Spring 6-1-2016

Activation of Target Gene Expression in Neurons by the C. elegans RFX Transcription Factor, DAF-19

Katherine P. Mueller

Lawrence University, katherine.p.mueller@lawrence.edu

Follow this and additional works at: https://lux.lawrence.edu/luhp

Part of the Congenital, Hereditary, and Neonatal Diseases and Abnormalities Commons, Disease Modeling Commons, Genetic Processes Commons, Laboratory and Basic Science Research Commons, Molecular and Cellular Neuroscience Commons, Molecular Genetics Commons, and the Nervous System Diseases Commons

Activation of Target Gene Expression in Neurons by the *C. elegans* RFX Transcription Factor, DAF-19

Katherine Mueller

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

May 2016

ABSTRACT

DAF-19, the only RFX transcription factor found in *C. elegans*, is required for the formation of neuronal sensory cilia. Four isoforms of the DAF-19 protein have been reported, and the m86 nonsense (null) mutation affecting all four isoforms has been shown to prevent cilia formation. Transcriptome analyses employing microarrays of L1 and adult stage worms were completed using RNA from *daf-19(m86)* worms and an isogenic wild type strain to identify additional putative DAF-19 target genes. Using transcriptional fusions with GFP, we compared the expression patterns of several potential gene targets using fluorescence confocal microscopy. Expression patterns were characterized in various genetic backgrounds in order to determine isoform-specific expression patterns. Additionally, we completed rescue experiments using cDNAs encoding specific DAF-19 isoforms in a *daf-19* null genetic background. Our data indicate that several new genes are activated by DAF-19 in both ciliated and non-ciliated neurons. We are currently developing isoform-specific CRISPR/Cas9 mutants to further explore the precise mechanisms by which different DAF-19 isoforms regulate their target genes.

Table of Contents

INTRODUCTION

The mysteries of the brain have long been one of the most elusive areas of medicine. As our understanding of both genetics and anatomy improves, one of the primary questions of neurobiology asks where these two areas intersect - that is, how does gene expression affect neuronal function? Recently, the urgency of this question has increased as we uncover evidence that failures of gene function and regulation underlie many of the most insidious diseases of the nervous system, including Alzheimer's, Huntington's, and Parkinson's disease. The nematode model organism *Caenorhabditis elegans* has proved particularly useful for addressing such questions, due to its well-mapped connectome and cell lineage.

In the present study, we seek to identify and characterize unknown gene targets of the *C. elegans* transcription factor DAF-19, a protein that we believe plays a role in maintaining synaptic protein levels. Mutants of this gene show various signs of neurodegeneration, which may prove an effective model for conditions such as Alzheimer's disease. We have found a variety of novel targets of DAF-19 whose expression patterns in the nervous system are affected by mutations in this transcription factor. Furthermore, we have evidence that two of the protein's lesser-understood isoforms may be responsible for the gene's potential role in synaptic protein level maintenance and neuronal gene expression. The present study seeks to identify and characterize these targets, and also uses CRISPR/Cas9 genome engineering to develop better tools with which to understand the isoform-specific functions of DAF-19.

C. elegans as a model organism

Caenorhabditis elegans is a small, free-living roundworm that thrives in rotting vegetation around the world. Nobel laureate Sydney Brenner first established the use of *C. elegans* as a model organism in the 1960s, after a search for a simple eukaryotic system in which

to study organismal development and the nervous system. Brenner chose *C. elegans* as his preferred model organism for several traits that make it uniquely useful for biological research. The worm is easy to culture thanks to its ability to self-fertilize, its short generation time, and its capacity to thrive on a food source of *E. coli*. Genetic manipulations in *C. elegans* are quite easy to carry out, as hermaphrodites can either be mated with males or maintained through self-propagation, following traditional Mendelian genetics. Further genetic changes can be obtained through mutagenesis screens or the use of site-directed mutagenesis techniques, including TALENs and the CRISPR-Cas9 system. *C. elegans* has an entirely transparent body, which makes it highly useful for anatomical studies, particularly those involving fluorescent imaging of specific structures (Corsi et al., 2015; Figure 1).
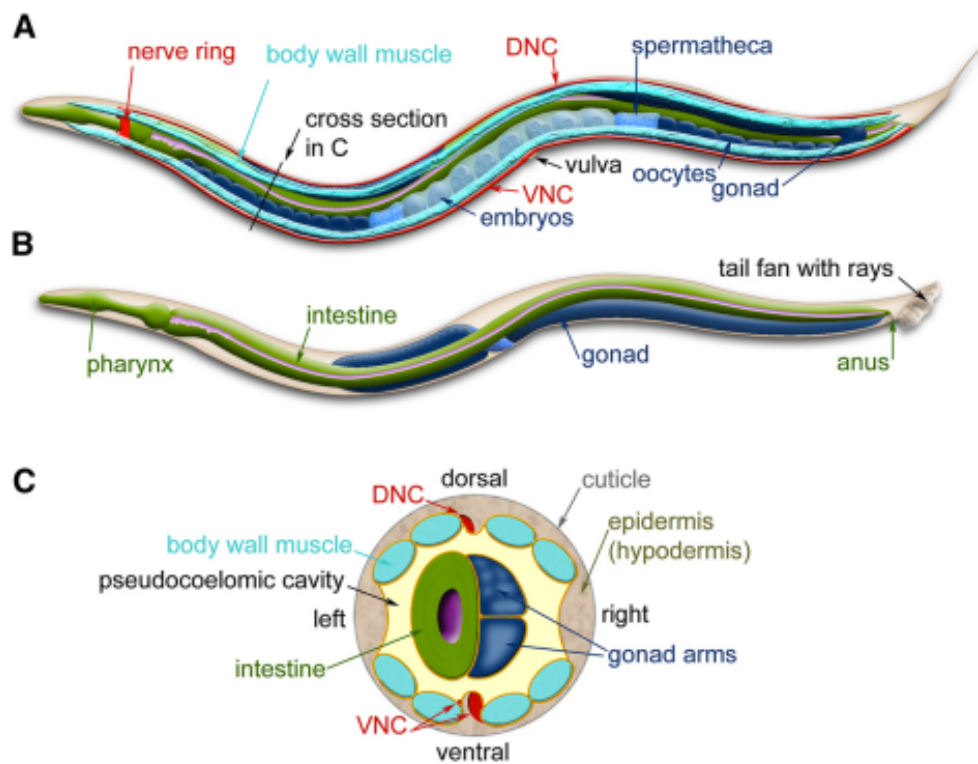

Figure 1: *C. elegans* anatomy. Lateral views of the adult hermaphrodite (A) and the adult male (B), as well as a cross section of the adult worm (C). Structures of note include the dorsal and ventral nerve cords (DNC and VNC), the pharynx, nerve ring, and gonads. Figure adapted from Corsi et al., 2015.

C. elegans adults are approximately 1 mm in length, and grow from egg to sexually mature adult in ~3 days when grown at 20°C. The worm has two sexes: hermaphrodites, which comprise over 99% of wild type populations, and males, which constitute approximately 0.1-0.2% of the population and occur when meiotic non-disjunction events occur (Corsi et al., 2015). The hermaphrodite possesses ovotestis which first produce sperm and then oocytes, which are fertilized from the worm's own spermatheca. The embryos develop independently of their mothers within tough eggshells, and are laid at the 24-cell stage; the full embryogenesis process takes approximately 16 hours at 20°C (Corsi et al., 2015). Once hatched, the worms grow through four larval stages (L1, L2, L3, and L4) with molts between each, before reaching maturity (Figure 2). In crowded, food-limited, or otherwise stressful environments, L2 larvae will forgo the usual developmental path and enter an alternative life stage known as dauer after the second molt. Dauer worms have no oral opening as the cuticle exoskeleton completely covers the animal, and they can survive without eating for several months, until a new food source is found. At this point, the larva will shed its mouth plug and continue development (Corsi et al., 2015).

C. elegans are particularly useful organisms for the study of developmental processes and the neuronal connectome, due to their extremely well-mapped and invariant cell lineage. Adult hermaphrodites have 959 somatic cells, while adult males have 1081 (Sulston & Horvitz, 1977). Our thorough understanding of this cell lineage has led to a much better picture of the early events of embryogenesis, including asymmetric cell division and programmed cell death. Additionally, studies of the connectome have allowed researchers to study the first completely defined neural network. Adult hermaphrodites contain 302 neuronal cells, many of which connect to a structure known as the nerve ring, which serves as the animal's brain (Sulston &

Horviz, 1977). This connectome is simple enough to have been simulated with robots

programmed to move based on representative nerve connections, which mimic simple nematode

behaviors (Szigeti et al., 2014). Our thorough understanding of the worm's small nervous

system ultimately provides an excellent model through which to understand the neural functions

of higher order animals.
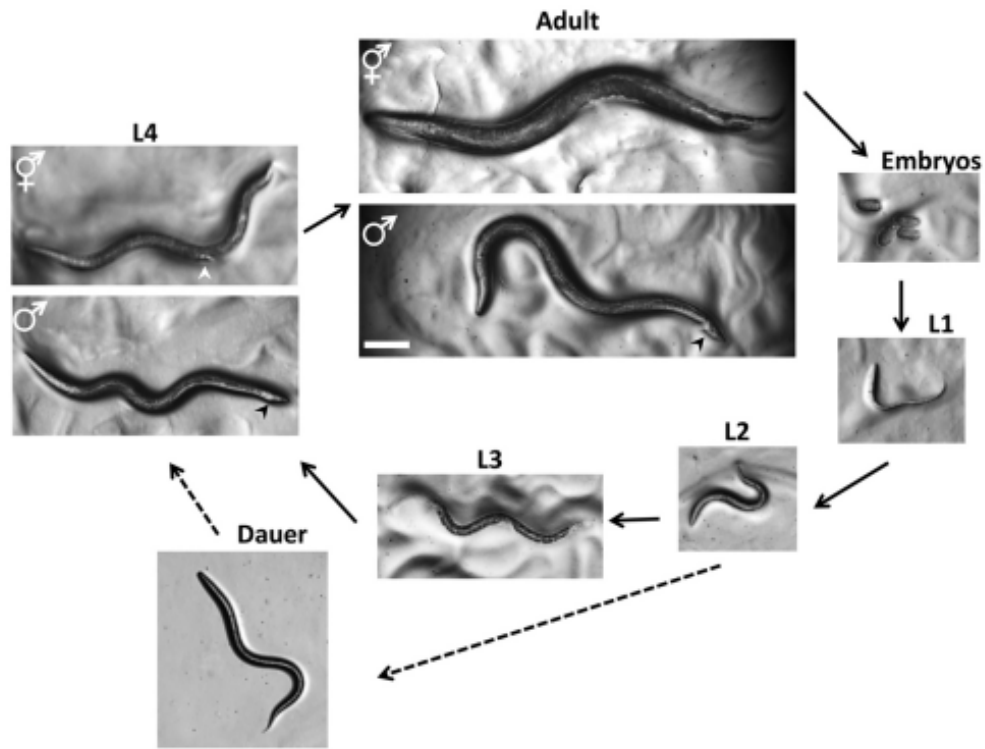
Figure 2: *C. elegans* life cycle. Worms mature from embryo to adult over the course of approximately 3 days, after going through four molts. Starved or otherwise stressed L2 larvae have the capacity to enter the dauer stage, and can survive without a food source for several months. Image adapted from Corsi et al. (2015).
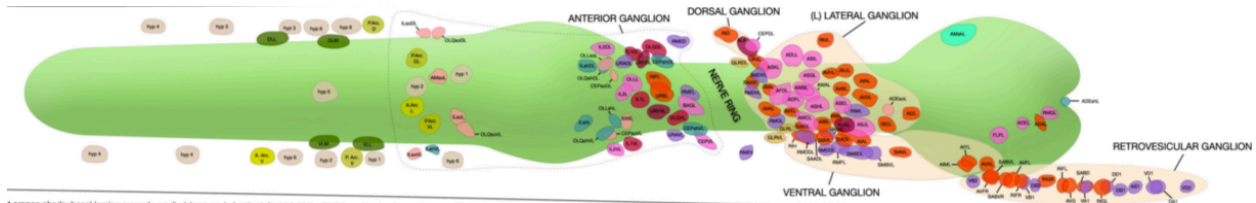
Figure 3: Head neurons in L1 larvae. Image adapted from *wormatlas.org*. The tip of the worm's nose is oriented to the left, with the two pharyngeal bulbs shown in green.

The RFX Transcription Factor DAF-19

The *C. elegans* transcription factor DAF-19 was first identified by Peter Swoboda and colleagues in 2000. Swoboda found that DAF-19 is an RFX-type transcription factor, so named for its characteristic DNA binding domain. It is the only known RFX transcription factor in *C. elegans*. Swoboda's research found that *daf-19* is expressed in all ciliated sensory neurons, and that the gene plays a critical role in ciliogenesis, to the point that *daf-19* null mutants contain no cilia. Additionally, *daf-19* mutants are dauer constitutive, meaning that all L2 stage worms enter the dauer pathway regardless of environmental conditions (Swoboda et al., 2000).

RFX transcription factors belong to the winged-helix family of transcription factors, and were originally identified in the mammalian immune system when the human RFX-1 gene was found to regulate the Major Histocompatibility Complex class II (MHCII) genes (Reith et al., 1988). RFX transcription factors control expression of MHCII and other genes by binding the 14 base pair x-box motif found in the target gene's promoter region. RFX transcription factors are widely conserved, and have been identified in various organisms including humans and mice (Reith et al., 1993), yeast (Wu & McLeod, 1995; Emery et al., 1996), *C. elegans*, *Drosophila*, *Xenopus*, and *D. Rerio* (Emery et al., 1996). They possess a unique DNA binding domain, a dimerizing domain near the carboxy-terminal end, and a flexible "hinge" region near the amino terminus (Reith et al., 1990). The human RFX genes are expressed in various tissues, including (but not limited to) the brain (RFX-1 and RFX-4), the testis (RFX-2 and RFX-4), pancreatic tissues (RFX-6), the kidneys (RFX-5 and RFX-7), various immune tissues (RFX-1 and RFX-5), and in some cases in all ciliated cells (RFX-3) (Aftab et al., 2008). These genes play a major role in ciliogenesis, among other functions, and mutations in these genes have been shown to cause significant medical conditions called ciliopathies. Studies in mice have demonstrated that

the RFX3 transcription factor directs nodal cilium development and left-right symmetry specification, and mutants are highly prone to early embryonic lethality and situs inversus (the reversal of the left-right organ pattern) (Bonnafe et al., 2004). Other major ciliopathies result from mutations in RFX genes, such as Bardet-Biedl Syndrome and primary ciliary diskinesia (Chen et al., 2006; Bonnafe et al., 2004). These pathologies cause a wide variety of symptoms, ranging from renal failure to chronic bronchial inflammation and sperm defects (omim.org). A better understanding of the functions of RFX transcription factors is important for understanding these diseases, many of which are poorly characterized. By studying the function and targets of DAF-19 in *C. elegans*, we may be able to identify novel targets that have human homologues, which may lead us to a better understanding of these disease mechanisms.

One of our primary questions about DAF-19 surrounds its four known splice isoforms. Senti and Swoboda identified three such isoforms, two of which (DAF-19A and DAF-19B) are significantly longer than the third (DAF-19C) (Senti & Swoboda, 2008). A fourth isoform (DAF-19M), found only in male worms, was characterized by Wang et al. in 2010 (Figure 4). Senti and Swoboda gave the first evidence that DAF-19 expression is not restricted to ciliated sensory neurons, as was previously reported (Swoboda et al., 2000). They used RNase protection and N and C-terminal specific antibodies to show that three isoforms of different lengths occur in hermaphrodite worms, and demonstrated with antibody staining that the short form DAF-19C isoform is found exclusively in ciliated sensory neurons, while the long form isoforms A and B are found only in non-ciliated neurons. This finding first suggested that the long form isoforms might play a role unrelated to ciliogenesis.
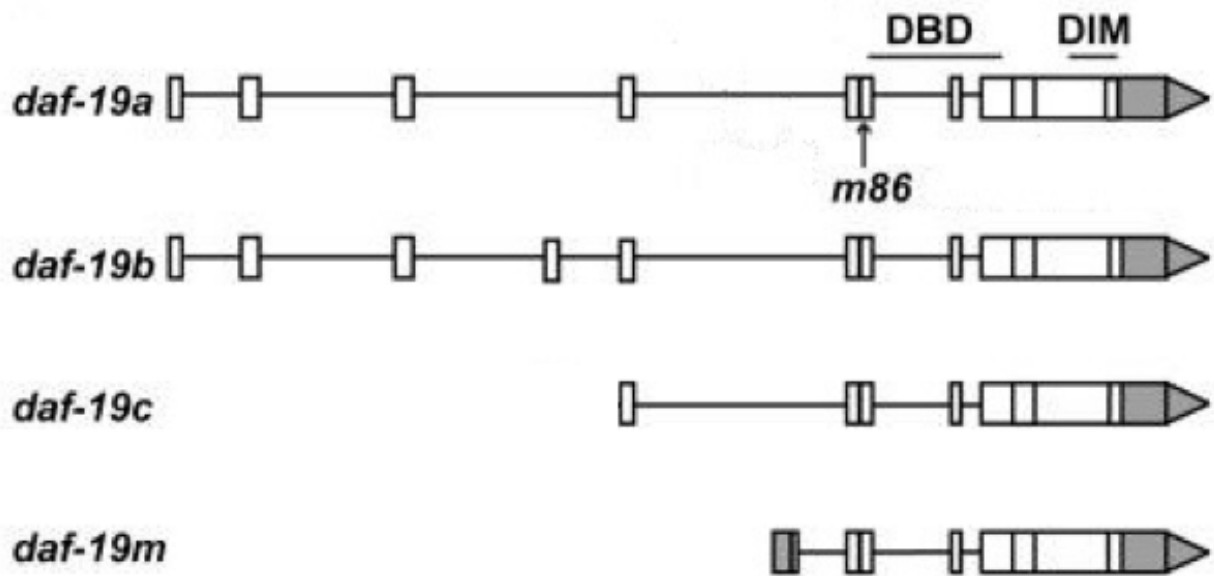
Figure 4: Isoforms of *daf-19*. Exons are depicted by boxes, and introns are depicted by the connecting lines. The DNA binding domain (DBD) and dimerization domain (DIM) are both shown, as is the *m86* null mutation affecting all isoforms. Figure adapted from Wang et al. (2010).

Senti and Swoboda further examined the behavioral patterns of *daf-19 m86* mutants (null for all isoforms), and found defects in dwelling and roaming behaviors. When placed on an agar plate seeded with bacteria, wild type worms alternate between "dwelling" behaviors where they stay in one part of the plate to feed (about 80% of the time) and "roaming" behaviors" where they rapidly crossing the entire bacterial lawn (Fujiwara et al., 2002). Senti and Swoboda found that in *m86* null mutants, worms did not show roaming behavior (Figure 5). This is consistent with ciliary defects, as worms without cilia in their sensory neurons are unable to properly smell food; thus, they do not seek it out. Interestingly, when Senti and Swoboda performed rescue experiments with isoform-specific rescue constructs, they found that isoforms A and C were each able to partially rescue dwelling and roaming behavior, but only worms with a full length genomic construct containing all three isoforms showed a completely rescued phenotype. This

suggests that the behavioral defects were not only caused by a lack of cilia, but rather reflect

additional neurodegenerative effects.
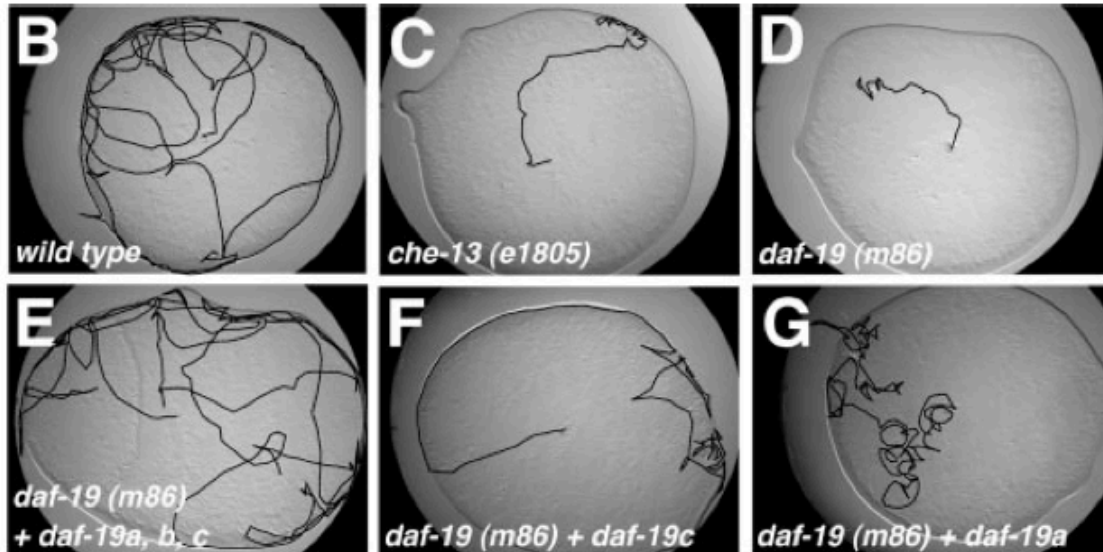
Figure 5: Dwelling and roaming assay. Wild type worms (B) show both behaviors, and completely cover the plate. *m86* (D) and *che-13* (C) mutants, which prevent cilia growth, fail to roam across the plate. Rescue experiments with genomic constructs of *daf-19C* (F) and *daf-19A* (G) each partially rescue the roaming phenotype, and the full length genomic construct (E) fully rescues the behavior. Figure adapted from Senti & Swoboda (2008).

Further assays with the paralytic drugs aldicarb and levamisole showed the possibility of

either a presynaptic or postsynaptic defect that affected the worm's ability to process

acetylcholine (a neurotransmitter). Additionally, Senti and Swoboda found decreased expression

of synaptic vesicle proteins in *daf-19* mutants, including UNC-64/syntaxin, IDA-1, UNC-17,

SNB-1, and SNT-1. However, none of these proteins appear to be directly regulated by DAF-19,

as each lacks an x-box in its promoter region, and they also showed no difference in transcript

levels between wild type and *m86* worms. Thus, while isoforms A and B appear to play a

significant role in maintaining synaptic proteins, they appear to do so indirectly (Figure 6; Senti

& Swoboda, 2008).
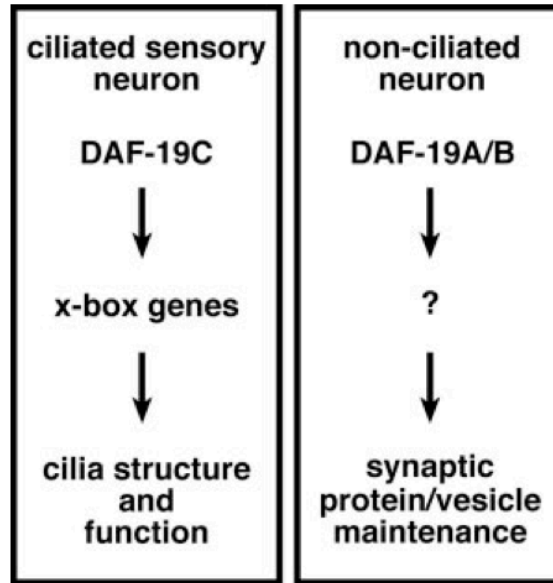
Figure 6: Potential functions of DAF-19 isoforms. Daf-19C is known to regulate ciliogenesis by binding the x-box sequences of target genes. DAF-19A/B appear to play a role in maintaining synaptic protein levels, although the mechanism for this is unknown. Figure adapted from Senti & Swoboda (2008).

De Stasio Lab: Past work

Since Senti and Swoboda's 2008 finding that the isoforms of DAF-19 may play different regulatory roles, our lab has been attempting to determine the targets of and mechanisms by which this process occurs. Elizabeth De Stasio first completed a microarray comparing relative gene expression levels across the transcriptome in adult worms that were either wild type or *m86* mutants of *daf-19*. She used worms that had been adults for two days due to their fully developed cilia, and because the loss of synaptic proteins had been seen most dramatically at this age (Senti & Swoboda, 2008). Multivariate analysis of the data resulted in a list of 180 genes with 1.5-fold or greater difference in expression between *m86* and wild type worms. From this list, various genes were chosen for further study based on several criteria, including protein domains suggesting a neuronal function or expression pattern, the presence of a human ortholog, or ready availability of mutant alleles. This list complements two additional microarrays

completed by Prasad Phirke, which examines differential expression in *m86* or wild type worms across the transcriptome in L1 larvae and 3-fold stage embryos. From the three lists, a total of 30 genes were selected for further analysis. Transcriptional fusions with green fluorescent protein (GFP) were generated for each of these genes, such that a plasmid containing the promoter region of the gene in question immediately precedes the open reading frame of GFP. An additional second marker such as mCherry was included in the plasmid as a positive control to identify transgenic worms. These were injected separately into two lines of worms, one wild type for *daf-19* and one containing the *m86* mutation. Since then, we have identified over ten genes that show visible *daf-19* dependence based on these transcriptional fusions, including several for which gene expression appears to be dependent on one or more isoforms of DAF-19.

CRISPR Mutagenesis

In order to study the different effects of the various isoforms of DAF-19, it is useful to be able to observe the expression patterns of putative target genes in *daf-19* isoform-specific mutants. As previously stated, our lab currently has several such mutants that have null alleles of *daf-19a* and *daf-19b*; however, we do not yet have any mutants that are specific to *daf-19c*. Of course, as all of the exons in DAF-19C are conserved in DAF-19A and DAF-19B, it is challenging to create a mutant that knocks out only the short isoform without affecting its two longer cousins. Such potential mutations are highly specific, and would be unlikely to arise from traditional mutagenesis screens; as such, we have need for a site-directed mutagenesis approach. Various tools have been described to create such specific mutations, including TALENS, first identified in 2007 by Römer et al., and zinc finger nucleases, first created by Kim, Cha, and Chandrasegaran in 1995. Today, the gold standard of genome editing is the CRISPR/Cas9 system, originally described by Jinek et al. (2012).

CRISPR, or "clustered regularly interspaced short palindromic repeats," is a genetic feature of some bacteria and archaea that provides adaptive immunity against viruses and plasmids (Wiedenheft et al., 2012). CRISPRs consist of short repeated elements of approximately 29 nucleotides separated by 32 nucleotide 'protospacer' DNA. These sequences contain DNA from viruses or plasmids with which the bacteria has previously come in contact, and serve as a genetic record of such prior encounters that functions as a kind of genetic vaccine for the bacteria. In addition to the CRISPR sequences, the bacteria also produce proteins from the *Cas* gene family of nucleases, which flank the CRISPR loci (Wiedenheft et al., 2012). When the CRISPR loci are transcribed, the various Cas proteins can pair with these sequences, which then recognize their complements in invading viruses or plasmids. The Cas proteins then induce double strand breaks in the invader's DNA, thus eliminating the threat (Figure 7).
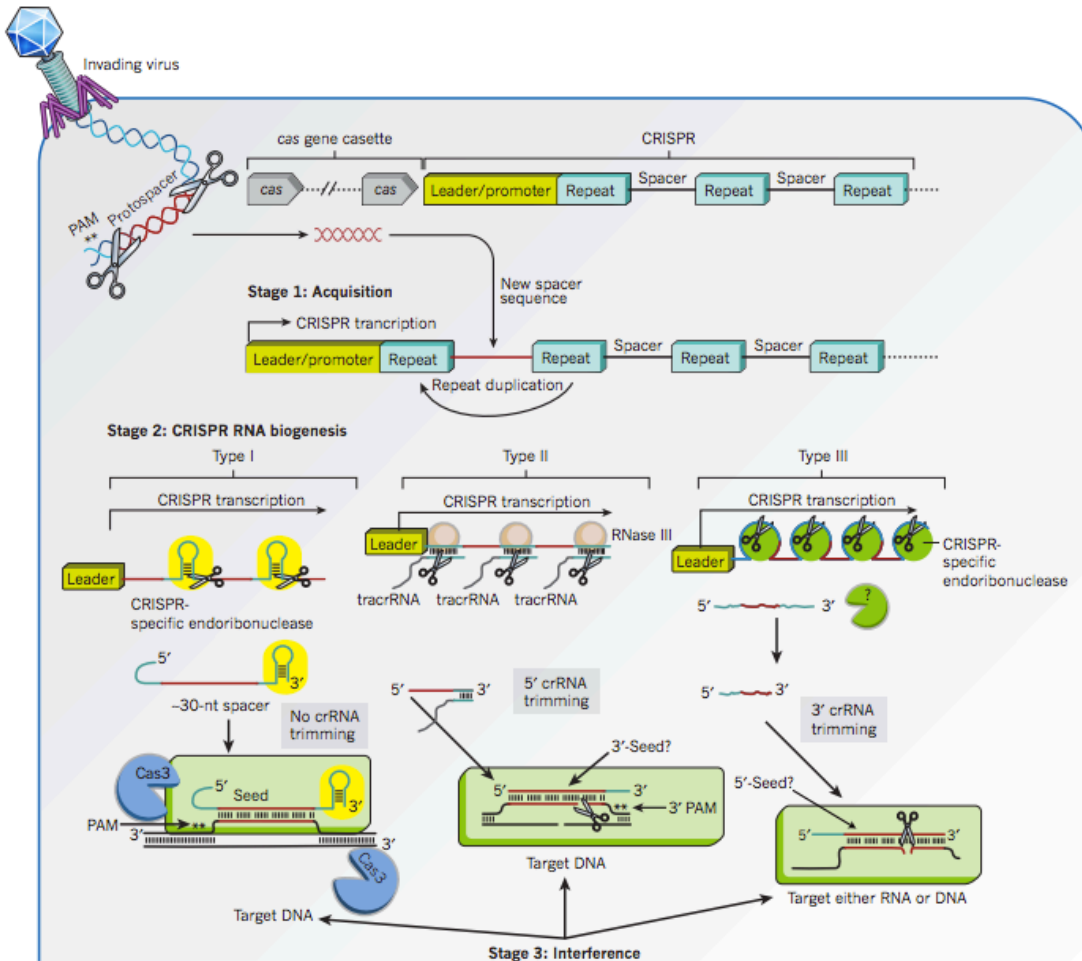
Figure 7: CRISPR-mediated adaptive immune systems in archaea and bacteria. Various CRISPR systems in the bacterial genome address invading threats by recognizing viral or plasmid DNA. Figure adapted from Wiedenheft et al. 2012.

CRISPRs were described as genome editing tools by the Doudna lab, which showed in 2012 that the system can be engineered to produce double strand breaks at nearly any desired locus (Jinek et al., 2012). Together with Emanuelle Charpentier, Doudna showed that a CRISPR RNA (called a crRNA) identifies the region of DNA to cleave, known as the protospacer. This crRNA base pairs to a trans-activating crRNA (tracrRNA), which directs the Cas protein (in this case Cas9) to the protospacer. Cas9 recognizes a short 5'-NGG-3' motif adjacent to the protospacer, known as the PAM sequence, which consists of any nucleotide (N) followed by two guanine nucleotides. The PAM sequence is necessary for Cas9 to bind and

cleave target DNA. Doudna and colleagues showed that it is possible to engineer a single RNA chimera containing both crRNA and tracrRNA which can be used to direct Cas9 to cleave DNA at specific loci as desired. crRNAs can be designed to target nearly any part of the genome, thus allowing a wide range of mutagenic applications for this technology.

After Cas9 induces a double strand break in the DNA, the body will attempt to repair the damage by ligating the ends of the broken fragments together through one of two pathways. The first, non-homologous end joining (NHEJ), was discovered in 1980s. In this pathway, the ends of the cleaved DNA are resected and then ligated together, and there are often deletions of a few bases that may affect the function of the gene products. As such, NHEJ is often used to generate knockout mutants. However, to produce specific mutants, the homologous recombination pathway can be used to make precise changes in the modified genome (Figure 8). In this pathway, a similar strand of DNA (for example, on the sister chromatid) is used as a template for the broken section. In CRISPR mutagenesis, this pathway can be exploited by introducing a repair template such as a plasmid or a small oligonucleotide, which spans the area with the break and has matching sequences known as homology arms on either end; the desired mutation is built into the middle of the template (Addgene, 2016). This process can be used to generate a wide range of mutants, from single base pair changes to whole gene insertions.
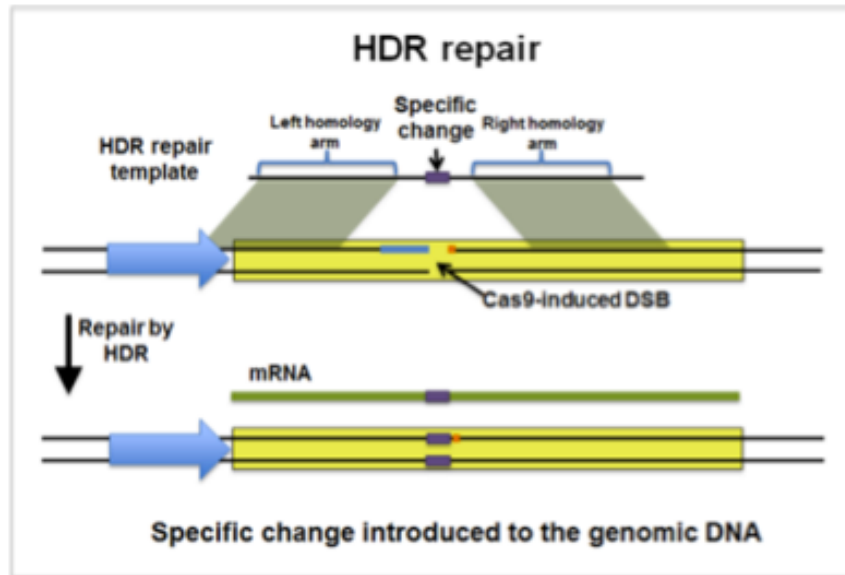
Figure 8: Homology-directed repair in a CRISPR mutant. Figure adapted from Addgene (2016).

In the past few years, CRISPR technology has become a revolutionary tool for the scientific community, as it provides an efficient method of in vivo site-directed mutagenesis. CRISPR was first adapted for use in *C. elegans* in 2013 by Friedland and colleagues. In order to introduce mutations into the germline, the worm gonad is injected with a mix of Cas9 protein, crRNAs, and a homology repair template, if this pathway is to be exploited. However, the process of screening for progeny with the desired mutation can be labor intensive. In order to mitigate this issue, a co-conversion strategy developed by the Seydoux lab can be used to improve screening efficiency (Paix et al., 2015). Based on a protocol by Arribere et al. (2014), Paix and colleagues created a cloning-free approach yielding high efficiency edits in *C. elegans* (Figure 9). In order to streamline the screening process, they designed a CRISPR strategy that creates two separate mutations in the genome: one at the locus of interest, and one on a separate chromosome that produces an easily visible dominant phenotype at the *dpy-10* gene. *dpy-10* mutants display a characteristic roller behavior pattern which is easy to identify on sight. As

such, progeny with at least one successful mutant can be identified without requiring labor-

intensive screening. These worms are far more likely to have received sufficient Cas9 protein,

crRNAs, and repair template copies for a successful mutagenesis at the site of interest, with a

success rate as high as 70% (Paix et al., 2015). Additionally, their protocol uses a direct-delivery

approach that requires no cloning, as the tracrRNA, crRNA, and repair oligonucleotide are all

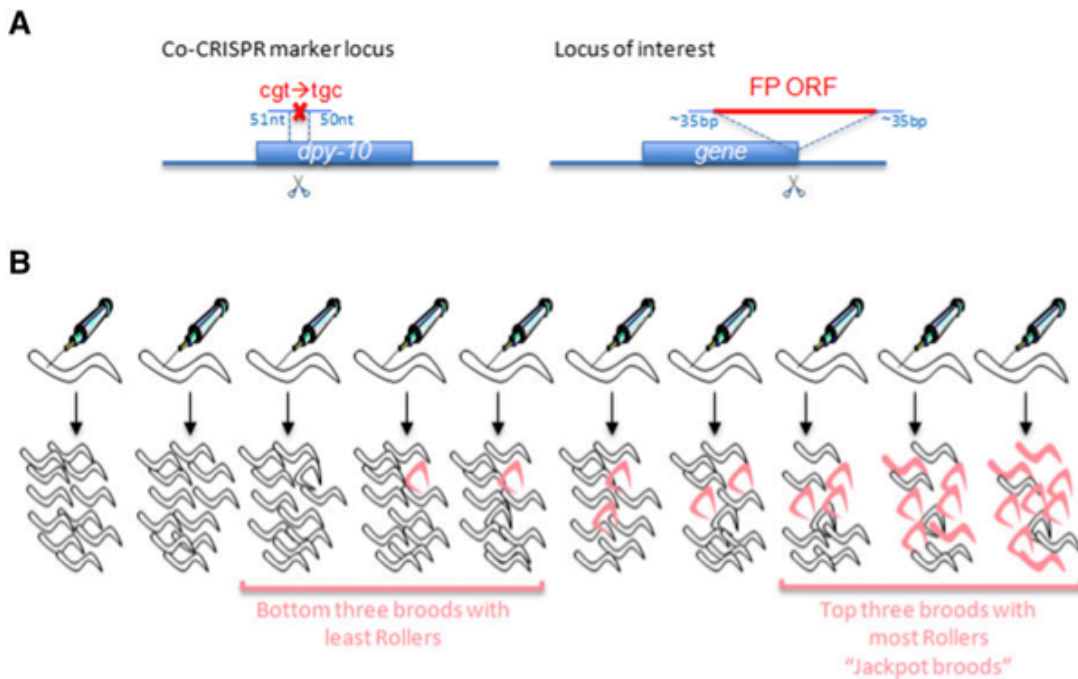synthesized chemically and injected together with Cas9.

Figure 9: Co-CRISPR strategy. (A) The marker mutation at *dpy-10* is shown along with the mutation at the locus of interest (Insertion of a fluorescent protein at the desired site). (B) Parent worms are injected and screened for roller progeny; those with high proportions of roller worms are termed "jackpot broods." Figure adapted from Paix et al., 2015.

Present Work

The present study seeks to continue our work in identifying gene targets of DAF-19, with a particular emphasis on determining isoform specificity. Five previously uncharacterized genes identified by previous microarray experiments, T01B11.2, T07F10.1, *del-4*, *srd-61*, and *decr-1.1* were tested for DAF-19 dependence in a variety of mutant backgrounds. Prasad Phirke and others generated transcriptional fusions of each of these genes with GFP, and injected them into worms that were either *m86* or wild type for *daf-19*. We then generated isogenic strains for each line by mating each injected strain with either the wild type or *m86* strains, such that each respective injection led to a pair of worm strains. GFP expression patterns of each target gene were examined using confocal microscopy, with an emphasis on neuronal expression. If a particular gene demonstrated differential expression patterns between wild type and *m86* backgrounds, further strains were created that contained isoform-specific mutants. We obtained a *daf-19a/b* mutant called *tm5562* from the Mitani lab, which contained an 865 bp deletion of exon 2 and a portion of the intron flanking sequence. Two further mutants were created by Deborah Sugiaman at the Swoboda lab. Of these, one (*of5)* affects both DAF-19A and DAF-19B via a 12 bp deletion with a +1 shifted ATG start codon, while the other (*of6*) contains a 53 bp deletion of exon 4 that affects only DAF-19B. These mutants were used to determine which isoform was responsible for the differential expression observed in the *m86* mutant (Figure 10).
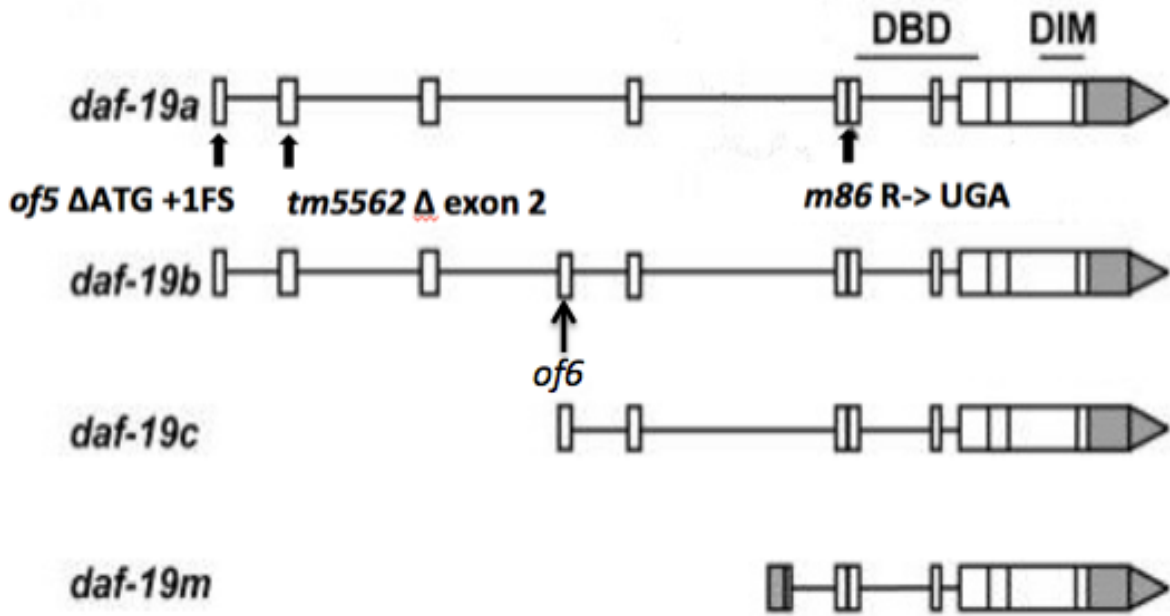
Figure 10: *daf-19 isoforms.* The locations of the *m86, of5, tm5562,* and *of6* mutations are shown with black arrows. Figure adapted from Wang et al. (2010).

In addition to characterizing the expression patterns of the five genes, we sought to develop an isoform-specific mutant of *daf-19* that would knock out DAF-19C only. As each exon of DAF-19C is conserved in DAF-19A/B, this is a somewhat difficult proposition. We designed a CRISPR mutagenesis strategy based on the previously described co-CRISPR approach developed by Paix and colleagues (2015). We elected not to use *dpy-10* as the co-conversion marker mutation, as it is on the same chromosome as *daf-19* and would therefore be difficult to mate out of successful mutants. Instead, we chose to insert a copy of GFP into the open reading frame of *gtbp-1*, a non-essential protein that is nevertheless expressed in most tissues. To generate the CRISPR mutant, we designed two approaches intended to knock out DAF-19C without affecting the function of the two long form isoforms. The first approach seeks to replace the ATG start codon in exon 5 (the first exon of DAF-19C) with the codon for alanine instead of methionine. We selected alanine for its small size and nonpolar character, which we

expect to play a chemical role similar to that of methionine in the two long-form isoforms; however, we hope that the lack of the initial start codon will be sufficient to stall translation of DAF-19 C (Figure 11). In the event that this approach is unsuccessful, our second design introduces a double frameshift mutation that creates a nonsense mutation in DAF-19C that is subsequently corrected for in DAF-19A (Figure 11). In addition to our CRISPR design, we optimized a purification protocol for Cas9 protein based on that provided by Paix and colleagues. Ultimately, we hope that our new mutant will allow us to more thoroughly understand which genes are activated and repressed by DAF-19A/B, in order to better understand their roles in the synapse.
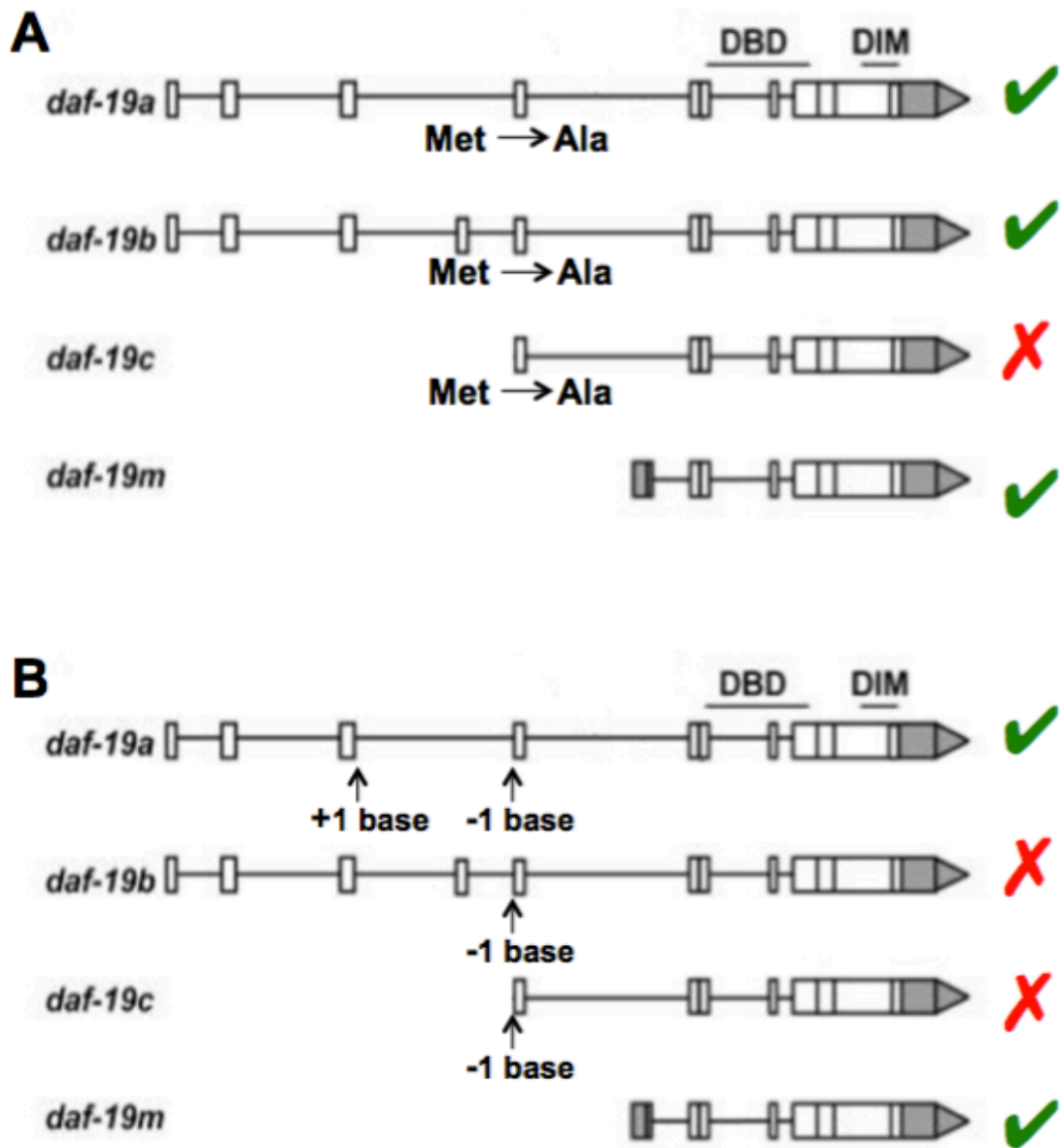
Figure 11: CRISPR mutagenesis approaches. Figure A (mutagenesis #1) shows the location of a single amino acid change from methionine to alanine at the start codon of exon five, which is expected to stall translation of isoform C without significantly affecting the other isoforms. Figure B (mutagenesis #2) shows the location of a double frameshift mutation that removes a single base from the beginning of exon five and the addition of that same base at the end of exon three. This is expected to result in single frameshift mutations that will knock out isoforms B and C while preserving the function of isoform A. Green checkmarks note hypothesized functional isoforms, while red X's show the presence of null alleles. Figure adapted from Wang et al. 2010.

MATERIALS AND METHODS

C. elegans strains and maintenance

All strains were maintained in a 15º or 20º incubator on 5 mm nematode growth medium

(NGM) agar plates, which were seeded with an *E. coli* OP50 lawn (Stiernagle, 1999) To

maintain each strain, hermaphroditic worms were transferred with a sterilized platinum pick to a

fresh agar plate approximately every five days, or by "chunking" a section of populated agar with

a sterilized spatula. Plates were wrapped with Parafilm strips to prevent drying or contamination

Lab	Strain	Genotype	Parental Strain	Constructed By
LU	495	pT07F10.1::GFP; *daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V*	BC14205	Baillie Lab & Yagmur Esemen
LU	496	pT07F10.1::GFP; *daf-19+; daf-12(sa204)X; him-5(e1490)V*	BC14025	Yagmur Esemen
LU	627	pC01B4.5::GFP; *daf-19+; daf-12(sa204)X; him-5(e1490)V*	OE4133	Katie Mueller
LU	630	pT01B11.2::GFP; *daf-19+; daf-12(sa204)X; him-5(e1490)V*	OE3895	Katie Mueller
LU	628	*daf-19(tm5562)II*, 6x backcrossed	LU621	Savannah Vogel and Kristen Bischel
LU	641	pT01B11.2::GFP; *daf-19(tm5562)II; daf-12(sa204)X*	OE3895	Katie Mueller
LU	642	pT07F10.1::GFP; *daf-19(tm5562)II; daf-12(sa204)X*	LU495	Katie Mueller
LU	646	pT01B11.2::GFP; *daf-19(of5)II; daf-12(sa204)X*	OE3895	Katie Mueller
LU	648	pT07F10.1::GFP; *daf-19(of5)II; daf-12(sa204)X*	LU495	Katie Mueller
LU	653	pT01B11.2::GFP; *daf-19(of6)II; daf-12(sa204)X*	OE3895	Katie Mueller
LU	655	pT28B8.5::GFP; *daf-19+; daf-12(sa204)X; him-5(e1490)V*	OE3912	Katie Mueller
LU	663	*daf-19c::daf-19(m86); daf-12(sa204)X*	OE3492	Brian Piasecki
LU	675	pT07F10.1::GFP; *daf-19a::daf-19(m86); daf-12(sa204)X*	LU495	Katie Mueller
LU	676	pT07F10.1::GFP; *daf-19c::daf-19(m86); daf-12(sa204)X*	LU495	Katie Mueller

LU	680	pT01B11.2::GFP; *gcy-32p*::mCherry; *daf-12(sa204)X*	LU630	Katie Mueller
LU	682	pT07F10.1::GFP; *gcy-32p*::mCherry; *daf-12(sa204)X*	LU495	Katie Mueller
LU	683	pT01B11.2::GFP, *daf-19c::daf-19(m86)*; *daf-12(sa204)X*	OE3895	Katie Mueller
LU	685	pT01B11.2::GFP, *daf-12(sa204)X, pha-1(e2123)III; him-5(e1490); otIs544 (cho-1::SL2::mCherry pha-1+)*	LU630	Sophie Scholtz
OE	3199	*daf-19a::daf-19(m86); daf-12(sa204)X*	OE3492	Gabi Senti
OE	3492	*daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V*	N/A	Swoboda Lab
OE	3507	pF43C11.3:GFP; *daf-12(sa204)X*	N/A	Swoboda Lab
OE	3705	pT01B11.2::GFP; *daf-19+; daf-12(sa204)X*	N/A	Swoboda Lab
OE	3738	*daf-19+; daf-12(sa204)X; him-5(e1490)V*	N/A	Swoboda Lab
OE	3869	pF43C11.3:GFP; *daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V*	N/A	Swoboda Lab
OE	3895	pT01B11.2::GFP; *daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V*	N/A	Swoboda Lab
OE	3912	pT28B8.5::GFP; *daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V*	N/A	Swoboda Lab
OE	4124	pC01B4.5::GFP; *daf-12(sa204)X*	N/A	Swoboda Lab
OE	4133	pC01B4.5::GFP; *daf-19(m86)II; daf-12(sa204)X; him-5(e1490)V*	N/A	Swoboda Lab
OF	5	*daf-19(of5)II*	N/A	Deborah Sugiaman
OF	6	*daf-19(of6)II*	N/A	Deborah Sugiaman
AX	2419 dbEx864	*gcy-32p*::mCherry	N/A	De Bono Lab
OH	13636	*pha-1(e2123)III; him-5(e1490); otIs544 (cho-1::SL2::mCherry pha-1+)*	N/A	Hobert Lab

Table 1: *C. elegans* **strains used and analyzed.**

Strain construction

All transgenic LU strains were constructed as isogenic pairs, meaning two strains with
different chromosomal mutations contain identical transgene arrays. All OE strains were
produced by separate microinjections in either *daf-19* mutant or wild type backgrounds (Table
1). To produce crosses, 12 males from one parent strain and three L4 hermaphrodites from the

second parent strain were picked to a fresh NGM plate and allowed to mate. Strains containing

the *him-5(e1490)V* allele produced high numbers of males to allow for easy mating; other male

populations were generated by heat shocking individual L4 hermaphrodites for 4-5 hours at 30º

C, allowing them to self-sex, and mating resulting male progeny with hermaphrodites of the

same strain.

To produce isogenic WT strains, hermaphrodites from transgenic *daf-19(m86)* strains

were mated with OE 3738 males. F1 progeny were dye filled, and transgenic dye filling

hermaphrodites were singly picked and allowed to self-sex. This step was repeated until 100%

of progeny were dye filling, indicating a homozygous *daf-19+/+* allele. A similar method was

used to produce isogenic *daf-19(tm5562)II*, *daf-19(of5)*, and *daf-19(of6)* strains, using males

with each respective mutation instead of OE 3738. Each of these mutants dye fill, so the same

assay was used to isolate progeny carrying the desired allele. This approach was also used to

create double transgenic rescue and marker strains. Several example crosses created to study the

gene T01B11.2 are shown in Table 2.

Strain	Crosses	Genotype
LU630	OE3895 ☿ x OE3738 ♂	pT01B11.2::GFP; *daf-19+; daf-12(sa204)X; him-5(e1490)V*
LU641	OE3895 ☿ x LU628 ♂	pT01B11.2::GFP; *daf-12(sa204)X; daf-19(tm5562)II*
LU646	OE3895 ☿ x OF5 ♂	pT01B11.2::GFP; *daf-12(sa204)X; daf-19(of5)II*
LU653	OE3895 ☿ x OF6 ♂	pT01B11.2::GFP; *daf-12(sa204)X; daf-19(of6)II*
LU680	LU630 ☿ x AX2419 ♂	pT01B11.2::GFP; *daf-12(sa204)X; gcy-32p*::mCherry
LU683	OE3895 ☿ x LU663 ♂	pT01B11.2::GFP, *daf-19c::daf-19(m86); daf-12(sa204)X*

Table 2: Example strain construction. Three L4 hermaphrodites and 12 adult males were mated for each
respective cross. Dye filling assays were used to isolate the desired phenotype, and were repeated until a
homozygous generation was produced.

Dye filling assay

DiI staining, based on the protocol by Tong and Burglin (2010), stains amphid and phasmid neurons in worms with functional cilia; thus, it can identify worms with either *daf-19+/+, daf-19(tm5562), daf-19(of5),* or *daf-19(of6)* alleles. Strains containing the *daf-19(m86)* allele show no staining of the amphid or phasmid neurons; as such, this is a useful tool for determining *daf-19* genotype in new strains or to identify worms containing an extrachromosomal *daf-19c* rescue construct, which restores ciliary development. Finally, dye filling can be used as an amphid and phasmid neuronal marker during confocal analysis, providing a useful positional marker when attempting to discern GFP expression patterns in transgenic worms. See Appendix A for details.

Confocal microscopy and anatomical analysis

Worm pads were created by melting 2% agarose and adding sodium azide to 10 mM. Approximately 50 μl of solution was pipetted onto a slide and covered by a second slide that was elevated slightly by two pieces of tape on adjacent slides (Figure 12). Worm pads were kept in a humid chamber. Worms from an unstarved plate were then washed off an agar plate with 1 mL of M9 buffer and allowed to settle in an Eppendorf tube. 3 μl of worms were transferred to a worm pad with 3 μl of extra supernatant and the slide was tilted to evenly distribute the worms before covering the pad with a cover slip. Transgenic expression patterns were analyzed using a Leica SP5 confocal microscope and LAS AF SP5 software. Images were taken with a 40X lens at zoom 2.0 for L4 and adult worms, zoom 3.0 for L3 worms, and zoom 4.0 for L1 and L2 worms (settings in Table 3). Images were compared to images and diagrams from *wormatlas.org* in order to elucidate expression patterns.

Type of Signal	Settings
GFP alone	Argon laser on at ~20% power HeNe 543 laser ~10% HyD2 detector 495 – 595 nm for GFP
GFP with Mcherry/DiI/DS red	Argon laser on at ~20% power HeNe 543 on at ~20% 594 laser at ~10% 488 laser ~10% HyD2 detector 495 – 535 nm for GFP HyD4 detector 650 – 790 nm for DiI or Mcherry

Table 3: Confocal settings for GFP and DiI/mCherry/DS red images. All images were collected using a Leica SP5 confocal microscope and LAS AF SP5 software using the 40X lens.
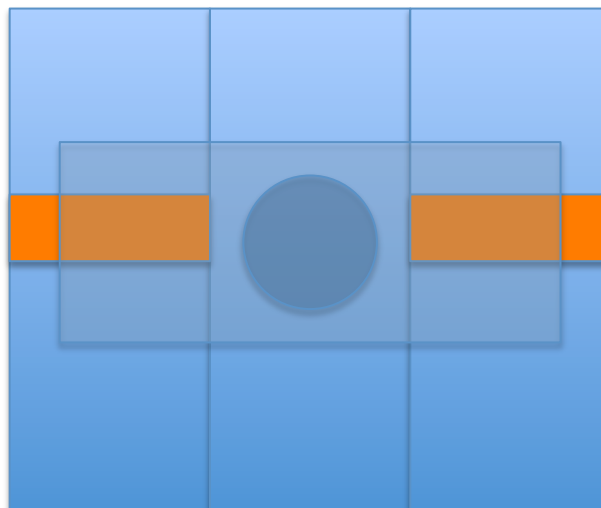
Figure 12: Worm pad preparation. A drop of molten 2% agarose with 10 mM sodium azide was pipetted onto one slide, then covered immediately with a second slide elevated by two adjacent pieces of tape.

CRISPR design and Cas9 purification

Two CRISPR mutagenesis designs were created based on the work of Paix and colleagues

(2015), which are detailed further in the results section of this thesis. Cas9 protein was purified

for use in CRISPR mutagenesis according to a protocol supplied by the Seydoux lab using the

nm2973 plasmid in *E. coli*, which they generously provided (Paix et al., 2015). This plasmid

contained a T7 viral promoter upstream of the Cas9 coding sequence, in addition to a

carbenicillin resistance gene that could be used to select for plasmid-containing cells. The

bacteria were induced to produce an excess of Cas9 protein using IPTG, which was then purified

using Ni affinity and Sepharose chromatography and Pierce Cassette dialysis. Slight

modifications to their protocol are noted with asterisks in Appendix B.

RESULTS

The present study seeks to identify putative DAF-19 target genes and determine specifically which isoforms of DAF-19 control each target. Beth De Stasio and Prasad Phirke identified genes that were differentially expressed in *daf-19(WT)* vs. *daf-19(m86)* backgrounds via transcriptome analyses of 3-fold embryonic, L1 larval, and adult worms (Phirke et al., 2011). From the lists of genes generated from L1 and adult worms, 30 genes were selected as likely targets of regulation by DAF-19, and transcriptional fusions of the promoter region of each gene with GFP were created and separately microinjected into the gonad of worms with a wild type *daf-19* gene. The present study seeks to characterize the expression patterns of five of these genes: T01B11.2, T07F10.1, *del-4*, *srd-61*, and *decr-1.1*. For each transgene, isogenic *daf-19+/+* (WT) and *daf-19(m86)* worms were examined via confocal microscopy (Table 4). Those that showed differential expression in these backgrounds were further analyzed by mating the transgenes into worms containing one of three *daf-19* mutations: *tm5562*, *of5* and *of6*, all of which affect the various isoforms of *daf-19* differently (see Figure 4). These strains could then be used to determine whether differential expression patterns were dependent on one or more DAF-19 isoform. Furthermore, rescue experiments were done for several genes that were shown to be DAF-19 dependent, by mating *m86* worms with strains that contained cDNAs encoding either the DAF-19A or DAF-19C isoforms. These reciprocal sets of experiments helped us to determine how DAF-19 regulates gene expression.

Representative images from *wormatlas.org* were used to identify anatomical features in which GFP was expressed, with a particular focus on identifying specific neurons. A minimum of 20 worms of all ages were imaged and analyzed for each strain, in order to establish representative phenotype profiles. The *C. elegans* nervous system has been extremely well

mapped, and many functions of individual neurons have been elucidated. Thus, this study sought to identify neuronal expression patterns as precisely as possible, in the hopes that the known functions of these neurons would provide insights into potential new roles for *daf-19,* particularly in synapse maintenance and other neuronal functions. In order to narrow down potential neuronal expression patterns, the amphid neurons of WT worms were stained with a lipophilic dye (DiI), thus providing anatomical landmarks for neuronal identification. Additionally, several double transgenic lines were created by crossing the relevant strains with worms containing a transgenic array that individually labeled a particular neuron, or set of neurons, with the fluorescent protein mCherry. These were used to verify neuronal identity by looking for co-localization of GFP and mCherry in neurons.

Gene name	Fold change (adult)	Fold change (L1 larvae)	Fold change (3-fold embryo)	Presence of x-box binding motif
T01B11.2	1.16	0.62	0.89	Present
T07F10.1	0.43	1.0	0.64	Not present
del-4	1.3	0.4	1.14	Present
srd-61	1.48	2.2	1.83	Not present
decr-1.1	2.6	0.41	0.46	Not present

Table 4: Relative transcription levels and sequence data for genes of interest. In most cases, genes were selected for further analysis due to a fold change of >1.5 or < 0.5 in at least one array. Fold change is defined as the relative transcription level of the gene in *daf-19(m86)* worms compared to wild type worms. Numbers shown are averages of three microarray experiments per strain. Interestingly, many of these differentially expressed genes lacked the characteristic x-box binding motif through which *daf-19* is known to control gene expression.

Characterizations of gene expression patterns

T01B11.2 is activated by DAF-19

T01B11.2 was identified as a putative *daf-19* target gene from Prasad Phirke's microarray analysis of L1 stage worms only. It is an ortholog of the human genes ETNPPL and PHYKPL, both of which are phospho-lyases. Phospho-lyases are a subcategory of lyases, enzymes that catalyze cleavage reactions (Lehninger et al., 2008); phospho-lyases act

specifically on phosphorus atoms. No clinical phenotypes are associated with either human homolog. T01B11.2 is known to be involved in embryonic development, and RNAi screens have shown that a lack of T01B11.2 is embryonic lethal; however, little else is known about its function or mechanisms of control (Simmer et al., 2003). Based on protein domain information, it is predicted to bind pyridoxal phosphate (PLP), and likely also has transaminase activity (wormbase.org). Transaminases catalyze the transfer of amino groups from α-amino to α-keto acids, an important step in catabolism. Thus, this gene may play a role in amino acid metabolism. Additionally, it has been shown to interact with the SGT-1 protein, which plays a role in embryonic and larval development, as well as reproductive processes (Li et al., 2004).

Transcriptome analysis of T01B11.2 by Elizabeth De Stasio and Prasad Phirke showed a 1.16 fold change in adult worms (up-regulation in *daf-19 m86* worms), a 0.62 fold change in L1 larvae, and a 0.89 fold change in 3-fold embryos (down-regulation in *daf-19 m86* worms). These numbers were below the thresholds that were considered to be significantly differentially expressed, as most genes were only analyzed further if they showed a fold change of >1.5 or <0.5 in at least one array. However, T01B11.2 contains an x-box motif, which has been shown to be a site of *daf-19* mediated gene expression in various other genes. As such, it was considered to be a reasonable candidate for further analysis, and a transcriptional fusion with GFP was generated by our collaborator, Prasad Phirke.

Expression of T01B11.2::GFP in a daf-19(WT) genetic background

Two separately injected lines of *daf-19(WT)* worms containing the T01B11.2::GFP transgene were analyzed for GFP expression via confocal microscopy (injections by Prasad Phirke). One of these lines (OE3705) was injected directly into a *daf-19* WT background, while the other (LU630) is the isogenic partner strain to OE3895 (T01B11.2::GFP in a *daf-19*(*m86*)

background) created as part of the present study. Anatomical features showing GFP expression were identified via comparison with representative images from *wormatlas.org* (Figure 13). Both strains show consistent expression of GFP in the intestine, hypodermis, body wall muscle, and pharynx (Figure 14; Tables 5 and 6). Additionally, six muscle arms were visible in the isthmus region of the pharynx. One to two pairs of neurons were visible in the dorsal region of the isthmus, typically midway between the anterior and posterior pharyngeal bulbs. Consistently, one of these pairs showed brighter GFP expression, and in some cases the second pair of neurons was not visible. Their morphology showed a long dendrite projecting anteriorly to the tip of the nose, with a second shorter dendrite extending ventrally across the isthmus. Finally, inconsistent expression of GFP appeared in one tail neuron and several non-neuronal cells, which were identified to be rectal glands surrounding the intestinal-rectal valve. Transgene expression did not appear to be age dependent.
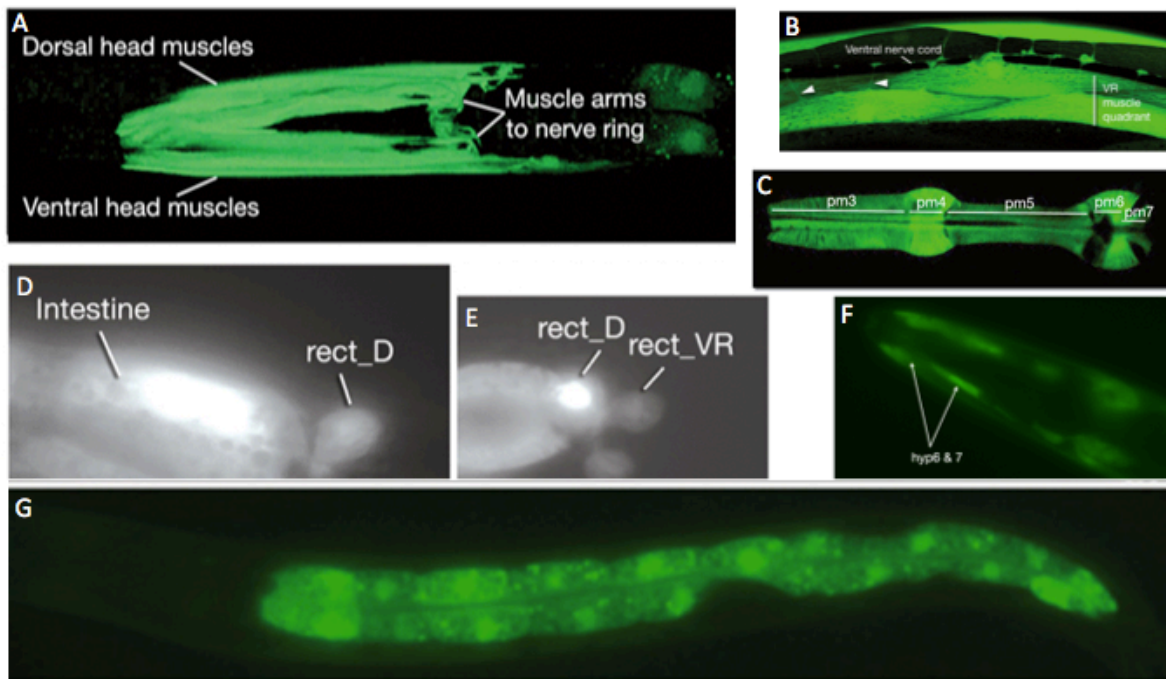
Figure 13: Representative images of anatomical structures identified in T01B11.2::GFP worms in a WT background. (A) Muscle arms connecting to the nerve ring across the isthmus. (B) Body wall muscle cells, with muscle arms connecting to the ventral nerve cord. (C) Pharyngeal muscle cells. (D – E) Rectal glands at the base of the intestine. (F) Hypodermal cells of the head. (G) Intestinal cells. All images adapted from *wormatlas.org*.
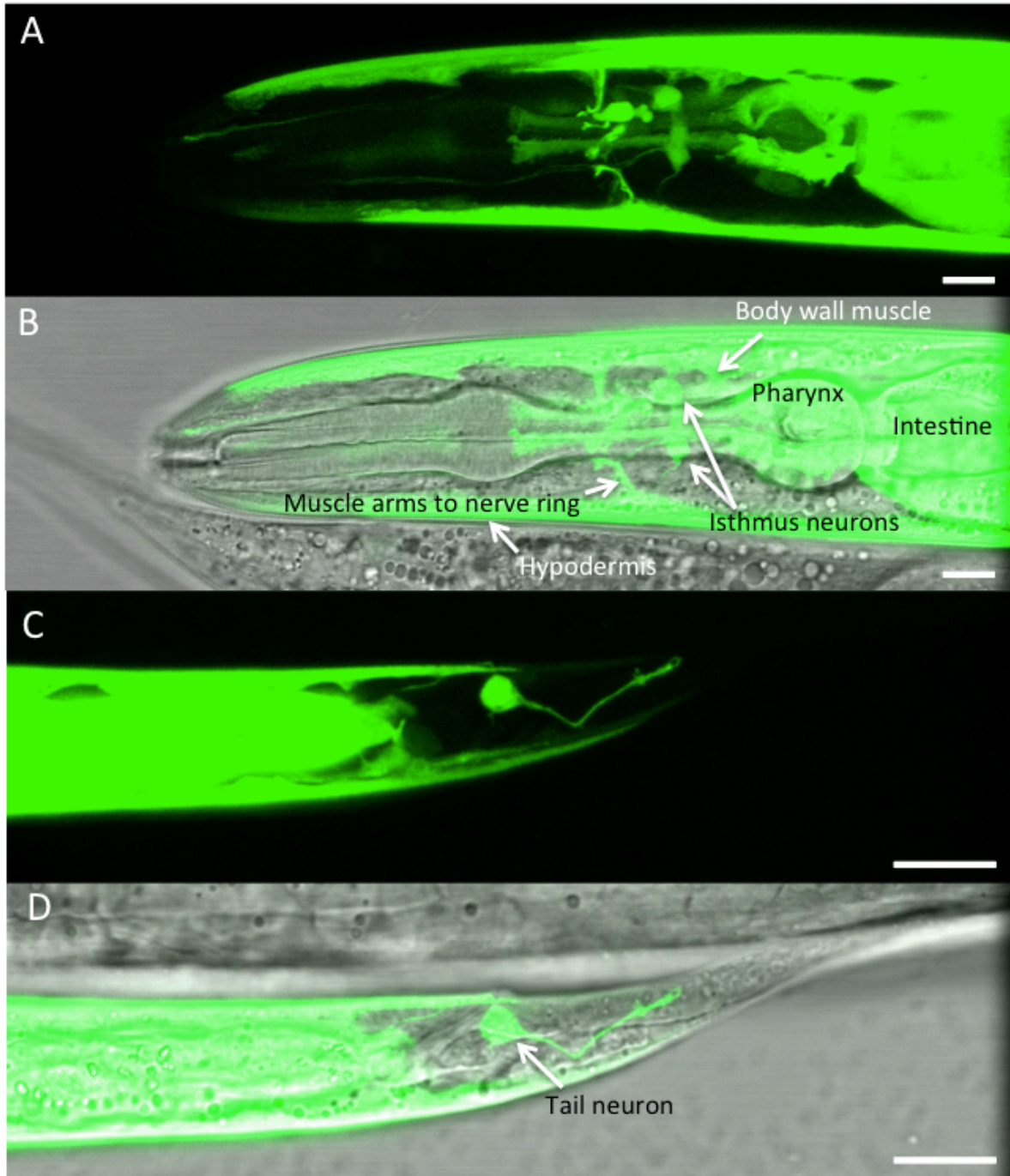
Figure 14: Expression of T01B11.2::GFP in a *daf-19(WT)* background. (A – B) High GFP expression was observed in the hypodermis, body wall muscle, intestine, and pharyngeal muscle, as well as 1-2 pairs of neurons in the isthmus. (C – D) Intermittent expression was also visible in a single tail neuron and rectal glands (not shown). Worms in this image and all images hereafter are oriented according to the graphic shown at right, and all scale bars are 10 μm. Figures A and B show an adult worm, while figures C and D show an L1 worm.
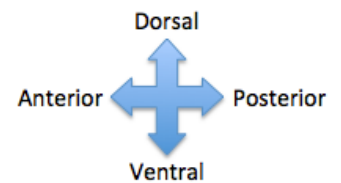
Strain: LU630	Number of *daf-19(WT)* animals with observed expression of T01B11.2::GFP						
Age	N =	Isthmus neurons	Tail neuron	Rectal glands	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	7	7	3	6	7	7	7	7
L3	7	7	3	3	7	7	7	7
L4	8	8	1	0	8	8	8	8
Adult	9	9	1	0	9	9	9	9
Total animals with indicated phenotype (%)	31	31 (100%)	8 (29%)	9 (29%)	31 (100%)	31 (100%)	31 (100%)	31 (100%)

Table 5: Localization of T01B11.2::GFP expression in a *daf-19(WT)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in 1-2 pairs of isthmus neurons, pharyngeal muscle, body wall muscle, intestine, and hypodermal tissues, with intermittent GFP expression in one tail neuron and rectal gland cells.

Strain: OE3705	Number of *daf-19* WT animals with observed expression of T01B11.2::GFP						
Age	N =	Isthmus neurons	Tail neuron	Rectal glands	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	3	3	3	2	3	3	3	3
L3	5	5	3	1	5	5	5	5
L4	4	4	2	0	4	4	4	4
Adult	8	8	1	0	8	6	6	8
Total animals with indicated phenotype (%)	20	20 (100%)	9 (45%)	3 (15%)	20 (100%)	18 (90%)	18 (90%)	20 (100%)

Table 6: Localization of T01B11.2::GFP expression in a separately injected *daf-19(WT)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in 1-2 pairs of isthmus neurons, pharyngeal muscle, body wall muscle, intestine, and hypodermal tissues, with intermittent GFP expression in one tail neuron and rectal gland cells. Expression levels in this strain were comparable to those observed in LU630.

Expression of T01B11.2::GFP in a daf-19(m86) genetic background

A single line of *daf-19(m86)* worms containing the T01B11.2::GFP transgene was analyzed for GFP expression via confocal microscopy (Strain name: OE3895). Injections were completed by Prasad Phirke, and the strain is isogenic to LU630 (previously described). Anatomical features showing expression were identified via comparison with representative

images from *wormatlas.org*. Consistent GFP expression was observed in the intestine, hypodermis, body wall muscle, pharynx, and muscle arms (Figure 15; Table 7). However, the two pairs of neurons present in the isthmus were absent in 90% of worms imaged, indicating differential expression. Inconsistent expression of GFP appeared in the rectal glands and one tail neuron, which showed similar morphology to the neuron that was observed in the two WT strains. Transgene expression did not appear to be age dependent.

Strain: OE3895	Number of *daf-19(m86)* animals with observed expression of T01B11.2::GFP						
Age	N =	Isthmus neurons	Tail neuron	Rectal glands	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	13	1	5	12	13	13	13	13
L3	6	1	2	5	6	6	6	6
L4	4	1	1	2	4	4	3	4
Adult	7	0	2	3	7	7	6	7
Total animals with indicated phenotype (%)	30	3 (10%)	10 (33%)	22 (73%)	30 (100%)	30 (100%)	28 (93%)	30 (100%)

Table 7: Localization of T01B11.2::GFP expression in a *daf-19* (*m86*) background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in pharyngeal muscle, body wall muscle, intestine, and hypodermal tissues, with intermittent GFP expression in one tail neuron and rectal gland cells. Notably, they lacked expression in isthmus neurons, indicating differential expression. Expression levels in this strain were comparable to those observed in LU630.
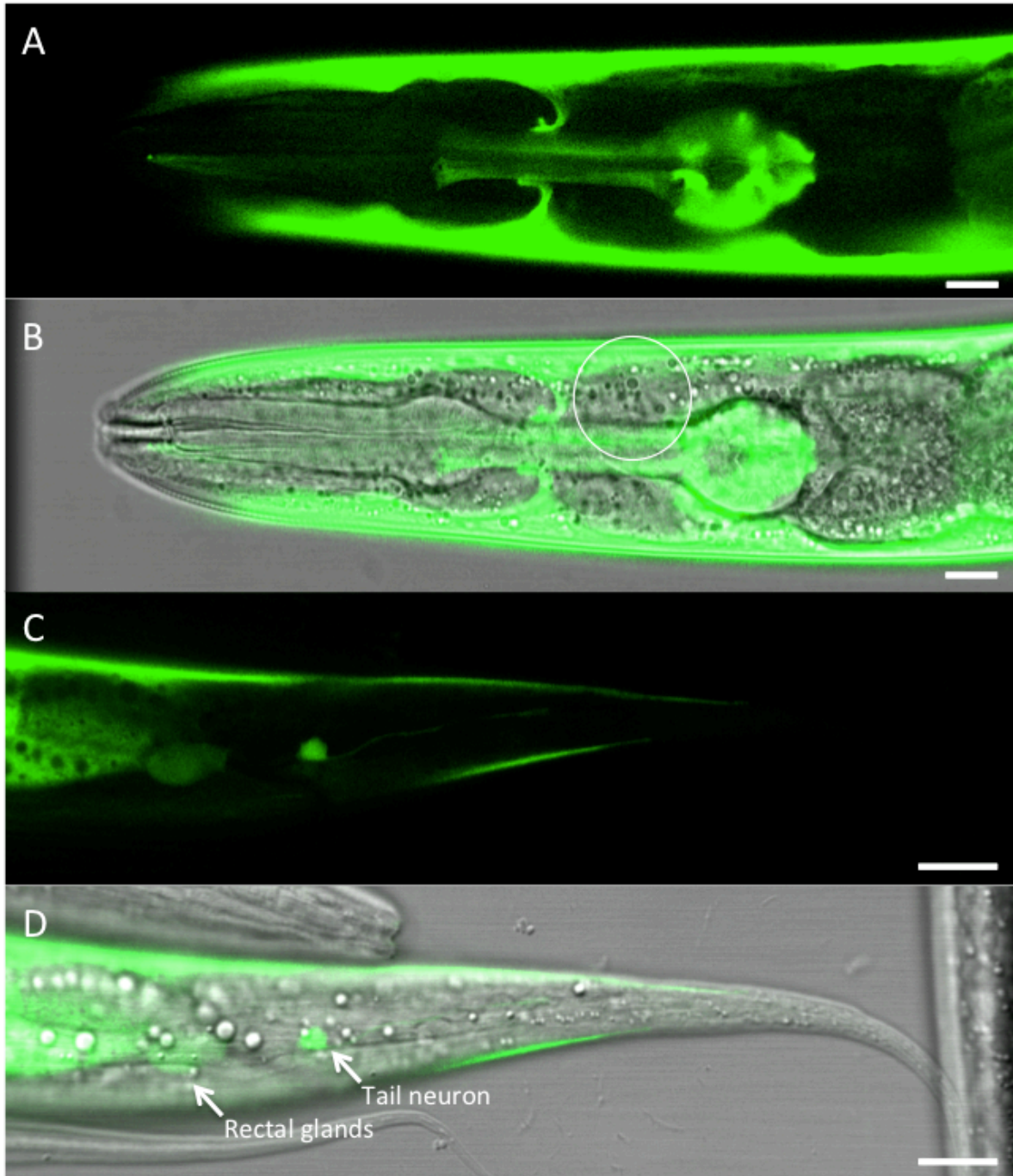
Figure 15: Expression of T01B11.2::GFP in a *daf-19(m86)* background. (A – B) High GFP expression was observed in the hypodermis, body wall muscle, intestine, and pharyngeal muscle, but was absent from the neurons observed in the isthmus in the isogenic WT strain (indicated by white circle). (C – D) Intermittent expression was also visible in a single tail neuron and rectal glands. All images show adult worms, and all scale bars are 10 μm.
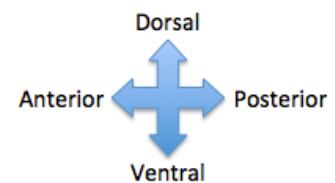
As DAF-19 dependency was established for T01B11.2, it was of interest to us to specifically identify the neurons in which T01B11.2 was differentially expressed using WT strains in which gene expression was more widespread. In order to narrow down candidate neurons, worms from both transgenic WT strains were treated with DiI in order to stain the amphid neurons ASK, ADL, ASI, AWB, ASH, and ASJ (Figure 16). Stained worms were then imaged via confocal microscopy, and the isthmus neurons of interest were analyzed for co-localization of DiI with GFP (Figure 17). Co-localization was not observed; as such, all six neurons were ruled out as potential candidates.
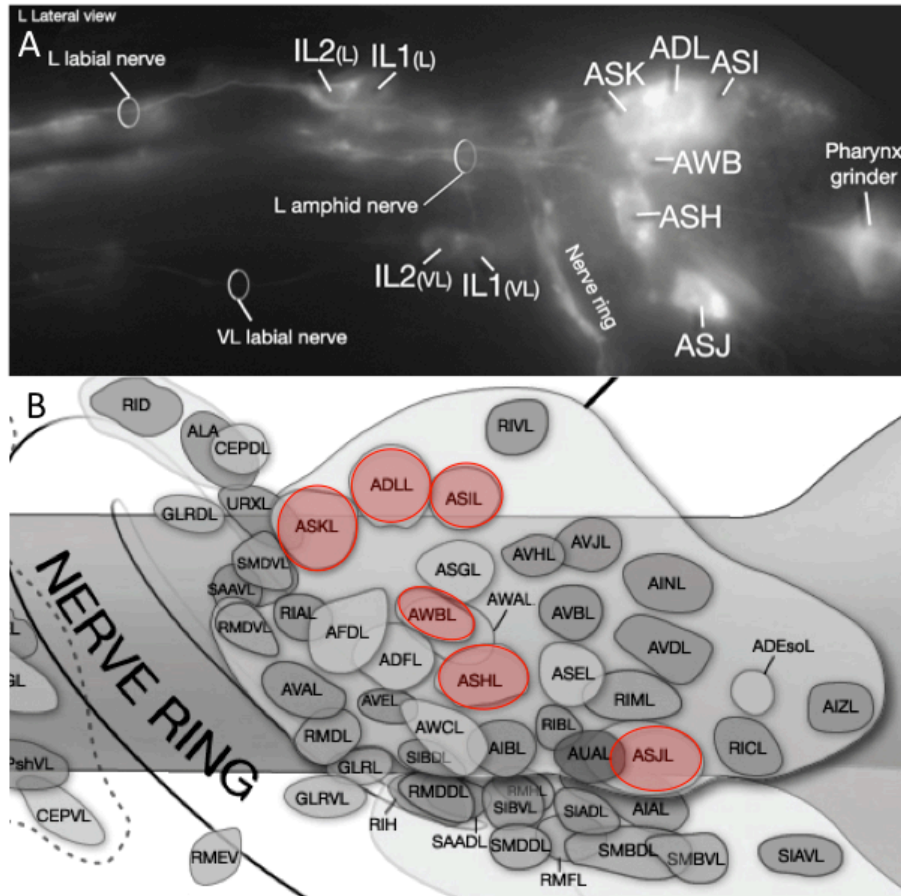
Figure 16: *C. elegans* amphid neurons. (A) Six amphid neurons dye fill when treated with DiI (ASK, ADL, ASK, AWB, ASH, ASJ). Confocal image from *http://www.wormatlas.org/EMmethods/DiIDiO.htm*. (B) Schematic adapted from *wormatlas.org*, with amphid neurons shown in red.
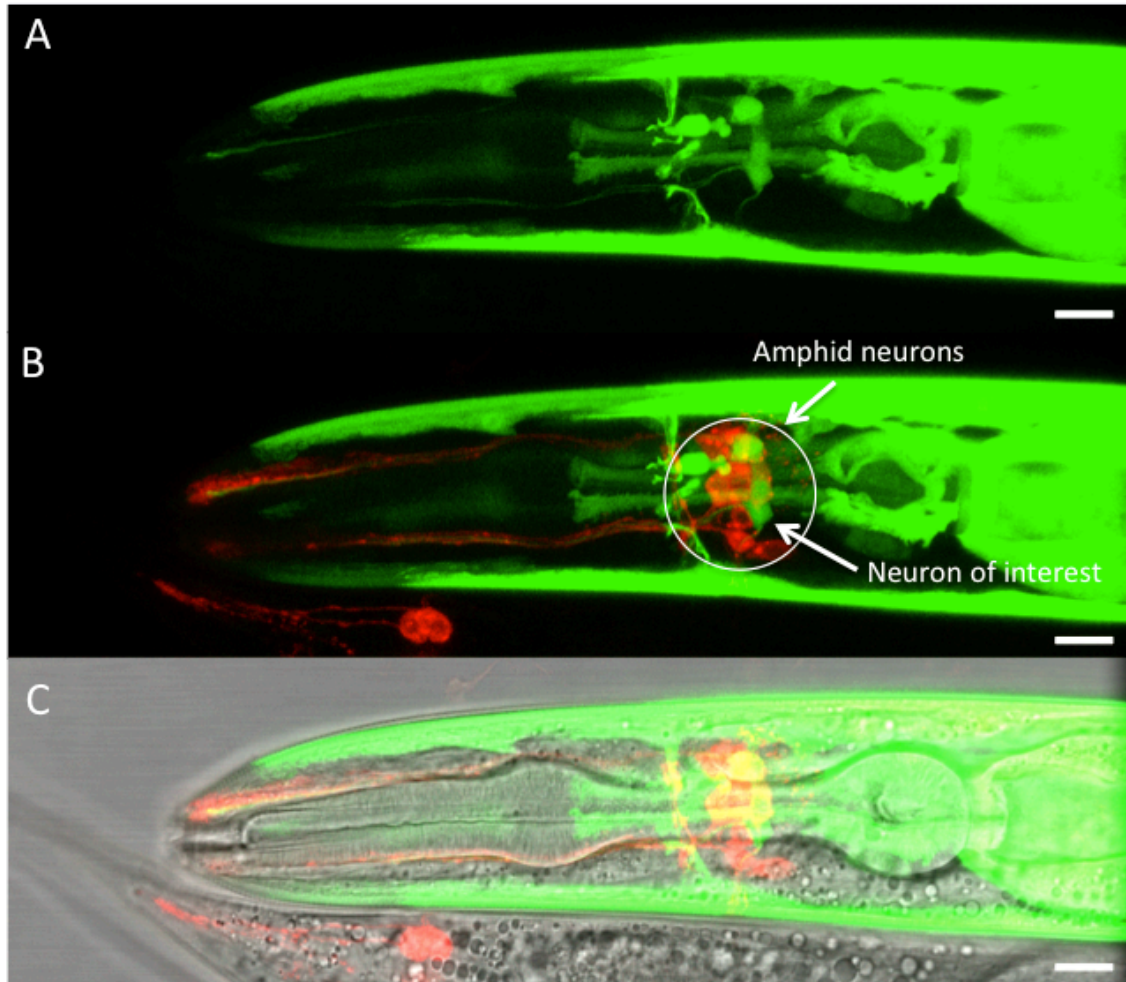
Figure 17: Expression of T01B11.2::GFP in a *daf-19(WT)* background with amphid neurons stained by DiI. The neuron of interest (A) did not co-localize with any dye-filling amphid neurons (shown in red, B – C). Images show adult worms, and scale bars are 10 μm.
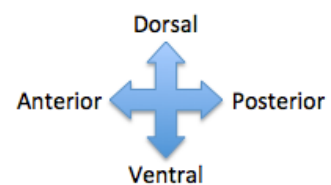
The amphid cell bodies were used as positional markers to identify other neurons in the

vicinity expressing GFP. While the precise positions of the GFP-expressing neurons varied

slightly, they generally appeared immediately anterior or posterior to the dorsal-most dye-filling

neurons. Based on their position and morphology, it was hypothesized that the brightest neuron

pair consisted of the right and left URX neurons, as this neuron appears in the mid-dorsal region

of the isthmus and has similar neuronal projections (Figure 18). Notably, it is one of few

neurons in the vicinity that does not connect to the nerve ring. As nerve ring expression was not

definitively observed in either WT strain, this suggested URX to be a promising candidate as a

site of differential T01B11.2 expression.
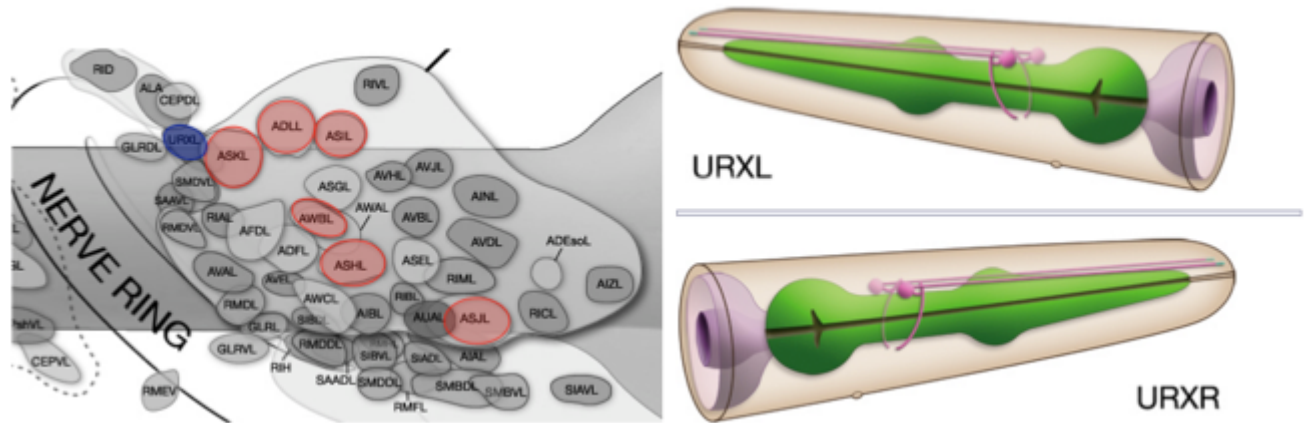
Figure 18: Potential neuron of interest. The differentially expressed isthmus neuron present in T01B11.2::GFP (WT) worms was hypothesized to be URX, shown in blue in the schematic to the left. All images adapted from *wormatlas.org*.

In order to confirm the identity of URX, a strain of worms containing a fluorescent

mCherry marker (*gcy-32p::mCherry*) in the URX neuron was procured from the De Bono lab.

These worms were crossed with *daf-19* WT worms (LU630) in order to produce double

transgenic worms containing both the URX marker and the T01B11.2::GFP transgene. Confocal

analysis revealed that the initial hypothesis was incorrect, as the mCherry marker did not co-

localize with GFP (Figure 19).
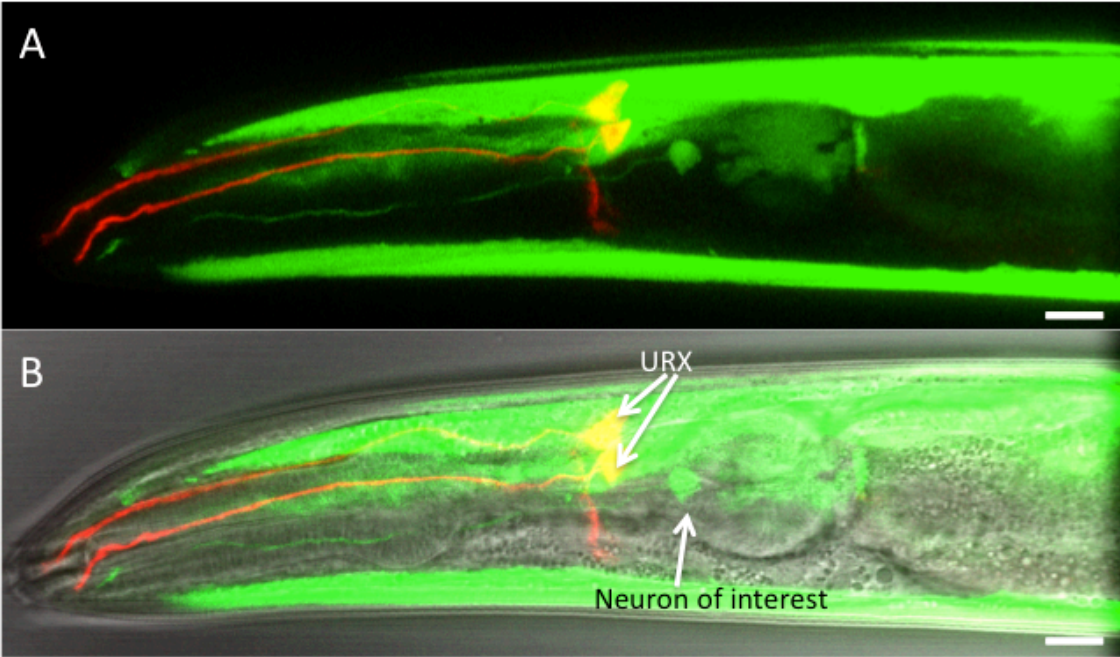
Figure 19. Expression of T01B11.2::GFP and gcy-32p::mCherry in a *daf-19+/+* background. (A – B) mCherry expression in the URX neuron did not colocalize with GFP expression in the isthmus neuron of interest. Images show adult worms, and scale bars are 10 μm.
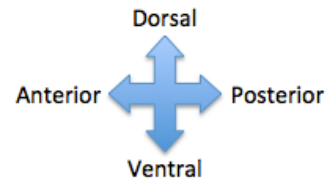
Once URX was ruled out as a candidate for the neuron of interest, various other neurons were considered as potential candidates. Based on a few images with decreased non-neuronal expression, we determined that the neurons of interest did in fact connect to the nerve ring. Additionally, we created a second double transgenic marker using a strain of worms that expressed mCherry in the nuclei of all cholinergic neurons (strain by Sophie Scholtz). This was used to confirm that the GFP-expressing neurons were not cholinergic, as they did not co-localize with mCherry (Figure 20). With these data, we were able to narrow down the possible range of neurons to two pairs (ASG and AWA), based on their location and morphology (Figure 20). Both have cell bodies in the mid dorsal region of the isthmus, are quite close to the dye-filling amphid neurons, are non-cholinergic, and are posterior to URX, all features shared with

the neurons of interest. As such, we currently hypothesize that T01B11.2 is activated in ASG and AWA by DAF-19.
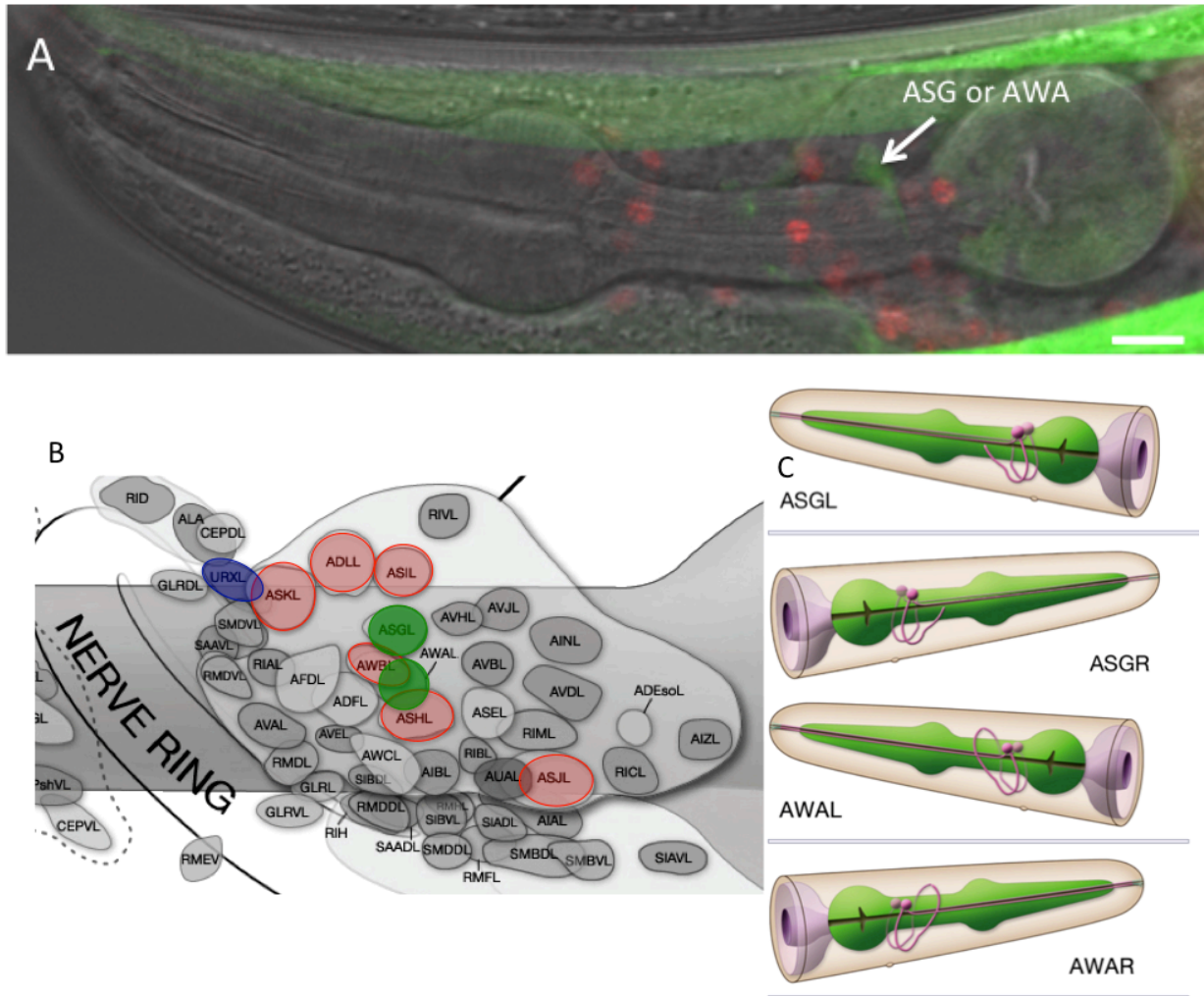
Figure 20: T01B11.2::GFP does not co-localize with mCherry expressed in cholinergic neurons. (A) A *cho-1::SL2::mCherry pha-1* marker was used to rule out the cholinergic neurons as potential sites for T01B11.2 expression, eliminating all but two cells as potential candidates. Image was taken by Sophie Scholtz, and shows an adult worm, with scale bar showing 10 µm. (B) The isthmus neurons expressing T01B11.2::GFP in WT worms were hypothesized to be ASG and AWA, shown in green. URX is shown in blue, and dye-filling amphid neurons are shown in red. (C) Individual schematics of the hypothesized neurons. All images adapted from *wormatlas.org*.

Expression of T01B11.2::GFP in other daf-19 mutant worms

In addition to identifying specific neurons in which T01B11.2 was differentially expressed in WT and *m86* backgrounds, we were interested in elucidating how these expression patterns were regulated. In order to determine whether particular DAF-19 isoforms are

responsible for transcriptional regulation of T01B11.2, we produced three isogenic strains

containing the T01B11.2::GFP transgene in different genetic backgrounds for *daf-19*. Each of

these mutants affected DAF-19A, DAF-19B, or both, but not DAF-19C; as such, all three

displayed dye-filling phenotypes. In order to produce the isogenic strains, we mated *daf-19(m86)* hermaphrodites with males that contained each respective mutation, and then selected

for dye-filling worms. The first of these was *of5*, a *daf-19a/b* null mutant produced by Debora

Sugiaman, which contained a 12 bp deletion with a +1 shifted ATG start codon that leaves only

DAF-19C functional (strain name: LU646). These worms showed an expression pattern that was

very similar to the two wild type strains that were analyzed, with consistent GFP expression in

the pharynx, body wall muscle, intestine, and hypodermis (Table 8). More importantly, 100% of

the worms analyzed expressed GFP in the isthmus neurons observed in the WT strains (Figure

21). This result indicates that T01B11.2 expression is not dependent on the presence of isoforms

A and/or B, as the transgene is expressed identically in WT and *of5* genetic backgrounds. Thus,

these preliminary data suggest that DAF-19C regulates expression of T01B11.2.

Strain: LU646	Number of *daf-19(of5)* animals with observed expression of T01B11.2::GFP							
Age	N =	Isthmus neurons	Tail neuron	Rectal glands	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	11	11	8	11	11	11	11	11
L3	4	4	2	3	4	4	4	4
L4	4	4	1	1	4	4	4	4
Adult	4	4	0	1	4	4	4	4
Total animals with indicated phenotype (%)	23	23 (100%)	11 (48%)	16 (76%)	23 (100%)	23 (100%)	23 (100%)	23 (100%)

Table 8: Localization of T01B11.2::GFP expression in a *daf-19(of5)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed GFP expression in 1-2 pairs of isthmus neurons, pharyngeal muscle, body wall muscle, intestine, and hypodermal tissues, with intermittent fluorescence in one tail neuron and rectal gland cells.

Figure 21: Expression of T01B11.2::GFP expression in a *daf-19*(*of5*) background.
(A – B) High GFP expression was observed in the hypodermis, body wall muscle, intestine, and pharyngeal muscle, as well as in the isthmus neurons hypothesized to be ASG and AWA in the isogenic WT strain (indicated by white circle). (C – D) Intermittent expression was also visible in a single tail neuron and rectal glands. All images show adult worms, and all scale bars are 10 μm.
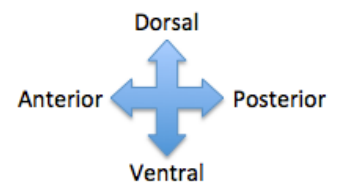
In order to further elucidate isoform-specific control mechanisms, we also generated a strain of transgenic T01B11.2::GFP worms in a *tm5562* background via the same process described above (strain name: LU641). This mutant, obtained from the Mitani lab, contained a large deletion in exon 2 which affected DAF-19A/B. Confocal analysis of this strain showed a very similar expression pattern to the *of5* mutant and both WT strains, with 95% of worms expressing GFP in the isthmus neurons of interest (Figure 22). GFP expression in the pharynx, body wall muscle, intestine, and hypodermis was also consistent with a WT phenotype (Table 9). As with the *of5* strain, these data suggest that T01B11.2 expression in these neurons is regulated by DAF-19C.

Strain: LU641	Number of *daf-19(tm5562)* animals with observed expression of T01B11.2::GFP						
Age	N =	Isthmus neurons	Tail neuron	Rectal glands	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	5	5	0	4	5	5	5	5
L3	6	6	2	2	6	6	6	6
L4	2	1	1	0	2	2	2	2
Adult	8	8	4	0	8	8	8	8
Total animals with indicated phenotype (%)	21	20 (95%)	7 (33%)	6 (29%)	21 (100%)	21 (100%)	21 (100%)	21 (100%)

Table 9: Localization of T01B11.2::GFP expression in a *daf-19*(*tm5562*) background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed GFP expression in 1-2 pairs of isthmus neurons, pharyngeal muscle, body wall muscle, intestine, and hypodermal tissues, with intermittent fluorescence in one tail neuron and rectal gland cells.
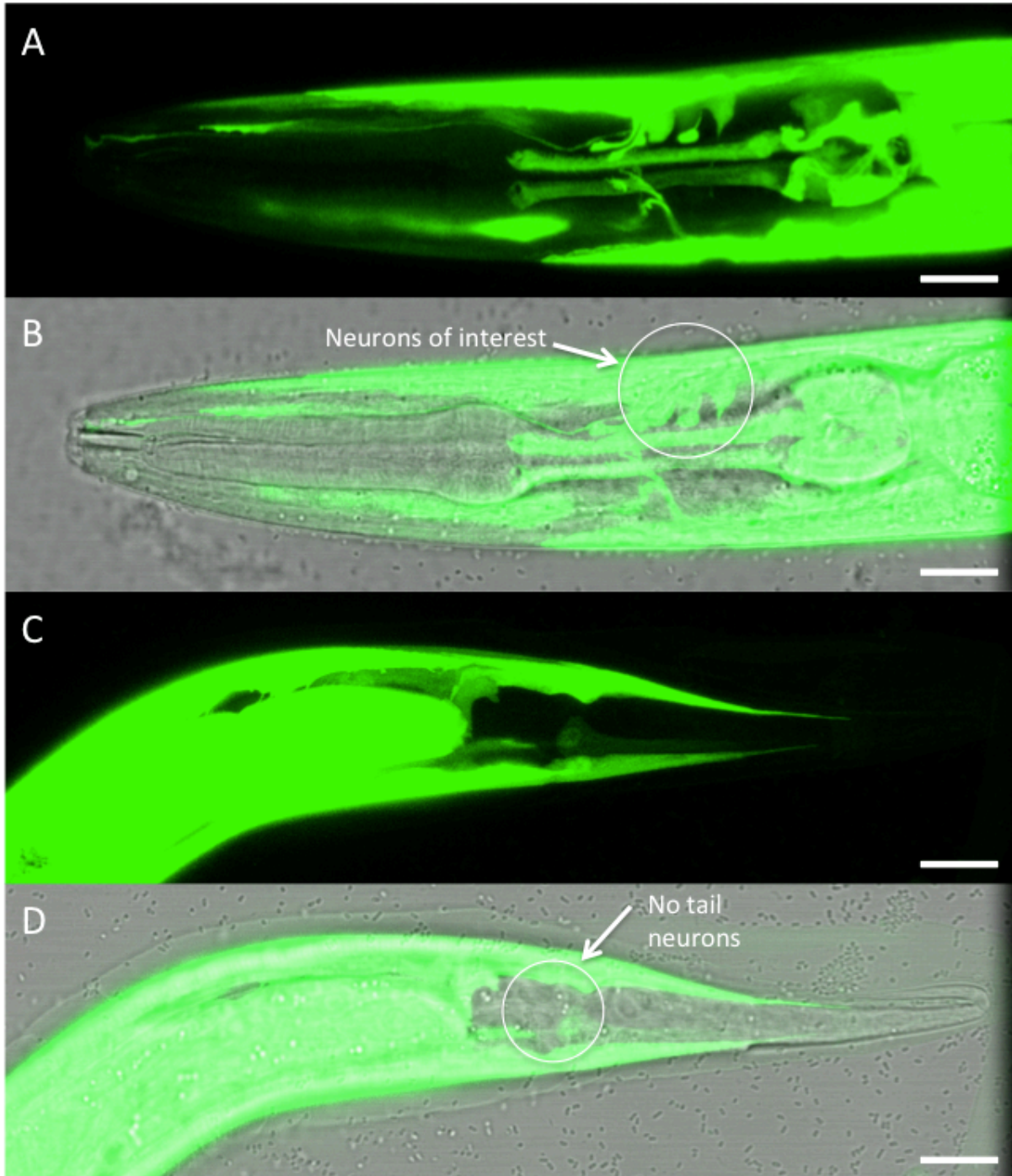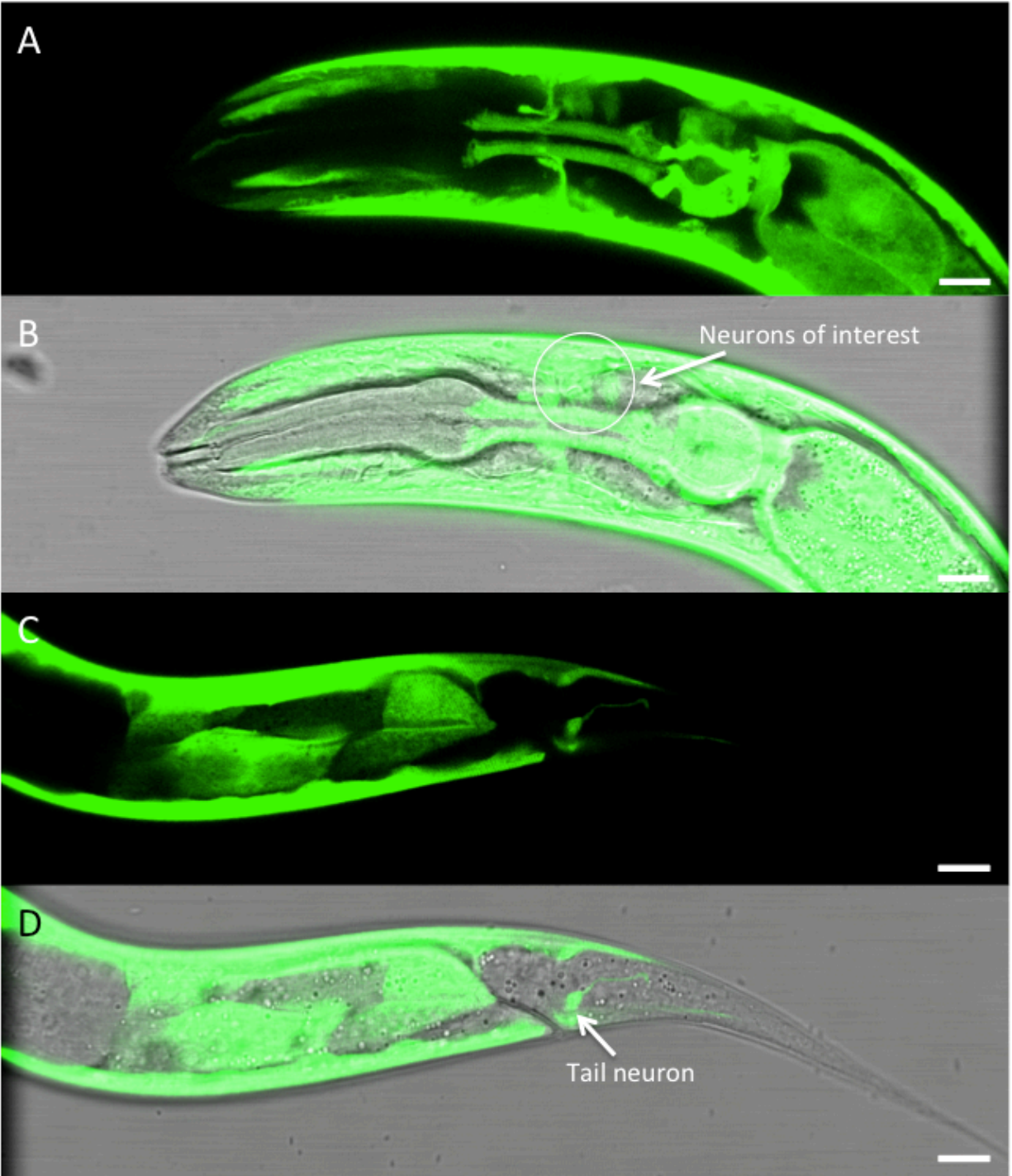
Figure 22: Expression of T01B11.2::GFP in a *daf-19*(*tm5562*) background. (A –
B) High GFP expression was observed in the hypodermis, body wall muscle,
intestine, and pharyngeal muscle, as well as in the isthmus neurons hypothesized
to be ASG and AWA in the isogenic WT strain (indicated by white circle). (C – D)
Intermittent expression was also visible in a single tail neuron and rectal glands.
All images show adult worms, and all scale bars are 10 μm.
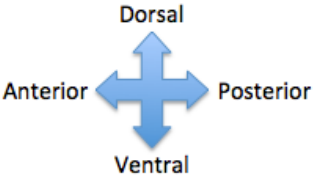
We generated an additional strain of transgenic T01B11.2::GFP worms in a *daf-19(of6)* background via the same method described above. The *of6* mutation, produced by our collaborator Debora Sugiaman, consists of a 53 bp deletion in exon 4 that renders the protein nonfunctional. However, since worms homozygous for the *daf-19(of6)* allele do dye fill and are not dauer constitutive, this mutation appears not to affect DAF-19C. Unlike with the *tm5562* and *of5* mutants that affect DAF-19A/B, this mutation negatively affects only DAF-19B, as exon 4 is uniquely translated in this isoform. As such, DAF-19A and DAF-19C are left intact. Interestingly, this strain showed extremely inconsistent expression across a variety of tissues, including the neurons of interest (Tables 9 and 10). Of the worms imaged, 25% showed a phenotype that resembled WT worms, featuring expression in the hypodermis, intestine, body wall muscle, pharynx, and isthmus neurons (Figure 23 A – B). Another 50% expressed a phenotype that resembled the *m86* strain, which lacked the GFP expression in the isthmus neurons of interest (Figure 23 C – D). Additionally, 25% of the worms imaged showed either more or less GFP expression than either of the previously described tissues, in a variety of phenotypes. This included worms with an absence of GFP in hypodermal and intestinal cells which virtually always display fluorescence in other transgenic strains (Figure 23 E – F), as well as worms with significantly increased neuronal overexpression in the isthmus in up to eight neurons (Figure 23 G – H). Finally, tail expression patterns ranged from no GFP expression (Figure 23 I – J) to neuronal GFP expression showing morphology that had previously not been observed in any strain (Figure 23 K – L). Thus, no single definitive expression pattern for T01B11.2::GFP *(of6)* worms could be established.
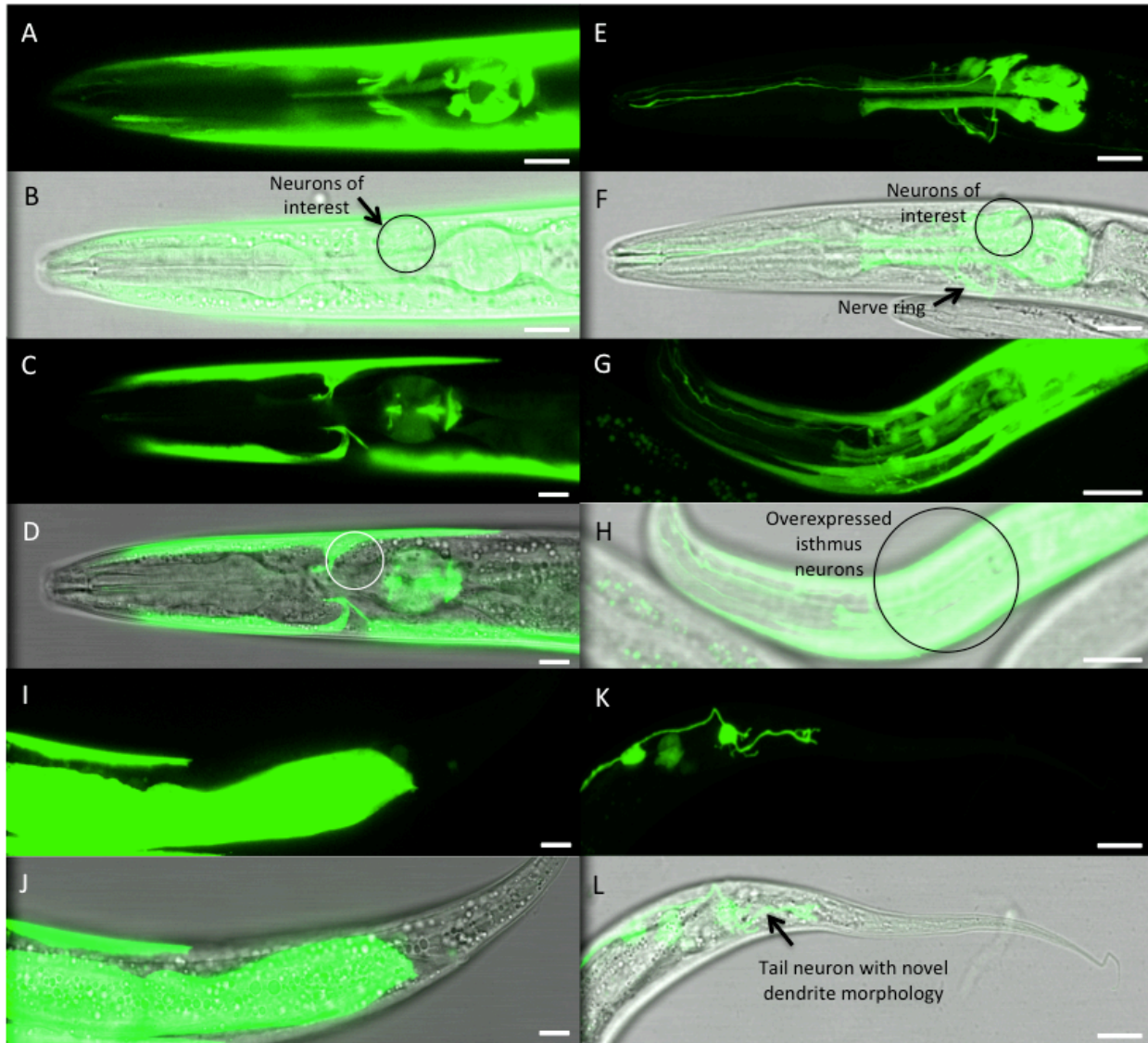
Figure 23: Expression of T01B11.2::GFP in a *daf-19*(*of6*) background. (A – B) A WT–like phenotype was observed in 25% of worms imaged, featuring GFP expression in the hypodermis, body wall muscle, intestine, and pharyngeal muscle, as well as in the isthmus neurons hypothesized to be ASG and AWA (indicated by black circle). (C – D) An *m86*–like phenotype was observed in 50% of worms imaged, with no GFP expression in the isthmus neurons of interest. (E – F) 25% of worms showed novel phenotypes that had previously not been seen in any strain. Panels E – F depict a worm that showed GFP expression in neuronal and pharyngeal tissue only. (G – H) An additional novel phenotype showed considerable overexpression in neuronal tissue in the isthmus region. (I – L) Tail expression patterns ranged from no neuronal expression (panels I – J) to increased neuronal expression of GFP featuring novel dendrite morphology (panels K – L). Worm ages range from L1 to adult; all scale bars are 10 μm.
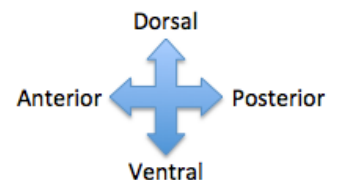
Strain: LU653	Number of *daf-19(of6)* animals with observed expression of T01B11.2::GFP						
Age	N =	Isthmus neurons	Tail neuron	Rectal glands	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	16	8	5	10	16	14	14	15
L3	4	2	1	1	3	3	3	3
L4	0	-	-	-	-	-	-	-
Adult	4	0	0	0	4	4	4	4
Total animals with indicated phenotype (%)	24	10 (42%)	6 (25%)	11 (46%)	23 (96%)	21 (87%)	21 (87%)	22 (92%)

Table 9: Localization of T01B11.2::GFP expression in a *daf-19(of6)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms showed GFP expression in the isthmus neurons of interest 42% of the time.

Number of *daf-19(of6)* animals with each representative phenotype				
Age	N =	WT	*m86*	Other phenotype
L1/L2	16	5	8	4
L3	4	1	1	2
L4	0	-	-	-
Adult	4	0	3	0
Total animals with indicated phenotype (%)	24	6 (25%)	12 (50%)	6 (25%)

Table 10: Frequency of each representative phenotype in *daf-19*(of6) worms. Worms showed highly variable expression patterns, including a WT-like phenotype (25%), an *m86* – like phenotype (50%), and several novel phenotypes (25%).

While data from the *tm5562* and *of5* strains strongly suggested that T01B11.2 is regulated by DAF-19C in isthmus neurons, the *of6* result led us to question the certainty of that statement. As we had no mutant that affected only *daf-19c*, we elected to do a rescue experiment using cDNAs expressing only *daf-19c*. Complementary DNA (cDNA) is double-stranded DNA synthesized from an mRNA template which can be used to reintroduce a gene transcript into an animal for which that gene has previously been knocked out. Gabi Senti generated cDNAs for *daf-19c*, which were incorporated into a plasmid that was then microinjected into a line of *m86* worms (Senti & Swoboda, 2008). This plasmid is overexpressed; as such, worms containing this transgene will have more DAF-19C expressed throughout the body than WT worms, and any

phenotypic data should be considered accordingly. The *daf-19c* rescue strain was crossed with

worms containing the T01B11.2 transgene in an *m86* background in the same manner described

above (strain name = LU683). Worms were dye filled to select for the rescue transgene, as the

dye-filling phenotype depends on the presence of functional DAF-19C. These double transgenic

worms were then analyzed using confocal microscopy. Preliminary data have indicated that

these worms display a WT phenotype, with GFP expression clearly visible in the isthmus

neurons of interest, in addition to hypodermis, body wall muscle, intestine, and pharyngeal

muscle (Figure 24; Table 11). When taken in context with the data from *tm5562* and *of5*

backgrounds, these data convincingly suggest that T01B11.2 is activated by DAF-19C.

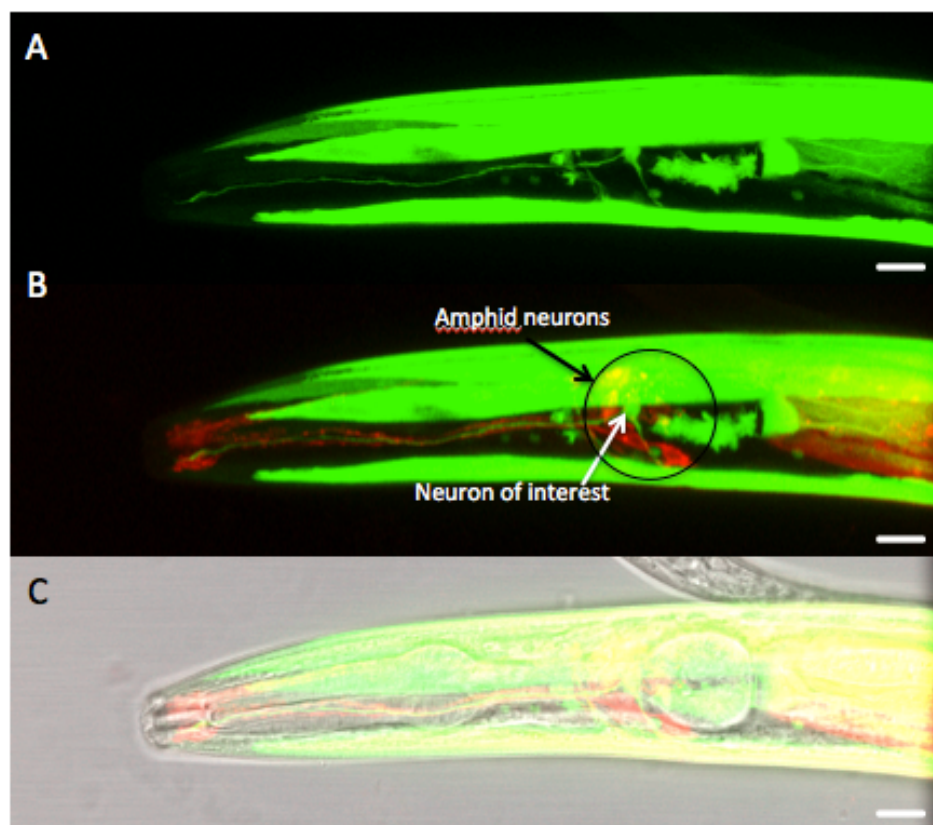**Figure 24: T01B11.2::GFP expression after transformation rescue of the DAF-19C isoform.** (A – C) High GFP expression was observed in the hypodermis, body wall muscle, intestine, and pharyngeal muscle. Addition of DAF-19C to *daf-19 (m86)* worms rescues expression in the isthmus neurons of interest, hypothesized to be ASG and AWA. Dye-filled amphid neurons shown in red. All worms are adults, and all scale bars are 10 μm.
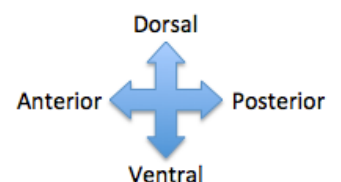
Strain: LU683	Number of *daf-19c* rescue animals with observed expression of T01B11.2::GFP							
Age	N =	Isthmus neurons	Tail neuron	Non-neuronal tail cells	Pharyngeal muscle	Body wall muscle	Intestine	Hypodermis
L1/L2	0	-	-	-	-	-	-	-
L3	0	-	-	-	-	-	-	-
L4	4	3	0	0	4	4	4	4
Adult	4	4	0	0	4	4	4	4
Total animals with indicated phenotype (%)	8	7 (87%)	0 (0%)	0 (0%)	8 (100%)	8 (100%)	8 (100%)	8 (100%)

Table 11: Localization of T01B11.2::GFP expression in a *daf-19C::daf-19(m86)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues.

T07F10.1 expression may be partially regulated by DAF-19C

T07F10.1 was identified as a putative *daf-19* target gene from Elizabeth De Stasio's microarray analysis of adult worms only, which showed a 0.43 fold change (down-regulation in *daf-19(m86)* worms). It is an ortholog of the human genes ANPEP (alanyl membrane aminopeptidase), ERAP1 (endoplasmic reticulum aminopeptidase 1) and other aminopeptidases such as LNPEP (Shaye & Greenwald, 2011). Aminopeptidases are enzymes that catalyze protein degradation by hydrolyzing amino-terminal residues from short peptides (Lehninger et al., 2008). They are critically important for a wide variety of biological processes including embryogenesis, antigen presentation, inflammation, and neuropeptide processing, and some alleles of the human orthologs have been associated with pathologies such as hypertension (Yamamoto et al., 2002). T07F10.1 localizes to the plasma membrane, and based on homology, it is predicted to have metallopeptidase and zinc ion-binding activity (*wormbase.org*). Notably, one of its human homologues (ANPEP) is thought to be involved in metabolizing regulatory peptides in various cell types, including synaptic membranes in the CNS (*omim.org*). T07F10.1 is predicted to interact with *tsp-7*, a neuronally expressed protein that plays a role in

morphogenesis, embryonic development, and mitotic spindle organization (Lee et al., 2008). No

information has been reported regarding its regulation, and interestingly, it lacks the

characteristic x-box motif through which *daf-19* has been shown to regulate other genes.

Using a strain of transgenic worms produced by the Baillie lab, Alex Hurlburt reported

that T07F10.1::GFP is expressed in the pharynx, intestine, excretory system, and nervous system

in *daf-19* WT worms (Hurlburt 2014; Table 12, Figure 25). Hurlburt was able to specifically

identify two head neurons, (URX and SABD) which showed GFP expression in 47% and 89% of

worms imaged, respectively. Additionally, he identified three pairs of tail neurons that expressed

GFP 100% of the time; these are PLM, PDA, and DVA/DVB. Hurlburt compared these

expression patterns with worms containing the T07F10.1::GFP transgene in an *m86* background,

and found that URX and SABD only showed GFP expression 17% and 40% of time, respectively

(Table 13; Figure 26).

Hurlburt's conclusions were verified by comparing data he collected with representative

anatomical images from *wormatlas.org* (Figures 27 and 28). Additionally, we were able to

definitively confirm his identification of the URX neuron as expression T07F10.1::GFP by

generating a double transgenic strain of worms containing a fluorescent mCherry marker (*gcy-32*p::mCherry) in the URX neuron (generous gift of the De Bono lab). Confocal analysis of this

strain showed GFP in an isthmus neuron clearly co-localizing with mCherry, indicating that

Hurlburt's initial identification was correct (Figure 29). Hurlburt was uncertain as to whether

this difference in expression frequency demonstrated true DAF-19 dependence; as such, we

elected to study the transgene's expression in various other mutant backgrounds of *daf-19* in

order to better elucidate whether it regulates T07F10.1, and if so, by which isoforms.

Strain: LU496		Number of *daf-19* WT animals with observed expression of T07F10.1::GFP						
Age	N =	SABD	URX	Tail neurons (#)	Excretory system	Vulval muscle	Gonad	Intestine
L1/L2	3	3	2	3 (2-3)	3	0	1	0
L3	9	8	5	9 (3-4)	9	0	2	0
L4	10	9	5	10 (2-5)	10	0	2	1
Adult	23	20	9	23 (2-5)	23	3	4	6
Total animals with indicated phenotype (%)	45	40 (89%)	21 (47%)	45 (100%)	45 (100%)	3 (7%)	9 (20%)	9 (20%)

Table 12: Localization of T07F10.1::GFP expression in a *daf-19(WT)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in the excretory system and three tail neurons, with intermittent GFP expression in the URX and SABD neurons. Occasional expression was visualized in the gonad, intestine, and vulval muscle. All data were collected by Alex Hurlburt in 2014.

Strain: LU495		Number of *daf-19* (*m86*) animals with observed expression of T07F10.1::GFP						
Age	N =	SABD	URX	Tail neurons (#)	Excretory system	Vulval muscle	Gonad	Intestine
L1/L2	3	0	0	3 (3-4)	3	0	0	0
L3	7	2	1	7 (3-5)	7	0	1	0
L4	14	10	5	14 (3-5)	13	0	1	0
Adult	23	7	2	23 (3-5)	23	3	4	0
Total animals with indicated phenotype (%)	47	19 (40%)	8 (17%)	47 (100%)	46 (98%)	3 (6%)	6 (13%)	0 (0%)

Table 13: Localization of T07F10.1::GFP expression in a *daf-19*(*m86*) background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in the excretory system and three tail neurons, with decreased GFP expression in the URX and SABD neurons as compared to WT worms. Occasional expression was visualized in the gonad and vulval muscle. All data were collected by Alex Hurlburt in 2014.
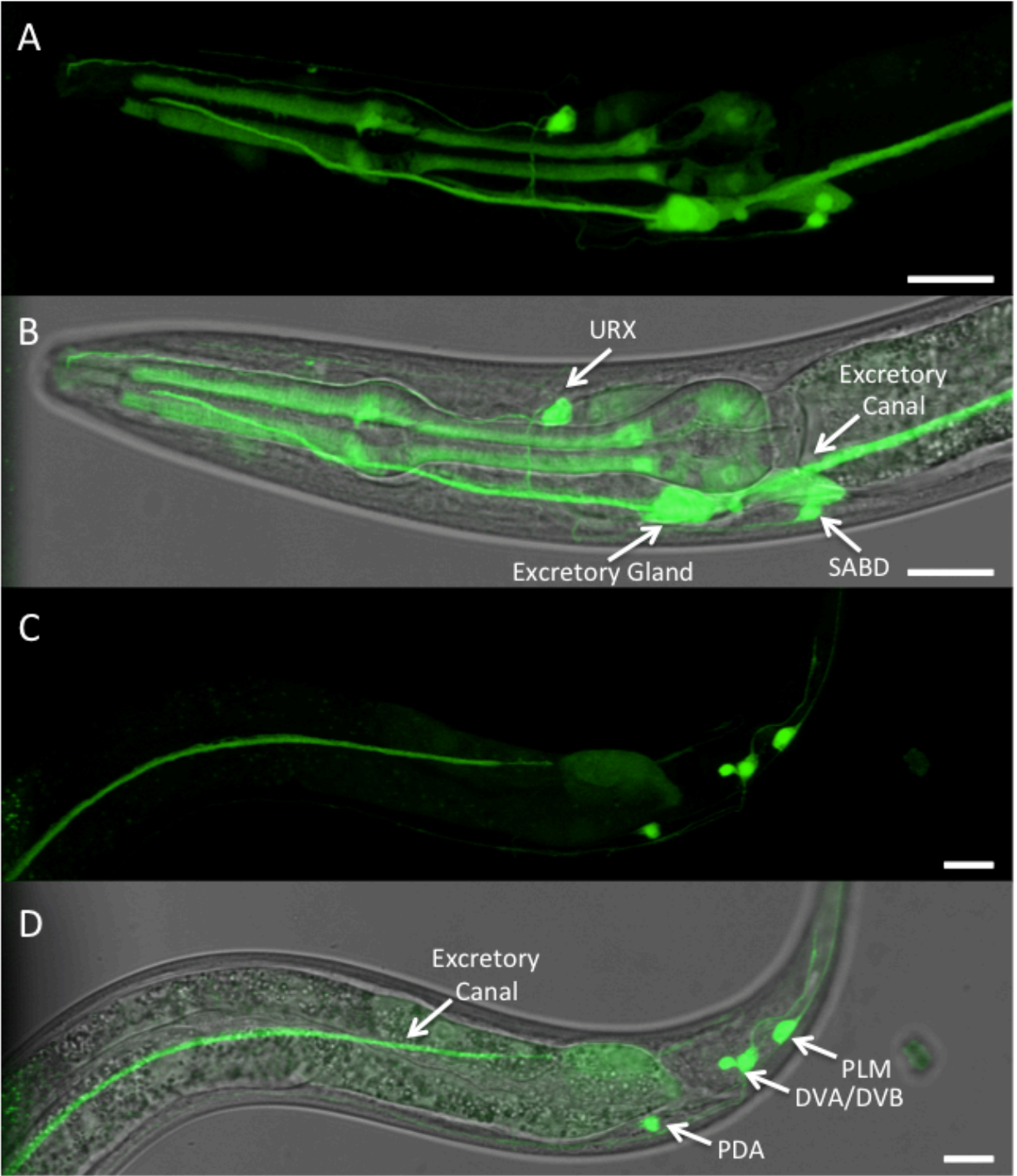
Figure 25. Expression of T07F10.1::GFP in a *daf-19(WT)* background. (A – D) Consistent GFP expression was observed in the pharyngeal muscle, excretory canal and gland, and three pairs of tail neurons (PDA, PLM, DVA/DVB). 89% of worms imaged expressed GFP in the SABD neuron, and 47% of worms expressed GFP in the URX neuron. Occasional GFP expression was seen in the intestine, gonad, and vulval muscle. All images were taken by Alex Hurlburt in 2014. All worms shown are adults, and scale bars are 10 μm.
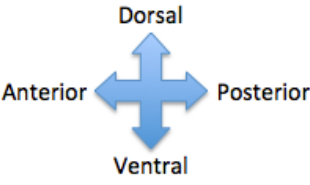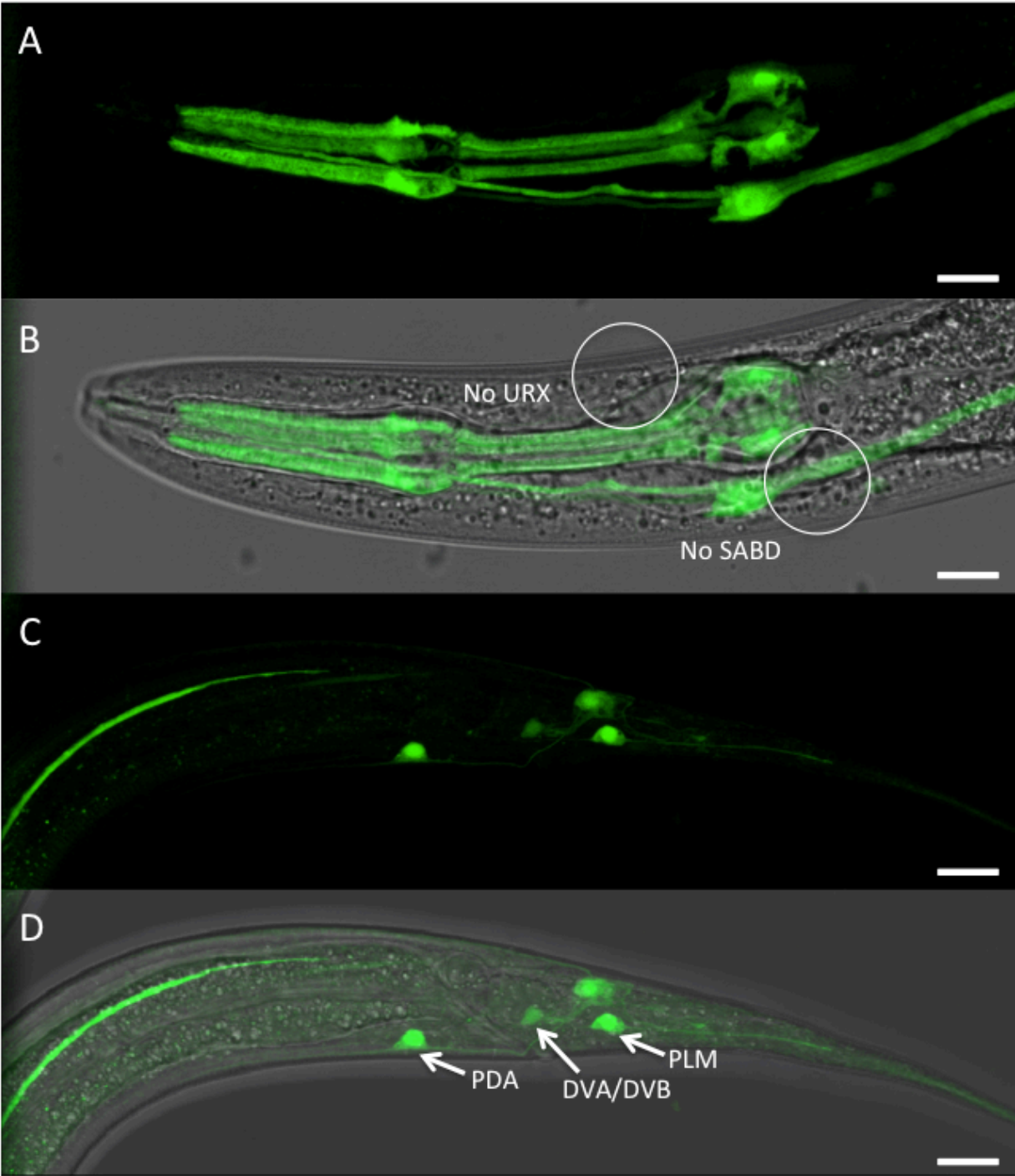
Figure 26. Expression of T07F10.1::GFP in a *daf-19*(*m86*) background. (A – D) Consistent GFP expression was observed in the pharyngeal muscle, excretory canal and gland, and three pairs of tail neurons (PDA, PLM, DVA/DVB). GFP was expressed less frequently in the SABD and URX neurons as compared to WT worms (40% and 17% respectively). Occasional GFP expression was seen in the gonad and vulval muscle. All images were taken by Alex Hurlburt in 2014. All worms shown are adults, and scale bars are 10 μm.
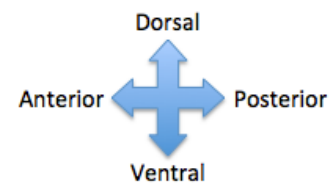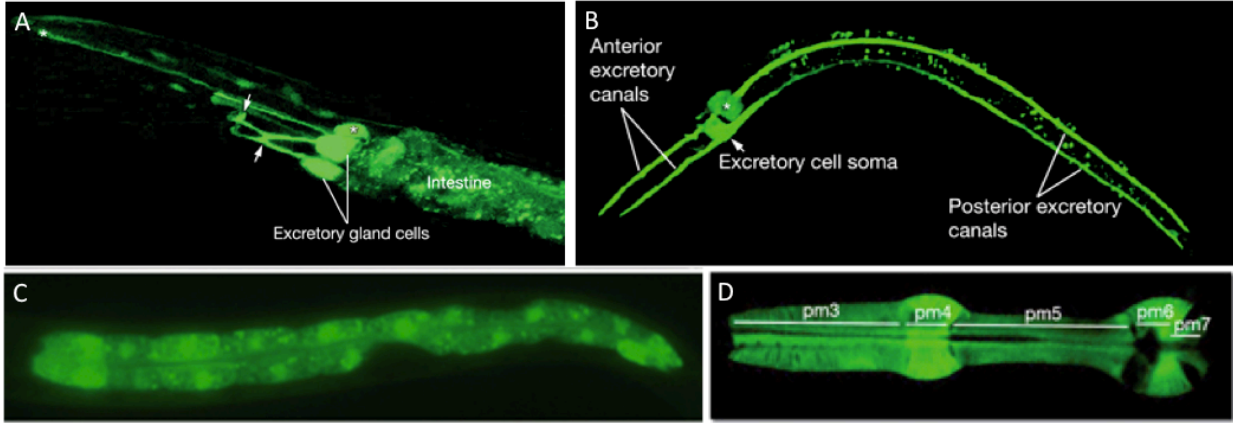
Figure 27: Representative images of non-neuronal anatomical structures identified in T07F10.1::GFP worms in both WT and *daf-19* (*m86*) backgrounds. (A – B) Excretory gland cells and excretory canals. (C) Intestinal cells. (D) Pharyngeal muscle cells. All images adapted from *wormatlas.org*.
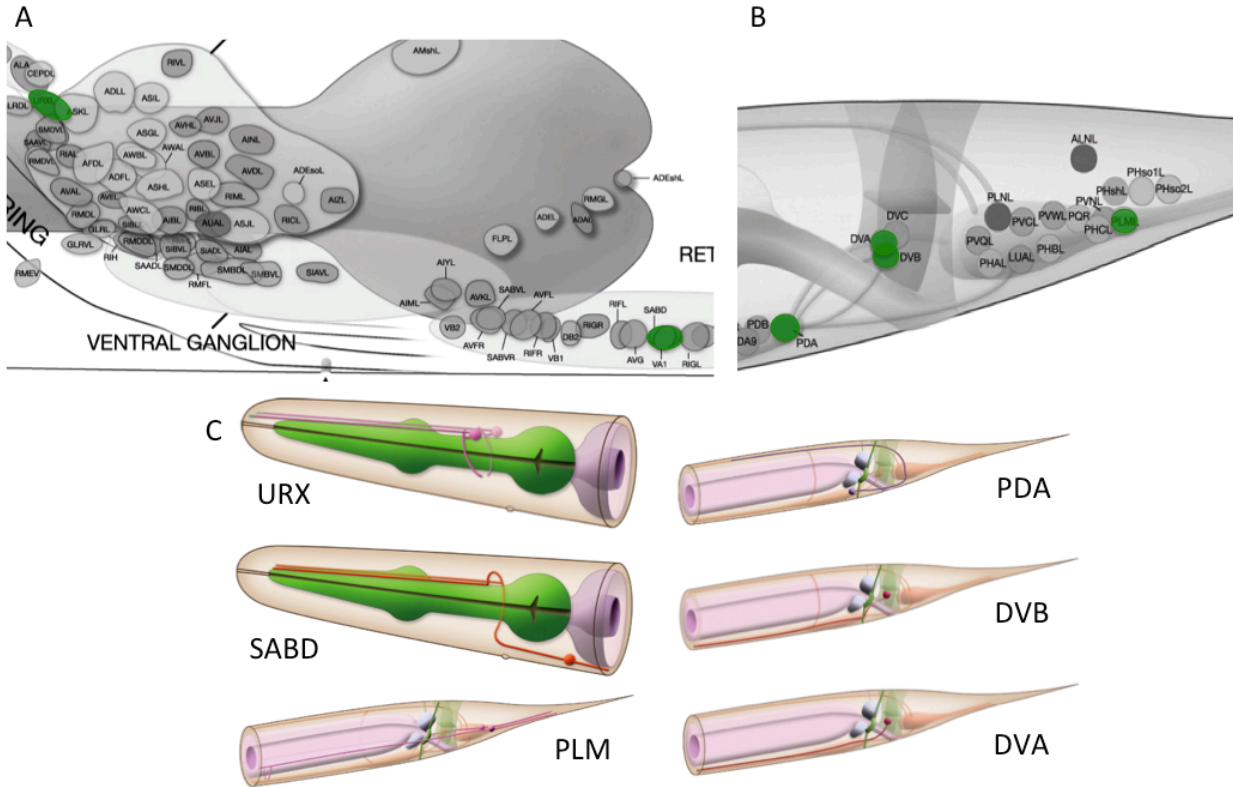
Figure 28: Schematics of neurons identified in T07F10.1::GFP worms. (A) Head neurons URX and SABD shown in green. (B) Tail neurons PLM, PDA, and DVA/DVB show in green. (C) Schematics of all neurons. All images adapted from *wormatlas.org*.
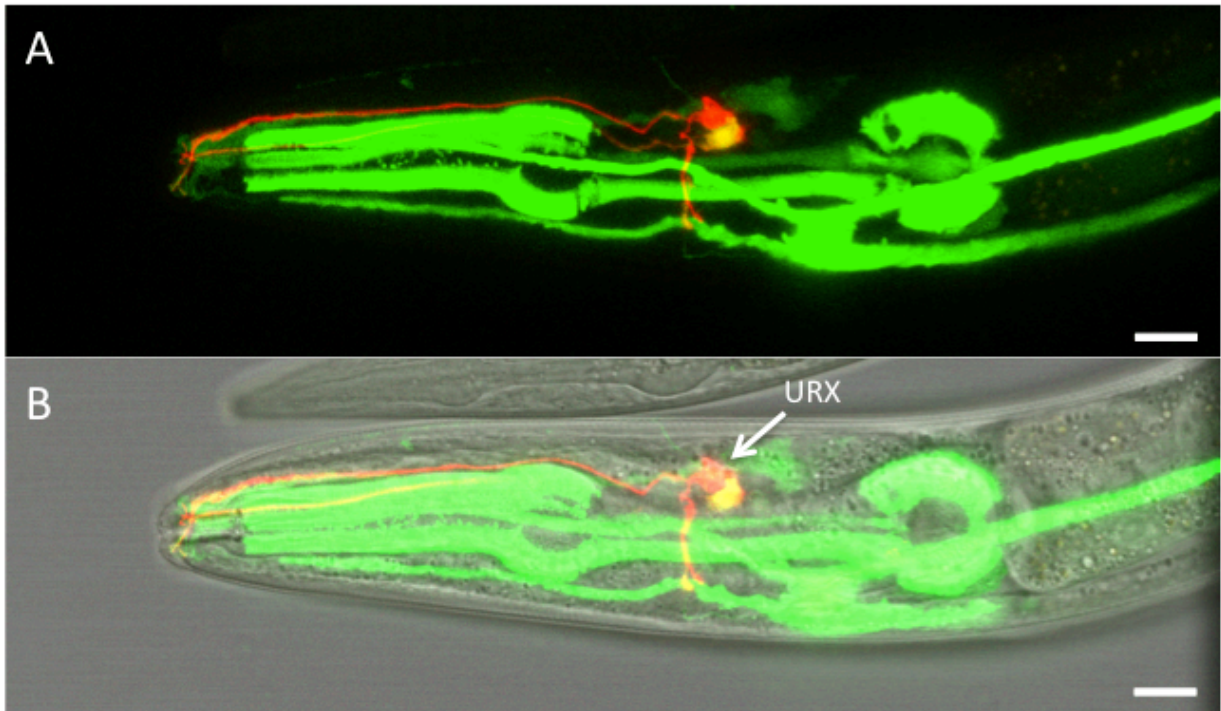
Figure 29: Confirmation of T07F10.1::GFP expression in the URX neuron. (A – B) GFP expressed in the isthmus neuron of interest co-localizes with mCherry in the URX neuron. Images depict an adult worm, and all scale bars are 10 μm.

Expression of T07F10.1::GFP in a daf-19(of5) genetic background

In order to assess the possibility that regulation of T07F10.1 might be isoform-specific, we generated an isogenic strain containing the T07F10.1::GFP transgene in an *of5* background, which lacks functional copies of DAF-19A/B. We created the strain by mating *daf-19*(*m86*) hermaphrodites with *daf-19*(*of5*) males, and selecting for dye-filling worms in the F1 generation (strain name: LU648). Confocal analysis revealed a GFP expression pattern similar to WT worms, with consistent expression in the excretory system, pharyngeal muscle, and tail neurons, for which the PDA and PLM pairs most frequently expressed GFP (Figure 30, Table 3). Inconsistent GFP expression was also observed in the intestine and in several unidentified tail cells. Notably, worms in this strain expressed GFP in the URX and SABD neurons 75% and 95% of the time, respectively. As these values are slightly higher than those previously reported by Hurlburt for WT worms, this was sufficient to suggest a WT–like phenotype as compared to

expression in the *m86* strain. As this phenotype was observed in worms lacking functional DAF-19A/B, these preliminary data would suggest that DAF-19C regulates expression of T07F10.1 in either URX or SABD.

Strain: LU648	Number of *daf-19*(*of5*) animals with observed expression of T07F10.1::GFP								
Age	N =	SABD	URX	Excretory system	PDA	PLM	DVA/ DVB	Additional tail cells	Intestine	Pharynx
L1/L2	10	10	8	10	9	10	7	5	9	10
L3	5	4	4	5	4	5	3	3	4	5
L4	2	2	2	2	2	2	1	2	1	2
Adult	3	3	1	3	1	2	1	1	3	3
Total animals with indicated phenotype (%)	20	19 (95%)	15 (75%)	20 (100%)	18 (90%)	19 (95%)	12 (60%)	11 (55%)	17 (85%)	20 (100%)

Table 14: Localization of T07F10.1::GFP expression in a *daf-19*(*of5*) background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in the excretory system and three tail neurons, with expression levels of GFP in the URX and SABD neurons that were comparable to WT worms. Inconsistent expression was observed in the intestine and several unidentified tail cells.

Figure 30: Expression of T07F10.1::GFP in a *daf-19*(*of5*) background. (A – D) Consistent GFP expression was observed in the pharyngeal muscle, excretory canal and gland, and three pairs of tail neurons (PDA, PLM, DVA/DVB). GFP was expressed in the SABD and URX neurons at a frequency slightly exceeding that of WT worms (95% and 75% respectively). Occasional GFP expression was seen in the gonad and vulval muscle, as well as in unidentified tail cells. Panels A–B depict an L3 larval worm, and panels C–D depict an adult worm. All scale bars are 10 μm.

Expression of T07F10.1::GFP in a daf-19(tm5562) genetic background

In order to further assess whether the DAF-19A/B isoforms have any effect on transcriptional regulation of T07F10.1, we generated an isogenic strain containing the T07F10.1::GFP transgene in a *tm5562* background. We created the strain by mating *daf-19(m86)* hermaphrodites with *daf-19(tm5562)* males, and selecting for dye-filling worms in the F1 generation (strain name: LU642). Confocal analysis revealed an expression pattern similar to WT worms, with consistent expression in the excretory system, pharyngeal muscle, and tail neurons, for which the PDA and PLM pairs most frequently expressed GFP (Figure 31, Table 15). Inconsistent GFP expression was also observed in the intestine and in several unidentified tail cells. Notably, worms of this strain expressed GFP in the URX and SABD neurons 65% and 91% of the time, respectively. As these values are slightly higher than those previously reported by Hurlburt for WT worms, this was sufficient to suggest a WT phenotype. Thus, it does not appear that DAF-19A is required to activate T07F10.1. Overall, these preliminary data suggest that DAF-19C may regulate expression of T07F10.1.

Strain: LU642	Number of *daf-19(tm5562)* animals with observed expression of T07F10.1::GFP									
Age	N =	SABD	URX	Excretory system	PDA	PLM	DVA/ DVB	Additional tail cells	Intestine	Pharynx
L1/L2	14	14	11	14	11	13	10	6	11	14
L3	3	3	1	3	3	2	2	0	1	3
L4	3	1	1	3	2	3	0	1	0	3
Adult	3	3	2	3	3	3	0	1	3	3
Total animals with indicated phenotype (%)	23	21 (91%)	15 (65%)	23 (100%)	19 (83%)	21 (91%)	12 (52%)	8 (35%)	15 (65%)	23 (100%)

Table 15: Localization of T07F10.1::GFP expression in a *daf-19(tm5562)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed fluorescence in the excretory system and three tail neurons. Expression levels of GFP in the URX and SABD neurons were comparable to WT worms. Inconsistent expression was observed in the intestine and several unidentified tail cells.

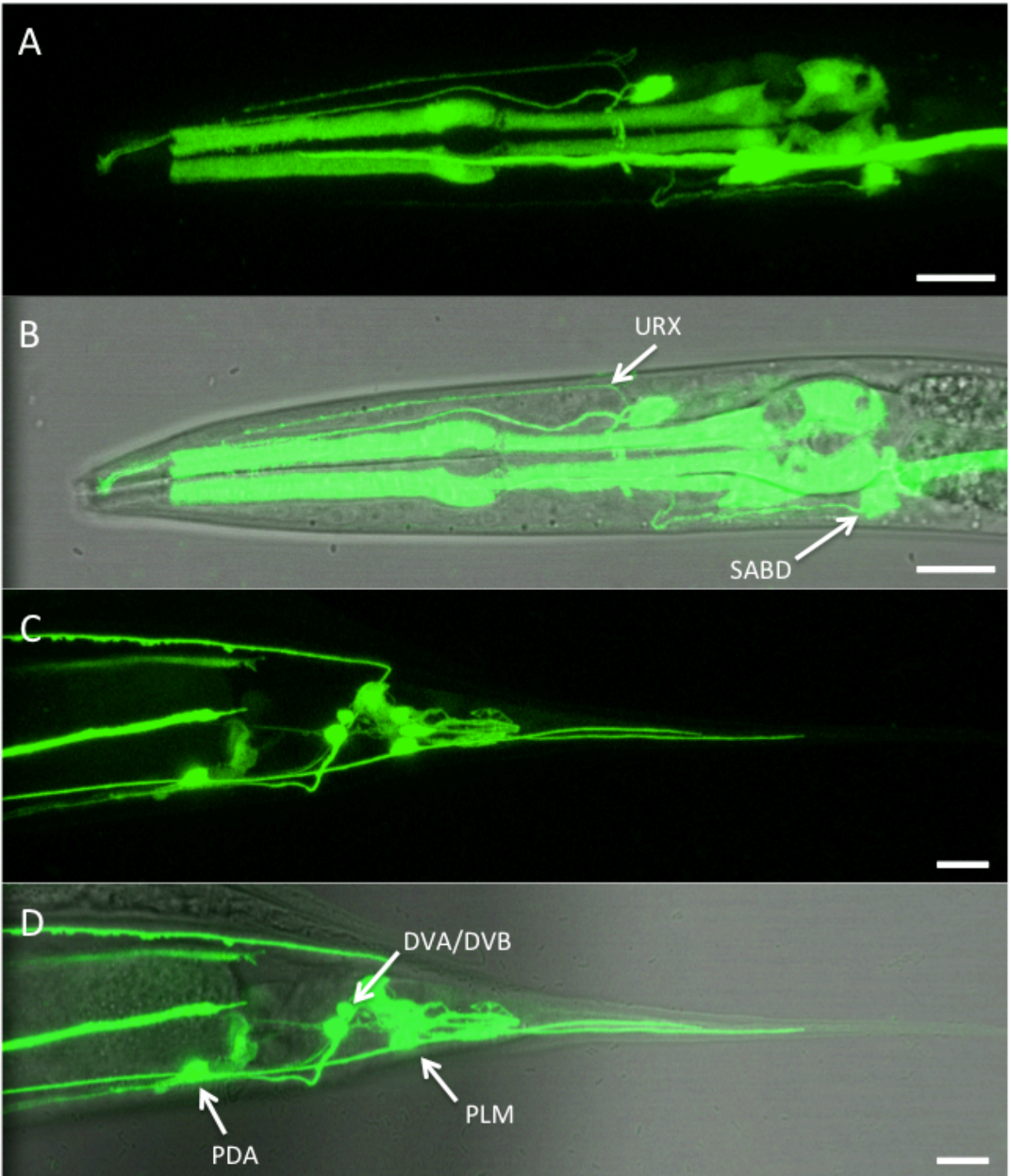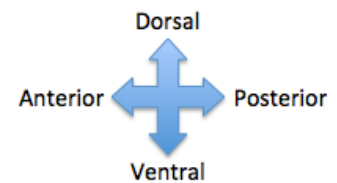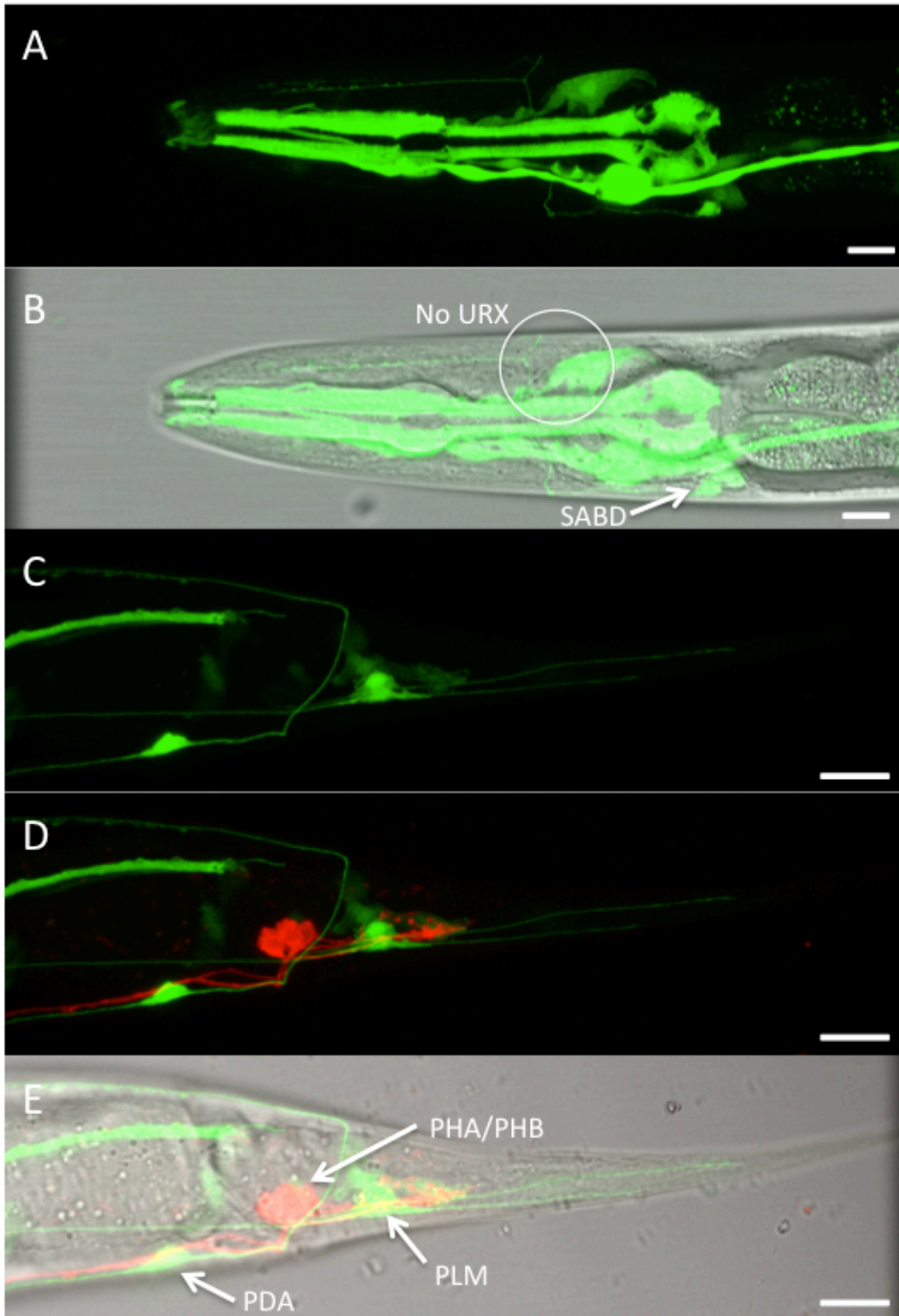**Figure 31: Expression of T07F10.1::GFP in a *daf-19*(*tm5562*) background.** (A – E) Consistent GFP expression was observed in the pharyngeal muscle, excretory canal and gland, and three pairs of tail neurons (PDA, PLM, DVA/DVB). GFP was expressed in the SABD and URX neurons at a frequency slightly exceeding that of WT worms (91% and 65% respectively). Occasional GFP expression was seen in the gonad and vulval muscle, as well as in unidentified tail cells. Panels D – E show dye-filling phasmid neurons in red as a positional marker. All worms shown are adults, and all scale bars are 10 μm.
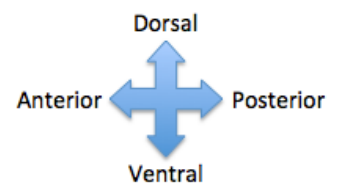
T07F10.1::GFP expression in a daf-19(m86) background is rescued by daf-19c expression

While preliminary data suggested that T07F10.1 is regulated by DAF-19C, we were uncertain as to whether the changes in GFP fluorescence identified between WT and *m86* strains constituted true differential expression. As we lacked a mutant specific to *daf-19c*, we elected to perform transformation rescues using cDNAs for *daf-19a* and *daf-19c*. Gabi Senti generated a *daf-19c* rescue strain by injecting an expression plasmid containing cDNAs for *daf-19c* into *m86* worms. An isogenic strain of double transgenic worms was then created by crossing worms from this line with worms containing the T07F10.1::GFP transgene in an *m86* background (strain name: LU676). This plasmid is overexpressed; thus, worms of this strain have an abundance of DAF-19C, but no other isoforms of the protein. Consistent GFP expression was observed in the pharynx, tail neurons, and excretory system, with occasional expression in unidentified tail cells. The two neurons of interest, URX and SABD, each expressed GFP in 94% of worms imaged (Table 16; Figure 32). These expression frequencies were significantly higher than those reported by Hurlburt in either WT or *m86* worms. As DAF-19C is overexpressed in these worms, this result is consistent with WT-like phenotypes observed in the *of5* and *tm5562* strains, which supports the hypothesis that T07F10.1 is regulated by DAF-19C.

Strain: LU676	Number of *daf-19c* rescue animals with observed expression of T07F10.1::GFP							
Age	N =	SABD	URX	Excretory system	PDA	PLM	DVA/DVB	Additional tail cells	Pharynx
L1/L2	20	20	19	19	17	20	18	1	20
L3	6	6	5	6	6	6	6	1	6
L4	1	1	1	1	1	1	1	0	1
Adult	5	3	5	4	4	5	5	1	5
Total animals with phenotype	32	30 (94%)	30 (94%)	30 (94%)	28 (87%)	32 (100%)	30 (94%)	3 (9%)	32 (100%)

Table 16: Localization of T07F10.1::GFP expression in a *daf-19C::daf- 19(m86)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues.
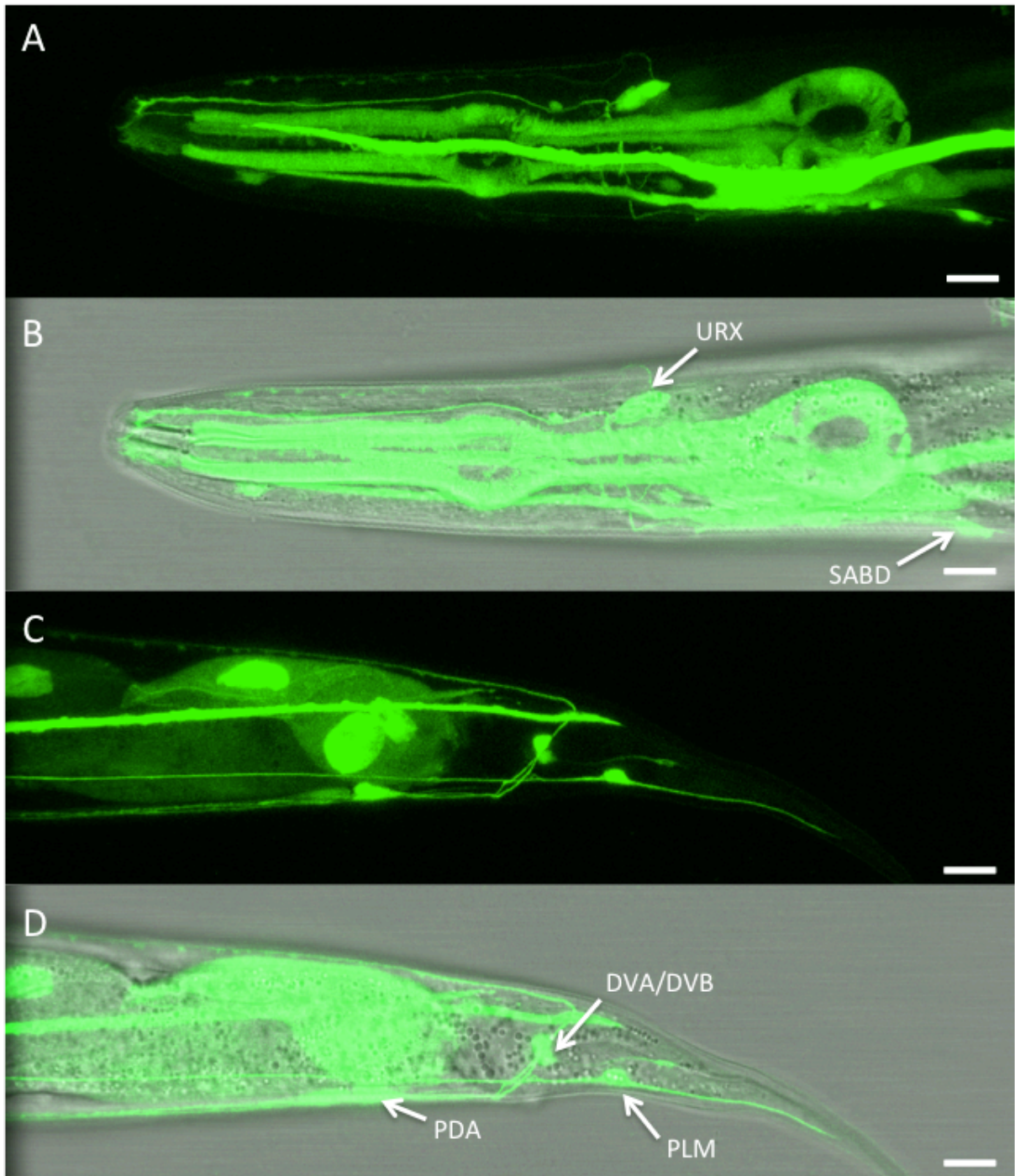
Figure 32: T07F10.1::GFP expression after transformation rescue of the DAF-19C isoform. (A – B) URX and SABD were expressed in 94% of worms imaged. Consistent GFP expression was observed in pharyngeal muscle and excretory tissues, with intermittent expression in unidentified tail cells. All worms pictured are adults, and all scale bars are 10 μm.

As a final experiment, we performed another transformation rescue experiment using *daf-19a*. As described previously, Gabi Senti generated a transgenic rescue line by microinjecting *m86* worms with a plasmid incorporating cDNAs for *daf-19a*. We then created a double transgenic rescue line using the method detailed above, which overexpressed DAF-19A only (strain name: LU675); we then expressed T07F10.1::GFP expression using confocal microscopy. Consistent GFP expression was observed in the pharynx and tail neurons, with frequent expression in the excretory system, and occasional expression in unidentified tail cells. The two neurons of interest, URX and SABD, expressed GFP in 68% and 72% of worms imaged, respectively (Table 17; Figure 33). These expression frequencies were significantly higher than those reported by Hurlburt in *m86* worms (see Table 13), but lower than those reported in WT, *daf-19c* rescue, *tm5562*, and *of5* worms. These results provide weak evidence that T07F10.1 may be regulated by DAF-19C; however, given the variability of the expression patterns between these various strains, these data are insufficient to conclusively prove that T07F10.1 is solely regulated by DAF-19.

Strain: LU675	Number of *daf-19a* rescue animals with observed expression of T07F10.1::GFP								
Age	N =	SABD	URX	Excretory system	PDA	PLM	DVA/DVB	Additional tail cells	Pharynx
L1/L2	7	7	7	6	7	7	6	3	7
L3	5	5	4	4	5	5	5	0	5
L4	6	4	5	5	6	6	6	0	6
Adult	7	2	1	6	7	6	6	0	7
Total animals with indicated phenotype (%)	25	18 (72%)	17 (68%)	21 (84%)	25 (100%)	24 (96%)	23 92%	3 (12%)	25 (100%)

Table 17: Localization of T07F10.1::GFP expression in a *daf-19A::daf- 19(m86)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues.
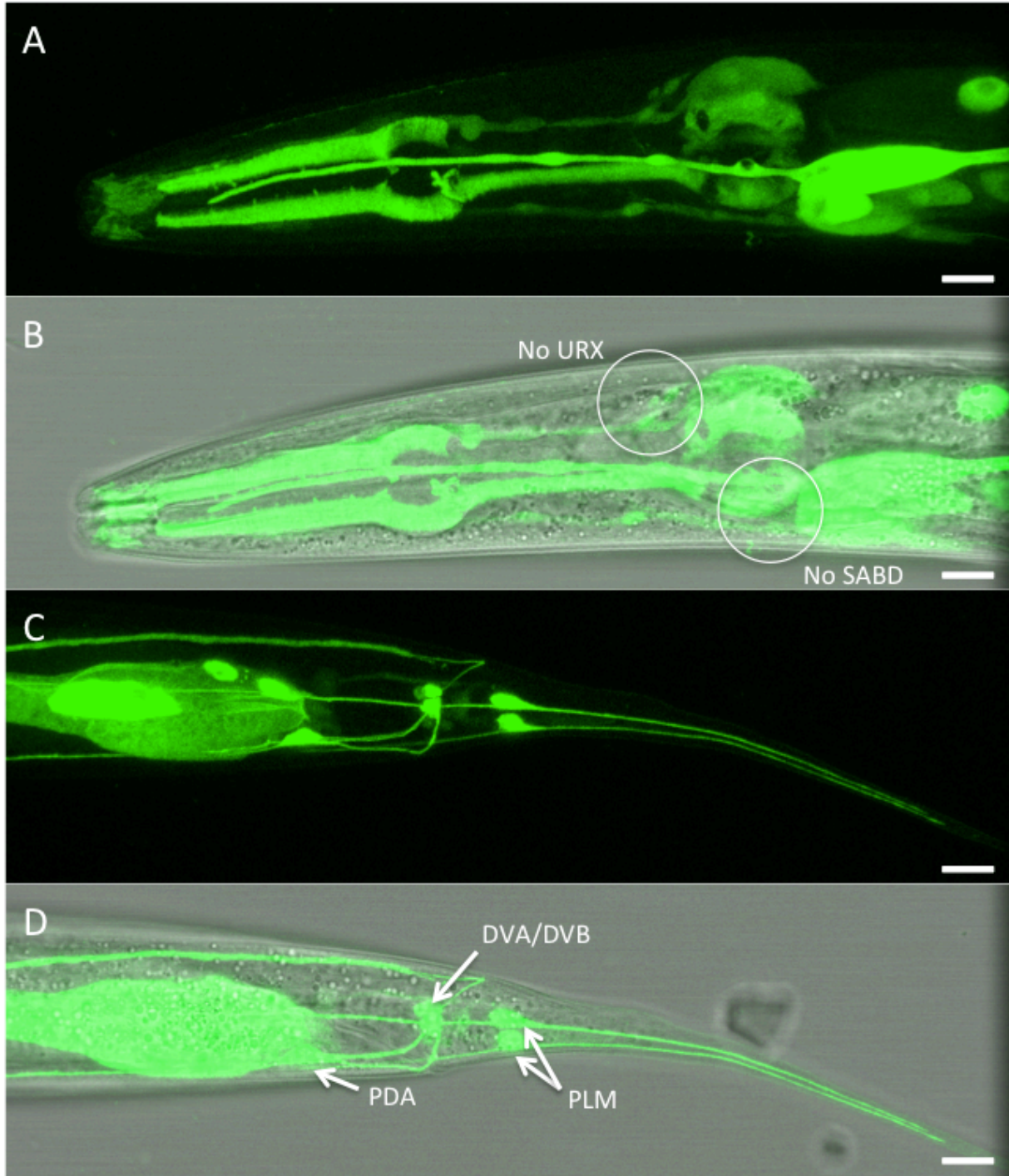
Figure 33: T07F10.1::GFP expression after transformation rescue of the DAF-19A isoform. (A – D) URX and SABD were expressed in 68% and 72% of worms imaged, respectively. Consistent GFP expression was observed in pharyngeal muscle and excretory tissues, with intermittent expression in unidentified tail cells. All worms pictured are adults, and all scale bars are 10 μm.

srd-61 expression is independent of DAF-19

      srd-61 was identified as a putative DAF-19 target gene from Prasad Phirke's microarray analysis of L1 larvae and 3-fold stage embryos, which showed a 2.2 and 1.8 fold change respectively (up-regulation in *daf-19(m86)* worms). Elizabeth De Stasio's microarray analysis of adult worms also showed a 1.48 fold change. Combined, these three microarrays suggested that *srd-61* expression is likely to be regulated by DAF-19, although *srd-61* lacks the characteristic x-box motif through which DAF-19 is known to regulate other genes. No data have been reported regarding the regulation of *srd-61* or its interaction with any other proteins; however, *srd-61* expression has been previously identified in several *C. elegans* neurons, including ASH, ASI, PHA, and PHB (Colosimo et al., 2004).

      srd-61 is a seven transmembrane G-protein-coupled receptor (7TM GPCR). This gene family consists of proteins that span the plasma membrane via a 7TM domain. The 7TM domain is comprised of seven helices embedded within the membrane that attach to an extracellular receptor, which can bind particular ligands. Upon binding, the 7TM domain will undergo a conformational change, which activates a cytosolic GTP-binding protein (G-protein) specific to the particular receptor.   The inactivated form of this G-protein carries a GDP molecule in one of its domains, and upon receiving a signal from the activated 7TM domain, it exchanges this GDP molecule for GTP. This leads to a signaling cascade that may affect various cellular processes, including gene transcription (Lehninger et al., 2008). *srd-61* has no known human homologs, and thus is associated with no specific disease states. However, humans have around 750 different GPCR genes, many of which have been associated with endocrine diseases (Vassart & Costagliola 2011). *srd-61* is a

member of the serpentine receptor class d (Srd) family, which contains various genes that regulate chemoreception (Robertson & Thomas, 2006). Chemoreception, or the ability of an organism to sense its environment through the presence of small molecules, is a critical element of olfaction. As DAF-19 has previously been shown to regulate gene expression in ciliated sensory neurons (which are responsible for olfaction), we hypothesized that DAF-19 might regulate *srd-61* in these tissues.

Expression of *srd-61::GFP* in a *daf-19(WT)* genetic background

Prasad Phirke generated a transcriptional fusion of the *srd-61* promoter region with GFP, and injected this separately into WT and *m86* worms. Confocal analysis of the WT strain (OE4124) showed consistent expression of GFP in 1-2 pairs of tail neurons as well as in a single pair of isthmus neurons that connected to the nerve ring. The latter neurons only expressed GFP in 51% of worms imaged (Table 18, Figure 34). A dye-filling assay was performed to determine the identity of these neurons. The isthmus neuron was clearly identified to be ASH, as GFP expression co-localized with red DiI in this pair of neurons. Additionally, the two pairs of tail neurons were identified to be PHA and PHB, the two dye-filling phasmid neurons (Figure 35).

Strain: OE4124	Number of *daf-19(WT)* animals with observed expression of *srd-61::GFP*		
Age	N =	ASH neurons	PHA/PHB neurons
L1/L2	39	27	39
L3	4	2	4
L4	8	3	8
Adult	23	6	23
Total animals with indicated phenotype (%)	74	38 (51%)	74 (100%)

Table 18: Localization of *srd-61::GFP* expression in a *daf-19(WT)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in the PHA/PHB neurons, and intermittent expression in the ASH neurons (51% of worms imaged).

Figure 34: Expression of *srd-61::GFP* in a *daf-19(WT)* background. (A) Consistent GFP expression was observed in a single pair of isthmus neurons, which fluoresced in 51% of worms imaged. (B – C) A dye-filling assay was used to identify this pair of neurons as ASH. (D) Consistent GFP expression was observed in 1-2 tail neurons. (E – F) A dye-filling assay was used to identify these neurons to be PHA/PHB. Panels A – C show an L2 stage worm, and panels D – E show an adult. All scale bars are 10 μm.

Figure 35: Schematics of neurons identified in *srd-61::GFP* worms. (A) Schematic of the amphid neuron ASH. (B) Schematic of the phasmid neurons PHA and PHB. All images adapted from *wormatlas.org*.

Expression of srd-61::GFP in a daf-19(m86) genetic background and an isogenic WT strain

A single line of *daf-19(m86)* worms containing the *srd-61::GFP* transgene was analyzed for GFP expression via confocal microscopy (OE4133). These worms demonstrated expression in the same head and tail neurons observed in the WT strain; however, the isthmus neuron pair fluoresced in 100% of worms imaged (Table 19, Figure 36). Given this variability in expression frequency, the preliminary data suggested that *srd-61* might be regulated by DAF-19 in the ASH neurons. To confirm this, the isogenic strain for OE4133 was created by mating *daf-19(m86)* worms containing the *srd-61::GFP* transgene to WT worms, and selecting for dye-filling progeny. This resulted in a new strain (LU627), which was similarly analyzed via confocal microscopy. This strain showed *srd-61::GFP* expression in the ASH and PHA/PHB neurons in all worms imaged; thus, the phenotype was identical to its isogenic *m86* strain (Table 20, Figure 37). These data indicate that *srd-61* is not regulated by DAF-19.

Strain: OE4133	Number of *daf-19* (*m86*) animals with observed expression of *srd-61*		
Age	N =	ASH neuron	PHA/PHB neurons
L1/L2	11	11	11
L3	3	3	3
L4	10	10	10
Adult	7	7	7
Total animals with indicated phenotype (%)	31	31 (100%)	31 (100%)

Table 19: Localization of *srd-61::GFP* expression in a *daf-19(m86)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in the ASH neurons and the PHA/PHB neurons.

Strain: LU627	Number of *daf-19* WT animals with observed expression of *srd-61*		
Age	N =	ASH neuron	PHA/PHB neurons
L1/L2	27	27	27
L3	8	8	8
L4	5	5	5
Adult	10	10	10
Total animals with indicated phenotype (%)	50	50 (100%)	50 (100%)

Table 20: Localization of *srd-61::GFP* expression in a *daf-19(m86)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms consistently showed expression in the ASH neurons and the PHA/PHB neurons.
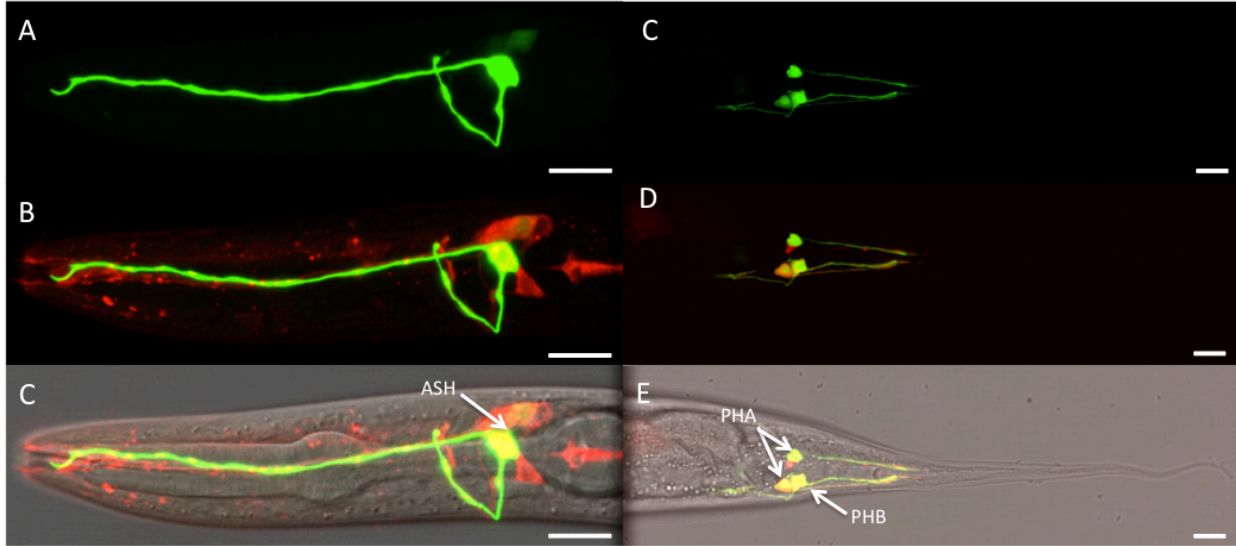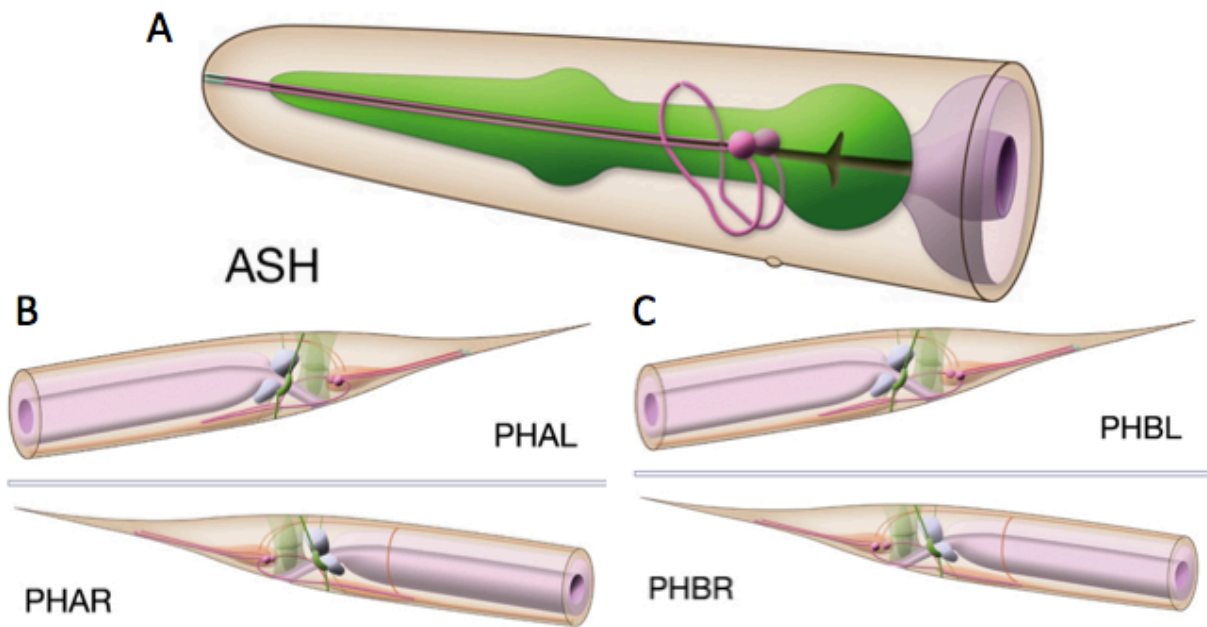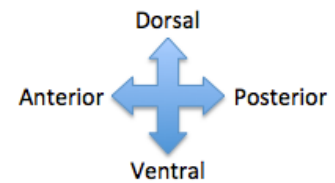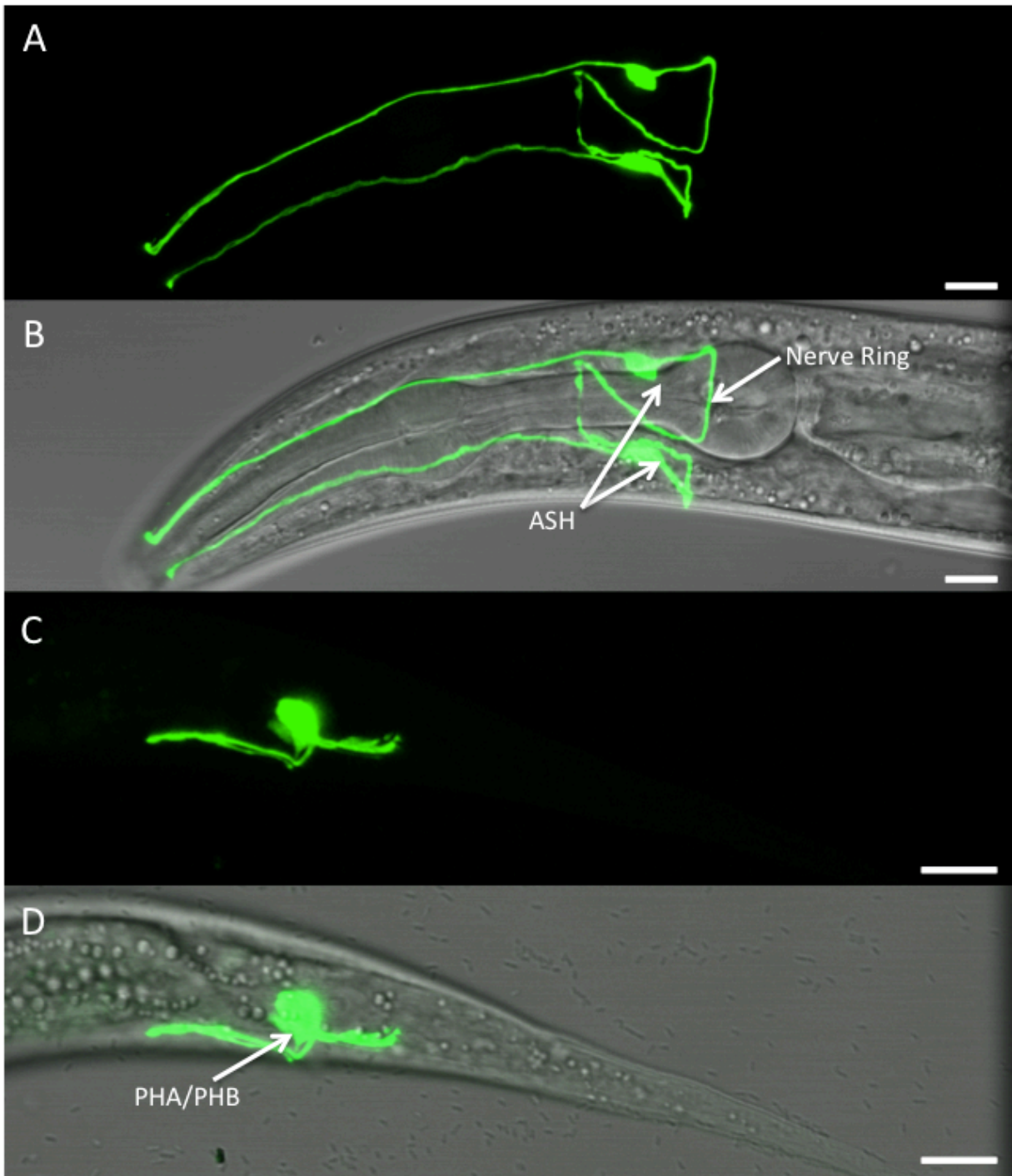
Figure 36: Expression of *srd-61::GFP* in a *daf-19(m86)* background. (A – B) Consistent GFP expression was observed in the ASH neuron pair, which fluoresced in 100% of worms imaged. (B – C) Consistent GFP expression was observed in the PHA/PHB neurons. It should be noted that this strain is not isogenic to OE4124. Panels A – B show an adult worm, and panels C – D show an L3 stage worm. All scale bars are 10 μm.
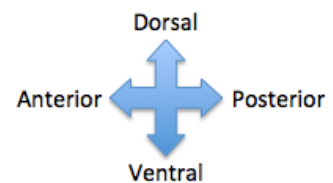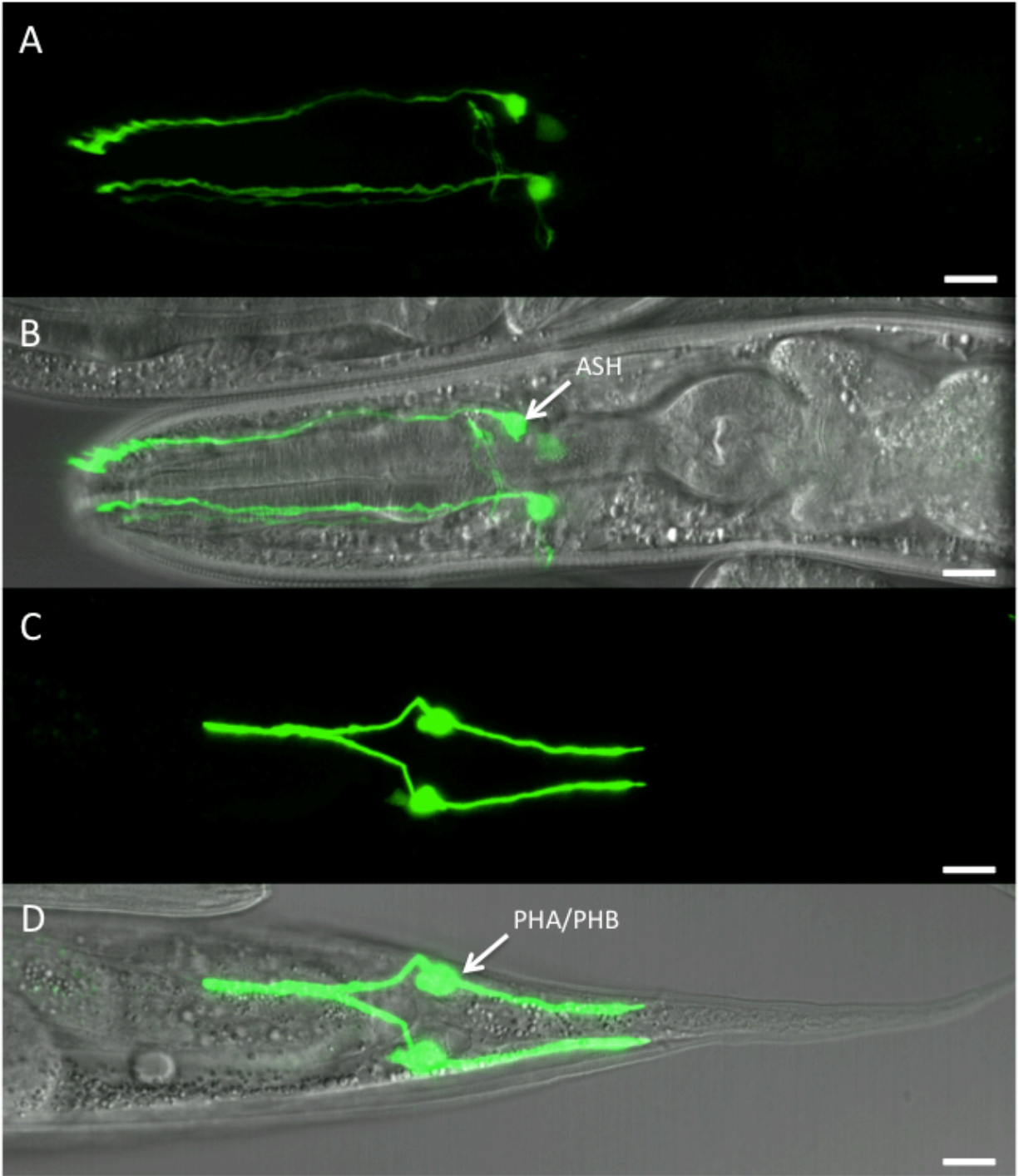
Figure 37: Expression of *srd-61::GFP* in a second line of *daf-19(WT)* worms. (A – B) Consistent GFP expression was observed in the ASH neuron pair, which fluoresced in 100% of worms imaged. (B – C) Consistent GFP expression was observed in the PHA/PHB neurons. It should be noted that this strain is isogenic to OE4133. All worms shown are adults, and all scale bars are 10 μm.
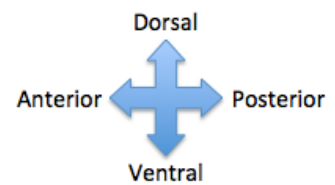
del-4 expression appears age-dependent in the absence of DAF-19

del-4 was identified as a putative *daf-19* target gene from Prasad Phirke's microarray

analysis of L1 larval worms. *del-4* is classified as a degenerin-like gene, a group of proteins

which participate in neuronal mechanotransduction, or the process by which neurons convert a

mechanical stimulus to electrical activity (Kellenberger & Schild, 2002). The 'del' gene

designation indicates that defects in these genes result in degeneration of sensory neurons

(Kellenberger & Schild, 2002). *del-4* is an ortholog of various members of the human non-

voltage gated sodium channel family, including SCNN1B, SCNN1D, SCNN1A, and SCNN1G

(NCBI.gov). These four genes code for epithelial sodium channels, which are transmembrane

proteins that selectively allow sodium ions to permeate the plasma membrane. Clinical

phenotypes associated with defects in these genes include Liddle's syndrome, a disease

characterized by hypertension resulting from abnormally high levels of sodium ions in the

kidneys (Kellenberger & Schild, 2002). Additionally, *del-4* is thought to be a potential model for

bronchiectasis, a condition characterized by chronic bronchial inflammation (Azad et al., 2009).

Using a whole-genome interactome approach, Zhong and Sternberg predicted that *del-4* interacts

with F58G6.7, a copper ion transporter, and T16G1.5, a protein of unknown function with

kinase-like domains (2006).

Transcriptome analysis of *del-4* by Elizabeth De Stasio and Prasad Phirke showed a 0.4

fold change in L1 larvae (down-regulation in *daf-19(m86)* worms) a 1.3 fold change in adult

worms, and a 1.14 fold change in 3-fold embryos (no significant expression in *daf-19(m86)*

worms). *del-4* contains the characteristic x-box motif which has been shown to be a site of *daf-*

19 mediated gene expression. As such, it was considered to be a promising candidate for further

analysis. Prasad Phirke found *del-4* expression in four neurons using fluorescence microscopy,

which he believed to be ASE, AIN, PHA, and PQR (unpublished). Additionally, Echtberger and colleagues identified *del-4* expression in the ASE gustatory neuron, which is necessary for taste reception (2007).

Expression of *del-4::GFP* in a *daf-19*(WT) genetic background

Prasad Phirke generated a transcriptional fusion of the *del-4* promoter region with GFP, and injected this into *daf-19(m86)* worms (OE3912). Hermaphrodites of this strain were mated with male worms with a *daf-19(WT)* genetic background to produce the isogenic WT strain (LU655). Confocal analysis of this strain showed mosaic expression in 0–8 head neurons and 1–3 tail neurons (Figure 38, Table 21). The most commonly observed neurons in the head included a bright pair of mid-isthmus neurons with a long dendrite reaching the tip of the nose, and a pair of neurons near the base of the nerve ring. Other neurons that appeared less frequently included an additional cell body located underneath the posterior bulb, and several fainter cell bodies in the mid-isthmus region that occasionally extended long dendrites toward the nose.

A dye-filling assay was performed to identify as many of these neurons as possible. While no GFP-expressing neurons were also found to dye fill, the dye-filling amphid neurons, these served as useful positional markers for neuronal identification. Based on its location and data from Prasad Phirke and Echtberger et al. (2007), the bright mid-isthmus neuron appearing in 60% of worms imaged is hypothesized to be ASE (Figure 39). The neurons appearing at the base of the isthmus were much more difficult to narrow down. Current hypotheses include the RMDD, AIB, RMF, RMH, and AIA neuron pairs. The tail neurons could not be definitively identified without some kind of co-localizing marker; however the most likely candidates are PQR, PHA and PHB, of which the former two neurons were hypothesized by Prasad Phirke (Figure 40).

Figure 38: Expression of *del-4::GFP* in a *daf-19(WT)* background. (A – C) Mosaic GFP expression was observed in up to 8 head neurons, with the most frequent expression occurring in the two indicated cells. A dye-filling assay was used to identify these neurons, which are hypothesized to be ASE and a second pair at the base of the nerve ring, possibly RMDD, AIB, RMF, RMH, or AIA. ASE fluoresced in 60% of worms imaged, and the ventrally located neuron fluoresced in 88% of worms imaged. (D – E) Up to 3 tail neurons located just posterior to the anus fluoresced; these are thought to be PQR, and either PHA or PHB. Panels A – C depict an adult male worm, and panels D – E depict an L3 larval worm. All scale bars are 10 μm.
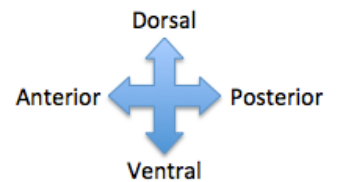
Strain: LU655	Number of *daf-19* WT animals with observed expression of *del-4::GFP*			
Age	N =	Head neurons (#)	Tail neurons (#)	Intestine
L1/L2	9	9 (1-8)	5 (1-2)	6
L3	2	2 (2-4)	1 (2)	2
L4	0	-	-	-
Adult	14	14 (2-4)	6 (1-3)	7
Total animals with indicated phenotype (%)	25	25 (100%)	12 (48%)	15 (60%)

Table 21: Localization of *del-4::GFP* expression in a daf-19(WT) background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms showed mosaic expression of GFP in up to 8 head neurons and up to 3 tail neurons, as well as intermittent intestinal expression.
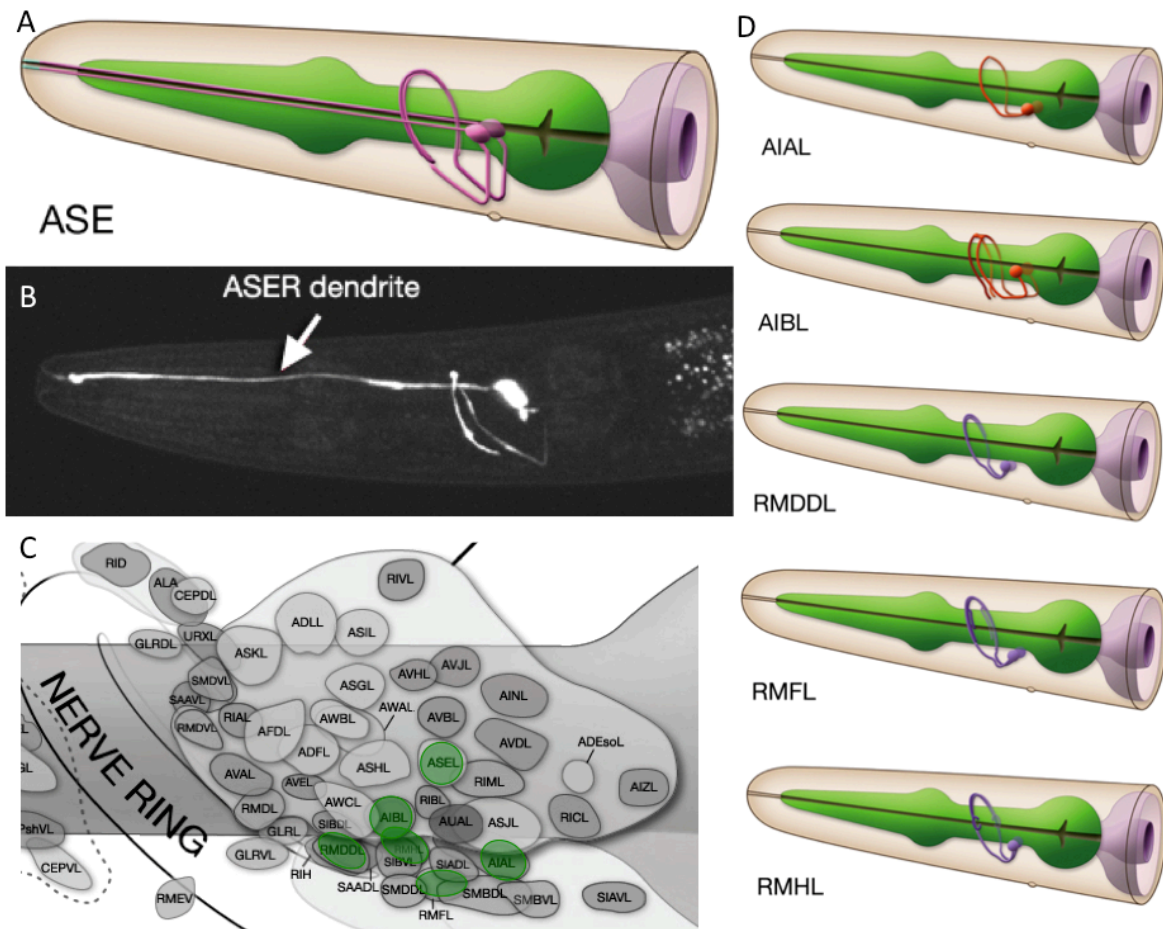
Figure 39: Schematics of neurons identified in *del-4::GFP* worms. (A) Schematic of the amphid neuron ASE. (B) Confocal image of the amphid neuron ASE. (C) Schematic of the hypothesized head neurons which may express *del-4::GFP*. (D) Schematics of five possible neurons that may be expressing GFP near the base of the nerve ring. All images adapted from *wormatlas.org*
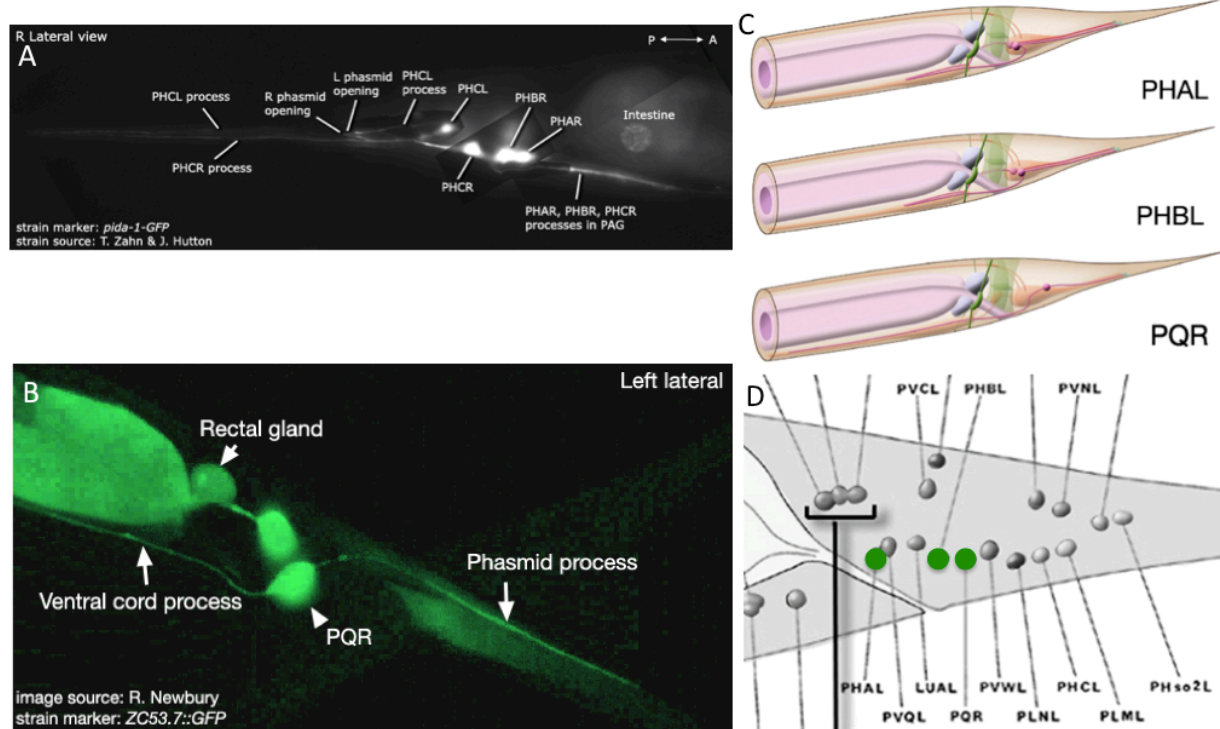
Figure 40: Schematics of neurons identified in *del-4::GFP* worms. (A) Confocal image showing the PHA and PHB neuron. (B) Confocal image showing the PQR neuron. (C – D) Schematics of the hypothesized tail neurons which may be expressing *del-4::GFP*. All images adapted from *wormatlas.org*.

Expression of *del-4::GFP* in a *daf-19(m86)* genetic background

An isogenic strain of *del-4::GFP* worms in a *daf-19(m86)* genetic background was

analyzed for GFP expression via confocal microscopy (OE3912). These worms showed a

similarly mosaic pattern of *del-4::GFP* expression in 1–6 isthmus neurons and 0–4 tail neurons

(Figure 41; Table 22). As with the WT strain, the neurons which most frequently expressed GFP

were the mid isthmus neuron pair predicted to be ASE, and the second pair immediately

posterior to the base of the nerve ring (likely to be RMDD, AIB, RMF, RMH, or AIA).

Interestingly, these worms showed age-dependent expression, as L1 and L2 larvae expressed

GFP in an average of 3.58 neurons, while adults expressed GFP in an average of 1.67 neurons

(Figure 42). This age-dependent expression was not visible in WT worms. In spite of the

variability in age-related expression between the WT and *m86* strains, the overall similarity in

phenotypes suggests that *del-4* is not regulated by *daf-19*, and further experiments will be

required to determine whether the age-dependent expression in *m86* worms is significant to our
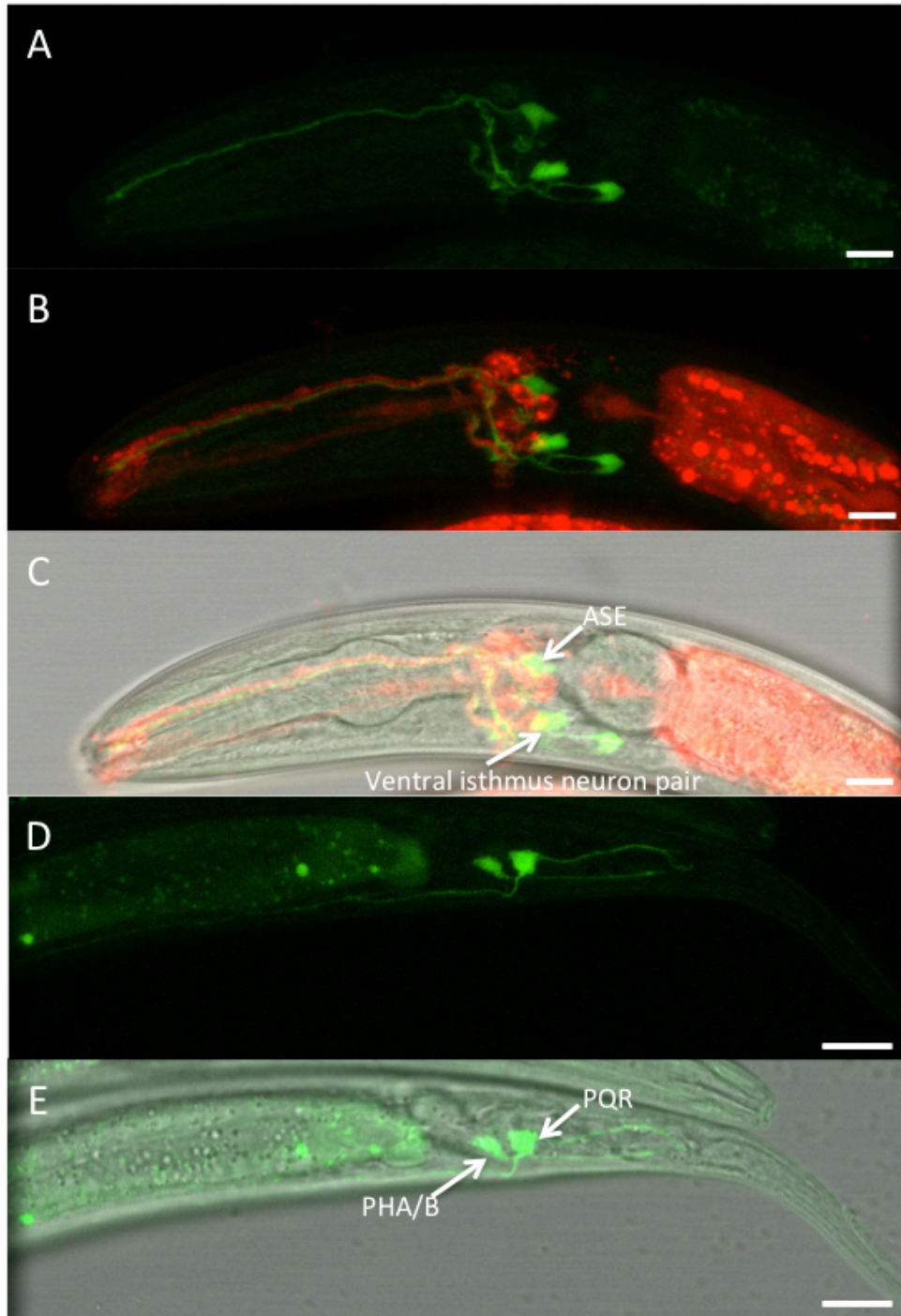
study.
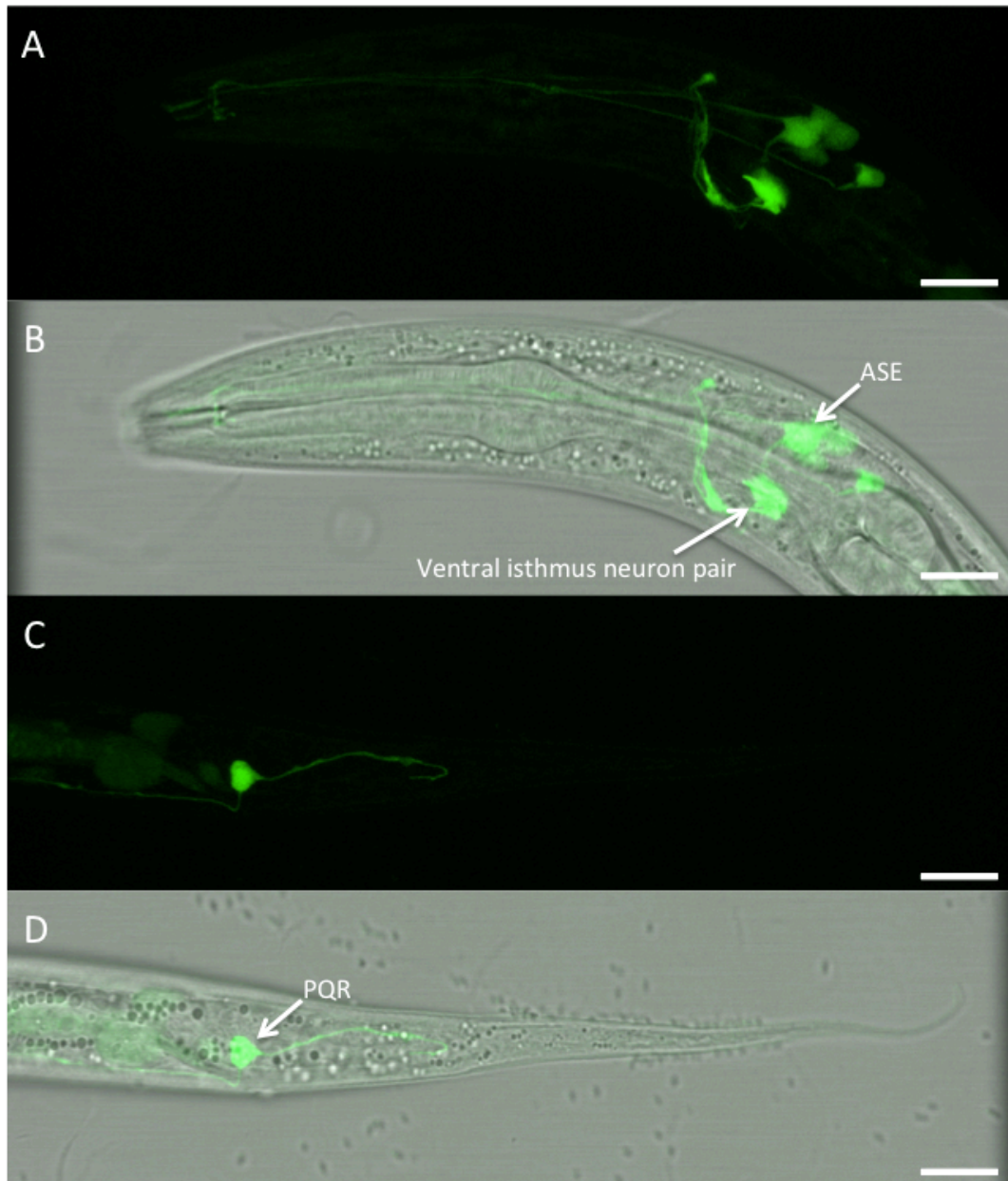
Figure 41: Expression of *del-4::GFP* in a *daf-19(m86)* background. (A – B) Mosaic GFP expression was observed in up to six head neurons, with the most frequent expression occurring in the two indicated cells, hypothesized to be ASE and a second unidentifiable pair. ASE fluoresced in 65% of worms imaged, and the ventrally located neuron fluoresced in 95% of worms imaged. (C – D) Up to four tail neurons located just posterior to the anus fluoresced; these are thought to include PQR, and either PHA or PHB. All images show L2 worms, and all scale bars are 10 μm.
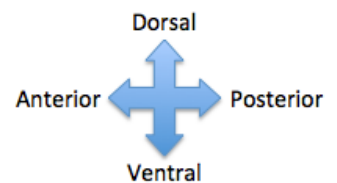
Strain: OE 3912	Number of *daf-19(m86)* animals with observed expression of *del-4::GFP*			
Age	N =	Head neurons (#)	Tail neurons (#)	Intestine
L1/L2	17	17 (2-6)	9 (1-3)	14
L3	10	10 (1-4)	4 (1-4)	8
L4	3	3 (1-3)	0	2
Adult	7	6 (1-2)	3 (1-3)	4
Total animals with indicated phenotype (%)	37	36 (97%)	16 (43%)	28 (76%)

Table 22: Localization of *del-4::GFP* expression in a *daf-19(m86)* background. Hermaphrodite worms of indicated ages were observed with confocal microscopy, with expression appearing in the indicated tissues. Worms showed mosaic expression of GFP in up to six head neurons and up to four tail neurons, as well as intermittent intestinal expression.
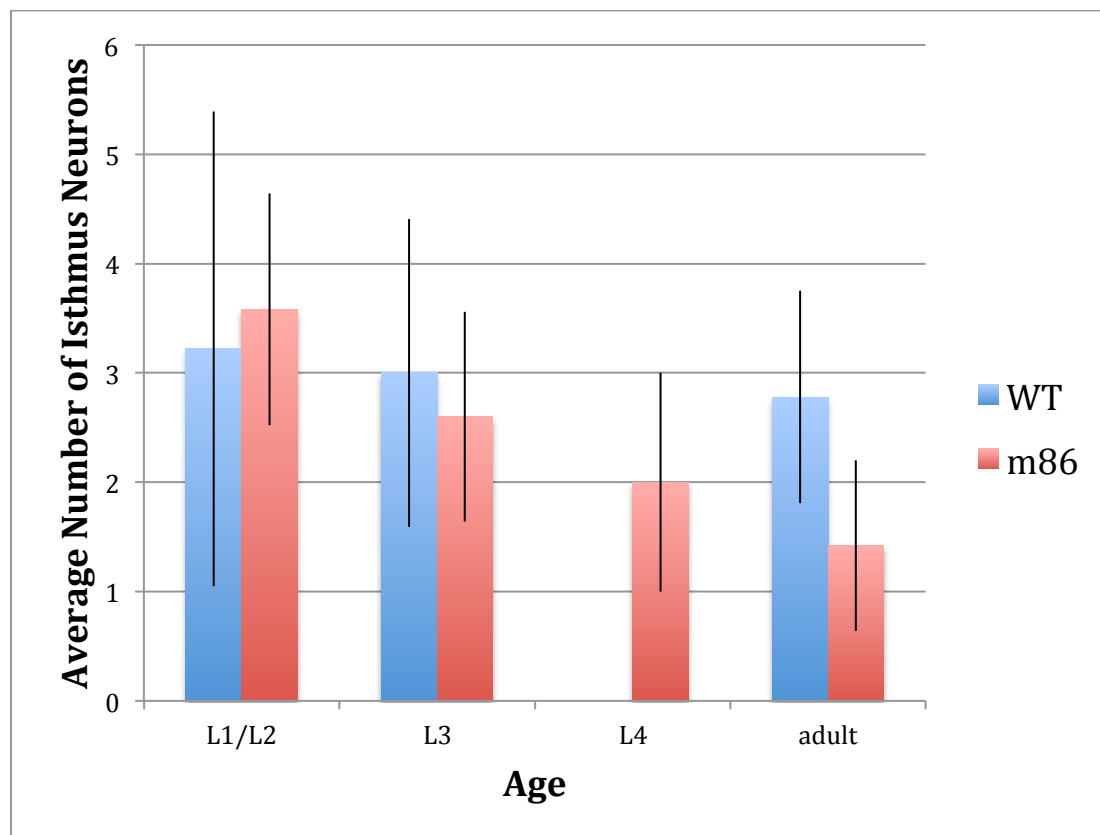

Figure 42: Average isthmus neuron expression of *del-4::GFP* is age-dependent a *daf-19(m86)* background. *m86 w*orms showed GFP expression in progressively fewer neurons as they aged, with a statistically significant difference between L1/L2 larvae and adult worms. No statistically significant difference was seen in WT worms. Error bars show 1 standard deviation.

***decr-1.1* expression was not observed in *daf-19(WT)* or *daf-19(m86)* worms**

decr-1.1 is a member of the DiEnoyl-CoA reductase family of mitochondrial genes. While little is known about *decr-1.1*, other *C. elegans* genes in this family are orthologous to the human DECR1 protein, which plays a role in metabolizing fatty enoyl-CoA esters (Helander et al., 1997). *decr-1.1* has no known interactions with other proteins, and no data have been reported regarding its regulation. It was identified as a putative DAF-19 target gene by Prasad Phirke's microarray analysis of L1 larvae and 3-fold stage embryos and Elizabeth De Stasio's microarray analysis of adult worms, with fold changes of 0.41, 0.46, and 2.6, respectively. This indicates that *decr-1.1* is up-regulated in adult worms and down-regulated in embryonic and larval worms with a *daf-19(m86)* genetic background. While *decr-1.1* lacks the characteristic x-box motif through which DAF-19 is known to regulate other genes, the degree of differential expression in WT vs. *m86* worms suggests that it is a probable candidate for DAF-19 dependency.

A transcriptional fusion of the *decr-1.1* promoter region with GFP was generated and injected separately into *daf-19(WT)* and *daf-19(m86)* worms (Prasad Phirke). Both strains were analyzed for GFP expression via confocal microscopy. Surprisingly, neither strain showed visible GFP expression in any tissues (Table 23). As such, it was impossible to determine whether *decr-1.1* is regulated by DAF-19, as no phenotype could be established. Each strain expressed an mCherry transgenic marker that was also incorporated into the plasmid, so the problem does not appear to lie with faulty microinjections. Rather, the *decr-1.1* promoter region may not be fully contained by the transcriptional fusion. Additionally, it is possible that levels of *decr-1.1* expression are simply too low to be visualized with GFP in either genetic background. Two additional lines of worms have been injected with the *decr-1.1::GFP* transgene and will be

analyzed for GFP expression. If these fail to show fluorescence, it may be necessary to create a new transcriptional fusion incorporating a larger portion of the sequence upstream of *decr-1.1*, in order to better capture the promoter region.

Genetic Background	*daf-19(m86)*		*daf-19(WT)*	
Age	N =	Number of worms showing GFP expression	N=	Number of worms showing GFP expression
L1/L2	14	0	12	0
L3	15	0	9	0
L4	13	0	13	0
Adult	8	0	16	0
Total animals with indicated phenotype (%)	50	0 %	50	0 %

Table 23: Transgenic expression of *decr-1.1::GFP* in *daf-19(WT)* and *daf-19(m86)* genetic backgrounds. GFP was not observed in worms expressing the *decr-1.1::GFP* transgene in either WT or *m86* genetic backgrounds for *daf-19.*

CRISPR Design

Of the five genes studied, at least one, T01B11.2, appears very likely to be regulated by DAF-19C, and T07F10.1 may be as well. More importantly, none of the genes studied appear to be regulated by DAF-19A/B. In order to confirm that DAF-19C is responsible for the expression patterns observed, it is important to be able to show not only that worms containing functional DAF-19C in a *daf-19(m86)* background show a WT phenotype, but also that worms lacking only this isoform show an *m86* phenotype. While rescue experiments using *daf-19a* cDNAs can mimic this genetic background, there are confounding factors that could affect the accuracy or viability of these experiments. Firstly, the amount of DAF-19 expressed in a rescue experiment depends on the number of copies of the transgene that are contained in each worm, which will result in different levels of expression than in worms expressing chromosomal *daf-19*. Secondly, *daf-19a* rescue experiments were not possible in T01B11.2::GFP transgenic worms, as they express GFP too brightly to see the fluorescent marker that identifies *daf-19A::daf-19(m86)*

worms. These worms are dye-filling defective, so we could not use a dye-filling assay to identify them as we did with *daf-19C* cDNA in *daf-19(m86)* worms. As T01B11.2 shows the strongest evidence for DAF-19C dependent expression of all the genes studied, we wished to find a new way to test transgenic expression in a *daf-19c -/-* background. As such, we designed two CRISPR mutagenesis strategies intended to knock out DAF-19C while leaving DAF-19A/B intact.

It should be noted that the precise sequence of the DAF-19C protein is not known, as conflicting evidence suggests that this isoform begins either in exon 4 or in exon 5 (Figure 43). *daf-19c* was originally thought to begin in exon 4, based on RNase protection assays completed by Gabi Senti in 2008. However, two separate transcripts for a potential short form *daf-19c* isoform are reported on *wormbase.org*, one of which begins with exon 4 and the other with exon 5. Additionally, data from the *daf-19(of6)* mutant, which affects exon 4, suggests that this exon is not included in DAF-19C and that translation begins with the ATG at the start of exon 5. The *of6* mutant, generated by Debora Sugiaman, has a WT dye-filling phenotype. This suggests that *of6* worms express functional DAF-19C, as the dye-filling phenotype depends on the presence of functional sensory cilia in the amphid neurons. These cilia only develop in the presence of DAF-19C, which is responsible for activating genes required for ciliogenesis. As such, the *of6* mutant may not affect *daf-19c* at all, which would support evidence for a transcript beginning in exon 5. While other lab members are currently addressing this question, the subsequent designs all assume that *daf-19c* does in fact begin with exon 5. Implementation of these designs will be carried out this summer by Debora Sugiaman.
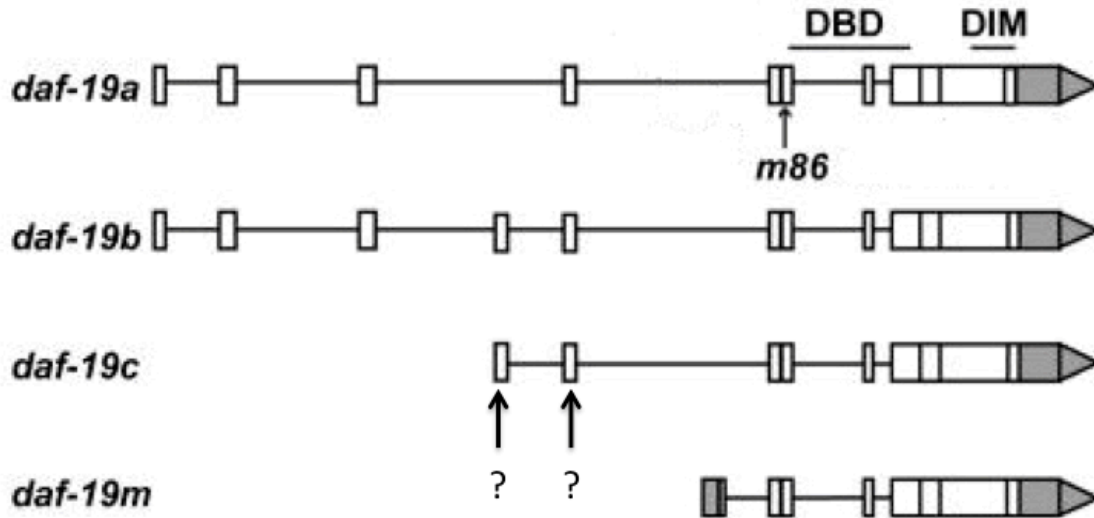
Figure 43: Potential *daf-19c* start sites. Conflicting evidence from Gabi Senti and Debora Sugiaman suggests that *daf-19c* may actually begin in exon 5. Image adapted from Wang et al. 2010.

Design and Workflow Summary

The two CRISPR designs detailed here are based on work done by Paix and colleagues (2015), who present an efficient protocol for mutagenesis that should ideally take only a few weeks to complete. In the first design, we propose to replace the start codon of exon 5 with an alanine codon, in the hopes that the change will not affect the functionality of DAF-19A/B, while stalling translation of DAF-19C. The second approach incorporates a double frameshift mutation that will knock out DAF-19B/C, but should be corrected in DAF-19A (Figure 44).

The approach that Paix and colleagues describe combines a cloning-free protocol with the co-CRISPR method developed by Arribere et al. (2014). In this method, worms are injected with an *in vitro*-synthesized Cas9-crRNA-tracrRNA complex, thus eliminating the need to clone either the guide RNA or the repair template into an expression vector. In the co-CRISPR strategy, two edits are made: one for the desired mutation, and one to produce a dominant marker mutation that is easily visible in both homozygous and heterozygous worms. This marker identifies worms that have received sufficient quantities of Cas9 protein, crRNA, tracrRNA, and

homology repair template for at least one successful mutation, which is a strong predictor of a second successful mutation. As such, only those populations that contain large numbers of worms with the marker mutation need be screened for the desired edit (Figure 45). That marker mutation can then be selected against when picking worms to initiate subsequent generations, as it will most often be heterozygous. Paix and colleagues found that up to 70% of worms with the marker mutation contained the second desired edit; thus, the method is highly robust. They used a dominant *dpy-10* (roller) mutation as their marker. However, as both *daf-19* and *dpy-10* are located on chromosome II, this is not ideal for our project. In such cases, Paix and colleagues recommend tagging the constitutively expressed (but nonessential) gene *gtbp-1* with *eGFP* to create a fluorescent marker for screening which can be easily outcrossed.
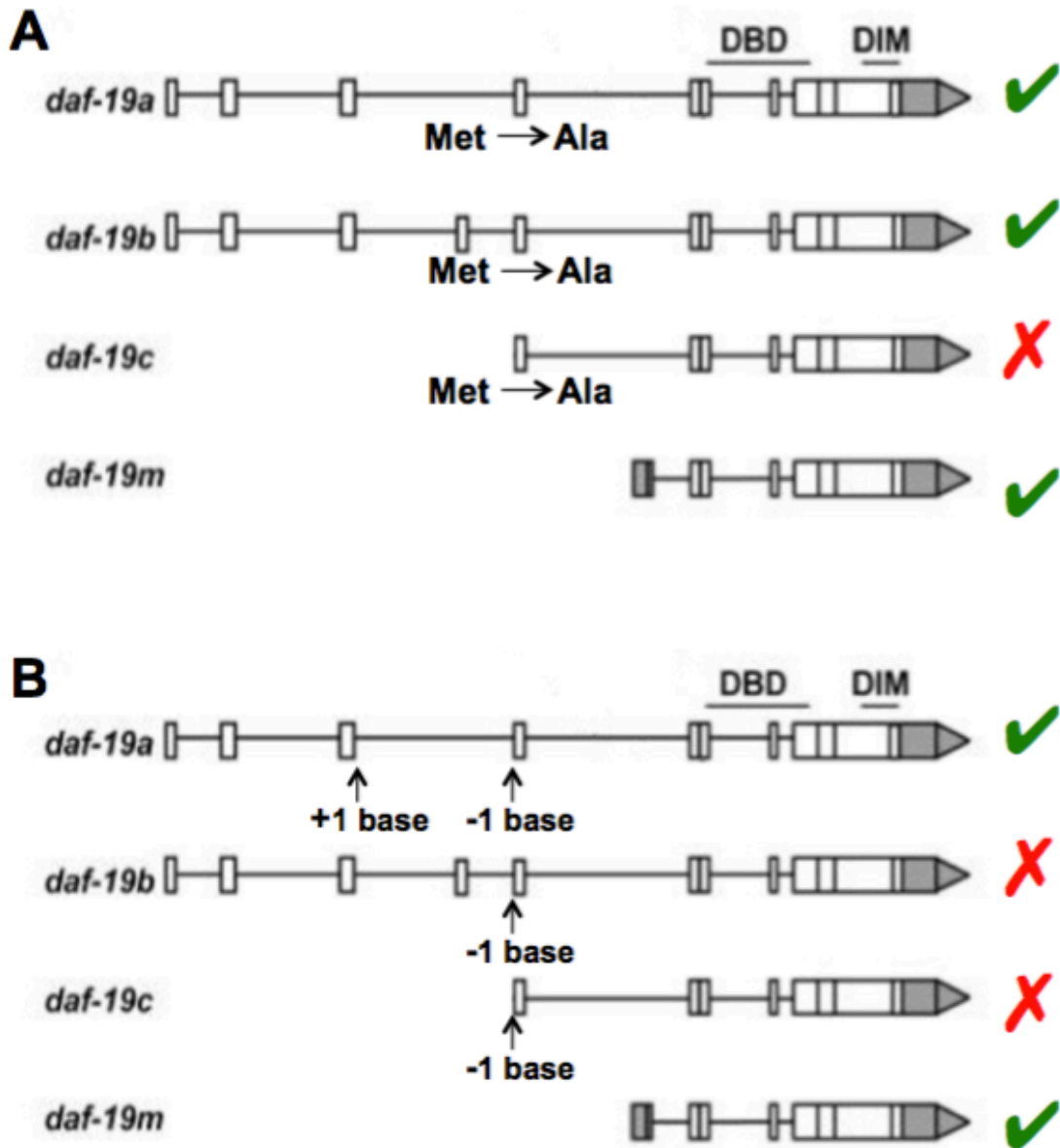
Figure 44: CRISPR mutagenesis approaches. Figure A (mutagenesis #1) shows the location of a single amino acid change from methionine to alanine at the start codon of exon 5, which is expected to stall translation of DAF-19C without significantly affecting the other isoforms. Figure B (mutagenesis #2) shows the location of a double frameshift mutation that removes a single base from the beginning of exon 5 and adds the same base to the end of exon 3. This is expected to result in two frameshift mutations that will knock out DAF-19B/C while preserving the function of DAF-19A. Green checkmarks note hypothesized functional isoforms, while red X's show the presence of null alleles. Figure adapted from Wang et al. 2010.
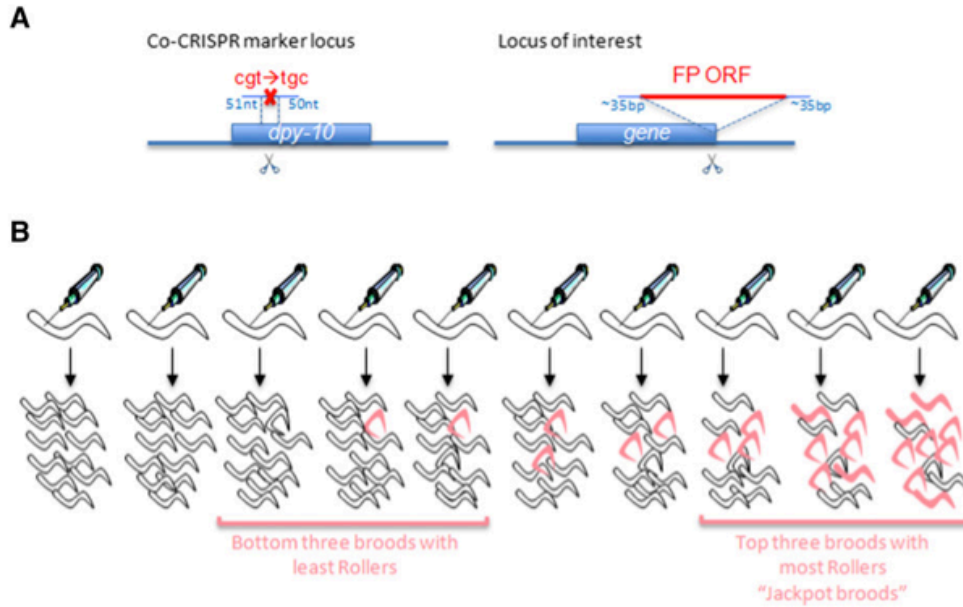
Figure 45: Co-CRISPR strategy. (A) The marker mutation at *dpy-10* is shown along with the mutation at the locus of interest (Insertion of a fluorescent protein at the desired site). (B) Parent worms are injected and screened for roller progeny; those with high proportions of roller worms are termed "jackpot broods." Figure adapted from Paix et al. (2015).

CRISPR Approach #1:

We chose to replace the start codon of exon 5 (Figure 46) with alanine instead of methionine, as both residues are nonpolar, and alanine has a very small side chain. As such, alanine will hopefully behave in a chemically similar way to methionine when incorporated into the fully folded protein. We used the *crispr.mit.edu* website to design guide RNAs near the exon 5 start codon. The program designed four possible guides (Figure 47; Table 24).

gagattattcgtaggtcgaaaacaacaatctagcaccgcgtttaatatttcatgagcactttgagttgaagaatgaagagaa**ataatg gaattttgaggagaaggatgattaaaaatttaaatttcgaaaattttcagaa**<mark style="background-color:red">ATG</mark><mark style="background-color:yellow">GAGGTCATCCAACAC TCGACAGACGATCCGAA</mark>TGGCACGCGAGAGGAATTCGACTATAATCAAATT GAATATGGAAAgtgagttgtgaaatataattggggagtctgaaacgtgaagtcttaaaaataaataatgaatataggcaa agaag

Figure 46: *daf-19* exon 5 and adjacent nucleotides. Exon 5 is depicted in capital letters and yellow highlighting, while lowercase letters show the immediately adjacent intron regions. The start codon is highlighted in red. Letters notated in bold show the sequence of nucleotides that was scanned for possible guide RNAs by the MIT CRISPR design tool.
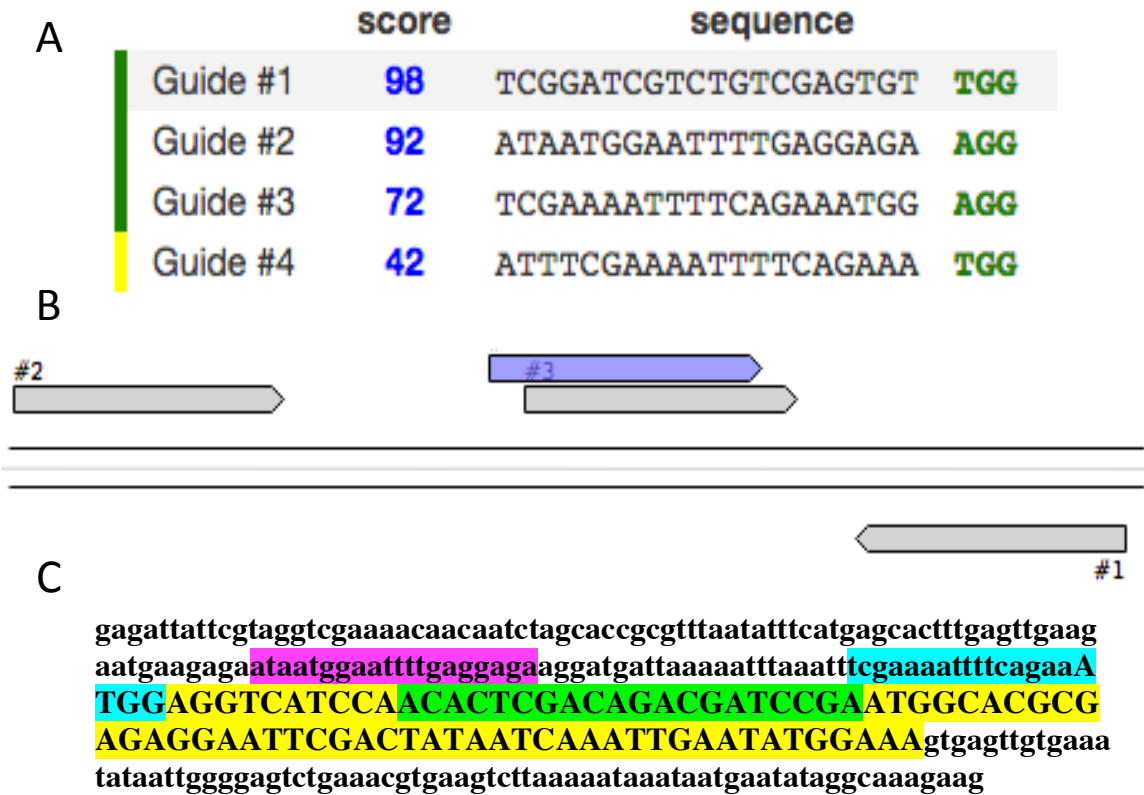
Figure 47: Sequence and location of potential crRNAs. (A) All guide RNAs were designed by the *crispr.mit.edu* tool, with scores reflecting the predicted relative accuracy of each guide. Sequences are shown 5' – 3', with adjacent PAM sequences in green letters. According to the program, each of the first three guides is considered high quality. (B) The schematic shows the relative location and direction of each crRNA. (C) Complementary sequences of the guides are shown with exon 5 and the immediately surrounding introns. Exon 5 is shown in capital letters and yellow highlights. Guide #1 is henceforth shown in green, guide #2 is shown in pink, and guide #3 is shown in blue.

crRNA	Guide sequence	PAM	Protospacer (coding strand, 5' – 3')	Off-target sites	Off-target sites in genes
1	UCGGAUCGUCUGU CGAGUGU	TGG	ACACTCGACAGACG ATCCGA	6	6
2	AUAAUGGAAUUUU GAGGAGA	AGG	ATAATGGAATTTTGA GGAGA	48	25
3	UCGAAAAUUUUCA GAAAUGG	AGG	TCGAAAATTTTCAGA AATGG	167	78

Table 24: Potential crRNA sequences. The sequence of each crRNA is shown with the corresponding protospacer sequence and PAM sequence, as well as the number of predicted off-target sites.

Various recommendations have been made with regards to designing efficient crRNAs. These include the following:

1. Several papers recommend designing crRNAS with 50-75% GC content (Doench et al., 2014, Gagnon et al., 2014). Guide #1 meets this criterion (55%), while guides 2 and 3 do not (30% each).

2. The presence of a GG dinucleotide immediately upstream of the PAM sequence significantly improves crRNA efficiency. A G is acceptable, but a C should not be used in this position (Farboud & Meyer, 2015). Guide #3 meets this criterion.

3. The proximity of the cleavage site to the desired edit site is a significant predictor of crRNA efficiency; Paix and colleagues recommend choosing crRNAs <10 bases away from the edit site. Guide #3 meets this criterion.

4. It is preferable to use crRNAs with few predicted off-target sites. For these sites, it is best to have at least 3 mismatches (preferably close to the PAM sequence) to minimize the risk of an additional cut. For the three potential crRNAs, guide #1 is by far the best choice in this regard (only 6 off target sites, all of which have 4 mismatches). Guide #2 has a significant number of off target sites, all of which have 3-4 mismatches. Guide #3 has the greatest potential for off-target effects; however, only two of the listed loci have fewer than 3 mismatches.

Due to its proximity to the desired mutation locus and the presence of a GG dinucleotide, we propose guide #3 to be the best candidate for CRISPR; however, we have created homology repair template designs for both guides #1 and #3, as described below. Whichever crRNA is used should immediately precede the universal sequence GUUUUAGAGCUAUGCUGUUUUG, which binds to the tracrRNA in order to guide the Cas9 protein to the desired site. Thus, if guide

RNA #1 is to be used, the full crRNA will read:

UCGGAUCGUCUGUCGAGUGUACACTCGACAGACGATCCGA.

Homology repair templates

Homology repair templates were designed based on recommendations from several sources, which include the following suggestions:

1. Repair templates should have homology arms that are ~35 nucleotides long, and should ideally terminate with a C or G (Paix et al., 2014).

2. Repair templates should be designed to avoid hairpins whenever possible (Paix et al. 2014). These can be tested for using an oligonucleotide property calculator.

3. Mutations should be introduced to the repair template to disrupt the crRNA sequence or the PAM sequence, so as to prevent a second cleavage event. Paix and colleagues recommend creating 2-4 mutations to disrupt the crRNA sequence, and note that changes to the PAM sequence are most effective.

4. When making edits, codon bias for *C. elegans* should be taken into account, and should match the original codon as closely as possible. Additionally, it is important to make only silent mutations. Paix and colleagues also recommend avoiding changes to noncoding regions, so as to avoid unknown regulatory motifs.

Repair template for crRNA # 3:

A repair template was designed to accompany crRNA #3 (Figure 48). This template is 75 bp long (~35 on each side of the edit), and ends with a guanine nucleotide on each end, as recommended by Paix and colleagues (2014). Potential hairpins, while not completely eliminated, were minimized as much as possible. The PAM sequence and protospacer region are

modified by a total of 8 mutations, so as to avoid repeated cleavage events in the same location.

Exon-specific changes are reflected in Table 25 with their respective codon bias. Higher codon

bias values indicate that a codon will be used more frequently; as such, any introduced codons

should ideally be used at least as frequently as the original codons.

A

ggatgattaaaaatttaaatttcgaaaattttcagaaATGGAGGTCATCCAACACTCGACAGACGATCCGAATGG

B

gtttgattaaaaatttaaatttcggatcctttcagaaGCTGAAGTCATCCAACACTCGACAGACGATCCGAATGG

Figure 48: Proposed homology repair template for crRNA #3. (A) The original genomic sequence covered by the repair template. crRNA is shown in blue, and exon 5 is shown in yellow highlighting and capital letters. (B) Proposed homology repair template. crRNA is shown in blue, and mutations are shown in red.

		Original codon	New codon
Mutation #1	Codon	ATG	GCT
	Amino acid	Methionine	Alanine
	Codon bias	1.0	0.36
Mutation #2	Codon	GAG	GAA
	Amino acid	Glutamic Acid	Glutamic Acid
	Codon bias	0.38	0.62

Table 25: Codon changes for repair template #3. Codon biases were calculated using the Codon Usage Database at *http://www.kazusa.or.jp/codon*.

Among the changes built into the intron region of the protospacer is a BAM HI restriction

site (GGATCC; Figure 49). This restriction site is intended to allow easy genotyping of worms

containing the repair template by isolating the sequence immediately around the edit, amplifying

it with PCR, and performing a restriction digest. This is shown in the sequence below, with the

restriction site depicted in gray, exon 5 shown in yellow/all caps, and forward and reverse primer

sites shown in red. After PCR and restriction digest, we would expect to see a single 325 bp

fragment in worms without a successful CRISPR edit, and two fragments (210 and 115 and bp,

respectively) in worms with a successful CRISPR edit.

*ccgcttatctgtctttttcctctgtgatttgatttttttgtgttaaacttaatttttttttgatggtgtacgggaagcggtaaaaggtag
accaattagaagaaggcctttgtgctctttatagagccaagcggctgtcgcttataaaataaagcttcacactaattatcta
ttatctatttcctcccccctc* <mark style="background:red">*ctctgcctttgtctgctcca*</mark>*aacggtaattattgggtttttgaggcagccggctctgtctctacctg
agattattcgtaggtcgaaaacaacaatctagcaccgcgtttaatatttcatgagcactttgagttgaagaatgaagaga
ataatggaatttttgaggagaagtttgattaaaaatttaaatttcggatcc* <mark style="background:gray">*ttt*</mark>*cagaa* <mark style="background:yellow">**GCTGAAGTCATCCAAC
ACTCGACAGACGATCCGAATGGCACGCGAGAGGAATTCGACTATAATCAAATTG
AATATGGAAA**</mark>*gtgagttgtgaaatataatt* <mark style="background:red">*gggggagtctgaaacgtgaagt*</mark>*cttaaaaataaat*

Figure 49: Mutated sequence for restriction digest. The 5'–3' sequence including exon 5 (yellow/capital letters) and the surrounding intron region after homology repair is shown. This sequence includes a single new BAH MI restriction site (shown in gray) that can be used to identify the mutated sequence with PCR and restriction digest. PCR primers surrounding the sequence of interest are shown in red.

Repair template for guide # 1:

A second repair template was designed to accompany crRNA #1 (Figure 50). This

template is 86 bp long (~40 bp on each side of the edit), and ends with a guanine nucleotide at

the 3' end, as recommended by Paix and colleagues (2014). Potential hairpins, while not fully

eliminated, were minimized as much as possible (data not shown). While the PAM sequence is

not mutated, the protospacer region is modified by a total of 4 mutations, so as to avoid repeated

cleavage events in the same location. Exon-specific changes are reflected in Table 26 with their

respective codon bias. As with the previous template design, a BAM HI restriction site is built

into the template to allow for easy genotyping (Figure 51). After PCR and restriction digest, we

would expect to see a single 325 bp fragment in worms without a successful CRISPR edit, and

two fragments (191 and 134 bp, respectively) in worms with a successful CRISPR edit.

A

*tcgaaaattttcagaa***ATGGAGGTCATCCA**ACACTCGACAGACGATCCGA**ATGGCACG CGAGAGGAATTCGACTATAATCAAATTG**

B

*tcg**gatcc**tttcagaa***GCT**GAGGTCATCCA**ACACTC**T**ACT**GA**T**GATCC**TA**ATGGCACGC GAGAGGAATTCGACTATAATCAAATTG**

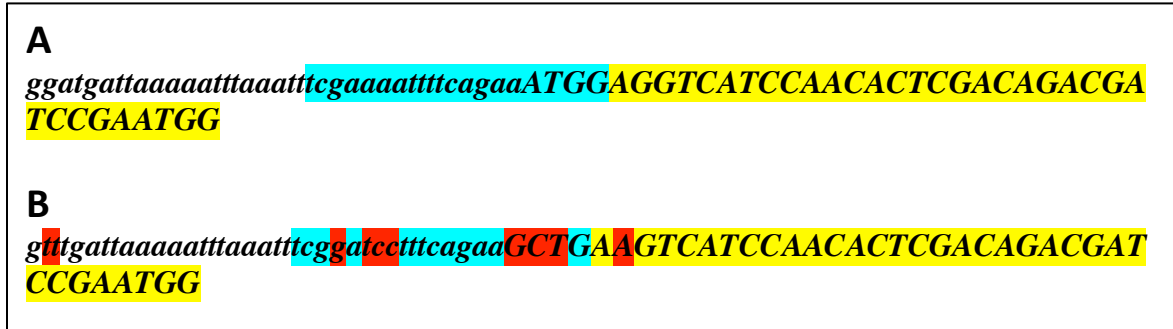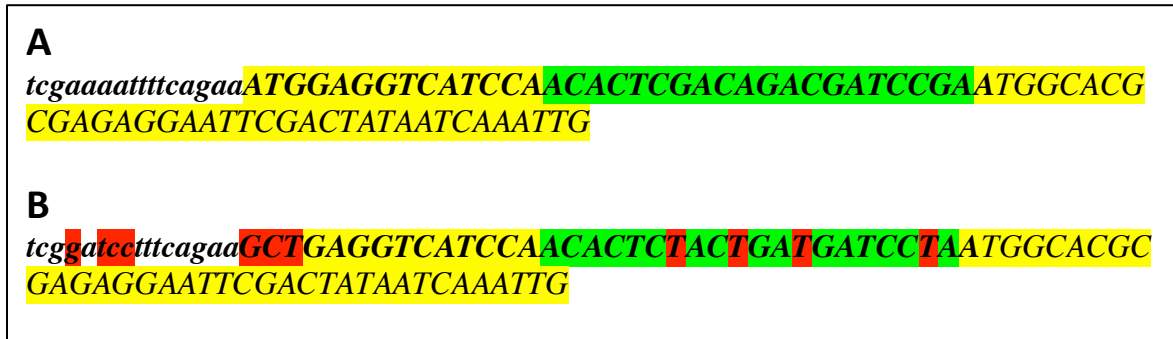**Figure 50: Proposed homology repair template for crRNA #1.** (A) The original genomic sequence covered by the repair template. crRNA is shown in green, and exon 5 is shown in yellow highlighting and capital letters. (B) Proposed homology repair template. crRNA is shown in green, and mutations are shown in red.

		Original codon	Change
Mutation #1	**Codon**	ATG	GCT
	Amino acid	Methionine	Alanine
	Codon bias	1.0	0.36
Mutation #2	**Codon**	TCG	TCT
	Amino acid	Serine	Serine
	Codon bias	0.15	0.21
Mutation #3	**Codon**	ACA	ACT
	Amino acid	Threonine	Threonine
	Codon bias	0.34	0.33
Mutation #4	**Codon**	GAC	GAT
	Amino acid	Aspartic Acid	Aspartic Acid
	Codon bias	0.33	0.68
Mutation #5	**Codon**	CCG	CCT
	Amino acid	Proline	Proline
	Codon bias	0.2	0.18

Table 26: Codon changes for repair template #1. Codon biases were calculated using the Codon Usage Database at *http://www.kazusa.or.jp/codon*.

*ccgcttatctgtctttcctctgtgatttgattttttgtgttaaacttaattttttttgatggtgtacgggaagcggtaaaaggtagac caattagaagaaggcctttgtgctctttatagagccaagcggctgtcgcttataaaataaagcttcacactaattatctattatc tatttcctcccccctc**ctctgcctttgtctgctcca**aacggtaattattgggttttgaggcagccggctctgtctctacctgagattat tcgtaggtcgaaaacaacaatctagcaccgcgtttaatatttcatgagcactttgagttgaagaatgaagagaataatgga attttgaggagaagtttgattaaaaaatttaaattt**tcggatcctttcagaa**GCTGAGGTCATCCAACACTCTA CTGATGATCCTAATGGCACGCGAGAGGAATTCGACTATAATCAAATTGAATATGG AAAgtgagttgtgaaatataatt**ggggagtctgaaacgtgaagt**cttaaaaataaat*

Figure 51: Mutated sequence for restriction digest. The 5' – 3' sequence including exon 5 (yellow/capital letters) and the surrounding intron region after homology repair is shown. This sequence includes a single new BAM HI restriction site (shown in gray) that can be used to identify the mutated sequence with PCR and restriction digest. PCR primers surrounding the sequence of interest are shown in red.

CRISPR Approach #2

A second CRISPR mutagenesis approach was designed to knock out *daf-19b* and *daf-19c* by introducing a 1 bp frameshift mutation starting in exon five. This is corrected for in *daf-19a* with an additional 1 bp frameshift mutation in exon 3, which adds back the deleted base and cancels out the first mutation (see Figure 44). As DAF-19B contains an additional exon (exon 4) between these two sites, the mutation in exon 3 will result in a frameshift that affects the integrity of exon 4, thereby likely causing a null mutation for this isoform. The sequences below show the specific reading frames and effects of these two mutations in DAF-19A/B/C, respectively.

1. The WT open reading frame for the end of exon 3 and the beginning of 5 (*daf-19a*) is shown below (exon 3 in yellow, exon 5 in blue):

==GAT GGT ACC GTG GGA GAT G====AA ATG GAG GTC ATC CAA CAC TCG ACA GAC==

2. Worms with both successful mutations will add an adenine nucleotide to the end of exon 3, and remove it from the beginning of exon 5; thus, the open reading frame will read as follows:

==GAT GGT ACC GTG GGA GAT GA====A ATG GAG GTC ATC CAA CAC TCG ACA GAC==

3. The WT isoform B open reading frame spanning exons 3 (yellow) and 4 (purple) reads as follows:

==GAT GGT ACC GTG GGA GAT G====CG TCA GTG ATG TTA GAT CCT ACA AAG ATA==

4. The exon 3 mutation will change this reading frame to the following sequence, thus changing the amino acids in exon 4 and likely rendering the protein nonfunctional:

GAT GGT ACC GTG GGA GAT GAC GTC AGT GAT GTT AGA TCC TAC AAA GAT

5. The WT open reading frame for the beginning of exon 5 (from isoform C, shown in green) is as follows:

ATG GAG GTC ATC CAA CAC TCG ACA GAC GAT CCG AAT GGC ACG CGA GAG

6. The exon 5 mutation would change this reading frame to the following sequence, thus changing all future amino acids in the isoform and rendering the protein product nonfunctional. Additionally, it is unlikely that the protein would be translated at all, given the lack of a start codon.

TGG AGG TCA TCC AAC ACT CGA CAG ACG ATC CGA ATG GCA CGC GAG AGG

When designing this set of mutations, splice sites were taken into account in order to avoid producing unwanted effects due to splice errors. According to Blumenthal and Meyer (1997), *C. elegans* introns obey the canonical GU-AG rule during splicing. This is to say, the 3' end of the intron will end in an AG dinucleotide, and the 5' end of the intron will begin with a GU dinucleotide. Additionally, *C. elegans* has a highly conserved extended consensus sequence at the 3' splice site (UUUUCAG). Figure 52 shows each exon with its surrounding splice sequences.

Figure 52: Exons 3 and 5 in WT *daf-19A* and *daf-19C*. Exons are shown in yellow/capital letters, and splice sequences are shown in red letters. The ATG start codon of *daf-19c* exon 5 is shown in bold. Two nucleotides at the beginning of exon 5 that are spliced differently in *daf-19a* and *daf-19c* are shown in blue.

It should be noted that there are two extra bases included in exon 5 in the *daf-19a* transcripts; thus, the splice site for this isoform is two bases earlier than in exon 5. While this conforms with the conserved 3' splice sequence, the splice site in *daf-19c* is non-canonical. As such, it is unclear whether splicing occurs immediately prior to the ATG start codon in exon 5 for all *daf-19c* transcripts, or whether the actual splice event occurs in the same place for both isoforms, and the ATG codon is simply identified by the ribosome at the start of translation, thereby determining the reading frame. Introns with 3' AA splice sites (such as the one adjacent to exon 5 in *daf-19c*) have been reported (e.g. Aroian et al., 1993), which would suggest that the transcript from WormBase is accurate as written. However, Zhang and Blumenthal (1996) found evidence suggesting that the UUUC portion of the 3' splice site may be particularly important for splice site identification; this could mean that this AA dinucleotide is present in *daf-19c* mRNAs, and is simply skipped by the ribosome in favor of the adjacent start codon. We proceed with the

assumption that the splice sites listed on WormBase are correct and will not be compromised by the addition or deletion of a base directly adjacent to them; however, this potential discrepancy is an important consideration.

crRNA Design

For the mutation adjacent to exon 5, it is possible to use the same crRNA that was created in the previous mutagenesis approach, as the mutation will occur in approximately the same location (see CRISPR approach #1). For the mutation adjacent to exon 3, the same approach detailed in the previous CRISPR design was used to generate an appropriate crRNA. A sequence of ~250 bp immediately surrounding the end of exon 3 was analyzed using the *crispr.mit.edu* design tool. 17 potential guides were reported, and the six best sequences were considered for predicted accuracy (Figure 53; Table 27).

A

	score	sequence	
Guide #1	99	AGTCAAGAATCTGCGCGAGC	TGG
Guide #2	99	TAACCTAAATGATGGTACCG	TGG
Guide #3	98	CAATGTTCTAACCTAAATGA	TGG
Guide #4	98	AACCTAAATGATGGTACCGT	GGG
Guide #5	98	AATGATGGTACCGTGGGAGA	TGG
Guide #6	98	TTTCTGGCAATGCCAACAGG	TGG

B

GCTCAAGAAATCATTGCCAGTCTCATTTCAATGTTCTAACCTAAATGATGGTACCGTGGGAGATGgtgagtttcaatcaaccacctgttggcattgccagaaatgatgaggccagctcgcgcagattcttgactaattagttttcgacgcagtatacccgtaaaatttaaaagaaaataaatatgagataacggacggatagataatcaaag

Figure 53: Sequence and location of potential crRNAs. (A) All crRNAs were designed by the *crispr.mit.edu* tool, with scores reflecting the predicted relative accuracy of each guide. Sequences are shown 5' – 3', with adjacent PAM sequences in green letters. According to the program, each of the first six guides is considered high quality. (B) Complementary protospacer sequences of the guides are shown with exon 5 and the immediately surrounding introns. Exon 5 is shown in capital letters/yellow highlights. Guide #1 is henceforth shown in green, guide #2 is shown in pink, and guide #6 is shown in blue (guides 3/4/5 not labeled).

Guide	Guide sequence	PAM	Protospacer (coding strand, 5' – 3')	Off-target sites	Off-target sites in genes
1	AGUCAAGAAUCUGC GCGAGC	TGG	GCTCGCGCAGATT CTTGACT	4	4
2	UAACCUAAAUGAUG GUACCG	TGG	TAACCTAAATGAT GGTACCG	3	2
3	CAAUGUUCUAACCU AAAUGA	TGG	CAATGTTGTAACC TAAATGA	9	7
4	AACCUAAAUGAUGG UACCGU	GGG	AACCTAAATGATG GTACCGT	8	5
5	AAUGAUGGUACCGU GGGAGA	TGG	AATGATGGTACCG TGGGAGA	6	3
6	UUUCUGGCAAUGCC AACAGG	TGG	CCTGTTGGCATTG CCAGAAA	8	6

Table 27: Potential crRNA sequences. The sequence of each crRNA sequence is shown with the corresponding coding strand sequence and PAM sequence, as well as the number of predicted off-target sites.

We recommend using crRNA #6, as it has a GG dinucleotide directly adjacent to the PAM sequence, has 50% GC content, is within 20 bp of the desired mutation, and has a low number of off-target sites, all of which have at least 4 mismatches (see criteria described in CRISPR approach #1).

Homology repair template for mutation #1 (adjacent to exon 5):

A repair template was designed to accompany the crRNA designed in the previous CRISPR approach, which can be used here for the exon 5 mutation (Figure 54). This template is 74 bp long (~37 bp on either side of the edit), and ends with a guanine nucleotide on either end, as recommended by Paix and colleagues (2015). A single base was deleted between the ATG start codon and the end of the 3' splice site. Our hope is that splicing will still occur in the same location, although it is possible that the splice event may one or two bases upstream, and that one or two adenine nucleotides will be included adjacent to the TGG motif; if this occurs, then the start codon may remain intact, and the mutation will fail to produce a frameshift.

An additional mutation was made to the PAM sequence in order to prevent repeated Cas9 binding. In the *daf-19a* double mutants, this will result in a silent mutation to a glutamic acid codon (Table 28). Furthermore, a BAM HI restriction site was built into the intron portion of the protospacer, both to disrupt binding and to create a simple marker for screening with restriction digest. The same primers from the first CRISPR design can be used to screen for this template.

Figure 54: Proposed homology repair template for crRNA (exon 5 mutation). (A) The original genomic sequence covered by the repair template. crRNA is shown in blue, and exon 5 is shown in yellow highlighting and capital letters. The splice consensus sequence preceding the exon is shown in red letters (B) Proposed homology repair template. crRNA is shown in blue, and mutations are shown in red highlights. The mutation of significance is a single base deletion at the beginning of exon 5. The splice consensus sequence preceding the exon is shown in red letters.

		Original codon	**Change**
Mutation #1	Codon	ATG	TG
	Amino acid	Methionine	Frameshift
Mutation #2	Codon	GAG	GAA
	Amino acid	Glutamic Acid	Glutamic Acid
	Codon bias	0.38	0.62

Table 28: Codon changes for the exon 5 repair template. Codon biases were calculated using the Codon Usage Database at *http://www.kazusa.or.jp/codon*.

Homology repair template for mutation #2 (adjacent to exon 3):

An additional repair template was designed to accompany crRNA #6 (Figure 55). This template is 75 bp long (~37 bp on either side of the edit). A single base (adenine) was added between the end of exon 3 and the adjacent splice site. Additional mutations were made to the PAM sequence and protospacer within the intron; these serve to disrupt repeated binding and to

create a BAM HI restriction site for screening (Figure 56). After PCR and restriction digest, we

would expect to see a single 489 bp fragment in worms without a successful CRISPR edit, and

two fragments (336 and 153 bp, respectively) in worms with a successful CRISPR edit.

A

AATGTTCTAACCTAAATGATGGTACCGTGGGAGATGgtgagtttcaatcaaccacctgttgg cattgccagaaa

B

AATGTTCTAACCTAAATGATGGTACCGTGGGAGATGAgtgagtttcaatcaacggatccttg gcattgccagaaa
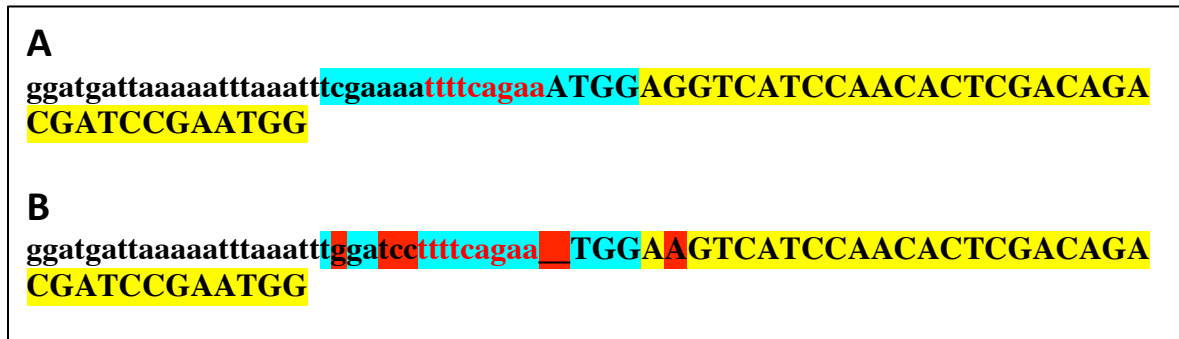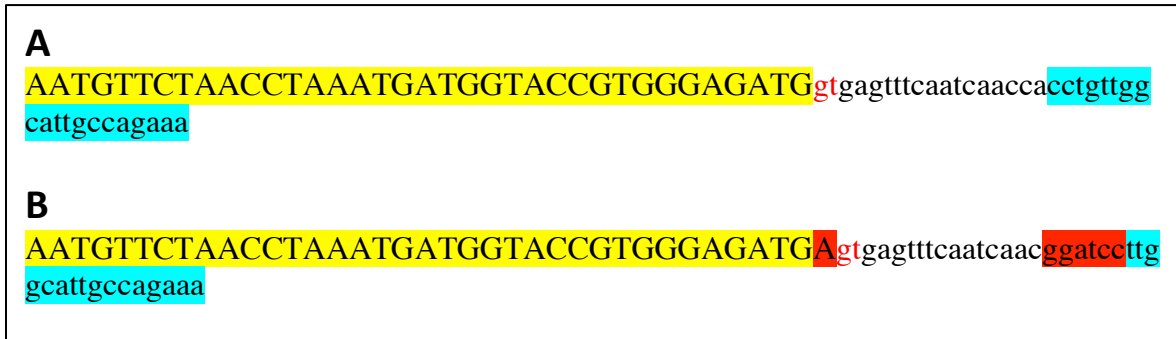
Figure 55: Proposed homology repair template for crRNA (exon 3 mutation). (A) The original genomic sequence covered by the repair template. crRNA is shown in blue, and exon 5 is shown in yellow highlighting and capital letters. The splice consensus sequence following the exon is shown in red letters (B) Proposed homology repair template. crRNA is shown in blue, and mutations are shown in red highlights. The mutation of significance is a single base addition at the beginning of exon 3. The splice consensus sequence following the exon is shown in red letters.
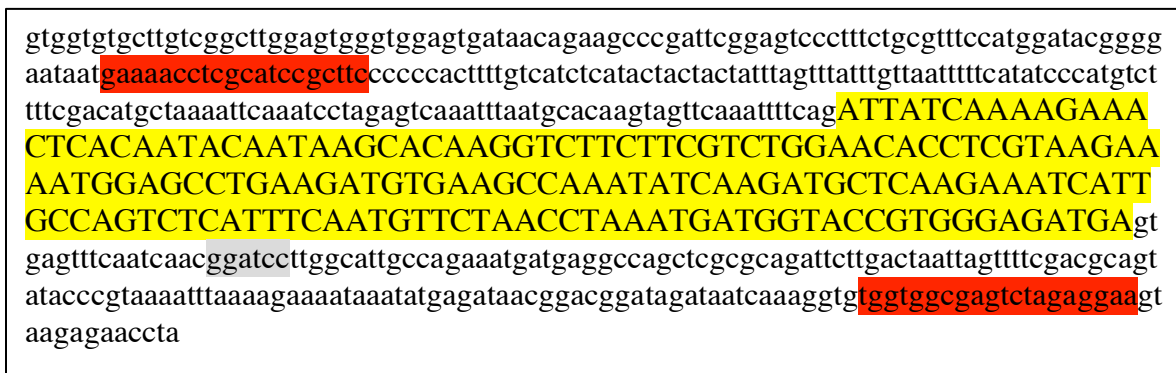
gtggtgtgcttgtcggcttggagtgggtggagtgataacagaagcccgattcggagtcccttctgcgtttccatggatacgggg aataatgaaaacctcgcatccgcttccccccactttttgtcatctcatactactactatttagtttatttgttaattttcatatcccatgtct tttcgacatgctaaaattcaaatcctagagtcaaatttaatgcacaagtagttcaaattttcagATTATCAAAAGAAA CTCACAATACAATAAGCACAAGGTCTTCTTCGTCTGGAACACCTCGTAAGAA AATGGAGCCTGAAGATGTGAAGCCAAATATCAAGATGCTCAAGAAATCATT GCCAGTCTCATTTCAATGTTCTAACCTAAATGATGGTACCGTGGGAGATGAgt gagtttcaatcaacggatccttggcattgccagaaatgatgaggccagctcgcgcagattcttgactaattagttttcgacgcagt atacccgtaaaatttaaaagaaaataaatatgagataacggacggatagataatcaaaggtgtggtggcgagtctagaggaagt aagagaaccta

Figure 56: Mutated sequence for restriction digest. The 5' – 3' sequence including exon 3 (yellow/capital letters) and the surrounding intron region after homology repair is shown. This sequence includes a single new BAM HI restriction site (shown in gray) that can be used to identify the mutated sequence with PCR and restriction digest. PCR primers surrounding the sequence of interest are shown in red.

In summary, this design requires two simultaneous CRISPR mutagenesis events that will

create two frameshift mutations, in which the first adenine nucleotide in exon 5 is deleted from

its present location and added to the end of exon 3. This will allow transcripts of *daf-19a* to

retain the same amino acid sequence as in WT worms, while resulting in nonfunctional DAF-19C, and likely DAF-19B as well. The obvious disadvantage to this approach is that it requires two simultaneous edits; however, work by Paix et al. (2015) and others has shown that effective crRNAs may be efficient enough to produce the desired double mutagenesis, particularly when a co-CRISPR approach is used (Arribere et al., 2014). Additionally, this frameshift mutation will almost certainly knock out DAF-19C, assuming that the splice sites listed on *wormbase.org* for *daf-19c* are accurate. With CRISPR approach #1, it is possible that replacing the first start codon with another amino acid will be insufficient to knock out the protein, as there are other in-frame methionine codons in close proximity that could potentially be used as alternate start sites. While this would result in a somewhat truncated protein product, there is no guarantee that it will results in a null mutant. Thus, if the two crRNAs created for CRISPR approach #2 prove to be efficient, the double frameshift approach may be more likely to produce the desired phenotype.

Cas9 purification optimization

Although the previously described CRISPR designs will be implemented this summer by Debora Sugiaman, we had originally planned to complete these experiments at Lawrence. As such, we optimized a Cas9 purification protocol described by Paix and colleagues (2015) to produce the Cas9 protein necessary for CRISPR mutagenesis (Appendix B). With a few modifications, their procedure proved to be efficient and relatively easy to implement.

Members of the Seydoux lab generously provided us with a strain of *E. coli* that had been transfected with a plasmid (nm2973) that contained the coding sequence for Cas9 protein, a T7 viral promoter just upstream of the coding sequence, and a carbenicillin resistance gene. These bacteria can be grown in media containing carbenicillin (an antibiotic), to select for only those cells that contained the desired plasmid. The T7 promoter allows selective induction of the gene

in the presence of IPTG, causing the bacteria to produce an overabundance of Cas9 protein

(Figure 57). After allowing the bacteria to proliferate in carbenicillin/luria broth media to the

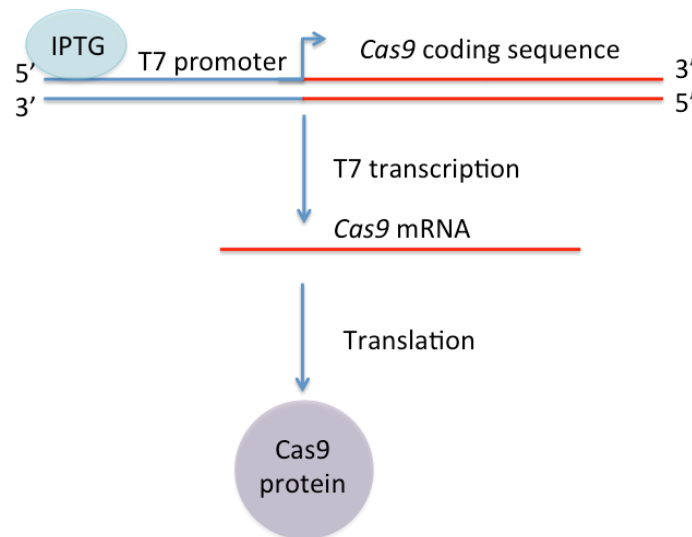appropriate population density, IPTG was added and the bacteria were allowed to grow

overnight.

Figure 57: Cas9 induction by T7 promoter. IPTG binds to the T7 promoter to induce transcription of the *Cas9* coding sequence that follows, resulting in overproduction of the protein.

After allowing the bacteria to produce Cas9, the culture was pelleted and resuspended,

and sonication in the presence of protease inhibitors was used to lyse the cells. The cleared

lysate was batch bound with Ni-agarose beads for 25 minutes. While Paix and colleagues

suggest batch binding for 45 minutes, preliminary experiments indicated that an excess of other

proteins bound to the beads, resulting in a more laborious purification procedure. As such, we

recommend shortening the time spent on this step.

Following the batch binding step, Ni affinity chromatography was used to isolate the

Cas9 protein. Ni affinity chromatography works by binding proteins with histidine tags to nickel

beads while allowing other contaminants to run through the column. The Cas9 coding sequence

in our plasmid contains a 6X histidine tag; thus, when this Cas9 comes in contact with the beads, it sticks to the column while other proteins are washed away with excess buffer (Figure 58). We found that it was unnecessary to wash the beads with 100 mL of buffer during this step, as recommended by Paix and colleagues; rather, half this volume was sufficient. The Cas9 protein was then eluted in 1 mL fractions with a buffer containing a high concentration of imidazole. Imidazole serves to outcompete the 6X histidine tag on the Cas9 protein, thus causing it to come off the beads and flow through the column into the eluent. We found that this step worked best when we doubled the concentration of imidazole suggested by Paix et al., as it caused the Cas9 protein to come off the beads all at once within a few fractions, rather than trickling off slowly. Individual fractions were then tested for the presence of the protein using SDS-page, which confirmed the presence of a protein that was the correct size (Figure 59).

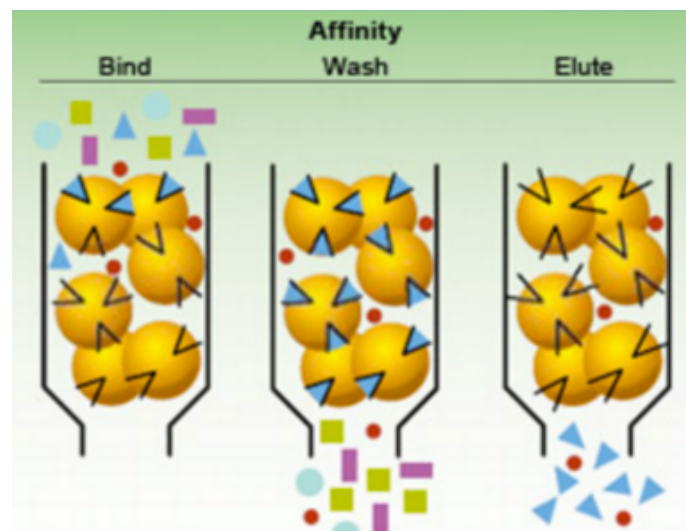**Figure 58: Ni-affinity chromatography.** The target protein (blue arrows) binds preferentially to the Ni beads, shown in yellow. Other proteins flow through the column during the wash step. The target protein is then eluted using a buffer containing imidazole, which is better able to bind the Ni beads and thus outcompetes it. Figure adapted from *bio-rad.com/en-us/applications-technologies/introduction-affinity-chromatography*.
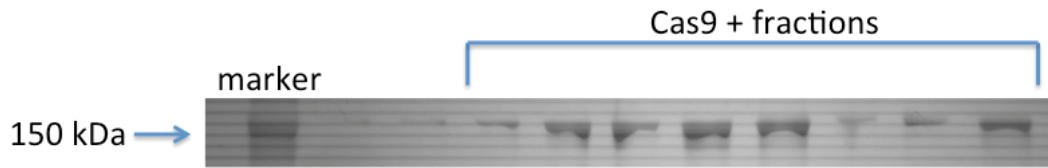
Figure 59: Cas9-containing fractions after elution. SDS-page was used to show the presence of a protein ~160 kDa in size, thus demonstrating the presence of Cas9.

After elution, fractions containing Cas9 were pooled and run over a Q Sepharose column to remove contaminating nucleic acids. Q Sepharose beads are positively charged, and thus bind to the negative phosphate backbone characteristic of DNA and RNA. As Cas9 protein is positively charged, it flows directly through the column. Following this step, the protein was dialyzed using a Pierce Cassette (Figure 60). These cassettes are comprised of a membrane-bound sac that is permeated with pores to allow material smaller than 100 kDa to pass through. The eluent was injected into the sac and allowed to dialyze overnight in buffer, then transferred to a second cassette and allowed to dialyze again. During dialysis, fluids and small particles pass through the membrane by osmosis until each side of the membrane contains an isotonic solution. The two-step dialysis process allows the vast majority of small contaminants (e.g. partially degraded proteins, etc.) to diffuse out of the Cas9-containing eluent, which is then removed with a syringe.
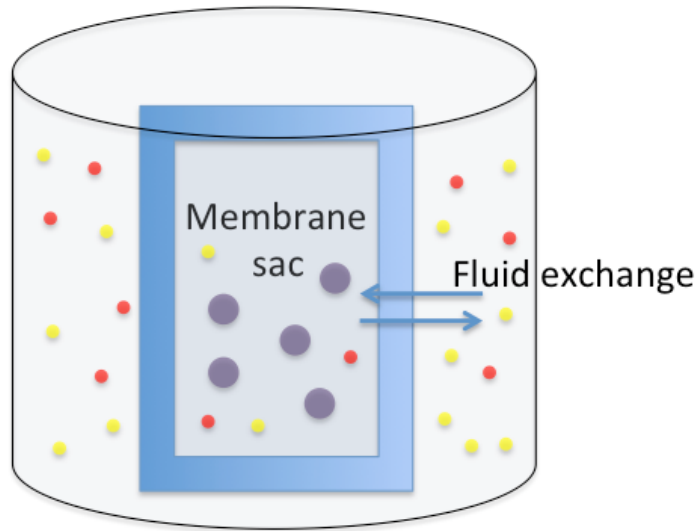
Figure 60: Dialysis by Pierce Cassette. The cassette membrane is permeated with pores that allow material smaller than 100 kDa in size to pass through (shown in red and yellow circles). The Cas9 protein (purple circles), at 160 kDa, is too large to flow through.

As a final step, the Cas9 protein was concentrated to ~10 mg/ml and frozen in 5 μl aliquots at -80ºC. Our final run of the protocol yielded 37 aliquots of 5 μl at a concentration of 9.84 mg/ml. As such, we are able to confirm the efficacy of the procedure outlined in Paix and colleagues (2015). Purification is considerably more cost-effective than using commercial Cas9, which can cost hundreds of dollars for a single injection; as such, this procedure may prove useful for future CRISPR experiments done at Lawrence.

DISCUSSION

The primary aim of this study was to identify new gene targets of DAF-19, with a particular focus on genes that are involved in neuronal processes. We wished to go beyond previous studies that have focused on the role of DAF-19 in ciliogenesis to explore targets that might play a role in synaptic maintenance. Additionally, we were interested in identifying which isoforms of DAF-19 control each of these targets. This focus stemmed from work done by Senti and Swoboda (2008), which found a novel indirect role for DAF-19 in regulating synaptic proteins.

In order to identify putative targets, Elizabeth De Stasio and Prasad Phirke performed three microarray analyses of adult, L1 larvae, and 3-fold stage embryos in *daf-19(WT)* and *daf-19(m86)* backgrounds. Previous studies aiming to identify DAF-19 target genes focused on the identification of x-box containing genes. Our study used a less biased approach, beginning with a comparative transcriptome analysis. This resulted in a list of 177 differentially expressed genes, of which 30 were selected from this list for further analysis. Gene selection was based on several factors, such as the degree to which the gene was differentially expressed (< 0.5 or >1.5 fold change), whether it had any known neuronal functions, the presence of an x-box motif in its promoter region, and homology to human genes. Transcriptional fusions of each gene's control region with GFP were made, so as to explore each gene's expression pattern. The present study sought to analyze five of these 30 genes for DAF-19 dependence.

Each transcriptional fusion was carefully analyzed for GFP expression using confocal microscopy. Isogenic strains were studied first in WT and *m86* backgrounds to confirm the presence of differential GFP expression before determining isoform-specific mechanisms of control. Once these expression patterns had been characterized, a series of new isogenic strains

in different *daf-19* genetic backgrounds were generated, including *tm5562* and *of5*, (both of which would affect DAF-19A/B production) and *of6* (affecting DAF-19B). Additionally, rescue experiments using *daf-19a* and/or *daf-19c* cDNAs were completed to better elucidate the effects of these respective isoforms on gene regulation. Of the five genes studied, one (T01B11.2) is clearly regulated by DAF-19C, while two additional genes (T07F10.1 and *del-4*) show some DAF-19 dependence; however, we conclude that this is not the sole regulating factor for either gene. A fourth gene, *srd-61,* showed expression patterns independent of DAF-19, while a final gene, *decr-1.1* showed no definitive expression pattern in either genetic background.

Confidence in reporter expression and DAF-19 dependence

While transgenic reporters are efficient and useful tools for determining gene expression patterns, there are various limitations and potential confounding factors that should be considered when interpreting the data reported here. Firstly, the transgenes themselves are not genomic DNA; rather, each constitutes an extrachromosomal plasmid array constructed in the worm gonad following microinjection. The level of transgene expression depends on the number of plasmid copies that assemble into a transgene array in the worm, and the degree to which they are stably inherited. As such, expression levels and plasmid integrity for each gene may vary from injection to injection. Stability issues with certain transgenes were quite apparent, as some were inherited much more readily than others from one generation of worms to the next. To minimize the problem of mosaicism, we created isogenic strains in the desired *daf-19* backgrounds by mating hermaphrodite worms from a single transgenic line with male *daf-*19 mutants; thus, the same transgenic array in all strains was inherited from a single microinjected worm. This served to normalize expression across *daf-19* phenotypes, although the results are not necessarily representative of the strain's "true" expression pattern.

In addition to the question of isogenicity, it should be noted that other factors might affect control of transgenic expression. For instance, transcriptional fusions were made based on the assumption that each gene's control region was contained within a region 2kb upstream of that gene's start codon, and that any regulation by DAF-19 would affect this region only. This is generally considered the standard for identifying control regions near a coding sequence; however, distal sequences that bind enhancer or repressor elements have been found far upstream or downstream of start codons for many genes (Wenick & Hobert, 2004). It is possible that some of the genes studied might be regulated by *daf-19* via sequences that were not contained within the transcriptional fusions, thus causing us to underreport the regulatory effects of DAF-19. Furthermore, plasmid-based expression patterns do not reflect the epigenetic conditions in which genes are regulated *in vivo*. Thus, it cannot be assumed that these experiments reflect all aspects of gene regulation.

It should be noted that all strains used, whether designated WT or mutant for *daf-19*, were studied in a *daf-12-/-*, and sometimes *him-5-/-* genetic background. *daf-19(m86)* mutants are dauer constitutive, meaning the vast majority of worms will enter the dauer developmental pathway as L2 larvae rather than maturing to adulthood. As this would prevent us from completing the vast majority of experiments, it was necessary to study worms in a *daf-12-/-* background, which prevents worms from entering the dauer pathway. This was standardized across all strains in all backgrounds for *daf-19*. Additionally, many strains were *him-5-/-*, a mutation that produces an abundance of males; these were used to complete matings. For specific genotypes of each strain, see Table 1 (methods). While neither of these mutations is known to affect *daf-19* behavior or any of the target genes studied, it remains a possibility that they may have skewed our results.

Finally, it should be noted that without the presence of a known co-localizing marker, all tissue identification relied on subjective judgment. We used representative images from *wormatlas.org* that included neuronal schematics and confocal images as anatomical references. While some tissues (e.g. intestines and pharyngeal muscle) were very easy to identify, others were considerably less clear. The neuronal cell bodies in the isthmus region of the pharynx proved particularly challenging, as many of these cells with similar morphology are located in a compact space, and are easily mistaken for one another. Furthermore, neuronal positions are not absolute, as cell bodies can migrate to slightly different locations during development. GFP expression levels also varied considerably due to mosaicism, such that it was at times difficult to determine how many neurons were fluorescing in a particular image.

Expression pattern of T01B11.2

Many tissues expressing T01B11.2::GFP could be confidently identified in both *daf-19(WT)* and *daf-19(m86)* worms; these included the intestine, pharyngeal muscle, hypodermis, muscle in the body wall and arms to the nerve ring, and rectal glands. Additionally, clear differential expression in two to four neurons in the isthmus region of the pharynx was visible when comparing expression of the same transgene array in *daf-19(WT)* and *daf-19(m86)* strains. However, identification of these neurons these proved quite challenging, in part because the level of GFP expression in other tissues often obscured these neurons. It was not always clear how many of these neurons were fluorescing, as they varied in intensity and were often obscured by the much brighter hypodermis; however, it seems most likely that the four cell bodies observed constitute two neuron pairs.

Once differential expression was identified in *daf-19(WT)* and *daf-19(m86)* backgrounds, we created a series of isogenic transgenic strains in the *daf-19(tm5562)*, *daf-19(of5)*, and *daf-*

19(of6) backgrounds. The first two strains showed a phenotype resembling *daf-19(WT)* worms, indicating that the DAF-19 isoforms affected by each of these alleles were not responsible for the differential expression of T01B11.2::GFP. As *tm5562* and *of5* both affect DAF-19A/B, this suggests that T01B11.2 is regulated specifically by DAF-19C. Additionally, recent data from a rescue experiment adding a *daf-19c* cDNA construct to worms with an *m86* background showed a WT phenotype. It should be noted that only eight worms from this strain have been imaged so far; however, the preliminary data are consistent with findings from earlier strains. It was not possible to repeat the same experiment with *daf-19a* cDNAs due to the overabundance of fluorescence in non-neuronal tissue, which obscured the green transgenic marker used to identify *daf-19a::daf-19(m86)* worms. However, this experiment is not critically important, given the considerable evidence indicating DAF-19C dependence.

In spite of the apparent clarity of these findings, data from the *daf-19(of6)* strain yielded highly unexpected results. Of the worms imaged, 25% showed a WT phenotype, 50% showed an *m86* phenotype, and 25% showed a variety of novel expression patterns. These included a complete lack of GFP expression in tissues such as the hypodermis and intestines (which nearly always expressed the transgene in other strains), and significant up-regulation of the transgene in other isthmus neurons. While only 24 worms from this strain were imaged, the degree of variability observed was striking. It is possible that this indicates a role for DAF-19B in the expression of T01B11.2, as this is the only isoform affected by the *of6* mutation. However, if this were truly the case, we would expect to see a more clear-cut WT or *m86* phenotype, rather than a mix of the two. A second possibility is that we have uncovered a previously unknown control region in exon 4, which is disrupted by the mutation. This could produce highly variable *daf-19* transcription, thus affecting its downstream targets and producing the various expression

patterns. More worms should be imaged in order to clarify these phenotypes, and it would be useful to compare them with *of6* expression patterns observed in other DAF-19 dependent genes, to see whether the unusual degree of phenotypic variability is restricted to this strain.

Identification of differentially expressed neurons

A dye-filling assay confirmed that GFP did not co-localize with any of the dye-filling amphid neurons, thus eliminating these six candidates as the site of T01B11.2::GFP expression; however, the neurons of interest frequently appeared either directly anterior or posterior to the three dorsal-most dye filling neurons (ASK, ADL, and ASI). This narrowed the range of possible neurons to 10-15 nearby cells. As initial images suggested that the neurons in question did not connect to the nerve ring, URX was considered to be a strong candidate, given its location just anterior to ASK and its distinctive processes; however, this possibility was eliminated through use of a double transgenic strain containing an mCherry marker in the URX neuron, which did not co-localize with GFP. Additionally, one image collected from the *daf-19(of6)* strain, in which the transgene had been lost or silenced in most non-neuronal tissues, allowed us to more clearly visualize the morphology of the two pairs of neurons. Interestingly, this image showed both pairs connecting to the nerve ring, a fact obscured by GFP expressed in the body wall muscle arms present in this region of most other T01B11.2::GFP worms. Additionally, as evidence indicated that T01B11.2 is very likely to be controlled by *daf-19c*, we realized that the cells of interest must be ciliated sensory neurons, as these alone express the *daf-19c* isoform (Senti & Swoboda, 2008). This allowed us to narrow the field to 14 ciliated sensory neurons in the mid isthmus region. The dendrite morphology of these candidates was assessed, and neurons with processes that did not match those seen in our images were eliminated. Finally, a second marker strain was generated by crossing worms expressing an mCherry marker

in all cholinergic neurons, and we were able to confirm that the neurons of interest were not cholinergic, as GFP did not co-localize with mCherry. This left us with three possible neuron pairs: ASG, AWA, and CEP. Of these, ASG and AWA were considered to be the best choices, as their positions relative to the dye-filling amphid neurons were closest to those observed for the neurons of interest.

Known functions of the ASG and AWA neurons

The glutamatergic ASG neuron pair consists of two sensory ciliated neurons, which play a role in chemosensation and taste (Pocock & Hobert, 2010). Additionally, they have been shown to play a role in lifespan regulation, as worms lacking functional ASG neurons live longer (Alcedo & Kenyon, 2004). The AWA neurons are also ciliated; these play an odorsensory role (Bargmann et al., 1993). Additionally, AWA is one of three key pairs of sensory neurons needed for sexual attraction in males (White & Jorgensen, 2012). Many of these functions (chemosensation, olfaction, taste) rely on the presence of cilia; as such, *daf-19* dependent expression in these neurons is consistent with the protein's known roles in ciliogenesis.

Hypothetical roles for T01B11.2 in a DAF-19 mediated gene pathway

T01B11.2 codes for a protein that is not known to be involved in ciliogenesis or cilia maintenance. Very little is known about its function, although is it thought to have transaminase activity. Such activity is critical for amino acid metabolism, which occurs in cells throughout the body; we therefore find no clear reason why T01B11.2 would be differentially expressed in these particular neurons. As T01B11.2 expression appears to be linked to DAF-19C, this result does not provide evidence to support Senti and Swoboda's hypothesis that DAF-19A/B regulate neuronal functions in non-ciliated neurons. It is possible that we have uncovered a novel role for

DAF-19C in regulating genes unrelated to either ciliogenesis or synapse maintenance; however, this is merely speculation. It may be the case that DAF-19 and T01B11.2 do not interact at all, and T01B11.2 is actually activated by an unknown downstream target of DAF-19C, which would also be differentially expressed in *daf-19(WT)* vs. *daf-19(m86)* backgrounds.

Expression pattern of T07F10.1

T07F10.1::GFP expression was confidently determined through comparison of representative confocal images with diagrams from *wormatlas.org.* Alex Hurlburt previously identified expression in the excretory system, pharynx, and several neurons in the head and tail, which he identified to be URX, SABD, PDA, PLM, and DVA/DVB. These identifications were confirmed through comparison with *wormatlas.org* resources, and URX expression in particular was definitively confirmed by generating a double transgenic strain that included an mCherry marker expressed in the URX neuron.

Between the *daf-19(WT)* and *daf-19(m86)* strains, GFP expression in the URX neuron decreased by 30 percentage points, while expression in the SABD neuron decreased by 49 percentage points. Although Alex Hurlburt was not convinced that these changes indicated true *daf-19* dependence, we elected to further study the gene in various *daf-19* backgrounds in order to see whether this gap in expression rates persisted. We found that both neurons expressed GFP more frequently in *daf-19(tm5562)* and *daf-19(of5)* backgrounds than in WT worms, indicating that the *daf-19a/b* isoforms were not responsible for activating expression of the transgene in these cells. Additionally, the *daf-19c* rescue strain showed GFP expression in 94% of the SABD and URX neurons. These preliminary data suggested the possibility that this gene is regulated by DAF-19C. However, the *daf-19a* rescue strain showed GFP expression in URX 68% of the time and SABD 72% of the time, a much higher proportion than was seen in the *daf-19(m86)* strain.

One interpretation of these data is that T07F10.1 is partially regulated by DAF-19C, but DAF-19A over-expression can also increase transgene expression. It is also possible that the differences we see in transgene expression in the *m86* and WT backgrounds are just variability in gene expression and not due to the genetic background.

It should be noted that due to time constraints, a relatively small number of worms was imaged for each strain, and it is unlikely that the proportions of GFP expression reported for each *daf-19* background are truly representative. In order to more confidently assess the possible regulatory effects of DAF-19 on T07F10.1, more data are needed to establish representative expression profiles. Additionally, we have not ruled out the possibility that DAF-19B is responsible for regulating T07F10.1 in these neurons. This will be addressed by creating a transgenic strain in a *daf-19(of6)* background, which is null for *daf-19b* only. If these worms express an *m86*-like phenotype (significantly decreased expression in URX and SABD), this would suggest that DAF-19B regulates expression of T07F10.1. It is also possible that this gene is regulated by both DAF-19 and some other unknown transcription factor, which could explain the variation in expression levels. However, at this time we are unable to conclusively state that T07F10.1 is solely DAF-19 dependent.

Expression pattern of srd-61

Of all the genes imaged, *srd-61* allowed for the most confident tissue identification. A single pair of neurons in the isthmus and two pairs of tail neurons fluoresced in 100% of worms imaged in both the *m86* strain and its isogenic WT partner. These were easy to identify, as all three neuron pairs co-localized with dye-filling neurons. Based on their position relative to the other amphid neurons, the cells observed in the head were determined to be the ASH pair. Additionally, two dye-filling phasmid neurons expressed GFP. As the tail has only two such

pairs of neurons, it was immediately apparent that the GFP-expressing neurons must be PHA and

PHB. As expression patterns were identical between WT and *m86* worms, we conclude that *srd-*

61 is not DAF-19 dependent. This is inconsistent with the microarray data reported for this gene,

which showed significant up-regulation in *daf-19(m86)* worms. However, a wide variety of

factors could have affected this result, including unknown mutations or epigenetic differences in

the particular worms used in the microarray. Ultimately, our results reveal how unreliable such

data are without more rigorous verification.

Additionally, our preliminary results from this gene show how important it is to use only

isogenic strains when comparison expression patterns. Initially, we imaged two lines of worms

which had been separately injected into either *daf-19(WT)* or *daf-19(m86)* backgrounds. While

these showed fluorescence in the same neuron, the frequency of expression was variable, and we

initially believed we had found a new DAF-19 target. However, when an isogenic strain for one

of these lines was produced, it immediately became clear that this was not the case, as ASH

fluoresced 100% of the time in both lines. This discrepancy is very likely to have arisen from

variable stability or inheritance of the transgenic plasmid, and highlights the fact that

transcriptional fusions are by nature subjective. As such, it is important to replicate initial results

in a second line.

Expression pattern of del-4

Expression patterns for *del-4* were highly mosaic, and could not be definitively

elucidated. In both *daf-19(WT)* worms and the isogenic *daf-19(m86)* strain, neuronal expression

was observed in up to eight neurons in the isthmus region of the pharynx, and up to four tail

neurons. Additionally, faint intestinal expression was observed in 60% of WT animals and 76%

of *m86* animals. It was frequently difficult to determine how many neuronal cell bodies were

fluorescing in an animal, as they exhibited highly variable levels of GFP expression. While these judgments were by nature subjective, efforts were made to control for this by repeatedly reanalyzing data for all images without reference to previous records, and averaging the perceived number of fluorescent neurons from separate days.

One of the most commonly visualized neurons included a pair in the mid-isthmus region of the pharynx that extended a long dendrite to the nose and connected to the nerve ring; additionally, GFP expression in this cell did not co-localize with any dye-filling neurons. This morphology resembled the ASE neuron pair, a hypothesis supported by Echtberger and colleagues, who previously identified *del-4* expression in this cell (2007). ASE is a ciliated sensory neuron that plays a role in taste sensation. As *daf-19* regulates ciliogenesis, it would be reasonable to expect its expression in this neuron. However, it should be noted that without a co-localizing marker, it is impossible to conclusively determine whether this identification is correct.

An additional pair of neurons, located just posterior to the base of the nerve ring, also frequently fluoresced in WT and *m86* backgrounds. GFP did not co-localize with any amphid neurons, and the cell's identity was impossible to definitively elucidate. The range of possibilities was reduced to five neurons with similar characteristics; these were AIA, AIB, RMDD, RMFL, and RMH. However, no further conclusions can be drawn at this time regarding the identity of these cells.

Intermittent GFP expression in up to four tail neurons was also observed in ~50% of WT and *m86* worms. These were hypothesized by Prasad Phirke to be PQR and PHA, based on preliminary fluorescence microscopy. PQR neurons extend a short ventral cord process and an additional phasmid process in opposite directions away from the cell body, and are located just

posterior to the anus. As one of the tail neurons frequently exhibited this distinctive morphology, it was identified as such with relative confidence. The other neurons resembled either PHA or PHB, two dye-filling phasmid neurons with a single process extending posteriorly. These should have been relatively easy to identify with a dye-filling assay; however, we failed to find any worms expressing GFP in these neurons on days when dye-filled animals were imaged. The experiment would need to be repeated to confirm co-localization; however, as the neurons do not exhibit DAF-19 dependent expression of GFP, this identification is not a priority.

There was no clear difference in expression between WT and *m86* strains with regards to which neurons fluoresce, and as such, the gene is unlikely to be DAF-19 dependent. However, it should be noted that the *daf-19(m86)* strain exhibited statistically significant age-dependent expression levels. L1/L2 worms showed GFP expression in an average of 3.58 isthmus neurons (standard deviation = 1.06), while each progressively older group showed less expression, with adults demonstrating fluorescence in an average of 1.67 neurons (standard deviation = 0.78). However, WT worms did not show the same trend. The relatively low numbers of worms observed in this strain (N=25) may have obscured this effect, and significantly more animals need to be imaged to draw further conclusions. However, the preliminary finding is quite interesting, as *del-4* defects have been shown to result in neurodegeneration. It is possible that this gene is misregulated in *daf-19(m86)* worms, resulting in a degenerative phenotype. While an absence of GFP fluorescence does not necessarily indicate neurodegeneration, there is a definite possibility that the change in fluorescence could indicative of neuronal damage. Interestingly, the microarray data actually found lower expression of *del-4* in L1 worms, indicating that the gene is down-regulated in young animals. This could indicate that DAF-19 ordinarily suppresses *del-4* expression in young worms, and that neurodegenerative phenotypes

result from its misregulation. We recently generated two new isogenic strains containing *del-4::GFP* in *daf-19(tm5562)* and *daf-19(of5)* backgrounds, which will be used to further investigate the possibility that *daf-19* mediates age-dependent expression of *del-4* in the isthmus neurons.

CRISPR design

One of our primary challenges in determining isoform-specific DAF-19 dependence stems from our lack of a *daf-19c* knockout mutant. Such a tool would prove extremely useful in determining whether particular genes are regulated by this isoform, as opposed to DAF-19A/B. We would expect such worms to demonstrate an *m86* phenotype if a target gene is regulated by DAF-19C, and a WT phenotype if the target gene is regulated by DAF-19A/B. However, as the sequence for *daf-19c* is entirely conserved in the two long-form isoforms, any mutation specific to *daf-19c* must avoid changing any of the shared exons in a manner that will affect expression of *daf-19a/b*. We designed two mutagenesis approaches to circumvent this problem. The first seeks to stall translation of DAF-19C by replacing its initial methionine with alanine, in the hopes that this will prevent the ribosome from recognizing a start site. While the change to alanine will produce a single nucleotide polymorphism (SNP) in *daf-19a/b*, we hope that alanine's small size and nonpolar quality will not significantly affect the structure of the protein.

The second mutagenesis approach relies on two frameshift mutations: one at the beginning of exon 5, and a second shift at the end of exon 3 to correct the first mutation. It should be noted that while this will entirely disrupt the *daf-19c* isoform, it is possible that translation could still occur in a different reading frame, beginning at a second methionine codon that occurs 52 residues into the protein. While a deletion of this size seems likely to still result in a nonfunctional protein, this cannot be determined for certain. *daf-19b* will also be disrupted, as

the exon 3 frameshift will produce a new stop codon at the end of exon 4, thus severely truncating the protein. Following mutagenesis, it will be necessary to confirm these phenotypes. The presence of functional DAF-19C could be determined with a simple dye-filling assay, as worms with a null mutation in this isoform will lack cilia and thus be unable to dye fill. Of course, the challenge here will be determining whether DAF-19A retains its function. As we currently have no assay to identify this, it will be necessary to generate and sequence cDNAs from spliced mRNA transcripts of this isoform. If the sequence from the mutant exactly matches the sequence of this transcript in WT worms, the protein should function normally. A similar approach could be used to characterize the DAF-19B protein, which we expect to be truncated at the end of the 4th exon.

Once these mutations are successfully created, it will be interesting to see what effects they have on genes we have previously identified to be DAF-19 dependent. In T01B11.2::GFP worms, we would expect these mutations to result in an *m86* phenotype, as evidence currently points toward DAF-19C mediated expression. This experiment could be done in place of the *daf-19a* rescue that we were unable to complete for this strain, due to its excessive GFP expression. We would not expect these mutations to have any effect on *srd-16::GFP* worms, as the transgene is not regulated by DAF-19. The T07F10.1 and *del-4* transgenes both showed inconclusive expression patterns, and it would be very interesting to see whether these mutations produce any effect. If T07F10.1 does in fact turn out to be regulated by DAF-19C, we would expect to see significantly decreased expression in the URX and SABD neurons. The effects on *del-4* cannot be determined until further data regarding a possible age-dependent phenotype are collected in *tm5562* and *of5* genetic backgrounds.

Future directions:

A variety of further experiments are needed to better understand the role that *daf-19* plays in regulating the five genes studied. As the initial strains containing the *decr-1.1::GFP* transgene showed no GFP expression in either *daf-19(WT)* or *daf-19(m86)* worms, we will begin by imaging a second pair of strains which were separately microinjected. If these show no fluorescence, it may be necessary to redesign the transgene, as the promoter region may not be fully captured. We would also like to generate strains containing the T07F10.1 and *del-4* reporters in an *of6* background, both to determine whether DAF-19B has any effect on expression and to look for any unusual fluorescence patterns (possibly indicative of a disrupted control region). Additionally, we wish to better identify neuronal expression with additional co-localization markers. We recently obtained a new strain that fluorescently labels the nuclei of cholinergic neurons using mCherry. This could prove to be a highly useful marker, either for co-localization with hypothesized neurons, or as a positional reference. If our hypothesized neurons do not show mCherry co-localization, this will allow us to rule out the cholinergic neurons as candidates, thus greatly reducing the number of possible cells.

In addition to better characterizing expression patterns and neuronal identification, it would be useful to approach the question of DAF-19 dependence with a new approach, such as chromatin immunoprecipitation (ChIP). This technique involves crosslinking a protein of interest (in this case DAF-19) with any DNA to which it binds, shearing the bound DNA into fragments, selecting for those bound to the protein via immunoprecipitation, and then sequencing the resultant DNA to determine which genes were bound (Collas, 2010). This experiment would confirm whether the target genes that we have identified are in fact targets of DAF-19, or whether they are actually regulated indirectly, that is, downstream of this transcription factor.

Additionally, it may reveal novel genes that interact with DAF-19, which were not identified by transcriptome analysis.

Finally, it would be interesting to explore the effects of null mutations in the genes found to be regulated by DAF-19. Many of these genes are minimally characterized, and further data regarding their null phenotypes could further elucidate whether they play a role in synaptic maintenance or other neuronal processes, and if so what this role entails. These data could be used to elucidate the *daf-19* gene pathway proposed by Senti and Swoboda (2008), which predicts that DAF-19 plays an indirect role in mediating neuronal function.

CONCLUSIONS

Transgenic GFP reporters for the genes T01B11.2, T07F10.1, *srd-61*, *del-4*, and *decr-1.1* were studied for *daf-19* dependent expression. Of these, T01B11.2 showed clear *daf-19* dependence in the ASG and AWA sensory neurons. T07F10.1 and *del-4* are partially regulated by *daf-19*, although we conclude that this is not the sole factor controlling their expression. T07F10.1 exhibited potentially differential expression in the URX and SABD neurons, and the *del-4* reporter showed age-dependent expression in ASE and several other neurons in the isthmus region of the pharynx in a *daf-19(m86)* background. *srd-61* expression is entirely independent of *daf-19,* and no conclusions can be drawn about the relationship between *daf-19* and *decr-1.1* due to a lack of *decr-1.1::GFP* expression in all imaged strains.

The T01B11.2 and T07F10.1 reporters were studied in various *daf-19* mutant strains in which the DAF-19A/B isoforms are presumed to be defective, and rescue experiments for these transgenes were completed using cDNAs for *daf-19a/c*. We can conclude that T01B11.2 is regulated by DAF-19C in particular, and it is likely that T07F10.1 is also partially regulated by this isoform. We also report two new CRISPR mutagenesis designs for a DAF-19C-specific knockout mutant, as well as the successful optimization of a Cas9 purification protocol provided by the Seydoux Lab. While the data provided here do not support Senti and Swoboda's hypothesis that DAF-19A/B play a role in regulating synaptic proteins, we can report DAF-19C dependence in at least one and possibly two novel genes which are not known to be related to ciliary function. Interestingly, both of these genes play roles in amino acid metabolism. While future work may link their functions to ciliary maintenance, it is also possible that we have uncovered an entirely new role for DAF-19C.

ACKNOWLEDGMENTS

REFERENCES

Aftab., S., Semenec, L., Chu, J.S., & Chen, Nansheng. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evolutionary Biology, 8,* 226.

Aroian, R., Levy, A., Koga, M., Oshima, Y., Kramer, J., & Sternberg, P. (1993). *Molecular and Cellular Biology, 13,* 626-637.

Arribere, J., Bell, R., Fu, B., Artiles, K., & Hartman, P. (2014). Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics, 198,* 837-846.

Azad, A., Rauh, R., Vermeulen, F., Jaspers, M., Korbmacher, J., Boisser, B.,...Cuppens, H. (2009). Mutations in the amiloride-sensitive epithelial sodium channel in patients with cystic-fibrosis-like disease. *Human Mutation, 30,* 1093-1103.

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,* 4417-4427.

Chen, N., Mah, A., Blacque, O., Chu, J., Phgora, K., Bakhoum, M.,...Stein, L. (2006). Identification of ciliary and ciliopathy genes in *Caenorhabditis elegans* through comparative genomics. *Genome Biology, 7.*

Collas, P. (2010). The Current State of Chromatin Immunoprecipitation. *Molecular Biotechnology, 45,* 87-100.

Colosimo, M., Brown, A., Mukhopadhyay, S., Gabel, C., Lanjuin, A., Samuel, A., & Sengupta, P. (2004). Identification of Thermosensory and Olfactory Neuron-Specific Genes via Expression Profiling of Single Neuron Types. *Current Biology, 14,* 2245-2251.

Corsi, A., Wightman, B., & Chalfie, M. A Transparent window into biology: A primer on *Caenorhabditis elegans*. *WormBook Genetics,* 1-31.

Doench, J., Hartenian, E., Grahm, D., Tothova, Z., Mudra, H., Smith, I.,...Root, D. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology, 32,* 1262-1267.

Echtberger, J., Lorch, A., Sleumer, M., Zapf, R., Jones, S., Marra, M.,...Hobert, O. (2007). The molecular signature and *cis*-regulatory architecture of a *C. elegans* gustatory neuron. *Genes and Development, 21,* 1653-1674.

Emery, P., Durand, B., Mach, B., & Reith, W. (1996) RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Research, 24,* 803-807.

Farboud, B., & Meyer, B. (2015). Dramatic Enhancement of Genome Editing by CRISPR-Cas9 Through Improved Guide RNA Design. *Genetics, 199,* 959-971.

Friedland, A., Tzur, Y., Esvelt, K., Colaiácovo, M., Church, G., & Calarco, J. (2013). Heritable genome editing in *C. elegans via a* CRISPR-Cas9 system. *Nature Methods, 10,* 741-743.

Fujiwara, M., Sengupta, P., & McIntire, S. (2002). Regulation of Body Size and Behavioral State of *C. elegans by* Sensory Perception and the EGL-4 cGMP-Dependent Protein Kinase. *Neuron, 36,* 1091-1102.

Gagnon, J., Valen, E., Thyme, S., Huang, P., Ahkmetova, L., Pauli, A.,...Schier, A. (2014). Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. *PLOS One, 9,* 1-8.

Helander, H., Koivuranta, K., Horelli-Kuitunen, N., Palvimo, J., Palotie, A., & Kalervo Hiltunen, J. (1997). Molecular Cloning and Characterization of the Human Mitochondrial 2,4-Dienoyl-CoA Reductase Gene (DECR). *Genomics, 46,* 112-119.

Hobert, O., & Wenick, A.S. (2004). Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. *Developmental Cell, 6,* 757-70.

Hurlburt, A. (2014). A Secret to the Synapse: Investigating the Targets of the Transcription Factor DAF-19 in the Nervous System of Caenorhabditis Elegans. Honors thesis, Lawrence University.

Jinek., M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, M,, & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science, 337,* 816-821.

Kellenberger, S, & Schild, L. (2002). Epithelial Sodium Channel/Degenerin Family of Ion Channels: A Variety of Functions for a Shared Structure. *Physiology Review, 82, 735-*767.

Kenyon, C., & Alcedo, J. (2004). Regulation of *C. elegans* Longevity by Specific Gustatory and Olfactory Neurons. *Neuron, 41,* 45-55.

Kim, Y., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: Zinc finger fusions to *Fok* I cleavage domain. *Proceedings of the National Academy of Science, 93,* 1156-1160.

Lee, I., Lehner, B., Crombie, C., Wong, W., Fraser, A.G., & Marcotte, E. (2008). A single gene network accurately predicts phenotypic effects of gene pertubation in *Caenorhabditis elegans*. *Nature Genetics, 40,* 181-188.

Lehninger, A., Nelson, D., & Cox, M. (2008). *Principles of Biochemistry* (5th ed.). New York, NY:W. H. Freeman and Company.

Li, S., Armstrong, C., Bertin, N., Ge, H., Milstein, S., Boxem, M.,...Vidal, Marc. (2004). A Map of the Interactome Network of the Metazoan *C. elegans*. *Science, 303,* 540-543.

McLeod, M., & Wu, S.Y. (1995). The sak1+ gene of Schizocaccharomyces pombe encodes an RFX family DNA-binding protein that positively regulates cyclic AMP-dependent protein kinase-mediated exit from the mitotic cell cycle. *Molecular and Cellular Biology*, *15,* 1479-1488.

Paix, A., Wang, A., Smith, H. Lee, C., Calidas, D., Lu, T.,...Seydoux, G. (2014). Scalable and Versatile Genome Editing Using Linear DNAs with Microhomology to Cas9 Sites in *Caenorhabditis elegans*. *Genetics, 198,* 1347-1356.

Paix., A., Folkmann, A., Rasoloson, D., & Seydoux, G. (2015) High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. *Genetics, 201,* 47-54.

Phirke, P., Efimenko, E., Mohan, S., Burghoorn, J., Crona, F., Bakhoum, M.,...Swoboda, P. (2011). Transcriptional profiling of *C. elegans* DAF-19 uncovers a ciliary base-associated protein and a CDK/CCRK/LF2p-related kinase required for intraflagellar transport. *Developmental Biology, 357,* 235-247.

Pocock, R., & Hobert, O. (2010). Hypoxia activates a latent circuit for processing gustatory information in *C. elegans*. *Nature Neuroscience, 13,* 610-614.

Reith, W., Satola, S., Herrero Sanchez, C., 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,* 897-906.

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 and Development, 4,* 1528-1540.

Reith., W., Ucla, C., Barras, E., Gaud, A., Duand, C., Herrero-Sanchez, C.,...Mach, B. (1993). 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,* 1230-1244.

Riddle, D., Blumenthal, T., Meyer, B., & Priess, J. (1997). *Introduction to C. elegans*. Cold Spring Harbor Press.

Robertson, H., & Thomas, J. (2006). The putative chemoreceptor families of *C. elegans*. WormBook.

Römer, I., Hahn, S., Jordan, T., Bonas, U., & Lahaye, T. (2007) Plant pathogen recognition

mediated by promoter activation of the pepper Bs3 resistance gene. *Science*, *318*, 645-648.

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*, 5517-5528.

Shaye, D., & Greenwald, I. (2011). OrthoList: A Compendium of *C. elegans* Genes with Human Orthologs. *PLOS One*, *6*, 1-12.

Simmer, F., Moorman, C., Van der Linden, A.M., Kujik, E., Van der Berghe, P.V.E., Kamath, R.,...Plasterk, R. (2003). Genome-wide RNAi of *C. elegans* Using the Hupersensitive *rrf-3* Strain Reveals Novel Gene Functions. *PLOS Biology*, 1, 077-084,

Stiernagle, T., & Hope, I.A. (1999) *C. elegans: A Practical Approach*. Oxford: Oxford University Press.

Sulston, J., & Horvitz, H. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Developmental biology*, *56*, 110-156.

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*, 411-421.

Szigeti, B., Gleeson, P. Vella, M., Khayrulin, S., Palyanov, A., Hokanson, J.,...Larson, S. (2014). OpenWorm: an open-science approach to modeling *Caenorhabditis elegans*. *Frontiers in Computational Neuroscience*, 8, 1-7.

Tong, Y., % Bürglin, T. (2010)., Conditions for dye-filling of sensory neurons in *Caenorhabditis elegans*. *Journal of Neuroscience Methods*, *188*, 58-61.

Vassart, G., & Costagliola, S. (2011). G Protein-coupled receptors: mutations and endocrine diseases. *Nature Reviews*, *7*, 362-372.

Wang, J., Schwartz, H., & Barr, M. (2010). Functional specialization of sensory cilia by an RFX Transcription Factor Isoform. *Genetics*, *186*, 1295-1307.

Wiedenheft, B., Sternberg, S., & Doudna, J. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, *482*, 331-338.

White, J., & Jorgensen, E. (2012). Sensation in a Single Neuron Pair Represses Male Behavior in Hermaphrodites. *Neuron*, *75*, 593-600.

Yamamoto, M., Nakayama, J., Yamakawa-Kobayashi, K., Hamaguchi, H., Miyazaki, R., & Arinami, T. (2002). Identification of 33 polymorphisms in the adipocyte-derived leucine aminopeptidase (ALAP) gene and possible association with hypertension. *Human Mutation*, *19*, 251-257.

Zhang, H., & Blumenthal, T. (1996). Functional analysis of an intron 3' splice site in
Caenorhabditis elegans. *RNA, 2,* 380-388.

Zhong, W., & Sternberg, P. (2006). Genome-Wide Prediction of *C. elegans* Genetic Interactions.
Science Reports, 311, 1481-1484.

CRISPR 101: A Desktop Resource. (2016). Addgene, addgene,org.

Online Mendelian Inheritance in Man, OMIM. McKusick-Nathans Institute of Genetic Medicine,
Johns Hopkins University (Baltimore, MD). http://omim.org.

WormAtlas, Altun, Z.F., Herndon, L.A., Worlkow, C.A., Crocker, C., Lints, R., & Hall, D.H.
(2002-2016). http://www.wormatlas.org

APPENDICES

Appendix A: Dye filling Assay

1. Wash worms off plate with 1 mL M9 buffer, and allow them to settle in an Eppendorf tube.
2. Remove supernatant and rinse worms with 1 mL M9 buffer. Repeat this step if the plate has excessive bacterial growth.
3. Remove supernatant. Add 0.5 mL M9 buffer and 5.0 mL DiI. Invert tube several times to mix.
4. Cover tube to avoid light exposure. Place tube on shaker table set for ~60 rpm, and allow to shake for 1-3 hours.
5. Remove tube from shaker table and pipette off supernatant. Rinse in 1 mL M9 buffer. Repeat at least once.
6. Pipette worms onto an NGM streak plate. Allow 12-24 hours before imaging, if worms are to be used for confocal analysis.
7. Dye filling worms will have visibly red amphid and phasmid neurons, clearly distinguishable by their dendrites. Non-dye filling worms will show only intestinal fluorescence.

Appendix B: Cas9 Purification

Unless otherwise noted, all steps should be done on ice or at 4ºC.

1. Prepare the following buffers prior to purification. All buffers are made with Milli-Q water. Dry ingredients should be added and dissolved first in Milli-Q, and the solution should be brought to the correct pH prior to adding glycerol and bringing the buffer to its final volume
    a. Buffer A (100 mL): 20mM Tris HCl pH 8.0, 250 mM KCl, 20 mM imidazole, 10% glycerol, 1 mM TCEP.
    b. Buffer B (100 mL): 20mM Tris HCl pH 8.0, 800 mM KCl, 20 mM imidazole, 10% glycerol, 1 mM TCEP.
    c. Buffer C* (100 mL): 20mM Hepes pH 8.0, 500 mM KCl, 500 mM imidazole*, 10% glycerol
    d. Buffer D (2L): 20mM Hepes pH 8.0, 500 mM KCl, 20% glycerol
    e. Buffer R1* (50 mL): 1M KCl, 1M imidazole
    f. Buffer R2* (50 mL): 1M KCl, 1M imidazole, 2M urea
2. Plate DE3 GOLD (Agilent, #230132) cells with nm2973 plasmid (Fu et al. 2014) and plate on LB + 50 μg/mL Carbenicillin. Grow overnight at 37º C; plate can then be kept for several weeks at 4ºC.
3. Inoculate 25 mL LB + 50 μg/mL Carbenicillin with a single colony from the fresh plate. Incubate at 37ºC overnight.
4. Transfer 5 mL of overnight culture to 1L LB + 0.1% glucose + 50 μg/mL Carbenicillin and grow at 25ºC (220 rpm). Grow culture to OD$_{600}$ ≈ 0.5.
5. Shift culture to 18ºC for 15-25 minutes, then add IPTG to 0.2 mM. It is best to measure out the IPTG into a small Eppendorf tube and add Milli-Q water to dissolve it just before adding it to the culture. Incubate overnight.
6. Make a solution of 100 mM PMSF in 1 mL EtOH.
7. Pellet culture in sterile 250 mL centrifuge bottles using F14S – 6X250Y rotor at 5000 rpm and 4ºC. Obtain wet weight, and resuspend each pellet at ~6 mL/g cells with Buffer A +

protease inhibitor (4 tablets/10 mL, Roche, #11836170001) + 1mM PMSF solution. Combine all samples in two 50 mL falcon tubes.

8. Sonicate cells 6 x 45s with a 1 second pulse/2 second pause at 30% amplitude. Alternate falcon tubes, and allow for 1 minute cooling between each cycle. Tubes should be kept on ice or in an ice bath, and excess bubble production should be avoided by keeping the sonicator tip off the tube wall to maximize sonication efficiency.

9. Recombine lysate and split evenly between two sterile Oakridge tubes. Spin lysate for 30 minutes at 16000 rpm and 4°C in SS34 rotor.

10. During spin step, place a sterile filter in the base of a 5 ml column and seal the side openings of the stopcock with Parafilm. Pipette 5 mL of Ni-agarose beads (50% slurry) into two small falcon tubes, and equilibrate with buffer A by filling the falcon tube to the top with buffer, spinning the beads down for 60 s at 3000 rpm, and removing the supernatant. Repeat equilibration two more times

11. Transfer clarified lysate to a fresh falcon tube and add beads. Batch bind Ni beads for ~25 minutes*.

12. Pipette the nickel beads into the column evenly so that they form a flat layer in the base of the column. Wash these with 50-100 mL* of Buffer B.

13. Elute protein with buffer C, collecting 25 1mL fractions in separate Eppendorf tubes.

14. Run fractions on SDS page gel to determine which fractions contain Cas9 protein, which appears as a large band at 160 kDa. Pool fractions containing Cas9 protein.

15. Prepare a second 5 mL column with a sterile filter and sealed stopcock. Pipette 10 mL Sepharose beads (50% slurry) into the column, and equilibrate with 1M KCl (25 mL) to charge the column, followed by buffer C (25 mL).

16. Flow eluent over Sepharose column and collect in a falcon tube. Dialyze in 1L of Buffer D for at least 5 hours* using a Pierce Cassette. Transfer the cassette to fresh Buffer D (1L) and dialyze overnight.

17. Concentrate protein to ~10mg/mL using a 100K centrifugal filter (Milipore, UFO910024). Centrifuge the eluent for 10 minutes at a time at 5,000 rpm and 4°C until the desired concentration is reached. A Nanodrop 2000 or Bradford Assay can be used to determine final concentrations.

18. Pipette 5 μl aliquots into PCR tubes and flash-freeze the concentrated protein using dry ice bath and store at -80°C.

19. Note that Ni beads can be regenerated as follows:
    a. Run R1 buffer over the column (50 mL)
    b. Run R2 buffer over the column (50 mL)
    c. Run sterile H$_2$O over the column (50 mL)
    d. Run 20% EtOH over the column (50 mL)
    e. Store beads in EtOH at 4°C.