5-30-2018

Mapping Out Head Country: Expression of DAF-19C in the Head Neurons of C. elegans

Hanwenheng (Billy) Liu
lhwh1208@hotmail.com

Mapping Out Head Country: Expression of DAF-19C in the Head Neurons of *C. elegans*

Hanwenheng (Billy) Liu

A Thesis Submitted in Candidacy for Honors at Graduation from Lawrence University
May 2018

I hereby reaffirm the Lawrence University Honor Code

Abstract

C. elegans has been used as a model system for studies of nervous system development and function for over 50 years due to its relative simplicity and strong similarity to other animals. One of such similarities is the presence of an RFX transcription factor, which is responsible for controlling gene expression in a handful of important organs, including the brain. RFX proteins are a family of highly evolutionarily conserved proteins that usually function by binding to a specific region (called the X-box) in the promoter of their target genes. A single RFX-coding gene has been found in *C. elegans*, called *daf-19*. The function of DAF-19 proteins thus far identified include regulating innate immunity and ciliogenesis. The *daf-19* gene encodes four related proteins, or isoforms: DAF-19A, B, C, and M. In this study, we fully characterize the expression pattern of the DAF-19C isoform through the use of fluorescently labeled protein constructs, confocal fluorescence microscopy, and neuron maps. We report that DAF-19C is expressed in at least 48 ciliated neurons and eight non-ciliated neurons in the head.

Table of Contents

INTRODUCTION

Caenorhabditis elegans

Caenorhabditis elegans is a nematode (round worm) that has been used as a model organism for almost 50 years. In the 1960s, Sydney Brenner, foreseeing a near end of "classical problems" of molecular biology, proposed a shift of biological research into developmental and neuronal biology (Brenner, 1963 & 1988). He initially proposed *Caenorhabditis briggsae* as a new, simpler model organism, because of its

> short life cycle, [easy cultivation], and [ability] to be handled in large numbers, like a micro-organism. It [has] relatively few cells, so that exhaustive studies of lineage and patterns can be made, and [is] amenable to genetic analysis." (Brenner 1963)

C. elegans was preferred over *C. briggsae* later for the former's higher survival rate in lab conditions and its ease of sectioning for electron microscopy. The worm has since been one of the standard model organisms by over 1,200 labs around the world today (Corsi et al., 2015; WormBase, 2018).

C. elegans populations consist mostly of self-fertilizing hermaphrodites. Males, having only a single X chromosome, occur in nature with a frequency of less than 0.2% (Corsi et al., 2015). Specific mutations, such as those in the gene *him-5*, can affect disjunction in meiosis and can thus increase the occurrence of males to 18% to 37% (Meneely et al., 2012). Heatshock may also cause nondisjunction (separation of homologous chromosomes during meiosis) and produce males, though at a lower rate. See Box 1 for an explanation of gene and protein naming conventions used throughout this text.

An average *C. elegans* worm has a life cycle of around 55 hours at 22ºC with an abundant food source (Figure 1), after which time the young adult hermaphrodite gains the ability to produce eggs. Sperm of hermaphrodites are produced first and stored in the spermatheca during

the L4, or fourth larval stage, and are used for self-fertilization later. But, if mated with a male, sperm of males will be used preferentially (Ward & Carrel, 1979). This evolutionary trait allows for the creation of new lab strains via mating.

Box 1 Genetic Nomenclature in *C. elegans*

daf-19	Names of gene that encodes DAF-19 proteins
DAF-19, DAF-19C	Names of the functional protein products of the gene with the same name. "C" indicates a particular isoform.
daf-19c::gfp	Insertion of *gfp* (**G**reen **F**luorescent **P**rotein) gene into the *daf-19c* gene
ofEX[daf-19c]	*daf-19c* is on an extrachromosomal (EX) transgene array; *of* designates the lab of production
daf-19(m86)II	An allele of *daf-19* that is named *m86*, and is on the second chromosome (II)

When under environmental stress, such as starvation or crowding, *C. elegans* will enter a non-reproductive dauer stage after the L1 larval molt. Dauers are thinner and longer than L2 and L3 larvae. As the main dispersal form in nature, they are extremely motile in response to mechanical stimulation, though immobile most of the time, and may stand on their tail and wave their body in the air ("nictating behavior", Lee et al., 2011; Reviewed by Corsi et al., 2015). Dauer animals have mouth plugs as part of their cuticle, providing stronger resistance against unfavorable substances and conditions in the environment (reviewed by Corsi et al., 2015). The metabolism of dauer animals is also reduced for survival in the absence of food (Riddle & Albert, 1997). Amphid neurons, especially AWC and AFD, responsible for chemotaxis and thermotaxis, respectively, are required for dauer formation (Bargmann & Mori, 1997; Riddle & Albert, 1997).

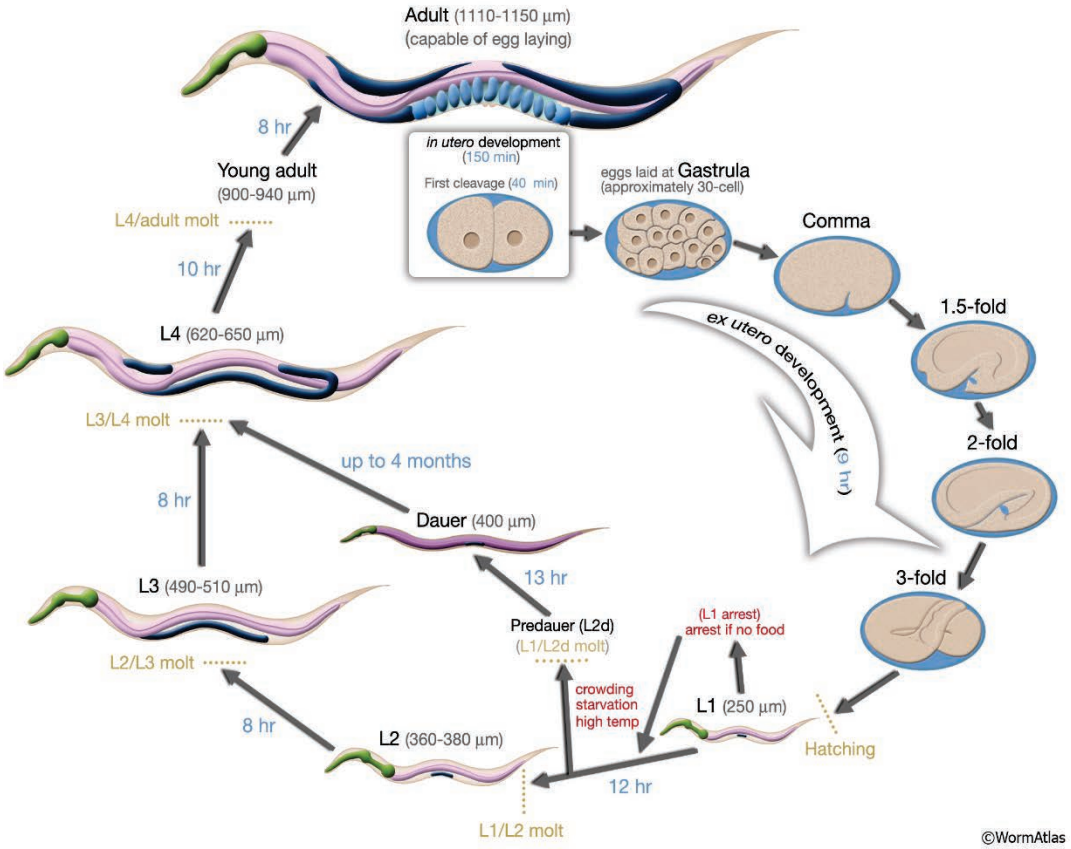


Figure 1 *C. elegans* life cycle at 22ºC. 0 min is fertilization. Numbers in blue along the arrows indicate the length of time the animal spends at a certain stage. First cleavage occurs at about 40 minutes post-fertilization. The length of the animal at each stage is marked next to the stage name in micrometers (µm). Original figure and figure legend from Wormatlas.org, Altan and Hall, 2005.

In laboratory, the layer of the mouth plug makes dauer worms unable to dye-fill. Dye-filling, which is the process of allowing the worms take up a lipophilic dye through ciliated dendrites that are usually exposed to the environment, is a method for detecting ciliary function and for neuron identification (Figure 9A; Hedgecock et al., 1985). Dauer larvae are also undesirable for strain maintenance due to possible epigenetic changes induced by starvation that can persist over two generations (Rechavi et al., 2014). Transferring dauer animals to a fresh nematode growth medium (NGM) plate usually induces development to the fourth larval stage (L4). For strains genetically modified to be dauer-constitutive, such as strains with *daf-19* mutations, less than a third of the animals can escape dauer arrest in a 15ºC environment (Swoboda et al., 2000).

One major advantage of *C. elegans* as a model organism is its transparency; life stages and

organs of live animals can be easily discerned under a dissection microscope (Corsi et al., 2015). Further, expression of transgenic markers can also be observed in live animals under fluorescent microscopes.

Corsi et al. (2015) summarized the medical significance of *C. elegans* research in their review:

> At least 38% of the *C. elegans* protein-coding genes have predicted orthologs in the human genome (Shaye & Greenwald, 2011), 60–80% of human genes have an orthologue in the *C. elegans* genome (Kaletta & Hengartner, 2006), and 40% of genes known to be associated with human diseases have clear orthologs in the *C. elegans* genome (Culetto & Sattelle, 2000). Thus, many discoveries in *C. elegans* have relevance to the study of human health and disease.

Regulatory Factor X Proteins

Regulatory Factor X-box binding proteins (RFX) are a family of winged-helix DNA-binding proteins that are highly conserved in eukaryotes; they consist of a DNA binding domain (DBD) and multiple conserved dimerization and extended dimerization domains (Emery et al., 1996a; Gajiwala et al., 2000; Aftab et al., 2008). As transcription factors, RFX proteins bind exclusively to a region within the promotor of their target genes. The region, termed the X-box (hence the name for the RFX proteins), is highly conserved among the eukaryotes (reviewed by Chu et al., 2012). X-box DNA sequence motifs have been used to find novel RFX factor-regulated genes.

RFX-1, RFX-2, and RFX-3 were first identified in human and mouse genomes (Reith et al., 1990 & 1994). Later, RFX-4 and RFX-5 surfaced, both in human and mouse, followed by SAK-1 in *Schizosaccharomyces pombe*, ScRFX in *Saccharomyces cerevisiae* (the later CRT-1), and CeRFX in *C. elegans* (now called DAF-19) (reviewed by Emery et al., 1996a). More recently discovered members of the family are RFX-6 and RFX-7 in humans (Aftab et al., 2008).

The RFX proteins, along with orthologs identified in many other eukaryotes, all share a

highly conserved DBD structure that is made up of three α-helices (H), β-strands (S), and connecting loops (L), with pairwise identities between 36% and 90% (Gajiwala et al., 2000; Figure 2A). Two strands and a loop form the wing-helix DBD that is shorter than that of other known DNA-binding proteins (Gajiwala et al., 2000). When interacting with an X-box region in the promoter, two RFX proteins dimerize by aligning in an inverse-parallel fashion, and flank the two sides of the X-box. The wing on the protein interacts with the major groove of the adjacent X-box half-site. A side chain of one of the helices (H3) on the RFX protein then interact with the minor grooves of the other X-box half-site (Gajiwala et al., 2000; Figure 2B). Fascinatingly, the homodimer complex has no direct protein-protein interactions, but the proteins cooperate through their interaction with DNA by hydrogen-bonding (Gajiwala et al., 2000).

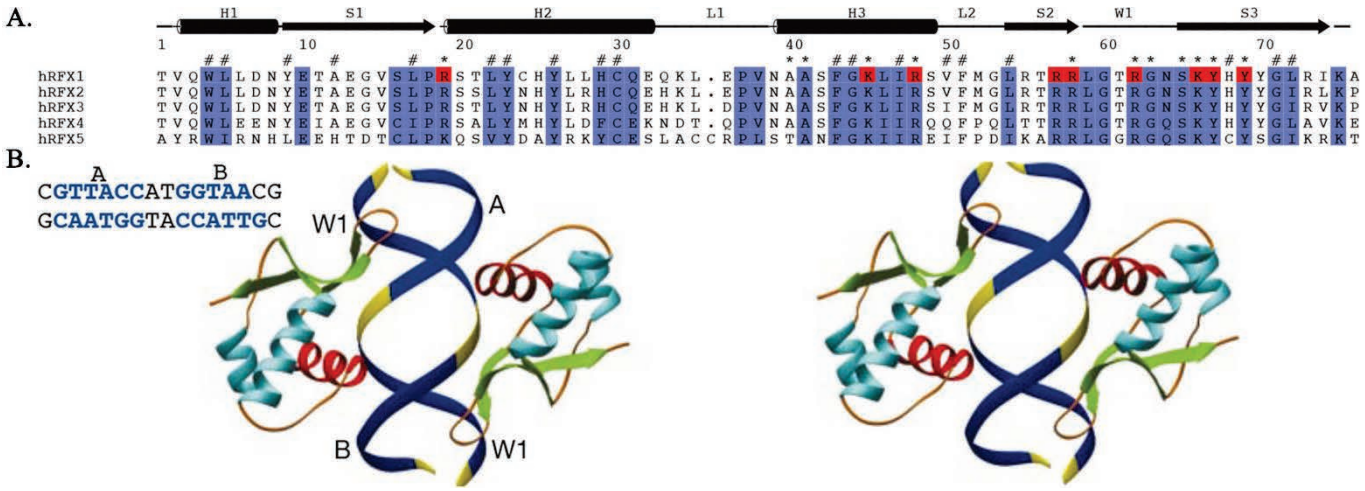


Figure 2 A. Amino acid sequence alignment of the DBDs of five human RFX proteins. Conserved amino acids are highlighted in purple. Hash signs denote hydrophobic core-comprising residues. Asterisks indicate side-chains buried within DNA grooves during binding. **B. Stereo diagram of hRFX-1-DNA 2:1 complex.** A and B indicate the two X-box half-sites. W1 denotes the wing-helix. Figures adapted from Gajiwala et al., 2000.

Along with the DBD, other regions of RFX-1 are also found to be orthologous to sequences of other RFX proteins (Emery et al., 1996a; Figure 3). These regions include extended dimerization domains (EDD) that aid in dimerization and the main dimerization domain (reviewed by Aftab et al., 2008; Figure 3). EDD mediates the formation of alternative dimerization complexes from conserved regions (Katan-Khaykovich et al., 1999). RFX-5,

however, does not have any dimerization domains. There are also regions of the RFX proteins that are rich in proline, glutamine, or acidic amino acids that are characteristic of transcription factors (Figure 3). RFX-4 and RFX-6 also have the conserved regions, in addition to the DBD, however, RFX-7, like RFX-5, lacks the dimerization domain (Figure 4). It is thought that the more closely related RFX-5 and RFX-7 play a different role in transcription than the other proteins and their orthologs. Indeed, from phylogenetic analysis based on the DBDs using CRT-1 from *S. cerevisiae* as an out-group, RFX-5 and RFX-7 and their orthologs are found to occupy the same subgroup in the RFX family, while the other subgroups contain RFX-1, -2, -3, and

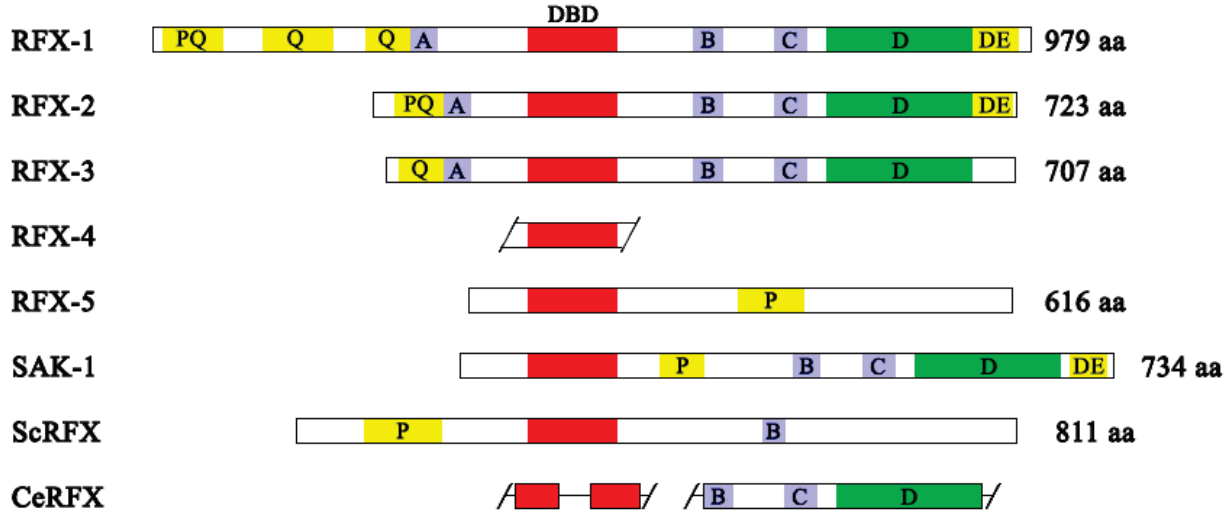


Figure 3 Alignment of RFX orthologs. Sequences are aligned by DNA-binding domains (DBD) in red. Yellow regions are rich in proline (P), glutamine (Q), or acidic amino acids (DE). Light blue indicates conserved regions A, B, and C. Green indicates the dimerization domain (D). Protein lengths are indicated after the sequence. ScRFX size is predicted from an open reading frame present. Image remastered from Emery et al., 1996a.

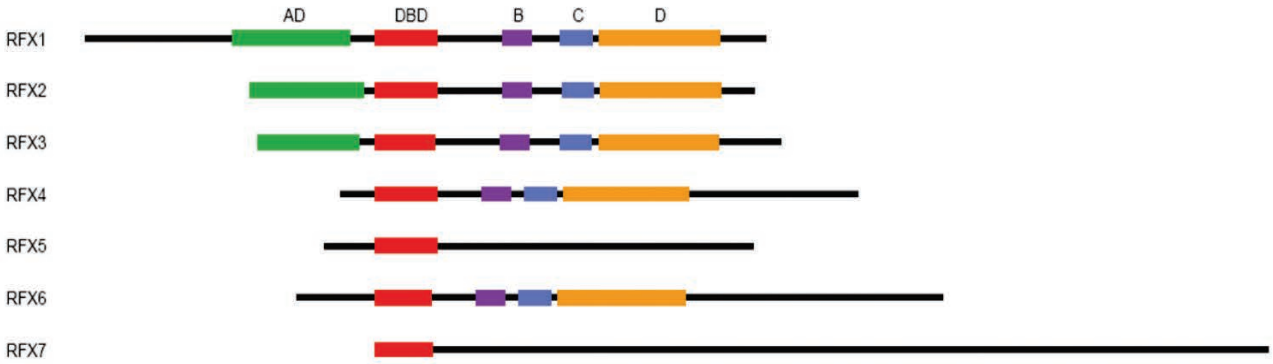


Figure 4 Alignment of functional motifs from known human RFX genes. Sequences are aligned by DBD, and the functional domains AD, DBD, B, C, and D are color-coded green, red, purple, blue, and yellow, as labeled. Domain length and positions are proportional to actual lengths. Figure and legend adapted from Aftab et al., 2008.

RFX-4 and -6. Aftab et al. (2008) speculated that the RFX-4/6 and RFX-5/7 subgroups are results of two gene duplication events prior to differentiation of the RFX-1 subgroup (Figure 5). *daf-19*, the only *C. elegans* RFX transcription factor gene, shares a common ancestor with the RFX-1/2/3 subgroup, but seems to have differentiated before RFX-1, -2, and -3.

The X-box motif is the general DNA sequence to which RFX proteins bind. The X-box region recognized by RFX proteins is identified as a 12 to 15-bp long nucleotide sequence of 5'-GTNRCC/n/RGYAACNN-3', in which "/n/" denotes a spacer of zero to three random nucleotides (Emery, 1996b; see **Box 2** for nucleotide codes). The X-box sequence was further refined to a sequence of 5'-GTHNYY/AT/RRNAAC-3' and averaged to 5'-RTHNYY/WT/RRNRAC-3' specifically for *C. elegans*. This sequence has been used in several bioinformatics studies to identify new target genes of RFX transcription factors (RFX TF) (Swoboda et al., 2000; Efimenko, 2005; Senti & Swoboda, 2008; Chu et al., 2012; Xie et al., 2013; Tammimies et al., 2016). At least in *C. elegans*, the X-box regions cluster most frequently between 60 bp and 130 bp upstream of the translation start site (Burghoorn et al., 2012).

Box 2 Nucleotide Codes. Bioinformatics. (n.d.). *IUPAC Codes*. Retrieved from https://www.bioinformatics.org/sms/iupac.html

IUPAC nucleotide code	Base
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
R	A or G
W	A or T
Y	C or T
H	A or C or T
N	any base

Despite the conservation of X-box nucleotide sequences, different RFX proteins bind to different specific X-box sequences (Emery et al., 1996a). Because of this specificity, multiple X-box regions may be present in a promoter, and cooperatively recruit RFX TFs to regulate the target gene expression. Both *S. cerevisiae* and *C. elegans* contain examples of motif pairs, in which a "strong" motif (referring to the motif driving stronger gene expression) and a less conserved, "weak" motif exist within the same promoter. The "weak" motif may attract and

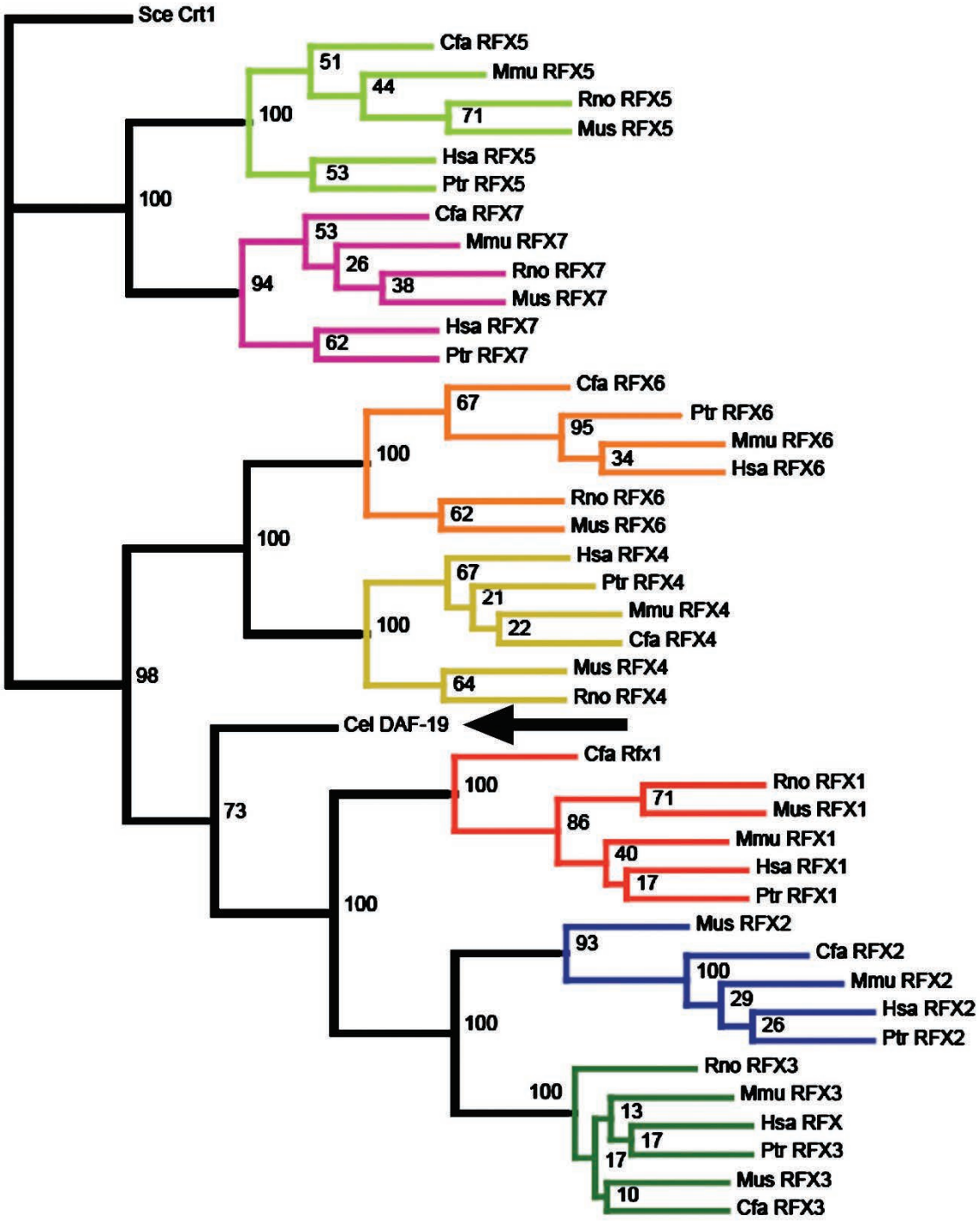


Figure 5 Phylogenetic analysis of mammalian RFX transcription factors based on DBDs. RFX proteins of six mammalian species and *C. elegans* are compared using CRT-1 as the out-group. Bootstrapped 100 times with numbers at each internal node as bootstrap value. Different colors indicate different orthologue groups. Arrow indicates DAF-19 from *C. elegans*. Six mammalian species: Mus–mouse (*Mus musculus*); Rno–Rat (*Rattus norvegicus*); Cfa–dog (*Canis familiaris*); Ptr–chimpanzee (*Pan troglodytes*); Mmu–monkey (*Macaca mulatta*) and Hsa–human (*Homo sapiens*). Figure adapted from Aftab et al., 2008.

therefore limit the “strong” motif’s access to RFX proteins, yielding an additive effect of enhancement or suppression, with the two X-box motifs working synergistically or antagonistically, respectively (Chu et al., 2012).

There are exceptions to X-box binding. As elaborated below, target genes regulated by certain isoforms of *daf-19*, *i.e.*, *daf-19a/b* and *daf-19m*, do not contain X-box motifs in their promoters (Senti & Swoboda, 2008; Wang et al., 2010).

Once bound to X-box sequence, some members of the mammalian RFX family can regulate processes related to the immune system. RFX-1 was first found to be important for the expression of major histocompatibility complex (MHC) class II genes (Reith et al., 1988), and later found to activate hepatitis B enhancer I (Reith et al., 1994). More recent studies also suggest its essential role in the initial development stages (Feng et al., 2009; reviewed by Choksi et al., 2014). RFX-5 is the most prominent immune system-related protein, forming a complex with a suite of transcription factors, such as RFXANK and CIITA to activate important MHC class II genes (Steimle et al., 1995; Kern et al., 1995). Lack of *RFX-5* expression may lead to bare lymphocyte syndrome, which disables the body’s immune system severely. DAF-19, the sole orthologue in the nematode *C. elegans*, cooperates with ATF-7 to trigger innate microbial immune response, not unlike RFX proteins in humans (Xie et al., 2013).

The other function of the RFX family, the focus of this study, is its regulation of cilia-related genes, specifically those involved in ciliogenesis. Cilia are ancient and complex organelles that serve sensory, motility, and developmental functions. They are hair-like organelles, protruding out of cells or the end of dendrites, and are present in nearly all unikont organisms. RFX proteins are present only in Unikonts, and appear to have been derived after ciliary genes evolved (Piasecki et al., 2010). RFX TF-mediated ciliogenesis is found only in animals, suggesting the

function of RFX TF is derived early in the animal lineage (Piasecki et al., 2010).

There is redundancy in RFX functions related to ciliogenesis. For example, *RFX-1* and its orthologs are expressed primarily in the brain (specifically cerebral cortex and Purkinje cells), in addition to "white blood cells, heart, eye, testis, and cancerous cells" (Aftab et al., 2008; Choksi et al., 2014). Though it is less involved in cilia-related gene regulations than other members of its subgroup, RFX-1 regulates the expression of *ALMS-1*, a gene that encodes a protein in the basal body of the cilia, either by itself as a homodimer or in concert with RFX-2 as a heterodimer (Purvis et al., 2010; reviewed by Choksi et al., 2014).

RFX-2 is expressed preferentially in the brain, but is also massively expressed in the testis (Aftab et al., 2008; Choksi et al., 2014). It is responsible for motile cilia in cells such as the epidermis of *Xenopus* larvae and multiciliated cells in mammalian airways (reviewed by Choksi et al., 2014), and is a key regulator of cilia/flagella-specific genes in mice. *RFX-2* mutants experience early apoptosis of germ cells, leading to spermatogenesis dysfunction (Kistler et al., 2015; Wu et al., 2016). RFX-2 is also responsible for ciliogenesis in the Kupffer's vesicle in zebrafish embryos, which is essential for left-right asymmetry of the body plan (Bisgrove et al., 2012).

Similarly, *RFX-3* is expressed in the embryonic node, and is essential for left-right asymmetry (Bonnafe et al., 2004), though it is expressed earlier than *RFX-2* (Bisgrove, et al. 2012). RFX-3 is also present in differentiating multiciliated cells of *Xenopus* (frogs), one of many expression patterns the gene shares with *RFX-2*. Lack of *RFX-3* may lead to shortened or fewer cilia in the node and pancreas, or inversely overproduced cilia in the node or the subcommissural organ (SCO) (reviewed by Choksi et al., 2014). Furthermore, RFX-1, RFX-2, and RFX-3 interact with each other to upregulate or downregulate *DYX1C1* and *DCDC2*

expression in the brain, maintaining cilia development, which otherwise would lead to developmental dyslexia (Tammimies et al., 2016).

RFX-4 is expressed in the testis and the brain (reviewed by Choksi et al., 2014), and most of its mutant phenotypes are identified in mice. For example, heterozygous *RFX-4* mutant mice exhibit severe hydrocephalus and lack of the SCO. Homozygous *RFX-4* mutants die shortly after birth, with an under-developed brain and ventricle (reviewed by Choksi et al., 2014). These phenotypes are thought to be a result of loss of cilia integrity caused by the mutation.

Interestingly, *RFX-4* and *RFX-6* do not have a glutamate-rich activation domain, and are thought to function by dimerizing with other *RFX* proteins (Aftab et al., 2008). *RFX-5* has not yet been found to have ciliary-related functions. The other member of its subgroup, *RFX-7*, is hypothesized to function by interacting with other non-RFX transcription factors (Aftab et al., 2008).

DAF-19 and Ciliogenesis

Even though multiple orthologs of RFX TF genes have been discovered in mammals, only one such gene is found in *C. elegans* (Swoboda et al., 2000), called *CeRFX* (Emery et al., 1996a), or, *daf-19*. *daf-19* was first found to regulate ciliogenesis in sensory neurons (Swoboda et al., 2000), and later found also to regulate the nematode's innate immunity similar to RFX-1 (Xie et al., 2013). It is also the first known regulator of synaptic maintenance (Senti & Swoboda, 2008). The effect on synaptic maintenance appears to be due to the action of the longer DAF-19A isoform (Senti & Swoboda, 2008). *daf-19* null alleles have defective roaming and dwelling behavior, that is rescued by transgene expression of DAF-19A alone (Senti & Swoboda, 2008); similarly, a mutation that deletes exon two of *daf-19* specific to the DAF-19A isoform, confers

the same dwelling and roaming defects as well as resistance to levamisole, an anti-parasitic drug (De Stasio et al., 2018). In contrast, the DAF-19C isoform regulates ciliogenesis in sensory neurons through direct X-box control (Gajiwala et al., 2000; Senti & Swoboda, 2008; Burghoorn et al., 2012). The smallest isoform, DAF-19M is expressed only in males where it controls specification of cilia in male-specific neurons such as HOA and HOB tail neurons (Wang et al., 2010).

The protein sequence of DAF-19 is most closely related to the RFX-1/2/3 subgroup (Figure 5). It contains the extended dimerization domain in addition to a conserved DNA binding domain. However, like the RFX-4/6 proteins, it lacks an activation domain that would interact with a portion of the RNA polymerase complex or basal transcription factors (Swoboda et al., 2000; Figure 6). It is also curious that *C. elegans* has only one copy of an RFX TF gene, since the subgroups diversified before the separation of nematodes and mammals (Aftab et al., 2008; Figure 5). If the phylogenetic analysis is correct, *C. elegans* should have orthologs of the other two subgroup as well. Aftab et al. (2008) speculate that the other orthologs were either lost to

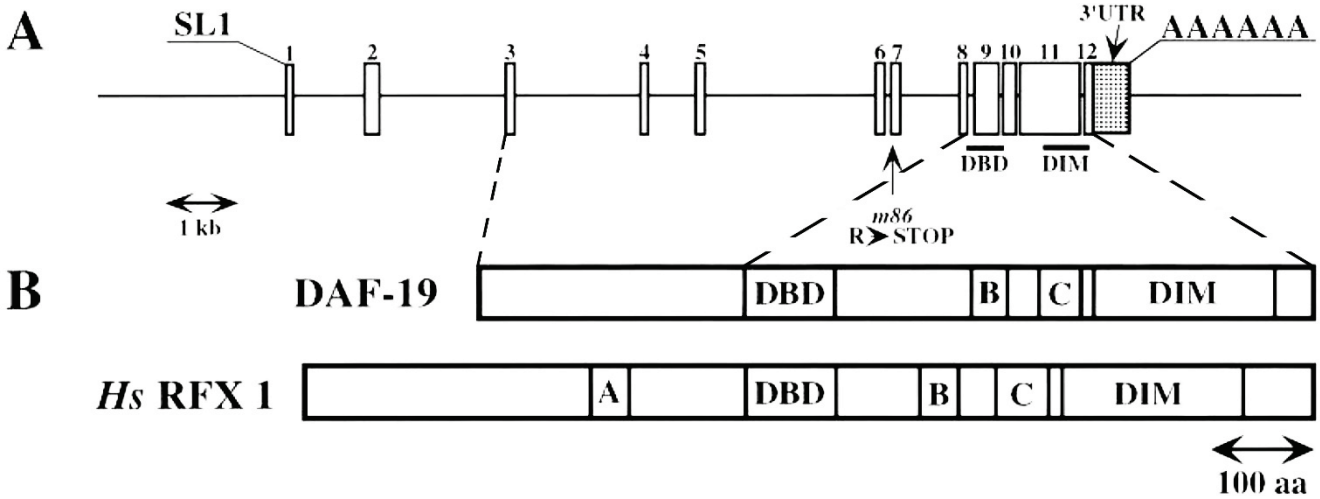


Figure 6 A. DNA structure of *DAF-19*. Numbered boxes depict exons, lines depict untranslated regions, and the dotted box depicts the **u**n**t**ranslated **r**egion at the 3' end. *DAF-19* is SL1 *trans*-spliced at the 5' end. The *m86* allele is a C to T point mutation, and results in an early stop codon. DBD: DNA Binding Domain; DIM: dimerization domain; AAAAAA: poly-A tail at the 3' end. **B. Amino acid sequence of DAF-19C, compared with that of human RFX-1.** Aligned at DBD. Each amino acid sequence as labeled. Dotted line aligns the protein with the gene sequence. The designation of the lettered domains remain consistent with previous figures. Figure adapted from Swoboda et al., 2000.

evolution, or never derived as a response to more complicated mammalian functions.

The *daf-19* gene has a total of twelve exons (Figure 6A) and three identified promoter regions. The longer isoforms can be alternatively spliced to generate a total of four different DAF-19 isoforms (Figure 7), each with its own functions. In the order of length, the four isoforms are *daf-19b*, *daf-19a*, *daf-19c*, and *daf-19m* (Senti & Swoboda, 2008; Wang et al., 2010; Figure 7). All four isoforms share the same DBDs and dimerization domains.

daf-19a and b are the two longer isoforms made from the *daf-19* gene. They are similar in exon composition except that *daf-19a* lacks exon 4 (Figure 7). *daf-19a/b* are thought to be expressed in around 240 non-ciliated head neurons, and have no cilia-related functions (Senti & Swoboda, 2008). Senti and Swoboda (2008) also found that *daf-19a* is necessary for synaptic transmission, and indirectly regulates expression of *unc-64*, *ida-1*, *unc-17*, *snb-1*, and *snt-1* at the protein level (mRNA levels of these genes were found to be stable). *daf-19a* mutants have reduced expression of synaptic vesicle components and other synaptic components, leading to reduction of synaptic function that worsens with age. The mechanism for this regulation remains unknown, since the aforementioned genes do not contain X-boxes in their promoters. It is also not known whether *daf-19a* and *daf-19b* work in concert or separately, or whether *daf-19b* has a specific function of its own at all.

daf-19c is a shorter isoform. Its protein begins at either exon 4 or 5 as both begin with a methionine codon. *daf-19c*-specific promoter elements reside in introns 3 and 4 (Senti & Swoboda, 2008; Figure 7). A mutation of exon 3 methionine (start codon) does not yield expected ciliogenesis phenotypes (De Stasio, unpublished). Thus the function of protein likely begins in exon 4. It is believed that *daf-19c* is expressed specifically in sensory head neurons and a selected number of tail neurons, but, like *daf-19a/b*, the detailed expression pattern isn't

characterized. Initial studies by Senti and Swoboda (2008) identified *daf-19c* as directly associated with ciliogenesis; *osm-5* and *bbs-7* are two identified target genes of *daf-19c*, both of which are "well-characterized, direct *daf-19* target genes that are expressed in ciliated sensory neurons and function in cilia formation," and are regulated through X-box binding. *daf-19c* also upregulates a suite of genes within the head neurons such as *asic-2* and *del-4* (predicted sodium channels), *spg-20* (orthologue of spartin, which inhibits synaptic growth), *ddn-1* (unknown), *eppl-1* (transaminase activity), and *mapk-15* (**m**itogen-**a**ctivated **p**rotein **k**inase) (De Stasio et al., 2018).

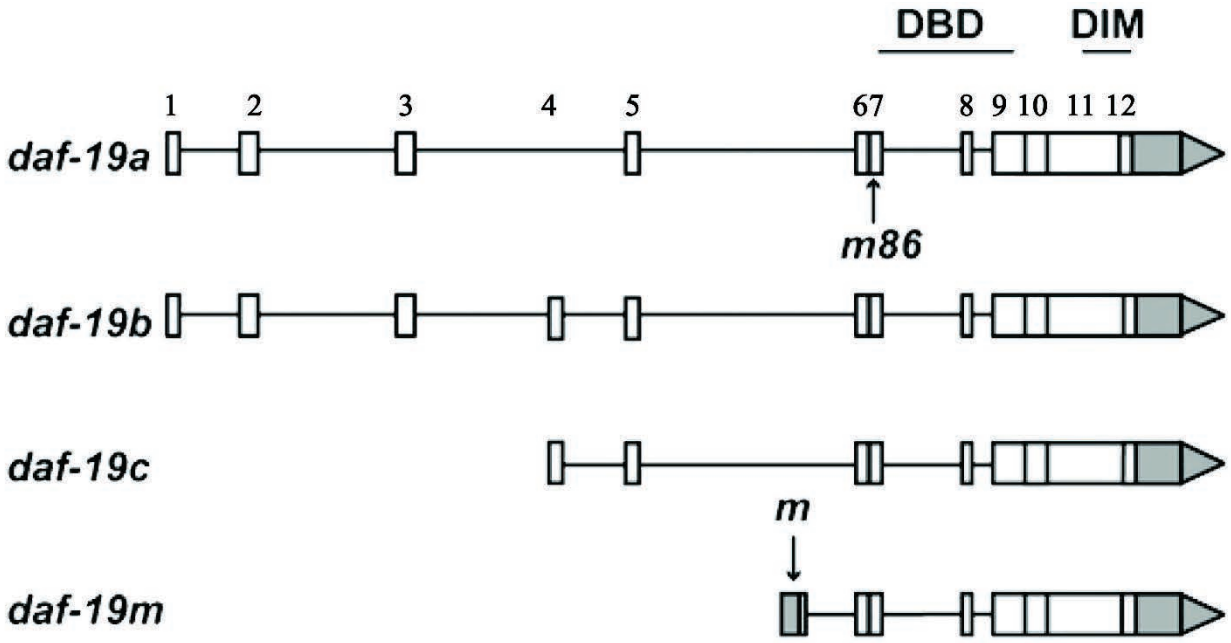


Figure 7 Genomic representation of the *daf-19* isoforms. The boxes depict exons and the line in between depict introns. The colored boxes depict untranslated 3' end region. Exons are numbered above; intron numbers follow the exons before it. *m86* and arrow points to the point mutation position in exon 7. *m* and arrow points to the modular promoter elements of *daf-19m* within intron 5. The four isoforms are as labeled, with the DBD spanning exon 7 to 9, and the dimerization domain encoded by parts of exon 11 and the entire exon 12. *daf-19a* exons: 1-3, 5-12; *daf-19b* exons: 1-12; *daf-19c* exons: 4-12; *daf-19m* exons: 6-12. Figure adapted from Wang et al., 2010.

daf-19m is the shortest isoform of *daf-19* and the most recently discovered. It consists only of exons 6 to 12, with a *daf-19m*-specific modular promoter element within intron 5 (Wang et al., 2010; Figure 7). Wang et al. (2010) reported that *daf-19m* is responsible for male-specific mating behavior. It is expressed in head neurons IL2 and male-specific CEM neurons and tail neurons

HOB and RnB. One of the roles of DAF-19M is to properly localize the ion channel, PKD-2, needed for proper male mating behavior. Though *daf-19m* is involved in mating behavior and is expressed in ciliated sensory neurons, it does not regulate ciliogenesis, but rather controls functional specialization of sensory cilia; intriguingly, its target genes, like those of *daf-19a/b*, do not contain an X-box in their promoters (Wang et al., 2010).

It is important to study these four isoforms individually since they share a number of exons that are fully sequenced. One way to study them is to perform knockouts and rescue experiments. This study takes the advantage of a point mutation in the seventh exon that is shared by all four isoforms, *m86* (Figure 7), which results in complete knockout of all four isoforms encoded by the *daf-19* gene, and allows rescue experiments with individual isoforms by means of transgene expression. This mutation was first found in a screen for dauer-constitutive larvae (Perkins et al., 1986), in which the worms were arrested in dauer stage after L1 larval stage. A third of *m86* dauer worms can escape dauer stage when in a 15ºC environment, but a more efficient practice for strain maintenance and experiment is to cross *daf-19(m86)* worms with *daf-12(sa204)* strain. Without expression of the *daf-12* gene, worms are unable to enter the non-reproductive dauer stage.

Research Aim and Approach

The aim of this research is to fully characterize the expression pattern of *daf-19c*. Though studies have been conducted on its functions, insufficient data exist regarding the identification of individual neurons that expresses *daf-19c* (Senti & Swoboda, 2008). We are interested in this question because we wonder whether DAF-19 isoforms control the expression of each other and thus help to specify neuronal cell fates.

We carried out our study by expressing only fluorescently labeled DAF-19C in an otherwise *daf-19* null genetic background. Using confocal fluorescent microscopy, we compared the *daf-19c*-expressing neurons to the cholinergic and glutamatergic neuron maps made by Pereira et al. (2015) and Serrano-Saiz et al. (2013), and with dye-filling neurons (Figure 9A; Hedgecock et al., 1985). Strains with cholinergic and glutamatergic marker genes were also crossed into the same *daf-19* mutant background. Here we report the number and identity of neurons that express *daf-19c* in worms of varying ages.

RESULTS

Approach

Expression of the gene *daf-19* produces several related proteins, including the protein DAF-19C (Figure 7). In this study, we aimed to identify the neurons that produce this protein. Our approach to the problem was to use a fluorescent marker, green fluorescent protein (GFP), to report the presence of the DAF-19C protein. To do so, we used a plasmid with *daf-19c::gfp* construct, pGG14 (Senti & Swoboda, 2008). pGG14 is a plasmid which contains the genomic DNA of *daf-19* starting from intron 3, thus encodes DAF-19C from its own promoter (see **Figure S2** for the plasmid structure). The *gfp* gene was attached just upstream of the *daf-19c* stop codon on the plasmid, such that a chimeric protein composed first of DAF-19C, then GFP would be produced in any cell that expresses *daf-19c* gene. The plasmid was injected into the gonad of a *C. elegans* strain

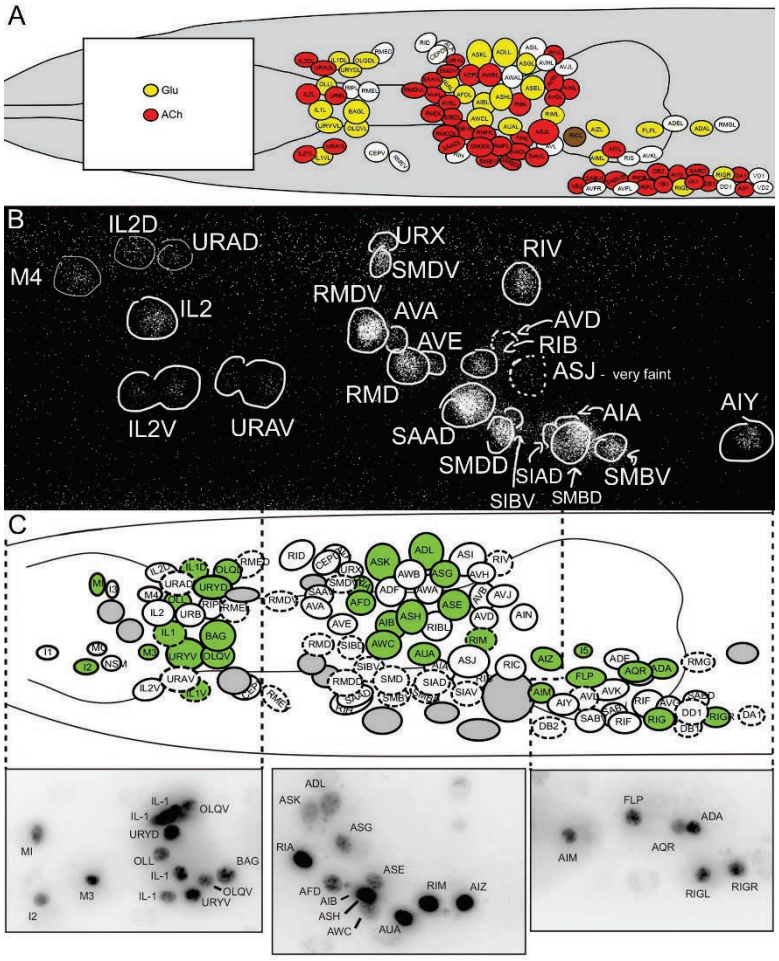


Figure 8 A. Map of fluorescent markers. Red indicates cholinergic neurons that express *cho-1::mCherry* fluorescent marker, and yellow indicates glutamatergic neurons that express *eat-4::mCherry* fluorescent marker. Adapted from Pereira et al., 2015. **B. Confocal image of cholinergic neurons expressing *cho-1::mCherry*.** Not all neurons are identified; identified neurons are sufficient for the purpose of this study. **C. Map of *eat-4::mCherry* fluorescent markers.** Green indicates glutamatergic neurons that express *eat-4::mCherry* fluorescent marker. Confocal image shows corresponding neurons in different sections of the pharynx. Adapted from Serrano-Saiz et al., 2013.

with *daf-19(m86)* mutant background. Because DAF-19C is known to be a transcription factor and therefore is transported to the nucleus; neurons that expresses DAF-19C will fluoresce green in their nuclei. We identified these neurons by comparing the location of the green fluorescence to that of pre-defined fluorescent marker expression and that of the dye-filling neurons. One of the fluorescent markers was *cho-1::mCherry*, expressed in all 52 cholinergic neuron classes (Figure 8B; Pereira et al., 2015); the other marker was *eat-4::mCherry*, expressed in all 38 glutamatergic neuron classes (Figure 8C; Serrano-Saiz et al., 2013). Dye-filling neurons were six amphid neuron classes: ASK, ADL, ASI, AWB, ASH, and ASJ (Figure 9A; Hedgecock et al., 1985).

Different isoforms of a gene product may interact with each other, or one isoform may change the expression pattern of other isoforms. To eliminate this possible interference, we generated strains with different fluorescent markers in the *daf-19(m86)* mutant background. This genetic background does not produce any form of DAF-19 protein, allowing us to examine expression patterns of any single DAF-19 isoform by injecting its gene into this background. LU663 is such a strain in the *m86* background, injected with *daf-19c::gfp*. LU663 also has the gene *elt-2::gfp* injected, which exhibits green polka dot pattern in the intestine under fluorescence microscopes and is used to distinguish transgenic offspring from non-transgenic siblings. Our marker strains, LU724 (with *cho-1::mCherry*) and LU725 (with *eat-4::mCherry*), also have the *daf-19(m86)* genetic background (see **Methods** section and **Table S2**) to avoid interference after mating.

We identified neurons by co-localization of the red and green fluorophores or by their relative positions using data from confocal microscopy. A confocal microscope collects a set of images at predefined, consecutive focal lengths throughout the sample, called a z-stack. Using

appropriate software (see **Methods** section), the z-stack can be compiled into either a 3D model or a maximum projection image (ProjMax). The 3D model is produced by aligning each image in the z-stack at their respective focal lengths. It is rotatable along a predefined axis, allowing for a spatial analysis of data. For example, in most occasions, a three-dimensional context is needed to ensure a genuine fluorescent co-localization. The 3D projection model, because of its rotatable axis, is helpful in determining whether the co-localization is genuine or caused by overlap of nearby fluorescent neurons. The maximum projection image is compiled by projecting the brightest pixel in each of the images of the z-stack onto a single image. This reconstructs a sharper image of the data in its entirety.

An example of the identification of *daf-19c*-expressing neurons AWC, AIB, and AUA is shown in Figure 9. We exposed these worms to fluorescent DiI as shown in figure 9A (Altun with permission). Based on the orientation of dendrites and location of the neurons, we first determined the anterior-posterior and dorsal-ventral axes. Then we determined the identity of each red dye-filled neuron according to the map (Figure 9; for full map, see Figure 15). ASK, ADL, and ASI are on the dorsal side, one next to each other, front anterior to posterior, forming an arch; sitting in the hollow of the arch is AWB; below that is ASH, and the big bottom-most neuron is ASJ. All of these neurons are in pairs: one on the left and one on the right. Neural nuclei expressing DAF-19C protein are false-colored blue (Figure 9C).

From the confocal maximum projection, we can see co-localization of red and blue in AWB; to the right of ASH is a neuron that is most likely ASE, judging by its position and distance (see pan-neuronal map, Figure 9B). Immediately left of ASJ can only be AUA, because of its proximity; and the neuron left of AUA and below ASH is most likely AIB. The neuron to the left of ASH can be either AVE or AWC, but because AWC is adjacent to AIB this neuron is more

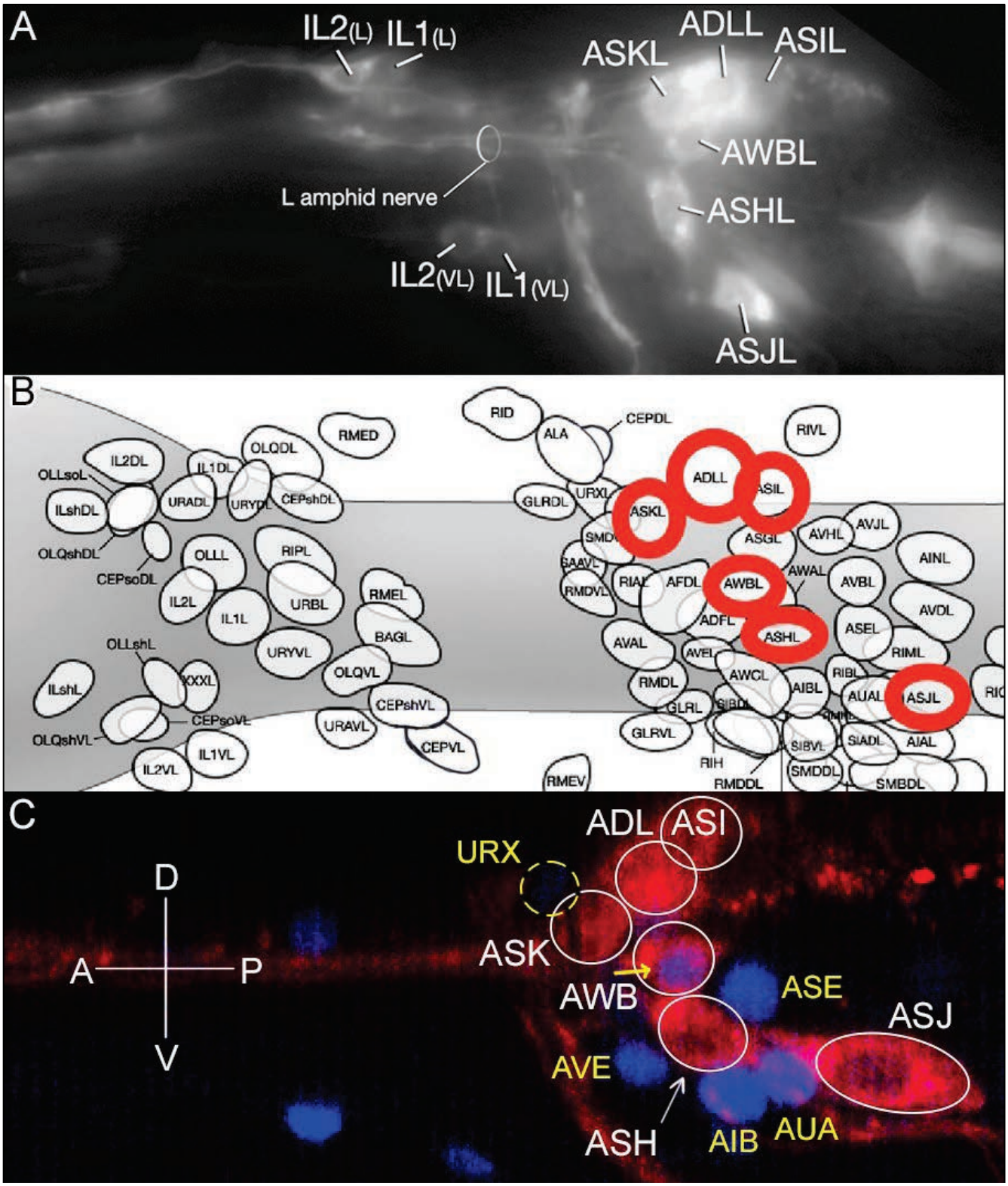


Figure 9 A. Dye-filling neurons in calcium acetate with DiI. DiI stains the entire neuron, including the axons and dendrites. Reproduced with permission from Zeynep Altun **B. Part of pan-neuronal map of *C. elegans*.** Red: Neurons that can be dye-filled with DiI in M9. **C. Maximum projection of a DiI dye-filled worm.** White circles, arrow, and names denote the red dye-filling neurons. Yellow dashed circle and names denote the blue neurons expressing DAF-19C::GFP. Yellow arrow points out co-localization of the two fluorophores.

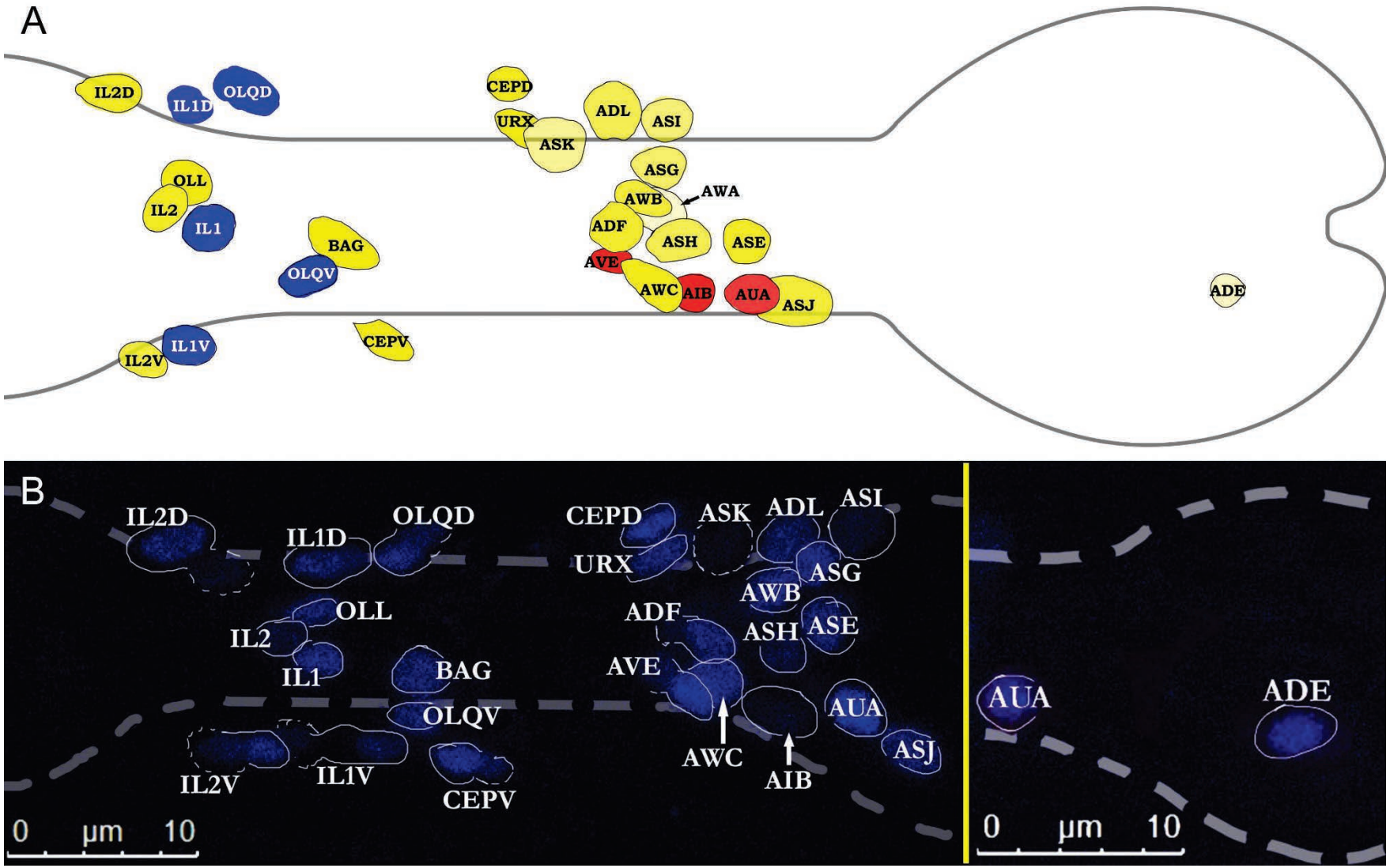


Figure 10 DAF-19C Expression in *C. elegans* Head Neurons. Pharyngeal bulbs outlined in grey for all images. **A. Schematic map of DAF-19C expression.** Neuron functions are color coded as follows: blue – polymodal neurons, yellow – sensory neurons, and red – interneurons. Transparency and saturation of each neuron directly reflect its frequency of DAF-19C expression. **B. Maximum projection of confocal microscopy of neurons with DAF-19C::GFP expression.** Neurons expressing green fluorescence attached to DAF-19C are false-colored blue. Yellow line separates the two samples. Scale bars as shown.

likely AVE. The small dash of blue on top of ASK should be URX. Because we have no further anterior markers, we cannot identify the three neurons further anterior.

The expression pattern of gene *daf-19c* was analyzed in this manner in a total of 27 worms. We separated the worms into three groups according to their stage of post-embryonic development: adult (13 worms, four of which were males), L4 to young adults (eight worms, two of which were males), and L1 to L3 (six worms, no males). Because dye-filling only identifies six neurons in the posterior pharyngeal bulb, we did not include results from dye-filled worms in our statistical analysis; these worms were used for initial neuron identification and confirmation instead.

DAF-19C is expressed mainly in cephalic ciliated sensory neurons

We analyzed the total number of worms that expressed DAF-19C in specific neurons across our sample. Neurons in which DAF-19C expression was seen in fewer than five worms (<20%) were discarded as noise (**Table S1**). We identified 28 neuron pairs that express DAF-19C, and compiled them into a single neuron map (Table 1; Figure 10). Twenty sensory neurons, five polymodal neurons, and three interneurons were identified (Figure 11). No motor neurons were confirmed to express DAF-19C (Table 1). Twenty-four of the 28 *daf-19c*-expressing neuron pairs were ciliated neurons; four (URX, AIB, AUA, and AVE) were not ciliated. The cephalic sensory neurons, FLP (non-ciliated) and AQR (ciliated), were never

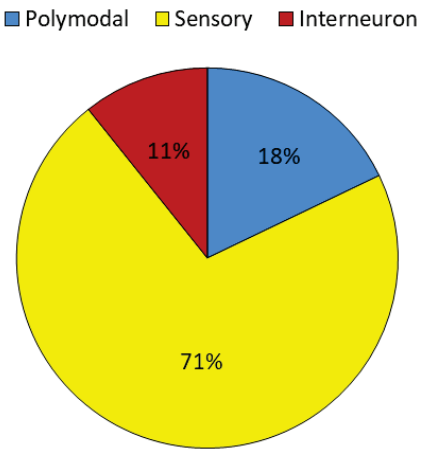


Figure 11 Percentages of neuron types with DAF-19C expression. 28 neurons express gene *daf-19c::gfp*: 20 sensory neurons (in yellow), five polymodal neurons (in blue), and three interneurons (in red).

observed to express DAF-19C (Table 2). Ciliated sensory neuron AFD was observed to express DAF-19C, but only in three worms (Table 1; **Table S1**). Because of our focus on the head neurons, we cannot determine whether DAF-19C was expressed in ciliated neurons located in the body or tail (PDE, PHA, PHB, PQR, and PVR). We have also observed and identified additional DAF-19C expression in male individuals. However, we were unable to identify these neurons, except for the CEM pairs ($n \leq 2$; **Figure S1**).

We analyzed the consistency of DAF-19C expression in these 28 neuron pairs. For every neuron pair:

$$Frequency\ (F) = \frac{number\ of\ worms\ in\ which\ the\ neuron\ expresses\ DAF\text{-}19C}{total\ number\ of\ worms\ in\ the\ sample\ size} \times 100\%$$

We also defined F≥80% as "consistent," 80%>F≥30% as "frequent," and F<30% as "inconsistent." Of the 28 neuron pairs, 15 consistently expressed DAF-19C, 11 had frequent expression, and two had inconsistent DAF-19C expression (Table 2 & 3). Among the consistent DAF-19C-expressing neuron pairs, seven were ciliated sensory neuron pairs: AWC, BAG, CEPV, CEPD, IL2D, IL2, and OLL; five were ciliated polymodal neuron pairs: IL1D, IL1, IL1V, OLQD, and OLQV; three are non-ciliated, a sensory neuron—URX—and two interneurons—AIB and AVE (Table 1, 2, & 3). Frequent DAF-19C-expressing neuron pairs include ten ciliated sensory neurons—ADF, ADL, ASE, ASG, ASH, ASI, ASJ, ASK, AWB, and IL2V—and one non-ciliated interneuron—AUA (Table 1, 2, & 3). The two neuron pairs with consistent DAF-19C expressions are ciliated sensory neurons AWA and ADE (Table 1 & 2). There are 33 neurons (16 pairs and one single neuron) that are deemed to show only infrequent DAF-19C expression, and thus discarded as noise, and include only one ciliated sensory neuron, AFD. The rest are non-ciliated neuron pairs: eight interneurons (AIA, AVA, AVJ, RIA, RIB, RIC, RIR, and URB), four motor neurons (M5, RMED, SMDV, and URAD), three polymodal neurons (ALA, SAAV, and

SIAD), and two sensory neurons (URAV and URY) (Table 1; **Table S1**)

Table 1 Ciliated and Non-ciliated DAF-19C-Expressing Neurons. 48 ciliated and 8 non-ciliated neurons were observed to have DAF-19C expression in our current population sample (n=27). Infrequent expression indicate neurons with DAF-19C gene expression in five or fewer worms. Refer to **Table S1** for more data.

Frequent Expressions				Infrequent Expressions			
Ciliated Neurons	**Number**	**Non-ciliated Neurons**	**Number**	**Ciliated Neurons**	**Number**	**Non-ciliated Neurons**	**Number**
ADE	2	URX	2	AFD	2	ALA	2
ADF	2	AIB	2			SAAV	2
ADL	2	AUA	2			SIAD	2
ASE	2	AVE	2			AIA	2
ASG	2					AVA	2
ASH	2					AVJ	2
ASI	2					RIA	2
ASJ	2					RIB	2
ASK	2					RIC	2
AWA	2					RIR	2
AWB	2					URB	2
AWC	2					M5	1
BAG	2					RMED	2
CEPV	2					SMDV	2
CEPD	2					URAD	2
IL2D	2					URAV	2
IL2	2					URY	2
IL2V	2				2 (1 pair)		33 (16 pairs + 1)
OLL	2			**Color Code**			
IL1D	2			Sensory neurons			
IL1	2			Polymodal neurons			
IL1V	2			Interneurons			
OLQD	2			Motor neurons			
OLQV	2						
	48 (24 pairs)		8 (4 pairs)				

Table 2 Consistency of DAF-19C Expression in Head Neurons Across Sample

Frequency	Number of Neurons	Neurons
>=0.8 (Consistent)	15	IL1D, IL1, IL1V, OLQD, OLQV, AWC, BAG, CEPV, CEPD, IL2D, IL2, OLL, URX, AIB, AVE
0.3=<x<0.8 (Frequent)	11	ADF, ADL, ASE, ASG, ASH, ASI, ASJ, ASK, AWB, IL2V, AUA
<0.3 (Inconsistent)	2	ADE, AWA

Table 3 Frequency of *daf-19c* Expression in *C. elegans* Head Neurons at Different Post-embryonic Developmental Stages. Table contains all neurons that in which *daf-19c* gene expression was identified five or more times in our sample. For neurons that expressed DAF-19C in five or fewer worms within our total sample ("noise"), refer to **Table S1**.

	Polymodal neurons					Sensory Neurons				
	IL1D	IL1	IL1V	OLQD	OLQV	ADE	ADF	ADL	ASE	ASG
Adult (n=13)	76.9%	84.6%	92.3%	92.3%	92.3%	38.5%	53.8%	38.5%	53.8%	46.2%
L4 & young adult (n=8)	100.0%	100.0%	100.0%	100.0%	100.0%	25.0%	100.0%	87.5%	100.0%	75.0%
L1 to L3 (n=6)	100.0%	100.0%	100.0%	100.0%	100.0%	0.0%	100.0%	100.0%	100.0%	66.7%
Total percentage (n=27)	88.9%	92.6%	96.3%	96.3%	96.3%	25.9%	77.8%	66.7%	77.8%	59.3%

	Sensory Neurons									
	ASH	ASI	ASJ	ASK	AWA	AWB	AWC	BAG	CEPV	CEPD
Adult (n=13)	53.8%	38.5%	53.8%	30.8%	7.7%	53.8%	69.2%	92.3%	69.2%	69.2%
L4 & young adult (n=8)	75.0%	62.5%	87.5%	50.0%	37.5%	87.5%	87.5%	87.5%	100.0%	100.0%
L1 to L3 (n=6)	83.3%	83.3%	100.0%	66.7%	16.7%	100.0%	100.0%	100.0%	100.0%	100.0%
Total percentage (n=27)	66.7%	55.6%	74.1%	44.4%	18.5%	74.1%	81.5%	92.6%	85.2%	85.2%

	Sensory Neurons					Interneurons		
	IL2D	IL2	IL2V	OLL	URX	AIB	AUA	AVE
Adult (n=13)	76.9%	69.2%	61.5%	92.3%	84.6%	76.9%	61.5%	69.2%
L4 & young adult (n=8)	75.0%	87.5%	75.0%	100.0%	100.0%	87.5%	75.0%	100.0%
L1 to L3 (n=6)	100.0%	100.0%	100.0%	83.3%	100.0%	83.3%	100.0%	100.0%
Total percentage (n=27)	81.5%	81.5%	74.1%	92.6%	92.6%	81.5%	74.1%	85.2%

DAF-19C expression differs in adults

To examine whether DAF-19C expression is consistent in different age groups, we separated our sample into three groups based on their developmental stage: adults, L4 larvae and young adults, and L1 to L3 larvae. L1 to L3 larvae are sexually immature, and from L4 developmental

stage to young adult stage, worms begin to become sexually active but cannot produce eggs. This grouping strategy was also convenient for identifying worm age based on physical features. For the larvae, the defining structure in identification was the future uterus, in the appearance of a void space at the intestine (Figure 12A). If the space was present, and the worm was slim, then it was an L4 larva. Any worms smaller in size and without the void space were classified as L1 to L3 larvae, with L1 worms being the smallest; worms bigger than L4 worms with a wider space but had no eggs were classified as young adults. Any worm with at least one egg was classified as an adult. Adults were usually much bigger under the microscope. Males were identified from

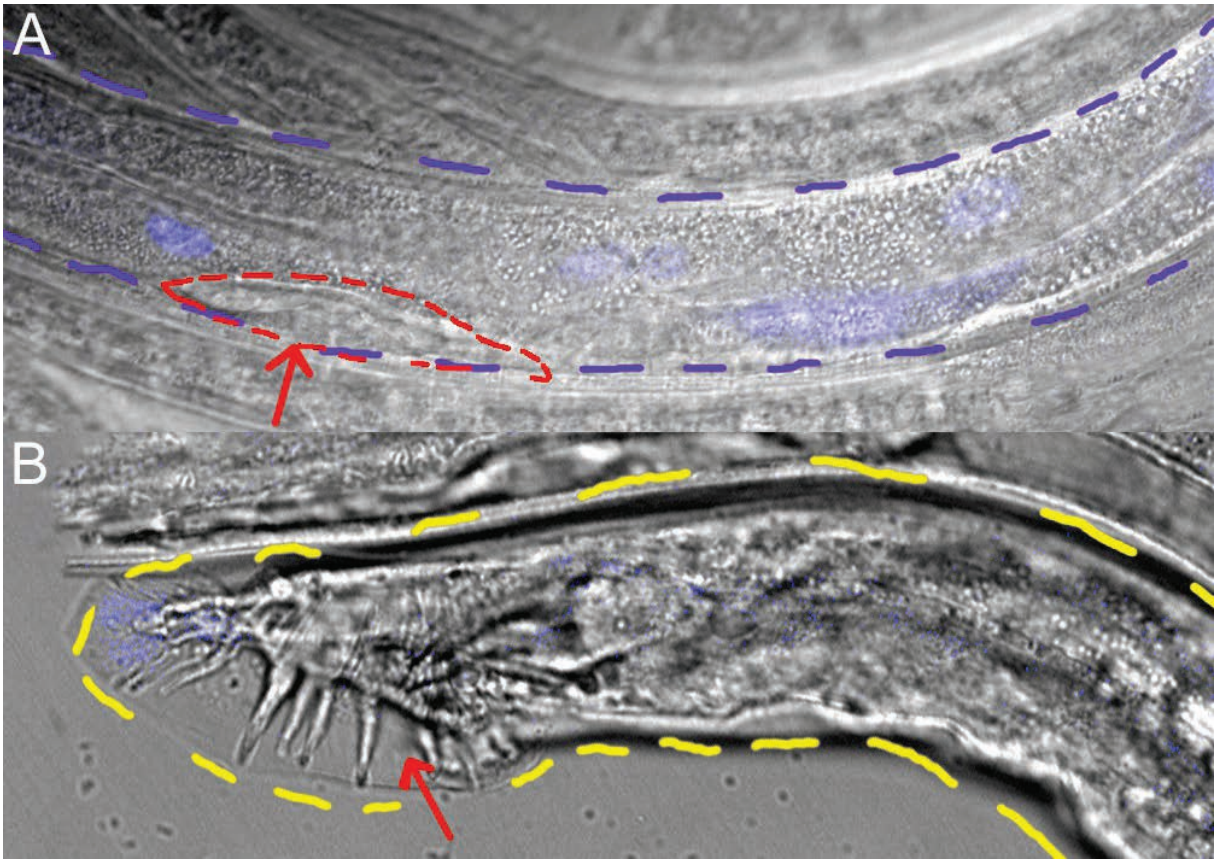


Figure 12 Bright-field microscopy image of *C. elegans* identifying structures. A. The intestine of an L4 larva. Red arrow and dashed line indicate the "void space" that is the future uterus and vulva site. Blue dashed line outlines the worm of interest. **B. The barbed tail of a male young adult *C. elegans*.** Red arrow indicates the tail. Yellow dashed line outlines the worm of interest.

their characteristic tails (Figure 12B), which were not discernable until the young adult stage. Thus worm sex was only identified in young adults and adults.

We hypothesized that in different developmental stages, there would be a difference in DAF-19C expression in the head neurons. We first described our data with average frequencies ($\overline{F}$) of DAF-19C expression in different neural function groups (polymodal, sensory, and interneurons), by defining $\overline{F}$ as

$$\overline{F} = \frac{\sum(Frequecy\ (F)\ of\ every\ neuron\ pair\ in\ the\ neuron\ function\ group)}{the\ number\ of\ neurons\ in\ the\ function\ group}$$

For example, for the interneuron functional group,

$$\overline{F}_{interneurons} = \frac{76.9\% + 61.5\% + 69.2\%\ (\textbf{the sum of frequecy of AIB, AUA, and AVE})}{3\ (\textbf{the number of interneurons})},$$

which is 69.2% for the adult group. We found that DAF-19C expression was less frequent in adults for neurons of all three functional groups when compared to the other two age groups (Table 3; Figure 13). There was noticeably less frequent DAF-19C expression in the L4 and young adult group compared to the L1 to L3 group (Figure 13). Similar results were found when we compared DAF-19C expression frequencies in individual neurons between different age groups. Except for neurons ADE and AWA, in every neuron, DAF-19C was less frequently expressed in the adult group than the other two age groups (Figure 14). DAF-19C expression in ADE was most frequently seen in adults, followed by L4 and young adults, but non-existant in younger L1 to L3 larvae (Figure 14). In AWA and ASG, L4 and young adult worms expressed DAF-19C most frequently, followed by L1 to L3 larvae, and least frequently in the adults (Figure 14). DAF-19C expression frequency in ASK follows the general pattern, but is noticibly lower in

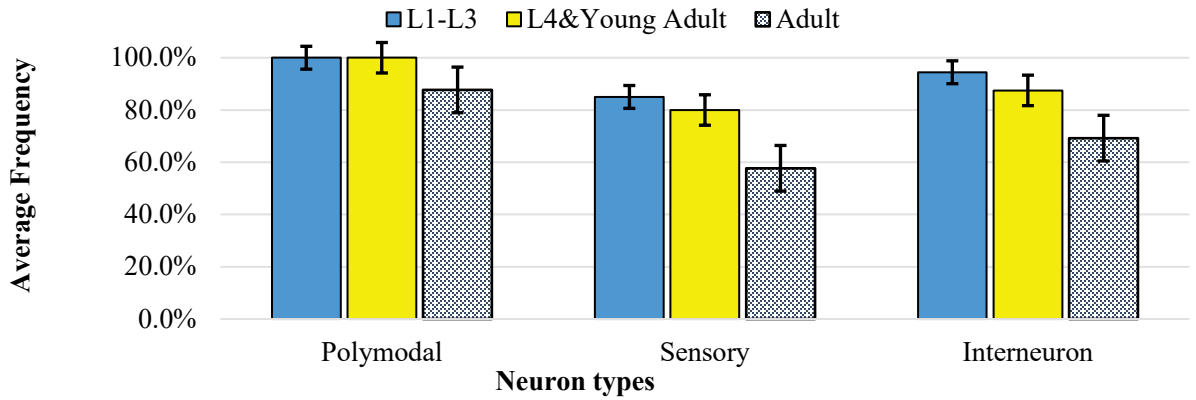


Figure 13 Frequency of DAF-19C expression in different neuron groups according to developmental stage. Averages are taken from the frequency of DAF-19C expression in polymodal, sensory, and interneurons across samples. Error bars show standard error. $n_{polymodal} = 5, n_{sensory} = 20, n_{interneuron} = 3; N_{adults} = 13, N_{L4\,\&\,young\ adults} = 8,\ N_{L1-L3s} = 6$. For $\overline{F}$ values, refer to Table 4.

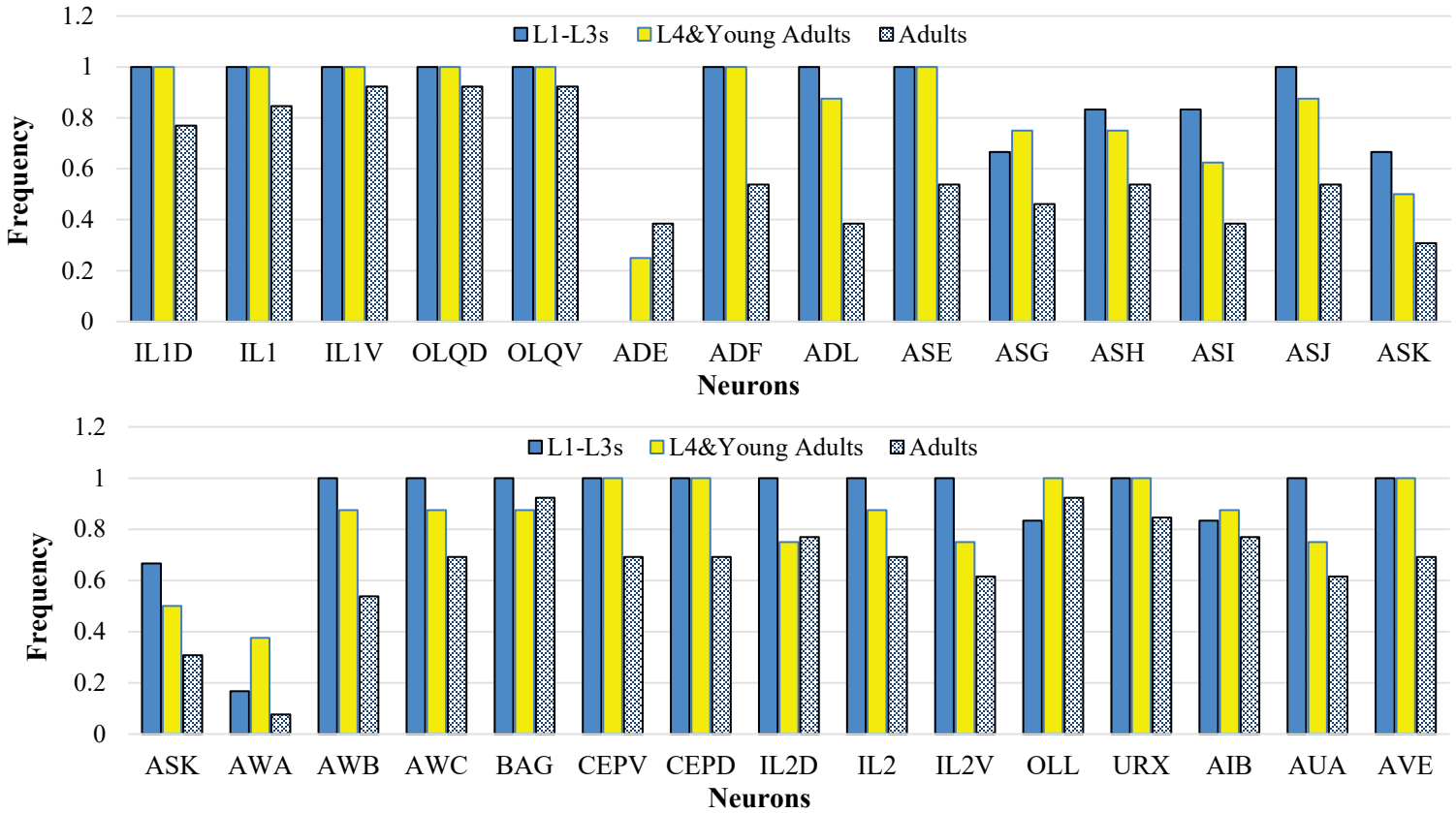


Figure 14 Frequency of DAF-19C expression in individual neuron pairs at different developmental stages. $N_{adults} = 13, N_{L4\,\&\,young\ adults} = 8,\ N_{L1-L3s} = 6$. For F values, refer to Table 3.

all three age groups when compared to that in the other neurons (Figure 14).

We further analyzed our data with statistical tests. Because our dataset had only one independent variable (developmental stage) that was separated into three categories, and our numerical dependent variables (frequency of DAF-19C's presence in a neuron across the sample) did not follow normal distribution, we used Kruskal-Wallis test for equal medians to test our hypothesis, and Mann-Whitney pairwise test to identify differences between age groups. We found that there is a significant difference in DAF-19C expression frequency among the age groups ($H = 23.97,\ p = 3.04 \times 10^{-6} \ll 0.05$; Box 2). *Post hoc* analysis with Mann-Whitney pairwise test further confirmed previous observations in Figure 13 and Figure 14, that the adult group expresses DAF-19C less frequently in their head neurons ($p_{Adults/L4\&YoungAdults} = 3.51 \times 10^{-4}, p_{Adults/L1-L3s} = 2.90 \times 10^{-6}$; Box 2).

Box 2 Results of Statistical Analysis Test. Kruskal-Wallis Test for Equal Medians and Mann-Whitney Pairwise for average DAF-19C expression frequency in head neurons in different ages. $N_{adults} = 13, N_{L4\ \&\ young\ adults} = 8,\ N_{L1-L3s} = 6$. Tests done through PAST 3.

Kruskal-Wallis Test for Equal Medians

H (chi2):	23.97
Hc (tie corrected):	25.41
p (same):	3.04E-06

There is a significant difference between sample medians

Mann-Whitney Pairwise Test

	Adults	L4 & Young Adults	L1 - L3s
Adults		3.51E-04	2.90E-06
L4 & Young Adults	3.51E-04		0.09583
L1 - L3s	2.90E-06	0.09583	

*Raw p values, uncorrected significance

DISCUSSIONS

DAF-19C is expressed in both ciliated and non-ciliated neurons, primarily in sensory neurons

To determine the cell-specific expression pattern of the C-isoform of the *daf-19* locus, we expressed the transgene *daf-19c::gfp* in worms with a *daf-19(m86)* genetic background. We used the dye-filling method along with fluorescently marked reporter genes, *cho-1::mCherry* and *eat-4::mCherry,* to identify neurons that express DAF-19C. We found 28 pairs of head neurons with DAF-19C expression. Twenty-four of the 28 neuron pairs were ciliated neurons, which included five polymodal neuron pairs and 19 sensory neuron pairs. Four of the 28 neuron pairs were non-ciliated neurons, which included three interneurons and one sensory neuron. We did not examine DAF-19C expression in tail neurons.

Our study broadly agrees with conclusions from previous research on *daf-19* expression. Senti & Swoboda (2008), using immunostaining, found that DAF-19C is "restricted to a small set of neurons in the head and the tail, a pattern reminiscent of ciliary sensory neurons." Our study shows that this is mostly correct. A huge portion of DAF-19C expression is located in ciliated neurons, and all of these are either sensory neurons or have sensory functions. The five polymodal neuron pairs, IL1D, IL1, IL1V, OLQD, and OLQV, all have mechanosensory functions (Altun and Hall). However, our study contradicts Senti & Swoboda's conclusion that DAF-19C is expressed only in ciliated sensory neurons, and that together with DAF-19A/B-expressing non-ciliated neurons exhibits "basically a pan-neuronal expression pattern of DAF-19" (2008). We found DAF-19C expression in non-ciliated sensory neurons URX, which is also

a ring interneuron, and in three neuron pairs, AIB, AUA, and AVE, that are currently cataloged as non-ciliated interneurons (Altun & Hall).

DAF-19C expression in non-ciliated neurons could explain feeding/roaming behaviors

Senti and Swoboda (2008) found that *daf-19(m86)* mutants in *C. elegans* have abnormal feeding and roaming patterns, and these abnormal behaviors could only be partially rescued if only one *daf-19* isoform is expressed. They conclude that, because the presence of DAF-19A could rescue the feeding/roaming defect, DAF-19 may have a new role in the maintenance of synaptic neurotransmission. Indeed, Senti and Swoboda found DAF-19A/B isoforms maintains synaptic protein expression (2008).

We would like to push this further by suggesting that DAF-19C also has additional functions in synaptic neurotransmission. DAF-19C has been proven to be involved in ciliogenesis (Senti & Swoboda, 2008), however, our characterization of DAF-19C expression shows that it is also expressed in non-ciliated neurons involved in feeding (AIB, AUA, and URX) and locomotion or sensory-based locomotion (AVE and URX). Because non-ciliated neurons do not have cilia and thus have no requirement for ciliary-related proteins, DAF-19C must have another function unrelated to ciliogenesis and cilia maintenance when expressed in these neurons. Because DAF-19A maintains synaptic protein expression and can partially rescue feeding/roaming behavior alone, and DAF-19C can also partially rescue this behavior, we believe DAF-19C has a similar synaptic-related function in non-ciliated neurons, and perhaps in a few ciliated polymodal

neurons. This hypothesis is in line with the study by De Stasio et al. (2018), who identified a downstream regulation of *spg-20* by *daf-19c*. *spg-20* is an orthologue of human spartin, whose product is involved in endosomal trafficking (transport of molecules between organelles within a cell) and interacts with microtubules, and therefore synaptic growth and possibly neurotransmission.

AWA, ADE, and ASK show noticeably lower DAF-19C expression frequency

We calculated the frequency of DAF-19C expression in the identified neurons within our sample size, and grouped the neurons accordingly as consistent, frequent, and inconsistent. We found two neuron pairs that inconsistently expressed DAF-19C: AWA ($F = 18.5\%$) and ADE ($F = 25.9\%$), and one neuron, ASK, in which the expression frequency is lower than 50%.

We believe the observation of low frequency expression in AWA and ADE are due to human error and our methodology. AWA neurons are neither cholinergic nor glutamatergic nor one of the dye-filling neurons, meaning that none of our available fluorescent markers are able to tag its location. We could only identify AWA from its position relative to our markers. Most of our AWA identifications came from our dye-filling samples, and its position between AWB and ASH meant that it was easy to mistake other neurons such as ASE and ASG with AWA (Figure 15). However, we have at least two instances in which we identified DAF-19C expressions in AWA neurons alongside AWB and ASH neurons. ADE is located in the far posterior of the posterior

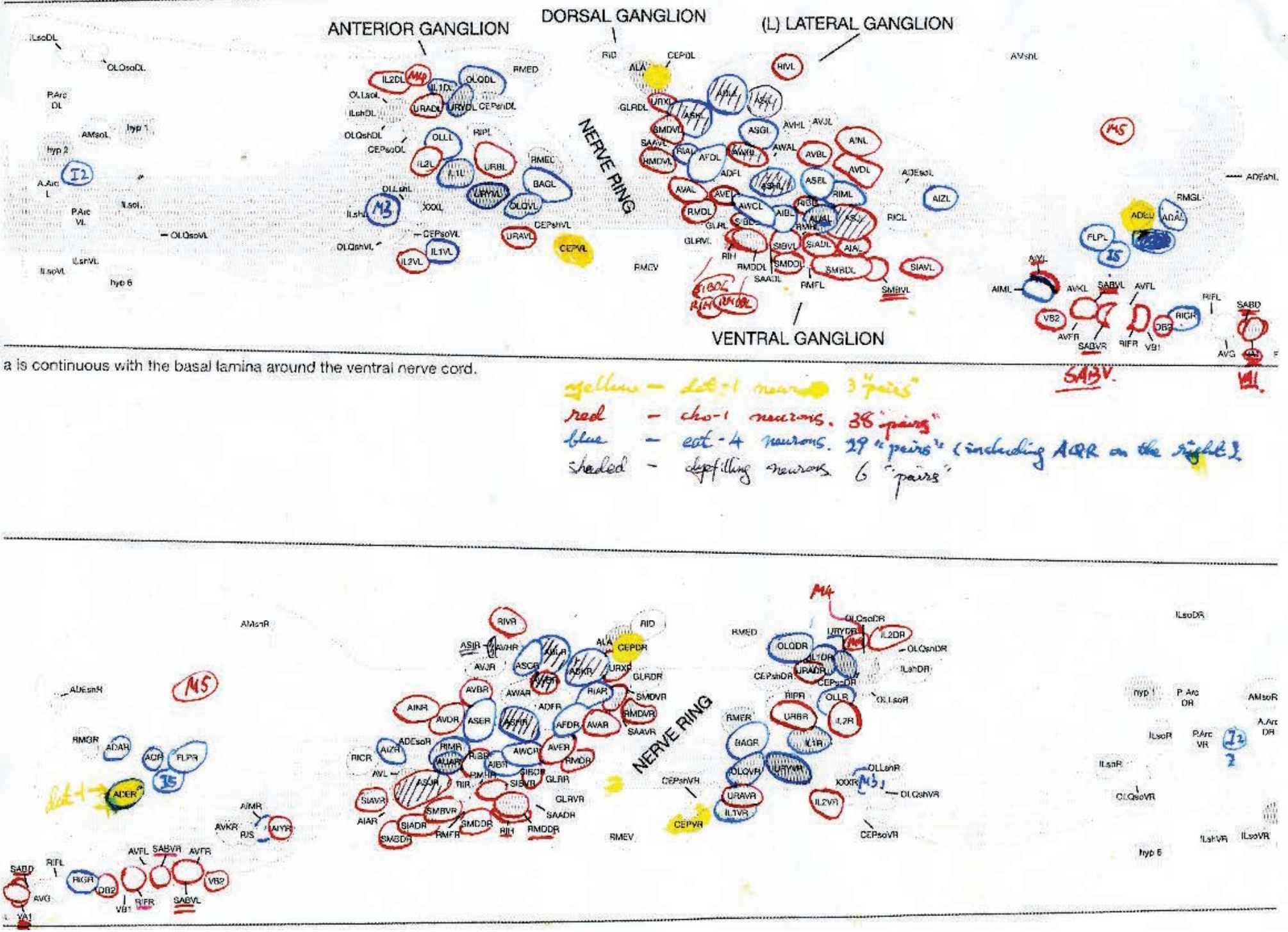


Figure 15 Identification of neurons expressing the gene of interest. Compiled map of cholinergic, glutamatergic, dopaminergic, and dye-filling neurons according to previous studies (Pereira et al., 2015; Serrano-Saiz et al., 2013; Sulston, Dew, & Brenner, 1975; Hedgecock et al., 1985). Adapted from Altun et al.

pharyngeal bulb (Figure 15) and is a dopaminergic neuron (Sulston et al., 1975). Because of our focus on the majority of neurons between the anterior and posterior pharyngeal bulbs, our imaging usually did not capture neurons located around the posterior pharyngeal bulb. The low frequency of image capture in the proximity of ADE contributed to its low frequency of identification, which in turn resulted in the observation of inconsistent DAF-19C expression.

DAF-19C expression in AWA and ADE could be confirmed with appropriate fluorescent markers, such *osm-9* for AWA (Colbert et al., 1997) and *dat-1* for ADE (Nass et al., 2002). Colbert et al. (1997) reported that a region close to the *osm-9* promoter is sufficient to drive OSM-9 expression in AWA. Thus, a fluorescent marker, *osm-9::mCherry*, could be synthesized and injected into LU663 to for fluorescence co-localization observation. Similarly, matings between a strain with *dat-1::mCherry* fluorescent transgenic marker and LU663 could confirm DAF-19C expression in ADE. We could not conduct this mating because our own strain of *dat-1::mCherry* reporter gene is co-injected with *elt-2::gfp*, which interferes with the marker of LU663 during offspring screening.

The low expression frequency of DAF-19C in ASK differs from that of AWA and ADE in that it is unlikely to be an artifact of imaging or neuronal identification. ASK is one of the six dye-filling neurons (Figure 9A; Hedgecock et al., 1985), and with green fluorescent protein attached to the protein of interest, expression of DAF-19C should be easy to spot in ASK if present. However, more than half of our samples did not have DAF-19C::GFP observed in ASK. Human error is clearly not the reason for its low expression frequency. We have also eliminated

the possible effect of mosaicism from the loss of transgene. When cells divide, the two daughter cells may receive an unequal amount of free plasmid. This process would repeat over time and result in an eventual loss of the plasmid in certain cells randomly. However, ASK is not the last DAF-19C-related neuron to arise from cell division (Sulston et al., 1983): all five polymodal neuron pairs arise after ASK, yet these are classified as consistent DAF-19C-expressing neurons (Figure 14). The proposed mosaicism, if indeed is the case, could be amended by integrating the gene of interest, *daf-19c::gfp*, into the genome of a strain in the *daf-19(m86)* background, and then observe through dye-filling for co-localization. Therefore, low DAF-19C expression frequency in ASK implies a biological significance. Because DAF-19C is responsible for ciliogenesis and cilia maintenance, our observation in ASK may point to the fact that ASK neurons do not require a constantly high amount of DAF-19C for ciliary maintenance. It is possible that *daf-19c* is not constitutive in ASK neurons, though the exact mechanism and reason is unknown.

DAF-19C expression decreases with age

We compared DAF-19C expression frequency between different developmental stages and found significantly lower expression of DAF-19C only in adults. The expression frequency in L4 larvae and young adults is observably lower than in L1 to L3 worms, but not statistically significant. This might be that, during the ciliogenesis process, once the main structure of the cilium is assembled, less DAF-19C amount would be needed for growth and maturation. This

would account for the observable insignificant decrease of DAF-19C expression frequency in L4 and young adults. For adults, after the cilium is mature, it is likely that less DAF-19C is required for the continual cilium maintenance. Cilium maturation would coincide with reproductive maturation at the adult stage; more energy would be directed for reproduction, instead of structural maintenance. This would lead to significantly less frequent *daf-19c* activity; *daf-19c* would be activated only when needed.

However, it is crucial to keep in mind that the frequency of DAF-19C expression is not the same amount as that of DAF-19C expression within a cell, or the activity of *daf-19c* gene. To measure the amount of DAF-19C expression, a test of fluorescence intensity should be conducted. We could not conduct an intensity test for this study because we did not keep a consistent excitation wavelength intensity during image collection with confocal microscopy. Our experimental worm strains suffer from mosaicism; the transgene and its fluorescent marker were lost occasionally during cell division, resulting in inconsistent expression amount across our sample size. Our excitation wavelength intensity and smart gain value had to be adjusted based on every sample in order to be sure that we could see expression in every neuron.

The test for fluorescent intensity should be conducted with most preferably integrated *daf-19c::gfp* gene for consistent expression and thus a consistent excitation wavelength intensity and smart gain value. The maximum projection image should then be analyzed with ImageJ for fluorescent intensity in each neurons. The brighter the fluorescence in a neuron, the more robust its DAF-19C expression, and the more amount of DAF-19C protein there is within the cell.

Comparing the brightness across different age groups would yield more detailed expression relationships between DAF-19C and age.

Moving forward: interactions between different *daf-19* isoforms

As aforementioned, our study adds to previous research on the expression pattern of *daf-19c* and provides more specific characterization of the expression pattern. Using antibodies against the N- and C-terminals of DAF-19 isoforms, Senti & Swoboda (2008) found a mutually exclusive expression pattern of the A/B and C isoforms of DAF-19: DAF-19C is expressed only in ciliated sensory neurons and DAF-19A/B is expressed in non-ciliated neurons. Also noted before was that in *daf-19(null)* mutant worms, whose feeding/roaming behavior is impaired, injection of either A/B isoform or C isoform would partially rescue the behavioral defect. Based on these results and ours, we hypothesize that there is an interaction between the two DAF-19 isoforms that confers repression of the C isoform expression in certain non-ciliated neurons, if not other ciliated neurons as well.

We suggest one experiment for future studies in order to examine the interaction between the longer and the shorter DAF-19 isoforms. Express both DAF-19A and DAF-19C::GFP together in a *daf-19(m86)* genetic background, and observe the expression patterns of the C isoform under this condition.

MATERIALS & METHODS

All *C. elegans* strains (**Table S2**) used were cultured following standard procedures (Brenner, 1974).

Dye-filling

Dye-filling was used to stain the amphid neurons, either as a marker for the amphid neurons or to determine whether a strain had functional neuronal cilia.

Dye-filling protocol was adapted from Worm Atlas. Dye solution was prepare with 1.5 ml M9 solution and 15 µl of 2.0 mg/ml DiO (Molecular Probes, catalog #D275) or DiI (Molecular Probes, catalog #282) in dimethylformamaide before every dye-filling. Worms were washed from agar plates into a 1.5 ml Eppendorf tube and spun in a centrifuge until a pellet was formed. The supernatant was removed with a vacuum and the entire dyeing solution was added. The tube was left on a rocker for one hour at medium rotation, and was spun again on a centrifuge until a pellet was formed. The supernatant was removed and 1.5 ml M9 solution was added, vortexed, and spun down in a centrifuge. After the supernatant was removed again (the washing with M9 was repeated if needed), the worm pellet was left on a fresh NGM plate for 5 to 8 hours before mounting on worm pads.

Mutant Strain

To analyze the expression pattern of DAF-19C, a strain expressing a fluorescently labeled

DAF-19C, but no other DAF-19 proteins was used. This strain. LU663, was constructed through microinjection of a plasmid expressing DAF-19C::GFP (**Figure S2**) from its endogenous promoter into hermaphrodites containing a knockout allele called *daf-19(m86ts)*. However, the *m86* allele confers a constitutive dauer phenotype in which animals do not reach reproductive age. This was remedied by the *daf-12(sa204)* knockout: without a functional copy of *daf-12*, the worms cannot enter dauer, allowing for reproduction and strain maintenance.

A co-injection marker gene, *elt-2::gfp*, was used to identify transgenic worms. *elt-2* expressed only in the gut (Sommermann et.al, 2010). Expression of this gene confers either a polka dot pattern in the gut when tagged with *gfp* or an entire red gut when tagged with *mCherry*. Thus these two markers do not interfere with the expression pattern of our gene of interest, *daf-19c*.

Neuronal Markers

The two markers used in the study were *cho-1::mCherry* (in LU724) and *eat-4::mCherry* (in LU725) (Pereira et al., 2015; Serrano-Saiz et al., 2013). Both markers were integrated into a chromosome (Unpublished data).

LU724 were constructed from a cross between OH13646, kindly provided by Oliver Hobert's lab from Columbia University, and LU602 (with *unc122::mCherry* as transgenic marker), constructed by Rosie Bauer at Lawrence University. UNC-122 is expressed mainly in coelomocytes (Loria et al., 2004), and the *unc-122::mCherry* marker shows up as two clusters of

two red dots along the ventral side of the worm. To generate DAF-19C-expressing strains that also carries a fluorescent neuronal marker, offspring from the initial mating (F1) were screened for both the presence of both red *cho-1* neurons and red coelomocytes. These were singly transferred onto new plates. At the F2 generation, worms were screened for red neurons only, and were again singly transferred. The F3 generation was subjected to a two-way screening. These plates were first screened for homozygous *cho-1::mCherry* genotype through fluorescent microscopy, and then screened for *daf-19(m86)* genetic background through DiI dye-filling. Dye-filling defective individuals with red neurons were singly plated again for further confirmation.

LU725 was constructed in a similar fashion (Figure 16). However, due to the fewer numbers

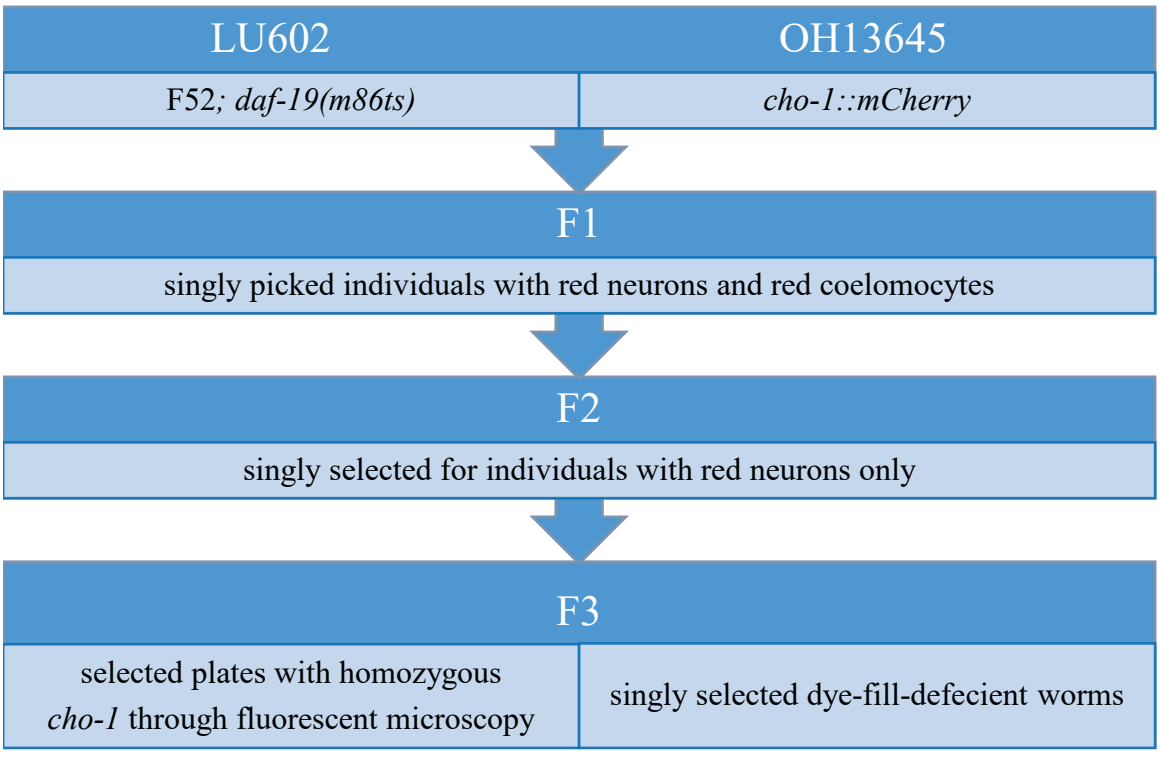


Figure 16 Flowchart for the procedure of making the marker strain LU724.

of glutamatergic neurons (Serrano-Saiz et al., 2013) and thus less-to-no observable fluorescence, screening of offspring was conducted using confocal microscopy. From the F1 generation, hermaphrodites that could dye-fill and without red coelomocytes were singly picked onto numbered plates. Dye-filling defective offspring from F1 were singly picked onto numbered plates. Once L1 larvae (the F3 generation) were observed on the plate, the F2 hermaphrodites were then mounted onto dotted worm pads (described below) for confocal microscopy. Worms were screened for the presence of red glutamatergic neurons. Only plates derived from worms with red neurons were then kept for the next step. F3 worms were further randomly picked and placed singly onto fresh NGM plates. When the plates become near-starved, half of each plate was washed and dye-filled with DiO to screen for *daf-19(m86)* background. These worms were then washed and mounted onto numbered worm pads to check for homozygous *eat-4::mCherry* genotype. Homozygote checks were repeated until the desired traits were confirmed.

The strains on which we performed the expression characterization were LU728, LU730, and LU663 itself. LU728 was a cross from LU725 and LU663, which helped us measure *daf-19c* neurons against the cholinergic neuron markers in a *daf-19(m86)* background. LU730 was a cross from LU725 and LU663, similar to LU728 except for having *eat-4::mCherry* as neuron marker. The construction procedures were similar to that of *cho-1::mCherry* marker strain construction, except that in these crossings, we singly picked the appropriate worms with green polka dot patterns in the guts as well, and did not have to screen for *m86* allele.

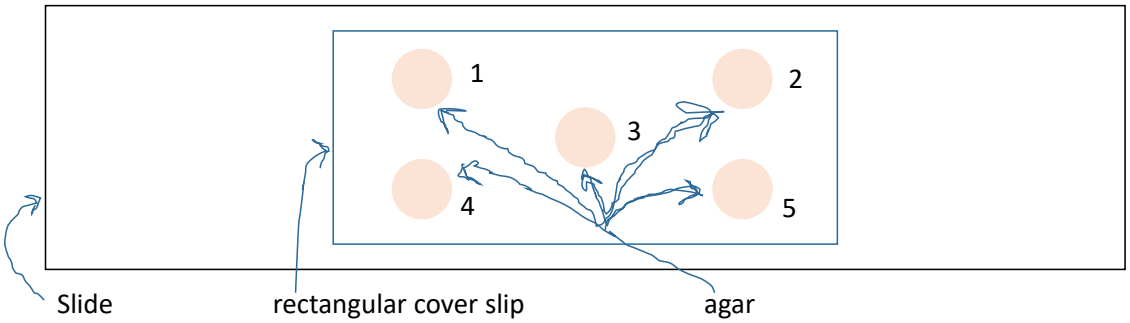


Figure 17 Dotted worm pads. Made with five 60 µl dots of 2% agar/NaN_3 solution on a microscope slide, covered with #2 coverslip.

Confocal Microscopy Preparation

All worm pads were made of 2.0% agarose according to standard procedure (**Appendix**) except for the dotted worm pad used during *eat-4::mCherry* screening.

Dotted worm pads were prepared from 750 µl of 2% agar and 15 µl of NaN_3. A total of 60 µl of prepared agar was dripped onto a 75x25x1 mm microscope slide to form 5 dots of agar (Figure 17). Each agar dot was numbered on the underside of the slide, and worms were mounted singly to each dot using a worm pick. The mounted worm pads were then covered with rectangular #2 cover slips for confocal microscopy.

Confocal Microscopy

Confocal microscopy was done on Leica TCS SP5 II confocal microscope at Lawrence University, using HCX PL APO CS 40x/1.3 Oil objective lens. Raw data were collected with Leica Application Suite Advanced Fluorescence version 2.7.3.9723 (LAS AF), with 17% argon

laser, HeNe 543 and HeNe 594 lasers active. HyD2 scanhead detected Leica/EGFP with a spectral range between 488-533 nm, and HyD4 scanhead detected Leica/mCherry between 656 and 696 nm. The PMT transmission collected DIC images. The wavelengths were set as following: 458 nm at 5%, 476 nm at 17%, 543 nm at 21%, and 594 nm at 19%. These were changed accordingly to the fluorescence of individual worms for brightness. During collection, the worms were first oriented with their heads in horizontal direction, and then DIC images of the worm's abdomen was taken at zoom 1.7 to determine dorsal-ventral orientation and age. Z-stacks were acquired using a scanning frequency of 5000 Hz with line average of 8. Pinhole size was adjusted from specimen to specimen for adequate neuron visualization. A zoom of 4 was used for worms aged from L4 to adult, while a zoom of 6 was used for worms aged from L3 and under. Only rarely a magnification of 5 was used. Other settings were at default (z-step size were default 0.17 μm).

The collected images were edited using software LAS AF and FIJI. 3D projection movies were 60 frames from -90º to 90º rotating along the x-axis when produced in LAS AF, and 360 frames from 0º to 360º rotating along x-axis when in FIJI. Movies were exported at 20 fps as .avi format.

Neuron Identification

The procedure of the identification of particular neurons was done as described in the **Results** section. The following programs were used in the process: LAS AF and FIJI for image

comparison, Windows Media Player (WMP) and Media Home Player – Black Edition (MHP-BE) for ProjMax movie playback, and Adobe Photoshop for image marking and conversion to PDF. Our own reference map (Figure 15) was prepared from a pan-neuron map from WormAtlas (Altun & Hall), blanked out through Photoshop, and then marked with known cholinergic (Pereira et al., 2015), glutamatergic (Serrano-Saiz et al., 2013), dopaminergic (Sulston, Dew, & Brenner, 1975), and dye-filling neurons (Hedgecock et al., 1985). Composite images and 3D maximum projection movies of worms were made using either LAS AF of FIJI. These were then identified as previously described using Photoshop.

CONCLUSION

We characterized DAF-19C expression in the head neurons of *C. elegans*. It is one component of basic research to understand the function of the DAF-19 gene. DAF-19C is an orthologue of the human RFX transcription factors, and an important transcription factor for *C. elegans*, targeting a suite of genes within its head neurons—some of the genes' functions are yet to be characterized. We hope this study lays down another piece of the foundation for future neuronal research, and paves the way for more to come.

ACKNOWLEDGEMENTS

I would like to give my thanks first to my off-campus friends and family who openly shared their support for me (and their distaste for my model system) through long-distance communication, and to my on-campus friends who did the same upfront. I would like to give my sincere appreciation and gratitude to the stockroom staff, JoAnn Stamm, Wayne Krueger, Melissa Eilbes, and Emily Hirn, for their help with the preparation of equipment, reagents, and NGM plates. This study is possible because of their care and support. I would like to thank Chelsea Gosney for critically reviewing my introduction, and Professor Jodi Sedlock for encouraging me throughout the winter term and my introduction writing process. I would like to thank members of the De Stasio and Piasecki labs who offered support, insights, and companionship in the past year, who kept me upbeat and well throughout my time in and outside the lab: Ricky Peña, Elisa Carloni, Nabor Vasquez, Alex Lessenger, Mikaela Hentz, and Shane Farrell; also thanks to members of Dickson lab, who shared lab space with me and offered no less than what my own lab members gave: Xinrui (Sam) Yang and Sophie Dion-Kirschner; also I would like to thank profusely are Professor Brian Piasecki, Rosie Bauer, Sophie Bice, and Aaron Schroeder, who gave me support from the start, and stepping in for my training when my mentor Professor Elizabeth De Stasio was on sabbatical.

Last but not least, I wish to express my sincere, profuse, and upmost gratitude to Professor Elizabeth De Stasio for her patient guidance and academic support. Thank you for taking me into your lab and opening my eyes on the subject of genetics and the field of *C. elegans*. Thank you

for your patient discussions, edits, and timely responses to all the questions I had. Thank you also for your encouragement throughout the year. I’m honored to be your student. This project would not have existed without your mentorship.

REFERENCES

Aftab, S., Semenec, L., Chu, J., & Chen, N. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evolutionary Biology, 8*(1), 226. doi:10.1186/1471-2148-8-226

Altun, Z.F. DiI and DiO Staining in *C. elegans*. In *Wormatlas*. http://wormatlas.org/neurons/Images/ADLDiI.jpg

Altun, Z. F. and Hall, D. H. (2005). Handbook of *C. elegans* Anatomy. In *WormAtlas*. http://www.wormatlas.org/ver1/handbook/contents.htm

Altun, Z.F. and Hall, D.H. (2009). Introduction. In *WormAtlas*. doi:10.3908/wormatlas.1.1. Retrieved from http://wormatlas.org/hermaphrodite/introduction/Introframeset.html

Bargmann C.I., and Mori I. (1997). Chemotaxis and Thermotaxis. In: Riddle DL, Blumenthal T, Meyer BJ et al., editors. *C. elegans II.* 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1997. Chapter 25. Available from: https://www.ncbi.nlm.nih.gov/books/NBK19978/

Bisgrove, B. W., Makova, S., Yost, H. J., & Brueckner, M. (2012). RFX2 is essential in the ciliated organ of asymmetry and an RFX2 transgene identifies a population of ciliated cells sufficient for fluid flow. *Developmental Biology, 363*(1), 166–178. doi:10.1016/j.ydbio.2011.12.030

Bonnafe, E., Touka, M., AitLounis, A., Baas, D., Barras, E., Ucla, C., … Reith, W. (2004). The Transcription Factor RFX3 Directs Nodal Cilium Development and Left-Right Asymmetry Specification. *Molecular and Cellular Biology, 24*(10), 4417–4427. doi:10.1128/MCB.24.10.4417-4427.2004

Burghoorn, J., Piasecki, B. P., Crona, F., Phirke, P., Jeppsson, K. E., & Swoboda, P. (2012). The *in vivo* dissection of direct RFX-target gene promoters in *C. elegans* reveals a novel cis-regulatory element, the C-box. *Developmental Biology, 368*(2), 415–426. doi:10.1016/j.ydbio.2012.05.033

Brenner, S. (1963). “Excerpt from Proposal to the Medical Research Council, October, 1963”. *How the Worm Got Started*. Retrieved from http://hobertlab.org/how-the-worm-got-started/

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

Brenner, S. (1988). Foreword. *The Nematode* Caenorhabditis elegans, edited by W. B. Wood. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Choksi, S. P., Lauter, G., Swoboda, P., & Roy, S. (2014). Switching on cilia: transcriptional networks regulating ciliogenesis. *Development, 141*(7), 1427–1441. doi:10.1242/dev.074666

Chu, J. S. C., Tarailo-Graovac, M., Zhang, D., Wang, J., Uyar, B., Tu, D., … Chen, N. (2012). Fine tuning of RFX/DAF-19-regulated target gene expression through binding to multiple sites in *Caenorhabditis elegans*. *Nucleic Acids Research, 40*(1), 53–64. doi:10.1093/nar/gkr690

Colbert, H. A., Smith, T. L., & Bargmann, C. I. (1997). OSM-9, a novel protein with structural similarity to channels, is required for olfaction, mechanosensation, and olfactory adaptation in *Caenorhabditis elegans*. *The Journal of Neuroscience*, *17*(21), 8259–8269. doi:10.1523/JNEUROSCI.17-21-08259.1997

Corsi, A. K., Wightman, B., & Chalfie, M. (2015). A Transparent Window into Biology: A Primer on *Caenorhabditis elegans*. *Genetics*, *200*(2), 387–407. doi:10.1534/genetics.115.176099

Culetto, E., & Sattelle, D. B. (2000). A role for *Caenorhabditis elegans* in understanding the function and interactions of human disease genes. *Human Molecular Genetics*, *9*(6), 869–877. doi:10.1093/hmg/9.6.869

De Stasio, E. A., Mueller, K. P., Bauer, R. J., Hurlburt, A. J., Bice, S. A., Scholtz, S. L., … Swoboda, P. (2018). An expanded role for the RFX transcription factor DAF-19, with Dual Functions in Ciliated and Non-ciliated Neurons. *Genetics*, genetics.300571.2017. doi:10.1534/genetics.117.300571

Efimenko, E. (2005). Analysis of xbx genes in C. elegans. *Development, 132*(8), 1923–1934. doi:10.1242/dev.01775

Emery, P., Durand, B., Mach, B., & Reith, W. (1996a). RFX Proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Research*, *24*(5), 803–807. doi:10.1093/nar/24.5.803

Emery, P., Strubin, M., Hofmann, K., Bucher, P., Mach, B., & Reith, W. (1996b). A consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity. *Molecular and Cellular Biology, 16*(8), 4486–4494.

Feng, C., Xu, W., & Zuo, Z. (2009). Knockout of the regulatory factor X1 gene leads to early embryonic lethality. *Biochemical and Biophysical Research Communications, 386*(4), 715–717. doi:10.1016/j.bbrc.2009.06.111

Gajiwala, K. S., Chen, H., Cornille, F., Roques, B. P., Reith, W., Mach, B., & Burley, S. K. (2000). Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature, 403*(6772), 916–921. doi:10.1038/35002634

Hedgecock, E. M., Culotti, J. G., Thomson, J. N., & Perkins, L. A. (1985). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes. *Developmental Biology*, *111*(1), 158–170. doi:10.1016/0012-1606(85)90443-9

Kaletta, T., & Hengartner, M. O. (2006). Finding function in novel targets: *C. elegans* as a model organism. *Nature Reviews Drug Discovery*, *5*(5), 387–399. doi:10.1038/nrd2031

Katan-Khaykovich, Y., Spiegel, I., & Shaul, Y. (1999). The dimerization/repression domain of RFX1 is related to a conserved region of its yeast homologues crt1 and sak1: a new function for an ancient motif. Edited by A. Klug. *Journal of Molecular Biology, 294*(1), 121–137. doi:10.1006/jmbi.1999.3245

Kern, I., Steimle, V., Siegrist, C.-A., & Mach, B. (1995). The two novel MHC class II transactivators RFX5 and CIITA both control expression of HLA-DM genes. *International Immunology, 7*(8), 1295–1299.

Kistler, W. S., Baas, D., Lemeille, S., Paschaki, M., Seguin-Estevez, Q., Barras, E., … Reith, W. (2015). RFX2 is a major transcriptional regulator of spermiogenesis. *PLOS Genetics, 11*(7), e1005368. doi:10.1371/journal.pgen.1005368

Lee, H., Choi, M., Lee, D., Kim, H., Hwang, H., Kim, H., … Lee, J. (2011). Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nature Neuroscience*, *15*(1), 107–112. doi:10.1038/nn.2975

Loria, P. M., Hodgkin, J. A., & Hobert, O. (2004). A Conserved postsynaptic transmembrane protein affecting neuromuscular signaling in *Caenorhabditis elegans*. *Journal of Neuroscience*, *24*(9), 2191–2201. doi:10.1523/JNEUROSCI.5462-03.2004

Meneely, P. M., McGovern, O. L., Heinis, F. I., & Yanowitz, J. L. (2012). Crossover distribution and frequency are regulated by *him-5* in *Caenorhabditis elegans*. *Genetics*, *190*(4), 1251–1266. doi:10.1534/genetics.111.137463

Nass, R., Hall, D. H., Miller, D. M., & Blakely, R. D. (2002). Neurotoxin-induced degeneration of dopamine neurons in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences*, *99*(5), 3264–3269. doi:10.1073/pnas.042497999

Piasecki, B. P., Burghoorn, J., & Swoboda, P. (2010). Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals. *Proceedings of the National Academy of Sciences*, *107*(29), 12969–12974. doi:10.1073/pnas.0914241107

Pereira, L., Kratsios, P., Serrano-Saiz, E., Sheftel, H., Mayo, A. E., Hall, D. H., … Hobert, O. (2015). A cellular and regulatory map of the cholinergic nervous system of *C. elegans*. *ELife, 4*. doi:10.7554/eLife.12432

Perkins, L. A., Hedgecock, E. M., Thomson, J. N., & Culotti, J. G. (1986). Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Developmental Biology, 117*(2), 456–487.

Purvis, T. L., Hearn, T., Spalluto, C., Knorz, V. J., Hanley, K. P., Sanchez-Elsner, T., … Wilson, D. I. (2010). Transcriptional regulation of the Alström syndrome gene ALMS1 by members of the RFX family and Sp1. *Gene*, *460*(1–2), 20–29. doi:10.1016/j.gene.2010.03.015

Rechavi, O., Houri-Ze'evi, L., Anava, S., Goh, W. S. S., Kerk, S. Y., Hannon, G. J., & Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell*, *158*(2), 277–287. doi:10.1016/j.cell.2014.06.020

Reith, W., Satola, S., Sanchez, C. H., Amaldi, I., Lisowska-Grospierre, B., Griscelli, C., … Mach, B. (1988). Congenital immunodeficiency with a regulatory defect in MHC class II gene expression lacks a specific HLA-DR promoter binding protein, RF-X. *Cell, 53*(6), 897–906. doi:10.1016/S0092-8674(88)90389-3

Reith, W., Herrero-Sanchez, C., Kobr, M., Silacci, P., Berte, C., Barras, E., … Mach, B. (1990). MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally

independent dimerization domain. *Genes & Development, 4*(9), 1528–1540. doi:10.1101/gad.4.9.1528

Reith, W., Ucla, C., Barras, E., Gaud, A., Durand, B., Herrero-Sanchez, C., … Mach, B. (1994). RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. *Molecular and Cellular Biology*, 14(2), 1230–1244.

Riddle, D. L., and Albert, P.S. (1997). The Dauer State. In: Riddle DL, Blumenthal T, Meyer BJ et al., editors. *C. elegans II*. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1997. Chapter 26, Section 2. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK20053/

Senti, G., & Swoboda, P. (2008). Distinct isoforms of the RFX transcription factor DAF-19 regulate ciliogenesis and maintenance of synaptic activity. *Molecular Biology of the Cell, 19*(12), 5517–5528.

Serrano-Saiz, E., Poole, R. J., Felton, T., Zhang, F., De La Cruz, E. D., & Hobert, O. (2013). Modular control of glutamatergic neuronal identity in *C. elegans* by distinct homeodomain proteins. *Cell, 155*(3), 659–673. doi:10.1016/j.cell.2013.09.052

Sommermann, E. M., Strohmaier, K. R., Maduro, M. F., & Rothman, J. H. (2010). Endoderm development in *Caenorhabditis elegans*: The synergistic action of ELT-2 and -7 mediates the specification→differentiation transition. *Developmental Biology*, *347*(1), 154–166. doi:10.1016/j.ydbio.2010.08.020

Steimle, V., Durand, B., Barras, E., Zufferey, M., Hadam, M. R., Mach, B., & Reith, W. (1995). A novel DNA-binding regulatory factor is mutated in primary MHC class II deficiency (bare lymphocyte syndrome). *Genes & Development, 9*(9), 1021–1032. doi:10.1101/gad.9.9.1021

Sulston, J., Dew, M., & Brenner, S. (1975). Dopaminergic neurons in the nematode *Caenorhabditis elegans*. *The Journal of Comparative Neurology*, *163*(2), 215–226. doi:10.1002/cne.901630207

Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology*, *100*(1), 64–119. doi:10.1016/0012-1606(83)90201-4

Swoboda, P., Adler, H. T., & Thomas, J. H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Molecular Cell, 5*(3), 411–421.

Tammimies, K., Bieder, A., Lauter, G., Sugiaman-Trapman, D., Torchet, R., Hokkanen, M.-E., … Swoboda, P. (2016). Ciliary dyslexia candidate genes *DYX1C1* and *DCDC2* are regulated by Regulatory Factor X (RFX) transcription factors through X-box promoter motifs. *The FASEB Journal, 30*(10), 3578–3587. doi:10.1096/fj.201500124RR

Wang, J., Schwartz, H. T., & Barr, M. M. (2010). Functional specialization of sensory cilia by an RFX transcription factor isoform. *Genetics, 186*(4), 1295–1307. doi:10.1534/genetics.110.122879

Ward, S., & Carrel, J. S. (1979). Fertilization and sperm competition in the nematode *Caenorhabditis elegans*. *Developmental Biology*, *73*(2), 304–321. doi:10.1016/0012-1606(79)90069-1

WormBase (2018). http://www.wormbase.org/resources/laboratory#012--10

Wu, Y., Hu, X., Li, Z., Wang, M., Li, S., Wang, X., … Han, C. (2016). Transcription factor RFX2 is a key regulator of mouse spermiogenesis. *Scientific Reports, 6*(1). doi:10.1038/srep20435

Xie, Y., Moussaif, M., Choi, S., Xu, L., & Sze, J. Y. (2013). RFX transcription factor DAF-19 regulates 5-HT and innate immune responses to pathogenic bacteria in *Caenorhabditis elegans*. *PLoS Genetics, 9*(3), e1003324. doi:10.1371/journal.pgen.1003324

SUPPLEMENTARY MATERIALS

Table S1 Neurons in which DAF-19C Expression is Regarded as Noise. Five or fewer worms in our sample size expressed DAF-19C in the neurons listed below. These expressions were discarded as noise. Samples include dye-filled worms, where most occurrences happen.

	Polymodal neurons			Sensory Neurons		Interneurons			
	ALA	SAAV	SIAD	AFD	URYV	AIA	AVA	AVJ	RIA
Adults (n=22)	0	0	1	2	1	3	0	0	1
L4 & young adults (n=11)	0	0	0	0	1	0	2	1	2
L1 to L3s (n=15)	3	0	0	1	1	0	3	0	0
Total occurrence (n=48)	3	0	1	3	3	3	5	1	3
Percentage	6.7%	0.0%	2.2%	6.7%	6.7%	6.7%	7.4%	2.2%	6.7%

	Interneurons				Motor Neurons				
	RIB	RIC	RIR	URB	M5	RMED	SMDV	URAD	URAV
Adults (n=22)	0	1	1	0	0	0	0	0	0
L4 & young adults (n=11)	1	0	0	1	1	1	1	0	1
L1 to L3s (n=15)	0	1	0	0	0	0	0	0	0
Total occurrence (n=48)	1	2	1	1	1	1	1	0	1
Percentage	2.2%	4.4%	2.2%	2.2%	2.2%	2.2%	2.2%	0.0%	2.2%

Table S2. *C. elegans* Strains Used in this Study

Lab	Strain	Genotype	Phenotype	Parental Strain	Comments
LU	602	*daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V; ofEX1076[pF52D2.2, unc-122::dsRed]*	ds red in coelomocytes	OE4476 - remade from LU649, OE3738, and OE3492	Bonnie Arbuckle, 2013. Remade by Rosie Bauer
LU	663	*daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V; lrEX176[pGG14, elt-2::gfp]*	transgenic worms dye-fill, non-transgenic do not, green speckled intestine	LU3492 injected with pGG14, isolate 3	Brian Piasecki & Elizabeth De Stasio
LU	724	*daf-19(m86); daf-12(sa204); him-5(e1490); cho-1::mCherry*	isolate 8A2	OH13646 x LU602 (non-transgenic)	Billy Liu
LU	725	*daf-19(m86); daf-12(sa204); him-5(e1490); eat-4::mCherry*	dye-fill deficient and dauer deficient	OH13645 x LU602 (non-transgenic)	Billy Liu
LU	726	*daf-19(m86); daf-12(sa204); him-5(e1490); cho-1::mCherry*	isolate 6B1	OH13646 x LU602 (non-transgenic)	Billy Liu
LU	728	*daf-19(m86); daf-12(sa204); him-5(e1490); cho-1::mCherry; otIs544[cho-1(fosmid)::SL2::mCherry::H2B]*	*cho-1 + daf-19C*	LU663 x LU724	Billy Liu
OH	13645	*pha-1(e2123)III; him-5(e1490); otls518[eat-4::SL2::mCherry pha-1+]*	integrated transgenic line, glutamatergic mCherry marker		From Oliver Hobert Lab
OH	13646	*pha-1(e2123)III; him-5(e1490); otls544[cho-1::SL2::mCherry+pha-1+]*	integrated transgenic line, cholinergic mCherry marker		From Oliver Hobert Lab

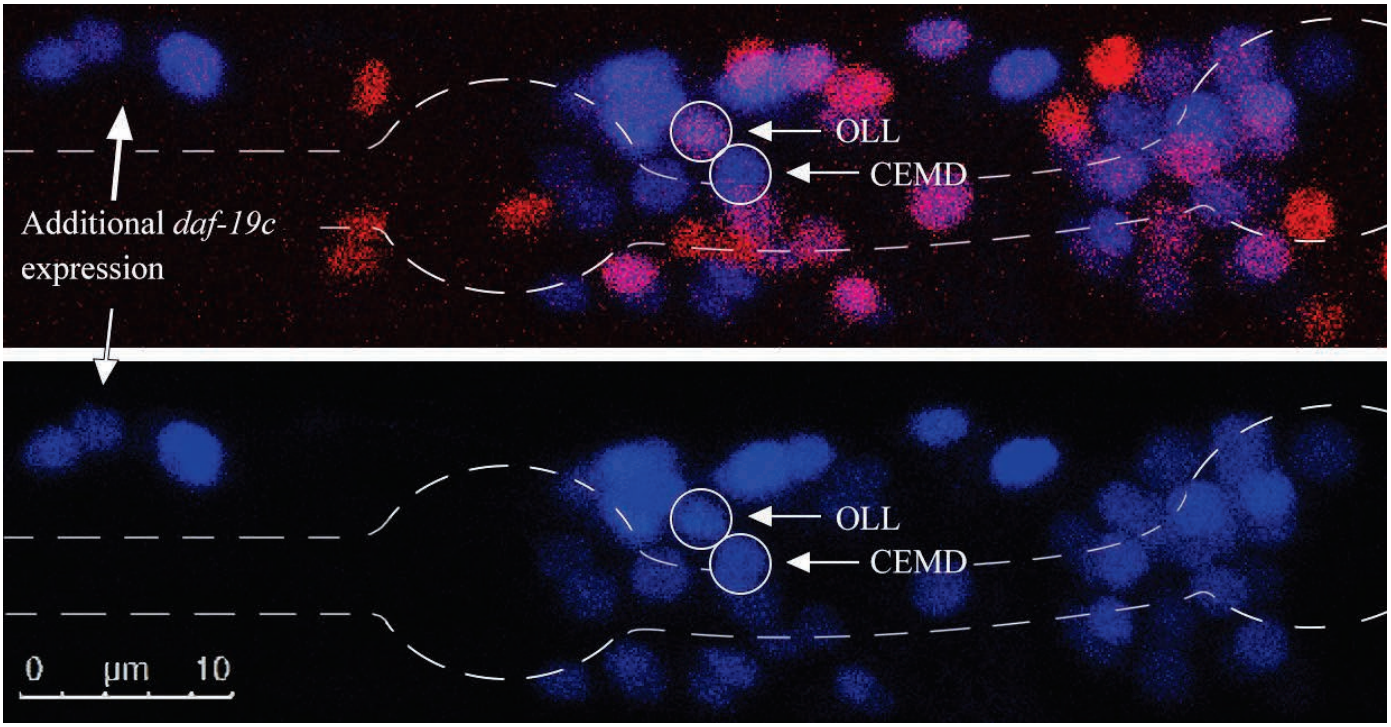


Figure S1 Maximum projection image of a young adult male worm showing CEMD with red glutamatergic neuron marker. Red indicates eat-4::mCherry marker neurons; blue indicates DAF-19C::GFP expression. White dotted line indicates location of the pharynx. Additional unidentifiable DAF-19C-expressing neurons are indicated by arrows. Scale bar as shown.

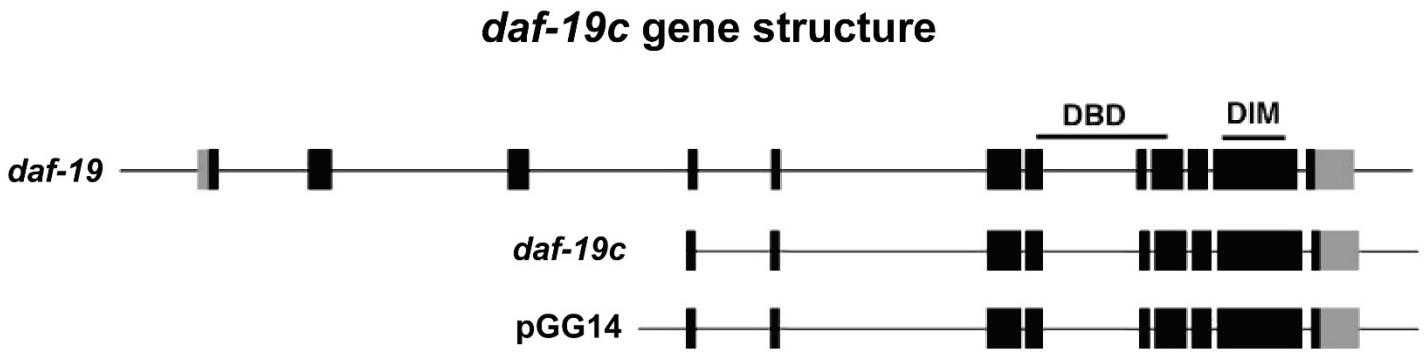


Figure S2 Structure of daf-19 gene in plasmid in LU663. pGG14 starts at the third intron of the genomic daf-19 gene, thus contains all the nucleotide sequences needed for DAF-19C expression. Figure adapted from Senti & Swoboda, 2008.

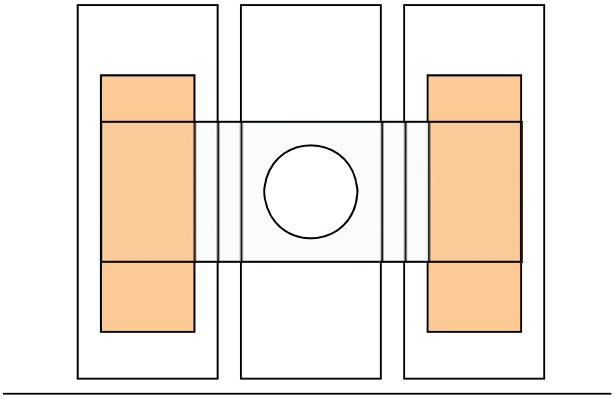
APPENDIX

(Courtesy of Brian Piasecki)

Worm Pad & Slide Preparation (Perform 30-45 minutes before your scheduled microscope time)

- Place 0.75ml tube of 2% agarose in 95C heat block (cap with cover to prevent from opening).
- After agarose has completely melted, transfer the tube to a 45C heat block to begin to cool.
- Wait 1-2 minutes and then add 15µl of 1M sodium azide (20mM solution total).
- Mix the tube by quickly inverting the tube a few times and return it to the heat block (be careful not to allow the agarose to solidify) – repeat if necessary.
- Prepare worms by washing them from plates into a 1.5ml disposable screw-cap tube as follows:
    - Add 1.5 to 2ml of fresh M9 growth medium to a plate.
    - Pipette the liquid up and down several times, using the stream coming out of the pipette tip to dislodge the worms on the plate.
    - Quickly transfer the liquid containing the worms into the tube.
- Wait for worms to settle completely, and then remove most of the liquid (leave only about 200 µl in the bottom) by pipetting the liquid off the top of the pellet.
- When worms are ready and the agarose is prepared, transfer the agarose solution in the heat block from the upstairs to the basement room with the confocal microscope (likely best to have the person working with you meet you in lab first). Bring your worms along with you too!

Use the following image as a guide for making a worm pad



- Line up 3 slides, the two outside slides should contain a strip of labeling tape down the middle.
- Use the P1000 with a blue tip to pipette a drop of agarose (measuring does not work well)

onto the center of a slide that is sandwiched between two slides with tape.

- Quickly place another slide on top so that it rests on the taped slides (the tape prevents the pad from being squeezed too thin).
- After 10-30 seconds, remove the top (or bottom) slide leaving a pad on one of the two of them.

  The remaining slide without agarose can be reused.
- Place 3-5 µl of concentrated worms onto the pad and quickly affix a coverslip.
- Worms will only be good for imaging for ~15minutes.