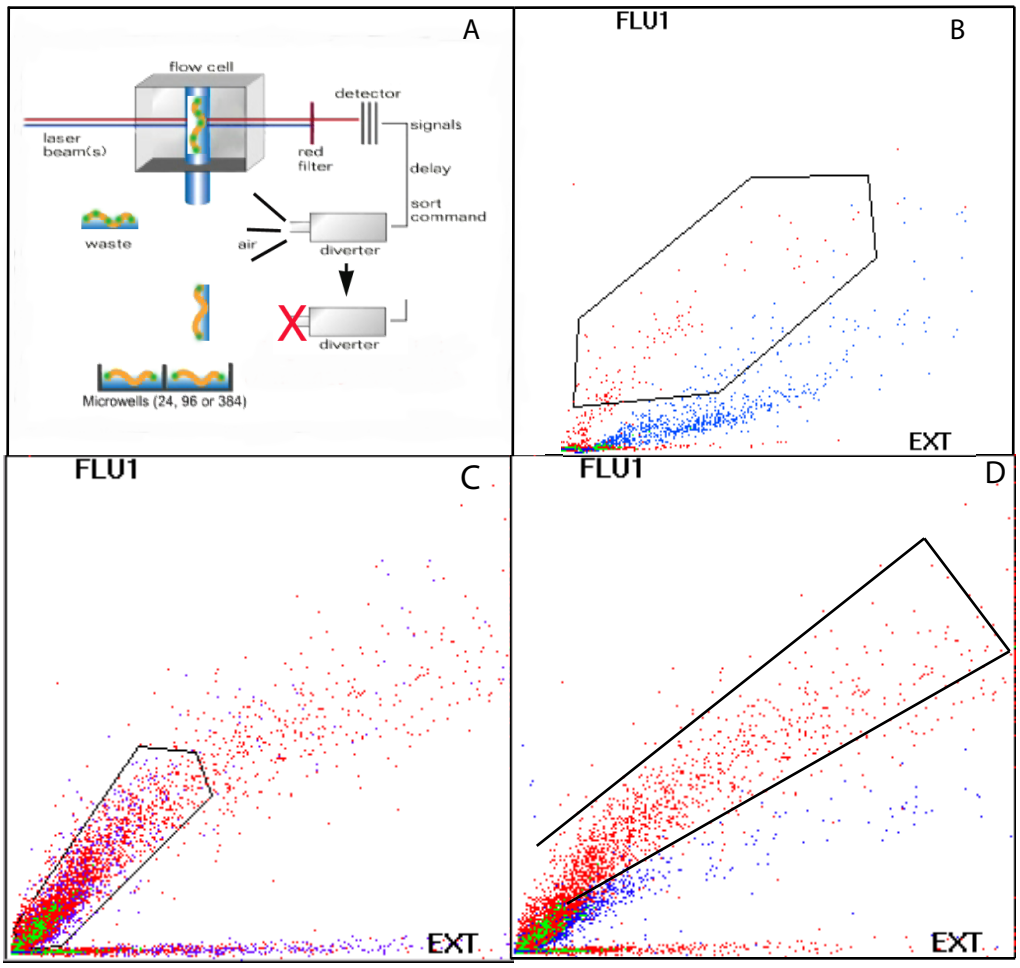


Figure 2.2 The Union Biometric COPAS Biosort isolates mutants that affect *acr-5*::YFP expression.

A. The Biosort can identify animals based on size (forward scatter) and GFP intensity (fluorescence). The schematic demonstrates the Biosort selecting animals with decreased expression.
B. WT *acr-5*::YFP animals (red) show a higher flourescent intensity than background (blue).
C. WT *acr-5*::YFP animals (red) are indistinguishable from *unc-4; acr-5*::YFP animals (blue).
D. WT *acr-5*::YFP (red) animals are easily separated from *wd54; acr-5*::YFP animals (blue)

Table 2.1 Mutant categories uncovered in intial *acr-5*::YFP visual screen.	
Mutant Phenotype	**Number identified**
Sterile Unc	173
Ectopic YFP near vulva	8
Dauer mutant	2
Male mutant	1
unc-4 alleles	2
decreased acr-5::YFP	7

BioSort gates for similar mutants (Fig 2.2D). [The BioSort could not distinguish *unc-4* (*e120*) mutants from wildtype and we therefore did not attempt to isolate animals with increased or ectopic *acr-5*::YFP expression in the ventral nerve cord (Fig 2.2C)]. The progeny of ~6000 mutagenized F1s were analyzed with these settings to isolate 14 independent mutants with decreased *acr-5*::YFP expression.

The combined results of these genetic screens identified several classes of mutants with altered *acr-5*::YFP expression (Table 2.1). Animals that failed to produce progeny, the so called 'Sterile Uncs' were the most highly represented class in this population (Horvitz, Sternberg *et al*. 1983; O'Connell, Leys *et al*. 1998). In these animals, cell division is blocked in multiple postembryonic lineages including the P-cell lineages that give rise to *acr-5*::YFP-expressing VB motor neurons. We therefore decided against further characterizing animals with this phenotype in order to focus on mutations with potentially specific roles in motor neuron differentiation. Additional classes of mutants with global defects in overall development were also discarded. For example, mutations that generate a constitutive dauer phenotype (Riddle and Albert 1997) or that affect the sex of the animal (e.g., *him* mutations) (Hodgkin 1983) were detected in this screen (Table 2.1). Dauer and male animals have additional motor neurons in the VNC (Sulston, Albertson *et al*. 1980) in comparison to the wildtype hermaphrodite, so the isolation of these mutants was not surprising. Several independent mutations that result in a single ectopic *acr-5*::YFP positive neuron near the vulva were isolated but also resulted in sterility and were therefore not characterized further. As expected, the visual screen revealed two new *unc-4* alleles in which *acr-5*::YFP is ectopically expressed in A-class motor neurons (Fig 2.3A). Complementation tests and DNA sequencing confirmed that these mutations

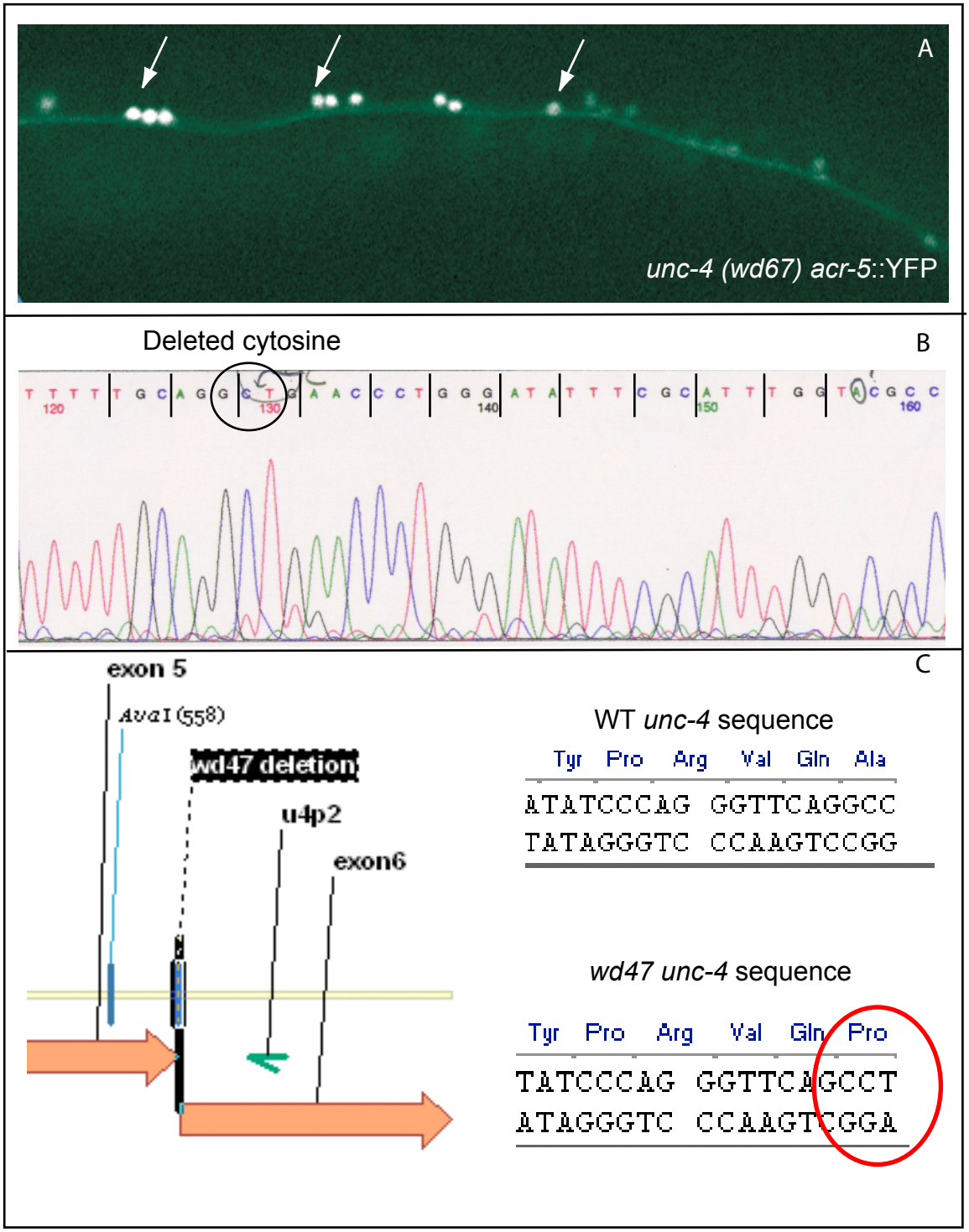


Figure 2.3 Screen for changes in motor neuron fate reveals a new allele of *unc-4*

a. *unc-4 (wd47)* was isolated in a visual screen of mutants that affect *acr-5*::YFP. *acr-5*::YFP, which is normally restricted to DB and VBs, is now expressed ectopically in DA and VA motor neruons.

b. Sequencing *unc-4* in NC596 (*wd47*) revealed that the molecular lesion was a deletion of a cytosine near the exon 5 and exon 6 splice junction.Note that the sequencing primer (u4p2) (C) is reading from the 3' to 5' direction. Arrows highlight ectopic motor neurons.

c. A schematic showing the consequence of the *wd47* deletion and a sequencing image that shows an amino acid change from an alanine to a proline.

disrupt *unc-4* function (Fig 2.3B). Our efforts to define the final class of mutants showing decreased *acr-5*::YFP expression are described in the next section.

Characterization of animals with reduced *acr-5*::YFP expression

A total of 21 mutants with reduced expression in the VNC were isolated from the combined results of the manual and automated screens (Table 2.2). These mutants could be divided into 3 subtypes based on movement and ventral cord morphology. We reasoned that genes that function in specifying late aspects of motor neuron identity would have a full complement of motor neurons yet could produce defects in movement. Type I mutants were uncoordinated, but DIC optics revealed that these mutants had an abnormal number of neurons in the VNC (Fig 2.4A). Type II mutants showed reduced *acr-5*::YFP expression, normal VNC architecture, and wildtype movement. Type III mutant animals appeared to have an intact VNC but showed reduced *acr-5*::YFP expression and uncoordinated movement. We therefore focused on this group of mutations as potential regulators of motor neuron fate. As described below, genetic and molecular analysis revealed that one of these strains, NC682, includes a new *rpm-1* allele.

NC682 shows a synthetic Unc trait arising from two independent mutations

NC682 (*wd54*) mutants are forward Unc and show no expression of *acr-5*::YFP in the L2 stage (Fig 2.4C); animals are also slightly Dumpy but the VNC architecture is grossly normal. As B-class motor neurons are known to control forward movement, we speculated that the forward Unc phenotype displayed by these mutants results from a change in B-class motor neuron fate. Genetic characterization of NC682 revealed that

Table 2.2 Summary of mutants that affect *acr-5*::YFP expression						
strain nar	**mutant**	**movement**	**acr-5 expression**	**VNC architecture**	**head and tail neurons**	**rpm-1 pathway?**
NC603	mut 20	WT	VCs in adult (average expression), several cell types in L2 (weak)	WT	WT	
NC605	mut 21	unc	decreased acr-5	missing cells in VNC	WT	
NC635	mut 23	unc	WT and missing acr-5::YFP	poor, cells exist, though		
NC627	mut 27	WT	some Bs are YFP+ -- very few, though	WT	WT	**rpm-1 allele**
NC616	mut-28	forward unc	very dim	WT	WT	**rpm-1, original isolate**
NC610	mut 29	WT	VNC is drastically reduced -- mostly expression in DBs (some VBs)	WT	WT	
NC611	mut 30	WT	mostly off in VBs, on in DBs -- reduced	WT	WT	
NC649	mut 33	WT	slightly reduced -- many are wildtype -- some cells are YFP -	WT	WT	
NC668	mut 35	WT -- backs?	decreased in some VBs, ectopic in DAs (weak)	WT	WT	
NC675	mut 36	unc and WT	decreased	cell bodies lump together	WT	
NC728	mut 38	unc	missing some YFP	missing some cells	?	
NC665	mut 39	WT	decrease in YFP expression -- B-class still express	WT	WT	
NC669	mut 41	WT	reduced acr-5, much stronger than rpm-1, off in posterior	abnormal cell bodies	WT	**rpm-1 allele**
NC679	mut 42	WT	overall decrease in acr-5::YFP -- b-class expression	WT	decreased	
NC666	mut 43	WT	reduced expression throughout	WT	WT	
NC680	mut 46		few YFP+ cells	missing cells in VNC		
NC670	mut 47	WT	WT, maybe a few have decreased acr-5 -- not many, though	WT	WT	
NC729	mut 48		ectopic cells once in a while	sketchy -- missing cells?		
NC730	mut 51	WT	ectopic YFP near vulva	WT	WT	
NC731	mut 52	WT and backward unc	expressed in Bs -- off in some, weak in others	missing some cells	WT	
NC651	mut 53	WT	reduced overall -- a few YFP - cells -- a few ectopic cells	WT	WT	

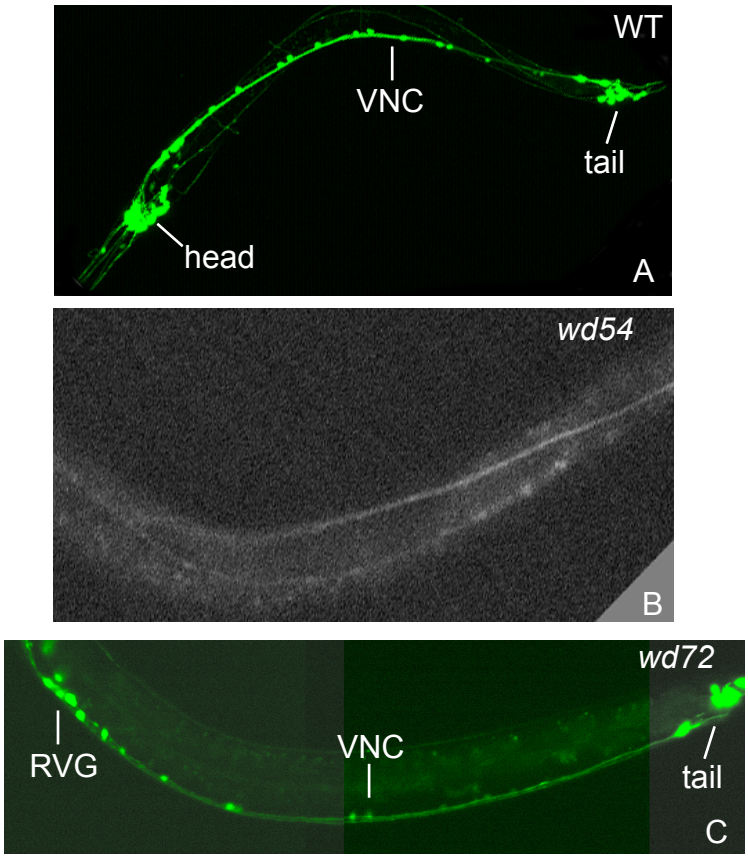


Figure 2.4 A synthetic Unc phenotype arises from the combined mutations of *wd54* and *wd72*

A. Confocal image of *acr-5*::YFP expression in the ventral nerve cord (VNC).
B. *acr-5*::YFP expression is largely absent in *wd54, wd72* double mutants.
C. *acr-5*::YFP expression is decreased but not eliminated in a *wd72* background

this strain includes two independent mutations (*wd54* and *wd72*) that act synergistically to eliminate *acr-5*::YFP expression in the VNC. Evidence for this possibility was initially detected during outcrossing when the Unc trait was recovered at a much lower frequency (<1/10) than the Mendelian ratio of ¼ expected for a single gene mutation (data not shown). The *wd72* allele was isolated from the F2 progeny of this outcross and showed a less severe reduction in *acr-5::*YFP expression than the parent strain and moved normally (Fig 2.4 C). A survey of additional motor neuron markers (Table 2.3) revealed that *unc-129::*GFP is ectopically expressed in VB motor neurons in *wd72* and in *rpm-1* mutants (Fig 2.5). This trait was useful for genetic experiments demonstrating that *wd72* is an *rpm-1* allele (see below).

wd72* as an allele of *rpm-1

Linkage data and deficiency mapping placed *wd72* in a small genetic interval on chromosome V (Fig 2.6C) (see Methods). Wormbase (WS100) showed six open reading frames (ORFs) in the likely genetic interval containing *wd72* (+1.64 to +1.88). Additional ORFs flanking this region were considered as candidates for the *wd72* locus due to the uncertain physical endpoints of deficiencies used to define this interval. A *mod-1* mutation complemented *wd72* and was therefore ruled out as was *unc-23*, which has a "bent head" phenotype that is not observed in *wd72*. Mutations in *rpm-1* (+1.59) (Fig 2.7A) result in visible synaptic defects in GABAergic motor neurons. I therefore used the *unc-25*::SNB-1::GFP marker to show that *wd72* also exhibits a similar defect in SNB-1::GFP localization in in GABAergic motor neurons (Fig. 2.7B). Furthermore, we showed that *rpm-1* (*ok364*) fails to complement the *unc-129*::GFP expression defect of

Table 2.3 GFP reporters tested for regulation by *rpm-1*		
Promoter GFP markers or antibodies	**Normal VNC expression pattern**	**wd72 expression pattern**
acr-2	DA, VA, DB, VB	DA, VA, DB, VB
acr-5	DB, VB	**off**
lin-11	VC	VC
unc-3	DA, VA, DB, VB, AS	DA, VA, DB, VB, AS
unc-4	DA, VA	DA, VA
unc-25	DD, VD	DD, VD
unc-129	DA, DB	**VB**
del-1	VA, VB	VA, VB
unc-53	DA, AS	DA, AS
vab-7	DB	DB
unc-47	DD, VD	DD, VD

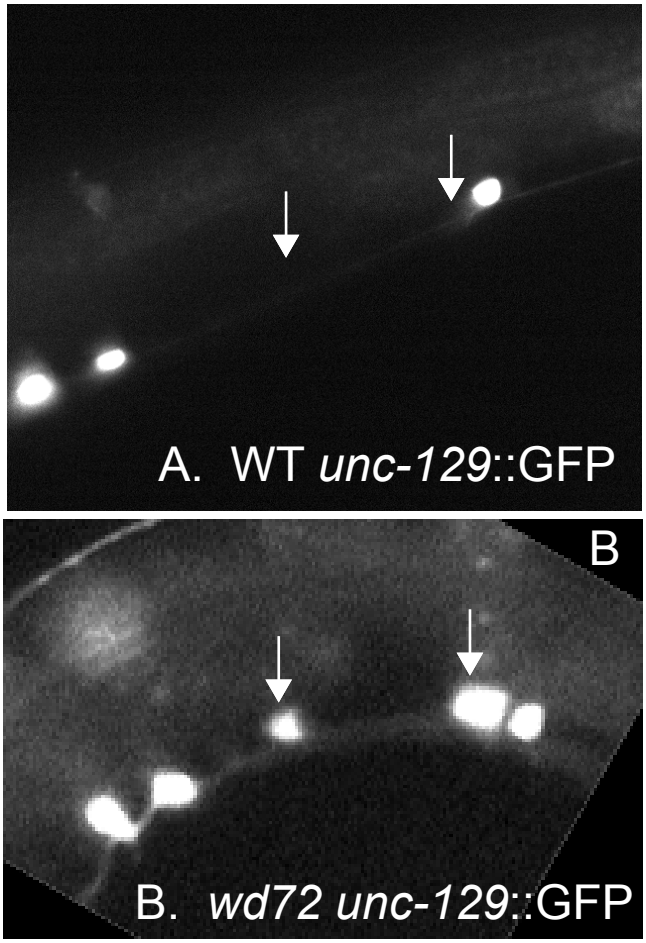


Figure 2.5 *wd72* regulates the expression of the TGF-beta-like molecule *unc-129*::GFP

A. WT *unc-129*::GFP is expressed in DA and DB motor neurons.

B. In *wd72* mutants, *unc-129*::GFP is ecopically expressed in VB motor neurons.

*Arrow indicate positions of VB motor neuron cell bodies.

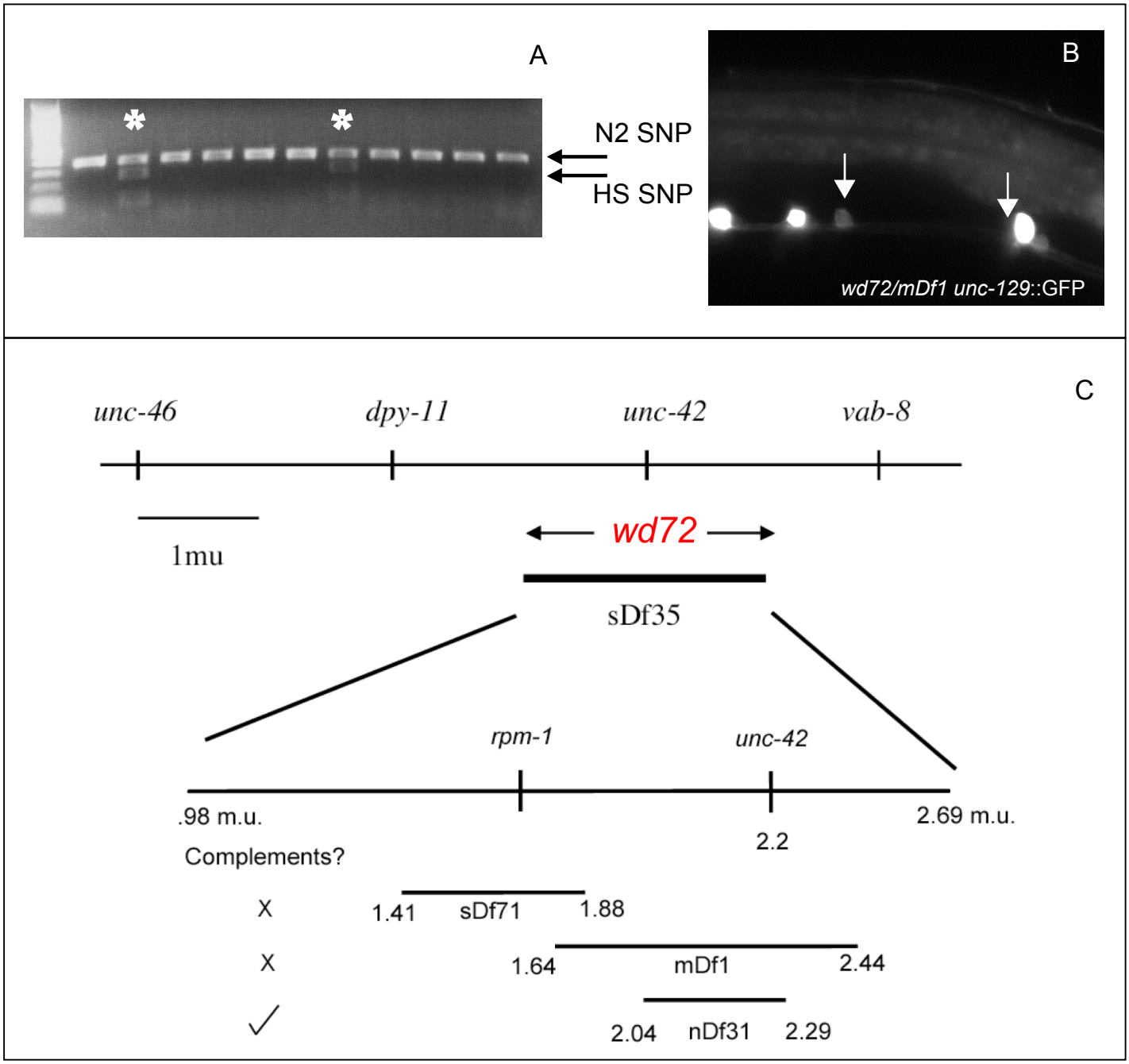


Figure 2.6 *wd72* maps near the center of chromosome V

A. Snip-SNP mapping links *wd72* near the SNP *F32D8* which is positioned at +3 on chromosome V. Lanes 3 and 7 (*) contain recombinant mutants with both Hawaiian (HS) and Bristol (N2) single nucleotide polymorphisms (SNPs).

B. *mDf1* (+1.64 to +2.44) fails to complement *wd72*.
Arrows = ectopic expression of *unc-129*::GFP in VB motor neurons.

C. Summary of deficiency mapping data showing that *wd72* maps near *rpm-1*.

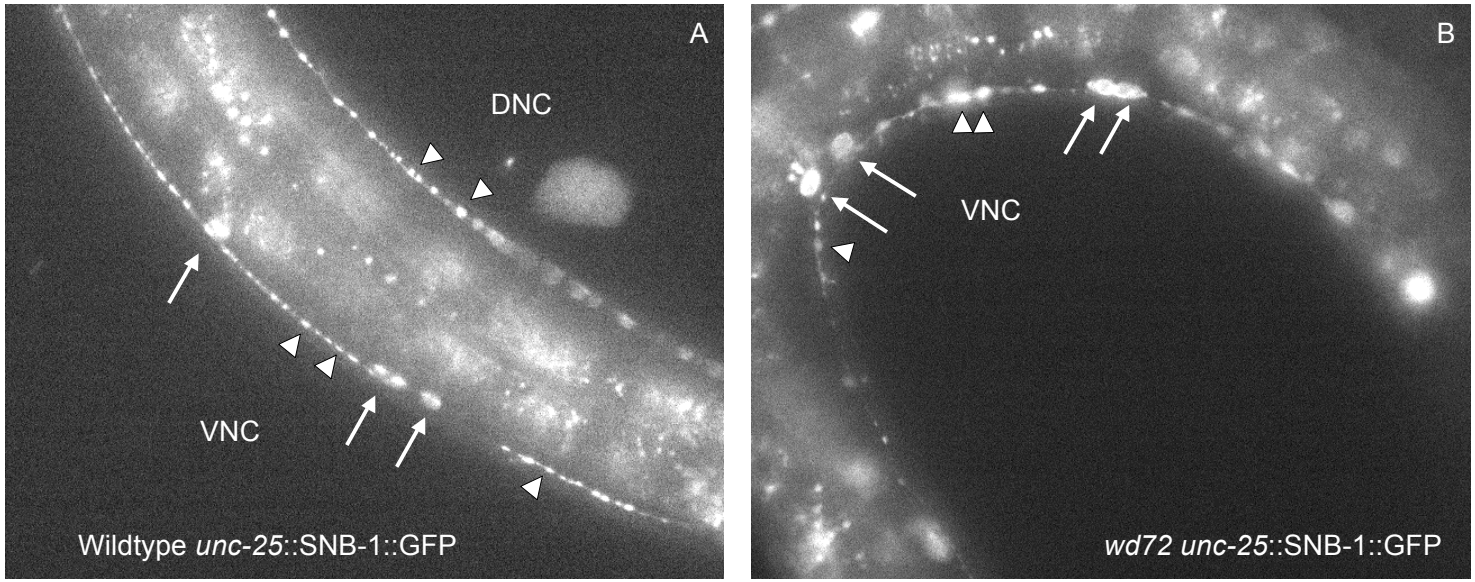


Figure 2.7 *wd72* disrupts GABAergic motor neuron synaptic morphology

A. The ventral and dorsal nerve cords of wildtype *unc-25*::SNB-1::GFP animals show small, punctate, evenly distributed GFP spots. These spots represent GABA synapses within the motor neuron circuit.

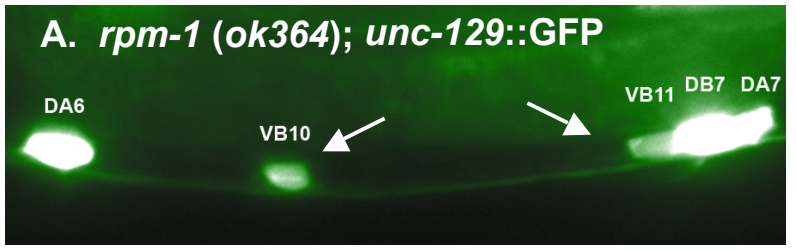
B. *wd72* phenocopies *rpm-1* mutants. In *rpm-1* mutants, synapses are large and unevenly distributed.

Arrowheads = synapses Arrows = GABA neuron cell bodies

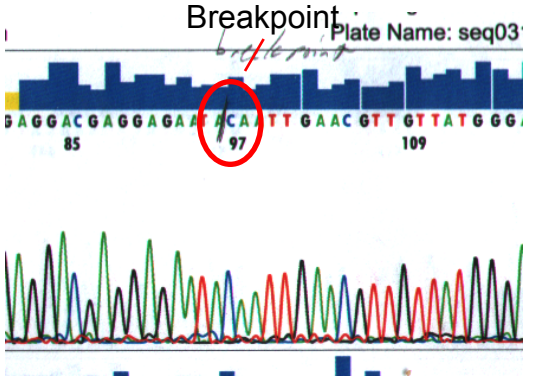
wd72 and that *ok364* and *rpm-1(ju41)* show this phenotype as a homozygous mutant (Figure 2.8A, B). *rpm-1 (ju41)* is a point mutation in the E3 ubiquitin ligase domain of *rpm-1* that results in a truncated protein (Zhen, Huang *et al*. 2000). The DNA sequence of the *rpm-1* (*ok364*) deletion allele indicates that *ok364* also likely to represent an *rpm-1* null allele (Fig 2.8D-F). Complementation tests with other Type III mutants isolated in this screen revealed two additional *rpm-1* alleles, *wd55* and *wd67* (Fig 2.8C). Taken together, these data indicate that *wd55, wd67* and *wd72* are loss-of-function mutations in the *rpm-1* locus and suggest the intriguing possibility that the synaptic protein RPM-1 regulates neuronal gene expression.

Discussion

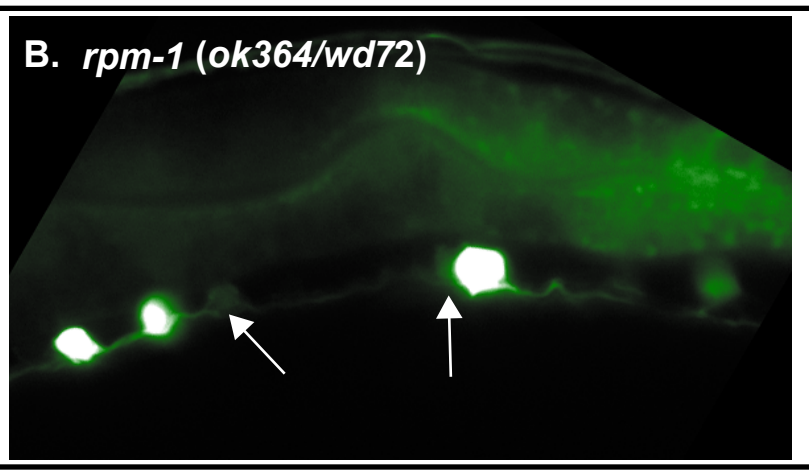
A genetic screen designed to detect mutations in genes that specify motor neuron fate led to the isolation of 21 mutants with altered expression of the B-class motor neuron marker *acr-5::YFP*. Two of these mutations are new alleles of *unc-4*, a homeodomain transcription factor that is known to regulate *acr-5* and to specify synaptic inputs to VA motor neurons (Miller, Shen *et al*. 1992; White, Southgate *et al*. 1992; Winnier, Meir *et al*. 1999; Von Stetina, Fox *et al*. 2007). Most of the mutants isolated in this screen remain uncharacterized but may affect genes with interesting roles in motor neuron fate. For example, in *wd49* animals, the loss of *acr-5*::YFP expression in B-class motor neurons is accompanied by ectopic *acr-5*::YFP expression in VC motor neurons. VC motor neurons normally innervate vulval muscles to regulate egg-laying activity (Li and Chalfie 1990). The *wd49* phenotype is similar to that of *pag-3* mutants in which VB



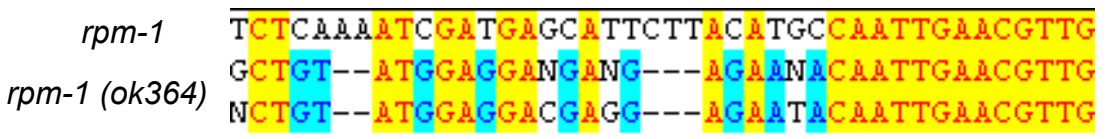
D

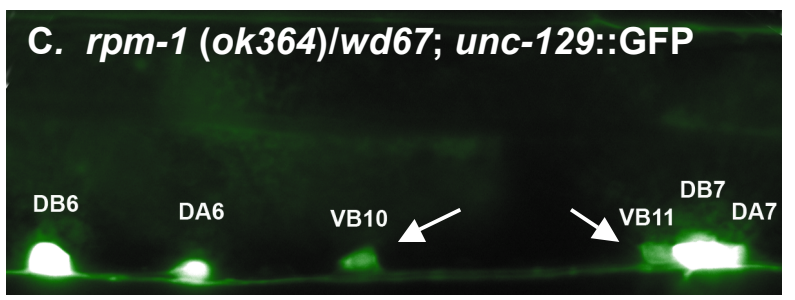


rpm-1 (ok364) sequence



E





F

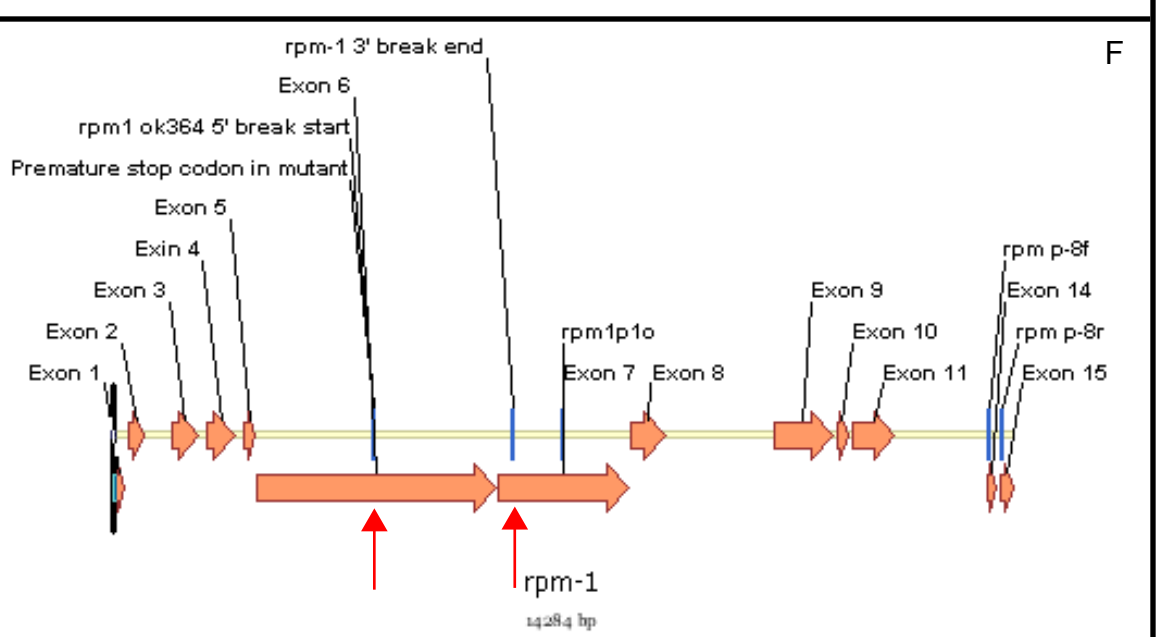


Figure 2.8 *wd72* is an allele of *rpm-1*

A. *rpm-1 (ok364)* display ectopic *unc-129*::GFP in VB motor neurons.
B. *wd72* fails to complement the *rpm-1 (ok364)* deletion allele.
C. *rpm-1 (ok364)* fails to complement a second mutation pulled from the *acr-5*::YFP screen.
D,E. Sequencing of the *rpm-1 (ok364)* deletion reveals the endpoints.
F. Schematic of the *rpm-1* deletion. *rpm-1 (ok364)* removes a portion of the gene spanning the two red arrows and results in a premature stop in Exon 7.
White arrows = ectopic VB motor neurons

motor neurons throughout the ventral nerve cord assume VC motor neuron characteristics (Cameron, Clark *et al*. 2002). In the future, it will be interesting to determine if *wd49* affects a *pag-3*-dependent pathway. Also of note, whereas the BioSort was unable to identify mutants that result in ectopic expression of *acr-5*::YFP (e.g. *unc-4*), additional experiments did demonstrate that mutations in *unc-4* could be isolated with the marker *del-1*::GFP (Fig 2.9). *del-1* encodes a degenerin-like channel subunit and is ectopically expressed in *unc-4* mutants (Winnier, Meir *et al*. 1999). Thus, a BioSort using this *del-1::*GFP reporter could potentially reveal new mutations in other *unc-4* pathway genes.

Genetic tests demonstrated that three independent mutations that disrupt *acr-5*::YFP expression are alleles of *rpm-1*, a conserved E3 ubiquitin ligase that regulates synaptic morphology and function (Schaefer, Hadwiger *et al*. 2000; Wan, DiAntonio *et al*. 2000; Zhen, Huang *et al*. 2000; Burgess, Peterson *et al*. 2004; D'Souza, Hendricks *et al*. 2005). These findings suggest the intriguing possibility that a protein degradation event that depends on *rpm-1* function regulates transcription of downstream genes that contribute to synaptic assembly. This model is consistent with the recent evidence that RPM-1/Highwire in *C. elegans* and *Drosophila* negatively regulate a MAP Kinase signaling cascade by targeting MAPKKK for degradation at the synapse. The additional finding that a presumptive downstream target of the MAP Kinase cascade in *Drosophila*, the conserved transcription factor, Fos, mediates Highwire function, offers additional strong support for this model. Although my genetic screen identified two genes, *acr-5* and *unc-129*, that are presumptivly regulated by *rpm-1*, genetic evidence has ruled out a likely functional role for these in synaptic assembly. Thus, these results motivated my

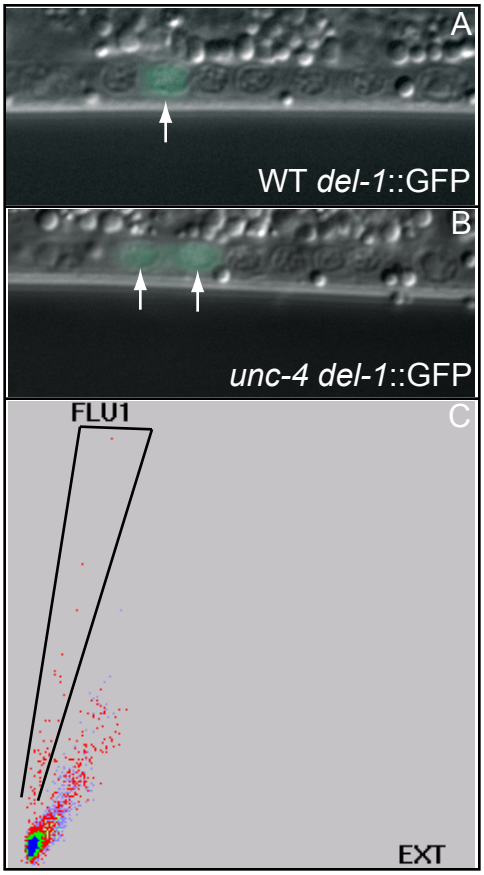


Figure 2.9 The Biosort detects ectopic expression of *del-1*::GFP in *unc-4* mutants.

A. In wildtype L2 animals, *del-1*::GFP is expressed in VB motor neurons.
B. In *unc-4* mutants, *del-1*::GFP is etopically expressed in VA motor neurons.
C. The Union Biometric COPAS biosort can selectively differentiate between wildtype animals (blue) and *unc-4* mutant animals (red) expressing *del-1*::GFP (boxed area).

goal of optimizing cell specific microarray technology for the purpose of additional candidate genes for downstream effectors of *rpm-1* regulation of synaptic assembly.

CHAPTER III

DEFINING DOWNSTREAM TARGETS OF RPM-1

Introduction

As mentioned in the previous chapter, *rpm-1* encodes an E3 ubiquitin ligase. E3 ubiquitin ligases act in conjunction with F-box proteins to target specific proteins for proteasomal degradation (Willems, Schwab *et al*. 2004). This finding led to the hypothesis that RPM-1 regulates synaptic morphology by controlling the levels of target proteins at the synapse. Two lines of evidence support this model. Initially, Liao, *et al*, identified *fsn-1*, an F-box protein that, when mutated, phenocopies the GABAergic synaptic defects of *rpm-1*. FSN-1 physically interacts with RPM-1, and the synaptic defects of double mutants are similar to mutations in either gene alone (Liao, Hung *et al*. 2004). These data suggest that RPM-1 and FSN-1 function in a complex to target proteins for degradation. Secondly, the Jin lab conducted a sensitized genetic screen to identify other targets of RPM-1. *rpm-1* mutants exhibit no obvious movement defects (Zhen, Huang *et al*. 2000; Nakata, Abrams *et al*. 2005). However, the combination of *rpm-1* alleles with certain other mutations (such as *syd-2* – see chapter 1) that affect synaptic morphology results in a synergistic uncoordinated phenotype (Nakata, Abrams *et al*. 2005). Since neither mutation is severely Unc alone, additional mutations that affect either *rpm-1* or this second pathway should "suppress" the Unc trait of the double mutant strain. This strategy detected mutations in three members of the MAP Kinase cascade, *pmk-3*, *mkk-4*, and *dlk-1*, as strong dominant suppressors of *rpm-1* (Nakata, Abrams *et al*.

2005). This effect is likely to be related to synaptic function since mutations in these MAP Kinase pathway components restore normal synaptic morphology to *rpm-1* mutants. *dlk-1* encodes a MAPKKK, *mkk-4* encodes a MAPKK, and *pmk-3* encodes a p38 MAP Kinase. *dlk-1* is expressed at the synapse and is likely to activate *mkk-4*, which, in turn, activates *pmk-3*. Biochemical analysis has shown that DLK-1 is ubiquitinated in an *rpm-1* dependent manner *in vitro* (Nakata, Abrams *et al*. 2005). This finding suggests that *rpm-1* normally functions to negatively regulate this MAP Kinase signaling cascade by promoting degradation of DLK-1, the first component of the pathway.

MAP Kinases play a well-established role in synaptic plasticity. In Aplysia, application of the neurotransmitter serotonin (5-HT) in successive bursts induces long term facilitation (LTF) (Kandel 2004). LTF, in turn, is associated with the growth of new synaptic connections. LTF is activated by the MAPK ERK and requires both the synthesis of new proteins and active transcription. Hence, the plasticity of a synapse is dependent on both MAPK signaling and transcription. Interestingly, the proteasome is also associated with LTF (see chapter 2), again linking protein degradation with transcription and synaptic development.

While MAPK signaling plays a prominent role in synaptic plasticity, the role of MAP Kinase during synaptic development is less well understood. However, genetic evidence in *C. elegans* substantiates the claim that synaptic development may depend on MAP Kinase signaling in a pathway independent of *rpm-1*. *In vivo* protein localization assays place the MAPKKK *nsy-1* at the synapses of an asymmetric bilateral pair of sensory neurons (AWC) during late embryonic development (Chuang and Bargmann

2005). This time period corresponds to a critical period in *C. elegans* synaptic development. While both AWC neurons synapse on each other, one of the pair, AWCR, is dominantly pre-synaptic. The loss of *nsy-1* disrupts the asymmetry of AWC neurons (Wes and Bargmann 2001), and may affect the asymmetric distribution of synapses. Therefore, the MAP Kinase cascade may indirectly regulate synaptic development by controlling cell identity.

Although the mammalian proteins downstream of the *rpm-1* homologue Phr-1 have not been defined (Burgess, Peterson *et al*. 2004), recent evidence indicates that a MAP Kinase signaling pathway is also regulated by *Drosophila* Highwire. Mutations in wallenda, a MAPKKK, can fully suppress Highwire mutant synaptic defects (Collins, Wairkar *et al*. 2006). In this case, genetic evidence also suggests that the transcription factor Fos functions downstream of the Highwire-regulated MAP Kinase signaling cascade (Collins, Wairkar *et al*. 2006). This finding in Drosophila offers a potential mechanism to explain our independent discovery that *rpm-1* regulates specific transcript levels in the *C. elegans* nervous system. In the following, we use a genomic approach to identify the transcriptional targets of *rpm-1*. A profile of the entire C. elegans nervous system generated from animals in the 2nd larval stage identified 558 enriched transcripts in the *rpm-1* mutant dataset. An additional 168 transcripts were downregulated in an *rpm-1* background. These data support the idea that *rpm-1*/Highwire controls synaptic development via RPM-1-dependent MAP Kinase signaling pathways that regulate gene transcription. Here, I describe our efforts to confirm this hypothesis by identifying RPM-1 regulated transcripts that are involved in synaptogenesis.

Methods and Materials

Microarray data generation and analysis

Worm harvesting, RNA isolation, amplification and hybridization are described in detail in Chapter 4 and are identical for both wildtype and *rpm-1 (ok364)* animals expressing the pan neural mRNA tagging transgene, *F25B3.3*::FLAG::PAB-1. Hybridization intensities for each experiment were scaled to a global average signal intensity and normalized by RMA (Robust Multi-Array analysis) (see Fox *et al*, 2005). *rpm-1*-regulated transcripts were identified by comparing RMA (Robust Multi-Array analysis) normalized intensities from the *rpm-1* profile *vs* RMA-normalized intensities from wildtype. Transcripts showing statistically significant differences were identified using Significance Analysis of Microarrays (SAM, Chapter 4). A two-class unpaired analysis of the data was performed to identify genes that differ by > 1.5X (enriched) or <1.5X (depleted) vs the Pan-neural reference dataset at a False Discovery Rate of <1%.

Molecular Biology

To construct *unc-4*::TBB-6::CFP (pJW5), *tbb-6* was amplified from genomic DNA using the primers: pSC392TBB6F1 (5'TCTAGAATGAAAGAAATTATTAACG TTCAAGTTGG) and pSC392TBB6R1 (5'GGTACCGCTGAATGAACTTTC TCATATTGTTGG). The resulting fragment was gel purified and transformed into pCR2.1-TOPO as per the Invitrogen TOPO-TA protocol. pJW5 was digested with XbaI to produce a 1520 bp fragment containing the TBB-6 genomic fragment. To express TBB-6 in GABAergic neurons, the *unc-25* promoter was digested out of pSC392 with

XbaI and the 1520 bp fragment from the pJW5 digestion was ligated into pSC392 to make pJW6 (*unc-25*::TBB-6::CFP). pSV9-TOPO was then digested with XbaI and HindIII, producing a 3019 bp fragment that contained the *unc-4* promoter. An XbaI HindIII double digest of pJW6 and subsequent purification of a 5775 bp band removed the *unc-25* promoter. The 3019 bp fragment and the 5775 bp fragment from the previous two digestions were then ligated to produce pJW7 (*unc-4*::TBB-6::CFP).

RNAi experiments

unc-129::GFP; *eri-1* (*mg366*) doubles were used to assess the effect of the RNAi of *rpm-1* on gene expression and to test candidate depleted microarray targets for regulation of gene expression. Similarly, *unc-129*::GFP; *eri-1*; *rpm-1* triples were used to check the effectiveness of RNAi of *pmk-3* on gene expression and, subsequently, to screen the enriched microarray targets of *rpm-1* for effects on gene expression. *eri-1* (*mg366*) *unc-25*::VAMP; *lin-15B* (*n744*) triples and *eri-1 unc-25*::VAMP, *rpm-1*; *lin-15B* animals were created to assess whether RNAi of *rpm-1* regulated genes affects synaptic structure. Movement phenotypes were examined using the *eri-1*; *syd-2 (ju37) lin-15B* triple and the *eri-1*; *rpm-1*; *syd-2 (ju37) lin-15B* quadruples.

All RNAi experiments were performed by the feeding method using dsRNA expression strains from the Ahringer RNAi library (Kamath and Ahringer 2003; Kamath, Fraser *et al*. 2003). Cultures were grown from a single colony overnight in 2mL of LB + 50μg/mL of ampicillin. 3-4 mL LB was innoculated with 30-40 μL of the overnight cultures. The bacteria were grown to an absorbance of 0.8 (~43/4 hr) at 37° C). 3 mL of LB +IPTG (40 mM final concentration) was added to each culture to induce dsRNA

production. After 4 hours at 37° C, bacteria were pelleted at ~4000 rpm in a desktop centrifuge. Bacterial pellets were resuspended in 250 μL of M9 spiked with 10 μL of IPTG (1 M). The bacterial suspension was applied to unseeded NGM plates and allowed to dry. Three L4 larvae were added to each plate. F1 progeny of these animals were scored for RNAi effects on the fourth or fifth day after initiating the treatment.

Real-Time (RT) PCR experiments

mRNA generated from IP experiments was purified using the Qiagen RNeasy mini kit as above, with the optional DNase step described in the Qiagen protocol, and resuspended in a final volume of 50 μL and the concentration of the purified mRNA was determined using a Pharmacia GeneQuant II spectrophotometer. The Protoscript Kit (New England Biolabs) was used to reverse transcribe 100 ng of sample RNA with 2 μL of random primers (15 μM), 4 μL of dNTP mix (2.5 mM each). The reaction was heated at 70° C for 5 minutes and placed on ice. 2 μL of 10X RT Buffer, 1 μL of RNase Inhibitor, and 1 μL of M-MuLV Reverse Transcriptase were added to each reaction, incubated at 42° C for one hour, and then inactivated at 95° C for 5 minutes. The reaction was cooled for 2 minutes on ice. 5 μL of RNase H (10 units) was added to the reaction and incubated at 37° C for 20 minutes. The enzyme was then inactivated at 95° C for 5 minutes. The reaction volume was brought to 200 μL with the addition of DEPC-treated water.

Primers for real-time PCR were designed to span an intron near the 3' end of each transcript and to generate a ~150 bp amplified fragment. Primers for the following genes were used: *gpd-1* (*gpd-1* f1 5'GCCGATGGACCAATGAAGGG, *gpd-1* r1

5'CCGATGAGGTCGACAACACGG), *F57F4.4* (*F57F4.4* f1 5' GCTTGCGCTGGATCGTGCGCC, *F57F4.4* r1 5' GCATCCGGAGAGTCCGCCG), *tbb-6* (*tbb-6* f1 5' GCTGAACAAATCATCAGTGTGG, *tbb-6* r1 5' CTCCGACTCTTCCATCGTG), *dlk-1* (*dlk-1* f1 5' GCCACCCGGTCCGATGGGC, *dlk-1* r1 5' CGGACTGCTCCGGCATCGTC), *act-1* (*act-1* f1 5' GGGTATGGAGTCCGCCGG, *act-1* r1 5' GGGAAGCGAGGATAGATCCTC) Samples were assayed in triplicate. 0.5 μL of cDNA was used for each experiment. 12.5 μL of IQ SYBR Green supermix was added to each sample, along with the following: 0.5 μL of each primer and 11 μL of sterile water (per manufacturers instructions). All reactions were performed in an Biorad Icycler using the following conditions: cycle 1, 5 min at 95° C; cycle 2, 30 s at 94° C, 30s at 60° C, 1 min at 72° C, and 10 s at 78° C, 40 times; cycle 3, 1 min at 55° C; cycle 4, 55° C for 10 s with 0.5° ramping each repeat, 80 repeats. Results were normalized to GAPDH (*gpd-1*) and actin (*act-1*), whose transcript levels were also monitored during real-time PCR.

Nematode Strains and Genetics

Nematodes were grown as described (Brenner 1974). All genetics were performed at 20° C. *unc-129*::GFP; *rpm-1* (*ok364*) males were crossed into *pmk-3 (ok169)* hermaphrodites. *unc-129*::GFP-positive hermaphrodite cross progeny were picked for selfing. F2 animals showing wildtype *unc-129*::GFP expression were allowed to self and homozygous *rpm-1* (*ok364*); *pmk-3* (ok169) off spring were verified by PCR [Primers - OK169_internal_f: ttttcactgcgtctcaatcg, OK169_internal_b:

tttcaaatttgcaggtgtgc, same PCR conditions as *rpm-1* deletion (Chapter 2 methods) except a 3 minute extension time].

Heterozygous *unc-129*::GFP animals were crossed into *syd-1* (*ju82*). *syd-1*(*ju82*) coils ventrally when prodded backwards. Motor neurons of F2 *syd-1* animals expressing *unc-129*::GFP were then scored.

The deletion mutant *tbb-6 (tm2004)* was obtained from the Japanese knockout consortium (Gengyo-Ando and Mitani 2000). *tbb-6* (tm2004) *unc-25*::VAMP::GFP was created. Homozygous *tm2004* animals were detected by PCR (Primers – tm2004if: CGTTGGAACTTGTATCATGC, tm2004ir: TTGCGACGGAACAGGCCTGT). PCR of the *tbb-6* deletion is optimized for use with Opti-Prime buffer 5 (10 mM Tris-HCL, pH8.8, 1.5 mM $MgCl_2$, and 25 mM KCL), and the PCR cycle is identical to *rpm-1* deletion detection described in chapter 2 except the 72° extension time is 3:15, see Single Worm PCR in Chapter 2). *tbb-6* maps close to *rpm-1* so, in order to generate a *tbb-6 rpm-1* double mutant, it was necessary to pick a recombinant chromosome. *unc-4*2 maps to the right (2.16 M.U.s) of *rpm-1* (1.6 M.U.s) and to the left of *tbb-6* (+4 M.U.s). A recombination event between an *rpm-1 unc-4*2 double mutant and *tbb-6* should generate *rpm-1 tbb-6*. *unc-129*::GFP *rpm-1 unc-4*2 heterozygous double mutants were crossed into *tbb-6* (tm2004). Non-Unc *rpm-1* mutants were selected and homozygosed for the *tbb-6* deletion.

acdh-1 (*ok1489*) was obtained from the CGC. Homozygous *acdh-1* (*ok1489*); *unc-25*::VAMP::GFP animals were confirmed by PCR (primers – *OK1489*_external_f: gtcacctcaaaccaagggaa, *OK1489*_external_b: ggtgggatgtacggtaggag, all PCR conditions are identical to *rpm-1* (*ok364*) except the extension time is 3:15. *rpm-1* (*ok364*) *unc-*

25::VAMP::GFP animals were mated into *acdh-1*, homozygosed for *rpm-1*, and then homozygosed for *acdh-1* using PCR. In order to test *acdh-1* for synaptic defects in a genetically sensitized background, *rpm-1* (*ok364*) *unc-4*2 *unc-25*::VAMP::GFP animals were mated into *acdh-1* (*ok1489*). *acdh-1* was homozygosed by PCR, and animals with wildtype movement (*rpm-1* heterozygotes) were scored for synaptic defects.

Results

syd-1 functions in a parallel pathway and does not affect *unc-129*::GFP expression

Mutations in *syd-1*, a PDZ domain containing protein, result in moderately uncoordinated animals with synaptic morphology defects (Hallam, Goncharov *et al*. 2002). *syd-1*; *rpm-1* double mutants are severely uncoordinated, suggesting that these two proteins function in separate pathways. This model is also consistent with the finding that *syd-1* does not affect *unc-129*::GFP expression (data not shown). Additionally, *syd-1* also does not enhance *rpm-1*-dependent *unc-129*::GFP expression defect. These results suggest that *syd-1* controls synaptic morphology via a separate downstream pathway that may not depend on the regulation of gene expression in VB class motor neurons and also indicate that changes in synaptic morphology alone are not enough to change gene expression.

The MAP Kinase cascade regulates gene expression downstream of RPM-1

Mutations in *pmk-3* (MAP Kinase) rescue *rpm-1* synaptic defects thereby suggesting that *pmk-3* functions downstream of *rpm-1* to control synaptogenesis (Nakata,

Abrams *et al*. 2005). To determine if *rpm-1* also controls gene expression via *pmk-3*, I asked if the *pmk-3* mutation can restore normal *unc-129*::GFP expression in *rpm-1* mutants. This experiment clearly demonstrates that *pmk-3* mutations dominantly rescue the *rpm-1* dependent *unc-129*::GFP expression defect (Figure 3.1A, B). This result led to the hypothesis that *rpm-1* specifies synaptic morphology by regulating gene expression (Figure 3.1C).

A genetic screen for downstream components of the RPM-1 pathway

Our finding that *pmk-3* suppresses the *unc-129*::GFP defect observed in *rpm-1* mutants, suggested an effective strategy for isolating mutations in additional *rpm-1* pathway genes. As opposed to the *rpm-1* synaptic defect, which is difficult to score and must be observed in a compound microscope ($\geq$ 630 X), ectopic expression of *unc-129*::GFP in the ventral cord can be readily observed with low power magnification ($\leq$ 200 X) on a fluorescence stereodissecting microscope. I worked with Sarah Holt (Vanderbilt undergraduate) and Manisha Tripathi (IGP rotation student) to conduct an F1 screen for dominant suppressors (*pmk-3* is a dominant suppressor) of ectopic *unc-129*::GFP expression in *rpm-1(ok364)*. We isolated 6 independent Rpm-1 suppressor mutations. (Figure 3.2A). Two isolates of *sorp-1* (*wd74* and *wd75*) (suppressor of *rpm-1*) are semi-dominant, map to the same chromosome, and are allelic. Based on the map location these alleles are likely mutations in the previously characterized MAPKKK *dlk-1*. Two other isolates (*sorp-3, wd79 and wd80*) are allelic and map to a region near *pmk-3* (Figure 3.2C, data not shown). However, sequencing of the *pmk-3* coding region and 1.6 kb of upstream promoter from *sorp-3* (*wd79*) did not detect a mutation (data not shown). These mutants

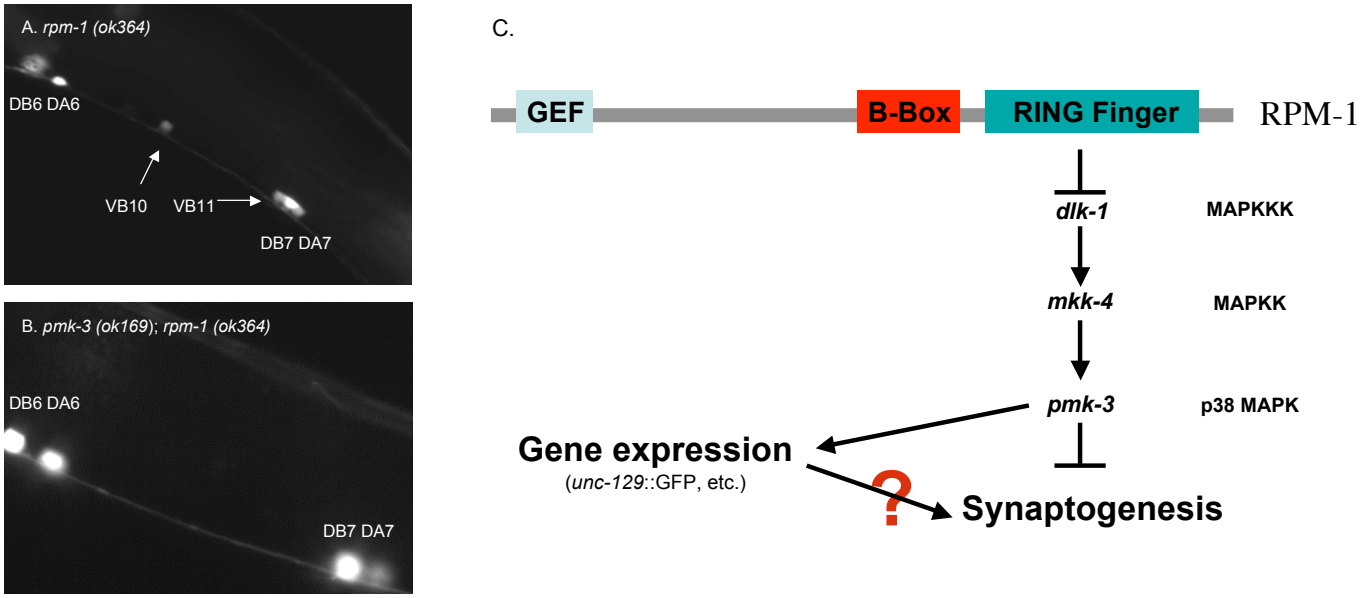


Figure 3.1 RPM-1 regulates transcription through MAP Kinase signaling

A. *rpm-1* (*ok364*) mutants ectopically express *unc-129*::GFP in VB motor neurons.

B. Ectopic expression of *unc-129*::GFP in an *rpm-1 (ok364*) mutant (A) is suppressed in an *rpm-1 (ok364) pmk-3 (ok169)* double mutant.

C. RPM-1 functions at the synapse and regulates MAP Kinase signaling to control gene expression. We speculate that RPM-1-regulation of gene expression is necessary for proper synaptic morphology.

Table of isolated Suppressors Of RPm-1 (*sorp*)			
	Chromosome	**alleles**	**Phenotype**
sorp-1	I	wd74 and wd75	semi-dominant
sorp-3	IV	wd79 and wd80	recessive
sorp-5	Not III, IV, V, X	wd78	recessive
sorp-6	Not I, IV, V, X	wd81	recessive
sorp-7		wd82	dominant
sorp-8		wd83	dominant

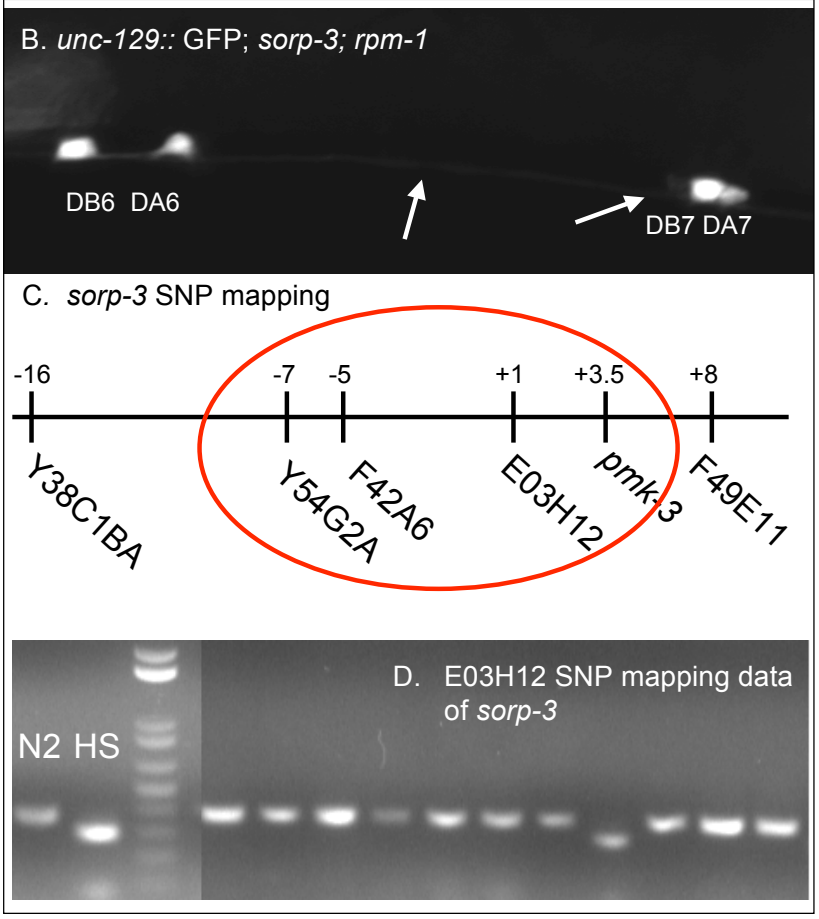


Figure 3.2 An *rpm-1* suppressor screen identifies 8 potential *rpm-1* pathway components.

A. A table of all suppressors isolated in the mutant screen.
B. *sorp-3* partially suppresses the Rpm-1 gene expression phenotype.
C. *sorp-3* maps to Chromosome IV near *pmk-3*.
D. Linkage of *sorp-3* to the center of chromosome IV using SNP markers.

are also weaker Rpm-1 suppressors than the *pmk-3* deletion allele and therefore could potentially affect another nearby *rpm-1* pathway component. Finally, one of the *rpm-1* suppressor mutations, *sorp-6*, is not linked to any of the known MAP Kinase genes. However, further genetic mapping experiments to refine the map location of this gene were difficult and eventually discontinued in order to focus on candidate *rpm-1* regulated genes revealed by the microarray data.

Microarray experiments reveal RPM-1-regulated transcripts in *C. elegans* neurons

We used microarray technology to determine the full complement of genes regulated by RPM-1 in the *C. elegans* nervous system. Because *rpm-1*-dependent synaptic defects are observed in GABA motor neurons in the L2 larval stage, we selected this developmental period for the profiling experiment (Zhen, Huang *et al*. 2000). The recently developed method of mRNA tagging (Roy, Stuart *et al*. 2002; Kunitomo, Uesugi *et al*. 2005; Pauli, Liu *et al*. 2005; Yang, Edenberg *et al*. 2005; Von Stetina, Fox *et al*. 2007) (Von Stetina, Watson, *et al*., *in press*) (See Chapter 4) was used to isolate transcripts from all *C. elegans* neurons. Wildtype vs *rpm-1* mutant profiles were compared to identify candidate RPM-1-regulated genes.

Reproducible microarray results were obtained from three independent RNA preparations generated from *rpm-1* (*ok364*) (Figure 3.3A). Scatter plots of these data versus a microarray profile obtained from the wildtype nervous system detected transcripts with altered intensity values (Figure 3.3B). Transcripts showing statistically significant differences in abundance were determined (see Methods). This treatment identified 523 transcripts with elevated expression ($\geq$ 1.5 X) in *rpm-1* mutants (False

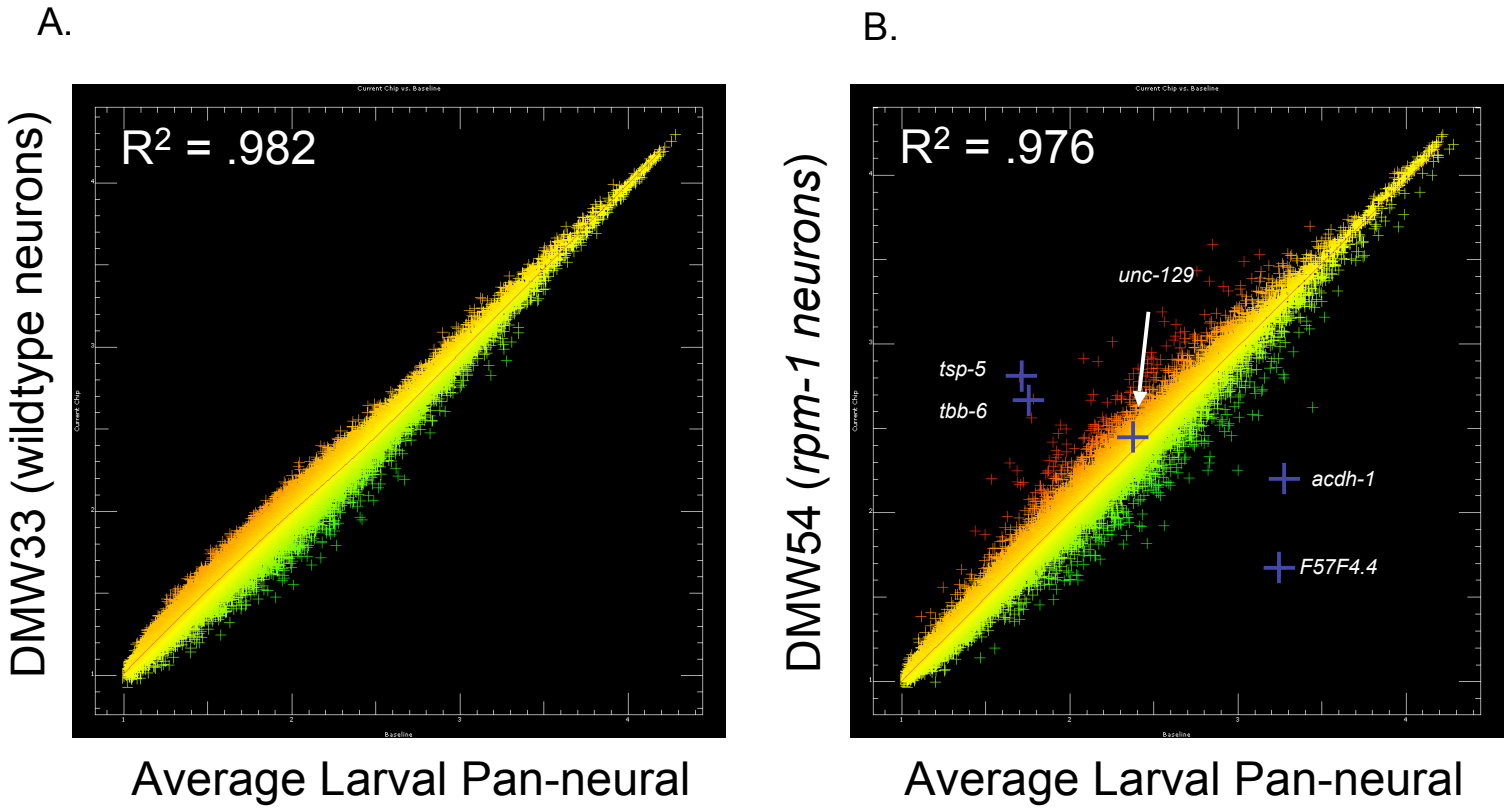


Figure 3.3 Profiles of the nervous system identify RPM-1-regulated genes.
A. A single wildtype neuronal profile is compared against the average wildtype neuronal profile. Note the tight correlation.
B. A single *rpm-1* neuronal profile is compared against the average wildtype neuronal profile. While these datasets are highly correlated, significantly enriched genes (red) and depleted genes (blue) are visible.

Discovery Rate (FDR) < 10%). As expected, *unc-129* is detected in this list of upregulated transcripts and therefore serves as a positive control for this data set. An additional 155 genes show decreased transcript levels (≤ 1/1.5 X) relative to wildtype in the *rpm-1* mutant background.

Gene families regulated by RPM-1

Genes from the *rpm-1* enriched and *rpm-1* depleted datasets were organized into groups based on KOGs (Figure 3.4). A large number of transcription factors are enriched in the *rpm-1* dataset, including 17 nuclear hormone receptors (NHR). In *C. elegans*, one member of the NHR family, *unc-55*, is necessary for defining appropriate synaptic inputs in GABAergic VD motor neurons. While *unc-55* is not regulated by *rpm-1*, the large number of NHRs regulated by *rpm-1* suggests that this family could play a role in synaptic development.

Surprisingly, the only synaptic protein regulated by RPM-1 is the SNARE syntaxin. Syntaxin plays a critical role in vesicle fusion. Increased levels of syntaxin could affect the rate of synaptic vesicle fusion, but previous experiments suggest that the overexpression of syntaxin does not affect synaptic structure or function. It seems unlikely, then, that the overexpression of syntaxin alone is responsible for *rpm-1* synaptic defects.

Approximately 150 genes are depleted in the *rpm-1* mutant background. This dataset includes a large number of transporter proteins (8), in addition to a large number of collagens (8). Rpm-1 mutants are shorter and fatter than normal animals (Dpy), and

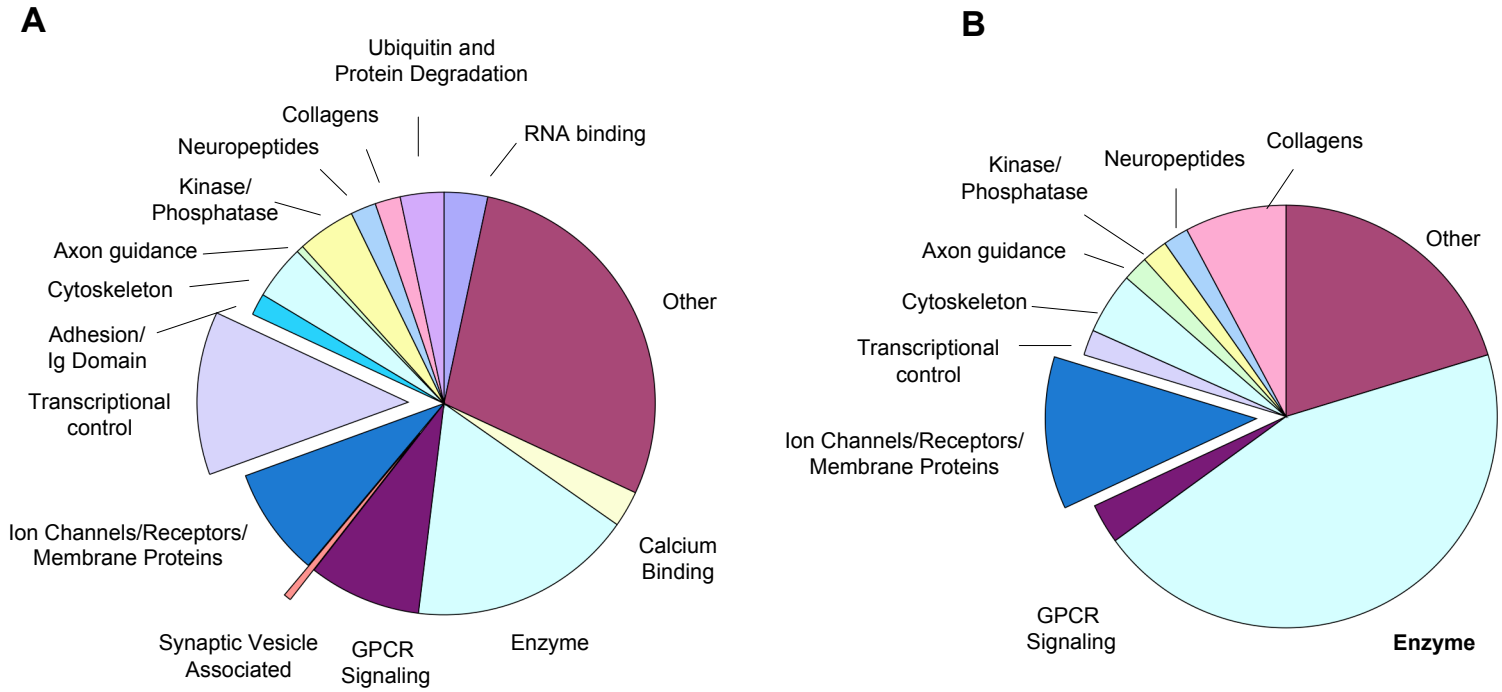


Figure 3.4 Gene families regulated by RPM-1

A. RPM-1 negatively regulates 523 genes. Dividing the genes into 15 subfamilies reveals the transcriptional regulation of 4 F-box proteins and 4 E3 ligases. Transcription factors are also highly enriched in *rpm-1* mutants (40), and almost half of these (17) are hormone receptor proteins. Very few known synaptic vesicle associated proteins are enriched.

B. RPM-1 positively regulates 155 genes. A large number of transporters are depleted in this dataset (8 of 12 Receptor proteins). Collagens are also depleted, consistent with body-size changes observed in rpm-1 mutants. A large number of enyzmes are also regulated.

changes in the expression of cuticle collagen could disrupt body size. A large number of enzymes are also depleted in the *rpm-1* dataset.

Confirmation of *rpm-1*-regulated transcripts with RT-PCR assays and with GFP reporter genes

Selected *rpm-1*-regulated transcripts (Table 3.1) in these microarray-derived lists were assayed by real-time PCR and/or promoter::GFP reporters to provide independent confirmation of RPM-1-dependent expression. Nilesh Kashikar, a rotation student in the Miller lab, assisted with these experiments. Actin (*act-1*) and GAPDH (*gpd-1*), two genes not regulated by *rpm-1* and broadly expressed in the worm, were selected as controls in order to normalize the data. Our microarray results indicate that the β-tubulin-encoding gene, *tbb-6*, is elevated ~15 fold in *rpm-1* neurons versus wild type (Fig 3.3A) (Table 3.1). Real-time PCR experiments confirmed a high level (40 fold) of enrichment for the *tbb-6* transcript in the *rpm-1* mutant neurons (Figure 3.5A). A similar result was obtained with a *tbb-6*::GFP reporter gene which shows increased intensity as well as ectopic expression in selected neurons in an *rpm-1* mutant. For example, in wildtype animals, neurons in the restrovesicular ganglion (RVG) do not express *tbb-6*::GFP (Fig 3.5B). In contrast, in *rpm-1 tbb-6*::GFP mutants, many RVG neurons are GFP positive (Fig 3.5C). *tbb-6*::GFP is also more highly expressed in all motor neurons in the ventral nerve cord (data not shown). These combined results strongly support the model that *rpm-1* functions as a negative regulator of *tbb-6* expression in the nervous system. In contrast to *tbb-6*, the novel transcript *F57F4.4* is highly depleted (~1/40x) in the microarray profile of *rpm-1* mutant neurons (Fig 3.3B). This result was also confirmed by Real Time

Table 3.1 Summary of GFP reporter and Real-time PCR data					
Gene	**Common name**	**Molecular identity**	**Microarray enrichment**	**Real-time PCR enrichment**	**GFP regulated by RPM-1**
T05C12.10	*qua-1*	Hedgehog-related	1.8	-	no
C27A2.6	*dsh-2*	Dsh	2.0x	-	no
K09A9.1	-	Protein Kinase related	2.3	-	no
R03D7.4	-	Transcription elongation factor	1.8	-	no
T02C12.1	*hum-5*	Myosin Heavy Chain	2.2	-	no
T04H1.9	*tbb-6*	Beta-tubulin	15x	40x	yes (Fig 3.4)
F57F4.4	-	unknown	0.25x	2000x	no (Fig 3.4)
C53D6.2	*unc-129*	TGF-Beta	1.6	-	yes (Fig 3.1)
K03F8.2	*acr-5*	Acetylcholine receptor	-	-	yes (Fig 2.5)
F33E2.2	*dlk-1*	MAP KKK	1x	1.6x	not tested
T09F3.3	*gpd-1**	GAPDH	1x	1x	-
T04C12.6	*act-1**	actin	1x	1x	-
	**Real time PCR controls*				

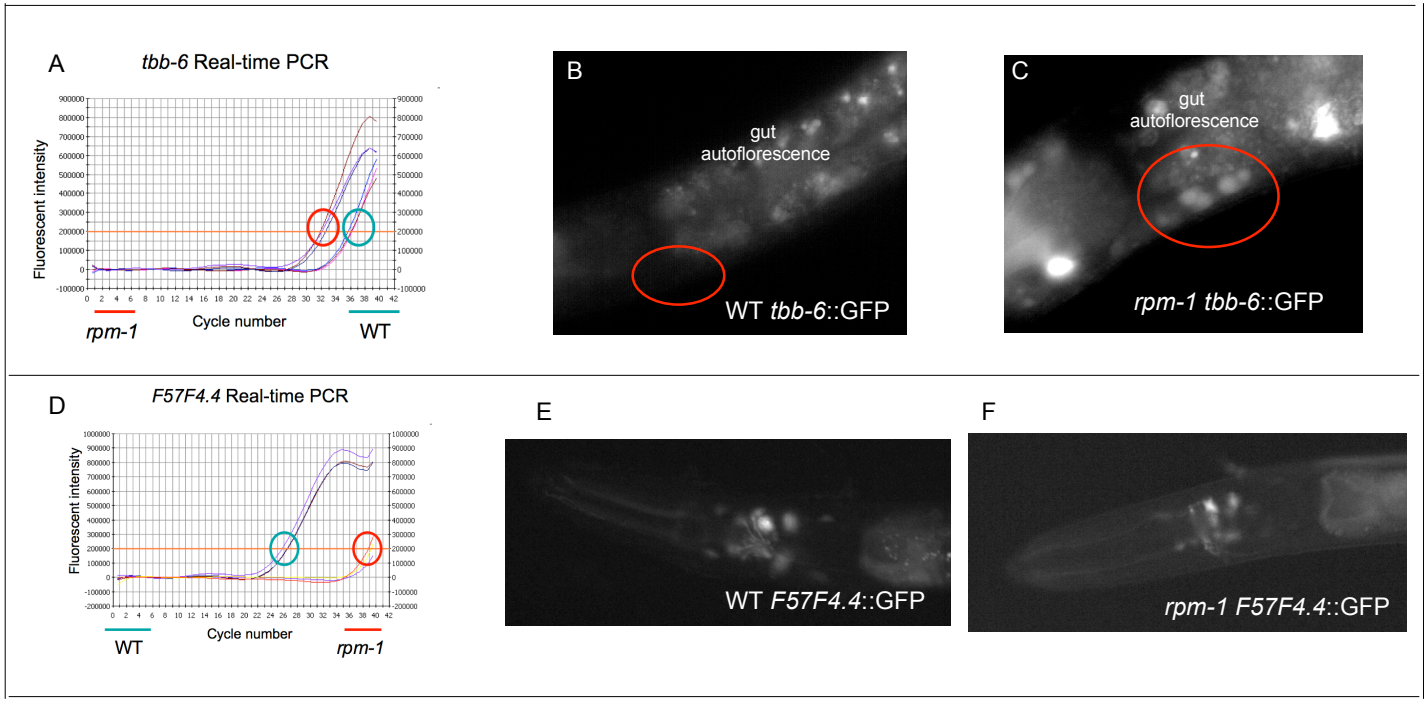


Figure 3.5 Microarray analysis identifies transcriptional targets of *rpm-1.*

A. Real-time PCR experiments confirm increased levels of *tbb-6* transcript in *rpm-1* (red) vs wildtype (blue) neural RNAs.
B. Wildtype *tbb-6*::GFP is not expressed in the RVG (red circle), and only weakly expressed in head and tail neurons (not shown).
C. tbb-6::GFP is ectopically expressed in the RVG (red circle), the VNC, and head and tail neurons.
D. Real-time PCR experiments confirm the positive regulation of *F57F4.4* by RPM-1.
E and F. *F57F4.4*::GFP expression is not changed in an *rpm-1* background. Head neurons are shown.

PCR, which shows that *F57F4.4* mRNA is decreased 2000 fold in *rpm-1* relative to the wildtype sample (Figure 3.5D). Expression levels of *F57F4.4*::GFP, however, were not reduced in an *rpm-1* mutant background (Figure 3.5E, F). This negative result could mean that key gene regulatory elements are not included in the *F57F4.4*::GFP construct. Although the MAPKKK *dlk-1*, shows modest enrichment (~1.6X) in the *rpm-1* microarray dataset, real-time PCR did not confirm this result. Promoter::GFP constructs for 5 additional candidate *rpm-1*-regulated transcripts (Table 1) also failed to show either enhanced or ectopic expression in the nervous system in *rpm-1(ok364)*. It may be significant that these negative results were obtained for candidate *rpm-1* regulated transcripts showing < 2.3 X enrichment in a data set with a relatively high false discovery rate (FDR $\leq$ 10%). In the future, it will be important to use both RT-PCR and GFP reporter genes to test additional candidate *rpm-1* regulated transcripts detected in the microarray experiments.

tbb-6 **is not required for synaptic defects in** ***rpm-1*** **mutant GABAergic motor neurons**

Our findings indicate that *rpm-1* is likely to regulate transcription of the β-tubulin encoding gene, *tbb-6.* This result is intriguing because other evidence indicates that Tubulin plays a pronounced role in synaptic development (Roos, Hummel *et al*. 2000; Zhang, Bailey *et al*. 2001; Broadie and Richmond 2002; Franco, Bogdanik *et al*. 2004; Ruiz-Canada, Ashley *et al*. 2004; Trotta, Orso *et al*. 2004; Gogel, Wakefield *et al*. 2006). MAP1B, the human homolgue of the microtubule binding protein Futsch, has been shown, in vitro, to bind microtubules and slow their depolyerization (Vandecandelaere, Pedrotti *et al*. 1996; Halpain and Dehmelt 2006). Mutations that disable Futsch, result in

reduced arborization and increase the size of the synapse in *Drosophila* (Roos, Hummel *et al*. 2000). Mutations in the *rpm-1* homolog, Highwire, on the other hand, increase arbor branching and decrease bouton size (Wan, DiAntonio *et al*. 2000). Since *tbb-6* is highly enriched in the *rpm-1* dataset, I hypothesized that this increase could produce the *rpm-1* synaptic defects. This model predicts that a loss-of-function mutation in *tbb-6* should rescue *rpm-1* synaptic defects in *C. elegans*. A *tbb-6* deletion allele was obtained from the *C. elegans* knockout consortium to test this idea; *tbb-6 (tm2004)* is viable and shows no obvious movement defects. An assay with a *unc-25*::SNB-1::GFP indicates that *tbb-6* mutants maintain normal GABA neuron synaptic number and morphology (Fig 3.6A, B). These GABAergic synapses are disrupted in *rpm-1* mutants (Zhen, Huang *et al*. 2000). This Rpm-1 mutant defect is not suppressed in *rpm-1 tbb-6* double mutant animals, however (data not shown). This result suggests that overexpression of *tbb-6* is not required for the *rpm-1* GABAergic motor neuron synaptic defects.

Overexpression of TBB-6 does not disrupt SAB axon morphology

Mutations in *rpm-1* lead to ectopic axon branching in touch neurons and in SAB motor neurons (Schaefer, Hadwiger *et al*. 2000). A potential role for *tbb-6* in this mutant phenotype is suggested by the observation that *tbb-6*::GFP is ectopically expressed in *rpm-1* mutant animals in a ganglion (RVG) that includes the SAB motor neurons. We speculated that the branching defects observed in *rpm-1* mutants (Schaefer, Hadwiger *et al*. 2000) could be due to the over expression of TBB-6. A transgenic line expressing full length CFP::TBB-6 in SAB motor neurons was constructed to test this idea with the

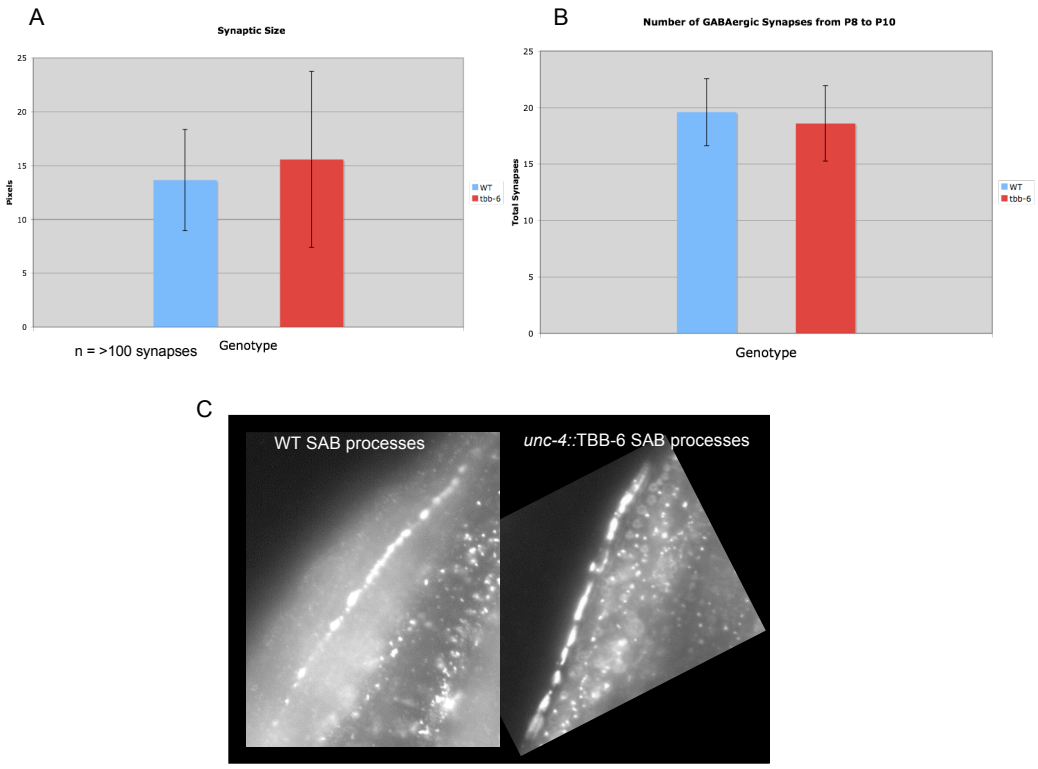


Figure 3.6 The *tbb-6* (*tm2004*) deletion mutant does not affect GABAergic or SAB synapses.

A. Synaptic size (in pixels) is unaffected in *tbb-6* (*tm2004*).
B. Synaptic number in the dorsal cord (p8 to p10) is also unaffected.
C. SAB synaptic morphology and process placement is unaffected in *tbb-6* mutants.

assistance of Kathie Watkins, an RAII in the Miller lab. As shown in Figure 3.6C, no detectable difference was observed between wildtype and *rpm-1* SAB processes. This negative result is difficult to interpret, however, because immunostaining was necessary to detect CFP expression in these neurons, which could mean that *tbb-6* levels are not sufficiently elevated in this transgenic line to produce an effect. It should be possible to resolve this question by asking if the deletion allele, *tbb-6 (tm2004)* suppresses the SAB axon branching defect in *rpm-1* mutants.

RNAi experiments reveal that the candidate *rpm-1* target gene, *acdh-1*, regulates *unc-129*::GFP expression

An RNAi screen was conducted to ask if other candidate *rpm-1*-regulated genes from the microarray experiment function downstream of *rpm-1* to control gene expression *in vivo*. This analysis was limited to transcripts that were regulated $\geq$ 2X in *rpm-1* mutants and also available in the Ahringer RNAi feeding library (78 enriched and 25 downregulated genes.) We reasoned that RNAi knockdown of transcripts with elevated expression in an *rpm-1* mutant background should rescue the Rpm-1 gene expression defect (*e.g.* ectopic *unc-129*::GFP) if they encode components (*e.g.* transcription factors) required for *rpm-1*-dependent transcription. Conversely, RNAi knockdown of *rpm-1*-depleted transcripts should mimic the Rpm-1 gene expression phenotype in a widltype animal. A mutation in the gene *eri-1*, which enhances sensitivity to RNAi was used to perform these experiments (Kennedy, Wang *et al*. 2004). The sensitivity of this assay to *rpm-1* function was confirmed by showing that RNAi knockdown of *rpm-1* results in ectopic *unc-129::*GFP expression (Fig 3.7A-C). However,

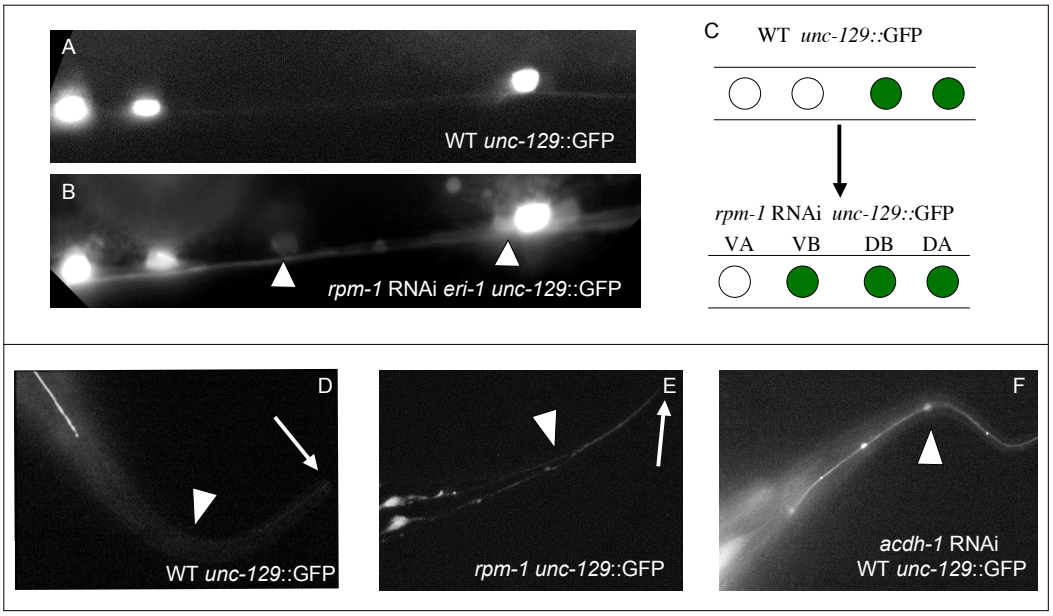


Figure 3.7 RNAi of the RPM-1-regulated gene *acdh-1* produces Rpm-1-like phenotypes

A, B, and C. RNAi of *rpm-1* produces gene expression defects. Arrowheads highlight ectopic VB neurons.

D and E. In addition to affecting *unc-129*::GFP expression in VBs, *rpm-1* also disrupts expression of *unc-129*::GFP in an unidentified tail neuron. The arrow points to the tip of the tail. Note that GFP expression in *rpm-1* mutants extends past the marked arrowhead and to the tip of the tail.

F. RNAi of *acdh-1* phenocopies *rpm-1*.

RNAi of upregulated genes did not suppress this defect in an *rpm-1* mutant and RNAi of only one gene in the depleted list, *acdh-1*, showed a weakly penetrant Rpm-1-like *unc-129::*GFP expression defect (~5% or 2/43 animals examined) (Figure 3.7D-F).

Mutations in *acdh-1* do not affect GABAergic motor neuron synaptic morphology

acdh-1 encodes a conserved short chain acyl-CoA dehydrogenase, a key enzyme in lipid metabolism (Wanders 2004) and is therefore a plausible candidate for a regulator of membrane-dependent events in the nervous system such as synaptic assembly. The *acdh-1* transcript is robustly downregulated (0.07X) in the *rpm-1* microarray dataset (Table 3.1). Since *rpm-1* and *acdh-1* negatively regulate *unc-129*::GFP expression, we speculated that *acdh-1* mutants could mimic the RPM-1 synaptic defect. However, the deletion mutant, *acdh-1 (ok1489)* does not disrupt GABA motor neuron synaptic morphology in a wildtype background. *acdh-1(ok1489)* also showed no effect in genetic background sensitized for reduced *rpm-1* function with the heterozygous strain, *rpm-1*(*ok364*)/+ (Figure 3.8). Given the weak effect of *acdh-1* on the *rpm-1* phenotype in VB motor neurons (*i.e.* ectopic *unc-129*::GFP) this negative result is not surprising. Since the RNAi effect of *acdh-1* on *unc-129*::GFP expression is incompletely penetrant (~5%) we speculate that *acdh-1* may function in combination with other genes to regulate synaptic morphology. For example, *acdh-1* is a member of a multigene family of closely related short chain acyl-CoA dehydrogenases (WormBase) with which it could potentially share redundant functions.

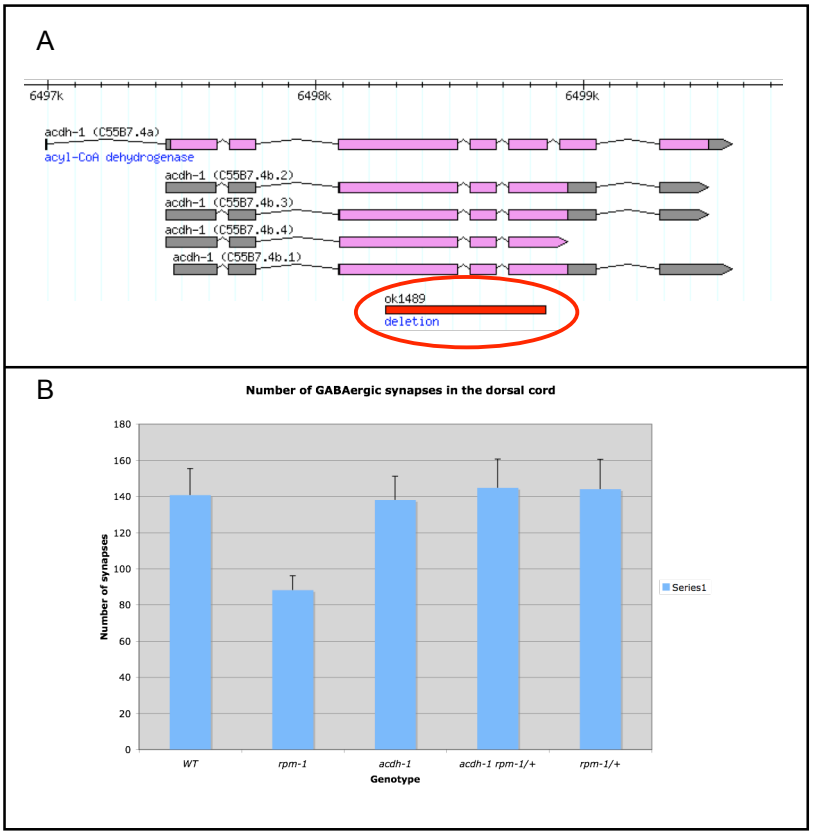


Figure 3.8 A deletion allele of *acdh-1* does not affect GABAergic synaptic morphology.

A. *acdh-1 (ok1489)* disrupts the *acdh-1* locus.

B. No synaptic defects are seen in *acdh-1 (ok1489)*. Also, *acdh-1* fails to produce synaptic defects in *rpm-1* heterozygotes.

A behavioral RNAi screen for *rpm-1* pathway genes

Although the *unc-129*::GFP reporter provides a useful method for assessing *rpm-1*-dependent gene expression, this assay is potentially insensitive to *rpm-1* target genes with direct roles in synaptic assembly and therefore may not have detected RNAi effects for these *rpm-1*-regulated transcripts in our microarray data set. As a complement to this assay, I developed a behavioral RNAi screen for *rpm-1* pathway genes. This approach is based on the synthetic "Unc" phenotype described above for *rpm-1 syd-2* double mutants. Second site mutations that reverse the downstream effects of the *rpm-1* mutation result in improved locomotion for this strain. Genetic suppressors of the *rpm-1* phenotype, the MAP Kinase components *pmk-3*, *dlk-1*, and *mkk-4*, for example, were isolated from a similar genetic background (i.e. *rpm-1 syd-1*) (Nakata, Abrams *et al*. 2005). We therefore reasoned that RNAi of genes enriched in an *rpm-1* microarray data set that are also required for the *rpm-1* synaptic defect should rescue the movement defect of *rpm-1 syd-2* double mutants. To perform these experiments, I constructed an RNAi sensitive line (*eri-1*; *rpm-1*; *syd-2 lin-15B*) (Kennedy, Wang *et al*. 2004; Sieburth, Ch'ng *et al*. 2005; Wang, Kennedy *et al*. 2005). As expected, RNAi of *pmk-3* suppressed the severe Unc phenotype of this strain resulting in worm tracks throughout the bacterial lawn (Figure 3.9B). I then tested most of the *rpm-1* transcriptional targets using the Ahringer RNAi library (523 enriched genes, 421 in the library, 396 tested) for suppression of the movement defect. RNAi of only one of the tested genes, *dlk-1* (MAP KKK), resulted in suppression of the *rpm-1* movement defect (Figure 3.9C).

To test the down-regulated dataset, I constructed an *eri-1*; *syd-2 lin-15B* triple mutant. In this case, we reasoned that RNAi of genes that are required for normal *rpm-1*

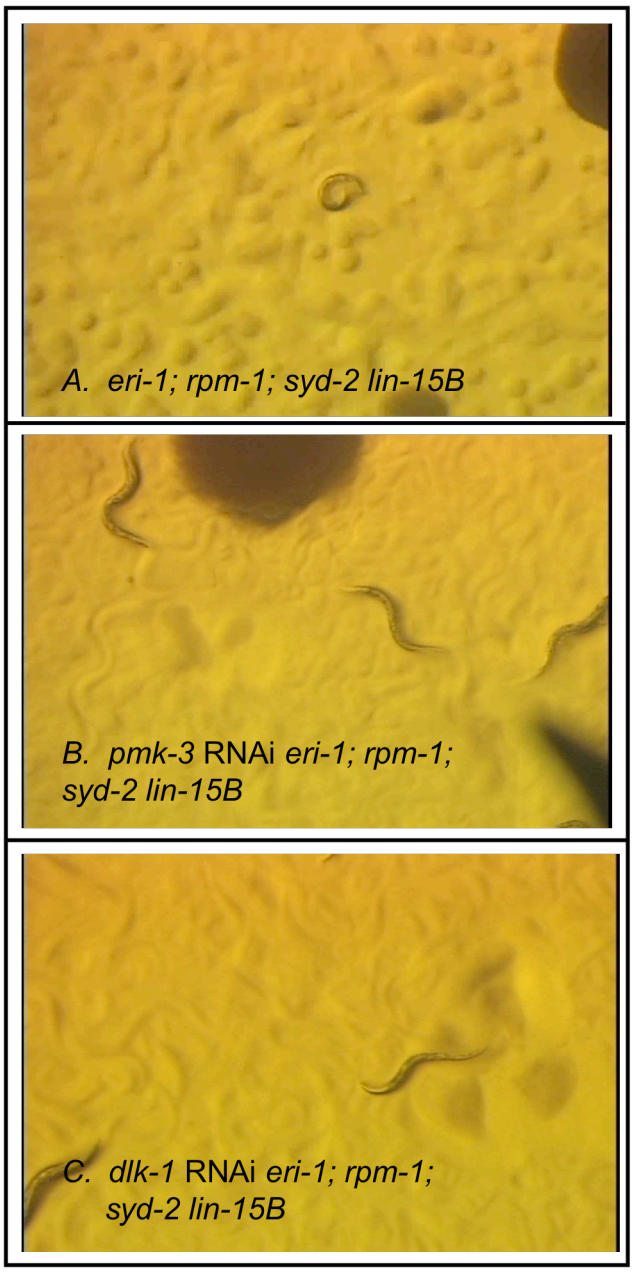


Figure 3.9. RNAi of *pmk-3* and *dlk-1* rescue *rpm-1 (ok364) syd-2 (ju37)* movement defects.

A. *rpm-1 syd-2* double mutants are severely uncoordinated.

B. RNAi of *pmk-3* rescues the Uncoordinated movement phenotype.

C. RNAi of transcripts regulated by *rpm-1* identified *dlk-1* as a downstream target.

pathway function should phenocopy the synthetic Unc phenotype of *rpm-1 syd-2* double mutants. (A duplicate RNAi assay of the *eri-1*; *lin-15B* line was used to exclude uncoordinated phenotypes that were not related to the *rpm-1* pathway.) These experiments were performed with the assistance of an IGP rotation student, Logan Dumitrescu. None of the genes tested in this assay (155 depleted, 101 in library, 79 tested) however, produced a synthetic Unc phenotype. This result rules out a role for these downregulated transcripts in the *rpm-1* dependent traits that interact with *syd-2* but leaves open the possibility that these transcripts could function in an *rpm-1* pathway in neurons that do not affect movement.

The *C. elegans fos* homolog is not required for *rpm-1* function

As described above, the transcription factor Fos is required for the *highwire* mutant phenotype in Drosophila. This result suggests that highwire indirectly inhibits fos activity by downregulating MAP Kinase signaling (Collins, Wairkar *et al*. 2006). We speculated that the *C. elegans fos* homologue, *fos-1*, could be similarly regulated by *rpm-1*. However, RNAi of *fos-1* failed to suppress the synthetic Unc phenotype of the RNAi sensitive *rpm-1 syd-2* mutant strain (data not shown). This result suggests that *fos-1* is unlikely to function in the *rpm-1* pathway and is in fact consistent with the observation that the highwire MAP Kinase pathway includes JNK MAPK which regulates fos whereas *rpm-1* controls p38 MAPK which typically regulates other classes of downstream transcription factors. Perhaps a genome wide RNAi screen of *C. elegans*

transcription factors genes (Reece-Hoyes, Deplancke *et al*. 2005) with the *rpm-1 syd-2* line could identify the presumptive downstream transcriptional regulators in the *rpm-1* pathway.

Discussion

Microarray analysis of *rpm-1* mutants identifies transcriptional targets

This work was motivated by our surprising discovery that *rpm-1* mutants perturb expression of specific GFP reporter genes in C. elegans neurons. This observation suggests the obvious possibility that RPM-1 regulates gene transcription to control synaptic assembly. We used new microarray based technology and identified 523 upregulated genes and 155 downregulated transcripts in the nervous system of *rpm-1* mutants. These results confirmed that the native *unc-129* transcript is upregulated, as originally suggested by our observation that *unc-129*::GFP is ectopically expressed in additional neurons in *rpm-1* mutants. A variety of gene families are regulated by *rpm-1* and indicate that, potentially, the need for a diverse set of downstream pathways for *rpm-1* dependent regulation of synaptic assembly.

We used Real Time PCR to confirm the regulation of two of the most highly regulated genes (*tbb-6* 40X up, *F57F4.4* 0.0005X down). Upregulation of *tbb-6* was confirmed with a GFP reporter. Genetic experiments, however, failed to confirm that *tbb-6* is required for neither the *rpm-1* GABAergic synaptic defects nor the axon branching defects observed in SAB neurons. *tbb-6* is still an attractive candidate because of the well established role for tubulin in synaptic morphology (Hummel, Krukkert *et al*.

2000; Roos, Hummel *et al*. 2000; Zhang, Bailey *et al*. 2001; Franco, Bogdanik *et al*. 2004; Ruiz-Canada, Ashley *et al*. 2004; Gogel, Wakefield *et al*. 2006; Yan and Broadie 2007). We need to look for *rpm-1* dependent defects in other classes of neurons in which *tbb-6* is ectopically expressed in *rpm-1* mutants

Since the *rpm-1* synaptic defect in GABA neurons is not sensitive to RNAi, we used other indirect strategies to screen genes identified by the microarray analysis. We initially used *unc-129*::GFP as a read out for *rpm-1* affects. RNAi of *acdh-1* weakly phenocopied *rpm-1* gene expression defects but the effect was very weak (~5% penetrance) and a genetic test with the *acdh-1* deletion mutant failed to confirm a role in *rpm-1* dependent synaptic assembly. The use of this screen is limited since it can only identify *rpm-1* pathway genes that control transcription and therefore would be insensitive to genes with diret roles in synaptic assembly. To overcome this limitation, we performed a functional screen based on genetic interactions with *syd-2*. This screen confirmed a role for *dlk-1* in the *rpm-1* pathway, but did not detect other genes in the microarray profile. We hypothesize that *rpm-1* may regulate *dlk-1* at both the protein and transcriptional level. In this model, *dlk-1* could function in a positive feedback loop by activating its own transcription during critical periods of synaptic development. A GFP reporter of *dlk-1* should clarify whether RPM-1 regulates transcriptional levels of *dlk-1*. As the first component of the MAP Kinase signaling pathway that mediates *rpm-1* function, dlk-1 is expected to show a strong genetic interaction with *syd-2* in this assay.

The large number of candidate *rpm-1* regulated transcripts in our microarray data sets suggests *rpm-1* controls a broad range of genes downstream of this MAPK pathway. Thus, genetic ablation of any single one of these potential RPM-1 effector genes might

not result in a strong synaptic phenotype. We also note, although RPM-1/Highwire/phr proteins are broadly expressed in neurons, genetic analysis suggests that the downstream targets of the regulatory pathways that they control may be specific to particular neuron classes (Schaefer, Hadwiger *et al*. 2000; Wan, DiAntonio *et al*. 2000; Zhen, Huang *et al*. 2000; Nakata, Abrams *et al*. 2005; Patel, Lehrman *et al*. 2006). For example, GABAergic synapses are clearly disrupted in *rpm-1* mutants whereas the neuromuscular synapses of nearby cholinergic motor neurons are only mildly affected (Nakata, Abrams *et al*. 2005). This evidence suggests that *rpm-1* regulated transcripts that we have identified in our microarray data sets may be exclusively regulated in subsets of specific neurons and that it will be necessary to observe these synapses in order to evaluate the potential roles of these genes in synaptic assembly.

Additional work is also needed to identify the specific transcription factor(s) that we hypothesize are functioning downstream of *rpm-1*. In *Drosophila*, mutations in the transcription factor fos rescue the synaptic defects of the *rpm-1* homologue Highwire (Collins, Wairkar *et al*. 2006). RNAi of the *C. elegans* orthologue of fos, however, does not rescue Rpm-1 synaptic defects. The finding that Fos does not regulate *rpm-1* dependent traits in *C. elegans* is consistent with the observation that *rpm-1* and Highwire regulate different members of the MAP Kinase family (Nakata, Abrams *et al*. 2005; Collins, Wairkar *et al*. 2006). For example, rather than inhibiting a p38 MAP Kinase, Highwire negatively regulates the MAP Kinase JNK which is an established activator of fos (Kyosseva 2004). While p38 has been associated with fos activation, p38 is also known to activate many additional transcription factors (Kyosseva 2004). In *C. elegans*, p38 may activate an alternative transcriptional cascade to modulate synaptogenesis.

CHAPTER IV

GENE EXPRESSION PROFILES OF *C. ELEGANS* NEURONS

This chapter has been provisionally accepted for publication in ***BMC Genome Biology.*** I am co-first author with Steve Von Stetina. My work on the paper involved generating and amplifying RNA for all mRNA-tagging lines, optimizing the mRNA-tagging procedure, scoring GFP lines generated from the dataset, constructing gene expression databases from wormbase, and a large portion of the data analysis. Steve initiated the project, generated the LA mRNA-tagging line, and was involved in almost all other aspects of the paper. Rebecca Fox generated the MAPCeL microarray data.

Introduction

The nematode *C. elegans* is a widely used model system for developmental studies. The major tissues of complex metazoans, e.g. muscle, intestine, nervous system, skin, etc. are represented in the worm, but the entire animal is comprised of fewer than 1,000 somatic cells. Owing to this simplicity and to the rapid development of the *C. elegans* body plan, the anatomy of every adult cell has been described and the patterns of division giving rise to each one are known (Sulston and Horvitz 1977; Sulston, Schierenberg *et al*. 1983). The *C. elegans* genome is fully sequenced (Consortium 1998; Hillier, Coulson *et al*. 2005) and encodes over 20,000 predicted genes. Thus, *C. elegans* offers a unique opportunity to identify specific combinations of genes that define the differentiation and structure of specific cell types. In principle, microarray profiles can

provide this information. In order to implement this approach, however, the small size of *C. elegans* (length = 1mm) has required the development of specialized methods for extracting mRNA from specific cell types. In one approach, MAPCeL (Micro-Array Profling of *C. elegans* cells), GFP-labeled cells are isolated by Fluorescence Activated Cell Sorting (FACS) from preparations of dissociated embryonic cells (Fox, Von Stetina *et al*. 2005). This method has now been utilized to profile global gene expression in specific subsets of neurons and muscle cells (Zhang, Ma *et al*. 2002; McKay, Johnsen *et al*. 2003; Colosimo, Brown *et al*. 2004; Blacque, Perens *et al*. 2005; Cinar, Keles *et al*. 2005; Fox, Von Stetina *et al*. 2005)(Fox *et al*, submitted). An alternative technique, mRNA-tagging (Roy, Stuart *et al*. 2002), can be utilized to profile larval cells which are not readily accessible for FACS. In this approach, an epitope-tagged mRNA binding protein (FLAG-PAB) is expressed transgenically with a specific promoter (Fig. 4.1). FLAG-PAB-bound transcripts are then immunoprecipitated for microarray analysis. mRNA-tagging profiles have been reported for two major tissues, body wall muscles and the intestine (Roy, Stuart *et al*. 2002; Pauli, Liu *et al*. 2005).

Here, we apply the MAPCeL and mRNA-tagging strategies to provide a comprehensive picture of gene expression in the embryonic and larval nervous systems. This analysis reveals ~2,500 transcripts that are significantly elevated in neurons versus other *C. elegans* cell types during these developmental periods. The enrichment in these datasets of transcripts known to be expressed in neurons, as well as newly-created GFP reporters from previously uncharacterized genes in these lists, confirmed the tissue specificity of our results. The “Pan-neural” transcripts detected in these datasets encode proteins with a wide array of molecular functions including ion channels,

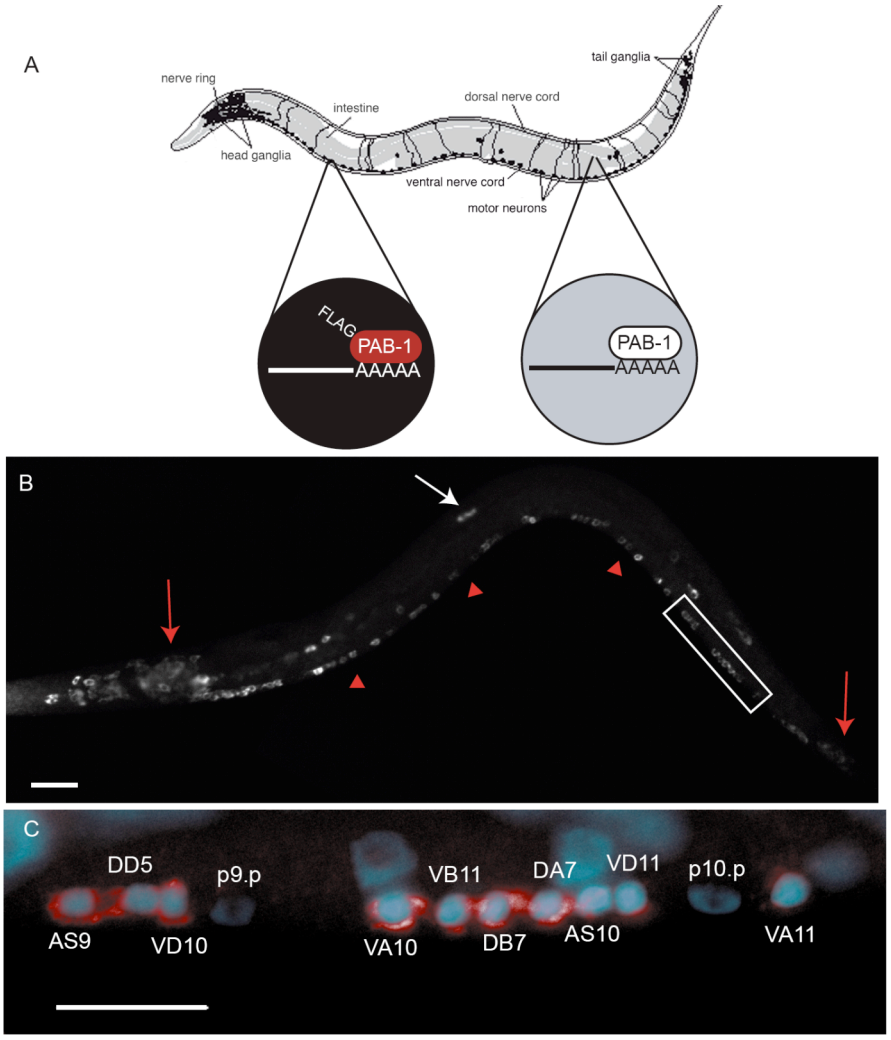


Figure 4.1 mRNA-tagging isolates neural specific transcripts.
A. The mRNA-tagging strategy for profiling gene expression in the *C. elegans* nervous system. A pan-neural promoter drives expression of FLAG-tagged poly-A binding protein (*F25B3.3*::FLAG-PAB-1) in neurons (black). Native PAB-1 is ubiquitously expressed in all cells (gray). Neural-specific transcripts are isolated by co-immunoprecipitation with anti-FLAG antibodies (artwork courtesy of Erik Jorgensen).
B. Immunostaining detects FLAG::PAB-1 expression in neurons in head and tail ganglia (red arrows), ventral nerve cord motor neurons (red arrowheads), and touch neurons (white arrow). Lateral view of L2 larvae. Anterior to left.
C. Close-up view of posterior ventral cord (boxed image in B), showing anti-FLAG staining (red) in cytoplasm surrounding motor neuron nuclei (e.g. AS9, DD5, etc.) stained with DAPI (blue). Note that hypodermal blast cells (P9p and P10p) do not show anti-FLAG staining.
Anterior is left, ventral is down. Scale bars = 10 μm.

neurotransmitter receptors and transcription factors. The discovery of 27 uncharacterized human homologs enriched in both embryonic and larval neurons, suggests that these profiles have uncovered novel genes with potentially conserved function in the nervous system.

In order to identify transcripts that are selectively expressed in a specific neural cell type, we used the mRNA-tagging strategy to fingerprint a subset of motor neurons (A-class) in the ventral nerve cord of L2 stage larvae. This A-class dataset contains ~400 significantly enriched genes. Approximately 25% of these transcripts are not detected in the profile of the entire nervous system. This finding suggests that individual neurons may express rare transcripts that are likely to be restricted to specific neuron types. The application of the mRNA-tagging strategy to profile a specific class of larval neurons complements earlier work in which this method was used to profile larval ciliated neurons (Kunitomo, Uesugi *et al*. 2005) and also experiments in which MAPCeL and other FACS-based approaches have been applied to selected embryonic neurons (Zhang, Ma *et al*. 2002; McKay, Johnsen *et al*. 2003; Colosimo, Brown *et al*. 2004; Blacque, Perens *et al*. 2005; Cinar, Keles *et al*. 2005; Fox, Von Stetina *et al*. 2005). Thus, this work demonstrates the utility of complementary profiling strategies that can now be applied to catalog gene expression in specific *C. elegans* neurons throughout development.

Materials and Methods

Nematode strains

Nematodes were grown as described (Brenner 1974). Strains were maintained on nematode growth media (NGM) plates inoculated with the *E. coli* strain OP50(Stiernagle). Strains used to isolate transcripts via mRNA-tagging were N2 (wildtype), SD1241 (*gaIs153*, *F25B3.3*::FLAG::PAB-1), NC694 (*wdEx257*, *unc-4*::3XFLAG::PAB-1)(Von Stetina, Fox *et al*. 2007). GFP-tagged embryonic neurons were isolated from NW1229 (*evIs111*, *F25B3.3*::GFP)(Altun-Gultekin, Andachi *et al*. 2001) (J. Culotti, personal communication) for MAPCeL analysis.

Molecular Biology

To create pPRSK29 (*F25B3.3*::FLAG::PAB-1), 4 kb of the F25B3.3 promoter upstream of the predicted ATG start was amplified using the following primers: Dp-5 (5' – GTC AAC TAG TGT ATG ATT CCT CG-3') and Dp-3 (5'-TCG GGG TAC CTA TCG TCG TCG TCG TCG ATG CCG TCT TCA CGA-3'). The predicted ATG start of *F25B3.3* was replaced with an Asp718 site in the 3' primer. This PCR fragment was cloned into pCR2.1-TOPO (Stratagene) to generate pPRSK29.1. pPRSK29.1 was digested with BamH1 and Asp718 to obtain the promoter fragment. pPRSK9 (*myo-3*::FLAG::PAB-1) (Roy, Stuart *et al*. 2002)was digested with Asp718 and SacI to obtain the FLAG::PAB-1 fragment. pBluescript SK was digested with SacI and BamHI, and a 3-way ligation was performed to obtain pPRSK29 (*F25B3.3*::FLAG::PAB-1).

Biolistic transformation and Transgenic Generation

pPRSK29 (60 ng/μL) was co-injected with pTG99 (*sur-5*::GFP, 20 ng/μL) using standard injection protocols(Mello and Fire 1995). The resulting transgenic array was integrated using a Stratalinker (Stratagene) at 300 Joules/m^2 (Evans (ed.)) (Shohei Mitani, personal communication).

Microparticle bombardment was conducted as described (Fox, Von Stetina *et al*. 2005). GFP reporters were selected at random from a subset of plasmids received from the Promoterome project (Dupuy, Li *et al*. 2004). Microparticle bombardment was conducted as described (Fox, Von Stetina *et al*. 2005).

Generating synchronized populations of L2 larvae for mRNA-tagging

Strains were grown to "starvation" (i.e., all dauer larvae) on ten 60 mm NGM plates at 25°C. Half of each 60 mm plate was split into four pieces and placed on a 150 mm 8P plate (Schachat, Garcea *et al*. 1978) inoculated with the *E. coli* strain Na22. The resultant twenty 8P plates were incubated at 25°C until a majority of the food was depleted and most animals were gravid adults (a "line" of worms is usually found at the retreating edge of the bacteria). The worms were removed from the plates with ice-cold M9 buffer (22mM KH_2PO_4, 22mM Na_2HPO_4, 85mM NaCl, 1mM $MgSO_4$) and collected by centrifugation. Washes were repeated until the supernatant was clear of bacteria. A sucrose float (30 ml ice cold M9 buffer, 20 ml cold 70% sucrose) was performed to create an axenic nematode suspension. Animals were washed 2X in ice-cold M9 buffer, then resuspended in 75 ml bleach solution (15 ml Chlorox, 3.75 ml 10N NaOH, 56.25 ml water). Worms were transferred to a 125 ml glass beaker with a stir bar and incubated for

5-6 min while stirring rapidly (solution turns a dark yellow when nearing completion). When a majority of adults burst, the solution was passed through a 53 μm nylon mesh (Fisher #08-670-201) to separate intact embryos from worm carcasses. Embryos were harvested by centrifugation and washed at least 3X with M9 buffer. Embryos were resuspended in RT M9 buffer and incubated on a nutator for 12-16 hours at 20°C to allow L1 larvae to hatch and arrest.

Arrested L1 larvae were collected by centrifugation. Animals were resuspended in 1 ml RT M9 buffer and split equally over six 150mm 8P plates. L1s were grown at 20°C for 22-25 hours to reach mid-L2, as evidenced by the appearance of the post-deirid sensory organ (~80%) (Sulston and Horvitz 1977). L2s (~0.3 ml - 1 ml) were harvested from 8P plates and sucrose floated as above. Worms were resuspended in 30 ml cold M9.

mRNA-tagging

Methods are identical to those in Roy, *et al*. 2002 with the following modifications. Synchronized L2 larvae were resuspended in 2-3 ml homogenization buffer (HB) [50 mM HEPES, pH 7.6; 150 mM NaCl; 10 mM $MgCl_2$; 1 mM EGTA, pH 8.0; 15 mM EDTA, pH 8.0; 0.6 mg/ml Heparin; 10% glycerol] and passed through a French press at 6000 psi. Total RNA was isolated from 100 μl of lysate. An amount of lysate equivalent to 200 μg total RNA was used for co-immunoprecipitation. Following co-immunoprecipitation, beads were washed 3x by brief treatment with 2 ml low-salt homogenization buffer (LSHB) [20 mM HEPES, pH 7.6; 25 mM NaCl; 1 mM EGTA, pH 8.0; 1 mM EDTA, pH 8.0; 0.6 mg/ml Heparin; 10% glycerol]. Beads were then washed 3x for 30 minutes in 2 ml LSHB. The LSHB treatment substantially reduced

non-specific RNA binding to the agarose beads (data not shown). Elution and mRNA extraction were performed as in Roy, *et al*., 2002. (See detailed protocol in Additional File 20).

Isolation of RNA from embryonic neurons for MAPCeL analysis

In the MAPCeL (Micro-Array Profiling of *C. elegans* Cells) method, GFP cells are isolated by FACS for microarray analysis. Primary cultures of embryonic cells were prepared (Christensen, Estevez *et al*. 2002) from a transgenic line expressing GFP in all neurons, NW1229 (*evIs111*, *F25B3.3*::GFP) (Altun-Gultekin, Andachi *et al*. 2001) (Joe Culotti, personal communication). After 24 hour in culture, GFP-labeled neurons were obtained by FACS and total RNA isolated as described (Fox, Von Stetina *et al*. 2005). Muscle profiling data used in Figs. 4.4 and 4.7 were obtained by MAPCeL of embryonic muscle cells after 24 hours in culture (M24 dataset) as described in Fox *et al*., submitted. The top 50 enriched genes in this dataset were selected on the basis of statistical rank.

RNA amplification and microarray data analysis

A *C. elegans* Affymetrix chip was used for all microarray experiments (See http://www.affymetrix.com for probe set information.). For mRNA-tagging experiments, 25 ng of co-immunoprecipitated RNA was amplified and labeled as previously described (Fox, Von Stetina *et al*. 2005). Larval Pan-neural (*F25B3.3*::FLAG::PAB-1) profiles were obtained in triplicate. Four independent Larval A-class motor neuron (*unc-4*::3XFLAG::PAB-1) profiles were obtained. Reference profiles were generated from low levels of non-specifically bound RNA obtained from mock IPs of synchronized

populations of wildtype (N2) L2 larvae. Five independent Reference datasets were obtained. 100 ng of total RNA was amplified and labeled for the MAPCeL sample, *F25B3.3*::GFP, isolated in triplicate. A previously obtained profile of total RNA isolated from all viable embryonic cells in culture was used as a MAPCeL Reference (Fox, Von Stetina *et al*. 2005).

Hybridization intensities for each experiment were scaled by reference to a global average signal from the same array (Additional Files 25, 26) and normalized by RMA (Robust Multi-Array Analysis) (Additional Files 27, 28). We identified transcripts in two categories: (1) Expressed Genes (EGs), or transcripts that are reliably detected in a given sample; (2) Enriched genes or transcripts with intensity values that are significantly higher than Reference samples. EGs were estimated for the mRNA-tagging samples as follows. Expressed transcripts in the *F25B3.3*::FLAG::PAB-1 (Larval Pan-neural) and the *unc-4*::3XFLAG::PAB-1 (Larval A-class motor neurons) were initially identified on the basis of a "Present" call in a majority (e.g. 2/3) of experiments as determined by Affymetrix MAS 5.0. In this approach, genes are called "Absent" and therefore excluded when the Mismatch (MM) value exceeds the Perfect Match (PM) intensity for a given gene. This analysis initially identified 8084 "Present" transcripts in the Larval Pan-neural sample and 7578 transcripts in the Larval A-class motor neuron sample (Additional File 21). These lists, however, are likely to include mRNAs that are non-specifically bound to the anti-FLAG sepharose beads at low levels relative to *bona fide* neuronal transcripts (see below). We reasoned that transcripts included in the Experimental samples that are actually derived from this non-specific pool should be generally detected in the Reference sample at higher intensity values. Therefore, to exclude these non-specific mRNAs from

the list of predicted neuronal genes, the average RMA-normalized intensity for each transcript in the Reference sample was subtracted from the RMA value of the corresponding gene in the Experimental sample. Transcripts with resultant positive values were considered EGs whereas transcripts with negative values after this operation were removed. In a final adjustment, a limited number of transcripts that are detected as neuronally enriched (see below) but not scored as Present by MAS 5.0 were restored to the lists. This treatment identified 4033 EGs in the Larval Pan-neural dataset and 3320 EGs in the Larval A-class motor neuron profile. EGs (7953) for the MAPCeL Embryonic Pan-neural dataset were identified as previously described(Fox, Von Stetina *et al*. 2005).

To detect neuronally-enriched transcripts, RMA-normalized intensities for Experimental vs Reference samples were statistically analyzed using Significance Analysis of Microarrays software (SAM, Stanford). A two-class unpaired analysis of the data was performed to identify genes that differ by $\geq$ 1.5 fold from the Reference at a False Discovery Rate (FDR) of < 1% for the Larval Pan-neural, Embryonic Pan-neural, and Larval A-class motor neuron datasets. These genes were considered significantly enriched.

RMA normalized intensity values for all datasets were imported into GeneSpring GX 7.3 to generate the line graphs shown in Figs 4.4 and 4.7. Each Experimental dataset was paired to its corresponding Reference dataset for these diagrams.

Annotation of datasets

We utilized Perl scripts and hand annotation to identify all known neuronally expressed *C. elegans* transcripts [WormBase Release 146 (WS146)]. First, WormMart was used to identify all transcripts with expression patterns. This list was filtered for genes represented on the Affymetrix microarray. For genes that have multiple spots on the microarray, only one representative spot was kept in the list (3044). Genes with expression patterns with no spatial information or exclusive to males were eliminated (2837). Each gene was then placed into two categories based on its known expression pattern: neural (1612) vs non-neural (1225) using the following criteria: We used a Perl script ('keyword_search.pl', Additional File 22) to search descriptions of 2837 genes with known expression patterns for genes with defined neural expression. To reduce the number of false positives identified, we first searched under the term 'cell group', which provides simple, but clear, spatial expression information. Using this strategy, the majority of neuronally expressed genes were separated from the full dataset. Several genes in WormBase, however, had no cell group, or contained insufficient data in the cell group description to determine neural expression. Therefore, WormBase was also searched for terms associated with neuronal expression. This list was hand-annotated to ensure its validity (for a full list of search terms, see Additonal File 23).

Hypergeometric calculations

Overlap statistics were calculated using web-based software designed by Jim Lund (University of Kentucky). The website is:

http://elegans.uky.edu/MA/progs/overlap_stats.html. The number of genes in the genome was set at 18666 (total number of genes represented on the *C. elegans* Affymetrix array). When using this calculation, a representation factor below 1.0 indicates under-representation, while a value above 1 indicates over-representation.

Microscopy and identification of GFP expressing cells.

GFP expressing animals were visualized by differential interference contrast (DIC) and epifluorescence microscopy using either a Zeiss Axioplan or Axiovert compound microscope. Digital images were recorded with CCD cameras (ORCA I, ORCA ER, Hammatsu Corporation, Bridgewater, NJ).

C. elegans interactome

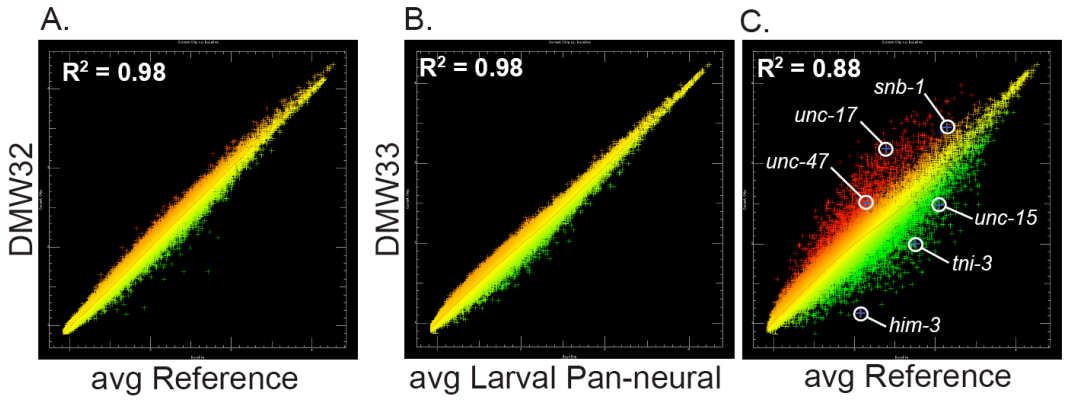
Genes enriched in both the larval and embryonic pan-neural datasets were used to seed the *C. elegans* Interactome (Li, Armstrong *et al.* 2004) (http://vidal.dfci.harvard.edu/interactomedb/i-View/interactomeCurrent.pl). The map was trimmed to exclude genes that with one interacting partner. The initial dataset consisted of 710 genes, of which 17% (124) were listed in the Interactome database. One large cluster of 34 interactors was identified, containing 17 proteins from the original seed. The additional 17 genes were categorized as enriched, expressed, or not present in the Pan-neural datasets. Genes were assigned to categories based on their known or predicted functions in *C. elegans* and other organisms.

Results

Neuronal mRNA-tagging yields reproducible microarray expression profiles

To profile gene expression throughout the nervous system, we generated a stable, chromosomally integrated transgenic line expressing an epitope-tagged Poly-A binding protein (FLAG::PAB-1) in all neurons. Pan-neuronal expression was confirmed by immunostaining with a FLAG-specific antibody (Fig. 4.1). We selected the 2nd larval stage (L2) to test the application of the mRNA-tagging method. At this stage, the nervous system is largely in place and therefore should express a broad array of transcripts that define the development and function of most neurons. Sub-microgram quantities of mRNA isolated by the mRNA-tagging method were amplified and labeled for application to an Affymetrix chip representing ~90% of predicted *C. elegans* genes. Neuron-enriched transcripts in these samples were detected by comparison to a Reference profile of all larval cells (See Methods). We reasoned that this approach should detect a significant fraction of known neuronal transcripts and thus provide an initial test of the specificity of this strategy.

Comparisons of independently derived datasets for both the Experimental (Larval Pan-neural) and Reference samples showed that individual replicates for each condition are highly reproducible (Fig. 4.2 A,B). For example, an average coefficient of determination (R^2) of ~0.96 was calculated from pairwise combinations of each individual Reference dataset (Fig. 4.2D). The Pan-neural datasets were similarly reproducible (R^2 ~ 0.96) (Fig. 4.2E). The overall concurrence of these data are graphically illustrated in the scatter plots shown in Figures 4.2A and 4.2B.



D. Reference hybridizations

	DMW15	DMW20	DMW21	DMW32
DMW15				
DMW20	0.97			
DMW21	0.96	0.98		
DMW32	0.95	0.97	0.96	
DMW41	0.95	0.97	0.97	0.97

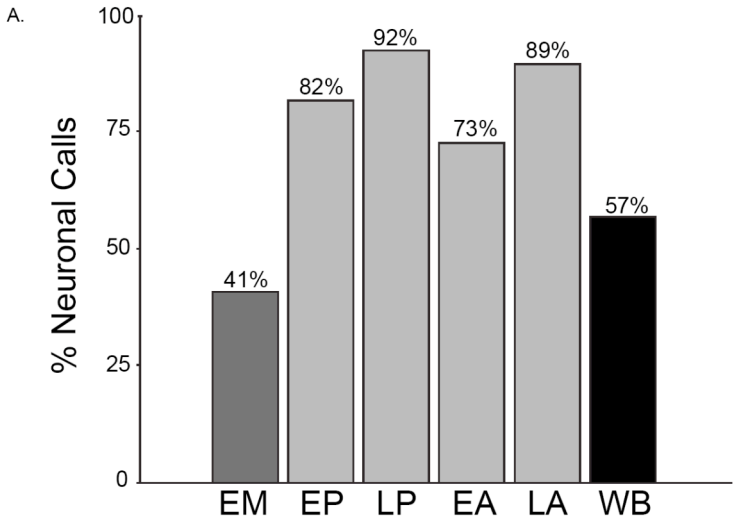
E. Larval Pan-neural hybridizations

	DMW33	DMW42
DMW33		
DMW42	0.95	
DMW43	0.95	0.98

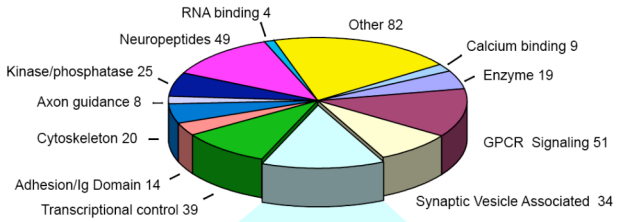
Figure 4.2. Microarray profiles reveal transcripts enriched in *C. elegans* neurons.
A. Scatter plot of intensity values (log base 2) for representative hybridization (DMW32) of RNA isolated from all larval cells (Reference) by mRNA-tagging compared to the average intensity of the Reference dataset.
B. Scatter plot of a representative Larval Pan-neural hybridization (DMW33) compared to the average intensities for all three Larval Pan-neural hybridizations.
C. Results of a single Larval Pan-neural hybridization (DMW33) (red) compared to average Reference intensities (green) to identify differentially expressed transcripts. Known neural genes *snb-1* (Synaptobrevin, all neurons), *unc-17* (VAChT, cholinergic neurons), and *unc-47* (VGAT, GABAergic neurons) are enriched (red). Depleted genes include two muscle-specific transcripts (*unc-15*, paramyosin and *tni-3*, troponin) and a germline-specific gene (*him-3*) (green).
D-E. Pairwise comparisons of individual hybridizations. Coefficient of determination (R^2) values for **D.,** all pairwise combinations of Reference hybridizations and for **E.,** all pairwise combinations of Larval Pan-neural hybridizations indicate reproducible results for both Reference and Experimental samples.

Transcripts detected by neuronal mRNA-tagging are expressed in neurons

Scatter plots comparing Larval Pan-neural vs Reference data revealed a substantial number of transcripts with significant differences in hybridization intensities (Fig. 4.2C). Statistical analysis detected 1562 transcripts (≥1.5X, ≤1% (False Discovery Rate, FDR) with elevated expression in the Larval Pan-neural sample (Additional File 1). Strikingly, we found that 92% of the 443 genes with known expression patterns included in the Larval Pan-neural enriched dataset (409/443) are listed in WormBase as neuronally expressed (Fig. 4.3A) (Additional File 1). By contrast, only ~57% of all genes (1612/2837) with defined expression patterns in WormBase (www.wormbase.org) are annotated as expressed in neurons (See Methods) (Fig. 4.3A) (Additional Files 2, 3). Moreover, transcripts with key roles in neuronal function are highly-represented in this list. For example, 55 transcripts encoding ion channels, receptors or membrane proteins with known expression in the *C. elegans* nervous system are enriched (Fig. 4.3B) (Additional File 7). The enrichment of transcripts known to be expressed in neurons demonstrates that the Larval Pan-neural profile is largely derived from neural tissue. This conclusion is also substantiated by the finding that transcripts that are highly expressed in other cell types are preferentially excluded from this dataset (e.g., Fig. 4.2C). For example, microarray profiling experiments identified a total of 1926 transcripts enriched in either larval germline, muscle or intestinal cells (GMI) (Additional File 5) (Pauli, Liu *et al*. 2005). This set of genes is significantly under-represented (97/1562) in the Larval Pan-neural dataset (representation factor = 0.6, $p<2.033e^{-9}$, a representation factor less than 1 indicates under-representation, see methods). Of the 97 genes that intersect our Larval Pan-neural profile and the GMI set, 35 have a previously characterized spatial



B.



Ion Channel/Receptor/Membrane Protein 55

Gene	Description	Fold Change
glr-5	Ionotropic glutamate receptor	8.8
dat-1	Dopamine transporter	6.7
F10E7.9	Na+/K+ symporter	5.8
sol-1	CUB domain protein	5.5
des-2	nAChR	4.2
glr-1	Ionotropic glutamate receptor	4.2
F08A10.1a	Ca++ activated K+ channel	3.9
R13A5.9	Predicted transporter	3.9
mtd-1	Novel transmembrane protein	3.8
glr-4	Ionotropic glutamate receptor	3.4
deg-3	nAChR	3.2
nmr-2	Ionotropic glutamate receptor	3.2
twk-29	Twik K+ channel	3.0
inx-4	Innexin	3.0
mec-2	Stomatin	2.9
glc-4	glutamate gated chloride channel	2.9
unc-2	Calcium channel	2.9
mod-1	Ligand-gated ion channel	2.9
clh-2	Chloride channel	2.9
cho-1	Choline transporter	2.8

Figure 4.3 Microarray profiles detect known *C. elegans* neural genes.
A. Histogram showing fraction of annotated genes in microarray datasets with known *in vivo* expression in neurons for each neuronal microarray dataset. The list of annotated genes used for this comparison includes all genes with known cellular expression patterns listed in WormBase (see Methods). Note significant enrichment for neuronal genes in microarray datasets obtained from neurons (73-92%) relative to the fraction of all annotated genes in WormBase (57%) and embryonic muscle (41%) that show some expression in the nervous system. Microarray datasets are: EM = Embryonic Muscle datasets; EP = Embryonic Pan-neural datasets; LP = Larval Pan-neural datasets; EA = Embryonic A-class motor neuron datasets; LA = Larval A-class motor neuron datasets; WB = WormBase.
B. The Larval Pan-neural enriched dataset (LP) contains 443 transcripts previously annotated as expressed in neurons in WormBase. Genes were grouped according to functional categories characteristic of neurons. The top 20 enriched ion channel/receptor/membrane proteins are featured. (Additional File 7)

expression pattern. Of these, 89% (31/35) are also expressed in neurons. A comparison of the top 50 most significantly enriched transcripts in a MAPCeL profile of embryonic body wall muscle cells (Fox *et al*., submitted) detects only four transcripts that also show elevated expression in the Larval Pan-neural profile (Fig. 4.4A) (Additional File 6). Independent results have confirmed that at least one of these, the acetylcholine receptor subunit *acr-16*, is expressed in both muscle and neurons (Francis, Evans *et al*. 2005; Touroutine, Fox *et al*. 2005). The apparent low frequency of false positives empirically defined by these comparisons is consistent with the estimated FDR $\leq$ 1% for this dataset. The stringent exclusion of non-neuronal transcripts has been achieved, however, while retaining sensitivity to transcripts that may be expressed in limited numbers of neurons (Fig. 4.5). For example, our methodology identifies genes that are expressed in only two neurons; *daf-7* (TGF-beta-like peptide expressed in ASIL and ASIR) (Ren, Lim *et al*. 1996) and *gcy-8* (guanylate cyclase expressed in AFDL and AFDR) (Inada, Ito *et al*. 2006)(Fig. 4.5).

The strong enrichment of known neuronal genes in the Larval Pan neural dataset indicates that other previously uncharacterized transcripts in this list are also likely to be expressed in the nervous system. To test this prediction, we evaluated GFP reporter genes for representative transcripts in this profile. As shown in Table 4.1, all but one of the transgenic lines (24 of 25) derived from these promoter GFP fusions show expression in neurons (Fig. 4.6). 56% of the GFP reporters tested (14/25) are exclusively detected in neurons. For example, the stomatin gene, *sto-4*, is highly expressed in ventral cord motor neurons, touch neurons and in head and tail ganglia (Table 4.1) (Fig. 4.6D, H). Our

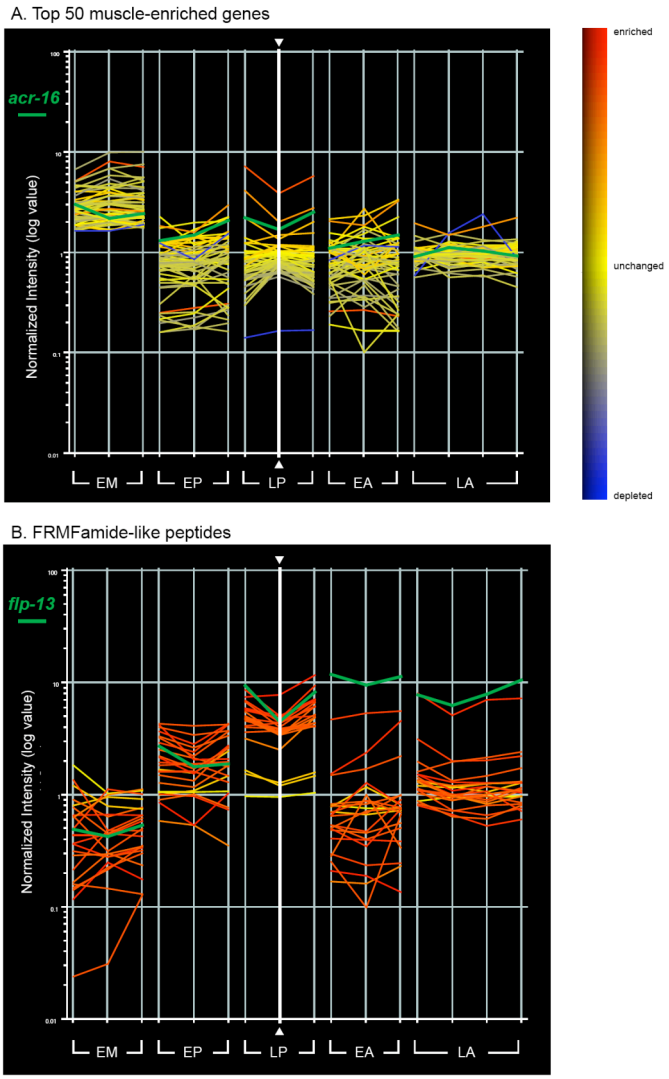
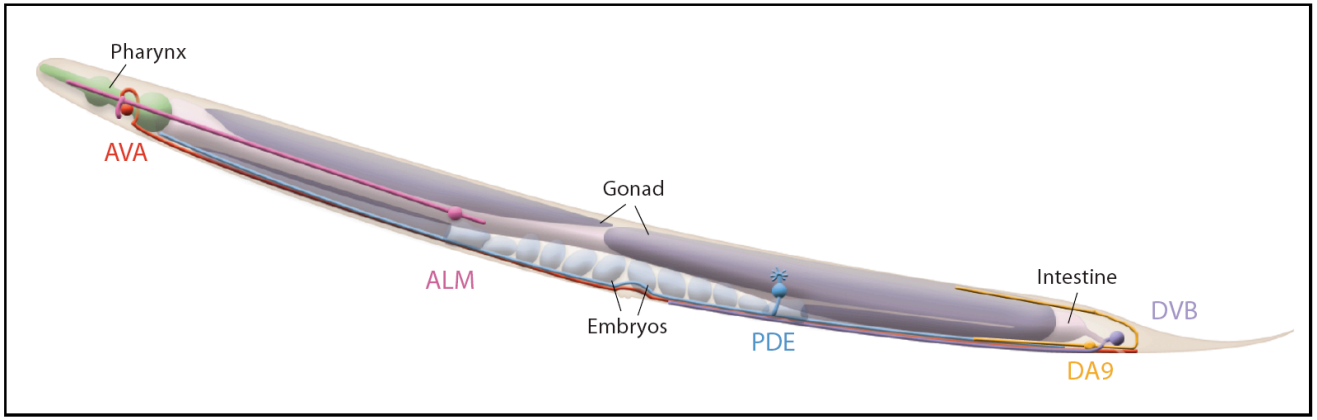
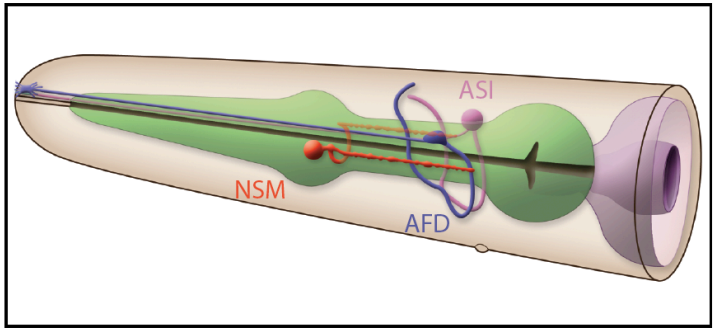


Figure 4.4 Neuropeptides are highly represented in profiles of neural cells while transcripts highly enriched in body wall muscle are excluded. Line graphs display qualitative relative intensity values (Experimental/Reference) for selected genes based on the log of the normalized intensity ratio (experimental/control) for each feature on the *C. elegans* Affymetrix array (see Methods). Each vertical line represents an individual replicate for each Experimental sample. Thus, trends in expression levels for a particular gene or sets of genes can be visualized across all datasets. Abbreviations as defined in Fig. 3A. Horizontal lines are colored (see heat map at right) according to relative enrichment of a single Larval Pan-neural (LP) replicate (vertical white line with arrowheads). The heat map on the right denotes the colors of the horizontal lines; red indicates high intensity ratio, green is a low intensity ratio, and yellow denotes no change. The bright green vertical line marks one larval pan-neural dataset, which defines the heat map upon which all the other datasets are based.
A. The top-50 ranked genes from embryonic muscle show limited enrichment in neuronal datasets. One exception is *acr-16*, marked by the horizontal green line, which is highly enriched in the LP dataset. *acr-16* encodes a nicotinic acetylcholine receptor that is expressed in both muscle cells and neurons [14].
B. FRMF-amide-like peptides (*flp*) are enriched in neurons. A majority (20/23) of *the 23 defined flp* transcripts are enriched in the LP dataset, while specific subsets of *flp* transcripts are enriched in the other neuronal datasets (EP, EA, LA). Importantly, most are excluded from the muscle (EM) dataset. The horizontal green highlights *flp-13*, which is the most highly enriched *flp* transcript in the A-class motor neuron (EA, LA) datasets.





Neuron	Gene name	Function
AFD	*gcy-8**	Thermosensory
ASI	*daf-7**	Chemosensory
NSM	*eat-4*	Serotonergic
AVA	*glr-1*	Glutamatergic
PDE	*cat-1*	Dopaminergic
DVB	*unc-25*	GABAergic
DA9	*unc-17*	Cholinergic
ALM	*mec-2*	Mechanosensory

Figure 4.5 Pan-neural datasets detect neuron-specific transcripts. A representation of transcripts enriched in the larval pan-neural dataset and a subset of the neurons in which those transcripts are expressed. **A.** Lateral view of an adult worm depicting selected neurons labeled and their morphologies. The U-shaped germline, spermatheca and embryos are colored various shades of gray. Ventral is down, anterior is to the left. **B.** Close-up of the adult head, showing the serotonergic neuron NSM and the cell bodies and processes for two sensory neurons, AFD and ASI. For simplicity, only one of the two pairs of neurons is diagrammed. The pharynx is colored green and the anterior end of the intestine is gray. **C.** Table displaying representative genes enriched in the Larval Pan-neural dataset (LP) and expressed in each indicated neuron. * denotes exclusive expression in the listed cell type. (Artwork courtesy of Zeynep Altun, Chris Crocker and David Hall, www.WormAtlas.org).

Table 4.1. Expression of promoter-GFP reporters for transcripts enriched in larval Pan-neural or A-class motor neuron data sets.

			Pan-neural			A-class	
Cosmid	**Gene**	**Protein**	**EP Fold Change**	**LP Fold Change**	**In Neurons?**	**Fold Change**	**UNC-4 neuron(s)**
C01G6.4		Predicted E3 ubiquitin ligase	1.8	--	√	--	**VA, DA**
VF11C1L.1	*ppk-3*	PIP kinase	1.8	--	√	--	**VA, DA**
C25D7.8		novel	1.9	--	√	--	**VA, DA**
F08G12.1			3.0	--	√	--	**VA, DA**
M79.1	*abl-1*	Abelson kinase	2.3	--	√	--	**VA, DA**
F25G6.4	*acr-15*	Acetylcholine receptor	--	4.9	√	--	**VA, DA**
T27A1.6[1]	*mab-9*	Transcription factor	--	1.7	√	--	**DA**
F39G3.8[1]	*tig-2*	TGF-β	--	1.8	√	--	**VA, DA**
T19C4.5[1]		Novel	--	2.0		--	
CC4.2[1,3]	*nlp-15*	neuropeptide	--	6.5	√	--	
C18H9.7[1]	*rpy-1*	rapsyn	--	2.7	√	--	**DA**
Y71D11A.5		Ligand-gated ion channel	2.1	1.8	√	--	
C04E12.7		Phospholipid scramblase	--	3.2	√	1.8	**VA, DA**
F36A2.4[1,4]	*twk-30*	K+ channel	--	2.1	√	5.1	**VA, DA**
Y71H9A.3	*sto-4*	stomatin	--	3.0	√	1.6	**VA**
F29G6.2[1]		Novel	--	3.2	√	1.6	**VA, DA, SAB, I5, AVF**
C44B11.3	*mec-12*	Alpha-tubulin	--	5.9	√	1.9	**VA, DA**
T23D8.2[1]	*tsp-7*	tetraspanin	--	3.5	√	4.8	**VA, DA**
T05C12.2[1]	*acr-14*	Acetylcholine receptor	--	1.5	√	3.1	**DA**
F33D4.3[1,2]	*flp-13*	neuropeptide	--	7.1	√	7.9	**I5**
C11D2.6[1]	*nca-1*	Ca++ channel	--	2.3	√	2.2	**VA, DA**
E03D2.2[1,3]	*nlp-9*	neuropeptide	--	3.1	√	2.5	**VA**
F55C12.4[1]		novel	--	3.5	√	2.1	**DA**
F43C9.4[1,5]	*mig-13*	CUB domain	--	1.8	√	2.8	**VA, DA**
F39B2.8		Predicted membrane protein	1.7	3.5	√	2.1	**VA, DA**
K02E10.8[1]	*syg-1*	Ig domain	1.8	1.8	√	1.8	**VA, DA**

ZC21.2[1]	*trp-1*	Ca++ channel	1.9	2.2	√	1.9	**VA, DA**
Y47D3B.2a [1,6]	*nlp-21*	neuropeptide	3.9	8.3	√	3.7	**VA, DA**
F09C3.2[1]		phosphatase	1.9	2.7	√	1.7	**VA, DA**
T27E9.9		Ligand-gated ion channel	2.3	4.0	√	3.1	
Y34D9B.1	*mig-1*	Frizzled-like	--	--	√	1.6	**VA, DA**

GFP expression in neurons (check mark), in A-class motor neurons (DA, VA, SAB, I5). GFP expression was typically determined in L2 larvae. Full expression pattens can be found in Additional File 17. Expression patterns for some of these GFP reporters have been previously reported: 1. Fox, RM et al. 2005. 2. Kim, K and LI, C, 2004. 3. Li, C et al., 1999. 4. Salkoff L, et al., 2001. 5. Sym, M et al., 1999. 6. Nathoo, AN et al., 2001.

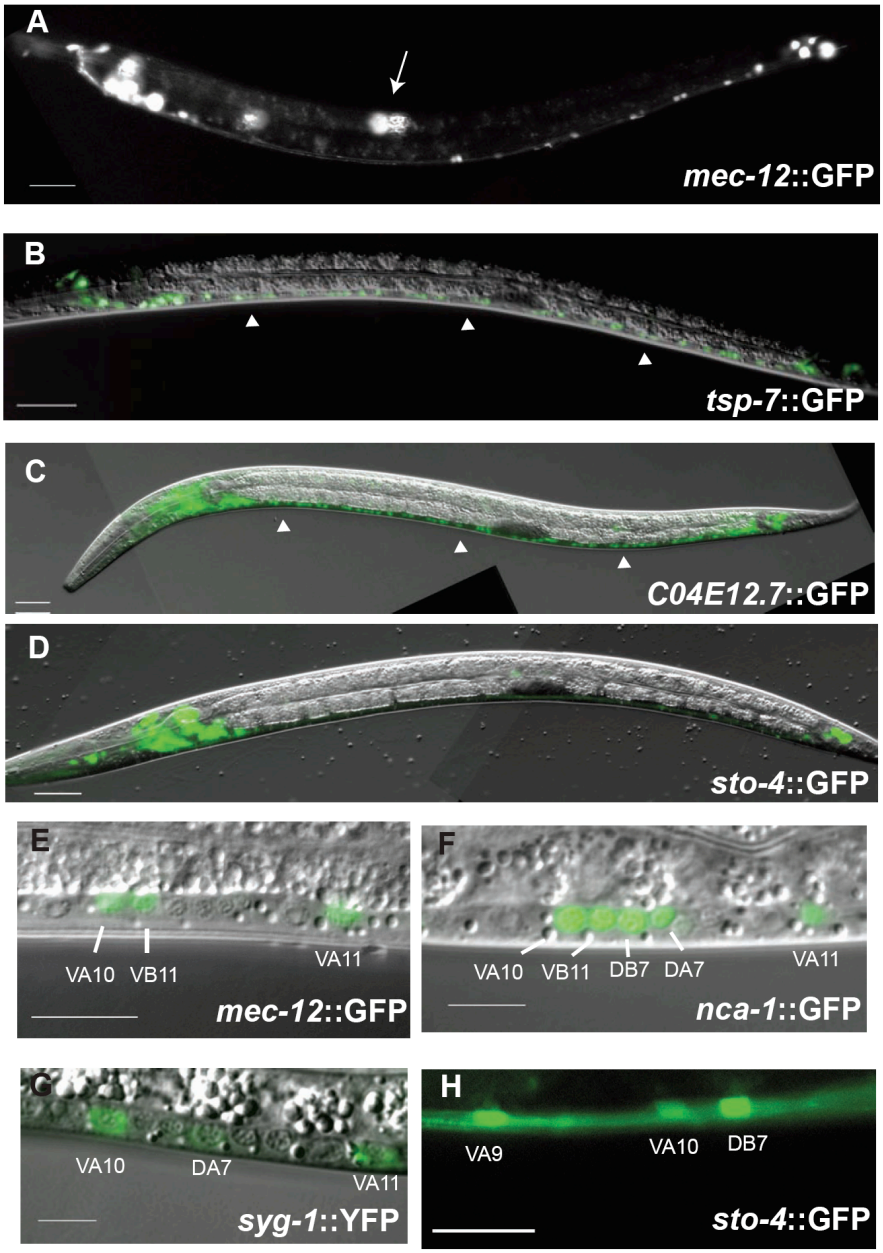


Figure 4.6 GFP reporters validate neuronal microarray datasets. Transgenic animals expressing GFP reporters for representative genes detected in neuron-enriched microarray datasets. Anterior to left, ventral down. GFP images are combined with corresponding matching DIC micrographs for panels B-G. **A, E.** *mec-12*::GFP is expressed in touch neurons (arrow) and in specific ventral cord motor neurons **(E)** at the L2 stage. **B,C.** *tsp-7*::GFP and *C04E12.7*::GFP are widely expressed in the nervous system with bright GFP in head and tail ganglia and in motor neurons of the ventral nerve cord (arrow heads). **(D, F, G. H)** Note expression of GFP reporters for *sto-4*, *nca-1*, and *syg-1* in A-class (DA, VA) and other B-class (DB, VB) ventral cord motor neurons.

GFP-reporter analysis demonstrates that the remaining 11 genes tested are expressed in other tissues in addition to neurons. For example, the GFP reporter for *C04E12.7* (phospholipid scramblase), which is expressed widely throughout the nervous system, is also expressed in muscle cells (Table 4.1) (Fig. 4.6C). Thus, these results inducate that the genes identified in the Larval Pan-neural profile largely fall into two classes; those that are exclusively expressed in neurons, and those that are expressed in multiple tissues, including neurons. Our finding of neuronal GFP expression for transcripts exhibiting a wide range of enrichment (1.5 to 8.3 fold) predicts that the majority of genes in this list that have not been directly tested are also likely to be expressed in neurons. Together, these results demonstrate that our Pan-neural mRNA-tagging approach enriches for *bona fide* neuronally-expressed transcripts and effectively excludes transcripts expressed exclusively in other tissues.

Gene families enriched in neurons of *C. elegans* larvae

Protein-encoding genes in the enriched Larval Pan-neural profile were organized into groups on the basis of KOGs and other descriptions that identify functional or structural categories (Table 4.2) (Additional File 4)(Tatusov, Fedorova *et al*. 2003). Over half (880/1562) are homologous to proteins in at least one other widely diverged eukaryotic species (i.e. KOGs and TWOGs), 49 of which are classified as uncharacterized conserved proteins. Homologs for an additional 225 Pan-neural enriched proteins are limited to other nematode species (i.e. LSEs).

Transcripts encoding proteins with fundamental roles in neuronal activity or signaling are highly represented in this dataset. (For a comprehensive list see Additional File 4). For

Table 4.2. Transcripts enriched in *C. elegans* neurons.

Category	Embryonic Pan-neural	Larval Pan-neural	Embryonic A-class	Larval A-class
Ion Channels/Receptors/Membrane Proteins	**122**	**156**	**60**	**41**
Acetylcholine Receptors	*13*	*24*	*9*	*9*
GABA Receptors	*1*	*4*		*3*
Glutamate Receptors	*8*	*8*	*1*	*2*
Potassium Channels	*11*	*24*	*8*	*10*
Calcium Channels	*8*	*10*	*7*	*4*
DEG/ENaC Channels	*3*	*10*	*1*	*1*
Stomatins	*3*	*7*	*2*	*1*
Other Ligand-Gated Ion Channels	*6*	*13*	*2*	*2*
Gap junction proteins (innexins)	*4*	*4*	*1*	*1*
Symporters/Exchangers/Transporters	*24*	*27*	*12*	*3*
Other Membrane Proteins	*41*	*25*	*17*	*5*
Axon Guidance	**4**	**8**	**8**	**3**
Adhesion/Ig Domain	**6**	**17**	**10**	**11**
Cytoskeleton-related	**33**	**34**	**16**	**5**
Transcriptional Control	**90**	**91**	**38**	**10**
Homeobox	*8*	*28*	*3*	*3*
Hormone Receptors	*24*	*15*	*5*	*1*
Aryl-hydrocarbon Receptors	*1*	*3*	*1*	
SMADs		*3*	*1*	*1*
HMG box	*5*	*5*		
HLH factors	*2*	*4*	*1*	*1*
Other transcription factors	*32*	*25*	*13*	*4*
General Factors	*18*	*8*	*14*	
Kinase/Phosphatase	**82**	**79**	**51**	**18**
GPCR Signaling	**107**	**169**	**42**	**25**
G-protein Coupled Receptors	*85*	*137*	*33*	*18*
G-proteins	*8*	*10*	*3*	*3*
Regulators of G-protein signaling (GTPases, GEFs, GRKs)	*7*	*8*	*4*	*2*
Adenylate/Guanylate Cyclases	*7*	*14*	*2*	*2*
Rab/Rho/Rac GTPase signaling	**17**	**7**	**7**	**2**
Neuropeptides	**39**	**58**	**11**	**13**
FMRFamide-like (flp)	*13*	*20*	*4*	*5*
Neuropeptide-like (nlp)	*13*	*18*	*3*	*4*
Insulin-like	*9*	*11*	*2*	*1*
TGF-beta	*1*	*3*	*1*	*1*
Pro-protein convertases	*3*	*6*	*1*	*3*
Calcium Binding	**18**	**26**	**12**	**9**
Synaptic Vesicle Associated	**38**	**53**	**25**	**17**
RNA Binding	**22**	**14**	**22**	**3**
Ubiquitin Associated	**39**	**19**	**12**	**3**
Enzymes	**199**	**103**	**111**	**30**

Collagens	**2**	**1**	**5**	**24**
Other	**297**	**205**	**174**	**44**
Unnamed/Uncharacterized	**159**	**127**	**161**	**56**
Unclassified	**363**	**395**	**230**	**98**
total	*1637*	*1562*	*995*	*412*

example, in addition to the 34 Synaptic Vesicle (SV) associated transcripts from Fig. 4.3B (Additional File 7), transcripts for 17 proteins with potential roles in synaptic vesicle function are identified (Fig. 4.7). These include six members of the synaptotagmin family of calcium-dependent phospholipid binding proteins (*snt-1, snt-4, snt-5, snt-6*, DH11.4, T10B10.5), only one of which, *snt-1*, has been previously shown to function in neurons (Nonet, Grundahl *et al*. 1993). Expression of the additional synaptotagmin genes in the nervous system may account for the residual synaptic vesicle function of *snt-1* mutants (Nonet, Grundahl et al. 1993; Xu, Mashimo et al. 2007). Three members of the copine family (B0495.10*, tag-64,* T28F3.1), a related group of calcium-binding proteins with potential roles in synaptic vesicle fusion (listed as part of endocytosis machinery in Fig. 4.7), are also enriched (Nakayama, Yaoi *et al*. 1998).

In addition to genes with general functions in synaptic vesicle signaling, the Larval Pan-neural profile includes transcripts encoding proteins with roles specific to particular neurotransmitters. For example, the plasma membrane and vesicular transporters for choline and acetylcholine (*cho-1* and *unc-17*), GABA (*snf-11* and *unc-46*, *unc-47*), dopamine (*dat-1* and *cat-1*), and glutamate (*glt-3* and *eat-4*) are included (Fig. 4.7) (Jayanthi, Apparsundaram *et al*. 1998; Okuda, Haga *et al*. 2000; Rand, Duerr *et al*. 2000; Jiang, Zhuang *et al*. 2005; Matthies, Fleming *et al*. 2006; Mullen, Mathews *et al*. 2006). The corresponding families of neurotransmitter-specific ligand-gated ion channels are highly represented, including 22 members of the ionotropic nicotinic ACh receptor family (Additional File 4). Other classes of ion channels with key neural functions are also abundant such as potassium channels (24), voltage-gated calcium channels (10) and DEG/ENaC sodium channels (10) (Table 4.2).

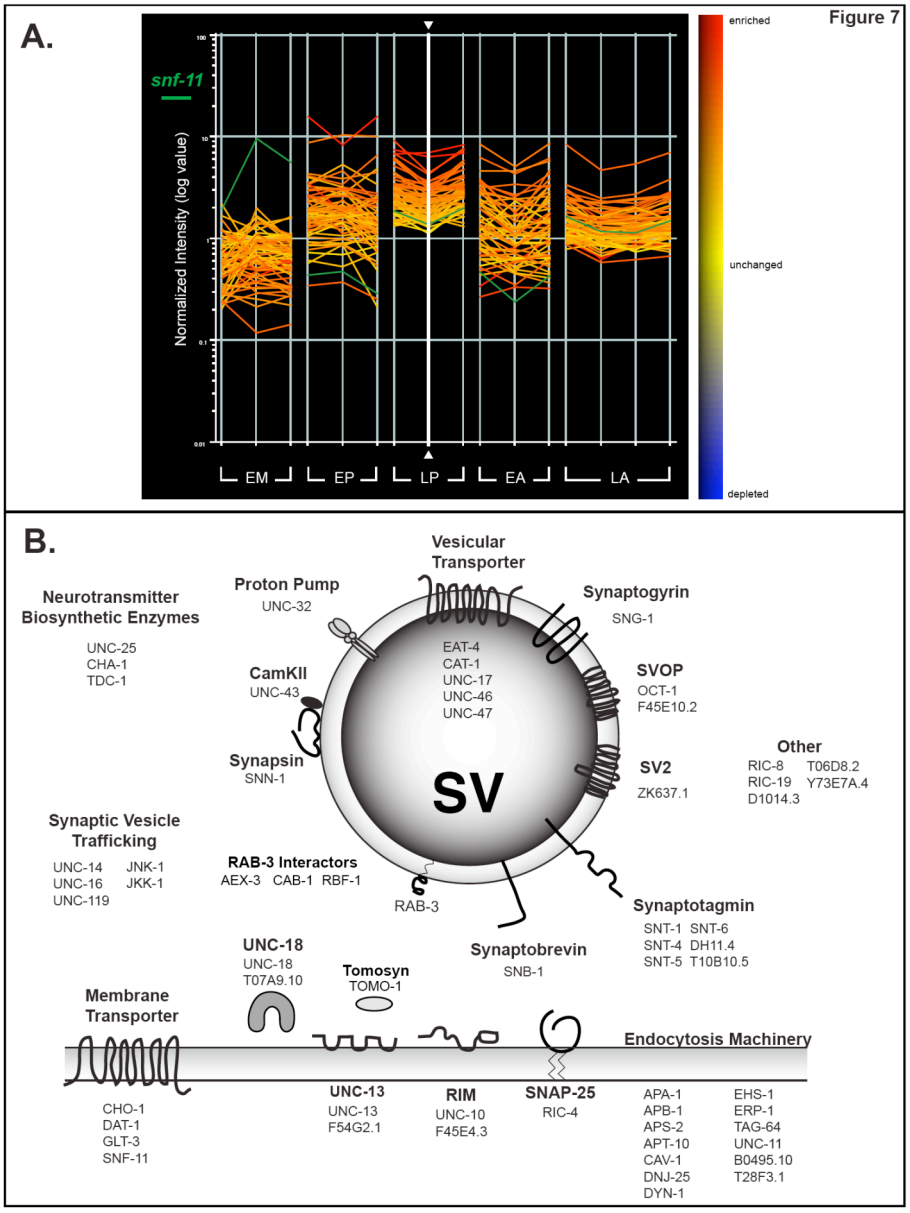


Figure 4.7 Transcripts encoding proteins that function in synaptic transmission are enriched in the neural datasets but largely excluded from muscle. A. The line graph depicts 60 synaptic transmission genes that are enriched in the Larval Pan-neural (LP) dataset (colors from heat map at right are defined by LP sample denoted by vertical white line with arrowheads). Most of these transcripts are also enriched in other neuronal datasets (EP, EA, and LA) datasets but not in embryonic muscle (EM). An exception is *snf-11* (horizontal green line), the membrane-bound GABA transporter, which is significantly elevated in the EM and LP datasets, consistent with its known expression in muscle and neurons. **B.** Many of the proteins encoded by the 60 LP-enriched synaptic transmission genes are localized to synaptic vesicles (SV) (center circle) or to the plasma membrane (shaded rectangle). Other proteins are predicted to play known roles in other related functions such as the synthesis of neurotransmitters and/or vesicular trafficking.

The wide range of neurotransmitter-specific genes in the Larval Pan-neural dataset reflects the diverse array of neuron types in *C. elegans* (Fig. 4.5). This point is underscored by the detection of a large number of transcription factors with established roles in neuronal specification (Table 4.3). These include UNC-86, the POU homeodomain protein that regulates the differentiation of a broad cross-section of neuron classes (Finney and Ruvkun 1990; Sze, Zhang *et al*. 2002; Sze and Ruvkun 2003) as well as transcription factors that define specific neuronal subtypes such as the canonical LIM homeodomain MEC-3 (mechanosensory neurons) (Way and Chalfie 1988; Chalfie and Au 1989; Duggan, Ma *et al*. 1998) and the UNC-4 homeodomain (A-class ventral cord motor neurons, see below) (Miller, Niemeyer *et al*. 1993; Miller and Niemeyer 1995; Winnier, Meir *et al*. 1999; Von Stetina, Fox *et al*. 2007). Transcription factors with undefined roles in the nervous system are also identified. Of particular note are fifteen members of the Nuclear Hormone Receptor (NHR) family, only one of which, *fax-1*, has been previously shown to regulate neuronal differentiation (Much, Slade *et al*. 2000).

A striking example of the power of this profiling approach is revealed by the presence of a large number of genes involved in peptidergic signaling. Neuropeptides are potent modulators of synaptic transmission. A combination of genetic and pharmacologic experiments have assigned specific neuromodulatory roles to FMRFamide and related peptides (FaRPs) encoded by members of the "*flp*" (FMRFamide like peptides) gene family (Li 2005). Examples include *flp-13* (cell excitability)(Rogers, Franks *et al*. 2001), *flp-1* (locomotion) (Nelson, Rosoff *et al*. 1998) and *flp-21* (feeding behavior) (Rogers, Reale *et al*. 2003). The enriched status of the majority of *flp* genes (20/23) in the Larval

Table 4.3. Major transcription factor families enriched in *C. elegans* neurons

Transcription Factor Families		Fold Change			
		Embryonic Pan-neural	Larval Pan-neural	Embryonic A-class	Larval A-class
Cosmid Name	Common Name				
Homeobox					
C40H5.5	*ttx-3*	1.6	1.5		
C33D12.1	*ceh-31*		1.6		
D1007.1	*ceh-17*		1.6		
K02B12.1	*ceh-6*		1.6		1.6
T13C5.4		3.3	1.6		
T26C11.7	*ceh-39*		1.6		
ZC64.3	*ceh-18*		1.6		
C18B12.3			1.7		
C28A5.4	*ceh-43*		1.7		
F56A12.1	*unc-39*		1.8		
W03A3.1	*ceh-10*		1.8		
C10G8.6	*ceh-34*		1.9		
C17H12.9			2		
F55B12.1	*ceh-24*		2		
ZC123.3		1.6	2		
C30A5.7	*unc-86*		2.1		
T26C11.5	*ceh-41*	1.6	2.1		
C39E6.4	*mls-2*		2.3		
F01D4.6	*mec-3*		2.6		
R08B4.2			2.6		
B0564.10	*unc-30*	2.2	2.7		
W05E10.3	*ceh-32*		2.7		
F26C11.2	*unc-4*		2.8	13.2	9.0
C37E2.4	*ceh-36*		2.9		
R07B1.1	*vab-15*		2.9		
Y54F10AM.4	*ceh-44*		3.3		2.1
F58E6.10	*unc-42*	2.4	3.8		
ZC247.3	*lin-11*		5.2		
C07E3.5		1.7			
F46C8.5	*ceh-14*	1.7			
W06A7.3	*ret-1*			1.8	
Y113G7A.6	*ttx-1*			2.8	
Hormone Receptor					
Y94H6A.1			1.5		
F47C10.3			1.6		
F47C10.7			1.6		
F56E3.4			1.7		
H01A20.1	*fax-1*		1.7		
R09G11.2	*nhr-3*	1.8	1.7	1.9	
T03G6.2	*nhr-1*		1.7		
Y39B6A.17	*nhr-40*		1.7		
C24G6.4	*nhr-95*	1.6	1.8	1.8	
K06B4.8	*nhr-47*	2	2		
K06B4.1			2.1		
K06B4.2	*nhr-51*	2.5	2.1		
F21D12.1	*nhr-52*		2.2		
C49F5.4	*nhr-21*		2.9		
F07C3.10			4		
C06C6.4	*nhr-63*	1.8			
C08F8.8	*nhr-67*	1.5			
C17E7.8	*nhr-124*	2.3			
F09C6.9	*nhr-116*	3.3		1.9	
F16B4.9		1.9			

F31F4.12		1.6			
F41B5.9		2.5		1.7	
F44C8.11	*nhr-96*	1.8			
F44C8.9		1.7			
F48G7.11		2			
F59E11.8		1.8			
K06B4.10					1.8
K08A2.5	*nhr-88*	1.5			
K11E4.5	*nhr-71*	1.8			
R07B7.15		1.7			
R11E3.5	*nhr-104*			1.8	
T07C5.4		3.2			
T19A5.4	*nhr-44*	1.9			
T27B7.1	*nhr-59*	1.7			
T27B7.4	*nhr-115*	2.9			
Y17D7A.3	*nhr-65*	1.5			
Y67D8B.2		1.8			
Aryl-hydrocarbon Receptors					
C25A1.11	*aha-1*		1.7		
C41G7.5	*ahr-1*		1.7		
C56C10.10		2.2	2	2.3	
SMADs					
F01G10.8	*daf-14*		1.7		
F25E2.5	*daf-3*		1.9		
F37D6.6	*tag-68*		2.2	2.3	2.0
HMG box					
F40E10.2	*sox-3*	1.6	1.5		
T22B7.1	*egl-13*		1.6		
T05A7.4	*hmg-11*	2.3	2.1		
F47G4.6			2.3		
K08A8.2	*sox-2*	2.5	2.9		
C12D12.5		4.7			
Y17G7A.1	*hmg-12*	1.8			
HLH factors					
C43H6.8	*hlh-15*	4.4	1.6		
F58A4.7	*hlh-11*		1.6		
Y16B4A.1	*unc-3*		2.9	3	4.9
F48D6.3	*hlh-13*		3		
W02C12.3		1.8			

Pan-neural profile (Fig. 4.4B) parallels immunostaining and GFP reporter results showing expression of this gene family in the *C. elegans* nervous system (Kim and Li 2004). Transcripts encoding insulin-like peptides (*ins*) and neuropeptide-like genes (*nlp*) are among the most highly enriched mRNAs in the Pan-neural dataset (Additional File 4). Neuropeptide activating proteases such as the proprotein convertase, *egl-3,* and the carboxypeptidase, *egl-21*, are also elevated (Jacob and Kaplan 2003). Finally, we detect 136 members of the G-protein coupled Receptor (GPCR) family including four GPCRs (*npr-1, npr-2, npr-3* and T19F4.1) that have been either directly identified as neuropeptide receptors or implicated in neuropeptide-dependent behaviors (Keating, Kriek *et al*. 2003; Rogers, Reale *et al*. 2003; Mertens, Meeusen *et al*. 2005) (E. Siney, A. Cook, N. Kriek, L. Holden Dye, personal communication). The strong representation of diverse neuropeptidergic components in the Larval Pan-neural profile is suggestive of a nervous system that is richly endowed with complex signaling pathways for modulating function and behavior.

A comparison of larval neuronal expression to an embryonic neuronal profile generated by MAPCeL

To complement the profile of the larval nervous system obtained by the mRNA-tagging method, a Pan-neural GFP reporter gene (Altun-Gultekin, Andachi *et al*. 2001)(J. Culotti, personal communication) was used to mark embryonic neurons for MAPCeL analysis. Comparisons of independent replicates showed that these data are highly reproducible (Additional File 8). 1637 enriched genes ($\geq$1.5X, FDR $\leq$ 1%) were identified versus a Reference dataset obtained from all embryonic cells (Additional File 1). The majority (82%) of transcripts in this list with known expression patterns are

expressed in neurons (Fig. 4.3A). All of the promoter GFP fusions (10/10) created from previously uncharacterized genes in the enriched Embryonic Pan-neural dataset showed expression in neurons, further validating this MAPCeL profile (Table 4.1). A comparison of the embryonic (MAPCeL) and larval (mRNA-tagging) profiles reveals considerable overlap with ~45% of transcripts (710/1637) enriched in the embryonic neurons also elevated in larval neurons (Fig. 4.8A). The intersection of these two datasets is significantly enriched (96%) for known neuron-expressed genes (representation factor 5.2, $p<1e^{-325}$). Of particular interest are 27 genes in this core group of neuronal transcripts that encode uncharacterized conserved proteins with human homologs (Additional File 9).

As an additional test of the similarities between these independent datasets, we examined the Embryonic and Larval Pan-neural profiles for elevated expression of gene families with roles in synaptic vesicle function (Fig. 4.7A). Both the Embryonic and Larval Pan-neural datasets were enriched for many of these components. In contrast, the majority of these transcripts are not upregulated in a MAPCeL profile of embryonic muscles (Fox *et al*., submitted). Interestingly, the one exception to this correlation, the GABA transporter *snf-11*, is known to be expressed in body wall muscle in addition to neurons (Mullen, Mathews *et al*. 2006).

Examination of the Embryonic and Larval Pan-neural datasets also revealed a large number of genes that function in the dauer pathway. The dauer larva adopts an alternative developmental program to withstand stressful conditions (e.g. starvation, overcrowding, high temperature). The decision to adopt the dauer state is regulated by the

Figure 8

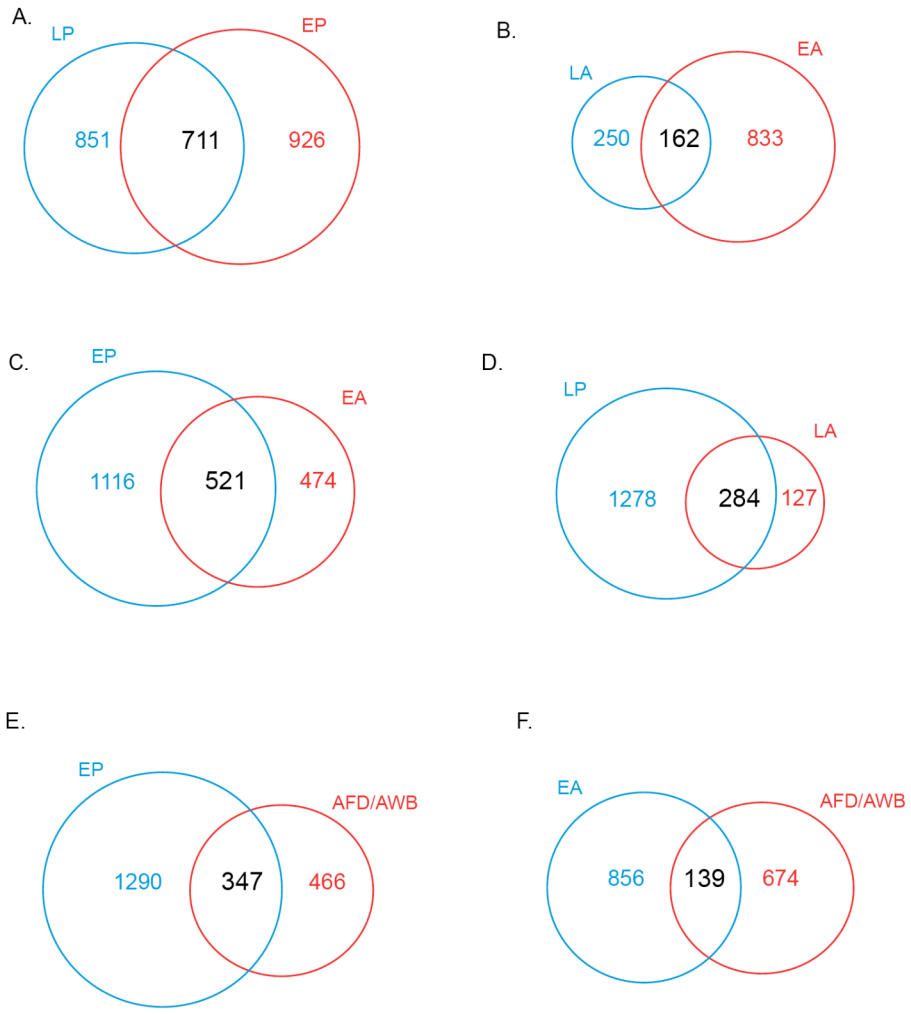


Figure 4.8 Venn Diagrams comparing transcripts from profiled cell types at specific stages of development. A. Larval Pan-neural (LP) and Embryonic Pan-neural (EP) datasets are enriched for many common transcripts, but also contain transcripts exclusive to either developmental stage. **B**. Larval A-class (LA) and Embryonic A-class (EA) identify 162 shared transcripts. Transcripts exclusively enriched in either cell type may contribute to the unique morphologies of DA vs VA motor neurons (see Fig. 10). **C, D.** The depth of the pan-neural datasets (EP, LP) is underscored by substantial overlap with the A-class motor neuron profiles (EA, LA). Additionally, the many transcripts exclusive to the EA and LA datasets emphasize the power of cell-specific profiling. **E-F.** Comparisions between the embryonic neural specific datasets (EP, EA) described in this paper versus the embryonic profile of specific thermosensory neurons, (AFD and /AWB described by Colosimo et al. 2004. The AFD/AWB profile shows greater overlap with the Embryonic Pan-neural (EP) dataset (**E**) than with the Embryonic A-class motor neuron (EA) profile (**F**). See Additional Files 10, 11 for lists of genes identified in each comparison.

nervous system and is triggered during the L1/L2 transition in response to environmental cues (Riddle and Albert 1997; Swoboda, Adler *et al*. 2000; Murakami and Johnson 2001; Ailion and Thomas 2003; Li, Kennedy *et al*. 2003; Beckstead and Thummel 2006). Fig. 4.9 graphically represents the dauer pathway genes identified in the combined Pan-neural datasets.

Transcription factors constitute the largest gene family that is differentially enriched between the Embryonic and Larval Pan-neural profiles (Table 4.3). For example, the combined Pan-neural datasets detect a total of 30 nuclear hormone receptors (NHRs). However, 16 NHRs are exclusively detected in embryonic neurons, whereas only six are enriched solely in larval neurons. Homeodomain transcription factors are also unequally distributed across the two datasets. Of 32 enriched homeoproteins, 24 are exclusive to the Larval Pan-neural profile, whereas only 4 are selectively elevated in the Embryonic Pan-neural dataset (Table 4.3). The relative lack of enrichment of homeodomain mRNAs in the Embryonic Pan-neural profile was initially surprising given strong genetic evidence for the widespread role of the members of this transcription factor class in embryonic neural development (Chalfie and Au 1989; Hallam, Singer *et al*. 2000; Altun-Gultekin, Andachi *et al*. 2001; Esmaeili, Ross *et al*. 2002; Von Stetina, Treinin *et al*. 2006). A likely explanation for this finding is that many homeobox transcripts are dynamically expressed in multiple cell types in the embryo but are increasingly restricted to neurons during larval development (Ahringer 1996; Esmaeili, Ross *et al*. 2002). This view is consistent with our observation that a majority (22/28) of homeodomain genes that are enriched in the Larval Pan-neural dataset are in fact also detected as Expressed Genes (EGs) in the Embryonic Pan-neural profile (see below).

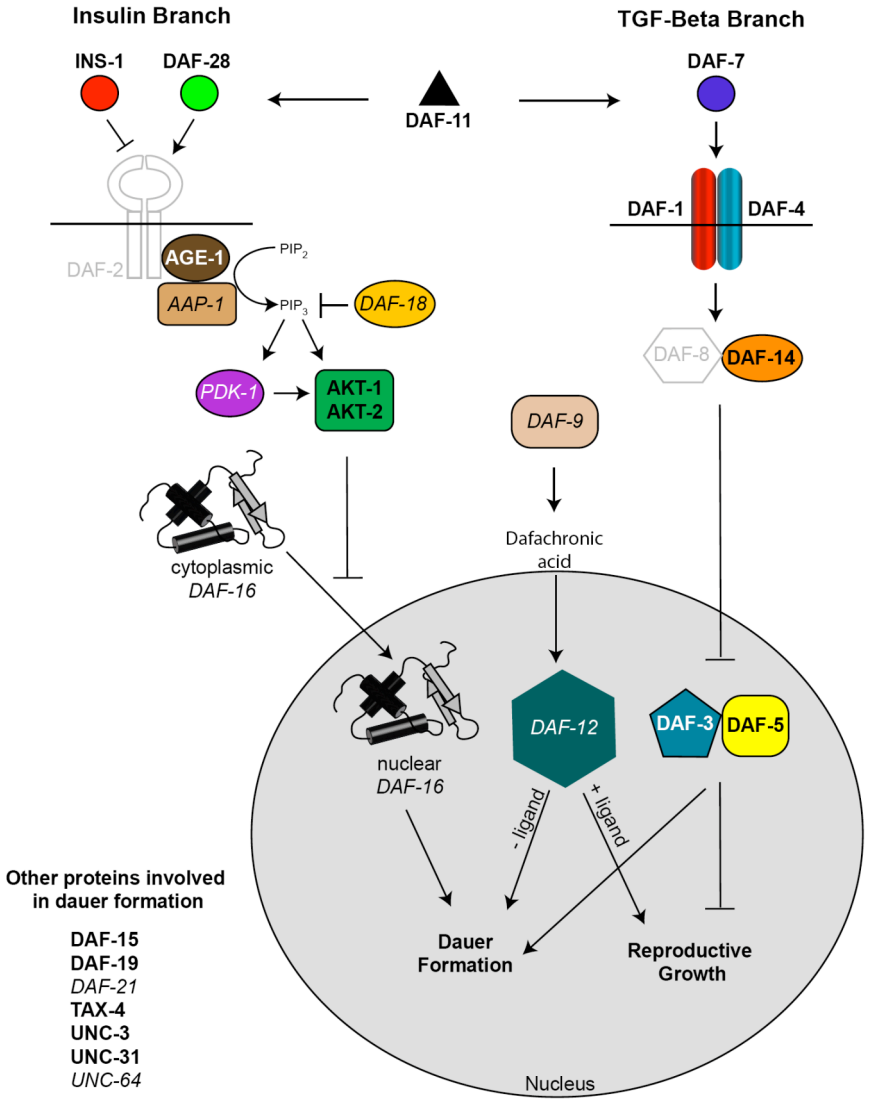


Figure 4.9 A majority of dauer pathway genes are highly enriched in the larval (LP) and embryonic (EP) pan-neural datasets. The decision to enter an alternative life cycle, the dauer larval, is made by the nervous system. Two pathways influence the decision to dauer, an alternative developmental pathway adopted in unfavorable conditions. During normal growth, the DAF-28 insulin-like molecule activates the DAF-2 insulin receptor to initiate a signal transduction pathway that prevents the translocation of the DAF-16 Forkhead transcription factor inato the nucleus, thus preventing entry into dauer formation. In a parallel pathway, DAF-7/TGF-Beta activates the activin receptors DAF-1 and DAF-4 to inhibit the Smad/Sno complex DAF-3/DAF-5, thereby promoting reproductive growth. The guanylyl cyclase DAF-11 promotes the expression of DAF-28 and DAF-7. During reproductive growth, the CYP2 cytochrome P450 enzyme DAF-9 is active and produces the DAF-12 ligand dafachronic acid. In the presence of its ligand, the nuclear hormone receptor DAF-12 promotes normal development. In the absence of its ligand, DAF-12 instead promotes dauer formation. Other proteins function independently of these pathways (e.g. DAF-19 is a transcription factor that specifies ciliated neurons that detect exogenous dauer inducing signals). Bold lettering denotes enriched transcripts and italics marks EGs detected in at least one of the pan-neural datasets. Gray letters refer to transcripts not found in either EP or LP dataset.

C. elegans Interactome identifies neuronal genes potentially involved in synaptic function

The *C. elegans* Interactome documents ~5,500 protein-protein interactions derived from yeast-2-hybrid results, from interologs (i.e. interactions between protein homologs in other species) and from functional interactions described in the literature (Li, Armstrong *et al*. 2004). To gain further insight into the functional significance of prospective neural genes identified by these microarray datasets, we looked for evidence of interactions among proteins encoded by these genes in the Interactome database (see Methods). The 711 transcripts enriched in both the Embryonic and Larval Pan-neural datasets were uploaded for this analysis (Fig. 4.8A). This search generated an interaction map with a single prominent cluster. Most of the transcripts in this group (30/34) are detected in at least one of the Pan-neural datasets (Fig. 4.10). Our finding that the majority of the genes in this Interactome group are expressed in the nervous system favors the idea that these networks reflect authentic interactions in neurons. We note that 13 of the proteins in this list (yellow circles in Fig. 4.10) have not been previously assigned to the nervous system. Annotation of this Interactome map with functional data for each corresponding protein revealed two distinct subclusters featuring roles in either synaptic transmission or nucleic acid binding. For example, the JIP3/JSAP1 JNK scaffolding protein, UNC-16, interacts with KLC-2 (kinesin light chain) to regulate vesicular transport in neurons (Sakamoto, Byrd *et al*. 2004). Other members of this interacting complex, MKK-4 (MAP Kinase Kinase) and JNK-1 (Jun Kinase) are also required for maintaining normal synaptic structure (Byrd, Kawasaki *et al*. 2001; Nakata, Abrams *et al*. 2005). These findings suggest that additional proteins in this subcluster

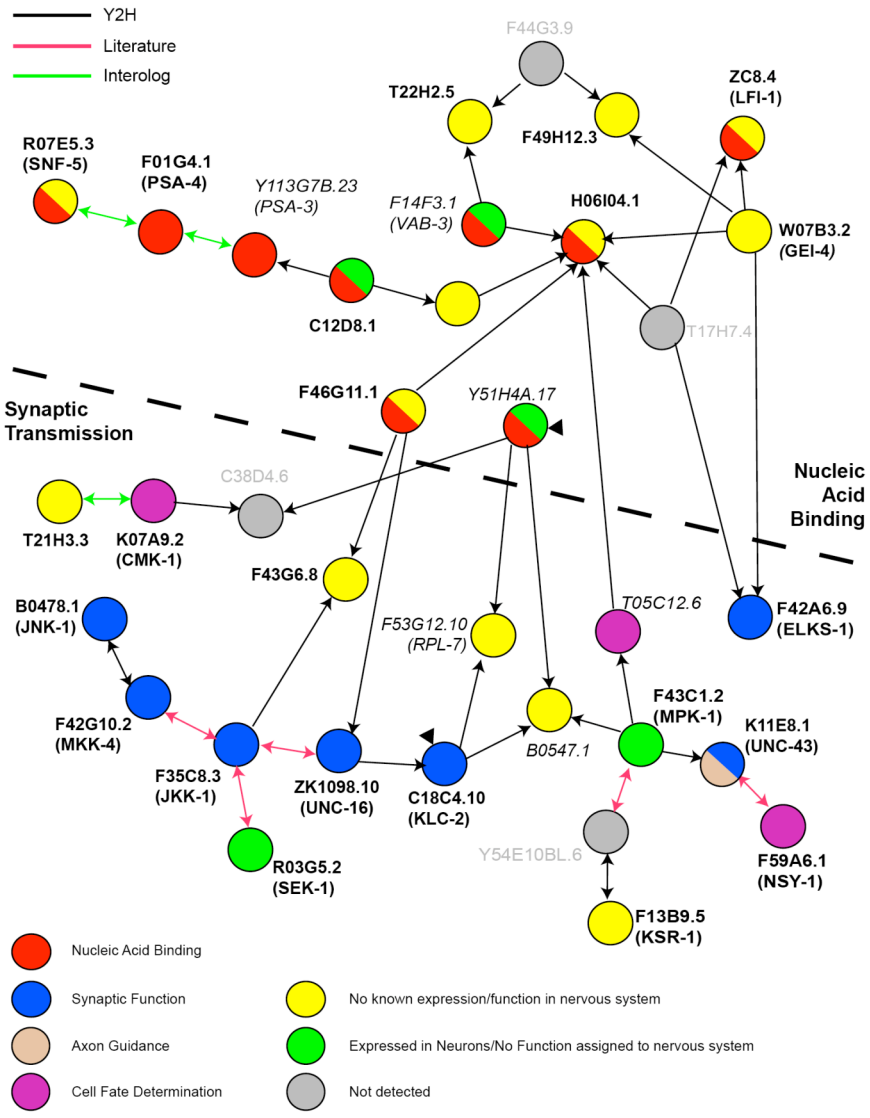


Figure 4.10 Interactome map of genes detected in the EP and LP datasets. Bold lettering denotes enriched transcripts, and italics marks EGs detected in at least one of the pan-neural datasets. Gray letters refer to transcripts not found in either dataset. Colored lines connecting the proteins represent datasets from which the interactions were isolated. Black lines represent interactions isolated by yeast-2-hybrid (Y2H), red lines depict known interactions listed in worm PD (literature), and green lines denote *in silico* searches against orthologous pairs (interolog). Black arrows point from bait to prey. Arrowheads indicate self-interactions (refer to proteins that homodimerize). Colored circles refer to protein categories. Multi-colored colors circles represent proteins that fit into multiple categories. Proteins are segregated into 2 groups (Nucleic Acid Binding vs Synaptic Transmission) as denoted by a dashed black line. Proteins above the large black dotted line belong to the nucleic-acid binding group. Proteins below this line are a part of the synaptic output group.

may function at the synapse. F43G6.8 (E3 Ubiquitin Ligase) and B0547.1 (COP-9 Signalosome subunit) are attractive possibilities as synaptic development and function are regulated by ubiquitin-dependent protein degradation (Patrick 2006). As more phenotypic data are compiled, this analysis can be extended to encompass data derived from RNAi experiments which may yield models for molecular machines that function in neurons (Gunsalus, Ge *et al*. 2005).

An mRNA-tagging transcriptional profile of a small subset of neurons

Although our gene expression profiles of the embryonic and larval nervous systems provide a comprehensive list of transcripts that function in neurons, these data lack the spatial resolution to identify the specific neurons in which these transcripts are expressed. For example, the dopamine transporter, *dat-1*, is highly enriched (15.9X) in the Larval Pan-neural dataset, but *dat-1* expression is limited to eight dopaminergic neurons (Nass, Hall *et al*. 2002). Other transcripts that are also restricted to a small number of neurons, however, might not be detected in a global profile of the entire nervous system. For example, the genes *gcy-5* and *gcy-6* (guanylate cyclase) are each expressed in single neurons, ASER and ASEL (Chang, Johnston *et al*. 2003), respectively, and neither is enriched in the Larval Pan-neural dataset. The application of the mRNA-tagging strategy to individual classes of neurons should therefore correlate gene expression with specific neurons as well as detect low abundance transcripts with potential key functions in these cells. To test this idea, we used the *unc-4* promoter to express FLAG-PAB-1 in only the subset of neurons in the ventral nerve cord that express the UNC-4 homeodomain protein. In the L2 larva, *unc-4*::GFP and *unc-4*::LacZ reporters show strong expression in a total

of 18 neurons: VA motor neurons (12), SAB motor neurons (3), the I5 pharyngeal motor neuron (1) and AVF interneurons (2) (Miller and Niemeyer 1995; Lickteig, Duerr *et al*. 2001). Weaker, sporadic expression is observed in 9 embryonically derived DA motor neurons at this stage. (*unc-4* is strongly expressed in the DAs in the embryo and in L1 larvae.) To increase the sensitivity of the mRNA-tagging method for profiling these neurons, PAB-1 was labeled with three tandem repeats of the FLAG epitope (3XFLAG). Fig. 4.11A and 4.11B show a mid-L2 larval animal (NC694) expressing the *unc-4*::3XFLAG::PAB-1 transgene in VA, SAB, and I5 motor neurons and in AVF interneurons; less intense expression is seen in the DA motor neurons. Because most (24/27) of the neurons in this group are members of the "A-class" of ventral cord excitatory motor neurons (VA, SAB, DA), we will refer to the mRNA-tagging data obtained from this transgene as the "Larval A-class motor neuron" (LA) profile (Fig. 4.9).

As previously observed for the Larval Pan-neural data (Fig. 4.2), independent hybridizations resulted in highly reproducible data for the Larval A-class motor neuron profile. (Additional File 8). A comparison of the A-class hybridization data to the Reference sample of mRNA from the average larval cell detected 412 enriched genes (see Methods). Of the 114 genes in this list with known expression patterns, 102 (~90%) are found in neurons (Fig. 4.3A). 96 of these genes have detailed spatial information, and 76 (~80%) of these show annotated expression in regions that also contain UNC-4-expressing neurons (Additional File 1). Of particular note, the native *unc-4* transcript, which is selectively expressed in these neurons *in vivo*, is the most highly enriched (8x) mRNA in this dataset. Other known A-class-motor neuron genes in this list

include the vesicular acetylcholine transporter (VAChT) *unc-17* and the Olf/EBF transcription factor *unc-3* (Fig. 4.11c) (Prasad, Ye *et al*. 1998; Lickteig, Duerr *et al*. 2001). In contrast, transcripts known to be restricted to other cell types, such as muscle (*myo-2*, *unc-22*) or GABAergic neurons (*unc-25*) are depleted from the A-class neuronal profile (Figs 4.4A, 4.11C). For example, <2% of transcripts selectively expressed in larval germ line, intestine, or muscle (30/1926) are enriched in the Larval A-class motor neuron profile (Additional File 5) (Pauli, Liu *et al*. 2005).

All of the GFP reporter lines (19/19) constructed for A-class enriched transcripts (Table 4.1) (Additinoal File 17) are expressed in UNC-4 neurons. For example, in the mid-L2 stage ventral nerve cord, *mec-12*::GFP is expressed in DA, VA, VB and VD motor neurons (Fig. 4.6A, E) and *syg-1*::GFP (Ig domain) is detected in DA and VA motor neurons among others (Fig. 4.6G). These results strongly suggest that most of the genes in the UNC-4 neuron enriched dataset are expressed in these cells *in vivo*. Thus, these data indicate that the mRNA-tagging method can produce a reliable profile of subsets of neurons in *C. elegans*.

A subset of Pan-neural genes are expressed in Larval A-class motor neurons

Nearly 70% of the Larval A-class enriched transcripts (282/412) are also elevated in the Larval Pan-neural dataset (representation factor 8.2, $p<2.9e^{-209}$) (Additional File 10). As expected, genes with known functions in all neurons are highly represented in this group (Table 4.2). Synaptic vesicle associated transcripts that are widely expressed in

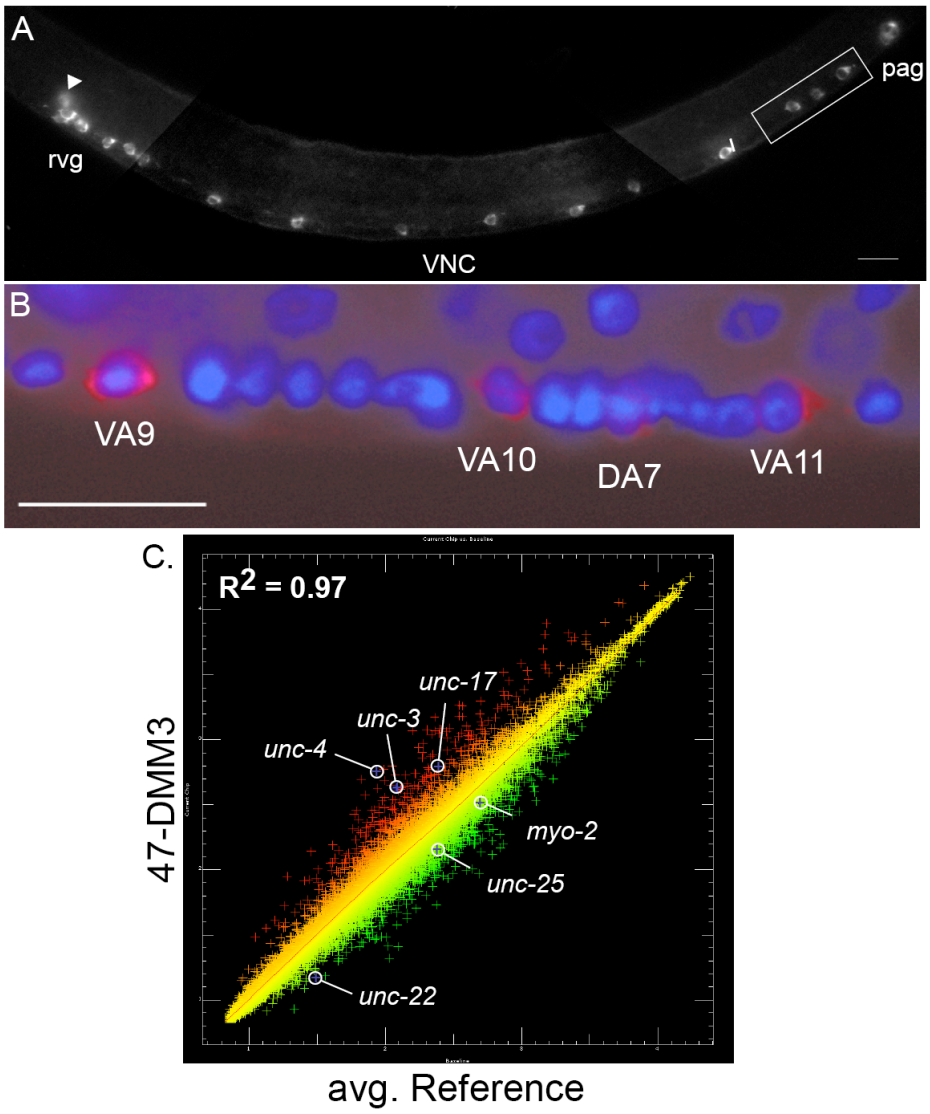


Figure 4.11 Larval A-motor neuron enriched transcripts are revealed by mRNA-tagging with *unc-4*::3XFLAG::PAB-1.
A. Antibody staining detects FLAG::PAB-1 expression in A-class neurons in the retrovesicular ganglion (RVG), ventral nerve cord (VNC), and pre-anal ganglion (PAG) and in the I5 pharyngeal neuron (arrowhead). Lateral view of L2 larva, anterior is to left, ventral down.
B. Close-up of posterior ventral nerve cord (boxed image in A), showing that anti-FLAG staining (red) is restricted to cytoplasm of A-class motor neurons. DAPI (blue) marks cell nuclei. (compare to Fig. 1, where all motor neurons show anti-FLAG staining). Anterior is left, ventral is is down. Scale bars = 10 µm
C. Results of a single Larval A-class hybridization (47-DMM3) (red) compared to average Reference intensities (green) to identify differentially expressed transcripts. The known A-class genes *unc-4* (homeodomain, A-class neurons), *unc-3* (O/E transcription factor, cholinergic VNC motor neurons), and *unc-17* (VAChT, cholinergic neurons) are enriched (red) in 47-DMM3. Genes expressed in other classes of neurons (*unc-25*, GAD, GABAergic neurons) or other tissues (*myo-2*, pharyngeal muscle myosin; *unc-22*, body wall muscle structural protein) are depleted (green) relative to the Reference profile.

the nervous system such as *rab-3* (G-protein), *snt-1* (synaptotagmin) and *snb-1* (synaptobrevin) are enriched in both datasets. Absences from the Larval A-class profile are correlated with class-specific functions in neurons. For example, the 60 transcripts encoding proteins involved in synaptic transmission enriched in the Larval Pan-neural dataset include vesicular transporters for GABA (*unc-47*), glutamate (*glt-3*), dopamine/serotonin (*cat-1*) and acetylcholine (*unc-17*) (Fig. 4.7B)(Rand, Duerr *et al*. 2000). The selective enrichment of the vesicular acetylcholine transporter, *unc-17*, in the Larval A-class profile is consistent with the known cholinergic signaling capacity of A-class motor neurons (Lickteig, Duerr *et al*. 2001). In another striking example of neuron-specific gene expression, the "*mec*" genes, which are required for normal differentiation or function of mechanosensory neurons, are highly represented in the Larval Pan-neural dataset but are not detected in the Larval A-class profile (Table 4.4) (Goodman and Schwarz 2003). The one exception is the alpha-tubulin encoding gene, *mec-12*, for which enriched expression in A-class neurons is confirmed with a GFP reporter gene (Fig. 4.6A,E). As described above, most of the known *flp* genes are enriched in the Pan-neural dataset(Li 2005). A subset of five *flp* genes is found in the A-class dataset (*flp-2,4,5,12,13*), providing enhanced spatial resolution for the expression repertoire of this large family of neuropeptide transmitters (Fig. 4.4B).

The A-class profile includes ~130 transcripts that are not detected in the Larval Pan-neural dataset (Additional File 10). Interestingly, ~20% of these genes (23/127) encode collagen-like proteins for which neural functions are largely undefined. *cle-1*, a type XVIII collagen, and the one member of this protein family that does have a

Table 4.4. Genes expressed in mechanosensory neurons are differentially detected in the larval Pan-neural dataset versus the larval A-class dataset.

Cosmid name	Gene name	Fold Change		Description
		Pan-neural	A-class	
F14D12.4	*mec-2*	2.9		Prohibitins and stomatins of the PID superfamily
F01D4.6	*mec-3*	2.6		Transcription factor, contains HOX domain
T01C8.7	*mec-4*	2.7		Non voltage-gated ion channels (DEG/ENaC family)
W02D3.3	*mec-6*	1.9		Unnamed protein
ZK154.3	*mec-7*	3.6		Beta tubulin
F16F9.5	*mec-10*	1.8		Non voltage-gated ion channels (DEG/ENaC family)
F57H12.7	*mec-17*	7.6		Uncharacterized conserved protein
C52B9.9	*mec-18*	3.1		Acyl-CoA synthetase
C44B11.3	*mec-12*	5.9	1.9	Alpha tubulin

documented role in the nervous system (Ackley, Kang *et al*. 2003), is enriched in both the Larval Pan-neural (LP) and A-class (LA) datasets. We speculate that post-embryonic motor neurons may secrete collagens and other extracellular matrix (ECM) components for assembly into the basement membrane that envelopes the ventral nerve cord (White, Southgate *et al*. 1986). Indeed, our data confirm that UNC-6 (netrin), a critical ECM signal that steers migrating cells and neuronal growth cones is highly expressed in Larval A-class motor neurons (Fig. 4.12) (Wadsworth, Bhatt *et al*. 1995).

Comparison of transcripts enriched in embryonic vs. Larval A-class motor neurons.

We have previously used the MAPCeL strategy to profile embryonic motor neurons marked with *unc-4*::GFP (Fox, Von Stetina *et al*. 2005). These include 12 Embryonic A-class motor neurons (9 DA and 3 SAB) and a single pharyngeal neuron, I5 (Fox, Von Stetina *et al*. 2005). The Embryonic A-class motor neurons are similar to the post-embryonic VAs in that they express *unc-4*, are cholinergic, extend anteriorly directed axons, and receive inputs from the command interneurons AVA, AVD, and AVE (White, Southgate *et al*. 1986). The strong overlap of these distinct morphological and functional traits as well as some residual larval expression of *unc-4* in Embryonic A-class motor neurons (Fig. 4.11B) are consistent with the observation that ~40% of transcripts enriched in the Larval A-class motor neuron dataset (162/412) are also elevated in the Embryonic A-class motor neuron MAPCeL profile (representation factor 7.4, $p<3.1e^{-99}$) (Fig. 4.8B) (Additional File 10). Transcripts from the cholinergic locus, *cha-1* (choline acetyl transferase) and *unc-17*(Vesicular ACh transporter), which are essential for the biosynthesis and packaging of acetylcholine (ACh) into synaptic vesicles are enriched in both A–class motor neuron profiles (Rand, Duerr *et al*. 2000). In addition to these gene

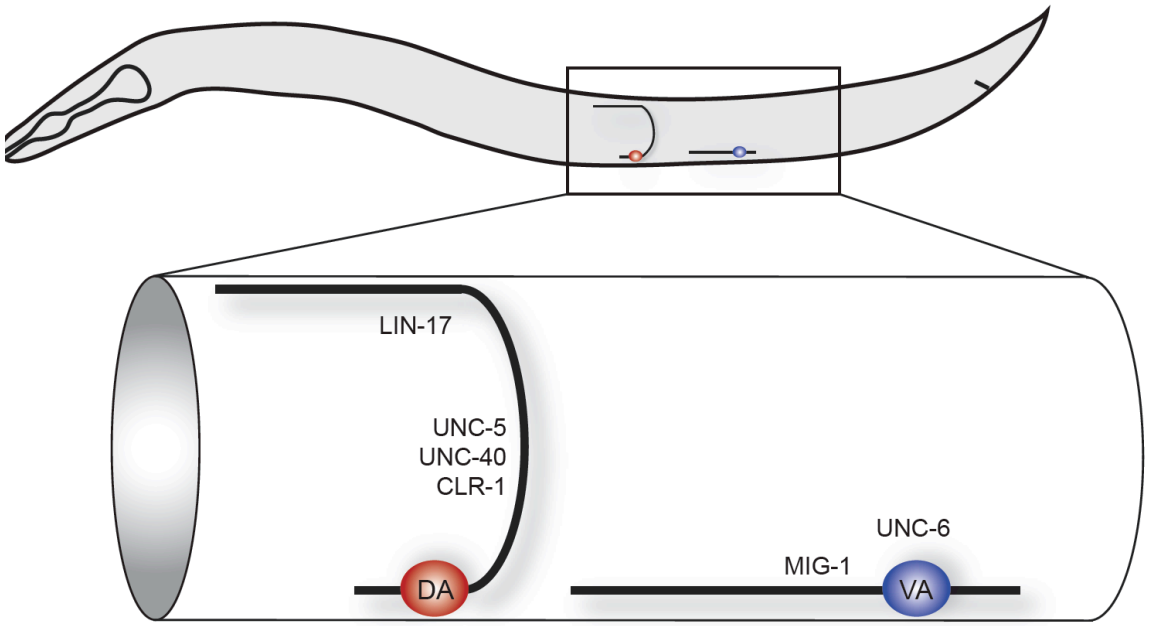


Figure 4.12 Differential expression of axon guidance cues and receptors in A-class motor neurons. DA motor neurons extend commissures to innervate muscles on the dorsal side whereas VA motor axons are retained in the ventral nerve cord. Embryonic DA and larval VA motor neuron cell bodies reside in the ventral cord and extend processes that are guided by specific molecular cues. The neurites of DAs, in response to a ventrally-localized Netrin (UNC-6) source, are repelled dorsally by the action of the netrin receptors UNC-40 and UNC-5, and the CLR-1 receptor tyrosine phosphatase. The VAs do not express these receptors and thus project ventrally directed axons. Enrichment of the Wnt receptors *lin-17* in the Embryonic A-class motor neuron (EP) dataset and *mig-1* in the Larval A-class motor neuron (LA) profile could be indicative of a Wnt-dependent mechanism for directing anterior outgrowth of DA and VA motor axons.

families, several others are enriched in both embryonic and Larval A-class motor neurons (Additional File 19). Acetylcholine signaling depends on the synaptic vesicle cycle and genes with key roles in this mechanism are elevated in both datasets: these include *unc-18*, *snt-1* (Syntaxin), *snn-1* (Synapsin), *ric-4* (SNAP-25), *sng-1* (Synaptogyrin), *unc-2* (Calcium channel), *rab-3,* and *unc-11* (Clathrin component). In addition, genes with either established or likely roles in the G-protein coupled signaling pathways that modulate ACh release from these motor neurons (*dop-1, pkc-1, kin-2, gar-2, rgs-1, rgs-6, gpc-2)* are common to both enriched datasets (Reynolds, Schade *et al*. 2004; Fox, Von Stetina *et al*. 2005). The general role of A-class motor neurons in both releasing and responding to a broad range of neuroactive signals is underscored by the embryonic and larval enrichment of multiple neuropeptides (i.e. *flp-2, flp-4, flp-5,* and *flp-13)* (Fig. 4.4B). Shared ionotropic receptors include the nAChR subunits, *acr-12*, *acr-14* and *unc-38* which lead to excitatory responses as well as the recently described Acetylcholine Gated Chloride subunit, *acc-4* (T27E9.9) which should mediate acetylcholine-induced inhibition of motor neuron activity (Putrenko, Zakikhani *et al*. 2005). Together, these data support the proposal that *C. elegans* A-class motor neurons utilize complex mechanisms for integrating signals originating as either paracrine or autocrine stimuli (Fox, Von Stetina *et al*. 2005).

Other transcripts that are highly enriched in both embryonic and Larval A-class datasets with potential roles in specifying shared characteristics of this motor neuron class include: *syg-1*, an Ig-domain membrane protein that localizes the presynaptic apparatus of the HSN motor neuron in the egg laying circuit (Fig. 4.6G) (Shen and

Bargmann 2003); *rig-6*, the nematode homolog of contactin, a membrane protein with extracellular fibronectin and Ig domains that organizes ion channel assemblages(Ranscht 1988; Boyle, Berglund *et al*. 2001); and *cdh-11*, the homolog of calsyntenin, a novel cadherin-like molecule that is highly localized to postsynaptic sites (Vogt, Schrimpf *et al*. 2001). Finally we note that of the 25 genes that encode innexin gap junction components(Starich, Sheehan *et al*. 2001), only one, *unc-9*, is enriched in both of the A-class motor neuron datasets. This finding points to the UNC-9 protein as a likely component of gap junctions that couple A-class motor neurons with command interneurons that drive motor circuit activity in the ventral nerve cord (Von Stetina, Fox *et al*. 2007).

In addition to genes that are enriched in both embryonic and Larval A-class motor neurons, we also detected transcripts that are selectively elevated in one or the other dataset (Additional File 10). Transcription factors comprise the largest group of differentially expressed genes. Of 24 transcription factor genes enriched in Embryonic A-class motor neurons, only two, *unc-3* and *unc-4*, are also included in the separate list of 10 transcription factors enriched in Larval A-class motor neurons (Table 4.3). UNC-3 (O/E HLH protein) and UNC-4 (homeodomain protein) have been previously shown to specify shared characteristics of embryonic and Larval A-class motor neurons (Prasad, Ye *et al*. 1998; Winnier, Meir *et al*. 1999; Lickteig, Duerr *et al*. 2001). Roles for the remaining transcription factors in the differentiation of these motor neuron subtypes are unknown. For example, members of the POU (*ceh-6*) and CUT (*ceh-44*) classes of homeodomain protein families, which are well-established determinants of neuronal fate (Latchman 1999; Nepveu 2001), are selectively enriched in the Larval A-class list.

Conversely, five members of the nuclear hormone receptor family (*nhr-3, nhr-95, nhr-104, nhr-116* and F41B5.9) are preferentially expressed in embryonic A-type motor neurons. The extent to which these different combinations of transcription factors account for characteristics that distinguish embryonic and Larval A-class motor neurons can now be explored by genetic analysis.

A key morphological feature that distinguishes DA from VA motor neurons is clearly linked to differential levels of specific transcripts in Embryonic vs Larval A-class datasets. During embryonic development, DA motor neurons extend commissures that circumnavigate the body wall to innervate dorsal muscles. The dorsal trajectory of DA motor neuron outgrowth depends on the UNC-6/Netrin receptors, *unc-5, unc-40,* and the Receptor Protein Tyrosine Phosphatase (RPTP) *clr-1* genes (Hedgecock, Culotti *et al.* 1990; Chang, Yu *et al*. 2004), all three of which are enriched in the Embryonic A-class (EA) dataset (Fig. 4.12). In contrast, *unc-5*, *unc-40* and *clr-1* are not elevated in larval VA motor neurons which consequently innervate muscles on the ventral side. Guidance cues that govern the anteriorly-directed outgrowth of motor axons the dorsal and ventral nerve cords, respectively, are not known. However, a likely candidate to direct axonal outgrowth along the *C. elegans* anterior-posterior axis is Wingless (Wnt) signaling (Maloof, Whangbo *et al*. 1999; Whangbo and Kenyon 1999; Zinovyeva and Forrester 2005). In this regard, it is interesting that a comparison of the embryonic and Larval A-class motor neuron transcripts identifies two different Wnt receptors that are selectively enriched in either the DA (*lin-17*) or VA (*mig-1*) motor neurons. In addition, the transcript for the Wnt ligand *cwn-1* shows elevated expression in the Embryonic A-class dataset.

Comparisons to microarray profiles of *C. elegans* sensory neurons identify differentially expressed transcripts.

Colosimo, *et al*. (2004) used MAPCeL to profile the sensory neurons AFD and AWB (Colosimo, Brown *et al*. 2004). We find that < 20% of AFD/AWB enriched transcripts also show elevated expression in embryonic A-type motor neurons (Fig. 4.8F) (Additional File 11), a finding consistent with the distinct roles of these neuron classes in *C. elegans*. For example, the AFD-specific guanylate cyclase genes, *gcy-8* and *gcy-23*, are excluded from the enriched embryonic A-type motor neuron dataset (EA), whereas the A-class specific transcription factor, *unc-4*, is not found in the AFD/AWB profile (Additional File 11). In contrast, a significantly larger fraction (~43%) of AFD/AWB enriched transcripts, including *gcy-8* and *gcy-23*, are elevated in the Embryonic Pan-neural profile (Fig. 4.8E) (Additional File 11). Similar results were obtained when comparing the Larval Pan-neural and A-class datasets to a larval profile of chemosensory neurons (Kunitomo, Uesugi *et al*. 2005) (data not shown). These findings confirm the reliability of these neuron-specific profiling methods for identifying differentially expressed transcripts and confirm that the Pan-neural profiling approach is sufficiently sensitive to detect genes expressed in diverse cell types throughout the *C. elegans* nervous system.

Microarray profiles are consistent with gene expression topographic maps

We compared our data to a topographic map derived from 553 microarray experiments in which genes are assigned to specific "mountains" based on similarities in gene expression (Kim, Lund *et al*. 2001). In some instances, co-regulated genes were

grouped into specific functional subsets, thereby defining the "name" of the mountain. For example, mountain 6 contains many genes that are known to function in neurons. Neuronal transcripts identified in all four of our neuronal microarray experiments (Embryonic and Larval Pan-neural, Embryonic and Larval A-class) are significantly over-represented in the neuromuscular mountain (mountain 1) and one of the neuronal mountains (mountain 6). In contrast, transcripts in the embryonic muscle dataset are significantly under-represented in mountains 1 and 6 but are over-represented in the muscle mountain (mountain 16) (Fox, *et al*. submitted). These data further validate our neuronal expression profiles.

Detection of Expressed Genes (EGs)

We limited the analysis above to transcripts that show elevated expression in neurons relative to other cell types in order to focus on genes that may function predominantly in the nervous system. These microarray data, however, also include intensity values for a larger group of transcripts that may be broadly expressed in neurons as well as in other tissues. We define these transcripts as "Expressed Genes" or "EGs." In this work, we identified 7953 EGs in the MAPCeL profile of embryonic neurons using criteria that exclude transcripts that are likely to originate from the small fraction (~10%) of non-GFP cells in the FACS preparation (Fox, Von Stetina *et al*. 2005) (Additional Files 12). Similar considerations were employed to identify EGs in the datasets obtained with the mRNA-tagging method. First, the intensity values for each EG must be called "present" in a majority (e.g. 2/3) of microarray replicates. A second criterion takes into account low amounts of non-specific mRNA bound to the immunoprecipitating sepharose beads (See

Materials and Methods). We adopted a simple computational approach to filter out genes from the list of potential neuronal EGs that might be detectable in this background RNA because of high levels of expression in non-neuronal cells; transcripts with average normalized intensity values in the Reference dataset greater than corresponding average values in the neuronal mRNA-tagging profiles were removed from this list (See Materials and Methods). These considerations identified a total of 4033 EGs in the Larval Pan-neural dataset and 3320 in the Larval A-class profile (Additional File 13). This treatment is relatively stringent as it is likely to exclude at least some transcripts that may be ubiquitously expressed (e.g. "housekeeping" genes) or potentially more highly expressed in another tissue relative to the nervous system. This prediction is consistent with the finding that ~20% (509/2422) (Additinoal File 15) of transcripts identified in independent microarray experiments as highly enriched in GMIc (germline, muscle, or intestine plus the genes common to all three groups) remain in the list of Larval Pan-neural EGs (Additional File 13). In contrast, 48% (1172/2422) (Additional File 15) of transcripts enriched in these other tissues are included in the list of 6342 EGs in the larval Reference dataset (Additional File 13).

A comparison of all EGs in the larval and embryonic datasets described in this paper (i.e. Reference, Pan-neural, A-class motor neurons), in addition to the previously described Embryonic A-class dataset(Fox, Von Stetina *et al*. 2005), reveals a total of ~12,000 unique transcripts or 63% of the predicted genes represented on the *C. elegans* Affymetrix Gene Chip (Additional File 14). We note that ~1,600 of these EGs correspond to transcripts that have not been previously confirmed by ESTs (Additional File 16); a subset of 336 transcripts from this group are enriched in at least one of the

neuronal datasets, suggesting that they may have specific functions in *C. elegans* neurons.

Discussion

We have used two complementary microarray-based strategies to obtain comprehensive gene expression profiles of developing *C. elegans* neurons. In the MAPCeL method, GFP-labeled embryonic neurons were isolated by FACS for microarray profiling(Fox, Von Stetina *et al*. 2005). Because postembryonic neurons are not readily available for sorting(Christensen, Estevez *et al*. 2002), we used an alternative strategy, the mRNA-tagging method, to profile the larval nervous system(Roy, Stuart *et al*. 2002). In this approach, neuronal mRNAs were purified by immunoprecipitation from transgenic animals expressing an epitope-tagged RNA binding protein (FLAG-PAB-1) in larval neurons. Together, these microarray datasets identify 2488 transcripts that show elevated expression in the *C. elegans* nervous system relative to other tissues in at least one developmental stage (i.e. embryonic or larval) (Additional File 10). A bioinformatic query of WormBase confirmed an enrichment of known neural transcripts in these datasets (Fig. 4.3A). In addition, analysis of a representative group of newly constructed GFP reporters has confirmed *in vivo* neural expression of > 90% of previously uncharacterized genes on these lists (Table 4.1). We therefore conclude that these "Pan-neural" profiles provide accurate representations of gene expression in the *C. elegans* embryonic and larval nervous systems. These transcripts encode a broad array of functions. For example, as expected, ion channels, neurotransmitter receptors and synaptic vesicle components are highly represented (Fig. 4.7, Table 4.2, Additional File

4). In a striking indication of the complex signaling capacity of the *C. elegans* nervous system, most of the known peptide neurotransmitter genes (e.g. 20 of 23 FMRFamide genes or "flps") are enriched in the Larval Pan-neural dataset (Fig. 4.4, Additional File 4) (Li, Nelson *et al*. 1999). Neural functions for previously uncharacterized members of these gene families can now be assigned by genetic or RNAi analysis. With this possibility in mind, we tested the applicability of these expression data for predicting *in vivo* functions for genes in this dataset that are also included in a genome-wide interaction map or "Interactome" for *C. elegans* proteins (Li, Armstrong *et al*. 2004). This analysis revealed that proteins encoded by a subset of Pan-neural transcripts are linked to identified components of the synaptic vesicle cycle and therefore predicts that genetic or RNAi perturbation of these genes should result in neurotransmitter signaling defects (Fig. 4.10). In addition to finding transcripts that may have shared roles in both the embryonic and larval nervous system, these Pan-neural profiles have also identified a significant number of genes (71%, 1777/2488) that are differentially enriched in either embryonic or larval neurons. In the future, it will be interesting to determine if these genes define stage specific features of the developing nervous system.

The mRNA-tagging method can be used to generate gene expression profiles of specific neurons

In addition to detecting transcripts that are broadly expressed throughout the nervous system (i.e. synaptic vesicle components) the Pan-neural profiles also include genes that are selectively expressed in specific neurons. In most instances, these known assignments are based on promoter-GFP reporter constructs for a limited number of genes in a given neuron and are therefore incomplete. To test the applicability of the mRNA-tagging

strategy for obtaining a comprehensive gene expression profile of a specific subset of neurons, we utilized this approach to fingerprint a group of 18 larval cells largely comprised of A-type motor neurons(Miller and Niemeyer 1995; Lickteig, Duerr *et al*. 2001). This experiment revealed > 400 transcripts with enriched expression in these cells (Additional File 1). Although the majority (70%) of these transcripts also show elevated expression in the Larval Pan-neural profile (Fig. 4.8), a significant fraction of these mRNAs are exclusively enriched in the A-class dataset in this comparison and are therefore likely to represent genes with limited expression in the nervous system. These results indicate that the mRNA-tagging strategy can now be applied to monitor gene expression in specific *C. elegans* neurons and that this approach should detect neuron-specific genes with potential key roles in the specification or function of individual neuron types. Our findings confirm an earlier study in which a neuron specific promoter was used in conjunction with the mRNA-tagging strategy to identify transcripts that are highly expressed in a group of ~50 sensory neurons from *C. elegans* (Kunitomo, Uesugi *et al*. 2005). Our work provides the important technical advance, however, of substantially enhancing the sensitivity of this method; we show that reliable profiles can be obtained by amplifying nanogram quantities of mRNA whereas the method of Kunitomo *et al*. (2005) required micrograms of starting mRNA.

Despite the successful use of mRNA-tagging for these cell-specific profiling experiments, additional improvements in this method would be helpful. For example, with any given promoter, we sometimes observe FLAG-1::PAB-1 staining in the expected cell types as well as in additional ectopic locations (data not shown). This problem is unlikely to result from gene expression domains in the transgenic PAB-1

construct because the substitution of PAB-1 cDNA to remove all possible genomic PAB-1 regulatory sites did not rectify this problem (Von Stetina *et al*., unpublished data). Our solution has been to generate multiple transgenic lines for each construct until we obtain at least one line in which FLAG-PAB-1 expression is limited to the cells of choice. A second problem with this method is pull-down of non-specific mRNA bound to the anti-FLAG sepharose beads. We have reduced this background by including a stringent wash step with a low salt buffer, but additional treatments to remove this extraneous mRNA would enhance the sensitivity of this method (See Methods). Lastly, some promoters result in subviable transgenic lines or unpredictable genetic interactions that limit profiling experiments (Von Stetina, Fox *et al*. 2007) (data not shown). The biological mechanisms of these effects are unknown but have also been observed for PAB-1 mRNA-tagging lines in *Drosophila* (Yang, Edenberg *et al*. 2005).

Applications of cell specific microarray profiling methods

The mRNA-tagging strategy has been used to generate robust gene expression profiles of major *C. elegans* tissues (i.e., muscles, intestine, nervous system) (Roy, Stuart *et al*. 2002; Pauli, Liu *et al*. 2005) (this paper). By exploiting promoter elements with more limited expression, it has also been possible to extend this approach to specific subsets of neurons. These results suggest that mRNA-tagging can now be exploited to obtain gene expression profiles in a broad array of cell types at precisely defined developmental intervals. For example, an mRNA-tagging approach could be used to identify key genes that direct the rewiring of the motor neuron circuit during early larval development (White, Albertson *et al*. 1978). The combined profiling results reported in

this paper identify a set of 177 transcription factors showing enriched expression in neurons. Genetic analysis has established that many of these transcription factors regulate key aspects of neuronal differentiation and function (Chalfie and Au 1989; Miller, Shen *et al*. 1992; Prasad, Ye *et al*. 1998; Hallam, Singer *et al*. 2000; Altun-Gultekin, Andachi *et al*. 2001; Esmaeili, Ross *et al*. 2002; Hobert; Von Stetina, Treinin *et al*. 2006). Both the MAPCeL and mRNA-tagging approaches can now be utilized to generate comparisons of mutant vs wildtype profiles that should reveal transcription factor-regulated genes in specific neurons (Zhang, Ma *et al*. 2002; Von Stetina, Fox *et al*. 2007). Microarray profiling of mutants for other classes of proteins could also be utilized to reveal unexpected gene regulatory roles. For example, a comparison of mRNA-tagging datasets obtained from mutant vs wildtype animals, indicates that the conserved synaptic protein, RPM-1/Highwire, regulates gene expression throughout *C. elegans* nervous system (J. Watson, S. Von Stetina, D. Miller, unpublished results). The *C. elegans* nervous system is uniquely well-defined with a wiring diagram denoting chemical synapses and gap junctions among all 302 neurons. It should now be possible to exploit these cell-specific microarray profiling methods to define genes expressed in each type of neuron in this circuit. In turn, novel computational methods could be exploited to link specific subsets of these genes to roles in defining the connectivity architecture of this network (Kaufman, Dror *et al*. 2006; Varadan, Miller *et al*. 2006).

Towards defining the Transcriptome

A comparison of the three larval datasets described in this work [Reference, Larval Pan-neural, Larval A-class motor neuron] reveals that 1424 Expressed Genes

(EGs) are shared and are therefore likely to represent transcripts that function in a broad array of cell types. In contrast, a smaller number of transcripts are uniquely detected in either the Larval Pan-neural (1189) or Larval A-class motor neuron (435) datasets. The three embryonic datasets (Reference, Embryonic Pan-neural, Embryonic A-class motor neuron) commonly express 4995 EGs, with 280 EGs unique to Embryonic A-class motor neurons and 480 mRNAs selectively detected in the Embryonic Pan-neural profile. These findings suggest that microarray-based strategies to confirm *in vivo* expression of all predicted *C. elegans* genes or to identify new, previously unknown transcripts (e.g. tiling array profiles)(Manak, Dike *et al*. 2006), will require extraction of mRNA from a variety of specific cells and tissues with methods similar to those described here.

Conclusion

~9000 *C. elegans* genes represented on the Affymetrix array have annotated human homologs (Additional File 3). Roughly 5% (525) of these genes encode uncharacterized conserved proteins. Our combined microarray data have revealed that 108 of these transcripts are enriched in neurons (Additional File 24). The high conservation of this subset of genes from nematodes to humans indicates that the encoded proteins may play pivotal roles in neuronal function or specification. The MAPCeL and mRNA-tagging strategies provide sufficient temporal information to pinpoint the developmental period during which a gene may function, as well as the spatial resolution to define the neuron in which it is expressed. With the powerful molecular and genetic tools available to *C. elegans* researchers, it should now be possible to delineate the roles of these novel targets in the nervous system.

AUTHORS CONTRIBUTIONS

SEV initiated the Larval Pan-neural and A-class profiling experiments, generated the A-class PAB-1 transgenic, aided in data analysis, scored GFP reporters, and helped draft the manuscript. JDW generated the Reference profiles, helped generate the Larval Pan-neural and A-class profiles, compiled expression data from WormBase, scored GFP reporters, aided in data analysis, and helped draft the manuscript. RMF generated the embryonic pan-neural dataset and critiqued the manuscript. KLO generated the Perl scripts used to cull expression patterns from WormBase and helped analyze the data downloaded from WormBase. WCS generated transgenic animals expressing GFP reporters. PJR generated the Pan-neural PAB transgenic line and critiqued the manuscript. DMM oversaw all aspects of project and helped draft the manuscript.

ACKNOWLEDGMENTS

We thank Stuart Kim for support and for comments on the manscript; Denis Dupuy and Marc Vidal for GFP reporter plasmids; Cathy Alford and Jim Price of the VU Flow Cytometry Special Resource Center; Susan Barlow for technical assistance; Christian Schaffer and Marilyn Ritchie for updating Perl scripts used to annotate the data; Braden Boone, John Mote, Phillip Dexheimer and Shawn Levy of the Vanderbilt Microarray Shared Resource (VMSR) for help with microarray experiments and using GeneSpring software. This work was supported by NIH grants R01 NS26115 and P01 DK58212 (DMM), F31 NS043068 (SEV), F31 NS049743 and T32 MH64913 (JDW), F31

NS046293 (RMF), T32 HD07502 (SEV and RMF), P30 CA68485, P60 DK20593, P30 DK58404, HD15052, P30 EY08126, and P01 HL6744 and the Beckman Institute (PJR). Additional support for the microarray experiments was provided by DK58749 (A.L. George).

CHAPTER V

COMPARISONS OF WT-PICO AND IVT AMPLIFICATIONS EXPAND THE C. ELEGANS NEURAL TRANSCRIPTOME

Introduction

Cell-specific profiles can be generated from a variety of different sources. The tissue can be physically dissected, but this poses a problem if the desired cell types are rare or integrated within other cell types (Yang, Edenberg *et al*. 2005). Also, cells isolated in this way may be damaged. Labeled-tissue can be extracted using laser capture microdissection, which avoids damaging the tissue and provides the means to perform comparative molecular analysis on adjacent tissue (Emmert-Buck, Bonner *et al*. 1996; Bonner, Emmert-Buck *et al*. 1997; Dolter and Braman 2001; Kube, Savci-Heijink *et al*. 2007). FACS sorting is also a viable method, and can be used to significantly enrich for a target tissue type (Fox, Von Stetina *et al*. 2005; Nelson, Hempel *et al*. 2006; Von Stetina, Fox *et al*. 2007). However, FACS sorted cells are removed from their natural environment for an extended amount of time, and mRNA obtained and hybridized to a microarray may not represent the *in vivo* transcriptome. mRNA-tagging, on the other hand, provides a way to generate RNA from a specific tissue *in vivo* (Roy, Stuart *et al*. 2002; Kunitomo, Uesugi *et al*. 2005; Pauli, Liu *et al*. 2005; Yang, Edenberg *et al*. 2005; Von Stetina, Watson *et al*. 2007).

Several different methods have been used to profile *C. elegans* tissue. For instance, whole worm profiles of normal animals compared against mutant animals lacking a gonad have generated useful data that help define the development of the sex

organ(Reinke, Smith *et al*. 2000; Miller, Ruest *et al*. 2003). However, this method is not specific enough to detect changes when only a limited number of cells are affected by a specific mutation. FACS sorting has been used to successfully profile several different cell types, including a number of neuron subclasses (Zhang, Ma *et al*. 2002; Cinar, Keles *et al*. 2005; Fox, Von Stetina *et al*. 2005). The technique, however, is limited in that it can only obtain RNA from embryonically derived tissue (no method has yet been demonstrated that can separate the outer cuticle from the nervous system). Our lab and others have demonstrated the power of the mRNA-tagging technique to define a profile from a specific stage in *C. elegans* development (Roy, Stuart *et al*. 2002; Kunitomo, Uesugi *et al*. 2005; Pauli, Liu *et al*. 2005; Yang, Edenberg *et al*. 2005; Von Stetina, Watson *et al*. 2007). In this method, an epitope-tagged PolyA binding protein (PAB) is expressed in a specific tissue type. PAB is then cross-linked to the RNA via formaldehyde fixation, immunoprecipitated, and then the cross-link is removed, allowing separation of the RNA and PAB. We have used this method to profile the entire nervous system, in addition to profiling a subclass of neurons (A-class motor neurons) (see chapter 4). As only a limited quantity of mRNA can be isolated using mRNA-tagging, we have used a two round IVT RNA amplification method to generate sufficient material to hybridize to an array. IVT is a T7 based method in which mRNA is reverse transcribed into cDNA and subsequently in vitro transcribed into aRNA (Nygaard and Hovig 2006). Our efforts using this method indicate a strong enrichment for neural genes. However, a selection of genes that are known to be expressed exclusively in neurons were not enriched by this method. In this paper, we evaluate an alternative method of RNA amplification, WT-PICO, that requires significantly less time, fewer

steps, and less starting material than the traditional IVT amplification protocol (Singh, Maganti *et al*. 2005). WT-Pico uses a different strategy to amplify RNA (Figure 5.1). First, RNA is reverse transcribed into cDNA using a DNA/RNA chimeric primer. The RNA is then partially degraded by heating the reaction mixture. The fragments of RNA act as a primer for DNA polymerase, which displaces the RNA fragments. During the Ribo-SPIA step, RNase H, RNA/DNA chimeric primers, and DNA polymerase are added to the reaction. RNA duplexed with DNA is selectively digested by RNase H, which allows new RNA/DNA chimeric primers to bind the cDNA and facilitate the action of DNA polymerase. We have used this technique to profile neurons from the 2nd larval stage and compared our results to those obtained using IVT. We have confirmed that WT-PICO amplifies more efficiently than IVT and that its cDNA target results in a greater dynamic range and a cleaner overall signal with more present calls than with labeled aRNA obtained by IVT amplification. Our results confirm that the gene expression profile generated with WT-PICO is highly enriched for neuronal transcripts. The sum of microarray data from the pan neural-derived samples amplified by the IVT and WT-PICO methods identifies 2,134 transcripts with elevated expression in the *C. elegans* nervous system. The transcripts included in this neural enriched sample encode proteins with a broad array of predicted functions in the *C. elegans* nervous system. Only ~50% of these transcripts, however, are detected as enriched by both methods of RNA amplification. On the basis of this result, we suggest that both IVT and WT-PICO methods show significant nucleotide sequence bias and therefore that, where possible,

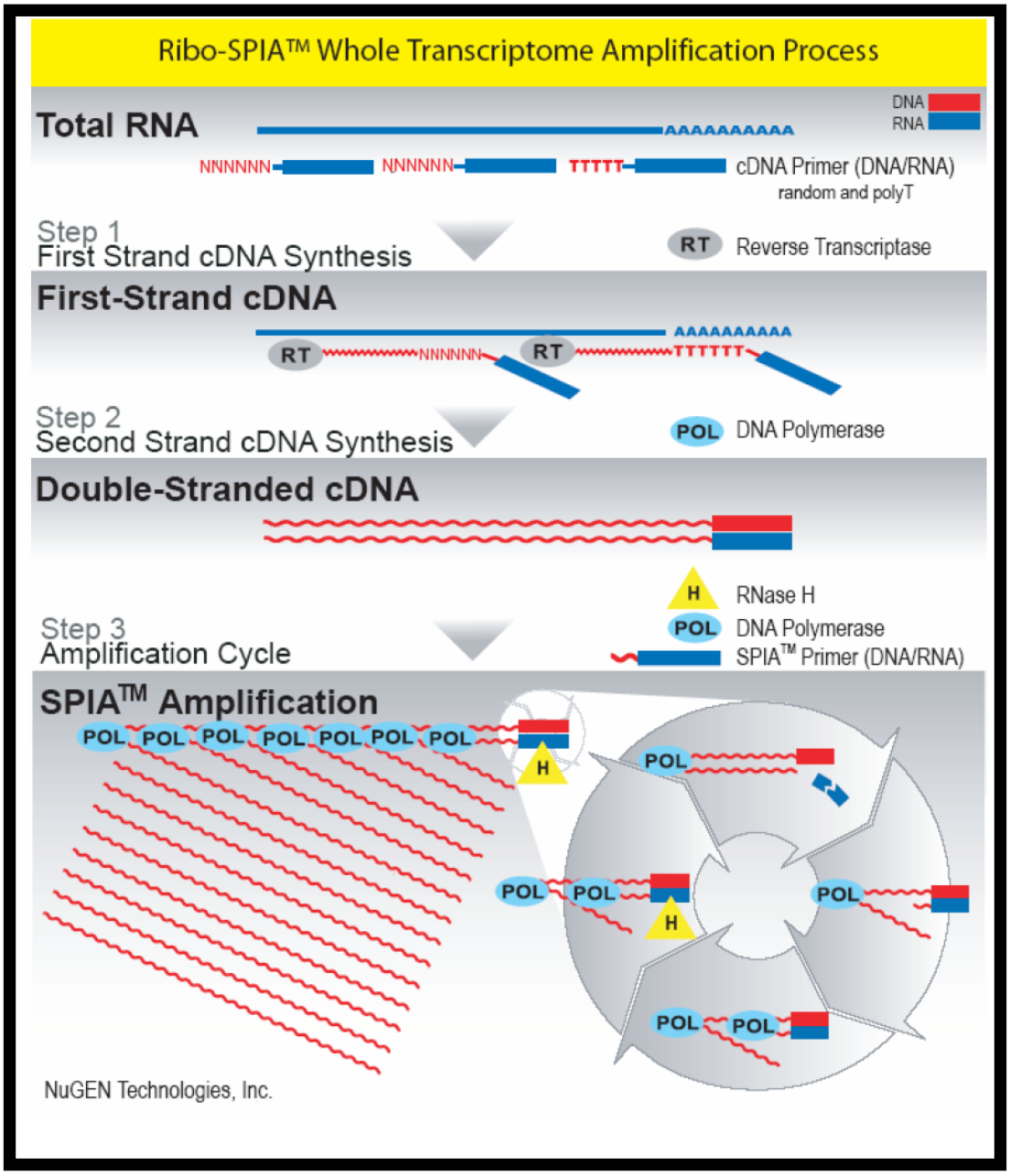


Figure 5.1 Diagram of the WT-Pico amplification method
In WT-Pico, chimeric cDNA(red)/RNA(blue) primers bind throughout the length of mRNA. The mRNA is reverse transcribed into cDNA. A heating step partially degrades the RNA strand. The fragments of RNA are used to prime the synthesis of the 2nd strand. SPIA amplification requires a series of steps. RNase H (yellow) selectively degrades duplexed RNA. Additional cDNA/RNA primers bind to sscDNA and are, in turn, extended by DNA polymerase. The reaction continues to cycle.

comprehensive gene expression profiles should be based on more than one method of RNA amplification.

Materials and Methods

Nematode strains

Nematodes were grown as described (Brenner 1974). Strains were maintained on nematode growth media (NGM) plates inoculated with the *E. coli* strain OP50. Strains used to isolate transcripts via mRNA-tagging were N2 (wildtype Bristol strain) and SD1241 (*gaIs153*, *F25B3.3*::FLAG::PAB-1).

Transgenic Generation

GFP reporters were selected at random from a subset of plasmids received from the Promoterome project. Microparticle bombardment was conducted as described (see above).

mRNA-tagging and RNA amplification

Worm stocks were grown as described (see above). Methods for mRNA-tagging are identical to those published in Von Stetina, Watson, *et al*. RNA from 3 Pan-neural lines and 5 N2 lines were amplified using IVT (described in detail in Von Stetina, Watson, *et al*). This same RNA was amplified by WT-Pico. 2 ng of starting material was amplified using version 1 of the WT-Pico Ribo-SPIA system as described in Singh, *et al* (the pRS method). 3 Pan-neural samples and 3 reference samples showed efficient

amplification. However, two reference samples failed to amplify. We then obtained RNA from two independent cultures of SD1241. This RNA amplified efficiently. Combined with the previous data, we generated a total of 5 Pan-neural and 3 N2 samples using WT-Pico.

Microarray data analysis

Analysis of enriched genes was performed as described with the following exceptions: a 3% FDR was used to generate the WT-Pico enriched list.

Annotation of datasets and additional data analysis

Annotation was performed as previously described except all *C. elegans* transcripts were annotated using WormBase Release 170 (WS170.wormbase.org). Affymetrix GeneChip Operating Software (GCOS) was used to calculate the average number of present calls in each sample. Genes with consistent present calls were determined using a perl script (consensus.pl). A gene was considered consistently present only if it was observed in all replicates.

RMA normalized intensities for consistently present genes were generated as previously described and then averaged across replicates. Pan-neural/reference intensities were calculated for WT-Pico and IVT data and $\log_2$ transformed. The Coefficient of determination (R^2) value for the resulting scatter plot was calculated in Microsoft Excel.

Mismatch intensities were compared against Perfect match intensities using the Bioconductor affy package. This generated a percentage of transcripts with higher

mismatch match than perfect match intensities. Line diagrams were generated as described previously.

Microscopy and identification of GFP expressing cells

GFP expressing animals were visualized by differential interference contrast (DIC) and epifluorescence microscopy using either a Zeiss Axioplan or Axiovert compound microscope. Digital images were recorded with CCD cameras (ORCA I, ORCA ER, Hammatsu Corporation, Bridgewater, NJ).

Results and Discussion

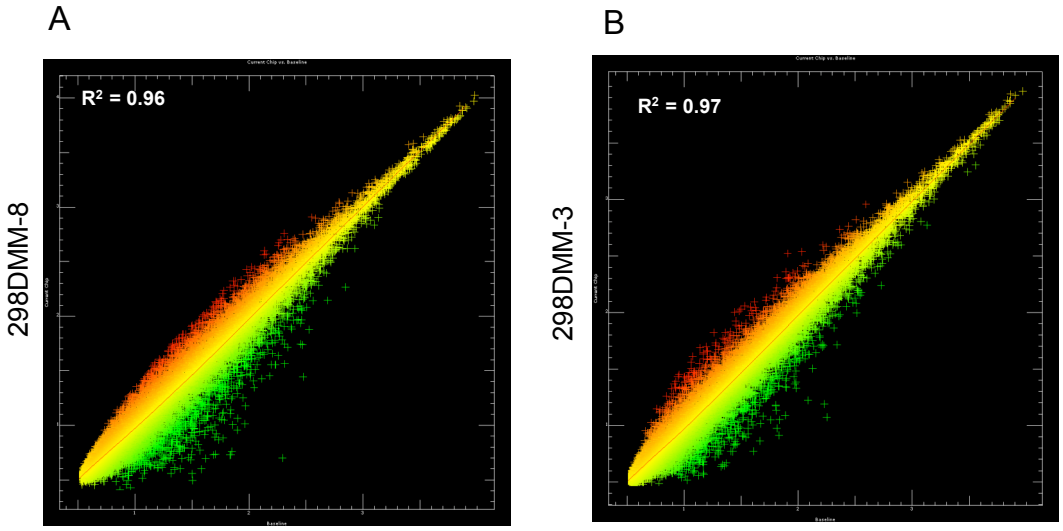
A comparison of two amplification methods, WT-PICO and IVT

We used the WT-PICO method to amplify RNA obtained from all *C. elegans* neurons ("Pan Neural") by the mRNA-tagging method; microarray data were generated from five independent Pan Neural RNA samples. A companion Reference data set was obtained with three replicates of RNA from all *C. elegans* cells. These results were compared to microarray data previously obtained from IVT-amplified samples (Von Stetina, Watson *et al*., 2007). A sum of eight Pan Neural and Reference data sets were produced for each amplification method (Table 1). All WT-PICO amplification reactions were performed with 2 ng of starting RNA whereas the IVT amplifications utilized 25 ng of sample RNA. Our results indicate that WT-PICO is more efficient than 2 rounds of IVT with a yield of ~4.6 ug of amplified cDNA/ng of starting RNA for WT-PICO vs ~1.6 ug of aRNA/ng of sample RNA for IVT ($p < 0.001$) (Table 1). Comparisons of signal intensities generated

from independent replicates showed that the WT-PICO-amplified Pan neural and Reference samples are reproducible (Fig 5.2). For example, the coefficient of correlation for five Pan Neural samples $R^2 = 0.96$ compares favorably to an $R^2 = 0.98$ for the three IVT-generated Pan-Neural profiles. It is important to note that the variance here includes the combined effects of independent sample preparation as well as RNA amplification reactions. Thus, these R^2 values are indicative of highly reproducible data sets.

We measured other parameters derived from the microarray data to compare the performance of the WT-PICO vs IVT-amplified targets. The *C. elegans* Affymetrix Gene Chip includes 22,499 probe sets. Overall, hybridizations with the WT-PICO amplified target sample resulted in a greater number of Present calls than the IVT target (Table 5.1). For example, 56% of the probe sets were scored as Present in Pan Neural profile obtained with the WT-PICO target vs 41% Present calls in the IVT generated Pan Neural dataset (Table 5.1) ($p < 0.05$). We then identified transcripts that are consistently present in all datasets for a group of replicates. We see a similar trend to that described above. For example, WT-Pico identified 9198 present genes while IVT identified 7382 present genes (Table 5.2). When probe sets to common transcripts are consolidated, these results identify ~17% more Present genes (7409) in the WT-PICO profile than in the IVT generated microarray data (6354) (Table 2). The greater number of Present calls derived from the WT-PICO data sets is correlated with the finding that the WT-PICO target results in relatively less mismatch hybridization than the IVT sample. On the Affymetrix Gene Chip, each Perfect Match (PM) oligonucleotide is paired with a MisMatch (MM) probe that includes a single base pair substitution. The hybridization intensity of each

Pico-amplified scatterplots



Reference R^2 values

	DMM6	DMM8
DMM6		
DMM8	0.89	
DMM9	0.88	0.88

Pan-neural R^2 values

	DMM2	DMM3	DMM4	DMM10
DMM2				
DMM3	0.92			
DMM4	0.89	0.92		
DMM10	0.91	0.93	0.92	
DMM11	0.91	0.93	0.92	0.97

Figure 5.2 WT-Pico amplified samples are well correlated

A. Scatter plot of normalized intensity values for a single reference sample (298-DMM-8) compared to the average of all reference samples.

B. Scatter plot of a single Pan-neural sample compared against the average of all Pan-neural samples. Pairwise comparisons (R^2) of individual hybridizations are listed below the figures.

Table 5.1 Hybridization and amplification summaries for WT-Pico and IVT amplifications							
	Starting Material (ng)	**Average Yield (ug)**	**Average yield/ng**	**No. of chips**	**Scale Factor**	**Average Intensity Values**	**Affymetrix average present calls/chip**
IVT neural	25	44.8	1.8	3	4.4	817.0	41.4
IVT reference	25	38.8	1.6	5	5.4	898.6	42.1
Pico neural	2	10.5	5.3	5	9.6	841.2	56.1
Pico reference	2	7.6	3.8	3	16.4	991.2	46.7

Table 5.2 Total number of transcripts identified by RNA amplification			
	Total number of present probesets	**Total number of present genes**	
IVT neural	7382	6354	
IVT reference	7325	6302	
Pico neural	9198	7409	
Pico reference	7771	6238	

MM probe is subtracted from that of the paired PM probe to correct for stray signal. An overall PM vs MM discrimination score for the probe set is calculated from these values to distinguish between Present, Marginal or Absent transcripts. For the combined IVT-amplified Pan Neural and reference samples, we find that 29 $\pm$ 0.5% of MM signals exceed the paired PM value whereas only 24 $\pm$ 1% of the WT-PICO derived signals show MM>PM ratios. These differences are statistically significant ($p < 0.01$) (Fig 5.3). Similar results have been noted previously and attributed to the finding that mismatched RNA:DNA heteroduplexes are thermodynamically more stable than comparable DNA:DNA hybrids (Singh, Maganti et al. 2005; Eklund, Turner et al. 2006).

Neuron-enriched transcripts are identified by both the WT-PICO and IVT amplification methods

To test the ability of the WT-PICO amplified sample to detect differentially expressed transcripts, the Pan Neural data set was compared to the Reference profile obtained from all cells (see Methods). As expected, scatter plots reveal significant differences between these datasets with 1625 transcripts showing elevated intensity values in the pan neural sample vs 1325 depleted mRNAs (Fig 5.4). As an independent test of the validity of these data, the list of enriched genes was compared to WormBase to identify the subset of transcripts previously described as expressed in neurons (Von Stetina, Watson *et al*., submitted). This analysis revealed 520 transcripts in the WT-PICO-amplified data set with known expression patterns *in vivo*. Of these, 85% are annotated in WormBase as expressed in neurons (Fig 5.5). This finding is comparable to the observation that 90% of the 518 transcripts in the enriched IVT-amplified pan neural profile with expression data

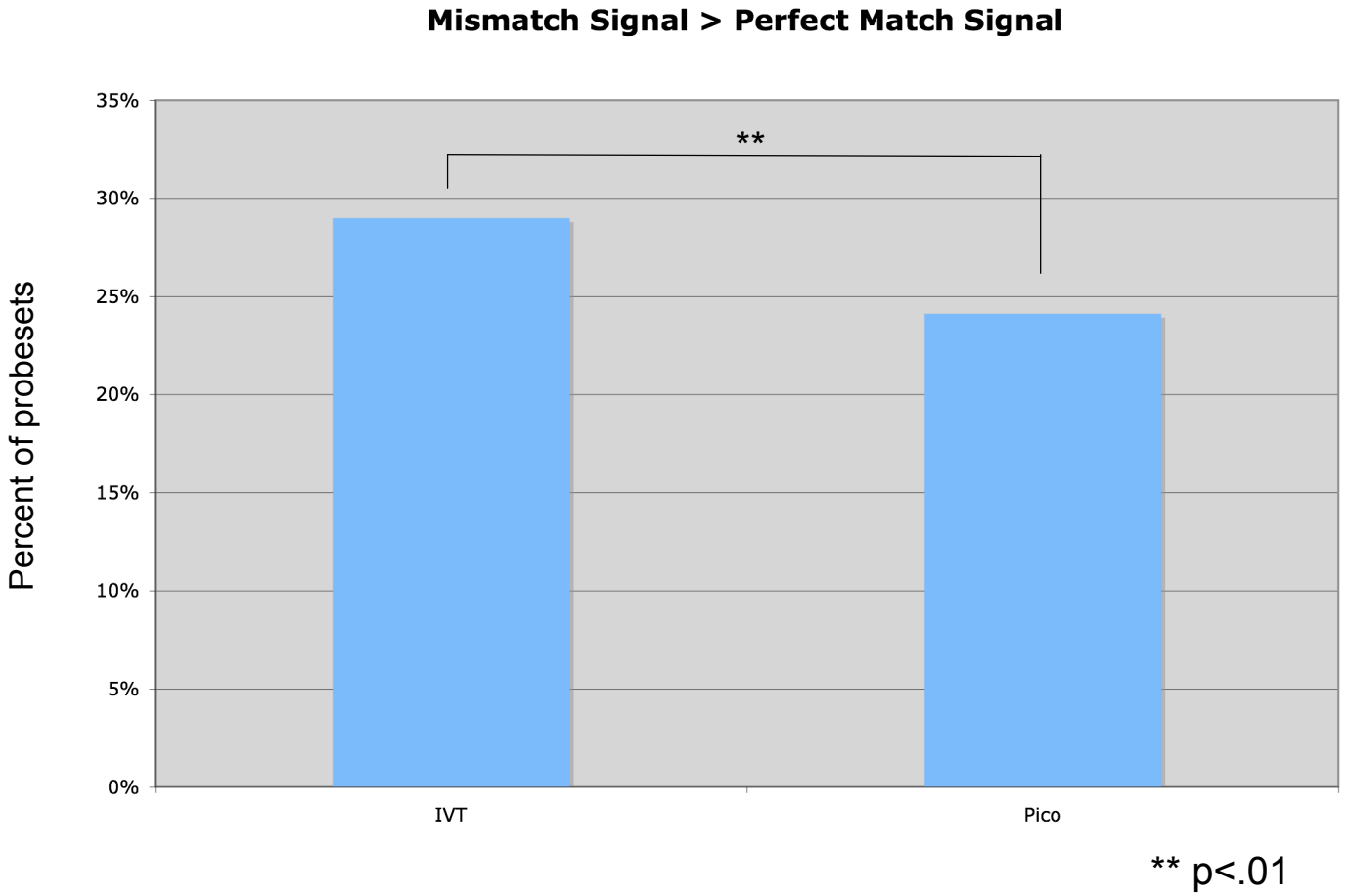


Figure 5.3 WT-Pico amplifications reduce non-specific binding
Mismatch intensities were higher than perfect match intensities more often in IVT than in WT-Pico.

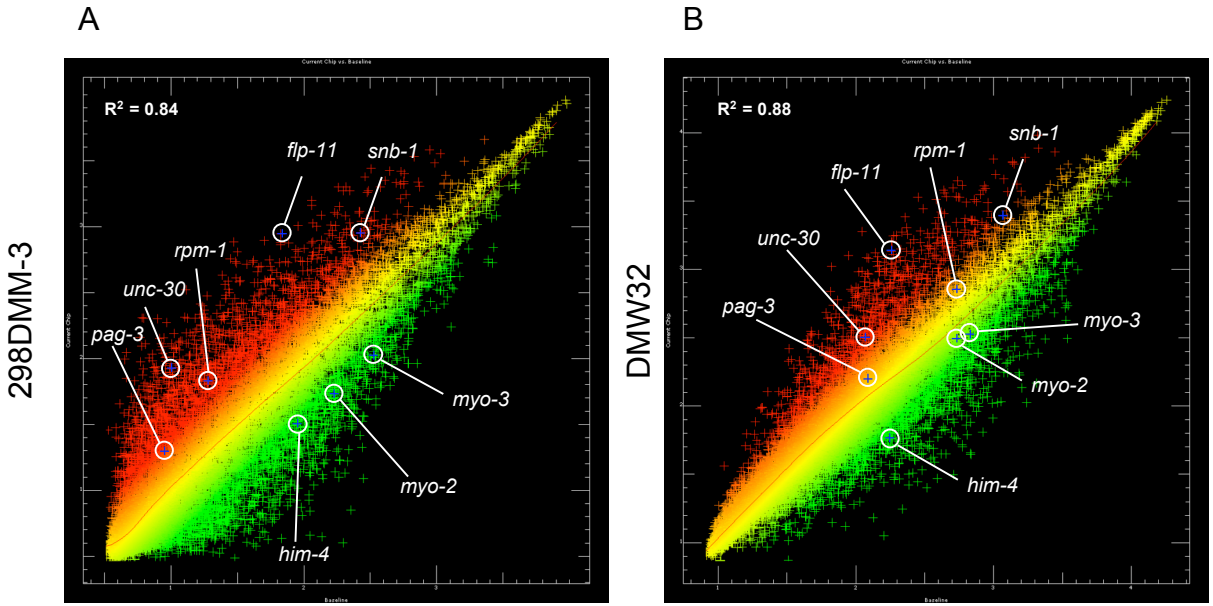
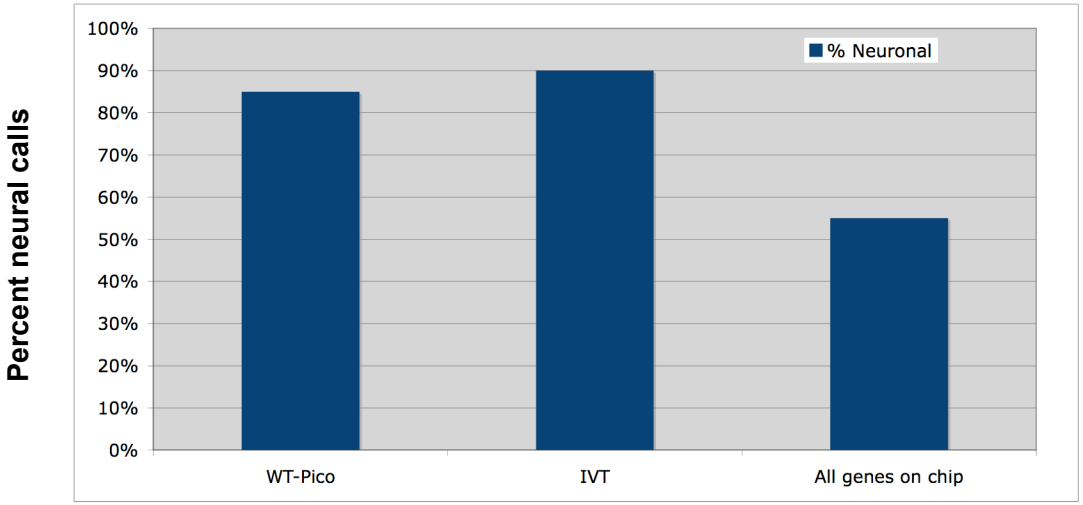


Figure 5.4 Scatter plots reveal significant differences between WT-Pico amplified neural datasets and reference

A. A single WT-Pico Pan-neural sample (298-DMM-3) compared against the average reference. Transcripts in red are enriched, whereas green transcripts are depleted.

B. IVT amplification of Pan-neural samples also identifies enriched and depleted transcripts. Selected genes enriched in both WT-Pico and IVT include *snb-1* (synaptobrevin), *flp-11* (FMRF-like peptide), and *unc-30* (homeodomain transcription factor). Genes selectively enriched in WT-Pico include *rpm-1* (E3 ubiquitin ligase) and *pag-3* (Zn-finger transcription factor). Note the wider dynamic range in WT-Pico.

A



B

	WT-Pico	IVT	All genes on chip
Total Transcripts	1625	1560	18696
Known expression patterns	520	518	3524
% Neuronal	85%	90%	55%
FDR	3%	1%	

Figure 5.5 WT-Pico identifies neural transcripts.
A. Known neural transcripts are highly represented in both the WT-Pico and IVT enriched datasets when compared against all genes on the microarray chip with known expression data.
B. A summary of WT-Pico and IVT expression data.

in WormBase are also detected in neurons. In both cases, the microarray profiles show a significant bias for authentic neuronal transcripts as only 55% of all genes with expression patterns listed in WormBase are neuronal (Fig 5.5). These findings confirm that both the WT-PICO and IVT amplification methods detect transcripts that are differentially expressed in the *C. elegans* nervous system.

To estimate the overall concordance of these data, we compared normalized intensity values for differentially expressed transcripts identified by each method. Log_2 of the IVT pan neural/Reference ratio was plotted versus that of the WT-PICO/Reference (Fig 5.6) for probe sets with Present calls in all of the Pan Neural samples (see methods). The R^2 value of 0.74 is indicative of significant correlation between these two amplification methods for these data, similar to that seen in previous comparisons of Ribo-SPIA and IVT (Singh, Maganti *et al*. 2005).

We expanded this comparison to consider all probe sets on the *C. elegans* chip. These results are depicted in Fig 5.7 in the form of a line graph in which the intensity values for all three of the IVT pan neural replicates and for the five WT-PICO Pan-neural samples are normalized against the corresponding average Reference intensities. Lines are color coded as enriched (red), depleted (blue) or unchanged (yellow) relative to the Reference. Colors for each gene are fixed by the relative values of sample #3 (vertical green line) in the WT-PICO data set. This global analysis shows an overall trend in which most of transcripts detected by both methods show similar patterns of differential expression. For example, 53 transcripts enriched in the IVT derived pan neural sample encode proteins with established or likely functions in neurotransmitter release at the synapse (Von Stetina, Watson, *et al*. 2007). 37 (70%) of these genes are also enriched in

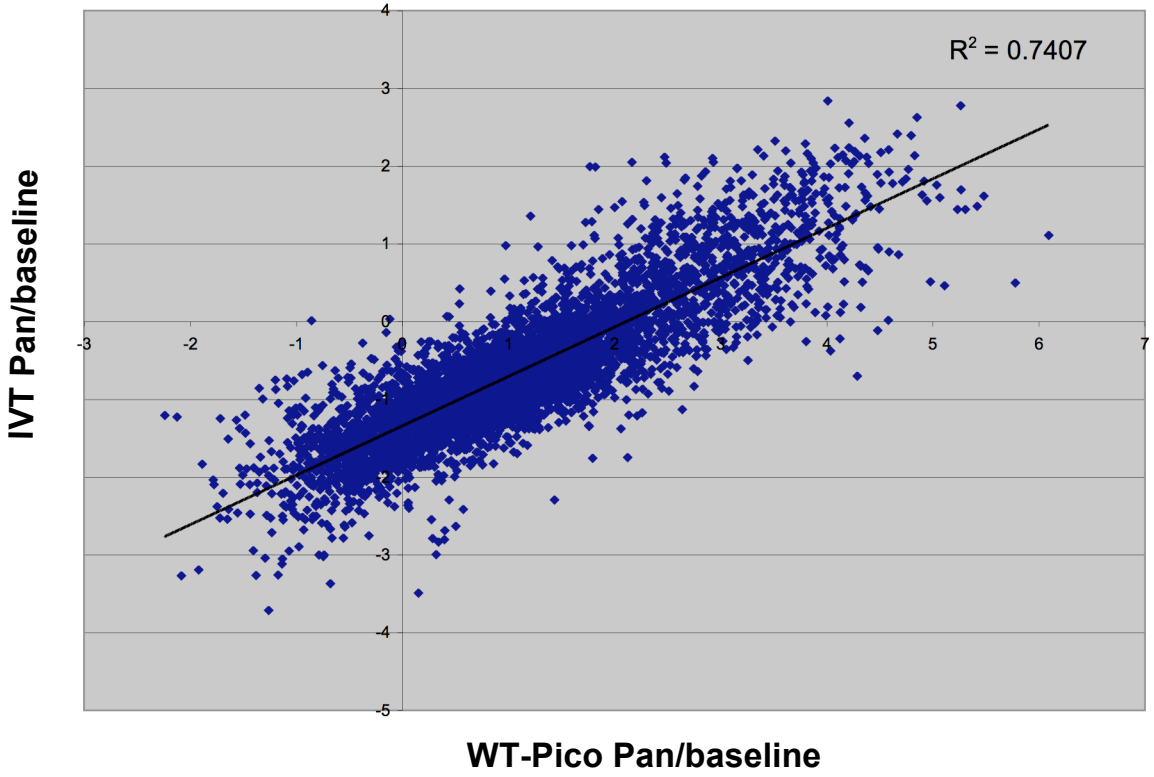


Figure 5.6 WT-Pico and IVT amplifications show significant correlation.
Ratios (log_2 of Pan-neural/Reference) of RMA normalized intensity values are plotted for genes identified as present in all Pan-neural amplifications.

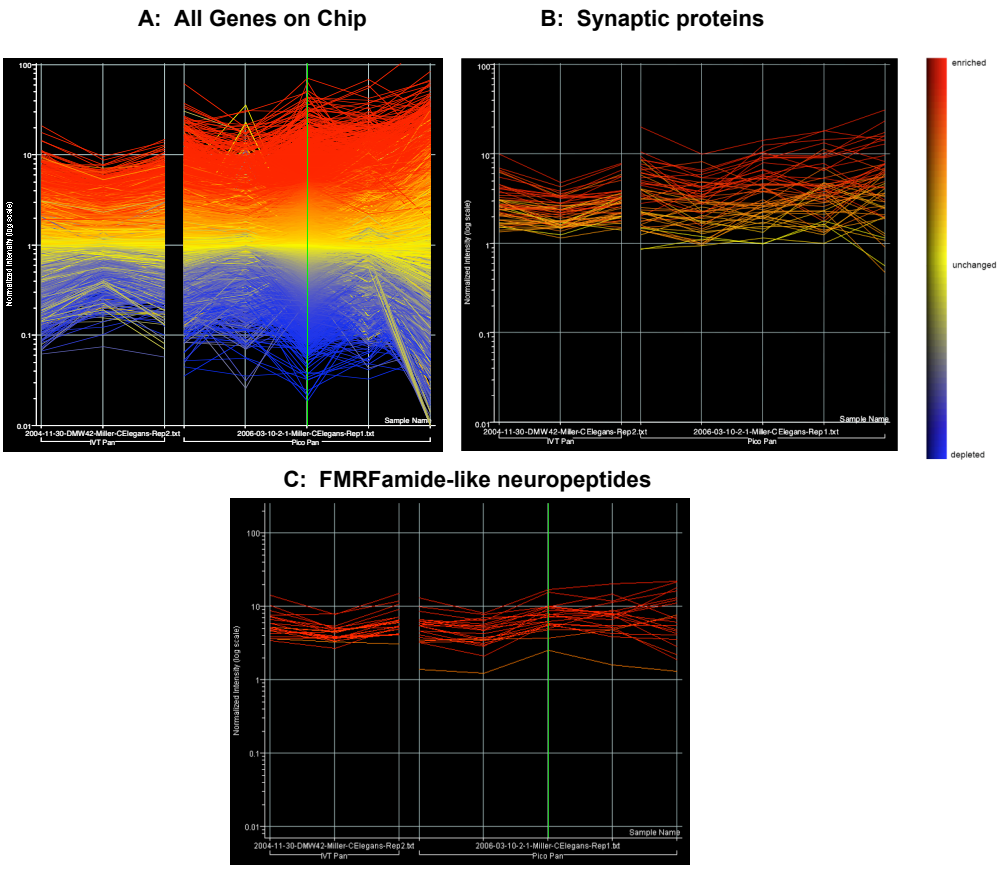


Figure 5.7 Similar patterns of differential expression are observed in WT-Pico versus IVT

A. A line graph depicting expression changes for all spots in each Pan-neural replicate relative to its baseline indicates that expression changes are conserved.
B. Many of the 53 synaptic proteins identified as enriched by IVT are also enriched in the WT-Pico dataset.
C.22 of 23 FMRFamide-like neuropeptides (flps) identified as enriched by IVT are also enriched in the WT-Pico dataset.
Red lines = enriched relative to baseline, yellow = unchanged, blue = depleted

the WT-PICO pan neural dataset and essentially all of these transcripts show intensity values that greater than or equal to the Reference (Fig 5.7B). Similar results were obtained for transcripts encoding FMRFamide-like proteins (flps), a large family of peptide neurotransmitters that are largely restricted to the *C. elegans* nervous system (Fig 5.7C) (Li, Kim *et al*. 1999). In addition to identifying similar trends in the relative intensity values of specific transcripts obtained by both methods, the line graphs also reveal a difference in the apparent overall spread of hybridization signals with the WT-PICO results showing a significantly larger dynamic range of differential expression vs the IVT data set (Fig 5.7A,B). A similar result was obtained in previous comparisons of IVT vs WT-PICO derived microarray data (Singh, Maganti *et al*. 2005).

Pico and IVT identify distinct neural transcripts

A comparison of the Pan neural enriched transcripts detected in these microarray experiments identifies a core group of 1050 genes that are detected by both the IVT and WT-PICO methods (Fig 5.8A). This analysis also revealed, however, that a comparable number of transcripts is selectively enriched in either the WT-PICO or IVT derived data sets. For example, 607 transcripts are detected as enriched in the WT-PICO pan neural sample but not in the IVT data set whereas 477 genes are specifically enriched in the IVT pan neural profile but not in the WT-PICO pan neural data set. These findings were validated by comparison to independently derived data that measures the expression and function of these genes in the *C. elegans* nervous system *in vivo*. First, we established that a majority of genes in either the IVT-only or WT-PICO only pan neural enriched data sets with known expression patterns listed in WormBase are annotated as expressed

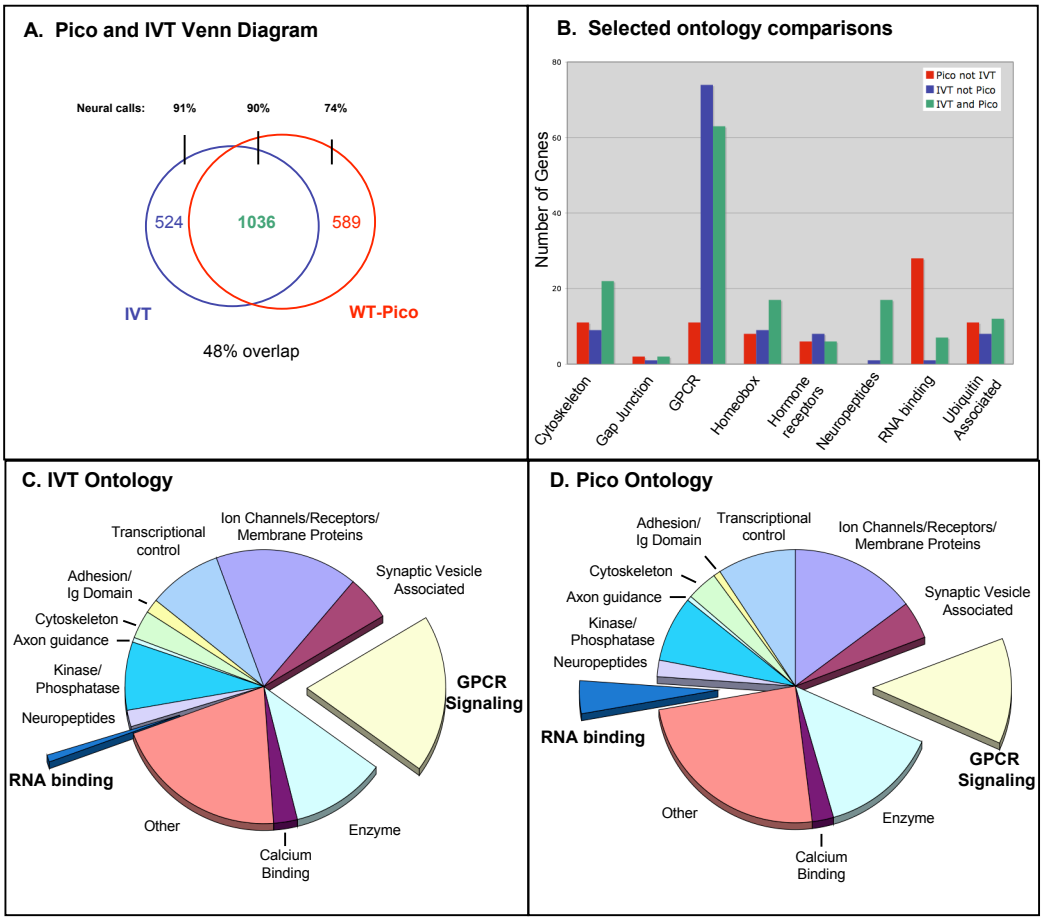


Figure 5.8 WT-Pico and IVT are selectively enriched for certain transcripts.

A. Venn Diagram highlights the similarities and differences between IVT and WT-Pico. As pictured, all three datasets are enriched for neural genes.

B. An ontological comparison of WT-Pico and IVT reveals subclasses of genes highly enriched in either WT-Pico or IVT. For instance, a large number of GPCRs are only identified by IVT (blue). Conversely, a large number of RNA binding proteins are only identified by WT-Pico (red). Neuropeptides are largely identified by both (green).

C and D. Pie charts highlighting protein class similarities. Note that GPCRs and RNA-binding proteins are selectively enriched by one amplification method (see above).

in neurons (Fig 5.8A). Additional genetic data have established specific neural functions for a subset of these differentially detected genes. For example, the WT-PICO-only subset of Pan Neural enriched transcripts includes *rpm-1*, an E3 ubiquitin ligase that regulates synaptic assembly. Similarly, the transcripts encoding the wnt receptor mig-13, which affects migration of the Q neuroblast and its descendants, is enriched exclusively in the IVT dataset. These findings suggest that each method of RNA amplification may result in the detection of a unique subset of *bona fide* pan neural enriched genes. We tested this idea by constructing GFP reporters for a representative set of genes listed in either the WT-PICO or IVT pan neural enriched data sets (Table 3). In this approach, the upstream promoter or regulator region of a specific gene is fused to GFP and reintroduced into the organism to monitor expression in the intact animal (see Methods). Nine transgenic lines were constructed from the WT-PICO only dataset. Neuronal GFP expression was confirmed in all 9 of these lines with reporters for two genes showing GFP expression exclusively in neurons (Fig. 5.9). Similarly, all eight of the GFP reporters built for genes in the IVT-only neural enriched dataset show expression in *C. elegans* neurons *in vivo* (Table 5.3) (Fig 5.9). Thus, these results strongly support the conclusion that the pan neural enriched data sets generated by each of these methods are reliably detecting transcripts expressed in the *C. elegans* nervous system.

Transcripts enriched in the IVT-only and WT-PICO-only pan neural data sets were classified according to Gene Ontology terms. A majority of the identified classes of proteins are equally represented in both data sets (Fig 5.8 C, D). Two gene families, however, are preferentially detected by one of these methods (Fig 5.8B). G-protein

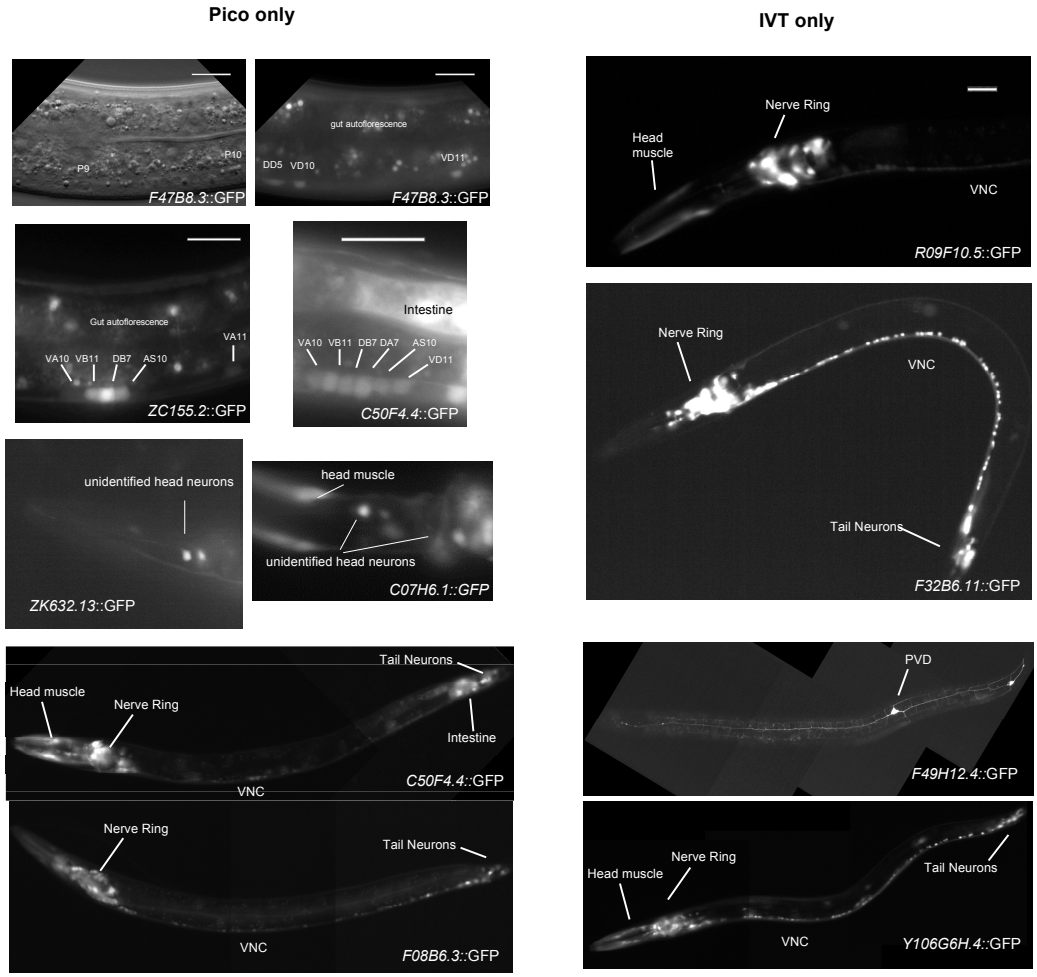


Figure 5.9 Both WT-Pico and IVT identify novel neural transcripts.

A. 9 transcripts were selected at random from the WT-Pico only dataset and promoter::GFP lines were created. All 9 lines showed expression in the nervous system.

B. 8 transcripts were selected at random from the IVT-only dataset and GFP lines were generated. 7 lines showed expression in the nervous system. One line failed to show GFP expression.

Table 5.3. Expression of promoter-GFP reporters for transcripts enriched in larval pan neural or WT-Pico amplified larval Pan-neural datasets.

Cosmid	Gene	KOG (other description)	IVT Fold Change	WT-PICO Fold Change	In Neurons?
IVT only					
F30F8.2		Glutaminase	1.6	--	√
F32B6.11		Unnamed Protein	2.5	--	√
F49H12.4		(novel)	2.6	--	√
H01A20.1	*nhr-3*	Nuclear Hormone Receptor	1.7	--	√
R09F10.5		(novel)	2.3	--	√
T19C4.5		(novel)	2.0	--	No expression
W01H2.3	*rab-37*	(Rab GTPase)	6.9	--	√
Y106G6H.4		Unnamed Protein	1.9	--	√
WT-PICO only					
C50F4.4		(novel)	--	2.0	√
F08B6.3		Reticulocalbin, calumenin, DNA supercoiling factor, and related Ca2+-binding proteins of the CREC family (EF-Hand protein superfamily)	--	2.7	√
K10B2.4		Predicted membrane protein	--	1.6	√
ZK632.13	*lin-52*	Uncharacterized conserved protein	--	2.9	√
C07H6.1	*lig-4*	ATP-dependent DNA ligase IV	--	2.9	√
F08G12.1		GTPase Rab1/YPT1, small G protein superfamily, and related GTP-binding proteins	--	3.6	√
F47B8.3		Glutaredoxin-related protein	--	2.0	√
T20G5.10		General control of amino-acid synthesis 5-like 1	--	2.7	√
ZC155.2		(putative nucleosome assembly factor)	--	2.8	√

GFP expression was typically determined in L2 larvae.

Coupled Receptors (GPCRs) represent one of the largest groups of structurally related proteins encoded by the *C. elegans* genome (GPCR-like genes comprise more than 2% of the genome in *C. elegans*). 74 GPCR transcripts are enriched in the IVT-only dataset whereas a separate group of 11 GPCRs are selectively enriched in the PICO-only pan neural profile. Conversely, RNA binding proteins are highly represented in the PICO-only data set with 28 members but not in the IVT-only profile which includes a single RNA binding protein transcript. We suggest that a careful analysis of the nucleotide sequences of these transcripts and their corresponding probe sets on the *C. elegans* Affymetrix Gene Chip could potentially identify the molecular basis for differential detection of these genes by these specific methods of RNA amplification.

Conclusions

While genetic screens can reveal key elements in a variety of biological pathways, many phenotypes can be masked by genetic redundancy and lethality. In order to understand the full complement of genes that specify a neuron at a particular time during development, a genomic strategy is needed. We have used two different amplification methods to generate the neural transcriptome for *C. elegans* at the 2nd larval stage. Our work indicates that both IVT and WT-Pico generate very similar datasets. However, the WT-Pico amplification is more robust than IVT. In addition, WT-Pico is technically less challenging, and can provide results within a single day. While many transcripts are identified as enriched in both amplification methods, we also observed as substantial number of genes identified by only one. We have confirmed that both the WT-Pico only and IVT-only datasets are enriched for neural transcripts. We have shown that a

combination of amplification methods provides enhanced resolution when determining expression profiles for the nervous system. We speculate that these data are applicable to many different types of tissue.

CHAPTER VI

GENERAL DISCUSSION AND FUTURE DIRECTIONS

My work has revealed tantalizing evidence for a novel mechanism linking synaptic development with transcription. These data support the hypothesis that the very large and highly conserved synaptic protein RPM-1 regulates transcription and that this activity is required for normal synaptic assembly. To identify candidate RPM-1 target genes, I implemented new microarray-based technology to profile gene expression in the *C. elegans* nervous system. In addition to revealing hundreds of candidate RPM-1 regulated genes, our approach has also generated the first comprehensive description of gene expression in *C. elegans* neurons. In collaboration with Stephen Von Stetina, we utilized these methods to profile the larval A class motor neurons. We then confirmed that these data sets are accurate representations of *C. elegans* gene expression in these cells. Finally, to faciliate the application of microarray profiling methods to single neuron types, I have demonstrated the utility of a new powerful method of RNA amplification, WT-PICO, that can generate reliable profiles from limited amounts of starting RNA. This discussion will focus on the power of using genomic approaches to profile neurons, and will also describe the limitations of our experimental design.

Transcription in Synaptic Development

Studies in *Drosophila, Aplysia* and *C. elegans* emphasize the importance of transcription during neural partner selection and synaptogenesis (Chapter 1). In

Drosophila, genetic approaches have revealed transcriptional regulators that function in conjunction with Highwire to shape synaptic development. In one of these pathways, BMP signals from muscle bind to presynaptic receptors, initiating a signaling cascade that ultimately activates the SMAD transcription factors MAD and Medea (McCabe, Marques et al. 2003; McCabe, Hom et al. 2004). Double mutants of *highwire* and *medea* rescue *highwire* arborization defects. However, this rescue is only partial, as synaptic boutons remain small, a phenotype characteristic of *Highwire* mutants. This could indicate that BMP signaling functions downstream of Highwire. It seems likely that the two pathways share transcriptional targets necessary for controlled arborization, but a separate pathway downstream of Highwire may regulate bouton size. Further work is needed to address these questions.

The discovery that Fos acts downstream of Highwire suggests that RPM-1 may act through a similar pathway (Collins, Wairkar et al. 2006) (Figure 6.1). Our data indicate that Fos is not downstream of *rpm-1*. The presumptive transcription factor downstream of RPM-1 has not yet been identified. Our microarray data contain many transcription factors, but these targets are transcriptionally regulated. Since transcriptional changes in *rpm-1* mutants are mediated by MAP Kinase, the transcription factor downstream of RPM-1 may be activated through phosphorylation rather than transcription. Further genetic screens, similar to the *unc-129*::GFP screen described in chapter 3, could reveal this target.

We have demonstrated the regulation of transcription by the E3 ubiquitin ligase RPM-1. This observation, and the discovery that RPM-1 acts through controlling MAP Kinase signaling, led to the hypothesis that RPM-1 regulates synaptic development via

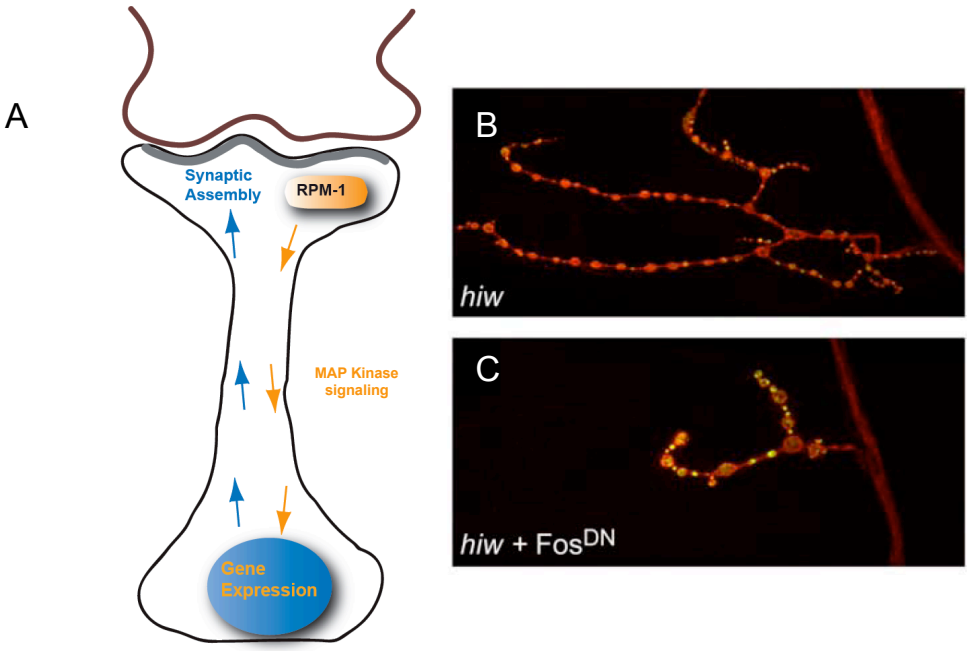


Figure 6.1 Model of the RPM-1 pathway

A. RPM-1 regulates transcription indirectly by modulating MAP Kinase signaling. We speculate that genes regulated transcriptionally by rpm-1 control synaptic assembly.
B. In Drosophila, mutations in highwire disrupt NMJ synapses.
C. Reducing Fos levels rescues highwire NMJ defects. Since rpm-1 and highwire are highly conserved, we speculate that transcription may influence synaptic development downstream of rpm-1.

transcription. I initially observed *rpm-1* regulation of gene expression using two readily available promoter::GFP reporter genes. However, since the random selection of GFP reporters is an inefficient way to ID downstream components, I opted to use a genomics approach to identify additional transcriptional targets involved in synaptogenesis. Comparing wildtype neuronal profiles against *rpm-1* neuronal profiles led to the discovery that *rpm-1* regulates approximately 700 transcripts. We identified one attractive candidate, the β-tubulin *tbb-6*, and confirmed its regulation by RT-PCR and with a *tbb-6*::GFP reporter that showed broad upregulation in the nervous system of *rpm-1* mutants. It was surprising that genetic analysis did not reveal a role for *tbb-6* in the *rpm-1* phenotype in either GABA motor neurons or in a specific subset of cholinergic motor neurons that innervate the muscles of the head. Given the critical role for tubulin in synaptic development, we think it is likely that RPM-1 regulation of *tbb-6* is important for synaptic development or axonal architecture in *C. elegans* neurons. The difficulty remaining is to identify the neurons in which these effects occur. This model is reasonable to consider given the well-established and context-dependent role of RPM-1 in different classes of neurons. For example, GABA neuron synapses are typically more severely affected than cholinergic synapses. Thus, in the future, we could identify the classes of neurons in which *tbb-6* is ectopically expressed, use neuron specific promoters to visualize their synapses and processes, and determine if *tbb-6* functions downstream of *rpm-1* in those specific neurons.

Another approach that could identify *rpm-1* transcriptional targets would be to profile specific neuron classes. We have successfully profiled the cholinergic Larval A-class motor neurons (Chapter 4). We can use this same strategy to identify *rpm-1*-

regulated transcripts in heavily affected neural classes. As mentioned above, GABAergic synapses are especially affected in *rpm-1* mutants. Sarah Anthony, a graduate student in our lab, has generated a GABA specific mRNA-tagging line that could reveal cell specific targets of *rpm-1*. This limited spatial resolution could help define the *rpm-1* pathway in a single cell type.

How is *rpm-1* itself regulated? This question remains to be answered. In the introduction, I mentioned that the synaptic guidepost *syg-2* is necessary to position HSN synapses at the appropriate position anterior to the vulva (Shen, Fetter et al. 2004). The synaptic defects observed in *rpm-1* mutants, suggest that RPM-1 may localize to the developing synapse early during synaptic development. The molecular identity of RPM-1 suggests that it may serve as a sensing mechanism to modulate protein levels at the developing synapse. It is therefore reasonable to suppose that assembly of RPM-1-regulated targets in the developing synapse could inhibit RPM-1 in a negative feedback loop that coordinates synaptic development with gene transcription.

Expression Profiles of the *C. elegans* nervous system

In order to identify the transcriptional targets of RPM-1, we first needed to profile the *C. elegans* nervous system. The first method we tested was mRNA-tagging (Roy, Stuart et al. 2002), which has the advantage of providing stage specific *C. elegans* profiles. One important outcome of this analysis was the discovery that previous mRNA profiles generated using this method were contaminated with non-specific mRNA (Roy, Stuart et al. 2002; Kunitomo, Uesugi et al. 2005; Pauli, Liu et al. 2005). We devoted a considerable amount of effort to reduce non-specific mRNA in these samples. The result

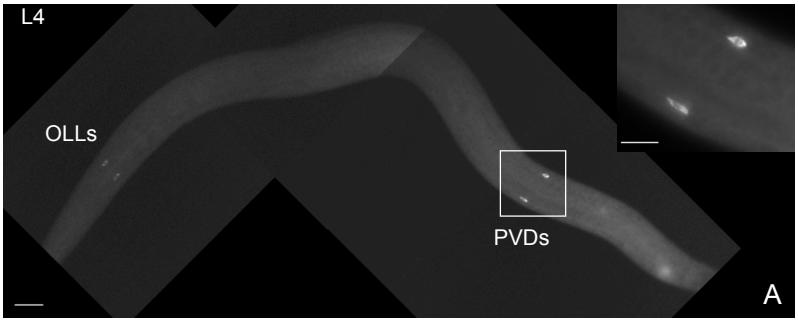
was the generation of a robust data set describing gene expression throughout the nervous system. We also demonstrated the application of this technique to a subset of neurons. Interestingly, additional work by Von Stetina, et al (2007) using the LA and EA datasets have revealed transcriptional targets of the homeodomain transcription factor UNC-4. One of these targets, *ceh-12*, functions downstream of *unc-4* to regulate synaptic choice. These data confirm that microarray profiles generated from mRNA-tagging can be used to answer complex biological questions.

We have also expanded the Pan-neural dataset using a relatively new amplification technique, WT-Pico (Singh, Maganti et al. 2005). A comparison of enriched genes between WT-Pico and IVT profiles revealed a significant amount of overlap, but also revealed subpopulations that were identified by only one of the amplification methods. These data have generated an additional ~500 neural transcripts. Work in our lab has shown the utility of these datasets. In an RNAi screen, Laurie Earls, a graduate student in the lab, has used a profile of embryonic GABA neurons to identify a member of the cell death pathway (Earls, Miller, unpublished data). Analyzing these regulated transcripts in more detail will likely shed light on additional biological paradigms.

We are now using the WT-Pico method to amplify RNA from limited cell types. For instance, Sarah Anthony and I are collaborating on developing a profile of larval stage ventral cord GABAergic neurons (Figure 6.2). We have also used this technique to develop a preliminary profile of the PVD mechnosensory neuron, which has elaborate processes and is thought to function in nociception (JDW, DMM, MT, unpublished data).

These techniques should reveal genes with previously uncharacterized functions and provide the C. elegans community with a repository of gene expression data.

ser-2prom3::3XFLAG-PAB-1



unc-25::GFP

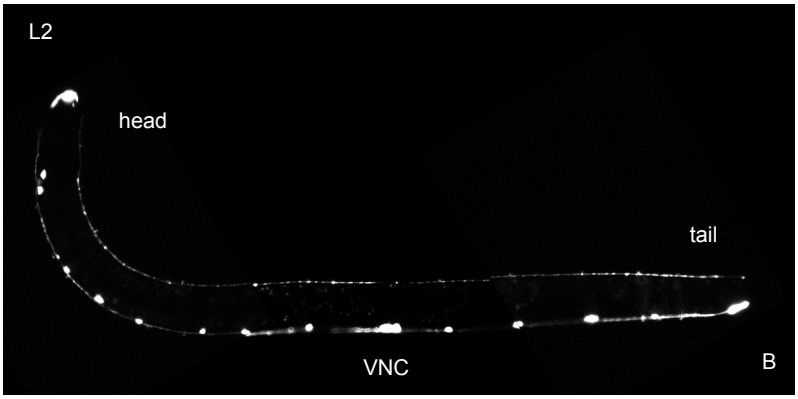


Figure 6.2 Profiling specific cell lines may provide the spatial resolution needed to identify functional RPM-1pathway components.

A. Line expressing a 3XFLAG-PAB-1 specifically in OLL interneurons and PVD mechanosensory neurons. The effect of rpm-1 on these neurons is not known.

B. A line expressing GFP in all GABA neurons. A 3XFLAG-PAB-1 line has been generated in our lab.

REFERENCES

Aberle, H., A. P. Haghighi, et al. (2002). "wishful thinking encodes a BMP type II receptor that regulates synaptic growth in Drosophila." Neuron **33**(4): 545-58.

Ackley, B. D., S. H. Kang, et al. (2003). "The basement membrane components nidogen and type XVIII collagen regulate organization of neuromuscular junctions in Caenorhabditis elegans." J Neurosci **23**(9): 3577-87.

Ahringer, J. (1996). "Posterior patterning by the Caenorhabditis elegans even-skipped homolog vab-7." Genes Dev **10**(9): 1120-30.

Ailion, M. and J. H. Thomas (2003). "Isolation and characterization of high-temperature-induced Dauer formation mutants in Caenorhabditis elegans." Genetics **165**(1): 127-44.

Altun-Gultekin, Z., Y. Andachi, et al. (2001). "A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans." Development **128**(11): 1951-69.

Arber, S., B. Han, et al. (1999). "Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity." Neuron **23**(4): 659-74.

Ataman, B., J. Ashley, et al. (2006). "Nuclear trafficking of Drosophila Frizzled-2 during synapse development requires the PDZ protein dGRIP." Proc Natl Acad Sci U S A **103**(20): 7841-6.

Ataman, B., V. Budnik, et al. (2006). "Scaffolding proteins at the Drosophila neuromuscular junction." Int Rev Neurobiol **75**: 181-216.

Beckstead, R. B. and C. S. Thummel (2006). "Indicted: worms caught using steroids." Cell **124**(6): 1137-40.

Blacque, O. E., E. A. Perens, et al. (2005). "Functional genomics of the cilium, a sensory organelle." Curr Biol **15**(10): 935-41.

Bonner, R. F., M. Emmert-Buck, et al. (1997). "Laser capture microdissection: molecular analysis of tissue." Science **278**(5342): 1481,1483.

Boyle, M. E., E. O. Berglund, et al. (2001). "Contactin orchestrates assembly of the septate-like junctions at the paranode in myelinated peripheral nerve." Neuron **30**(2): 385-97.

Brenner, S. (1974). "The genetics of Caenorhabditis elegans." Genetics **77**(1): 71-94.

Broadie, K. S. and J. E. Richmond (2002). "Establishing and sculpting the synapse in Drosophila and C. elegans." Curr Opin Neurobiol **12**(5): 491-8.

Burgess, R. W., K. A. Peterson, et al. (2004). "Evidence for a conserved function in synapse formation reveals Phr1 as a candidate gene for respiratory failure in newborn mice." Mol Cell Biol **24**(3): 1096-105.

Byrd, D. T., M. Kawasaki, et al. (2001). "UNC-16, a JNK-signaling scaffold protein, regulates vesicle transport in C. elegans." Neuron **32**(5): 787-800.

Cameron, S., S. G. Clark, et al. (2002). "PAG-3, a Zn-finger transcription factor, determines neuroblast fate in C. elegans." Development **129**(7): 1763-74.

Chalfie, M. and M. Au (1989). "Genetic control of differentiation of the Caenorhabditis elegans touch receptor neurons." Science **243**: 1027-1033.

Chang, C., T. W. Yu, et al. (2004). "Inhibition of netrin-mediated axon attraction by a receptor protein tyrosine phosphatase." Science **305**(5680): 103-6.

Chang, Q. and R. J. Balice-Gordon (2000). "Highwire, rpm-1, and futsch: balancing synaptic growth and stability." Neuron **26**(2): 287-90.

Chang, S., R. J. Johnston, Jr., et al. (2003). "A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans." Genes Dev **17**(17): 2123-37.

Chilton, J. K. (2006). "Molecular mechanisms of axon guidance." Dev Biol **292**(1): 13-24.

Christensen, M., A. Estevez, et al. (2002). "A primary culture system for functional analysis of C. elegans neurons and muscle cells." Neuron **33**(4): 503-14.

Chuang, C. F. and C. I. Bargmann (2005). "A Toll-interleukin 1 repeat protein at the synapse specifies asymmetric odorant receptor expression via ASK1 MAPKKK signaling." Genes Dev **19**(2): 270-81.

Cinar, H., S. Keles, et al. (2005). "Expression profiling of GABAergic motor neurons in Caenorhabditis elegans." Curr Biol **15**(4): 340-6.

Colavita, A., S. Krishna, et al. (1998). "Pioneer axon guidance by UNC-129, a C. elegans TGF-beta." Science **281**(5377): 706-9.

Collins, C. A. and A. DiAntonio (2007). "Synaptic development: insights from Drosophila." Curr Opin Neurobiol **17**(1): 35-42.

Collins, C. A., Y. P. Wairkar, et al. (2006). "Highwire restrains synaptic growth by attenuating a MAP kinase signal." Neuron **51**(1): 57-69.

Colosimo, M. E., A. Brown, et al. (2004). "Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types." Curr Biol **14**(24): 2245-51.

Consortium, T. C. e. S. (1998). "Genome Sequence of the nematode *C. elegans*: A platform for investigating biology." Science **282**: 2012-2018.

D'Souza, J., M. Hendricks, et al. (2005). "Formation of the retinotectal projection requires Esrom, an ortholog of PAM (protein associated with Myc)." Development **132**(2): 247-56.

Dai, Y., H. Taru, et al. (2006). "SYD-2 Liprin-alpha organizes presynaptic active zone formation through ELKS." Nat Neurosci **9**(12): 1479-87.

Dalma-Weiszhausz, D. D., J. Warrington, et al. (2006). "The affymetrix GeneChip platform: an overview." Methods Enzymol **410**: 3-28.

Dolter, K. E. and J. C. Braman (2001). "Small-sample total RNA purification: laser capture microdissection and cultured cell applications." Biotechniques **30**(6): 1358-61.

Drier, E. A., M. K. Tello, et al. (2002). "Memory enhancement and formation by atypical PKM activity in Drosophila melanogaster." Nat Neurosci **5**(4): 316-24.

Duggan, A., C. Ma, et al. (1998). "Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes." Development **125**(20): 4107-19.

Dunah, A. W., E. Hueske, et al. (2005). "LAR receptor protein tyrosine phosphatases in the development and maintenance of excitatory synapses." Nat Neurosci **8**(4): 458-67.

Dupuy, D., Q. R. Li, et al. (2004). "A First Version of the Caenorhabditis elegans Promoterome." Genome Res **14**(10B): 2169-75.

Eklund, A. C., L. R. Turner, et al. (2006). "Replacing cRNA targets with cDNA reduces microarray cross-hybridization." Nat Biotechnol **24**(9): 1071-3.

Emmert-Buck, M. R., R. F. Bonner, et al. (1996). "Laser capture microdissection." Science **274**(5289): 998-1001.

Esmaeili, B., J. M. Ross, et al. (2002). "The C. elegans even-skipped homologue, vab-7, specifies DB motoneurone identity and axon trajectory." Development **129**(4): 853-62.

Evans (ed.), T. C. Transformation and microinjection. Wormbook. T. C. e. R. Community, WormBook.

Finney, M. and G. Ruvkun (1990). "The unc-86 gene product couples cell lineage and cell identity in C. elegans." Cell **63**: 895-905.

Fox, R. M., S. E. Von Stetina, et al. (2005). "A gene expression fingerprint of C. elegans embryonic motor neurons." BMC Genomics **6**(1): 42.

Francis, M. M., S. P. Evans, et al. (2005). "The Ror receptor tyrosine kinase CAM-1 is required for ACR-16-mediated synaptic transmission at the C. elegans neuromuscular junction." Neuron **46**(4): 581-94.

Franco, B., L. Bogdanik, et al. (2004). "Shaggy, the homolog of glycogen synthase kinase 3, controls neuromuscular junction growth in Drosophila." J Neurosci **24**(29): 6573-7.

Gengyo-Ando, K. and S. Mitani (2000). "Characterization of mutations induced by ethyl methanesulfonate, UV, and trimethylpsoralen in the nematode Caenorhabditis elegans." Biochem Biophys Res Commun **269**(1): 64-9.

Gogel, S., S. Wakefield, et al. (2006). "The Drosophila microtubule associated protein Futsch is phosphorylated by Shaggy/Zeste-white 3 at an homologous GSK3beta phosphorylation site in MAP1B." Mol Cell Neurosci **33**(2): 188-99.

Goodman, M. B. and E. M. Schwarz (2003). "Transducing touch in Caenorhabditis elegans." Annu Rev Physiol **65**: 429-52.

Griffith, L. C. and V. Budnik (2006). "Plasticity and second messengers during synapse development." Int Rev Neurobiol **75**: 237-65.

Guan, Z., J. H. Kim, et al. (2003). "p38 MAP kinase mediates both short-term and long-term synaptic depression in aplysia." J Neurosci **23**(19): 7317-25.

Gunsalus, K. C., H. Ge, et al. (2005). "Predictive models of molecular machines involved in Caenorhabditis elegans early embryogenesis." Nature **436**(7052): 861-5.

Haghighi, A. P., B. D. McCabe, et al. (2003). "Retrograde control of synaptic transmission by postsynaptic CaMKII at the Drosophila neuromuscular junction." Neuron **39**(2): 255-67.

Hallam, S., E. Singer, et al. (2000). "The C. elegans NeuroD homolog cnd-1 functions in multiple aspects of motor neuron fate specification." Development **127**(19): 4239-52.

Hallam, S. J., A. Goncharov, et al. (2002). "SYD-1, a presynaptic protein with PDZ, C2 and rhoGAP-like domains, specifies axon identity in C. elegans." Nat Neurosci **5**(11): 1137-46.

Halpain, S. and L. Dehmelt (2006). "The MAP1 family of microtubule-associated proteins." Genome Biol **7**(6): 224.

Hedgecock, E. M., J. G. Culotti, et al. (1990). "The unc-5, unc-6, and unc-40 genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in C. elegans." Neuron **4**(1): 61-85.

Hillier, L. W., A. Coulson, et al. (2005). "Genomics in C. elegans: so many genes, such a little worm." Genome Res **15**(12): 1651-60.

Hobert, O. (2005). "Specification of the nervous system." *Wormbook*.

Hodgkin, J. (1983). "Male phenotypes and mating efficiency in Caeneorhabditis elegans." Genetics **103**: 43-64.

Horvitz, H. R., P. W. Sternberg, et al. (1983). "Mutations that affect neural cell lineages and cell fates during the development of the nematode Caenorhabditis elegans." Cold Spring Harb Symp Quant Biol **48 Pt 2**: 453-63.

Hummel, T., K. Krukkert, et al. (2000). "Drosophila Futsch/22C10 is a MAP1B-like protein required for dendritic and axonal development." Neuron **26**(2): 357-70.

Inada, H., H. Ito, et al. (2006). "Identification of guanylyl cyclases that function in thermosensory neurons of Caenorhabditis elegans." Genetics **172**(4): 2239-52.

Jacob, T. C. and J. M. Kaplan (2003). "The EGL-21 carboxypeptidase E facilitates acetylcholine release at Caenorhabditis elegans neuromuscular junctions." J Neurosci **23**(6): 2122-30.

Jayanthi, L. D., S. Apparsundaram, et al. (1998). "The Caenorhabditis elegans gene T23G5.5 encodes an antidepressant- and cocaine-sensitive dopamine transporter." Mol Pharmacol **54**(4): 601-9.

Jiang, G., L. Zhuang, et al. (2005). "A Na+/Cl- -coupled GABA transporter, GAT-1, from Caenorhabditis elegans: structural and functional features, specific expression in GABA-ergic neurons, and involvement in muscle function." J Biol Chem **280**(3): 2065-77.

Jin, Y. Synaptogenesis. Wormbook. T. C. e. R. Community, WormBook.

Jin, Y., E. Jorgensen, et al. (1999). "The Caenorhabditis elegans gene unc-25 encodes glutamic acid decarboxylase and is required for synaptic transmission but not synaptic development." J Neurosci **19**(2): 539-48.

Kamath, R. S. and J. Ahringer (2003). "Genome-wide RNAi screening in Caenorhabditis elegans." Methods **30**(4): 313-21.

Kamath, R. S., A. G. Fraser, et al. (2003). "Systematic functional analysis of the Caenorhabditis elegans genome using RNAi." Nature **421**(6920): 231-7.

Kandel, E. R. (2004). "The molecular biology of memory storage: a dialog between genes and synapses." Biosci Rep **24**(4-5): 475-522.

Kaufman, A., G. Dror, et al. (2006). "Gene Expression of Caenorhabditis elegans Neurons Carries Information on Their Synaptic Connectivity." PLoS Comput Biol **2**(12): e167.

Kaufmann, N., J. DeProto, et al. (2002). "Drosophila liprin-alpha and the receptor phosphatase Dlar control synapse morphogenesis." Neuron **34**(1): 27-38.

Keating, C. D., N. Kriek, et al. (2003). "Whole-genome analysis of 60 G protein-coupled receptors in Caenorhabditis elegans by gene knockout with RNAi." Curr Biol **13**(19): 1715-20.

Kemphues, K. (2000). "PARsing embryonic polarity." Cell **101**(4): 345-8.

Kennedy, S., D. Wang, et al. (2004). "A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans." Nature **427**(6975): 645-9.

Kim, K. and C. Li (2004). "Expression and regulation of an FMRFamide-related neuropeptide gene family in Caenorhabditis elegans." J Comp Neurol **475**(4): 540-50.

Kim, S. K., J. Lund, et al. (2001). "A gene expression map for Caenorhabditis elegans." Science **293**(5537): 2087-92.

Kittel, R. J., C. Wichmann, et al. (2006). "Bruchpilot promotes active zone assembly, Ca2+ channel clustering, and vesicle release." Science **312**(5776): 1051-4.

Koushika, S. P., J. E. Richmond, et al. (2001). "A post-docking role for active zone protein Rim." Nat Neurosci **4**(10): 997-1005.

Kube, D. M., C. D. Savci-Heijink, et al. (2007). "Optimization of laser capture microdissection and RNA amplification for gene expression profiling of prostate cancer." BMC Mol Biol **8**: 25.

Kummer, T. T., T. Misgeld, et al. (2006). "Assembly of the postsynaptic membrane at the neuromuscular junction: paradigm lost." Curr Opin Neurobiol **16**(1): 74-82.

Kunitomo, H., H. Uesugi, et al. (2005). "Identification of ciliated sensory neuron-expressed genes in Caenorhabditis elegans using targeted pull-down of poly(A) tails." Genome Biol **6**(2): R17.

Kyosseva, S. V. (2004). "Mitogen-activated protein kinase signaling." Int Rev Neurobiol **59**: 201-20.

Latchman, D. S. (1999). "POU family transcription factors in the nervous system." J Cell Physiol **179**(2): 126-33.

Lee, S. K. and S. L. Pfaff (2001). "Transcriptional networks regulating neuronal identity in the developing spinal cord." Nat Neurosci **4 Suppl**: 1183-91.

Li, C. (2005). "The ever-expanding neuropeptide gene families in the nematode Caenorhabditis elegans." Parasitology **131 Suppl**: S109-27.

Li, C. and M. Chalfie (1990). "Organogenesis in C. elegans: positioning of neurons and muscles in the egg-laying system." Neuron **4**(5): 681-95.

Li, C., K. Kim, et al. (1999). "FMRFamide-related neuropeptide gene family in Caenorhabditis elegans." Brain Res **848**(1-2): 26-34.

Li, C., L. S. Nelson, et al. (1999). "Neuropeptide gene families in the nematode Caenorhabditis elegans." Ann N Y Acad Sci **897**: 239-52.

Li, S., C. M. Armstrong, et al. (2004). "A map of the interactome network of the metazoan C. elegans." Science **303**(5657): 540-3.

Li, W., S. G. Kennedy, et al. (2003). "daf-28 encodes a C. elegans insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway." Genes Dev **17**(7): 844-58.

Liao, E. H., W. Hung, et al. (2004). "An SCF-like ubiquitin ligase complex that controls presynaptic differentiation." Nature.

Lickteig, K. M., J. S. Duerr, et al. (2001). "Regulation of neurotransmitter vesicles by the homeodomain protein UNC-4 and its transcriptional corepressor UNC-37/groucho in Caenorhabditis elegans cholinergic motor neurons." J Neurosci **21**(6): 2001-14.

Ling, D. S., L. S. Benardo, et al. (2002). "Protein kinase Mzeta is necessary and sufficient for LTP maintenance." Nat Neurosci **5**(4): 295-6.

Llamazares, S., G. Tavosanis, et al. (1999). "Cytological characterisation of the mutant phenotypes produced during early embryogenesis by null and loss-of-function alleles of the gammaTub37C gene in Drosophila." J Cell Sci **112 (Pt 5)**: 659-67.

Maloof, J. N., J. Whangbo, et al. (1999). "A Wnt signaling pathway controls hox gene expression and neuroblast migration in C. elegans." Development **126**(1): 37-49.

Manak, J. R., S. Dike, et al. (2006). "Biological function of unannotated transcription during the early development of Drosophila melanogaster." Nat Genet **38**(10): 1151-8.

Marques, G. (2005). "Morphogens and synaptogenesis in Drosophila." J Neurobiol **64**(4): 417-34.

Mathew, D., B. Ataman, et al. (2005). "Wingless signaling at synapses is through cleavage and nuclear import of receptor DFrizzled2." Science **310**(5752): 1344-7.

Matthies, D. S., P. A. Fleming, et al. (2006). "The Caenorhabditis elegans choline transporter CHO-1 sustains acetylcholine synthesis and motor function in an activity-dependent manner." J Neurosci **26**(23): 6200-12.

McCabe, B. D., S. Hom, et al. (2004). "Highwire regulates presynaptic BMP signaling essential for synaptic growth." Neuron **41**(6): 891-905.

McCabe, B. D., G. Marques, et al. (2003). "The BMP homolog Gbb provides a retrograde signal that regulates synaptic growth at the Drosophila neuromuscular junction." Neuron **39**(2): 241-54.

McKay, S. J., R. Johnsen, et al. (2003). "Gene expression profiling of cells, tissues, and developmental stages of the nematode C. elegans." Cold Spring Harb Symp Quant Biol **68**: 159-69.

Mello, C. and A. Fire (1995). "DNA transformation." Methods in Cell Biology **48**: 451-482.

Mertens, I., T. Meeusen, et al. (2005). "Molecular characterization of two G protein-coupled receptor splice variants as FLP2 receptors in Caenorhabditis elegans." Biochem Biophys Res Commun **330**(3): 967-74.

Miller, D. M., 3rd and C. J. Niemeyer (1995). "Expression of the unc-4 homeoprotein in Caenorhabditis elegans motor neurons specifies presynaptic input." Development **121**(9): 2877-86.

Miller, D. M., 3rd, C. J. Niemeyer, et al. (1993). "Dominant unc-37 mutations suppress the movement defect of a homeodomain mutation in unc-4, a neural specificity gene in Caenorhabditis elegans." Genetics **135**(3): 741-53.

Miller, D. M., M. M. Shen, et al. (1992). "C. elegans unc-4 gene encodes a homeodomain protein that determines the pattern of synaptic input to specific motor neurons." Nature **355**(6363): 841-5.

Miller, K. E., J. DeProto, et al. (2005). "Direct observation demonstrates that Liprin-alpha is required for trafficking of synaptic vesicles." Curr Biol **15**(7): 684-9.

Miller, M. A., P. J. Ruest, et al. (2003). "An Eph receptor sperm-sensing control mechanism for oocyte meiotic maturation in Caenorhabditis elegans." Genes Dev **17**(2): 187-200.

Misgeld, T., T. T. Kummer, et al. (2005). "Agrin promotes synaptic differentiation by counteracting an inhibitory effect of neurotransmitter." Proc Natl Acad Sci U S A **102**(31): 11088-93.

Much, J. W., D. J. Slade, et al. (2000). "The fax-1 nuclear hormone receptor regulates axon pathfinding and neurotransmitter expression [In Process Citation]." Development **127**(4): 703-12.

Mullen, G. P., E. A. Mathews, et al. (2006). "The Caenorhabditis elegans snf-11 gene encodes a sodium-dependent GABA transporter required for clearance of synaptic GABA." Mol Biol Cell **17**(7): 3021-30.

Murakami, S. and T. E. Johnson (2001). "The OLD-1 positive regulator of longevity and stress resistance is under DAF-16 regulation in Caenorhabditis elegans." Curr Biol **11**(19): 1517-23.

Nakata, K., B. Abrams, et al. (2005). "Regulation of a DLK-1 and p38 MAP kinase pathway by the ubiquitin ligase RPM-1 is required for presynaptic development." Cell **120**(3): 407-20.

Nakayama, T., T. Yaoi, et al. (1998). "N-copine: a novel two C2-domain-containing protein with neuronal activity-regulated expression." FEBS Lett **428**(1-2): 80-4.

Nass, R., D. H. Hall, et al. (2002). "Neurotoxin-induced degeneration of dopamine neurons in Caenorhabditis elegans." Proc Natl Acad Sci U S A **26**: 26.

Nelson, L. S., M. L. Rosoff, et al. (1998). "Disruption of a neuropeptide gene, flp-1, causes multiple behavioral defects in Caenorhabditis elegans." Science **281**(5383): 1686-90.

Nelson, S. B., C. Hempel, et al. (2006). "Probing the transcriptome of neuronal cell types." Curr Opin Neurobiol **16**(5): 571-6.

Nepveu, A. (2001). "Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development." Gene **270**(1-2): 1-15.

Ngo, S. T., P. G. Noakes, et al. (2007). "Neural agrin: a synaptic stabiliser." Int J Biochem Cell Biol **39**(5): 863-7.

Nonet, M. L., K. Grundahl, et al. (1993). "Synaptic function is impaired but not eliminated in C. elegans mutants lacking synaptotagmin." Cell **73**: 1291-1305.

Nygaard, V. and E. Hovig (2006). "Options available for profiling small samples: a review of sample amplification technology when combined with microarray profiling." Nucleic Acids Res **34**(3): 996-1014.

O'Connell, K. F., C. M. Leys, et al. (1998). "A genetic screen for temperature-sensitive cell-division mutants of Caenorhabditis elegans." Genetics **149**(3): 1303-21.

Odden, J. P., S. Holbrook, et al. (2002). "Drosophila HB9 is expressed in a subset of motoneurons and interneurons, where it regulates gene expression and axon pathfinding." J Neurosci **22**(21): 9143-9.

Ohtsuka, T., E. Takao-Rikitsu, et al. (2002). "Cast: a novel protein of the cytomatrix at the active zone of synapses that forms a ternary complex with RIM1 and munc13-1." J Cell Biol **158**(3): 577-90.

Okuda, T., T. Haga, et al. (2000). "Identification and characterization of the high-affinity choline transporter." Nat Neurosci **3**(2): 120-5.

Packard, M., E. S. Koo, et al. (2002). "The Drosophila Wnt, wingless, provides an essential signal for pre- and postsynaptic differentiation." Cell **111**(3): 319-30.

Panicker, A. K., M. Buhusi, et al. (2003). "Cellular signalling mechanisms of neural cell adhesion molecules." Front Biosci **8**: d900-11.

Patel, M. R., E. K. Lehrman, et al. (2006). "Hierarchical assembly of presynaptic components in defined C. elegans synapses." Nat Neurosci **9**(12): 1488-98.

Patrick, G. N. (2006). "Synapse formation and plasticity: recent insights from the perspective of the ubiquitin proteasome system." Curr Opin Neurobiol **16**(1): 90-4.

Pauli, F., Y. Liu, et al. (2005). "Chromosomal clustering and GATA transcriptional regulation of intestine-expressed genes in C. elegans." Development.

Powell, C. M., S. Schoch, et al. (2004). "The presynaptic active zone protein RIM1alpha is critical for normal learning and memory." Neuron **42**(1): 143-53.

Prasad, B. C., B. Ye, et al. (1998). "unc-3, a gene required for axonal guidance in Caenorhabditis elegans, encodes a member of the O/E family of transcription factors." Development **125**(8): 1561-8.

Putrenko, I., M. Zakikhani, et al. (2005). "A family of acetylcholine-gated chloride channel subunits in Caenorhabditis elegans." J Biol Chem **280**(8): 6392-8.

Rand, J. B., J. S. Duerr, et al. (2000). "Neurogenetics of vesicular transporters in C. elegans." Faseb J **14**(15): 2414-22.

Ranscht, B. (1988). "Sequence of contactin, a 130-kD glycoprotein concentrated in areas of interneuronal contact, defines a new member of the immunoglobulin supergene family in the nervous system." J Cell Biol **107**(4): 1561-73.

Reece-Hoyes, J. S., B. Deplancke, et al. (2005). "A compendium of Caenorhabditis elegans regulatory transcription factors: a resource for mapping transcription regulatory networks." Genome Biol **6**(13): R110.

Reinke, V., H. E. Smith, et al. (2000). "A global profile of germline gene expression in C. elegans." Mol Cell **6**(3): 605-16.

Ren, P., C. S. Lim, et al. (1996). "Control of C. elegans larval development by neuronal expression of a TGF-beta homolog." Science **274**(5291): 1389-91.

Reynolds, N. K., M. A. Schade, et al. (2004). "Convergent, RIC-8 Dependent G{alpha} Signaling Pathways in the C. elegans Synaptic Signaling Network." Genetics.

Riddle, D. L. and P. S. Albert (1997). Genetic and environmental regulation of dauer larva development. *C. elegans* II. D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess. Plainview, NY, Cold Spring Harbor Laboratory Press**:** 739-768.

Rogers, C., V. Reale, et al. (2003). "Inhibition of Caenorhabditis elegans social feeding by FMRFamide-related peptide activation of NPR-1." Nat Neurosci **6**(11): 1178-85.

Rogers, C. M., C. J. Franks, et al. (2001). "Regulation of the pharynx of Caenorhabditis elegans by 5-HT, octopamine, and FMRFamide-like neuropeptides." J Neurobiol **49**(3): 235-44.

Roos, J., T. Hummel, et al. (2000). "Drosophila Futsch regulates synaptic microtubule organization and is necessary for synaptic growth." Neuron **26**(2): 371-82.

Roy, P. J., J. M. Stuart, et al. (2002). "Chromosomal clustering of muscle-expressed genes in Caenorhabditis elegans." Nature **418**(6901): 975-9.

Ruiz-Canada, C., J. Ashley, et al. (2004). "New synaptic bouton formation is disrupted by misregulation of microtubule stability in aPKC mutants." Neuron **42**(4): 567-80.

Ruiz-Canada, C. and V. Budnik (2006). "Synaptic cytoskeleton at the neuromuscular junction." Int Rev Neurobiol **75**: 217-36.

Sagasti, A., N. Hisamoto, et al. (2001). "The CaMKII UNC-43 activates the MAPKKK NSY-1 to execute a lateral signaling decision required for asymmetric olfactory neuron fates." Cell **105**(2): 221-32.

Sakamoto, R., D. T. Byrd, et al. (2004). "The C. elegans UNC-14 RUN Domain Protein Binds to the Kinesin-1/UNC-16 Complex and Regulates Synaptic Vesicle Localization." Mol Biol Cell.

Sanyal, S. and M. Ramaswami (2006). "Activity-dependent regulation of transcription during development of synapses." Int Rev Neurobiol **75**: 287-305.

Schachat, F., R. L. Garcea, et al. (1978). "Myosins exist as homodimers of heavy chains: demonstration with specific antibody purified by nematode mutant myosin affinity chromatography." Cell **15**(2): 405-11.

Schaefer, A. M., G. D. Hadwiger, et al. (2000). "rpm-1, a conserved neuronal gene that regulates targeting and synaptogenesis in C. elegans." Neuron **26**(2): 345-56.

Sharma, S. K. and T. J. Carew (2004). "The roles of MAPK cascades in synaptic plasticity and memory in Aplysia: facilitatory effects and inhibitory constraints." Learn Mem **11**(4): 373-8.

Shen, K. and C. I. Bargmann (2003). "The immunoglobulin superfamily protein SYG-1 determines the location of specific synapses in C. elegans." Cell **112**(5): 619-30.

Shen, K., R. D. Fetter, et al. (2004). "Synaptic specificity is generated by the synaptic guidepost protein SYG-2 and its receptor, SYG-1." Cell **116**(6): 869-81.

Shi, S. H., L. Y. Jan, et al. (2003). "Hippocampal neuronal polarity specified by spatially localized mPar3/mPar6 and PI 3-kinase activity." Cell **112**(1): 63-75.

Shirasaki, R. and S. L. Pfaff (2002). "Transcriptional codes and the control of neuronal identity." Annu Rev Neurosci **25**: 251-81.

Sieburth, D., Q. Ch'ng, et al. (2005). "Systematic analysis of genes required for synapse structure and function." Nature **436**(7050): 510-7.

Singh, R., R. J. Maganti, et al. (2005). "Microarray-based comparison of three amplification methods for nanogram amounts of total RNA." Am J Physiol Cell Physiol **288**(5): C1179-89.

Speese, S. D. and V. Budnik (2007). "Wnts: up-and-coming at the synapse." Trends Neurosci **30**(6): 268-75.

Starich, T., M. Sheehan, et al. (2001). "Innexins in C. elegans." Cell Commun Adhes **8**(4-6): 311-4.

Stiernagle, T. Maintenance of C. elegans. Wormbook. T. C. e. R. Community, WormBook.

Stuart, J. M., E. Segal, et al. (2003). "A gene-coexpression network for global discovery of conserved genetic modules." Science **302**(5643): 249-55.

Sulston, J. E., D. G. Albertson, et al. (1980). "The Caenorhabditis elegans male: postembryonic development of nongonadal structures." Dev Biol **78**(2): 542-76.

Sulston, J. E. and H. R. Horvitz (1977). "Post-embryonic cell lineages of the nematode, Caenorhabditis elegans." Dev Biol **56**(1): 110-56.

Sulston, J. E., E. Schierenberg, et al. (1983). "The embryonic cell lineage of the nematode Caenorhabditis elegans." Dev Biol **100**(1): 64-119.

Swoboda, P., H. T. Adler, et al. (2000). "The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans." Mol Cell **5**(3): 411-21.

Sze, J. Y. and G. Ruvkun (2003). "Activity of the Caenorhabditis elegans UNC-86 POU transcription factor modulates olfactory sensitivity." Proc Natl Acad Sci U S A **100**(16): 9560-5.

Sze, J. Y., S. Zhang, et al. (2002). "The C. elegans POU-domain transcription factor UNC-86 regulates the tph-1 tryptophan hydroxylase gene and neurite outgrowth in specific serotonergic neurons." Development **129**(16): 3901-11.

Tatusov, R. L., N. D. Fedorova, et al. (2003). "The COG database: an updated version includes eukaryotes." BMC Bioinformatics **4**: 41.

Thaler, J., K. Harrison, et al. (1999). "Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9." Neuron **23**(4): 675-87.

Touroutine, D., R. M. Fox, et al. (2005). "acr-16 encodes an essential subunit of the levamisole-resistant nicotinic receptor at the Caenorhabditis elegans neuromuscular junction." J Biol Chem **280**(29): 27013-21.

Trotta, N., G. Orso, et al. (2004). "The hereditary spastic paraplegia gene, spastin, regulates microtubule stability to modulate synaptic structure and function." Curr Biol **14**(13): 1135-47.

Vandecandelaere, A., B. Pedrotti, et al. (1996). "Differences in the regulation of microtubule dynamics by microtubule-associated proteins MAP1B and MAP2." Cell Motil Cytoskeleton **35**(2): 134-46.

Varadan, V., D. M. Miller, 3rd, et al. (2006). "Computational inference of the molecular logic for synaptic connectivity in C. elegans." Bioinformatics **22**(14): e497-506.

Vogt, L., S. P. Schrimpf, et al. (2001). "Calsyntenin-1, a proteolytically processed postsynaptic membrane protein with a cytoplasmic calcium-binding domain." Mol Cell Neurosci **17**(1): 151-66.

Von Stetina, S. E., R. M. Fox, et al. (2007). "UNC-4 represses CEH-12/HB9 to specify synaptic inputs to VA motor neurons in C. elegans." Genes Dev **21**(3): 332-46.

Von Stetina, S. E., M. Treinin, et al. (2006). "The motor circuit." Int Rev Neurobiol **69**: 125-67.

Von Stetina, S. E., J. D. Watson, et al. (2007). "Gene expression profiling of C. elegans neurons by mRNA-tagging." submitted.

Wadsworth, W. G., H. Bhatt, et al. (1995). "Neuroglia and pioneer axons express UNC-6 to provide global and local netrin cues for guiding migrations in *Caenorhabditis elegans*." Neuron **16**: 35-46.

Wagh, D. A., T. M. Rasse, et al. (2006). "Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in Drosophila." Neuron **49**(6): 833-44.

Wan, H. I., A. DiAntonio, et al. (2000). "Highwire regulates synaptic growth in Drosophila." Neuron **26**(2): 313-29.

Wanders, R. J. (2004). "Peroxisomes, lipid metabolism, and peroxisomal disorders." Mol Genet Metab **83**(1-2): 16-27.

Wang, D., S. Kennedy, et al. (2005). "Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants." Nature **436**(7050): 593-7.

Way, J. C. and M. Chalfie (1988). "mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans." Cell **54**: 5-16.

Wes, P. D. and C. I. Bargmann (2001). "C. elegans odour discrimination requires asymmetric diversity in olfactory neurons." Nature **410**(6829): 698-701.

Whangbo, J. and C. Kenyon (1999). "A Wnt signaling system that specifies two patterns of cell migration in C. elegans." Mol Cell **4**(5): 851-8.

White, J. G., D. G. Albertson, et al. (1978). "Connectivity changes in a class of motoneurone during the development of a nematode." Nature **271**: 764-766.

White, J. G., E. Southgate, et al. (1992). "Mutations in the Caenorhabditis elegans unc-4 gene alter the synaptic input to ventral cord motor neurons." Nature **355**(6363): 838-41.

White, J. G., E. Southgate, et al. (1976). "Structure of the ventral nerve cord of *Caenorhabditis elegans*." Phil. Trans.R. Soc. Lond. **B275**: 327-348.

White, J. G., E. Southgate, et al. (1976). "The structure of the ventral nerve cord of Caenorhabditis elegans." Philos Trans R Soc Lond B Biol Sci **275**(938): 327-48.

White, J. G., E. Southgate, et al. (1986). "The structure of the nervous system of the nematode *Caenorhabditis elegans*." Philosophical Transactions of the Royal Society of London **B314**: 1-340.

Wicks, S. R., R. T. Yeh, et al. (2001). "Rapid gene mapping in Caenorhabditis elegans using a high density polymorphism map." Nat Genet **28**(2): 160-4.

Willems, A. R., M. Schwab, et al. (2004). "A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin." Biochim Biophys Acta **1695**(1-3): 133-70.

Winnier, A. R., J. Y. Meir, et al. (1999). "UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in Caenorhabditis elegans." Genes Dev **13**(21): 2774-2786.

Wixon, J., M. Blaxter, et al. (2000). "Caenorhabditis elegans." Yeast **17**(1): 37-42.

Xu, J., T. Mashimo, et al. (2007). "Synaptotagmin-1, -2, and -9: Ca(2+) Sensors for Fast Release that Specify Distinct Presynaptic Properties in Subsets of Neurons." Neuron **54**(4): 567-81.

Yan, Y. and K. Broadie (2007). "In vivo assay of presynaptic microtubule cytoskeleton dynamics in Drosophila." J Neurosci Methods **162**(1-2): 198-205.

Yang, Z., H. J. Edenberg, et al. (2005). "Isolation of mRNA from specific tissues of Drosophila by mRNA tagging." Nucleic Acids Res **33**(17): e148.

Zhang, Y., C. Ma, et al. (2002). "Identification of genes expressed in C. elegans touch receptor neurons." Nature **418**(6895): 331-5.

Zhang, Y. Q., A. M. Bailey, et al. (2001). "Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function." Cell **107**(5): 591-603.

Zhen, M., X. Huang, et al. (2000). "Regulation of presynaptic terminal organization by C. elegans RPM-1, a putative guanine nucleotide exchanger with a RING-H2 finger domain." Neuron **26**(2): 331-43.

Zhen, M. and Y. Jin (1999). "The liprin protein SYD-2 regulates the differentiation of presynaptic termini in C. elegans." Nature **401**(6751): 371-5.

Zinovyeva, A. Y. and W. C. Forrester (2005). "The C. elegans Frizzled CFZ-2 is required for cell migration and interacts with multiple Wnt signaling pathways." Dev Biol **285**(2): 447-61.