


Spring 5-29-2013

Investigating Potential Target Genes of the RFX Transcription Factor DAF-19 in *Caenorhabditis elegans*

He Zhang

LAWRENCE UNIVERSITY, soukhz@gmail.com

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

 Part of the [Bioinformatics Commons](#), [Biology Commons](#), [Genetics Commons](#), [Molecular and Cellular Neuroscience Commons](#), and the [Molecular Genetics Commons](#)

© Copyright is owned by the author of this document.

Recommended Citation

Zhang, He, "Investigating Potential Target Genes of the RFX Transcription Factor DAF-19 in *Caenorhabditis elegans*" (2013).
Lawrence University Honors Projects. 39.
<https://lux.lawrence.edu/luhp/39>

This Honors Project is brought to you for free and open access by Lux. It has been accepted for inclusion in Lawrence University Honors Projects by an authorized administrator of Lux. For more information, please contact colette.brautigam@lawrence.edu.

**INVESTIGATING POTENTIAL TARGET GENES OF THE RFX TRANSCRIPTION
FACTOR DAF-19 IN *CAENORHABDITIS ELEGANS***

He Zhang

Faculty Advisor: Elizabeth De Stasio

Biology Department

Lawrence University

Appleton, Wisconsin

May, 2013

**This work is presented for Honors at Graduation in Independent Study at
Lawrence University, Appleton, Wisconsin**

I Hereby Reaffirm the Lawrence University Honor Code

Table of Contents

I. Acknowledgements	ii
II. Abstract	iii
III. Introduction	1
Background	1
Transcription Factor DAF-19	4
<i>Caenorhabditis elegans</i>	5
Previous work	7
Research Questions	10
Goals and Hypothesis	14
IV. Materials and Methods	16
PCR	16
Restriction Digestion	21
Ligation and Transformation	22
Lysis PCR	23
Restriction Mapping	24
Microinjection	26
Worm maintenance	30
Dye-filling assay	31
Confocal imaging	31
V. Results	35
Transcriptional:: <i>gfp</i> construct	36
Ligation and transformation	38
Confirming the presence of recombinant plasmid	39
Transgenic worms	46
Confocal images	49
VI. Discussion	61
Choice of the recombinant plasmids	61
Obtaining transgenic worms	62
Gene expression studies	64
Conclusion	67
VII. References	69
VIII. Supplemental Table	

Acknowledgements

I would like to thank the project advisor Elizabeth De Stasio for her kind and enthusiastic help in experimental designs, data collection, data interpretation, and comments on the writing.

I would like to thank Brian Piasecki for his generous help in microinjection and training on the confocal microscope.

I would like to thank Yagmur Esemen, Gaelen Lombard-Knapp and Ethan Landes for their co-operation in data collection and confocal imaging, as well as their critical comments.

I would like to thank Kimberly Dickson for her help in modifying protocols of PCR and restriction digestion.

I would like to thank JoAnn Stamm, Wayne Krueger and Bonnie Arbuckle for making the research possible.

Finally, I would like to thank all the members of my reviewing committee for their patience and passion.

Abstract

Neurodegenerative diseases, such as Alzheimer's disease, are characterized by an age-related decrease in the synaptic activity of the patient's brain. Previous research suggested that a RFX transcription factor DAF-19 in the nematode *Caenorhabditis elegans* (*C. elegans*) may be involved in the maintenance of synaptic protein levels. Particularly, worms that were DAF-19A/B defective showed reduced synaptic activities when compared to their age-matched controls.

This study investigated the role of DAF-19A/B isoforms in the *C. elegans* nervous system. Three genes, *F46G11.3*, *F57B10.9*, and *F58E2.3* were selected as potential downstream targets of DAF-19A/B based on their potential neuronal expression. Among the three, *F46G11.3* and *F57B10.9* were expressed in the neurons. The gene expression pattern of *F57B10.9* in a wild-type animal and a DAF-19 defective animal was shown to be different, suggesting the gene may be a target of DAF-19A/B.

Investigating Potential Target Genes of the RFX Transcription Factor DAF-19 in *Caenorhabditis elegans*

Neurodegenerative Disease

Neurodegenerative diseases are characterized by progressive loss of neuronal function often followed by neuronal cell death. Many diseases, including Parkinson's, Huntington's, and Alzheimer's belong to this category (Ross and Poirier, 2004). Symptoms of each disease include memory loss, impaired cognition, behavioral changes, and loss of mental function (Ross and Poirier, 2004). While effective treatments for neurodegenerative diseases are very new or yet to be discovered, recent advances in genetics enable us to associate genetic mutations with such diseases. A genetic mutation, or a change of the nucleotide sequence in the genome, can result in erroneous protein production or loss of gene products which, either directly or indirectly, lead to changes of an organism's phenotype.

Alzheimer's disease is an irreversible brain disease caused by abnormal deposition of particular malformed proteins in neuronal cells (Alzheimer, 1907). One such protein is amyloid precursor protein (APP), a highly conserved membrane protein primarily expressed in the brain (Zheng and Koo, 2006). The formation of APP plaques due to aberrant of proteolytic cleavage of the protein is partially caused by the single-gene mutations on human chromosome 1, 14, and 21, as each of these mutations is responsible for changes in the way that APP is broken down (Goate et al., 1989; St George-Hyslop et al., 1992). When APP is

cleaved wrongly due to these mutations, it produces fibrous and insoluble proteins called beta amyloids, which can be deposited inside nerve cells and exert harmful effects due to its malfunction. Another protein, called *tau*, also aberrantly aggregates inside neurons in neurodegenerative disease states. *Tau* protein is a microtubule-related protein that modulates the stability of neurons. However, *tau* may undergo a chemical change called hyperphosphorylation and, as a result, self-tangles and deforms the microtubule structure, eventually causing neuron death (Goedert, 1993).

Alternative models of neurodegeneration

Another model for neurodegenerative diseases involves the ubiquitin proteasome system, known as the cellular protein degradation factory. The proteasome is crucial because it controls the level of cellular proteins by cutting and degrading large proteins when needed and, thus, keeps the right concentration of certain proteins in a cell. Although a direct connection between the proteasome system and neurodegenerative diseases remains unclear, past research has shown that the proteasome has a role in controlling the level of some synaptic proteins, such as UNC-13, to change the extent of neurotransmission (Aravamudan and Broadie, 2003). This finding is crucial because it proposes an alternative model of neurodegenerative disease that does not necessarily, but could, involve the aggregation of mis-folded proteins in and around neurons.

With the premise that neurodegeneration may be related to changes in protein degradation, we propose a neurodegenerative disease model that involves Regulatory Factor X (RFX) type transcription factors. A transcription factor is a protein that binds to a DNA sequence and regulates the process of transcription, in which mRNA is produced from the DNA template. The production of mRNA enables cells to produce specific proteins through the process of translation. Thus, a transcription factor acts as a switch by which genes are expressed or unexpressed. RFX transcription factors contain a highly conserved DNA binding domain and they are involved in biological processes ranging from control of the cell cycle to brain development (Efimenko et al., 2005; Emery et al., 1996). Though the RFX proteins are found in humans, it is crucial to elucidate their conserved functions using a model organism. In this study, we studied the protein DAF-19, the only RFX transcription factor in the nematode, *Caenorhabditis elegans*.

Supplemental Information 1: Genetic Notation

The name of both a gene and the protein it encodes are shown in a ‘3-letter plus number’ notation. (displayed below)

The letters, or gene name, often represents the mutant phenotype conferred by a loss-of-function allele of the gene. (e.g. *daf* indicates “dauer formation defective”)

The number indicates the *x*th gene that is designated as conferring such a phenotype.

Gene name: *daf-19*

Protein product of the gene: DAF-19

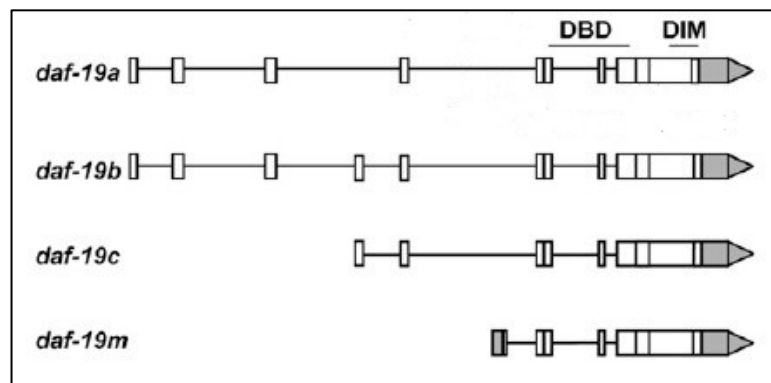
Transcription Factor DAF-19

Due to their highly conserved DNA binding domain, all RFX transcription factors are known to bind to a specific 14 base-pair DNA sequence called the X-box (Efimenko et al., 2005). It was the presence of X-box that shone light on the relationship between DAF-19 and neurons. Swoboda and colleagues (2000) found that the absence of DAF-19 in the nematode, *Caenorhabditis elegans*, results in a loss of cilia, organelles found in eukaryotic cells that project from the larger cell body. Thus, DAF-19 is important in maintaining ciliary structures. Deficiencies in cilia can result in a group of diseases called ciliopathies, which includes polycystic kidneys, neurosensory impairment, and Bardet-Biedl syndrome (Badano et al., 2006). The group further investigated the genes that are involved in ciliogenesis, the production of ciliary structures, and discovered that some genes contain the X-box in their promoter region (Blaque et al., 2005; Burghoorn et al., 2012; Efimenko et al., 2005). X-box deletion leads to loss of ciliary gene expression and, thus, RFX transcription factors may bind to the X-box motif to enhance ciliary gene expression. Thus, DAF-19 plays a crucial role in ciliogenesis and, since cilia are involved in sensory functions, the study also finds an interaction between DAF-19 and the nervous system.

The *daf-19* gene encodes at least four related proteins, or isoforms, of DAF-19. The *daf-19* gene can generate different protein products through processes such as alternative splicing and the use of alternative promoters to initiate transcription. The isoforms include

DAF-19A, B, C and M (Figure 1). Previous research showed that DAF-19C is the major player in ciliary functions (Burghoorn et al., 2012; Senti and Swoboda, 2008). DAF-19M is also known to act specifically in ciliated sensory neurons where it is involved in generating male mating behavior (Wang et al., 2010). The specific function of DAF-19A and B is not clear; however, Senti and Swoboda (2008) suggested that these related isoforms be involved in synaptic maintenance in non-ciliated neurons.

Figure 1. Structure of *daf-19* transcripts. Each box represents an exon and the gaps between each exon are introns. Transcripts *a* and *b* have only one exon difference (exon 4). The *daf-19c* isoform is specific to ciliated sensory neurons. Transcript *m*, the shortest of the four known



isoforms, is responsible for male mating behaviors. All transcripts share a conserved DNA binding domain. DBD = DNA Binding Domain; DIM = Dimerization domain, which could be important for interacting with other transcription factors. Figure from Wang et al. (2010).

Model organism: *Caenorhabditis elegans*

Caenorhabditis elegans (*C. elegans*) is a free-living nematode that can be found in soil compost. Sydney Brenner first proposed that it could be used as a model organism in the late 1960s and, since then, *C. elegans* has been used widely in the field of biology, especially in genetic studies. A number of elements make *C. elegans* a prominent model organism. (1) It

has a fast reproduction rate. *C. elegans* has a short life cycle that lasts for about three days in a 20°C environment. (2) Most *C. elegans* are hermaphrodites, which means that they can produce both eggs and sperm, allowing self-fertilization. Thus, it is easy to create and maintain a large number of homozygous mutants that are useful in comparing the genetics of a mutant to that of wild type. (3) Because of their rapid reproduction, large sample sizes of offspring useful for genetic analysis are generated in a short time period. Mutated alleles are often recessive and, thus, very hard to find in a normal setting. However, this issue can be resolved by studying the vast populations of homozygous mutant *C. elegans*. A larger sample size theoretically provides more consistency and provides a more accurate representation of the phenotype's characteristics. (4) It is a eukaryote, which means it possesses cellular and molecular systems that resemble those of humans. (5) *C. elegans* has a relatively small sequenced genome. Reverse genetics approaches can be used to determine whether a mutation in a gene leads to a certain phenotype. Also, a completely mapped genome means that it is easier to carry out bioinformatics approaches, in which DNA sequences of multiple genes can be clustered together to see whether they have anything in common. (6) *C. elegans* is transparent throughout its life-cycle, making its anatomy easily observable. In addition, researchers can use fluorescent protein labeling to determine the location and movement of cellular components. This provides both temporal and spatial component to the worm's anatomy, which enhances our understanding of the cellular and molecular dynamics of *C.*

elegans. (7) Finally, *C. elegans* is easily cultured in the lab. They can be placed on an agar plate that contains *E. coli* as their food source, and they can be cultured at normal temperatures (15°C-25°C).

Previous Work

Question: What differentiates DAF-19A/B from DAF-19C isoform?

One recent study of the role of DAF-19 (Senti and Swoboda, 2008) in ciliary formation and maintenance, functions ascribed exclusively to DAF-19C isoform, also elucidated potential functions of DAF-19 A/B isoforms. Using antibodies specific to the N- and C-termini of the DAF-19 proteins, they found that while *daf-19c* isoform is present in ciliated sensory neurons, DAF-19A/B isoforms are expressed in over 200 non-ciliated neurons but are not found in ciliated neurons. They also showed that, although worms expressing only DAF-19C had normal sensory-related behaviors, their synaptic transmission was defective. These animals' nervous system appeared normal; however, when compared to age-matched controls, DAF-19A/B defective worms showed a reduction in levels of certain proteins involved in synaptic vesicle transportation (Senti and Swoboda, 2008). Interestingly, such expression patterns could be rescued by DAF-19A expression. Given their potential role in synaptic maintenance, this study tries to unravel the specific roles of the DAF-19A and B isoforms.

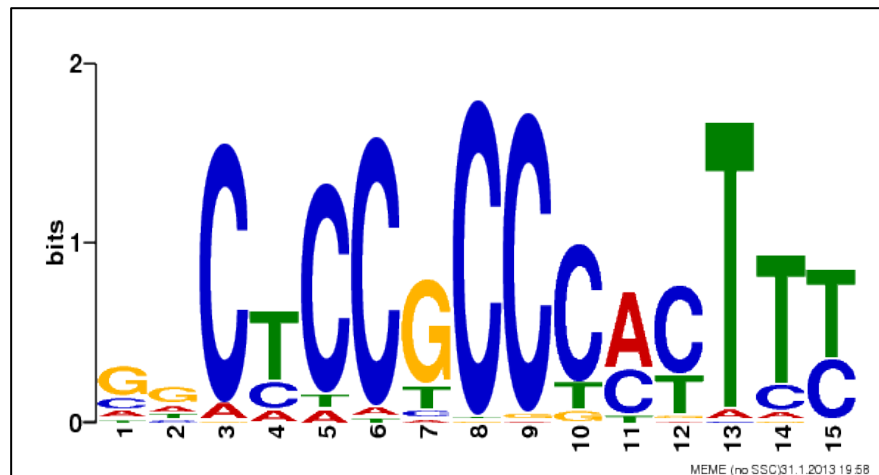
The P.I., Dr. Elizabeth De Stasio, previously investigated the differences in the transcriptome of wild type animals and *daf-19* mutants using microarray technology. Dr. De Stasio investigated a total of 22,500 transcripts extracted from both the wild-type and *daf-19* mutant worms. The transcripts were harvested from identical populations of 2-day-old, wild-type and *daf-19* mutant adult worms. This experiment generated a list of 177 genes that showed more than 1.5-fold differences in expression levels between the wild-type and the *daf-19* mutant worms.

Question: Do the 177 genes have anything in common?

Chelsey Sand (2011) investigated putative control regions of each of the 177 genes, by collecting and comparing one thousand base pairs of DNA sequence upstream of the ATG start codon of each gene. Sand took a bioinformatics approach where she searched DNA consensus sequences that might serve to bind transcription factors that are common in the 177 genes using the program MEME (Multiple Em for Motif Elicitation). MEME aligns multiple DNA sequences and searches for statistically significant similarities in the sequences (Bailey et al., 2006). The use of MEME enabled Sand to generate a pictorial representation of a novel 15 base pair consensus DNA sequence that, including its variants, was found in 50 genes among the list of 177 genes (Figure 2). This means that nearly 30% of the genes contained a putative novel DNA binding sequence, and this sequence may be a potential binding motif of

DAF-19, used to regulate the genes' expression. This novel sequence was named the "Sand-box" after its discoverer.

Figure 2. A consensus Sand-box motif. A novel DNA motif found in the promoter region of 50 out of 177 genes selected by Dr. Elizabeth De Stasio. (Sand, 2011) The height (y-axis) of each nucleotide base (ACGT) indicates the relative frequency at the given



position (x-axis) in the putative motif. E-value (the probability that the generated sequence is found by chance in the data set) equals $7.6 \cdot e^{(-92)}$, indicating that the motif is extremely unlikely to occur at random in the set of sequences ranging from 370 to 1300 base pairs that she used. Image generated by MEME program.

(Program found at: <http://meme.nbcr.net/meme/cgi-bin/meme.cgi>).

Supplemental Information 2: The necessity of *daf-12* gene mutation

One setback in using *daf-19* mutant worms is that the mutants undergo dauer larva formation at standard temperature (25°C).

Dauer formation means a developmental arrest. Therefore, there are virtually no adult *daf-19* mutant worms.

To resolve this issue, another gene in the dauer formation pathway, *daf-12*, is mutated. A *daf-12* mutant does not form dauer larva and, since it functions earlier in the dauer formation pathway than does *daf-19*, a *daf-19; daf-12* double mutant does not undergo dauer formation. Thus, an adult *daf-19* mutant can be observed under microscopes.

To make sure that the only dependant variable is the presence of *daf-19*, *daf-12* is also mutated in the control strain.

‘Wild-type’ genotype (control): *daf-19* +/+ ; *daf-12* -/-

***daf-19* mutant genotype:** *daf-19* -/- ; *daf-12* -/-

Current Research Questions and Approaches

Question 1: Among the 177 genes of interest, which genes should be prioritized?

In order to determine whether the 177 genes are expressed in non-ciliated neurons along with the DAF-19A isoform, we first prioritized research on several genes based on: (1) a connection with the nervous system established by previous research; (2) the presence of a gene homolog in humans; (3) the presence of a Sand-box; or (4) an established gene expression pattern in the nervous system of *C. elegans*. In this study, genes *F46G11.3*, *F57B10.9*, and *F58E2.3* were investigated to determine whether they are expressed in neurons and whether their expression levels truly differ in wild-type and *daf-19* mutant animals. (Supplemental Table 1)

F46G11.3 is homologous to a Cyclin-G-associated kinase (GAK) gene in *Homo sapiens*. GAK is known as a regulator of clathrin-mediated membrane trafficking, a process that is required for endocytosis, the intake of material from outside the cell by that cell (Sieburth et al., 2005; Zhang et al., 2005). This is crucial because neurotransmission relies on the process of endocytosis to recapture released neurotransmitters. If endocytosis is inhibited, the cell will receive less input of neurotransmitters and, eventually, cell death occurs. Thus, this gene meets criteria (1) and (2) above. Its expression pattern in worms is unknown and it does not contain an identified Sand-box motif in 2000 base pairs upstream of the coding region.

F57B10.9 also has a human ortholog, called SPARTIN. Mutation of this gene in humans is associated with Hereditary Spastic Paralegia, generally known as Troyer Syndrome. The disease is characterized by numerous neurological disorders, especially progressive weakening of the lower half of the body. *F57B10.9* is an interesting target primarily because of its importance in the nervous system. This gene comported with all criteria stated above; it contains a consensus Sand-box motif located 315 base pairs upstream of the ATG of the gene (Sand, 2011). Furthermore, Dupuy and colleagues (2007) have already determined that *F57B10.9* is expressed in the *C. elegans* nervous system in wild-type animals. Korzynski (2012) compared *F57B10.9* gene expression in both wild-type and *daf-19* mutants using confocal microscopy and found that the two strains showed a difference in gene expression in

the anterior, or head neurons. This study continues confocal imaging of these worm strains to ensure that these observations are valid.

F58E2.3 is a Sand-box-containing gene that does not have a known *H. sapiens* ortholog. It was chosen to test whether it is expressed in neurons primarily to test the validity of the Sand-box. Thus, it meets criterion (3). Through use of the MAST program (program can be found at: <http://meme.nbcr.net/meme/cgi-bin/mast.cgi>), which discovers the presence of a user-defined DNA motif in specific DNA sequences (Bailey et al., 2006), I showed that the putative control region of *F58E2.3* contains the consensus Sand-box sequence (GGCTCCGCCCACTTC, Figure 1) located 359 base pairs upstream of the ATG of the gene (Sand, 2011). It is significant because the transcriptional regulatory sequences in *C. elegans* genes are known to be located within 2000 base pairs upstream of the ATG of the gene (Blackwell and Walker, 2006). Therefore, the presence of the consensus Sand-box sequence may imply DAF-19 mediated regulation of the gene.

Question 2: Are the genes of interest expressed in the C. elegans nervous system?

Given that previous research has suggested a role for DAF-19A/B isoforms in the nervous system of *C. elegans*, specifically, in non-ciliated neurons, it is crucial to visually identify each gene's expression pattern in the nervous system. To do so, we fused the putative control region of each gene of interest to a plasmid vector that contains Green Fluorescent

Protein (GFP), a protein that emits green fluorescence when a laser is shot at it at a certain wavelength (488nm). This construct is called transcriptional::*gfp* fusion and any cell that has the molecule(s) to bind to the putative control region of the gene will emit green fluorescence because instead of the actual gene product, GFP is produced. Once these transcriptional fusion constructs are obtained, they are inserted into the gonad of a hermaphrodite worm via microinjection so the expression of the construct can be followed in live animals. To screen transgenic worms expressing added DNA, we co-inject another construct named *elt-2>::mCherry*, which allows us to select the transgenic worms based on the expression of red fluorescent protein in their intestines. When injected with the constructs, the worm's offspring may form an extrachromosomal array and utilize the injected DNA (Evans, 2006). This enables transcription and translation of GFP and mCherry proteins and, thus, transgenic worms display green and red fluorescence in particular cell types. Since the inheritance of the array is unstable, we allow the transgenic worms to reach the F2 generation offspring before undertaking visual analysis of gene expression patterns.

When the inheritance of an extrachromosomal array is stable, we proceed with confocal imaging, through which we can observe the specimen at multiple focal planes. This technique can thus generate a three-dimensional reconstructed image of the expression pattern of the gene. Since the *C. elegans* nervous system is well understood to the level of cell bodies, a three-dimensional picture allows us to see whether the expression pattern of the

gene is neuronal.

Question 3: Is the gene expression DAF-19 dependent?

After observing the expression pattern of the gene of interest and confirming it is neuronal, we compare the gene expression pattern in wild-type and *daf-19* mutant animals. To do so, we set up matings between the injected strain and the corresponding control strain and allow the same extrachromosomal array to be expressed in the other strain as well. Therefore, we generate two strains of animals that have the same array of transgenes, but that differ in the presence of DAF-19 protein. Three dimensional image analyses via confocal microscope will be used again to assess the difference in gene expression between the wild-type and *daf-19* mutant animals. Thus, we can examine whether expression of the gene is, either directly or indirectly, controlled by the DAF-19 protein.

Goals and Hypothesis

Experimental Goal: Determine the relevance of Sand-box motif and neural gene expression.

In addition to the two genes that contain a consensus Sand-box motif in their putative control regions (*F57B10.9* and *F58E2.3*), another subset of genes from De Stasio's microarray-generated list were tested in the lab of Dr. Peter Swoboda. Dr. Swoboda selected a number of genes and created transgenic worms with a wild-type genotype (*daf-19* +/+; *daf-12* +/+). We selected genes studied by Dr. Swoboda's lab that contain a consensus

Sand-box motif in their putative control region. We hope to elucidate whether genes that have a Sand-box motif are expressed in neurons and whether those genes are differentially expressed in wild-type and mutant animals.

Hypothesis: The Sand-box motif is present in genes that are associated with the nervous system and, based on the previous microarray results, three-dimensional image analysis will confirm that genes with a Sand-box motif in their putative control region are expressed differently in wild-type and *daf-19* ^{-/-} mutant animals.

Broader Goal: Determine whether DAF-19A/B is a key to understanding neurodegenerative diseases.

Genes that are regulated by DAF-19A/B proteins and that show a significant difference in their expression levels in wild-type and *daf-19* ^{-/-} mutant animals can be analyzed and evaluated for their connection with certain cellular pathways. Our final goal is to elucidate whether DAF-19A/B and their downstream effectors can be targets for curing neurodegenerative diseases such as Alzheimer's disease, given that they are involved in maintaining synaptic protein levels.

Materials and Methods

Primer Design

To investigate how selected genes are regulated, we used polymerase chain reaction (PCR) to amplify putative control regions located upstream of each gene's translational start site. The putative promoters were ligated with a plasmid vector, pPD95.75, which encodes green fluorescent protein (GFP).

Primers are short chains of a single stranded nucleic acid that bind to single stranded DNA to initiate DNA replication. In order to amplify a promoter region that is double stranded, we designed two primers that each correspond to the 5' or 3' end of the promoter. The upstream primer contains a sequence that is identical to the 5' end of the target gene's promoter. The downstream primer contains a sequence complementary to the 3' end of the template DNA plus strand (Table 1). Amplification of the two complementary sequences enables us to obtain a double stranded promoter region.

Followed by amplification, the amplicons have to be cloned into the expression plasmid. To facilitate cloning, each primer also contains a restriction enzyme digest site that can be cut by the same restriction enzymes used to cut the plasmid vector. Restriction digestion creates sticky ends on the plasmid vector that are complementary to those of the cut amplicons; these ends can be ligated together.

Gene	Direction	T _m (°C)	Primer Sequence
F46G11.3	Foward	62.8	GTCGCAA GTCGAC AAGTTGAAGATAATCGAAAGACTACAG
F46G11.3	Reverse	63.6	CGACAGC GGATCCCTTGGAATTGCAGATTTATAGT
F58E2.3	Forward	60.8	CTCAGG T CGAC ATGCCACTGATAACTTGTATTC
F58E2.3	Reverse	59.6	GCAGCAC GGATCCCTTGGAATCTGAAATTTTTTG

Table 1. Primers designed to amplify the putative control regions of *F46G11.3* and *F58E2.3* genes. Black colored sequence indicates the annealing sequence. The red color indicates either of the two restriction sites used (Sall or BamHI). The blue nucleotides were added to stabilize restriction enzyme binding, and for adjusting the melting temperature of the primers. T_m = Melting Temperature.

Potential primers were characterized using the OligoAnalyzer (ver. 3.1) (<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>) to see if the primers met the following criteria:

1. Annealing sequence of 20-30 base pairs in length.
2. G/C content between 40% and 60%.
3. Forward and reverse primers with similar annealing temperature
(different by $\leq 5^{\circ}\text{C}$).
4. The primers will not dimerize or form hairpin structures.

The expected PCR product when using the *F46G11.3* putative control region as a template is 1193 base pairs in length, while for *F58E2.3*, the expected amplicon size is 499 base pairs.

Polymerase Chain Reaction (PCR)

Through PCR, we can generate a large number of copies of a certain region of DNA from a small amount of template DNA. Self-designed primers from Integrated DNA Technologies were used in PCR to generate large number of copies of each desired promoter region. Genomic DNA of wild type (N2) worms was used as the template in each PCR.

PCR repeats a cycle that consists of three parts: denaturation, annealing, and elongation. During denaturation, the temperature is raised to 95°C, which breaks the hydrogen bonds between the DNA bases. This step separates the double stranded DNA molecules into two single stranded DNAs, which means the bases are exposed for primers to bind. During the annealing step, the temperature is lowered for 30 seconds so that the primers may anneal to the desired template strands. Since each primer pair has a different optimum annealing temperature, we first tested two annealing temperatures (58°C and 60°C) for each PCR reaction. For later PCR reactions that involved the primers we designed, 58°C was used as the annealing temperature. In the elongation step, the temperature was raised to 72°C for 45 seconds, based on previous successful PCR amplification protocol. This temperature allows TAQ DNA polymerase to add nucleotides that are complementary to the template strands. A PCR repeats the cycle 35 times to amplify the targeted promoter region, and is further followed by a 72°C final elongation step for 3 minutes. The purpose of the last step is to provide a sufficient time to add nucleotides to any fragments that are incompletely

extended, and prevents any unexpected sticky ends in the amplicon.

The MgCl_2 concentration was adjusted to achieve optimum reaction conditions for each primer pair, because Mg^{2+} ions act as a co-factor that helps the activity of DNA polymerase and affects primer binding. We modified the MgCl_2 concentration in the PCR reaction mix to $3\mu\text{M}$ for amplification of the F46 promoter and to $4\mu\text{M}$ for that of F58. In summary:

PCR recipe for 100 μL Reaction Volume

2 μL *C. elegans* N2 strain (wild type) template DNA (85ng/ μL)
10 μL Forward Primer (10 μM)
10 μL Reverse Primer (10 μM)
5 μL dNTPs (10 μM)
10 μL 10x Mg Free Taq Buffer (New England Biolabs)
12 μL (F46)/ 16 μL (F58) MgCl_2 (25mM, New England Biolabs)
0.8 μL Taq polymerase enzyme (5000 U/mL, New England Biolabs)

PCR cycles

1. Initialization: 95°C – 3 minutes
2. Denaturation: 94°C – 45 seconds
3. Annealing: 58°C – 30 seconds
4. Elongation: 72°C – 45 seconds
($\times 35$ of cycles of step 2-4)
5. Final Elongation: 72°C – 3 minutes
6. Hold: 8°C

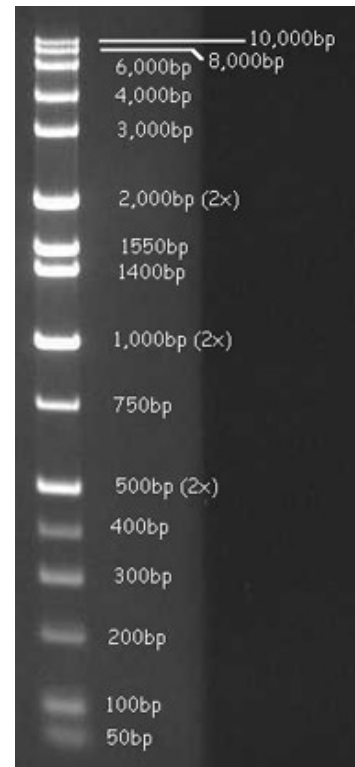
Gel Electrophoresis

The PCR products were analyzed using gel electrophoresis, in which the size of the amplicons was measured using the Hi-Lo DNA ladder (Minn. Molecular). We employed 1.0% agarose gels and 95V electric potential to drive the electrophoresis. Since the phosphate backbone of DNA is negatively charged, we can apply electrical current to force DNA to move towards the positive charge. The DNA separation step lasted for about 60min, and the gel was stained in ethidium bromide (EtBr) for 15 minutes before its visualization under UV light. EtBr is a flat fluorescent molecule that will intercalate into the DNA molecule, which means it can place itself between the planar nucleotide bases. Since EtBr is fluorescent under the UV light, the intercalation allows us to see how far the DNA travels in the gel. A larger DNA fragment will move less than a smaller DNA fragment because a large fragment cannot pass through the pores of the gel very easily and, thus, requires more time to travel down the gel. The Hi-Lo DNA ladder contains multiple DNA fragments that are different in size, which means it displays multiple bands on the gel that can be compared to our experimental bands (Figure 3).

Figure 3. A sample Hi-Lo DNA ladder visualized on a 1.0% agarose gel. Each band contains DNA fragments of different sizes (labeled on the right side of the bands) and by visualizing the ladder and unknown DNA fragments simultaneously on the gel, the size of the unknown DNA can be determined.

(Image taken from: COSMO BIO CO., LTD.

<http://www.cosmobio.co.jp>)



Restriction Digestion

After confirming that amplicons of the expected size were generated through PCR, we used restriction enzymes to cut the two ends of the amplicons. Restriction enzymes are proteins that recognize and cut covalent phosphodiester bonds at a certain DNA sequence. Bacteria produce restriction enzymes to counteract viruses that insert DNA sequences into their genome. Since the target sequence of each restriction enzyme is specific and unique, we used these enzymes to cut both the amplicons and the plasmid vector pPD95.75, so that the amplicons and the vector have complementary sticky ends. This step ensures that the amplicon is inserted into the vector in the right orientation.

A restriction digestion reaction includes the restriction enzymes, associated buffer, and DNA. The reaction vial was placed at 37°C for 60 minutes, and the results were analyzed with gel electrophoresis. We then performed a PCR purification using QIAquick PCR

Purification Kit[®] in order to eliminate the small DNA fragments that are separated from the original amplicon. The same purification step was applied to digested vector pPD95.75.

Ligation and Transformation

After both the amplicons and the cloning vector pPD95.75 were digested, the restriction sites became “sticky ends”. The amplicons were fused with the linearized pPD95.75 vector using Quick Ligase (New England Biolab). Ligase is an enzyme that connects the two DNA strands by catalyzing the formation of phosphodiester bonds between the sticky ends and reconnecting the backbone of the two strands.

The fused plasmid was transformed into DH5α *E. coli* competent cells. Transformation means genetic alteration of a cell by taking up foreign genetic material. Competent cells had been exposed to CaCl₂ so that the transfer of genetic materials happens much easier. They are capable of incorporating foreign DNA especially when they are given a ‘shock.’ We induced a heatshock by placing the cells in a 42°C for exactly 30 seconds and then put them back on ice. After 60 minutes of incubation at 37°C in liquid LB medium, the cells were plated on LB agar which contained antibacterial ampicillin (100µg/mL) and incubated at 37°C overnight.

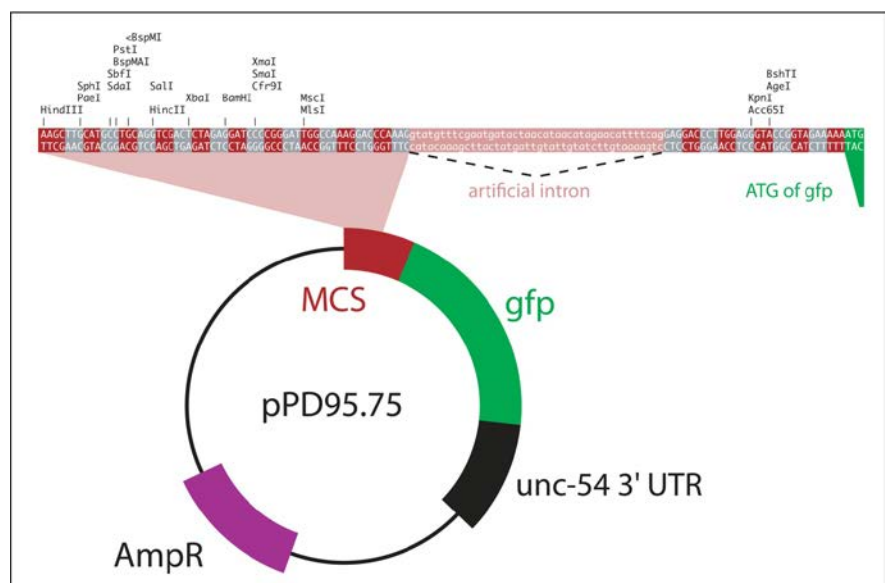
If a competent *E. coli* cell has incorporated a successfully ligated plasmid, the cell will express the ampicillin resistance gene in the pPD95.75 plasmid vector, and it will form a

colony. As a negative control, we transformed linearized pPD95.75 vector, which was cut with both restriction enzymes, into competent cells and plated them under the same conditions. The negative control should show little or no colony formation, indicating the cut pPD95.75 vector is unable to self-ligate.

Lysis PCR

In order to further verify the presence of the intended amplicon in the transformed plasmid, lysis PCR was performed. We used two primers that bind to specific sites upstream and downstream of the multiple cloning site (MCS) in the vector pPD95.75 to carry out this reaction (Figure 4). When a PCR is carried out using the original pPD95.75 vector as the template DNA, the two primers will generate a band that is 347 base pairs in length. However, if the vector contains the insert in between the two restriction sites, the PCR product will be 347 base pairs + the size of the amplicon in length.

Figure 4. Simplified plasmid map for the cloning vector pPD95.75. The AmpR gene sequence allows the transformed cells to express ampicillin resistance gene. The sites recognized by each specific restriction enzyme are displayed above. Image from Addgene (<http://www.addgene.org>)



Colonies were randomly selected and each colony was lysed using a lysis buffer [10X] and the PCR reaction was carried out using the plasmid inside the cell as the template DNA. When colonies contained the desired product, we streaked each colony and then extracted its DNA by following the QIAprep Miniprep Kit protocol.

Restriction Mapping

After the presence of insert was confirmed, we investigated the orientation of the insert using the restriction enzyme HindIII for *F46G11.3* and MfeI for *F58E2.3*. Restriction digests were completed as described above. The common feature of the two enzymes was that they cut the plasmid at two sites: one in the insert, and another in the pPD95.75 backbone. The estimated size of pPD95.75 with insert is 5674 base pairs for *F46G11.3* and 4980 base pairs for *F58E2.3* (sequence retrieved and modified from <http://www.addgene.org/1494/>), and we arbitrarily assigned the HindIII site as the first base pair of the plasmid.

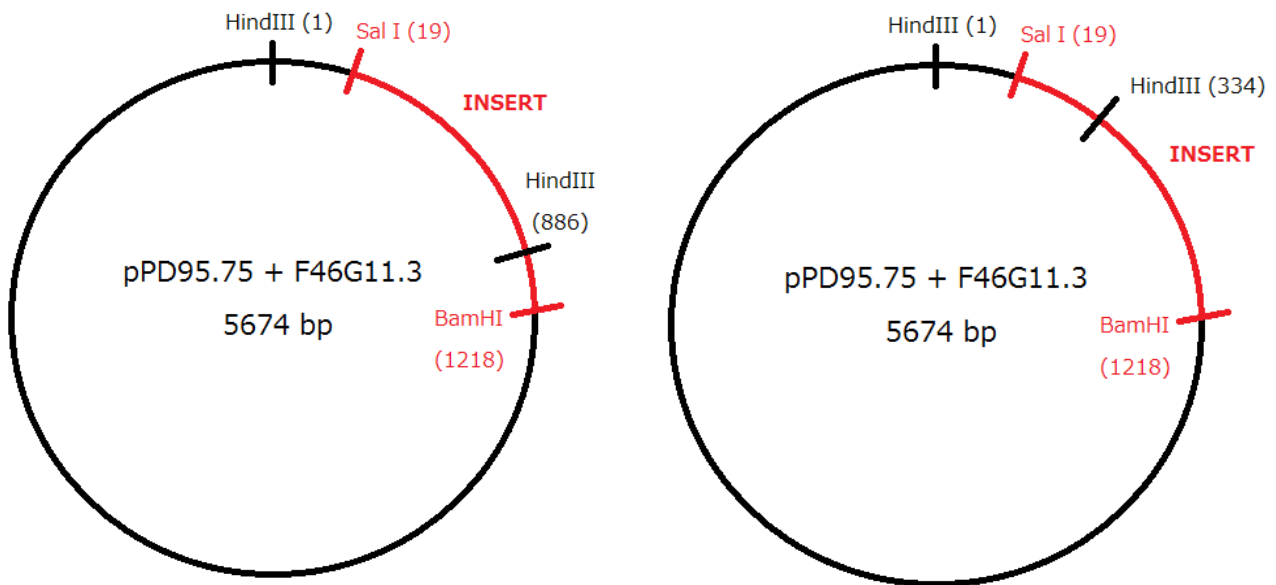


Figure 5. Restriction mapping of the plasmid pPD95.75 fused with the promoter of gene *F46G11.3*. Left: the plasmid map if the insert is in the correct direction. Restriction digestion using HindIII will generate two fragments that are 885 base pairs and 4819 base pairs in length. Right: plasmid map if the insert is in the wrong direction. HindIII will cut the circular plasmid into two fragments that are 333 base pairs and 5341 base pairs in length.

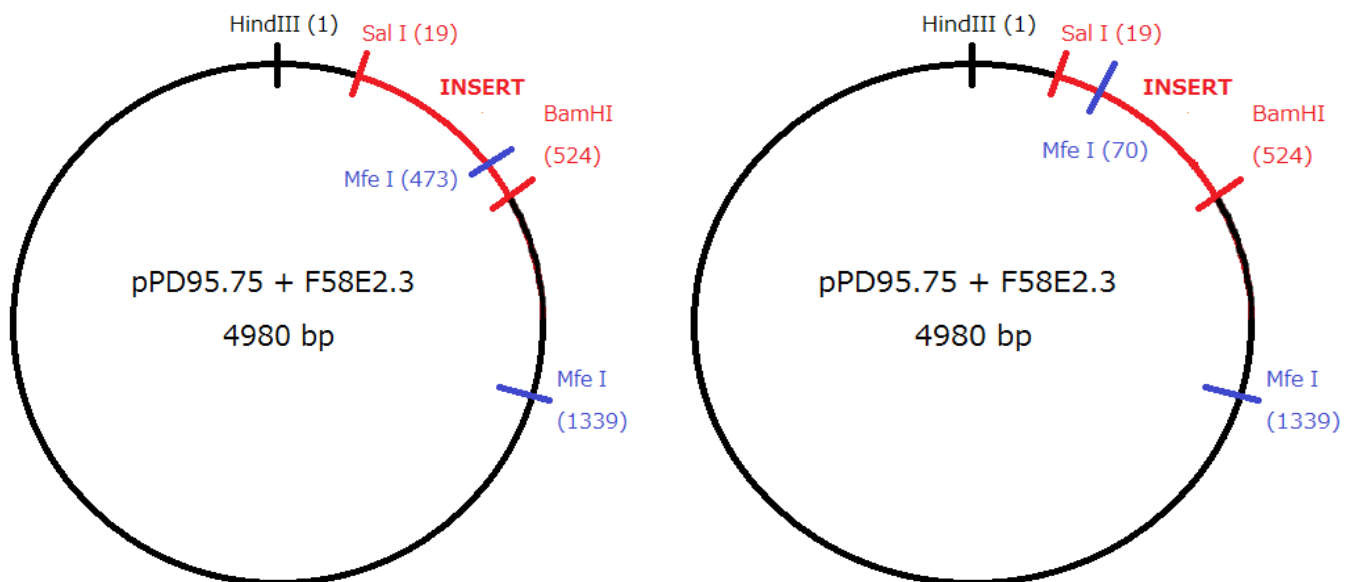


Figure 6. Restriction mapping of the plasmid pPD95.75 fused with the promoter of gene *F58E2.3*. Left: plasmid map if the insert is in the correct orientation. Restriction digestion using MfeI will generate two fragments that are 866 base pairs and 4114 base pairs in length. Right: plasmid map if the insert is in the wrong orientation. MfeI will generate two fragments that are 1279 base pairs and 3701 base pairs in length.

Since the ligated plasmid is circular, enzymes such as HindIII and MfeI that cut the plasmid twice, separate the plasmid into two fragments. These fragments can be visualized on a gel and we can compare our result to our expected results based the DNA sequences retrieved from addgene and wormbase.org

After a successful restriction mapping, we sent a mixture of the constructed plasmid and one of the two self-designed primers to the DNA Analysis Facility on Science Hill at Yale University for DNA sequencing. One primer will initiate the sequencing process by facilitating the nucleotides to bind to constructed plasmid. The sequencing result was aligned with our expected DNA sequence retrieved from wormbase.org via the Clustal Omega tool (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).

Microinjection: choosing the worm strain

Transcriptional::*gfp* fusions were injected directly into the gonads of worms that were either wild-type or *daf-19* *-/-* mutant. The choice of the host strain was based on the results of the microarray analysis (E. De Stasio, unpublished). With the premise that transcriptional::*gfp* fusion expression will reflect the fold-change in gene expression seen in *daf-19* *-/-* and *daf-19* *+/+* populations, the construct was injected into the strain that is expected to generate more gene expression in order to obtain a clearer picture of possible GFP distribution. We injected two transcriptional::*gfp* fusions that contained putative control

regions of genes *F46G11.3* or *F58E2.3*.

Gene name: *F46G11.3*

Fold-change in gene expression in *daf-19* mutant animals: 0.66

Interpretation: More gene expression in **wild-type** adult worms

Experimental design: Inject *F46G11.3::gfp* construct into wild-type animals
(OE3738: *daf-12* -/-; *him-5* -/-)

Gene name: *F58E2.3*

Fold-change in gene expression in *daf-19* mutant animals: 0.60

Interpretation: More gene expression in **wild-type** adult worms

Experimental design: Inject *F46G11.3::gfp* construct into wild-type animals
(OE3738: *daf-12* -/-; *him-5* -/-)

Microinjection

Once we confirmed that the plasmid construction was successful, we microinjected the plasmid into the gonads of adult hermaphrodite worms to create transgenic offspring. We created an injection ‘mix’ containing three kinds of plasmids: 1. plasmid pPD95.75 with either the *F46G11.3* or *F58E2.3* promoter; 2. *elt-2::mCherry* as a co-injection marker; 3. pBR322 plasmid. *elt-2::mCherry* is another reporter construct that is expressed using the *elt-2* gene promoter and creates a red fluorescent protein. Since *elt-2* is only expressed in the worm’s intestine and the color under fluorescent microscopy is red, we can easily recognize any worm that expresses the transgene. pBR322 is an *E. coli* cloning vector and has no effect in *C. elegans*. It is added to the injection mix in order to decrease the toxicity of the newly added foreign plasmid by diluting the overall expression plasmid concentration, as well as to

increase the chance for the worm to form a new pseudo-chromosomal array in its cells. This array ensured that whenever *elt-2::mCherry* was incorporated in the worm, our transcriptional::*gfp* construct was also present in the worm. For injection mix, each plasmid's concentration was adjusted to the following:

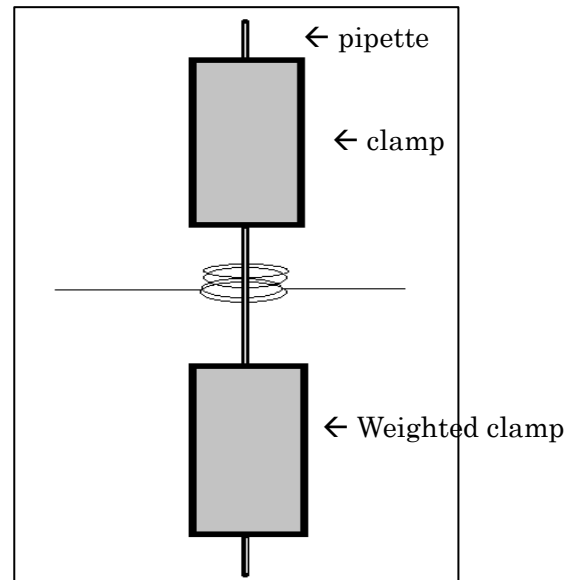
<i>F46G11.3/F58E2.3 promoter::GFP</i>	40ng/ μ L
<i>elt-2::mCherry</i>	10ng/ μ L
pBR322	50ng/ μ L
<hr/>	
Total	100ng/ μ L

We added M9 buffer to bring the total volume up to 50 μ L, and we centrifuged the mixture at 14,000rpm for 10 minutes immediately before microinjecting to spin down any particulate matter in the mix.

Microinjection needles were made by melting a glass capillary. Once the capillary was placed on the 'puller,' heat was applied to the center of the capillary and eventually the weight of the arm of the puller would drag down a half of the molten capillary (Figure 7). Therefore, two microinjection needles can be made from one glass capillary. 0.75 μ L of microinjection mix was carefully taken from the top layer of the solution and pipetted into the needle opening. After approximately three to five minutes, the injection mix falls to the needle tip and the needle can be loaded into the Narishige microinjector. The needle tip was then broken open either by scratching the tip across the 2.0% agarose pad, or by hitting the tip with a thin glass capillary tube, which is thinned by heating and stretching it. A successful

break in the needle can be verified by observing a small bubble of liquid coming out of the needle when hitting the injection button on the injector.

Figure 7. Schematic representation of a pipette puller. A glass pipette is inserted through the two pullers (gray) and an electric current (19amps) flows through the coiled wire. The resistance of the wire produces heat and melts the glass pipette, which forms two needles.



While one experimenter broke the needle, another experimenter picked young adult hermaphrodite mutant worms (*daf-12* *-/-*; *him-5* *-/-*) from a plate and placed them onto an NGM agar plate lacking a bacterial lawn. Since contamination is always an issue in maintaining *C. elegans*, we attempt to reduce contamination as much as possible. Then a single worm was placed on a 2.0% agarose pad in a small drop of mineral oil, used to marginally hydrate the worm while it is stuck to the agarose pad for an accurate injection.

When the needle was set right next to the worm, the height of the needle and the focal plane was adjusted so that the tip of the needle is on the same plane as the gonad in the worm's body and at about a 45° angle to the body of the worm. After pushing the needle through the worm cuticle, the injection mix was injected into the gonad of the worm. The

swelling of the gonad due to the addition of fluid signified a successful injection. Some oocytes in the gonad took up the plasmids and the plasmids formed an extrachromosomal array that contained all 3 plasmids. A drop of M9 buffer was applied to the worm to enhance its ability to recover from the trauma. The injected worm was then rescued from the agarose pad onto an OP50-streaked plate and another drop of M9 buffer was added.

Transgenic Worms Culture and Strain Maintenance

Worms that were injected with the plasmid mixture were observed daily under the fluorescent microscope for screening. Transgenic worms were singly picked onto an NGM agar plate based on the expression of red fluorescent protein mCherry in their intestines. Worm pick was created by fusing a platinum wire with a glass Pasteur pipette. The experimenter first dipped the worm pick into the OP50 *E. coli* growing on a NGM plate and then gently touched a worm so that the worm adhered to the *E. coli*. After the worm was transferred to a new plate, the plate's edge was sealed with Parafilm to preserve the humidity and the plate was stored in the 20°C incubator. After three to five days when one to two life cycles of the worms had passed, the same process was repeated for the downstream generations until each plate contained a large number of transgenic worms.

Dye-filling Assay

A fluorescent lipophilic dye, DiI, was employed to stain certain ciliated neurons in *daf-19* ^{+/+} wild-type animals. The dye-filling assay provides us a landmark of what neuronal cells we are observing and, thus, is important during image analysis.

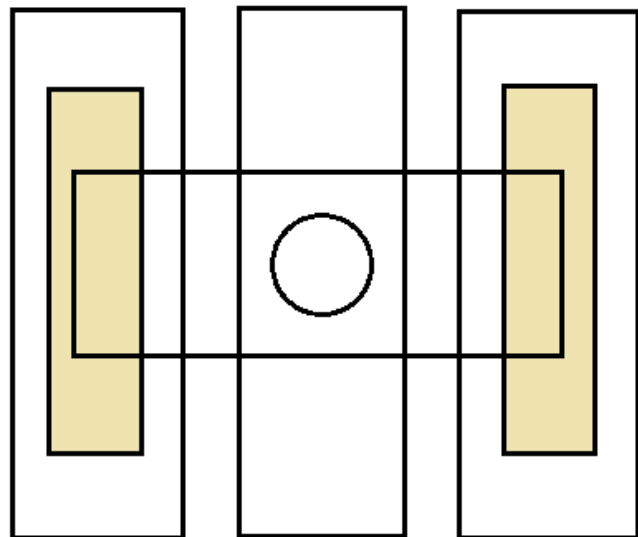
The DiI staining solution was prepared by adding 5 μ L DiI stock solution to 0.5mL M9 buffer. The final concentration was 20 μ g/mL concentration. The worms were rinsed with 2.0 mL of M9 buffer and transferred into a 1.5mL Eppendorf tube. After approximately 5 minutes, worms would settle down at the bottom of the Eppendorf tube. The excess supernatant was pipetted out and the 0.5mL DiI staining solution was added to the Eppendorf tube that contained the worms. The Eppendorf tube was covered with aluminum foil and rotated for 1 hour. The worms were rinsed twice with 1.0 mL of M9 buffer and they were plated on a new NGM plate so that they can eat and excrete excess DiI that was left in their body for at least two hours. DiI can be visualized using the same excitation laser as for visualizing mCherry.

Preparation for Confocal Imaging

Before acquiring images, worm pads were prepared on Corning 1mm thick microscope glass slides so that the worms could be pipetted onto the pad and be observed under the confocal microscope. A bottle of 2.0% agarose was prepared with 0.40g of agarose

mixed with 20mL of M9 buffer. The mixture was heated with microwave at 1500W for 20 seconds. 500 μ L of the heated mixture was transferred to a 1.5mL Eppendorf tube and placed at 95°C in a heat block. 10 μ L of 1M sodium azide (in DMF, a common solvent) was added so that the worms could be anaesthetized during the imaging. A drop of the agarose mixture was pipetted onto a microscope slide and another slide was placed perpendicular to bottom slide so that the agarose drop spread in a thin pad (Figure 8).

Figure 8. A schematic diagram for creating a worm pad. The colored part indicates a layer of labeling tape, which prevents the worm pad from being crushed by the top slide. The drop of agarose is located at the middle.



To collect a population of worms for imaging, worms were washed from an NGM plate with 1.5mL of M9 buffer and transferred to a 1.5mL Eppendorf tube. The worms were left for 5 minutes so that they could settle at the bottom of the tube. The worms were pipetted from the Eppendorf tube onto the worm pad with a pipette and a #1 (0.13mm thick) cover slide was placed over the pad.

Confocal Imaging

All confocal images were acquired using the Leica SP5 X confocal system. To view GFP expression, we excited the fluorophore with the argon laser at a wavelength of 488nm. This allows the excited fluorophore to emit green fluorescence and the confocal microscope captured fluorescence of wavelengths between 495nm and 590nm using the HyD detector. Live images were acquired with two distinct settings: 1) “zoomed setting”, using the 40x oil-immersed objective lens and a zoom factor of 2.0; and 2) “body-view setting”, using the 10x objective lens and a zoom factor of 1.7. The zoomed setting was employed to acquire images of a certain part of the worm, such as the head, the vulva region and the tail, with greater detail. When using the zoomed setting, a drop of immersion oil was dripped onto the 40x objective lens prior to placing the slide in the holder. The body-view setting was used to acquire a whole-body picture of a fluorescent worm. This setting provided us an overview of the GFP expression pattern in each worm and further image acquisitions using the zoomed setting could be conducted.

The target worm was first spotted using differential interference contrast (DIC) light and the focal point was adjusted to the plane that showed clearest image. All image acquisition was conducted with the room light turned off. The image resolution was set to 1024 pixels for X dimension and 256 pixels for Y dimension. The scan field was rotated so that the worm faces left from the observer’s perspective and the ventral and dorsal sides of

the worm were aligned with the top and the bottom of the scan field, respectively. Settings for laser output and general hardware settings are displayed in Table 2.

Laser Settings	
AOTF (488nm)	10.00%
HyD4 (Hydrogen Light)	Active
HyD4 gain	10.00
Laser (Argon, visible)	12%
Scan Field Settings	
X dimension	1024
Y dimension	256
Field Rotation Angle	Varies between $\pm 100^\circ$

Table 2. General hardware settings for the Leica SP5 X confocal system. AOTF indicates the wavelength of the laser used to excite the fluorophore.

All confocal images were characterized and analyzed on Leica Application Suite Advanced Fluorescence simulator software. Images taken at each focal plane were max-projected to produce a planar view of GFP expression in the worm. DIC images of the same scan field were also acquired to determine where the GFP expression was localized. For worms that had GFP expression in the body wall, a movie that shows images from each focal plane was created to look for greater details.

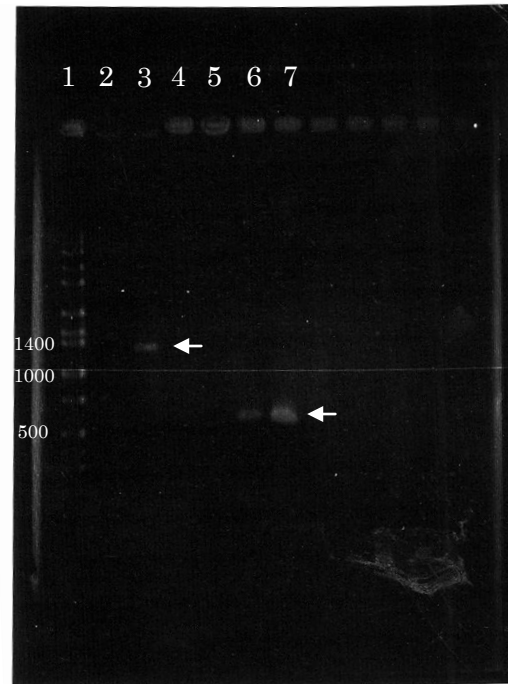
Results

A set of experiments were designed to elucidate whether any of the 177 genes identified by transcriptome analysis as potentially being regulated by DAF-19 (De Stasio, unpublished) are actually controlled by the DAF-19 transcription factor. Because DAF-19 is known to be expressed in the hypodermis and neuronal tissue of *C. elegans*, our first goal was to determine whether (1) any of the putative DAF-19 controlled genes show neuronal expression; and (2) whether expression of any of these genes is truly dependent on DAF-19A/B *in vivo*. To achieve these goals, we first created a recombinant plasmid that contained the putative control region of a desired gene followed by a GFP coding region. In this way, when the plasmid is incorporated into cells of the worm, any cell that has the capability of driving transcription of the desired gene will display GFP expression. Two of the 177 genes putatively controlled by DAF-19 are *F46G11.3* and *F58E2.3*. Neither genes have yet been characterized genetically or biochemically, but are known to be transcribed. A large-scale study of genes required for synapse structure (Sieburth et al., 2005) identified *F46G11.3* as a gene that has a vital role in the synaptic vesicle cycle and endocytosis, both of which are important in neurotransmission. *F58E2.3* was previously shown to encode for an F-box protein, a family of proteins that contains a specific motif that functions as the site of protein-protein interaction (Kipreos and Pagano, 2000).

Construction of Transcriptional Reporters

In order to clone the putative control regions of genes known as *F46G11.3* and *F58E2.3*, we employed PCR to amplify the region upstream of each coding region using wild-type worm (N2) DNA as the template. Each primer included a restriction enzyme site used for later cloning. PCR is a sensitive process in which an optimum reaction condition must be provided for amplification of DNA templates. Since each PCR primer pair has a different optimum annealing temperature, we first tested two annealing temperatures (58°C and 60°C) for the PCR reactions. These two parameters were chosen because they are close to the melting temperature of the four primers used (Materials and Methods, Table 1). We also modified the MgCl₂ concentration in the PCR reaction mix to determine ideal ion concentration. PCR done with a 60°C annealing temperature resulted in a smear from the *F58E2.3* template when using 4μM MgCl₂ (not shown); while reactions using the 58°C annealing temperature resulted in positive results for both amplicons (Figure 9).

Figure 9. DNA fragments from a PCR with annealing temperature of 58°C. The first lane shows a Hi-Lo DNA marker (Minn. Molecular) with DNA fragments of known size (indicated at left). Lane 3 includes DNA fragments generated using primers specific to the control region of *F46G11.3* (arrow); lanes 6 and 7 show the fragments generated using primers specific to the putative control region of *F58E2.3* (arrow).

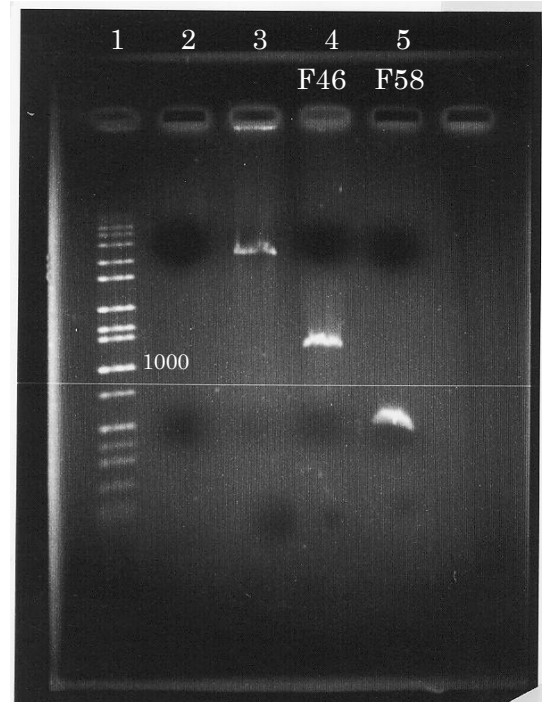


To amplify the putative control region of *F46G11.3*, 3 μ M MgCl₂ was the optimum ion concentration. The resulting amplicon size was between 1000 and 1400 base pairs (Figure 9, lane 3). Since the expected length of the PCR product was 1199 base pairs, the gel electrophoresis result indicated that the amplicon size met our expectation. The expected amplicon of *F58E2.3* was produced when 3 or 4 μ M MgCl₂ was added to the reaction mix. The resulting fragment size was about 500 base pairs, which again indicates an amplicon of the correct size (499 base pairs) was generated.

After determining the optimum annealing temperature, a 100 μ L PCR reaction was used to prepare a sufficient amount of PCR product for cloning. A QIAquick PCR Purification Kit[®] was used to purify the PCR products from other PCR reaction ingredients such as primers, and excess dNTPs. After purification, the amplicons were cut with restriction enzymes Sall and BamHI in order to produce sticky ends complementary to those of the

cloning vector, which was also cut by the same enzymes. After the restriction digestion, digested amplicons were again purified with the Purification Kit and visualized on a gel (Figure 10).

Figure 10. Purified PCR product after restriction digestion. The first lane contains the Hi-Lo DNA ladder; a fragment of 1000 base pairs is indicated. Lane 3 contains the digested linear vector pPD95.75. Lanes 4 and 5 shows the purified, correct size PCR amplicons of the putative control regions of *F46G11.3* and *F58E2.3*, respectively. Each is the expected length.



Ligation and transformation

If a competent *E. coli* cell contains a successfully ligated plasmid, the cell will express the ampicillin resistance gene from the plasmid and will form a colony on an agar plate containing ampicillin. Therefore, we expected to see a large number of colonies forming on the agar plate after transformation and overnight incubation. We observed over 50 colonies on the LB + ampicillin plate and, thus, our observation supported our expectation. As a negative control, we transformed linearized pPD95.75 vector into competent cells and plated them under the same conditions. The negative control should show little or no colony

formation, indicating that the cut pPD95.75 vector was unable to ligate due to incompatible sticky ends. The number of colonies formed in the negative control varied from 0 to 4 colonies per plate, which indicated that cells containing a ligated plasmid lacking an insert occurred at an extremely low rate.

Confirming the Presence of a Recombinant Plasmid through Lysis PCR

To determine whether a recombinant plasmid contained the appropriately sized insert, PCR using primers that flank the multiple cloning site (MCS) of the vector were used. When a PCR is carried out using the original pPD95.75 vector as the template DNA, the two primers will generate a band that is 347 base pairs in length. However, if the vector contains the insert in between the two restriction sites, the PCR product length will be 347 base pairs + the size of the amplicon. Therefore, we expected a 1540 base pair fragment to be formed when the control region of gene *F46G11.3* was cloned (347 base pair + 1193 base pair insert), and a 846 base pair fragment when *F58E2.3* was cloned (347 base pair + 499 base pair insert).

Our results confirmed that three out of twenty colonies tested contained the desired plasmid that incorporated the putative control region of *F46G11.3*. We also observed that three colonies contained a plasmid that included the putative control region of *F58E2.3* (Figure 11).

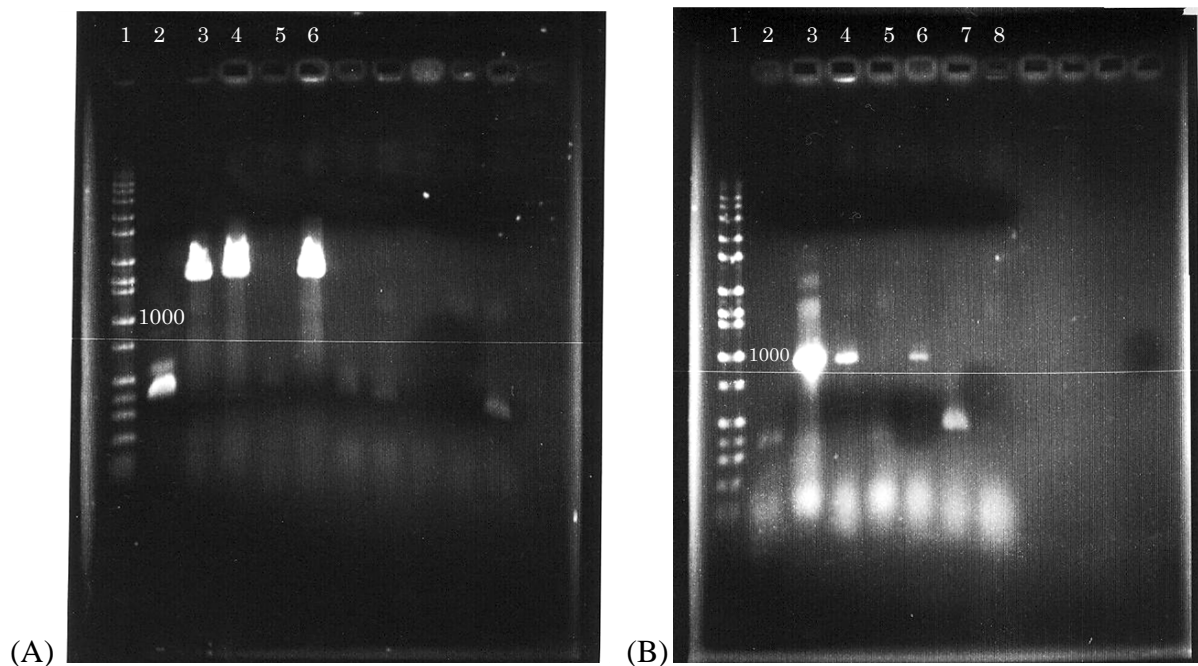


Figure 11. (A) Lysis PCR results for *F46G11.3* visualized on 1.0% agarose gel. Lane 1 contains the Hi-Lo DNA ladder. Lane 2 shows a fragment that is about 400 base pairs in length, indicating the insert is not present. Lanes 3, 4, and 6 contain fragments that are between 1550 and 2000 base pairs in length, indicating the insert is present. (B) Lysis PCR results for *F58E2.3* visualized on 1.0% agarose gel. Lane 1 contains the Hi-Lo DNA ladder. Lane 2 shows a fragment that indicates no insert is present. Lanes 3, 4, and 6 contain fragments that are slightly less than 1000 base pairs in length, indicating the insert is present. The smears at the bottom of each lane represent the residual primers used in PCR. Amplification was not observed in lanes 5 and 8.

Confirming Insert Orientation Using Restriction Mapping

After confirming the appropriate size of the cloned insert, restriction mapping was used to test whether the insert was in the desired orientation. This is crucial because a reversed insert results in no transcription, which means zero GFP expression in worms. With the expected fragment sizes known (Materials and Methods, Figure 5 & 6), two restriction enzymes, *Hind*III and *Mfe*I, were used to cut the plasmids that contained the putative control region of *F46G11.3* and *F58E2.3*, respectively. Restriction digestion of both plasmids

generated a band that is between 750 base pairs and 1000 base pairs in size (Figure 12). Both digestions should yield a band that is approximately 900 base pairs in size when the insert is in the correct orientation (Materials and Methods, Figure 5 and 6). Thus, these results suggest that the desired plasmids were created.

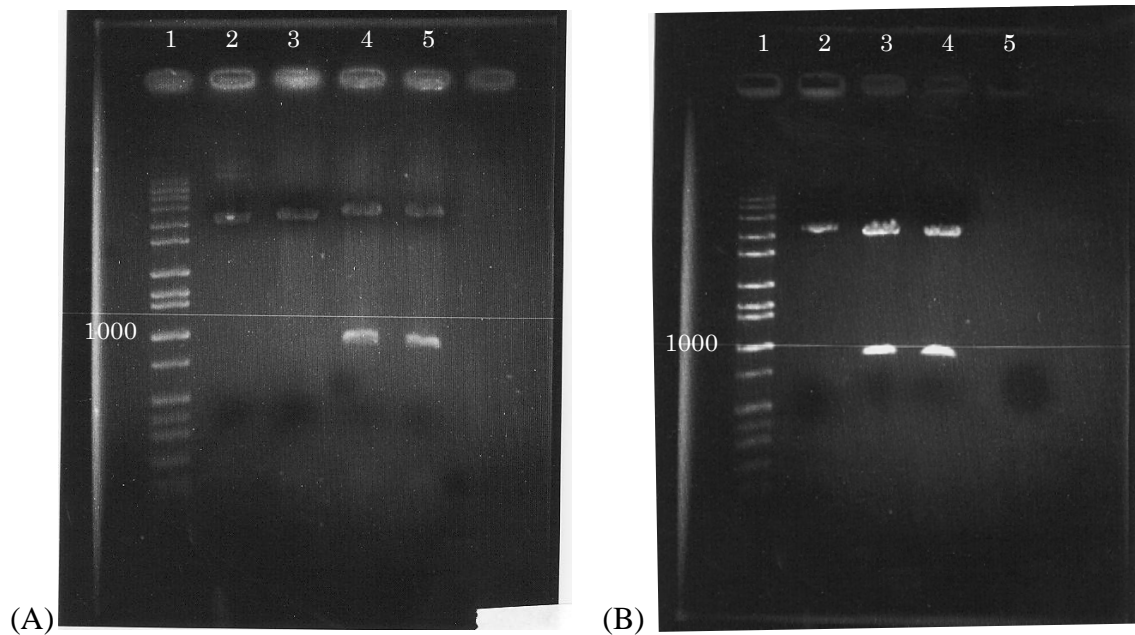


Figure 12. Restriction mapping visualized on 1.0% agarose gel. Lane 1 in both images contain Hi-Lo DNA ladder. (A) *F46G11.3*: lanes 4 and 5 contain a DNA fragment that is less than 1000 base pairs, indicating the insert was in the right direction. (B) *F58E2.3*: lanes 3 and 4 contain a DNA fragment that is less than 1000 base pairs, indicating the insert was in the right direction. Each lane contains a sample obtained from different colonies.

Confirmation of Cloned Sequences

Although restriction mapping and PCR-based measurement of the insert size suggest that the expected DNA was cloned in the proper orientation, we sought additional evidence that our cloning was precise. At this stage, we sent the sample plasmids to a DNA analysis facility to sequence portions of the plasmid. Two samples prepared from different colonies

were sequenced for each construction. In addition to our plasmids we also sent primers used to direct the sequencing reactions from each side of the inserted sequences. Sequencing can thus start at the restriction sites of the insert and it will be carried out from both the forward and the reverse directions.

Both sequencing results showed that the inserts consist of appropriate DNA sequences (Figure 13 and 14). From our results we chose to continue our experiments with sample 3 for the pPD95.75 + *F58E2.3* recombinant plasmid, and sample 15 for the pPD95.75 + *F46G11.3* plasmid. These plasmids were amplified and prepared to be injected into wild-type worms (genotype: *daf-12* *-/-*; *him-5* *-/-*).

Sall R.S.

```

pPD95_75_F58      AAGCTTGCATGCCTGCAGGTCGACatgccactgataacttgattcttcaatctcttctt
Sample_3_FWD      -----cttcat
Sample_5_FWD      -----ctcgtctt
                                     * ** *
                                     ↑  ↑  ↑

pPD95_75_F58      ttatctttcgttcccatatgactcatalcagcaatatgggatatgtcataaatttggtca
Sample_3_FWD      ttatctttcgttcccatatgactcatalcagcaatatgggatatgtcataaatttggtca
Sample_5_FWD      ttatctttcgttcccatatgactcatalcagcaatatgggatatgtcataaatttggtca
*****

pPD95_75_F58      caatacttcttgctaaagtgtttggccccaccattcggacagtttctccgccctttttc
Sample_3_FWD      caatacttcttgctaaagtgtttggccccaccattcggacagtttctccgccctttttc
Sample_5_FWD      caatacttcttgctaaagtgtttggccccaccattcggacagtttctccgccctttttc
*****
                                     Sand-box

pPD95_75_F58      ttatgtattatagcatctccggtggactgcttttcatctttttacaataaagttctac
Sample_3_FWD      ttatgtattatagcatctccggtggactgcttttcatctttttacaataaagttctac
Sample_5_FWD      ttatgtattatagcatctccggtggactgcttttcatctttttacaataaagttctac
*****

pPD95_75_F58      tacctgacccaaagccttctttaatagtttcacttcaacattttgaattttttaaat
Sample_3_FWD      tacctgacccaaagccttctttaatagtttcacttcaacattttgaattttttaaat
Sample_5_FWD      tacctgacccaaagccttctttaatagtttcacttcaacattttgaattttttaaat
*****

pPD95_75_F58      tttaccgcattctggcaaatttcgatgattttaacaataagcaacttatataatgataa
Sample_3_FWD      tttaccgcattctggcaaatttcgatgattttaacaataagcaacttatataatgataa
Sample_5_FWD      tttaccgcattctggcaaatttcgatgattttaacaataagcaacttatataatgataa
*****

pPD95_75_F58      aactgagttttttctgtagctgtggtttcgaaaaagtttagtttttatccctattact
Sample_3_FWD      aactgagttttttctgtagctgtggtttcgaaaaagtttagtttttatccctattact
Sample_5_FWD      aactgagttttttctgtagctgtggtttcgaaaaagtttagtttttatccctattact
*****
                                     ↑

pPD95_75_F58      aatctccggcgctacccaaaaattaatttctcacaaccaaatctgtatttacaattgca
Sample_3_FWD      aatctccggcgctacccaaaaattaatttctcacaaccaaatctgtatttacaattgca
Sample_5_FWD      aatctccggcgctacccaaaaattaatttctcacaaccaaatctgtatttacaattgca
*****

pPD95_75_F58      acgttttcaatcactacctaatttcaaaaaatttcagattcca-GGATCCCGGGATTGG
Sample_3_FWD      acgttttcaatcactacctaatttcaaaaaatttcagattccaaggatccccgggattgg
Sample_5_FWD      acgttttcaatcactacctaatttcaaaaaatttcagattccaaggatccccgggattgg
*****
                                     BamHI R.S.

```

Figure 13. Sequencing result for pPD95.75+F58E2.3 recombinant plasmid. The line “pPD95_75_F58” shows the expected sequence. 469 out of 499 base pairs of the insert were sequenced. The two colored underlines show the two restriction sites and the asterisk marks under each alignment indicate that the sequence results match with the expected sequence. Sand-box motif was included in the insert and is indicated by the brown-colored underline. Any potentially erroneous sequence is indicated by an arrow. Gene expression studies were undertaken using sample 3.

A

SaII R.S.

```
pPD95_75      AAGCTTGCATGCCTGCAGGTCGACaagttgaagataatcgaaagactacagaaaaaacgc
Sample13FWD    -----
Sample15FWD    -----

pPD95_75      aaaaaagataaaagtaatgttacataaagtttccggtaaaactggttgagcgaaaaatc
Sample13FWD    -----gtaatgttacataaagtttccggtaaaactggttgagcgaaaaatc
Sample15FWD    -----gtaatgttacataaagtttccggtaaaactggttgagcgaaaaatc
                *****

pPD95_75      acaaaagatgtttgttttacgagagctgcaactggttccacgttggttacatggtgtgag
Sample13FWD    acaaaagatgtttgttttacgagagctgcaactggttccacgttggttacatggtgtgag
Sample15FWD    acaaaagatgtttgttttacgagagctgcaactggttccacgttggttacatggtgtgag
                *****

pPD95_75      gttagcaaatgCGTcagcagccaccacttggcgcgtttcgtgtgCGTtccccacacccaaa
Sample13FWD    gttagcaaatgCGTcagcagccaccacttggcgcgtttcgtgtgCGTtccccacacccaaa
Sample15FWD    gttagcaaatgCGTcagcagccaccacttggcgcgtttcgtgtgCGTtccccacacccaaa
                *****

pPD95_75      aaagggcgcatgtatgcaacatttgcgttttttcattgattttcctaatttattacaac
Sample13FWD    aaagggcgcatgtatgcaacatttgcgttttttcattgattttcctaatttattacaac
Sample15FWD    aaagggcgcatgtatgcaacatttgcgttttttcattgattttcctaatttattacaac
                *****↑*****

pPD95_75      gtgaatttatttttaaacagaaatcagtacactaacattcacagtcttgcaataaatgct
Sample13FWD    gtgaatttatttttaaacagaaatcagtacactaacattcacagtcttgcaataaatgct
Sample15FWD    gtgaatttatttttaaacagaaatcantacactaacattcacagtcttgcaataaatgct
                *****↑*****

pPD95_75      gcactgacaaacaataagtttgcactgaaaaaaatcgtttttgcattgcaaaaaatt
Sample13FWD    gcactgacaaacaataagtttgcactgaaaaaaatcgtttttgcattgcaaaaaatt
Sample15FWD    gcactgacaaacaataagtttgcactgaaaaaaatcgtttttgcattgcaaaaaatt
                *****

pPD95_75      ttcagtttgctgtaaaaattaaatgaaacctttaaccgttttcttataaaaagaaagttt
Sample13FWD    ttcagtttgctgtaaaaattaaatgaaacctttaaccgttttc-----
Sample15FWD    ttcag-----
                *****
```

B

BamHI R.S.

```
pPD95.75+F46_RVS      CCTCctgaaaatggttctatggtatgtagtatcattcgaaacatacCTTTGGGTCCTTTG
Sample13RVS            -----
Sample15RVS            -----

pPD95.75+F46_RVS      GCCAATCCCAGGGATCCcttggaaattgcagatttatagtagaatgagtcattaaaaat
Sample13RVS            -----t
Sample15RVS            -----

pPD95.75+F46_RVS      aataaaatgtagatggtttatgagcggtagcatcaatatgaaaaaggaaaaaaatatatga
Sample13RVS            aataaaatgtagatggtttatgagcggtagcatcaatatgaaaaaggaaaaaaatatatga
Sample15RVS            -----tggtagatggtttatgagcggtagcatcaatatgaaaaaggaaaaaaatatatga
                        *****

pPD95.75+F46_RVS      aaaagtataagaccaggataaaaaatgtacaatttacaatttaattcaaaggctcaggt
Sample13RVS            aaaagtataagaccaggataaaaaatgtacaatttacaatttaattcaaaggctcaggt
Sample15RVS            aaaagtataagaccaggataaaaaatgtacaatttacaatttaattcaaaggctcaggt
                        *****

pPD95.75+F46_RVS      ggtatctataacataatTTTTTcAAAATaagaattggacttatgaagaacattcttt
Sample13RVS            ggtatctataacataatTTTTTcAAAATaagaattggacttatgaagaacattcttt
Sample15RVS            ggtatctataacataatTTTTTcAAAATaagaattggacttatgaagaacattcttt
                        *****

pPD95.75+F46_RVS      ctcaattacatatcttgcttctcaaaaattattctaaatttgcaaatgatgTTTTaca
Sample13RVS            ctcaattacatatcttgcttctcaaaaattattctaaatttgcaaatgatgTTTTaca
Sample15RVS            ctcaattacatatcttgcttctcaaaaattattctaaatttgcaaatgatgTTTTaca
                        *****

pPD95.75+F46_RVS      cgTTTcaatcaataagaaatcgcttaagccttctatctTTTaaagcttcatctgcatag
Sample13RVS            cgTTTcaatcaataagaaatcgcttaagccttctatctTTTaaagcttcatctgcatag
Sample15RVS            cgTTTcaatcaataagaaatcgcttaagccttctatctTTTaaagcttcatctgcatag
                        *****

pPD95.75+F46_RVS      aatataaagtacacatttccattagcgaagactaacggcaagaaattcacattattgt
Sample13RVS            aatataaagtacacatttccattagcgaagactaacggcaagaaattcacattattgt
Sample15RVS            aatataaagtacacatttccattagcgaagactaacggcaagaaattcacattattgt
                        *****

pPD95.75+F46_RVS      caggaataggatagattccatctcgcatctactgaaatgactctcgtggtcgaaaaatc
Sample13RVS            caggaataggatagattccatctcgcatctactgaaatgactctcgtggtcgaaaaatc
Sample15RVS            caggaataggatagattccatctcgcatctactgaaatgactctcgtggtcgaaaaatc
                        *****

pPD95.75+F46_RVS      ctgagagaagcggagtattgggcagcagcagcgaagtggggtgtcggttggcaacgga
Sample13RVS            ctgagagaagcggagtattgggcagcagcagc-----
Sample15RVS            ctgagagaagcggagtattgggcagcagcagcgaagtggggtgtcggttggcaacgga
                        *****
```

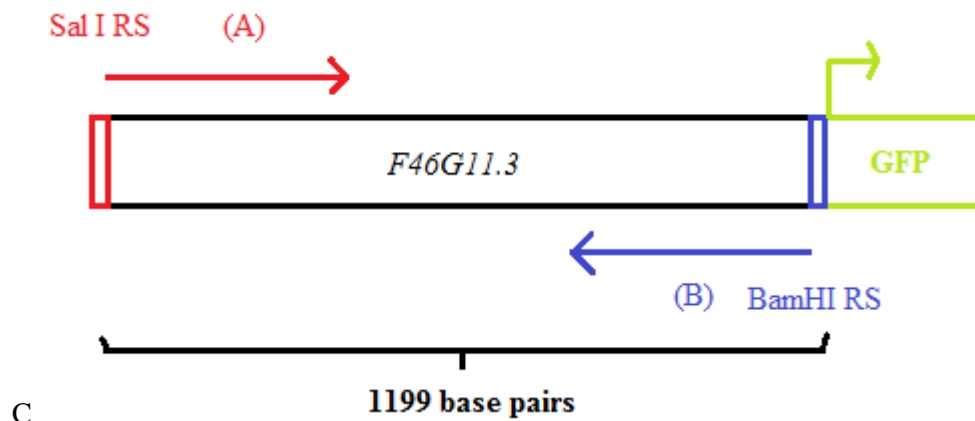


Figure 14. Sequencing result for pPD95.75+F46G11.3 recombinant plasmid using forward primer (A) and reverse primer (B). Note that due to the large amplicon size, accurate sequencing did not cover the entire insert. The line “pPD95.75+F46” and “pPD95.75+F46_RVS” show the expected sequences of the recombinant plasmid. The two colored underlines show the two restriction sites and the asterisk marks under each alignment indicate that the sequence results match with the expected sequence. Any potentially erroneous sequence is indicated by an arrow. (C) A schematic diagram of sequencing. The forward primer (A) was used to sequence from the SalI RS, while the reverse primer (B) was used to sequence the opposite strand from the BamHI RS. Green colored region indicates the translational start site and the coding region of GFP. Approximately 800 out of 1193 base pairs of the insert were sequenced. Sample 15 was used for subsequent gene expression studies.

Production of Transgenic Worms

After confirming that we obtained the desired recombinant plasmids, we mixed the constructs with a co-injection marker (*elt-2::mCherry*) and carrier DNA (pBR322) and microinjected the mix into the gonad of a young adult hermaphrodite wild-type worm (*daf-19* +/+; *daf-12* -/-; *him-5* -/-). The microinjection enables the offspring to take up the DNA mixture and form an extrachromosomal array that contains both the recombinant plasmid and the co-injection marker.

At approximately 48 to 72 hours after microinjection, the worms were observed

under the fluorescent microscope to see whether transgenic offspring were produced. We successfully generated transgenic worm strains and we selected transgenic worms based on their red fluorescence due to the presence of *elt-2::mCherry* (Figure 15).

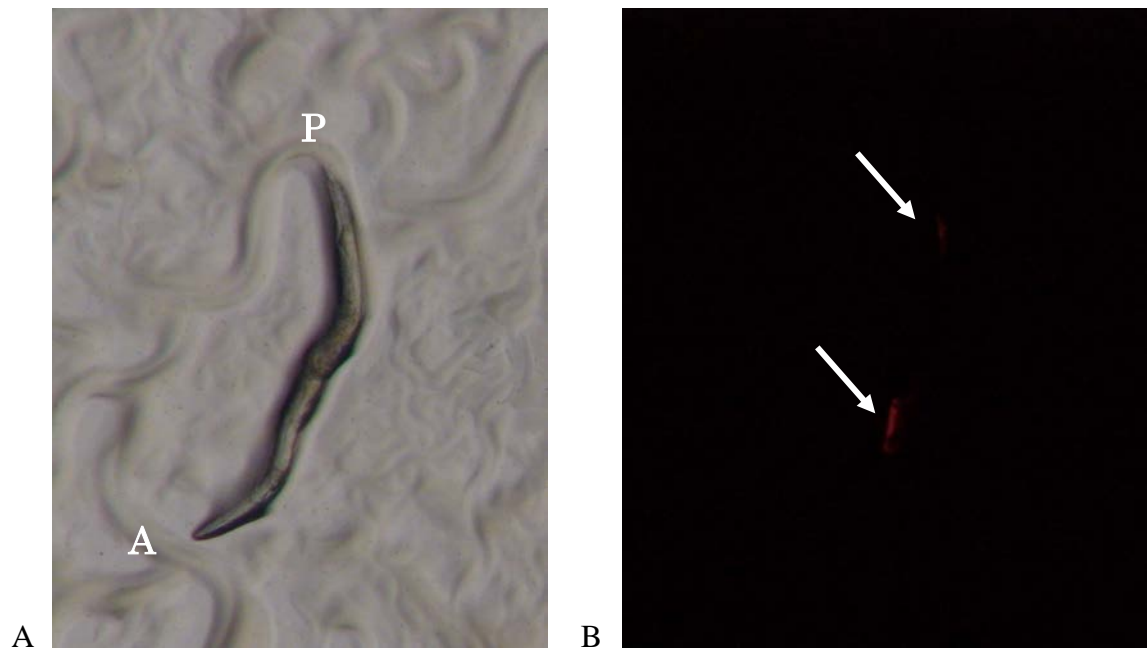


Figure 15. A transgenic LU600 (*daf-19* *+/+*) young adult hermaphrodite observed under the fluorescent microscope. (A) DIC light microscopy image of the worm. A = anterior; P = posterior. (B) *elt-2::mCherry* expressed in the intestine of the same worm as in (A). Expression is indicated by the arrow.

Supplemental Information 3: Assigning Strain Names

Each worm strain is assigned a name that is specific to the lab to which it belongs. Strain names are composed of two parts: (1) capital letters that represent the owner or the institution; The De Stasio lab uses “LU”, which stands for Lawrence University; (2) numerical codes that represent the order of strain creation.

e.g. LU491, meaning the 491th *C. elegans* strain created at Lawrence University.

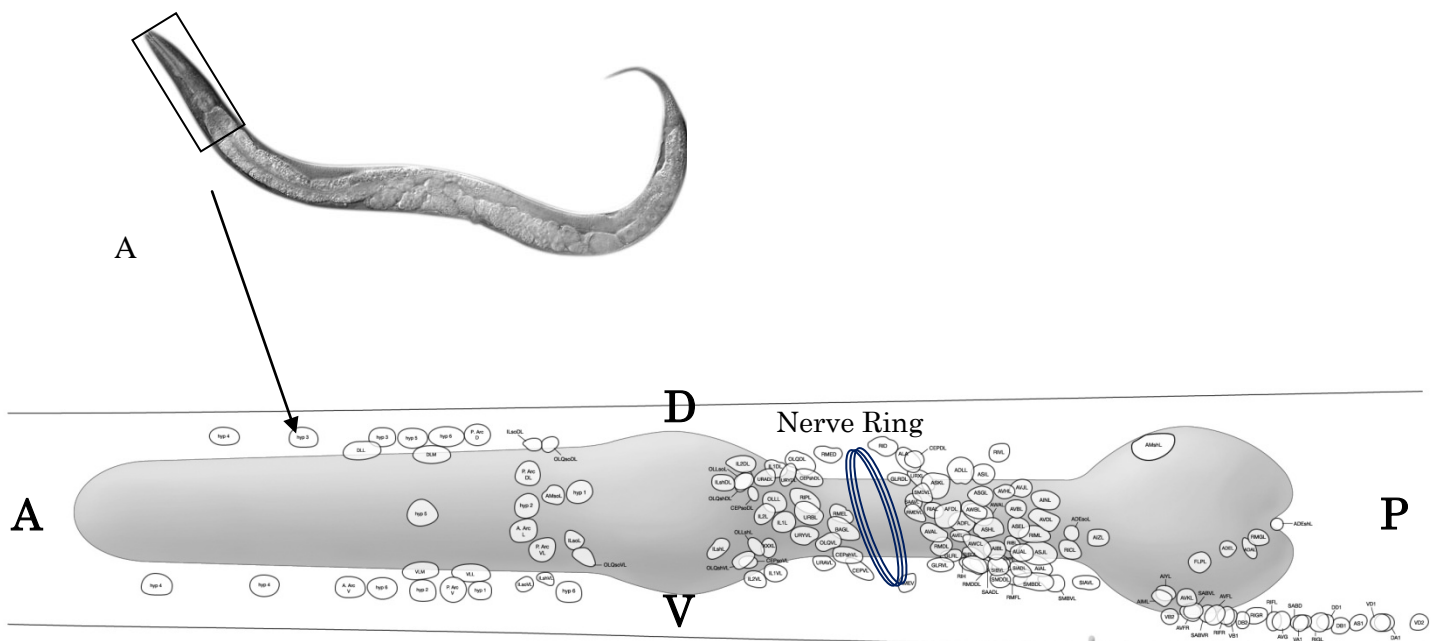
Strains that appear in this thesis include:

LU479: *F57B10.9* in *daf-19* $-/-$; *daf-12* $-/-$ **mutant** animal

LU491: *F57B10.9* in *daf-19* $+/+$; *daf-12* $-/-$ wild-type animal

LU497: *F58E2.3* in *daf-19* $+/+$; *daf-12* $-/-$ wild-type animal

LU600: *F46G11.3* in *daf-19* $+/+$; *daf-12* $-/-$ wild-type animal



B

Figure 16. (A) A DIC light microscopy image of a young adult hermaphrodite. Image from Universiteit Utrecht: <http://www.testweb.science.uu.nl>. (B) A schematic image of the *C. elegans* head neurons. The gray area represents the pharynx and the circles show where neuronal cell bodies are located. The nerve ring, a structure that consists of concentrated nervous tissue is located between the two terminal bulbs. A = anterior; P = posterior; D = dorsal; V = ventral.

Assessing GFP Expression via Confocal Imaging

In order to determine which cells and tissues express the target gene in *C. elegans*, we observed the worm strains LU479, LU491, LU497, and LU600 using confocal microscopy. All images displayed in this study are standardized as follows unless otherwise noted: the target worm's anterior region is located on the image's left side and the worm's dorsal side is aligned with the image's upper edge. With high magnification images that captured detail in the head, vulval, or tail regions, a magnification of 800X was used. With low magnification pictures that captured the worm's whole body, a magnification of 200X was used. Zoom was set at 2.0 and image capture at 1024×256 was used (Materials and Methods, Table 2). Fluorescence images displayed in this thesis are maximum projections of a Z-stack, unless otherwise noted.

F58E2.3::gfp Expression

The GFP expression pattern was characterized for the strain LU497 (*daf-19* +/+; *daf-12* -/-) in hermaphrodite animals; expression was predominantly intestinal. High magnification (800X) images were acquired and we observed scattered GFP expression in worm intestines (n = 10) especially in or near the gut granules (Figure 17). No GFP expression was found in neuronal cells. The images were compared with a negative control (*daf-19* +/+; *daf-12* -/-; *him-5* -/-) that contained no extrachromosomal array and shows no

GFP expression in the anterior region of the worm (Figure 18). Whether the expression pattern was due to GFP or due to auto-fluorescence was determined by this comparison.

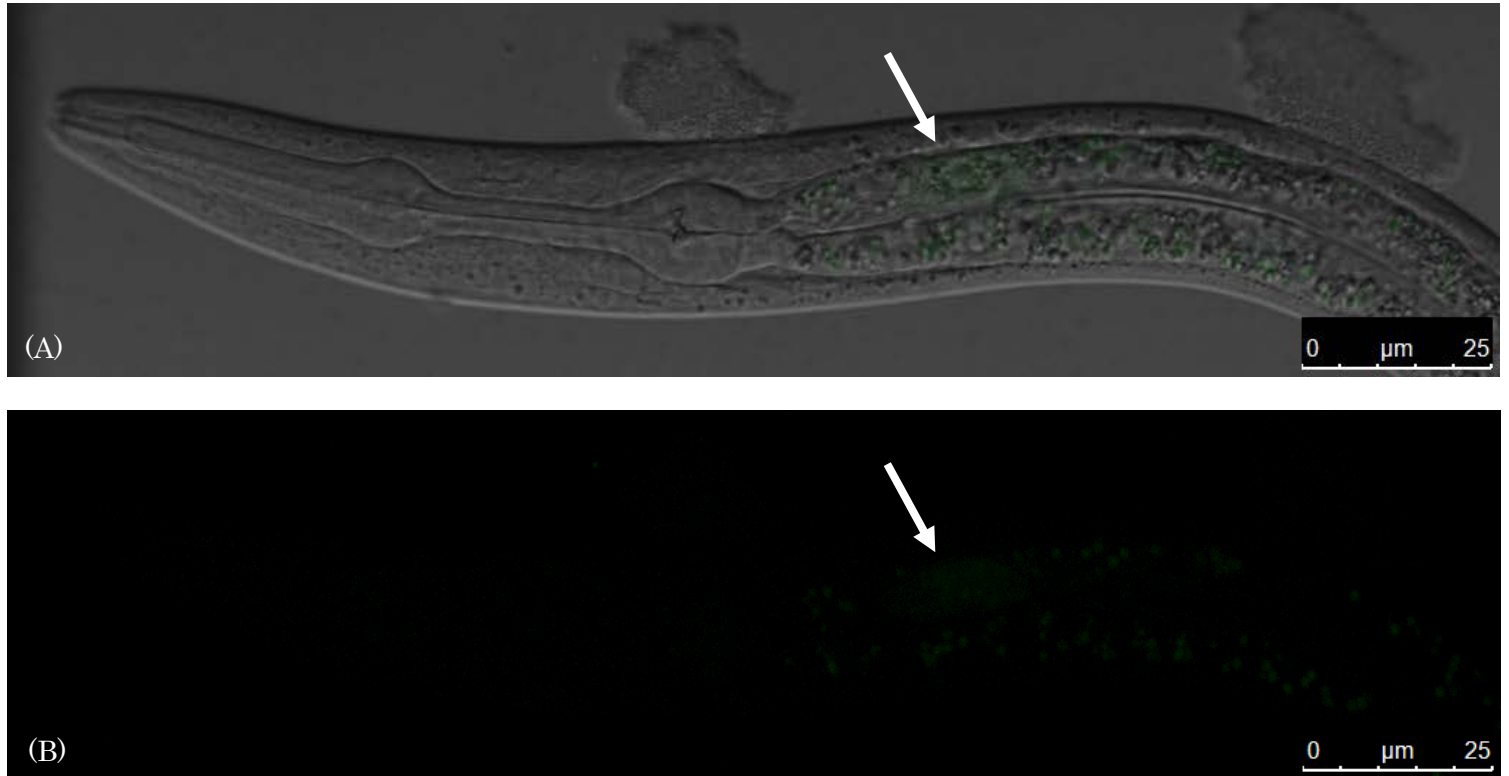


Figure 17. Expression of *F58E2.3::gfp*. Representative confocal images of the anterior of a *daf-19* *+/+*; *daf-12* *-/-* L3 hermaphrodite expressing *F58E2.3::gfp*. (A) Superimposed images of GFP expression and a DIC light microscopy. GFP expression is indicated by the arrow and is located in the intestine. (B) A maximum-projection GFP expression image acquired from the same worm as in (A).

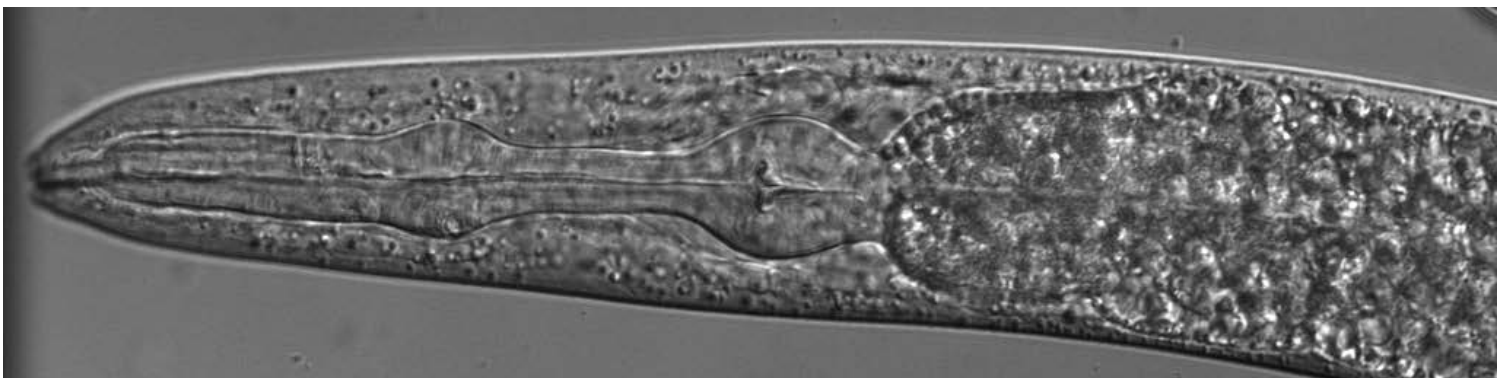


Figure 18. Superimposed images of fluorescence and a DIC light microscopy image, taken from a non-transgenic *daf-19* *+/+*; *daf-12* *-/-*, L3 hermaphrodite. Note that there is no GFP expression in the anterior region of the animal.

F57B10.9::gfp Expression

F57B10.9::gfp expression in *daf-19* *+/+* wild type (LU491) worms was compared to that of *daf-19* *-/-* mutants (LU479) using confocal imaging. Both high (800X) and low (200X) magnification images were acquired. We observed GFP expression in both ciliated and non-ciliated neurons (Figure 19 B&C, Figure 20) around the nerve ring that is located in between the anterior and posterior bulbs of the pharynx. This observation applied to both *daf-19* *+/+* animals (LU491) and *daf-19* *-/-* animals. However, we observed that *daf-19* *+/+* animals showed stronger GFP expression in sensory cilia, the terminal structures of dendrites emanating from the cell bodies are located anterior to the nerve ring. On the other hand, while only 6% of the *daf-19* *+/+* animals showed GFP expression in their body wall muscle, more than half of the *daf-19* *-/-* animals showed GFP expression at the same location (Figure 19C, Table 4).

Age-related GFP expression patterns were also compared between the two strains. GFP expression in the ventral and dorsal nerve cords appear to decline as the animal ages (80% in young larvae versus 25% in adults, Table 4). However, this pattern was not found in *daf-19* *+/+* animals, for which the percentage of worms showing GFP expression did not differ between the young larvae and the adults (40% versus 33%, relatively).



Figure 19. Expression of *F57B10.9::gfp*. Representative high magnification (800X) confocal images of the anterior of LU479 (*daf-19 -/-*) and LU491 (*daf-19 +/+*) hermaphrodite expressing *F57B10.9::gfp*. (A) Superimposed images of GFP expression and a DIC light microscopy acquired from LU491 strain L4 hermaphrodite. (B) A maximum-projection GFP expression image acquired from the same worm as in (A). (C) A maximum-projection GFP expression image acquired from LU479 strain L3/L4 hermaphrodite. The presence of *daf-19* genotype is indicated on each image. The circles indicate the difference in GFP expression in the amphid neurons of the two strains.

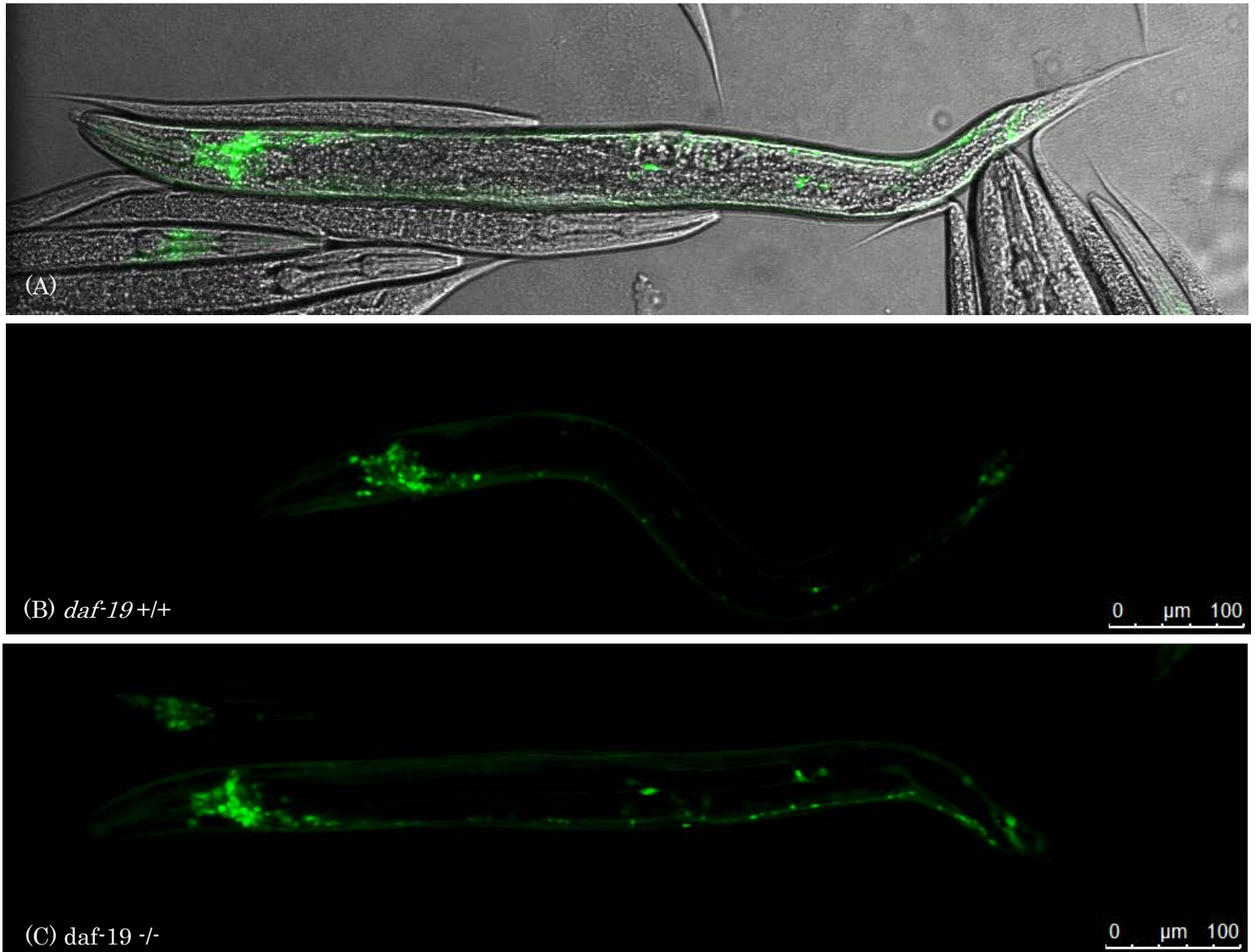


Figure 20. Representative low magnification (200X) confocal images of the whole body of LU479 (*daf-19 -/-*) and LU491 (*daf-19 +/+*) young adult hermaphrodites. (A) Superimposed image of GFP expression and a DIC light microscopy acquired from LU479 strain adult hermaphrodite. (B) A maximum-projection GFP expression image acquired from LU491 strain adult hermaphrodite. (C) A max-projected GFP expression image acquired from LU479 strain adult hermaphrodite. The *daf-19* genotype is indicated on each image.

GFP expression location	Strain	L1/L2	L3	L4	Adult	Total % expressed
	LU479	n = 5	n = 2	n = 3	n = 8	n = 18
	LU491	n = 5	n = 2	n = 3	n = 6	n = 16
Neurons						
Head Neurons	LU479	5 (100%)	2 (100%)	3 (100%)	8 (100%)	100%
	LU491	5 (100%)	2 (100%)	3 (100%)	6 (100%)	100%
Nerve Cords	LU479	4 (80%)	0	0	2 (25%)	33%
	LU491	2 (40%)	2 (100%)	2 (66%)	2 (33%)	50%
Body Neurons	LU479	3 (60%)	0	0	6 (75%)	50%
	LU491	3 (60%)	1 (50%)	3 (100%)	5 (83%)	75%
Tail Neurons	LU479	-	-	-	-	0%
	LU491	-	-	-	-	0%
Muscles						
Pharyngeal muscle	LU479	-	-	-	-	0%
	LU491	-	-	-	-	0%
Body Wall Muscle	LU479	2 (40%)	1 (50%)	3 (100%)	4 (50%)	56%
	LU491	1 (20%)	0	0	0	6%
Vulval Muscle	LU479	-	-	-	-	0%
	LU491	-	-	-	-	0%
Other						
Intestine	LU479	-	-	-	-	0%
	LU491	-	-	-	-	0%
Hypodermis	LU479	5 (100%)	2 (100%)	3 (100%)	7 (88%)	94%
	LU491	5 (100%)	1 (50%)	3 (100%)	6 (100%)	88%
Reproductive Tissue	LU479	-	-	-	-	0%
	LU491	-	-	-	-	0%

Table 4. Summary of tissue-specific expression of *F57B10.9::gfp* in both LU491 (*daf-19* +/+) and LU479 (*daf-19* -/-) hermaphrodites. Each column shows the number and the percentage of observed worms expressing GFP in the indicated tissues. Male worms were extremely rare and, thus, not assessed. Notable differences in GFP expression between the two strains are indicated in the red box. LU479 (*daf-19* -/-): n= 18; LU491 (*daf-19* +/+): n= 16.

F46G11.3::gfp Expression

F46G11.3::gfp expression in the strain LU600 (*daf-19* +/+; *daf-12* -/-; *him-5* -/-) was characterized using a higher magnification setting (800X) with an emphasis on the anterior region of the worm. *F46G11.3::gfp* was expressed in two pairs of likely neuronal cell bodies near the posterior terminal bulb of the pharynx (Figure 23B&C). By comparing the location of these cell bodies with the known location of neurons that take up the lipophilic dye, DiI (Hall and Russell, 1991; Figure 21A), we were able to tentatively identify the GFP-positive neurons (Figure 21B, Figure 23). We hypothesize that the cell body pair that is closer to the terminal bulb is interneuron pair either RIC, AIY, or AIM based on their location relative to the posterior bulb of the pharynx (Figure 22). An interneuron is a neuron that forms a connection between other neurons. These interneurons extend their axons between the two cell bodies and form a 'relay station' between two or more neuronal cells. Interneurons provide variety in the type of neurotransmitters a neuronal circuit can transmit and thus more excitatory and inhibitory switches are created within the circuit.

RIC is associated with the neurotransmitter octopamine (Alkema et al., 2005), an invertebrate equivalent of epinephrine. AIY is associated with the signaling of acetylcholine, a vital neurotransmitter that is found widely in the animal kingdom (Li and Kim, 2008). AIM is associated with neurotransmitter serotonin, which is known to be involved in smooth muscle regulation in humans (Li and Kim, 2008). The posterior cell body pair may be either

interneuron RIF or RIG, based on the ring-like structure made by the axons of these neurons

(Figure 22).

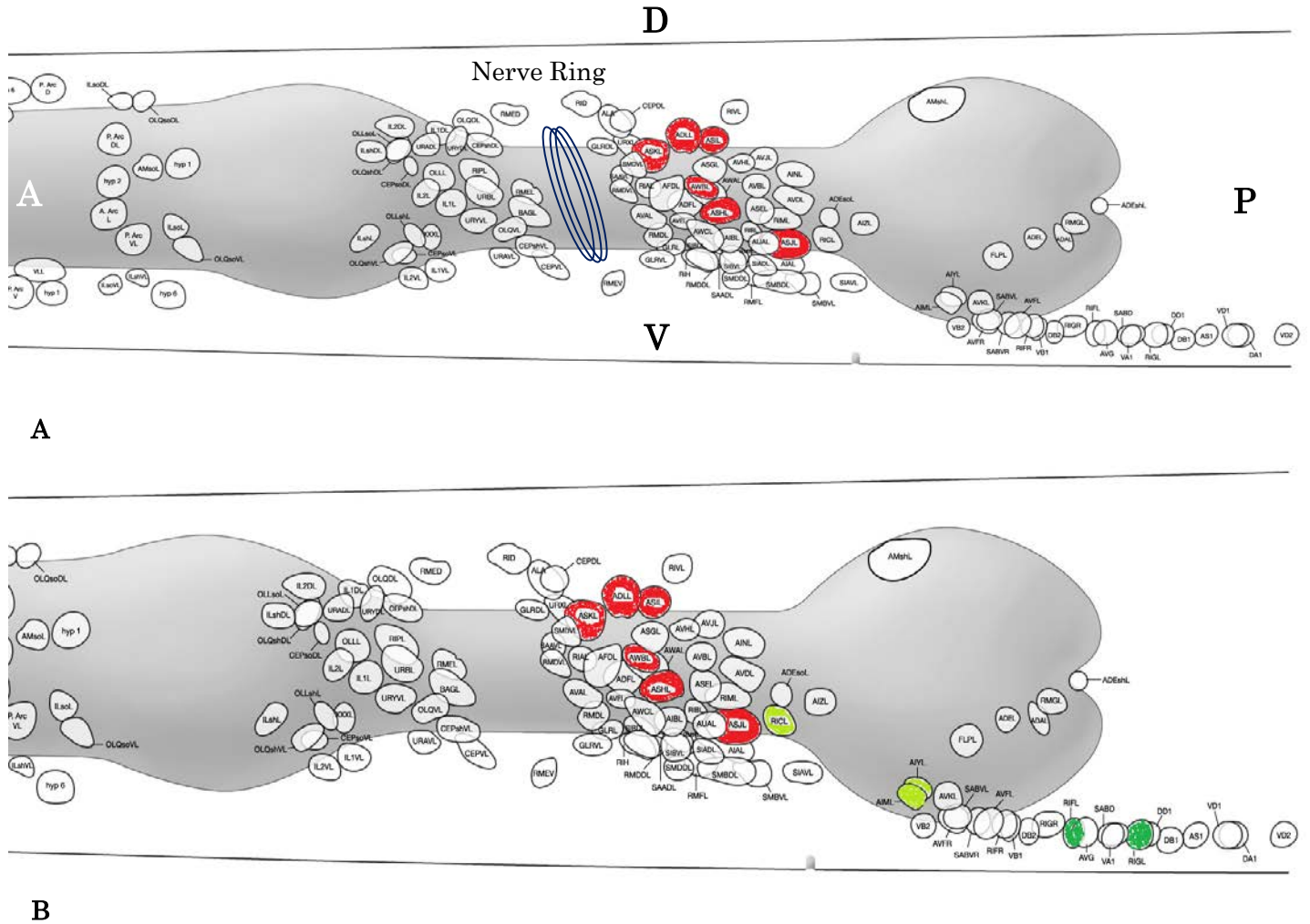


Figure 21. Schematic maps of the neurons that are stained by lipophilic dye DiI. (A) In hermaphrodite animals, there six neurons affected by DiI: ASK, AWB, ADL, ASI, ASH, and ASJ. All of them are colored in red. The orientation of the worm is also indicated with the letter A, P, V, and D: A = Anterior; P = Posterior; V = Ventral; D = Dorsal. (B) Estimated GFP expression pattern in interneuron RIC, AIY, AIM, (three colored in light green) RIF, and RIG (two colored in darker green).

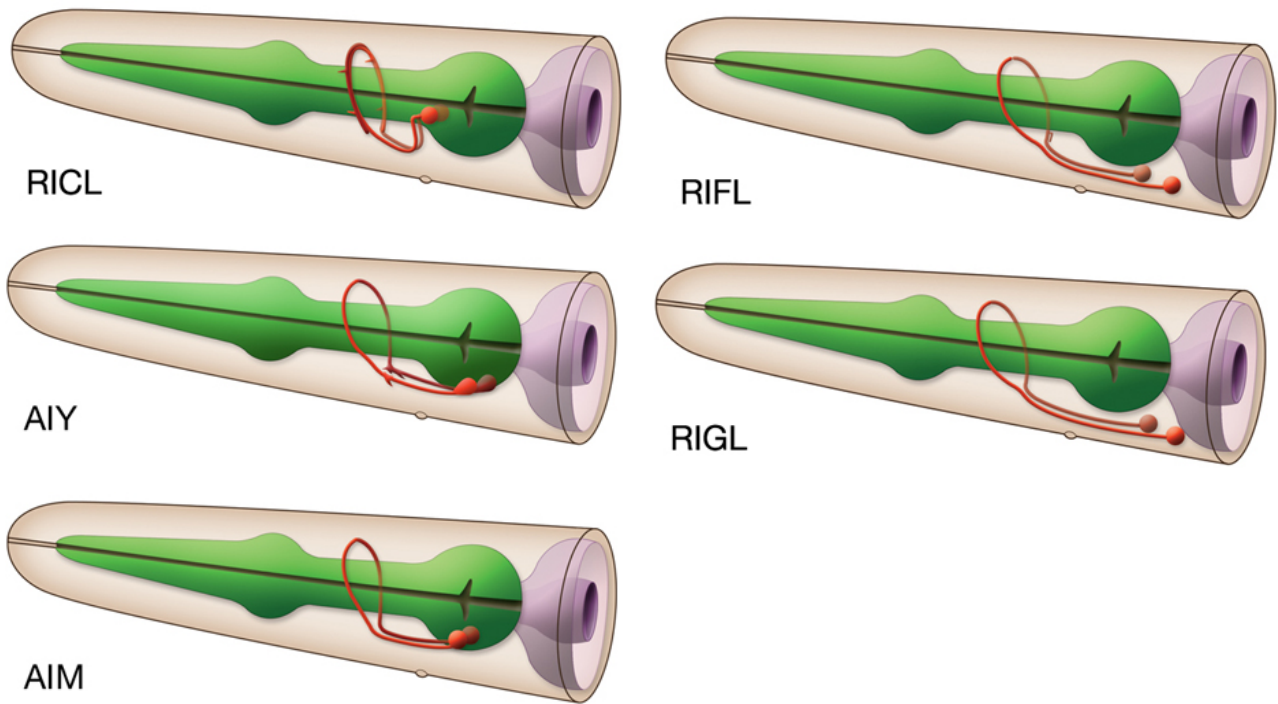


Figure 22. A schematic diagram of interneuron RIC, AIY, AIM (left), RIF, and RIG (right). The neuron cell body and axon are colored in red. The green colored region shows the anterior and posterior bulbs of the pharynx. L = Left. Images taken from WormAtlas (<http://www.wormatlas.org>).

The presence of the cell body pairs was constant in worms of each age group and we observed roughly half of the worms expressing GFP in their nerve cords, body neurons (Figure 24) and intestine (Table 5). All of the young larvae (L1/L2) expressed GFP in their intestine while only 33% showed intestinal GFP expression in both L4 and adult worms. Thus, intestinal GFP expression may be age-related. GFP expression in other tissues did not appear to be correlated with the age of the animal.

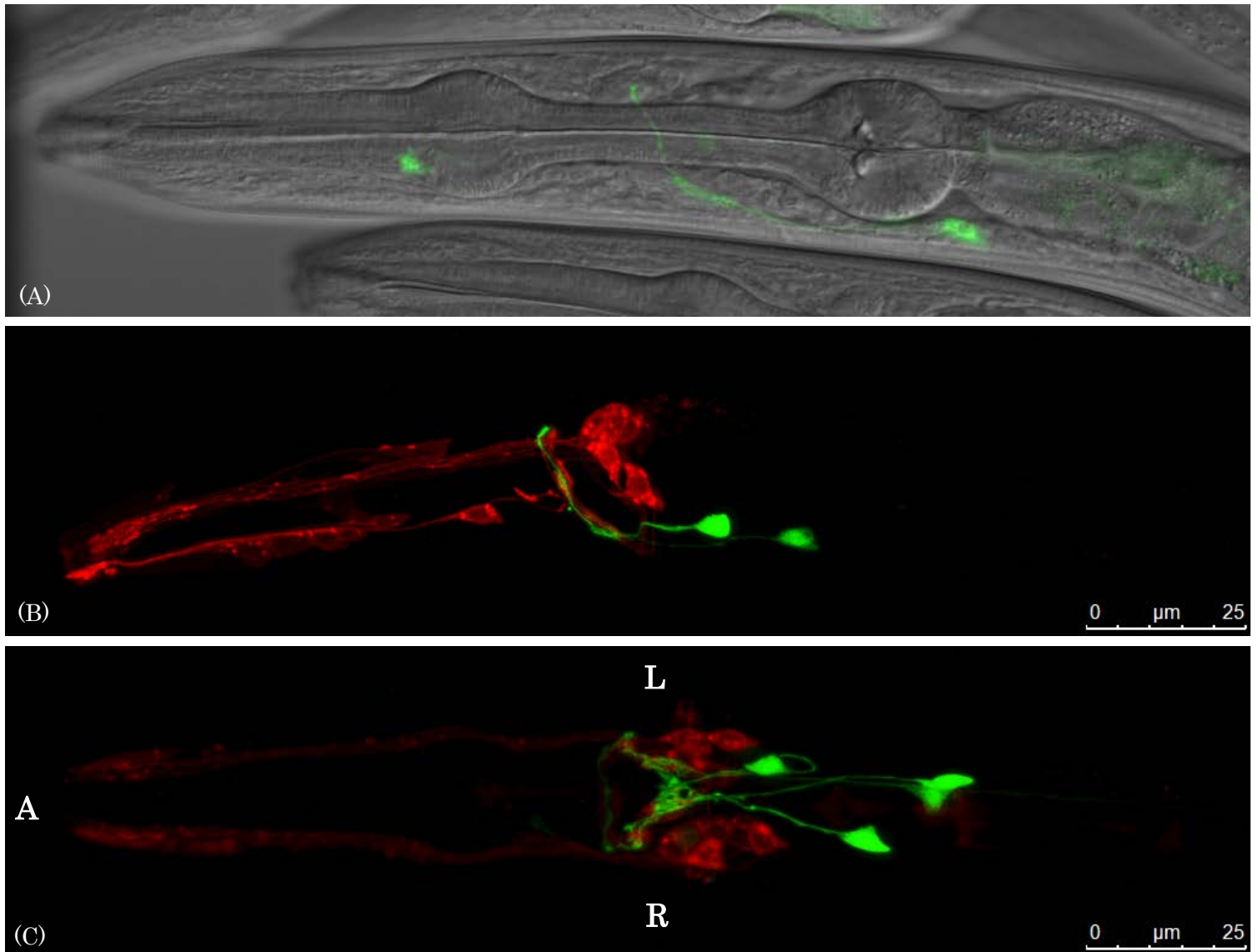


Figure 23. Expression of *F46G11.3::gfp*. Representative high magnification (800X) confocal image of the anterior of a LU600 (*daf-19* *+/+*; *daf-12* *-/-*) hermaphrodite. (A) Superimposed images of GFP expression and a DIC microscopy acquired from a LU600 young adult hermaphrodite. (B) A superimposed image of maximum-projection GFP expression and DiI fluorescence acquired from a young adult hermaphrodite. (C) Superimposed images of maximum-projection GFP expression and DiI fluorescence showing the view from the **ventral** side of a young adult hermaphrodite. The orientation of the specimen is indicated as following: A = Anterior; L = Left; R = Right.



Figure 24. Expression of *F46G11.3::gfp* near the vulval region of a LU600 (*daf-19* +/+; *daf-12* -/-) young adult hermaphrodite. GFP expression (indicated by the arrow) and a DIC microscopy image are superimposed. The orientation of the worm is consistent with the standard image display: Left side of the image shows the anterior of the worm; the top of the image shows the dorsal side of the worm.

GFP expression location	L1/L2 (n = 5)	L3 (n = 1)	L4 (n = 3)	Adult (n = 3)	Male (n = 2)	Total (n = 14)
Neurons						
Head Neurons	5 (100%)	1 (100%)	3 (100%)	3 (100%)	2 (100%)	100%
Nerve Cords	2 (40%)	1 (100%)	2 (67%)	1 (33%)	0	43%
Body Neurons	2 (40%)	1 (100%)	1 (33%)	1 (33%)	N/A	42%
Tail Neurons	-	-	-	-	-	0%
Muscle						
Pharyngeal muscle	-	-	-	-	-	0%
Body Wall Muscle	-	-	-	-	-	0%
Vulval Muscle	-	-	-	-	N/A	0%
Other						
Intestine	5 (100%)	0	1 (33%)	1 (33%)	1 (50%)	57%
Hypodermis	4 (80%)	N/A	N/A	N/A	1 (50%)	36%
Reproductive Tissue	-	-	-	-	-	0%

Table 5. Summary of tissue-specific expression of *F46G11.3::gfp* (LU600, *daf-19* +/+). Each column shows the number of worms expressing GFP in the indicated tissue. N/A indicates that the image of the noted region was not acquired. Hyphens indicate no GFP expression was observed even though the image the noted region was acquired.

Discussion

The current study was conducted to unravel the following questions:

1. *Are the genes of choice expressed in neuronal tissue?*
2. *Is the gene expression DAF-19 dependent?*
3. *Is the presence of the Sand-box motif associated with neuronal gene expression?*

Previous work by Senti and Swoboda (2008) suggests that the DAF-19A/B isoform is expressed in non-ciliated neurons, hypodermis, and possibly the intestine (Swoboda, unpublished). The location of expression of these isoforms differentiates its putative function from that of the DAF-19C isoform, which is involved in cilia formation and maintenance. Thus, if the selected gene is a downstream target of DAF-19A/B, it should co-localize with DAF-19A/B. Its expression pattern should differ in a *daf-19* ^{-/-} animal, particularly if non-ciliated neurons, hypodermis and intestine are the sites where DAF-19A/B is localized. A gene that is activated by DAF-19A/B will show decreased gene expression in the *daf-19* ^{-/-} mutant animals; while a gene that is repressed by DAF-19A/B will show increased expression in the mutant animals.

Choice of Recombinant Plasmid

Two recombinant plasmid samples were sequenced for each recombinant plasmid that we created (*F58E2.3::gfp* - sample 3&5; *F46G11.3::gfp* – sample 13&15) and we

obtained satisfactory results as we successfully cloned the desired putative regions of the two genes and confirmed the cloning by measuring fragment size, restriction mapping and sequencing. Within the 469 base pairs that were sequenced for *F58E2.3::gfp* construct, we found three erroneous bases, two of which were in sample 5. Furthermore, one of these errors in sample 5 was located approximately 120 base pairs upstream of the translational start site (ATG). Although the general transcriptional regulatory regions in *C. elegans* are often less than 2000 base pairs (Blackwell and Walker, 2006), the transcription factors bind to motifs such as a TATA box, which are located less than 100 nucleotides away from ATG. In addition, the average number of nucleotide upstream of translational start site for the Sand-box motif was 377 (Sand, 2011). Based on these points, an erroneous base near translational start site can be more detrimental to the process of transcription. Thus, sample 3 was chosen as the *F58E2.3::gfp* construct to minimize potential transcriptional errors.

In the case of the *F46G11.3::gfp* construct, although sample 15 contained two erroneous nucleotides at 906 and 972 base pair upstream of the translational site, sample 13 contained an error at 248 base pair upstream of the translational site. Thus, sample 15 was chosen for pursuing further gene expression studies. However, it has to be noted that only two thirds of the insert was sequenced and the region that was not sequenced may contain errors that can affect transcription.

Confidence in Obtaining Transgenic Worms

The expression of the co-injection marker *elt-2::mCherry* in the intestine of transgenic worms suggested that we generated transgenic worms that also took up our transcriptional::*gfp* constructs (Results, Figure 15) since the transcriptional::*gfp* construct was added in higher concentration in the injection mixture. The success in visualizing *F46G11.3::gfp* expression further validates the formation of the repetitive extrachromosomal arrays in the transgenic worms (Results, Figure 21). The expression pattern of *F46G11.4::gfp* (in LU600) was fairly consistent in *daf-19 +/+* animals. Any variation, such as only the posterior neuronal cell body pair showing GFP expression, may be due to genetic mosaicism. A genetic mosaic is an individual in which the extrachromosomal array was lost during mitotic cell division during development. Thus, sets of cells have different genotypes when the extrachromosomal element fails to be transferred to a daughter cell. As a result, the lineage of that cell will not carry the extrachromosomal element and, thus, a variety in the expression pattern between individual worms is to be expected.

F58E2.3::gfp expression was definitely not neuronal; however, its expression pattern in the intestine was consistent. A comparison between the transgenic worm and the non-transgenic *daf-19 +/+* negative control (Figure 18) shows that auto-fluorescence is not consistent in the non-transgenic negative control animals. Since we observed 100% GFP expression in the intestine of the transgenic worms, the observed fluorescence is not likely to

be auto-fluorescence. However, it is also difficult to conclude that the GFP expression was from the gut granules because the granules are organelles, not cells and they have no capability of driving transcription and translation alone. One possibility that explains the likely GFP expression in the intestine is the systematic error occurred during image acquirement. The wavelengths at which the laser was collected ranged from 495nm to 595nm and the mCherry emission spectra overlap with these values (580nm to 660nm). Therefore, the confocal microscope may have collected the fluorescence of *elt-2::mCherry* and projected the data using GFP fluorescent scheme. This explains why the expression is limited to the intestine, since *elt-2::mCherry* is only expressed in the intestinal area. To correct this error, future image acquisition of GFP expression in transgenic worms should only collect laser emissions that range between 495nm and 570nm.

Expression Patterns - F57B10.9::gfp

Three major differences were observed through comparing *daf-19* *+/+* animals and *daf-19* *-/-* mutant animals. These include GFP expression in (1) ciliated neurons and (2) body wall muscles, and (3) age-related GFP expression in nerve cords. *daf-19* *+/+* individuals, compared to *daf-19* *-/-* mutant animals, showed much stronger GFP expression in the cell bodies of ciliated sensory neurons that are located anterior to the nerve ring. Note that this observation is based on the morphology of the neuronal cells that express GFP expression

and previous investigation by Korzynski (2012) suggests that there is no co-localization of these neurons and the dye-filling neurons. Our observation agrees with previous findings that the *daf-19* *-/-* mutants show elimination of ciliogenesis (Senti and Swoboda, 2008). A strong GFP expression in the ciliated neurons also means that the gene *F57B10.9* is localized more in those cells than other neurons that are located posterior to the nerve ring.

The difference in GFP expression in the body wall muscles of the two strains support Dr. De Stasio's previous transcriptome analysis. Dr. De Stasio's microarray experiment has found that the expression of *F57B10.9* in *daf-19* *-/-* mutant animals were 1.58 times higher than that in wild-type animals. Our observation suggests that only the *daf-19* *-/-* individuals express *F57B10.9* in their body wall muscle cells at older stages (Results, Table 4). The reason for this observation is unclear. However, we may conclude that DAF-19 is indirectly regulating the gene *F57B10.9* since DAF-19A/B is not localized in the body wall muscle cells to our current knowledge. Further studies may include behavioral assays using an *F57B10.9* null mutant and the effect of *F57B10.9* in motor control of the worm may be unraveled.

An age-related decline in GFP expression in the nerve cords was only observed in *daf-19* *-/-* mutants. As Senti and Swoboda (2008) suggested, DAF-19A/B defective worms show decreased level of some synaptic proteins. This result may indicate that *F57B10.9* is involved in synaptic vesicle transportation, which is a function that has never been

characterized. The current work should be followed by a translational::gfp fusion, in which the entire transcript is incorporated into the plasmid so that the localization of *F57B10.9* gene product within the cells can be visualized.

Determining the Neuronal Cell Bodies Expressing F46G11.3::gfp

Using the DiI dye-filling map (Results, Figure 21A), five interneurons were listed as candidates for the two neuronal cell body pairs that express *F46G11.3::gfp*. Interneuron RIC, AIY, and AIM were selected to be the putative neuron cell that is located anterior relative to the other cell body pair. RIC (Results, Figure 23) was chosen because of its axon structure. Comparing the schematic figure and Figure 21 A&B, we can see a curvature of axon that the three images have in common. The axons of the interneuron AIY also made a ring-like structure and it is selected as a candidate because it is involved in the signaling of neurotransmitter acetylcholine (Li and Kim, 2008). As Sieburth and colleagues (2005) have shown, *F46G11.3* was one of the 185 genes of which the RNA interference (RNAi) of the gene resulted in a decrease in acetylcholine secretion. Thus, it is logical to consider that the expression of *F46G11.3* is co-localized with a cholinergic neuron. The final candidate, AIM, is involved in the transmission of serotonin. Interestingly, two studies have shown that interneurons AIM and RIF form the only neuron pair that is subject to developmental neurite pruning in *C. elegans* (Hayashi et al., 2009; Kage et al., 2005). A neurite is a projection of a

neuron and it can be either an axon or a dendrite. Studies have found that the excessive neurites are eliminated during larval stages in the wild-type (N2) animals. Although we did not observe any age-related pattern in *F46G11.3::gfp* expression (Results, Table 5), a larger sample number may generate more statistically significant results and a change in the gene expression pattern along the developmental stages of the worm. In addition, a comparison between the *daf-19* *+/+* animals (LU600) and *daf-19* *-/-* animals may reveal a potential connection between DAF-19 and age-related neurites pruning in *C. elegans*.

The cell body pair that is located posterior to the terminal bulb of the pharynx was identified as either RIF or RIG. The notable difference between the two interneurons is the presence of gap junctions (White et al., 1986). A gap junction of a neuron is also referred as an electrical synapse, in which the signal between the two neurons is transmitted mechanically and electrically, not through neurotransmitters or neuropeptides. However, the presence of gap junctions only explains the type of neurons that these two interneurons relay. One possible way of determining the cell body that expresses *F46G11.3* may be cell transfection. We may transfect the *F46G11.3::gfp* construct into the cell culture of either RIF or RIG and observe whether the cells display any GFP expression.

Summary and Conclusion

This study investigated whether the genes *F46G11.3*, *F57B10.9* and *F58E2.3* are the

downstream targets of RFX transcription factor DAF-19. Gene expression studies revealed that *F46G11.3* and *F57B10.9* are expressed in neuronal tissues and we observed notable difference in the *F57B10.9::gfp* expression in the *daf-19* *+/+* animals and *daf-19* *-/-* mutant animals. Our study with a Sand-box containing *F58E2.8::gfp* construct did not show any neuronal expression. Thus, the correlation between the presence of Sand-box motif in the gene's putative control region and neuronal expression is inconclusive. More analysis will be undertaken with genes studied by Dr. Swoboda's lab that contain a Sand-box motif and, ideally, the association between the Sand-box motif and neuronal gene expression can be confirmed.

References

- Alkema, M.J., Hunter-Ensor, M., Ringstad, N., and Horvitz, H.R. (2005). Tyramine functions independently of octopamine in the *Caenorhabditis elegans* nervous system. *Neuron*, 46, 247-260.
- Alzheimer, A. (1907). Über eine eigenartige Erkrankung der Hirnrinde. *Allgemeine Zeitschrift für Psychiatrie und Psychisch-gerichtliche Medizin*, 64, 146-148.
- Aravamudan, B., and Broadie, K. (2003). Synaptic *Drosophila* UNC-13 is regulated by antagonistic G-protein pathways via a proteasome-dependent degradation mechanism. *Journal of Neurobiology*, 54, 417-438.
- Badano, J.L., Mitsuma, N., Beales, P.L., and Katsanis, N. (2006). The ciliopathies: an emerging class of human genetic disorders. *Annual Review of Genomics and Human Genetics*, 7, 125-148.
- Bailey, T.L., Williams, N., Misleh, C., and Li, W.W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research*, 34, W369-W373.
- Blackwell, T.K., and Walker, A.K. ed. (2006) Transcription mechanisms. *Wormbook, ed. The C. elegans Research Community*.
- Blacque, O.E., Perens, E.A., Boroevich, K.A., Inglis, P.N., Li, C., Warner, A., Khattra, J., Holt, R.A., Ou, G., Mah, A.K., McKay, S.J., Huang, P., Swoboda, P., Jones, S.J., Marra, M.A., Baillie, D.L., Moerman, D.G., Shaham, S., and Leroux, M.R. (2005). Functional genomics of the cilium, a sensory organelle. *Current Biology*, 15, 935-941.
- Burghoorn, J., Piasecki, B.P., Crona, F., Phirke, P., Jeppsson, K.E., Swoboda, P. (2012). The *in vivo* dissection of direct RFX-target gene promoters in *C. elegans* reveals a novel cis-regulatory element, the C-box. *Developmental Biology*, 368, 415-426.

- Dupuy, D., Bertin, N., Hidalgo, C., Venkatesan, K., Tu, D., Lee, D., Rosenberg, J., Svrikapa, N., Blanc, A., Carnec, A., Carvunis, A., Pulak, R., Shingles, J., Reece-Hoyes, J., Hunt-Newbury, R., Viveiros, R., Mohler, W., Tasan, M., Roth, F., and Le Peuch, C. (2007). Genome-scale analysis of *in vivo* spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nature Biotechnology*, *25*, 663-668.
- Efimenko, E., Bubb, K., Mak, H.Y., Holzman, T., Leroux, M.R., Ruvkun, G., Thomas, J.H., and Swoboda, P. (2005). Analysis of *xbx* genes in *C. elegans*. *Development*, *132*, 1923-1934.
- Emery, P., Durand, B., Mach, B., and Reith, W. (1996). RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Research*, *24*, 803-807.
- Evans, T.C., ed. (2006) Transformation and microinjection. *Wormbook, ed. The C. elegans Research Community*.
- Goate, A.M., Owen, M.J., James, L.A., Mullan, M.J., Rossar, M.N., Haynes, A.R., Farrall, M., Lai, L.Y.C., Roques, P., Williamson, R., and Hardy, J.A. (1989). Predisposing locus for Alzheimer's disease on chromosome 21. *The Lancet*, *333*, 352-355.
- Goedert, M. (1993). Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Trends in Neurosciences*, *16*, 460-465.
- Hall, D.H., and Russell, R.L. (1991). The posterior nervous system of the nematode *Caenorhabditis elegans*: serial reconstruction of identified neurons and complete pattern of synaptic interactions. *The Journal of Neuroscience*, *11*, 1-22.
- Hayashi, Y., Hirotsu, T., Iwata, R., Kage-Nakadai, E., Kunitomo, H., Ishihara, T., Iino, Y., and Kubo, T. (2009) A trophic role for Wnt-Ror kinase signaling during developmental pruning in *Caenorhabditis elegans*. *Nature Neuroscience*, *12*, 981-987.

- Kage, E., Hayashi, Y., Takeuchi, H., Hirotsu, T., Kunitomo, H., Inoue, T., Arai, H., Iino, Y., and Kubo, T. (2005). MBR-1, a novel helix-turn-helix transcription factor, is required for pruning excessive neurites in *Caenorhabditis elegans*. *Current Biology*, 15, 1554-1559.
- Kipreos, E.T., and Pagano, M. (2000). The F-box protein family. *Genome Biology*, 1, REVIEWS3002.
- Korzynski, J. (2012). Investigating neuronal protein regulation by *daf-19*: a genetic approach. *Honors Dissertation, Lawrence University*.
- Li, C., and Kim, K., ed. (2008) Neuropeptides. *Wormbook, ed. The C. elegans Research Community*.
- Ross, C.A., and Poirier, M.A. (2004). Protein aggregation and neurodegenerative disease. *Nature Medicine*, 10, S10-S17.
- Sand, C.A. (2011). Novel Motif Discovery in a Set of Co-regulated Genes. *Honors Dissertation, Lawrence University*.
- Senti, G., and 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.
- Sieburth, D., Ch'ng, Q., Dybbs, M., Tavazoie, M., Kennedy, S., Wang, D., Dupuy, D., Rual, J. F., Hill, D., Vidal, M., Ruvkun, G., and Kaplan, J. M. (2005). Systematic analysis of genes required for synapse structure and function. *Nature*, 436, 510-517.

- St George-Hyslop, P., Haines, J., Rogaev, E., Mortilla, M., Vaula, G., Pericak-Vance, M., Foncin, J-F., Montesi, M., Bruni, A., Sorbi, S., Rainero, I., Pinessi, L., Pollen, D., Polinsky, R., Nee, L., Kennedy, J., Macciardi, F., Rogaeva, E., Liang, Y., Alexandrova, N., Lukiw, W., Schlumpf, K., Tanzi, R., Tsuda, T., Farrer, L., Cantu, J-M., Duara, M., Amaducci L., Bergamini, L., Gusella, J., Roses, A., and McLachlan, D.C. (1992). Genetic evidence for a novel familial Alzheimer's disease locus on chromosome 14. *Nature Genetics*, 2, 330-334.
- Swoboda, P., Adler, H.T., and Thomas, J.H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Molecular Cell*, 5, 411-421.
- Wang, J., Schwartz, H.T., and Barr, M.M. (2010). Functional specialization of sensory cilia by an RFX transcription factor isoform. *Genetics*, 184, 1295-1307.
- White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B, Biological Sciences*, 314, 1-340.
- Zhang, C.X., Engqvist-Goldstein A.E.Y., Carreno, S., Owen, D.J., Smythe, E., and Drubin., D.G. (2005). Multiple roles for cyclin G-associated kinase in clathrin-mediated sorting events. *Traffic*, 6, 1103-1113.
- Zheng, H., and Koo, E.H. (2006). The amyloid precursor protein: beyond amyloid. *Molecular Neurodegeneration*, 1, doi:10.1186/1750-1326-1-5.

Gene Name	Neuronal Expression	Fold Change in <i>daf-19</i> ^{-/-} Animals	Human Ortholog	Sand-box Motif Present?	Location of Sand-box	Main Expression Pattern
<i>F46G11.3</i>	✓	0.66	Cyclin-G-associated Kinase (GAK)	×	N/A	Two cell body pairs found near the posterior terminal bulb of the pharynx
<i>F57B10.9</i>	✓	1.58	SPARTIN	✓	315 base pairs upstream of ATG	Multiple head neurons
<i>F58E2.3</i>	×	0.6	N/A	✓	359 base pairs upstream of ATG	Possibly intestinal

Supplemental Table 1. Summary of the genes studied.