"The problem of neurology is to understand man himself."

-Dr. Wilder Penfield

Dedicated to my family.

REGULATION AND FUNCTION OF THE *Lhx* GENE, *lin-11*, IN

CAENORHABDITIS ELEGANS NERVOUS SYSTEM DEVELOPMENT

REGULATION AND FUNCTION OF THE *Lhx* GENE, *lin-11*, IN *CAENORHABDITIS ELEGANS* NERVOUS SYSTEM DEVELOPMENT

By **Siavash Amon**,

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2017)

Hamilton, Ontario (Department of Biology)

TITLE: REGULATION AND FUNCTION OF THE *Lhx* GENE, *lin-11*, IN *CAENORHAB-DITIS ELEGANS* NERVOUS SYSTEM DEVELOPMENT

AUTHOR: **Siavash Amon** (McMaster University)

SUPERVISOR: Dr. Bhagwati GUPTA

NUMBER OF PAGES: ix, 137

Abstract

Lhx genes are a sub-family of Hox genes that play important roles in animal development. In *Caenorhabditis elegans* there are seven *Lhx* genes, including the founding family member *lin-11*. The *lin-11* gene is necessary for the specification of neuronal and reproductive tissues. My thesis work has involved understanding the mechanism of *lin-11* regulation and its function in these tissues. To this end, I addressed two distinct but complementary questions, one of which focused on how transcriptional regulation of *lin-11* occurs and the second on the role of LIN-11 protein domains/regions.

My work on the transcriptional regulation has uncovered important roles of two of the largest *lin-11* introns, intron 3 and intron 7. These introns promote *lin-11* expression in non-overlapping sets of amphid neurons. Based on gene expression patterns and behavioural assays, intron 3 is capable of restoring *lin-11* function in *lin-11(n389)* null mutant allele. Comparison of intron 3-driven reporter expression in the neuronal cell types between *C. elegans* and *C. briggsae* has revealed *cis* and *trans* evolutionary changes in *lin-11* regulation between the two species. Functional dissection of the introns in *C. elegans* has led to the identification of three distinct non-overlapping enhancers, each specific for a single amphid neuron, i.e., RIC, AIZ, and AVG. I have also identified four transcription factors, SKN-1, CEH-6, CRH-1, and CES-1, that act through these enhancers to regulate neuronal expression of *lin-11*.

Furthermore, I have characterized the function of the LIM domains and a proline-rich (PRR) *C*-terminus region of LIN-11 in the specification of neuronal and reproductive tissues. My work shows that while the LIM domains are required for LIN-11 function in these tissues, the PRR region is dispensable. I have also examined the functional conservation of *lin-11* domains using two other *Lhx* genes, *Drosophila melanogaster* (*dLim1*) and *Mus musculus* (*Lhx1*), and found that both of these genes were able to rescue *lin-11* defects. Together, my work has significantly advanced our understanding of transcriptional regulation of *lin-11*, the importance of LIM domains in tissue formation, and functional conservation of *Lhx* genes across phyla.

Acknowledgements

There are a number of people who have contributed to my progress as a student over the years. First, I would like to thank my advisor, Dr. Bhagwati Gupta, for his guidance and support. You have always challenged me and pushed me to become a better researcher and to think critically.

My committee members, Dr. Rama Singh and Dr. Roger Jocobs, have been supportive, encouraging and provided helpful suggestions on experiments and my projects. Their advice and knowledge have been extremely important in guiding me and making graduate school a learning environment.

I would like to thank Dr. Lesley MacNeil for her support, motivation, and providing comments on my thesis and manuscript. I am always amazed at her ability to recall volumes of scientific knowledge about *C. elegans*. Also, I would also thank Andrea Tench, Sabhi Rashid and Romy Pabla for providing comments on my thesis.

I am grateful to my colleagues in the Gupta lab, both past and present, for their encouragements and help during those days when lab work was not going right.

I would like to express my gratitude to my family for their love and support throughout my life. They have encouraged me to dream big and pursue my goals in life. Thanks to my father for constantly reminding me to stay positive and the value of hard work. Thanks to my mother for teaching me to be patient and to remind me to not forget the brighter side of life. Thanks to my brothers, Jaihoon and Mansoor, for believing in me and motivated me through the ups and downs of graduate school. I will be forever grateful to you all.

Finally, I would like to thank my friends for reminding me to enjoy small things in life when I got too wrapped up in my work.

Contents

List of Figures

List of Tables

Acronyms

Ap - Apterous

C.el - Caenorhabditis elegans

C.br - Caenorhabditis briggsae

C.re - Caenorhabditis remanei

C.bre - Caenorhabditis brenneri

C.ne - Caenorhabditis negoni

C.se - Caenorhabditis senica

D. melanogaster - Drosophila melanogaster

M.musculus - Mus musculus (Mice)

Dfd - Deformed family homeodomain proteins

CFP - Cyan fluorescent protein

dsRED- *Discosoma sp.* red fluorescent protein

Egl - Egg-laying defect

GABA - γ-Aminobutyric acid

GFP - Green fluorescent protein

HD - Homeodomain (protein)

Hox - Homeobox (gene)

LIM - Conserved domain (*<u>L</u>in-11*, *<u>I</u>sl-2*, *<u>M</u>ec-3*)

Pvl- Protruding vulva

PRR - Proline-rich region

RNAi - RNA interference

TF - Transcription Factor

Utse - Uterine-seam cell

UTR - Untranslated region

VC - Ventral cord

VPC - Vulval precursor cell

Chapter 1.0

Introduction

1.1 Development of the nervous system

The brain is often described as the most sophisticated organ in an organism, and scientists have been fascinated by its complexity for decades. The human brain is composed of trillions of neurons that communicate using neurotransmitters, which pass between these neurons through hundreds of trillions of interconnections called synapses. All neurons are physically connected by cellular processes (axons/dendrites) that receive and transmit information via neurotransmitters in the presynaptic and postsynaptic regions respectively. One of the challenges in the field of neuroscience is to understand the molecular mechanisms that underlie the generation, positioning, and connection of neurons during development. The intricacy of the human brain presents a difficult system to utilize to study nervous system development; thus, a simpler model like *Caenorhabditis elegans (C. elegans)*, is required (Herculano-Houzel 2009; Koch 1999). *C. elegans* have a simple nervous system composed of only 302 neurons in adult hermaphrodites, and 388 neurons in adult males (Emmons 2005; Sulston and Horvitz 1977; Sulston et al. 1983). The 302 neurons in a hermaphrodite can be subdivided into 118 different neuronal classes based on their position, morphology, neurite projection patterns and synaptic connectivity, making them a simple system for understanding neurogenesis on small circuit levels (Hobert 2016). Together, the action of numerous transcription factors (TFs) is required for the development and function of the nervous system. The

expression of these transcription factors and their target genes is tightly regulated during development. Differentiation of progenitor cells toward a neural cell fate, and the formation of connections between these cells and other established neurons requires the coordination of a large number of genes, many of which encode transcription factors, e.g., Hox proteins (Avraham et al. 2009; Mallo and Alonso 2013; Zheng et al. 2015).

1.2   Genes important for nervous system development

Studies on the development of nervous system in eukaryotes have shown the involvement of many genes and gene families that are conserved across phyla. During neurogenesis, neuronal cells follow a highly stereotypic path of differentiation to give rise to functional neurons at various developmental stages. The homeobox genes are one family of genes that play key roles during differentiation and patterning of the nervous system. The Hox genes were discovered in *Drosophila melanogaster* (Bridges et al. 1923), and were later shown to exist in two separate gene clusters: the *Antennapedia* (ANT-C) and *Bithorax* (BX-C) complexes (Figure 1.1) (Kaufman et al. 1980). Ed Lewis' pioneering work on BX-C complex in *D. melanogaster* revealed that mutations in anterior and posterior genes of the BX-C complex affected anterior larval structures, whereas mutations in the posterior region of this complex affected posterior larval patterning. Lewis called this phenomenon, in which the morphology of a given appendage is transformed into the likeness of another, homeotic transformation (Lewis 1978). The sequential activation of Hox genes in the cluster is important, because premature or delayed Hox gene activation has been shown to produce phenotypic alterations in organisms (Kondo et al. 1998; Zákány et al. 1997). There are many examples that demonstrate homeotic transformation in *D. melanogaster*, including flies with feet structures in place of mouthparts, extra pairs of wings, and legs replacing the antennae (Aplin and Kaufman 1997; Beachy 1990; Carroll et al. 1995; Lewis 1978; Shimell et al. 1994).

Genetic studies, molecular cloning (Bender et al. 2004) and reporter expression analyses demonstrated that in *D. melanogaster*, Hox genes are expressed in a collinear manner along the antero-posterior axis of the embryo (Akam 1987; Duboule and Dollé 1989; Hobert et al. 1998). The collinearity of the Hox genes plays important roles in development, because the genes in a specific complex become sequentially activated generally from 3' to 5' during embryogenesis (Kmita and Duboule 2003; Mallo et al. 2010). In addition to the arrangement of the Hox genes influencing the pattern of their transcription, many *cis*-regulatory elements that modulate Hox gene expression have been identified, which allowed for better understanding of their regulatory mechanisms (Akbari et al. 2006). Furthermore, it has been shown that the structure of Hox cluster can vary amongst different species. For example, the Hox clusters in *D. buzzattii* and *D. melanogaster* differ, but the overall expression pattern of genes within the clusters remains the same (Negre et al. 2005). This finding may suggest that the integrity of the Hox gene complex may not necessarily be required for the establishment of Hox expression patterns. While much work has been done to understand the Hox genes in *D. melanogaster*, Hox genes have also been identified in vertebrate and invertebrate organisms, including *Mus musculus*, *Xenopus laevis*, and *C. elegans* (Duboule and Dollé 1989; Hobert and Westphal 2000).

The Hox proteins, encoded by the Hox genes, are characterized by the presence of conserved 60 amino acid homeodomain (HD) composed of helix-turn-helix motifs (Qian et al. 1989). HD containing proteins are transcription factors (TFs) that bind to DNA sequences in the major groove of the double helix to regulate the expression of their target genes (McGinnis and Krumlauf 1992; Passner et al. 1999). Specifically, in vitro, Hox proteins can bind as monomers and multimers to specific DNA sequences (McGinnis and Krumlauf 1992). Identification and characterization of Hox genes over the years have revealed the presence of several subfamilies, including LIM, POU, PAX, PRD, SIX, TALE, CUT, PROS, ZF HOX, NK and many others (Bürglin and Affolter 2016). These subfamily members are involved in the development of a wide variety of neurons (Hobert and Westphal 2000; Larroux et al. 2007; Mallo and Alonso 2013; Srivastava et al. 2010; Wollesen et al. 2014). This work will focuses on one of the Hox subfamily, namely *Lhx*

that encodes LIM-Homeodomain (LIM-HD) proteins (Srivastava et al. 2010).

1.3 The structural and classification of Lhx proteins

The acronym LIM is derived from the first letter of the three founding LIM-HD family members *lin-11* (Freyd et al. 1990), *Isl1* (Karlsson et al. 1990) and *mec-3* (Way and Chalfie 1988). All Lhx transcription factors have a common protein structure, composed of two *N*-terminal LIM domains and a *C*-terminal DNA-binding homeodomain (Gong and Hew 1994; Sánchez-García et al. 1993; Schmeichel and Beckerle 1997). The homeodomains (HD) of Lhx proteins can bind both to distinctive DNA sequences and mediate protein-protein interactions (Gong and Hew 1994; Sánchez-García et al. 1993). The cysteine-histidine-rich LIM domains consist of approximately 55 amino acids that form double zinc fingers (Figure 1.2) (Deane et al. 2004; Pérez-Alvarado et al. 1994; Schmeichel and Beckerle 1997). Structural studies of the LIM domains by NMR spectroscopy and X-ray crystallography demonstrated that each LIM domain is packed in antiparallel β-hairpins (Pérez-Alvarado et al. 1994; Schmeichel and Beckerle 1997). The LIM domains mediate protein-protein interaction between Lhx proteins and their cofactors that are required for transcriptional regulation (Bach 2000; Khurana et al. 2002). There are 30 known protein-protein interaction pairs for LIM domains have been described (Bach 2000; Khurana et al. 2002). Based on amino acid sequence similarities between the HD, the *Lhx* genes are further subdivided into six sub-groups (Figure 1.3).

1.4 Function of *Lhx* genes in the vertebrate nervous system

The vertebrate central nervous system is divided into two defined modules: the brain and the spinal cord. The brain is more complex, as it contains many neuronal cells of diverse cell types, which form more complex neural circuitry compared to the spinal cord. The *Lhx* genes play an important role in the differentiation, axonal pathfinding, the formation of critical signaling centers, regulation of early progenitor cell specification,

and proliferation of developing neurons in the brain. Although the spinal cord may be less complex in comparison to the brain, it too contains many different types of neurons. During development, these neurons are generated and then precisely connected, within the spinal cord and potentially to neurons extending out into the periphery, to form circuitry that is essential for normal sensory and motor functions. This process also involves intricate transcriptional regulation by multiple Lhx TFs.

1.4.1   The brain

There are twelve known *Lhx* genes in vertebrates, which are important in the development of multiple brain regions. Many of the *Lhx* genes have been found to be essential in forebrain development in the mice (summarized in Table 1.1). For instance, *Lmx1b* has been reported in the dopaminergic neurons in the developing adult mesencephalon (Smidt et al. 2000). In *Lmx1b* knockout mice, the dopaminergic neurons are generated, but fail to express the mesencephalon specific paired-like-HD gene, *Ptx3*, and thus, do not differentiate during later stages in development (Smidt et al. 2000). Furthermore, *Lmx1b* is also required for the serotonergic neurons in the developing hindbrain (Cheng et al. 2003; Ding et al. 2003). In the *Lmx1b* knockout mutants, serotonergic neuron precursor cells fail to migrate to their final locations (Cheng et al. 2003; Ding et al. 2003).

The expression of the *Lhx5* gene is observed in the neural progenitor cells lining the medial wall of the telencephalon, the site of the developing hippocampus (Zhao et al. 1999). In *Lhx5* knockout mice, the neural progenitor cells in the developing hippocampus fail to exit the cell cycle on time during cell division (Zhao et al. 1999). Interestingly, those cells that can exit the cell cycle do not differentiate or migrate to the appropriate layers in the developing hippocampus (Zhao et al. 1999). Similarly, *Lhx6* and *Lhx8* are expressed in the ventral telencephalon (Matsumoto et al. 1996; Grigoriou et al. 1998). In *Lhx8* knockout mice, the cholinergic neuron progenitors still form in the forebrain, but their terminal differentiation is defective (Zhao et al. 1999; Mori et al. 2004). Knock-down of *Lhx6* expression using RNA interference (RNAi) revealed that

γ-aminobutyric acid (GABAergic) interneurons are generated, but their migration from the ventral telencephalon to the cortex is blocked (Alifragis et al. 2004).

1.4.2   Spinal cord

Similar to the brain, the function of *Lhx* genes was first reported in the developing spinal cord of vertebrates, more specifically chick embryos (Tsuchida et al. 1994). Combinations of the four *Lhx* genes (*Islet1*, *Islet2*, *Lhx1*, and *Lhx3*) were reported to play important roles in the location and axonal projection of various types of motor neurons in the spinal cord (Bonanomi and Pfaff 2010). Interestingly, motor neuron cell transplantation in the zebrafish spinal cord revealed that the gene expression profile, morphology, and axonal projection of these neuronal cells were altered and resembled those of neurons residing at the implant site (Appel et al. 1995). These findings suggested that the differentiation of motor neuron in the spinal cord of zebrafish is controlled by combinatorial expression of several LIM-HD proteins (Appel et al. 1995).

Similarly, in mice, *Islet1* is required for the differentiation of postmitotic precursor cells that give rise to all subclasses of motor neurons soon after their exit from the cell cycle (Ericson et al. 1992). In *Islet1* knockout mice, motor neurons failed to differentiate and undergo apoptosis (Pfaff et al. 1996). Similarly, in *Islet2* knockout mice, the visceral motor neurons failed to migrate and project their axons to appropriate targets (Tsuchida et al. 1994; Thaler et al. 2002). The two other *Lhx* genes, *Lhx3* and *Lhx4*, are important in all subclasses of motor neurons in the spinal cord that extend their axons ventrally from the neural tube when they exit the cell cycle (Zhadanov et al. 1995; Yamashita et al. 1997). In *Lhx3* and *Lhx4* knockout mice, motor neurons were generated, but the axons were extended dorsally instead of ventrally (Sharma et al. 1998). Together, these findings suggest that Lhx transcription factors play an important role in neuronal differentiation and axon migration during the early developmental processes. The functional analysis of *Lhx1* and *Lmx1b* further provided evidence to support the role of Lhx proteins in motor neuron axonal pathfinding. For instance, *Lhx1* is expressed in a subclass of motor neurons of the spinal cord that innervate the dorsal limb muscle (Tsuchida et al. 1994)

and *Lmx1b* is expressed in cells in the dorsal limb mesenchyme (Riddle et al. 1995; Vogel et al. 1995). In knockout mice for both *Lhx1* or *Lmx1b*, the motor neurons are specified, but their axonal projections into the limb are randomized (Kania et al. 2000). The functional studies of *Lhx* genes in the spinal cord of *Xenopus* have also led to similar conclusions (Segawa et al. 2001). For example, *Islet2* overexpression in zebrafish showed defects in cell position, neurotransmitter expression, axonal outgrowth and pathfinding of sensory and motor neurons (Segawa et al. 2001). Together, a common theme emerges amongst the *Lhx* genes in the nervous system of all organism, that their functions are required for generation, differentiation, and axonal projections of neurons, and all are expressed in complex and largely non-overlapping domains.

1.5 The function of the *Lhx* genes in the development of the invertebrate nervous system

Extensive functional studies in two leading invertebrates, the fruit fly *D. melanogaster* and the nematode *C. elegans*, have revealed important roles of Lhx transcription factors in the specification of the developing nervous system. The gain or loss of individual *Lhx* gene functions can result in a complete block of differentiation and subsequent loss of an entire group of neurons, or cause subtle changes in various aspects of neuronal development such as cell migration, process outgrowth, axonal projection, or neurotransmitter synthesis. Such roles are mediated in-part by dynamic regulation of *Lhx* gene expression. Many *Lhx* genes are expressed within regions of the developing nervous system, but some have additional function in non-neuronal tissues, such as muscle, reproductive system, and endocrine tissue. Continued expression of the *Lhx* genes in tissues suggests a role in the maintenance of the differentiated cell phenotypes.

The Lhx proteins are required for neuronal differentiation in both *C. elegans* and *D. melanogaster*. In both these organisms, it has been demonstrated that Lhx transcription factors play essential roles in the specification and differentiation of neurons as well as in the proper guidance of axons in the developing nervous system (Hobert et al. 1998;

Lilly et al. 1999a; Sarafi-Reinach et al. 2001). Mutations in the *Lhx* genes result in morphological and behavioural defects.

1.5.1 *Drosophila melanogaster*

In *D. melanogaster*, all six *Lhx* genes, are implicated in many developmental processes (summarized in Table 1.1) (Srivastava et al. 2010). One of these genes, *Apterous (Ap)*, is expressed in a subset of interneurons in the ventral cord neurons (VCN), where it is required for proper axonal pathfinding, fasciculation of neurons, formation of the wing blade and the notum, and type of neurotransmitters neurons produce(Benveniste et al. 1998; Lundgren et al. 1995; Herzig et al. 2001). Mutation in the *Ap* gene causes variations in wing morphology, ranging from mild defects in the hinge region to complete lack of wings (Stevens and Bryant 1985; Cohen et al. 1992). Furthermore, *Ap* is required for normal initiation of neuropeptide expression by the Tv neurons. However, its function is not required for the survival or morphological differentiation of the neuron themselves (Benveniste et al. 1998).

The *D. melanogaster*, *dLim3* and *Islet* genes function to specify the identity and axonal projection pattern of motoneurons and the dopaminergic and serotonergic interneurons in the VNC (O'Keefe et al. 1998; Das et al. 2011; Lilly et al. 1999a; Winchell and Jacobs 2013). Loss of *dLim3* or *Islet* function does not affect the generation or survival of these neurons, but does disrupt their axonal pathfinding and cell fate (Thor and Thomas 1997; Thor et al. 1999). Interestingly, some motoneurons express both *dLim3* and *Islet* and project their axons in intersegmental nerve b (ISNb) branch to muscles of the embryonic abdominal hemisegment (Thor et al. 1999). In *dLim3* knockout animals, the intersegmental nerve b (ISNb) neurons project axon processes incorrectly towards the ISNd neuron target sites. Similarly, when *dLim3* is misexpressed in ISNd neurons, they project axonal processes incorrectly to the ISNb target sites (Thor et al. 1999). These findings suggested that *dLim3* and *Islet* have a combinatorial role during axon migration in the VNC for certain motor neurons (Thor et al. 1999). Together, *Lhx* genes

8

play important roles in differentiations of cells, specially neuronal cells, and manipulation of an invertebrate system like *D. melanogaster* has allowed for better understanding significance of these transcription factors.

1.6 *C. elegans* as a model organism to understand the regulation and function of *Lhx* genes in nervous system development

In *C. elegans*, there are seven genes encoding Lhx proteins (summarized in Table 1.1), all of which are involved in the differentiation of specific subsets of neuronal cells (Srivastava et al. 2010). Four of the seven genes (*lim-7*, *lim-6*, *lim-4* and *lin-11*) have also been shown to play roles outside of the nervous system, in specification and differentiation of non-neuronal tissues (Hall et al. 1999; Hobert et al. 1999; Newman et al. 1999; Newman and Sternberg 1996).

1.6.1 Mechanosensory neurons

Touch and other mechanical senses are critically important for sensory perception in nematodes (Goodman 2006). In *C. elegans*, there are 46 predicted sensory neurons implicated in the detection of attractive or aversive touch stimuli (White et al. 1986). Specifically, 14 neurons have been implicated by cell ablation experiments in the detection of harsh body touch (Way and Chalfie 1989). Roughly, 26 neurons are involved in sensing mechanical stimuli around the nose, including bilaterally symmetric ASH neurons, polymodal nociceptors that detect aversive chemical, osmotic as well as mechanical stimuli (Kaplan and Horvitz 1993). In addition, many neurons, including certain classes of motor neurons and interneurons, are thought to be proprioceptive.

The *mec-3* gene was one of the first *Lhx* gene identified in *C. elegans*. It is expressed in ten mechanosensory neurons; four of these neurons are paired, the ALM (L/R), PLM (L/R), FLP (L/R), and PVD (L/R) neurons, and two unpaired neurons, the AVM and PVM. *mec-3* is also expressed in the interneuron BDU(L/R) and the sensory neuron

FLP neurons (L/R) (Way and Chalfie 1989). *mec-3* is key for terminal differentiation of these neurons to the mechanosensory fate; thus, animals with mutations in these genes are defective in mechanosensation (Way and Chalfie 1989). The role of *mec-3* in differentiation of mechanosensory neurons is further supported by morphological studies of PVD in *mec-3* mutant animals, where extensive branching failed to occur, suggesting the neuron was not fully differentiated (Tsalik et al. 2003).

1.6.2   Thermosensory neurons

Genes involved in *C. elegans* thermotaxis have been identified by mutant screens (Hedgecock and Russell 1975). When placed in a temperature gradient, wild-type animals move toward the temperature at which they have been cultivated (Hedgecock and Russell 1975). The neuronal thermotactic response pathway has been defined by laser-ablating individual neuronal cells and observing the animals' behaviour in temperature gradient environment (Mori et al. 1995). Three *Lhx* genes (*lin-11*, *ttx-3* and *ceh-14*) are required for the differentiation and maintenance of functional integrity of several neurons that are part of the thermosensory network. This network consists of the sensory neuron AFD and interneurons AIY and AIZ (Aoki and Mori 2015; Cassata et al. 2000; Hobert et al. 1998). The *ceh-14*, *ttx-3* and *lin-11* genes are expressed in the AFD, AIY, and AIZ neurons respectively; mutating each of these genes results in defective terminal differentiation of the respective neuron and causes thermosensory behavioral abnormalities (Aoki and Mori 2015; Cassata et al. 2000; Hobert et al. 1998). For example, *ttx-3* mutants are cryophilic; they move independently of their cultivation temperature toward lower temperatures and *lin-11* mutants are thermophilic (Hobert et al. 1997; Hobert et al. 1998).

1.6.3   Chemosensory neurons

Lhx genes have also been implicated in the differentiation of the chemosensory neurons in *C. elegans*. There are 11 pairs of chemosensory neurons in the adult worm that are

responsible for sensing odors in the environment (Bargmann and Cornelia 2006). The AWA and AWC neurons detect attractants, the AWB neuron detects repellent odours, and ASG neurons sense pheromones (Hart and Chao 2009; Peckol et al. 2001). In wild-type animals, *lim-4* is required for the differentiation of AWB neuron; however, in the *lim-4* mutant the AWB neurons adopt the AWC function and fate (Sagasti et al. 1999). Similarly, *lin-11* is required for differentiation of AWA, AWC, and ASG neurons (Sarafi-Reinach and Sengupta 2000). In *lin-11* mutants, the AWA neuron adopts an AWA/AWC hybrid morphology (Sarafi-Reinach and Sengupta 2000). Interestingly when *lin-11* is ectopically expressed in the ciliated neurons, some neurons adopt the ASG fate (Sarafi-Reinach et al. 2001). Thus, the function of Lhx is important for the differentiation of distinct chemosensory neurons in *C. elegans*.

1.6.4   Other neurons

Lhx genes are also important in differentiation of other neurons including GABAergic and pioneering neurons. The *C. elegans* nervous system contains 26 GABAergic neurons, and *lim-6* is involved in the differentiation of five of these: RME (L/R), AVL, RIS and DVB (Hobert et al. 1999; McIntire et al. 1993). RME (L/R), AVL and DVB are motor neurons, whereas RIS is an interneuron (White et al. 1986). *lim-6* is also expressed in three additional neurons, PVT and RIG (L/R), which express the neuropeptide FMR-Famide (Hobert et al. 1999; Schinkmann and Li 1992). In *lim-6* mutants, the PVT and RIG neurons fail to form proper axonal projections and lack expression of the enzyme required for the synthesis of the γ-aminobutyric acid (GABA) neurotransmitter (Hobert et al. 1999). Furthermore, neuronal asymmetry is also affected in *lim-6* mutants. The two sensory neurons required for taste, ASE (Left) and ASE (Right) are bilaterally symmetric, however, both have distinctive functions and gene expression profiles (Hobert et al. 1999). *lim-6* is required to repress expression of the ASER-specific guanylyl-cyclase gene, *gcy-5*, in ASEL (Pierce-Shimomura et al. 2001; Cheng et al. 2003). Furthermore, the function of *lim-6* is also required for the ASEL neuron to distinguish sodium from

chloride ions in the environment (Pierce-Shimomura et al. 2001). Together, the function of *Lhx* genes are required for differentiation and specification of different classes of non-overlapping neurons.

1.6.5   Regulation and function of *Lhx* genes in *C. elegans* nervous system development

The seven Lhx proteins in *C. elegans*, mentioned above, fall into six subclasses based on Hox sequence alignments (Figure 1.3): TTX-3 is in the apterous group, LIM-4 is in the LHX6/7 group, LIM-7 is in the ISLETS group, LIM-6 is in the LMX group, CEH-14 is in the LIM-3 group, and MEC-3 and LIN-11 are both in the LIN-11 group. All of the *C. elegans Lhx* genes appear to be expressed after the last cell division, and are thought to be involved in the specification of neurons into particular neuronal subtypes. For instance, loss-of-function mutations in *mec-3*, which is expressed post-mitotically in touch neurons, caused cells that would have become touch receptor neurons to become interneuron-like cells (Way and Chalfie 1988). Similarly, *ttx-3*, *lin-11*, *lim-4*, *lim-7* and *lim-6* are first expressed in neuronal lineages after the final cell division and their expression is maintained throughout adulthood; thus, these genes are required for the specification of neural cells, but not their genesis. All seven *C. elegans Lhx* genes are regulated by a complex network of transcription factors involved in neuronal differentiation process.

1.6.6   *mec-3*

In *C. elegans*, the best characterized transcriptional network for a *Lhx* gene is that elucidated for *mec-3* in mechanosensory neurons (Figure 1.4). As stated earlier, *mec-3* is expressed in ten mechanosensory neurons; four of which are paired ALM (L/R), PLM (L/R), FLP (L/R), and PVD (L/R) and two unpaired AVM and PVM, as well as in the interneuron BDU(L/R) and the FLP (L/R) sensory neuron (Duggan et al. 1998; Way and Chalfie 1989). As shown in Figure 1.4A, in the mechanosensory neurons, a bHLH transcription factor formed through the heterodimerization of LIN-32 and HLH-2, coupled with a muscle segment homeobox protein, VAB-15, activate expression of a

POU-HD protein, *unc-86*, which subsequently activates *mec-3* expression (Duggan et al. 1998; Way and Chalfie 1989). UNC-86 and MEC-3 form a heterodimer, which further activates *mec-3* expression and activates the touch receptor neuron-specific genes such as *mec-4* and *mec-7*. Expression of a muscle segment homeobox protein, *pag-3*, is also activated by UNC-86 and MEC-3, which further activates PAG-3. However, in order for the interneuron BDU to acquire the correct neuronal fate, *pag-3* is expressed at low levels and represses *mec-3* function, which prevents the BDUs from becoming touch receptor neurons (Figure 1.4B) (Aamodt et al. 2000; Jia et al. 1996). In the FLP neuron, the *mec-3* regulatory network is initiated, but the touch receptor-specific genes are repressed by *egl-46* (Figure 1.4C). It is important to note that there are many other genes that are predicted to be targets of *mec-3*, and yet to be tested (see Appendix A - Table 1). Nevertheless, the overall regulation of *mec-3* within a specific neuronal cell is very complex, and requires the co-ordination of numerous TFs.

1.6.7   *ttx-3*

The expression of *ttx-3* has been reported in one pair of interneuron AIY (L/R), and the regulatory pathway of *ttx-3* has been well studied in this neuron pair (Figure 1.5) (Altun-Gultekin et al. 2001). AIY (L/R) development and function is initiated by CEH-10 (paired-like class of homeodomain proteins), which activates the expression of *ttx-3*. TTX-3 in turn activates the expression of other target genes that are required for the function of these neurons, including synthesis and packaging of acetylcholine, neuropeptide receptors, various types of ion channels and many others (Altun-Gultekin et al. 2001; Hobert et al. 1997; Wenick and Hobert 2004). *ttx-3* mutants are cryophilic, a phenotype that is caused by a failure to balance two opposing warm and cold temperature-processing pathways integrated via the AIY and AIZ interneurons, respectively (Figure 1.6). Since the AIY neurons are defective in *ttx-3* mutants, this results in an imbalance in AIZ interneuron activity leading to the displayed cryophilic behaviour. Moreover, it is not known if the same regulatory pathway is also required for the differentiation of the remaining four neurons (ASI, AIA, NSM, and ADL), or if additional proteins are

needed. There are many transcription factors that are predicted to regulate *ttx-3* in the nervous system, though their involvement in *ttx-3* regulation has yet to be tested (Appendix A - Table 1). Together, this suggests that tissue specific regulatory network of *ttx-3* is very complex.

1.6.8   *ceh-14*

ceh-14 expression in the head was observed in the sensory neurons AFD (L/R), in the interneurons BDU (L/R), and in the asymmetrical interneuron ALA, and in the tail it was observed in eight neurons. However, most studies have focused on deciphering the regulatory network of *ceh-14* in AFD differentiation (Cassata et al. 2000; Kagoshima and Kohara 2015). Mutations in *ceh-14* result in the failure of AFD fate specification, leading to abnormal morphology in the ciliary structure at the tip of AFD dendrites, and the inability of mutant animals to properly respond to temperature. *ceh-14* mutants show athermotactic (no response to temperature) behaviour (Figure 1.6) (Cassata et al. 2000; Hedgecock and Russell 1975). The AFD differentiation process is initiated by the transcription factor UNC-86 (POU-Homeodomain), which activates the expression of *ceh-14* (Figure 1.5) (Kagoshima and Kohara 2015). In the AFD neurons, CEH-14 activates many cell specific target genes that encode for receptor-type guanylyl cyclase, GCY-8 and GCY-18, glutamate transporter, *eat-4*, and immunoglobulin superfamily of transmembrane protein, ZIG-1 (Figure 1.5) (Kagoshima and Kohara 2015; Kuhara et al. 2011; Serrano-Saiz et al. 2013; Wang et al. 2013). Together, activation of the *ceh-14* regulatory pathway allows the AFD neuron to acquire proper neuronal fate. There are many other transcription factors that are likely to play roles in the differentiation of AFD neurons, however, it has yet to be discerned where they are precisely involved in the *ceh-14* regulatory pathway (Appendix A - Table 1).

1.6.9   *lim-6*

lim-6 expresses in a set of neurons, including the chemosensory, interneuorns and motor neurons, as well as other tissues including epithelial cells of the uterus and the excretory system (Figure 1.5) (Hobert et al. 1999; Pierce-Shimomura et al. 2001). Reporter gene expression in the nervous system has been observed to begin late in embryogenesis, after these neuronal cells have been generated; thus, the function of *lim-6*, like all *Lhx* genes, is not required for the genesis of these neurons (Sulston et al. 1983). Most of the neurons where *lim-6* is expressed are GABAergic neurons, namely RME (L/R), AVL, RIS and DVB (Hobert et al. 1999). Closer examination of these neuronal cells identified RME (L/R), AVL and DVB as motorneurons, RIS is an interneuron (White et al. 1986). And in nematodes GABA acts primarily at neuromuscular synapses to relax the body muscles during locomotion and foraging and to contract the enteric muscles during defecation (Jorgensen 2005).

At present, there are no known regulators of *lim-6* in the neurons, however, there are several target genes that LIM-6 regulates during RIS interneuron differentiation (Figure 1.5). LIM-6 activates five target genes in the RIS interneuron, *dop-1* (D1 like dopamine receptor), *ser-4* (5-HT1 metabotropic serotonin receptors), *glr-1* (AMPA-type ionotropic glutamate receptor), *zig-5* (immunoglobulin superfamily of proteins) and *unc-25* (GABA neurotransmitter biosynthetic enzyme) (Tsalik et al. 2003). Similar to the other *Lhx* genes discussed above, there are several predicted *lim-6* interacting genes that regulate its expression in specific tissues, including in the neuronal cells; however, these have yet to be investigated further (Appendix A - Table 1).

1.6.10   *lim-4*

The regulation of *lim-4* in neuronal differentiation is not well understood, however, several studies have reported the overall expression of *lim-4*. *lim-4* expression has been observed in the AWB, RME, RMEV, RMD, RMD (L/R), RID, RIV, SAA and SIA neurons (Kim et al. 2015; Sagasti et al. 1999; Zheng et al. 2005). LIM-4 promotes the

AWB cell fate while repressing the AWC fate (Figure 1.7) (Sagasti et al. 1999). In *lim-4* mutants, AWB adopts many molecular, morphological, and functional characteristics of AWC. Interestingly, ectopic expression of LIM-4 from the *odr-3*, AWC specific, promoter is sufficient to cause the AWC cells to take on several morphological and molecular aspects of the AWB fate (Sagasti et al. 1999). In another study, it has been shown that LIM-4 regulates the ADF serotonergic phenotype (Zheng et al. 2005). In *lim-4* mutants, ADF specification occurs, however, the neurons fail to express a serotonin secretory phenotype and exhibit defective cilia properties (Zheng et al. 2005) further demonstrated that *lim-4* is regulated by separable modules; more specifically, the *cis*-elements within introns are necessary and sufficient to direct *lim-4* to confer the serotonergic phenotype on ADF neurons, whereas its 5'-upstream sequence directs *lim-4* function in AWB neurons (Zheng et al. 2005). Several genes have been predicted to potentially interact with *lim-4*; however, future studies are required to test and identify their specific interactions (see Appendix A - Table 1).

1.6.11 *lim-7*

lim-7 is one of seven *Lhx* encoding genes and the sole *Islet* ortholog in *C. elegans* (Figure 3). *lim-7* deletion causes early larval lethality with terminal phenotypes including uncoordination, detached pharynx, constipation and morphological defects (Voutev et al. 2009). The *lim-7p::GFP* reporter shows that *lim-7* is expressed in the four distal pairs of sheath cells surrounding each hermaphrodite gonad arm and in the four URA motorneurons in the head (Hall et al. 1999; Voutev et al. 2009). However, the precise role of *lim-7* in these specific tissue is still unknown, and currently only two articles have been published that report expression of *lim-7* (Hall et al. 1999; Voutev et al. 2009). Furthermore, while a number of genes have been identified that are important for sheath specification or function, including *emb-30*, *lin-9*, *lin-26* and *unc-52*, how these genes interact with *lim-7* still is still not clear (Hall et al. 1999). In addition, there are various other genes that are predicted to interact with *lim-7* during development, but they have yet to be tested, and their interaction is currently unclear (Appendix A - Table 1).

1.6.12 *lin-11*

The expression of *lin-11* is reported in a subset of neurons in the head ganglia and the lumbar ganglion in the tail (Table 2) (Hobert et al. 1998; Sarafi-Reinach and Sengupta 2000). These neurons belong to many different classes: thermosensory, chemosensory, interneurons and pioneer neurons, which together integrate and output a specific behaviour. The following paragraphs will describe the known roles of *lin-11* in each of these classes of neurons.

In the thermosensory system, *lin-11* is required for differentiation of the AIZ interneuron, which is part of the thermoregulatory network and responds to high temperatures (Figure 1.6). Thus, *lin-11* mutants are known to display thermophilic behaviours (Hobert et al. 1998). *lin-11* expression is maintained throughout adulthood, suggesting it has a role in the maintenance of cellular function of the AIZ neuron in the thermoregulatory circuit (Hobert et al. 1998). However, the regulatory mechanisms controlling *lin-11* in the thermosensory neuron, AIZ, is not known.

lin-11 also plays a role in the specification of chemosensory neurons, which are collectively required to detect volatile and water-soluble attractant or repellent chemicals. As shown in Figure 1.8, UNC-130 (Forkhead domain) acts in parallel with ALR-1 (*Aristaless/Arx*) and regulate *lin-11* expression in the precursor cell of the AWA and ASG neurons (Nash et al. 2000; Sarafi-Reinach and Sengupta 2000). In the AWA olfactory neurons, transient expression of *lin-11* during late embryonic/early larval stages is necessary to initiate *odr-7* expression, whereas in the ASG neurons, *lin-11* expression is maintained through all postembryonic stages (Figure 1.8) (Sagasti et al. 1999; Sengupta et al. 1994). This temporal regulation of *lin-11* expression in these neurons is critical for correct fate specification, as continued expression of *lin-11* in the AWA neurons causes defects (Sarafi-Reinach and Sengupta 2000). Together, these findings suggest that strict temporal control of *lin-11* expression is critical for correct AWA and ASG fate specifications. Other chemosensory neurons such as ADF, ADL and ASH also require the function of *lin-11* to maintain proper cellular function, but the regulation of *lin-11* in

these specific neurons is not known.

Pioneer neurons are the first neurons to extend axons in the developing embryo. In the ventral cord of *C. elegans*, the major axon tract on the right side is pioneered by the AVG axon from the anterior end, whereas the left axon tract is pioneered by the PVP axon from the posterior end. The left axon tract consists of only 4 axons, whereas the right axon tract contains about 50 axons in adult animals (Hutter 2003). *lin-11* is required for specification of both the AVG and PVP pioneer neurons (Hutter 2003). Specifically, as illustrated in Figure 1.9, LIN-11 inhibits the function of *unc-86* in the AVG/RIR precursor cells, which allows for AVG to acquire proper fate, it is not known if other proteins also facilitate this process (Baumeister et al. 1996; Hobert et al. 1998). LIN-11 regulation in the PVP pioneer neuron is not known. Furthermore, laser ablation studies revealed that the removal of the AVG neuron from a wildtype embryo, resulted in the right ventral cord axon tract becoming partially disorganized, especially the motor neurons such as the DA and DB neurons (Durbin 1987). Thus, a similar axonal routing defect are also observed in *lin-11* mutants, however, there are no known consequences that result from axon disorganizations in the right and left tracts in *lin-11* mutants (Durbin 1987).

lin-11 has also been reported to be expressed in the RIC, AVJ, AVA, and AVE interneurons and required for their differentiation (Sarafi-Reinach and Sengupta 2000). The functions of the RIC and AVJ interneurons in *C. elegans* are not clear, but AVA is known as a command interneuron and it synapses with motor neurons that are responsible for forward and backward locomotion, allowing for rapid withdrawal from a noxious stimulus (Piggott et al. 2011). The AVE interneurons have similar overlapping function as AVA by controlling backwards locomotion, however, it is not known if their function is as significant as the AVA (Piggott et al. 2011). The important role of LIN-11 in different types of neurons suggests the involvement of multiple upstream regulators and interacting factors. However, the genetic and transcriptional network of *lin-11* regulation in these neurons is not well understood.

1.7   *Caenorhabditis briggsae* as model for comparative studies

C. briggsae and *C. elegans* diverged from a common ancestor an estimated ~30 million years ago, however, *C. briggsae* shares nearly identical developmental and behavioural traits with *C. elegans*, and both are commonly used for comparative evolutionary studies (Cutter 2008; Stein et al. 2003). As summarized below, there are many cellular and developmental processes conserved between these two sister species, however studies have also reported functional divergences in various other processes as well. For example gene expressed during excretory duct cell development, sex determination, male tail development and development of the vulva (Gupta et al. 2007). Vulval morphogenesis in particular is a well established model for comparative developmental studies between *C. elegans* and *C. briggsae.* Though vulva formation is morphologically identical between these two species, there are key differences in the underlying regulation of the genes controlling this process (Félix 2007; Sharanya et al. 2012). In addition, closely related species of *Caenorhabditis* also offer advantages for the study of the *Lhx* genes in the nervous system. In this regard, sequence comparison between *C. elegans* and *C. briggsae* has revealed that roughly one-third of the genes lack one-to-one orthologous relationships, making *C. briggsae* a good model for identifying conserved as well as divergent mechanisms that underlie the specification of neuronal processes (Gupta et al. 2007; Gupta et al. 2003). Thus, comparative studies of these two nematode species offer advantages for evolutionary similarities and differences in signaling and regulation of genes in developmental processes.

1.8   Goals of this thesis

The work described here provides for the first time a detailed analysis of *lin-11* regulation and function in the nervous system of two related *Caenorhabditis* species: *C. elegans* and *C. briggsae.* I have performed a structure and function analysis of the *lin-11* gene to decipher its role in organ formation and neuronal specification. I have utilized

a series of molecular and genetic assays to determine the precise role of *lin-11* in these developmental processes. I formulated four goals that I used as a tool to guide my experiments: First, to understand the role of the intron 3 and intron 7 regions in regulating neuronal *lin-11* expression, which is discussed in Chapter 3. Second, to identify the putative TF binding site(s) contained within the *lin-11* intronic sequences that mediate *lin-11* function in the amphid neurons (also discussed in Chapter 3). Third, to perform a functional analysis of the conserved domains of LIN-11 and finally rescue the *lin-11* mutation with other *Lhx* orthologs that are discussed in Chapter 4. Towards these four goals, I set several specific aims: first, I dissected the intronic regions of *lin-11* to identify the *cis*-regulatory elements for neuronal expression. Next, I used these *cis*-elements to identify the potential putative TF binding sites that may mediate *lin-11* function in the nervous system. I combined this information with RNA interference (RNAi) and genetic approaches to test several putative *lin-11* regulators, and I then quantified their impact on *lin-11* expression patterns. Furthermore, I performed functional analysis of the conserved domains of LIN-11, including LIM1, LIM2, and the proline-rich regions, and identified their roles in neuronal and reproductive tissue specification. Finally, I tested the functional conservation between *lin-11* and its orthologs in the fruit fly *D. melanogaster* (*dLim1*) and mouse *Mus musculus* (*Lhx1*) by performing transgene rescue of the reproductive system and neuronal defects in the *C. elegans lin-11* null mutants using *dLim1* and *Lhx1* genes.

1.9   The major findings of this thesis

Previously, *lin-11* mutants were shown to have defects in many behavioral processes including thermosensation, electrosensation and chemosensation (Hobert et al. 1998; Salam et al. 2013; Sarafi-Reinach and Sengupta 2000). Consistent with this, *lin-11* is expressed in a subset of amphid sensory neurons, however, how *lin-11* expression is regulated in these neurons is not well understood. To this end, I have studied the regulation of *lin-11* and report identification of new enhancer elements that reside within

the two introns, 3 and 7, that are necessary for neuronal cell fate specifications. I demonstrate that these intronic enhancers direct reporter gene expression in amphid sensory neurons. Using a transgenic approach, I examined functional conservation of the introns between *C. elegans* and *C. briggsae*. The cross-species analysis revealed that while intron 7 enhancer-driven expression is conserved, intron 3-driven expression shows significant divergence. Rescue experiments in *C. elegans* showed that intron 3 is necessary and sufficient for *lin-11* function in neurons using both GFP reporters as well as behavioral assays. Sequence comparison with three other *Caenorhabditis* species revealed that both introns possess evolutionarily conserved regions that may function to activate *lin-11* expression in neurons. Dissection of these regions identified elements in intron 3 that are specific for RIC (L/R) and in intron 7 for AVG and RIF (L/R). Using bioinformatics, I identified several TF binding sites located within conserved regions that may regulate *lin-11* expression in these subset of neurons. Further experiments confirmed the involvement of four of these TFs, *skn-1* (*bZIP* family), *crh-1* (*CREB* family), *ceh-6* (POU Homeodomain) and *ces-1* (C2H2-zinc finger), in *lin-11*-mediated neuronal specification process.

Furthermore, I have characterized the function of the conserved domains in LIN-11 in the nervous system and reproductive system. My work shows that the LIM domains are required for the specification of the neuronal and reproductive tissues, since mutations affecting these domains result in defective phenotype (see Chapter 3 and 4). However, the proline-rich region is not necessary for LIN-11 function in the tissues examined since truncating this region had no impact on the *lin-11* phenotype. However, it is plausible that this region cooperates with LIM domains to mediate protein-protein interactions. More experiments will be required to test this possibility. Finally, I demonstrate the degree of conservation of function between *Lhx* genes, specifically the *D. melanogaster*, (*dLim1*), and *M. musculus*, (*Lhx1*), homologs of *lin-11*. Sequence alignments of *lin-11*, *dLim1* and *Lhx1* show strong conservation within the LIM domains, HD domain and proline-rich region. As expected, my work showed that *D.melanogaster* and mouse homologs could to replace LIN-11 function in *lin-11* null mutants.

In conclusion, the work described in this thesis significantly advances our understanding of the mechanism of *lin-11* regulation, function, and evolution in the neuronal differentiation process.
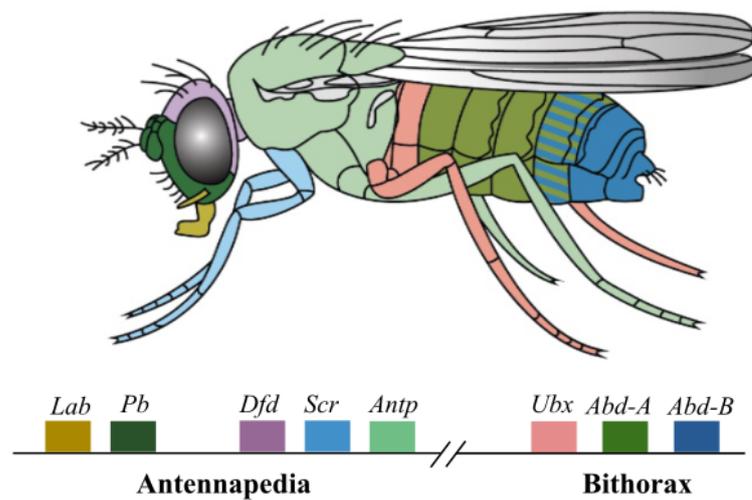
FIGURE 1.1: The patterns of expression compared to the chromosomal locations of the genes of the Hox complex. The sequence of genes in each of the two subdivisions of chromosomal complex corresponds to the spatial sequence complex corresponds to the spatial sequence in which the genes are expressed. *lab: labial, pb: proboscipedia, Dfd: deformed, Scr: sex combs reduced, Antp: antennapedia, Ubx: ultrabithorax, abdA: abdominal A, abdB: abdominal B.* Adapted from Abate-Shen (2002) and modified from Hox genes of fruit fly, public domain.
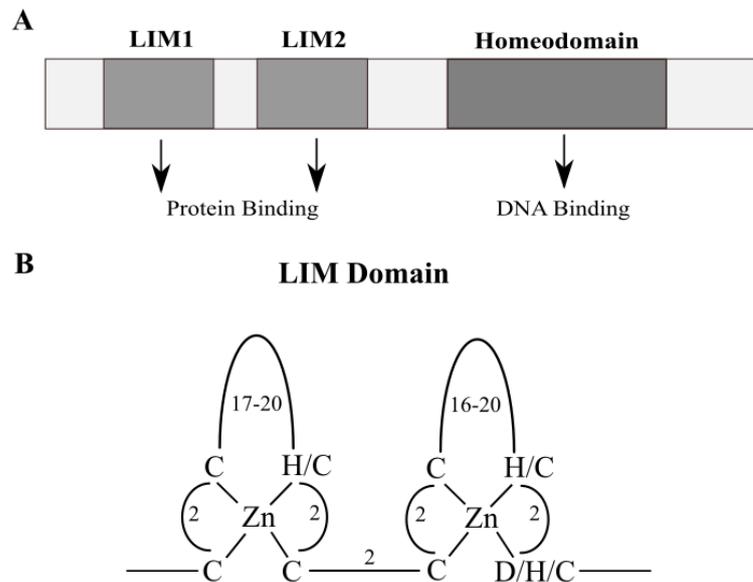
FIGURE 1.2: (A) Schematic illustration showing major functional motifs of a LIM-HD protein. (B) The LIM domain is a zinc-binding, cysteine-rich motif consisting of two tandemly repeated zinc fingers. The classic LIM consensus sequence includes a $C_{X2}C_{X17-20}H/C_{X2}C_{X2}C_{X2}C_{X16-20}H/C_{X2}C/H/D$. The numbers represent the residues that are located between the conserved cysteine and histidine of the LIM domains. D/H/C abbreviation for Aspartate/Histidine/Cysteine residues respectively. Adapted from Dawid et al. (1998).
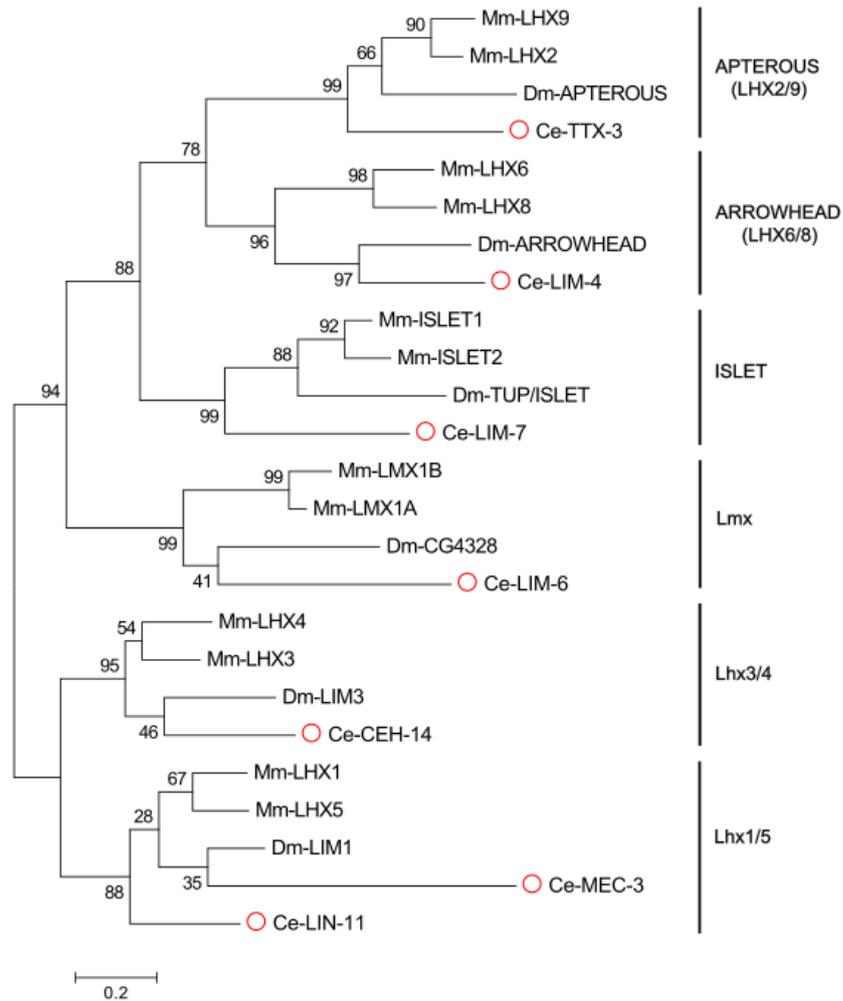
FIGURE 1.3: Phylogeny of *Lhx* genes. *Lhx* genes are divided into six sub-groups. *C. elegans* have seven known *Lhx* genes (red circles), one representing gene in each of these six sub-groups, and two in Lhx1/5. *D. melanogaster* have six known *Lhx* genes, one each sub-class; and mice have 12 *Lhx* genes and at least one in each sub-class. This phylogenetic tree was assemble using protein sequences from the NCBI and using MEGA7 software (neighbor-joining). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. *Ce = C. elegans*, *Dm = D. melanogaster*, and *Mm = M. musculus*.
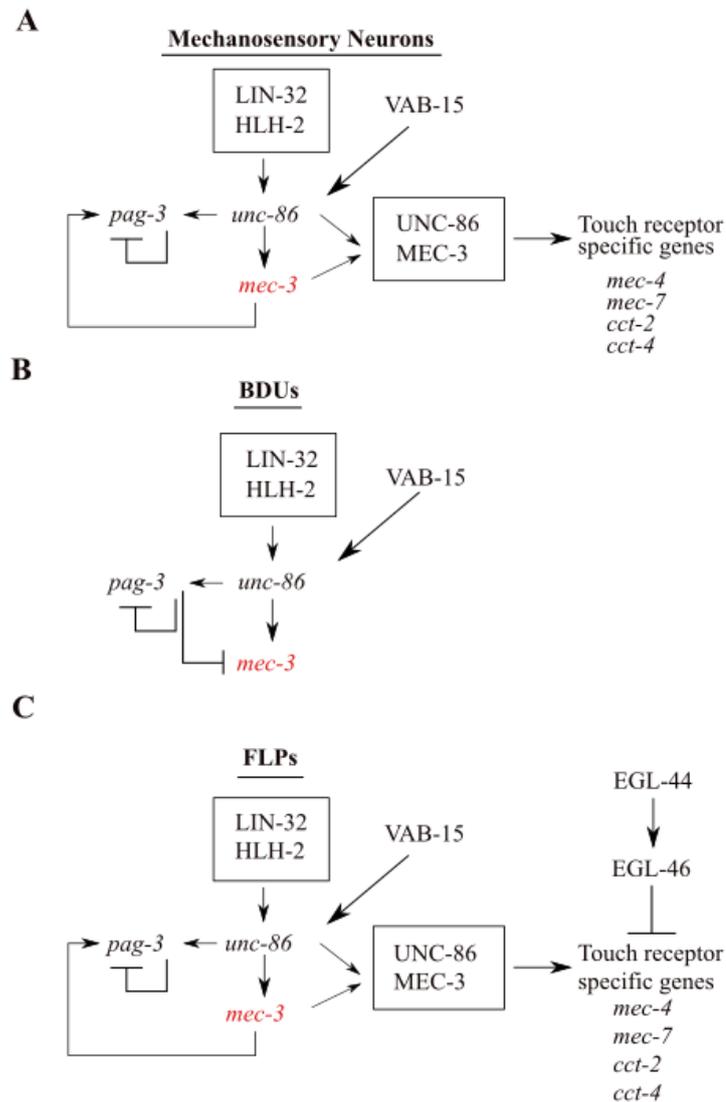
FIGURE 1.4: The genetic regulation of *mec-3* in neuronal development. (A) The *unc-86* gene is expressed in touch-cell precursors (Finney and Ruvkun 1990) and is necessary for the generation of the presumptive touch cells. UNC-86 forms heterodimers with MEC-3 and the heterodimers activate *mec-3* expression (Way and Chalfie 1989; Xue et al. 1993). UNC-86::MEC-3 hetero-oligomers directly activate four downstream genes, *mec-4*, *mec-7*, *cct-2* and *cct-4*, that are necessary for touch-cell function. (B) In interneuron BDU, the core regulation of *mec-3* pathway is similar to mechanosensory neurons, however, the target genes are not known. (C) In the sensory neuron FLP, the EGL-44 activates EGL-46 that inhibits the mechanosensory specific genes to prevent the FLP acquiring the mechanosensory like cell fate.

AIY(L/R) Interneuron	AFD (L/R) Sensory Neuron	RIS Interneuron
ceh-10 ↓ *ttx-3* ↓ *kal-1* *ser-2* *unc-17* *ceh-23*	*unc-86* ↓ *ceh-14* ↓ *gcy-8* *gcy-18* *eat-4* *zig-1*	*lim-6* ↓ *dop-1* *ser-4* *glr-1* *zig-5* *unc-25*
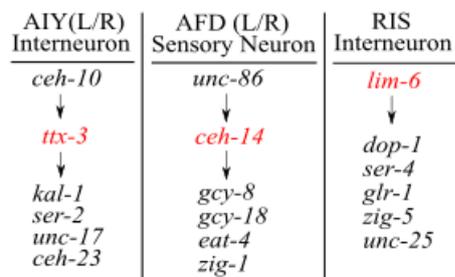
FIGURE 1.5: Genetic regulation of *Lhx* genes (*ttx-3*, *ceh-14* and *lim-6*) in neuronal development. (A) In the interneuron AIY, CEH-10 activates TTX-3 and constitutes a regulatory cascade of transcription factors that controls all subtype specific features of the AIY interneurons, by activating various target genes. (B) In AFD fate specification, the transcription factor UNC-86 activates CEH-14, and activate several specific target genes that define its fate. (C) Transcription factor LIM-6 activates various target genes, that specify the interneuron, RIS, cell fate. The regulators of *lim-6* is not known.
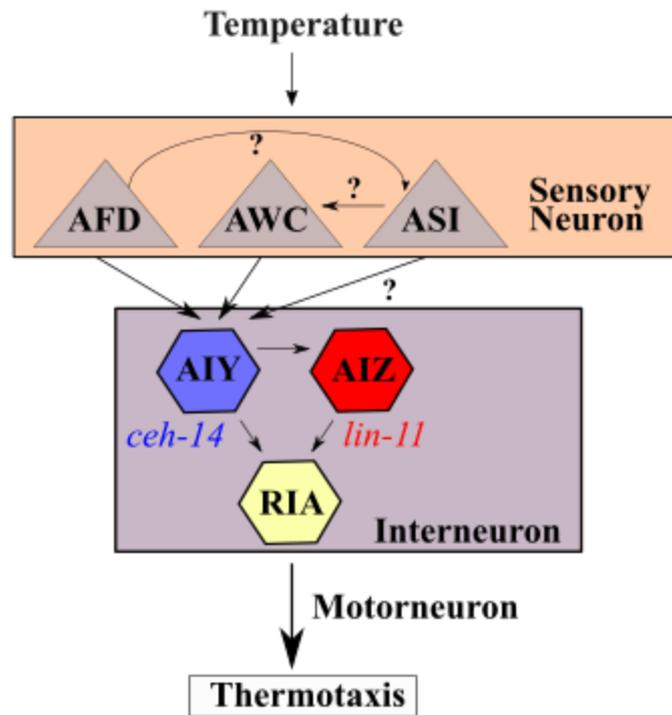
FIGURE 1.6: The neuronal thermoregulatory pathway in *C. elegans*. This pathway consists of two temperature modules the low (blue) and high (red). *lin-11* is expressed in the AIZ interneuron, which is part of the thermoregulatory network and responds to high temperatures; thus, *lin-11* mutants display thermophilic behaviours. *lin-11* expression is maintained throughout adulthood, suggesting that *lin-11* plays a role in the functional maintenance of this thermoregulatory circuit. The low temperature circuit consists of the sensory neuron AFD and the interneruon AIY. The TTX-1 and CEH-14 is required for the differentiation of the sensory neuron AFD and TTX-3 is required for differentiation of the interneuron AIY adapted from Aoki and Mori (2015).
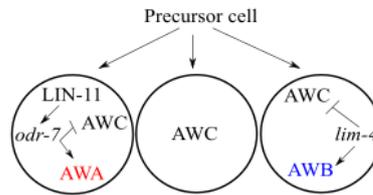
FIGURE 1.7: Olfactory cell fate specification by *Lhx* genes, *lin-11* and *lim-4*, in *C. elegans.* LIN-11 activates the *odr-7* expression, which inhibits the AWC fate and the cell acquires the AWA fate. In the absence of ODR-7 and LIM-4, these cells will develop like AWC, a potential olfactory default state. The AWB fate adopted by cells that express *lim-4*. Adapted from Sagasti et al. (1999)
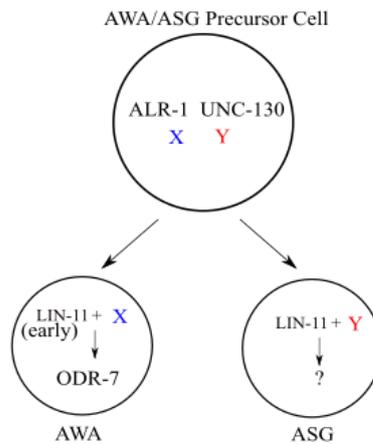
FIGURE 1.8: Genetic regulation of *lin-11* in AWA/ASG precursor cell. The transcription factors, ALR-1 and UNC-130 together function in the AWA/ASG precursors to asymmetrically allocate the unidentified determinants X and Y to the AWA and ASG neurons respectively. LIN-11 is expressed early in AWA, where it may act with X to initiate ODR-7 expression. ODR-7 then autoregulates as LIN-11 expression decreases. LIN-11 is expressed throughout adulthood in ASG, where it may act with Y to establish an ASG fate. Adapted from Sarafi-Reinach and Sengupta (2000)
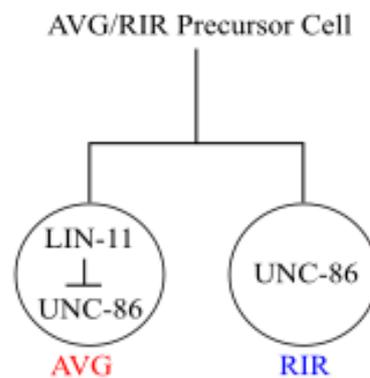
FIGURE 1.9: LIN-11 controls the asymmetry of UNC-86 expression in the AVG/RIR precursor cell lineages that generate RIR and AVG neurons. LIN-11 inhibits the function of transcription factor UNC-86 in the AVG neuron and prevents the cell acquiring the RIR fate. However, in the interneuron RIR, the function of UNC-86 allows the cell to acquire correct cell fate (Baumeister et al. 1996)
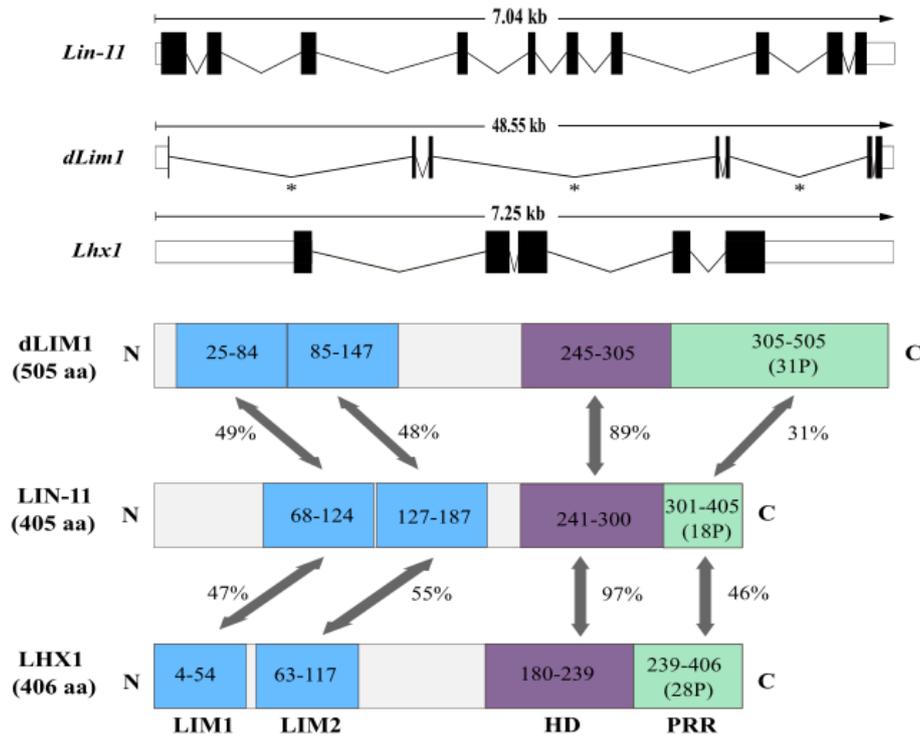
FIGURE 1.10: Comparison of *lin-11* gene and protein structure of *C. elegans* with its orthologs, *dLim1* and *Lhx1*. *D. melanogaster dLim1* is ~47 kb in size and located in chromosome X, and *M. musculus Lhx1* is ~4.6 kb in length and located in chromosome 11. Key structural differences include that in both *D. melanogaster*, *dLim1*, and mice, *Lhx1*, genes have fewer exons and introns compared to *C. elegans lin-11*. Furthermore, *D. melanogaster*, *dLim1*, have three large introns (~16kb, ~19kb, and ~9 kb respectively) marked with asterisks (*) and a 20% larger protein (505 aa). The number of proline (P) that are found in the proline-rich region (PRR) are indicated for each protein. Scale bar indicates 1kb.

TABLE 1.1: List of *Lhx* genes in different species.

Species	Gene	Sub-family	Function	Reference
C. elegans	*mec-3*	Lhx3/4	Required for differentiation of mechanosensory neurons	Duggan et al. 1998; Way and Chalfie 1989
	lin-11	Lhx1/5	Vulva development, neuronal development and fate specification, utse cell differentiation and fertility	Freyd et al. 1990; Hobert et al. 1998; Newman et al. 1999; Sarafi-Reinach et al. 2001; Gupta et al. 2003
	lim-4	Lhx1/8	Required for differentiation of chemosensory and motor neurons, changes in movement and foraging behaviour	(Kim et al. 2015; Sagasti et al. 1999; Zheng et al. 2005)
	lim-6	Lmx	Affects differentiation of GABAergic neurons, uterine morphogenesis	(Hobert et al. 1999; Pierce-Shimomura et al. 2001)
	lim-7	Islet	Larval lethality, by locomotion and morphological defects, a detached pharynx, body cavity vacuolization, and constipation	(Voutev et al. 2009)
	ttx-3	Lhx2/9	Required for the differentiation of the thermoregulatory interneurons, AIY(L/R), AIA, ADL, ASI, NSM	(Hobert et al. 1997; Altun-Gultekin et al. 2001; Wenick and Hobert 2004)
	ceh-14	Lhx3/4	Required for function of the AFD, AIZ, and AIY neurons, cryophilic thermotaxis behaviour	(Cassata et al. 2000; Aurelio et al. 2003)
D. melanogaster	*Apterous*	Lhx2/9	Wing, haltere, muscle and leg development, axon pathfinding and neurotransmitter identity, juvenile hormone synthesis, sexual functioning	(Ringo et al. 1991; Cohen et al. 1992; Lundgren et al. 1995; Benveniste et al. 1998; Ghazi et al. 2000; Pueyo et al. 2000)
	Arrowhead	Lhx6/8	Establishment of abdominal histoblasts and salivary gland imaginal rings	(Curtiss and Heilig 1995; Curtiss and Heilig 1997)
	dLim1	Lhx1/5	Leg and antenna development, neuronal subclass specification in the ventral nerve cord	(Lilly et al. 1999a; Tsuji et al. 2000)
	dLim3	Lhx3/4	Motorneuron axon guidance and regulation of cell cycle	(Thor et al. 1999)
	Islet/Tailup	Islet	Required for motorneuron pathfinding and neurotransmitter identity	(Thor and Thomas 1997)
	CG4328	Lmx	Nothing is known about this gene	Predicted by bioinformatics (Flybase)
Mammals	*Lhx1*	Lhx1/5	Head induction, axis formation, intermediate mesoderm differentiation	(Shawlot and Behringer 1995; Tsang et al. 2000; Kinder et al. 2001)
	Lhx2	Lhx2/9	Early cortical development, eye, erythrocyte development, embryonic lethal	(Porter et al. 1997; Xue et al. 1993; Bulchand et al. 2001; Monuki et al. 2001)
	Lhx3	Lhx3/4	Anterior pituitary cell types missing, required for projections of specific motorneurons, prenatal lethal, pituitary gland organogenesis	(Sheng et al. 1996; Sharma et al. 1998; Thor et al. 1999)
	Lhx4	Lhx3/4	Hypoplastic anteior pituitary, required for projection of specific motoneruons, perinatal lethal	(Li et al. 1994; Sheng et al. 1997; Sharma et al. 1998; Raetzman et al. 2002)
	Lhx5	Lhx1/5	Morphogenesis of huppocampus, postnatal lethal	(Zhao et al. 1999)
	Lhx6	Lhx6/8	Expresses in first branchial arch and basal forbrain, striatal interneuron specification	(Grigoriou et al. 1998)
	Lhx8	Lhx6/8	Expresses in first branchial arch and basal forbrain, striatal interneuron specification	(Grigoriou et al. 1998)
	Lhx9	Lhx2/9	Expressed in pioneer neurons in the cortex, limbs, gonad formation	(Rétaux et al. 1999; Bertuzzi et al. 1999; Birk et al. 2000)
	Islet1	Islet	Required for certain motor and interneurons, pancreatic mesenchyme and islet cells, and pituitary cell differentiation	(Pfaff et al. 1996; Ahlgren et al. 1997; Takuma et al. 1998)
	Islet2	Islet	Expressed in brain and spinal cord, motor neurons	(Tsuchida et al. 1994; Kania et al. 2000; Shirasaki and Pfaff 2002)
	Lmx1a	Lmx	Expressed in pancreas, brain, and spinal cord, neuronal migration in the cortex	(Millonig et al. 2000; Costa et al. 2001; Riddle et al. 1995; Vogel et al. 1995)
	Lmx1b	Lmx	Development of the skull, limb, kidney, anterior segment of the eye, mesencephalic dopaminergic neurons, podocyte differentiation, motorneruon axon trajectory, dorasal limb	(Dreyer et al. 1998; Rohr et al. 2002; Miner et al. 2002)

32

TABLE 1.2: List of *lin-11* specific neurons and their functions (Sarafi-Reinach et al. 2001).

Location	Name	Type	Function
Head ganglia	AVA	Interneuron	Backward locomotion
	AVE	Interneuron	Backward locomotion
	RIC	Interneuron	Octopamine release in the absence of food
	AIZ	Interneuron	Thermosensory
	ADF	Sensory neuron	Control entry into dauer stage, chemotactic response to cAMP, biotin, Cl- and Na+
	ASH	Sensory neuron	Response to nose touch, hyperosmolarity and volatile repellent chemicals
	ADL	Sensory neuron	Avoidance behaviour from Cd2+, octonol and Cu2+
	AWA	Sensory neuron	Chemotaxis to diacetyl, pyrazine, trimethylthiazole
	ASG	Sensory neuron	Chemotaxis to lysine, cAMP, biotin, Cl- and NA+
	AVG	Pioneer neuron	Pioneer the right tract of the ventral nerve cord via UNC-6/Nitrin
Lumbar ganglia	PVP	Pioneer neuron	Pioneers the left tract of the ventral cord
	PVQ	Pioneer neuron	Pioneering of the left tract of the ventral cord
	DVA	Interneuron	Mechanosensory integration of anterior or posterior touch circuit
	DVC	Interneuron	Mechanosensory integration of anterior or posterior touch circuit

Chapter 2.0

Materials and Methods

2.1 Strains and general methods

All strains were maintained at 20 °C, unless otherwise noted, using standard culture methods (Brenner, 1974; Wood, 1988). The following strains were used in this study: wildtype strains: N2 (*C. elegans*), AF16 (*C. briggsae*), mutants: PS2821(*lin-11(n389)*) (Ferguson and Horvitz 1985) and *unc-119(ed4)* (Maduro and Pilgrim 1995). The DNA microinjection procedure was followed as described previously (Mello et al. 1991). To eliminate an array-specific bias in GFP expression or phenotypic rescue studies, more than one transgenic line per construct was examined and one line with typical pattern was quantified and examined in detail.

All transgenic extrachromosomal arrays specific to Chapter 3 experiments have been described in the Chapter 3 methods sections. However, here are additional strains that were generated for Chapter 4 specific experiments, and all of the plasmids were microinjected into *lin-11(n389)* background, unless stated otherwise:

bhEx150 [pVH10.17(odr-2::CFP)], lin-11(ps1); bhEx157 [pVH10.17(odr-2::CFP)], lin-11(ps1); bhEx158 [pVH10.17(odr-2::CFP)], unc-119(tm4063); bhEx170[pVH10.17(odr-2::CFP) + pGLC93(lin-11::YFP) + unc-119(+)], bhEx159[pGF50 (FL-lin-11gDNA)+ pV10.17(odr-2::CFP)], bhEx160[pGF50 (FL-lin-11gDNA)+ pV10.17(odr-2::CFP)], bhEx237 (pGLC89 [(lin-11::GFP) + pGLC90

(lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx238 (pGLC89 [(lin-11::GFP) + pGLC90 (lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx239 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx240 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx253 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx254 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx255 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx257 (pGLC89 [(lin-11::GFP) + pGLC90 (lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx258 (pGLC89 [(lin-11::GFP) + pGLC90 (lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx261 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx262 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx263 (pGLC94 [(lin-11-5'UTRp::Lhx1cDNA) + pGLC97 (lin-11-int3p::Lhx1cDNA) + pGLC102 (odr-2::dsRED)], bhEx264 (pGLC96 [(lin-11-5'UTRp::dLim1cDNA) + pGLC98 (lin-11-int3p::dLim1cDNA) + pGLC102 (odr-2::dsRED)], bhEx265 (pGLC96 [(lin-11-5'UTRp::dLim1cDNA) + pGLC98 (lin-11-int3p::dLim1cDNA) + pGLC102 (odr-2::dsRED)], bhEx266 (pGLC96 [(lin-11-5'UTRp::dLim1cDNA) + pGLC98 (lin-11-int3p::dLim1cDNA) + pGLC102 (odr-2::dsRED)], bhEx267 (pGLC89 [(lin-11::GFP) + pGLC90 (lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx268 (pGLC88 [(lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx269 (pGLC88 [(lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx270 (pGLC88 [(lin-11::GFP) + pGLC102 (odr-2::dsRED)], bhEx271 (pGLC96 [(lin-11-5'UTRp::dLim1cDNA) + pGLC98 (lin-11-int3p::dLim1cDNA) + pGLC102 (odr-2::dsRED)], bhEx272[pGLC102(odr-2::dsRED)]

2.2   Molecular biology and transgenic

It has been described in Chapter 3 in the methods section in more detail. Briefly, the transgenic animals carrying extrachromosomal arrays were generated by standard micro-injection technique (Mello et al. 1991). *C. elegans unc-119* (Maduro and Pilgrim 1995) was used as a rescue marker in most cases. *pVH10.17 (odr-2b::CFP)* and *pVH13.05 (glr-1::dsRED)*, kindly provided by Hutter lab, were used as neuronal markers. The plasmid pGF50 is a 19.5 kb subclone of ID6 cosmid that contains the entire *lin-11* coding region and regulatory sequences (Freyd 1991). The concentrations of plasmids ranged from 20 ng/μL to 60 ng/μL. The total injection mix was less than 200 ng/μL. The following plasmids were made as part of this thesis.

Some of these have been described in the published study (see Chapter 3). These are: pGLC58 (*Cel-lin-11-int-7p::GFP*, Figure 2.1), pGLC59 (*Cel-lin-11-int-3p::GFP*, Figure 2.2), pGLC60 (*Cbr-lin-11-int-7p::GFP*, Figure 2.3), pGLC61(*Cbr-lin-11-int-3p::GFP*, Figure 2.4), pGLC65 (*Cel-lin-11-int-7p::GFP*, Figure 2.5), pGLC66 (*Cel-lin-11-int-3p::GFP*, Figure 2.6), pGLC67 (*Cel-lin-11-int-7p::GFP*, Figure 2.7), pGLC87 (*Cel-lin-11-int7p::lin-11cDNA*, Figure 2.8), pGLC92 (*Cel-lin-11-int3p::lin-11cDNA*, Figure 2.9), pGLC93 (*Cel-lin-11-int-7p::GFP*, Figure 2.10).

pGLC88 (LIN-11ΔPro). This plasmid was created in two steps. First, GL832 (5' - attgcatgcgtcttgcatcgctacggacca- 3') and GL833 (5'ggtctgcaggacgattcgcagtacgattatttg 3') was used to amplify 4.6 kb fragment that stretches from 5'UTR until middle of intron 2. This fragment was digested with SphI and PstI and cloned in the pPD95.69 vector. Next, GL834(5'gtcctgcagacccactcctggatatattgtag 3') and GL835 (5'ttatcccgggcatgaaa-cattcgggaacatatc 3') was used to amplify 4.7 kb fragment that starts from intron 2 of *lin-11* until the end of exon 8. This fragment was digested with PstI and XmaI and cloned into the first vector. See Figure 2.11 for a map.

pGLC89 (LIN-11ΔLIM; neuron specific). A 708 bps fragment was amplified using the primers, GL938 (5'- atactgcagatggaaggcaatcgattcgtg- 3') and GL939 (5'tatcccgggctac-catgaaaccggagttggtttt 3') and digested with PstI and XmaI. This fragment was cloned

into pGLC97 plasmid backbone. See Figure 2.12 for a map.

pGLC90 (LIN-11ΔLIM; vulva specific). A 708 bps fragment was amplified using the primers, GL938 (5'- atactgcagatggaaggcaatcgattcgtg- 3') and GL939 (5'tatcccgggctac-catgaaaccggagttggttttt 3') and digested with PstI and XmaI. This fragment was cloned into pGLC94 plasmid backbone that contain *lin-11* 5'UTR. See Figure 2.13 for a map.

pGLC91 (*lin-11*- exons 8-10). GL548 (5'attctgcaggtgtggtttcaaaaccgccgaag 3') and GL939 (5'tatcccgggctaccatgaaaccggagttggttttt 3') was used to amplify 360bps of *lin-11* cDNA (exon 8-10). This fragment was digested with XmaI and PstI and ligated into pGLC90 backbone. pGLC90 was digested with PstI and XmaI and 8.1 kb fragment obtained. See Figure 2.14 for a map.

pGLC94 (*lin-11-5'UTR::Lhx1cDNA*).This plasmid was generated in two steps. First, the *lin-11* 5'UTR (3.6 kb) was amplified using the primers, GL832 (5'attgcatgcgtcttg-catcgctacggacca 3') and GL849 (5'atactgcaggtgcactgtgcgggctgcaaaag 3') and digeste with SphI and PstI, then cloned into fire vector pPD95.69. Next, primers, GL850 (5'tatg-gatccctaccacacggctgcctcgttcattt 3') and GL851 (5'agaatgcatgtactacacaccagccagcacaatatcg 3') was used to clone mice *Lhx1* cDNA (1.2 kb) from PYX-Asc plasmid that was ac-quired from addgene. This fragment was digested with PstI and BamHI and ligated into the plasmid from step one. See Figure 2.15 for a map.

pGLC96 (*lin-11-5'UTR::cDNAdLim1*). pGLC94 was digested with PstI and BamHI to remove the mouse *Lhx1* cDNA. *dLim1* was cloned using, GL851 (5'agaatgcatgtac-tacacaccagccagcacaatatcg 3') and GL859 (5'tatggatccttaccaaacgagcccttcattttgc 3') and digested with with NsiI and BamHI. Both the fragment and backbone was ligated, in this plasmid the PstI and NsiI sites do not exist, because both fused together. See Figure 2.16 for a map.

pGLC97 (*lin-11-int3p::Lhx1cDNA*) This plasmid was generated first by digesting the pGLC94 with SphI and PstI to remove the 3.6 kb *lin-11* (5'UTR) and 1.4 kb and *lin-11* intron 3 was inserted instead. Intron 3 was amplified using GL11(5'aggcatgccaggctaagc-ctcttctc 3') and GL921 (5'tatctgcagacacacgaatcgattgccttccatgatata 3') and digested with

SphI and PstI, and ligated into the backbone that contains *lin-11* intron 3. See Figure 2.17 for a map.

pGLC98 (*lin-11-int3p::dLim1cDNA*). First, pGLC97 was digested with PstI and BamHI and 5.9kb backbone was acquired. GL851 (5' agaatgcatgtactacacaccagccagca-caatatcg 3') and GL852 (5'tatggatccttaccaaacgagcccttcattttgc 3') was used to amplify 1.5 kb *dLim1* cDNA, and this fragment was digested with NsiI and BamHI, this fragment was ligated with the backbone. Since NsiI and PstI sites were fused, this site is destroyed. See Figure 2.18 for a map.

pGLC102 (*odr-2::dsRED*). pVH13.05 was digested with BamHI and SpeI to remove the dsRED and cloned into pVH10.07 vector. See Figure 2.19 for a map.

2.3   Microscopy

see Chapter 3 - Methods section.

2.4   Behavioral assay

see Chapter 3 - Methods section.

2.5   Bioinformatics

see Chapter 3 - Methods section.

2.6   RNAi

see Chapter 3 - Methods section

2.7   Phylogenetic tree

Molecular phylogenetic analysis by maximum likelihood method. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al. 1992). The protein sequences for *Lhx* genes were obtained from the NCBI website (https://www.ncbi.nlm.nih.gov/). The tree with the highest log likelihood (-5075.3042) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 25 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 129 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016).

1 HindIII (1)
7 SphI (1)
Exon7 13..42
483 BsaBI (1)
Intron7 43..1326

Exon8 1327..1400
1385 EcoRV (1)
GL439 1379..1400
1401 PstI (1)
1407 SalI (1)
1413 XbaI (1)
1419 BamHI (1)
1424 XmaI (1)
1424 SmaI (1)
1432 MscI (1)

pGLC58
5920 bp

1512 SacI (1)
1554 AgeI (1)

4552 Eam1105I (1)

1732 NcoI (1)

4332 FspI (1)

1830 XhoI (1)

Ampicillin 3770..4630
4185 PvuI (1)

1940 BsaAI (1)
GFP 1573..2429

4074 ScaI (1)

2116 BstZ17I (1)

2279 MfeI (1)

AmpR_Pro 3700..3728

3572 AflII (1)
3510 BstAPI (1)
3377 PfoI (1)
3319 ApaI (1)
3319 Bsp120I (1)
3314 BsiWI (1)
3310 EagI (1)
3304 SpeI (1)

2963 BtrI (1)

FIGURE 2.1: Map of pGLC58

FIGURE 2.2: Map of pGLC59
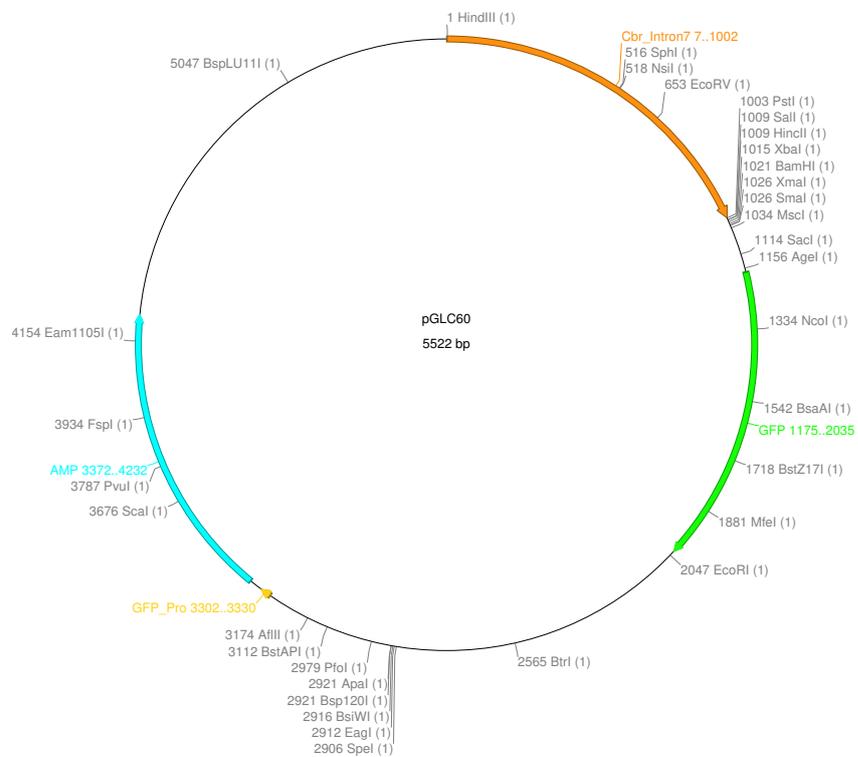
FIGURE 2.3: Map of pGLC60
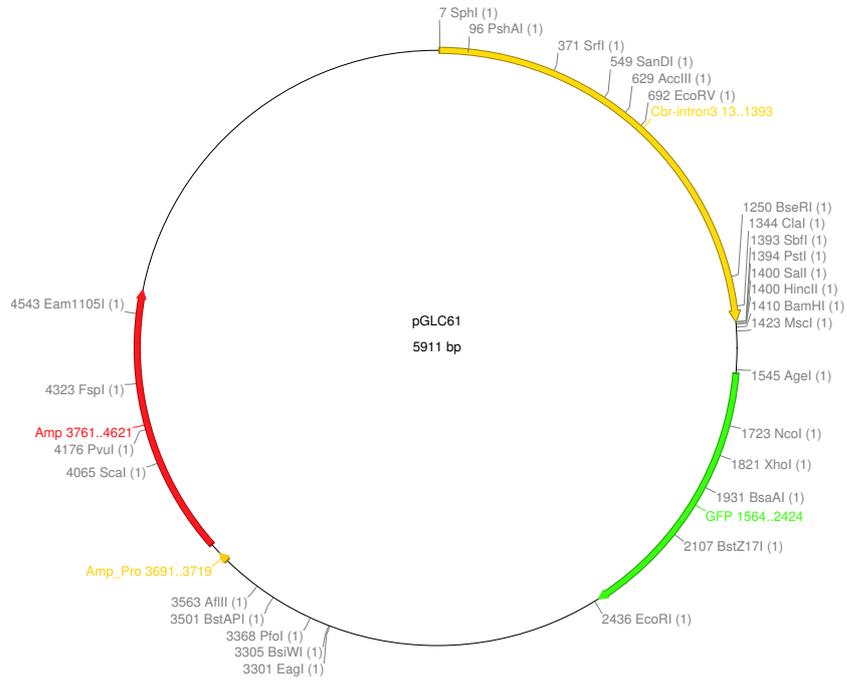
FIGURE 2.4: Map off pGLC61
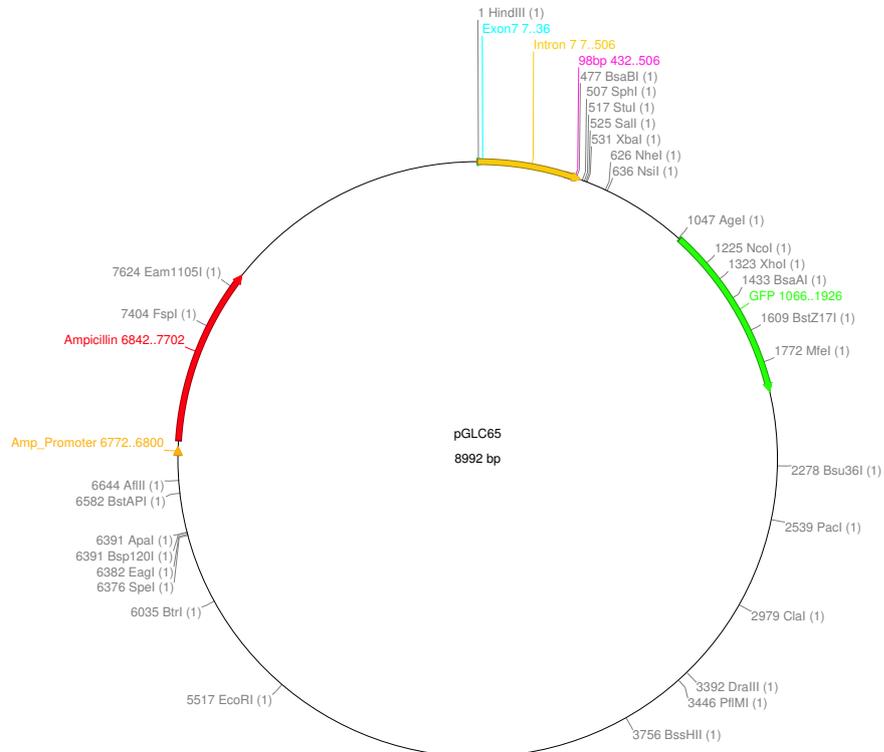
FIGURE 2.5: Map of pGLC65
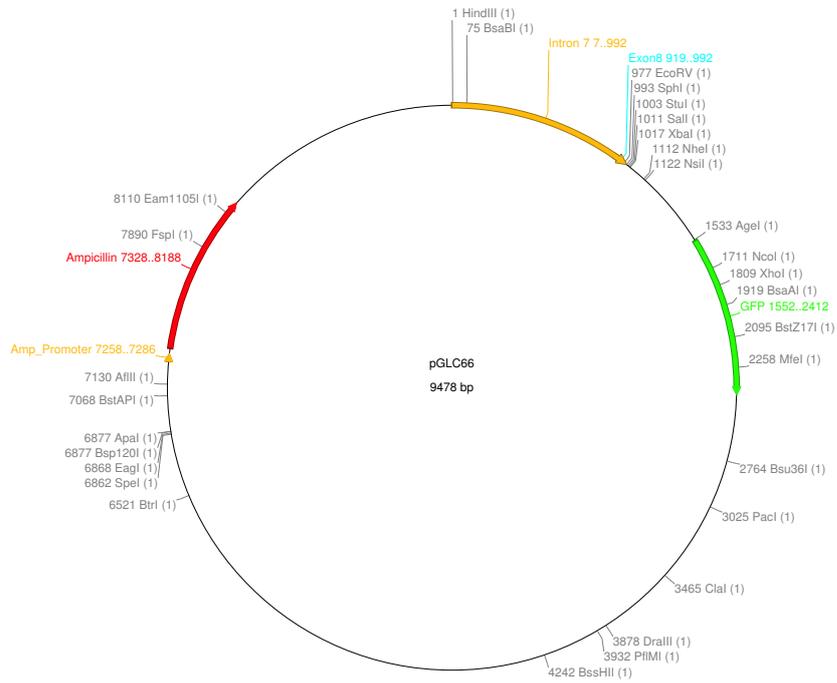
FIGURE 2.6: Map of pGLC66
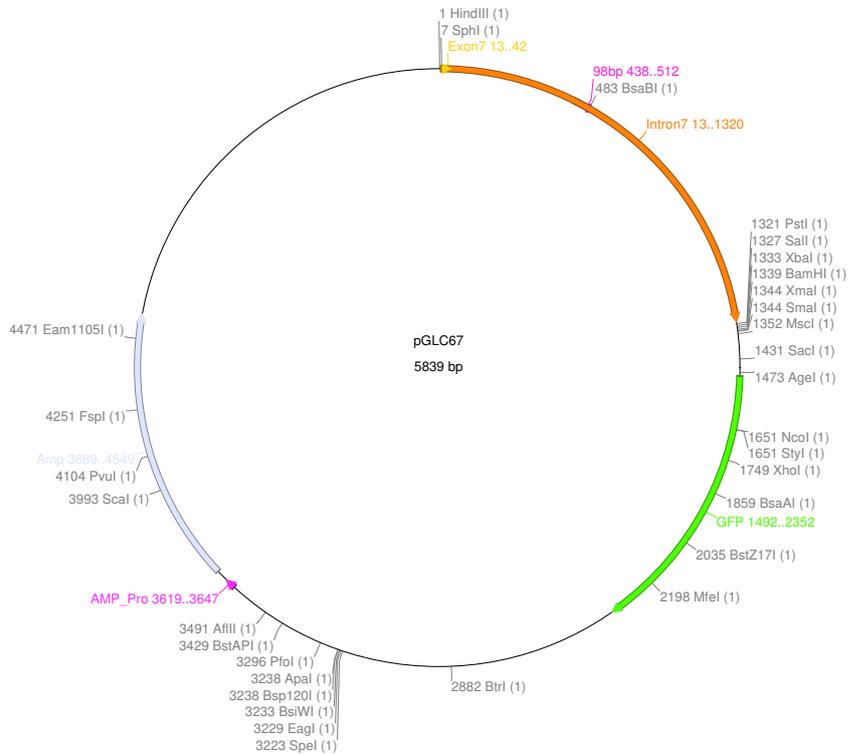
FIGURE 2.7: Map of pGLC67
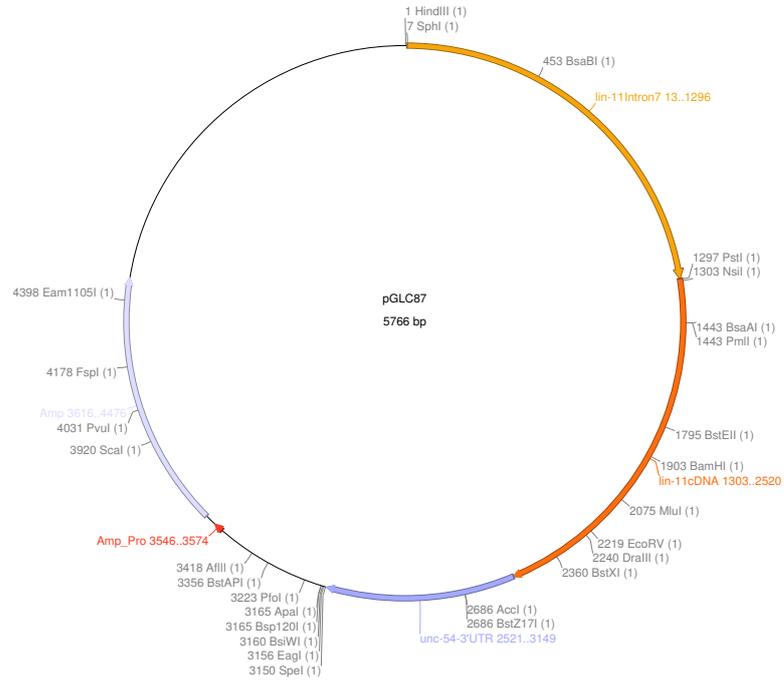
FIGURE 2.8: Map of pGLC87
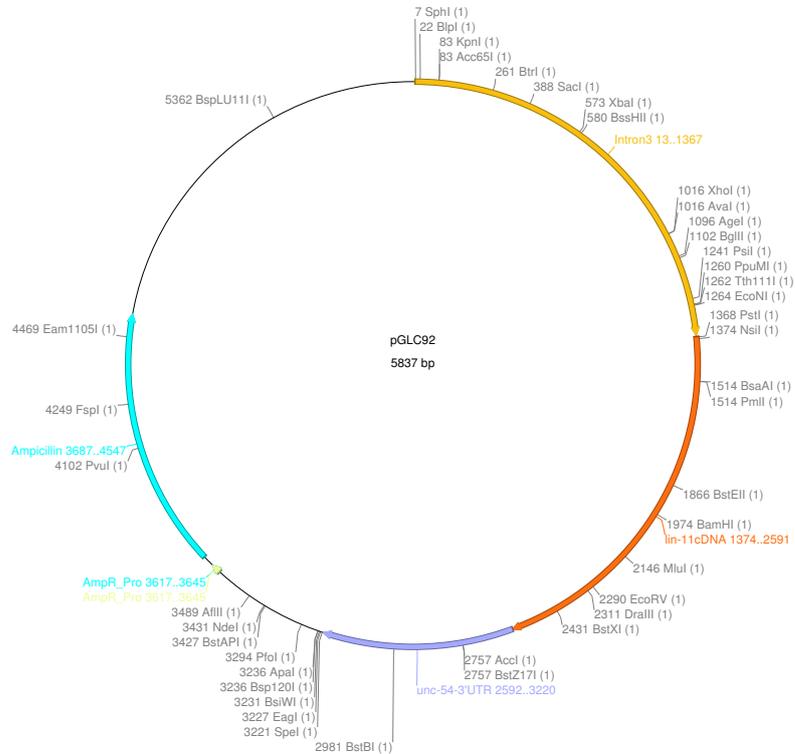
FIGURE 2.9: Map of pGLC92
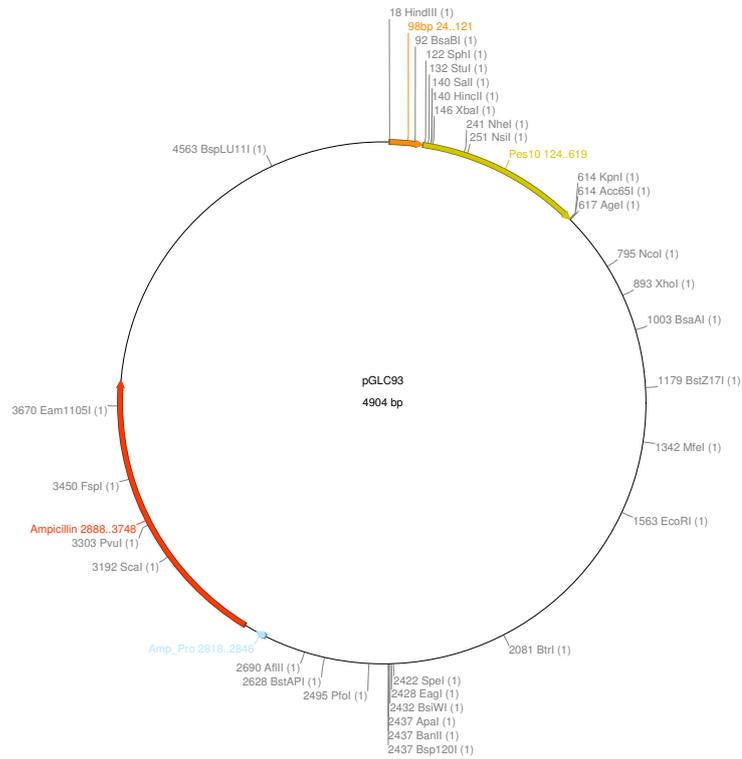
FIGURE 2.10: Map of pGLC93
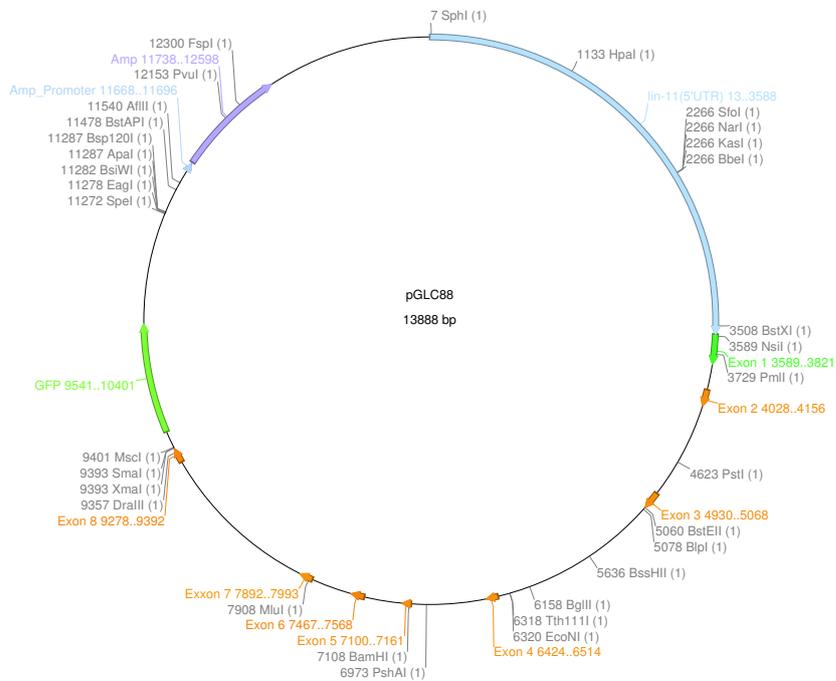
FIGURE 2.11: Map of pGLC88
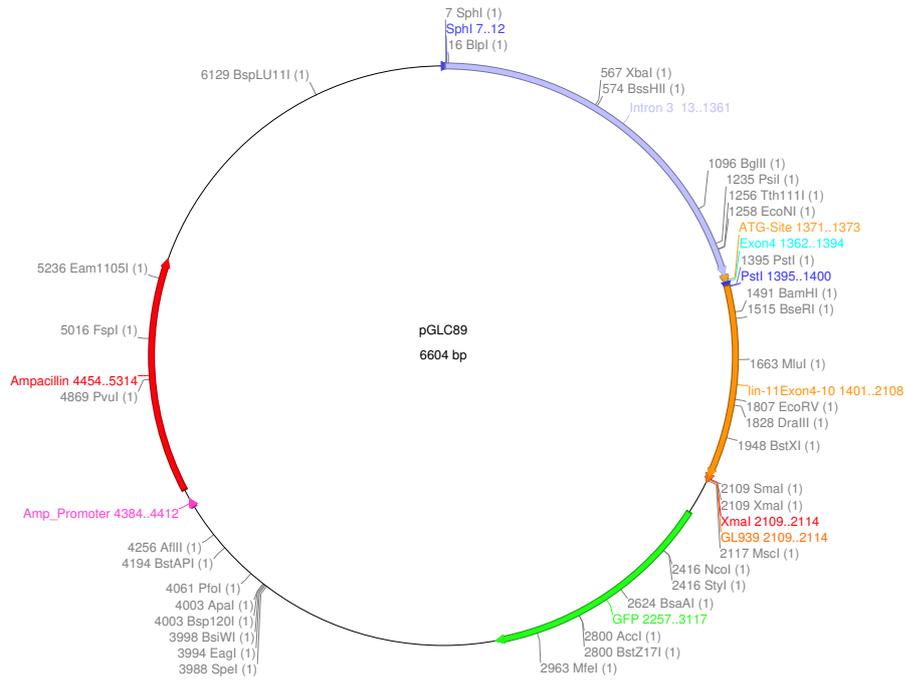
FIGURE 2.12: Map of pGLC89
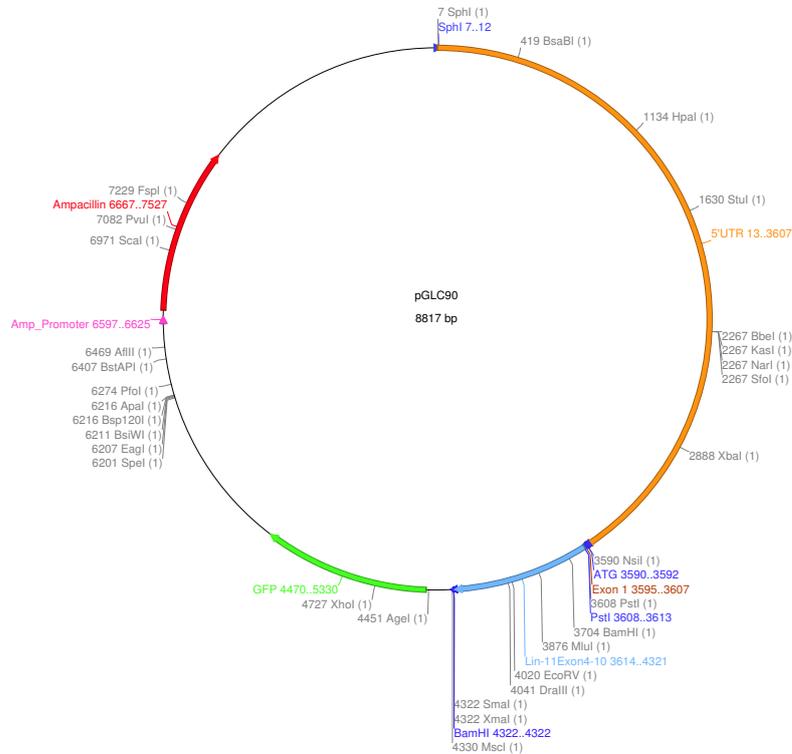
FIGURE 2.13: Map of pGLC90
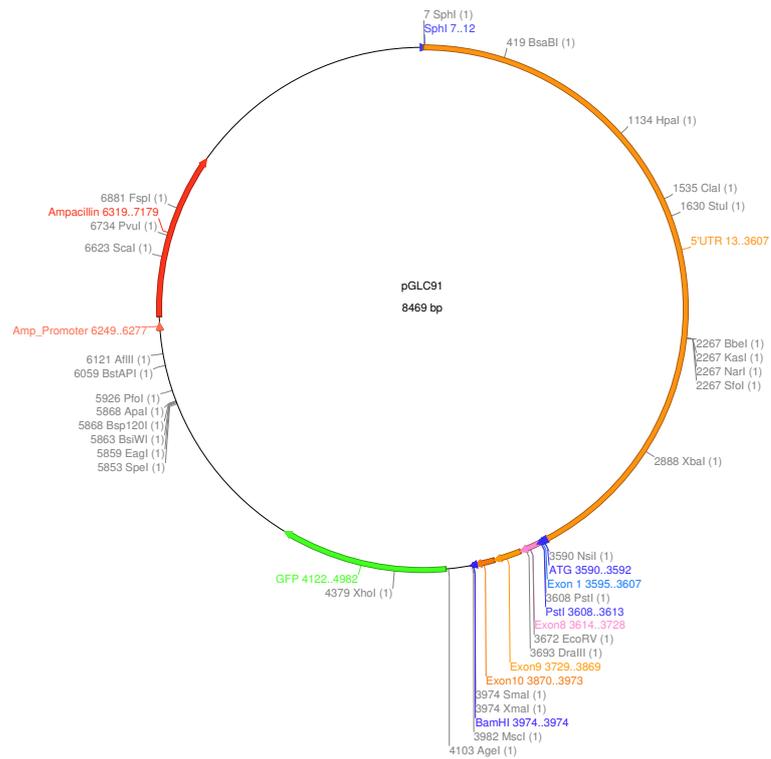
FIGURE 2.14: Map of pGLC91
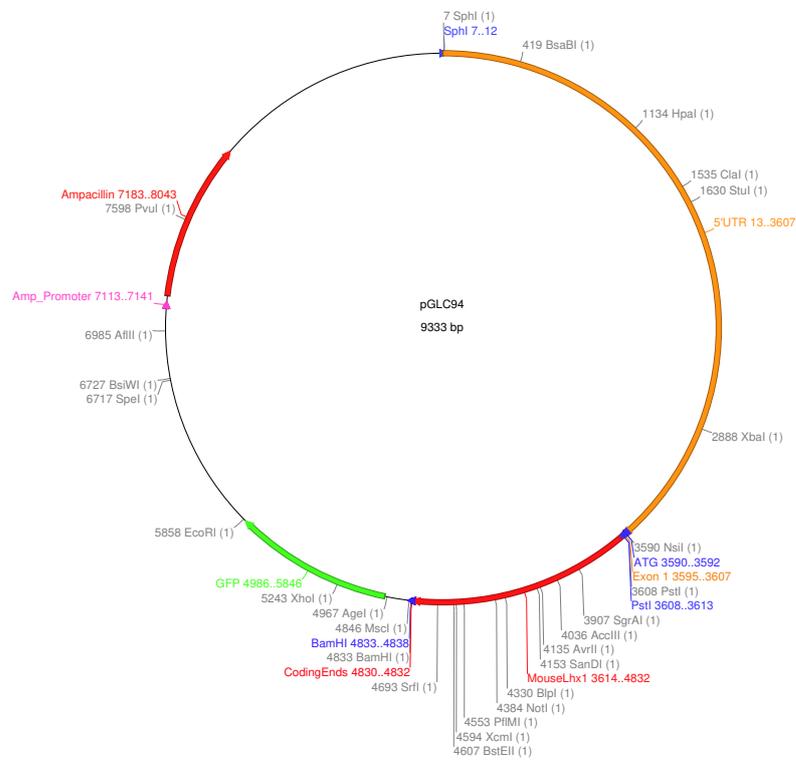
FIGURE 2.15: Map of pGLC94
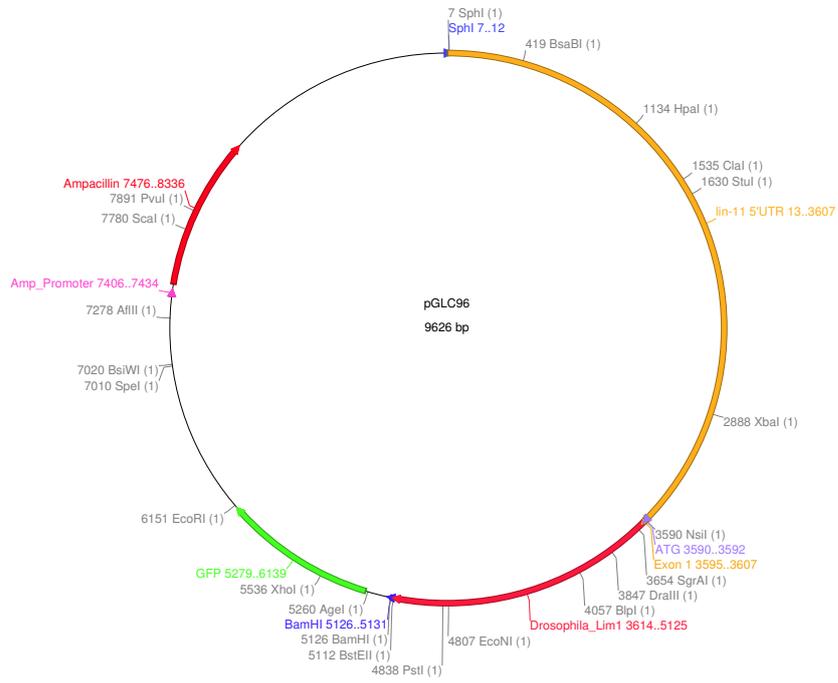
FIGURE 2.16: Map of pGLC96
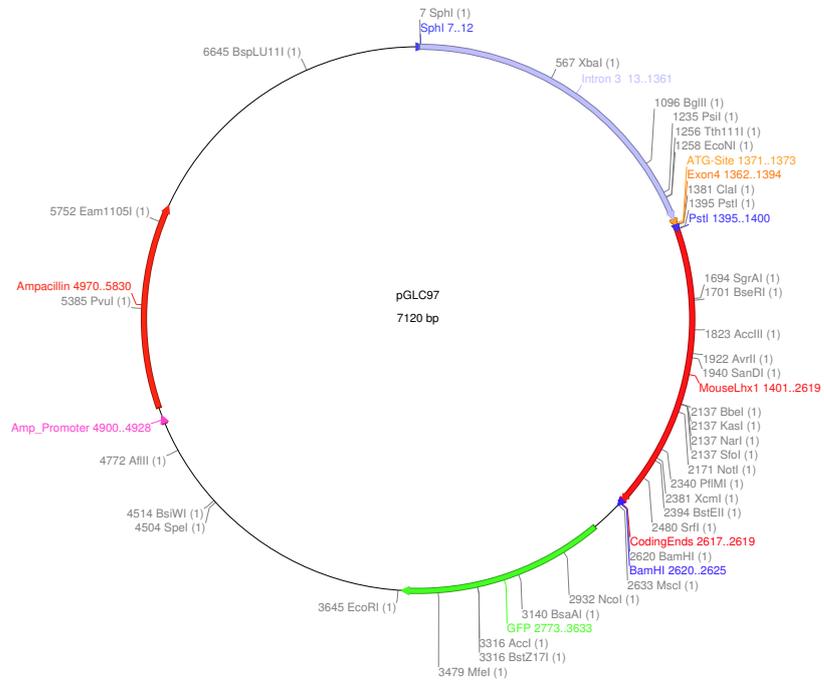
FIGURE 2.17: Map of pGLC97
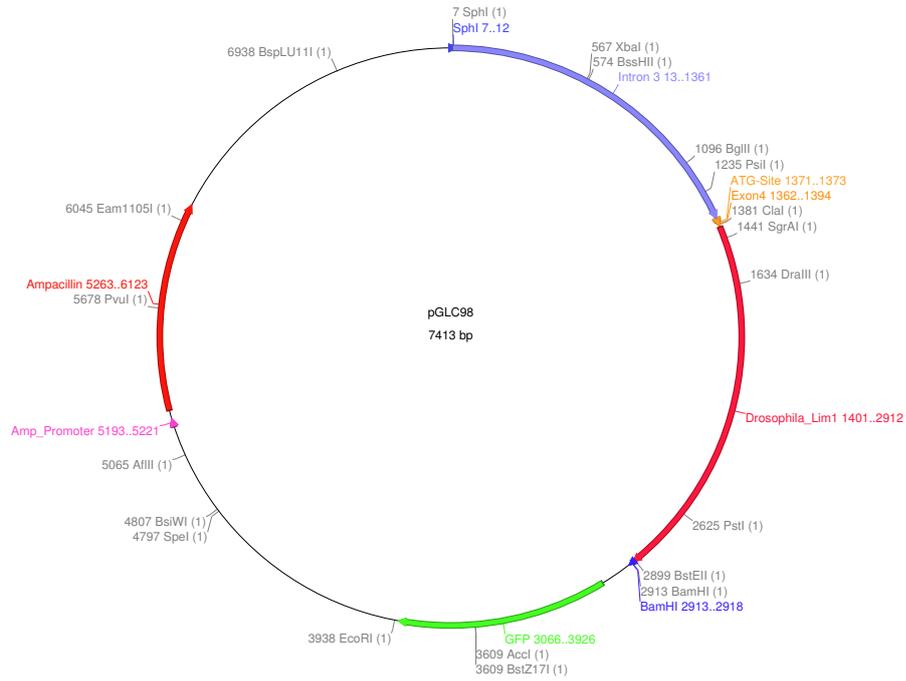
FIGURE 2.18: Map of pGLC98
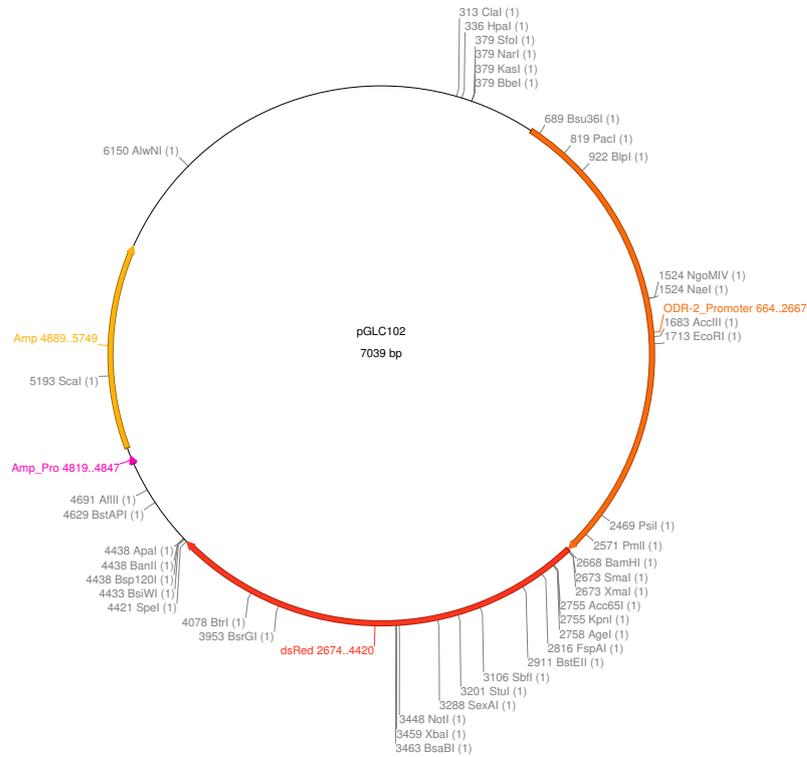
FIGURE 2.19: Map of pGLC102
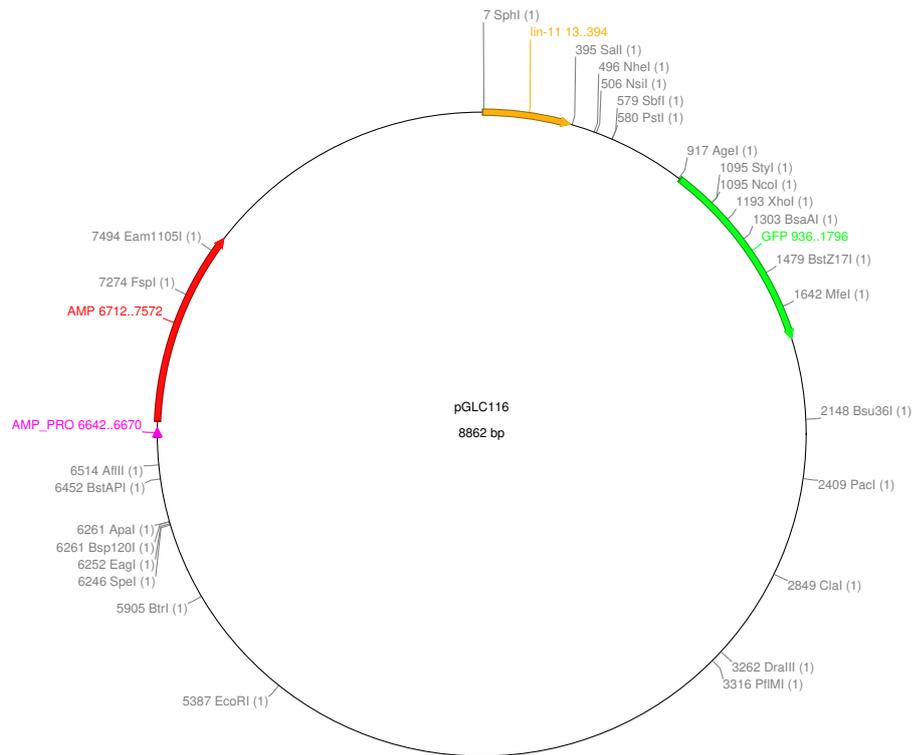
FIGURE 2.20: Map of pGLC116

FIGURE 2.21: Map of pGLC122

Chapter 3.0

Intron-specific patterns of divergence of *lin-11* regulatory function in *C. elegans* nervous system

3.1 Summary

The work described in this chapter has been published in three different journals.

3.1.1 Amon and Gupta (2017) - Developmental Biology

Amon, S., & Gupta, B. P. (2017). Intron-specific patterns of divergence of *lin-11* regulatory function in the *C. elegans* nervous system. *Developmental Biology*, 424(1), 90-103.

In this paper I have described the dissection of *lin-11* intronic regions to understand the mechanism of transcriptional regulation. I have identified two of the *lin-11* introns that activate reporter gene expression in a subset of sensory neurons. Four-way comparison of *lin-11* enhancers using orthologs from *C. briggsae*, *C. remanei*, and *C. brenneri* has uncovered multiple sub-regions, which narrowed the sequences that are likely to play important roles in *lin-11* regulation. The roles of some of these regions have been tested experimentally, revealing the mechanism of *lin-11* regulation in certain neurons. I have identified four transcription factors that are predicted to bind to these sub-regions within the introns to regulate *lin-11* expression. Three of these transcription factors

(*skn-1*, *crh-1*, and *ceh-6*) are negative regulators of *lin-11* and bind within a 35 bps sub-region of intron 3 and regulate its expression in the interneuron RIC. *ces-1* is predicted to bind within the sub-region of intron 7 and acts as positive regulator of *lin-11* in the pioneer neuron AVG. A working model has been proposed that illustrates the complex regulation of *lin-11* in neuronal differentiation.

All experiments in this manuscript were carried out by me. Asad Siddiqi and Aleem Adatia helped with the construction of six plasmids (pGLC 58, 59, 60, 61, 65 and 66). I contributed to the writing of the manuscript.

See page 54 for the published work.

3.1.2   Amon and Gupta (2017) - Data In Brief

Amon, S., & Gupta, B. P. (2017). Multi-species alignments of *C. elegans lin-11* intronic sequences and putative transcriptional regulators. Data in Brief, 12, 87-90.

This manuscript includes supplementary data related to the above Developmental Biology publication (Amon and Gupta 2017a). It includes sequence comparison of *lin-11* introns with other *Caenorhabditis* species, and also include list of TFs that are predicted to bind introns 3 and 7.

See page 68 for the published work.

3.1.3   Salam et al. (2013) - Worm

Salam, S., Ansari, A., **Amon, S**., Rezai, P., Selvaganapathy, P. R., Mishra, R. K., & Gupta, B. P. (2013). A microfluidic phenotype analysis system reveals function of sensory and dopaminergic neuron signaling in *C. elegans* electrotactic swimming behavior. In Worm (Vol. 2, No. 2, p. e24558). Taylor & Francis.

This paper includes characterization of electrosensory behavior of *lin-11(n389)* animals. It was the first study that reported electrotaxis defects in *lin-11* mutants that

are likely to be caused by abnormal differentiation of amphid neurons. My work showed that while few *lin-11* worms responded to the electric field and moved at a significantly slower speed towards cathode, majority of animals remained immobile or moved in the wrong direction, i.e., towards anode.

The published work is included in the Appendix B - Page 107.

ELSEVIER

Intron-specific patterns of divergence of *lin-11* regulatory function in the *C. elegans* nervous system

CrossMark

Siavash Amon, Bhagwati P. Gupta*

Department of Biology, McMaster University, Hamilton, ON, Canada L8S-4K1

ARTICLE INFO

ABSTRACT

The diversity of neurons in the nervous system is specified by many genes, including those that encode transcription factors (TFs) and play crucial roles in coordinating gene transcription. To understand how the spatiotemporal expression of TF genes is regulated to generate neuronal diversity, we used one member of the LIM-Hox family, *lin-11*, as a model that is necessary for the differentiation of amphid neurons in the nematode *C. elegans* and a related species *C. briggsae*. We characterized transcriptional regulation of *lin-11* and uncovered regulatory roles of two of the largest introns, intron 3 and intron 7. These introns promote *lin-11* expression in non-overlapping sets of neurons. Phenotypic rescue experiments in *C. elegans* revealed that intron 3 is capable of restoring *lin-11* function based on gene expression patterns and behavioral assays. Interestingly, intron 3-driven reporter expression showed differences in neuronal cell types between *C. briggsae* and *C. elegans,* indicating evolutionary changes in *lin-11* regulation between the two species. Reciprocal transformation experiments provided further evidence consistent with functional changes in both *cis* and *trans* regulation of *lin-11*. To further investigate transcriptional regulation of *lin-11*, we dissected the intronic regions in *C. elegans* and identified cell-specific enhancers. These enhancers possess multiple sequence blocks that are conserved among *Caenorhabditis* species and possess TF binding sites. We tested the role of a subset of predicted TFs and discovered that while three of them (SKN-1, CEH-6, and CRH-1) act via the intron 3 enhancer to negatively regulate *lin-11* expression in neurons, another TF (CES-1) acts positively via the intron 7 enhancer. Overall, our findings demonstrate that neuronal expression of *lin-11* involves multiple TF regulators and regulatory modules some of which have diverged in *Caenorhabditis* nematodes.

1. Introduction

The nervous system of an animal consists of many different types of neurons that function to modulate behavior through control of perception, locomotion, and physiology. The identity and function of these neurons are specified by the coordinated action of many genes whose spatiotemporal expression is regulated by transcription factors (TFs). The nematode *C. elegans* is an established model organism for studying how neurons develop and acquire distinct identities. *C. elegans* hermaphrodites contain a small set of neurons (302) (White et al., 1986), making it relatively easy to manipulate these cells at the level of small circuits. This simple nervous system contains diverse neuronal cell types that are characterized by the neurotransmitter produced and their axonal trajectory (Bargmann, 2006; Hobert, 2013, 2016a, 2016b; Hobert et al., 2016; White et al., 1986). A number of TF genes are known to play important roles in the development of *C. elegans* neurons (Hobert, 2016a, 2016b). These include the LIM-Hox family member *lin-11* (Freyd et al., 1990). LIM-Hox genes are a

subfamily of the homeobox (Hox)-containing TF genes. The proteins encoded by these genes share a tandem zinc finger domain that mediates protein-protein interactions (Hobert and Westphal, 2000). The LIM domain was first identified as a conserved cysteine-rich region shared by LIN-11 (*C. elegans*) and two other proteins, ISL1 (mouse) and MEC-3 (*C. elegans*) (Freyd et al., 1990; Karlsson et al., 1990; Way and Chalfie, 1988).

While the LIM-Hox genes are involved in the development of numerous cells and organs, their roles in neuronal differentiation appear to be a common evolutionary theme. *lin-11* mutants show defects in behavioral processes such as thermosensation, electrosensation and chemosensation (Hobert et al., 1998; Salam et al., 2013; Sarafi-Reinach et al., 2001). Consistent with this, *lin-11* is expressed in a subset of amphid neurons (Hobert et al., 1998; Sarafi-Reinach et al., 2001).

Similar to *lin-11*, other LIM-Hox family members are also known to play important roles during neuronal development (Curtiss and Heilig, 1998; Dawid et al., 1998; Hobert and Westphal, 2000). For example,

dLim1 is required for the proper specification of motor neurons and a subset of interneurons (Lilly et al., 1999). In mouse, *isl1* is involved in the formation of the nervous system, including the differentiation of nociceptive neurons, cranial ganglion neurons, and basal forebrain cholinergic neurons (Elshatory and Gan, 2008; Liang et al., 2011; Sun et al., 2008). Other LIM-Hox genes such as *lhx2*, *lhx3* and *lhx4* are also necessary for the formation of neuronal tissues (Hirota and Mombaerts, 2004; Kolterud et al., 2004; Pinto do et al., 2002; Sharma et al., 1998). While *lhx2* promotes neuronal differentiation in human embryonic stem cells and mouse brain (Hou et al., 2013), *lhx3* and *lhx4* are required for the differentiation and migration of over-lapping subsets of motor neurons (Pfaff et al., 1996; Sharma et al., 1998; Sheng et al., 1996).

Closely related species of *Caenorhabditis* also offer advantages for the study of the LIM-Hox genes in the nervous system. In this regard, *C. briggsae* is particularly useful because of its morphological similarity to *C. elegans* and its amenability to genetic manipulation (Gupta et al., 2007). Sequence comparison between *C. elegans* and *C. briggsae* has revealed that roughly one-third of the genes lack one-to-one orthologous relationship, making *C. briggsae* a good model for identifying conserved as well as divergent mechanisms that underlie the specification of neuronal processes (Gupta et al., 2007; Gupta and Sternberg, 2003) (also see InParanoid8 http://inparanoid.sbc.su.se). For example, a study of *gcy* (guanylyl cyclase) receptor genes uncovered interesting differences in asymmetric expression of some of the family members in the ASE (L/R) pair of gustatory neurons (Ortiz et al., 2006).

Although a few genes that genetically interact with *lin-11* in amphid neurons have been identified (Schmid et al., 2006; Sze and Ruvkun, 2003), the regulatory mechanism of *lin-11* expression in these neurons is not understood. To this end, we have characterized *lin-11* non-coding regions and report the identification of new enhancer elements that reside within the two largest introns, 3 and 7. Intron 3 was previously shown to activate *GFP* reporter expression in certain amphid neurons, although its precise role in *lin-11* regulation was not studied (Yamada et al., 2012). We show that while intron 3 directs reporter gene expression in a total of seven amphid neurons, including three (ADL, ADF, and AVJ) observed in most animals, intron 7 primarily promotes expression in AVG. Our rescue experiments in *C. elegans*, using both *GFP* reporter and behavioral assays, confirmed that enhancer sequences from intron 3 are necessary and sufficient for *lin-11* function. Using transgenic strains, we examined the functional and sequence conservation of the intronic enhancers between *C. elegans* and *C. briggsae*. The results showed that while intron 7 enhancer-driven expression is observed in the same neuron in both species, intron 3 drives expression in different sets of neurons. These results, together with interspecies expression studies, provide the first evidence for both *cis* and *trans* evolution of intron 3-mediated neuronal expression of *lin-11*.

Using sequence comparison, we identified conserved regions within introns that may play important roles in *lin-11* regulation. Dissection of these regions in *C. elegans* identified elements in intron 3 that are specific for RIC neuron and in intron 7 for AVG and RIF neurons. Next, we used an *in silico* approach to find TF binding sites within the conserved regions that may regulate *lin-11* expression in subsets of amphid neurons. Further experiments confirmed the involvement of four TF genes, *skn-1* (bZIP family), *crh-1* (CREB family), *ceh-6* (POU Homeodomain) and *ces-1* (C2H2 Zn finger), in the *lin-11*-mediated neuronal specification process. Overall, our results demonstrate that the spatiotemporal regulation of *lin-11* is directed by distinct enhancer elements that respond to multiple transcription factors in a neuron-specific manner.

2. Materials and methods

2.1. Strains

All cultures were maintained at 20 °C using standard culture methods (Brenner, 1974). Worms were grown on lawns of *E. coli* OP50 (uracil auxotroph) bacteria. The strains used in this study are listed below.

C. elegans: Wild-type N2, bhIs8[pGLC58(*lin-11-int7p::GFP*)+*unc-119*(+)], bhIs10[pGLC58(*lin-11-int7p::GFP*)+*unc-119*(+)], bhEx97[pGLC59(*lin-11-int3p::GFP*)+*unc-119*(+)], bhEx98[pGLC59(*lin-11-int3p::GFP*)+*unc-119*(+)], bhEx113[pGLC60(*Cbr-lin-11-int7p::GFP*)+*unc-119*(+)], bhEx115[pGLC60(*Cbr-lin-11-int7p::GFP*)+*unc-119*(+)], bhEx127[pGLC66(*lin-11-int7p-986bp::GFP*)+*unc-119*(+)], bhEx128[pGLC66(*lin-11-int7p-986bp::GFP*)+*unc-119*(+)], bhEx150 [pVH10.17(*odr-2::CFP*)], bhEx159[pGF50 (lin-11 genomic)+pVH10.17(*odr-2::CFP*)], bhEx160[pGF50 (lin-11 genomic)+pVH10.17(*odr-2::CFP*)], bhEx167[pGLC93(*lin-11-int7p-98bp::GFP*)+pVH10.17(*odr-2::CFP*)+*unc-119*(+)], bhEx168[pGLC93 (*lin-11-int7p-98bp::GFP*)+pVH10.17(*odr-2::CFP*)+*unc-119*(+)], bhEx176[pGLC92(*lin-11-int3p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx177[pGLC87 (*lin-11-int7p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx189[pGLC116(*lin-11-int3p-382bp::GFP*)+pGLC102(*odr-2::DsRed*)+*unc-119*(+)], bhEx190[pGLC116(*lin-11-int3p-382bp::GFP*)+pGLC102(*odr-2::DsRed*)+*unc-119*(+)], bhEx202[pGLC102(*odr-2b::dsRED*)+pVH10.17 (*odr-2::CFP*)], bhEx203[pGLC87(*lin-11-int7p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx204[pGLC87 (*lin-11-int7p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx211[pGLC92(*lin-11-int3p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx213[pGLC61(*Cbr-lin-11-int3p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx214[pGLC61(*Cbr-lin-11-int3p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx215[pGLC65(*lin-11-int7p-598bp::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx216[pGLC65(*lin-11-int7p-598bp::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx228[pGLC122(*lin-11-int3p-348bp:: GFP*)+pGLC132(*lin-11-int3p::DsRed*)], bhEx229[pGLC122(*lin-11-int3p-348bp::GFP*)+pGLC132(*lin-11-int3p::DsRed*)], bhEx234[pGLC92(*lin-11-int3p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx235[pGLC92(*lin-11-int3p::lin-11cDNA*)+pVH10.17(*odr-2::CFP*)], bhEx241[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::dsRed*)], bhEx243[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::dsRed*)], bhEx249[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::dsRed*)], bhEx250[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::dsRed*)], bhEx251[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::dsRed*)], crh-1(tz2), ceh-2(ch4), ceh-6(gk679), ces-1(n703), fax-1(gm83), lin-15b(n744), skn-1(ok2315); nT1[qIs51], sox-3(ok510), tbx-2(ut180), uIs60(unc-119::sid-1, unc-119::YFP).

C. briggsae: Wild-type AF16, bhEx197[pGLC61(*Cbr-lin-11-int3p::GFP*)+*unc-119*(+)], bhEx200[pGLC61(*Cbr-lin-11-int3p::GFP*)+*unc-119*(+)], bhEx205[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx206[pGLC59(*lin-11-int3p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx207[pGLC59(*lin-11-int3p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx208[pGLC58(*lin-11-int7p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx217[pGLC116(*lin-11-int3p-382bp::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx218[pGLC116(*lin-11-int3p-382bp::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx219[pGLC60(*Cbr-lin-11-int7p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx220[pGLC60(*Cbr-lin-11-int7p::GFP*)+pGLC102(*odr-2::DsRed*)], bhEx227[pGLC93(*lin-11-int7p-98bp::GFP*)+pGLC102(*odr-2::DsRed*)].

2.2. RNAi

RNAi was performed using the Ahringer lab bacterial feeding library. The protocol has been previously described (Timmons and Fire, 1998). All experiments were repeated at least three times, and batches with similar results were pooled and analyzed. Two controls, empty vector L4440 and *GFP* knockdown, were used in all batches. Synchronized L1 larvae were fed with bacteria containing double-

stranded RNA. Worms were synchronized as follows: gravid adults were treated with a hypochlorite solution for 4–5 min. Embryos were washed five times with M9 and allowed to hatch for 16–30 h at 20° with gentle agitation. The *sid-1* RNAi-hypersensitive strain, which expresses *sid-1* under the control of the *unc-119* promoter (*bhEx202*), was also used. This strain increases the efficiency of RNAi knockdown in the nervous system (Calixto et al., 2010).

2.3. Molecular biology and transgenics

The transgenic animals carrying extrachromosomal arrays were generated using a standard microinjection technique (Mello et al., 1991). *C. elegans unc-119* (Maduro and Pilgrim, 1995) was used as a rescue marker in most cases. pVH10.17 (*odr-2::CFP*) and pVH13.05 (*glr-1::DsRed*), kindly provided by the Hutter lab, were used as neuronal markers. The plasmid pGF50 is a 19.5 kb subclone of the ID6 cosmid, which contains the entire *lin-11* coding region and regulatory sequences (Freyd, 1991).

We found earlier that both introns 3 and 7 contain enhancer as well as promoter sequences since they are capable of activating *GFP* reporter when cloned into a Fire lab 'promoter-less' vector (from June 1995 kit). Therefore expression studies involving full-length intronic sequences were performed using pPD95-series vectors. However, fragments of the introns were sub-cloned into the 'enhancer-assay' vector pPD107.94 because they lacked endogenous promoter and were unable to activate reporter expression on their own.

The following plasmids were made as part of this study. The concentrations of plasmids ranged from 20 ng/μl to 60 ng/μl. The total injection mix was less than 200 ng/μl. To eliminate an array-specific bias in *GFP* expression or phenotypic rescue studies, more than one transgenic line per construct was examined, and one line with a typical pattern was studied in detail.

pGLC58 (*lin-11-int-7p::GFP*). A 1388 bp fragment of *C. elegans lin-11* containing the full-length intron 7 sequence was amplified using the primers 5′-caacttgcatgcgagacaggactcaacatgagag′-3′ and 5′-ggtcttctgcagcgggattgacgatatcctccaa-3′ and digested with *Sph*I and *Pst*I. This fragment was cloned into the Fire vector pPD95.67.

pGLC59 (*lin-11-int-3p::GFP*). A 1355 bp fragment of *C. elegans lin-11* containing the full-length intron 3 sequence was amplified using the primers 5′-aggcatgccaggctaagcctcttctc-3′ and 5′-atctgcaggaca-cacgaatcgattgccttcc-3′ and digested with *Sph*I and *Pst*I. This fragment was cloned into the Fire vector pPD95.73.

pGLC60 (*Cbr-lin-11-int-7p::GFP*). A 996 bp fragment of *C. briggsae lin-11* containing the full-length intron 7 sequence was amplified using the primers 5′-ccaagcttgggctgaacatgagagtcattcag-3′ and 5′-atctgcagccaccgtatcgaagctgtttc-3′ and digested with *Pst*I and *Hind*III. This fragment was cloned into the Fire vector pPD95.67.

pGLC61 (*Cbr-lin-11-int-3p::GFP*). A 1381 bp fragment of *C. briggsae lin-11* containing full-length intron 3 amplified using the primers 5′-ctgcatgctacgggtgatcaggtgagaatctc-3′ and 5′-tgctgcag-gagtcgatgtttttggtagccg-3′. The fragment was digested with *Sph*I and *Pst*I and cloned into the Fire vector pPD95.73.

pGLC65 (*lin-11-int-7p-500bp::GFP*). A 500 bp fragment of *C. elegans lin-11* containing part of the intron 7 sequence was amplified using the primers 5′-caacttaagcttgagacaggactcaacatgagag-3′ and 5′-gttgcatgcttctctcccttgaccactca-3′. The fragment was digested with *Hind*III and cloned into the Fire vector pPD107.94.

pGLC66 (*lin-11-int-7p::GFP*). A 986 bp fragment of *C. elegans lin-11* containing part of the intron 7 sequence was amplified using the primers 5′-atcaagcttggttgtgtttagagagctatacg-3′ and 5′-ggtcttgcatgccgggattgacgatatcctccaa-3′ and digested with *Hind*III and *Sph*I. This fragment was cloned into the Fire vector pPD107.94.

pGLC67 (*lin-11-int-7p::GFP*). A 1307 bp fragment of *C. elegans lin-11* containing part of the intron 7 sequence was amplified using the primers 5′-caacttgcatgcgagacaggactcaacatgagag-3′ and 5′-cacctgca-

gaatggaaatttatttggaaaacaaattaatgggag-3′ and digested with *Sph*I and *Pst*I. This fragment was cloned into the Fire vector pPD95.69.

pGLC87 (*lin-11-int7p::lin-11cDNA*). Full-length *C. elegans lin-11* cDNA was obtained by digesting the plasmid pYK452F7-1 (Gupta and Sternberg, 2002) with *Pst*I and *Spe*I. A 3.1 kb piece was subcloned into the plasmid pGLC67.

pGLC92 (*lin-11-int3p::lin-11cDNA*). Full-length *C. elegans lin-11* cDNA was obtained by digesting pYK452F7-1 plasmid (Gupta and Sternberg, 2002) with *Pst*I and *Spe*I. A 3.1 kb fragment was subcloned into the plasmid pGLC59.

pGLC93 (*lin-11-int-7p-98bp::GFP*). A 98 bp fragment of *C. elegans lin-11* intron 7 sequence was amplified using the primers 5′-gatgaagcttggttgtgtttagagagc-3′ and 5′-aatgcatgctctctcccttgaccactc-3′ and digested with *Sph*I and *Hind*III. This fragment was cloned into the Fire vector pPD136.64.

pGLC116 (*lin-11-int-3p-382bp::GFP*). A 382 bp fragment of *C. elegans lin-11* intron 3 sequence was amplified using the primers 5′-atagcatgccttgaatgtgtccccatcat-3′ and 5′-tatttattcattattttgaagt-caatcggtcgactat-3′ and digested with *Sph*I and *Sal*I. This fragment was cloned into the Fire vector pPD107.94.

pGLC122 (*lin-11-int-3p-348bp::GFP*). A 348 bp fragment of *C. elegans lin-11* intron 3 sequence was amplified using the primers 5′-aatcgcatgccttgaatgtgtccccatcatgagcccctgtcagttgcacctgcctcttaccga-gagctctctaccactcgtttttcttgaatagaagatgtcact-3′ and 5′-tatttattcattattttt-gaagtcaatcggtcgactat-3′ and digested with *Sph*I and *Sal*I. This fragment was cloned into the Fire vector pPD107.94.

2.4. Microscopy

Worms were mounted on 3% agar pads using 0.03% sodium azide as an anesthetic. DIC and fluorescence microscopy was performed using Nikon Eclipse 80i and Zeiss ApoTome2 microscopes. Images were captured with Hamamatsu ORCA-ER and Zeiss AxioImager D1 digital cameras and processed using the software Nikon NIS-Elements version 2.0 and Zeiss Zen version 2.0.

2.5. Behavioral assays

The experimental setup for the electrotaxis assay has been previously described (Rezai et al., 2010; Salam et al., 2013). Stage-synchronized animals were obtained by hypochlorite treatment. Briefly, gravid hermaphrodites were grown at 20 °C and were treated for 4–5 min with a hypochlorite solution containing commercial bleach and 4 N NaOH (3:2 ratio). Embryos were washed five times with M9 and then allowed to hatch in M9 buffer overnight. The L1s were placed on NG-agar plates and grown until adulthood (60 h at 20 °C). The young adults of N2, *lin-11*(*n389*) and *bhEx176* (*lin-11-int3p::lin-11cDNA*) with no eggs were selected for assays. The DC electric field was activated, and each animal was allowed to travel a minimum distance of 5000 μm. Electrotaxis responses of worms were recorded using a video camera. Ten or more animals were scored for each strain; for each animal, the response was monitored for 2–9 min and for up to 45,000 μm total distance. NIH ImageJ software was used to analyze the videos and to manually calculate the speed of each animal. Up to three turning events were used to determine the average response of each animal.

Single-worm thermotaxis assays were performed as previously described (Mori and Ohshima, 1995). Stage-synchronized L1 animals were placed on NG-agar plates and allowed to grow for two generations at 20 °C. Young adults of the F2 generation were placed on the outer edge of assay plates (one worm per plate). Prior to transferring the worm, the plate was kept on top of a vial of frozen acetic acid (2.7 cm in diameter) for an hour to establish a temperature gradient. The movement track of the animal was manually recorded and analyzed using three concentric circles drawn on the back of the plate, corresponding to different temperature zones (coldest in the innermost, cultivation

92

temperature in the middle ring, and near room temperature in the outermost ring).

2.6. Bioinformatics

The *lin-11* intronic sequences from *C. elegans*, *C. briggsae*, *C. brenneri* and *C. remanei* were aligned using MussaGL (version 1.1.0) (http://woldlab.caltech.edu/cgi-bin/mussa). The conservation threshold was set at a minimum of 70% (21 per 30-nucleotide sliding window) and a maximum of 90% (27 per 30-nucleotide sliding window).

To identify putative TF genes for introns 3 and 7, we used the CIS-BP database software (http://cisbp.ccbr.utoronto.ca/TFTools.php) (Weirauch et al., 2014). See Table 1 in Amon and Gupta (submitted for publication) for a list of predicted genes.

3. Results

LIN-11 is required for the proper specification of a subset of amphid neurons in *C. elegans* (Hobert et al., 1998; Sarafi-Reinach et al., 2001). We set out to identify regulatory regions in *lin-11* and *trans*-acting factors required for this specification. Previously, we demonstrated that the *lin-11* promoter contains distinct sequences required for response to the Wnt and LIN-12/Notch pathways in the developing vulva and uterine π lineage cells, respectively (Gupta and Sternberg, 2002; Marri and Gupta, 2009; Newman et al., 1999). However, this fragment, which contains the entire intergenic region between *lin-11* and the upstream gene, did not drive expression in amphid neurons.

3.1. lin-11 *expression in amphid neurons is regulated by enhancers located within introns 3 and 7*

To identify additional, neuronal-specific regulatory elements, we focused on the two largest introns of *lin-11*, intron 3 (termed 'int3') and intron 7 (termed 'int7'). These introns were part of the genomic fragments of *lin-11* that showed enhancer activities in neurons in transgene expression experiments (B. Gupta and P. Sternberg, unpublished). Therefore in this study we used full-length intronic sequences to characterize *GFP* expression patterns in detail. Transgenic *C. elegans* stable lines carrying *lin-11-int3p::GFP* and *lin-11-int7p::GFP* transgenes were generated, and *GFP* expression was analyzed in more than one independently isolated line (see Section 2). Examination of stable lines revealed that both *lin-11-int3p::GFP* and *lin-11-int7p::GFP* animals show distinct patterns of neuronal *GFP* expression.

The *lin-11-int3p::GFP* (pGLC59) transgenic lines that contain the entire intron 3 sequence showed *GFP* reporter expression in seven neurons at the larval stage L4. These include two sensory neurons (ADL and ADF), four interneurons (AVJ, RIC, AIZ, and RIF) and one pioneer interneuron (AVG) (see below). Expression was also detected in embryos. Because some of the *lin-11* neurons are specified in the embryo, *lin-11::GFP* was expected to be expressed in these neurons during their specification. Consistent with this, GFP fluorescence was observed as early as in the pre-gastrula stage in the presumptive head and tail regions (Fig. 1A and B). By the 3-fold embryonic stage, fluorescence could be seen in neuronal cells in the anterior region (Fig. 1E and F). Specific neurons were identified in larval stages based on their stereotypic positions and morphologies (see Fig. 1I and K for a mid-L4 stage animal). GFP fluorescence in ADL, ADF and AVJ was very strong and observed in almost all animals (Figs. 1I and 2A). Fluorescence was less frequently observed in RIC and AIZ and was rarely seen in RIF and AVG (Figs. 1 and 2). This pattern of *lin-11-int3p::GFP* expression agrees with what has been reported earlier by others using different transgenic strains (Hobert et al., 1998; Sarafi-Reinach et al., 2001; Yamada et al., 2012). Overall, these results show

that intron 3 plays an important role in regulating *lin-11* expression in several amphid neurons.

Unlike intron 3, intron 7-driven *GFP* reporter expression in *C. elegans lin-11-int7p::GFP* (pGLC58) transgenic animals was observed in fewer cells. Fluorescence was first detected in two head neurons in post-gastrulating embryos; based on reporter gene expression, one of those neurons is likely to be AVG (Fig. 3A and B). The identification of this neuron as AVG in larval stages was confirmed by co-localization with two reporters, *glr-1::DsRed* (Fig. 3I) and *odr-2::DsRed* (not shown) (Chou et al., 2001; Maricq et al., 1995). The other unidentified cell (Fig. 3I and K) had variable fluorescence.

Fluorescence in *lin-11-int-7p::GFP* animals persisted in both neurons throughout the larval stages, but the intensity decreased with time. Adults showed faint expression mostly in AVG. In summary, introns 3 and 7 show significant differences in enhancer activities in amphid neurons.

3.2. lin-11 *intron 3 can promote expression of* lin-11 *and rescue amphid neuron defects*

Our findings that introns 3 and 7 activate reporter gene expression in amphid neurons led us to further investigate their roles in neuronal specification. To this end, we carried out rescue experiments using a molecular null allele of *lin-11, n389*, that lacks the entire *lin-11* coding region. The *lin-11(n389)* animals have defects in amphid neurons and consequently show behavioral abnormalities such as inability to sense and remember the cultivation temperature and a lack of electrosensory response (Hobert et al., 1998; Salam et al., 2013).

We asked whether the enhancers within these introns (3 and 7) are capable of activating *lin-11* expression and rescuing the neuronal defect in *lin-11(n389)* animals. To address this, transgenic animals carrying full-length *C. elegans lin-11* cDNA under the control of intronic sequences (pGCL92 for intron 3 and pGLC87 for intron 7) were generated. The *odr-2::CFP* reporter was used as a cell fate marker to confirm the differentiation of *lin-11*-expressing neurons. The examination of CFP fluorescence in all four stable lines (see Section 2) revealed that pGLC92 was able to rescue RIF and AIZ specification defects in a majority of animals but had no significant impact on the AVG phenotype (Fig. 4).

In addition to the reporter assay, we tested transgenic lines for changes in behavioral phenotypes. *C. elegans* exhibit the tendency to move towards the cultivation temperature when placed in a thermal gradient, a phenomenon known as thermotaxis (Hedgecock and Russell, 1975; Mori and Ohshima, 1995). This thermotactic response depends upon input from two opposing temperature-sensing pathways and is mediated by the interneurons AIZ and AIY (Mori and Ohshima, 1995). The *lin-11* mutants are thermophilic due to the failure of AIZ specification (Hobert et al., 1998). Unlike wild-type animals, which spend most of their time in the region of the cultivation temperature (79%, n=56), very few *lin-11(n389)* animals did so (24%, n=52) (Fig. 5). We generated animals carrying the *lin-11-int-3p::lin-11* transgene in the *lin-11* null *(n389)* genetic background and observed a significant rescue of *lin-11* defects. Specifically, 68% of animals (56/82) showed improved thermotaxis response as judged by their ability to migrate to the zone of the cultivation temperature. These data demonstrate that intron 3-mediated expression of *lin-11* is sufficient to rescue the thermotactic behavior defect associated with AIZ interneurons.

Next, we examined animals for changes in their electrotaxis behavior. In the presence of a low-voltage direct-current (DC) electric field, *C. elegans* moves towards the cathode (Gabel et al., 2007; Rezai et al., 2010; Sukul and Croll, 1978). This response is mediated by sensory neurons, including ASJ and ASH (Gabel et al., 2007). When performed in a microfluidic channel device, the assay allows precise measurements of speed, body bend frequency, and reversals when exposed to the electrical stimulus (Rezai et al., 2010). We had earlier

93

Fig. 1. Neuronal expression of *lin-11* intron 3 promoter-driven *GFP* reporter in *C. elegans* and *C. briggsae*. (A,B,E,F,I,K) *C. elegans bhEx98*, (C,D,G,H,J,L) *C. briggsae bhEx200*. (A, B) Several cells are showing GFP fluorescence in a *C. elegans* embryo. In *C. briggsae* (C-D) a single neuron, AVA, is visible at a comparable stage. The fluorescence intensity is higher in cells by 3-fold (E,F) and pretzel (G-H) stages. (I-L) At least seven neurons (AVG, RIF, RIC, AIZ, AVJ, ADL and ADF) are identified in a *C. elegans* L4 stage larva and only one (AVA) in *C. briggsae*. Scale bars are shown (panel B for A-H and panel K for I-L).

reported defects in the electrotactic response of *lin-11(n389)* animals (Salam et al., 2013). *lin-11* animals either move very slowly in the channel or do not move at all (Fig. 6). We found that rescued lines for *lin-11* intron 3 (*bhEx176*) had significantly improved electrotactic movement, in that all animals (n=11) showed robust responses to the stimulus, with an average speed of 170 μm/s (Fig. 6). These results show that intron 3 of *lin-11* contains important regulatory sequences for promoting *lin-11* expression and the specification of the fates of amphid neurons.

Similar to those for intron 3, phenotypic rescue experiments were also carried out for intron 7. Examination of three different stable lines (*lin-11-int7p::lin-11cDNA*, pGLC87; see Section 2) revealed no rescue of the AVG specification defect. Specifically, *odr-2::CFP* fluorescence was absent in AVG in all animals examined (n=41 for *bhEx177*, n=38 for *bhEx203* and n=47 for *bhEx204*). We conclude that while intron 7 drives *GFP* reporter expression in AVG (Fig. 3), it does not appear to be sufficient to promote AVG fate specification.

3.3. Analysis of C. briggsae *introns reveals conservation and divergence in enhancer function*

To understand the evolutionary changes in the regulation of *lin-11* in amphid neurons, we examined corresponding intronic sequences in *C. briggsae*. Transgenic *C. briggsae* animals carrying the full-length *Cbr-lin-11* intron 7-driven *GFP* reporter, *Cbr-lin-11-int7p::GFP* (pGLC60), were generated. The *C. briggsae lin-11-int7p::GFP* animals showed an expression pattern identical to that of *C. elegans lin-11-int7p::GFP* (Fig. 3). Briefly, GFP fluorescence was detected in two neurons in post-gastrulating embryos (Fig. 3G and H). Based on the similar morphology and position of neurons in *C. elegans* larvae, one of these was identified as AVG (Fig. 3J and L). The expression in AVG persisted in adults. We conclude that intron 7 enhancer-mediated expression of *lin-11* is conserved in both nematode species.

In contrast to intron 7 results, GFP fluorescence in *C. briggsae bhEx197* transgenic animals carrying intron 3-driven GFP reporter, *Cbr-lin-11-int3p::GFP* (pGLC61), was almost always present in a

94
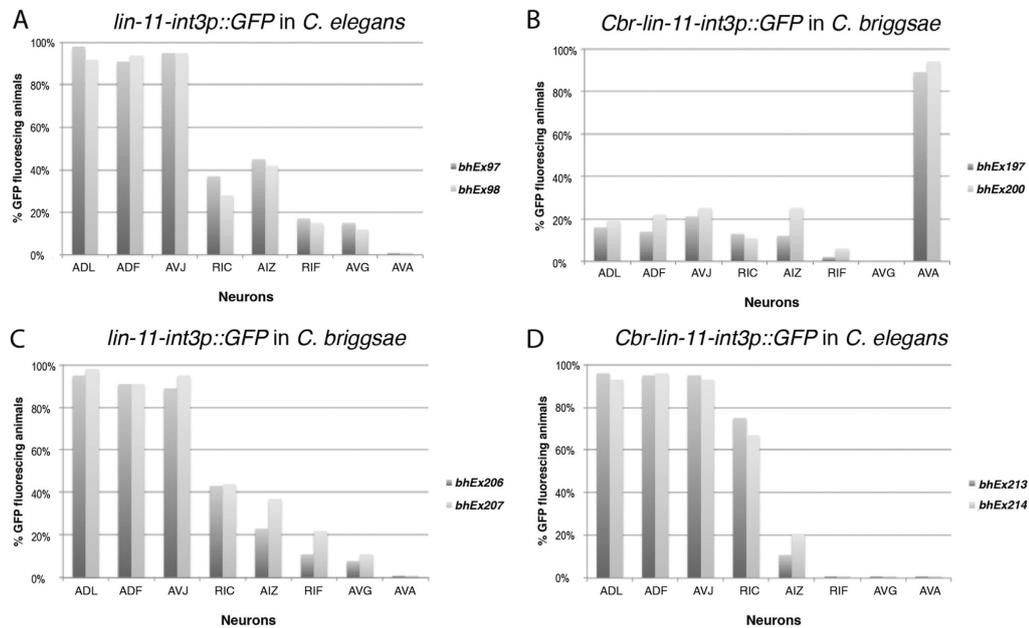
Fig. 2. Quantification of *lin-11* intron 3-driven *GFP* reporter fluorescing neurons in *C. elegans* and *C. briggsae*. Two transgenic strains were examined for each reporter construct. (A) GFP fluorescence in *C. elegans* is predominantly seen in ADL, ADF and AVJ, followed by RIC and AIZ (n=41 for *bhEx97* and 34 for *bhEx98*). Two other neurons, AVG and RIF, are observed in very few animals. (B) The corresponding intronic region from *C. briggsae* activates *GFP* transgene expression in *C. briggsae* in a single neuron, AVA (n=45 for *bhEx197* and 36 for *bhEx200*). Occasionally animals show expression in other amphid neurons as well. (C, D) expression patterns are qualitatively similar to that in A (n=53 for *bhEx206* and 45 for *bhEx207*, n=46 for *bhEx213* and 56 for *bhEx214*).

single pair of interneurons, namely the AVA neurons (Fig. 1J). This pattern was identical in another independently generated stable line (*bhEx200*; Fig. 2B). Neuronal quantification revealed that while in most animals (83%, n=81) expression was restricted to AVA, a small fraction (roughly 15%) also showed expression in additional neurons (Fig. 2B). We also performed temporal expression analysis, which revealed that fluorescence in AVA started in early embryos prior to gastrulation (Fig. 1C and D) and continued during later stages (see Fig. 1G and H for a pretzel-stage animal). The fluorescence intensity increased in larvae (Fig. 1J and L), and also persisted in adults (data not shown). Considering that in *C. elegans*, intron 3 promotes reporter gene expression in seven amphid neurons (Fig. 2A), these results show that the intron 3-mediated neuronal regulation of *lin-11* has diverged between the two species.

3.4. Cross-species genetic analysis reveals functional divergence for cis-encoded information in the intron 3 enhancer

The differences observed between *C. elegans* and *C. briggsae* for the intron 3 reporters may result from changes in the intronic sequences (*cis*) itself, with changes in TF binding sites, from differences in the expression of *trans*-acting factors (expression of key TFs), or both (Carroll, 2000; Gordon and Ruvinsky, 2012; Lynch and Wagner, 2008). To begin to develop some insight into the factors that mediated this evolutionary change, reciprocal transformations (swapping) of the regulatory sequences were carried out.

Transgenic *C. elegans* strains were generated carrying a *C. briggsae lin-11-int3p::GFP* construct. *GFP* expression analysis of two such independent lines (*bhEx213* and *bhEx214*) revealed bright fluorescence in ADL, ADF, AVJ and RIC (henceforth termed '*lin-11* primary' neurons) (Figs. 2D and 7A and C). The three neurons (AIZ, RIF and AVG) that were observed earlier at a lower frequency in *C. elegans* lines

expressing *GFP* under the control of endogenous intron 3 promoter (*lin-11-int3p::GFP*) (Figs. 1I and K and 2D) were not visible in any of the *bhEx213* and *bhEx214* animals (Fig. 2C). This pattern, combined with a lack of expression in AVA, allows us to conclude that the *Cbr-lin-11-intron 3* enhancer contains elements to promote expression in *C. elegans lin-11* primary neurons. These data suggest an evolutionary change in *trans*-acting factors in *C. briggsae* to regulate intron 3-mediated *lin-11* expression in amphid neurons.

If the above model of the evolution of *trans*-acting factors holds true, then we should expect that *C. elegans* intron 3 will promote expression in AVA when introduced into *C. briggsae*. To test this, we generated *C. briggsae* strains carrying *C. elegans lin-11-int3p::GFP* sequences (*bhEx206* and *bhEx207*). The earliest reporter expression was detected in several cells in embryos (data not shown). During larval stages, fluorescence was localized to *lin-11* primary neurons (Fig. 7B and D). No fluorescence was observed in AIZ, RIF, or AVG. Two different stable lines showed a similar pattern of expression (Fig. 2C). Such a neuronal pattern cannot be explained solely by *trans*-factor evolution and is most likely caused by changes in both *cis* and *trans* regulation. Thus, *C. briggsae* intron 3 may have acquired regulatory sequences for AVA-specific expression.

3.5. Introns 3 and 7 contain evolutionarily conserved sequences

The results presented thus far demonstrate that introns 3 and 7 play important roles in the specification of amphid neurons in *C. elegans* and *C. briggsae*. To further understand how these non-coding regions regulate *lin-11* expression in *C. elegans*, we searched for evolutionarily conserved sequences that may be recognized by transcription factors and function in a cell-specific manner. It is important to note that *lin-11* coding regions are highly conserved between *C. elegans* and *C. briggsae* and can be easily aligned (75–90% nucleotide
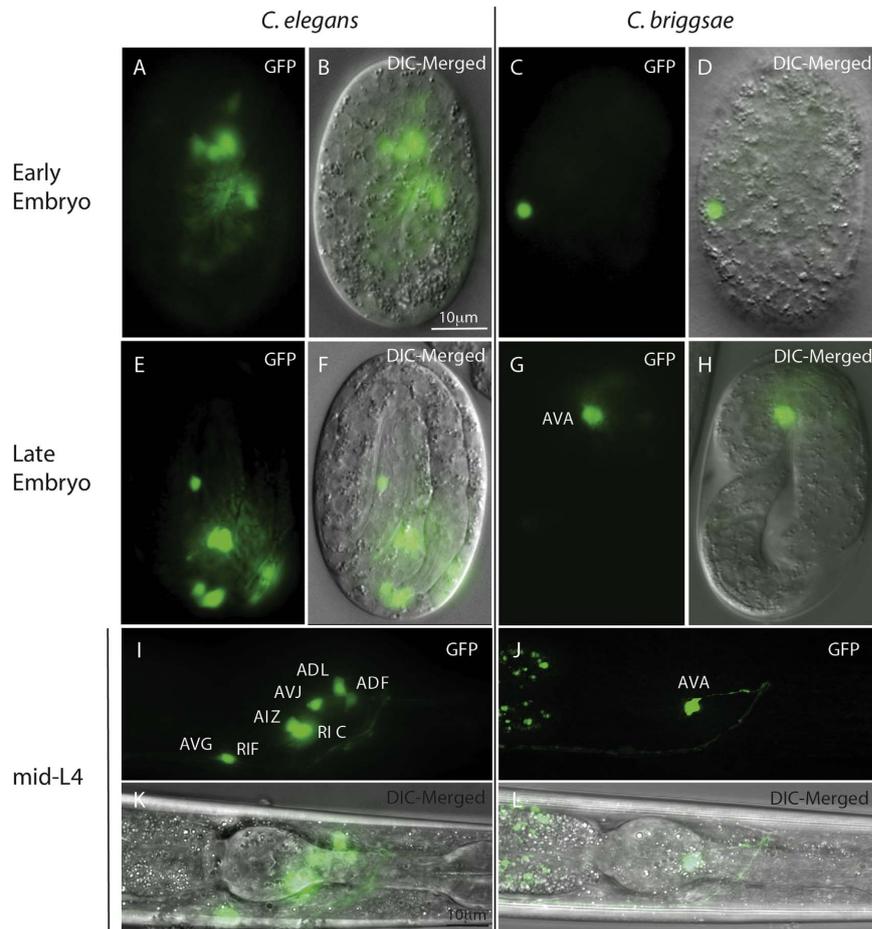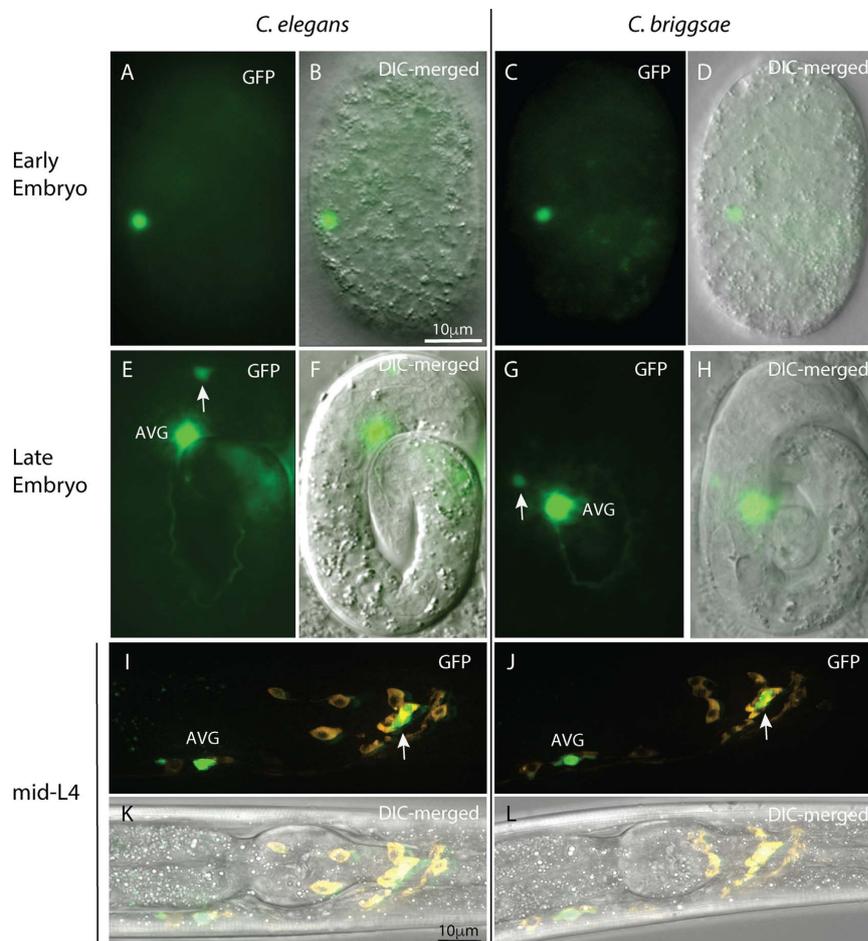
Fig. 3. Neuronal expression of *lin-11* intron 7 promoter-driven *GFP* reporter in *C. elegans* and *C. briggsae*. (A,B,E,F,I,K) *C. elegans bhIs8*, (C,D,G,H,J,K) *C. briggsae bhEx219*. (A-D) GFP fluorescence is clearly visible in AVG in an early stage embryo. (E-H) By 3-fold stage an additional neuron is also visible (marked with arrows). (I-L) Expression pattern during the L4 larval stage remains unchanged. *glr-1::DsRed* expressing neurons (orange) are also visible. Arrows mark an additional unknown neuron. Scale bars are shown (panel B for A-H and panel K for I-L).

identity for each exon, using Wormbase and NCBI BLAST tools) (Sharanya et al., 2012).

Intronic sequences from four *Elegans* subgroup species, *C. elegans, C. briggsae, C. remanei* and *C. brenneri,* were evaluated using MussaGL (multi-species sequence analysis; http://mussa.caltech.edu/mussa). MussaGL is an N-way sequence comparison algorithm that integrates similarities among three or more genomes, taking family relationships into consideration (Brown et al., 2002). MussaGL produces orientation-independent maps that highlight domains of similarity between the sequences, indicating the uniqueness of potential regulatory elements (http://mussa.caltech.edu/mussa). This alignment tool has previously been used successfully to identify conserved enhancer elements in many genes, including *lin-11, lin-39* and *ceh-13* (Kuntz et al., 2008; Marri and Gupta, 2009). The MussaGL alignments of *lin-11* intronic sequences revealed several conserved regions in both introns (Fig. 8A). Specifically, three stretches (referred to as C3-1, C3-2, and C3-3) were observed in intron 3 using a threshold of 80% identity within 30-nucleotide sliding windows (see Section 2)

(Fig. 8B). The C3-3 block is the largest conserved region (55 bp), followed by C3-1 (41 bp) and C3-2 (35 bp).

Similar 4-way alignments of intron 7 revealed two closely located conserved regions C7-1 and C7-2 that are 43 bp and 31 bp long, respectively (Fig. 8). Overall, multi-species sequence alignments of both introns identified five blocks of conserved sequences that may play important roles in *lin-11* regulation in neurons. In the next section, we describe experiments that test the *in vivo* functions of sequences corresponding to these regions.

3.6. Dissection of introns 3 and 7 reveals regulatory elements that include conserved regions

The C3-1 block overlaps with a region that was previously shown to activate *GFP* expression in the ADL, ADF, RIC and AIZ neurons (Yamada et al., 2012). To determine whether C3-1 possesses enhancer activity in any of these neurons, transgenic animals were generated carrying a *GFP* reporter under the control of a 382 bp fragment of

96

60
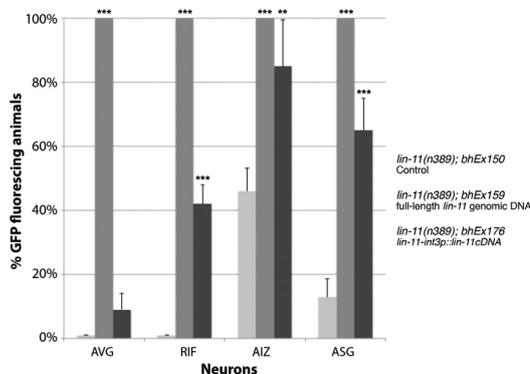
Fig. 4. Rescue of neuronal defects by *lin-11 intron 3-driven lin-11 cDNA* (*lin-11-int3p::lin-11cDNA*) and full length *lin-11* genomic region (pGF50). The *odr-2::CFP* expression in four neurons (AVG, RIF, AIZ and ASG) was quantified. While pGL50 animals were fully rescued for all four neurons (n=42), the *lin-11-int-3p::11cDNA* animals showed RIF, AIZ and ASG in most cases (n=30). Statistically significant responses are marked with stars (Student *t*-test, **: p value < 0.01, ***: p value < 0.001).
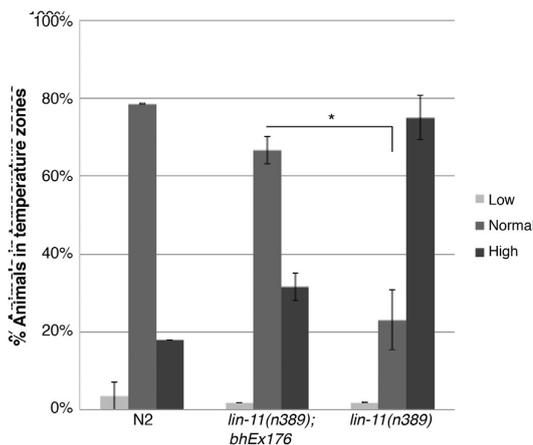
Fig. 5. Rescue of *lin-11* thermotaxis defect by *lin-11 intron 3-driven lin-11 cDNA*. Animals were placed in one of three temperature zones: low (preference for lower than cultivated temperature), normal (preference for cultivated temperature), and high (preference for higher than cultivated temperature). The *lin-11* rescued animals *lin-11(n389); bhEx176* behave similar to wild type (N2) control. The number of animals examined were 58 for N2, 58 for *lin-11(n389)*, and 56 for *lin-11(n389); bhEx176*. Statistically significant responses are marked with stars (Student *t*-test, *: p value < 0.05).

intron 3 (Fig. 9A). *GFP* expression in transgenic animals was observed in a single neuron pair, the RIC neurons (Figs. 9A and 10A, C and F). A similar expression pattern was also observed when the reporter was introduced into *C. briggsae* (Fig. 10B, D and F). As expected, deleting the C3-1 region resulted in a lack of fluorescence in RIC (Fig. 9A) (n=68 for *bhEx228* and n=53 for *bhEx229* in *C. elegans*), demonstrating the functional importance of these sequences in RIC specification.

The intron 7 was dissected into two overlapping fragments, each driving a *GFP* reporter (pGLC65 and pGLC66 in Fig. 9B). pGLC65 carries both conserved regions, i.e., C7-1 and C7-2. Examination of pGLC65 transgenic animals revealed GFP fluorescence in two neurons, namely AVG and RIF. The fluorescence in AVG was comparatively brighter and seen in many animals (55%, n=65), but RIF fluorescence was observed at a much lower frequency (34%, n=65). Another
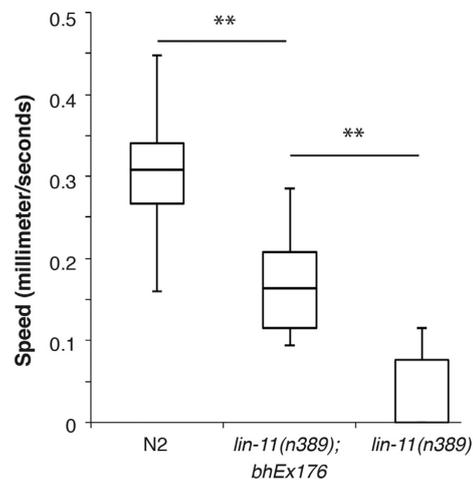
Fig. 6. Box plot showing the electrotaxis speed of wild-type N2, *lin-11* intron 3 rescued *lin-11(n389); bhEx176* and *lin-11(n389)* animals. The line inside each box marks the median of data sample, with lower and upper ends representing the 25th and 75th quartile, respectively. Median is zero for *lin-11(n389)*. The vertical lines represent the spread of data. The speed of rescued animals has improved compared to *lin-11(n389)*, although it is still slower than wild type. Statistically significant responses are marked with stars (Mann-Whitney *U* test, **: p value < 0.01).

unidentified neuron was also faintly visible in a few animals (data not shown). The pGLC66 transgenic animals showed no detectable expression (Fig. 9B), although a 98 bp fragment from this region (pGLC93, Fig. 9B) that overlaps with pGLC65 showed strong fluorescence in RIF neurons in almost all transgenic animals (Fig. 10E and G). Since RIF is not detected in animals carrying full-length intron 7 enhancer, it is plausible that pGLC66 region contains sequences to inhibit such ectopic reporter expression.

In conclusion, our enhancer dissection experiments together with sequence alignments have identified conserved regions in both intron 3 and intron 7 that are involved in regulating *lin-11* expression in subsets of amphid neurons.

3.7. Multiple transcription factors are involved in regulating neuronal expression of lin-11

Identification of functionally important regions in introns 3 and 7, as well as the presence of conserved sequences, led us to investigate the genetic network of *lin-11* regulation. To this end, we searched for putative TF that may act through intronic sequences to regulate *lin-11* transcription. Two different searches were performed, CIS-BP bioinformatics tool (http://cisbp.ccbr.utoronto.ca) for known TF binding sites within the conserved blocks C3-1, C7-1 and C7-2, and modENCODE ChIP-Seq data (http://www.modencode.org) for factors that show significant binding to specific regions of introns. Altogether, we identified 90 unique candidate TFs that include 35 for C3-1, 37 for C7-1, and 46 for C7-2 (Table 1, Fig. 11A) [also see Table 1 in Amon and Gupta (Submitted for publication)]. The TFs from the modENCODE dataset are predominantly for intron 7 (See Table 1 in Amon and Gupta (Submitted for publication)). To investigate the potential roles of a subset of these TFs in *lin-11* regulation, we examined changes in *lin-11::GFP* expression following inactivation of their function using mutation and RNAi approaches.

Of the 35 CIS-BP TF gene candidates for intron 3, we tested the requirements of seven (*ceh-2, ceh-6, crh-1, fax-1, skn-1, sox-3*, and *tbx-2*) that had high consensus binding scores (See Table 1 in Amon and Gupta (Submitted for publication)) and are expressed in head

C. elegans
Cbr-lin-11-int3p::GFP

C. briggsae
lin-11-int3p::GFP
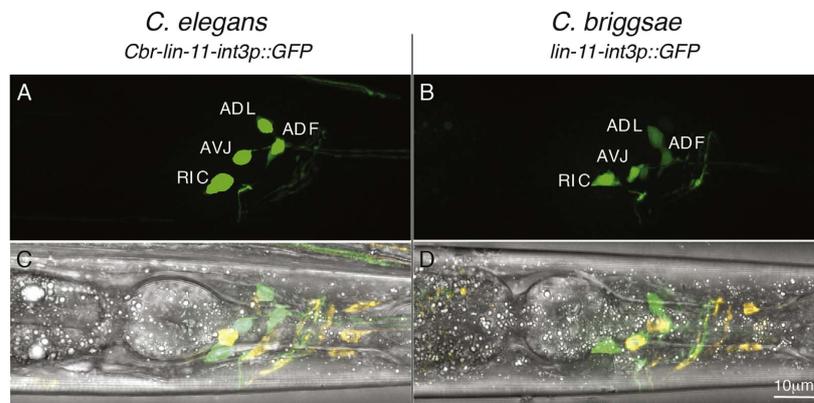
Fig. 7. Analysis of *lin-11* intron 3-driven *GFP* reporter expression in transgenic animals following reciprocal transformations. (A, C) *Cbr-lin-11-int3p::GFP bhEx214* in *C. elegans*, (B, D) *lin-11-int3p::GFP bhEx206* in *C. briggsae*. In all cases four neurons (ADL, ADF, AVJ and RIC) are observed in transgenic animals. See Fig. 2 for quantification of GFP fluorescence. Scale bar is shown in panel D.
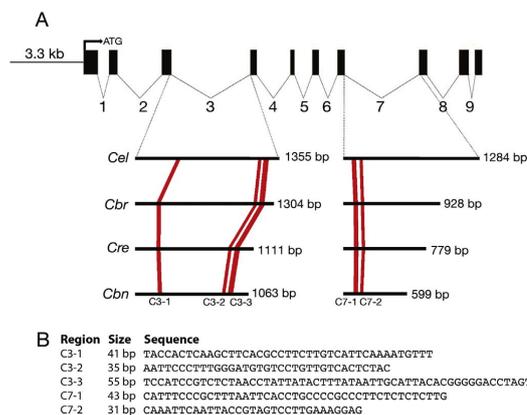
Fig. 8. (A) Schematic drawings of *lin-11* ORF and conserved blocks in introns 3 and 7 revealed by multi-species sequence alignments. *Cel: C. elegans, Cbr: C. briggsae, Cre: C. remanei, Cbn: C. brenneri.* (B) The conserved blocks in intron 3 and intron 7 along with their sequences are shown.
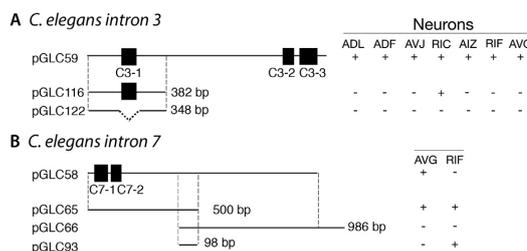
Fig. 9. Dissection of *lin-11* introns 3 and 7 to identify neuronal cell enhancers. The horizontal lines (not drawn to scale) represent intronic fragments used in *GFP* reporter expression studies. The black boxes represent conserved sequence blocks described in Fig. 8. The presence (+) or absence (-) of GFP fluorescence in neurons is indicated. pGLC59 and pGLC58 plasmids that contain entire intron 3 and intron 7 sequences, respectively.

neurons (Aspock et al., 2003; Burglin and Ruvkun, 2001; Kimura et al., 2002; Miyahara et al., 2004; Much et al., 2000; Vidal et al., 2015). These represent members of the homeobox family *(ceh-2* and *ceh-6)*, a CREB homolog *(crh-1)*, a nuclear hormone receptor family *(fax-1)*, a bZIP family/NRF homolog *(skn-1)*, the HMG-box family/mammalian SRY-box family *(sox-3)* and the T-box-like family *(tbx-2)* (Haerty et al., 2008; Reece-Hoyes et al., 2005).

The full-length *lin-11-int3p::GFP* stable line *(bhEx98)* was crossed into each of the TF-mutant strains, and neuronal cells were examined. Quantification of GFP fluorescence revealed that *ceh-2(ch4), fax-1(gm83), sox-3(ok510)* and *tbx-2(ut180)* animals showed no obvious changes in the *lin-11-int3p::GFP* pattern (Fig. 11B). However, the remaining three TF-mutant strains, *skn-1(ok2315), ceh-6(gk679)* and *crh-1(tz2),* showed significantly higher frequencies of some of the GFP-fluorescing amphid neurons compared to the controls (Fig. 11B). In a few cases, we observed ectopically fluorescing neurons whose identities could not be determined. In the case of *skn-1* mutants, this phenotype was particularly striking, as almost all animals showed at least two ectopically fluorescing neurons (94%; n=54). These results suggest that *skn-1, ceh-6* and *crh-1* negatively regulate *lin-11* expression to promote neuronal differentiation in *C. elegans*.

Of the 76 candidate TF genes that are predicted to regulate *lin-11* expression via the intron 7 enhancer (CIS-BP and modENCODE, see Table 1 in Amon and Gupta (Submitted for publication)), we tested the involvement of seven by RNAi (Table 1). We found that knocking down *ces-1* (C2H2 zinc finger/Snail family) by RNAi caused significantly fewer animals to show GFP fluorescence in AVG (40% reduction, n=20, p < 0.01). A similar observation was also made in *ces-1(n703)* animals (50% reduction, n=30, p < 0.01 value, Fig. 12). Interestingly, the predicted *ces-1* binding site (snail-family consensus sequence RCAGGTG) (Metzstein and Horvitz, 1999) is located within the C7-1 region, although it was not identified by CIS-BP. Considering that C7-1 maps within the AVG enhancer of intron 7 (pGLC65, Fig. 9B), CES-1 may act via C7-1 to regulate *lin-11* expression.

4. Discussion

This study focuses on the regulation of a LIM-HOX gene, *lin-11*, in the *C. elegans* nervous system and on evolutionary divergence between *C. elegans* and *C. briggsae*. Although *lin-11* was previously shown to be necessary for the differentiation of amphid neurons (Hobert et al., 1998; Sarafi-Reinach et al., 2001), the mechanism of its spatiotemporal regulation was unknown. Our work has identified enhancer regions
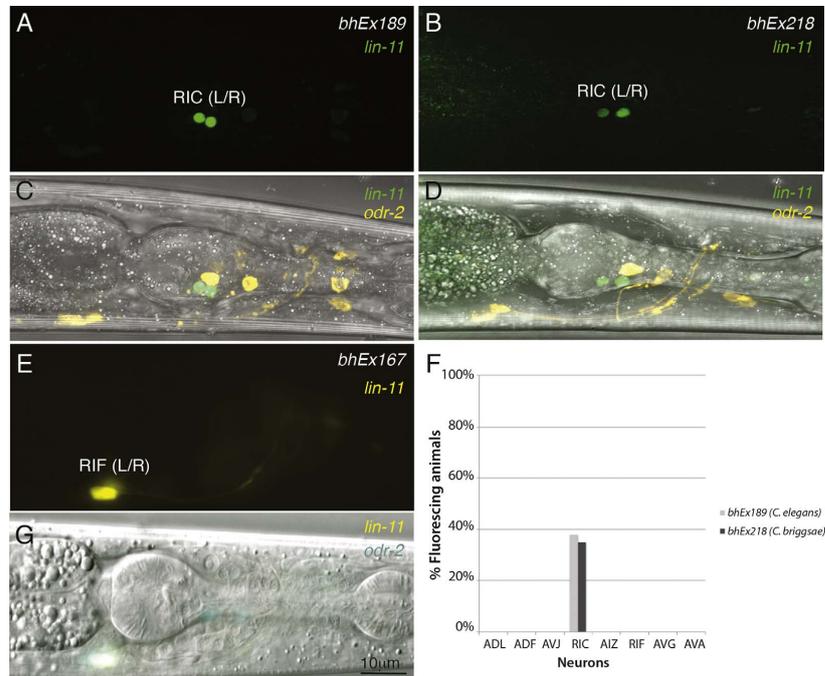
Fig. 10. Enhancer activities of *lin-11* intron sub-clones. (A, C, E, G) *C. elegans* and (B, D) *C. briggsae*. In panels A-D *lin-11* intron 3 reporter (pGLC116, 382 bp intron 3 subclone) fluorescence is in green and *odr-2* in yellow. In E and G, intron 7 reporter (pGLC93, 98 bp intron 7 subclone) is in yellow and *odr-2* reporter in cyan. (A,C) *bhEx189* animals carrying pGLC116 plasmid show fluorescence in a pair of RIC neurons. (B, D) Same pattern is observed in *bhEx218C. briggsae* animals carrying pGLC116. (E, G) The 98 bp region (pGLC93) that overlaps with pGLC65 and pGLC66 activates reporter expression specifically in a pair of RIF neurons. (F) Roughly 40% of *C. elegans* and *C. briggsae* transgenic animals show GFP fluorescence in RIC. Scale bar is shown in panel G.

located within introns 3 and 7 of *lin-11*. The intron 3 enhancer is necessary and sufficient for *lin-11* function in amphid neurons as judged by the expression of molecular markers and rescue of behavioral defects. Interestingly, our comparative studies in *C. briggsae* revealed that the function of intron 3 has diverged, indicating differences in the mechanism of neuronal differentiation between *C. elegans* and *C. briggsae*. Finally, we have identified four transcription factors (s*kn-1, ceh-6, crh-1* and *ces-1*) that play important roles in mediating *lin-11* expression in *C. elegans* amphid neurons.

4.1. Neuronal expression of lin-11 is regulated by enhancers located within intronic regions

Mutations in *lin-11* were previously recovered in genetic screens for genes that are involved in the development of the reproductive system (Ferguson and Horvitz, 1985). Subsequently, *lin-11* was also shown to play roles in neuronal differentiation (Ferguson and Horvitz, 1985; Garriga et al., 1993; Hobert et al., 1998; Sarafi-Reinach et al., 2001). *lin-11* mutants exhibit behavioral defects due to abnormalities in neuronal cell fates. For example, the animals are thermophilic, i.e., they prefer higher temperature due to the failure of AIZ interneurons to

differentiate correctly (Hobert et al., 1998). In addition, they also exhibit an electrosensory defect that is reflected in their inability to sense the direction of direct-current (DC) electric field (Salam et al., 2013). Consistent with these findings, *lin-11* is expressed in several amphid neurons (Hobert et al., 1998; Sarafi-Reinach et al., 2001).

Despite its important role in the nervous system, how *lin-11* is regulated in specific neuronal cell types during development is not understood. To address this question, we studied the enhancers and upstream transcriptional regulators of *lin-11*. Based on previous work by other groups as well as our own, we reasoned that the larger introns of *lin-11* may be important for amphid neuronal expression. In this work, we report that introns 3 and 7 contain regulatory elements that mediate *lin-11* expression in multiple neuronal cells. While the role of intron 3 in neuronal cells (ADL, ADF, AVJ, RIC and AIZ) was previously reported by Yamada and colleagues (Yamada et al., 2012), we have found that intron 7 is specific for a single pioneer neuron, AVG. We observed interesting temporal differences between *lin-11-int3p::GFP* and *lin-11-int7p::GFP* expression. The earliest GFP fluorescence in both these reporter lines was detected in late embryogenesis. In the case of intron 3, the fluorescence continued to become brighter with age and was also present in adults. However, intron 7 showed the

Table 1
Predicted and experimentally validated TF genes that regulate *lin-11* expression in amphid neurons.

Intron	Source	Number of TF genes	Validated TF genes	TF genes affecting reporter expression
Intron 3 (C3-1)	CIS-BP	35	*ceh-2, ceh-6, crh-1, fax-1, skn-1, sox-3, tbx-2*	*ceh-6, crh-1, skn-1*
Intron 7 (C7-1 and C7-2)	CIS-BP	71	*blmp-1, lin-39, mab-5, mec-3*	–
Intron 7	modENCODE	8	*blmp-1, ces-1, ham-1, nfya-1*	*ces-1*

99
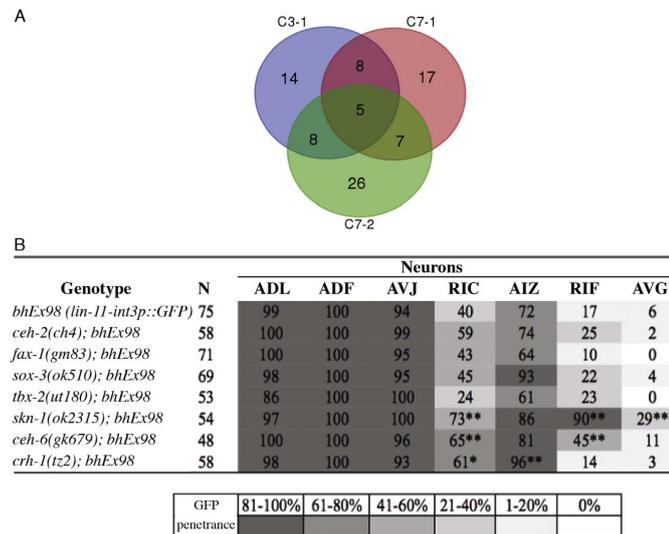
Fig. 11. Transcriptional regulation of *lin-11::GFP* in amphid neurons. (A) Venn diagram showing a partial overlap between CIS-BP predicted TF sites for C3-1, C7-1, and C7-2 conserved blocks of sequences. (B) Analysis of intron 3-driven GFP expression pattern in TF mutant backgrounds. The penetrance of GFP fluorescence is shown in the shades of grey. N represents number of animals for each genotype. Stars (*) mark numbers that are statistically significant compared to the *bhEx98* control (*p < 0.01, **p < 0.001).
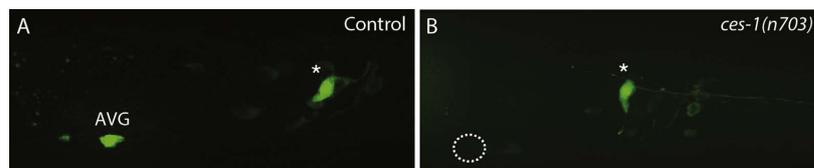
Fig. 12. *ces-1* mutants show defects in *lin-11* expression in the AVG neuron. The *lin-11* intron 7-driven *GFP* reporter in control *bhIs8* transgenic animals (A) is seen in AVG and an unknown neuron (marked with star). In *ces-1(n703); bhEx241* animals (B) GFP fluorescence is absent in AVG (dotted circle area).

opposite trend in GFP intensity, i.e., neuronal cells were brighter early on but became almost completely invisible in adults. Thus, the temporal requirements for these two introns in mediating *lin-11* function in neuronal differentiation appear to differ.

To further extend the reporter gene expression studies, we carried out functional rescue experiments. These results showed that intron 3-driven expression of *lin-11* cDNA can successfully rescue neuronal as well as behavioral defects. Thus, transgenic animals displayed normal electrosensory and thermosensory responses. It has been previously shown that ASH and ASJ are important for electrotactic behavior (Gabel et al., 2007), and although these neurons were not observed in our transgenic lines, it is worth noting that *lin-11::GFP* strains carrying the entire *lin-11* open reading frame are reported to have faint and occasional expression in ASH (Sarafi-Reinach et al., 2001). Thus, it is possible that *lin-11* is expressed in ASH below the detectable limit in our strains. Unlike intron 3, intron 7 failed to show any rescuing activity, suggesting that it may cooperate with other enhancer elements (e.g., intron 3) to ensure proper function of *lin-11* in AVG. Overall, our findings have revealed the involvement of multiple regulatory regions in *lin-11*-mediated neuronal function. Finally, it should be mentioned that Sarafi-Reinach and colleagues (Sarafi-Reinach et al., 2001) had reported at least three additional neurons beyond those described in this study, namely ASH (mechano- and chemosensory), AVE (inter-neuron) and ASG (chemosensory), in transgenic animals expressing a *GFP* reporter under the control of the entire *lin-11* genomic region. It may be that these neurons are also fluorescing in our transgenic lines

but at levels too faint to be detectable. Alternatively, additional regulatory sequences, beyond those in introns 3 and 7, may be required. Future studies are needed to investigate these possibilities in order to gain a deeper understanding of the mechanism of *lin-11* transcriptional regulation in neuronal differentiation processes.

4.2. Evolution of lin-11 regulation in amphid neurons

The work described here provides for the first time a detailed analysis of *lin-11* regulation in the nervous system of two *Caenorhabditis* species. Our results show that both intron 3 and 7 possess enhancer elements, and there are differences in intron 3-mediated *lin-11* expression between the two species.

The *lin-11* promoter-driven reporter gene expression in *C. briggsae* shows interesting similarities and differences from *C. elegans*. While intron 7 functions in a conserved manner in AVG, the role of intron 3 has diverged. We have also uncovered a novel AVA-specific expression in *C. briggsae* intron 3 transgenic lines, providing evidence for an evolutionary change in the regulation of *lin-11* between the two species. To further explore differences in intron 3-mediated *lin-11* regulation, we examined the expression of a *C. briggsae lin-11-int3p::GFP* construct in *C. elegans*. Additionally, we carried out the complementary experiment by transforming *C. briggsae* with *C. elegans lin-11-int3p::GFP*. We expected that if differences in intron 3-driven regulation are caused by *cis* changes, then a single neuron, i.e., AVA, will be visible in *C. elegans* transgenic worms, whereas all
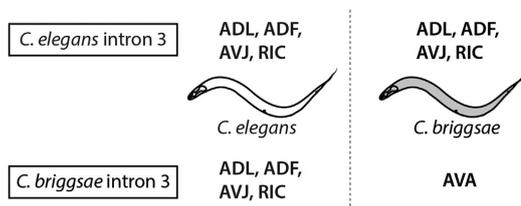
Fig. 13. Summary of *lin-11* intron 3-driven *GFP* reporter expression in *C. elegans* and *C. briggsae* transgenic animals. Except for *Cbr-lin-11-int3p::GFP* construct in *C. briggsae* that activates expression in AVA, in all other cases four *lin-11* primary neurons (ADL, ADF, AVJ, and RIC) are observed.

seven *lin-11*-specific amphid neurons would be visible in *C. briggsae*. However, we found that both types of transgenic worms had expression in only four (*lin-11* primary) neurons (Fig. 13). These findings are best explained by both *cis* and *trans* evolution of *lin-11* regulation. Thus, *Cel-lin-11* intron 3 is likely to contain enhancer sequences that serve as high-affinity binding sites for *trans*-acting factors in *C. elegans* but similar sequences in *Cbr-lin-11* intron 3 may serve as poor sites in *C. briggsae*, possibly due to changes in *trans*-acting factors. While it is unclear how these elements might have survived in the *C. briggsae* genome during evolution, one explanation could be that they have regulatory roles since we did observe a small fraction of *C. briggsae* transgenic animals showing fluorescence in *lin-11* primary neurons (~15%, see Section 3 and Fig. 2B). Although a trivial possibility could also be that these are functionless extant DNA sequences. The expression in AVA neurons in *C. briggsae*, but not in *C. elegans*, may result from an acquisition of *C. briggsae*-specific *cis* elements in *Cbr-lin-11*. This is supported by additional sequence alignments involving two *Caenorhabditis* species, *C. nigoni* and *C. sinica*, which are close relatives of *C. briggsae* (see Fig. 1 in Amon and Gupta (Submitted for publication)). Specifically, we found certain sequence blocks in intron 3 that are less conserved in *C. elegans* (e.g., see block 1, Fig. 1 panel J in Amon and Gupta (Submitted for publication), compare with panels H and I).

Comparative studies of gene regulation in nematodes have identified several examples of *cis* and *trans* evolution of gene expression. Ortiz and colleagues (Ortiz et al., 2006) characterized the regulation of several guanylyl cyclase genes. They showed that while the expression of one of these, *gcy-4*, in *C. elegans* is biased towards the ASE(R) chemosensory neuron, its ortholog in *C. briggsae* is bilaterally expressed in both ASE neurons (L/R). Cross-species experiments revealed that such a difference is likely to be caused by an evolutionary change in *trans*-acting factors.

Two other examples from nematodes provide evidence for both *trans* and *cis* evolution. One of these involves the regulation of neuronal and non-neuronal genes from five different *Caenorhabditis* species (Barriere and Ruvinsky, 2014). A majority of these genes (*unc-25, unc-46, unc-47, oig-1* and *acr-14*) are expressed in GABAergic neurons. Sequence comparisons showed that while the coding regions of all these genes are highly conserved, they exhibit significant variations in *cis* regulatory sequences. Nonetheless, expression patterns of orthologs were found to be very similar. More detailed analysis of transgenic lines revealed subtle changes in the types of neurons that showed reporter expression, suggesting divergence in gene regulation, possibly both at *cis* and *trans* levels. The other study involves the Ovo family of zinc finger transcription factors, specifically LIN-48 (Johnson et al., 2001; Wang et al., 2004). Although *lin-48* is conserved in *C. briggsae*, only in *C. elegans* does it act in excretory duct cell formation, for which it is regulated by CES-2 (zinc finger). By contrast, *lin-48* functions in the hindgut development in both these species. The analysis of the *lin-48* hindgut enhancer revealed binding sites for EGL-38 (PAX 2/5/8), but the relative importance of these sites has changed.

Studies on regulatory evolution in other eukaryotes have also provided evidence for similar changes in the mechanism of gene expression. In one case, researchers compared two different strains of a single species, the yeast *Saccharomyces cerevisiae* (Sung et al., 2009). They analyzed gene expression in BY (laboratory) and RM (wild) strains of *S. cerevisiae* and found that 232 genes showed divergent expression due to both *cis* and *trans* evolutionary changes. It was shown that while *cis* changes are important for expression differences, *trans* changes contributing to similar variations occurred roughly three times as frequently.

Formation of dorsal appendages (DAs) in *Drosophila* fruit fly species has also involved *cis-trans* evolution. Soma-dependent modulations contribute to the divergence of *rhomboid (rho)* expression during evolution of *Drosophila* eggshell morphology (Nakamura et al., 2007). Dissection of the *rho* enhancer, a gene that is involved in the formation of DAs in *D. melanogaster* (*Dmel*) and *D. virilis* (*Dvir*), showed *trans* regulatory changes as revealed by the *rho* activation patterns in heterologous species. Interestingly, the authors also found evidence for *cis* changes that contributed to a somewhat broader domain of *rho* expression (caused by extra rows of cells) in *Dvir rho*-carrying *D. melanogaster* transgenic lines. Yet another example of *cis-trans* evolution is the wing spots in the *Drosophila* subgroup (Werner et al., 2010). Wings in *D. guttifera* fruit flies contain 16 vein-associated spots and four inter-vein shades that are caused by activation of the *yellow* enhancer (*vein spot*) by *Wingless*. The *vein spot* from *D. deflecta*, a closely related species, directs a different pattern but when introduced in *D. guttifera*, it regenerated the endogenous spots, consistent with the *trans* evolution of *Wingless* function. In addition, *vein spot* from *guttifera* and *deflecta* also appeared to have undergone *cis* evolution, given that their enhancer activities appear to differ from the corresponding region in *D. melanogaster*. Overall, these studies, combined with our findings on species-specific changes in *lin-11* regulation in nematodes, demonstrate that a combination of *cis* and *trans* evolution often drives changes in gene expression, and might have played a major role in producing developmental diversity in eukaryotes.

4.3. Conserved enhancer elements promote lin-11 expression in subsets of amphid neurons

We dissected introns 3 and 7 of *lin-11* in order to identify the regulatory elements that may be important for the specification of amphid neurons. A fragment of intron 3, roughly one-third of its length (382 bp), efficiently activated reporter gene expression in a single neuron, RIC, suggesting that intron 3 possesses multiple distinct enhancers that function in a neuron-specific manner and respond to different transcription factors to promote cell differentiation. In the case of intron 7, we uncovered at least two distinct enhancers, one of which is located within a 98 bp region that appears to be involved in activating *lin-11* expression specifically in RIF. The AVG enhancer resides in a 598 bp stretch that contains the C7-1 (34 bp) conserved block (see below). It is important to note that we have not specifically tested the enhancer activities of conserved sequences but rather fragments that contain these sequences. Therefore, it is possible that certain non-conserved sequences, in addition to those identified here, may also be involved in promoting *lin-11* neuronal expression.

The presence of regulatory elements within the two introns led us to further dissect the mechanism of the *lin-11* transcriptional network. A four-way sequence comparison involving *C. elegans*, *C. briggsae*, *C. remanei* and *C. brenneri* revealed three conserved nucleotide blocks within intron 3 (C3-1, C3-2, and C3-3) and two within intron 7 (C7-1 and C7-2). These blocks show 80% identities among *Caenorhabditis* species, raising the possibility that they may be functionally important. In agreement with this, we found that the RIC enhancer in *C. elegans* maps within the C3-1 block of intron 3. This was further confirmed by the analysis of transgenic animals in which *GFP* reporter expression
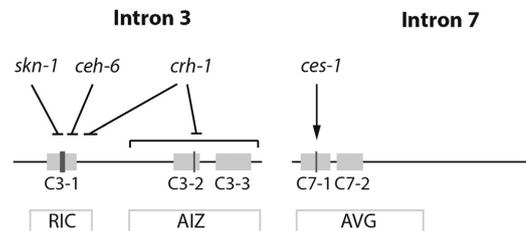
Fig. 14. A model for transcriptional regulation of *lin-11* in the *C. elegans* nervous system. The horizontal lines represent introns and solid grey boxes conserved blocks of sequences. The vertical lines inside boxes mark the locations of consensus binding sites for SKN-1, CEH-6, CRH-1, and CES-1. The approximate enhancer regions for RIC, AIZ and AVG neurons are also shown. Based on enhancer assays, SKN-1, CEH-6, and CRH-1 may act on a cluster of sites within C3-1 to negatively regulate *lin-11* in RIC differentiation. In addition, CRH-1 might recognize sequences outside of the C3-1 block, such as a C3-2 site, to affect expression in AIZ. The *lin-11* transcription in AVG is promoted by CES-1, possibly through a C7-1 site within the intron 7 enhancer.

correlated with the presence or absence of this sequence. For intron 7, the conserved blocks C7-1 and C7-2 are located within the AVG enhancer region. More experiments will need to be done to determine whether one or both of these blocks, or neither of these, are involved in AVG differentiation. The functional relevance of conserved intronic sequence blocks from other *Caenorhabditis* species is also an area for future investigation (see Fig. 1 in Amon and Gupta (Submitted for publication) for additional data).

4.4. Transcriptional regulatory network of lin-11 in neuronal cells

To understand the transcriptional network of *lin-11* regulation in the nervous system, we used an *in silico* tool (CIS-BP) as well as the modENCODE Chip-Seq database, to search for TF genes that might bind to the enhancer regions within the two introns. These two approaches identified a total of 90 TFs, of which 14 were tested experimentally. The results showed that three TF genes, *skn-1*, *ceh-6*, and *crh-1*, act via intron 3 and that a single gene, *ces-1*, acts via intron 7 to regulate *lin-11* in specific neurons. Fig. 14 summarizes our findings and a model of transcriptional regulation of *lin-11* in amphid neurons, incorporating the results of sequence conservation and reporter assays to test predicted TFs in specific neurons. According to the model, *lin-11* expression is mediated by a combination of enhancers located within introns 3 and 7. Three transcription factors (SKN-1, CEH-6 and CRH-1) with potential binding sites within the sequence of C3-1 negatively regulate *lin-11* expression in the interneuron RIC. Why three different TFs are involved in the specification of a single neuron is an open question. Furthermore, whether they interact with each other or function at different times is also unclear. It is worth pointing out that the binding sites for the factors are partially overlapping, which may indicate some competition for access to the sites. Additional experiments are required to determine the precise mechanisms of SKN-1, CEH-6 and CRH-1 functions in RIC development. CRH-1 may also act through other intron 3 sites to promote AIZ specification (Fig. 14). In the case of intron 7, a single positive regulator of *lin-11*, CES-1, has been identified, and it is likely to act via the C7-1 region to promote AVG differentiation (Fig. 14). There is also a distinct 98 bp region that specifically activates reporter expression in RIF neurons in smaller subclones, although transcriptional regulators for this region are yet to be identified. In conclusion, by dissecting the intronic regions of *lin-11*, we have revealed the role of at least three neuron-specific enhancers and a network of four TFs that are important for neuronal differentiation. These results form the basis of future work to understand how these TFs function to activate cell-specific expression of *lin-11*, as well as the roles of additional enhancers and factors in this regulatory process.

References

Amon, S., Gupta, B.P., 2017. Multi-species alignments of *C. elegans lin-11* intronic sequences and putative transcriptional regulators. Data in Brief. (Submitted for publication).

Aspock, G., Ruvkun, G., Burglin, T.R., 2003. The Caenorhabditis elegans ems class homeobox gene ceh-2 is required for M3 pharynx motoneuron function. Development 130, 3369–3378.

Bargmann, C.I., 2006. Chemosensation in *C. elegans*. In: WormBook (Ed). The *C. elegans* Research Community, WormBook. http://dx.doi.org/10.1895/wormbook.1.123.1, 〈http://www.wormbook.org〉.

Barriere, A., Ruvinsky, I., 2014. Pervasive divergence of transcriptional gene regulation in Caenorhabditis nematodes. PLoS Genet. 10, e1004435.

Brenner, S., 1974. The genetics of *Caenorhabditis elegans*. Genetics 77, 71–94.

Brown, C.T., Rust, A.G., Clarke, P.J., Pan, Z., Schilstra, M.J., De Buysscher, T., Griffin, G., Wold, B.J., Cameron, R.A., Davidson, E.H., Bolouri, H., 2002. New computational approaches for analysis of cis-regulatory networks. Dev. Biol. 246, 86–102.

Burglin, T.R., Ruvkun, G., 2001. Regulation of ectodermal and excretory function by the C. elegans POU homeobox gene ceh-6. Development 128, 779–790.

Calixto, A., Chelur, D., Topalidou, I., Chen, X., Chalfie, M., 2010. Enhanced neuronal RNAi in C. elegans using SID-1. Nat. Methods 7, 554–559.

Carroll, S.B., 2000. Endless forms: the evolution of gene regulation and morphological diversity. Cell 101, 577–580.

Chou, J.H., Bargmann, C.I., Sengupta, P., 2001. The Caenorhabditis elegans odr-2 gene encodes a novel Ly-6-related protein required for olfaction. Genetics 157, 211–224.

Curtiss, J., Heilig, J.S., 1998. DeLIMiting development. Bioessays 20, 58–69.

Dawid, I.B., Breen, J.J., Toyama, R., 1998. LIM domains: multiple roles as adapters and functional modifiers in protein interactions. Trends Genet. 14, 156–162.

Elshatory, Y., Gan, L., 2008. The LIM-homeobox gene Islet-1 is required for the development of restricted forebrain cholinergic neurons. J. Neurosci. 28, 3291–3297.

Ferguson, E.L., Horvitz, H.R., 1985. Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode Caenorhabditis elegans. Genetics 110, 17–72.

Freyd, G., 1991. Molecular Analysis of the Caenorhabditis elegans Cell Lineage Gene lin-11, Biology. Massachusetts Institute of Technology, Boston, 159.

Freyd, G., Kim, S.K., Horvitz, H.R., 1990. Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-11. Nature 344, 876–879.

Gabel, C.V., Gabel, H., Pavlichin, D., Kao, A., Clark, D.A., Samuel, A.D., 2007. Neural circuits mediate electrosensory behavior in Caenorhabditis elegans. J. Neurosci. 27, 7586–7596.

Garriga, G., Desai, C., Horvitz, H.R., 1993. Cell interactions control the direction of outgrowth, branching and fasciculation of the HSN axons of Caenorhabditis elegans. Development 117, 1071–1087.

Gordon, K.L., Ruvinsky, I., 2012. Tempo and mode in evolution of transcriptional regulation. PLoS Genet. 8, e1002432.

Gupta, B.P., Sternberg, P.W., 2002. Tissue-specific regulation of the LIM homeobox gene lin-11 during development of the Caenorhabditis elegans egg-laying system. Dev. Biol. 247, 102–115.

Gupta, B.P., Sternberg, P.W., 2003. The draft genome sequence of the nematode Caenorhabditis briggsae, a companion to C. elegans. Genome Biol. 4, 238.

Gupta, B.P., Johnsen, R., Chen, N., 2007. Genomics and biology of the nematode Caenorhabditis briggsae. In: WormBook (Ed.), The C. elegans Research Community, WormBook. http://dx.doi.org/10.1895/wormbook.1.136.1http://dx.doi.org/, 〈http://www.wormbook.org〉.

Haerty, W., Artieri, C., Khezri, N., Singh, R.S., Gupta, B.P., 2008. Comparative analysis of function and interaction of transcription factors in nematodes: extensive conservation of orthology coupled to rapid sequence evolution. BMC Genom. 9, 399.

Hedgecock, E.M., Russell, R.L., 1975. Normal and mutant thermotaxis in the nematode Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA 72, 4061–4065.

Hirota, J., Mombaerts, P., 2004. The LIM-homeodomain protein Lhx2 is required for complete development of mouse olfactory sensory neurons. Proc. Natl. Acad. Sci. USA 101, 8751–8755.

Hobert, O., 2013. The neuronal genome of *Caenorhabditis elegans*. In: WormBook (Ed.), The *C. elegans* Research Community, WormBook. http://dx.doi.org/10.1895/wormbook.1.161.1, 〈http://www.wormbook.org〉.

Hobert, O., 2016a. A map of terminal regulators of neuronal identity in Caenorhabditis elegans. Wiley Inter. Rev. Dev. Biol. 5, 474–498.

Hobert, O., 2016b. Terminal selectors of neuronal identity. Curr. Top. Dev. Biol. 116, 455–475.

Hobert, O., Westphal, H., 2000. Functions of LIM-homeobox genes. Trends Genet. 16, 75–83.

Hobert, O., Glenwinkel, L., White, J., 2016. Revisiting neuronal cell type classification in Caenorhabditis elegans. Curr. Biol. 26, R1197–R1203.

Hobert, O., D'Alberti, T., Liu, Y., Ruvkun, G., 1998. Control of neural development and function in a thermoregulatory network by the LIM homeobox gene lin-11. J. Neurosci. 18, 2084–2096.

Hou, P.S., Chuang, C.Y., Kao, C.F., Chou, S.J., Stone, L., Ho, H.N., Chien, C.L., Kuo, H.C., 2013. LHX2 regulates the neural differentiation of human embryonic stem cells via transcriptional modulation of PAX6 and CER1. Nucleic Acids Res. 41, 7753–7770.

Johnson, A.D., Fitzsimmons, D., Hagman, J., Chamberlin, H.M., 2001. EGL-38 Pax regulates the ovo-related gene lin-48 during Caenorhabditis elegans organ development. Development 128, 2857–2865.

Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., Edlund, T., 1990. Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. Nature 344, 879–882.

Kimura, Y., Corcoran, E.E., Eto, K., Gengyo-Ando, K., Muramatsu, M.A., Kobayashi, R., Freedman, J.H., Mitani, S., Hagiwara, M., Means, A.R., Tokumitsu, H., 2002. A CaMK cascade activates CRE-mediated transcription in neurons of Caenorhabditis elegans. EMBO Rep. 3, 962–966.

Kolterud, A., Wandzioch, E., Carlsson, L., 2004. Lhx2 is expressed in the septum transversum mesenchyme that becomes an integral part of the liver and the formation of these cells is independent of functional Lhx2. Gene Exp. Patterns 4, 521–528.

Kuntz, S.G., Schwarz, E.M., DeModena, J.A., De Buysscher, T., Trout, D., Shizuya, H., Sternberg, P.W., Wold, B.J., 2008. Multigenome DNA sequence conservation identifies Hox cis-regulatory elements. Genome Res. 18, 1955–1968.

Liang, X., Song, M.R., Xu, Z., Lanuza, G.M., Liu, Y., Zhuang, T., Chen, Y., Pfaff, S.L., Evans, S.M., Sun, Y., 2011. Isl1 is required for multiple aspects of motor neuron development. Mol. Cell Neurosci. 47, 215–222.

Lilly, B., O'Keefe, D.D., Thomas, J.B., Botas, J., 1999. The LIM homeodomain protein dLim1 defines a subclass of neurons within the embryonic ventral nerve cord of Drosophila. Mech. Dev. 88, 195–205.

Lynch, V.J., Wagner, G.P., 2008. Resurrecting the role of transcription factor change in developmental evolution. Evolution 62, 2131–2154.

Maduro, M., Pilgrim, D., 1995. Identification and cloning of unc-119, a gene expressed in the *Caenorhabditis elegans* nervous system. Genetics 141, 977–988.

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

Marri, S., Gupta, B.P., 2009. Dissection of lin-11 enhancer regions in Caenorhabditis elegans and other nematodes. Dev. Biol. 325, 402–411.

Mello, C.C., Kramer, J.M., Stinchcomb, D., Ambros, V., 1991. Efficient gene transfer in *C.elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 10, 3959–3970.

Metzstein, M.M., Horvitz, H.R., 1999. The C. elegans cell death specification gene ces-1 encodes a snail family zinc finger protein. Mol. Cell 4, 309–319.

Miyahara, K., Suzuki, N., Ishihara, T., Tsuchiya, E., Katsura, I., 2004. TBX2/TBX3 transcriptional factor homologue controls olfactory adaptation in Caenorhabditis elegans. J. Neurobiol. 58, 392–402.

Mori, I., Ohshima, Y., 1995. Neural regulation of thermotaxis in Caenorhabditis elegans. Nature 376, 344–348.

Much, J.W., Slade, D.J., Klampert, K., Garriga, G., Wightman, B., 2000. The fax-1 nuclear hormone receptor regulates axon pathfinding and neurotransmitter expression. Development 127, 703–712.

Nakamura, Y., Kagesawa, T., Nishikawa, M., Hayashi, Y., Kobayashi, S., Niimi, T., Matsuno, K., 2007. Soma-dependent modulations contribute to divergence of rhomboid expression during evolution of Drosophila eggshell morphology. Development 134, 1529–1537.

Newman, A.P., Acton, G.Z., Hartwieg, E., Horvitz, H.R., Sternberg, P.W., 1999. The lin-11 LIM domain transcription factor is necessary for morphogenesis of C. elegans uterine cells. Development 126, 5319–5326.

Ortiz, C.O., Etchberger, J.F., Posy, S.L., Frokjaer-Jensen, C., Lockery, S., Honig, B., Hobert, O., 2006. Searching for neuronal left/right asymmetry: genomewide analysis of nematode receptor-type guanylyl cyclases. Genetics 173, 131–149.

Pfaff, S.L., Mendelsohn, M., Stewart, C.L., Edlund, T., Jessell, T.M., 1996. Requirement for LIM homeobox gene Isl1 in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. Cell 84, 309–320.

Pinto do, O.P., Richter, K., Carlsson, L., 2002. Hematopoietic progenitor/stem cells immortalized by Lhx2 generate functional hematopoietic cells in vivo. Blood 99, 3939–3946.

Reece-Hoyes, J.S., Deplancke, B., Shingles, J., Grove, C.A., Hope, I.A., Walhout, A.J., 2005. A compendium of Caenorhabditis elegans regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol. 6, R110.

Rezai, P., Siddiqui, A., Selvaganapathy, P.R., Gupta, B.P., 2010. Electrotaxis of Caenorhabditis elegans in a microfluidic environment. Lab Chip 10, 220–226.

Salam, S., Ansari, A., Amon, S., Rezai, P., Selvaganapathy, P.R., Mishra, R.K., Gupta, B.P., 2013. A microfluidic phenotype analysis system reveals function of sensory and dopaminergic neuron signaling in C. elegans electrotactic swimming behavior. Worm 2, e24558.

Sarafi-Reinach, T.R., Melkman, T., Hobert, O., Sengupta, P., 2001. The lin-11 LIM homeobox gene specifies olfactory and chemosensory neuron fates in C. elegans. Development 128, 3269–3281.

Schmid, C., Schwarz, V., Hutter, H., 2006. AST-1, a novel ETS-box transcription factor, controls axon guidance and pharynx development in C. elegans. Dev. Biol. 293, 403–413.

Sharanya, D., Thillainathan, B., Marri, S., Bojanala, N., Taylor, J., Flibotte, S., Moerman, D.G., Waterston, R.H., Gupta, B.P., 2012. Genetic control of vulval development in Caenorhabditis briggsae. G3 2, 1625–1641.

Sharma, K., Sheng, H.Z., Lettieri, K., Li, H., Karavanov, A., Potter, S., Westphal, H., Pfaff, S.L., 1998. LIM homeodomain factors Lhx3 and Lhx4 assign subtype identities for motor neurons. Cell 95, 817–828.

Sheng, H.Z., Zhadanov, A.B., Mosinger, B., Jr., Fujii, T., Bertuzzi, S., Grinberg, A., Lee, E.J., Huang, S.P., Mahon, K.A., Westphal, H., 1996. Specification of pituitary cell lineages by the LIM homeobox gene Lhx3. Science 272, 1004–1007.

Sukul, N.C., Croll, N.A., 1978. Influence of potential difference and current on the electrotaxis of Caenorhabditis elegans. J. Nematol. 10, 314–317.

Sun, Y., Dykes, I.M., Liang, X., Eng, S.R., Evans, S.M., Turner, E.E., 2008. A central role for Islet1 in sensory neuron development linking sensory and spinal gene regulatory programs. Nat. Neurosci. 11, 1283–1293.

Sung, H.M., Wang, T.Y., Wang, D., Huang, Y.S., Wu, J.P., Tsai, H.K., Tzeng, J., Huang, C.J., Lee, Y.C., Yang, P., Hsu, J., Chang, T., Cho, C.Y., Weng, L.C., Lee, T.C., Chang, T.H., Li, W.H., Shih, M.C., 2009. Roles of trans and cis variation in yeast intraspecies evolution of gene expression. Mol. Biol. Evol. 26, 2533–2538.

Sze, J.Y., Ruvkun, G., 2003. Activity of the Caenorhabditis elegans UNC-86 POU transcription factor modulates olfactory sensitivity. Proc. Natl. Acad. Sci. USA 100, 9560–9565.

Timmons, L., Fire, A., 1998. Specific interference by ingested dsRNA. Nature 395, 854.

Vidal, B., Santella, A., Serrano-Saiz, E., Bao, Z., Chuang, C.F., Hobert, O., 2015. C. elegans SoxB genes are dispensable for embryonic neurogenesis but required for terminal differentiation of specific neuron types. Development 142, 2464–2477.

Wang, X., Greenberg, J.F., Chamberlin, H.M., 2004. Evolution of regulatory elements producing a conserved gene expression pattern in Caenorhabditis. Evol. Dev. 6, 237–245.

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

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., Zheng, H., Goity, A., van Bakel, H., Lozano, J.C., Galli, M., Lewsey, M.G., Huang, E., Mukherjee, T., Chen, X., Reece-Hoyes, J.S., Govindarajan, S., Shaulsky, G., Walhout, A.J., Bouget, F.Y., Ratsch, G., Larrondo, L.F., Ecker, J.R., Hughes, T.R., 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158, 1431–1443.

Werner, T., Koshikawa, S., Williams, T.M., Carroll, S.B., 2010. Generation of a novel wing colour pattern by the Wingless morphogen. Nature 464, 1143–1148.

White, J.G., Southgate, E., Thomson, J.N., Brenner, S., 1986. The structure of the nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. Lond. B Biol. Sci. 314, 1–340.

Yamada, K., Tsuchiya, J., Iino, Y., 2012. Mutations in the pqe-1 gene enhance transgene expression in Caenorhabditis elegans. G3 2, 741–751.

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Multi-species alignments of *C. elegans lin-11* intronic sequences and putative transcriptional regulators

CrossMark

Siavash Amon, Bhagwati P. Gupta*

Department of Biology, McMaster University, Hamilton ON L8S-4K1, Canada

ARTICLE INFO

ABSTRACT

This data article contains multi-species alignments of the regulatory region of *C. elegans* LIM-HOX gene *lin-11* and lists of transcription factors that are predicted to bind to *lin-11* enhancers and regulate expression in amphid neurons. For further details and experimental findings please refer to the article by Amon and Gupta in Developmental Biology (S. Amon, B.P. Gupta, 2017) [1].

© Published by Elsevier Inc.

Specifications Table

Subject area	*Biology*
More specific subject area	*Evolutionary Developmental Genetics*
Type of data	*Figure and Table*
How data was acquired	*Using software and web tools*
Data format	*Analyzed*
Experimental factors	*None*

68

Experimental features	*Genomic sequences were aligned and transcription factor binding sites were predicted*
Data source location	*Hamilton, Canada*
Data accessibility	*Data is with this article*

Value of the data

- *C. elegans lin-11* intron 3 possesses conserved sequence blocks that map within functionally defined neuronal enhancers.
- *C. briggsae lin-11* intron 3 possesses some sequences that are conserved in *C. nigoni* and *C. sinica* (two closest relatives of *C. briggsae*), but not in *C. elegans*.
- *In silico* analysis revealed putative transcription factor binding sites within conserved blocks of *C. elegans lin-11*. The functional relevance of these sites can be investigated to understand transcriptional regulation of *lin-11* in neuronal cell differentiation.

1. Data

Six *Caenorhabditis* species were used to perform sequence alignments of *lin-11* intron 3. These are *C. briggsae, C. sinica, C. nigoni, C. remanei, C. brenneri,* and *C. elegans*. MussaGL program (http://woldlab. caltech.edu/cgi-bin/mussa) was used at 70% and 80% window thresholds. Multiple alignments were carried out that included *C. briggsae* and *C. nigoni* (Fig. 1). In general, conservation decreases as the number of species and alignment threshold are increased. Four-way alignments reveal six distinct conserved blocks at 70% threshold. Some of these blocks are part of larger stretches in 2-way and 3-way alignments. At 80% threshold block 2 lacks conservation when either one of the *C. remanei, C. brenneri* and *C. elegans* species are included. Additionally, block 1 is lost in the case of *C. elegans*. Of the three sequence blocks described in the accompanied article [1], namely, C3-1, C3-2 and C3-3 that are conserved between *C. elegans, C. brenneri, C. remanei* and *C. briggsae*, block 3 corresponds to C3-1, block 5 to C3-2, and block 6 to C3-3 (Fig. 1).

We used a computational tool CIS-BP (http://cisbp.ccbr.utoronto.ca/TFTools.php) [2] to search for transcription factors (TFs) that may bind to conserved blocks in introns 3 (C3-1) and 7 (C7-1 and C7-2) and potentially *lin-11* expression in neurons. A total of 35 TF genes were identified for C3-1, 37 for C7-1, and 46 for C7-2 (Table 1). In addition, we searched for modENCODE dataset (http://www. modencode.org) and found eight TFs that bind to intron 7 sequences Table 1).

2. Experimental design, materials and methods

The *lin-11* intronic sequences from *Caenorhabditis* species were aligned using MussaGL (multi-species sequence analysis, version 1.1.0 for Mac OS X), an N-way sequence alignment software that was developed by Wold lab (Caltech, USA). The conservation threshold was set at 70% (21 per 30-nucleotide sliding window) and 80% (27 per 30-nucleotide sliding window).

To identify the putative TF genes for *C. elegans* introns 3 and 7, we used the CIS-BP database software. The setting included motif model 'PWMs-LogOdds' and Threshold 8. According to the website, this motif model option scores each position in each sequence with all position weight matrices, using a standard log odds scoring method. For more details see the help page on the website (http://cisbp.ccbr.utoronto.ca/help.html).
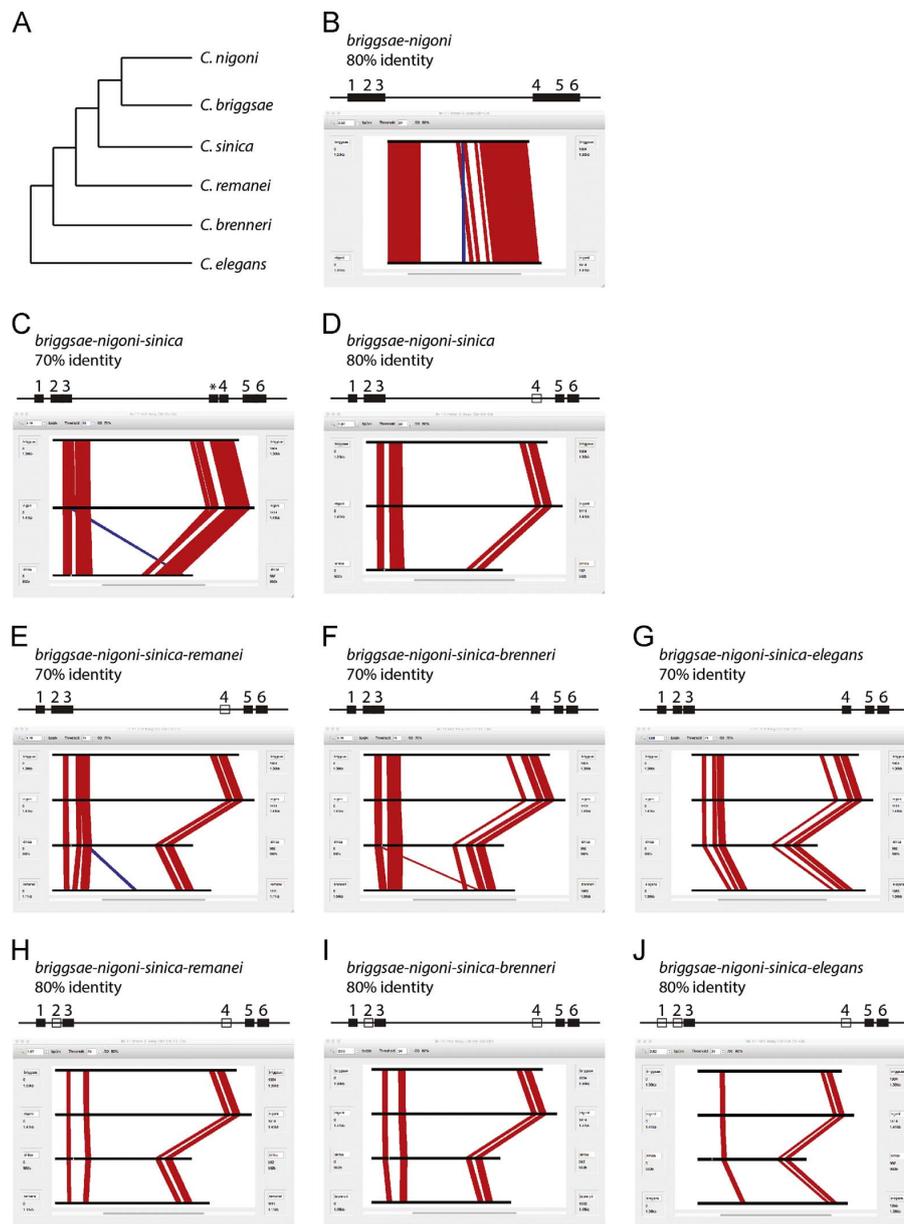
Fig. 1. Multi-species sequence alignments of *lin-11* intron 3 using MussaGL. (A) Phylogenetic relationship of a subset of *Caenorhabditis* species. *C. nigoni* is the closest known relative of *C. briggsae*, followed by *C. sinica*. (B–J) Screenshots of aligned sequences. Species names and alignment threshold are shown above the each screenshot. Alignments in the same orientation are shown in red whereas those in the opposite orientation are in blue. At least six distinct conserved blocks (solid black rectangles on the horizontal line, drawn above the each screenshot) are observed in intron 3, termed 1 to 6 (each roughly 30–50 nucleotides long), that appear to form two distinct clusters. The blocks are not always distinct and some are part of larger conserved regions in 2-way and 3-way alignments. An open rectangle indicates an absence of a conserved block. Star in C marks an extra block that is not present in any other alignment.

Conflicts of interest

None.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.03.027.

References

[1] S. Amon and B.P. Gupta, Intron-specific patterns of divergence of *lin-11* regulatory function in the *C. elegans* nervous system. Developmental Biology, 2017, 424(1), 90-103.
[2] M.T. Weirauch, et al., Determination and inference of eukaryotic transcription factor sequence specificity, Cell 158 (6) (2014) 1431–1443.

Chapter 4.0

Function and evolutionary conservation of LIN-11 domains in neuronal and reproductive tissue differentiation

4.1 Overview

LIN-11 is composed of two N-terminal LIM domains, a C-terminal DNA-binding home-odomain (HD), and a proline-rich region (PRR) (Figure 4.1) (Bach 2000; Hobert and Westphal 2000). The HD is known to bind DNA in a sequence-specific manner (Cohen et al. 1992; Diaz-Benjumea and Cohen 1993). The LIM domain, containing cysteines and histidines at conserved positions, is roughly 55 amino acids long and together form the double zinc fingers (Figure 1B, C) (Pérez-Alvarado et al. 1994; Schmeichel and Beckerle 1997). The LIM domain is required for mediating protein-protein interactions by inter-acting with certain cofactors, such as LIM domain-binding protein 1 (*Ldb1*) (Bach 2000; Khurana et al. 2002). To date, 30 protein-protein interaction pairs for LIM domains have been identified and described (Bach 2000; Khurana et al. 2002).

Unlike the HD and LIM domains, very little is known about the role of the C-terminus PRR. To understand the requirements of the LIN-11-conserved LIM domain and the PRR, I took a genetic approach to investigating and characterizing their function in the nervous and reproductive tissues, where the function of LIN-11 is required for proper cell differentiation (Hobert et al. 1998; Sarafi-Reinach et al. 2001). To this end,

I constructed several plasmids lacking the LIM domain, used them to rescue the *lin-11* phenotype in both the neuronal and reproductive tissues, and performed a thermotactic behavioral assay to test the rescued animals' nervous system function in a temperature gradient environment. Similarly, I constructed a plasmid lacking the PRR, used this plasmid to rescue the *lin-11* phenotype in the neuronal and reproductive tissues, and performed a thermotactic behavioural assay to test the function of the nervous system of the rescued animals in the temperature gradient environment. I also used *lin-11* orthologs from *D. melanogaster* (*dLim1*) and *M. musculus* (*Lhx1*) to test their functional conservation by examining their abilities to rescue the *lin-11* mutant phenotype. I performed thermotactic behaviour assays to determine whether the rescued animals' nervous systems displayed wildtype behavior in a temperature gradient environment.

Together, the work described in this chapter serves as in-depth analysis of the role that the LIM domain and proline-rich region (PRR) of LIN-11 play in mediating its function in the specification of the nervous and reproductive tissues. The inter-species rescue experiments demonstrate that LIM-HD proteins are evolutionarily and functionally conserved, and that the function of LIN-11 in *C. elegans* can be restored by its orthologous proteins *dLIM1* and *LHX1*, from *D. melanogaster* and mouse, respectively.

4.2 Sequencing of *lin-11* alleles reveals the nature of mutations affecting LIM domain, Homeodomain, and Proline-rich region

For my Master's thesis, I sequenced nine alleles of *lin-11* and reported on the nature of their mutations (see Table 5 in Amon 2012). In brief, four alleles (rh309, sy251, sy533, and ty6) possess intron/exon splicing deformities, and the n382 allele contains a nonsense mutation that affects the LIM2 domain. Two alleles (sy534 and sy634) contain point mutations that affect LIM1 and LIM2, respectively. The n1281 allele has a Tc1 transposon insertion that affects the HD domain, and the n566 is a deletion allele that affects the LIM2 domain. For one of the alleles, ps1, amphid neuronal phenotype was also examined, using *odr-2::CFP* (and more recently, *odr-2::dsRED*) cell fate markers,

and found to be defective (Figures 16 and 17 in Amon 2012). These results, together with the sequencing data, suggest that LIM, HD and PRR are essential for LIN-11 function in specification of neuronal and reproductive tissues.

The following sections describe new experiments to investigate the requirements of the LIM domain and PRR. For this, I constructed rescue plasmids that lack these regions and attempted to rescue the *lin-11* defects.

4.3   LIM domains of LIN-11 are necessary for the development of amphid neurons and reproductive tissues

In *C. elegans*, the function of the conserved LIM domains (LIM1 and LIM2) of LIN-11 in vulva and neuronal differentiation is not well understood. It has been demonstrated in other organisms whose LIM domains are zinc-finger structures, which function as a modular binding interface to mediate protein–protein interactions (Kadrmas and Beckerle 2004). Studies have shown, through the association of LIM domains with their protein-binding partners, that LIM proteins participate in an array of biological processes, including regulation of actin organization in the cytoplasm; neuronal pathfinding; integrin-dependent adhesion; and mediation of communication between the cytosolic and the nuclear compartments (Kadrmas and Beckerle 2004). To better understand the role of the conserved LIM domains of LIN-11 in vulva precursor cell and neuronal cell specification, I created two rescue plasmids, pGLC89 and pGLC90, that encode truncated LIN-11 lacking the LIM domains (LIN-11ΔLIM) (see Chapter 2). Specifically, pGLC89 and pGLC90 contain intron 3 (neuron) and 5'UTR (vulva) tissue-specific promoters in frame with *lin-11* cDNA that lack the LIM domains (Chapter 2, Figures 2.12 and 2.13). Together, these plasmids were co-expressed with the neuronal marker, *odr-2::dsRED*, in an attempt to rescue the *lin-11(n389)* mutant phenotype. Three transgenic extra-chromosomal arrays were generated (*bhEx257*, *bhEx258*, and *bhEx267*) using these three plasmids, and the nervous and reproductive tissues were analyzed and quantified in

rescue animals (Figures 4.2, 4.3 and 4.4; Table 4.1). In all three transgenic extrachromosomal arrays, the LIN-11ΔLIM did not rescue the *lin-11* specific phenotypes in the nervous or reproductive tissues. The transgenic animals exhibited Pvl, Egl, and locomotion defects, similar to *lin-11(n389)*. Previously, *lin-11(n389)* animals were shown to exhibit an egg-laying defective (Egl) phenotype that was linked to abnormalities in vulva development (visible as protruding vulva or Pvl in adults) and vulva-uterine connection (lack of uterine seam cell or utse) (Figure 4.3A).

In the nervous system, the rescued animals had a similar phenotype to the *lin-11(n389)* mutant animals, whereby the *odr-2::dsRED* reporter was rarely expressed in AIZ (~33%, n= 90) and never in ASG (0%, n = 90), AVG (0%, n = 90) and RIF (0%, n = 90) neurons (Figure 4.5B, Table 4.1). It is not surprising that LIN-11ΔLIM did not rescue the neuronal and the reproductive defects in the *lin-11* mutants, because it is known that, in other organisms, LIM domain function is a prerequisite for nuclear localization processes and protein-protein stability of protein complexes (Arber and Caroni 1996; Bach 2000; Kong et al. 1997).

Although the reporter expression in cells allows us to determine the expression of a gene, a more specific assay is required to fully understand the function of the cells. Since the expression of *odr-2::dsRed* was present in the AIZ interneuron in a small population (~34%, n = 90) of rescued animals, and the AIZ interneuron is part of the thermoregulatory network (Figure 1.6), I performed a thermotactic behavioral assay to test the behavioral output of the rescued animals' nervous symptoms in a temperature gradient environment. The majority (~64%, n = 30) of rescued animals (*bhEx257*) displayed thermophilic behavior similar to the *lin-11(n389)* (~63%, n = 30), while ~20% (n = 30) of the rescued animals displayed wildtype behavior by staying within the cultivated zone. The remaining ~16% (n= 30) displayed cryophilic behaviur (Figure 4.6). Together, defective thermotactic behavior and *odr-2::dsRED* quantification in the nervous system suggest that LIN-11ΔLIM does not restore the function of the AIZ neuron, although ~34% (n = 90) of the rescued animals showed *odr-2::dsRED* reporter expression in the AIZ neuron.

In conclusion, LIM domains of LIN-11 are important for its function in the specifications of the neuronal and reproductive tissues; mutations that affect the LIM domains will result in *lin-11* phenotype.

4.4 Deletion of the Proline-rich region of LIN-11 has no impact on the development of neurons and reproductive tissues

Proteins with repetitive short PRRs have been demonstrated to function in many different biological processes. For example, they may be required for the rapid recruitment of several proteins during processes such as initiation of transcription, signaling cascades, and cytoskeletal rearrangements (Liu et al. 2003; Sprenger-Haussels and Weisshaar 2000; Zarrinpar et al. 2003). In vertebrates, the PRR of SOS1 binds to the two SH3 domains of GRB2 and regulates cytoplasmic signaling, such as through the action of Ras protein (Egan et al. 1993; Gale et al. 1993; McDonald et al. 2009; Simon et al. 1993). The PRR is also required for nuclear transport. For example, the Borna disease virus (BDV) must be transported from the cytoplasm into the nucleus for replication and transcription to occur (Kobayashi et al. 2001; Shoya et al. 1998). This is accomplished via the nuclear localization signals (NLS) on BDV, which are present in both *N*-terminal and *C*-terminal regions of the protein (Kobayashi et al. 2001; Shoya et al. 1998). These two NLS can function independently, and both contain several proline residues that are essential for nuclear localization to occur (Kobayashi et al. 2001; Shoya et al. 1998).

In *C. elegans*, the function of the LIN-11 PRR in development is currently unclear. I constructed a truncated LIN-11 plasmid that lacks the PRR (LIN-11ΔPro), and using this plasmid, pGLC88, and the neuronal marker, *odr-2b::dsRED*, I rescued the *lin-11(n389)* phenotype. Three transgenic lines were generated (*bhEx268*, *bhEx269*, and *bhEx270*), and neuronal and reproductive tissues were analyzed in the rescue animals (Table 4.1). In all three transgenic arrays, LIN-11ΔPro rescued the *lin-11(n389)* mutations in both the neuronal and reproductive tissues. Approximately ~29% (n = 72) of

the rescued animals were able to lay eggs, whereas all of the *lin-11* mutants were egg-laying defective (Figure 4.2). A closer examination of the vulva organ showed that in a small population of the rescued animals (~29%, n = 72), the vulva structure displayed invagination, toroid finger protrusions, and uterine seam cells (utse) acquired the wild-type phenotypes (Figure 4.3D). The utse is an H-shaped cell within the hermaphrodite uterus that functions in attaching the uterus to the body wall (Newman et al. 1999). However, a large population (~71%, n = 72) of the rescued animals did not lay eggs, and examination of the vulva organ revealed defects in one or all of these categories (invagination, toroid finger protrusions, and utse) (Figure 4.2 D1-D3). For example, some rescued animals showed defects only in the utse, while others showed defects in both the utse and toroids finger protrusions. Together, these findings suggest that the function of the LIN-11 PRR is not essential in vulva cell specifications.

In the nervous system, LIN-11ΔPro rescued *lin-11(n389)* animals expressed the *odr-2::dsRED* reporter in *lin-11* specific neurons ASG (~14%, n = 69), AIZ (~47%, n = 69), AVG (~47%, n = 69), and RIF (~66%, n = 69) (Figures 4.4 and 4.5C). In *lin-11(n389)* mutants, these neurons do not express the *odr-2* reporter, because these specific neurons have not acquired the appropriate cell fates. Expression of *odr-2::dsRED* reporter in the rescued animals suggests that neuronal cells in these animals have acquired the appropriate fate, and express the correct terminal differentiation gene, *odr-2*, that determines the functional property of a neuron (Figure 4.4). However, it is unknown whether the remaining nine *lin-11* specific neurons (ADL, AVA, AVE, ASH, ADF, AVJ, AWA, ASG, and RIC) are also rescued by the function of LIN-11ΔPro. Thus, future experiments are required using additional neuron-specific promoters to test whether these neurons are also rescued in *lin-11* mutants (see 5.3.1 section for details).

lin-11 mutants display thermophilic behavior due to defective interneuron AIZ (Hobert et al. 1998). To further investigate this defective behavior, I performed a thermotaxis assay to test if the LIN-11ΔPro rescued animals are able to display wildtype behavior in a temperature gradient environment. Briefly, the worms are grown at an ambient

temperature (20 °C); when exposed to a temperature gradient environment, many remain in the cultivated temperature zone (Chapter 2). The rescued animals, *bhEx268*, exhibited significantly improved thermotactic behaviour (p < 0.05; n = 30) compared to the *lin-11(n389)* mutants (n = 30; Figure 4.6). Most rescued animals (~57%, n = 30) moved more frequently in the cultivated thermal gradient zone than did the wild type animals (~77%, n = 30) (Figure 4.6). Furthermore, ~33% (n = 30) of *bhEx268* displayed the *lin-11(n389)* mutant-like behavior (very little movement; remained in the high temperature zone), even though the *odr-2::dsRED* reporter expression was present in the AIZ interneuron (Hobert et al. 1998). This suggests that not all animals are rescued by LIN-11ΔPro, even though the *odr-2::dsRED* expression is present in the AIZ interneuron.

Together, rescued animals' thermotactic behaviour and neuronal quantifications suggest that the *lin-11* phenotype is rescued by LIN-11ΔPro. This work demonstrates that the PRR of LIN-11 is not essential for its overall function in the specification of the nervous and reproductive tissues. However, it cannot be ruled out that the PRR of LIN-11 may function in mediating protein-protein stabilization together with the LIM domains; additional experiments are required to test this possibility.

4.5 *Drosophila melanogaster*, *dLim1*, can restore LIN-11 function in *C. elegans*

To test functional conservation of *Lhx* genes from other invertebrates that can restore the function of *lin-11* in *C. elegans*, the dLim1 gene from *D. melanogaster* was an ideal candidate. Not only is *D. melanogaster* an established invertebrate model organism, but *dLim* is the ortholog of *lin-11*, and its function has been reported in neuron specification in ventral nerve cord, leg, and antenna development (Figure 1.3) (Lilly et al. 1999b; Tsuji et al. 2000)

A gene and protein comparison between *lin-11* and its ortholog *dLim1* revealed 34%

conservation in overall protein structure between these two species (Figure 1.10). Specifically, domain comparisons showed that LIM1 and LIM2 are 49% and 48% similar, respectively, while the HD and PRR are 89% and 31% similar in identities (Figure 1.10). At the gene level, it was found that *dLim1* has fewer exons and introns (7 and 6, respectively) compared to *lin-11* (10 and 9, respectively). On average, the *dLim1* exons are twice as large as *lin-11* exons, dLIM1 is 20% (505 aa) larger than LIN-11 (405 aa), and dLIM1 has more proline residues (31P) in the PRR compared to LIN-11 (18P). In development, the functions of the PRR in both of these organisms are unclear; one possibility is that they are required for mediating protein-protein stabilization.

To test whether dLIM1 can replace the function of LIN-11 and rescue the *lin-11* mutation, I constructed two tissue-specific plasmids, pGLC96 and pGLC98, that contain the *D. melanogaster dLim1* cDNA to be expressed in the vulva cells and neuronal cells, respectively (Figures 2.16 and 2.18). Three transgenic lines were generated (*bhEx264*, *bhEx265* and *bhEx26*6) using these rescue plasmids, and expression was analyzed in the nervous and reproductive tissues. In the reproductive tissue, ~35% (n = 88) of the rescued animals (*bhEx264*) laid eggs, compared to the *lin-11* mutants, which were all egg-laying defective (Figure 4.2). A closer examination of the vulva organ revealed that invagination, toroid finger protrusions, and the utse had acquired the wild type phenotypes (Figure 4.3 E). However, a large population (~65%, n= 88) of the rescued animals did not lay eggs, and displayed defective vulva structure as well (Figure 4.2 E1-E3). For example, in some rescued animals that did not lay eggs, the utse was defective, resulting in blockage of the passage of the egg exiting the uterus. In other rescued animals, defects were observed in both the utse and toroids finger protrusions. These variation patterns observed in the vulval phenotype could be the result of the mosaicism of extrachromosomal arrays, with the expression of foreign DNA expressed differently in every animal; alternately, *dLim1* may not rescue the *lin-11* phenotype in some animals. It is important to note that the GFP reporter expression in the vulval cells was extremely faint; it was therefore impossible to quantify individual cells in detail.

In the nervous system, *lin-11(n389)*-rescued animals showed expression of *odr-2::dsRED*

neuron reporter in *lin-11*-specific neuronal cells, ASG (~13%, n = 60), AIZ (~62%, n = 60), RIF (~72%, n = 60) and AVG (~17%, n = 60) (Figures 4.4 and 4.5; Table 4.1). In *lin-11(n389)*, these neurons do not express *odr-2::dsRED* reporter, because they have not acquired proper cell fate (Table 4.1). Thus, compared to the *lin-11(n389)* mutants, the rescued animals, *bhEx264*, showed significantly more *odr-2::dsRed* reporter expression in the *lin-11*-specific neurons (p < 0.05 - 0.001, n = 80). The expression of *odr-2::dsRED* in these *lin-11*-specific neurons indicates that neuronal cells have acquired proper cell fates by expressing the terminal differentiation gene, *odr-2*, which determines the functional properties of these neurons. It is important to note that *odr-2::dsRED* fluorescence in AVG and ASG neurons are very weak and very rare, so the identification of these neurons was based solely on cell positions. Furthermore, it is not clear whether the remaining eight *lin-11*-specific neurons (ADL, AVA, AVE, ASH, ADF, AVJ, AWA, and RIC) are also rescued by the function of dLIM1, since the *odr-2::dsRED* reporter is specific to AIZ, ASG, RIF, and AVG neurons. Thus, additional experiments are required using neuron-specific reporters to determine whether these eight *lin-11* specific neurons are rescued by *dLim1* (see 5.3.1 section for details).

Reporter expression in cells allows us to determine the expression of a gene, but a more specific assay is required to fully understand the cells' functions. *lin-11* mutants display thermophilic behavior due to defective AIZ interneurons (Hobert et al. 1998). To further investigate this defective behavior, I performed a thermotaxis assay to test whether the dLim1 rescued animals are able to display wildtype behavior in a temperature gradient environment. I found that ~53% (n = 30) of the rescued animals, *bhEx264*, displayed wildtype behavior by moving often and remaining within the cultivated temperature zone. While the majority (~57%, n = 30) of *lin-11(n389)* animals exhibited thermophilic behavior, the *bhEx264* rescued animals displayed significantly improved thermotactic behavior compared to the mutants (p < 0.01, n = 30) (Figure 4.7). These findings suggest that the thermoregulatory behavior that is disrupted by interneuron AIZ in *lin-11(n389)* mutants is rescued in *bhEx264* animals, because these animals display wildtype behavior in a temperature gradient environment. Furthermore, ~34% (n = 30) of the rescued animals showed thermophilic behaviors, and ~14% (n = 30) displayed cryophilic

behaviors, which could be the result of *dLim1* failing to rescue the defective thermotactic phenotype; or, alternately, the result of mosaicism of the extrachromosomal DNA (Figure 4.7).

In conclusion, the invertebrate ortholog *Lhx* gene, *dLim1*, can restore the function of *lin-11* in *C. elegans* in both the neuronal and reproductive tissues.

4.6 *Mus musculus*, *Lhx1*, can restore *lin-11* function in both the neuronal and reproductive tissues

The invertebrate *Lhx* gene, *dLim1*, restored the function of *lin-11* in *C. elegans*; however, it is not known whether vertebrate *Lhx* genes can perform the same function in *C. elegans*. To investigate this, I focused on the *lin-11*-orthologous gene in mice, *Lhx1*. In mice, *Lhx1* is required for head induction, axis formation, and intermediate mesoderm differentiation (Shawlot and Behringer 1995; Tsang et al. 2000; Kinder et al. 2001).

A gene and protein comparison between *lin-11* and its ortholog *Lhx1* revealed a 39% conservation in overall protein structure between these two species (Figure 1.10). Specifically, domain comparisons showed that LIM1 and LIM2 are 47% and 55% similar, respectively, and the HD and PRR are 97% and 46% similar in identities (Figure 1.10). At the gene level, *Lhx1* has fewer exons and introns (5 and 4, respectively) compared to *lin-11* (10 and 9, respectively). On average, the *dLim1* exons are twice as large as *lin-11* exons, and the protein LHX1 is similar in size (406 aa) to LIN-11 (405 aa). However, LHX1 has more proline (28P) residues in the PRR compared to LIN-11 (18P), and the functions of these regions are not clear in the development of these organisms.

To investigate the degree of functional conservation, I constructed two plasmids, pGLC94 and pGLC97, containing the cDNA sequence of the *Lhx1* gene under the endogenous *lin-11* promoters (Figure 2.15 and 2.17). The pGLC94 plasmid contains the vulval cell-specific promoter, *lin-11* (5'UTR), in frame with *Lhx1cDNA*. The pGLC97

plasmid, contains a neuron-specific promoter, *lin-11* (intron 3), in frame with *Lhx1cDNA*. The *lin-11* tissue-specific promoters were used in lieu of the native *Lhx1* promoter to prevent ectopic expression of *Lhx1* in *C. elegans*, and more importantly, to ensure the function of *Lhx1* was directed within the *lin-11*-specific tissues, in this case neuronal and reproductive tissues. Three transgenic extrachromosomal arrays were generated (*bhEx261*, *bhEx262*, and *bhEx263*) using these plasmids, and the neuronal and vulval cells were analyzed and quantified.

In *C. elegans*, the vulva-precursor cells acquire proper cell fate and together form the vulva organ (Figure 4.3A). In *lin-11* mutants, the vulva-precursor cells do not differentiate, and as a result do not form the vulva organ and exhibit the egg-laying defect phenotype (Egl) (Figure 4.3B). However, ~37% (n = 81) of the *Lhx1* rescued *lin-11(n389)* animals were able to lay eggs, and closer examination of the vulva organ structure revealed that the invagination, toroid finger protrusions had acquired the wildtype phenotypes (Figures 4.2 and 4.3F). Moreover, ~63% (n = 81) of the rescued animals displayed the Egl phenotype and defective vulva organ. A closer examination of the vulva organ revealed variation in the structure: in some animals only utse is defective, while in others, both utse and toroids finger protrusions are defective (Figure 4.2 F1-F3). This variation in the vulva phenotype could be the result of *Lhx1* unable to rescue the vulva phenotype, or of mosaicism of extrachromosomal DNA. It is important to note that the GFP reporter expression in the vulva cells is very faint, which made detailed analysis of the cells impossible.

In the nervous system, *Lhx1*-rescued *lin-11(n389)* animals expressed the *odr-2::dsRED* reporter in *lin-11*-specific neurons, ASG (~18%, n = 66), AIZ (~70%, n = 66), RIF (~92%, n = 66) and AVG (~19%, n = 66) (Figures 4.4 and 4.5). These neurons do not acquire proper cell fate in *lin-11(n389)* mutants, and as a result they do not express *odr-2::dsRED* reporter. Compared to the *lin-11(n389)* mutants, the *bhEx261* rescue animals show a significantly higher percentage of *odr-2::dsRED* expression in these neurons (p < 0.05 - 0.001) (Figure 4.4). These results suggest that in the rescued animals, *Lhx1*

restored the function of *lin-11*, and these neurons acquired proper cell fate. It is important to note that *odr-2::dsRED* fluorescence in AVG and ASG neurons is both weak and rare, which made the identification of these neurons very difficult. It is not known whether the remaining eight *lin-11*-specific neurons (ADL, AVA, AVE, ASH, ADF, AVJ, AWA, and RIC; Table 1.2) are also rescued by the function of *Lhx1*. Thus, additional experiments are required, using neuron specific reporters to test whether these neurons are also rescued in *lin-11* mutants by function of *Lhx1*.

Although reporter expression in cells allows us to determine the expression of a gene, a more specific assay is required to fully understand the function of the cells. I performed a thermotaxis assay on the rescued animals to determine whether the rescued animals are able to display wildtype behavior in a temperature gradient environment. Briefly, the worms are grown at an ambient temperature ($20\,°C$), and when exposed to a temperature gradient environment, many remained in the cultivated temperature zone (see Chapter 3). It is known that *lin-11* mutants display thermophilic behavior due to defects in interneuron AIZ (Hobert et al. 1998). As summarized in Figure 4.7, ~50% (n = 30) of the *bhEx261* rescued animals displayed the wildtype behavior by moving frequently and remaining within the cultivated temperature zone. The majority of *lin-11(n389)* mutants, ~57% (n = 30), displayed thermophilic behavior; ~27% (n = 30) displayed cryophilic behavior; and ~17% (n = 30) displayed wildtype behavior. Compared to the *lin-11(n389)* mutants, the rescued animals showed significantly (p < 0.05; n = 30) improved thermotactic behavior (Figure 4.7). These findings suggest that the thermoregulatory behavior disrupted by the defective AIZ interneuron in the *lin-11(n389)* mutants is rescued in the rescue animals; thus, the rescued animals display wildtype behavior in a temperature gradient environment. In addition, ~30% (n = 30) of the rescued animals displayed thermophilic behavior, and ~20% (n = 30) displayed cryophilic behavior; this could occur because some animals are not rescued by *Lhx1*, or due to mosaicism of extrachromosomal DNA(Figure 4.7).

Together, these findings suggest that the murins *Lhx* gene, *Lhx1*, can restore the function of *lin-11* in *C. elegans*, in both the neuronal and reproductive tissues. This

demonstrates that the domains of Lhx proteins are similar enough to mediate protein-protein and protein-DNA interactions similarly to the endogenoius LIN-11.

A

	LIM1	LIM2		HD	PRR
	68-124	127- 187		241-300	301-405

Protein Binding   DNA Binding   Unknown

B   LIM Domain

C

```
                10          20          30          40          50
        MHSSSSFIIT SLEEEEKKPP AHHLHQQSIE DVGSVTSSAT LLLLDSATWM
                60          70          80          90         100
        MPSSTTQPHI SEISGNE CAA CAQPILDRYV FTVLGKCWHQ SCLRCCDCRA
               110         120         130         140         150
        PMSMTCFSRD GLILCKTDFS RRYSQR CAGC DGKLEKEDLV RRARDKVFHI
               160         170         180         190         200
        RCFQCSVCQR LLDTGDQLYI MEGNRFVCQS DFQTATKTST PTSIHRPVSN
               210         220         230         240         250
        GSECNSDVEE DNVDACDEVG LDDGEGDCGK DNSDDSNSAK RRGPRTTIKA
               260         270         280         290         300
        KQLETLKNAF AATPKPTRHI REQLAAETGL NMRVIQVWFQ NRRSKERRMK
               310         320         330         340         350
        QLRFGGYRQS RRPRRDDIVD MFPNDQQFYP PPPPSNVQFF CDPYTTSPNN
               360         370         380         390         400
        PETIQMAPQF AVPTENMNMV PEPYTEQSAT PPEFNEDTFA CIYSTDLGKP

        TPVSW 405
```

FIGURE 4.1: (A) Schematic structure of LIN-11, which consists of two LIM domains that are made up of 57 and 61 amino acids (aa) respectively; a homeodomain (60 aa) and a PRR (105 aa). (B) The LIM domain is a zinc-binding, cysteine-rich motif consisting of two tandemly repeated zinc fingers. The LIN-11 LIM consensus sequence is $C_{X2}C_{X18}H_{X2}C_{X2}C_{X2}C_{X17}C_{X2}D$, where X signifies other amino acids. (C) LIN-11 consists of 405 aa; the LIM domains are depicted in blue, HD in red, and the proline-rich region (PRR) in green.
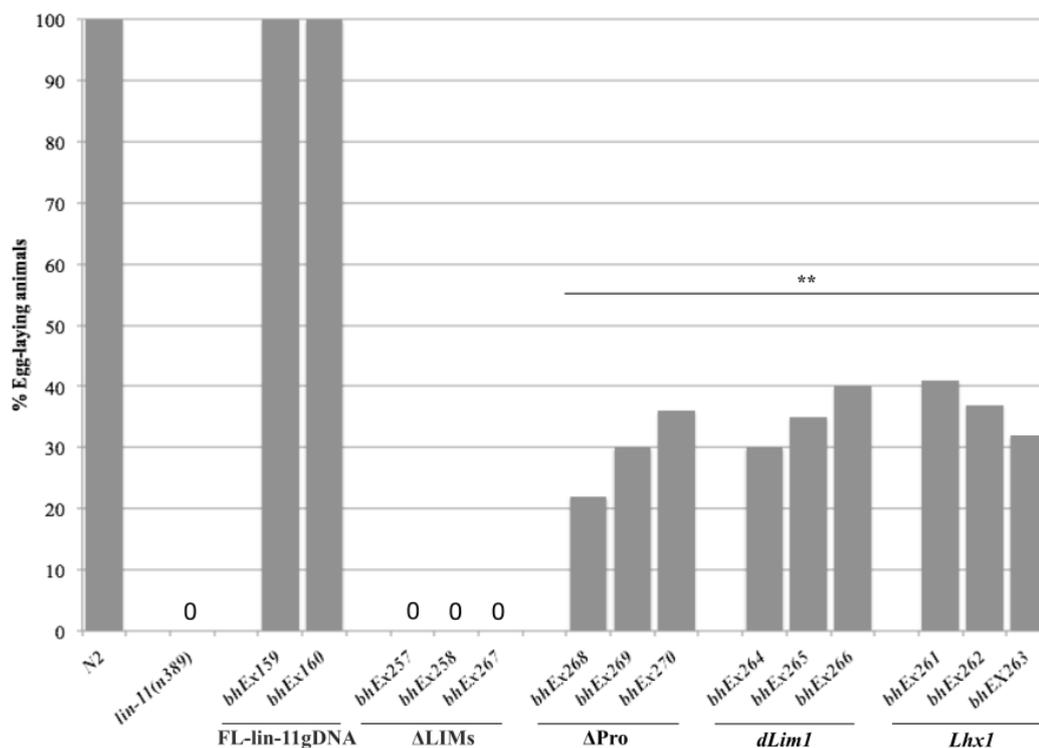
FIGURE 4.2: Quantifications of *lin-11(n389)*-rescued animals that lay-eggs. The *C. elegans* adult wildtype (N2) hermaphrodite lay ~300 eggs, while *lin-11(n389)* animals are egg-laying defective (Egl), because the vulval precursor cells are not differentiated and the vulva structure is not made. All ΔLIM rescued *lin-11(n389)* animals exhibit Egl phenotype and do not lay-eggs to suggest that LIM domains are required for proper LIN-11 function in vulval cell differentiation. By contrast, approximately 29% (n= 72) of three different transgenic arrays of the LIN-11ΔPro rescued *lin-11(n389)* transgenic arrays (*bhEx268*, *bhEx269*, and *bhEx270*) were able to lay eggs. In addition, 35% (n= 88) and 37% (n= 81) of *lin-11(n389)* mutants rescued with *lin-11* orthologs, *dLim1* and *Lhx1* laid eggs respectively.
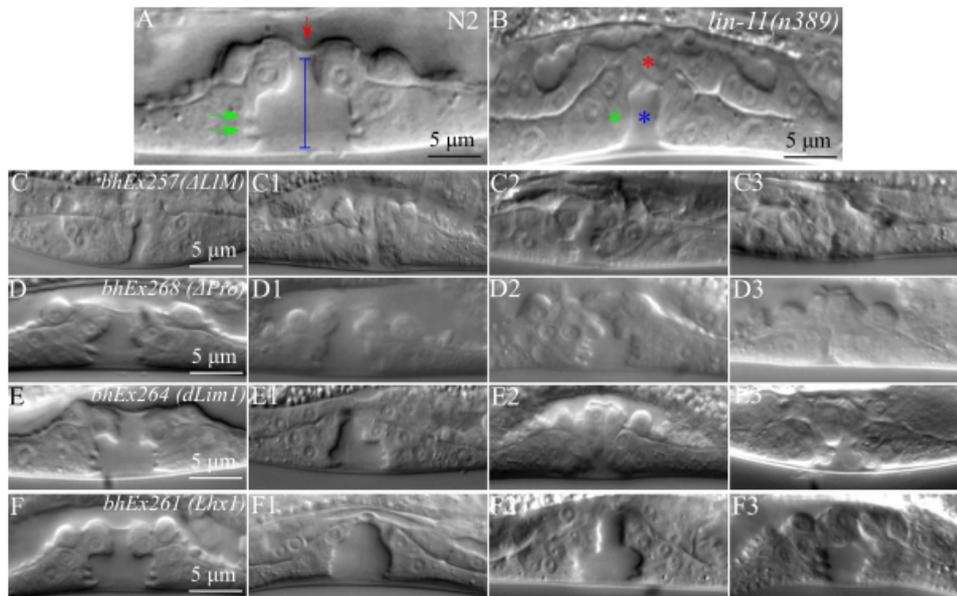
FIGURE 4.3: *lin-11(n389)* rescued vulval morphologies. (A) Illustrating the vulva morphology of the wildtype animals (N2). The vulva invagination is represented by the blue line, green arrows mark the toroid finger protrusions, and the red arrow indicates the uterine seam cells (utse). The *lin-11* null allele, n389 (B), displays defects in invagination (blue *), utse (red *) and toroid finger protrusions (green *), contributing to the egg-laying defective observed in these animals (Egl). *lin-11(n389)* animals rescued with LIN-11ΔLIM (C-C3) did not rescue the vulva phenotypes. These animals are Egl, and lack the overall vulva structure observed in the N2. *lin-11(n389)* mutants rescued by LIN-11ΔPro (D-D3) are partially rescued as ~29% (n= 72) laid eggs (see Figure 3), in these animals (D-D1) the vulva structure has proper invagination, toroid finger protrusions and utse. However, the majority of the rescued animals (~71%, n= 72) exhibit the Egl phenotype, and vulva structure is completely defective (D2-D3). *lin-11* ortholog *dLim1* rescued the *lin-11(n389)* vulva defect (E), and ~35% (n= 88) of animals laid eggs. However, the remaining ~65% (n= 88) of the rescued animals showed Egl phenotype and the overall vulva structure look similar to that of *lin-11(n389)* mutants (E1-E3). Similarly, *Lhx1*-rescued *lin-11(n389)* animals showed rescued vulva structure phenotype (F) and ~37% (n= 81) laid eggs. However, the majority (~63%; n= 81) of the rescued animals exhibited the defective vulva structure (F1-F3).
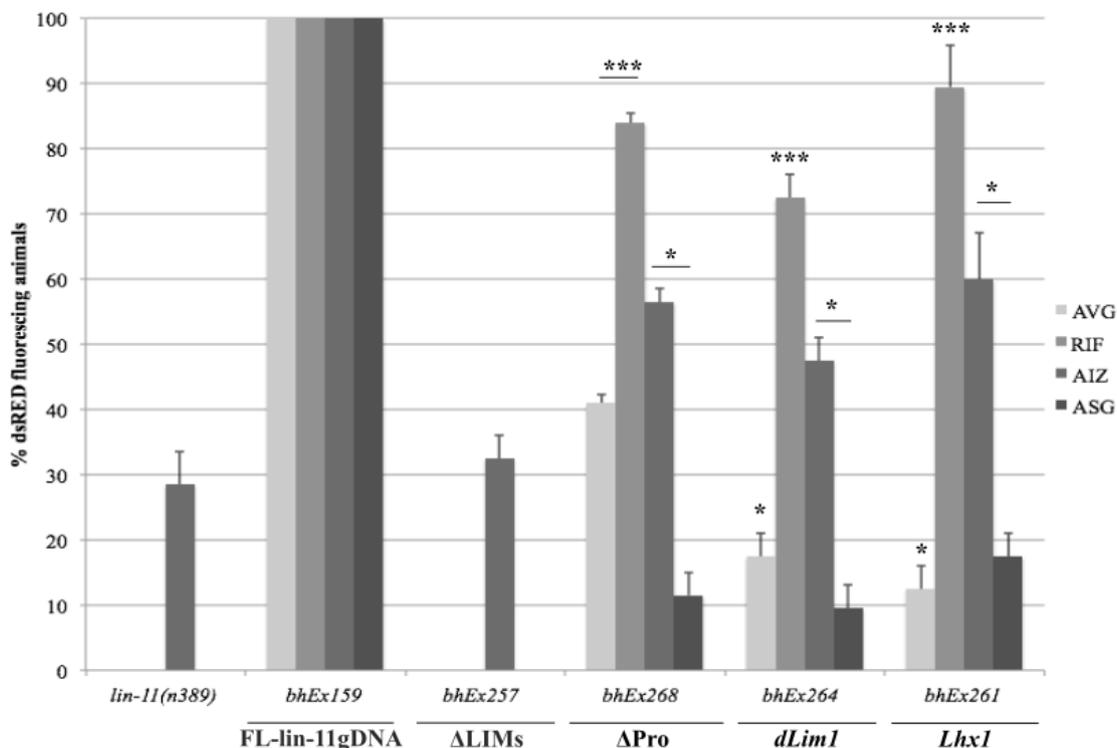
FIGURE 4.4: Quantification of truncated LIN-11 and its orthologs, *dLim1* and *Lhx1*, dsRED reporter fluorescing neurons in *lin-11* rescued animals. In *lin-11(n389)*, *odr-2::dsRED* fluorescence is present in one neuron, AIZ (~29%, n= 45), and it is not present in the AVG, RIF or ASG neurons. Similarly, *lin-11(n389)* mutant rescued with LIN-11ΔLIM, *bhEx257*, where the *odr-2::dsRED* fluorescence is present in the AIZ neuron in a small proportion (~30%, n= 40) of the animals. LIN-11ΔPro-rescued animals showed *odr-2::dsRED* reporter expression in AVG (~41%, n= 44), RIF (~84%, n= 44), AIZ (~57%, n= 44) and ASG (~3%, n= 44), which is significant compared to *lin-11(n389)* mutant animals. In addition, *lin-11* intron 3-driven *lin-11* ortholog *dLim1cDNA*, *bhEx264*, rescued *lin-11(n389)* animals and expressed the *odr-2::dsRED* reporter in AVG (~18%, n= 40), RIF (~73%, n= 40), AIZ (~48%, n= 40) and ASG (~3%, n= 40). Similarly, *lin-11* intron 3-driven *lin-11* ortholog *Lhx1cDNA*, *bhEx261*, rescued *lin-11(n389)* animals and expressed the *odr-2::dsRED* reporter in AVG (~13%, n= 40), RIF (~90%, n= 40), AIZ (~60%, n= 40) and ASG (~18%, n= 40). *: p< 0.05 and ***: p < 0.001.
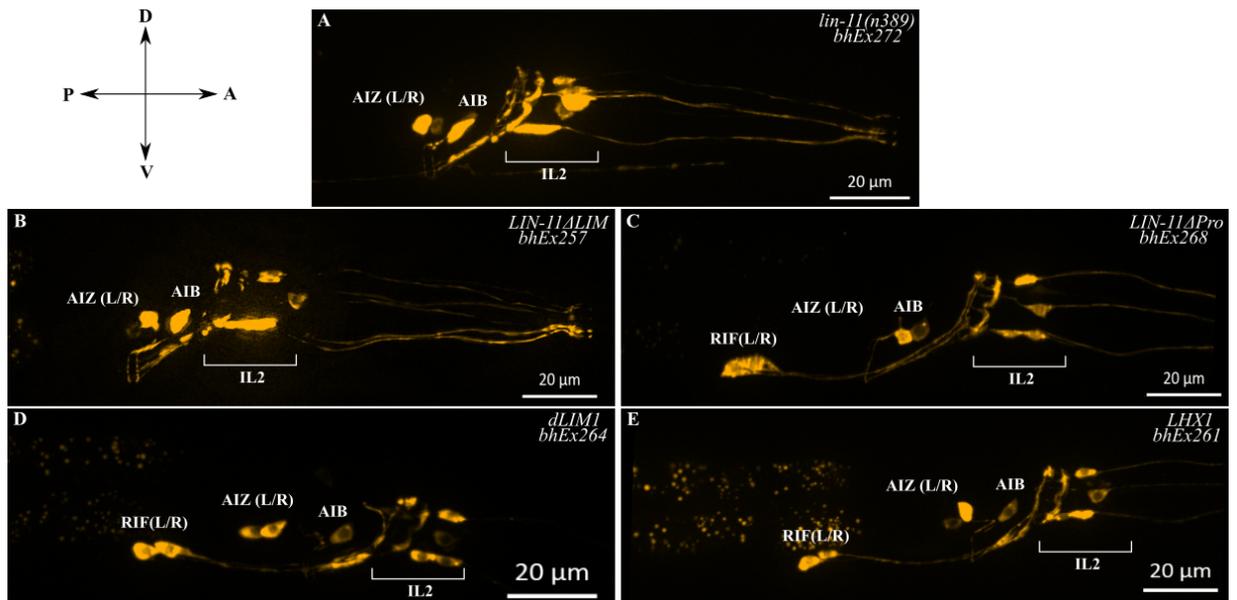
FIGURE 4.5: Rescue of neuronal defects by truncated LIN-11 and intron 3-driven *dLim1* (*int3p::dLim1cDNA*) and *Lhx1* (*int3p::Lhx1cDNA*) cDNAs. The *odr-2::dsRED* expression in four neurons (AVG, RIF, AIZ, and ASG) was quantified. (A) The expression of *odr-2* in *lin-11(n389)* is present in AIZ, but never in RIF, AVG or ASG (see Figure 4.4; Table 4.1).(B) Similarly, *lin-11(n389)* mutants rescued with LIN-11ΔLIM showed similar expression as the control, *bhEx272*, and reporter expression was present in AIZ neuron only. (C) By contrast, *lin-11(n389)* animals rescued with LIN-11ΔPro expressed the reporter strongly in AIZ and RIF, and weakly in ASG and AVG (see Figure 4.4; Table 4.1). Similarly, *lin-11(n389)* rescued with its orthologs dLIM1 (D) and LHX1 (E) express the *odr-2::dsRED* reporter strongly in the AIZ and RIF neurons, and weakly in ASG and AVG neurons (see Figure 4.4; Table 4.1). A - anterior, D - dorsal, P - posterior, and V - ventral.
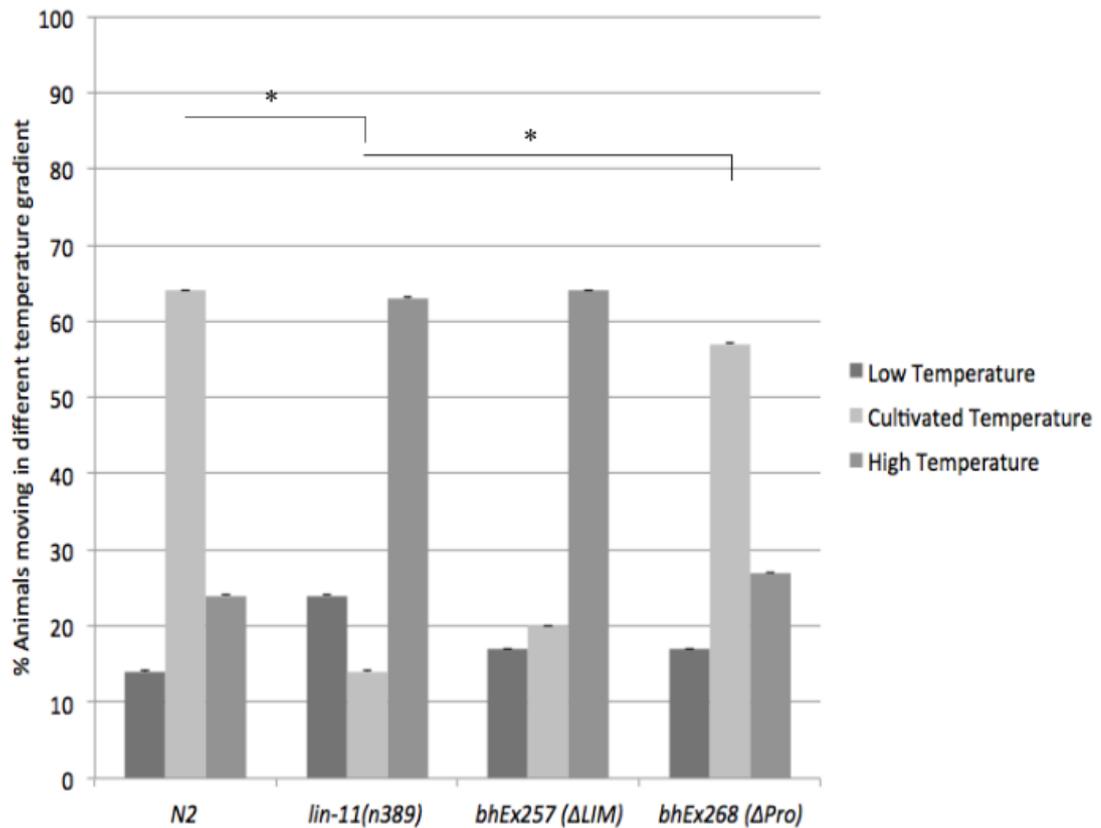
FIGURE 4.6: Rescue of *lin-11* thermotaxis defect by *lin-11* intron 3-driven truncated *lin-11* cDNA. Animals were placed at the edge of the assay plate and were scored after 1 hour. Animals were found in one of three temperature zones: low (preference for lower than cultivated temperature), cultivated temperature, and high (preference for higher than cultivated temperature). The wildtype, N2, were found in the cultivated temperature, while *lin-11(n389)* exhibited thermophilic behaviour. The *lin-11(n389)* rescued with LIN-11ΔLIM (*bhEx257*) displayed thermophilic behaviour as well. In contrast *lin-11(n389)* rescued with LIN-11ΔPro, *bhEx268*, behaved similarly to the wildtype (N2) control. 30 animals were examined for each strain. *: p< 0.05.
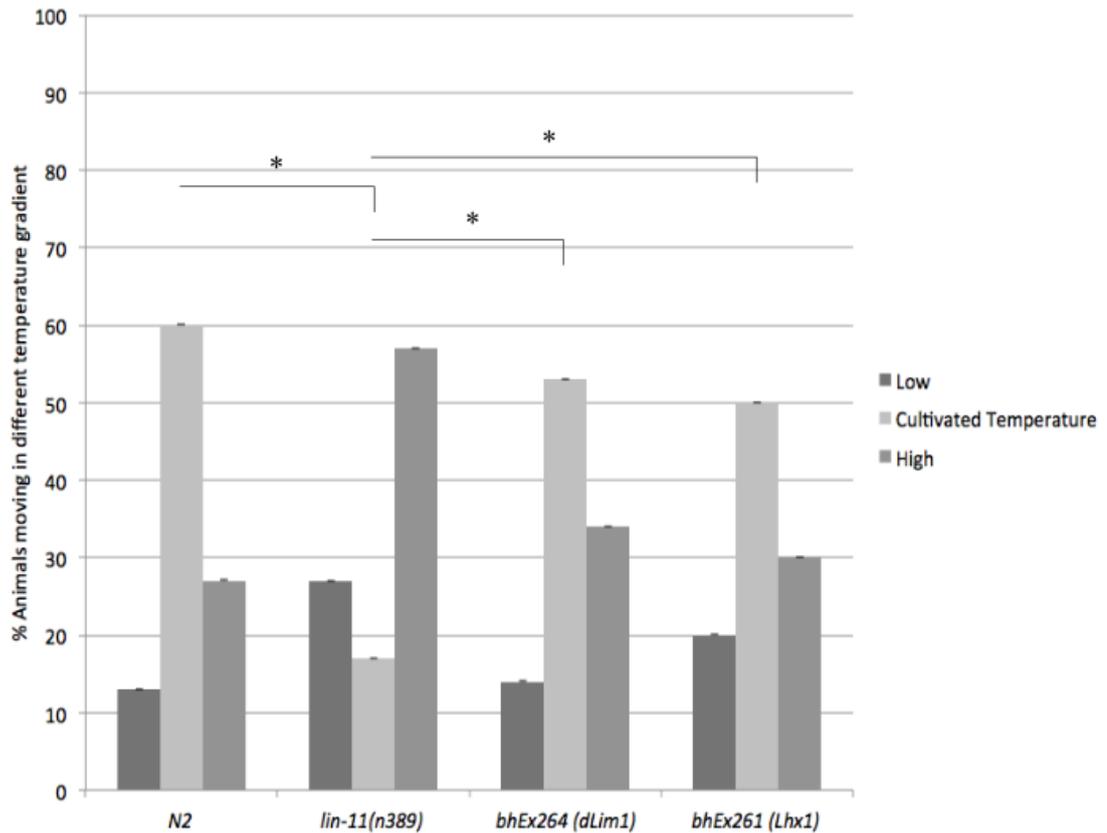
FIGURE 4.7: Rescue of *lin-11(n389)* thermotaxis defect by *lin-11* intron 3-driven *lin-11* orthologs, *dLim1* and *Lhx1*, cDNAs. Animals were placed at the edge of the assay plate and were scored after one hour. Animals were found in one of three temperature zones: low (preference for lower than cultivated temperature), cultivated temperature, and high (preference for higher than cultivated temperature). The *lin-11* rescued with *dLim*, *bhEx264*, and *Lhx1*, *bhEx261*, displayed improved thermotactic behaviour similar to the wildtype (N2) animals. For the quantifications, 30 animals were examined for each strain. *: p< 0.05
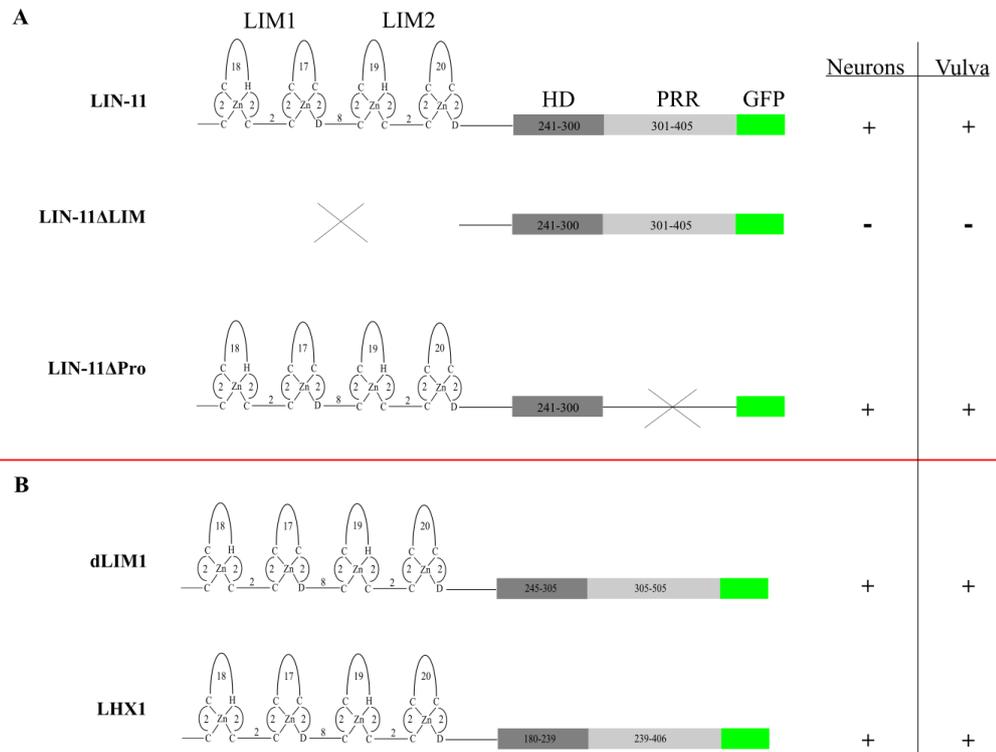
FIGURE 4.8: Schematic representation of transgenes tested for their ability to rescue the *lin-11* phenotype. Full-length, truncated, and *dLim1* and *Lhx1* cDNAs were cloned into plasmids that had neuronal (intron 3) or vulva-specific (5'UTR) promoters (see Chapter 2). The tissue specific-promoters enabled the cDNA to be expressed in *lin-11*-specific cells. (A) The LIN-11 protein is 405 amino acids (aa) in length, and can restore the *lin-11* mutation in both neuronal and vulval cells. LIN-11ΔLIM truncates both LIM domains by eliminating the amino acids 1-199 from the LIN-11 wildtype protein. LIN-11ΔLIM did not restore the function of LIN-11 in the neuronal and vulval tissues. LIN-11ΔPro eliminates aa 324-405 from the LIN-11 wildtype protein. LIN-11ΔPro restored the function of LIN-11 in the neuronal and vulval tissues. (B) *lin-11* orthologs, *dLim1* and *Lhx1*, are 502 and 405 aa in length respectively. Both of these proteins are able to restore the function of LIN-11 in the neuronal and vulval tissues.

TABLE 4.1: Quantifications of neuronal phenotype in *lin-11* rescued animals.

Species	Plasmid	Promoter	Insert Name	Background	Array	Neuronal marker	Egg-Laying	*odr-2* Reporter Expression			
								AVG	RIF	AIZ	ASG
C. elegans	pVH10.17	*odr-2*	CFP	*unc-119(tm4063)*	*bhEx170*	*odr-2::CFP*	100%	85% (n= 15)	85% (n =15)	90% (n= 15)	70% (n= 15)
	pGLC102	*odr-2*	dsRED	n389	*bhEx272*	*odr-2::dsRED*	0%	0% (n= 25)	0% (n= 25)	32% (n= 25)	0% (n= 25)
	pVH10.17	*odr-2*	CFP	ps1	*bhEx157*	*odr-2::CFP*	0%	0% (n= 25)	0% (n=25)	32% (n= 25)	0% (n= 25)
	pVH10.17	*odr-2*	CFP	ps1	*bhEx158*	*odr-2::CFP*	0%	0% (n=20)	0% (n= 20)	25% (n=20)	0% (n= 20)
	pVH10.17	*odr-2*	CFP	n389	*bhEx150*	*odr-2::CFP*	0%	0% (n=15)	0% (n= 15)	27% (n= 15)	0% (n= 15)
	pGF50	5' UTR and all introns	FL-gDNA	n389	*bhEx159*	*odr-2::CFP*	85% (n= 20)	100% (n= 15)	100% (n= 15)	100% (n= 15)	100% (n= 15)
				n389	*bhEx160*		75% (n= 15)	100% (n= 20)	100% (n= 20)	100% (n= 20)	100% (n= 20)
	pGLC88	5' UTR, All introns except 8 and 9	ΔProline	n389	*bhEx268*	*odr-2::dsRED*	22% (n= 27)	42% (n= 24)	83% (n=24)	58% (n= 24)	10% (n= 24)
				n389	*bhEx269*		30% (n= 20)	44% (n= 25)	56% (n= 25)	44% (n= 25)	12% (n= 25)
				n389	*bhEx270*		36% (n= 25)	55% (n=20)	60% (n= 20)	40% (n= 20)	20% (n= 20)
	pGLC89	Intron3	ΔLIM	n389	*bhEx257*	*odr-2::dsRED*	0% (n= 20)	0% (n =20)	0% (n =20)	30% (n= 20)	0% (n= 20)
	pGLC90	5'UTR		n389	*bhEx258*		0% (n= 20)	0% (n= 30)	0% (n= 30)	37% (n= 30)	0% (n= 30)
				n389	*bhEx267*		0% (n= 20)	0% (n= 20)	0% (n= 20)	35% (n= 20)	0% (n= 20)
M. musculus	pGLC94	5'UTR	*lin-11p::Lhx1cDNA*	n389	*bhEx261*	*odr-2::dsRED*	41% (n= 29)	10% (n = 20)	94% (n= 20)	65% (n= 20)	20% (n= 20)
	pGLC97	Intron3		n389	*bhEx262*		37% (n= 30)	25% (n = 16)	94% (n= 16)	69% (n= 20)	13% (n= 16)
				n389	*bhEx263*		32% (n= 22)	23% (n= 30)	87% (n= 30)	77% (n= 30)	20% (n 30)
D. melanogaster	pGLC96	5'UTR	*lin-11p::dLim1cDNA*	n389	*bhEx264*	*odr-2::dsRED*	30% (n= 23)	15% (n= 20)	70% (n= 20)	45% (n= 20)	15% (n= 20)
	pGLC98	Intron3		n389	*bhEx265*		35% (n= 20)	20% (n= 20)	85% (n= 20)	75% (n= 20)	15% (n= 20)
				n389	*bhEx266*		40% (n= 45)	15% (n= 20)	60% (n =20)	65% (n= 20)	10% (n= 20)

Chapter 5.0

Discussion and Future Directions

5.1 *lin-11* is regulated by multiple enhancers in specific neuronal cells

The work described in this thesis focuses on the regulation and function of *lin-11* in specification of amphid neurons. Expression analysis showed that two of the largest introns of *lin-11* (introns 3 and 7) could to drive the expression of a GFP reporter in a subset of amphid neurons. Specifically, intron 3 mediated expression was present in seven amphid neurons, ADL, ADF, AVJ, RIC, AIZ, RIF, and AVG, while intron 7 mediated expression was present in two neurons: AVG and a second neuron, whose fate was not clearly established. Phenotypic rescue experiments in *C. elegans* revealed that intron 3 restors *lin-11* function based on gene expression patterns and behavioral assays; however, intron 7 is unable to rescue *lin-11* function in the neurons. Intron 3-driven expression analysis has provided evidence of both *cis* and *trans* changes in *lin-11* regulation between *C. elegans* and *C. briggsae*. Additionally, sequence alignment between four *Caenorhabditis* species revealed three distinct non-overlapping enhancers within intron 3, and two within intron 7 (Chapter 3). I found that one of these conserved regions within intron 3, C3-1, contains an enhancer for interneuron RIC, and removal of this region prevents the expression of the GFP reporter in this neuron. Similarly, intron 7 harbors enhancers for the AVG neuron. Using an in silico tool (CIS-BP) as well as the modENCODE Chip-Seq database, I identified a total of 114 putative TFs that potentially bind to these conserved regions, of which 14 genes were experimentally

examined. Using both genetic and expression studies, four TFs were identified that act through these enhancers to regulate expression of *lin-11* in the nervous system, namely SKN-1, CEH-6, CRH-1, and CES-1 (Chapter 3). Specifically, I have shown that SKN-1, CEH-6, and CRH-1 are negative regulators of *lin-11*, and may bind to the C3-1 region to promote specification of the RIC interneuron; CES-1 is a positive regulator of *lin-11* with a potential binding site within C7-1, which promotes the AVG interneuron specification. My work has uncovered a complex regulatory genetic pathway that governs the expression of *lin-11* in specification of RIC and AVG interneurons neurons.

The expression of *lin-11* has been reported in various other amphid neurons, including in the specification of five amphid sensory neurons: ADL, ADF, ASG, AWA, and ASH (Figure 1.2) (Hobert et al. 1998; Sarafi-Reinach et al. 2001; Yamada et al. 2012). The complex regulation of *lin-11* in these sensory neurons is not well understood; however, several researchers have attempted to identify the regulatory pathway of *lin-11* in specification of these neurons. For example, one such study showed that in *lin-11* mutants, the AWA sensory neuron adopts the AWC-like fate, assessed by expressing the AWC-specific olfactory receptor gene *str-2* (Sarafi-Reinach and Sengupta 2000). However, it is not clear whether the remaining four sensory neurons (ASG, ADL, ADF, and ASH) undergo a similar cell fate change in *lin-11* mutants as well, and if so, which cell identity they acquire. Furthermore, the authors also report that the sensory neurons ASG and AWA are derived from the same neuroblast precursor cells; however, the function of *lin-11* is completely different in these specific neurons. For example, *lin-11* is sufficient for specification of the ASG neuron, but in AWA neuron LIN-11 activates *unc-86* (Figure 1.8) (Hobert et al. 1998; Melkman 2005; Sarafi-Reinach et al. 2001; Sarafi-Reinach and Sengupta 2000). This finding suggests that spatiotemporal regulation of *lin-11* is very complex and varies from cell to cell.

In addition to the sensory neurons, the expression of *lin-11* is also reported in the ventral cold motor (VC), pioneering, thermosensory, and olfactory neurons (Table 1.2). The 5'UTR region of *lin-11* harbors *cis*-regulatory sequences that drive GFP reporter

expression in the VC motoneurons; however, little is known regarding the purpose *lin-11* serves in these neurons(Marri and Gupta 2009). In the pioneering neurons, AVG and PVP, the expression of *lin-11* has also been reported (Hutter 2003; Hutter et al. 2005). In AVG neurons, LIN-11 inhibits the *unc-86* expression and prevents it from acquiring RIR cell fate, while the role of *lin-11* in PVP neurons is not clear (Figure 1.9). It has been reported that elimination of the pioneering neurons results in midline crossing defects exhibited by several interneurons and motoneurons (Durbin 1987). In thermosensory neurons, the function of *lin-11* is required for AIZ specification (Hobert et al. 1998; Sarafi-Reinach et al. 2001). *lin-11* mutants display thermophilic behaviour. Little is known about the regulatory pathway governing *lin-11* in AIZ interneurons.

Finally, the expression of *lin-11* is also reported in various other neurons including RIC, AVA, and AVE; however, little is known about the regulatory mechanism that governs *lin-11* in these neurons. Together, the regulation and function of *lin-11* in the nervous system of *C. elegans* is very complex and not well understood; thus, additional work is required to further our understanding in the regulatory mechanisms of *lin-11*. In section 5.4.2, I have summarized few proposed experiments.

5.2 LIM domains are essential however proline-rich region is dispensable for LIN-11 function

The LIM domain is a unique double-zinc finger motif found in many different proteins, including homeodomain transcription factors, kinases, and adaptors (Dawid et al. 1998; Gong and Hew 1994; Hobert and Westphal 2000). The LIM domain has conserved patterns of cysteine and histidine residues (Figure 4.1) and binds to Zn^{2+} ions for structural integrity, but it can also bind to Cd^{2+}, Co^{2+}, and Cu(I)- (Kosa et al. 1994; Sadler 1992). Proteins containing LIM domain are involved in a variety of biological processes including regulation of gene transcription, cytoskeleton organization, cell lineage specification, and organ development (Gong and Hew 1994).

In *C. elegans*, little is known regarding the function of the LIM domain of LIN-11 in tissue specification. To better understand their roles in tissue specification, I created two rescue plasmids, pGLC89 and pGLC90, that lacked the LIM domains and rescued the *lin-11(n389)* phenotype. Rescue experiments revealed that LIN-11ΔLIM is not capable of restoring *lin-11* function in either the neuronal or reproductive tissues based on gene expression patterns and behavioral assays (Figures 4.2, 4.4 and 4.6). A similar finding was also reported for *Lhx* gene *Apterous* in *D. melanogaster* (O'Keefe et al. 1998). Thus, these findings suggested that LIM domains of LIN-11 are important for its function in the specification of neuronal and reproductive tissues.

Proline-rich regions (PRR) are found in many proteins, and in many cases are thought to function as docking sites for signaling modules (Kay et al. 2000), and in recruitment of proteins during transcription (Sudol et al. 2001), signaling cascades (Ball et al. 2002; Tu et al. 1998), cytoskeletal rearrangements (Holt 2001; Renfranz and Beckerle 2002), and other key cellular processes (McPherson 1999; Zarrinpar et al. 2003). My work demonstrates that eliminating the PRR of LIN-11 does not affect its function in the nervous and reproductive tissues. Phenotypic rescue experiments revealed that LIN-11ΔPro is capable of restoring *lin-11* function, as demonstrated by gene expression patterns and behavioral assays (Chapter 4). A closer examination revealed that the rescue animals were able to lay eggs due to the proper formation of the vulva organ, which is defective in *lin-11* mutants. In the nervous system, the rescue animals expressed *odr-2::dsRED* reporter in a subset of *lin-11*-specific neurons, namely AVG, RIF, and AIZ. Furthermore, rescued animals show improved thermotactic behavioral responses, as judged by their ability to migrate to the zone of the cultivated temperature, when they are placed in a temperature gradient environment. Thermotactic behavioral quantification revealed that LIN-11ΔPro is able to restore the defect associated with the AIZ interneuron in *lin-11* mutant animals.

My experiments has shown that the functions of LIM domains are critical, while the function of PRR may not be as significant. Truncating the LIM domains did not restore the function of LIN-11 in the neuronal and reproductive tissues, whereas the removal of

the PRR does not show no visibleor obvious effects in these tissues. This might suggest that the PRR is nonessential in *C. elegans*. Additional experiments are required to test this possibility and to identify the exact role of PRR in development. Three popular hypotheses have been proposed in other eukaryotes that can be used to explain the role of LIN-11 PRR in development. These hypotheses are based on the assumption that the PRR tails of several proteins interact to form a complex with indefinite stoichiometry (Figure 5.1): (i) proteins with PRR termini may bind to multiple modules and enhance specificity, (ii) the PRR may bind to multiple recognition surfaces during protein-protein interaction and enhance specificity, (iii) PRR extends the interaction surface with the peptide to include residues beyond the PRR core (Zarrinpar et al. 2003; Williamson 1994). To this end, it is plausible that the PRR may enhance protein-protein binding specificity, and although this may not be essential but it could function together with the LIM domains in mediating protein-protein interactions.

5.3 *Drosophila melanogaster dLim1* and mouse *Lhx1* can replace *lin-11* function

High sequence conservation between *lin-11*, *dLim* and *Lhx1* suggested that *lin-11* orthologs may able to functionally replace *lin-11* role in *C. elegans*; likely due to their ability to interact with conserved gene network. To test this assumption, I created four plasmids that contained tissue specific *lin-11* promoters in frame with *Lhx1* or *dLim1* cDNA's (Chapter 4). Both, *dLim* and *Lhx1* restored the *lin-11* mutant phenotype in both the reproductive and neuronal tissues. A closer examination revealed that the rescue animals were able to lay eggs as a result of the proper formation of the vulva organ, which is defective in *lin-11* mutants. In the nervous system, the rescue animals expressed the *odr-2::dsRED* reporter in a subset of *lin-11* specific neurons, namely AVG, RIF, and AIZ. In *lin-11* mutants, these neurons do not acquire the proper cell fate and as a result, they do not express the *odr-2::dsRED* reporter. Furthermore, rescued animals show improved thermotactic behavioral responses, as judged by their ability to migrate to the zone of the cultivated temperature, when placed in a temperature

gradient environment. Thermotactic behavioral quantification revealed that both *Lhx1* and *dLim1* were sufficient to restore the defect associated with the AIZ interneuron in *lin-11* mutant animals. Together, these findings suggest that the function of *lin-11* can be restored in both neuronal and reproductive tissues by its orthologs, *dLim1* and *Lhx1*.

5.4 Future directions

As discussed above, my work has made significant contributions to understanding of the regulation and function of *lin-11* in the neuronal and reproductive tissues. While I have answered several questions, many others are still unresolved and summarized below. I have provided brief justification for these along with proposed experiments.

5.4.1 Examine the fates of *lin-11* neurons using additional neuronal reporters

In Chapter 3, I have shown that intron 3 expression is present in seven amphid neurons, namely ADL, ADF, AVJ, RIC, AIZ, RIF and AVG. Intron 3-driven *lin-11cDNA* rescues differentiation defects in three of these neurons (AIZ, RIF, and AVG), as judged by the expression of the *odr-2::dsRED* reporter. Since *lin-11* phenotypic rescue relied on a single neuronal reporter, it is unclear whether the *lin-11-int3::lin-11cDNA* is capable of rescuing the remaining four neurons (ADL, ADF, AVJ, and RIC) as well. Thus, additional neuron-specific reporters, such as *gpa-3* (ADL and ADF) (Hilliard et al. 2004), *ceh-10* (AVJ) (Forrester et al. 1998) and *fax-1* (RIC) (Wightman et al. 2005), should be used to test *lin-11* rescued phenotype.

Similarly, in Chapter 4, my work demonstrated that *lin-11* intron 3-driven *dLim1::cDNA and Lhx1::cDNA* are able to restore the *odr-2::dsRED* expression in AVG, RIF and AIZ neurons. However, it is unclear whether ADL, ADF, AVJ, and RIC neurons are also rescued by these proteins in *lin-11(n389)* mutants. Thus, the previously discussed neuron-specific reporters should be used in these studies as well. Together, these proposed experiments will reveal whether transgene expression (i.e., *lin-11*, *dLim1*, and *Lhx1*) in

lin-11-null mutants rescues defects in all *lin-11* amphid neurons, thereby enhancing our understanding of regulation and evolutionary conservation of *Lhx* genes.

5.4.2 Identify *cis*-regulatory sequences for *lin-11* neurons other than RIC and AVG

My work on the dissection of *lin-11* regulation has led to the identification of two introns (3 and 7) that promote expression in a total of seven amphid neurons. However, since *lin-11* is also expressed in many other neurons, namely AVA, AVE, AWA, ASH in the head and PVP, PVQ, and DVA/DVC neurons in the tail, more work is required to identify additional non-coding regions of *lin-11*, leading to the discovery of new *cis*-regulatory sequences for these *lin-11*-specific amphid neurons. For this, experimental approaches similar to those in Chapter 3 can be taken. For example, intronic sequences will need to be cloned into enhancer assay vectors (Fire lab from June 1995 kit) and expression examined in transgenic animals. Candidate sequences could then be dissected, aligned, and in silico analyzed, to reveal regulatory elements as well as potential transcriptional regulators. These experiments will broaden our understanding of the regulatory mechanisms (gene networks and pathways) that govern *lin-11* function in the nervous system. The results will allow us to explain the mechanistic basis of *lin-11* function in mediating behavioral processes such as thermotaxis, electrotaxis and chemotaxis.

5.4.3 Test additional transcription factors that may regulate *lin-11* expression

In Chapter 3, 114 TFs were predicted that bind to C3-1 and C7-1 regions, by in silico tool (CIS-BP) and ModEncode Chip-Seq database. Experimentally, I examined 16 of these genes and demonstrated function of four of these in *lin-11* regulation. A logical extension of this work would be to test the remaining 74 TFs in the list to identify additional *lin-11* regulators in the neuronal differentiation process. Beyond the list of 90 TFs included in this thesis (see Table 1 in Amon and Gupta (2017b)), several other TFs have also been reported to interact with *lin-11* (Appendix A - Table 1). One approach

would be to carry out RNAi experiments using a neuronal RNAi hypersensitive strain background like *eri-1*, *rrf-3*, *lin-15b* and *lin-35* (Kennedy et al. 2004; Simmer et al. 2003; Lehner et al. 2006). Genes that alter *lin-11* expression can then be followed to study their biological roles, as well as interactions with other regulators of *lin-11*.

5.4.4 Determine whether SKN-1, CRH-1 and CEH-6 directly regulate *lin-11* expression

I have demonstrated that *skn-1*, *crh-1* and *ceh-6* negatively regulate *lin-11* expression in RIC interneuron by binding to the C3-1 region and *ces-1*, positively activating the expression of *lin-11* in AVG neuron by binding to C7-1 region (Chapter 3). However, it is unclear whether these TFs directly or indirectly bind to the enhancers within the intron 3 or 7. With molecular techniques that are commonly used to investigate DNA-protein interactions such as yeast one-hybrid (Y1H), chromatin immunoprecipitation (ChIP), and DNA electrophoretic mobility shift assay (EMSA), one could test whether *lin-11* regulators bind to predicted TF sites.

The knowledge of binding of TFs to *lin-11 cis*-regulatory regions will allow for building a precise genetic model of gene regulation of *lin-11* in the nervous system. The data will form the basis of future experiments to investigate how *lin-11* interacts with other genes to regulate neuronal differentiation processes.

5.4.5 Test whether a single LIM domain of LIN-11 is biologically functional

My work demonstrated only that both LIM domains are required for the function of LIN-11, as removal of these domains results in a lack of rescue of neuronal and reproductive system defects. However, it is unclear whether both domains are necessary, or whether a single domain can mediate LIN-11 function. It is also unclear whether there is a preference for either one of these domains, i.e., LIM1 or LIM2. This proposed experiment can be tested by creating plasmids lacking one domain at a time and testing

rescue of *lin-11(n389)* phenotype in the neuronal and reproductive tissues.

5.4.6 Examine whether *lin-11* can functionally replace *D. melanogaster* *dLim1* and mouse *Lhx1*

My work has shown that *lin-11* orthologs, *dLim1* and *Lhx1*, are able to rescue the *lin-11* phenotype in the neuronal and reproductive tissues. However, reciprocal experiments, i.e. transforming *D. melanogaster* and mouse tissues with *lin-11*, have not been performed. It will be interesting to test whether *lin-11* rescues the phenotypes of *dLim1* in *D. melanogaster* and *Lhx1* in mice. This will require generating plasmids expressing *lin-11* under the control of *dLim1* and *Lhx1* promoters/enhancers. These experiments are significant in detailed analysis of the evolutionary conservation of *Lhx* genes among eukaryotes, more specifically whether the invertebrate *C. elegans lin-11* is capable of restoring the functions of other orthologous genes in invertebrate (i.e., *D. melanogaster* *dLim1*) and vertebrate (i.e., mouse *Lhx1*) species.
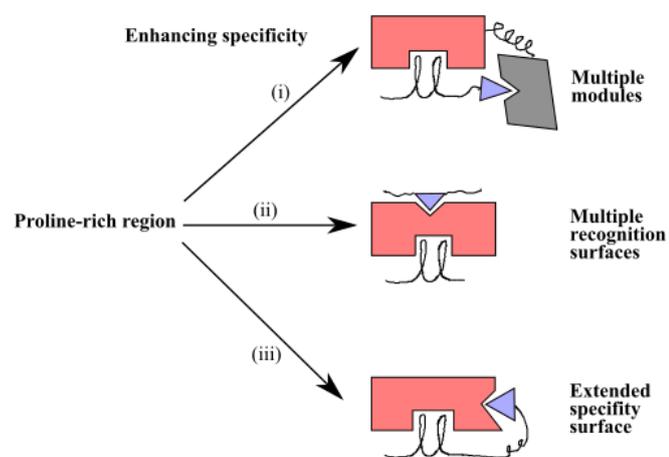
FIGURE 5.1: Proposed mechanisms for enhancing the specificity of protein-protein recognition by proline-rich region (PRR). Three common hypothesis are proposed to explain the function of PRR or a protein (Zarrinpar et al. 2003). (i) PRR may function in binding to multiple active modules and enhance specificity of protein-protein complex. (ii) PRR may recognize multiple sites on the surface of protein complex and (iii) extend the interaction surface with the peptide to include residues beyond the PRR core. Red and greay squares represent interacting proteins, purple triangle represent PRR and LIM domains are represented by two L-lines.

Appendix A

Supporting Materials

TABLE A1.1: List of predicted factors that may interact with *Lhx* genes in *C. elegans.*

Gene	Predicted Genes	Reference
lim-6	*ahr-1, ceh-36, cog-1, daf-7, die-1, gcy-5, ldb-1, lin-49, lsy-6, odr-7, unc-25, unc-37, zig-2, zig-3, zig-4*	(Zhong and Sternberg 2006) and wormbase
lim-7	*ceh-12, ceh-6, cfi-1, cwn-2, dbl-1, efl-1, grd-1, grd-2, hpl-1, ina-1, odb-1, let-23, ptp-3, unc-40, unc-86, vab-10, wrt-6*	(Zhong and Sternberg 2006) and wormbase
lim-4	*ceh-10, ceh-37, ceh-6, had-4, kal-1, lin-14, mab-5, odr-7, slt-1, str-1, str-2, hlh-1*	(Zhong and Sternberg 2006) and wormbase
ttx-3	*alr-1, cdc-42, cdh-4, ced-10, ceh-10, ceh-14, ceh-16, ceh-2, ceh-20, ceh-23, ceh-32, cog-1, crb-1, daf-7, dbl-1, egl-44, glp-1, grd-11, grd-2, hen-1, hlh-14, hlh-3, hmp-2, ina-1, kal-1, kin-1, ldb-1, let-23, let-60, lim-7, lin-11, lin-12, odr-7 ppk-1, ptp-3, ser-2, slt-1, sma-6, sra-11, tax-6, tbx-2, unc-17, unc-62, vab-1, vab-15, wrt-6*	(Zhong and Sternberg 2006) and wormbase
mec-3	*ceh-14, egl-5, flp-4, mec-17, mec-4, mec-7, odr-7, sem-4, unc-86, unc-97*	(Zhong and Sternberg 2006) and wormbase
lin-11	*cdh-3, ceh-16, ceh-32, cfi-1, cnd-1, crb-1, daf-7, dbl-1, glp-1, grd-11, grd-2, lag-1, ldb-1, let-23, let-60, lin-12, lin-17, lin-28, lin-29, lin-39, lin-4, mab-5, mls-1, odr-7, pop-1, sem-4, syg-1, syg-2, ttx-3, unc-86, vab-3, wrt-6*	(Zhong and Sternberg 2006) and wormbase
ceh-14	*ldb-1, lim-7, cdc-42, ced-10, ceh-6, dbl-1, glp-1, mec-3, odr-7, ptp-3, rfc-1, sax-3, smp-1, sth-1, ttx-3, unc-40, unc-6, zig-4, ceh-17*	(Zhong and Sternberg 2006) and wormbase

Appendix B

Supporting Materials

RESEARCH PAPER

Worm 2:3, e24558; July/August/September 2013; © 2013 Landes Bioscience

A microfluidic phenotype analysis system reveals function of sensory and dopaminergic neuron signaling in *C. elegans* electrotactic swimming behavior

Sangeena Salam,[1] Ata Ansari,[1,3] Siavash Amon,[1] Pouya Rezai,[2] P. Ravi Selvaganapathy,[2] Ram K. Mishra[3] and Bhagwati P. Gupta[1,*]

[1]Department of Biology; McMaster University; Hamilton, ON Canada; [2]Department of Mechanical Engineering; McMaster University; Hamilton, ON Canada; [3]Department of Psychiatry and Behavioral Neuroscience; McMaster University; Hamilton, ON Canada

Keywords: *C. elegans*, nematode, microfluidics, electrotaxis, neuronal signaling, dopamine signaling, neurodegeneration, neurotoxin, 6-OHDA, MPTP, Rotenone

The nematode (worm) *C. elegans* is a leading multicellular animal model to study neuronal-basis of behavior. Worms respond to a wide range of stimuli and exhibit characteristic movement patterns. Here we describe the use of a microfluidics setup to probe neuronal activity that relies on the innate response of *C. elegans* to swim toward the cathode in the presence of a DC electric field (termed "electrotaxis"). Using this setup, we examined mutants affecting sensory and dopaminergic neurons and found that their electrotactic responses were defective. Such animals moved with reduced speed (35–80% slower than controls) with intermittent pauses, abnormal turning and slower body bends. A similar phenotype was observed in worms treated with neurotoxins 6-OHDA (6- hydroxy dopamine), MPTP (1-methyl 4-phenyl 1,2,3,6-tetrahydropyridine) and rotenone (20–60% slower). We also found that neurotoxin effects could be suppressed by pre-exposing worms to a known neuroprotective compound acetaminophen. Collectively, these results show that microfluidic electrotaxis can identify alterations in dopamine and amphid neuronal signaling based on swimming responses of *C. elegans*. Further characterization has revealed that the electrotactic swimming response is highly sensitive and reliable in detecting neuronal abnormalities. Thus, our microfluidics setup could be used to dissect neuronal function and toxin-induced neurodegeneration. Among other applications, the setup promises to facilitate genetic and chemical screenings to identify factors that mediate neuronal signaling and neuroprotection.

Introduction

The survival and functioning of living organisms depend on their ability to continuously monitor their surrounding environment. This process is mediated by the nervous system that detects environmental stimuli and depending upon the context and the experience of the animal generates an appropriate response. Understanding how neurons respond to stimuli and initiate a specific behavioral outcome is complicated in vertebrates due to the presence of a very large number of neurons and complex interconnections. Invertebrate animal models, such as the nematode (worm) *C. elegans*, offer a simpler nervous system and a wide range of experimental techniques to dissect the neuronal basis of behavior. The adult *C. elegans* hermaphrodite contains 302 neurons, whose morphology and inter-connections are well established.[1,2] The animal has a short life cycle (2.5–3 d) making it possible to study developmental processes relatively quickly. In addition, it can be manipulated in the laboratory using a number

of powerful genetic and genomic tools. These features have greatly facilitated the study of neuronal cell fate specification and cell signaling.

The sensory neurons in *C. elegans* that mediate odor detection (chemosensory) and physical contacts (mechanosensory) have been studied in significant detail. Many of these have ciliated dendrites that are directly or indirectly exposed to the external environment and facilitate detection of environmental stimuli by activating signal transduction pathways. The signaling process is mediated by several proteins including G-protein coupled receptors (GPCRs) and transient receptor potential (TRP) channels.[3-5] In hermaphrodites eight of the mechanosensory neurons (two pairs of CEPs and one pair of ADE in the head region and one pair of PDE in the posterior region) produce dopamine (DA) neurotransmitter.[6] All of these neurons have ciliated endings that are embedded in the cuticle and their processes extend to the nose tip (CEPs) and along the anterior and posterior lateral midlines (ADE and PDE, respectively).[5] The dopaminergic neurons

*Correspondence to: Bhagwati P. Gupta; Email: guptab@mcmaster.ca
Submitted: 02/14/13; Revised: 04/02/13; Accepted: 04/04/13
http://dx.doi.org/10.4161/worm.24558

www.landesbioscience.com   Worm   e24558-1

107

modulate different behavioral responses of the animal such as foraging, locomotion rate, egg laying, defecation and swimming in a liquid environment.[7-11]

The traditional assays to measure behavioral activities of *C. elegans* involve qualitative and quantitative analyses of phenotypes using manual and semi-automated approaches. These methods are slow, labor-intensive and prone to errors. Consequently, they represent a significant bottleneck in analyzing a large number of animals in a rapid and unbiased manner. Recent advances in microfluidics research has made it possible to perform many of the routine and laborious experiments in *C. elegans* automatically, precisely and in a high-throughput manner.[12-15] Microfluidics offers unparalleled control over manipulation of worms because of its ability to precisely control microscopic volumes of fluids inside minute chambers and channels that are comparable to the size of worms. This allows efficient handling of worms and confers greater control over confinement, treatment and observation at a high resolution. Furthermore, microfluidics deals with tiny volumes of liquid resulting in reduced cost of media and chemical usage. Microfluidics devices have a wide variety of applications in research involving *C. elegans*. These include growth studies, mutant screening, nanosurgery, neuronal imaging and movement analysis.[16-19]

We recently demonstrated that low voltage direct current (DC) and pulsed DC electric fields inside a microfluidic channel cause *C. elegans* to swim toward cathode (termed "electrotaxis").[20,21] The electrotactic responses of some of the nematodes, including *C. elegans*, were examined earlier using an open gel surface Petri plate setup.[22-24] It was noted that although animals generally were attracted toward the cathode, they did not travel directly toward it along the imaginary axis from anode to the cathode. Instead, they were found to travel at an angle with respect to this line. The angle of motion was proportional to the strength of the electric field making the response more complex to analyze.[24] Unlike the open gel surface setup, the microchannel environment streamlines the electric field along its axis and stimulates worms to swim in a straight line with a characteristic sinusoidal pattern. This simplifies the movement analysis of animals by allowing precise quantification of parameters such as speed, body bend frequency and reversals. Furthermore, the electrotaxis response is instant, reversible and highly robust. As a result, abnormalities in neuronal signaling can be identified fairly quickly and reliably. Finally, microfluidics is amenable to parallelization and automation thereby accelerating the pace of discovery.

The precise role of electrotaxis behavior in worms is currently unknown but research involving plant and insect parasitic nematodes has suggested that it could facilitate host finding.[23,25] Studies in *C. elegans* have revealed that such a response in an open gel surface environment is mediated by amphid neurons.[24] In this study, we examined the electrotactic swimming behavior of worms in a microfluidic channel device using neuronal mutants and found that in addition to amphid sensory neurons dopaminergic neurons are also involved in this process. The role of dopaminergic neurons was further investigated by exposing worms to three different neurotoxins, 6-OHDA (6-hydroxy dopamine), MPTP (1-methyl 4-phenyl 1,2,3,6-tetrahydropyridine) and rotenone (a

pesticide). These toxins cause degeneration of DA neurons similar to that shown in mammalian models.[26-28] A common mode of their action appears to be the inhibition of mitochondrial respiratory chain and generation of reactive oxygen species (ROS).[29] We found that toxin-treated worms had abnormal swimming behavior. A comparison with plate-based behavioral assays revealed that the microfluidics assay is highly sensitive in detecting movement defects. The phenotypic analyses of worms were performed by quantifying parameters that eliminated subjectivity and bias. Lastly, we investigated the role of a neuroprotective compound acetaminophen and found that it suppresses the effect of toxins in the channel assay. Together, these results demonstrate that the microfluidic electrotaxis is a sensitive assay to study neuronal signaling and toxin-induced neuronal damage, as well as to identify neuroprotective drugs in *C. elegans*.

Results

The microfluidic device used in our electrotaxis assay (**Fig. 1**) provides an environment in which the electric field streamlines are confined in the axial direction of the channel. This results in a uniform stable field that stimulates worms to move along the channel length. Furthermore, the narrow diameter (300 μm) ensures motion in a near straight line fashion. We have earlier shown that this setup allows worms to swim toward the cathode when exposed to a low voltage DC electric field.[20] The electrotactic swimming response of *C. elegans* is robust, continuous and occurs at a characteristic speed. Changing the direction of the electric field causes the worm to stop transiently, turn toward the cathode and resume motion (**Fig. 1D and E**). Therefore, any deviation from these characteristics can be considered as an abnormality. As described in Materials and Methods, we characterized the electrotaxis behavior of worms using four parameters, namely electrotaxis speed, turn time, body bend frequency and electrotaxis time index (ETI).

Initially, we investigated the behavior of wild-type N2 worms in the channel in the absence of an electric field. The animals showed random swimming and turning activities with a mean body bend frequency of 1.3 (n = 6) (**Video 1**). Due to multiple reversals, they failed to cover long distance in any one direction. Thus, the electrical stimulus is required to propel worms in a directed manner.

The electrotactic swimming of *C. elegans* in the channel depends on its intact neuronal and muscular systems.[20] Consequently, defects in any of these components could alter the swimming behavior. While the electrosensory defect may result from the inability of sensory neurons to receive and process electrical signals, general locomotion defects could arise due to problems in the neuromuscular system that controls motor responses such as speed and amplitude of motion. Hence, electrosensory mutants will have difficulty in sensing the direction of the electric field. Such animals are expected to swim toward either pole while reversing direction frequently. However, the locomotory mutants should recognize the electric field polarity and move specifically toward cathode, albeit at a reduced pace. Our analysis of the neuronal mutants in following sections agrees with these

possibilities and demonstrates that the microfluidics setup can be used to identify and characterize new genes mediating electrotaxis behavior.

Amphid sensory neurons mediate electrosensory responses in a microfluidic channel device. Although the biological basis and the mechanism of electrotaxis is poorly understood, a subset of neurons in the anterior ganglion were found to respond to the electric field stimulus in an open gel surface Petri dish setup.[24] We used *osm-5* mutants to examine the role of sensory neurons in our channel assay. The *osm-5* gene encodes an intraflagellar transport protein that is homologous to human IFT88 and is required for cilia formation in sensory neurons including the amphids.[30] The *osm-5* animals appeared very active and in the absence of the electric field exhibited swimming behavior similar to wild-type N2 (body bend frequency of 1.8 Hz, n = 6; see above for N2 data) (**Fig. S1**). However, in the presence of the electric field, the animals showed severely defective electrotactic response. The average speed of *osm-5* was nearly 60% lower than the wild-type N2 (n = 11) (**Fig. 2**). Whereas N2 worms swam straight toward cathode without pausing (**Video 2**), the *osm-5* worms stopped and reversed direction many times (number of reversals for each animal ranged between 3–20, n = 11) (**Video 3**). Few of them became immobile after slight initial swimming and did not recover (18%, n = 11) (**Video 4**). In addition to spontaneous turning and lack of motion, *osm-5* worms also exhibited intermittent pauses, abnormal body postures and swimming in wrong direction (i.e., toward anode) indicating that they lacked a sense of direction. Consequently, the turn time of animals was highly variable (**Fig. 2**). To further demonstrate the electrotaxis defect in *osm-5* animals, we computed the time of all cathode-directed swimming events and determined ETI (see Materials and Methods). As expected, the ETI of *osm-5* was greatly reduced compared to N2 (p < 0.0001) (**Fig. 3**). As a control, we also tested a non-neuronal mutant *him-8(e1479)* that produces increased frequency of males due to defects in X-chromosome segregation.[31] The speed, turn time and ETI of *him-8* were comparable to N2 (**Figs. 2 and 3**).

We examined another neuronal mutant *lin-11* that is weakly uncoordinated and exhibits chemosensory and thermosensory defects.[32-34] *lin-11* is a founding member of the LIM Homeobox family of transcription factors.[35] It is expressed in a subset of neurons in the head ganglion (including sensory neurons ADF and ADL, interneurons AIZ, RIC, AVG, AVA and AVE, and chemosensory neuron ASG) and is necessary for their differentiation.[32,33] We found that *lin-11(n389)*-null mutants[35] had severely defective electrotactic response. Some animals showed no reaction to the stimulus (17%, n = 23) (**Videos 5 and 6**) whereas others moved in the channel but with multiple pauses and significantly reduced speed (76 μm/sec, n = 19, compared with N2 control: 141 μm/sec, n = 12) (**Fig. 2**). Additionally, *lin-11* worms were also defective in sensing the direction of the electric field. Upon switching the field polarity,
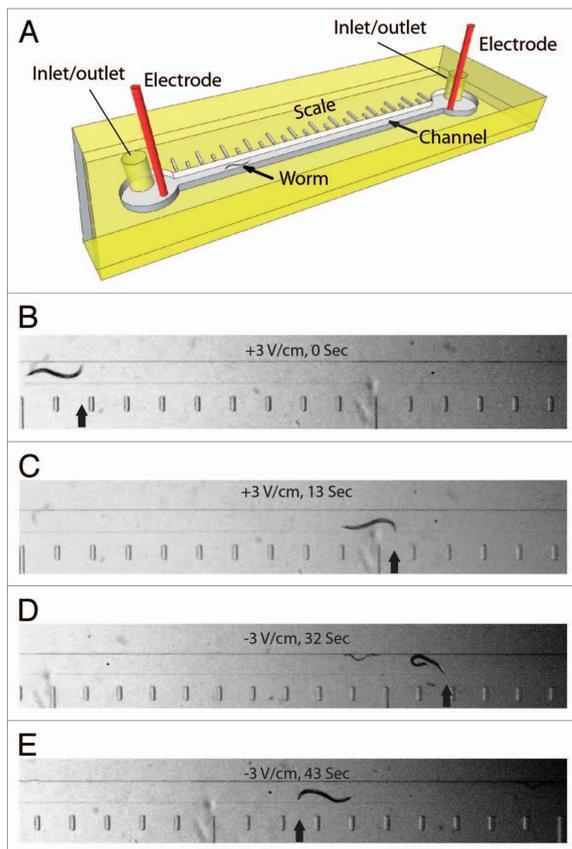
Figure 1. Microfluidic electrotaxis setup. (**A**) A detailed view of the microfluidic device. Worms are loaded and removed through inlet/outlet tubes. Electrotaxis is performed in the channel (a worm is shown). The scale along the length of the channel is used to determine the speed. (**B–E**) Snapshots of a worm in the channel during electrotaxis. Scale bar is visible on the bottom. The electric field voltage and time are shown in each panel. The head of the worm is marked by an arrow. Reversal of the electric field polarity (**D and E**) causes the worm to switch its direction of motion.

the animals either took a very long time to turn (average turn time 49.2 sec, n = 16, compared with 10.3 sec, n = 12 for control) (**Fig. 2**) or failed to turn at all (16%, n = 19). In agreement with these findings, the ETI of *lin-11* was greatly reduced (p < 0.0001) (**Fig. 3**). Taken together, these results demonstrate that *osm-5* and *lin-11* are required for the electrosensory response of *C. elegans* in the microfluidic channel assay.

Defects in dopamine signaling affect electrotaxis of animals. DA signaling is involved in modulating the locomotion of *C. elegans* in response to environmental stimuli.[7] Therefore, we examined the electrotaxis swimming response of DA pathway mutants in our microfluidic channel setup. Animals having mutation in

Figure 2. Electrotaxis speed and turn time responses of wild-type N2 and mutant animals. The lower and upper lines of each box represent the 25th and 75th quartile of data sample, respectivel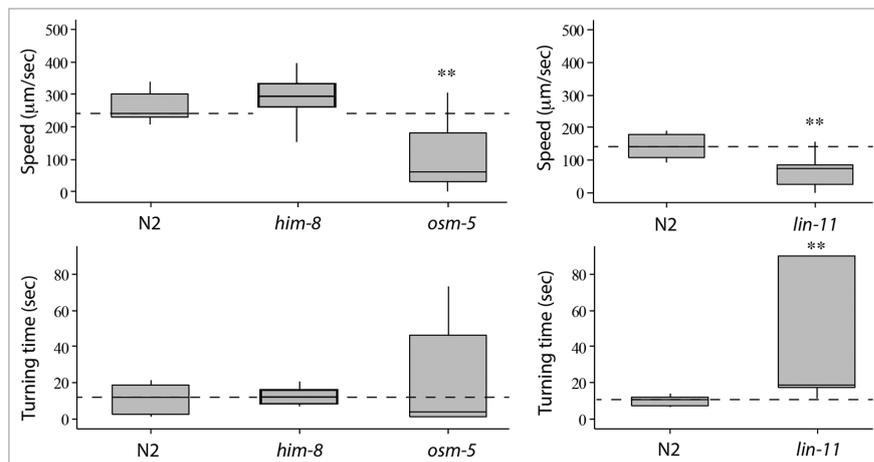y. The middle line inside the box marks the median. The end points of the vertical line (both top and botom) are the maximum and minimum data points of the sample, respectively. The dotted horizontal line corresponds to the median of the control. The *him-8* worms are comparable to wild-type, whereas *osm-5* and *lin-11* move slower than N2 and take much longer time to turn. Statistically significant responses are marked with stars (*: p < 0.05, **: p < 0.01, ***: p < 0.0001).
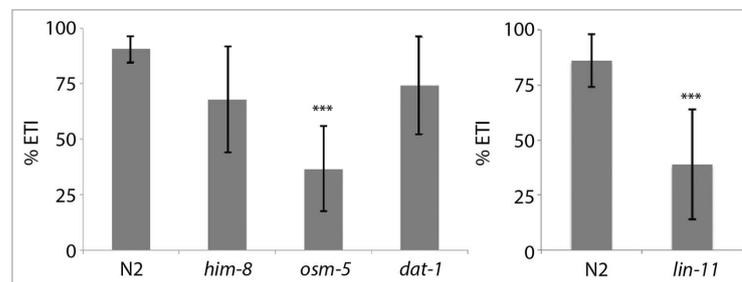

Figure 3. Electrotaxis time index (ETI) of N2 control and mutant animals. The *him-8* and *dat-1* are normal whereas *osm-5* and *lin-11* show a defective response (***: p < 0.0001).

the *cat-2* gene, the tyrosine hydroxylase required for DA biosynthesis,[36] showed normal response (average speed 296 μm/sec and turn time 4.1 sec, n = 10 for the *e1121* hypomorph); however, dopamine transporter mutants (*dat-1(ok157)*) showed a defective phenotype (average speed 189 μm/sec and turn time 26.1 sec, n = 10) compared with control animals (**Fig. 4**). Since DAT-1 mediates the reuptake of DA from the synaptic cleft back into the presynaptic terminal,[37] it is possible that in the absence of DAT-1 function, extracellular DA alters the activity of certain post-synaptic neurons, thereby causing reduced electrotaxis speed in the microfluidic channel. Additionally, DA could stimulate certain dopamine receptors on motor neurons resulting in a slower speed of animals. The lack of electrotaxis phenotype in *cat-2(e1121)* animals may be attributed to the fact that DA level

is not completely abolished (40% residual DA compared with wild type[38]). Because *dat-1* mutants show reduced speed but no impact on sensing the electric field polarity, it suggests that DA signaling modulates locomotion without affecting the electrosensory response of animals. This conclusion is supported by the normal ETI response of *dat-1* animals (p = 0.1253) (**Fig. 3**).

In addition to mutants, we also used three different chemical compounds 6-OHDA, MPTP and rotenone that are toxic to DA neurons. Previous work in vertebrates has shown that these chemicals cause degeneration of DA neurons in the substantia nigra region of the brain.[39,40] 6-OHDA is preferentially taken up by DA neurons via the DAT transporter. Once inside the DA neuron, it causes multiple reactions including inactivation of the mitochondrial respiratory chain leading to an increase in ROS level.[29,41,42]

In the case of MPTP, it is metabolized into an active toxic product MPP⁺ that enters DA neurons through the DAT-1 transporter. MPP⁺ induces inhibition of mitochondrial respiratory enzyme complex I and an increase in ROS production.[43,44] Exposure to rotenone in rat and *Drosophila melanogaster* models has shown to cause apoptosis and oxidative damage of DA neurons.[39,45] *C. elegans* DA neurons are equally sensitive to the above three neurotoxins and undergo degeneration upon exposure.[26,27]

To carry out the electrotaxis assay on worms exposed to 6-OHDA, MPTP and rotenone, we first optimized chemical exposure conditions (see Materials and Methods). Toxin-treated worms were placed inside the channel without any pre-selection and their electrotactic responses were analyzed. We found that animals had defects in swimming behavior. The phenotypes included slower speed, intermittent pauses and reduced sensitivity. Exposure to 6-OHDA (either 1 h or 4 h at 100 μM concentration) caused a significant reduction in the speed (40–60% slower, 126–174 μm/sec average speed) of animals (n = 12 for 1 h condition and 15 for 4 h condition) without altering the turn time (**Fig. 5**). Besides reduced speed, we also observed other defects in movement. Frequently, animals showed incoherent electrotaxis characterized by active sinusoidal motion followed by periods of slow responses or a lack of activity (**Videos 7 and 8**). Partial paralysis was also observed where the posterior half of the body was rigid such that the worm appeared to drag itself while moving (**Video 9**). In addition, we also detected phenotypes such as sudden freeze (**Video 10**), tremor (**Video 11**) and a lack of motion (**Video 12**). Not all phenotypes were observed in every animal and, furthermore, these were specific to microfluidic channel environment. Animals grown on Petri plates did not show any such phenotype. In few cases, we also measured body bend frequency and found it to be lower in exposed worms compared with controls (average frequency 0.2 Hz for 6-OHDA 4 h exposure, n = 11 animals and 1.8 Hz for N2, n = 10 animals).

Treatments of MPTP and rotenone for 8 h caused similar defects (roughly 40% slower speed in each case, MPTP: 178 μm/sec, rotenone: 160 μm/sec; n = 12 and 13 worms, respectively), albeit the reduction in speed was somewhat less compared with 6-OHDA (see above, **Fig. 5**). This is different from the 6-OHDA exposure that caused maximum defect in as little as 1 h (see above). The turn time was also affected (up to 2-fold slower for each toxin; MPTP: 19 sec; rotenone: 15 sec) (**Fig. 5**). Consistent with slow electrotaxis swimming speed, the body bend frequency of animals was also reduced (MPTP 8 h: 1.0 Hz, n = 3, compared with 1.6 Hz, n = 4 for N2 control; rotenone, data not shown).
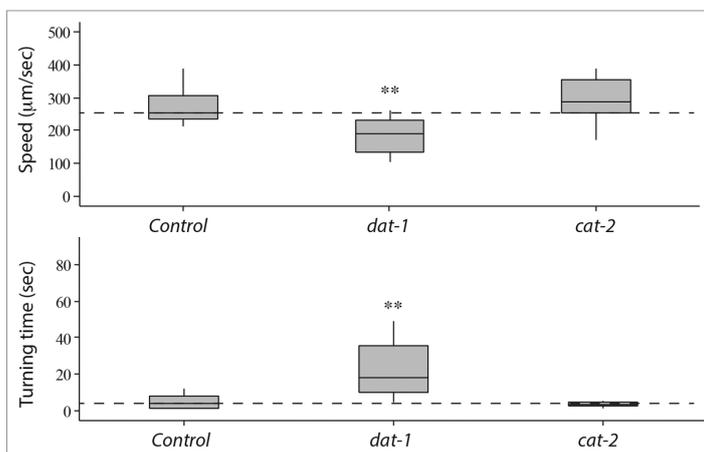

Figure 4. Electrotaxis speed and turn time responses of *dat-1* and *cat-2* mutants. Refer to **Figure 2** for an explanation of the plot structure and statistical significance of data. *dat-1* mutants move slower and have higher turn time compared with controls. However, *cat-2* animals show no such defect.

In summary, our experiments involving DA pathway mutants and neurotoxin-treated worms demonstrate the sensitivity of the microfluidic electrotaxis assay in detecting abnormalities in DA signaling. Quantification of swimming defects allowed us to compare phenotypes in different conditions as well as between different sets of animals. Together, these results showed that a reduction in DA signaling contributes to abnormalities in electrotaxis swimming behavior in the channel assay. We conclude that our microfluidics setup can be used to identify and study factors affecting DA signaling and neurodegeneration in worms.

Plate-based phenotypic analysis of mutants and toxin-treated animals. In addition to the electrotaxis phenotype, we also examined gross morphology and behavior of animals on Petri plates. This was done to investigate whether movement defects are specific to the microfluidic device or observed on open agar gel surface as well. For this, we measured the speed of *osm-5*, *dat-1* and *lin-11* mutants by following their tracks on bacterial lawns (see Materials and Methods). The average speed of *dat-1* and *osm-5* was 162 μm/sec (n = 23) and 183 μm/sec (n = 18), respectively, which is not significantly different from the wild-type N2 (190 μm/sec, n = 19) and *him-8* (188 μm/sec, n = 9) (**Fig. 6A**). However, *lin-11* animals showed a lower speed compared with N2 (40 h stage, average speed 53 μm/sec, n = 16, N2: 77 μm/sec, n = 14) (**Fig. 6A**). This agrees well with earlier studies showing that *lin-11* animals are weakly uncoordinated.[34] Overall, our results provide support to the conclusion that the microfluidic electrotaxis is a sensitive method to detect abnormalities in the swimming response of animals. Furthermore, they reveal that the lack of electrotaxis in some *lin-11* animals and spontaneous reversals in *osm-5* are unique to the channel environment and not observed in plate-based assay.

Next, we examined plate-level responses and cellular defects in toxin-treated worms. In general, worms were healthy and
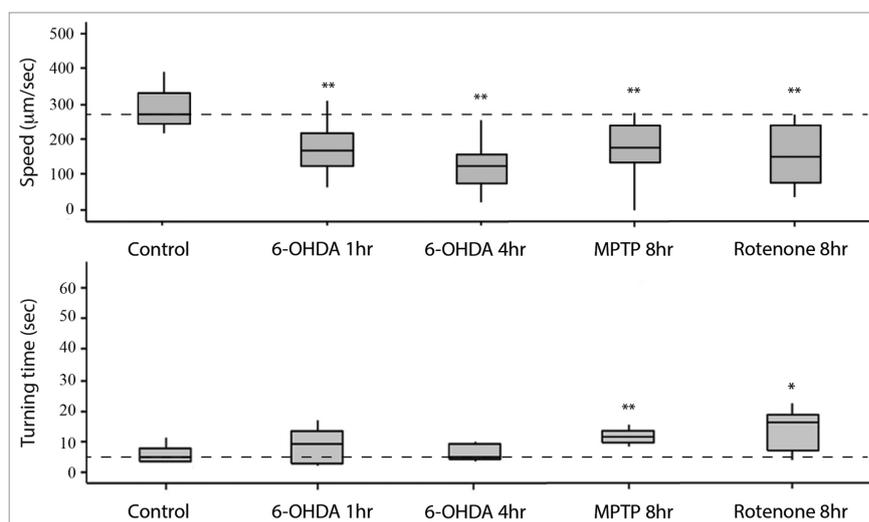
Figure 5. Electrotaxis speed and turn time responses of toxin-treated worms inside the channel. Refer to **Figure 2** for an explanation of the plot structure and other details. Control refers to unexposed wild-type N2 animals. In all cases, the speed of exposed worms is significantly different from the control. A similar phenotype is also observed for turn time of animals.

fertile with no obvious morphological defect at any stage. Except for 4 h 6-OHDA-exposure condition that caused weak uncoordinated movement (Unc) in roughly 40% of the population, in all other cases, animals moved well and were fairly active (**Fig. 6B**). In one case, 6-OHDA 1 h, we also determined the chemotaxis response to NaCl and found it to be similar to untreated N2 control animals (**Fig. 6C**). The number of animals examined for plate-based assays are: 6-OHDA - 605 for 1 h and 420 for 4 h, and N2 control - 220; MPTP - 502 and control - 192; Rotenone - 353 and control - 163. For chemotaxis assays the numbers are: 6-OHDA 1 h - 169 (NaCl) and 34 (water); control - 101 (NaCl) and 33 (water).

To correlate microfluidic behavioral defects in toxin-treated worms with DA neuronal function, the morphology of neurons was also investigated. For this, we used a *dat-1_p::YFP* transgenic strain *(bhEx120)* in which YFP expression is observed in DA neuronal cell bodies as well as their projections (**Fig. 7A**).[26,37] The synchronized L1 stage *bhEx120* animals were exposed to 100 µM 6-OHDA for 4 h and subsequently grown on a standard NG agar plate seeded with OP50 bacteria. The DA neurons were examined in adults. We found an age-dependent increase in neurodegeneration in toxin-treated animals (18%, n = 81 on day 3 and 26%, n = 42 on day 6) (**Fig. 7B–D**). The dendritic processes of CEPs showed variable degeneration such that YFP fluorescence had spotty appearance (**Fig. 7B**). In some cases, the entire dendritic processes were missing (**Fig. 7C**). These phenotypes are similar to that reported earlier[26] and are consistent with the electrotaxis defect in worms that we observed.

Acetaminophen confers neuroprotection against toxins in the microfluidics assay. If toxin-induced neuronal damage affects the electrotaxis behavior of worms, then the phenotype could be suppressed by protecting DA neurons. To examine this possibility, we tested a known neuroprotective compound acetaminophen. Studies in rats and *C. elegans* have shown that acetaminophen has a protective effect against MPP⁺, 6-OHDA and glutamate toxicity in DA neurons.[46-48] We found that pre-treatment of acetaminophen conferred significant protection on DA neurons against all three neurotoxins. The pre-exposed worms were phenotypically normal and had significantly faster speed compared with neurotoxin-treated worms (**Fig. 8**). However, the turn time of animals showed no improvement (data not shown). Because acetaminophen protects DA neurons,[46,48] these results further support the involvement of DA signaling in controlling the electrotaxis behavior of animals. Additionally, our findings show that the microfluidic channel-based assay can be used as a screening tool to identify new chemicals with neuroprotective properties.

Discussion

Microfluidic electrotaxis as a sensitive and reliable assay to detect neuronal activity. We have used a new microfluidics electrotaxis assay to examine neuronal signaling in worms. This unique method allows us to indirectly control the motion inside the channel. In the absence of an electric field, worms show random swimming behavior and cannot be guided in a desired direction. However, when exposed to the field, they move toward cathode in a stereotypic manner and with a constant speed.[20] Previous findings[20,24] and results presented in this paper demonstrate that this process depends on the function of neurons

and muscles. The electrotaxis response is highly stereotypic, reproducible and can be quantified.[20,21] This makes it possible to assess precisely the impact of alterations on worm's behavior by observing its swimming characteristics.

We have shown that the normal electrotaxis response relies on the intact electrosensory system. In the case of *osm-5* (human IFT88 homolog) and *lin-11* (LIM Homeobox family) mutants that affect amphid neurons, animals showed significantly reduced speed and frequently failed to detect the electric field polarity. While the involvement of amphid neurons and *osm-5* in mediating electrosensory behavior on open gel surface was previously reported,[24] our work has revealed novel turning and paralysis phenotypes in *osm-5* and *lin-11* animals in the microfluidic device assay.

We have provided the evidence for the role of DA neuron signaling in modulating electric field-induced swimming responses. This is based on results that DA transporter mutant *dat-1* has a reduced activity as judged by slower speed and higher turn time of animals. A similar phenotype was also observed in worms exposed to DA-specific neurotoxins (6-OHDA, MPTP and rotenone). The low doses of toxins (25–700 μM range, for a maximum duration of 8 h) caused electrotaxis defects in the channel without seriously impacting the growth and viability of animals. Thus, the microfluidic platform can be used as an effective and non-invasive tool to detect neuronal abnormalities in worms.

The effect of neurotoxins on *C. elegans* electrotaxis. Previous studies and our own work have shown that neurotoxins 6-OHDA, MPTP and rotenone cause movement and other abnormalities in *C. elegans*[27,28] (see Materials and Methods). Some of the phenotypes, such as lethality, could be non-specific and may result from exposure to high doses of toxins. Because DA neurons are not required for survival, the viability defect could result from disruption of other cellular processes. Consistent with this, we found that reducing toxin doses eliminated lethality.

Because worms exposed to low doses of toxins appear generally healthy and active on plates, a sensitive assay is needed to monitor DA signaling defects. While one could directly visualize neurons using a GFP reporter, such an approach is slow and subjective. Therefore, it is unlikely to be a sensitive measure of DA neuronal activity. Our work establishes microfluidic electrotaxis as a rapid and sensitive assay to monitor movement and its neuronal basis in worms. The quantification of movement parameters allows comparison between different groups of animals in a reliable manner and eliminates any bias associated with manual counting and judgment.

Aside from the reliable quantitative analysis, our assay provides a unique opportunity to investigate some of the phenotypes caused by DA neuronal loss. During electrotaxis experiments, we observed that toxin-treated worm often failed to coordinate the

Figure 6. Plate-level phenotypes of mutants and toxin-treated animals. (**A**) The results are plotted similar to **Figure 2**. The *him-8, osm-5* and *dat-1* animals are comparable to the N2 control but *lin-11* animals show reduced speed (**: p < 0.01). (**B**) Exposure to 6-OHDA, MPTP and rotenone causes weak Unc phenotype. Except for 6-OHDA 4 h cases that caused roughly 40% of animals to become Unc, all other treatments affected less than 3% of animals. (**C**) The chemotaxis responses of N2 and 6-OHDA 1 h-exposed animals toward NaCl chemoattractant. The toxin-treated worms show normal behavior.

swimming of different parts of their body relative to one another. This lack of synchrony translates into an overall gait abnormality that is characterized by short, staggered movements. We did not observe such a phenotype on the plate level. 6-OHDA-induced gait problems have been observed previously in the rat model but most importantly, this phenomenon relates closely with the shuffling gait that is a classical symptom of human Parkinson disease.[49] The execution of synchronized voluntary movement requires not only proprioceptive feedback from peripheral receptors but also higher-level supraspinal processing that allows for kinesthesia.[50] Defects in kinesthetic ability in toxin-treated worms could explain the etiology behind the short, staggered movements observed during electrotaxis. The gait abnormality observed in toxin-treated worms may result from a loss of proprioceptive ability due to the loss of dopaminergic neurons. Studies have shown that TRPN mechanosensitive ion channels are implicated in nematode proprioception.[51,52] Therefore, in the future it will be useful to examine TRPN function in the electrotaxis behavior.

The sudden freeze, tremor and partial paralysis phenotypes of toxin-treated worms in the channel assay may be reminiscent of bradykinesia in Parkinsonian subjects. In *C. elegans*, sinusoidal movement requires the out-of-phase contraction of dorsal and ventral musculature.[53] In Parkinson patients, D1 and D2 class of DA receptors are proposed as major contributing factors to the development of bradykinesia.[54] The corresponding
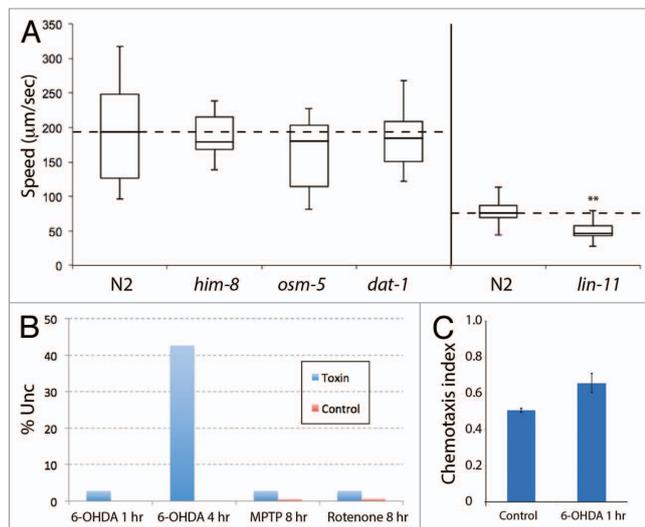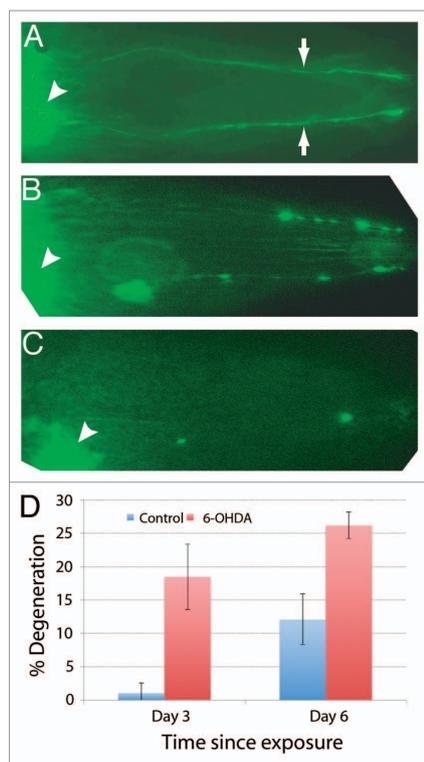
Figure 7. Degeneration of DA neurons in 6-OHDA-exposed worms visualized by *dat-1_p::YFP* expression. In all cases anterior is toward the right. (**A**) In wild-type *bhEx120* animals, YFP fluorescence can be observed in CEP cell bodies (arrowhead) and their processes (arrows). (**B**) Exposure to 6-OHDA causes spotty appearance of CEP neuronal processes indicating degeneration of neurons. (**C**) Another 6-OHDA-treated animal. The neuronal processes are almost completely missing. (**D**) Quantification of neuronal defects in 3-d- and 6-d-old control (untreated) and 6-OHDA-treated animals (sample size: 81 6-OHDA and 88 controls for day 3 set; 42 6-OHDA and 33 controls for day 6 set).

receptor family members in *C. elegans*, *dop-1* (D1) and *dop-3* (D2), are co-expressed on the cholinergic motor neurons of the ventral and dorsal cord and control locomotion.[55] Therefore, it is conceivable that reduced dopamine in toxin-treated worms might affect DOP receptor function resulting in abnormal biphasic activity of muscles. Analysis of mutants affecting these and two other DA receptors (*dop-4*, D1 class; *dop-2*, D2 class) in the microfluidics setup should reveal their function in mediating electrotaxis.

Electrotaxis-based behavioral assay as a tool to screen for neuroprotective compounds. Traditional chemical screening using *C. elegans* involves feeding worms with compounds of interest present in food (e.g., in a 96-well microtiter plate). The subsequent analysis of animals involves monitoring their growth,

viability and specific cellular and molecular markers. The existing methods of phenotypic analyses are slow, tedious and prone to human errors. The microfluidic system described here allows for objective and quantitative analysis of worm behavior following drug exposure. With some modifications, including development of automated video capturing and data analysis tools, the system can be used to screen for chemicals affecting neuronal function in a high-throughput manner. As an example, we tested the effect of acetaminophen, a compound with known neuroprotective response in *C. elegans*.[48] The results showed that acetaminophen protects worms from the toxic effect of 6-OHDA as judged by the normal electrotaxis response in the channel assay. Therefore, our microfluidic electrotaxis system holds promise in chemical screening and identification of potential therapeutic compounds with neuroprotective properties.

Materials and Methods

***C. elegans* cultures.** Worms were grown at 20°C on standard NG-agar plates containing *E. coli* OP50 culture as previously described.[56] The strains used in this study are: N2 (wild-type), DY328 *dat-1_p::YFP (bhEx120)*, CB1489 *him-8(e1489)*, PS2821 *lin-11(n389)*, PR813 *osm-5(p813)*, RM2702 *dat-1(ok157)* and CB1112 *cat-2(e1112)*.

Synchronized worms were used for all the assays and were prepared by bleach treatment.[20] Briefly, gravid hermaphrodites were treated with a solution containing commercial bleach and 4 N NaOH (3:2 ratio). The dead worms were washed with M9 buffer and incubated at room temperature for 24 h to allow fertilized embryos to hatch into L1 larvae.

Except for *lin-11(n389)* animals, all other plate-based and electrotaxis assays were done with 69 h young adults. This stage was chosen based on our finding that almost all synchronized wild-type L1s, when placed on NG-agar plates, reach to adulthood by 69 h at 20°C (97% adult and remaining younger stages, n = 1,002). The *lin-11* mutants were tested at 40 h stage because of their egg laying-defective (Egl) phenotype.[57]

Molecular biology and transgenics. The *dat-1_p::YFP* plasmid pGLC72 was made by amplifying a 710 bp fragment of *dat-1* 5' genomic region using primers GL563 (5'-AGGAAGCTTCCA GTTTTCACTAAAACGACTCATACACTTCTC-3') and GL564 (5'-ATGGGTACCGGCACCAACTGCATGGCT AAAAATTGTTGAG-3'). The resulting PCR product was digested with *Hind*III and *Kpn*I and subcloned into pPD136.64 (Fire lab vector, www.addgene.com). pGLC72 was injected into *unc-119(ed4)* animals to generate stable transgenic lines.

Toxin treatments and optimizations. All toxin treatments were done with synchronized L1 populations. Worms were exposed to toxins for different time periods with mild shaking on a rocking platform. Following exposure, tubes were briefly centrifuged and worm pellets were washed once with M9 buffer. Worms were transferred to NG-agar culture plates. Desired concentrations of 6-OHDA (100 μM) (Sigma Aldrich, 162957), MPTP (700 μM) (Toronto Research Chemicals, M325913) and rotenone (25 μM) (Sigma Aldrich, R8875) were prepared in M9 1 d before the assay and stored at –20°C. The

6-OHDA solution is sensitive to light therefore it was kept in the dark.

We modified toxin exposure protocols by lowering the concentration and exposure time. This was necessary because plate-based assays in the past used high doses of chemicals resulting in pleotropic defects such as delayed growth and lethality.[26,27,58] This precluded us from carrying out electrotaxis experiments. Additionally, we were concerned about non-specific effects due to prolonged exposures to high doses of chemicals that could affect neurons other than those involved in DA signaling. For example, L1 animals treated with 5 mM 6-OHDA for 30 min (one of the lowest concentrations reported) showed extreme sluggishness, uncoordinated (Unc) movement, protruding vulva, growth arrest and early larval lethality[58] (data not shown). We were unable to examine such worms in the channel because they were practically immobile and unresponsive to the electric field stimulus. Lowering the 6-OHDA concentration by 50-fold (100 μM) improved the overall health of worms, thereby allowing us to carry out electrotaxis assays.

We also optimized the MPTP and rotenone treatment protocols. It was reported earlier that worms grown in the presence of 1.4 mM MPTP for 3 d (starting L1 larval stage) were uncoordinated and extremely slow growing.[27] In the case of rotenone, a dose of 25 μM for 4 d caused lethality.[28] We found that reducing the rotenone exposure (25 μM rotenone for 12 h) allowed animals to survive but affected their growth and movement. However, none of the above toxin conditions could be used in the channel assay since worms were too sick to move and did not respond to the electric field stimulus. So, we further modified the toxin treatment conditions and found that animals exposed to 700 μM MPTP or 25 μM rotenone up to 8 h were generally healthy on plates and could be used to perform electrotaxis experiments.

Neuroprotection assay. The 10 mM stock solution of acetaminophen (Sigma Aldrich, A7085) was prepared fresh and diluted to a final concentration of 100 μM at the time of the assay. Synchronized L1 stage animals were incubated for 24 h in 100 μM drug containing M9 buffer. They were washed once with M9 and placed in toxin-containing solution for 1 h (100 μM 6-OHDA) or 8 h (700 μM MPTP and 25 μM rotenone). After an additional wash, the animals were transferred to NG-agar plates. The electrotaxis assays were performed on 69 h adults.

Microscopy. Worms were mounted on glass slides containing agar pads and observed using Zeiss AxioImager D1 and Nikon Eclipse Nomarski fluorescence microscopes. Epifluorescence was visualized using a GFP filter (HQ485LP, Chroma Technology USA). Degeneration of DA neurons was examined in 6-OHDA-treated *dat-1$_p$::YFP* transgenic worms. For this, L1 worms were exposed to 6-OHDA for 4 h, washed once with M9 and then plated on NG agar plates. Adults were examined for YFP fluorescence in dopaminergic neurons and their processes.

Plate-based assays. Well-fed synchronized worms (described above) were used to quantify their movement responses on agar plates. The stages were 69 h (post-L1) young adult for *osm-5,*
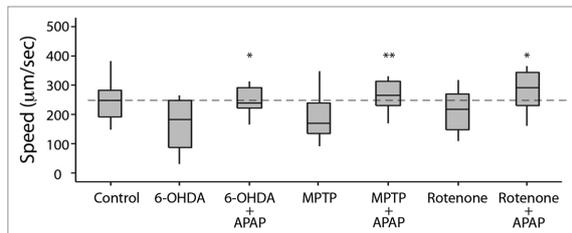

Figure 8. Electrotactic responses of worms treated with toxins and acetaminophen (para-acetylaminophenol, APAP). Control refers to untreated wild-type N2 animals. Refer to **Figure 2** for an explanation of the plot structure and other details. The speed of APAP-treated group is comparable to the control and higher than the corresponding toxin-treated group.

dat-1 and *him-8* and 40 h (post-L1) larvae for *lin-11*. The corresponding stages of N2 worms were used as controls. Worms were placed, one at a time, on a fresh 1-d-old thin bacterial lawn. Following acclimatization for 5 min. the animal's movement was observed for 30 sec. The beginning and end positions of tracks were marked and the image was captured using a digital camera (Point Grey, FL3-GE-13S2C) attached to Leica S8APO microscope. The movement speed was calculated by NIH ImageJ software (rsbweb.nih.gov/ij/).

Chemotaxis assay was performed as described earlier.[59,60] NaCl 100 mM solution was used as a chemo-attractant. 6-OHDA-treated animals were washed thoroughly and placed at the center of the agar plate. The plate contained a drop of NaCl at one end and water control at the other end. The assay was run for 1 h. After this, animals toward water and NaCl spots were counted and chemotaxis index (CI) was calculated.

Electrotaxis assays and data analysis. The experimental setup of microfluidic electrotaxis and fabrication of channels have been described earlier[20] (**Fig. 1**). Briefly, soft lithography technique was used to fabricate the microchannel (5 cm long and 300 μm wide) in polydimethylsiloxane (PDMS) (**Fig. 1A**). The DC electric field strength was set at 3 V/cm. Worms were loaded into the channel using a syringe pump and brought in the field of view. The electric field was activated and each animal was allowed to travel a minimum distance of 5,000 μm in one direction (toward cathode) (**Fig. 1B and C**). The electric field was reversed when the worm reached the end of the channel thereby causing it to turn and swim in the opposite direction (**Fig. 1D and E**). It should be mentioned here that the movement of the worm inside the microfluidic channel during electrotaxis assay is not affected by the electrokinetic flow.[20]

Electrotaxis responses of worms were recorded using a Nikon Coolpix digital camera (P5100) attached to a Leica stereomicroscope. Each data set typically consisted of 10–25 animals. For each animal, the response was monitored for 2–9 min duration and up to 45,000 μm total distance (in both directions). NIH ImageJ was used to analyze the captured videos.

Four swimming parameters, electrotaxis speed, turn time, body bend frequency and electrotaxis time index (ETI), were quantified manually. The electrotaxis speed is an average response of animals

115

that was obtained by dividing the total swimming distance by elapsed time. In this analysis, only movement toward cathode was used to calculate the distance. Any motion toward anode was ignored. The elapsed time is the duration of the assay. The turn time is the period to complete a U-turn following a switch in the electric field polarity. Up to three turning events were used to determine the average response of each animal. The body bend frequency is the average number of sine waves per second. This was obtained by dividing the total number of sine waves produced by an animal in an experiment by the duration of the assay (in seconds). Only those sine waves were counted that spanned at least half of the channel diameter. Finally, ETI is the measure of electrotaxis response. The ETI of an animal was calculated by dividing the sum of all cathode-directed swimming time by total time of assay. For a given genotype, a lower ETI indicates less cathode-directed motion of animals whereas higher ETI indicates greater amount of time spent in moving toward cathode.

The statistical significance of data was evaluated using unpaired Student's t-test and non-parametric Mann Whitney test. The p values less than 0.05 were considered statistically significant.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Supplemental Materials

Supplemental material may be found here:
www.landesbioscience.com/journals/worm/article/24558/

References

1. White JG, Southgate E, Thomson JN, Brenner S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. Philos Trans R Soc Lond B Biol Sci 1986; 314:1-340; PMID:22462104; http://dx.doi.org/10.1098/rstb.1986.0056.

2. Bargmann CI. Neurobiology of the *Caenorhabditis elegans* genome. Science 1998; 282:2028-33; PMID:9851919; http://dx.doi.org/10.1126/science.282.5396.2028.

3. Pedersen SF, Owsianik G, Nilius B. TRP channels: an overview. Cell Calcium 2005; 38:233-52; PMID:16098585; http://dx.doi.org/10.1016/j.ceca.2005.06.028.

4. Venkatachalam K, Montell C. TRP channels. Annu Rev Biochem 2007; 76:387-417; PMID:17579562; http://dx.doi.org/10.1146/annurev.biochem.75.103004.142819.

5. Bargmann CI. Chemosensation in *C. elegans*. WormBook 2006; ed. The *C. elegans* Research Community, WormBook; http://dx.doi.org/10.1895/wormbook.1.123.1, http://www.wormbook.org.

6. Sulston J, Dew M, Brenner S. Dopaminergic neurons in the nematode *Caenorhabditis elegans*. J Comp Neurol 1975; 163:215-26; PMID:240872; http://dx.doi.org/10.1002/cne.901630207.

7. Sawin ER, Ranganathan R, Horvitz HR. *C. elegans* locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. Neuron 2000; 26:619-31; PMID:10896158; http://dx.doi.org/10.1016/S0896-6273(00)81199-X.

8. Hills T, Brockie PJ, Maricq AV. Dopamine and glutamate control area-restricted search behavior in *Caenorhabditis elegans*. J Neurosci 2004; 24:1217-25; PMID:14762140; http://dx.doi.org/10.1523/JNEUROSCI.1569-03.2004.

9. Schafer WR, Kenyon CJ. A calcium-channel homologue required for adaptation to dopamine and serotonin in *Caenorhabditis elegans*. Nature 1995; 375:73-8; PMID:7723846; http://dx.doi.org/10.1038/375073a0.

10. Weinshenker D, Garriga G, Thomas JH. Genetic and pharmacological analysis of neurotransmitters controlling egg laying in *C. elegans*. J Neurosci 1995; 15:6975-85; PMID:7472454.

11. McDonald PW, Jessen T, Field JR, Blakely RD. Dopamine signaling architecture in *Caenorhabditis elegans*. Cell Mol Neurobiol 2006; 26:593-618; PMID:16724276; http://dx.doi.org/10.1007/s10571-006-9003-6.

12. Hulme SE, Shevkoplyas SS, Samuel A. Microfluidics: streamlining discovery in worm biology. Nat Methods 2008; 5:589-90; PMID:18587316; http://dx.doi.org/10.1038/nmeth0708-589.

13. Crane MM, Chung K, Stirman J, Lu H. Microfluidics-enabled phenotyping, imaging, and screening of multicellular organisms. Lab Chip 2010; 10:1509-17; PMID:20383347; http://dx.doi.org/10.1039/b927258e.

14. Rezai P, Salam S, Selvaganapathy PR, Gupta BP. Microfluidic systems to study the biology of human diseases and identify potential therapeutic targets in *Caenorhabditis elegans*. In: Iniewski K, ed. Integrated microsystems. Florida: CRC Press, 2012:581-608.

15. Ben-Yakar A, Chronis N, Lu H. Microfluidics for the analysis of behavior, nerve regeneration, and neural cell biology in *C. elegans*. Curr Opin Neurobiol 2009; 19:561-7; PMID:19896831; http://dx.doi.org/10.1016/j.conb.2009.10.010.

16. Chung K, Crane MM, Lu H. Automated on-chip rapid microscopy, phenotyping and sorting of *C. elegans*. Nat Methods 2008; 5:637-43; PMID:18568029; http://dx.doi.org/10.1038/nmeth.1227.

17. Chronis N, Zimmer M, Bargmann CI. Microfluidics for in vivo imaging of neuronal and behavioral activity in *Caenorhabditis elegans*. Nat Methods 2007; 4:727-31; PMID:17704783; http://dx.doi.org/10.1038/nmeth1075.

18. Rohde CB, Zeng F, Gonzalez-Rubio R, Angel M, Yanik MF. Microfluidic system for on-chip high-throughput whole-animal sorting and screening at subcellular resolution. Proc Natl Acad Sci USA 2007; 104:13891-5; PMID:17715055; http://dx.doi.org/10.1073/pnas.0706513104.

19. Park S, Hwang H, Nam SW, Martinez F, Austin RH, Ryu WS. Enhanced *Caenorhabditis elegans* locomotion in a structured microfluidic environment. PLoS One 2008; 3:e2550; PMID:18575618; http://dx.doi.org/10.1371/journal.pone.0002550.

20. Rezai P, Siddiqui A, Selvaganapathy PR, Gupta BP. Electrotaxis of *Caenorhabditis elegans* in a microfluidic environment. Lab Chip 2010; 10:220-6; PMID:20066250; http://dx.doi.org/10.1039/b917486a.

21. Rezai P, Salam S, Selvaganapathy PR, Gupta BP. Effect of pulse direct current signals on electrotactic movement of nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. Biomicrofluidics 2011; 5:44116-441169; PMID:22232698; http://dx.doi.org/10.1063/1.3665224.

22. Sukul NC, Croll NA. Influence of potential difference and current on the electrotaxis of *Caenorhabditis elegans*. J Nematol 1978; 10:314-7; PMID:19305860.

23. Bird AF. The attractiveness of roots to the plant parasitic nematode *Meloidogyne javanica* and *M. hapla*. Nematologica 1959; 4:322-35; http://dx.doi.org/10.1163/187529259X00534.

24. Gabel CV, Gabel H, Pavlichin D, Kao A, Clark DA, Samuel AD. Neural circuits mediate electrosensory behavior in *Caenorhabditis elegans*. J Neurosci 2007; 27:7586-96; PMID:17626220; http://dx.doi.org/10.1523/JNEUROSCI.0775-07.2007.

25. Shapiro-Ilan DI, Campbell JF, Lewis EE, Elkon JM, Kim-Shapiro DB. Directional movement of steinernematid nematodes in response to electrical current. J Invertebr Pathol 2009; 100:134-7; PMID:19041325; http://dx.doi.org/10.1016/j.jip.2008.11.001.

26. Nass R, Hall DH, Miller DM 3rd, Blakely RD. Neurotoxin-induced degeneration of dopamine neurons in *Caenorhabditis elegans*. Proc Natl Acad Sci USA 2002; 99:3264-9; PMID:11867711; http://dx.doi.org/10.1073/pnas.042497999.

27. Braungart E, Gerlach M, Riederer P, Baumeister R, Hoener MC. *Caenorhabditis elegans* MPP+ model of Parkinson's disease for high-throughput drug screenings. Neurodegener Dis 2004; 1:175-83; PMID:16908987; http://dx.doi.org/10.1159/000080983.

28. Ved R, Saha S, Westlund B, Perier C, Burnam LG, Sluder A, et al. Similar patterns of mitochondrial vulnerability and rescue induced by genetic modification of alpha-synuclein, parkin, and DJ-1 in *Caenorhabditis elegans*. J Biol Chem 2005; 280:42655-68; PMID:16239214; http://dx.doi.org/10.1074/jbc.M505910200.

29. Glinka Y, Gassen M, Youdim MB. Mechanism of 6-hydroxydopamine neurotoxicity. J Neural Transm Suppl 1997; 50:55-66; PMID:9120425; http://dx.doi.org/10.1007/978-3-7091-6842-4_7.

30. Haycraft CJ, Swoboda P, Taulman PD, Thomas JH, Yoder BK. The *C. elegans* homolog of the murine cystic kidney disease gene Tg737 functions in a ciliogenic pathway and is disrupted in osm-5 mutant worms. Development 2001; 128:1493-505; PMID:11290289.

31. Hodgkin J, Horvitz HR, Brenner S. Nondisjunction mutants of the nematode *Caenorhabditis elegans*. Genetics 1979; 91:67-94; PMID:17248881.

32. Sarafi-Reinach TR, Melkman T, Hobert O, Sengupta P. The lin-11 LIM homeobox gene specifies olfactory and chemosensory neuron fates in *C. elegans*. Development 2001; 128:3269-81; PMID:11546744.

33. Hobert O, D'Alberti T, Liu Y, Ruvkun G. Control of neural development and function in a thermoregulatory network by the LIM homeobox gene lin-11. J Neurosci 1998; 18:2084-96; PMID:9482795.

34. Ferguson EL, Horvitz HR. Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. Genetics 1985; 110:17-72; PMID:3996896.

35. Freyd G, Kim SK, Horvitz HR. Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene lin-11. Nature 1990; 344:876-9; PMID:1970421; http://dx.doi.org/10.1038/344876a0.

36. Lints R, Emmons SW. Patterning of dopaminergic neurotransmitter identity among *Caenorhabditis elegans* ray sensory neurons by a TGFbeta family signaling pathway and a Hox gene. Development 1999; 126:5819-31; PMID:10572056.

37. Jayanthi LD, Apparsundaram S, Malone MD, Ward E, Miller DM, Eppler M, et al. The *Caenorhabditis elegans* gene T23G5.5 encodes an antidepressant- and cocaine-sensitive dopamine transporter. Mol Pharmacol 1998; 54:601-9; PMID:9765501.

38. Sanyal S, Wintle RF, Kindt KS, Nuttley WM, Arvan R, Fitzmaurice P, et al. Dopamine modulates the plasticity of mechanosensory responses in *Caenorhabditis elegans*. EMBO J 2004; 23:473-82; PMID:14739932; http://dx.doi.org/10.1038/sj.emboj.7600057.

39. Betarbet R, Sherer TB, MacKenzie G, Garcia-Osuna M, Panov AV, Greenamyre JT. Chronic systemic pesticide exposure reproduces features of Parkinson's disease. Nat Neurosci 2000; 3:1301-6; PMID:11100151; http://dx.doi.org/10.1038/81834.

40. Bové J, Prou D, Perier C, Przedborski S. Toxin-induced models of Parkinson's disease. NeuroRx 2005; 2:484-94; PMID:16389312; http://dx.doi.org/10.1602/neurorx.2.3.484.

41. Jonsson G, Sachs C. Actions of 6-hydroxydopamine quinones on catecholamine neurons. J Neurochem 1975; 25:509-16; PMID:1173817; http://dx.doi.org/10.1111/j.1471-4159.1975.tb04357.x.

42. Glinka YY, Youdim MB. Inhibition of mitochondrial complexes I and IV by 6-hydroxydopamine. Eur J Pharmacol 1995; 292:329-32; PMID:7796873.

43. Chiba K, Trevor A, Castagnoli N Jr. Metabolism of the neurotoxic tertiary amine, MPTP, by brain monoamine oxidase. Biochem Biophys Res Commun 1984; 120:574-8; PMID:6428396; http://dx.doi.org/10.1016/0006-291X(84)91293-2.

44. Ali SF, David SN, Newport GD, Cadet JL, Slikker W Jr. MPTP-induced oxidative stress and neurotoxicity are age-dependent: evidence from measures of reactive oxygen species and striatal dopamine levels. Synapse 1994; 18:27-34; PMID:7825121; http://dx.doi.org/10.1002/syn.890180105.

45. Coulom H, Birman S. Chronic exposure to rotenone models sporadic Parkinson's disease in *Drosophila melanogaster*. J Neurosci 2004; 24:10993-8; PMID:15574749; http://dx.doi.org/10.1523/JNEUROSCI.2993-04.2004.

46. Maharaj DS, Saravanan KS, Maharaj H, Mohanakumar KP, Daya S. Acetaminophen and aspirin inhibit superoxide anion generation and lipid peroxidation, and protect against 1-methyl-4-phenyl pyridinium-induced dopaminergic neurotoxicity in rats. Neurochem Int 2004; 44:355-60; PMID:14643753; http://dx.doi.org/10.1016/S0197-0186(03)00170-0.

47. Casper D, Yaparpalvi U, Rempel N, Werner P. Ibuprofen protects dopaminergic neurons against glutamate toxicity in vitro. Neurosci Lett 2000; 289:201-4; PMID:10961664; http://dx.doi.org/10.1016/S0304-3940(00)01294-5.

48. Locke CJ, Fox SA, Caldwell GA, Caldwell KA. Acetaminophen attenuates dopamine neuron degeneration in animal models of Parkinson's disease. Neurosci Lett 2008; 439:129-33; PMID:18514411; http://dx.doi.org/10.1016/j.neulet.2008.05.003.

49. Cenci MA, Whishaw IQ, Schallert T. Animal models of neurological deficits: how relevant is the rat? Nat Rev Neurosci 2002; 3:574-9; PMID:12094213; http://dx.doi.org/10.1038/nrn877.

50. Maschke M, Gomez CM, Tuite PJ, Konczak J. Dysfunction of the basal ganglia, but not the cerebellum, impairs kinaesthesia. Brain 2003; 126:2312-22; PMID:12821507; http://dx.doi.org/10.1093/brain/awg230.

51. Wilson RI, Corey DP. The force be with you: a mechanoreceptor channel in proprioception and touch. Neuron 2010; 67:349-51; PMID:20696370; http://dx.doi.org/10.1016/j.neuron.2010.07.022.

52. Kang L, Gao J, Schafer WR, Xie Z, Xu XZ. *C. elegans* TRP family protein TRP-4 is a pore-forming subunit of a native mechanotransduction channel. Neuron 2010; 67:381-91; PMID:20696377; http://dx.doi.org/10.1016/j.neuron.2010.06.032.

53. Driscoll M, Kaplan J. Mechanotransduction. In: Riddle DL, Blumenthal T, Meyer BJ, Priess JR, eds. *C. elegans* II. Cold Spring Harbor (NY), 1997.

54. Korchounov A, Meyer MF, Krasnianski M. Postsynaptic nigrostriatal dopamine receptors and their role in movement regulation. J Neural Transm 2010; 117:1359-69; PMID:21076988; http://dx.doi.org/10.1007/s00702-010-0454-z.

55. Allen AT, Maher KN, Wani KA, Betts KE, Chase DL. Coexpressed D1- and D2-like dopamine receptors antagonistically modulate acetylcholine release in *Caenorhabditis elegans*. Genetics 2011; 188:579-90; PMID:21515580; http://dx.doi.org/10.1534/genetics.111.128512.

56. Brenner S. The genetics of *Caenorhabditis elegans*. Genetics 1974; 77:71-94; PMID:4366476.

57. Trent C, Tsuing N, Horvitz HR. Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. Genetics 1983; 104:619-47; PMID:11813735.

58. Marvanova M, Nichols CD. Identification of neuroprotective compounds of *Caenorhabditis elegans* dopaminergic neurons against 6-OHDA. J Mol Neurosci 2007; 31:127-37; PMID:17478886.

59. Rezai P, Salam S, Selvaganapathy PR, Gupta BP. Electrical sorting of *Caenorhabditis elegans*. Lab Chip 2012; 12:1831-40; PMID:22460920; http://dx.doi.org/10.1039/c2lc20967e.

60. Hart AC. Behavior. Wormbook 2006; ed. The *C. elegans* Research Community, WormBook; http://dx.doi.org/10.1895/wormbook.1.87.1, http://www.wormbook.org.

Bibliography

Aamodt, E., Shen, L., Marra, M., Schein, J., Rose, B., and McDermott, J. B. (2000). Conservation of sequence and function of the pag-3 genes from C. elegans and C. briggsae. *Gene* 243(1-2), 67–74.

Abate-Shen, C. (2002). Deregulated homeobox gene expression in cancer: cause or consequence? *Nature Reviews Cancer* 2(10), 777–785.

Ahlgren, U., Pfaff, S. L., Jessell, T. M., Edlund, T., and Edlund, H. (1997). Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature* 385(6613), 257.

Akam, M. (1987). The molecular basis for metameric pattern in the Drosophila embryo. en. *Development* 101(1), 1–22.

Akbari, O. S., Bousum, A., Bae, E., and Drewell, R. A. (2006). Unraveling cis-regulatory mechanisms at the abdominal-A and Abdominal-B genes in the Drosophila bithorax complex. en. *Dev. Biol.* 293(2), 294–304.

Alifragis, P., Liapi, A., and Parnavelas, J. G. (2004). Lhx6 regulates the migration of cortical interneurons from the ventral telencephalon but does not specify their GABA phenotype. *Journal of Neuroscience* 24(24), 5643–5648.

Altun-Gultekin, Z., Andachi, Y., Tsalik, E. L., Pilgrim, D., Kohara, Y., and Hobert, O. (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. en. *Development* 128(11), 1951–1969.

Amon, S. and Gupta, B. P. (2017a). Intron-specific patterns of divergence of lin-11 regulatory function in the C. elegans nervous system. *Developmental Biology* 424(1), 90–103.

Amon, S. and Gupta, B. P. (2017b). Multi-species alignments of C. elegans lin-11 intronic sequences and putative transcriptional regulators. *Data in Brief* 12, 87–90.

Aoki, I. and Mori, I. (2015). Molecular biology of thermosensory transduction in C. elegans. *Curr. Opin. Neurobiol.* 34, 117–124.

Aplin, A. C. and Kaufman, T. C. (1997). Homeotic transformation of legs to mouthparts by proboscipedia expression in Drosophila imaginal discs. *Mech. Dev.* 62(1), 51–60.

Appel, B., Korzh, V., Glasgow, E., Thor, S., Edlund, T., Dawid, I. B., and Eisen, J. S. (1995). Motoneuron fate specification revealed by patterned LIM homeobox gene expression in embryonic zebrafish. *Development* 121(12), 4117–4125.

Arber, S. and Caroni, P. (1996). Specificity of single LIM motifs in targeting and LIM/LIM interactions in situ. en. *Genes Dev.* 10(3), 289–300.

Aurelio, O., Boulin, T., and Hobert, O. (2003). Identification of spatial and temporal cues that regulate postembryonic expression of axon maintenance factors in the C. elegans ventral nerve cord. *Development* 130(3), 599–610.

Avraham, O., Hadas, Y., Vald, L., Zisman, S., Schejter, A., Visel, A., and Klar, A. (2009). Transcriptional control of axonal guidance and sorting in dorsal interneurons by the Lim-HD proteins Lhx9 and Lhx1. en. *Neural Dev.* 4, 21.

Bach, I. (2000). The LIM domain: regulation by association. en. *Mech. Dev.* 91(1-2), 5–17.

Ball, L. J., Jarchau, T., Oschkinat, H., and Walter, U. (2002). EVH1 domains: structure, function and interactions. en. *FEBS Lett.* 513(1), 45–52.

Bargmann, C. and Cornelia, B. (2006). Chemosensation in C. elegans. *WormBook.*

Baumeister, R., Liu, Y., and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the Caenorhabditis elegans POU gene unc-86 during neurogenesis. en. *Genes Dev.* 10(11), 1395–1410.

Beachy, P. A. (1990). A molecular view of the Ultrabithorax homeotic gene of Drosophila. *Trends Genet.* 6, 46–51.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., Lewis, E. B., and Hogness, D. S. (2004). Molecular Genetics of the Bithorax Complex in Drosophila Melanogaster. In: *Genes, Development and Cancer*, 287–302.

Benveniste, R., Thor, S., Thomas, J. B., and Taghert, P. H. (1998). Cell type-specific regulation of the Drosophila FMRF-NH2 neuropeptide gene by Apterous, a LIM homeodomain transcription factor. *Development* 125(23), 4757–4765.

Bertuzzi, S., Porter, F. D., Pitts, A., Kumar, M., Agulnick, A., Wassif, C., and Westphal, H. (1999). Characterization of Lhx9, a novel LIM/homeobox gene expressed by the pioneer neurons in the mouse cerebral cortex. *Mechanisms of development* 81(1), 193–198.

Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L., et al. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature* 403(6772), 909–913.

Bonanomi, D. and Pfaff, S. L. (2010). Motor axon pathfinding. *Cold Spring Harbor perspectives in biology* 2(3), a001735.

Bridges, C. B., Morgan, T. H., and Washington, C. I. o. (1923). *The third-chromosome group of mutant characters of Drosophila melanogaster.*

Bulchand, S., Grove, E., Porter, F., and Tole, S. (2001). LIM-homeodomain gene Lhx2 regulates the formation of the cortical hem. *Mechanisms of development* 100(2), 165–175.

Bürglin, T. R. and Affolter, M. (2016). Homeodomain proteins: an update. *Chromosoma* 125(3), 497–521.

Carroll, S. B., Weatherbee, S. D., and Langeland, J. A. (1995). Homeotic genes and the regulation and evolution of insect wing number. en. *Nature* 375(6526), 58–61.

Cassata, G., Kagoshima, H., Andachi, Y., Kohara, Y., Dürrenberger, M. B., Hall, D. H., and Bürglin, T. R. (2000). The LIM homeobox gene ceh-14 confers thermosensory function to the AFD neurons in Caenorhabditis elegans. en. *Neuron* 25(3), 587–597.

Cheng, L., Chen, C.-L., Luo, P., Tan, M., Qiu, M., Johnson, R., and Ma, Q. (2003). Lmx1b, Pet-1, and Nkx2. 2 coordinately specify serotonergic neurotransmitter phenotype. *Journal of Neuroscience* 23(31), 9961–9967.

Cohen, B., McGuffin, M. E., Pfeifle, C., Segal, D., and Cohen, S. M. (1992). apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins. en. *Genes Dev.* 6(5), 715–729.

Costa, C., Harding, B., and Copp, A. J. (2001). Neuronal migration defects in the Dreher (Lmx1a) mutant mouse: role of disorders of the glial limiting membrane. *Cerebral Cortex* 11(6), 498–505.

Curtiss, J. and Heilig, J. S. (1995). Establishment of Drosophila imaginal precursor cells is controlled by the Arrowhead gene. *Development* 121(11), 3819–3828.

Curtiss, J. and Heilig, J. S. (1997). Arrowhead encodes a LIM homeodomain protein that distinguishes subsets of Drosophila imaginal cells. *Developmental biology* 190(1), 129–141.

Cutter, A. D. (2008). Divergence times in Caenorhabditis and Drosophila inferred from direct estimates of the neutral mutation rate. en. *Mol. Biol. Evol.* 25(4), 778–786.

Das, S., Sadanandappa, M. K., Dervan, A., Larkin, A., Lee, J. A., Sudhakaran, I. P., Priya, R., Heidari, R., Holohan, E. E., Pimentel, A., et al. (2011). Plasticity of local GABAergic interneurons drives olfactory habituation. *Proceedings of the National Academy of Sciences* 108(36), E646–E654.

Dawid, I. B., Breen, J. J., and Toyama, R. (1998). LIM domains: multiple roles as adapters and functional modifiers in protein interactions. en. *Trends Genet.* 14(4), 156–162.

Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., and Matthews, J. M. (2004). Tandem LIM domains provide synergistic binding in the LMO4:Ldb1 complex. en. *EMBO J.* 23(18), 3589–3598.

Diaz-Benjumea, F. J. and Cohen, S. M. (1993). Interaction between dorsal and ventral cells in the imaginal disc directs wing development in Drosophila. en. *Cell* 75(4), 741–752.

Ding, Y.-Q., Marklund, U., Yuan, W., Yin, J., Wegman, L., Ericson, J., Deneris, E., Johnson, R. L., and Chen, Z.-F. (2003). Lmx1b is essential for the development of serotonergic neurons. *Nature neuroscience* 6(9), 933–938.

Dreyer, S. D., Zhou, G., Baldini, A., Winterpacht, A., Zabel, B., Cole, W., Johnson, R. L., and Lee, B. (1998). Mutations in LMX1B cause abnormal skeletal patterning and renal dysplasia in nail patella syndrome. *Nature genetics* 19(1), 47–50.

Duboule, D. and Dollé, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. en. *EMBO J.* 8(5), 1497–1505.

Duggan, A., Ma, C., and Chalfie, M. (1998). Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes. en. *Development* 125(20), 4107–4119.

Durbin, R. M. (1987). *Studies on the development and organisation of the nervous system of Caenorhabditis elegans.*

Egan, S. E., Giddings, B. W., Brooks, M. W., Buday, L., Sizeland, A. M., and Weinberg, R. A. (1993). Association of Sos Ras exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation. *Nature* 363(6424), 45–51.

Emmons, S. W. (2005). Male development. en. *WormBook*, 1–22.

Ericson, J., Thor, S., Edlund, T., Jessell, T. M., and Yamada, T. (1992). Early stages of motor neuron differentiation revealed by expression of homeobox gene Islet-1. *Science* 256(5063), 1555–1560.

Félix, M.-A. (2007). Cryptic quantitative evolution of the vulva intercellular signaling network in Caenorhabditis. en. *Curr. Biol.* 17(2), 103–114.

Ferguson, E. L. and Horvitz, H. R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode Caenorhabditis elegans. en. *Genetics* 110(1), 17–72.

Finney, M. and Ruvkun, G. (1990). The unc-86 gene product couples cell lineage and cell identity in C. elegans. *Cell* 63(5), 895–905.

Forrester, W. C., Perens, E., Zallen, J. A., and Garriga, G. (1998). Identification of Caenorhabditis elegans genes required for neuronal differentiation and migration. *Genetics* 148(1), 151–165.

Freyd, G., Gwen, F., Kim, S. K., and Robert Horvitz, H. (1990). Novel cysteine-rich motif and homeodomain in the product of the Caenorhabditis elegans cell lineage gene lin-II. *Nature* 344(6269), 876–879.

Freyd, G. A. (1991). Molecular analysis of the Caenorhabditis elegans cell lineage gene lin-11. PhD thesis. Massachusetts Institute of Technology.

Gale, N. W., Kaplan, S., Lowenstein, E. J., Schlessinger, J., and Bar-Sagi, D. (1993). Grb2 mediates the EGF-dependent activation of guanine nucleotide exchange on Ras. *Nature* 363(6424), 88–92.

Ghazi, A., Anant, S., and Raghavan, K. V. (2000). Apterous mediates development of direct flight muscles autonomously and indirect flight muscles through epidermal cues. *Development* 127(24), 5309–5318.

Gong, Z. and Hew, C. L. (1994). Zinc and DNA binding properties of a novel LIM homeodomain protein Isl-2. en. *Biochemistry* 33(50), 15149–15158.

Goodman, M. B. (2006). Mechanosensation. *WormBook: the online review of C. elegans biology*, 1–14.

Grigoriou, M., Tucker, A. S., Sharpe, P. T., and Pachnis, V. (1998). Expression and regulation of Lhx6 and Lhx7, a novel subfamily of LIM homeodomain encoding genes, suggests a role in mammalian head development. *Development* 125(11), 2063–2074.

Gupta, B. P., Johnsen, R., and Chen, N. (2007). Genomics and biology of the nematode Caenorhabditis briggsae.

Gupta, B. P., Wang, M., and Sternberg, P. W. (2003). The C. elegans LIM homeobox gene lin-11 specifies multiple cell fates during vulval development. *Development* 130(12), 2589–2601.

Hall, D. H., Winfrey, V. P., Blaeuer, G., Hoffman, L. H., Furuta, T., Rose, K. L., Hobert, O., and Greenstein, D. (1999). Ultrastructural features of the adult hermaphrodite gonad of Caenorhabditis elegans: relations between the germ line and soma. en. *Dev. Biol.* 212(1), 101–123.

Hart, A. and Chao, M. (2009). From Odors to Behaviors in Caenorhabditis Elegans. In: *Frontiers in Neuroscience*, 1–33.

Hedgecock, E. M. and Russell, R. L. (1975). Normal and mutant thermotaxis in the nematode Caenorhabditis elegans. en. *Proc. Natl. Acad. Sci. U. S. A.* 72(10), 4061–4065.

Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.* 3.

Herzig, M. C., Thor, S., Thomas, J. B., Reichert, H., and Hirth, F. (2001). Expression and function of the LIM homeodomain protein Apterous during embryonic brain development of Drosophila. *Development genes and evolution* 211(11), 545–554.

Hilliard, M. A., Bergamasco, C., Arbucci, S., Plasterk, R. H., and Bazzicalupo, P. (2004). Worms taste bitter: ASH neurons, QUI-1, GPA-3 and ODR-3 mediate quinine avoidance in Caenorhabditis elegans. *The EMBO journal* 23(5), 1101–1111.

Hobert, O., D'Alberti, T., Liu, Y., and Ruvkun, G. (1998). Control of neural development and function in a thermoregulatory network by the LIM homeobox gene lin-11. en. *J. Neurosci.* 18(6), 2084–2096.

Hobert, O., Tessmar, K., and Ruvkun, G. (1999). The Caenorhabditis elegans lim-6 LIM homeobox gene regulates neurite outgrowth and function of particular GABAergic neurons. en. *Development* 126(7), 1547–1562.

Hobert, O. and Westphal, H. (2000). Functions of LIM-homeobox genes. en. *Trends Genet.* 16(2), 75–83.

Hobert, O. (2016). Chapter Twenty-Five-Terminal Selectors of Neuronal Identity. *Current topics in developmental biology* 116, 455–475.

Hobert, O., Oliver, H., Ikue, M., Yukiko, Y., Hidehiro, H., Yasumi, O., Yanxia, L., and Gary, R. (1997). Regulation of Interneuron Function in the C. elegans Thermoregulatory Pathway by the ttx-3 LIM Homeobox Gene. *Neuron* 19(2), 345–357.

Holt, M. (2001). Cell motility: proline-rich proteins promote protrusions. *Trends Cell Biol.* 11(1), 38–46.

Hutter, H. (2003). Extracellular cues and pioneers act together to guide axons in the ventral cord of C. elegans. en. *Development* 130(22), 5307–5318.

Hutter, H., Wacker, I., Schmid, C., and Hedgecock, E. M. (2005). Novel genes controlling ventral cord asymmetry and navigation of pioneer axons in C. elegans. *Dev. Biol.* 284(1), 260–272.

Jia, Y., Xie, G., and Aamodt, E. (1996). pag-3, a Caenorhabditis elegans gene involved in touch neuron gene expression and coordinated movement. en. *Genetics* 142(1), 141–147.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. en. *Comput. Appl. Biosci.* 8(3), 275–282.

Jorgensen, E. M. (2005). GABA. en. *WormBook*, 1–13.

Kadrmas, J. L. and Beckerle, M. C. (2004). The LIM domain: from the cytoskeleton to the nucleus. *Nat. Rev. Mol. Cell Biol.* 5(11), 920–931.

Kagoshima, H. and Kohara, Y. (2015). Co-expression of the transcription factors CEH-14 and TTX-1 regulates AFD neuron-specific genes gcy-8 and gcy-18 in C. elegans. en. *Dev. Biol.* 399(2), 325–336.

Kania, A., Johnson, R. L., and Jessell, T. M. (2000). Coordinate roles for LIM homeobox genes in directing the dorsoventral trajectory of motor axons in the vertebrate limb. en. *Cell* 102(2), 161–173.

Kaplan, J. M. and Horvitz, H. R. (1993). A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences* 90(6), 2227–2231.

Karlsson, O., Olof, K., Stefan, T., Torbjörn, N., Helena, O., and Thomas, E. (1990). Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo-and a Cys-His domain. *Nature* 344(6269), 879–882.

Kaufman, T. C., Lewis, R., and Wakimoto, B. (1980). Cytogenetic Analysis of Chromosome 3 in Drosophila Melanogaster: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. en. *Genetics* 94(1), 115–133.

Kay, B. K., Williamson, M. P., and Sudol, M. (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. en. *FASEB J.* 14(2), 231–241.

Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. *Nature* 427(6975), 645–649.

Khurana, T., Khurana, B., and Noegel, A. A. (2002). LIM proteins: association with the actin cytoskeleton. *Protoplasma* 219(1-2), 0001–0012.

Kim, J., Yeon, J., Choi, S.-K., Huh, Y. H., Fang, Z., Park, S. J., Kim, M. O., Ryoo, Z. Y., Kang, K., Kweon, H.-S., Jeon, W. B., Li, C., and Kim, K. (2015). The Evolutionarily

Conserved LIM Homeodomain Protein LIM-4/LHX6 Specifies the Terminal Identity of a Cholinergic and Peptidergic C. elegans Sensory/Inter/Motor Neuron-Type. en. *PLoS Genet.* 11(8), e1005480.

Kinder, S. J., Tsang, T. E., Wakamiya, M., Sasaki, H., Behringer, R. R., Nagy, A., and Tam, P. P. (2001). The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* 128(18), 3623–3634.

Kmita, M. and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. en. *Science* 301(5631), 331–333.

Kobayashi, T., Kamitani, W., Zhang, G., Watanabe, M., Tomonaga, K., and Ikuta, K. (2001). Borna disease virus nucleoprotein requires both nuclear localization and export activities for viral nucleocytoplasmic shuttling. en. *J. Virol.* 75(7), 3404–3412.

Koch, C. (1999). Complexity and the Nervous System. *Science* 284(5411), 96–98.

Kondo, T., Zákány, J., and Duboule, D. (1998). Control of Colinearity in AbdB Genes of the Mouse HoxD Complex. *Mol. Cell* 1(2), 289–300.

Kong, Y., Flick, M. J., Kudla, A. J., and Konieczny, S. F. (1997). Muscle LIM protein promotes myogenesis by enhancing the activity of MyoD. en. *Mol. Cell. Biol.* 17(8), 4750–4760.

Kosa, J. L., Michelsen, J. W., Louis, H. A., Olsen, J. I., Davis, D. R., Beckerle, M. C., and Winge, D. R. (1994). Common metal ion coordination in LIM domain proteins. *Biochemistry* 33(2), 468–477.

Kuhara, A., Ohnishi, N., Shimowada, T., and Mori, I. (2011). Neural coding in a single sensory neuron controlling opposite seeking behaviours in Caenorhabditis elegans. *Nat. Commun.* 2, 355.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33(7), 1870–1874.

Larroux, C., Fahey, B., Degnan, S. M., Adamski, M., Rokhsar, D. S., and Degnan, B. M. (2007). The NK homeobox gene cluster predates the origin of Hox genes. en. *Curr. Biol.* 17(8), 706–710.

Lehner, B., Calixto, A., Crombie, C., Tischler, J., Fortunato, A., Chalfie, M., and Fraser, A. G. (2006). Loss of LIN-35, the Caenorhabditis elegans ortholog of the tumor suppressor p105Rb, results in enhanced RNA interference. *Genome biology* 7(1), R4.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. In: *Genes, Development and Cancer.* Springer, 205–217.

Li, H., Witte, D., Branford, W., Aronow, B., Weinstein, M., Kaur, S., Wert, S., Singh, G., Schreiner, C., Whitsett, J., et al. (1994). Gsh-4 encodes a LIM-type homeodomain, is expressed in the developing central nervous system and is required for early postnatal survival. *The EMBO journal* 13(12), 2876.

Lilly, B., O'Keefe, D. D., Thomas, J. B., and Botas, J. (1999a). The LIM homeodomain protein dLim1 defines a subclass of neurons within the embryonic ventral nerve cord of Drosophila. en. *Mech. Dev.* 88(2), 195–205.

Lilly, B., D O'Keefe, D., Thomas, J. B., and Botas, J. (1999b). The LIM homeodomain protein dLim1 defines a subclass of neurons within the embryonic ventral nerve cord of Drosophila. *Mechanisms of development* 88(2), 195–205.

Liu, G., Xia, T., and Chen, X. (2003). The Activation Domains, the Proline-rich Domain, and the C-terminal Basic Domain in p53 Are Necessary for Acetylation of Histones on the Proximal p21 Promoter and Interaction with p300/CREB-binding Protein. *J. Biol. Chem.* 278(19), 17557–17565.

Lundgren, S. E., Callahan, C. A., Thor, S., and Thomas, J. B. (1995). Control of neuronal pathway selection by the Drosophila LIM homeodomain gene apterous. *Development* 121(6), 1769–1773.

Maduro, M. and Pilgrim, D. (1995). Identification and cloning of unc-119, a gene expressed in the Caenorhabditis elegans nervous system. en. *Genetics* 141(3), 977–988.

Mallo, M. and Alonso, C. R. (2013). The regulation of Hox gene expression during animal development. *Development* 140(19), 3951–3963.

Mallo, M., Wellik, D. M., and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. en. *Dev. Biol.* 344(1), 7–15.

Marri, S. and Gupta, B. P. (2009). Dissection of lin-11 enhancer regions in Caenorhabditis elegans and other nematodes. en. *Dev. Biol.* 325(2), 402–411.

Matsumoto, K., Tanaka, T., Furuyama, T., Kashihara, Y., Mori, T., Ishii, N., Kitanaka, J.-i., Takemura, M., Tohyama, M., and Wanaka, A. (1996). L3, a novel murine LIM-homeodomain transcription factor expressed in the ventral telencephalon and the mesenchyme surrounding the oral cavity. *Neuroscience letters* 204(1), 113–116.

McDonald, C. B., Seldeen, K. L., Deegan, B. J., and Farooq, A. (2009). SH3 Domains of Grb2 Adaptor Bind to PXψPXR Motifs Within the Sos1 Nucleotide Exchange Factor in a Discriminate Manner. *Biochemistry* 48(19), 4074–4085.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. en. *Cell* 68(2), 283–302.

McIntire, S. L., Jorgensen, E., Kaplan, J., and Horvitz, H. R. (1993). The GABAergic nervous system of Caenorhabditis elegans. *Nature* 364(6435), 337–341.

McPherson, P. (1999). Regulatory Role of SH3 Domain-mediated Protein–Protein Interactions in Synaptic Vesicle Endocytosis. *Cell. Signal.* 11(4), 229–238.

Melkman, T. (2005). Regulation of chemosensory and GABAergic motor neuron development by the C. elegans Aristaless/Arx homolog alr-1. *Development* 132(8), 1935–1949.

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in C. elegans: extrachromosomal maintenance and integration of transforming sequences. *The EMBO journal* 10(12), 3959.

Millonig, J. H., Millen, K. J., and Hatten, M. E. (2000). The mouse Dreher gene Lmx1a controls formation of the roof plate in the vertebrate CNS. *Nature* 403(6771), 764–769.

Miner, J. H., Morello, R., Andrews, K. L., Li, C., Antignac, C., Shaw, A. S., and Lee, B. (2002). Transcriptional induction of slit diaphragm genes by Lmx1b is required in podocyte differentiation. *The Journal of clinical investigation* 109(8), 1065–1072.

Monuki, E. S., Porter, F. D., and Walsh, C. A. (2001). Patterning of the dorsal telencephalon and cerebral cortex by a roof plate-Lhx2 pathway. *Neuron* 32(4), 591–604.

Mori, I., Ikue, M., and Yasumi, O. (1995). Neural regulation of thermotaxis in Caenorhabditis elegans. *Nature* 376(6538), 344–348.

Mori, T., Yuxing, Z., Takaki, H., Takeuchi, M., Iseki, K., Hagino, S., Kitanaka, J.-i., Takemura, M., Misawa, H., Ikawa, M., et al. (2004). The LIM homeobox gene, L3/Lhx8, is necessary for proper development of basal forebrain cholinergic neurons. *European Journal of Neuroscience* 19(12), 3129–3141.

Nash, B., Colavita, A., Zheng, H., Roy, P. J., and Culotti, J. G. (2000). The forkhead transcription factor UNC-130 is required for the graded spatial expression of the UNC-129 TGF-beta guidance factor in C. elegans. en. *Genes Dev.* 14(19), 2486–2500.

Negre, B., Casillas, S., Suzanne, M., Sánchez-Herrero, E., Akam, M., Nefedov, M., Barbadilla, A., Jong, P. de, and Ruiz, A. (2005). Conservation of regulatory sequences and gene expression patterns in the disintegrating Drosophila Hox gene complex. en. *Genome Res.* 15(5), 692–700.

Newman, A. P., Acton, G. Z., Hartwieg, E., Horvitz, H. R., and Sternberg, P. W. (1999). The lin-11 LIM domain transcription factor is necessary for morphogenesis of C. elegans uterine cells. en. *Development* 126(23), 5319–5326.

Newman, A. P. and Sternberg, P. W. (1996). Coordinated morphogenesis of epithelia during development of the Caenorhabditis elegans uterine-vulval connection. *Proceedings of the National Academy of Sciences* 93(18), 9329–9333.

O'Keefe, D. D., Thor, S., and Thomas, J. B. (1998). Function and specificity of LIM domains in Drosophila nervous system and wing development. en. *Development* 125(19), 3915–3923.

Passner, J. M., Ryoo, H.-D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999). *Structure of the Homeotic UBX/EXD/DNA ternary complex.*

Peckol, E. L., Troemel, E. R., and Bargmann, C. I. (2001). Sensory experience and sensory activity regulate chemosensory receptor gene expression in Caenorhabditis elegans. en. *Proc. Natl. Acad. Sci. U. S. A.* 98(20), 11032–11038.

Pérez-Alvarado, G. C., Miles, C., Michelsen, J. W., Louis, H. A., Winge, D. R., Beckerle, M. C., and Summers, M. F. (1994). Structure of the carboxy-terminal LIM domain from the cysteine rich protein CRP. en. *Nat. Struct. Biol.* 1(6), 388–398.

Pfaff, S. L., Mendelsohn, M., Stewart, C. L., Edlund, T., and Jessell, T. M. (1996). Requirement for LIM homeobox gene Isl1 in motor neuron generation reveals a motor neuron-dependent step in interneuron differentiation. en. *Cell* 84(2), 309–320.

Pierce-Shimomura, J. T., Faumont, S., Gaston, M. R., Pearson, B. J., and Lockery, S. R. (2001). erratum: The homeobox gene lim-6 is required for distinct chemosensory representations in C. elegans. *Nature* 412(6846), 566–566.

Piggott, B. J., Liu, J., Feng, Z., Wescott, S. A., and Xu, X. Z. S. (2011). The Neural Circuits and Synaptic Mechanisms Underlying Motor Initiation in C. elegans. *Cell* 147(4), 922–933.

Porter, F. D., Drago, J., Xu, Y., Cheema, S. S., Wassif, C., Huang, S.-P., Lee, E., Grinberg, A., Massalas, J. S., Bodine, D., et al. (1997). Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development. *Development* 124(15), 2935–2944.

Pueyo, J. I., Galindo, M. I., Bishop, S. A., and Couso, J. P. (2000). Proximal-distal leg development in Drosophila requires the apterous gene and the Lim1 homologue dlim1. *Development* 127(24), 5391–5402.

Qian, Y. Q., Billeter, M., Otting, G., Müller, M., Gehring, W. J., and Wüthrich, K. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressors. *Cell* 59(3), 573–580.

Raetzman, L. T., Ward, R., and Camper, S. A. (2002). Lhx4 and Prop1 are required for cell survival and expansion of the pituitary primordia. *Development* 129(18), 4229–4239.

Renfranz, P. J. and Beckerle, M. C. (2002). Doing (F/L)PPPPs: EVH1 domains and their proline-rich partners in cell polarity and migration. en. *Curr. Opin. Cell Biol.* 14(1), 88–103.

Rétaux, S., Rogard, M., Bach, I., Failli, V., and Besson, M.-J. (1999). Lhx9: a novel LIM-homeodomain gene expressed in the developing forebrain. *Journal of Neuroscience* 19(2), 783–793.

Riddle, R. D., Ensini, M., Nelson, C., Tsuchida, T., Jessell, T. M., and Tabin, C. (1995). Induction of the LIM homeobox gene Lmx1 by WNT6a establishes dorsoventral pattern in the vertebrate limb. *Cell* 83(4), 631–640.

Ringo, J., Werczberger, R., Altaratz, M., and Segal, D. (1991). Female sexual receptivity is defective in juvenile hormone-deficient mutants of theapterous gene ofDrosophila melanogaster. *Behavior genetics* 21(5), 453–469.

Rohr, C., Prestel, J., Heidet, L., Hosser, H., Kriz, W., Johnson, R. L., Antignac, C., and Witzgall, R. (2002). The LIM-homeodomain transcription factor Lmx1b plays a crucial role in podocytes. *The Journal of clinical investigation* 109(8), 1073–1082.

Sadler, I. (1992). Zyxin and cCRP: two interactive LIM domain proteins associated with the cytoskeleton. *J. Cell Biol.* 119(6), 1573–1587.

Sagasti, A., Hobert, O., Troemel, E. R., Ruvkun, G., and Bargmann, C. I. (1999). Alternative olfactory neuron fates are specified by the LIM homeobox gene lim-4. en. *Genes Dev.* 13(14), 1794–1806.

Salam, S., Sangeena, S., Ata, A., Siavash, A., Pouya, R., Ravi Selvaganapathy, P., Mishra, R. K., and Gupta, B. P. (2013). A microfluidic phenotype analysis system reveals function of sensory and dopaminergic neuron signaling in C. elegans electrotactic swimming behavior. *Worm* 2(2), e24558.

Sánchez-García, I., Osada, H., Forster, A., and Rabbitts, T. H. (1993). The cysteine-rich LIM domains inhibit DNA binding by the associated homeodomain in Isl-1. en. *EMBO J.* 12(11), 4243–4250.

Sarafi-Reinach, T. R., Melkman, T., Hobert, O., and Sengupta, P. (2001). The lin-11 LIM homeobox gene specifies olfactory and chemosensory neuron fates in C. elegans. en. *Development* 128(17), 3269–3281.

Sarafi-Reinach, T. R. and Sengupta, P. (2000). The forkhead domain gene unc-130 generates chemosensory neuron diversity in C. elegans. en. *Genes Dev.* 14(19), 2472–2485.

Schinkmann, K. and Li, C. (1992). Localization of FMRFamide-like peptides in Caenorhabditis elegans. en. *J. Comp. Neurol.* 316(2), 251–260.

Schmeichel, K. L. and Beckerle, M. C. (1997). Molecular dissection of a LIM domain. en. *Mol. Biol. Cell* 8(2), 219–230.

Segawa, H., Miyashita, T., Hirate, Y., Higashijima, S., Chino, N., Uyemura, K., Kikuchi, Y., and Okamoto, H. (2001). Functional repression of Islet-2 by disruption of complex with Ldb impairs peripheral axonal outgrowth in embryonic zebrafish. en. *Neuron* 30(2), 423–436.

Sengupta, P., Colbert, H. A., and Bargmann, C. I. (1994). The C. elegans gene odr-7 encodes an olfactory-specific member of the nuclear receptor superfamily. *Cell* 79(6), 971–980.

Serrano-Saiz, E., Poole, R. J., Felton, T., Zhang, F., De La Cruz, E. D., and Hobert, O. (2013). Modular Control of Glutamatergic Neuronal Identity in C. elegans by Distinct Homeodomain Proteins. *Cell* 155(3), 659–673.

Sharanya, D., Thillainathan, B., Marri, S., Bojanala, N., Taylor, J., Flibotte, S., Moerman, D. G., Waterston, R. H., and Gupta, B. P. (2012). Genetic control of vulval development in Caenorhabditis briggsae. en. *G3* 2(12), 1625–1641.

Sharma, K., Sheng, H. Z., Lettieri, K., Li, H., Karavanov, A., Potter, S., Westphal, H., and Pfaff, S. L. (1998). LIM homeodomain factors Lhx3 and Lhx4 assign subtype identities for motor neurons. en. *Cell* 95(6), 817–828.

Shawlot, W. and Behringer, R. R. (1995). Requirement for Lim1 in head-organizer function. *Nature* 374(6521), 425.

Sheng, H. Z., Zhadanov, A. B., Mosinger Jr, B., Fujii, T., Bertuzzi, S., Grinberg, A., Lee, E. J., Huang, S. P., Mahon, K. A., and Westphal, H. (1996). Specification of pituitary cell lineages by the LIM homeobox gene Lhx3. en. *Science* 272(5264), 1004–1007.

Sheng, H. Z., Moriyama, K., Yamashita, T., Li, H., Potter, S. S., Mahon, K. A., and Westphal, H. (1997). Multistep control of pituitary organogenesis. *Science* 278(5344), 1809–1812.

Shimell, M., Simon, J., Bender, W., and O'Connor, M. (1994). Enhancer point mutation results in a homeotic transformation in Drosophila. *Science* 264(5161), 968–971.

Shirasaki, R. and Pfaff, S. L. (2002). Transcriptional codes and the control of neuronal identity. *Annual review of neuroscience* 25(1), 251–281.

Shoya, Y., Kobayashi, T., Koda, T., Ikuta, K., Kakinuma, M., and Kishi, M. (1998). Two proline-rich nuclear localization signals in the amino- and carboxyl-terminal regions of the Borna disease virus phosphoprotein. en. *J. Virol.* 72(12), 9755–9762.

Simmer, F., Moorman, C., Linden, A. M. van der, Kuijk, E., Berghe, P. V. van den, Kamath, R. S., Fraser, A. G., Ahringer, J., and Plasterk, R. H. (2003). Genome-wide RNAi of C. elegans using the hypersensitive rrf-3 strain reveals novel gene functions. *PLoS Biol* 1(1), e12.

Simon, M. A., Dodson, G. S., Rubin, G. M., Olivier, J. P., Egan, S. E., Giddings, B. W., Brooks, M. W., Buday, L., Sizeland, A. M., Weinberg, R. A., Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., Bowtell, D., Li, N., Gale, N. W., Kaplan, S., Lowestein, E. J., Schlessinger, J., Bar-Sagi, D., Buday, L., Downward, J., and Chardin, P. (1993). Tyrosine kinases to Ras. *Trends Cell Biol.* 3(7), 228.

Smidt, M. P., Asbreuk, C. H., Cox, J. J., Chen, H., Johnson, R. L., and Burbach, J. P. H. (2000). A second independent pathway for development of mesencephalic dopaminergic neurons requires Lmx1b. *Nature neuroscience* 3(4), 337–341.

Sprenger-Haussels, M. and Weisshaar, B. (2000). Transactivation properties of parsley proline-rich bZIP transcription factors. *Plant J.* 22(1), 1–8.

Srivastava, M., Larroux, C., Lu, D. R., Mohanty, K., Chapman, J., Degnan, B. M., and Rokhsar, D. S. (2010). Early evolution of the LIM homeobox gene family. en. *BMC Biol.* 8, 4.

Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H. A., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C., Willey, D., Wilson, R. K., Durbin, R., and Waterston, R. H. (2003). The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. en. *PLoS Biol.* 1(2), E45.

Stevens, M. E. and Bryant, P. J. (1985). Apparent genetic complexity generated by developmental thresholds: the apterous locus in Drosophila melanogaster. *Genetics* 110(2), 281–297.

Sudol, M., Sliwa, K., and Russo, T. (2001). Functions of WW domains in the nucleus. *FEBS Lett.* 490(3), 190–195.

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev. Biol.* 56(1), 110–156.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* 100(1), 64–119.

Takuma, N., Sheng, H. Z., Furuta, Y., Ward, J. M., Sharma, K., Hogan, B., Pfaff, S. L., Westphal, H., Kimura, S., and Mahon, K. A. (1998). Formation of Rathke's pouch requires dual induction from the diencephalon. *Development* 125(23), 4835–4840.

Thaler, J. P., Lee, S.-K., Jurata, L. W., Gill, G. N., and Pfaff, S. L. (2002). LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. *Cell* 110(2), 237–249.

Thor, S., Andersson, S. G., Tomlinson, A., and Thomas, J. B. (1999). A LIM-homeodomain combinatorial code for motor-neuron pathway selection. *Nature* 397(6714), 76–80.

Thor, S. and Thomas, J. B. (1997). The Drosophila islet gene governs axon pathfinding and neurotransmitter identity. *Neuron* 18(3), 397–409.

Tsalik, E. L., Niacaris, T., Wenick, A. S., Pau, K., Avery, L., and Hobert, O. (2003). LIM homeobox gene-dependent expression of biogenic amine receptors in restricted regions of the C. elegans nervous system. *Dev. Biol.* 263(1), 81–102.

Tsang, T. E., Shawlot, W., Kinder, S. J., Kobayashi, A., Kwan, K. M., Schughart, K., Kania, A., Jessell, T. M., Behringer, R. R., and Tam, P. P. (2000). Lim1 activity is required for intermediate mesoderm differentiation in the mouse embryo. *Developmental biology* 223(1), 77–90.

Tsuchida, T., Ensini, M., Morton, S., Baldassare, M., Edlund, T., Jessell, T., and Pfaff, S. (1994). Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. *Cell* 79(6), 957–970.

Tsuji, T., Sato, A., Hiratani, I., Taira, M., Saigo, K., and Kojima, T. (2000). Requirements of Lim1, a Drosophila LIM-homeobox gene, for normal leg and antennal development. *Development* 127(20), 4315–4323.

Tu, J. C., Xiao, B., Yuan, J. P., Lanahan, A. A., Leoffert, K., Li, M., Linden, D. J., and Worley, P. F. (1998). Homer binds a novel proline-rich motif and links group 1 metabotropic glutamate receptors with IP3 receptors. en. *Neuron* 21(4), 717–726.

Vogel, A., Rodriguez, C., Warnken, W., and Belmonte, J. C. I. (1995). Dorsal cell fate specified by chick Lmx1 during vertebrate limb development. *Nature* 378(6558), 716.

Voutev, R., Keating, R., Hubbard, E. J. A., and Vallier, L. G. (2009). Characterization of the Caenorhabditis elegans Islet LIM-homeodomain ortholog, lim-7. en. *FEBS Lett.* 583(2), 456–464.

Wang, D., O'Halloran, D., and Goodman, M. B. (2013). GCY-8, PDE-2, and NCS-1 are critical elements of the cGMP-dependent thermotransduction cascade in the AFD neurons responsible for C. elegans thermotaxis. en. *J. Gen. Physiol.* 142(4), 437–449.

Way, J. C. and Chalfie, M. (1988). mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans. en. *Cell* 54(1), 5–16.

Way, J. C. and Chalfie, M. (1989). The mec-3 gene of Caenorhabditis elegans requires its own product for maintained expression and is expressed in three neuronal cell types. en. *Genes Dev.* 3(12A), 1823–1833.

Wenick, A. S. and Hobert, O. (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. en. *Dev. Cell* 6(6), 757–770.

White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986). The Structure of the Nervous System of the Nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 314(1165), 1–340.

Wightman, B., Ebert, B., Carmean, N., Weber, K., and Clever, S. (2005). The C. elegans nuclear receptor gene fax-1 and homeobox gene unc-42 coordinate interneuron identity by regulating the expression of glutamate receptor subunits and other neuron-specific genes. *Developmental biology* 287(1), 74–85.

Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. *Biochem. J* 297(2), 249–260.

Winchell, C. J. and Jacobs, D. K. (2013). Expression of the Lhx genes apterous and lim1 in an errant polychaete: implications for bilaterian appendage evolution, neural development, and muscle diversification. *EvoDevo* 4(1), 4.

Wollesen, T., McDougall, C., Degnan, B. M., and Wanninger, A. (2014). POU genes are expressed during the formation of individual ganglia of the cephalopod central nervous system. en. *Evodevo* 5, 41.

Xue, D., Tu, Y., and Chalfie, M. (1993). Cooperative interactions between the Caenorhabditis elegans homeoproteins UNC-86 and MEC-3. *Science* 261(5126), 1324–1328.

Yamada, K., Tsuchiya, J.-I., and Iino, Y. (2012). Mutations in the pqe-1 gene enhance transgene expression in Caenorhabditis elegans. en. *G3* 2(7), 741–751.

Yamashita, T., Moriyama, K., Sheng, H. Z., and Westphal, H. (1997). Lhx4, a LIM homeobox gene. *Genomics* 44(1), 144–146.

Zákány, J., Gérard, M., Favier, B., and Duboule, D. (1997). Deletion of a HoxD enhancer induces transcriptional heterochrony leading to transposition of the sacrum. en. *EMBO J.* 16(14), 4393–4402.

Zarrinpar, A., Bhattacharyya, R. P., and Lim, W. A. (2003). The structure and function of proline recognition domains. *Homo* 332(80), 20.

Zhadanov, A. B., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Westphal, H. (1995). Genomic structure and chromosomal localization of the mouse LIM/homeobox gene Lhx3. *Genomics* 27(1), 27–32.

Zhao, Y., Sheng, H. Z., Amini, R., Grinberg, A., Lee, E., Huang, S., Taira, M., and Westphal, H. (1999). Control of hippocampal morphogenesis and neuronal differentiation by the LIM homeobox gene Lhx5. *Science* 284(5417), 1155–1158.

Zheng, C., Diaz-Cuadros, M., and Chalfie, M. (2015). Hox Genes Promote Neuronal Subtype Diversification through Posterior Induction in Caenorhabditis elegans. en. *Neuron* 88(3), 514–527.

Zheng, X., Chung, S., Tanabe, T., and Sze, J. Y. (2005). Cell-type specific regulation of serotonergic identity by the C. elegans LIM-homeodomain factor LIM-4. *Dev. Biol.* 286(2), 618–628.

Zhong, W. and Sternberg, P. W. (2006). Genome-wide prediction of C. elegans genetic interactions. *Science* 311(5766), 1481–1484.