Ph. D. Thesis
박사 학위논문

Neuronal specification of a motor neuron in *C. elegans*

Jinmahn Kim (김 진 만 金 珍 滿)

Department of
Brain and Cognitive Sciences

DGIST

2018

Ph. D. Thesis
박사 학위논문

Neuronal specification of a motor neuron in *C. elegans*

Jinmahn Kim (김 진 만 金 珍 滿)

Department of
Brain and Cognitive Sciences

DGIST

2018

Neuronal specification of a motor neuron in *C. elegans*

Advisor: Professor Kyuhyung Kim
Co-advisor: Professor Yang Hoon Huh

by

Jinmahn Kim
Department of Brain and Cognitive Sciences
DGIST

A thesis submitted to the faculty of DGIST in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Brain and Cognitive Sciences. The study was conducted in accordance with Code of Research Ethics[1]

05 17. 2018

Approved by

Professor (Advisor)	Kyuhyung Kim	(signature)
Professor (Co-Advisor)	Yang Hoon Huh	(signature)

[1] Declaration of Ethical Conduct in Research**:** I, as a graduate student of DGIST, hereby declare that I have not committed any acts that may damage the credibility of my research. These include, but are not limited to: falsification, thesis written by someone else, distortion of research findings or plagiarism. I affirm that my thesis contains honest conclusions based on my own careful research under the guidance of my thesis advisor.

Neuronal specification of a motor neuron in *C. elegans*

Jinmahn Kim

Accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

05. 17. 2018

Head of Committee	(signature)
	Prof. Yang Hoon Huh
Committee Member	(signature)
	Prof. Kyuhyung Kim
Committee Member	(signature)
	Prof. Sung Bae Lee
Committee Member	(signature)
	Prof. Yong-Seok Oh
Committee Member	(signature)
	Prof. Han Kyoung Choe

Ph. D./ BS
201145009

김 진 만. Jinmahn Kim. Neuronal specification of a motor neuron in *C. elegans.* Department of Brain and Cognitive Sciences. 2018. 70p. Advisors Prof. Kyuhyung Kim, Co-Advisors Prof. Yang Hoon Huh

Abstract

The expression of specific transcription factors determines the differentiated features of postmitotic neurons. However, the mechanism by which specific molecules determine neuronal cell fate and the extent to which the functions of transcription factors are conserved in evolution are not fully understood. In *C. elegans*, the cholinergic and peptidergic SMB sensory/inter/motor neurons innervate muscle quadrants in the head and control the amplitude of sinusoidal movement. Here I showed that the LIM homeobox protein LIM-4 determines neuronal characteristics of the SMB neurons. In *lim-4* mutant animals, expression of terminal differentiation genes, such as the cholinergic gene battery and the *flp-12* neuropeptide gene, is completely abolished and thus the function of the SMB neurons is compromised. LIM-4 activity promotes SMB identity by directly regulating the expression of the SMB marker genes via a distinct *cis*-regulatory motif. Two human LIM-4 orthologs, LHX6 and LHX8, functionally substitute for LIM-4 in *C. elegans*. Furthermore, *C. elegans* LIM-4 or human LHX6 can induce cholinergic and peptidergic characteristics in the human neuronal cell lines. My results indicate that the evolutionarily conserved LIM-4/LHX6 homeodomain proteins function in generation of precise neuronal subtypes.

.

Keywords : *lim-4*, SMB, transcription factors, terminal selector gene, *C. elegans*

List of Contents

List of tables

List of figures

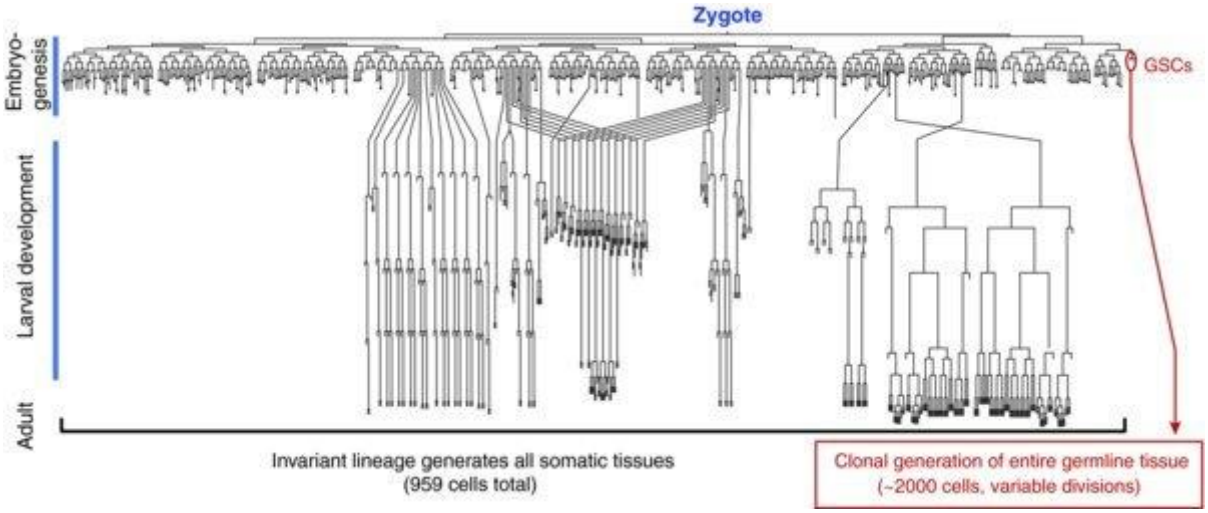
I. Introduction

The proper generation and maintenance of cells in the nervous system is essential for multi-cellular organisms. Each neuron achieves its identity by the acquisition of many distinct features, including appropriate synaptic contacts and expression of distinct sets of neurotransmitters. Fate determination and specification of neuronal cells largely relies on interactions between *trans*-acting transcription factors and *cis*-regulatory elements of their target genes [1, 2]. The same transcription factors may be used again after neuronal cell fate determination to maintain the neuron's integrity [3]. However, it has been challenging not only to discover the transcription factors, but also to identify their regulatory mechanisms and target genes critically associated with determination and maintenance of neuronal cell fate.

1.1 Historical background

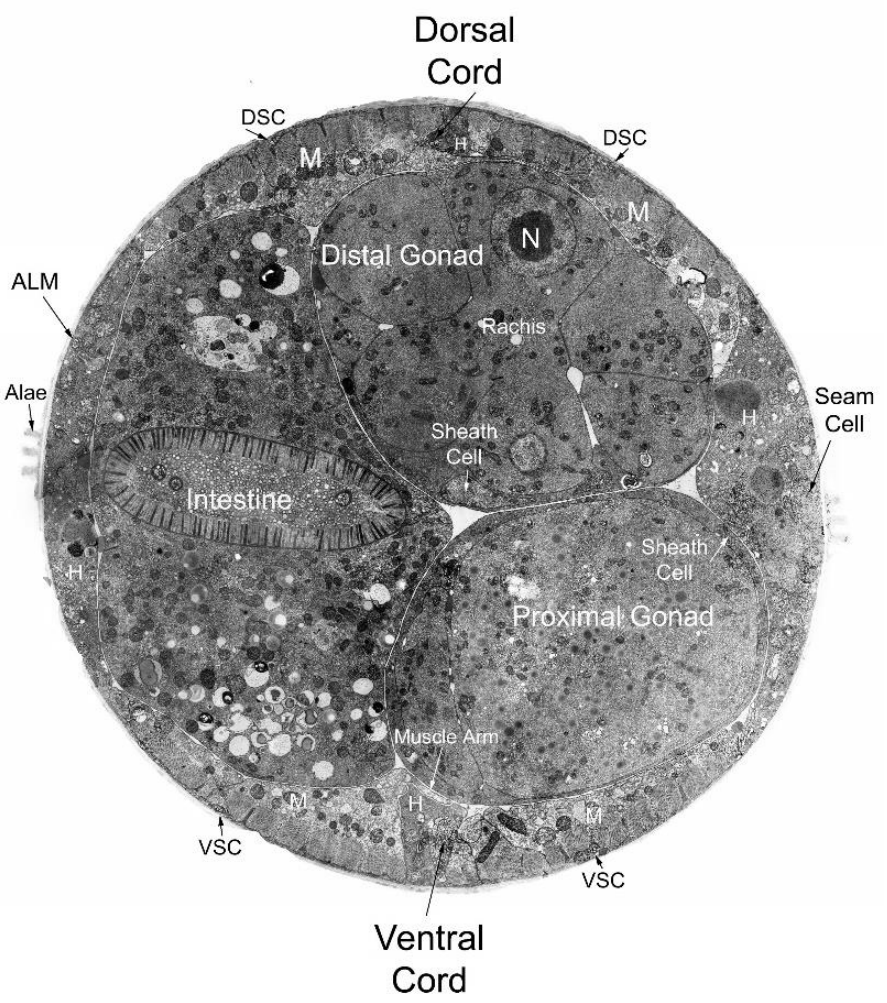
The nematode *Caenorhabditis elegans* was established as a model system by Sydney Brenner and its nervous system has been well-defined on the levels of cell-lineage, anatomy, and function [4]. In 1983, a comprehensive lineage map was completed of which describes the entire developmental cell lineage from the egg to the adult of the worms by using a light microscope with the Nomarski optics (Figure 1) [5]. At the same time, the neuronal structures and connectivity in nervous system were examined using electron microscopy (EM) which led to elucidate precise positions and shapes of individual neurons and construct wiring diagram of the 302 neurons (Figures 2) [6]. The knowledge of anatomical structure of neurons allowed specific ablation of neuronal cell bodies which revealed putative functions for most of sensory neurons, motor neurons and more than half of all interneurons (Table 1). Moreover, functions of each neurons have been confirmed partly by green fluorescent protein (*gfp*) reporter gene technology developed in *C. elegans* [7]. The expression pattern of numerous reporter genes within the nervous system have been analyzed at single neuron resolution and

showed spatial and terminal expression of neurotransmitter receptors, neurotransmitter synthesizing enzymes and transporters, ion channels, and neuronal signaling factors (Figure 3). Additionally, extensive genetic mutant screens via EMS mutagenesis identified mutant animals in which specific behaviors such as thermotaxis [8], mechanosensory [9], or chemotaxis behavior were affected [10, 11]. The molecular characterizations of these behavioral mutants showed that genes corresponding to these mutants encode transcription factors that regulate the functional properties of the neurons (Table 2) [12-20]. Mutations of the transcription factor genes broadly affected expressions of specific neuronal markers that determine terminal identity of the individual neurons and caused defects in behavioral phenotypes similar to those of the neuronal ablated worms (Table 1 and 2). Given the direct regulation of terminal identity features of a neuron, theses transcription factors have been termed as "terminal selectors" [2, 21].



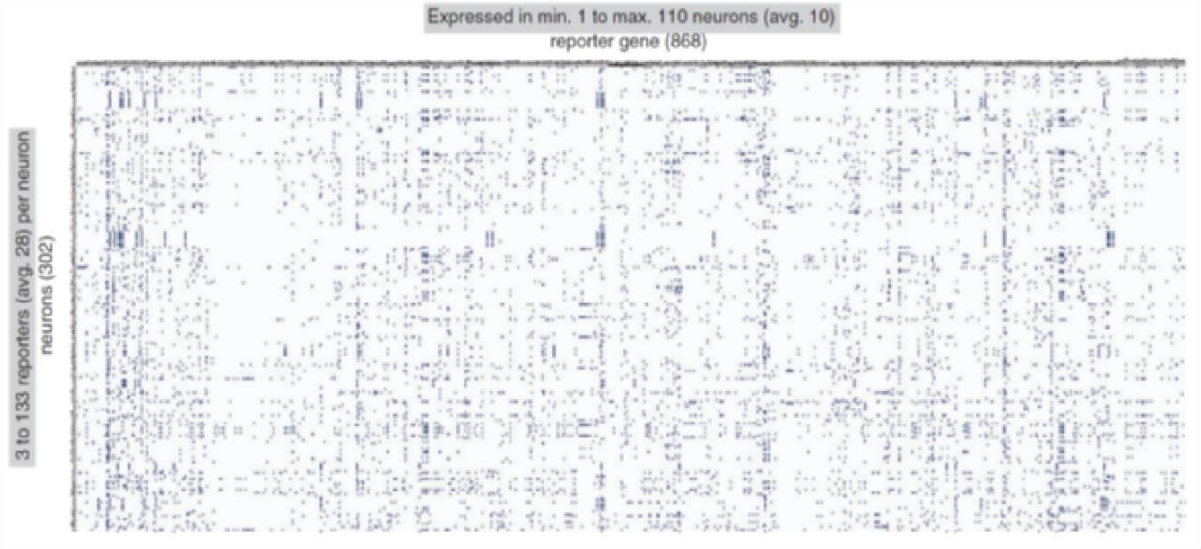
C. elegans germline stem cells and their niche in StemBook

Figure 1. A lineage map of the *C. elegans* nervous system. Lineage of cells generated in the embryo. X-axis indicates types of cells and tissues. Y-axis indicates developmental time. Germline stem cells subsequently divide variably to make the germline tissue during larval development and to maintain it in adults [22-24] (Adopted from a chapter of StemBook, 2008).



Wormatlas

Figure 2. An EM image of a cross-section through the midbody of *C. elegans*. This ultrastructural cross-section shows the location and morphology of a variety of anatomical structures. A thin tough cuticle surrounds the animal. The syncytial epidermis (or hypodermis; labeled H) is very thin where body wall muscle (M) is present. Several different processes bundle of the nervous system are visible including the ventral nerve cord, the dorsal nerve cord, the ventral and dorsal sublateral cords (DSC and VSC) and the mechanosensory process ALM. Four distinct quadrants of body wall muscle (M) are seen, two in ventral and two in dorsal sublateral positions. The intestine and the gonad fill most of the pseudocoelomic cavity. The diameter of an adult is approximately 60 nm (From www.wormatlas.org by David H. Hall).



Wormbase

Figure 3. Expression patterns of transgenes in the *C. elegans* nervous system. Manually curated expression patterns of transgenes throughout the *C. elegans* nervous system. Note that theses transgenes may not display the complete expression pattern of the respective gene, but they nevertheless serve as invaluable readouts of the regulatory state of a neuron (From www.wormbase.org by Daniela Raciti and Wen Chen and organized by Lori Glenwinkel).

Table 1. Functions of individual neurons in the *C. elegans* nervous system

Neuron type	Neuron	Function
Sensory Neuron	ADE	Mechanosensory [25, 26]
	ADF	Dauer formation [27]
	ADL	Olfaction (repulsive) [28]
	AFD	Thermosensation [29]
	ALA	Sleep, mechanosensory [30, 31]
	ALM	Mechanosensory [9]
	ALN	Aggregation, O_2 sensation [32, 33]
	AQR	Aggregation, O_2 sensation [32, 34]
	ASE	Gustation [35]
	ASG	Gustation [27, 35]
	ASH	Mechanosensory, gustation [36]
	ASI	Dauer formation, feeding state, Chemosensory, thermosensory [27, 35, 37]
	ASJ	Dauer formation and recovery [27]
	ASK	Gustation [35]
	AVM	Mechanosensory [31]
	AWA	Olfaction[35]
	AWB	Olfaction[35]
	AWC	Olfaction, thermosensation[35, 38]
	BAG	CO_2 Sensation [39]
	CEP	Mechanosensory [25, 26]
	FLP	Mechanosensory [36]
	IL1	Head bends, foraging [40]
	IL2	Nictation behavior [41]
	OLL	Pathogen avoidance behavior [42]
	OLQ	Mechanosensory, head bends, foraging [36, 40]
	PDE	Mechanosensory [25, 26]

	PHA	Gustation, mechanosensory [43, 44]
	PHB	Gustation, mechanosensory [43, 44]
	PHC	Heat avoidance [43, 45]
	PLM	Mechanosensory [9]
	PLN	Aggregation, O_2 sensation [32, 33]
	PQR	Aggregation, O_2 sensation [32, 34]
	PVD	Mechanosensory, thermosensory [46]
	PVM	Mechanosensory [9]
	URA	Unknown
	URB	Unknown
	URX	O_2 sensation [32, 34]
	URY	Mate-searching behavior of males, pathogen sensing [47, 48]
Interneuron	ADA	Unknown
	AIA	Processing sensory information [49]
	AIB	Reversal, pausing, integrating sensory information [49-51]
	AIM	Sexual attraction
	AIN	Unknown
	AIY	Locomotion, integrating sensory information [52, 53]
	AIZ	Locomotion, processing sensory information [52, 53]
	AUA	Aggregation [32]
	AVA	Command interneurons [52, 54]
	AVB	Command interneurons [52, 54]
	AVD	Command interneurons [52, 54]
	AVE	Command interneurons [52, 54]
	AVF	Egg-laying/ locomotion [55]
	AVG	Pioneer/ guidepost[56]
	AVH	Unknown
	AVJ	Unknown
	AVK	Unknown

	BDU	Mechanosensory [43]
	CAN	Fluid homeostasis
	DVA	Stretch reception [57]
	DVC	Unknown
	LUA	Pausing
	PVC	Command interneurons [52]
	PVN	Unknown
	PVP	Pioneer [58]
	PVQ	Pioneer [58]
	PVR	Unknown
	PVT	Pioneer/ guidepost [59]
	PVW	Unknown
	RIA	Thermosensory processing, quiescence [29, 60]
	RIB	Reversal behavior [50, 53]
	RIC	Feeding circuit [61]
	RID	Unknown
	RIF	Dwelling [62]
	RIG	Unknown
	RIH	Processes nose touch [63]
	RIP	Coupling touch and pharyngeal pumping [52]
	RIR	Unknown
	RIS	Quiescence [64]
	RIV	Omega turns [50]
	SAA	Unknown
	SDQ	Aggregation, O_2 sensation [32, 33]
Motorneuron	AS	Unknown
	AVL	Defecation [65]
	DA	Locomotion [52]
	DB	Locomotion [52]
	DD	Locomotion [65]

	DVB	Defecation [65]
	HSN	Egg-laying
	PDA	Unknown
	PDB	Unknown
	RIM	Reversal behavior, feeding circuit [50, 61]
	RMD	Head movement, foraging [40, 50]
	RME	Head movement [65]
	RMF	Unknown
	RMG	Aggregation [66]
	RMH	Unknown
	SAB	Head movement [67]
	SIA	Guidepost [68]
	SIB	Guidepost [68]
	SMB	Sinusoidal locomotion [50]
	SMD	Omega turns [50]
	VA	Locomotion [52]
	VB	Locomotion [52]
	VC	Egg-laying [69]
	VD	Locomotion [65]

Table 2. Terminal selectors regulate behavioral phenotype via the expression in the specific type of neuron

Terminal selector	Transcription factor type	Phenotype	Neuron class affected
UNC-3	EBF/COE	Uncoordinated locomotion [70]	Cholinergic ventral cord motoneurons [16, 67, 71]
UNC-30	Pitx-type homeodomain	Uncoordinated locomotion [70]	GABAergic ventral cord motoneurons [15]
			Cholinergic PVP interneurons [72]
UNC-42	Homeodomain	Uncoordinated locomotion [70]	Glutamatergic ASH sensory neurons [73]
			Glutamatergic AIB interneurons [73]
UNC-86	Brn3-type POU homeodomain	Mechnosensory defects [9]	Mechnosensory ALM/AVM/PLM/PVM neurons [74]
			Serotonergic HSN motor neurons [75]
			Serotonergic NSM neurons [76]
			Inner labial IL2 sensory-motor neurons [76]
			URA motor neurons [76]
			URB interneurons [76]
			Glutamatergic URX sensory neurons [73]
			Glutamatergic AQR/PQR sensory neurons [73]
			Glutamatergic PVR interneuron [73]
MEC-3	Lhx1/5 LIM homeodomain	Mechanosensory defects [9]	Mechnosensory ALM/AVB/PLM/PVM neurons [19, 74, 77]
CHE-1	C2H2 Zn finger	Chemotaxis defects [11]	Gustatory ASE sensory neurons [20, 29, 78]
TTX-1	Otx homeodomain	Thermotaxis defects [79]	Thermosensory AFD neuron [17]
TTX-3	Lhx2/7 LIM homeodomain	Thermotaxis defects [79]	Cholinergic AIY interneuron [29, 53, 80]
			Cholniergic AIA interneuron [75]
			Serotonergic NSM neuron [75]
			Glutamatergic ASK sensory neuron [73]
ODR-7	C4 Zn finger	Odortaxis defects [10]	Olfactory AWA neuron [35]

1.2 Terminal selector transcription factors coregulate expression of terminal differentiation genes

The neuronal identity is defined by unique expressions of multiple terminal differentiation genes that encode neurotransmitter type, neuropeptides, cytoskeletal proteins, and cell surface adhesion molecules (Figure 4). Expression of these genes are coregulated by terminal selectors through the shared *cis*-regulatory motifs. For example, The CHE-1 zinc finger transcription factor coregulates the expression of terminally differentiated features of ASE glutamatergic neuron including neuropeptide genes and putative chemoreceptors directly via AT-rich ASE motif (Figure 4A) [15, 81]. The EBF-type transcription factor *unc-3* controls the cholinergic gene battery in ventral nerve cord motor neurons by binding COE motif (Figure 4B) [71]. Additionally, the genes encoding the dopamine-synthesizing enzymes (*cat-2*, *bas-1*, and *bas-4)* and the dopamine transporters (*cat-1* and *dat-1*) share a common *cis*-regulatory motifs for *ast-1* and *ceh-43* transcription factors of which are required to drive expression in all dopaminergic neurons [82]. These *cis*-regulatory motifs may occur in multiple copies and are often, but not always, in close proximity (<1 kb) to the genes they regulate. In some cases, they have also been found many kilobases upstream of a gene or within introns. Other than neuronal type depending upon neurotransmitter expression, terminal selectors also coordinately regulate identity additional features of neurons, including sensory receptors, neuropeptides, ion channels, cytoskeletal proteins, and cell surface adhesion molecules (Figure 4). Upon knockout of the terminal selector, the expression of terminal makers is lost and the neuronal fate are not determined.

The specificity of individual neuronal features also could be determined by combinatorial codes of terminal selectors [83]. For example, the *ttx-3* LIM homeobox gene is a terminal selector for a cholinergic interneuron AIY, but is also expressed in several sensory neurons, where it has no apparent terminal selector function [80]. However, AIY neurons express an-

other gene, the *ceh-10* Prd homeobox gene, which is also expressed in several distinct neuron types [80]. But the expression of *ttx-3* and *ceh-10* overlaps exclusively in AIY, where the two proteins heterodimerize to cooperatively activate downstream terminal differentiation genes (Figure 4C) [84]. Taken together, the single or combination terminal selectors coregulate expression of terminal differentiation genes through the *cis*-regulatory motif for determining individual neuronal identity and generate diversification of neuronal type in the nervous system.

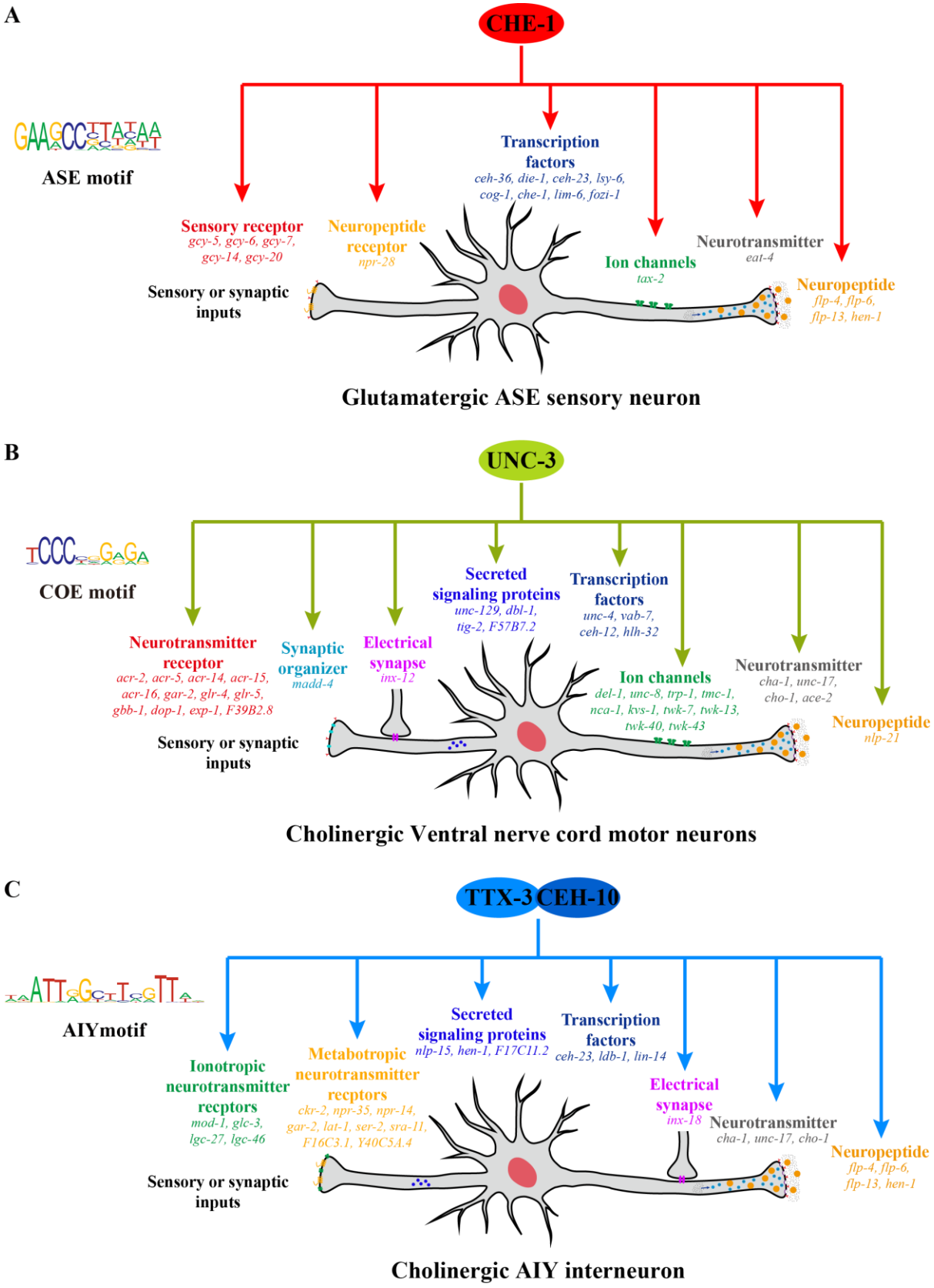

A
CHE-1
ASE motif
Transcription factors
ceh-36, die-1, ceh-23, lsy-6, cog-1, che-1, lim-6, fozi-1
Sensory receptor
gcy-5, gcy-6, gcy-7, gcy-14, gcy-20
Neuropeptide receptor
npr-28
Neurotransmitter
eat-4
Ion channels
tax-2
Neuropeptide
flp-4, flp-6, flp-13, hen-1
Sensory or synaptic inputs
Glutamatergic ASE sensory neuron
B
UNC-3
COE motif
Secreted signaling proteins
unc-129, dbl-1, tig-2, F57B7.2
Transcription factors
unc-4, vab-7, ceh-12, hlh-32
Neurotransmitter receptor
acr-2, acr-5, acr-14, acr-15, acr-16, gar-2, glr-4, glr-5, gbb-1, dop-1, exp-1, F39B2.8
Synaptic organizer
madd-4
Electrical synapse
inx-12
Ion channels
del-1, unc-8, trp-1, tmc-1, nca-1, kvs-1, twk-7, twk-13, twk-40, twk-43
Neurotransmitter
cha-1, unc-17, cho-1, ace-2
Neuropeptide
nlp-21
Sensory or synaptic inputs
Cholinergic Ventral nerve cord motor neurons
C
TTX-3 CEH-10
AIYmotif
Secreted signaling proteins
nlp-15, hen-1, F17C11.2
Transcription factors
ceh-23, ldb-1, lin-14
Ionotropic neurotransmitter receptors
mod-1, glc-3, lgc-27, lgc-46
Metabotropic neurotransmitter receptors
ckr-2, npr-35, npr-14, gar-2, lat-1, ser-2, sra-11, F16C3.1, Y40C5A.4
Electrical synapse
inx-18
Neurotransmitter
cha-1, unc-17, cho-1
Neuropeptide
flp-4, flp-6, flp-13, hen-1
Sensory or synaptic inputs
Cholinergic AIY interneuron

Figure 4. Terminal selectors induce expression of the terminally differentiated properties of mature neuron type. All genes shown here were shown to be direct targets of the indicated terminal selectors [71, 84, 85]. In addition to the genes shown here, scores of additional genes have been identified as being expressed in the respective neuron types and containing binding sites for the respective regulators, but these additional genes have not been validated for terminal for terminal selector-dependence.

1.3 Terminal selectors initiate and maintain the terminally differentiated state

A key defining feature of terminal selectors is that they are not only required to initiate terminal differentiation programs but are continuously expressed throughout the life of the neuron to maintain the differentiated state of a neuron [71, 82, 86, 87]. Such maintained expression has been explicitly shown to be mediated by direct autoregulation [78, 84, 88, 89]. For example, the LIM homeobox transcription factor *mec-3* autoregulates its expression in touch sensory neurons for their terminal differentiation [19]. The Zn finger transcription factors *che-1* and *odr-7* autoregulate their expression in the ASE and AWA sensory neurons, respectively [78, 90], and the LIM homeobox transcription factor *ttx-3* and the Prd-type homeobox transcription factor *ceh-10* jointly autoregulate their expression in the cholinergic AIY interneurons, aside from initiating the terminal differentiation program of AIY [84]. In all of the cases, the autoregulation is directly regulated by themselves. The *cis*-regulatory motifs of terminal selectors were found in their upstream region to maintain its own expressions [84, 91]. The regulatory factors that initiate terminal selector gene expression are transiently expressed in immature neurons or neuroblasts upon multiple regulatory signal inputs [91-94]. Then, the initiated terminal selectors autoregulate their expression and bind *cis*-regulatory motif to maintain terminal differentiation gene expressions (Figure 5).

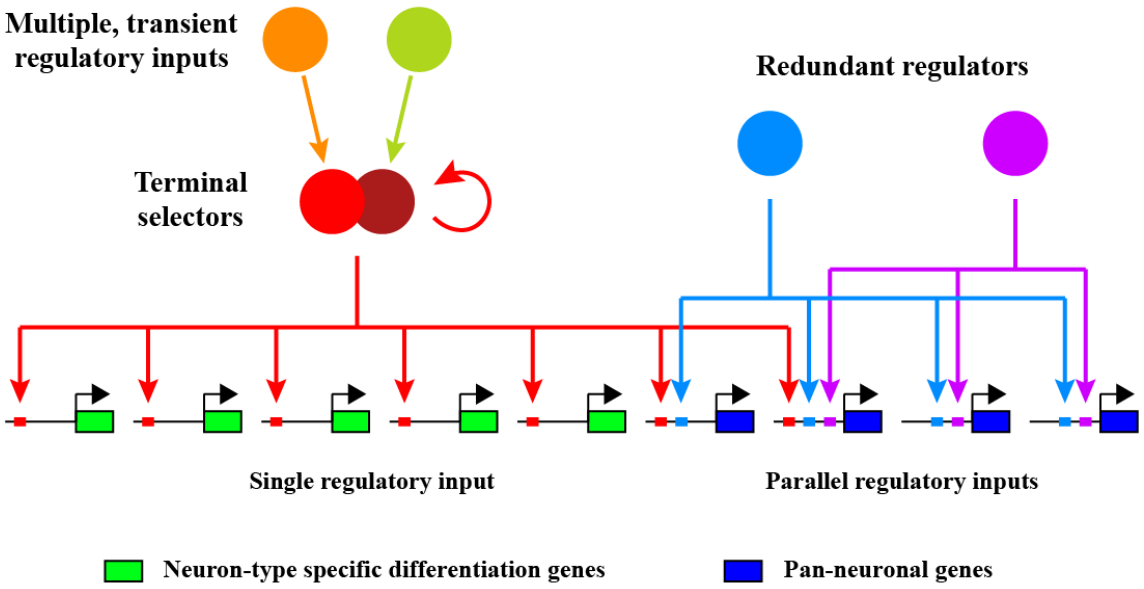


Figure 5. Regulation of neuron-type specific and pan-neuronal features. Multiple regulatory inputs (orange and green circles) are transiently expressed and activate terminal selectors. The terminal selectors (red circles) are usually form heterodimerization and autoregulate their expression. The terminal selectors regulate expressions of terminal differentiation genes by binding *cis*-regulatory motif (red rectangle). Pan-neuronal genes (blue rectangle) are regulated by redundant regulators (blue and purple circles) by binding the motif. *cis*-regulatory motifs of its terminal selector are indicated by corresponding color.

1.4 Regulations of pan-neuronal features

In *C. elegans*, the knockout of a terminal selector gene results in loss of the expression of many terminal identity of a neuron, but pan-neuronal genes are commonly unaffected [71, 78, 80, 82]. For example, sensory neurons continue to express pan-neuronal and pan-sensory features upon removal of the *che-1* gene, a terminal selector for gustatory neuron [20], or upon removal of *ast-1*, a terminal selector for dopaminergic neurons [82]. Components of terminal neuron identity that are not affected upon loss of these terminal selectors may be regulated by another regulatory factors acting in parallel. For example, many or most components of all ciliated sensory neurons, are coregulated via a common *cis*-regulatory motif (the X box) and a transcription factor DAF-19, phylogenetically conserved RFX-type factor [95]. In *unc-86* terminal selector mutants, the HSN motor neuron fails to adopt terminal identity features, but still possesses normal neuronal morphology [96]. Hence, there must be factors that act in parallel to terminal selectors to determine basic morphological features. In *Drosophila*, the function of transcription factors "morphology TF" (mTFs) has been reported to only define axonal/dendritic patterning features of motor neurons, but not other terminal features [97]. The mTFs is transiently expressed and regulates downstream genes involved in axon pathfinding or cell migration, suggesting that gene expression programs in terminally differentiated neurons may be controlled by parallel acting regulators that define distinct identity aspects. Taken together, terminal selectors are constitutively present in a neuron to initiate and maintain constitutively expressed terminal identity features of a neuron, while redundant regulators define more generic aspects of a neuron identity, such as pan-neuronal and morphological features in a distinct manner (Figure 5).

In the nematode *Caenorhabditis elegans*, about 40% of nervous system (~120 neurons) appear to be cholinergic, including a subset of motor neurons in the ventral nerve cord and

several sensory, motor and interneurons in the head, and many of these cholinergic neurons co-express neuropeptides [98, 99]. The SMB multimodal sensory/inter/motor neurons consist of two pairs of neurons that are located in the head and innervate the head neck muscles (Figure 6). Their processes, which run in ventral or dorsal sublateral cords to the tail and have electric and chemical synaptic contacts to other neurons in the head, were proposed to sense the stretch of body and regulate head locomotion [6]. In fact, laser ablation of the SMB neurons caused increased reversal frequency and wave amplitude of forward locomotion [50]. These neurons utilize at least two neurotransmitters, acetylcholine and a FMRFamide-related peptide, FLP-12 [99, 100]. Genes or molecules that are pivotal for the generation or differentiation of these SMB neurons have not been identified.

Here, I showed that the LIM homeodomain LIM-4 protein is necessary to drive expression of terminal differentiation genes, including the cholinergic gene battery and the *flp-12* neuropeptide gene, but not pan-neuronal genes in the SMB neurons; consequently, in *lim-4* mutants, the neuronal function of the SMB neurons is abolished. I found that LIM-4 maintains its own expression by autoregulation in the SMB neurons and ectopic expression of LIM-4 is sufficient to drive expression of the SMB marker in other cell types. Moreover, our promoter analyses and bioinformatic searches with the SMB marker genes identified a *cis*-regulatory motif that is necessary and sufficient to drive gene expression in the SMB neurons. I also showed that two *lim-4* human orthologs, *LHX6* and *LHX8,* functionally substitute for *lim-4* in *C. elegans*. Furthermore, expression of *C. elegans* LIM-4 or human LHX6 in the human neuroblastoma cell line induces cholinergic and peptidergic characteristics. I propose that there is an evolutionarily conserved role of *lim-4*/*LHX6*/*LHX8* LIM homeobox genes as terminal selectors to differentiate cholinergic and peptidergic neuronal cells and provide insight into how neuronal characteristics such as neurotransmitter identity are acquired via *trans*-acting and *cis*-regulatory mechanisms.

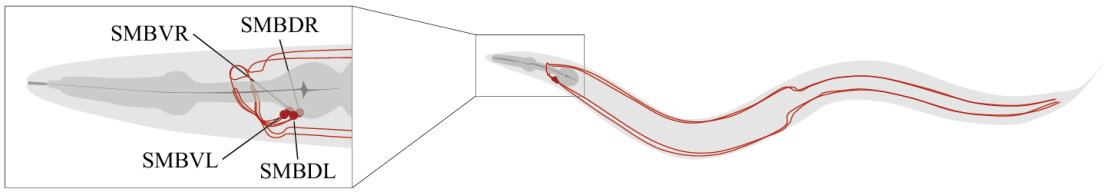


Figure 6. Schematic drawing of the SMB neurons in *C. elegans*. Four cell bodies are located in the head (DL: dorsal left, DR: dorsal right, VL: ventral left, VR: ventral right) and their processes run in sublateral cords to the tail.

II. Materials and Methods

Strains

N2 Bristol strain was used as wild-type strain. Mutant strain and transgenic strains used in this study are listed in S4 Table. All strains were maintained at 20°C.

Isolation of *lim-4* mutants

The *flp-12*p*::gfp(ynIs25)* integrated strain was used to performed EMS mutagenesis according to Sulston and Hodgkin (1988). Eight alleles (*yn19, lsk1, lsk2, lsk3, lsk4, lsk5, lsk6, lsk7*) in which GFP expression was completely abolished, were isolated from screening ~15,000 haploid genomes, found to be allelic to each other, and mapped on LG X. Based on three factor crosses using the double mutants *unc-6 dpy-6, dpy-8 unc-6,* and *unc-2 dpy-8, yn19* was located approximately 2.4 MU downstream of *unc-2* and 4.3 MU upstream of *dpy-8*. In this region, I did complementation tests with a *lim-4(ky403*) mutant and found that they were allelic. The molecular lesions were identified by sequencing amplification products of *lim-4*. All *lim-4* alleles were outcrossed with N2 at least five times before phenotypic analysis. To observe locomotion phenotypes of *lim-4* mutants (*ky403, yn19, lsk3, lsk5*), the integrated *flp-12*p*::gfp* array was removed by mating with N2 males.

Molecular biology and transgenic worms

For promoter analysis, promoter regions of *odr-2* and *unc-17* were amplified by PCR from N2 genomic DNA and were inserted into the pPD95.77 vector [101]. *lim-4*p*::gfp* [102] was gifted from Piali Sengupta. Promoter regions of each reporter construct were deleted with various digestion enzymes or PCR fusions. Mutagenesis was performed using QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol.

For the constructs used to test for rescue, the *hsp16.2* promoter was fused with the following cDNAs: *lim-4* cDNA (kind gift form Oliver Hobert), human *lhx6* cDNA (BC103937) and human *lhx8* cDNA (BC040321) (Thermo Fisher Scientific). The *lim-4*pΔ*3* or *odr-1* promoter were used to generate the l*im-4*pΔ*3::lim-4cDNA* or *odr-1*p::*lim-4cDNA* constructs, respectively.

To generate the *flp-7*p-*SMBmotif::gfp* construct, three copies of SMB motif oligomers that have SacI enzyme site at 5' and 3' ends were synthesized and inserted into SacI site at -353 bp region of the *flp-7* promoter.

otIs518; otIs534 (cho-1^{fosmid}*::yfp; eat-4*^{fosmid}*::mChOpti)* and the *eat-4*pΔ*5::gfp* construct were kind gifts from Oliver Hobert. To express *lim-4* cDNA in the subset of glutamatergic neurons, *lim-4* cDNA was replaced with *gfp* to generate *eat-4*pΔ*5::lim-4cDNA.* Then, 2.5 ng of *eat-4*pΔ*5::lim-4cDNA* was injected into the *otIs518; otIs534* strain with 50 ng *rol-6* as an injection marker.

To express *lim-4* cDNA in the AWC neurons during development stage, *ceh-36*p was fused with *lim-4* cDNA and 5 ng was injected into *vsIs48* (*unc-17*p*::gfp*) and *ynIs82 (flp-12*p*::gfp)* with 50ng of *odr-1*p*::dsRed* as an injection marker. As the control, 50 ng *odr-1*p*::dsRed* was injected into *vsIs48* (*unc-17*p*::gfp*).

*ceh-17*p::*dsRed* was kindly gifted from Satoshi Suo.

Heat shock treatment and phenotype analysis

Heat shocks were administered to fourth larva stage (L4) of the transgenic worms at 33°C twice for 30 minutes with an hour incubation at 20°C between heat shocks for recovery modified from [71]. After heat shocks, worms were incubated at 20°C for 14 hours to reach the young adult stage when *lim-4* phenotypes were assayed. To observe ectopic expression of *flp-12*p*::gfp* in other cell types, heat shocks were administered at the embryo 2- or 3-fold stage

two times at 37°C for 30 minutes with an hour incubation at 20°C between heat shocks. Heat shocks were administered at L4 animals to observe locomotion at days1, 3, and 6. The L4 stage was counted as day 0.

The level of *flp-12*p*::gfp* expression in the SMB neurons was quantified as strong, weak, off. Strong was determined as robust expression in the SMB neuronal cell bodies and processes. Weak was defined as faint expression in cell bodies and no expression in the processes. Off was defined as no *flp-12*p*::gfp* expression in either cell bodies and processes. To measure wave width and wavelength of the worm tracks, Leica microscope software (Leica Application Suite Advanced Fluorescence Lite 3.5. Ink) was used. Wavelength was defined as distance between one peak and the next corresponding peak, and wave width was distance from the peak to the trough of the sine wave. The average of six consecutive wave width and wavelength from each worm track was quantified as the individual data.

Laser ablation

Laser ablation experiments were performed as previously described [103]. L1 larvae of the integrated *lim-4*p*::gfp* (*oyIs35*) strain were anesthetized with 10 mM sodium azide and all four SMB or SAA neurons were killed by a nitrogen dye-pulsed laser (Photonic Instruments, St. Charles, IL). Animals were recovered for 3 days at 20°C. After performing locomotion assays, GFP expression of *lim-4*p*::gfp* reporter was observed to confirm laser ablations in the SMB or SAA neurons.

Bioinformatics analysis

The conservation of the *cis*-regulatory motif from the promoter analysis of the *flp-12, lim-4, odr-2,* and *unc-17* were examined by using the USCS genome browser (http://genome.ucsc.edu/). DNA sequences from the five different *Caenorhabditis* species

were obtained from the UCSC website, and aligned using the ClustalW2 in EBI (European Bioinformatics Institutes; http://www.ebi.ac.uk/Tools/msa/clustalw2/) to identify *cis*-regulatory regions including SMB motifs. The position frequency matrix (PFM) of LIM-4, LHX6, and LHX8 predicted binding sites were derived from a web based tool, PreMoTF (http://stormo.wustl.edu/PreMoTF). Predicted conserved motif sequence logo was obtained from the Seq2Logo website (http://www.cbs.dk/biotools/Seq2Logo/).

Microscopy

Fluorescent microscopic images were taken with a Zeiss LSM700 Confocal microscope and were obtained using ZEN 2009 Light Edition software. For light microscopic images of worms, Leica High-performance Fluorescence Stereomicroscopy M205FA was used and Leica Application Suite Advanced Fluorescence Lite 3.5 software was used to measure the phenotype.

Table 3. Transgenes and strains used in this study

Transgene	Genotype	Strain
ace-2 p*::gfp*	*otEx4432*	OH10850
*acr-2*p*::gfp*	*juIs14*	OH10851
*acr-5*p*::gfp*	*wdEx75*	NC216
*acr-14*p*::gfp*	*wdEx455*	NC1013
*ceh-17*p*::dsRed*	*lskEx237*	KHK246
	lskEx238	KHK247
*ceh-24*p*::gfp*	*ccIs4595*	PD4595
*ceh-36*p*::lim-4cDNA*	*lskEx369*	KHK387
	lskEx397	KHK421
*cho-1*p*::gfp*	*otIs323*	OH10101
cho-1^{fosmid}*::yfp*	*otIs534*	
cho-1^{fosmid}*::yfp; eat-4*^{fosmid}*::mChOpti*	*otIs534; otIs518*	
*cog-1*p*::gfp*	*syIs63*	PS3662
*eat-4*pΔ*5::lim-4cDNA*	*lskEx374*	KHK392
	lskEx375	KHK393
	lskEx376	KHK394
*flp-7*p^{-2446}*::gfp*	*ynIs66*	NY2066
*flp-7*p^{-2446}*-SMBmotif*$^{-353}$*::gfp*	*lskEx153*	KHK160
	lskEx154	KHK161
*flp-12*p^{-2652}*::gfp*	*ynIs25*	NY2025
	ynIs82	NY2082
*flp-12*p^{-1547}*::gfp*	*lskEx3*	KHK3
	lskEx4	KHK4
*flp-12*p^{-523}*::gfp*	*lskEx1*	KHK1
	lskEx2	KHK2
*flp-12*p^{-523}*-MUT*$^{-501}$*::gfp*	*lskEx96*	KHK100
	lskEx97	KHK101
*flp-12*p^{-523}*-MUT*$^{-441}$*::gfp*	*lskEx98*	KHK102
	lskEx99	KHK103
*flp-12*p^{-523}*-MUT*$^{-429}$*::gfp*	*lskEx100*	KHK104
	lskEx101	KHK105
*flp-12*p^{-523}*-MUT*$^{-429}$*::gfp*	*lskEx102*	KHK106
	lskEx103	KHK107
*flp-12*p^{-523}*-MUT*$^{-501-441-429-392}$*::gfp*	*lskEx88*	KHK92
	lskEx89	KHK93
*flp-12*p^{-483}*::gfp*	*lskEx13*	KHK13
	lskEx14	KHK14
*flp-12*p^{-436}*::gfp*	*lskEx15*	KHK15
	lskEx16	KHK16
*flp-12*p^{-393}*::gfp*	*lskEx20*	KHK20

	lskEx21	KHK21
*flp-12*p^{-393}-*MUT*$^{-370-361}$*::gfp*	*lskEx104*	KHK108
	lskEx105	KHK109
*flp-12*p^{-393}-*MUT*$^{-377}$*::gfp*	*lskEx106*	KHK110
	lskEx107	KHK111
*flp-12*p^{-393}-*MUT*$^{-383}$*::gfp*	*lskEx108*	KHK112
	lskEx109	KHK113
*flp-12*p^{-393}-*MUT*$^{-363}$*::gfp*	*lskEx110*	KHK114
	lskEx111	KHK115
*flp-12*p^{-393}-*MUT*$^{-341}$*::gfp*	*lskEx115*	KHK162
	lskEx116	KHK163
*flp-12*p^{-339}*::gfp*	*lskEx22*	KHK22
	lskEx23	KHK23
*flp-12*p^{-312}*::gfp*	*lskEx90*	KHK94
	lskEx91	KHK95
*flp-12*p^{-312}-*MUT*$^{-190}$*::gfp*	*lskEx76*	KHK80
	lskEx77	KHK81
*flp-12*p^{-312}-*MUT*$^{-193}$*::gfp*	*lskEx78*	KHK82
*flp-12*p^{-312}-*MUT*$^{-167}$*::gfp*	*lskEx80*	KHK84
	lskEx81	KHK85
*flp-12*p^{-312}-*MUT*$^{-243}$*::gfp*	*lskEx82*	KHK86
	lskEx83	KHK87
*flp-12*p^{-312}-*MUT*$^{-232}$*::gfp*	*lskEx84*	KHK88
	lskEx85	KHK89
*flp-12*p^{-312}-*MUT*$^{-243-232}$*::gfp*	*lskEx88*	KHK92
	lskEx89	KHK93
*flp-12*p^{-162}*::gfp*	*lskEx92*	KHK96
	lskEx93	KHK97
*flp-12*p^{-34}*::gfp*	*lskEx94*	KHK98
	lskEx95	KHK99
	lskEx21	KHK21
*flp-22*p*::gfp*	*lskEx247*	KHK256
	lskEx248	KHK257
hsp16.2::human lhx6 cDNA	*lskEx181*	KHK188
	lskEx183	KHK190
	lskEx220	KHK229
	lskEx221	KHK230
hsp16.2::human lhx8 cDNA	*lskEx225*	KHK234
	lskEx226	KHK236
	lskEx267	KHK276
	lskEx268	KHK277
hsp16.2::lim-4 cDNA	*lskEx51*	KHK52
	lskEx52	KHK53

lim-4	*ky403*	PY7485
	lsk1	KHK500
	lsk2	KHK501
	lsk3	KHK502
	lsk4	KHK503
	lsk5	KHK504
	lsk6	KHK505
	lsk7	KHK506
	yn19	PY3649
*lim-4*p^{-3583}*::gfp*	*oyIs35*	PY1958
	lskEx265	KHK274
	lskEx266	KHK275
*lim-4*p$^{-3379-1923}$*::gfp*	*lskEx32*	KHK32
	lskEx33	KHK33
*lim-4*p$^{-3379-1525}$*::gfp*	*lskEx59*	KHK59
	lskEx60	KHK60
*lim-4*p$^{-3379-1525}$*-MUT^{-64}::gfp*	*lskEx272*	KHK281
*lim-4*p$^{-3379-999}$*::gfp* (*lim-4*pΔ*3::gfp*)	*lskEx54*	KHK54
	lskEx55	KHK55
*lim-4*p$^{-3379-999}$*::lim-4cDNA* (*lim-4*pΔ*3::lim-4cDNA*)	*lskEx222*	KHK231
	lskEX223	KHK232
	lskEx224	KHK233
	lskEx251	KHK260
*lim-4*p$^{-3379-999}$*::mCherry* (*lim-4*pΔ*3::mCherry*)	*lskEx172*	KHK179
*lim-4*p$^{-3477-461}$*::gfp*	*lskEx11*	KHK11
	lskEx12	KHK12
*lim-4*p$^{-2288-735}$*::gfp*	*lskEx7*	KHK7
	lskEx8	KHK8
*lim-4*p$^{-1793-275}$*::gfp*	*lskEx38*	KHK38
	lskEx39	KHK39
*lim-4*p$^{-1793-247}$*::gfp*	*lskEx41*	KHK41
	lskEx42	KHK42
*lim-4*p$^{-1793-38}$*::gfp*	*lskEx46*	KHK46
	lskEx47	KHK47
*lim-4*p$^{-999-38}$*::gfp*	*lskEx128*	KHK132
	lskEx73	KHK77
*lim-4*p$^{-999-247}$*::gfp*	*lskEx133*	KHK138
	lskEx134	KHK139
*lim-4*p^{-3327}*::gfp*	*lskEx137*	KHK142
	lskEx138	KHK145
*lim-4*p^{-1923}*::gfp*	*lskEx140*	KHK148
	lskEx141	KHK149

*lim-4*p$^{-3379\text{-}1922}$*::gfp*	*lskEx145*	KHK152
	lskEx146	KHK153
*lim-4*p^{-82}*::gfp*	*lskEx122*	KHK126
	lskEx123	KHK127
	lskEx124	KHK128
	lskEx125	KHK129
*lim-4*p^{-82}*-MUT^{-53}::gfp*	*lskEx271*	KHK280
*lim-4*p^{-82}*-MUT^{-64}::gfp*	*lskEx143*	KHK150
	lskEx144	KHK151
*lim-4*p^{-38}*::gfp*	*lskEx270*	KHK279
*lim-4*p^{-999}*::gfp*	*lskEx191*	KHK201
*odr-1*p*::dsRed*	*lskEx370*	KHK388
	lskEx372	KHK390
	lskEx373	KHK391
*odr-1*p*::lim-4cDNA*	*lskEx138*	KHK143
	lskEx144	KHK144
*odr-2*p^{-2443}*::gfp*	*lskEx114*	KHK118
	lskEx115	KHK119
*odr-2*p^{-607}*::gfp*	*lskEx116*	KHK120
	lskEx117	KHK121
*odr-2*p^{-607}*-MUT^{-570}::gfp*	*lskEx162*	KHK169
	lskEx163	KHK170
*odr-2*p^{-607}*-MUT^{-548}::gfp*	*lskEx164*	KHK171
	lskEx165	KHK172
*odr-2*p^{-527}*::gfp*	*lskEx158*	KHK165
	lskEx159	KHK166
*odr-2*p^{-377}*::gfp*	*lskEx160*	KHK167
	lskEx161	KHK168
*odr-2*p^{-300}*::gfp*	*lskEx118*	KHK122
	lskEx119	KHK123
*rgef-1*p*::gfp*	*evIs111*	NW1229
*trp-1*p*::gfp*	*kyIs123*	OH1358
*unc-17*p^{-4410}*::gfp*	*vsIs48*	LX929
	lskEx199	KHK208
	lskEx212	KHK221
*unc-17*p^{-4410}*-MUT^{-4270}::gfp*	*lskEx273*	KHK282
*unc-17*p^{-4175}*::gfp*	*lskEx249*	KHK258
*unc-17*p^{-2820}*::gfp*	*lskEx215*	KHK224
*unc-17*p^{-2242}*::gfp*	*lskEx217*	KHK226
*unc-42*p*::gfp*	*gmEx104*	NG2591
*unc-119*p*::gfp*	*otIs45*	OH441

III. Results

3.1 Expression of a neuropeptide gene in the SMB neurons is abolished in *lim-4* mutants

To identify factors that specify the neuronal cell-fate of SMB, I performed a genetic screen to isolate animals in which the expression pattern of a terminal differentiation marker, *flp-12*p*::gfp* reporter, was disrupted exclusively in the SMB neurons. *flp-12* encodes a FMR-Famide-related neuropeptide and is expressed in a set of neurons that includes the SMB and SDQ neurons in adults [100]. Among mutants isolated from this screen, seven mutant alleles (named as *lsk1,2,4,5,6,7, yn19*) exhibited complete loss of *flp-12* expression, while one mutant allele (*lsk3*) showed weak expression of *flp-12* in all four SMB neurons at either adult (Figure 7 and 8; Table 4) or L1 larval developmental stage (Figure 7; Table 4 and 5) animals. By contrast, expression of *flp-12* in the SDQ and other neurons weakly expressing *flp-12* was unaffected in all eight mutants (Figure 7 and 8), indicating that expression of *flp-12* was specifically affected in the SMB neurons of these mutants.

From subsequent complementation test and three factor analysis, all mutants were found to be allelic to the previously identified *lim-4(ky403)* mutants. *lim-4* encodes a LIM homeodomain protein that is required for specification of AWB and ADF chemosensory neuron identity [102, 104]. In *lim-4*(*ky403*) null mutants, AWB cell fate is changed to that of the AWC chemosensory neurons, thereby causing dye-filling defects in the AWB neurons [102]. Expression of *flp-12* was also completely abolished in the SMB neurons of *ky403* mutants (Figure 7; Table 4). Like in the *ky403* mutants, the AWB neurons failed to dye-fill in the *lsk1-7* and *yn19* mutants (Figure 9; Table 6). The molecular lesions of all eight mutants mapped to the coding region of the *lim-4* gene (Figure 10). Five mutant alleles (*lsk1,4,5 6,7*) had nonsense mutations that resulted in premature translation stop, suggesting that these mutations are null alleles. *lsk2* had a mutation in the splice donor site after the 1st exon. *yn19* and *lsk3*

had missense mutations within the coding region of the second LIM domain, resulting in C199Y and E207K substitutions, respectively. The cysteine residue (C199) is critical for forming a zinc finger motif in the LIM domain [104]. The glutamate residue (E207) resides in the LIM domain and is highly conserved through evolution (Figure 10), suggesting that this residue is essential for LIM-4 function via protein-protein interactions. These findings indicate that LIM-4 has a role in regulating gene expression in the SMB neurons.

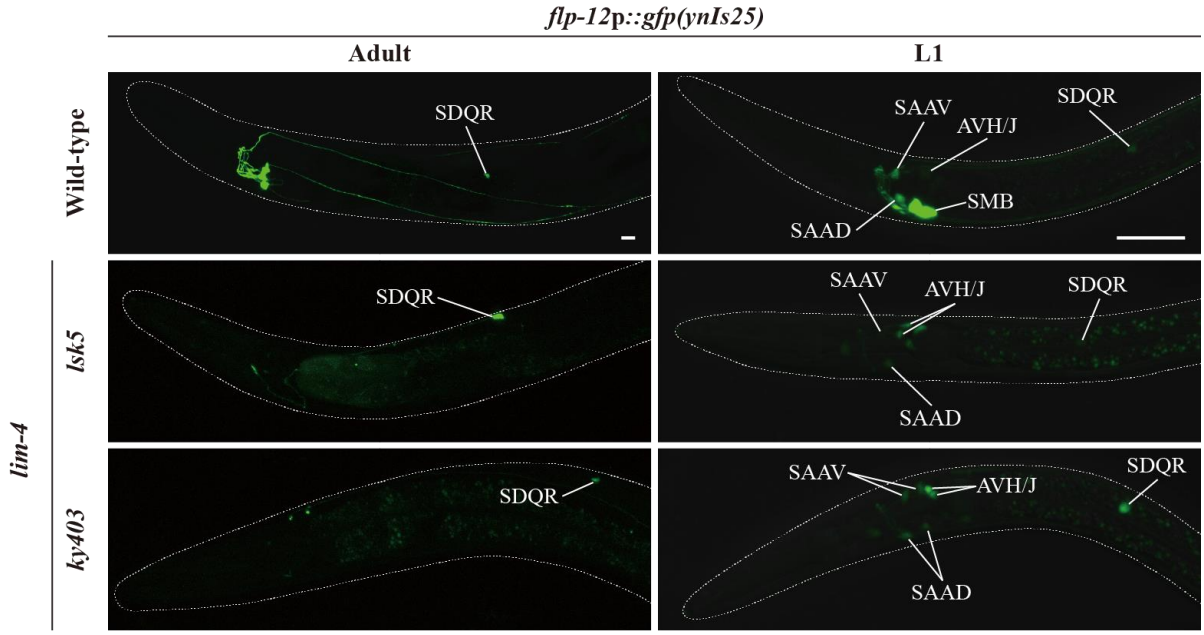
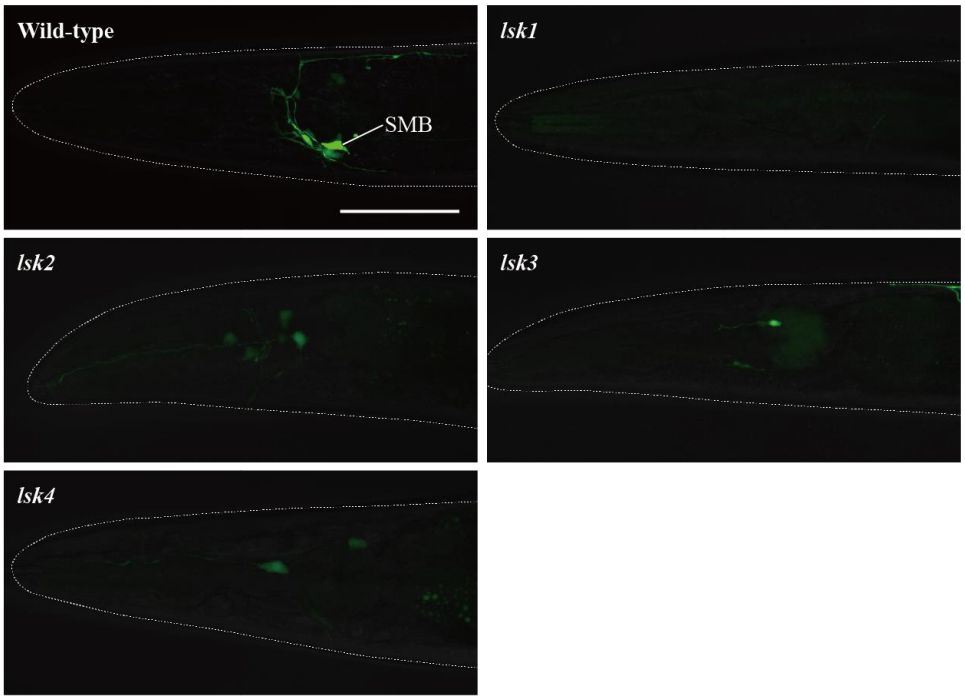


Figure 7. Expression of a *flp-12* neuropeptide reporter is abolished in the SMB neurons of *lim-4* mutants. Expression of the *flp-12* neuropeptide reporter is abolished specifically in the SMB neurons of *lim-4* mutants. GFP expression in *ynIs25* integrated strains is observed in the SMB and SDQ (L/R) neurons of wild-type adult (left column) or L1 larval (right column) stage animals while GFP expression in the SMB neurons is not detected in *lim-4* null mutant alleles (*lsk5*, *ky403*). Faint expression of the *flp-12*p*::gfp* reporter in the SAA and AVH/J neurons of L1 larvae is not altered in *lim-4* mutants. Anterior is at left in all images. Scale bars: 20 μm.

*flp-12*p::*gfp(ynIs82)*



*flp-12*p::*gfp(ynIs25)*

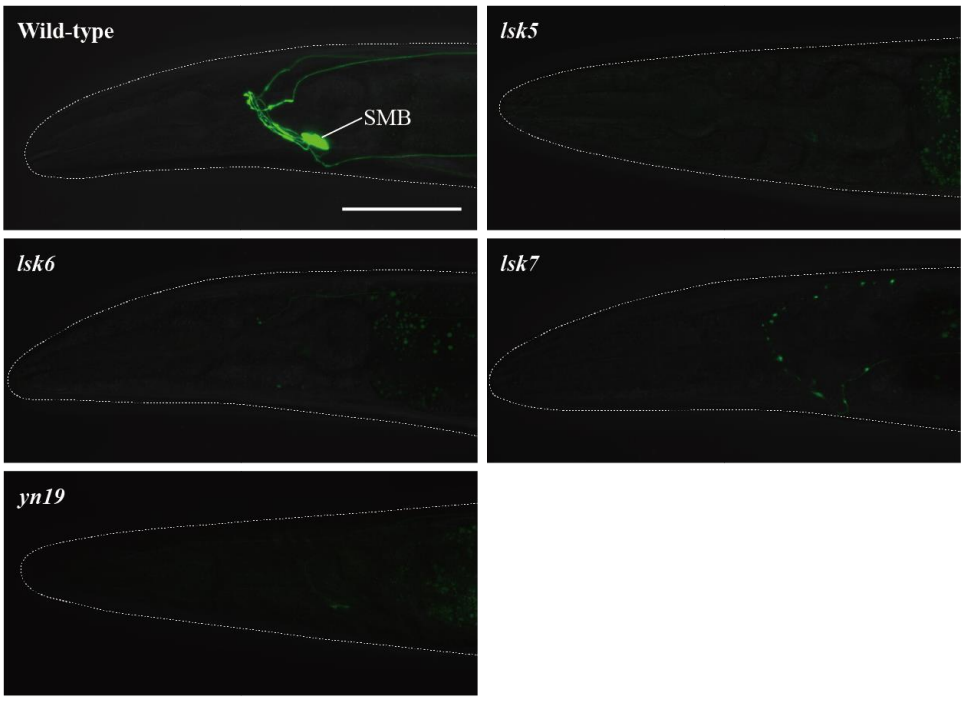


Figure 8. Expression of a *flp-12*p*::gfp* reporter is abolished in adult *lim-4* mutants. Two integrated strains expressing a *flp-12*p*::gfp* reporter (*ynIs82* or *ynIs25*) were used for EMS mutagenesis screens. Note that *ynIs82* integrated strains exhibit variable and weak GFP expression in a few neurons of the head. Anterior is at left in all images. Scale bars: 50 μm.

Table 4. Expression pattern of the SMB expressed markers in *lim-4* mutant animals

	Reporter construct	Gene product	Geno-type	% animals showing GFP expression in SMB		
				Off	Weak	Strong
SMB-specific	*flp-12*p*::gfp* (*ynIs25*)	FLP-12 Neuropep-tide	WT	0	0	100
			ky403	100	0	0
			yn19	100	0	0
			lsk3	77	20	3
			lsk5	100	0	0
			WT (L1)	0	0	100
			ky403 (L1)	100	0	0
	*odr-2*p*::gfp*	Ly-6 superfamily	WT	0	0	100
			ky403	100	0	0
	*trp-1*p*::gfp*	TRP channel fami-ly	WT	0	100	0
			ky403	100	0	0
	*lim-4*p*::gfp*	LIM homeodomain protein	WT	0	0	100
			lsk5	100	0	0
			WT (L1)	0	0	100
			lsk5 (L1)	0	6	92
Choliner-gic	*unc-17*p*::gfp*	vesicular acetyl-choline transporter	WT	0	0	100
			ky403	100	0	0
	*cho-1*p*::gfp*	choline transporter	WT	0	100	0
			ky403	93	7	0
Pan-neuronal	*rgef-1*p*::gfp*	Rap GRP	WT	0	0	100
			ky403	0	0	100
	*unc-119*p*::gfp*	chaperone	WT	0	80	20
			ky403	0	73	27

Expression pattern of GFP reporter constructs in the SMB neurons was observed in wide type (WT) or *lim-4 (ky403, yn19, lsk3, lsk5)* mutant animals at L1 larval stage (L1) or adult stage. Expression was observed at 630x or 400x magnification for L1 or young adults, respectively. Strong, weak or off expression is defined as GFP expression observed in both cell bodies and processes, observed in only cell bodies, or not observed either in cell bodies or processes, respectively. $n \geq 30$.

Table 5. Expression pattern of the *flp-12* in *lim-4* mutant animals at L1 stage

Reporter construct	Genotype	% animals showing GFP expression in SMB (L1)		
		No	Weak	Strong
*flp-12*p*::gfp (ynIs82)*	WT	0	0	100
	lsk3	20	70	10
*flp-12*p*::gfp (ynIs25)*	WT	0	0	100
	yn19	100	0	0
	lsk5	100	0	0

Expression pattern of *flp-12*p*::gfp* reporter constructs in the SMB neurons was observed in wild type (WT) or *lim-4* (*lsk3, lsk5, yn19*) mutant animals at L1 larval stage (L1). Expression was observed at 630x. Strong, weak or no expression is defined as GFP expression observed in both cell bodies and processes, observed in only cell bodies, or not observed either in cell bodies or processes, respectively. n≥50.

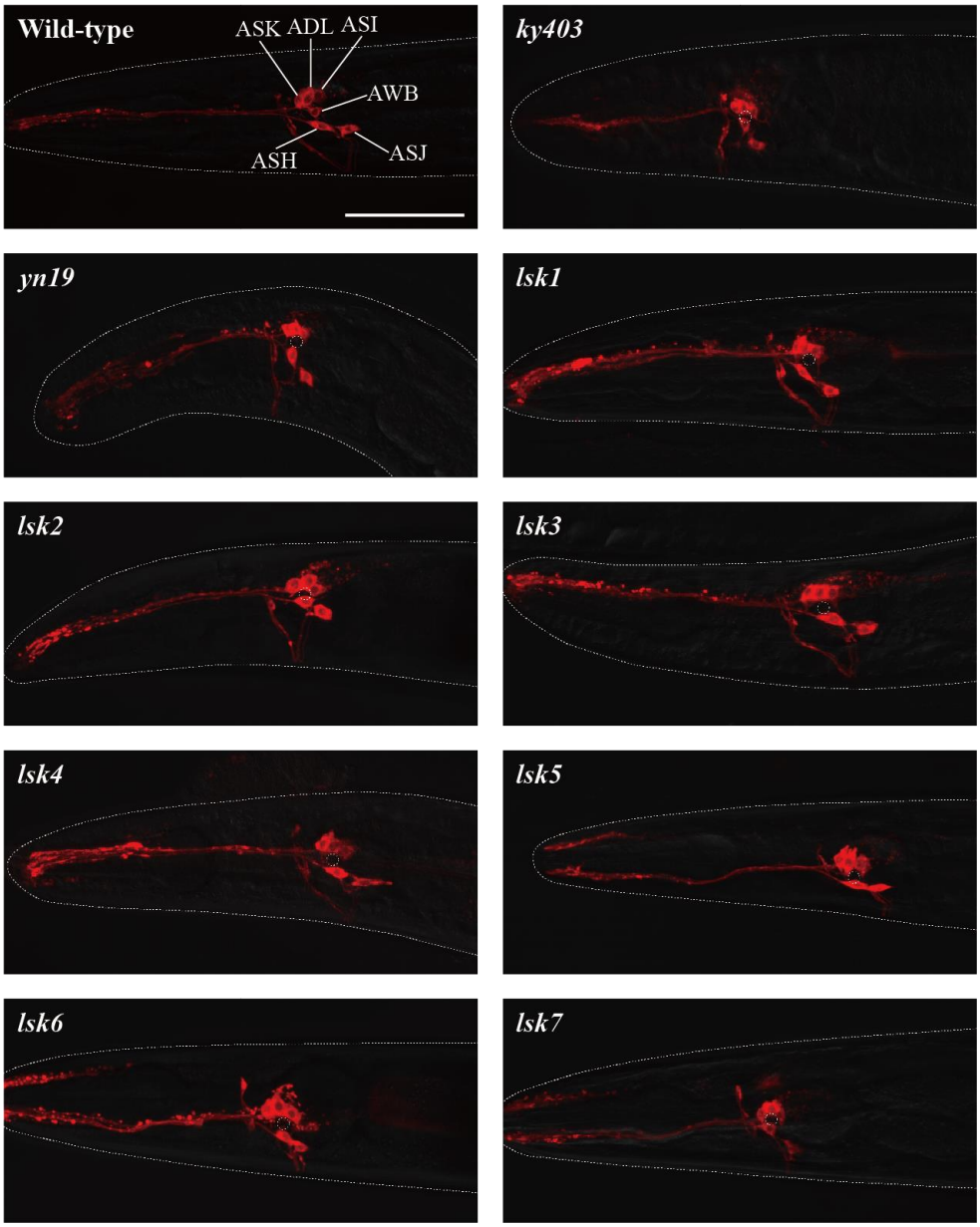


Figure 9. The AWB neurons fail to be dye-filled in *lim-4* mutants. The AWB neurons and additional five pairs of amphid neurons (ASK, ADL, ASI, ASH, ASJ) in the head of wild-type animals fill with lipophilic dye DiD [105]. In *ky403* and eight newly identified *lim-4* alleles, the AWB neurons fail to be dye-filled. Images are derived from z-stacks of confocal microscopy images taken for left-side amphid neurons. Dashed circles indicate position of the AWB neurons. Anterior is at left in all images. Quantitative analysis of these phenotypes is shown in S2 Table. Scale bar: 50 μm.

Table 6. Dye-filling defects in the AWB neurons of *lim-4* mutant animals

Genotype	% animals showing dye-filling in AWB/ADF
WT	100
ky403	23
yn19	3
lsk3	0
lsk5	57

Lipophilic dye DiD was used to observe dye-filling in the AWB neurons of *lim-4* mutant animals (*ky403, yn19, lsk3, lsk5*). In *lim-4* (*ky403, lsk5)* null mutants, the AWB neurons are not dye-filled, while the ADF neurons are ectopically dye-filled [106]. Dye staining in adult animals was observed at 400x. n≥30.

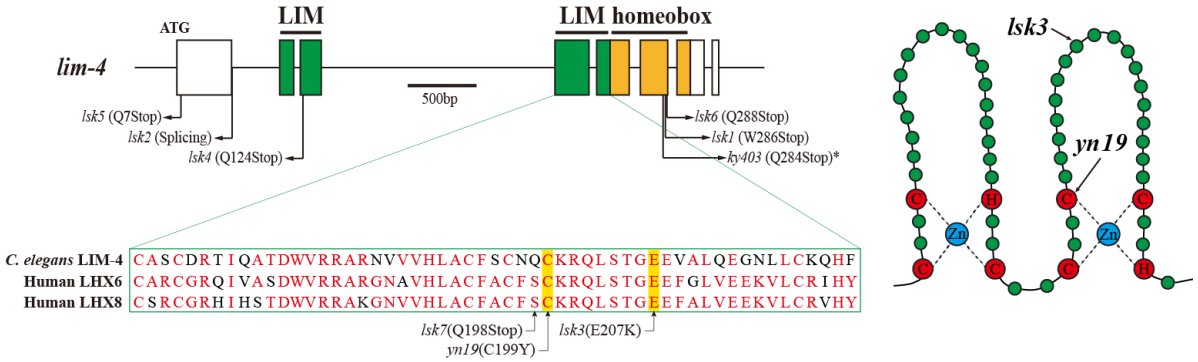


Figure 10. Genomic structure of *lim-4* (left) and schematic structure of the LIM domain of LIM-4 (right). The sequence alignment of the part of LIM domain of *C. elegans* LIM-4 and human LHX6 and LHX8 is shown. Identical residues in at least two proteins are shown in red. Molecular lesions of *lim-4* mutant alleles are indicated. *Mutation in *ky403* was previously reported [102]. LIM domain and homeodomain are labeled in green and in yellow, respectively.

3.2 Expression of terminally differentiated SMB markers including cholinergic genes is abolished in *lim-4* mutants

To determine the extent to which LIM-4 regulates gene expression in the SMB neurons, I examined additional SMB terminal differentiation genes, including *odr-2* GPI-anchored cell surface protein [107], *trp-1* TRPC channel [108], and cholinergic markers such as *unc-17* vesicular acetylcholine transporter (VAChT) [109] and *cho-1* choline transporter (ChT) (Figure 11) [110]. In order to locate the SMB cell bodies, I used expression of *ceh-17*p*::dsRed* in the cell bodies of SIAV as a marker that is directly adjacent to the cell bodies of SMBD (Figure 12) [111]. None of the SMB specific or cholinergic markers were expressed in the SMB neurons of *lim-4* mutants while expression in other neuron types was generally not affected (Figure 13; Table 4). Next, I tested expression of two well-characterized pan-neuronal gene markers, *rgef-1* Ras guanine nucleotide releasing protein and *unc-119* chaperone [71, 80]. Expression of these pan-neuronal genes was not altered in the SMB neurons of *lim-4* mutants (Figure 13; Table 4), indicating that the SMB cells may retain neuronal properties.

To determine whether the *lim-4* SMB neurons adopted different cell fate such as the structurally and/or functionally related sub-lateral nerve cord neurons including the SIA, SIB, and SMD neurons, I tested markers including *ceh-17* for SIA [111], *flp-22* for SMD [100] or *ceh-24* for SIA, SIB, and SMD [68, 112]. None of these markers were ectopically expressed in *lim-4* mutants (Figure 14), suggesting that the cell fate of the SMB neurons is not transformed to that of the structurally and/or functionally related cell types.

The SMB neurons are generated from ABalpapap (SMBDL, SMBVL) or ABarappap (SMBDR, SMBVR) precursors, and three of their sister cells undergo programmed cell death before hatching (Figure 15) [22]. I observed expression of pan-neuronal markers in SMB of adult *lim-4* mutants (Figure 13), suggesting that the SMB cells do not adopt the apoptotic fate of their sister cells. Based on these results, I concluded that LIM-4 activity does not initiate

neuronal cell fate, but specifies SMB cell fate by regulating expression of terminal differentiation genes, thereby acting as a terminal selector transcription factor in the SMB neurons.

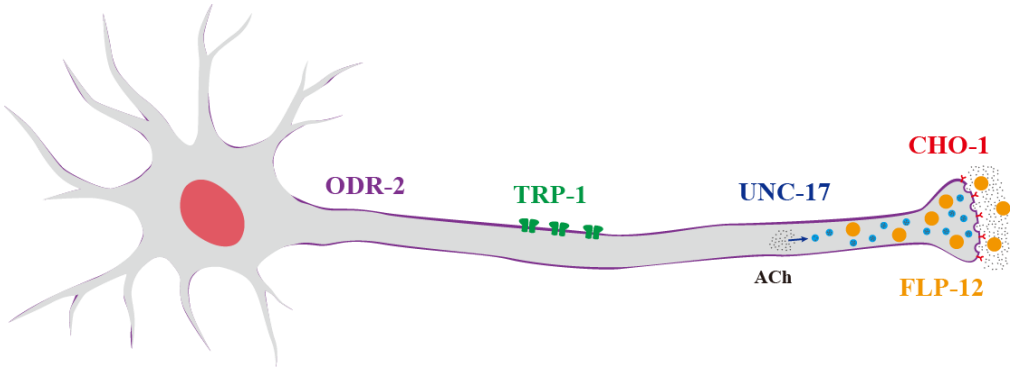


Figure 11. Schematic drawing of expressed genes in the SMB neurons; *odr-2* (GPI-anchored cell surface protein), *trp-1* (TRPC channel), *unc-17* (VAChT), *cho-1* (ChT), and *flp-12* (neuropeptide)

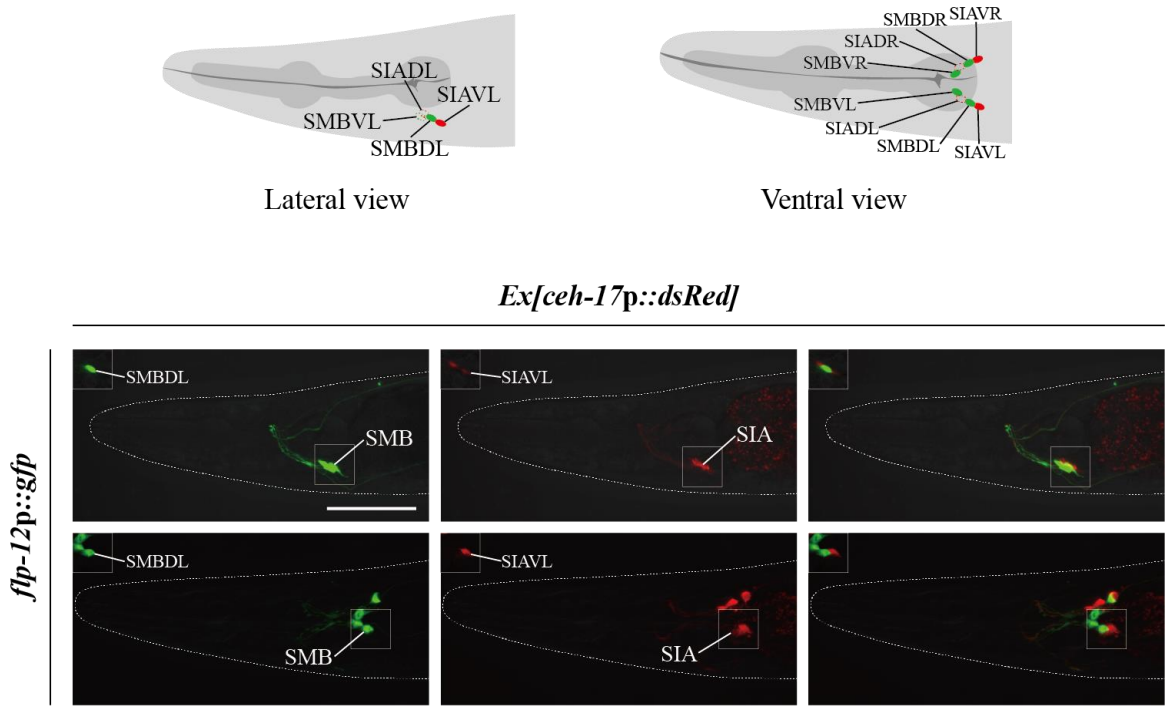


Figure 12. Identification of the SMB cell bodies. Shown are the images of expression of *flp-12*p*::gfp* (SMB) or *ceh-17*p*::dsRed* (SIA), the merged images, or schematic drawing of cell bodies in the merged images. Top images are from lateral view and bottom images are from ventral view of head of worms. Although positions of the cell bodies of SMBs or SIAs are variable, I consistently observed that the cell bodies of SMBDL/R and SIAVL/R are located immediately adjacent to each other at the same focal plane in all tested animals (n>30). Therefore, I identified SMBDL/R cell bodies via Nomarski optics by comparing with expression of *ceh-17*p*::dsRed* in the cell body of SIAVL/R. Images are derived from z-stacks of confocal microscopy images while images in the upper-left boxed regions are single focal plane confocal microscopy images. Anterior is at left in all images. Scale bar: 50 μm.

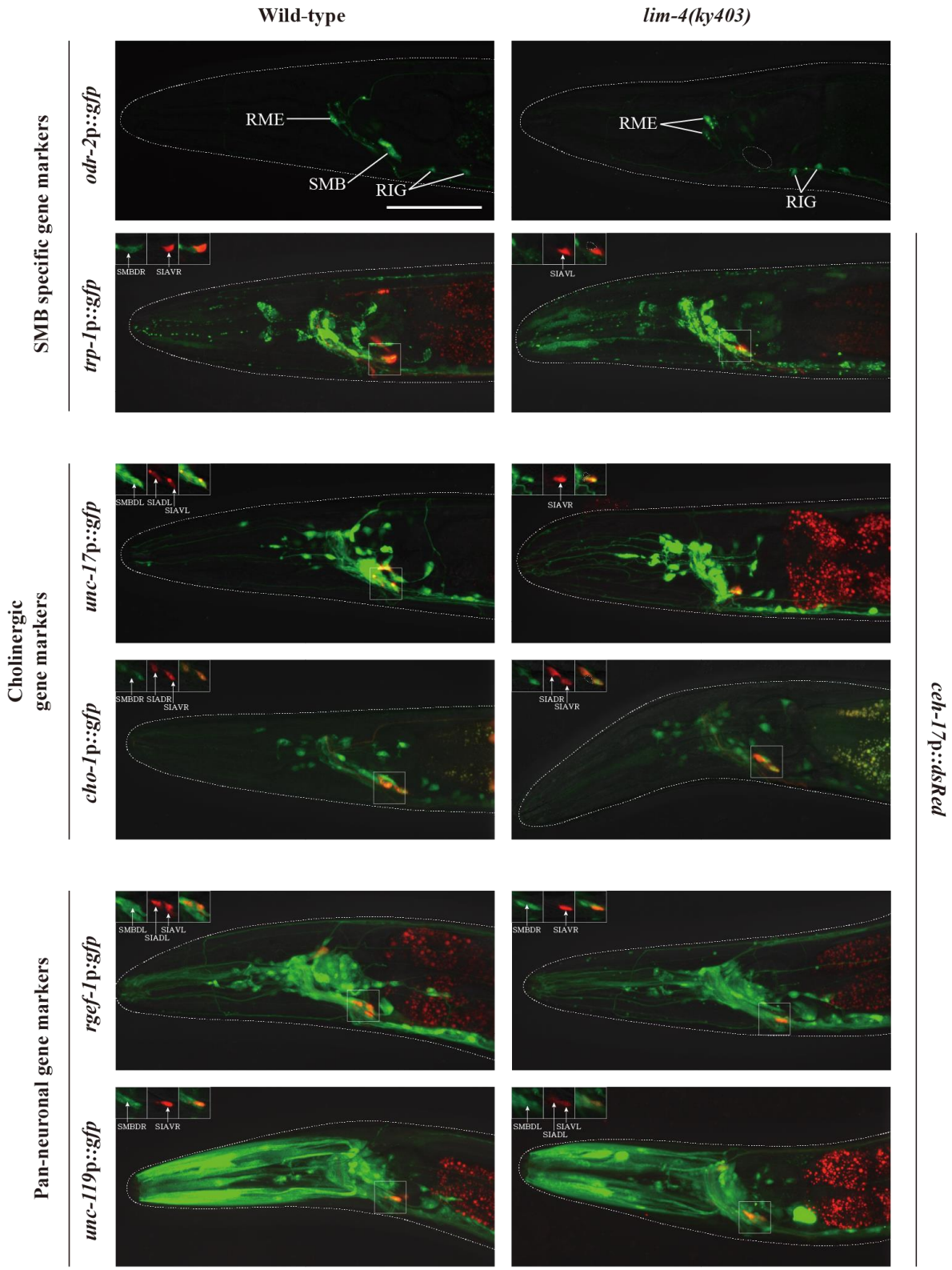

Wild-type
lim-4(ky403)
SMB specific gene markers
odr-2p::gfp
RME
SMB
RIG
RME
RIG
trp-1p::gfp
SMBDR SIAVR
SIAVL
Cholinergic gene markers
unc-17p::gfp
SMBDL SIADL SIAVL
SIAVR
cho-1p::gfp
SMBDR SIADR SIAVR
SIADR SIAVR
ceh-17p::dsRed
Pan-neuronal gene markers
rgef-1p:gfp
SMBDL SIADL SIAVL
SMBDR SIAVR
unc-119p::gfp
SMBDR SIAVR
SMBDL SIADL SIAVL

Figure 13. LIM-4 regulates expression of the terminally differentiated markers in the SMB neurons. Expression of the indicated reporter constructs is shown in wild-type (left column) or *lim-4(ky403)* mutant (right column) animals. Merged images with the *ceh-17*p*::mCherry* reporter expression in the SIA neurons (shown in red) were shown for *trp-1*, *unc-17*, *cho-1*, *rgef-1*, or *unc-119* promoter reporter for help in identification of the SMB neurons (see Figure 12). Images are derived from z-stacks of confocal microscopy images while images in the upper-left boxed regions are single focal plane confocal microscopy images. Quantitative analysis of these phenotypes is shown in Table 1. Anterior is to the left. Scale bar: 50 μm.

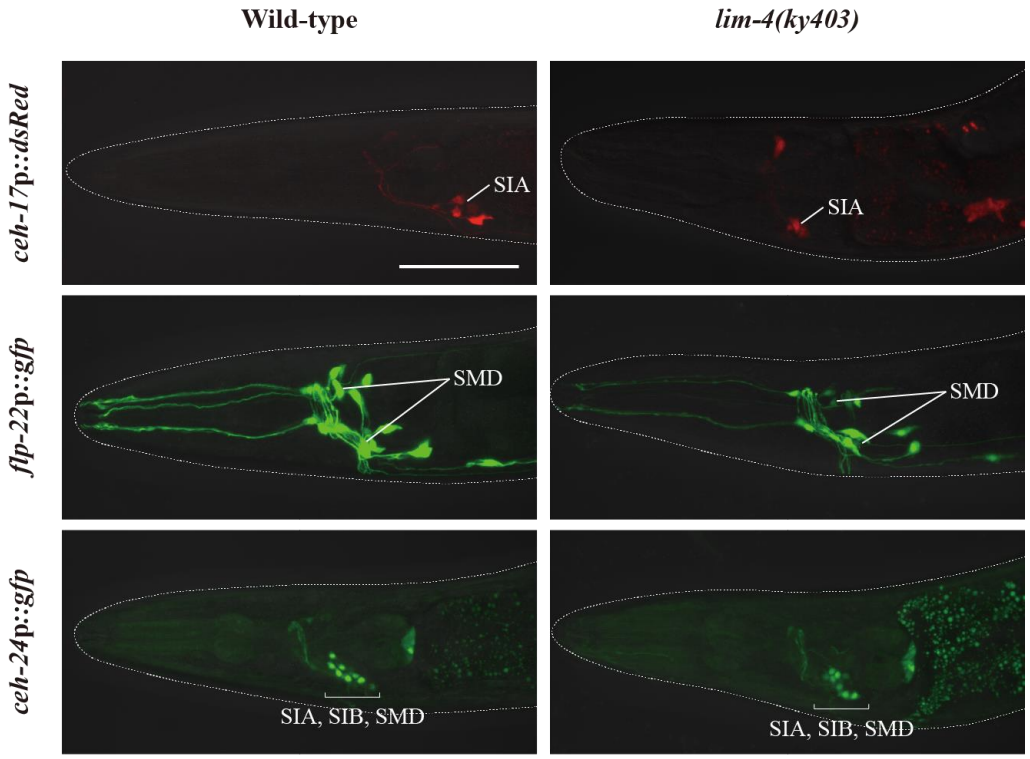


Figure 14. The SMB neurons in *lim-4* mutants do not adopt the cell-fate of the SIA, SIB, or SMD neurons. Expression of the indicated reporter constructs is shown in wild-type (left column) or *lim-4(ky403)* mutant (right column) animals. Expression of the *ceh-24*p*::gfp* reporter is detected in about 8 cells bodies in the head of worms including the SIA, SIB, and SMD neurons but not the SMB neurons. Images are derived from z-stacks of confocal microscopy images. Anterior is to the left. Scale bar: 50 μm.

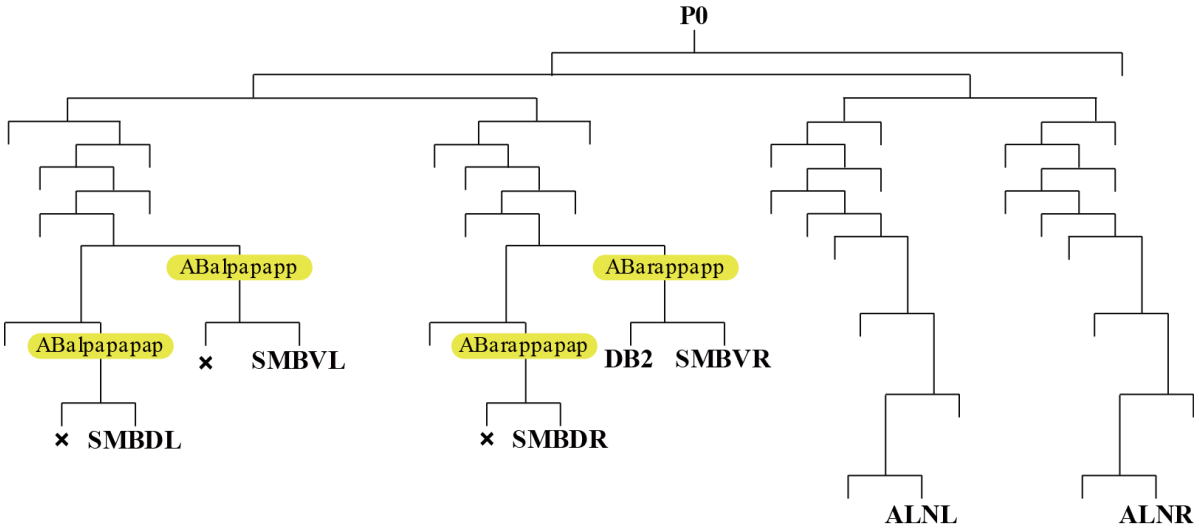


Figure 15. The cell lineage of the SMB neurons and ALN neurons in wild-type animals. Precursor cells of the SMB neurons are shown in yellow circles. x indicates the programmed cell death. Heat shocks to transgenic animals expressing *hsp::lim-4 cDNA* transgene induced *flp-12* expression in the ALN neurons (see Figure 24).

3.3 The function of the SMB neurons is compromised in *lim-4* mutants

Wild-type animals move in sinusoidal waves of a consistent wave width and wavelength (Figure 16) [70]. *lim-4* mutants move in a coiled or loopy fashion (Figure 16) [102]. To quantitate the loopy uncoordinated movement, the waveforms of these animals were measured by viewing tracks made in a bacterial lawn and compared to that of wild-type animals (Figure 17A). *lim-4* mutants had significantly accentuated waveforms (Figure 17B). While the average wavelength for *lim-4* null mutants (*ky403* or *lsk5*) is similar or mildly decreased compared to that of wild-type animals, the average wave width for *ky403* or *lsk5* mutants (*ky403*: 359.46$\pm$13.51 µm, n=30; *lsk5*: 374.27$\pm$16.47 µm, n=30) is about 70% higher than that of wild-type animals (N2: 194.54$\pm$4.28 µm, n=30) (Figures 16 and 17B). *yn19* and *lsk3* missense mutants similarly exhibited significantly larger wave width (*yn19*: 312.31$\pm$7.75 µm, n=30; *lsk3*: 365.43$\pm$11.97 µm, n=30) (Figures 16 and 17B), suggesting that *yn19* and *lsk3* mutations also fully eliminate the contribution of LIM-4 to locomotion.

To assess whether the loopy movement of *lim-4* mutants is due to a functional defect of the SMB neurons, I ablated the SMB neurons by laser microsurgery. Consistent with a previous study [50], killing the SMB neurons resulted in a loopy or coiled movement phenotype; the average wave width was increased by over 50% compared to control animals and similar to that of *lim-4* mutants (Figure 17C). *lim-4* is also expressed in the SAA neurons (see below) of which roles have been implicated in head locomotion [6]. Laser ablation of the SAA neurons did not cause a loopy or coiled movement, ruling out the possibility that defects of the SAA neurons result in movement defects of *lim-4* mutants (Figure 17C). These results indicate that the SMB neurons function to regulate locomotion by modulating the wave width of the animal and that the loopy phenotype of *lim-4* mutants is due to defects in the function of the SMB neurons.

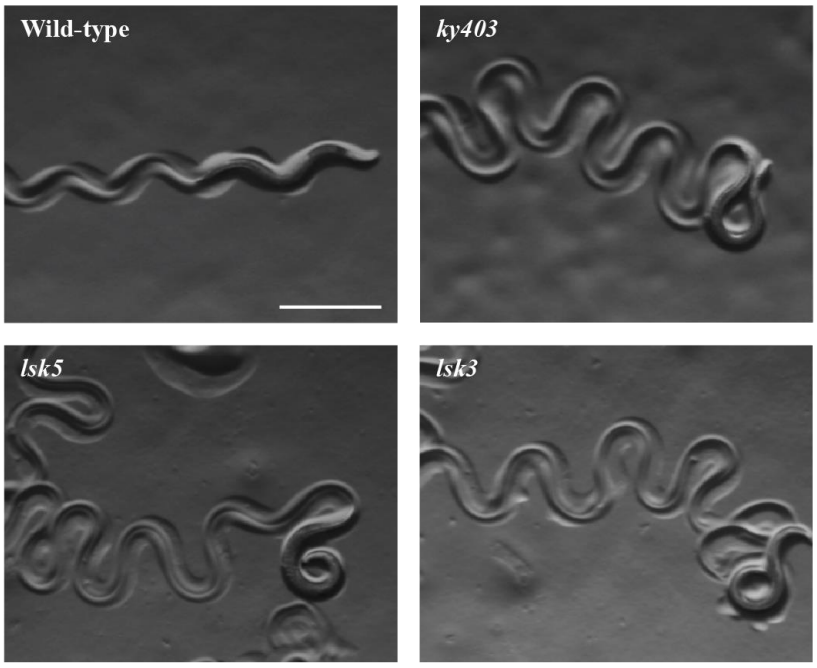


Figure 16. *lim-4* mutant animals moved in a coiled or loopy fashion due to the functional defects of the SMB neurons. Wild-type animals show a characteristic sinusoidal waveform whose tracks can also be observed in the bacterial lawn. *lim-4* (*ky403*, *lsk5*, *lsk3*) mutant animals showed an exaggerated waveform characterized by an increased wave width. Scale bar: 0.5 mm

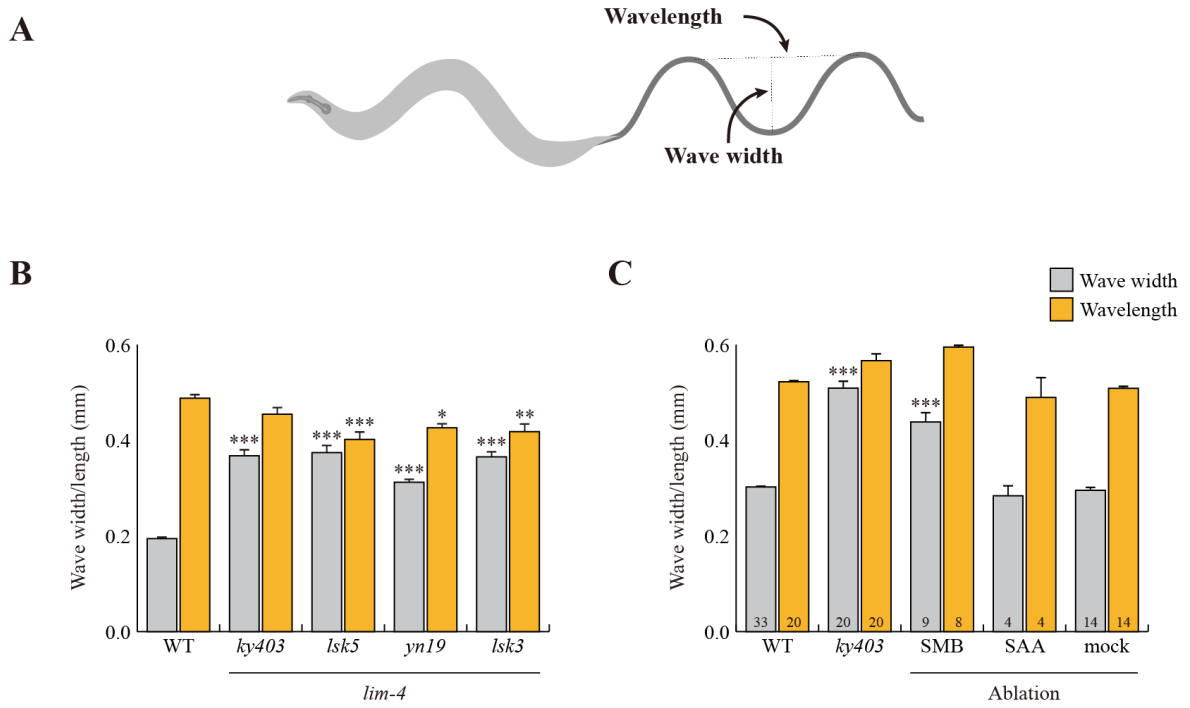


Figure 17. *lim-4* mutant and SMB ablated animals exhibited increased wave width but not wave length of sinusoidal waveform. (A) To quantitate locomotion, the waveforms of different animals were analyzed and compared to that of wild-type animals by viewing tracks made in a bacterial lawn under a microscope. Wave width and wavelength were measured and averaged as the distance from the peak to the trough of the sine wave and distance between one peak and the next corresponding peak, respectively. (B-C) Average of wave width and wavelength of the indicated genotypes. n≥30 for each. Error bars are the SEM. *, ** and *** indicate significantly different from wild-type at $p<0.05$, 0.01, and 0.001, respectively (one-way ANOVA test followed by the Tukey post-hoc test). The number of animals tested is indicated on the bars.

3.4 *lim-4* is expressed and functions in the SMB neurons to regulate their terminal specification

lim-4 has previously been shown to be expressed in several neuronal types in the head of postembryonic animals; these neurons include the AWB, SIA, SAA, RID, RIV, and RMD neurons, but not the SMB neurons [102]. Like the SMB neurons, the SIA neurons project their processes into the sub-lateral nerve cords and their cell morphology and position are similar to those of the SMB neurons [6]. To determine whether *lim-4* expression was mis-identified in the SIA neurons, I examined the expression pattern of *lim-4*p*::gfp* transgene (*oyIs35*) that includes 3.6 kb of upstream sequence [102, 113] and compared it to that of *ceh-17*p*::dsRed*, a SIA marker [111] or *flp-12*p*::mCherry*, a SMB marker [100], respectively (Figure 18). I observed co-localization of *lim-4* expression with that of *flp-12* but not of *ceh-17*, indicating that *lim-4* is expressed in the SMB neurons rather than the SIA neurons. In support of this re-assignment, the expression of other SMB markers such as *odr-2*p*::gfp* or *trp-1*p*::gfp* was completely abolished in the SMB neurons of *lim-4* mutants, whereas expression of *ceh-17* was not affected (Figures 13 and 14).

Expression of *lim-4* was previously shown to be autoregulated in the AWB neurons but not in the other LIM-4-expressing neurons [102]. I confirmed that *lim-4* expression in the AWB neurons was not seen in *lim-4* mutants at L1 larval stage, whereas the expression of *lim-4* in the other neurons including the SMB neurons, was detected (Figure 19). However, *lim-4* expression in the SMB neurons gradually decreased from the L1 larval stage until it became undetectable in the adult stage (Figure 19; Table 4). Hence, *lim-4* appears to be required to maintain its own expression in the SMB neurons but does not initiate its expression, further supporting its role as a terminal selector gene in the SMB neurons.

To determine whether *lim-4* acts cell-autonomously within SMB, I tried to rescue *lim-4* phenotypes by expressing a wild-type *lim-4* cDNA driven under the control of *lim-4*p*Δ3* pro-

moter. The *lim-4*p*Δ3* promoter includes minimal upstream regulatory sequences that drive transgene expression exclusively in the SMB neurons but not as strongly as the full promoter of *lim-4* and more dominantly in the SMBD than SMBV neurons (see below), and was used to identify expression in SMB of genes tested in this study (Figure 20). The gene expression and locomotion defects of *lim-4* mutants were partially restored, while the dye-filling defects were still present and the normal average wavelength was not altered, indicating that LIM-4 acts in the SMB neurons to affect locomotion and transmitter specification (Figures 21A and 21B; Table 7). Taken together, *lim-4* is expressed and acts in the SMB neurons to specify the SMB cell-fate.

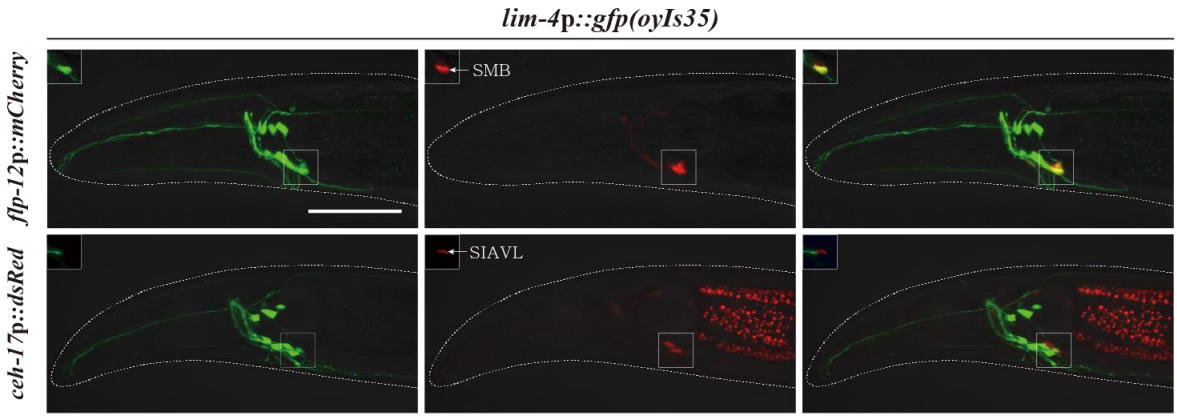


Figure 18. *lim-4* is expressed in SMB neurons. GFP expression of *oyIs35* animals carrying an integrated *lim-4*p*::gfp* reporter is overlapped with expression of the *flp-12*p*::mCherry* reporter (the SMB marker) but not with expression of the *ceh-17*p*::dsRed* reporter (the SIA marker). Images are derived from z-stacks of confocal microscopy images while images in the upper-left boxed regions are single focal plane confocal microscopy images. Anterior is to the left. Scale bar: 50 μm.

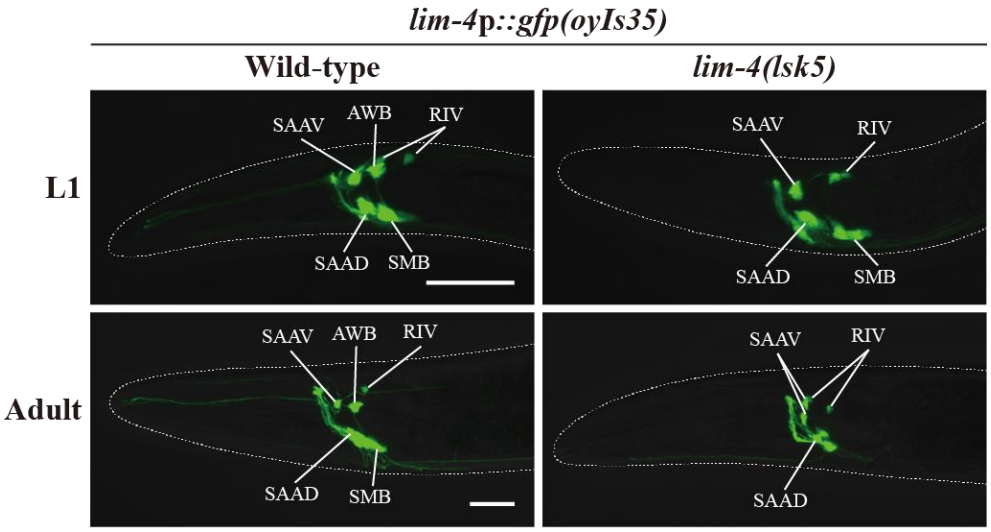
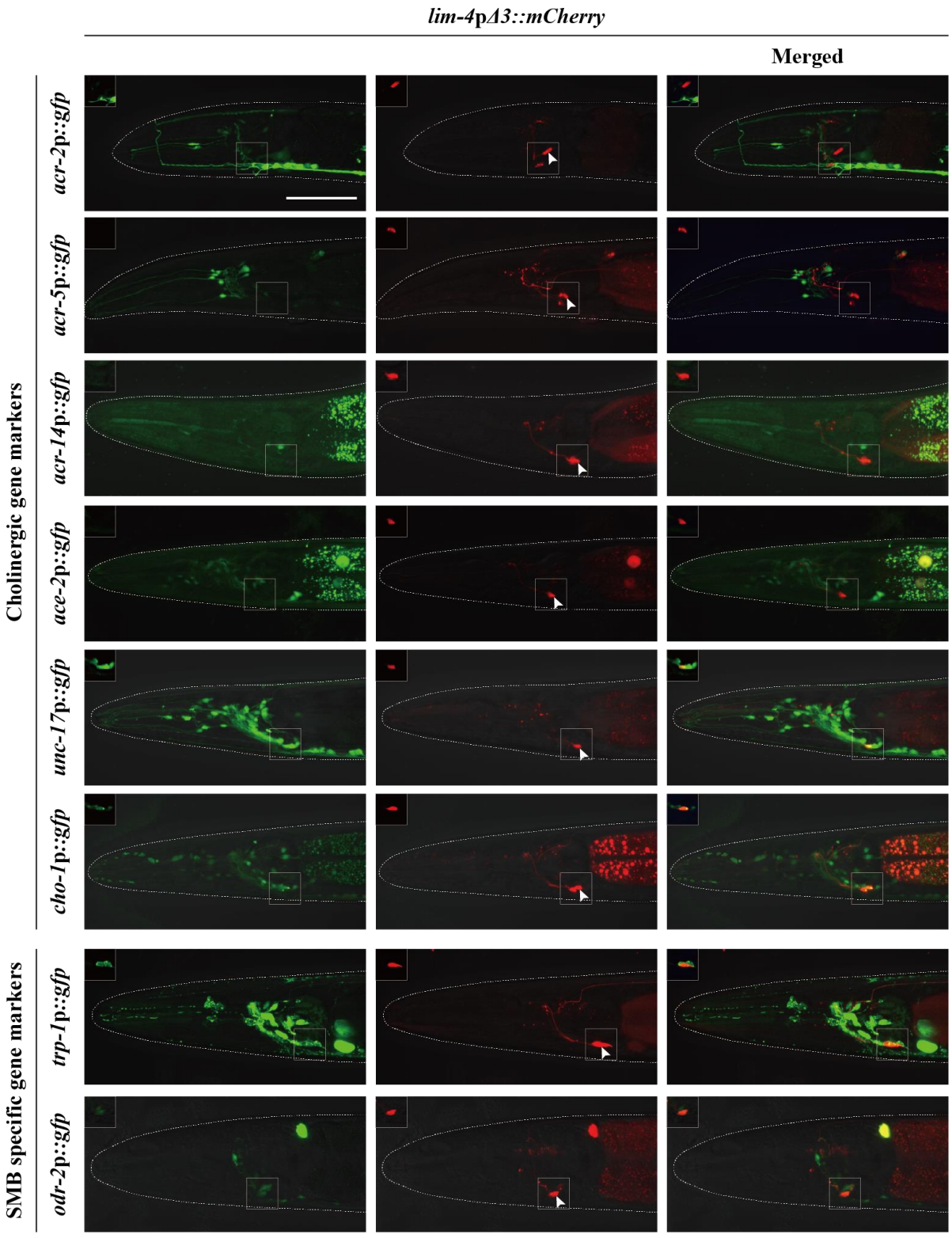


Figure 19. *lim-4* is expressed in and acts in the SMB neurons and its expression is auto-regulated. In *lim-4* mutants, expression of the *lim-4*p*::gfp* reporter is abolished in the cell bodies and processes of the SMB neurons in adult but not L1 larval stage animals. GFP expression of the *lim-4*p*::gfp* reporter is shown in wild-type (left column) or *lim-4* (*lsk5*) mutant (right column) animals at L1 larval (top) or adult stage (bottom). Note that the *lim-4*p*::gfp* reporter is not expressed in the AWB neurons of *lim-4* mutants [102]. Images are derived from z-stacks of confocal microscopy images. Quantitative analysis of these phenotypes is shown in Table 4. Anterior is to the left. Scale bar: 20 μm.


lim-4pΔ3::mCherry
Merged
Cholinergic gene markers
acr-2p::gfp
acr-5p::gfp
acr-14p::gfp
ace-2p::gfp
unc-17p::gfp
cho-1p::gfp
SMB specific gene markers
trp-1p::gfp
odr-2p::gfp

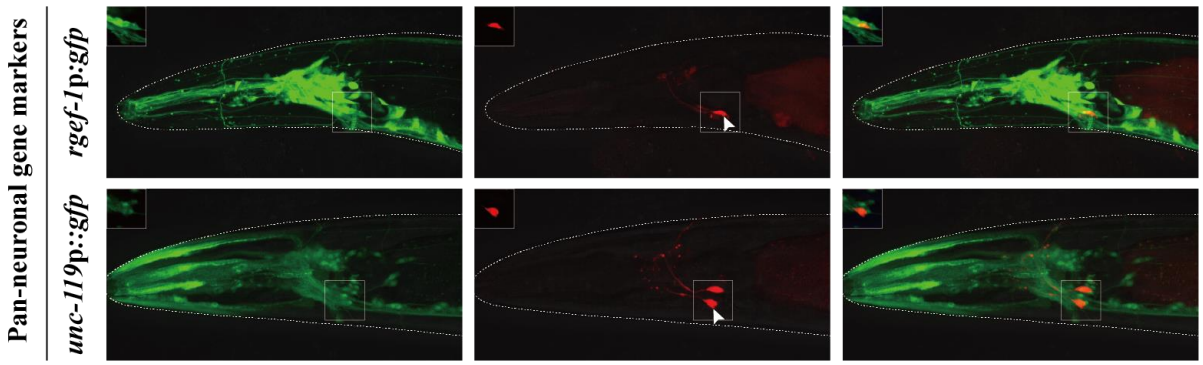


Figure 20. The SMB neurons express cholinergic or pan-neuronal markers. Expression of *gfp* reporter constructs under the control of cholinergic (*unc-17*, *cho-1*), SMB-specific (*trp-1*, *odr-2*) or pan-neuronal (*rgef-1*, *unc-119*) gene promoter is overlapped with expression of the *lim-4*p*Δ3::mCherry* reporter in the SMB neurons of wild-type animals. Expression of *acr-2*, *acr-5*, *acr-14*(AChRs), or *ace-2* (AChE) cholinergic marker is not detected in the SMB neurons. Expression of the indicated reporter constructs and of the *lim-4*p*Δ3::mCherry* reporter is shown in left or middle column, respectively, and the merged images are shown in right column. Images are derived from z-stacks of confocal microscopy images while images in the upper-left boxed regions are single focal plane confocal microscopy images. Anterior is at left in all images. Scale bar: 50 μm.

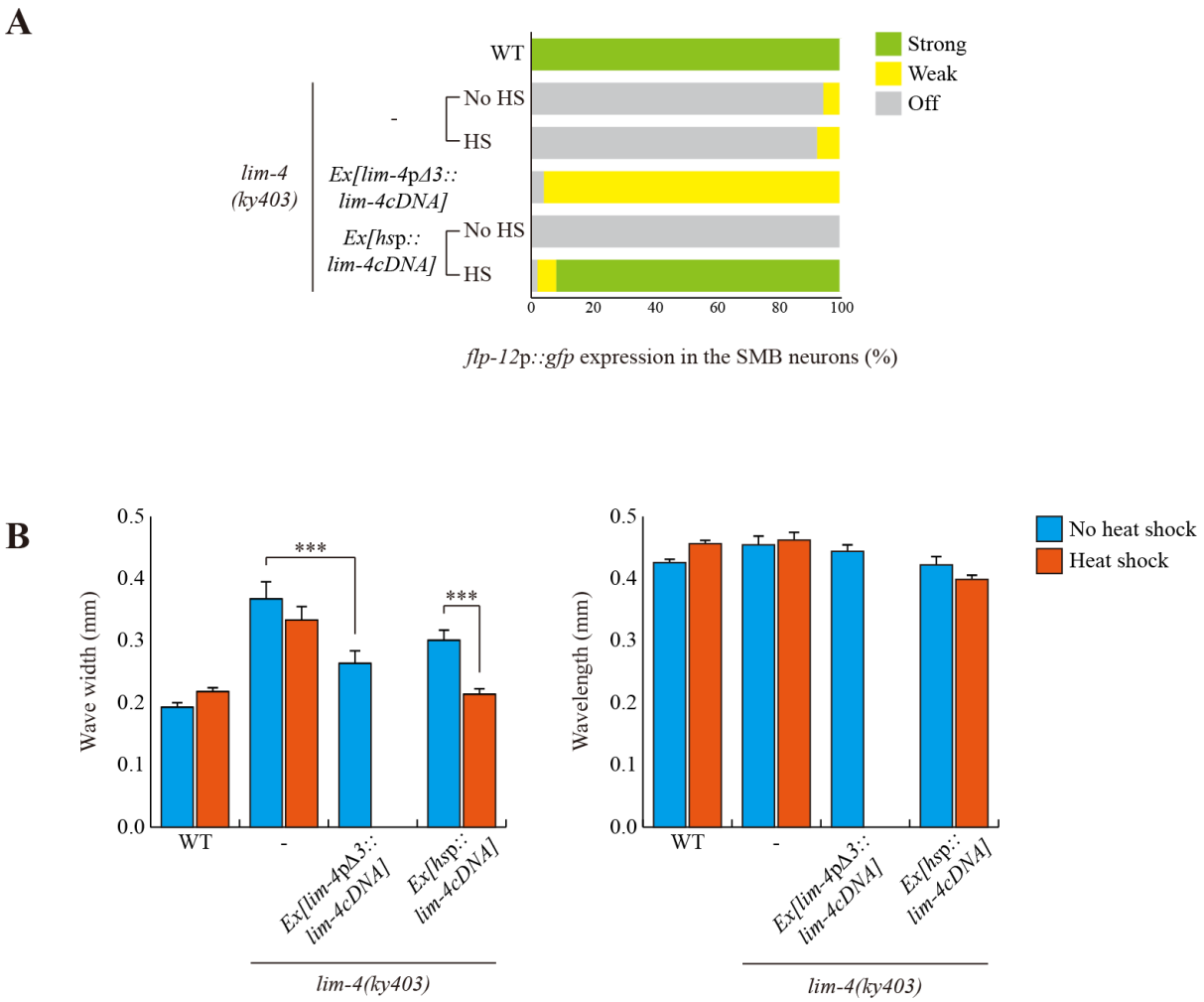


Figure 21. The gene expression and locomotion defects of *lim-4* mutants were partially restored. (A) Percentage of animals of the indicated genotypes expressing stably integrated *flp-12*p*:gfp* reporter (*ynIs25*) is shown. Strong, weak or off expression is defined as GFP expression observed at 400x magnification in both cell bodies and processes, in only cell bodies, or not observed either in cell bodies or processes, respectively. Heat shocks were treated to L4 larval stage animals at 33°C twice for 30 minutes and after 14 hours, phenotypes were analyzed. Over two independent transgenic lines were tested. n≥50 for each. (B) The average of wave width or wavelength of the indicated genotypes. n≥30 for each. Error bars are the SEM. *** indicates significantly different between indicated animals at $p<0.001$ (one-way ANOVA test followed by the Tukey post-hoc test). Heat shocks were applied to L4 larval stage animals at 33°C twice for 30 minutes and phenotypes were analyzed after 14 hours.

Table 7. Dye-filling defects of *lim-4* mutants are partially rescued by temporal expression of LIM-4

Genotype	Extrachromosomal array		% animals showing dye-filling in AWB/ADF
WT	-	No Heat shock	100
lim-4 (ky403)	-	No Heat shock	23
	-	Heat shock	17
	*Ex[lim-4*pΔ*3::lim-4cDNA]*	No Heat shock	33
	Ex[hsp::lim-4cDNA]	Heat shock	77
	Ex[hsp::lhx6cDNA]	Heat shock	53
	Ex[hsp::lhx8cDNA]	Heat shock	50

Lipophilic dye DiD was used to observe dye-filling in the AWB/ADF neurons of *lim-4* mutant animals. Dye staining in adult animals was observed at 400x. n≥30.

3.5 Postdevelopmental expression of LIM-4 is sufficient to restore the SMB-specific defects of *lim-4* mutants

To determine when the activity of LIM-4 is required for the expression of the SMB markers and proper locomotive movement, first I expressed *lim-4* with an inducible, ubiquitously expressed heat-shock promoter (*hsp16.2*) [101]. Upon transient supply of *lim-4* gene activity at the fourth larval stage (i.e., long after the SMB neurons have differentiated in the embryo), expression of *flp-12* was fully restored and the loopy phenotype of *lim-4* mutants was rescued (Figures 21A-21B, 22, and 23A). These results demonstrate that postdevelopmental expression of LIM-4 is sufficient to restore the expression of the SMB markers and the function of the SMB neurons in *lim-4* mutants. This data further indicate that the SMB neurons are not irreversibly switched to another cell-fate and demonstrate that loss of *lim-4* does not result in irreversible developmental defects. The dye-filling defects of the AWB/ADF neurons in *lim-4* mutants were partially rescued after multiple heat shocks (Table 7) [102, 106].

I also used the inducible rescue assay to corroborate the prediction that *lim-4* is continuously required to maintain the functional properties of the SMB neurons. To this end, I supplied *lim-4* activity via the heat-shock promoter at L4 stage and then analyzed the animals after 14 hours at the young adult stage. In these animals, I found the locomotory defects to be partially rescued (Figure 23B). When assayed after a long time interval at 3 and 6 days (i.e., older adult stages), the animals again displayed a mutant phenotype indistinguishable from the control (Figure 23B), suggesting that the transient rescuing ability of the *lim-4* gene activity has faded. These results demonstrate that *lim-4* does not only initiate but also maintains the expression of the SMB terminal differentiation genes and, hence, the function of the SMB neurons.

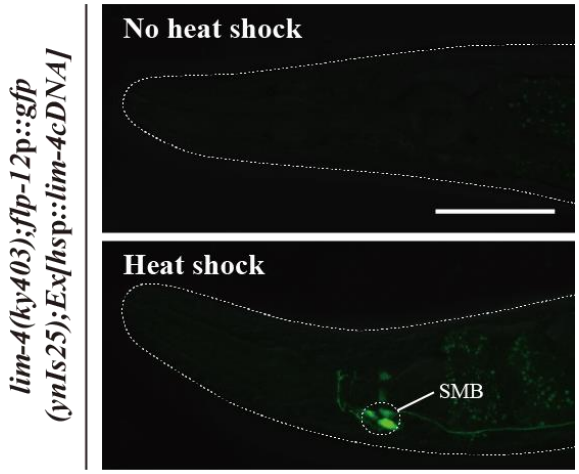


Figure 22. Expression of *flp-12* was fully restored by temporal expression of LIM-4. Shown are representative pictures of *lim-4* mutants containing the *hsp::lim-4cDNA* transgene with no heat shock or after heat shock treatment. Images are derived from z-stacks of confocal microscopy images. Scale bar: 50 µm.

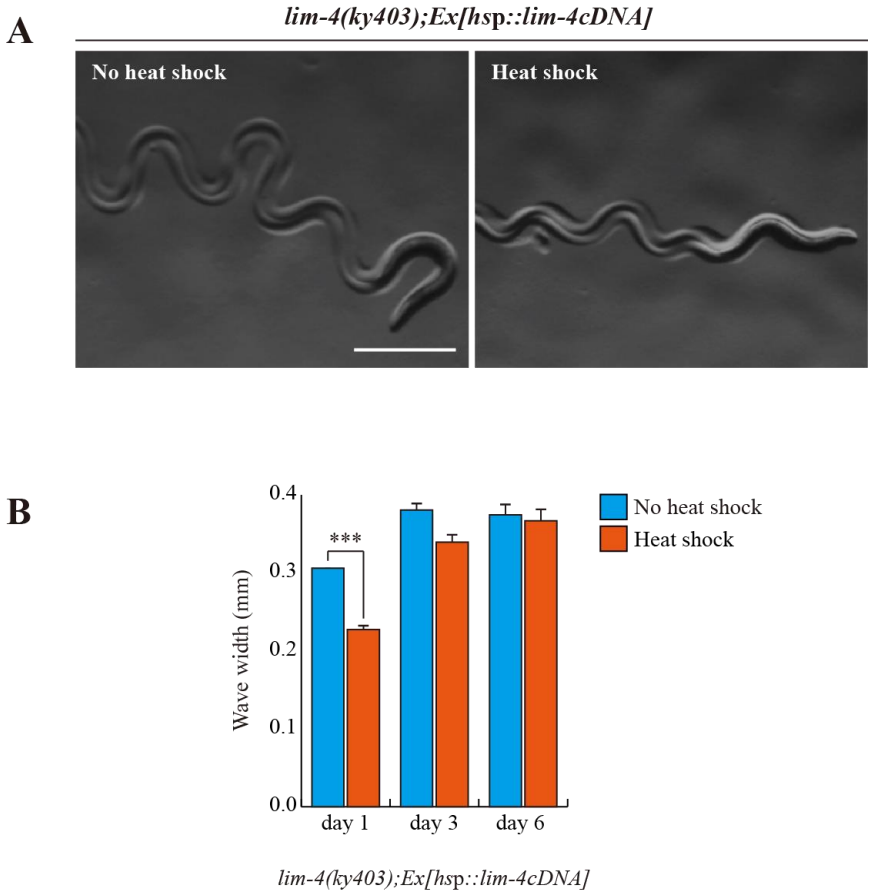


Figure 23. The loopy phenotype of *lim-4* mutants was rescued by temporal expression of LIM-4. Shown are representative pictures of *lim-4* mutants containing the *hsp::lim-4cDNA* transgene with no heat shock or after heat shock treatment. (A) Images are derived from a light microscopy image (F: Scale bar: 0.5 mm). (B) *lim-4* is required to maintain function of the SMB neurons. The average wave width was analyzed in *lim-4* (*ky403*); *Ex*[*hsp::lim-4cDNA*] at 1 day after, 3 day after or 6 day after heat shock treatment. n≥30 for each. Error bars are the SEM. *** Significantly different between no heat shock and heat shock treatment conditions at $p<0.001$ (one-way ANOVA test followed by the Tukey post-hoc test).

3.6 Expression of *lim-4* is sufficient to induce the SMB identity in other cell-types

To address whether expression of *lim-4* is sufficient to induce the SMB identity in other cell-types, first I examined ectopic *flp-12* expression upon transient supply of *lim-4* gene activity via the heat-shock promoter at the embryonic stage. Although the heat-shock promoter should drive ubiquitous LIM-4 expression, ectopic *flp-12* expression was not seen broadly elsewhere; interestingly, expression was limited in only one cell-type, the ALN neurons, in 25% of transgenic animals (n=50) (Figure 24). The ALN neurons are a pair of cholinergic oxygen-sensing neurons in the tail [6, 99, 114] that do not appear functionally or linearly related to the SMB neurons (Figure 15).

I next attempted to express LIM-4 in the glutamatergic chemosensory neurons AWC using the promoter of the *ceh-36* homeobox gene [113, 115]. Ectopic expression of *unc-17* VAChT was detected in AWC (Figure 25) while the *flp-12* was not ectopically expressed in AWC (Figure 26). Moreover I induced broader ectopic expression of LIM-4 in the subset of glutamatergic neurons under a specific *eat-4* glutamate transporter gene promoter that drives reporter expression in 11 (but not in AWC) out of 38 glutamatergic neuron classes in the hermaphrodites [73]. Expression of LIM-4 in these cells did not drive ectopic expression of *cho-1* ChT or affect expression of *eat-4* (Figure 27), suggesting that expression of LIM-4 alone is not sufficient to generally induce cholinergic cell fate in a subset of glutamatergic neurons. These results suggest that *lim-4* is partially sufficient to drive expression of the SMB markers in a context-dependent manner.

*lim-4(ky403);flp-12*p*::gfp(ynIs25);Ex[hs*p*::lim-4cDNA]*

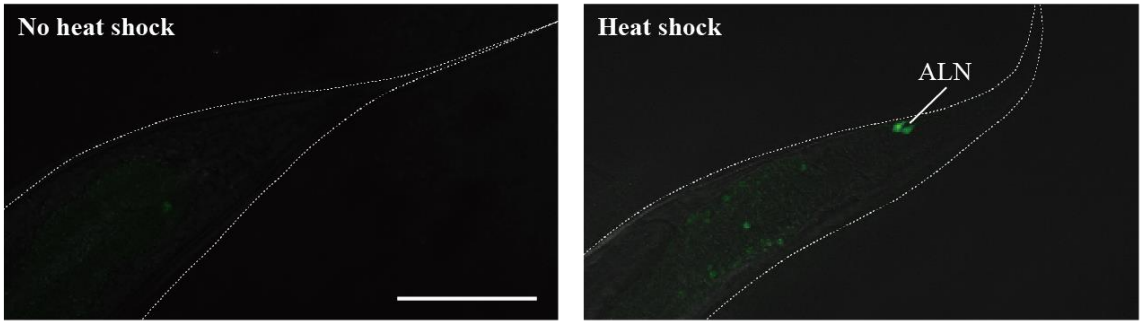


Figure 24. *lim-4* is sufficient to induce peptidergic marker expression in other cell-types. Ectopic expression of LIM-4 is sufficient to drive *flp-12* expression in the ALN cholinergic neurons. Heat shocks were applied to 2 or 3 fold stage of embryos at the 37°C twice for 30 minutes and phenotypes of adults were analyzed. Images are derived from z-stacks of confocal microscopy images. Posterior is at right. Scale bar: 50 μm.

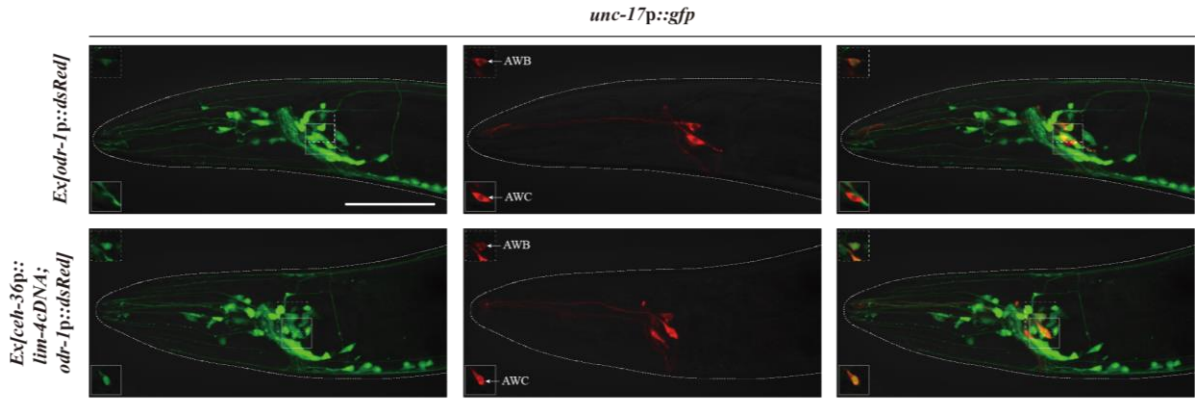


Figure 25. *lim-4* is sufficient to induce cholinergic marker expression in other cell-types. Ectopic expression of LIM-4 in AWC induces *unc-17* expression. *Ex[odr-1*p*::dsRed]* transgenic animals express dsRed in AWC and AWB [116]. Note that I observed *unc-17* expression in the AWB neurons of wild-type animals. Images are derived from z-stacks of confocal microscopy images while images in the upper or bottom-left boxed regions are single focal plane confocal microscopy images. Posterior is at right. Scale bar: 50 μm.

flp-12*p*::gfp(ynIs25);
Ex[ceh-36*p*::lim-4cDNA;odr-1*p*::dsRed]

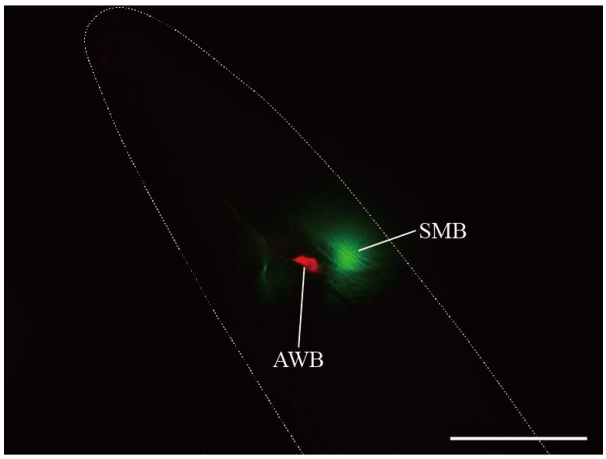


Figure 26. Ectopic expression of LIM-4 in the AWC neurons does not drive *flp-12* expression in the AWC neurons. *ceh-36* gene promoter drives LIM-4 expression in the AWC and ASE neurons [113]. *Ex[odr-1*p*::dsRed]* transgenic animals express dsRed in AWC and AWB [116]. Anterior is at left in all images. Scale bar: 50 μm.

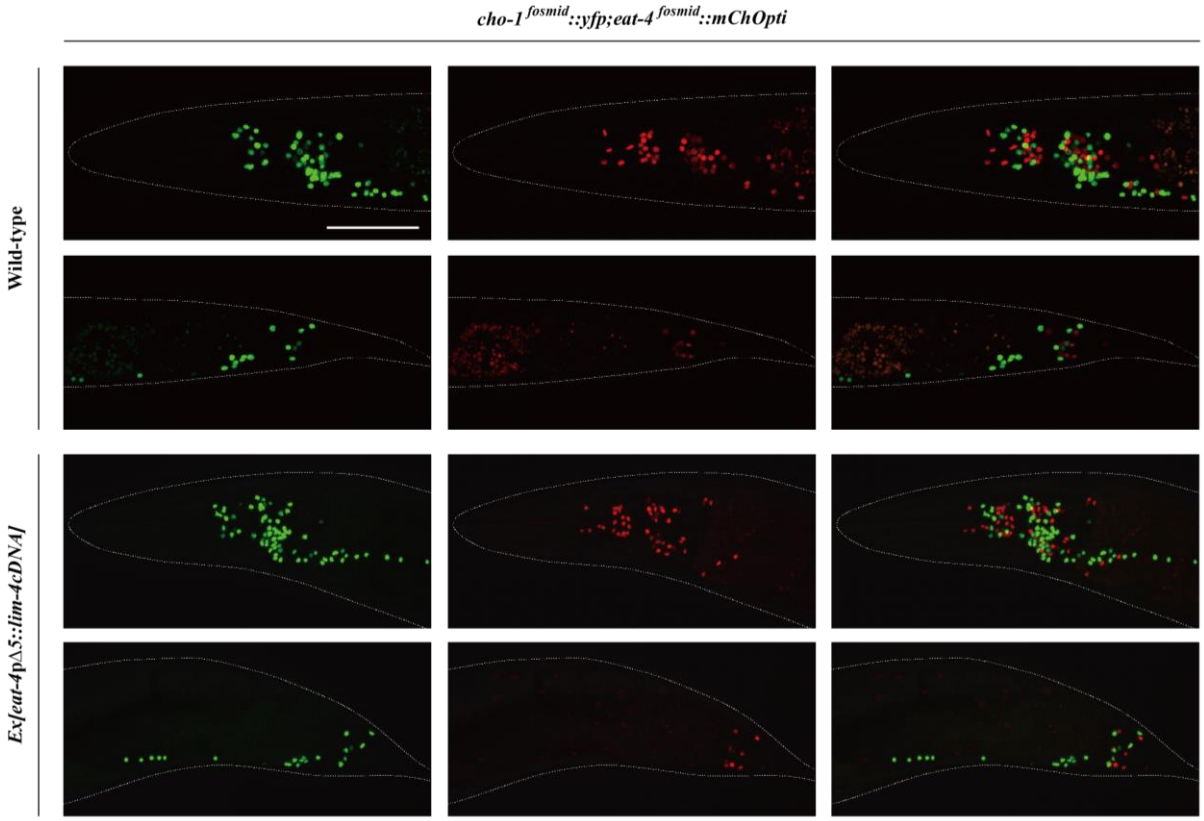


Figure 27. Ectopic expression of LIM-4 does not drive *cho-1* expression in the subset of glutamatergic neurons. *eat-4*p*Δ5* promoter drives LIM-4 expression in 11 out of 38 glutamatergic neuron types [73]. Images are derived from z-stacks of confocal microscopy images. Anterior is at left in all images. Scale bar: 50 μm.

3.7 LIM-4 regulates gene expression via a *cis*-regulatory motif in the SMB markers

Homeodomain transcription factors generally bind well-defined DNA sequences to control transcription of target genes[117]. Systemic analysis of homeodomain DNA-binding specificities allowed prediction of the recognition motif of each homeodomain protein, and a *cis*-regulatory motif containing the consensus TAAT core DNA sequences was predicted to be the binding site for the homeodomain of LIM-4 and its mammalian and *Drosophila* homologs (LHX6/8 and Arrowhead, respectively) (Figures 28 and 29) [118, 119]. Indeed, LHX6 and LHX8 have been shown to directly bind to the predicted DNA sequences (ATAATCA) in the promoter regions of the *Shh* gene [120].

To identify *cis*-regulatory motifs required to drive expression of the SMB markers in the SMB neurons, DNA sequences within the promoters of the SMB markers were serially deleted and the resultant transgenic animals were examined for altered expression patterns. From these analyses, first I determined a minimal region within the *flp-12* promoter for *flp-12* expression. Deletion of a 150 bp sequence located ~162 bp upstream of the translation start sequence caused decreased *gfp* expression in the SMB neurons (Figures 30 and 31). Within the 150 bp region, I next found four AT- rich DNA sequences that are fully conserved in the promoters of the *flp-12* orthologs in the related *Caenorhabditis* species (Figure 30). Mutations of two AT-rich DNA sequences resulted in an almost complete loss of *gfp* expression, while mutations of the other two sequences did not affect the *gfp* expression (Figure 30), indicating that the former two motifs are necessary for the expression of *flp-12* in the SMB neurons. DNA sequences of these motifs (AAAATTG and ACAATAG) resemble putative LIM-4 binding sequences, and will be referred to as SMB motifs (Figure 28). To test whether these SMB motifs are sufficient to drive gene expression in the SMB neurons, I inserted three copies of the SMB motifs in the promoter of *flp-7*, which is normally expressed in the several head neurons, but not in the SMB neurons (Figure 32) [100]. Transgenic animals expressing a *flp-*

*7*p-SMB motif*::gfp* reporter construct still exhibited *gfp* expression in *flp-7* expressing neurons, indicating that insertion of the SMB motifs within the regulatory region of *flp-7* does not alter the *flp-7* expression pattern. In addition, I observed consistent expression of *flp-7* in the SMB neurons in 100% transgenic animals (n=50) (Figure 32), suggesting that these SMB motifs are necessary and sufficient to drive gene expression in the SMB neurons. These results are consistent with the hypothesis that LIM-4 directly binds the SMB motifs to regulate expression of *flp-12* gene in the SMB neurons.

I next analyzed the promoters of additional SMB markers, *odr-2* and *unc-17* (note that expression of *trp-1* and *cho-1* in the SMB neurons is too weak to generate reliable results from their promoter analyses), and identified the regulatory motifs within the promoter regions of *odr-2* or *unc-17*. These motifs in *odr-2* or *unc-17* promoter resemble the SMB motifs found in *flp-12* promoter (Figure 33). I mutated these motifs in the context of the *odr-2* and *unc-17* reporter genes and found that these mutations reduced expression of the reporter genes (Figure 33). These results demonstrate that distinct terminal identity markers of a specific neuron are co-regulated by shared *cis*-regulatory motifs.

A *cis*-regulatory region in the *lim-4* promoter that is required for *lim-4* expression in the AWB neurons was previously identified [121]. I performed analogous deletion analysis experiments in transgenic animals to dissect the *lim-4* promoter to identify motifs required for *lim-4* expression in the SMB neurons (Figure 34). As proof-of-principle, I also identified the *lim-4* regulatory sequences for the AWB expression (Figure 34). However, I could not identify simple *cis*-regulatory motifs in the *lim-4* promoter required for the SMB expression (Figure 34). Instead, multiple regions in the *lim-4* promoter positively or negatively regulated LIM-4 expression in the SMB neurons, suggesting a complexity of *cis*-regulatory motifs in the *lim-4* promoter to ensure proper LIM-4 expression in the SMB neurons.

Figure 28. The predicted binding site of LIM-4 and SMB motif. Shown are the predicted binding site (top) for the homeodomain of LIM-4 from a web based tool, PreMoTF (http://stormo.wustl.edu/PreMoTF) [119] and the putative *cis*-regulatory SMB motif (bottom) identified from promoter analysis of *flp-12*, *odr-2*, and *unc-17* genes.

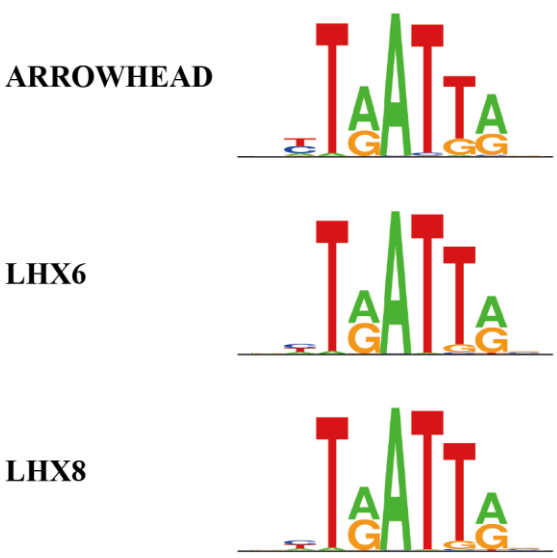


Figure 29. The predicted binding site of Arrowhead, LHX6, and LHX8. The predicted binding sites for the homeodomain of Arrowhead, LHX6, or LHX8. The binding sites are derived from a web based tool, PreMoTF (http://stormo.wustl.edu/PreMoTF) [119].

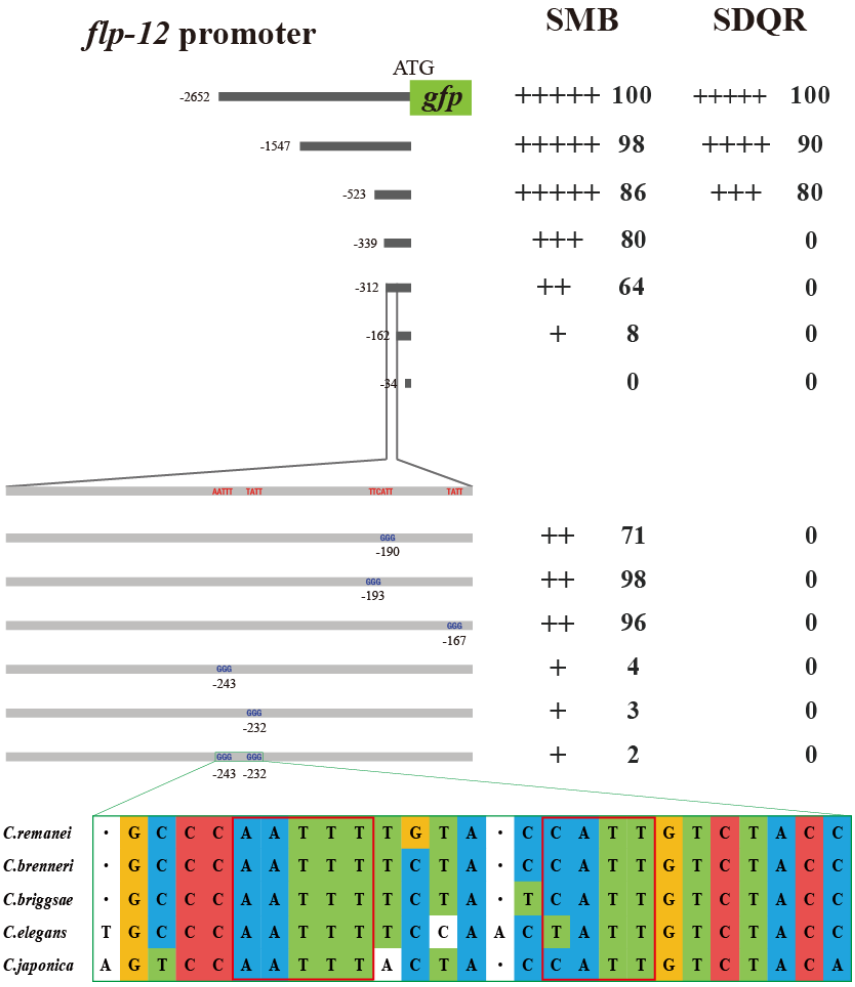


Figure 30. AT-rich two *cis*-regulatory motifs are required for the expression of *flp-12* in the SMB neurons. The percentage of transgenic animals expressing *flp-12*p*::gfp* reporter construct in the indicated neurons is shown. Strength of GFP expression is indicated by the number of + symbols. Wild-type nucleotides are indicated in red, mutated nucleotides in blue. At least two independent extrachromosomal lines for each construct were examined except a *flp-12* promoter (-2652) construct of which number was derived from one integrated line. n≥50 for each. Identified *cis*-regulatory sequences found in other *Caenorhabditis* species are shown below. Analysis between -523 bp and -339 bp upstream of *flp-12* promoter is shown in Figure 31.

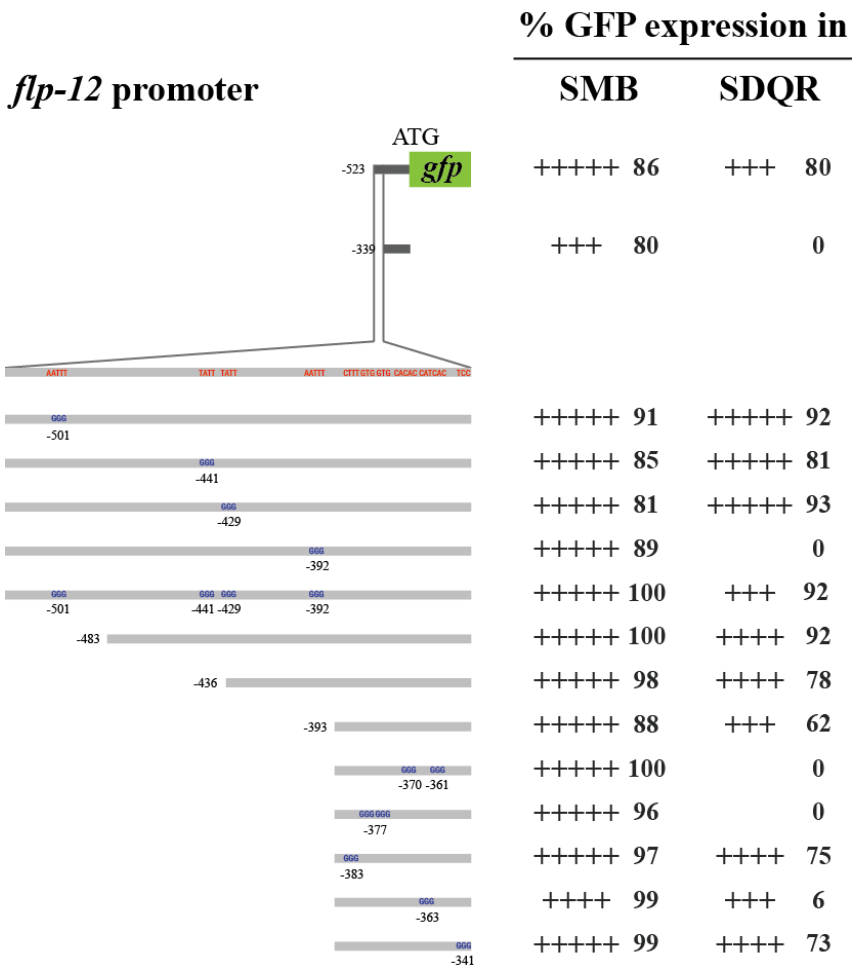


Figure 31. Analysis between -523 bp and -339 bp upstream of *flp-12* promoter. A *cis*-regulatory sequence for the SMB neurons was not identified in this promoter region. The percentage of transgenic animals expressing each *gfp* reporter construct in the indicated neurons is shown. Strength of GFP expression is indicated by the number of + symbols. Point mutated nucleotides are indicated as wild-type in red line. At least two independent extrachromosomal lines for each construct were examined. n≥50 for each.

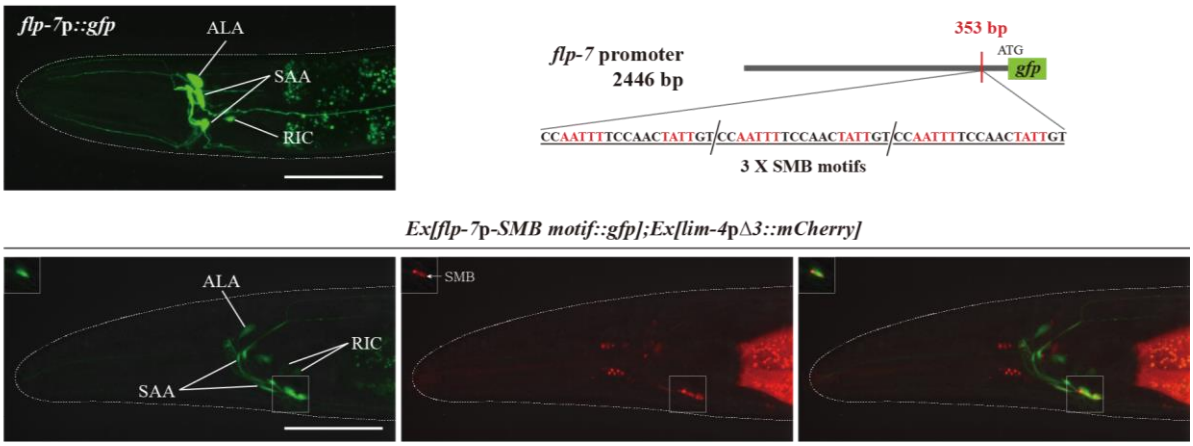
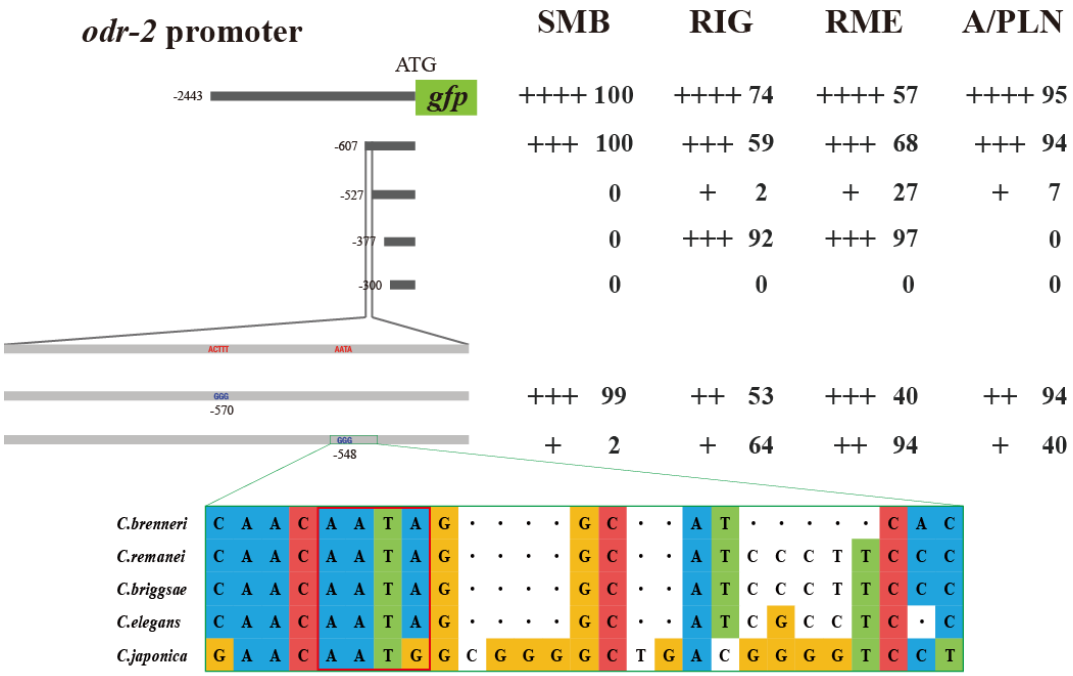


Figure 32. A *cis*-regulatory motif is necessary and sufficient to drive expression of terminal differentiated SMB markers. Insertion of the SMB motif into non-SMB expressed *flp-7* gene promoter induces *flp-7* expression in the SMB neurons. The inserted *cis*-regulatory sequences identified from *flp-12* promoter analysis and an inserted site in *flp-7* promoter are indicated. GFP expression is overlapped with expression of the *lim-4*p*Δ3::mCherry* reporter in the SMB neurons of wild-type animals. Images are derived from z-stacks of confocal microscopy images while images in the upper-left boxed regions are single focal plane confocal microscopy images. Anterior is to the left. Scale bars: 50 μm.


odr-2 promoter
SMB
RIG
RME
A/PLN
ATG
gfp
-2443
-607
-527
-377
-300
-570
-548
++++ 100
++++ 74
++++ 57
++++ 95
+++ 100
+++ 59
+++ 68
+++ 94
0
+ 2
+ 27
+ 7
0
+++ 92
+++ 97
0
0
0
0
+++ 99
++ 53
+++ 40
++ 94
+ 2
+ 64
++ 94
+ 40
C.brenneri C A A C A A T A G · · · · G C · · A T · · · · · C A C
C.remanei C A A C A A T A G · · · · G C · · A T C C C T T C C C
C.briggsae C A A C A A T A G · · · · G C · · A T C C C T T C C C
C.elegans C A A C A A T A G · · · · G C · · A T C G C C T C · C
C.japonica G A A C A A T G G C G G G G C T G A C G G G G T C C T

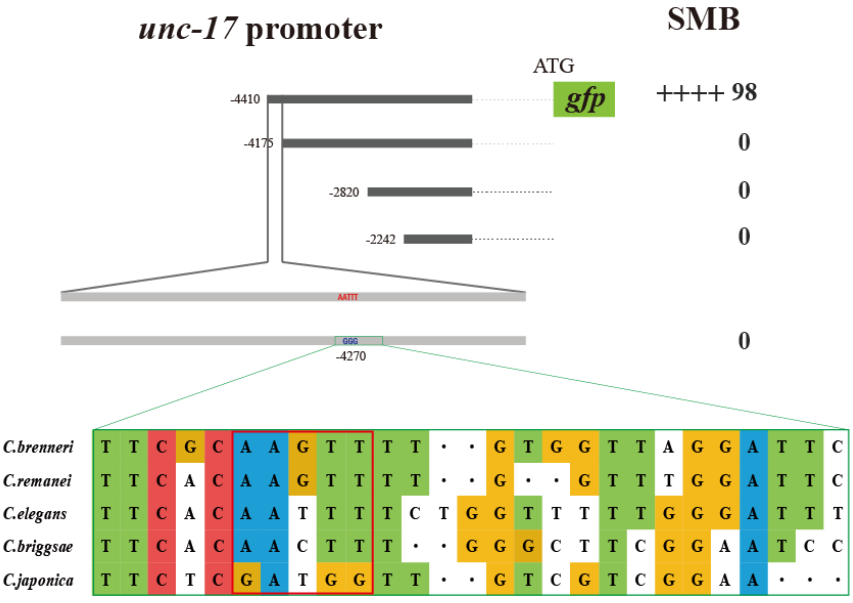

unc-17 promoter
SMB
ATG
gfp
-4410
-4175
-2820
-2242
-4270
++++ 98
0
0
0
0
C.brenneri T T C G C A A G T T T T · · G T G G T T A G G A T T C
C.remanei T T C A C A A G T T T T · · G · · G T T T G G A T T C
C.elegans T T C A C A A T T T T C T G G T T T T T G G G A T T T
C.briggsae T T C A C A A C T T T · · G G G C T T C G G A A T C C
C.japonica T T C T C G A T G G T T · · G T C G T C G G A A · · ·

Figure 33. *cis*-regulatory motifs is required for the expression of *odr-2* and *unc-17* in the SMB neurons. The percentage of transgenic animals expressing *odr-2*p*::gfp* or *unc-17*p*::gfp* reporter construct in the indicated neurons is shown. Strength of GFP expression is indicated by the number of + symbols. Wild-type nucleotides are indicated in red, mutated nucleotides in blue. At least two independent extrachromosomal lines for each construct were examined. n≥50 for each. Identified *cis*-regulatory sequences found in other *Caenorhabditis* species are shown below.

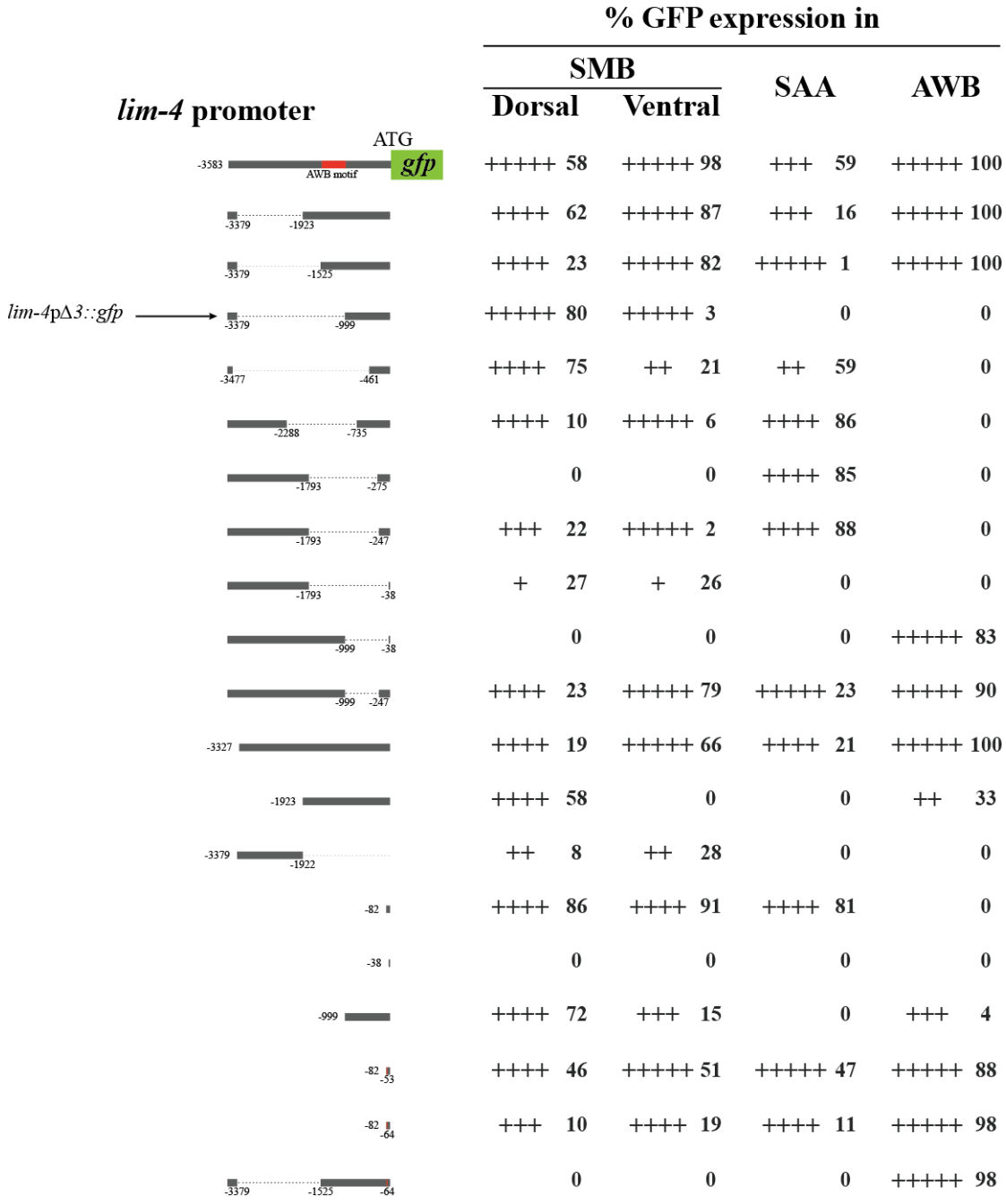


Figure 34. Analysis of the *lim-4* gene promoter. The percentage of transgenic animals expressing each *gfp* reporter construct in the indicated neurons is shown. Strength of GFP expression is indicated by the number of + symbols. Deleted regions in the promoter are indicated as a dotted line. Mutated nucleotides are indicated as wild-type in red line. The *cis*-regulatory sequence for the AWB expression of *lim-4* is indicated. At least two independent extrachromosomal lines for each construct were examined. n ≥50 for each.

3.8 The function of *lim-4* is conserved in human

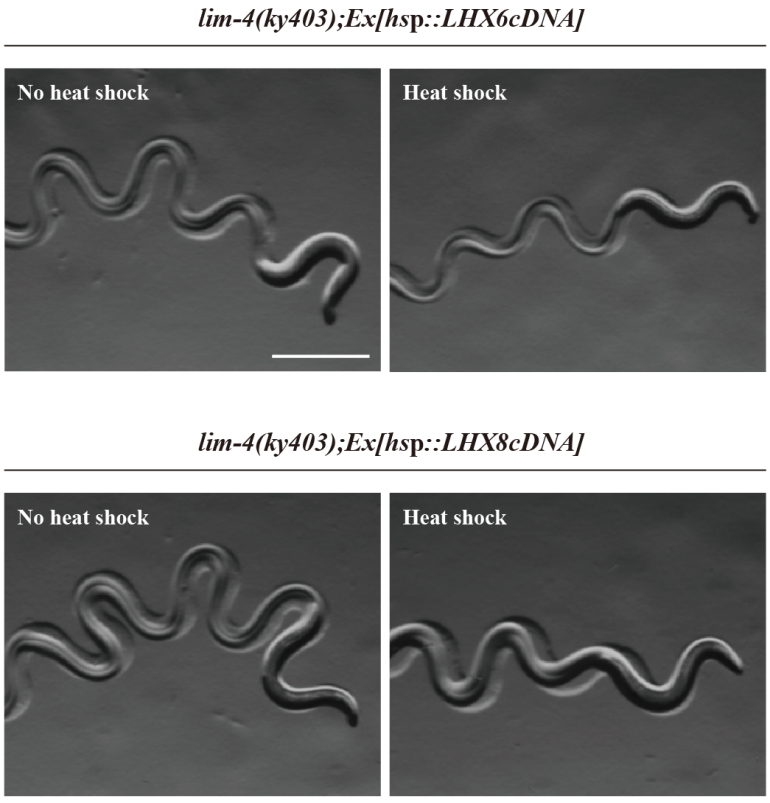
The mammalian genome contains two LIM-4 orthologs, *LHX6* and *LHX8*. In mice, these genes are largely expressed in the developing and adult striatum and orchestrate specification of interneuron identities; specifically, *LHX6* and *LHX8* are required to determine GABAergic/peptidergic and cholinergic interneuronal cell fate, respectively. In addition, these genes have redundant function to regulate expression of *shh* in MGE neurons [120]. LIM-4 exhibits high degree of protein sequence homology to LHX6 and LHX8 (in particular, 60 % identical in its homeodomain) (Figure 35A) [102], suggesting a functional conservation of these proteins.

To test for functional homology, I first tried to rescue *C. elegans lim-4* mutants by expressing human *LHX6 or LHX8* cDNA under the control of the heat shock promoter. Similar to *C. elegans lim-4* cDNA, human *LHX6 or LHX8* cDNA fully restored altered locomotion of *lim-4* mutants (Figure 35B-35C). Moreover, *LHX6* also fully rescued the defect of *flp-12* expression in *lim-4* mutants while *LHX8* did not rescue (Figure 36A-36B), indicating that LHX6 may have higher degree functional conservation to LIM-4 than LHX8.

A

C.elegans **LIM-4** TKRVRTTFAEDQLSVLQTYFNRDSNPDGADLEKIASMTGLSKRVTQVWFQNSRARQKKWH
Human LHX6 AKRARTSFTAEQLQVMQAQFAQDNNPDAQTLQKLADMTGLSRRVIQVWFQNCRARHKKHT
Human LHX8 AKRARTSFTADQLQVMQAQFAQDNNPDAQTLQKLAERTGLSRRVIQVWFQNCRARHKKHV

B



C

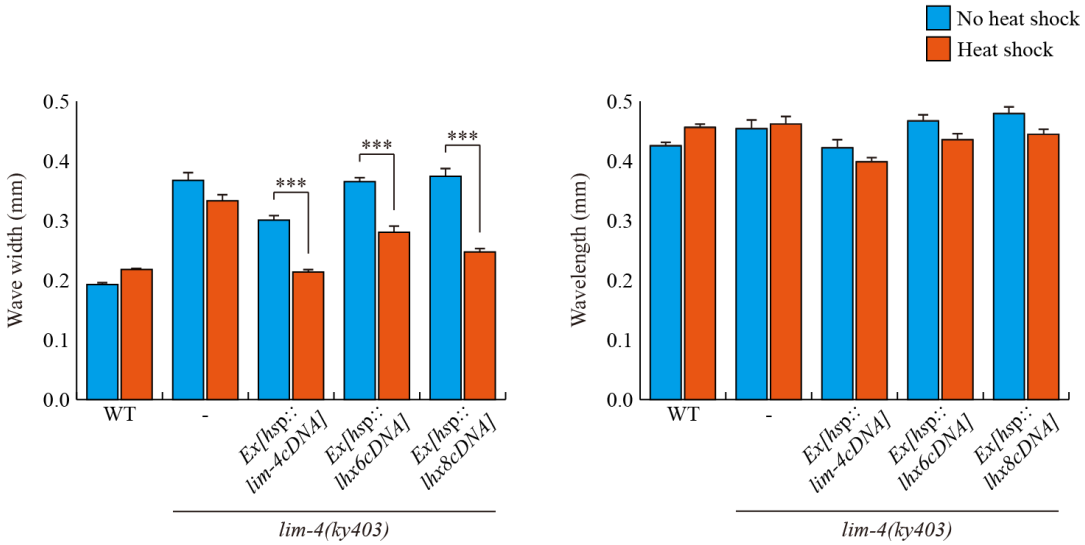
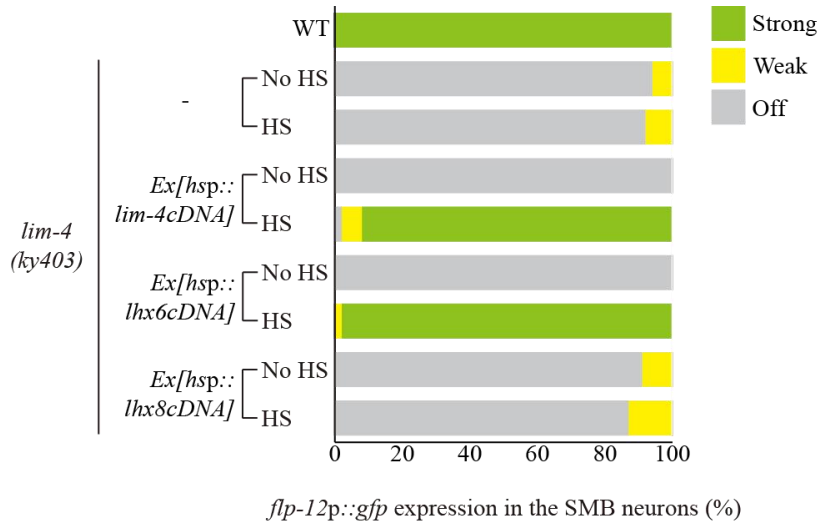


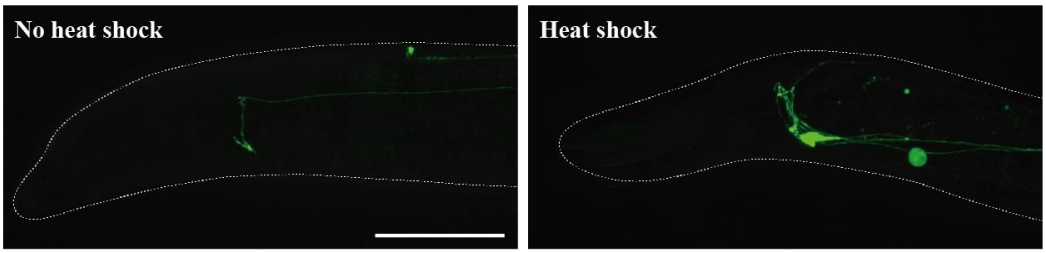
Figure 35. The loopy phenotype of *lim-4* mutants was fully restored by temporal expression of LHX6 or LHX8. (A) Shown is the sequence alignment of the homeodomain of *C. elegans* LIM-4 and human LHX6 and LHX8. Identical residues in at least two proteins are shown in red. (B) Representative pictures of *lim-4* mutants containing the *hs*p*::LHX6cDNA* or *hs*p*::LHX8cDNA* transgene with no heat shock (left) or after heat shock (right) treated conditions. Images are derived from a light microscopy image. Scale bar: 0.5 mm. (C) Average of wave width or wavelength of the indicated genotypes. n≥30 for each. Error bars are the SEM. *** indicates different between heat shock and no heat shock conditions at $p<0.001$ (student *t*-test). Heat shocks were applied to L4 larval stage animals at 33℃ twice for 30 minutes and phenotypes were analyzed after 14 hours. Data of WT, -, or *Ex[hs*p*::lim-4cDNA]* are from Figure 21B.

A



B

lim-4(ky403);flp-12p::gfp(ynIs25);Ex[hsp::LHX6cDNA]



lim-4(ky403);flp-12p::gfp(ynIs25);Ex[hsp::LHX8cDNA]

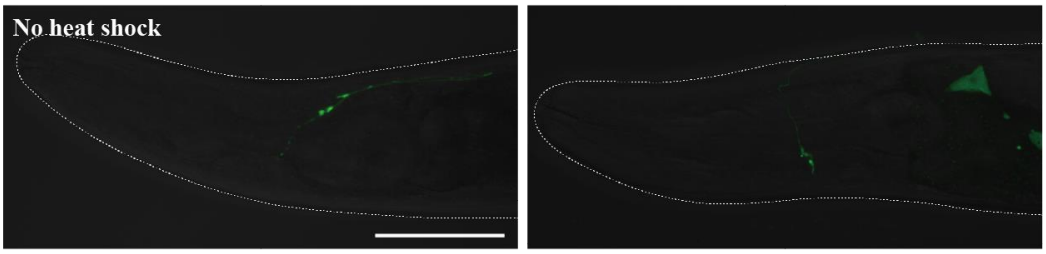


Figure 36. Human LHX6 fully rescued the defect of *flp-12* expression in *lim-4* mutant. (A) Percentage of animals of the indicated genotypes expressing stably integrated *flp-12*p*::gfp* reporter (*ynIs25*) is shown. Strong, weak or off expression is defined as GFP expression observed at 400x magnification in both cell bodies and processes, in only cell bodies, or not observed either in cell bodies or processes, respectively. Over two independent transgenic lines were tested. n≥50 for each. Data of WT, -, or *Ex[hsp::lim-4cDNA]* are from Figure 21A. (B) Representative pictures of *lim-4* mutants containing the *hsp::LHX6cDNA* or *hsp::LHX8cDNA* transgene with no heat shock (left) or after heat shock (right) treated conditions. Images are derived from z-stacks of confocal microscopy images. Scale bar: 50 μm.

IV. Discussions

In these studies, I have identified an important regulator that controls terminal differentiation of a distinct neuronal cell-type in *C. elegans*. The SMB neurons appear to have mixed neuronal functions. Their long, unbranched and synapse-free processes along the body may serve as proprioreceptors to sense body stretch. In addition, they contact over 20 sensory, inter, or motor neurons via chemical or electrical synapses and may integrate additional extrinsic or intrinsic cues to regulate head muscle contraction [6]. The SMB neurons co-express a unique combination of neurotransmitters, acetylcholine and FLP-12 FMRFamide-like neuropeptides. Thus, an intriguing question is how the SMB sensory/inter/motor neurons acquire their unique characteristics. My experiments show that the LIM-4 LIM homeodomain transcription factor is necessary and sufficient to promote and probably maintain SMB-specific properties and functions. In *lim-4* null mutants, the neurotransmitter identity and neuronal function of the SMB neurons are completely lost but pan-neuronal features are not affected. Transient LIM-4 expression in *lim-4* mutants not only restores the SMB characteristics and functions, but also induces ectopic expression of a SMB-expressed neurotransmitter in another cell-type. The promoter analysis suggests that LIM-4 directly regulates expression of the SMB terminal differentiation marker via well conserved homeodomain binding sequences and also controls its own expression. Hence, I propose that *lim-4* acts as a terminal selector gene to broadly specify the SMB neuronal identity [2, 21].

Terminal differentiation of the vertebrate nervous system is also mediated via the coordinate regulation of gene batteries by terminal selector transcription factors. The most notable examples are the Etv1, Nurr1&Pitx3, and Pet-1 transcription factors. Etv1 is a ETS domain transcription factor that terminally differentiate olfactory bulb dopaminergic neurons [82]. The homolog of Etv1 in *C. elegans* is *ast-1* that also functions as a terminal selector of dopaminergic neurons and regulates expressions of differentiation genes via conserved *cis*-

regulatory motif (Figure 37). In mice, Etv1 mediates not only specification of dopaminergic identity, but is also required for the proliferation and maintenance of bulbar dopaminergic neurons (Figure 37) [82, 122]. The Nurr1 orphan nuclear receptor and Pitx3 homeobox are also terminal selectors for dopaminergic neurons in the midbrain [123]. The heterodimer transcription factor Nurr1 & Pitx3 is required for the expression of dopamine pathway gene battery by inhibiting repressor activities and directly binding *cis*-regulatory control regions to control dopaminergic differentiation genes such as dopamine transporter and tyrosine hydroxylase (Figure 37) [123]. Nurr1 & Pitx3 are also sufficient to induce dopaminergic fate in an *in vitro* model [124-127]. Another example of vertebrate terminal selector is ETS domain transcription factor Pet-1 [94, 128, 129]. Pet-1 is expressed in most of serotonergic neurons and knockout of *pet-1* in mice results in severe differentiation defects of serotonergic neurons in the brain (Figure 37) [94]. The serotonergic differentiated genes such as the enzymes for serotonin synthesis, tryptophan hydroxylase, and serotonin transporter are directly controlled by the Pet-1 terminal selector, suggesting that Pet-1 drives serotonergic neuron differentiation [94, 128, 129]. These several studies have shown that, even in the complex nervous system of vertebrates, the neuronal identity is determined by the regulatory mechanisms that control the terminal neuron differentiation and that ensure the maintenance of the differentiated phenotype.

A few terminal selector genes that determine cholinergic cell-fate in *C. elegans* have been identified. For example, the Olf/EBF gene *unc-3*, the heterodimer of LIM homeobox gene *ttx-3* and Paired-like homeobox gene *ceh-10*, and POU homeobox gene *unc-86* regulate terminal differentiation of the A-, B-, and AS-type ventral nerve cord or SAB motor neurons, the AIY interneurons, and the IL2 sensory neurons, URA motor neurons and URB interneurons, respectively [67, 71, 76, 80, 84]. *ttx-3* also acts as a terminal selector in the AIA interneurons [76]. These genes regulate expression of not only the cholinergic gene battery, including *unc-*

17 (VAChT), but also other terminally differentiated cell-specific markers. Furthermore, these trans-acting factors appear to directly bind to the evolutionarily conserved *cis*-regulatory elements of most, if not all, their target genes. In the case of the *unc-17* promoter region, distinct *cis*-regulatory target sites, such as the COE motif for *UNC-3* and the AIY motif for TTX-3/ CHE-10, are systemically organized (Figure 38A). In this study, I identified an additional *cis*-regulatory element, called the SMB motif, in the *unc-17* gene (Figure 38A), suggesting that the elaborate *cis*-regulatory architecture ensures expression of cell-specific characteristics. Since additional terminal selector genes required for specification of over 50 uncharacterized cholinergic cell-types need to be identified, additional motifs must exist in the cholinergic gene battery such as *unc-17* (Figure 38A).

Recent work has shown that distinct combination of 13 different terminal selector genes defines identity of 25 different glutamatergic cell-types in *C. elegans*, suggesting that the combinatorial codes of terminal selector transcription factors are a general theme for determining cell-type specificity [73]. In fact, in the AIY cholinergic neurons, two terminal selector transcription factors, *ttx-3* and *ceh-10,* form a heterodimer that directly regulates expression of their target genes via a common *cis*-regulatory bipartite motif; mutations of each gene lead to complete loss of the AIY specific neuronal identity [80, 84]. I have tried to identify the putative binding partner(s) of LIM-4 in the SMB neurons by analyzing the expression pattern of the *flp-12* reporter construct in mutants for which genes were previously reported to be expressed in the SMB neurons such as the *fax-1* nuclear receptor and *cog-1* Nkx6-type homeobox transcription factor [130, 131]. None of these mutations affects *flp-12* expression in the SMB neurons (Table 8). Therefore, it is not yet clear which transcription factors work in combination with *lim-4* to control the terminal differentiation of the SMB neurons. Because the results demonstrate that *lim-4* expression is initiated at an early developmental stage and then autoregulated afterward, I also tested the possibility that expression of *fax-1* or

cog-1 in the SMB neurons may regulate *lim-4* expression in SMB. In chemosensory neuron types, the *lin-11* LIM homeobox, *ceh-37* Otx, *mls-2* HMX/NKX homeobox and *nhr-67* Tail-less/TLX genes control expression of terminal selector genes in the AWA, AWB, AWC, and ASE neurons, respectively [90, 113, 115, 132]. However, *lim-4* expression is not altered in *fax-1* or *cog-1* mutants (Table 8), indicating that these SMB-expressed transcription factors may have more specific roles in development or differentiation of the SMB neurons. It is still remained to be answer which factors functions as partners with LIM-4 or act upstream or downstream of LIM-4 to control terminal differentiation of the SMB neurons.

Previous studies show that *lim-4* determines the proper cell-type specification of the AWB and serotonergic ADF neuron types [102, 106]. The AWB and ADF neurons are two classes of amphidial chemosensory neurons in the head of worms that detect volatile chemical repellants and putative food signals, respectively [110]. In *lim-4* mutants, the AWB neurons lack AWB specific characteristics and functions, such as expression of putative 7-TM receptor *str-1*, AWB-specific cilia and axon morphology, and abilities to take up lipophilic dyes, and instead acquire features and functions of the AWC olfactory neurons, such as expression of the putative 7-TM receptor *str-2* and AWC-specific cilia and axon structures, suggesting that *lim-4* acts as a cell fate switch between the AWB and AWC neurons [102]. In case of ADF cell-fate specification, *lim-4* acts transiently in the precursor cells of ADF and regulates part of the terminal differentiation process; *lim-4* mutants lack expression of a set of serotonergic markers, including tryptophan hydroxylase *tph-1*, but do not affect expression of a putative 7-TM receptor *srh-142* [106]. The *ceh-37* Otx gene is required for expression of *lim-4* in AWB and *srh-142* in ADF, suggesting that *ceh-37* also differentially affects the AWB and ADF cell fates [113]. In addition, *lim-4* has a role in regulating axon morphology of the SAA neurons but not expression of terminal differentiation markers [102]. I conclude that the SMB neurons in *lim-4* mutants do not adopt a functionally or lineage-related cell fate, but

remained undifferentiated because they lose expression of most, if not all, terminal differentiation genes. Thus, *lim-4* plays distinct roles in neuronal development in a context-dependent manner and acts as a bona fide terminal selector for the differentiation of SMB (Figure 38B).

The LIM homeobox gene family has a high degree of structural conservation through evolution amongst the homeobox gene superfamily [133-135]. However, their functional conservation among distantly related species is relatively unexplored. I demonstrate that *C. elegans lim-4* and human *LHX6* and *LHX8* (also referred as L3 or LHX7) show striking functional similarity; LHX6 completely rescues locomotive defects and *flp-12* expression phenotypes of *lim-4* mutants, whereas human *LHX8* restores only locomotion but not *flp-12* expression in *lim-4* mutants. Furthermore, expression of either *LHX6* or *lim-4* is sufficient to drive cholinergic differentiation in human neuroblastoma cells (Appendix I). The role of LHX8 in cholinergic cell-fate determination in the mammalian nervous system has been well characterized [136-138]. Deletion of the murine *LHX8 (LHX7)*, causes a subtype of cholinergic interneurons to convert into another subtype of GABAergic interneurons [136]. However, the study of LHX6 function has focused on GABAergic fate specification [136, 139]. Also, the role of LHX6 in cholinergic cell fate determination in *C. elegans* or human neuroblastoma cells that acts as a terminal selector to control the differentiation of neuronal subtypes is revealed in this study (Appendix I and II). Cholinergic neurons in the mammalian forebrain have crucial roles in locomotive and cognitive functions and thus, understanding and manipulation of cholinergic cell fate specification may be beneficial to identify therapeutic targets and methods for neurodiseases resulting from cholinergic neuronal dysfunction.

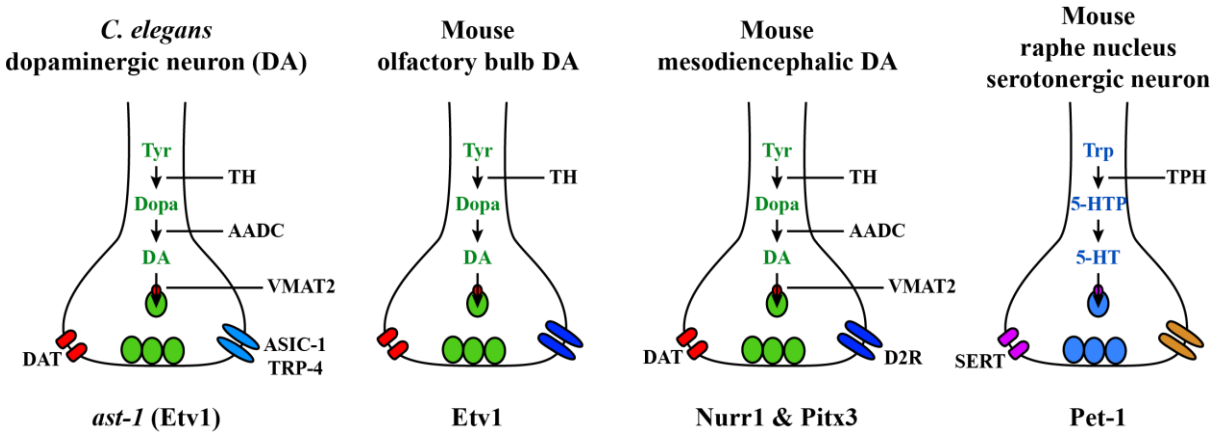


Figure 37. Coregulation and terminal selectors in Vertebrates. The essential transcription factors are shown under each neuron. The proteins whose genes are known to be regulated by the essential transcription factors are indicated. AADC, amino-acid decarboxylase; D2R, dopamine receptor D2; DAT, dopamine transporter; TH, tyrosine hydroxylase; TPH, tryptophan hydroxylase; VMAT2, vesicular monoamine transporter. ASIC-1, TRP-4, DAT and SERT are membrane transport or channel proteins. The neutrotransmitter-synthesis pathway is indicated in red inside each nerve terminal. Dopamine (DA) is synthesized from tyrosine (Tyr) via the intermediate Dopa. Serotonin (5-HT) is synthesized from tryptophan (Trp) via the intermediate 5-hydroxytryptophan (5-HTP).

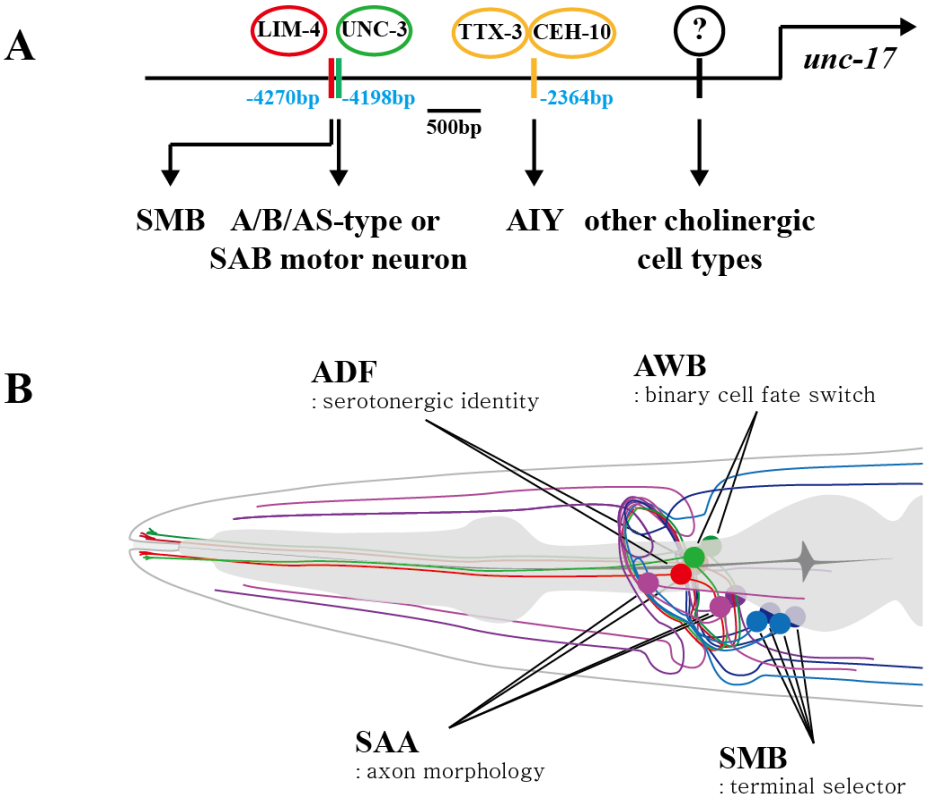


Figure 38. Model for the *cis*-regulatory architecture underlying cholinergic cell fate specification in *C. elegans*, and the roles of LIM-4 in neuronal development. (A) The *unc-17* (VChAT) promoter region is systemically organized with distinct *cis*-regulatory DNA sequences for cell-type specific *tans*-acting factors, such as the SMB motif for LIM-4 in the SMB neurons, COE motif for UNC-3 in the A/B/AS-type or SAB motor neurons and the AIY motif for TTX-3/CHE-10 in the AIY interneurons. Additional terminal selector genes and their target *cis*-regulatory elements for remaining uncharacterized cholinergic cell-types need to be identified. (B) LIM-4 has distinct roles in neuronal development in a context-dependent manner.

Table 8. The expression of the *flp-12* or *lim-4* in the SMB neurons is not altered in *fax-1* or *cog-1* mutant animals

Reporter construct	Genotype	% animals showing GFP expression in SMB		
		No	Weak	Strong
*Ex[flp-12*pΔ*1::gfp]*	WT	0	0	100
	fax-1 (gm83)	0	0	100
	cog-1 (sy275)	0	0	100
*lim-4*p*::gfp(oyIs35)*	WT	0	0	100
	fax-1 (gm83)	0	0	100
	cog-1 (sy275)	0	0	100

V. References

1. Guillemot, F., *Spatial and temporal specification of neural fates by transcription factor codes.* Development, 2007. **134**(21): p. 3771-80.
2. Hobert, O., *Regulation of terminal differentiation programs in the nervous system.* Annu Rev Cell Dev Biol, 2011. **27**: p. 681-96.
3. Deneris, E.S. and O. Hobert, *Maintenance of postmitotic neuronal cell identity.* Nat Neurosci, 2014. **17**(7): p. 899-907.
4. Hobert, O., *Neurogenesis in the nematode Caenorhabditis elegans.* WormBook, 2010: p. 1-24.
5. Sulston, J.E., *Neuronal cell lineages in the nematode Caenorhabditis elegans.* Cold Spring Harb Symp Quant Biol, 1983. **48 Pt 2**: p. 443-52.
6. White, J.G., et al., *The structure of the nervous system of the nematode Caenorhabditis elegans.* Philos Trans R Soc Lond B Biol Sci, 1986. **314**(1165): p. 1-340.
7. Chalfie, M., et al., *Green fluorescent protein as a marker for gene expression.* Science, 1994. **263**(5148): p. 802-5.
8. Hedgecock, L.W., *Resistance-enhancing compound from Mycobacterium tuberculosis.* Am Rev Respir Dis, 1975. **112**(5): p. 633-8.
9. Chalfie, M. and J. Sulston, *Developmental genetics of the mechanosensory neurons of Caenorhabditis elegans.* Dev Biol, 1981. **82**(2): p. 358-70.
10. Sengupta, P., H.A. Colbert, and C.I. Bargmann, *The C. elegans gene odr-7 encodes an olfactory-specific member of the nuclear receptor superfamily.* Cell, 1994. **79**(6): p. 971-80.
11. Lewis, J.A. and J.A. Hodgkin, *Specific neuroanatomical changes in chemosensory mutants of the nematode Caenorhabditis elegans.* J Comp Neurol, 1977. **172**(3): p. 489-510.
12. Baran, R., R. Aronoff, and G. Garriga, *The C. elegans homeodomain gene unc-42 regulates chemosensory and glutamate receptor expression.* Development, 1999. **126**(10): p. 2241-51.
13. Finney, M., G. Ruvkun, and H.R. Horvitz, *The C. elegans cell lineage and differentiation gene unc-86 encodes a protein with a homeodomain and extended similarity to transcription factors.* Cell, 1988. **55**(5): p. 757-69.
14. Hobert, O., et al., *Regulation of interneuron function in the C. elegans thermoregulatory pathway by the ttx-3 LIM homeobox gene.* Neuron, 1997. **19**(2): p. 345-57.
15. Jin, Y., R. Hoskins, and H.R. Horvitz, *Control of type-D GABAergic neuron differentiation by C. elegans UNC-30 homeodomain protein.* Nature, 1994. **372**(6508): p. 780-3.
16. Prasad, B.C., et al., *unc-3, a gene required for axonal guidance in Caenorhabditis elegans, encodes a member of the O/E family of transcription factors.* Development, 1998. **125**(8): p. 1561-8.
17. Satterlee, J.S., et al., *Specification of thermosensory neuron fate in C. elegans requires ttx-1, a homolog of otd/Otx.* Neuron, 2001. **31**(6): p. 943-56.
18. Sengupta, A. and S.K. Guha, *Multifactorial interactions in the aetiopathogenesis of EPH-Gestosis--a hypothesis.* Med Hypotheses, 1994. **43**(5): p. 322-6.
19. Way, J.C. and M. Chalfie, *mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in C. elegans.* Cell, 1988. **54**(1): p. 5-16.

20. Uchida, O., et al., *The C. elegans che-1 gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons.* Development, 2003. **130**(7): p. 1215-24.
21. Hobert, O., *Regulatory logic of neuronal diversity: terminal selector genes and selector motifs.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20067-71.
22. Sulston, J.E., et al., *The embryonic cell lineage of the nematode Caenorhabditis elegans.* Dev Biol, 1983. **100**(1): p. 64-119.
23. Kimble, J. and D. Hirsh, *The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans.* Dev Biol, 1979. **70**(2): p. 396-417.
24. Sulston, J.E. and H.R. Horvitz, *Post-embryonic cell lineages of the nematode, Caenorhabditis elegans.* Dev Biol, 1977. **56**(1): p. 110-56.
25. Sawin, E.R., R. Ranganathan, and H.R. Horvitz, *C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway.* Neuron, 2000. **26**(3): p. 619-31.
26. Hills, T., P.J. Brockie, and A.V. Maricq, *Dopamine and glutamate control area-restricted search behavior in Caenorhabditis elegans.* J Neurosci, 2004. **24**(5): p. 1217-25.
27. Bargmann, C.I. and H.R. Horvitz, *Control of larval development by chemosensory neurons in Caenorhabditis elegans.* Science, 1991. **251**(4998): p. 1243-6.
28. Troemel, E.R., et al., *Divergent seven transmembrane receptors are candidate chemosensory receptors in C. elegans.* Cell, 1995. **83**(2): p. 207-18.
29. Mori, I. and Y. Ohshima, *Neural regulation of thermotaxis in Caenorhabditis elegans.* Nature, 1995. **376**(6538): p. 344-8.
30. Van Buskirk, C. and P.W. Sternberg, *Epidermal growth factor signaling induces behavioral quiescence in Caenorhabditis elegans.* Nat Neurosci, 2007. **10**(10): p. 1300-7.
31. Sanders, J., et al., *The Caenorhabditis elegans interneuron ALA is (also) a high-threshold mechanosensor.* BMC Neurosci, 2013. **14**: p. 156.
32. Coates, J.C. and M. de Bono, *Antagonistic pathways in neurons exposed to body fluid regulate social feeding in Caenorhabditis elegans.* Nature, 2002. **419**(6910): p. 925-9.
33. Zimmer, M., et al., *Neurons detect increases and decreases in oxygen levels using distinct guanylate cyclases.* Neuron, 2009. **61**(6): p. 865-79.
34. Gray, J.M., et al., *Oxygen sensation and social feeding mediated by a C. elegans guanylate cyclase homologue.* Nature, 2004. **430**(6997): p. 317-22.
35. Bargmann, C.I., E. Hartwieg, and H.R. Horvitz, *Odorant-selective genes and neurons mediate olfaction in C. elegans.* Cell, 1993. **74**(3): p. 515-27.
36. Kaplan, J.M. and H.R. Horvitz, *A dual mechanosensory and chemosensory neuron in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1993. **90**(6): p. 2227-31.
37. Beverly, M., S. Anbil, and P. Sengupta, *Degeneracy and neuromodulation among thermosensory neurons contribute to robust thermosensory behaviors in Caenorhabditis elegans.* J Neurosci, 2011. **31**(32): p. 11718-27.
38. Biron, D., et al., *An olfactory neuron responds stochastically to temperature and modulates Caenorhabditis elegans thermotactic behavior.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 11002-7.
39. Hallem, E.A. and P.W. Sternberg, *Acute carbon dioxide avoidance in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 2008. **105**(23): p. 8038-43.
40. Hart, A.C., S. Sims, and J.M. Kaplan, *Synaptic code for sensory modalities revealed by C. elegans GLR-1 glutamate receptor.* Nature, 1995. **378**(6552): p. 82-5.

41. Lee, H., et al., *Nictation, a dispersal behavior of the nematode Caenorhabditis elegans, is regulated by IL2 neurons.* Nat Neurosci, 2011. **15**(1): p. 107-12.
42. Chang, H.C., J. Paek, and D.H. Kim, *Natural polymorphisms in C. elegans HECW-1 E3 ligase affect pathogen avoidance behaviour.* Nature, 2011. **480**(7378): p. 525-9.
43. Li, W., et al., *The neural circuits and sensory channels mediating harsh touch sensation in Caenorhabditis elegans.* Nat Commun, 2011. **2**: p. 315.
44. Hilliard, M.A., C.I. Bargmann, and P. Bazzicalupo, *C. elegans responds to chemical repellents by integrating sensory inputs from the head and the tail.* Curr Biol, 2002. **12**(9): p. 730-4.
45. Liu, S., E. Schulze, and R. Baumeister, *Temperature- and touch-sensitive neurons couple CNG and TRPV channel activities to control heat avoidance in Caenorhabditis elegans.* PLoS One, 2012. **7**(3): p. e32360.
46. Chatzigeorgiou, M., et al., *Specific roles for DEG/ENaC and TRP channels in touch and thermosensation in C. elegans nociceptors.* Nat Neurosci, 2010. **13**(7): p. 861-8.
47. Barrios, A., et al., *PDF-1 neuropeptide signaling modulates a neural circuit for mate-searching behavior in C. elegans.* Nat Neurosci, 2012. **15**(12): p. 1675-82.
48. Pradel, E., et al., *Detection and avoidance of a natural product from the pathogenic bacterium Serratia marcescens by Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2295-300.
49. Chalasani, S.H., et al., *Dissecting a circuit for olfactory behaviour in Caenorhabditis elegans.* Nature, 2007. **450**(7166): p. 63-70.
50. Gray, J.M., J.J. Hill, and C.I. Bargmann, *A circuit for navigation in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3184-91.
51. Bendena, W.G., et al., *A Caenorhabditis elegans allatostatin/galanin-like receptor NPR-9 inhibits local search behavior in response to feeding cues.* Proc Natl Acad Sci U S A, 2008. **105**(4): p. 1339-42.
52. Chalfie, M., et al., *The neural circuit for touch sensitivity in Caenorhabditis elegans.* J Neurosci, 1985. **5**(4): p. 956-64.
53. Tsalik, E.L. and O. Hobert, *Functional mapping of neurons that control locomotory behavior in Caenorhabditis elegans.* J Neurobiol, 2003. **56**(2): p. 178-97.
54. Piggott, B.J., et al., *The neural circuits and synaptic mechanisms underlying motor initiation in C. elegans.* Cell, 2011. **147**(4): p. 922-33.
55. Hardaker, L.A., et al., *Serotonin modulates locomotory behavior and coordinates egg-laying and movement in Caenorhabditis elegans.* J Neurobiol, 2001. **49**(4): p. 303-13.
56. Hutter, H., *Extracellular cues and pioneers act together to guide axons in the ventral cord of C. elegans.* Development, 2003. **130**(22): p. 5307-18.
57. Li, W., et al., *A C. elegans stretch receptor neuron revealed by a mechanosensitive TRP channel homologue.* Nature, 2006. **440**(7084): p. 684-7.
58. Wadsworth, W.G., H. Bhatt, and E.M. Hedgecock, *Neuroglia and pioneer neurons express UNC-6 to provide global and local netrin cues for guiding migrations in C. elegans.* Neuron, 1996. **16**(1): p. 35-46.
59. Aurelio, O., D.H. Hall, and O. Hobert, *Immunoglobulin-domain proteins required for maintenance of ventral nerve cord organization.* Science, 2002. **295**(5555): p. 686-90.
60. Nelson, M.D., et al., *The neuropeptide NLP-22 regulates a sleep-like state in Caenorhabditis elegans.* Nat Commun, 2013. **4**: p. 2846.
61. Greer, E.R., et al., *Neural and molecular dissection of a C. elegans sensory circuit that regulates fat and feeding.* Cell Metab, 2008. **8**(2): p. 118-31.
62. Flavell, S.W., et al., *Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in C. elegans.* Cell, 2013. **154**(5): p. 1023-1035.

63. Chatzigeorgiou, M. and W.R. Schafer, *Lateral facilitation between primary mechanosensory neurons controls nose touch perception in C. elegans.* Neuron, 2011. **70**(2): p. 299-309.
64. Turek, M., I. Lewandrowski, and H. Bringmann, *An AP2 transcription factor is required for a sleep-active neuron to induce sleep-like quiescence in C. elegans.* Curr Biol, 2013. **23**(22): p. 2215-23.
65. McIntire, S.L., et al., *The GABAergic nervous system of Caenorhabditis elegans.* Nature, 1993. **364**(6435): p. 337-41.
66. Macosko, E.Z., et al., *A hub-and-spoke circuit drives pheromone attraction and social behaviour in C. elegans.* Nature, 2009. **458**(7242): p. 1171-5.
67. Kratsios, P., et al., *Transcriptional coordination of synaptogenesis and neurotransmitter signaling.* Curr Biol, 2015. **25**(10): p. 1282-95.
68. Kennerdell, J.R., R.D. Fetter, and C.I. Bargmann, *Wnt-Ror signaling to SIA and SIB neurons directs anterior axon guidance and nerve ring placement in C. elegans.* Development, 2009. **136**(22): p. 3801-10.
69. Waggoner, L.E., et al., *Control of alternative behavioral states by serotonin in Caenorhabditis elegans.* Neuron, 1998. **21**(1): p. 203-14.
70. Brenner, S., *The genetics of Caenorhabditis elegans.* Genetics, 1974. **77**(1): p. 71-94.
71. Kratsios, P., et al., *Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene.* Nat Neurosci, 2011. **15**(2): p. 205-14.
72. Pereira, L., et al., *A cellular and regulatory map of the cholinergic nervous system of C. elegans.* Elife, 2015. **4**.
73. Serrano-Saiz, E., et al., *Modular control of glutamatergic neuronal identity in C. elegans by distinct homeodomain proteins.* Cell, 2013. **155**(3): p. 659-73.
74. Duggan, A., C. Ma, and M. Chalfie, *Regulation of touch receptor differentiation by the Caenorhabditis elegans mec-3 and unc-86 genes.* Development, 1998. **125**(20): p. 4107-19.
75. Desai, C., et al., *A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons.* Nature, 1988. **336**(6200): p. 638-46.
76. Zhang, F., et al., *The LIM and POU homeobox genes ttx-3 and unc-86 act as terminal selectors in distinct cholinergic and serotonergic neuron types.* Development, 2014. **141**(2): p. 422-35.
77. Zhang, Y., et al., *Identification of genes expressed in C. elegans touch receptor neurons.* Nature, 2002. **418**(6895): p. 331-5.
78. Etchberger, J.F., et al., *The molecular signature and cis-regulatory architecture of a C. elegans gustatory neuron.* Genes Dev, 2007. **21**(13): p. 1653-74.
79. Hedgecock, E.M. and R.L. Russell, *Normal and mutant thermotaxis in the nematode Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1975. **72**(10): p. 4061-5.
80. Altun-Gultekin, Z., et al., *A regulatory cascade of three homeobox genes, ceh-10, ttx-3 and ceh-23, controls cell fate specification of a defined interneuron class in C. elegans.* Development, 2001. **128**(11): p. 1951-69.
81. Eastman, C., H.R. Horvitz, and Y. Jin, *Coordinated transcriptional regulation of the unc-25 glutamic acid decarboxylase and the unc-47 GABA vesicular transporter by the Caenorhabditis elegans UNC-30 homeodomain protein.* J Neurosci, 1999. **19**(15): p. 6225-34.
82. Flames, N. and O. Hobert, *Gene regulatory logic of dopamine neuron differentiation.* Nature, 2009. **458**(7240): p. 885-9.
83. Jessell, T.M., *Neuronal specification in the spinal cord: inductive signals and transcriptional codes.* Nat Rev Genet, 2000. **1**(1): p. 20-9.

84. Wenick, A.S. and O. Hobert, *Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans.* Dev Cell, 2004. **6**(6): p. 757-70.
85. Etchberger, J.F. and O. Hobert, *Vector-free DNA constructs improve transgene expression in C. elegans.* Nat Methods, 2008. **5**(1): p. 3.
86. Doitsidou, M., et al., *A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in C. elegans.* Genes Dev, 2013. **27**(12): p. 1391-405.
87. Etchberger, J.F., et al., *Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in C. elegans.* Development, 2009. **136**(1): p. 147-60.
88. Xue, D., et al., *Regulation of the mec-3 gene by the C.elegans homeoproteins UNC-86 and MEC-3.* EMBO J, 1992. **11**(13): p. 4969-79.
89. Ptashne, M., *The chemistry of regulation of genes and other things.* J Biol Chem, 2014. **289**(9): p. 5417-35.
90. Sarafi-Reinach, T.R., et al., *The lin-11 LIM homeobox gene specifies olfactory and chemosensory neuron fates in C. elegans.* Development, 2001. **128**(17): p. 3269-81.
91. Bertrand, V. and O. Hobert, *Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in C. elegans.* Dev Cell, 2009. **16**(4): p. 563-75.
92. Craven, S.E., et al., *Gata2 specifies serotonergic neurons downstream of sonic hedgehog.* Development, 2004. **131**(5): p. 1165-73.
93. Krueger, K.C. and E.S. Deneris, *Serotonergic transcription of human FEV reveals direct GATA factor interactions and fate of Pet-1-deficient serotonin neuron precursors.* J Neurosci, 2008. **28**(48): p. 12748-58.
94. Liu, C., et al., *Pet-1 is required across different stages of life to regulate serotonergic function.* Nat Neurosci, 2010. **13**(10): p. 1190-8.
95. Swoboda, P., H.T. Adler, and J.H. Thomas, *The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans.* Mol Cell, 2000. **5**(3): p. 411-21.
96. Desai, P.N., *Medical ethics in India.* J Med Philos, 1988. **13**(3): p. 231-55.
97. Enriquez, J., et al., *Specification of individual adult motor neuron morphologies by combinatorial transcription factor codes.* Neuron, 2015. **86**(4): p. 955-970.
98. Li, C. and K. Kim, *Neuropeptides.* WormBook, 2008: p. 1-36.
99. Duerr, J.S., et al., *Identification of major classes of cholinergic neurons in the nematode Caenorhabditis elegans.* J Comp Neurol, 2008. **506**(3): p. 398-408.
100. Kim, K. and C. Li, *Expression and regulation of an FMRFamide-related neuropeptide gene family in Caenorhabditis elegans.* J Comp Neurol, 2004. **475**(4): p. 540-50.
101. Fire, A., S.W. Harrison, and D. Dixon, *A modular set of lacZ fusion vectors for studying gene expression in Caenorhabditis elegans.* Gene, 1990. **93**(2): p. 189-98.
102. Sagasti, A., et al., *Alternative olfactory neuron fates are specified by the LIM homeobox gene lim-4.* Genes Dev, 1999. **13**(14): p. 1794-806.
103. Chao, M.Y., et al., *Feeding status and serotonin rapidly and reversibly modulate a Caenorhabditis elegans chemosensory circuit.* Proc Natl Acad Sci U S A, 2004. **101**(43): p. 15512-7.
104. Dawid, I.B., R. Toyama, and M. Taira, *LIM domain proteins.* C R Acad Sci III, 1995. **318**(3): p. 295-306.
105. Hedgecock, E.M. and J.G. White, *Polyploid tissues in the nematode Caenorhabditis elegans.* Dev Biol, 1985. **107**(1): p. 128-33.

106. Zheng, X., et al., *Cell-type specific regulation of serotonergic identity by the C. elegans LIM-homeodomain factor LIM-4.* Dev Biol, 2005. **286**(2): p. 618-28.
107. Chou, J.H., C.I. Bargmann, and P. Sengupta, *The Caenorhabditis elegans odr-2 gene encodes a novel Ly-6-related protein required for olfaction.* Genetics, 2001. **157**(1): p. 211-24.
108. Colbert, H.A., T.L. Smith, and C.I. Bargmann, *OSM-9, a novel protein with structural similarity to channels, is required for olfaction, mechanosensation, and olfactory adaptation in Caenorhabditis elegans.* J Neurosci, 1997. **17**(21): p. 8259-69.
109. Alfonso, A., et al., *The Caenorhabditis elegans unc-17 gene: a putative vesicular acetylcholine transporter.* Science, 1993. **261**(5121): p. 617-9.
110. Okuda, T., et al., *Identification and characterization of the high-affinity choline transporter.* Nat Neurosci, 2000. **3**(2): p. 120-5.
111. Pujol, N., et al., *The homeodomain protein CePHOX2/CEH-17 controls antero-posterior axonal growth in C. elegans.* Development, 2000. **127**(15): p. 3361-71.
112. Harfe, B.D. and A. Fire, *Muscle and nerve-specific regulation of a novel NK-2 class homeodomain factor in Caenorhabditis elegans.* Development, 1998. **125**(3): p. 421-9.
113. Lanjuin, A., et al., *Otx/otd homeobox genes specify distinct sensory neuron identities in C. elegans.* Dev Cell, 2003. **5**(4): p. 621-33.
114. Chang, A.J., et al., *A distributed chemosensory circuit for oxygen preference in C. elegans.* PLoS Biol, 2006. **4**(9): p. e274.
115. Kim, K., R. Kim, and P. Sengupta, *The HMX/NKX homeodomain protein MLS-2 specifies the identity of the AWC sensory neuron type via regulation of the ceh-36 Otx gene in C. elegans.* Development, 2010. **137**(6): p. 963-74.
116. L'Etoile, N.D. and C.I. Bargmann, *Olfaction and odor discrimination are mediated by the C. elegans guanylyl cyclase ODR-1.* Neuron, 2000. **25**(3): p. 575-86.
117. Burglin, T.R., *Homeodomain subtypes and functional diversity.* Subcell Biochem, 2011. **52**: p. 95-122.
118. Noyes, M.B., et al., *Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites.* Cell, 2008. **133**(7): p. 1277-89.
119. Christensen, R.G., et al., *Recognition models to predict DNA-binding specificities of homeodomain proteins.* Bioinformatics, 2012. **28**(12): p. i84-9.
120. Flandin, P., et al., *Lhx6 and Lhx8 coordinately induce neuronal expression of Shh that controls the generation of interneuron progenitors.* Neuron, 2011. **70**(5): p. 939-50.
121. Nokes, E.B., et al., *Cis-regulatory mechanisms of gene expression in an olfactory neuron type in Caenorhabditis elegans.* Dev Dyn, 2009. **238**(12): p. 3080-92.
122. Cave, J.W., et al., *Differential regulation of dopaminergic gene expression by Er81.* J Neurosci, 2010. **30**(13): p. 4717-24.
123. Smidt, M.P. and J.P. Burbach, *A passport to neurotransmitter identity.* Genome Biol, 2009. **10**(7): p. 229.
124. Jacobs, F.M., et al., *Retinoic acid counteracts developmental defects in the substantia nigra caused by Pitx3 deficiency.* Development, 2007. **134**(14): p. 2673-84.
125. Jacobs, F.M., et al., *Identification of Dlk1, Ptpru and Klhl1 as novel Nurr1 target genes in meso-diencephalic dopamine neurons.* Development, 2009. **136**(14): p. 2363-73.
126. Jacobs, F.M., et al., *Pitx3 potentiates Nurr1 in dopamine neuron terminal differentiation through release of SMRT-mediated repression.* Development, 2009. **136**(4): p. 531-40.

127. Martinat, C., et al., *Cooperative transcription activation by Nurr1 and Pitx3 induces embryonic stem cell maturation to the midbrain dopamine neuron phenotype.* Proc Natl Acad Sci U S A, 2006. **103**(8): p. 2874-9.
128. Hendricks, T., et al., *The ETS domain factor Pet-1 is an early and precise marker of central serotonin neurons and interacts with a conserved element in serotonergic genes.* J Neurosci, 1999. **19**(23): p. 10348-56.
129. Hendricks, T.J., et al., *Pet-1 ETS gene plays a critical role in 5-HT neuron development and is required for normal anxiety-like and aggressive behavior.* Neuron, 2003. **37**(2): p. 233-47.
130. Wightman, B., et al., *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.* Dev Biol, 2005. **287**(1): p. 74-85.
131. Palmer, R.E., et al., *Caenorhabditis elegans cog-1 locus encodes GTX/Nkx6.1 homeodomain proteins and regulates multiple aspects of reproductive system development.* Dev Biol, 2002. **252**(2): p. 202-13.
132. Sarin, S., et al., *The C. elegans Tailless/TLX transcription factor nhr-67 controls neuronal identity and left/right asymmetric fate diversification.* Development, 2009. **136**(17): p. 2933-44.
133. Holland, P.W. and T. Takahashi, *The evolution of homeobox genes: Implications for the study of brain development.* Brain Res Bull, 2005. **66**(4-6): p. 484-90.
134. Hobert, O. and H. Westphal, *Functions of LIM-homeobox genes.* Trends Genet, 2000. **16**(2): p. 75-83.
135. Srivastava, M., et al., *Early evolution of the LIM homeobox gene family.* BMC Biol, 2010. **8**: p. 4.
136. Fragkouli, A., et al., *LIM homeodomain transcription factor-dependent specification of bipotential MGE progenitors into cholinergic and GABAergic striatal interneurons.* Development, 2009. **136**(22): p. 3841-51.
137. Cho, H.H., et al., *Isl1 directly controls a cholinergic neuronal identity in the developing forebrain and spinal cord by forming cell type-specific complexes.* PLoS Genet, 2014. **10**(4): p. e1004280.
138. Zhao, Y., et al., *The LIM-homeobox gene Lhx8 is required for the development of many cholinergic neurons in the mouse forebrain.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 9005-10.
139. Vogt, D., et al., *Lhx6 directly regulates Arx and CXCR7 to determine cortical interneuron fate and laminar position.* Neuron, 2014. **82**(2): p. 350-64.
140. Mori, T., et al., *The LIM homeobox gene, L3/Lhx8, is necessary for proper development of basal forebrain cholinergic neurons.* Eur J Neurosci, 2004. **19**(12): p. 3129-41.
141. Zhu, P., et al., *LIM-homeobox gene Lhx8 promote the differentiation of hippocampal newborn neurons into cholinergic neurons in vitro.* In Vitro Cell Dev Biol Anim, 2013. **49**(2): p. 103-7.
142. Shi, J., et al., *Lhx8 promote differentiation of hippocampal neural stem/progenitor cells into cholinergic neurons in vitro.* In Vitro Cell Dev Biol Anim, 2012. **48**(10): p. 603-9.
143. Li, H., et al., *Upregulation of Lhx8 increase VAChT expression and ACh release in neuronal cell line SHSY5Y.* Neurosci Lett, 2014. **559**: p. 184-8.
144. Lopes, F.M., et al., *Comparison between proliferative and neuron-like SH-SY5Y cells as an in vitro model for Parkinson disease studies.* Brain Res, 2010. **1337**: p. 85-94.

145. Biedler, J.L., et al., *Multiple neurotransmitter synthesis by human neuroblastoma cell lines and clones.* Cancer Res, 1978. **38**(11 Pt 1): p. 3751-7.
146. Kovalevich, J. and D. Langford, *Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology.* Methods Mol Biol, 2013. **1078**: p. 9-21.
147. Danks, K., et al., *Redistribution of F-actin and large dense-cored vesicles in the human neuroblastoma SH-SY5Y in response to secretagogues and protein kinase Calpha activation.* Brain Res Mol Brain Res, 1999. **64**(2): p. 236-45.

VI. Summary in Koreans

유사 분열 후의 신경 (post-mitotic neuron)들은 특정 전사 인자 (transcription factor)들의 발현에 따라 분화된 후의 특징들이 나타나게 된다. 그러나 신경 세포의 운명 (neuronal cell fate)을 결정하는데 관여하는 특정한 분자들과 이와 관련하여 진화 과정에서 보존된 전사인자들의 기능들에 대해 아직 잘 알려지지 않았다. 예쁜꼬마선충 (*C. elegans*)의 경우, 사분면에 걸쳐 발달된 머리 근육에 콜린성 (cholinergic)이자 펩티드성 (pepetidergic)인 감각/ 개재/ 운동신경 (sensory/inter/motor neuron)인 SMB 의 신경이 연결되어 있어서 이를 통해 사인 곡선 모양의 움직임의 폭 (wave width)을 조절한다. 이 논문에서는 SMB 가 신경으로 제대로 분화되고 신경세포로써의 특징을 가지게 하는데 LIM 호메오박스 (Homeobox) 단백질인 LIM-4 가 중요하다는 것을 밝혀냈다. *lim-4* 돌연변이에서 SMB 의 단말 분화 유전자들인 콜린성 유전자 배터리 (cholinergic gene battery)와 신경펩티드 유전자인 *flp-12* 의 발현이 억제되었으며, 또한 SMB 의 기능도 제대로 작동되지 않는 것을 확인하였다. LIM-4 단백질의 활성 정도에 따라 SMB 에 발현하는 유전자들의 발현이 조절된다는 것을 알아냈으며, 이는 *cis*-작용 부분 (*cis*-regulatory motif)을 통해 이루어지는 것을 밝혀냈다. 사람의 경우, LIM-4 의 상동 유전자 (orthologue)인 LHX6 와 LHX8 를 가지고 있으며, 예쁜꼬마선충에서 LIM-4 단백질을 대신하여 기능적으로 대체할 수 있는 것을 확인하였다. 이와 더불어, 예쁜꼬마선충 LIM-4 단백질과 인간 LHX6 단백질이 human 신경 세포들을 콜린성과 펩티드성 특징을 가진 신경으로 분화를 유도할 수 있다는 것을 밝혀냈다. 이 논문에서는 특정 신경으로 분화하기 위한

LIM-4/LHX6 호메오박스 단백질의 정밀한 조절 기능이 진화적으로 보존이 되었다는 것을 보여주었다.

핵심어 : *lim-4*, SMB, 전사 인자, 단말 선택 유전자, 예쁜꼬마선충

VII. Appendix

7.1. Overexpression of human LHX8 or *C. elegans* LIM-4 in the human neuroblastoma SH-SY5Y cells. (This experiment is performed by Jihye Yeon and Seong-Kyoon Choi)

LHX8 has been shown to be required for the development and maintenance of cholinergic neurons in mouse basal forebrain [138, 140]. Overexpression of LHX8 was sufficient to differentiate rat hippocampal neural stem cells or newborn neurons into cholinergic neuron types [141, 142] and induced expression of cholinergic markers in human neuroblastoma cell line [143]. Whether LHX6 has a similar role in specification of cholinergic cell fate has not been explored. Thus, we tested whether overexpression of human LHX6 or *C. elegans* LIM-4 could promote expression of cholinergic markers in the human neuroblastoma SH-SY5Y cells. These cells appear to mimic immature catecholaminergic neurons when untreated [144, 145], but can differentiate into various mature neuron-like phenotypes depending on the addition of differentiation-inducing agents [146]. We generated stable SH-SY5Y cell lines expressing either *LHX6-GFP*, *lim-4-GFP*, or empty vehicle and asked whether transfected cell lines express cholinergic markers. Expression of either *LHX6-GFP* or *lim-4-GFP* was detected predominantly in cell nuclei, supporting the action of LHX-6 and LIM-4 as transcription factors (Figures 39 and 40). Transfected cells were immunoreactive to VAChT or choline acetyltransferase (ChAT) antibodies and exhibited higher endogenous ChAT message levels, as assayed by quantitative reverse transcription polymerase chain reaction (RT-PCR), compared to that in cells transfected with an empty vehicle (Figures 39 and 40). Thus, human *LHX6* and even *C. elegans lim-4* are sufficient to promote expression of cholinergic markers in human cells. We also noted that cells expressing either *LHX6-GFP* or *lim-4-GFP* were morphologically different to cells expressing empty vehicle: empty vehicle expressing cells tended to grow in clusters and were round shape; by contrast, *LHX6-GFP* or *lim-4-GFP* expressing cells formed less clusters and appeared as spiky neuronal cells (Figures 39 and 40).

These results, therefore, indicate that LHX6 or LIM-4 can induce differentiation of SH-SY5Y cells into cholinergic as well as neuronal phenotypes.

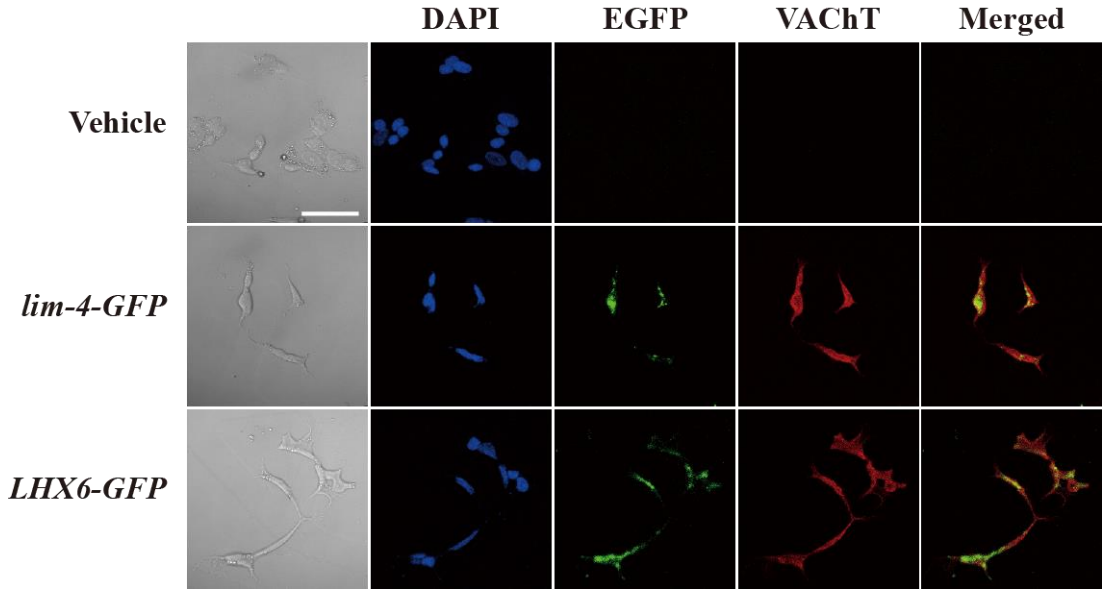


Figure 39. *C. elegans* LIM-4 or human LHX6 induces expression of cholinergic makers and neuronal characteristics in human neuroblastoma cells. Confocal images of SH-SY5Y human neuroblastoma cell line transfected by empty vector, *C. elegans lim-4* or human *LHX6* and immunostained with VAChT antibodies. Scale bar: 50 μm. Note that cells expressing either *LHX6-GFP* or *lim-4-GFP* exhibit spiky protrusions.

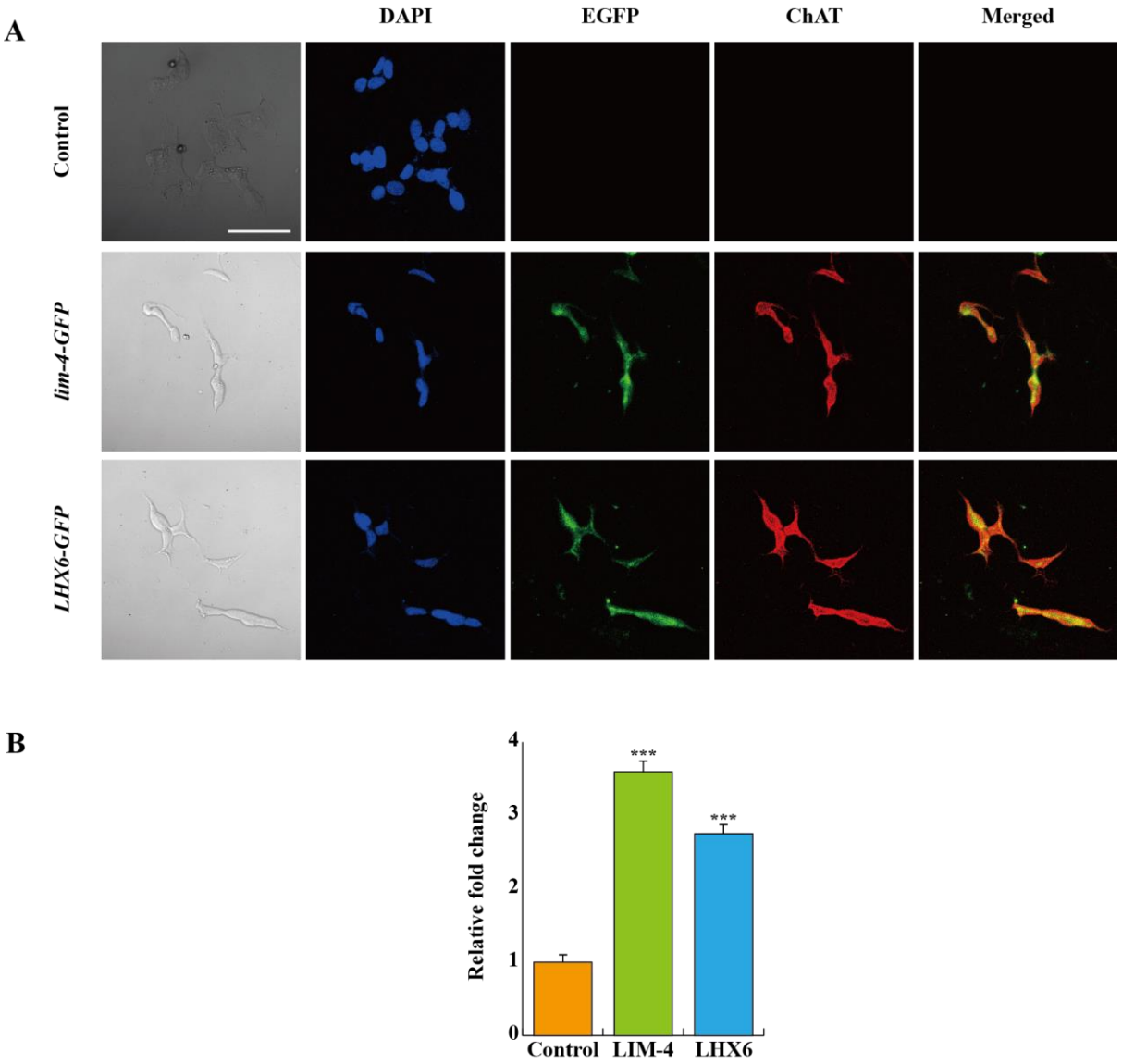


Figure 40. *C. elegans* LIM-4 or human LHX6 induces expression of ChAT and neuronal characteristics in human neuroblastoma cells. (A) Confocal images of SH-SY5Y human neuroblastoma cell line transfected by *C. elegans lim-4* or human *LHX6* and immunostained with ChAT antibodies. Scale bar: 50 μm. (B) Levels of ChAT transcripts are increased in SH-SY5Y human neuroblastoma cell line transfected by *C. elegans lim-4* or human *LHX6*. The relative fold change to housekeeping gene GAPDH is shown. *** indicates significantly different from control (untransfected cells) ($p<0.001$).

7.2. Ultrastructural analysis of LIM-4 or LHX6 transfected SH-SY5Y cell lines. (The images were obtained by Jihye Yeon and Prof. Huh Yang Hoon using electron microscopy)

We further examined the morphology of transfected cells by electron microscopy. Un-transfected or empty vesicle transfected SH-SY5Y cells exhibited typical shapes of mitochondria, endoplasmic reticulum (ER), and Golgi apparatus in the peri-nuclear region (Figure 41). In either *lim-4* or *LHX6* transfected cells, however, we observed additional ultrastructural components, such as 100nm large-dense core vesicles, which may contain neuropeptides, near the Golgi apparatus (Figure 41) [147]. Furthermore, in the spiky protruded region of *lim-4* or *LHX6* transfected cells, we also identified mitochondria and small synaptic vesicles that may represent axon terminals of neurons (Figure 41). These results further support that expression of LIM-4 or LHX6 may produce synaptic vesicles containing large neuropeptides and small molecule neurotransmitters in human cell lines.

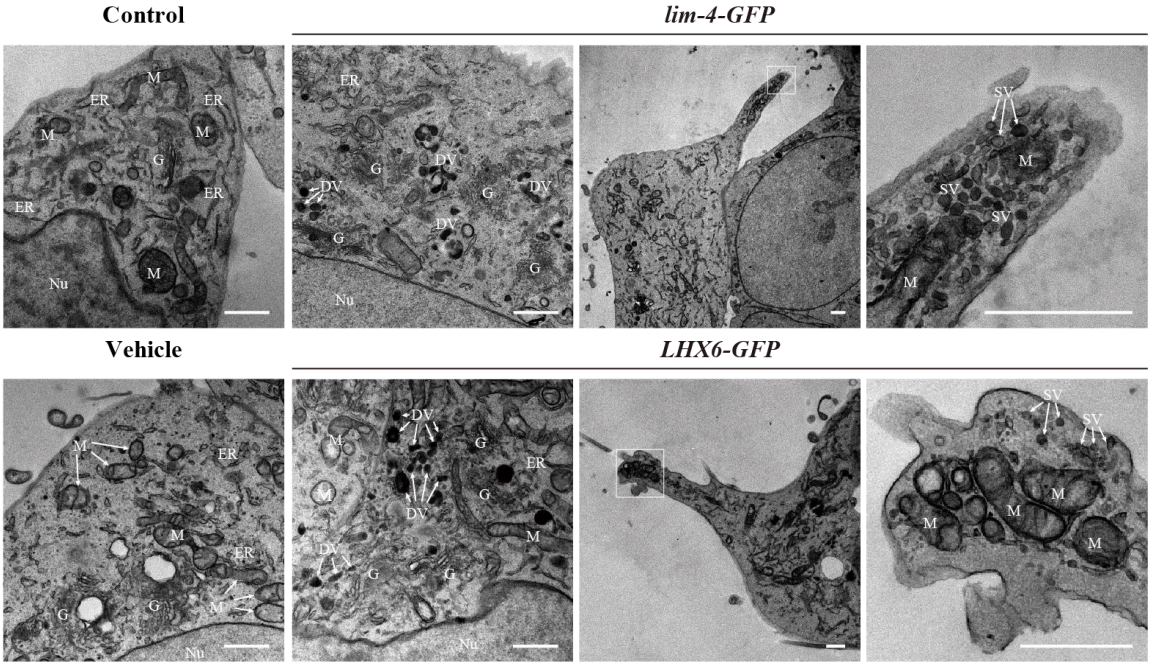


Figure 41. *C. elegans* LIM-4 or human LHX6 induces neuronal characteristics in human neuroblastoma cells. Ultrastructural analysis of LIM-4 or LHX6 transfected SH-SY5Y cell lines. Images from transmission electron microscope for untransfected (control), empty vector, *lim-4-GFP,* or *LHX6-GFP* transfected cells are shown. Left two columns are from peri-nuclear region and right two columns are from protruded region (far right images are higher magnification of the boxed area). Nu: nucleus, G: Golgi apparatus, M: mitochondria, ER: endoplasmic reticulum, DV: dense core vesicle, SV: synaptic vesicle. Scale bars: 1 μm. Confocal images of these control and transfected SH-SY5Y cells grown on the MatTek culture dish are shown in Figure 42.

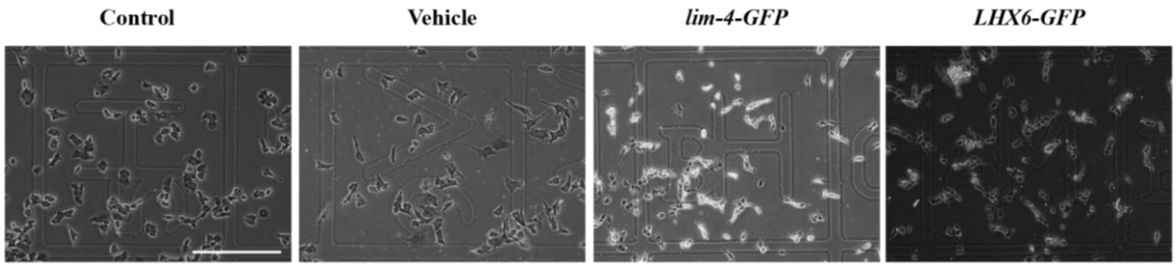


Figure 42. Confocal images of untransfected (control) and vehicle, *lim-4* or *LHX6* transfected SH-SY5Y cells grown on the MatTek culture dish are shown. Scale bar: 500 μm