Function of conserved transcription factors in the *C. elegans* somatic gonad and pharynx

BY

Lynn M. Clary

B.S., University of Illinois at Urbana-Champaign, 2001

THESIS

Submitted as partial fulfillment of the requirements

for the degree of Doctor of Philosophy in Biological Sciences
in the Graduate College of the
University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Jennifer Schmidt, Chair
Peter Okkema, Advisor
Aixa Alfonso
Don Morrison
Maxim Frolov, Biochemistry and Molecular Genetics

ACKNOWLEDGMENTS

First, I would like to thank my thesis advisor Peter Okkema, for his support and encouragement over the years. Pete is an excellent teacher and mentor and I will always appreciate his patience and willingness to listen to my questions and problems especially during the last two years.

I would also like to thank the members of my thesis committee: Jennifer Schmidt, Aixa Alfonso, Don Morrison and Maxim Frolov for their time and advice.

I would like to thank the past and present members of the Okkema laboratory for scientific and nonscientific discussions. Their friendships made the laboratory a great environment to work in. In addition I would like to thank all of my friends and colleagues in the Biological Sciences department for their help and support. I would like to acknowledge the worm community especially Michael Miller, Tim Schedl and David Greenstein for their advice and encouragement with my EGRH-1 project.

Finally, I would like to thank my family and especially my husband Jason for his endless support, patience and encouragement that allowed me to finish my PhD. I could never have succeeded without him.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Adenosine
bp	base pair
C	Cytosine
cDNA	complimentary DNA
cpm	counts per minute
cuEx	extrachromosomal array
DAPI	4',6-Diamidino-2-phenyindole
DIC	Differential interference contrast
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dsRNA	double stranded RNA
DTC	distal tip cell
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
G	Guanine
g	gram
gfp/GFP	green fluorescence protein
kb	kilo basepair
kD	kilodalton
L1	larval stage 1
M	molar
MAPK	Mitogen-activated protein kinase
mg	milligram
ml	milliliter
mM	millimolar
mRNA	messenger RNA
MSP	major sperm protein
N2	*C. elegans* Bristol strain
ng	nanograms
orf	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pOK	plasmid identification for Pete Okkema laboratory
RNA	ribonucleic acid
RNAi	RNA mediated interference
rpm	rotations per minute
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
T	Thymine
TAE	Tris-acetate-EDTA
UTR	untranslated region
µg	microgram
µl	microliter
°C	degree Celsius

SUMMARY

Mammalian EGR-family genes and T-box family genes are broadly expressed and mutants result in a variety of defects and diseases (O'Donovan et al., 1999). To aid in our understanding of how these transcription factor families regulate gene expression we analyzed the function of *C. elegans egrh-1* and *tbx-2* in the somatic gonad and pharynx respectively.

We found that the *C. elegans* Early Growth Response factor family member *egrh-1* inhibits oocyte maturation and ovulation until sperm are available. *C. elegans* is an important system for addressing the fundamental events of oocyte meiotic maturation, ovulation and fertilization, complementing studies in vertebrate systems (Hubbard and Greenstein, 2000). In *C. elegans* oocyte production, maturation and ovulation must be coordinated with sperm availability for successful fertilization. *egrh-1* mutants exhibit ectopic oocyte differentiation in the distal gonadal arm and accumulate abnormal and degraded oocytes proximally. These defects result in reduced brood size and partially penetrant embryonic lethality. Results of tissue-specific *egrh-1(RNAi)* experiments and genetic mosaic analyses revealed EGRH-1 function is necessary in the soma, and surprisingly this function is required in both the gut and the somatic gonad. Through transformation rescue experiments we show EGRH-1 in the somatic gonad inhibits oocyte maturation, ovulation and sperm recruitment.

The T-box family member TBX-2 is the sole *C. elegans* member of the conserved Tbx2 sub-family. It has been previously shown that *C. elegans* TBX-2 is required for the development of ABa-derived pharyngeal muscles; however it is not known how TBX-2 specifically regulates the development of these cells. We are interested in identifying TBX-2 targets that function in

SUMMARY (continued)

pharyngeal muscle development, and in determining if TBX-2 is a transcriptional activator or repressor. To identify targets of TBX-2, we have compared mRNA expression levels in wild-type and *tbx-2(bx59)* mutant embryos using Affymetrix microarrays. Of 19,885 probe sets examined, we found 980 mRNAs that were significantly up-regulated in *tbx-2(bx59)* relative to wild-type and 175 mRNAs that were significantly down-regulated. Using clustering analysis, comparisons to existing data sets on pharyngeal gene expression, and phylogenetic foot-printing to identify promoters with consensus T-box factor binding sites, we have analyzed a subset of genes and identified D2096.6 as a direct target of TBX-2. Our data indicate that D2096.6 is directly repressed by TBX-2 at a variant T-box binding site in the D2096.6 promoter.

In addition to D2096.6 we chose seven genes (*T25E4.1, pqn-71, myo-5, flp-1, flp-2, flp-11* and *flp-15)* for further analysis and tested if these candidates were targets of TBX-2. We obtained or constructed GFP promoter fusions for these seven genes and compared their expression in wild type and *tbx-2(RNAi)* animals. We found no alteration of expression for six of these seven reporters. *T25E4.1::gfp* expression was altered by *tbx-2(RNAi).* Our analysis of the T25E4.1 promoter region indicates this gene is downstream of TBX-2, but it is likely an indirect target.

I. GENERAL INTRODUCTION

1.1   *C. elegans* **as a model for developmental biology**

Developmental biology aims to understand the mechanisms underlying tissue and organ

formation during development. *C. elegans* is a free-living nematode that has proven to be a

useful model system in developmental biology. Specific biological processes as well as gene

families are conserved between mammals and *C. elegans,* and research in these worms has

provided valuable information on the molecular and cellular pathways underlying some

mammalian diseases (Kaletta and Hengartner, 2006).

The adult hermaphrodite contains 959 somatic cells, which form several tissues including

the intestine, muscle, neurons, hypodermis and neurons. It is about 1 mm in length, feeds on

bacteria such as *E. coli,* and is easily maintained in large numbers in a laboratory environment.

One of the major advantages of *C. elegans* is that the cell lineage is invariant and the complete

lineage and position of each cell is known (Sulston et al., 1983). In addition, worms are

transparent which allows for observations of a specific tissue or even a single cell at all stages of

development (Sulston and Horvitz, 1977; Sulston et al., 1983).

C. elegans has two sexes, a self-fertilizing hermaphrodite and a male. Hermaphrodites

possess two X chromosomes and five pairs of autosomes, while males possess one X

chromosome and five pairs of autosomes. Males are produced spontaneously in hermaphrodite

populations by meiotic non-disjunction at a frequency of approximately 1 in 500 (Hodgkin et al.,

1979). Development begins in the egg, and continues through four larval stages (Ll-L4) to the

adult (Brenner, 1974; Sulston et al., 1983) (Figure 1). Adult hermaphrodites produce 250-350

self progeny and the entire life cycle is complete in 3.5 days at 20°C (Byerly et al., 1976).
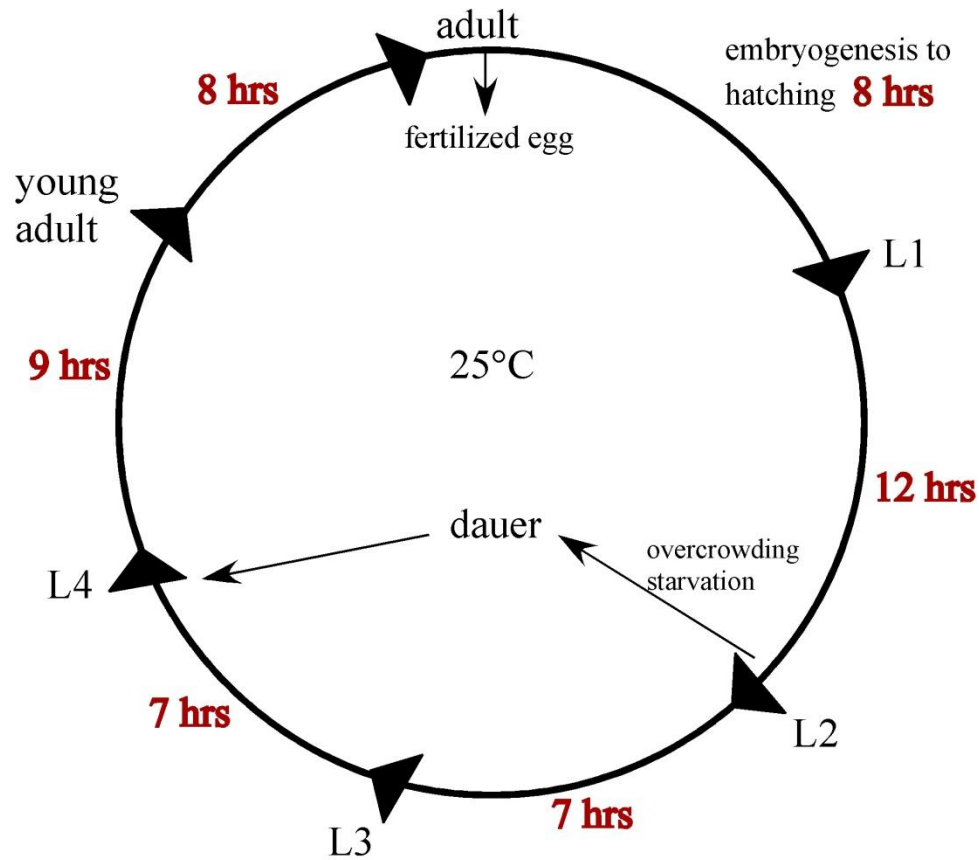
Figure 1: Diagram of *C. elegans* life cycle.

This diagram depicts *C. elegans* life cycle at 25ºC. Fertilization occurs at 0 minutes. Red numbers indicate the length of time the animal spends at each stage. 8 hours after a fertilized egg is laid embryos hatch into L1 larvae. There are four larval stages shown, including an alternate dauer stage in case of overcrowding and starvation.

1.2   Methods for genetic manipulations

C. elegans was the first multicellular organism to have its genome fully sequenced (Consortium, 1998). This along with its short life span has made it a popular model for genetic studies. Several methods have been discovered that allow for easy generation of mutant strains. Forward genetic screens have been very successful in *C. elegans* due to the ability to screen large numbers of worms and identify phenotypes induced by a mutagen. The completion of the genome sequencing has made reverse genetics such as RNAi very useful in identifying gene function. Gene function can be reduced specifically by injection of double stranded RNA (dsRNA) (Fire et al., 1998). Injection of dsRNA leads to specific degradation of the corresponding mRNA, a process that is referred to as RNA interference (RNAi) (Fire et al., 1998) . It has also been discovered that soaking worms in dsRNA or feeding worms bacteria which produce dsRNA results in an RNAi response (Tabara et al., 1998; Timmons and Fire, 1998). We utilized RNAi methods to aid in identification of EGRH-1 function in the somatic sheath cells and to identify targets of TBX-2 regulation. Self fertilization in hermaphrodites allows researchers to easily maintain mutants as either homozygous strains or heterozygous if the mutation is lethal or sterile.

The hermaphrodite distal gonadal arm is composed of a syncytial germ line which allows for easy generation of transgenic animals. When DNA is injected into a young hermaphrodite gonad, it recombines and forms an additional chromosome that is passed on to later generations, referred to as an extrachromosomal array (Mello et al., 1991). Extrachromosomal arrays undergo spontaneous mitotic loss and can be used in mosaic analysis to identify the specific cell or cells where a gene must be expressed for a specific phenotype (Yochem et al., 2000; Yochem and Herman, 2003). The invariant cell lineage of *C. elegans* is the key feature which makes

determining where in the lineage the extrachromosomal array is lost possible (Sulston et al., 1983). We utilized mosaic analysis to show that EGRH-1 function is required in the somatic sheath cells as well as the intestine which was a more surprising result as shown in chapter IV of this thesis. To maintain a strain of transgenic worms, easily identified co-injection markers are utilized, such as PRF4. This plasmid encodes a mutant collagen (*rol-6*) that induces a dominant phenotype, which causes animals to roll in circles (Kramer et al., 1990; Mello et al., 1991). Maintenance and availability of mutant strains has also been aided by the ability to freeze and recover worms (Brenner, 1974).

1.3   Transcriptional regulation

Microarray technology used in conjunction with bioinformatics provides a way to obtain large amounts of gene expression data and to identify targets of transcriptional regulators. The genome sequences of three *Caenorhabditis* species are available *elegans*, *briggsae* and *remanei* (Stein et al., 2003); http://genome.wustl.edu). These species are of the same genus, but they are more evolutionarily distant than human and mouse with their most recent common ancestor existing about 100 million years ago (Hillier et al., 2007). Comparisons of promoter regions between these three species allow identification of conserved cis-acting regulatory elements important in regulating gene expression. Bioinformatics analysis of known transcription factor binding sites upstream of coding regions has also been successful with the use of databases such as JASPAR (http://jaspar.genereg.net) and CONSITE (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/).

Transcriptional and translational reporter genes are used to study transcriptional regulation because the transparency of *C. elegans* allows for visualization of reporters such as GFP (Green Fluorescent Protein) and transgenic lines can be easily produced (Okkema and

Krause, 2005). Transcriptional reporter fusions have limitations and when constructed may lack some regulatory elements found in introns or the 3'UTR needed for proper expression (Wightman et al., 1993; Wenick and Hobert, 2004). However, they have been used successfully to analyze promoter regions and can be used in conjunction with microarray data to identify potential gene targets. We utilized phylogenetic analysis as well as GFP fusion proteins to analyze genes identified in our TBX-2 microarray and this allowed us to identify the first direct target of TBX-2 in *C. elegans* as shown in chapter VI of this thesis.

II. MATERIAL AND METHODS

2.1   Strains, plasmids and primers

Nematode strains, plasmids and primers are listed in Appendices E-G.

2.2   Handling of nematodes

C. elegans strains used in this work were grown under standard conditions (Lewis and Fleming, 1995). Transgenic lines were generated using techniques described by Mello and Fire (Mello and Fire, 1995). The plasmid pRF4 carrying a dominant *rol-6* allele was used as a transformation marker in transgenic lines (Mello et al., 1991).

2.3   Microscopy

Worms were visualized using a Zeiss Axioskop microscope operational for differential interference contrast (DIC) and fluorescence microscopy. Images were captured with an Axiocam camera and AxioVision software. MitoTracker and EGRH-1 antibody images were collected on a Zeiss Axiovert 200M equipped with a digital camera. Fluorescent images were collected as Z-stacks and subjected to 3-D deconvolution or directly collected.

2.4   DNA manipulations

2.4.1   Plasmid isolation: commercial kits

Commercial DNA isolation kits (Qiagen, Roche, and Sigma) were used for plasmid DNA miniprep and maxiprep purification, according to manufacturer supplied protocol.

2.4.2   DNA analysis by agarose gel electrophoresis

Plasmids were separated on 1-2% agarose gels base on size, in 1X TAE buffer (40mM Tris base, 30mM acetic acid, 5mM sodium acetate, 1mM EDTA) containing 0.3µg/ml ethidium bromide. HindIII/KpnI-digested lambda phage DNA was used as a molecular weight marker for

fragments over 1,000 base pairs. Commercial DNA molecular weight markers (Fermentas, New England Biolabs, and Invitrogen) were used for DNA fragments smaller than 1,000 base pairs.

2.4.3   DNA purification by agarose gel electrophoresis

Digested DNA was run on a 1-2% low melting point agarose gel in 1X TAE containing 0.3µl/ml ethidium bromide. Bands were visualized using long wavelength UV trans-illuminator. The gel slice containing the appropriate fragment of DNA was cut out and placed in a 1.5 ml microfuge tube with 0.4 ml 1X STOP (1M ammonium acetate, 10mM EDTA pH 8.0) and 1µl of 20µg/ µl glycogen. The gel slice was melted at 70ºC for 10 minutes with occasional agitation. Gel slices were extracted once with phenol pH 8.0 and centrifuged for 5 minutes at 14,000 rpm. The aqueous phase was then transferred to a new 1.5 ml tube and extracted with 1 volume of phenol/chloroform/isoamyl alcohol (25:24:1), and centrifuged for 5 minutes at 14,000 rpm. The aqueous phase was then transferred to a new 1.5 ml tube and extracted with 1 volume of chloroform/isoamyl (24:1), and centrifuged 2 minutes at 14,000 rpm. In a new 1.5 ml tube, 1 ml of EtOH was added to the aqueous phase from the previous extraction and the sample was centrifuged for 5 minutes at 14,000 rpm. All of the extractions were performed at room temperature. The pellet was washed with 500 µl absolute ethanol and centrifuged for 2 minutes. The supernatant was removed and the pellet was allowed to dry overnight. The dried pellet was then resuspended in the desired volume of TE (10mM Tris, 1mM EDTA) or dH$_2$O.

2.4.4   DNA ligation

DNA fragments were separated on a 1% low-melting temperature agarose gel, and fragments were excised from the gel (see DNA isolation by gel purification). Agarose slices were melted by incubation in 70ºC water bath for 2-5 minutes. 4 µl of vector and insert were added to tubes containing 2 µl of 5x Ligase buffer (Invitrogen), tubes were mixed and placed on

ice. 1 µl of T4 DNA ligase (Invitrogen) was pipetted into each tube and centrifuged for 10 seconds. Reactions were incubated at room temperature for 4 hours and then at 16ºC overnight. 50 µl 0.1 M CaCl$_2$ was added to samples and then the samples were heated at 65ºC for 5 minutes to melt the agarose. Ligations were transformed into DH5alpha competent cells.

DNA ligations were also performed in water using purified DNA. Ligations were prepared in 10 µl reactions containing vector, insert, 2 µl 5x ligase buffer (Invitrogen), 1 µl of T4 DNA ligase (Invitrogen) and water. The ligations were placed at 16ºC overnight and then transformed into DH5alpha competent cells.

2.4.5   DNA sequencing

Plasmids were sequenced at the University of Illinois at Chicago DNA sequencing facility. Plasmid templates were provided in 10 µl volumes at concentrations of 100 ng/µl.

2.5   Single worm PCR

Worm lysis was performed by picking worms (1 to 5 worms) into a PCR tube containing 2.5 µl 1x single worm lysis buffer (50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM MgCl$_2$, 0.45% Triton x-100, 0.45% Tween 20, 0.1% gelatin) and 120 µl/ml Proteinase K in a . One drop of mineral oil was added to each tube. In a thermal cycler tubes were incubated at 60ºC for 60 minutes, followed by 95ºC for 15 minutes to inactivate the Proteinase K. PCR was then performed using the appropriate primers. 25 µl of PCR mix (2.5 µl 10x PCR buffer, 0.75 µl 50 mM MgCl$_2$, 0.2 µl 25 mM dNTPs, 1 µl 100 ng/µl primer 1, 1 µl 100 ng/µl primer 2, 1 µl 100 ng/µl primer 3, 18.4 µl dH$_2$O, 0.25 µl Taq DNA polymerase) was added to each tube. A PCR cycle was designed based on the primers and PCR products being amplified, this cycle was repeated 39 times. PCR products were run on 1-2% agarose gel for analysis.

2.6   Handling of bacterial cells

2.6.1   <u>Preparation of competent DH5α *E. coli*</u>

An overnight culture was grown using a single DH5α colony in 3 ml 2xTY. The

following day 100 ml of 2xTY was added to a 500 ml flask and inoculated with 0.5 ml of the

overnight culture and shaken 1-2.5 hours until mid log phase A595=0.6. The culture was

transferred to 50 ml conical tubes and incubated on ice for 15 minutes. Cells were centrifuged at

4,000 rpm for 4 minutes at 4ºC to pellet. The supernatant was removed and the pellets were

resuspended in 2 ml cold Buffer 1 (10 mM MES pH 6.2, 100 mM RbCl2, 10 mM $CaCl_2$, 50 mM

$MnCl_2$). The resuspended pellets were then consolidated into one tube. 14 mls of Buffer 1 was

added to the resuspended pellets and incubated on ice for 15 minutes. Cells were pelleted at

4,000 rpm for 4 minutes at 4ºC and the supernatant was removed. The pellet was resuspended in

1.6 mls cold Buffer 2 (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl2, 15% glycerol) and

incubated on ice for 15 minutes. 200 µl aliquots were frozen on dry ice and stored at -70ºC until

needed.

Buffer 1 contained 0.390g MES dissolved in 100 ml dH_2O and the pH was adjusted to

6.2. In a separate flask 2.42g $RbCl_2$, 1.98g $MnCl_2$, and 2 ml 1M $CaCl_2$ was dissolved in 50 ml

dH_2O. The solutions were then mixed together, diluted to 200 mls and the pH was adjusted to

5.8 with 1M HOAc. The final solution was filter sterilized and stored at 4ºC.

Buffer 2 contained 0.419g MOPS, 15 mls 1 M $CaCl_2$, 0.242g $RbCl_2$, 30 mls glycerol, and

100 mls dH_2O. The solution was diluted to 200 mls and the pH adjusted to 6.5. The final

solution was filter sterilized and stored at 4ºC.

2.6.2   <u>Transformation of competent cells</u>

Plasmid DNA was mixed with 100 µl of competent cells on ice in 5 ml tubes and incubated on ice for 40 minutes. Cells were heat shocked for 20 seconds in a 42ºC water bath, and incubated on ice for 2 minutes. 100 µl of 2xTY was added to the cells and incubated for 1 hour shaking at 37ºC.   Cells were then spread on 2xTY plates with the appropriate antibiotic and incubated overnight at 37ºC

III. THE EGR-FAMILY GENE *EGRH-1* FUNCTIONS NON-AUTONOMOUSLY IN THE CONTROL OF OOCYTE MEIOTIC MATURATION AND OVULATION IN *C. ELEGANS*

3.1 Abstract

Oocyte production, maturation and ovulation must be coordinated with sperm availability for successful fertilization. In *C. elegans* this coordination involves signals from the sperm to the oocyte and somatic gonad, which stimulate maturation and ovulation. We have found that the *C. elegans* Early Growth Response factor family member EGRH-1 inhibits oocyte maturation and ovulation until sperm are available. In the absence of sperm, *egrh-1* mutants exhibit derepressed oocyte maturation marked by MAPK activation and ovulation. *egrh-1* mutants exhibit ectopic oocyte differentiation in the distal gonadal arm and accumulate abnormal and degraded oocytes proximally. These defects result in reduced brood size and partially penetrant embryonic lethality. We have found endogenous EGRH-1 protein and an *egrh-1::gfp* reporter gene are expressed in the sheath and distal tip cells of the somatic gonad, the gut and other non-gonadal tissues, and in sperm, but expression is not observed in oocytes. Results of tissue-specific *egrh-1(RNAi)* experiments and genetic mosaic analyses revealed EGRH-1 function is necessary in the soma, and surprisingly this function is required in both the gut and the somatic gonad. Based on transformation rescue experiments we hypothesize EGRH-1 in the somatic gonad inhibits oocyte maturation, ovulation and sperm recruitment.

3.2.   Introduction to the *C. elegans* reproductive system

    C. elegans is an important system for addressing the fundamental events of oocyte

meiotic maturation, ovulation and fertilization, complementing studies in vertebrate systems

(Hubbard and Greenstein, 2000). The amoeboid sperm of *C. elegans*, despite lacking an

acrosome and flagellum, carry out the same basic functions that are common to all sperm

(Singson, 2001). In many animals, including species of sponges, annelids, mollusks and

nematodes, sperm promote the resumption of meiosis in arrested oocytes (McCarter et al., 1999).

An advantage of *C. elegans* is the ability to isolate and maintain mutants that affect sperm and no

other cells. This is accomplished by selecting sterile hermaphrodites that cannot produce self

progeny but whose oocytes can be fertilized by wild type male sperm. In addition, the worms'

transparent cuticle allows for direct observation of gametogenesis, ovulation and fertilization

(Singson, 2001).

3.2.1   *C. elegans* reproductive system

    C. elegans is a self-fertilizing hermaphrodite and contains a reproductive system that

produces both sperm and oocytes. Hermaphrodites contain two symmetrical gonadal arms

surrounded by a basement membrane (Figure 2). Each gonad displays a distal-proximal axis

with proximal being toward the vulva. During the last larval stage hermaphrodites will produce

about 300 sperm and then switch to producing oocytes for the remainder of their lives.

Therefore, hermaphrodites are limited to an average of 300 self progeny and no longer produce

fertilized eggs once their sperm is depleted (Hirsh et al., 1976). At this point hermaphrodites can

resume fertilization of oocytes if they are mated to males. Once a hermaphrodite is mated to

males, they are capable of producing 1400 progeny (Kimble and Ward, 1988). Males occur

spontaneously at a frequency of approximately 1 in 500 worms (Hodgkin et al., 1979), and are

easily identified from females by their tail morphology. Males contain one gonad arm that continuously produces sperm throughout their life (Klass et al., 1976). Male populations can be maintained by setting up crosses with hermaphrodites and the resulting cross progeny will contain 50% males (Hirsh et al., 1976).
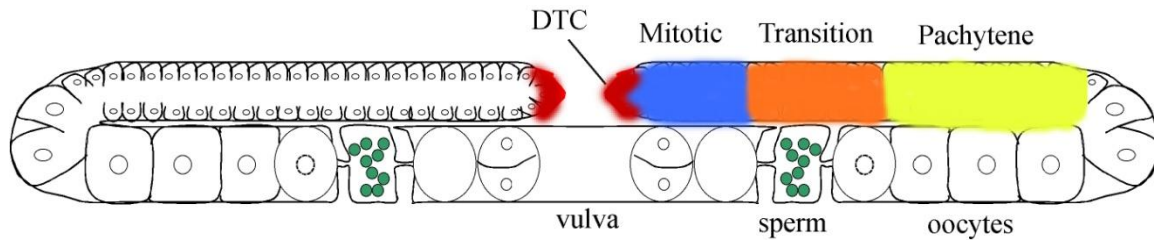
Figure 2: The adult hermaphrodite gonad and germline.

A drawing illustrating the adult hermaphrodite gonad and germline based on image from (Minasaki et al., 2009). The gonad consists of two identical arms. The distal tip of each arm contains the distal tip cell (DTC) shown in red. The distal arm of the gonad contains a syncytium of germline nuclei. Germline nuclei progress proximally and pass through stages of development that have been identified by three zones; mitotic proliferation (blue), transition (orange) and pachytene (yellow). As the nuclei enter meiosis and reach the bend of the gonad arm they begin to differentiate into oocytes. The proximal gonad arm contains large developing oocytes and the most proximal oocytes will undergo maturation and be ovulated into the spermatheca to be fertilized by sperm (green). The spermatheca from each gonad arm is connected by a common uterus where the fertilized embryos divide and are then laid through the vulva.

In hermaphrodites, the distal arm of the gonad is composed of a central undivided core of cytoplasm which is surrounded by a cylindrical layer of 1300 nuclei. The nuclei are in a syncytium surrounded by an incomplete plasma membrane (Hirsh et al., 1976). A population of mitotic germ cells is located at the distal end of the gonad and as germ cells progress proximally they enter meiosis (Hirsh et al., 1976). The proximal arm of the gonad contains a single row of oocytes which continue to enlarge and undergo maturation in an assembly line fashion. The proximal region of the gonad is surrounded by myoepithelial sheath cells (McCarter et al., 1997; Hall et al., 1999). Five sheath cell pairs form a thin layer, covering much of the gonad arm and the germ cells in adult hermaphrodites (Figure 3). Sheath cells 1 and 2 are the most distally located and contain the fewest myofilaments. Pairs 3 and 4 contain a large number of longitudinal filaments which function in contraction. The most proximal pair 5, contains a large number of both longitudinal and circumferential filaments which aid in contraction and pulling the spermatheca over the oocyte during ovulation (McCarter et al., 1997). The two most proximal pairs of sheath cells are the most important to ovulation because ablation of these cells results in defective ovulation (McCarter et al., 1997). In addition electron microscopy has shown that the proximal sheath pairs are able to communicate with oocytes through gap junctions (Hall et al., 1999). This is consistent with the phenotype of *ceh-18(mg57)* mutants which have reduced oocyte maturation and ovulation and were shown to have few or no gap junctions (Rose et al., 1997). Antibody staining with anti-myosin, which detects actin, reveals this filamentous network around the proximal gonad arm (Strome, 1986). Anti-myosin staining was used to visualize sheath cell morphology in studies of EGRH-1 mutants. Adjacent to the proximal gonad is the spermatheca (Figure 4). The spermatheca serves as the site of sperm storage and fertilization. The proximal gonad is separated from the spermatheca by a distal constriction and a

proximal constriction separates the spermatheca from the uterus (Hirsh et al., 1976; Ward and

Carrel, 1979). Ovulation occurs when the rate and intensity of sheath contractions increase and

pull the dilating distal spermatheca over the first oocyte. Spermathecae from both gonadal arms

are connected to a common uterus.
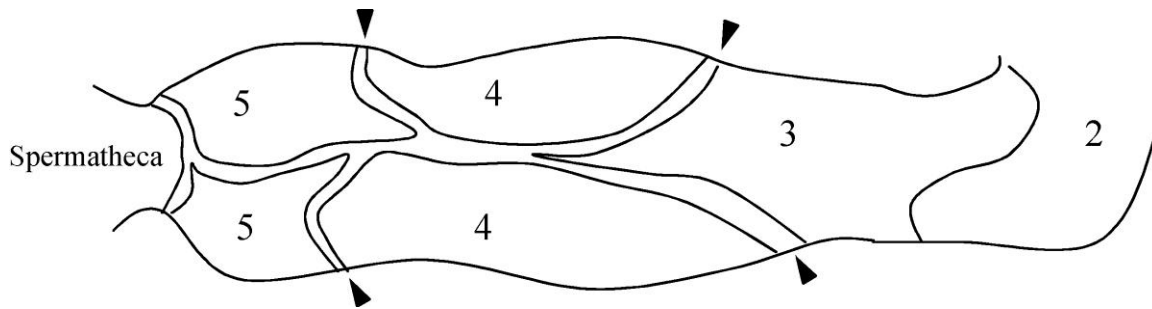
Figure 3: Arrangement of sheath cells in the proximal gonad of an adult hermaphrodite.

A drawing illustrating the proximal sheath boundaries in the proximal gonadal arm based on phalloidin staining and schematic from (McCarter et al., 1997) . Sheath cell pairs are labeled by numbers two through five; arrowheads indicate the boundaries between sheath cells. The proximal gonad is to the left.

3.2.2   <u>Oocyte growth</u>

The hermaphrodite gonad is composed of two U-shaped tubes and maturation of oocytes occurs in a distal to proximal manner. In late L4 larvae and adults, germ cells divide mitotically at the distal end of the gonad and enter meiotic prophase as they move proximally through the arm. Ablation of the distal tip cell (DTC) eliminates this population of dividing cells, showing that germline proliferation is dependent on DTC function (Kimble and White, 1981). Within the distal arm, germline cells have been categorized into three zones; the mitotic or proliferative, transition and pachytene zone (Figure 2) (Hirsh et al., 1976). These zones can be identified through DAPI staining based on nuclear morphology and distance from the DTC. The mitotic zone is distal to the transition zone; the transition zone is defined as the region where 80% of nuclei have the crescent-shaped morphology and ends where 80% of nuclei appear pachytene (Eckmann et al., 2004). The pachytene region is proximal to the transition zone and nuclei show a thread like appearance (Francis et al., 1995; MacQueen and Villeneuve, 2001). To determine if the progression of germ cell development is altered in *egrh-1* mutants, we compared the morphology and number of germ cell nuclei in different regions of the gonad in wild-type and *egrh-1(tm1736)* mutant females.

At the end of the distal arm closest to the loop region, the gonad core contains cytoplasmic components consisting of mitochondria, lipid drops and yolk bodies which will be enveloped into the developing oocytes (Hirsh et al., 1976). As oocytes develop in the proximal gonad they become more fully enclosed by membrane, increase in size, and chromosomes condense (Hirsh et al., 1976). The most proximal region of the gonad contains one to three fully developed oocytes that will undergo maturation and be ovulated into the spermatheca where they will be fertilized (Hirsh et al., 1976).

3.2.3   <u>Signaling from sperm promote oocyte maturation and sheath cell contractions</u>

The oocytes of most animal species arrest during meiotic prophase, and complete meiosis in response to intercellular signaling in a process called meiotic maturation. In *C. elegans*, each proximal gonadal arm contains 10-14 developing oocytes. All oocytes arrest at diakinesis of Meiotic Prophase I and resume meiosis in response to sperm (Hirsh et al., 1976). Oocyte meiotic maturation is defined by the transition between diakinesis and metaphase of Meiosis I and is characterized by nuclear envelope breakdown (NEBD) and cortical rearrangement resulting in the oocyte becoming more round. In *C. elegans*, the processes of meiotic maturation, ovulation, and fertilization are briefly linked. Sperm utilize the major sperm protein (MSP) as a hormone to initiate oocyte meiotic maturation and basal gonadal sheath cell contraction (Kuwabara, 2003) (Figure 4). Signaling by the VAB-1 Eph receptor and a somatic gonadal sheath cell-dependent pathway, defined by the CEH-18 POU-class homeoprotein, inhibit oocyte maturation, MAPK activation and ovulation when sperm are unavailable for fertilization. MSP binds VAB-1 and other receptor(s) in oocytes and sheath cells to inhibit VAB-1 and CEH-18 function. This promotes oocyte maturation, MAPK activation, and ovulation (Rose et al., 1997; Miller et al., 2003). In addition, MSP binds and activates VAB-1 on sheath cell membranes to stimulate the basal sheath cell contraction rates (Miller et al., 2001).
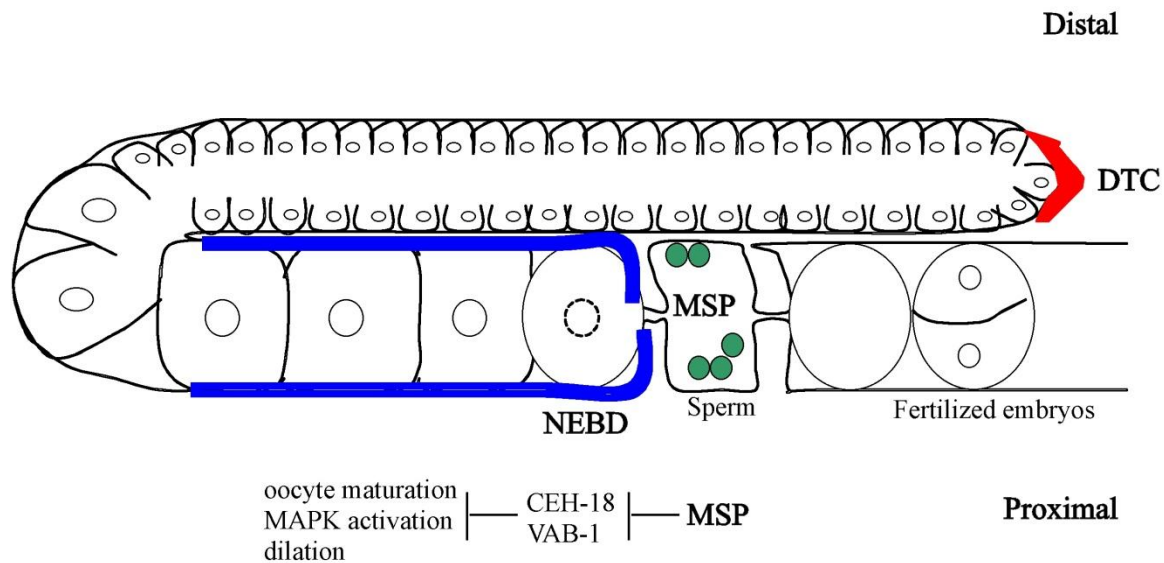
Figure 4: Signaling from sperm promotes oocyte maturation and ovulation.

A drawing of one gonad arm from an adult hermaphrodite based on an image from (Kuwabara, 2003). Germline development occurs from distal to proximal in an assembly line fashion; beginning with mitotic proliferation in the most distal end of the gonad followed by meiosis, oocyte differentiation, oocyte maturation and ovulation in the proximal gonad arm. Sheath pairs 3, 4 and 5 cover the proximal gonad arm and contract to aid in ovulation (region shown in blue). Just prior to ovulation, the proximal most oocyte undergoes meiotic maturation. Oocyte meiotic maturation is defined by the transition between diakinesis and metaphase of Meiosis I and is characterized by nuclear envelope breakdown (NEBD) and cortical rearrangement resulting in the oocyte becoming more round. In the absence of sperm VAB-1 and CEH-18 pathways function to inhibit oocyte maturation, MAPK activation and ovulation. Sperm signaling through major sperm protein (MSP) blocks VAB-1 and CEH-18 to promote oocyte maturation, MAPK activation, and spermathecal valve dilation.

3.2.4   A sperm activated oocyte signals its own ovulation

Ovulation occurs about every 23 minutes, and is supported by contractions from the proximal sheath pairs 3, 4 and 5 (McCarter et al., 1997). To complete successful ovulation, the proximal gonadal sheath cells contract rapidly, the distal constriction of the spermatheca dilates, and sheath cells pull the distal spermatheca over the mature oocyte. The maturing oocyte signals its own ovulation in two ways: it stimulates sheath contractions, which includes an increase in contraction rate and intensity during ovulation, and it induces spermathecal dilation during ovulation (McCarter et al., 1999); (Miller et al., 2001).

3.2.5   Emo oocytes result from defective maturation and/or ovulation

Mutations that lead to defective ovulation cause an endomitotic oocyte (Emo) phenotype (Iwasaki et al., 1996). When oocytes are trapped in the gonad arm due to defective ovulation, they undergo multiple rounds of nuclear envelope breakdown (M-phase entry) and S-phase, and become highly polyploid. An Emo phenotype can arise from a number of possible causes: defective dilation of the distal spermatheca; defective sheath cell contractions needed to progress the oocyte through the proximal gonad; defective sperm that are unable to provide the proper signals to arrested oocytes; and oocytes defective in meiotic cell cycle regulation (Iwasaki et al., 1996) (Yin et al., 2004). Genes with mutant alleles where defects in ovulation result in an Emo phenotype have been identified and characterized in *C. elegans*. We observed Emo oocytes in our EGRH-1 mutants and multiple factors including defects in meiotic cell cycle regulation and defective ovulation may be responsible for the embryonic lethality found in these mutants.

3.2.6   *egrh-1* is closely related to mammalian EGR genes

Early Growth Response (EGR) family genes are immediate early genes induced by mitogenic stimulation that were first identified in mammals [reviewed in (O'Donovan et al.,

1999)]. The human and mouse genomes each contain four EGR genes, which encode

transcription factors sharing three highly conserved C_2H_2 zinc-fingers that bind a GC-rich

consensus site (Swirnoff and Milbrandt, 1995; O'Donovan et al., 1999). EGR-family factors

regulate a variety of signaling pathways involved in cellular growth and differentiation, as well

as the transition from short-term to long-term memory (O'Donovan et al., 1999; Alberini, 2009).

Recent evidence suggests EGR genes function as tumor suppressors (Unoki and Nakamura,

2001; Joslin et al., 2007), and mutations in these genes have been associated with human

neuropathies [reviewed in (Houlden and Reilly, 2006)]. EGR gene mutant mice exhibit a variety

of defects, including infertility (Lee et al., 1996; Tourtellotte et al., 1999), defective nerve

myelination and abnormal hindbrain development (Schneider-Maunoury et al., 1993; Warner et

al., 1998), and sensory ataxia and an absence of muscle spindles (Tourtellotte and Milbrandt,

1998).

3.2.7   *egrh-1* functions in the somatic gonad to inhibit oocyte maturation, ovulation and sperm recruitment

        egrh-1 (<u>e</u>arly <u>g</u>rowth <u>r</u>esponse factor <u>h</u>omolog) is one of four *C. elegans* genes (*egrh-1*,

egrh-2, *egrh-3* and ZK337.2) closely related to mammalian EGR genes. Like mammalian EGR-

family members, the EGRH-1 protein contains three C_2H_2 zinc-fingers near the C-terminus. *In*

situ hybridization studies and SAGE analyses suggest that *egrh-1* mRNA is broadly expressed

from embryo through adult stages (Kohara, 2001; McKay et al., 2003). However, neither the

function nor expression pattern of *egrh-1* has been previously described. Here we characterize

EGRH-1 function and expression in adult hermaphrodites.

        Here we describe a role for EGRH-1 in regulating oocyte development. *egrh-1* loss

results in ectopic oocyte differentiation and accumulation of abnormal and degraded oocytes. In

the absence of sperm, *egrh-1* mutants exhibit MAPK activation and derepressed ovulation. *egrh-1* is expressed in the DTC and sheath cells of the somatic gonad, as well as in the gut, and we show *egrh-1* function is required in the somatic gonad and gut to regulate oocyte development. Based on our data we hypothesize EGRH-1 functions in the somatic gonad to inhibit oocyte maturation, ovulation and sperm recruitment.

3.3   Materials and Methods

3.3.1   <u>Nematode handling and strains</u>

     C. elegans were grown under standard conditions (Lewis and Fleming, 1995), and the

following strains were used: N2 [wild-type]; JK574 [*fog-2(q71)*] (Schedl and Kimble, 1988),

JK560 [*fog-1(q253)*] (Barton and Kimble, 1990); NL2099 [*rrf-3(pk1426)*] (Simmer et al., 2002);

NL2098 [*rrf-1(pk1417)*] (Sijen et al., 2001); OK0600 [*fog-1(q253); egrh-1(tm1736)*]; OK0639

[*fog-2(q71); egrh-1(tm1736)*]. OK0559 [*egrh-1(tm1736)*] was out-crossed from a strain

obtained from S. Mitani (National BioResource Project). Animals were genotyped for *egrh-*

1(tm1736) by single worm duplex PCR using primers PO761

(GGCACCATCATCGTCATACAAGG), PO762 (GGCAGCAGATAAGCCTGAAAATTC),

and PO762 (GCTGTAGTCATCCATTGGCTCGG) under conditions similar to those previously

described (Beaster-Jones and Okkema, 2004).

3.3.2   <u>General methods for nucleic acid manipulations</u>

     Standard methods were used to manipulate plasmid DNAs and oligonucleotides

(Ausubel, 1990), and all plasmid sequences are available from the authors. The *egrh-1*

containing cosmid C27C12 was kindly provided by Alan Coulson (Sanger Institute). Plasmid

pOK236.01 contains an *egrh-1(+)* genomic DNA fragment (bp 10647-19888 of C27C12; Acc#

Z69883). Plasmid pOK240.01 contains the same genomic fragment as pOK236.01 with the *gfp*

orf from pPD103.87 (kindly provided by A. Fire) digested with XbaI inserted in-frame into an

NheI site at bp 17454 in *egrh-1* exon 5. Plasmid pOK262.03 contains the *egrh-1* cDNA from

pOK165.12 digested with Xma1 and PspOMI ligated to XmaI and NotI digested pJM16 (kindly

provided by J. McGhee), which contains the *ges-1* promoter fragment. Plasmid pOK196.02

contains the *egrh-1* orf encoding amino acids 1-120 amplified from yk484c12/pOK165.12 and inserted into pMal2c (NEB).

3.3.3   Transgene analyses

All transgenic lines were generated by microinjection using pRF4 (100 ng/µl) carrying a dominant *rol-6* allele as a transformation marker (Mello and Fire, 1995). OK0601 *egrh-1(tm1736)*; cuEx503[C27C12(+)] contains cosmid C27C12 (2 ng/µl). OK0653 *egrh-1(tm1736)*; cuEx542[*egrh-1*(+)] contains plasmid pOK236.01 (2 ng/µl). OK0658 *egrh-1(tm1736);* cuEx547[*egrh-1::gfp*] contains plasmid pOK240.01 (2 ng/µl). For mosaic analyses, OK0636 *egrh-1(tm1736);* cuEx531[C27C12 + *sur-5::gfp*] contains cosmid C27C12 (2 ng/µl) and *sur-5::gfp* (100 ng/µl).

3.3.4   RNAi analyses

RNAi analyses were performed essentially as previously described (Fire et al., 1998). dsRNAs produced by in-vitro transcription (Ambion) from the *egrh-1* cDNA yk484c12/pOK165.12 which contains 2149 bp of EGRH-1 cDNA. The cDNA was cloned into a pBluescript cloning vector and amplified with primers PO3 and PO340 which recognize the T3 and T7 primer binding sites present in the Bluescript backbone. PCR program was as follows; (1) 94° 45 sec. (2) 56° 45 sec. (3) 72° 2 min. (4) got to step 1 35x (5) 12° hold . *egrh-1* dsRNA was injected at into the germline of young adult hermaphrodites, which were allowed to recover and transferred each day to freshly seeded plates. Phenotypes were scored in F1 progeny as young adults 24 and 48 hr post L4 stage. RNAi experiments in wild-type and mutant backgrounds were performed at 20˚C.

RNAi experiments in wild-type and *rrf-1* mutant backgrounds were performed at 20˚C, while those in *rrf-3(pk1426)* were performed at 16˚C.

3.3.5   Maturation rates

Oocyte maturation rates for *fog-2(q71)* and *fog-2(q71); egrh-1(tm1736)* females were

determined by counting total ovulations over a 4 hour time point at 25˚C (Miller et al., 2003).

Newly molted adults were used to avoid oocyte degradation found in older *fog-2(q71); egrh-*

1(tm1736) females.

3.3.6   Mating efficiency

To test the mating efficiency of *egrh-1(tm1736)* sperm, *fog-1(q253)* females were mated

with either N2 or *egrh-1(tm1736)* males. Single crosses were set up at 25˚C overnight, males

were then removed from plates and mated females were transferred once every 24 hours for 3

days. The total number of embryos and hatched larvae were counted.

3.3.7   Lethality and brood tests

Lethality tests were conducted as follows for wild type and mutant hermaphrodites. Two

L4 hermaphrodites were placed on 3 individual plates at 25˚C. After 24 hours moms were

transferred to new plates. Embryos from each plate were counted and transferred to new plates

at 25˚C. After 24 hours un-hatched embryos were counted and after 48 hours adults were

counted from each plate.

Brood tests were conducted as follows on wild type and *egrh-1* hermaphrodites. Six L4

egrh-1(tm1736) hermaphrodites and three L4 N2 hermaphrodites were individually plated and

placed at 25˚C. Every 6 hours moms were transferred to new plates and embryos were counted

until the original moms stopped laying embryos.

3.3.8   Cloning and bacterial transformation for protein expression

The MBP::EGRH-1 plasmid was constructed by a previous student (Tomas Vilimas).

The EGRH-1 C-terminus was amplified with primers PO695/PO696 and cloned into EcoRI-PstI

sites of pMal C2. To transform MBP::EGRH-1 into bacterial cells, 1 μl of plasmid DNA

pOK196.02 (MBP:EGRH-1 fusion) was mixed with 100 μl BL21 competent cells in 15 ml conical tubes

and placed on ice for 30 minutes. Cells were then heat shocked for 60 seconds at 42ºC in a water bath.

Cells were taken directly from the water bath and placed on ice long enough to add 100 μl 2xTY to each

tube. Cells were then shaken at 37ºC for 30 minutes. An ETOH sterilized glass spreader was used to

spread cells on 2xTY + ampicillin plates and plates were incubated overnight at 37ºC.

3.3.9   MBP::EGRH-1 expression

An individual colony from the transformation of pOK196.02 (MBP::EGRH-1 fusion) in

BL21 competent cells was picked and used to inoculate a 50 ml culture (50 ml LB + 50 μl

ampicillin) in a 150 ml flask. The culture was shaken overnight at 37ºC. Two percent of the

overnight culture was added to 1 Liter of LB (+ 1000 μl ampicillin) in a 2 liter flask. The culture

was placed at 37ºC with shaking until the OD reached 0.6-0.8. At this point, a 200 μl sample

was taken as the un-induced sample. This sample was pelleted with a short spin in the micro-

centrifuge (centrifuge 5804 R). Following centrifugation, the supernatant was removed and the

pellet was stored at -20ºC. The remaining culture was induced with 1M IPTG and placed at 37ºC

with shaking for 2 hours. After 2 hours a 200 μl sample was taken as the induced sample. This

sample was pelleted with a short spin in the micro-centrifuge (centrifuge 5804 R). Following

centrifugation, the supernatant was removed and the pellet was stored at -20ºC. The remaining

culture was split into four balanced conical tubes and centrifuged at 7,000 rpm for 1 hour at

4ºC(rotor F-34-6-38, centrifuge 5804 R). The supernatant was removed and the pellets were

stored at -20ºC.

The un-induced and induced pellet samples were then run on a 10% tris-HCl gel (BIO-

RAD Ready gel precast Gel for Polyacrylamide Electrophoresis). To run the samples on a gel

the pellets were resuspended in 20 µl dH$_2$0 by vortexing. Next, 20 µl 2x loading buffer was

added (25 ml 4x Tris-Cl, pH 6.8, 20ml glycerol, 4 g SDS, 3.1 g DTT, 1mg bromphenol blue and

dH$_2$0 to 100 ml) and the samples were boiled at 95ºC for 5 minutes. The samples were then

loaded in the gel along with a high and low molecular weight protein marker. The gel was run at

100 volts for 1-1.5 hours. After the gel finished running, it was removed and placed in

coomassie stain (0.5% coomassie, 25 ml methanol, 65 ml H20, 10 ml acetic acid) for 30 minutes

on a shaker at room temperature. Next, the coomassie stain was poured off and replaced with 50

ml of destaining solution (50 ml methanol, 130 ml dH$_2$0, 20 ml acetic acid). After 30 minutes

the destaining solution was poured off and replaced with 50 ml of fresh destaining solution. The

gel was washed with destaining solution until the protein bands were visible.

3.3.10 MBP::EGRH-1 solubility

After the protein was successfully expressed the cells were lysed. The pellets from the 1

liter of induced culture and the sonication binding buffer (100mM Tris-HCl pH7.5, 0.3M NaCl,

5% glycerol) were placed on ice. Once the cell pellets were thawed, the pellets were each

resuspended in 2 ml of sonication binding buffer by vortexing (sonication binding buffer volume

was 6x the pellet size) and then combined into one conical tube. 50 µl protease inhibitor cocktail

(Sigma) and lysozyme to a final concentration of 1 mg/ml was added to the tube and placed on

ice for 0.5 hours. Next, the culture was sonicated for 10 cycles 10 seconds each and the sample

was placed in ice between each cycle. Next, the lysate was transferred to an oak ridge tube and

50 µl of 10% Triton X-100 was added. The sample was placed on a rocker for 20 minutes at

4ºC. The sample was then centrifuged at 5,000 rpm for 45 minutes at 4ºC (rotor F-34-6-38,

centrifuge 5804 R). The supernatant was transferred to a new tube (this contains the soluble

fraction, pellet contains insoluble fraction). 100 µl of supernatant was collected for gel analysis

as total soluble protein fraction. All samples that were collected were run on a SDS-PAGE gel as previously described. For pelleted samples(insoluble) the pellet was scraped with the tip of a pipette and resuspended in 20 µl dH$_2$O and 20 µl 2x loading buffer.

3.3.11 MBP::EGRH-1 purification

To prepare the amylose resin, 3 ml of 50% amylose resin slurry (NEB) was washed with 6 ml MBP buffer (20mM Tris pH7.5, 200mM NaCl, 1mM EDTA) in a 10 ml Poly-Prep Chromatography Column (Bio-Rad #731-1550). The column was then centrifuged for 30 seconds at 3000 rpm (rotor F-34-6-38, centrifuge 5804 R) and the supernatant was removed. This wash was repeated 2 times. The soluble protein was transferred to the column containing the amylose resin and rocked for 4 hours at 4ºC to allow the protein to bind. The column was centrifuged for 30 seconds at 3,000 rpm at 4ºC. The supernatant was then transferred to a clean tube and 100 µl was collected as the unbound protein fraction. The column resin was washed with 5 ml cold MBP buffer for 3 minutes and centrifuge for 30 seconds at 3,000 rpm. The supernatant was discarded and this wash was repeated two times. The column resin was then resuspended in 1 volume MBP buffer. The following elution steps were performed at 4ºC. The column resin was allowed to settle and the liquid to flow until the meniscus reached the top of the amylose resin. The protein was eluted with 10 ml cold MBP buffer containing 10 mM maltose. One ml fractions were collected and stored on ice. To identify the fraction that contained the eluted protein 10 µl of each fraction was run on SDS-PAGE gel as previously described. Protein concentrations were estimated by Bradford Assay (BioRad).

3.3.12 EGRH-1 antibody production and purification

The *egrh-1* orf encoding amino acids 1-120 was amplified from yk484c12/pOK165.12 and inserted into pMal2c (NEB). Antibodies were raised in rabbits D1919 and D1920 (Open

Biosystems) against the MBP::EGRH-1 fusion protein purified from *E. coli* BL21 (Ausubel, 1990). EGRH-1 antibodies were purified by sequentially depleting against *E. coli* extracts expressing MBP and enriching against purified MBP::EGRH-1 bound to nitrocellulose filters (Duerr, 2006). D1919 reacted strongly with MBP::EGRH-1 on western blots and was used in these studies.

MBP::EGRH-1 antibody was purified as follows. MBP::EGRH-1 protein (used as antigen) and crude lysate of BL21 *E. coli* expressing MBP was bound to nitrocellulose membranes. To bind the protein two 90 mm petri plates were coated with 10-20 ml TBS-T-NFDM (150nM NaCl, 40mM Tris pH 7.4, 0.05% Tween 20, pH 8.0, 3% nonfat dairy milk) and placed on a shaker at low speed for 1 hour and 20 minutes. The plates were rinsed 3 times with 10-20 ml TBS-T (150mM NaCl, 40mM Tris pH7.4, 0.05% Tween 20, pH adjusted to 8.0) for 5 minutes on a shaker at room temperature. One mg of MBP::EGRH-1 and 3 mg of crude MBP protein was diluted to 15 ml with transfer buffer (20% methanol, 25mM Tris, pH 8.0), and added to individual petri dishes. Next, one Nitrocellulose membrane for MBP::EGRH-1 and crude MBP was placed in separate petri dishes with forceps and the plates were put on the shaker at room temperature for 1 hour. The nitrocellulose membranes were washed 3 times for 5 minutes with 10-20 ml TBS-T. The membranes were then incubated for 75 minutes in 10-20 ml TBS-T-NFDM on a shaker at room temperature. If needed the membranes were stored in a small volume of TBS-T-NFDM at -20°C.

To pre-elute the MBP::EGRH-1 and MBP membranes, the membranes were rinsed 3 times for 5 minutes each with 10-20 ml TBS-T. The membranes were then rinsed briefly with 10-20 ml dH$_2$O. 10 ml glycine buffer (5mM glycine, 0.01% BSA. 0.05% Tween-20, 0.5 M

NaCl. pH 2.3) was added for 60 seconds and then discarded. This process was performed for each membrane separately.

The MBP::EGRH-1 membrane was incubated with EGRH-1 antibody serum. The membrane was then rinsed 3 times for 5 minutes with 10-20 ml TBS-T on a shaker at room temperature. 10 ml of TBS-T and 1 ml of final bleed serum was added to the petri dish and placed on a shaker overnight at room temperature. The antibody solution was saved and checked on western to see what was not bound to the MBP::EGRH-1 membrane. The membrane then rinsed 3 times for 5 minutes with TBS-T. Next we eluted the EGRH-1 antibody from the membrane. A 15 ml tube was prepared that contained 400 µl of 1M Tris (pH 8.0) for the glycine rinse. The membrane was briefly rinsed with dH_2O and then 5 ml of fresh glycine buffer was added to the dish for 20 seconds. After 20 seconds the glycine buffer was collected in the prepared tube containing Tris pH 8.0. Immediately 5 ml fresh glycine buffer was added to the membrane for 20 seconds. The second rinse was collected in the tube containing the first rinse. Eluates were neutralized with 0.5-1 ml of 1 M Tris pH 8.0 to bring pH to 7.5. The final volume was 11 ml which contained the EGRH-1 antibody.

The MBP membrane was incubated with EGRH-1 antibodies collected from the previous enrichment. The MBP membrane was rinsed 3 times for 5 minutes in 10-20 ml TBS-T on a shaker at room temperature. The 11 mls of previously enriched EGRH-1 antibodies was added to the membrane and placed on a shaker at room temperature overnight. The supernatant was saved containing the EGRH-1 antibodies and checked on a western. The membrane was then discarded. This step removed any antibodies that bound to the MBP epitopes. The purified EGRH-1 antibodies were store at -80°C.

3.3.13 <u>Western blots</u>

EGRH-1::MBP and MBP proteins were induced as previously described and samples were run on a SDS-PAGE 10-12% Tris-HCl gel (BioRad). After running the samples the gel was cut in half. One side was used for the western and the other side used for coomassie staining. For the western the gel was placed in cold buffer (3g Tris base, 14.42g glycine, 15% methanol in 1 liter). Whatman paper and PVDF membrane transfer paper (Millipore) was cut to the gel size. PVDF membrane was activated by placing it in methanol for a few seconds and then washing it with dH$_2$O 3 times. The gel and filter papers were then assembled in the blot transfer tank so that the protein would transfer from the gel to the membrane paper. The gel was run at 2 amps per in^2 for 1 hour to complete the transfer. Once the gel is finished running the gel was removed and stained with coomassie to test the efficiency of the protein transfer. Next, the PVDF membrane was placed in 5% milk powder solution on a rocker to block for at least 1 hour at 4°C.

The PVDF membrane was incubated in EGRH-1 antibody diluted (1:60) with PBST (1x PBS, 0.05% Tween 20) for 1 hour rocking at 4°C. The membrane was then washed with 50 ml PBST at room temperature for 15 minutes, and this wash was repeated 3 times. Next, the membrane was incubated in secondary antibody (HRP-conjugated goat anti-rabbit, Chemicon) at the required dilution (1:3000) and rocked at 4°C for 1 hour. The membrane was washed with 50 ml PBST for 15 minutes and this wash was repeated 3 times. The membrane was then transferred to film for detection. To transfer to film, equal volumes of reagent 1 and 2 from ECL western Blotting Analysis system (Amersham Biosciences) was added to the membrane for 1-2 minutes and then removed. Finally, the membrane was covered in plastic wrap and exposed 5-20 seconds before being developed.

3.3.14 <u>Antibody Staining (freeze-crack/methanol/acetone fixation)</u>

To prepare for staining an etching pencil was used to draw a circle on one side of a microscope slide (Fisher Scientific). The slide was then coated with 30 μl of polylysine solution with in the area of the circle, and allowed to air dry. Mixed populations of wild type (N2) worms were washed off plates with cold dH_2O and transferred into 1.5 ml eppendorf tubes and centrifuged for 1 minute at 1,000 rpm (rotor F45-30-11, centrifuge 5804 R). This wash was repeated 2 times. Worms were then transferred to the polylysine coated slide, and allowed to settle for 1 minute. A coverslip (24mm X 50mm Fisher Scientific) was placed on the slide at an angle so that the edges of the coverslip extend over the edges of the slide. The slide was placed on a metal block, which was pre-frozen on dry ice, and slight pressure was applied to the coverslip with an etching pencil until the slide was frozen to burst adult worms. Once frozen the coverslip was snapped off of the slide and the worms were fixed through a methanol/acetone dilution series (5 min in 100% -20C methanol; 5 min in 100% -20C acetone; 2 min each in 95%, 70%, 50%, 30% methanol room temperature; 2 min 1x PBS, room temperature). The worms were incubated for 30 minutes in fresh PBT + 1% NGS (1X PBS + 0.1% Triton x-100, 0.1% BSA) to block. Next, excess buffer was wiped off of the slide and 50 μls (1:3 dilution) of primary antibody was added to the worms. The worms were then incubated overnight at room temperature in a humidity chamber. The following day the worms were washed 4 times for 10 minutes in TBS-T (150mM NaCl, 40mM Tris pH 7.4, 0.05% Tween 20 pH adjusted to 8.0). The excess liquid was wiped off of the slide and 50μls (1:500 dilution) of secondary antibody alexa 546 goat anti-rabbit was added to the worms. The worms were incubated for 4 hours at room temperature in a humidity chamber. The worms were then washed 4 times for 10 minutes in TBS-T. Excess liquid was wiped from the slide and 10 μls of aqueous mounting medium with

anti-fading agents (BioMeda M01) was added. Finally, a coverslip (24x50 Fisher Scientific) was placed over the worms and edges of the slide were sealed with nail polish.

3.3.15 Gonad/intestine immune-staining

Adult hermaphrodites were washed from plates with 1X egg buffer (118 mM NaCl, 48 mM KCl, 2 mM $CaCl_2$ $2H_2O$, 2 mM $MgCl_2$ $6H_2O$, 25 mM Hepes, pH adjusted to 7.3 with 1N NaOH and then filter sterilized). Worms were centrifuged in an eppendorf tube at 3,000 rpm for 1 minute (rotor F-34-6-38, centrifuge 5804 R) and washed with 1 ml 1X egg buffer. This wash was repeated 2 times. Worms were then transferred with a 1 ml pipette tip (coated with SurfaSil silconizing fluid, PIERCE) to a dissecting tray containing 200 µl 1X egg buffer + 0.2% levamisole. Only the number of worms that could be cut in approximately 15 minutes were transferred (~50). As paralysis set in, heads were cut off just posterior to the pharyngeal terminal bulb. The head of the worm was placed between two 26 gauge syringe needles and the needles were moved in a back and forth motion to sever the heads. At this point at least one gonad arm and the intestine were extruded from the body. Cut worms were transferred with a 1 ml pipette tip (coated with SurfaSil silconizing fluid, PIERCE) to a 1.5 ml presilconized eppendorf tube. The cut worms were then fixed with 1ml 1.25% fresh paraformaldehyde in egg buffer and samples were rocked at room temperature for 10 minutes. The worms were centrifuged at 3,000 rpm for 1 minute and the liquid was removed. Worms were then incubated in 1 ml PTB (1% BSA, 1X PBS, 0.1% Tween 20, 0.05% Na Azide, and 1mM EDTA) at room temperature for 1 hour with 2 buffer changes. Cut worms were again centrifuges at 3,000 rpm for 1 minute and the liquid was removed. A small amount of PTB was left in the tube and 30 µls of primary antibody was added. The worms were incubated overnight at 4°C with rocking. The following day the worms were washed for 4 hours at room temperature in PTC (0.1% BSA, 1X PBS, 0.1% Tween

20, 0.05% Na azide, 1mM EDTA) with 4 buffer changes of 1 ml each. Next, the worms were

incubated in PTB + 10% goat serum at room temperature for 30 minutes. A 1:800 dilution

secondary antibody Alexa 488 or Alexa 546 in PTB was added for 2 hours at room temperature.

Again the worms were washed for 2 hours at room temperature in PTC with four 1 ml buffer

changes and then an additional wash was done overnight with rocking at 4°C. The following day

worms were centrifuged at 3,000 rpm for 1 minute and the liquid was removed. Ten μls of

aqueous mounting medium with anti-fading agents (BioMeda) was added to the worms. The

worms were then mounted on pads of 2% agarose just before viewing. A coverslip was placed

on the worms and sealed with nail polish and the slides were stored at 4°C (protocol modified

from protocol by Barth Grant) (Grant and Hirsh, 1999),

For EGRH-1 antibody staining, adult hermaphrodites were dissected and fixed as

described. EGRH-1 was localized by indirect immunofluorescence using goat anti-rabbit

secondary antibodies labeled with Alexa 488 or Alexa 546. Worms were visualized using a

Zeiss Axioskop microscope equipped for DIC and fluorescence microscopy. Images were

captured using Axiocam camera and Axio vision software.

MAPK activation in dissected young adult hermaphrodites was detected using

monoclonal MAPK-YT antibodies (Sigma) (Miller et al., 2001) (1:2000 dilution) and Alexa 488

labeled goat anti-mouse secondary antibody 1:500 dilution. Young adult hermaphrodites and

females were dissected and fixed and visualized as previously described.

3.3.16 MitoTracker experiment

Protocol modified from (Kubagawa et al., 2006b)

MitoTracker Red CMXRos (Invitrogen) was used to label male sperm. L4 virgin males were

picked to a plate and placed at 16ºC overnight to age. Prior to use, MitoTracker was diluted in

DMSO to 1 mM concentration. Fifty males were placed in 300 µl M9 buffer in a watch glass and MitoTracker was added to a final concentration of 10 µM. The male worms were incubated in the dark for 2 hours at 25°C. Male worms were then transferred to a fresh plate and allowed to recover overnight at 16°C. Next, twenty-five males were placed with 6-8 anesthetized females/hermaphrodites (2.5 mM levamisole in TBS) on NGM plates with a 1 cm diameter drop of bacteria, OP50. Worms were allowed to mate for 18 hours at 20°C in the dark and then the mated females/hermaphrodites were mounted on 2% agarose pads for viewing.

3.3.17 <u>Mosaic analysis</u>

Mosaic analyses were carried out using *sur-5::gfp* expression to identify lineages lacking an *egrh-1*(+) transgene (Yochem and Herman, 2003). L4 and young adult animals from the strain OK0636 *egrh-1(tm1736); cuEx531[C27C12 + sur-5::gfp]* were screened using a fluorescence dissecting microscope for candidate mosaic animals that had lost the extrachromosomal transgene in the MS, E, and EMS lineages. Candidates were maintained for 24 hours, and 707 animals were examined using a compound microscope to identify the lineage that had lost the GFP(+) transgene and to score the EGRH-1 phenotype. Transgene loss in MS, E, and EMS lineages was verified by loss of GFP in spermatheca and posterior pharynx for MS, intestinal cells for E and loss in all of these cells for EMS. Animals were then scored in DIC for the presence of abnormal and degraded oocytes in the proximal gonad and an accumulation of oocytes in the distal gonad. In addition 1119 hermaphrodites were cloned and examined for loss of the array in the germline and presence of the mutant phenotype, but animals with a germline specific loss were not identified.

3.4 Results

3.4.1 *C. elegans* EGRH-1 is similar to EGR-family zinc-finger proteins

The *C. elegans* gene *EGRH-1* encodes a predicted 461 amino acid protein containing 3 C_2H_2 zinc-fingers near the C-terminus, and it is most closely related in reciprocal BLAST searches to mammalian Early Growth Response (EGR) family proteins and Drosophila Stripe. Within the zinc-finger regions EGRH-1 is 94% similar to mammalian EGR2 and 96% similar to Drosophila Stripe. (Figure 5 A, B) (Lee et al., 1995; O'Donovan et al., 1999). We have therefore named EGRH-1 for early growth response factor homology. EGRH-1 protein shares high sequence identity with EGR-family factors in the zinc-finger region, including residues that contact DNA (Wolfe et al., 2000), but it has no obvious similarity elsewhere. This high identity within the zinc-finger DNA-binding domain suggests EGRH-1 likely binds the same GC-rich sequence motif GCGKGGGCG recognized by mammalian EGR-family factors. This motif is found at 469 distinct sites in the *C. elegans* genome (Markstein et al., 2002), although the function of these sites has not been explored.
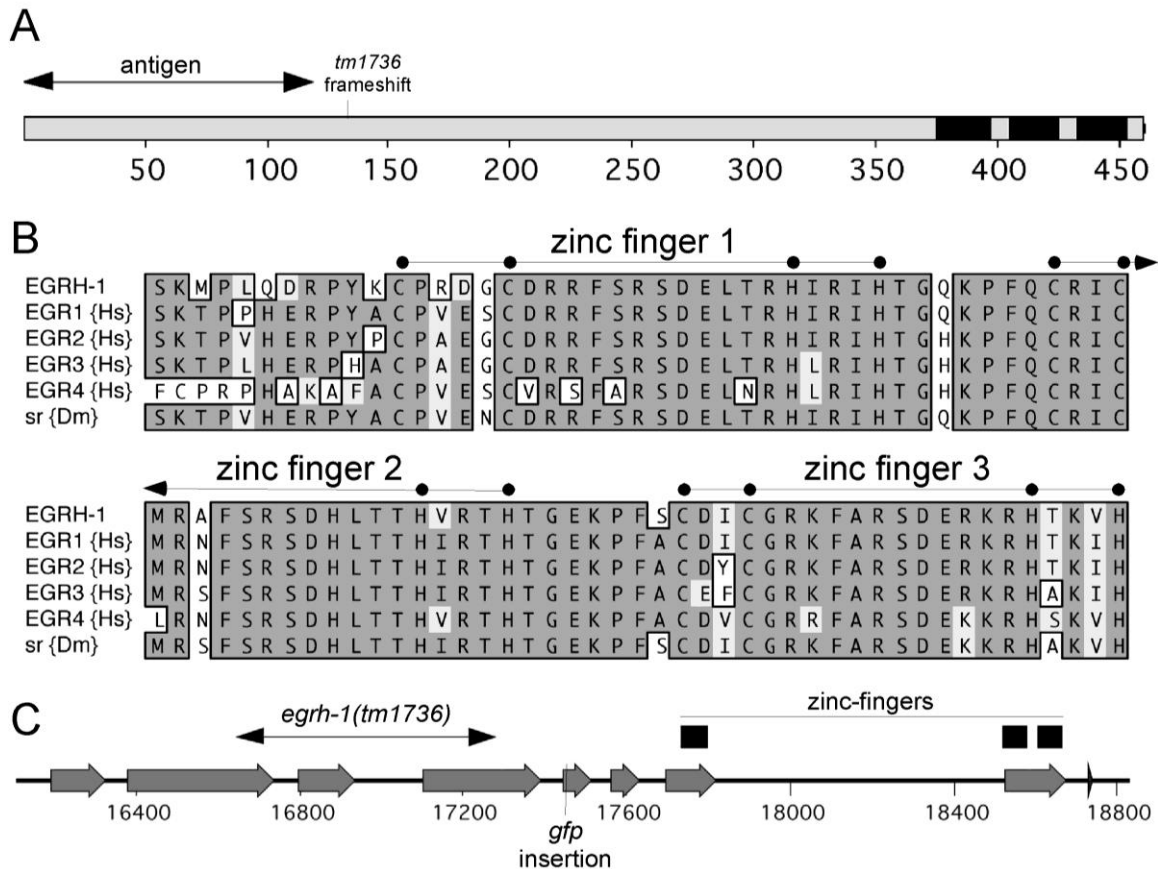
Figure 5: Organization of the *egrh-1* gene and protein

(A) Schematic of the EGRH-1 protein indicating the positions of the zinc-fingers (black boxes), the site of a frameshift mutation predicted in *egrh-1(tm1736)*, and the antigen used to raise an EGRH-1 antibody. (B) ClustalW alignment of the zinc-finger regions of *C. elegans* EGRH-1 (acc. CAA93744), human EGR1 (acc. P18146), EGR2 (acc. P11161), EGR3 (XP_944327), and Drosophila Stripe (CG7847-PB). (C) Schematic of the *egrh-1* gene numbered according to the C27C12 cosmid sequence (acc. # Z69883) with coding exons indicated as gray arrows. The regions encoding the zinc-fingers, the site of *gfp* orf insertion, and the region deleted by *tm1736* (bp. 16644-17282) are indicated.

3.4.2   *egrh-1* mutants exhibit oocyte defects, small broods and embryonic lethality

To characterize EGRH-1 function *in* vivo, we obtained the deletion mutant *egrh-1(tm1736)* from the National BioResource Project. This deletion removes 639 bp of genomic DNA, and is predicted to cause a frame shift mutation upstream of the EGRH-1 zinc fingers after amino acid 135 (Figure 5 A, C). Based on RNAi and rescue analyses described below, *egrh-1(tm1736)* appears to be a strong loss-of-function or null allele.

egrh-1(tm1736) mutants were morphologically normal and could be maintained as a homozygous, slow-growing strain, but they exhibited fertility defects and partially penetrant embryonic lethality. While wild type hermaphrodites produced an average of 277 self progeny (n=3; range 253-298) that included only 1% arrested embryos (n=146), *egrh-1(tm1736)* hermaphrodites produced an average of 70 self progeny (n=5; range 14-169) that included 25% arrested embryos (n=97). The arrested *egrh-1(tm1736)* embryos exhibited a variable terminal phenotype and some exhibited a round eggshell, which is observed in animals with ovulation defects (McCarter et al., 1997), but we have not characterized these embryonic phenotypes in detail.

Based on DIC images we found the gonadal morphology was normal in *egrh-1* mutants (Figure 6B). The somatic gonad contains 5 pairs of sheath cell nuclei that have distinct positions along the proximal gonad of each arm. The myoepithelial proximal sheath cell pairs 3 through 5 of the gonad contract to drive ovulation (Hirsh et al 1979; Strome 1986). DAPI and myosin heavy chain A and myosin heavy chain B antibodies were used to examine the number of sheath cell nuclei and filament organization in the somatic gonad of *egrh-1* mutants. We stained dissected gonads from wild type and *egrh-1* mutant hermaphrodites and observed 5 pairs of

sheath cell nuclei as well as correct myofilament organization in the *egrh-1* mutant (Figure 7).

To understand the basis of the *egrh-1* mutant fertility defect, we examined the germline of adult

egrh-1(tm1736) hermaphrodites. Immediately after molting to adults, the germline appeared

normal, but by 20 hours after molting 28% (n=64) of *egrh-1(tm1736)* animals contained

abnormal and degrading oocytes in the proximal gonadal arm and developing oocytes in the

distal arm, and these defects progressively worsened until 92% (n=63) of *egrh-1* animals

exhibited these phenotypes by 40 hours after molting (Figure 6; Table I). The degrading

proximal oocytes resembled the endomitotic oocytes described in mutants with defects in

oogenesis and ovulation (Iwasaki et al., 1996; McCarter et al., 1997; Rose et al., 1997), and

indeed we observed large, endomitotic nuclei in DAPI stained animals (Emo phenotype; data not

shown). The morphology of the distal oocytes in *egrh-1* mutants was similar to oocytes found in

the proximal gonadal arm of wild-type animals, but we never observed distal oocytes in wild

type hermaphrodites. We suggest that oocyte defects underlie both the low brood size and

embryonic lethality in *egrh-1(tm1736)*.

Because *egrh-1(tm1736)* may encode a truncated protein (Figure 5A), we asked if this

allele results in a loss of EGRH-1 function. We found that *egrh-1(+)* transgenes efficiently

rescued *egrh-1* mutants, and that reducing *egrh-1* activity by RNAi produced oocyte defects

nearly identical to *egrh-1(tm1736)* (Figure 6D, Table I). Importantly, we did not observe more

severe phenotypes in *egrh-1(RNAi)* animals in either a wild-type or an RNAi hypersensitive *rrf-3(pk1426)* mutant background. Taken together, these results indicate *egrh-1(tm1736)* is a strong

loss-of-function or null allele.
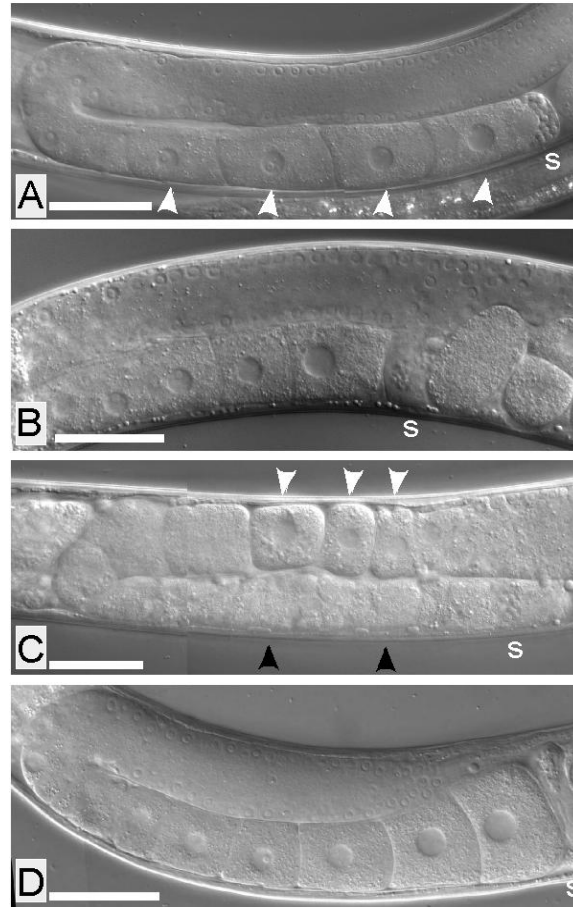
Figure 6: *egrh-1(tm1736)* **hermaphrodites accumulate degraded oocytes in the proximal gonad and differentiated oocytes in the distal gonad.**

DIC micrographs of adult hermaphrodite gonads: (A) wild type, 2 days post L4 (B) *egrh1(tm1736)*, 1 day post L4, prior to the onset of oocyte defects (C) *egrh1(tm1736)*, 2 days post L4, exhibiting degrading oocytes in the proximal gonadal arm and developing oocytes in the distal arm (D) *egrh-1(tm1736); cuEx503* carrying a wild type *egrh-1* transgene, 2 days post L4. In all cases, the distal arm is located on top, and the proximal spermatheca (s) is located in the lower right. Differentiated oocytes in the distal and proximal gonadal arms are marked by white arrowheads (A, C), and degrading oocytes in the proximal gonadal arm are marked by black arrowheads (C). Images are composites from two images to include the entire gonad, and the white bar indicates 40 µm.
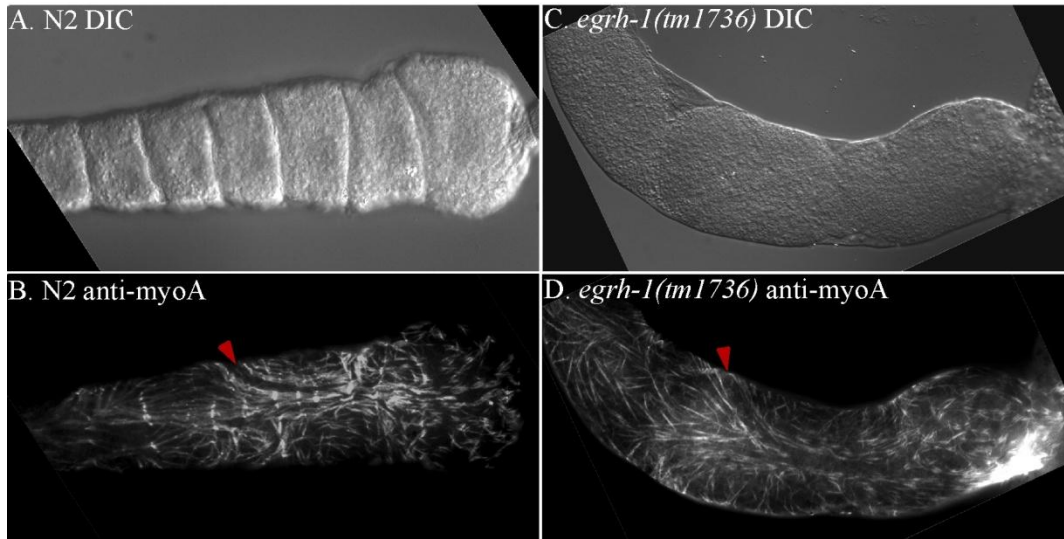
Figure 7: **Sheath filament organization appears normal in EGRH-1 mutants.**

Myosin heavy chain A antibodies were used to examine the filament organization in the somatic gonad of *egrh-1* mutants. (A,B) Image of a dissected gonad of a N2 adult hermaphrodite stained with anti-myoA to visualize wild type filament organization. Filaments cover the proximal gonad arm and are separated by distinct boundaries. Arrow head points to visible sheath cell boundary. (C,D) Image of a dissected gonad of a *egrh-1(tm1736)* adult hermaphrodite stained with anti-myoA shows filament organization looks wild type in *egrh-1(tm1736)* mutants. Filaments cover the proximal gonad arm and are separated by distinct boundaries. Arrow head points to visible sheath cell boundary. The proximal gonad is to the right in all images.

TABLE I: FREQUENCY OF ABNORMAL AND DISTAL OOCYTES

genotype	% animals exhibiting abnormal or distal oocytes (n)[a]
+/+	2% (57)
egrh-1(tm1736)	92% (63)
egrh-1(tm1736); cuEx503[C27C12(+)] [b,g]	17% (89)
egrh-1(tm1736); cuEx542[egrh-1(+)] [c,g]	18% (55)
egrh-1(tm1736); cuEx547[egrh-1::gfp] [d,g]	25% (53)
egrh-1(tm1736); cuEx531[C27C12 + sur-5::gfp] [e,g]	36% (98)
egrh-1(tm1736); cuEx562[ges-1::egrh-1(+)] [f,g]	52% (91)
egrh-1(RNAi)	42% (106)
egrh-1(RNAi); rrf-3(pk1426)	43% (203)
rrf-1(pk1417)	7% (60)
egrh-1(RNAi); rrf-1(pk1417)	12% (140)
fog-2(q71); egrh-1(tm1736)	82% (78)
fog-2(q71); egrh-1(tm1736); cuEx562[ges-1::egrh-1(+)] [f,g]	45% (61)

[a] hermaphrodites and females were scored 36-40 hr past L4 (25˚C).

[b] cuEx503 contains cosmid C27C12 (encoding egrh-1 and 6 additional transcripts) and pRF4

[c] cuEx542 contains plasmid pOK236.01 (with bp 10647-19888 of C27C12 containing only egrh-1) and pRF4

[d] cuEx547 contains the rescuing gfp translational fusion in plasmid pOK240.01 (identical to pOK236.01 with the gfp orf inserted in an NheI site in exon 5) and pRF4

[e] cuEx531 contains cosmid C27C12, sur-5::gfp, and pRF4

[f] cuEx562 contains plasmid pOK262.03 (with ges-1 promoter driving bp 658 to 2827 of EGRH-1 cDNA) and pRF4

[g] Rol transgenic animals were scored for the EGRH-1 phenotype

3.4.3 *egrh-1* mutant females exhibit constitutive MAPK activation and ovulation

In wild-type *C. elegans*, oocytes in the proximal gonad transiently arrest at the diakinesis stage of meiotic prophase I. In response to sperm-mediated signals, the proximal-most oocyte undergoes maturation, mitogen-activated protein kinase (MAPK) activation and ovulation (Miller et al., 2001; Page et al., 2001; Lee et al., 2007). MAPK is also activated in the distal gonadal arm as cells leave the pachytene stage, but this activation is sperm independent during young adult stage (Lee et al., 2007). We hypothesized that EGRH-1 may be involved in oocyte arrest, and the oocyte defects observed in *egrh-1* mutants results from inappropriate release from this arrest.

To test this hypothesis, we examined MAPK activation in the absence of sperm. Hermaphrodite sperm production is eliminated in *fog-2(q71)* and *fog-1(q253ts)* mutants, and these animals are functional females (Schedl and Kimble, 1988; Barton and Kimble, 1990). When unmated, these females' oocytes do not undergo MAPK activation (Miller et al., 2001), and their ovulation is delayed, leading to a visible accumulation of oocytes stacked in the proximal gonad. We compared MAPK activation in the double mutant strains *fog-2(q71); egrh-1(tm1736)* and *fog-1(q253ts); egrh-1(tm1736)* (which we term *egrh-1* mutant females) to that in *fog-2(q71)* and *fog-1(q253ts)* single mutants (termed wild-type females). *egrh-1* mutant females exhibited MAPK activation in proximal oocytes far more frequently than wild-type females, and in some cases activated MAPK was observed in more distal oocytes in the proximal arm (Figure 8; Table II). Sperm independent MAPK activation in the distal gonad occurred normally in both wild-type and *egrh-1* mutant females. Interestingly, *egrh-1* mutant females exhibited a delay in the accumulation of abnormal and distal oocytes compared to *egrh-1* mutant hermaphrodites and did not display these phenotypes until approximately 40 hours post L4. However, when mated

with males, this delay was eliminated, suggesting that sperm accelerate the onset of these oocyte defects.

We next asked if *egrh-1* mutant females ovulate in the absence of sperm derived signals. We attempted to observe ovulations directly, but in contrast to wild-type animals, *egrh-1* mutant females did not ovulate when mounted for microscopy. However, we found *egrh-1* females did not accumulate oocytes stacked in the proximal gonad (Figure 9C; Table II), suggesting *egrh-1* females ovulate in absence of sperm. Consistent with this conclusion, *egrh-1* mutant females contained more unfertilized oocytes ovulated into the uterus than wild-type females (Table II). When compared directly, the oocyte maturation rate for *fog-2(q71); egrh-1(tm1736)* females was significantly higher than that of *fog-2(q71); egrh-1(+)* females (Table II, $p<0.0001$). This increase is comparable to that of mutants in some negative regulators of oocyte meiotic maturation (Miller et al., 2003; Govindan et al., 2006).

Together these results indicate EGRH-1 negatively regulates oocyte MAPK activation and ovulation.
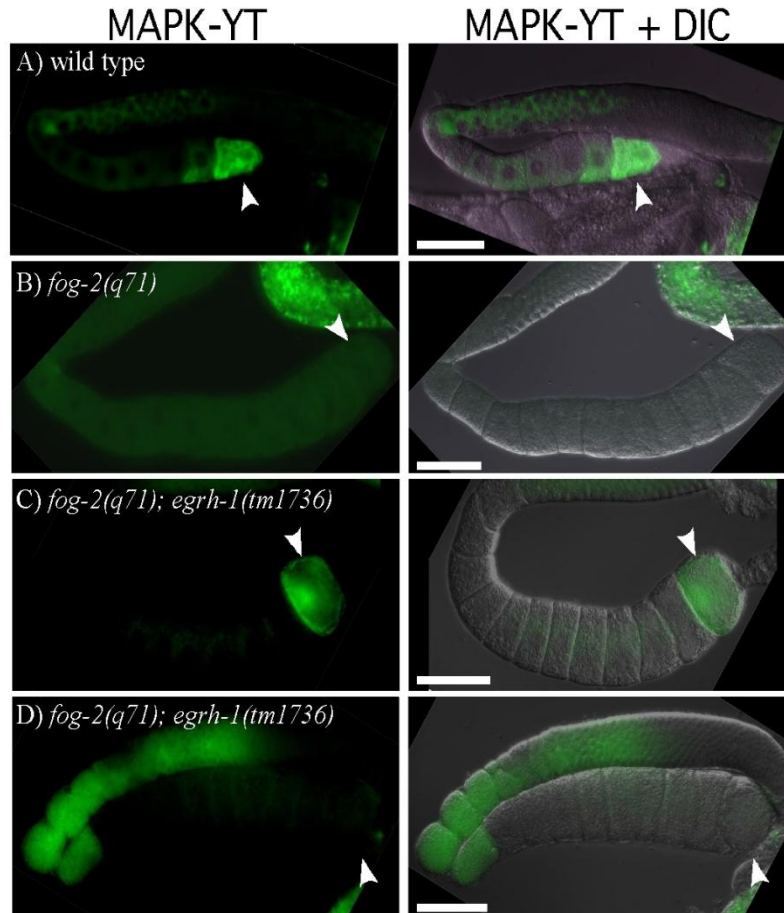
Figure 8: *egrh-1(tm1736)* **females activate MAPK in the absence of sperm.**

Fluorescence (left) and merged fluorescence/DIC (right) micrographs of dissected adult hermaphrodite gonads stained to detect the activated form of MPK-1 MAPK. Arrowheads mark the most proximal oocyte. (A) Wild-type hermaphrodite stained with MAPK-YT antibody shows activated MPK-1 MAPK in the most proximal oocytes. (B) Unmated *fog-2(q71)* females stained with MAPK-YT show no activated MAPK in oocytes. (C, D) Unmated *fog-2(q71); egrh-1(tm1736)* females show activated MAPK in the most proximal oocytes similar to wild-type hermaphrodites, as well as more distal oocytes not typically seen in wild type hermaphrodites. White bars indicate 40 µm.
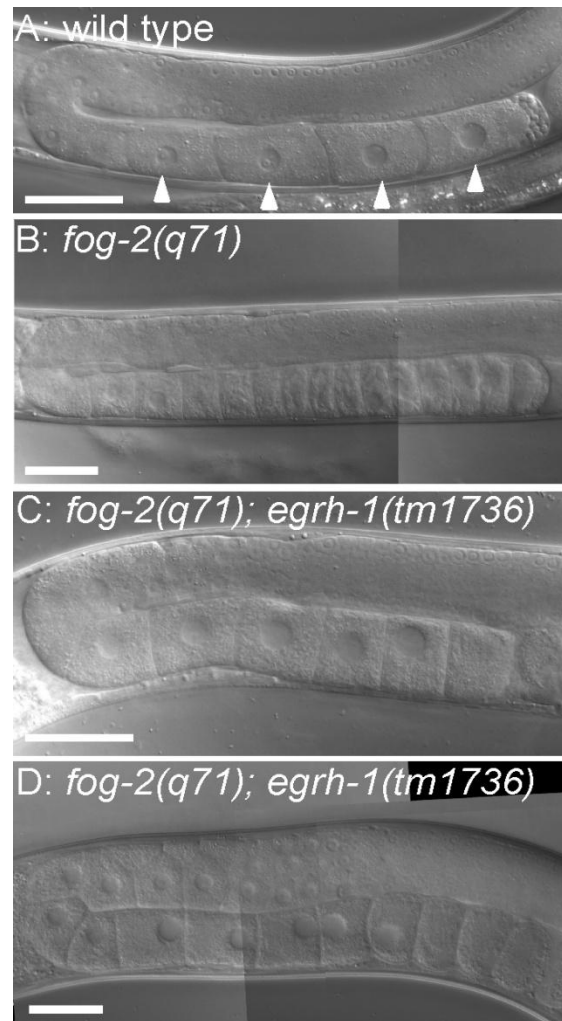
Figure 9: *fog-2(q71);egrh-1(tm1736)* **females ovulate in the absence of sperm.**

(A) Wild type hermaphrodite with oocytes in the proximal gonadal arm (arrowheads mark oocyte nuclei). **(B)** *fog-2(q71)* female with many oocytes stacked in the proximal gonadal arm. (C) *fog-2(q71); egrh-1(tm1736)* female at 1 day post L4. Unstacked oocytes appear similar to those of wild type hermaphrodites. (D) *fog-2(q71); egrh-1(tm1736)* female at 2 days post L4. The initial signs of degrading oocytes in the proximal gonadal arm and developing oocytes in the distal arm are visible. Images are composites from two images to include the entire gonad, and the white bar indicates 40 μm.

TABLE II: FREQUENCY OF OOCYTE MAPK ACTIVATION AND OVULATION

Genotype	% animals with activated oocyte MAPK (n)	% animals with oocytes stacked in proximal gonad (n)	% animals with 3 or more oocytes ovulated into uterus (n)	Oocyte maturations per gonad arm per hr
+/+	78% (49)	n.d	n.d.	n.d.
fog-2(q71)	1% (70)	100% (39)	1% (87)	0.14 ±0.10 (n=14)
fog-2(q71); egrh-1(tm1736)	67% (45)	9% (45)	86% (125)	0.62 ±0.19 (n=18)
fog-2(q71); egrh-1(tm1736); cuEx562[ges-1::egrh-1(+)] [a,b]	61% (97)	18% (33)	93% (27)	n.d.
fog-1(q253ts)	17% (24)	100% (48)	n.d.	n.d.
fog-1(q253ts); egrh-1(tm1736)	58% (90)	5% (43)	n.d.	n.d.

[a] *cuEx562* contains plasmid pOK262.03 (with the *ges-1* promoter driving expression of an cDNA) and pRF4

[b] Rol transgenic animals were scored for the EGRH-1 phenotype

n.d. not determined

3.4.4   *egrh-1* mutant germ cells progress through normal developmental stages

Germ cells are continuously produced in adult hermaphrodites, and a progression of distinct developmental stages can be recognized in DAPI stained preparations along the distal to proximal axis (Kimble and Crittenden, 2005; Hansen and Schedl, 2006). Near the distal tip, a region of mitotic germ cells is termed the proliferative zone. As germline nuclei progress proximally within the distal arm they exit the mitotic cell cycle and enter meiosis in the transition zone and subsequently enter pachytene stage of meiosis. As nuclei exit the pachytene region, MAPK is activated and these cells enlarge in the bend region as they leave the distal arm and eventually differentiate as oocytes arrested in meiotic prophase I in the proximal gonad arm.

To determine if the progression of germ cell development is altered in *egrh-1* mutants, we compared the morphology and number of germ cell nuclei in different regions of the gonad in wild-type and *egrh-1* mutant females. In general, germ cell progression was normal in *egrh-1* mutants. The morphology of germ cell nuclei appeared normal in each of the regions in *egrh-1(tm1736)* mutants, and MAPK activation occurred near the end of the pachytene region. However, *egrh-1* mutants exhibited more variability in the number of nuclei in the transition zone and pachytene region, and the average number of pachytene nuclei was reduced (Table III).

TABLE III: NUMBER OF GERMLINE CELL DIAMETERS IN THE DIFFERENT REGIONS OF THE GONAD

Genotype [a]	Proliferative Zone (n) [b]	Transition Zone (n) [b]	Pachytene region (n) [b]
fog-2(q71)	14 ± 2.2 (5)	8.2 ± 1.1 (5)	33.8 ± 1.9 (5)
fog-2(q71); egrh-1(tm1736)	12.8 ± 1.5 (5)	10.6 ± 3.0 (5)	26.8 ± 6.1 (5) [c]

Cell diameters were measured by counting the number of nuclei in a line along the distal-proximal axis

[a] Young adults18-20 hours post L4

[b] The Proliferative zone is distal to the Transition Zone, the Transition zone begins where 80% of nuclei have crescent shaped morphology and ends where 80% of nuclei appear pachytene (Eckmann et al., 2004), the Pachytene region is proximal to the Transition zone and show a thread like appearance (Francis et al., 1995; MacQueen and Villeneuve, 2001)

[c] This value is different from wild-type at 95% confidence (p=0.04)

3.4.5   <u>EGRH-1 is expressed in the distal tip cell and sheath cells of the somatic gonad</u>

To identify *egrh-1*-expressing tissues, we first examined transgenic animals bearing a rescuing *egrh-1::gfp* translational fusion (Figure 5C; Table I). *egrh-1::gfp* was expressed in embryos through adult stages in most somatic tissues, including the gut, pharynx, body wall muscle, and nervous system (data not shown), suggesting *egrh-1* is broadly expressed like mammalian Egr1 (McMahon et al., 1990; Watson and Milbrandt, 1990). To understand EGRH-1 function in regulating oocyte development, we examined *egrh-1::gfp* expression and endogenous EGRH-1 expression using an antibody targeting a unique region at the EGRH-1 N-terminus in dissected adult hermaphrodites. Using both techniques, EGRH-1 expression was observed in the sheath cells and distal tip cells of the somatic gonad, as well as in the intestine and sperm (Figure 10). EGRH-1 was localized to nuclei in all tissues except sperm where it appeared perinuclear. Importantly, no *egrh-1::gfp* expression or EGRH-1 antibody staining was observed in oocytes, suggesting EGRH-1 functions in other tissues to regulate oocyte maturation and ovulation. To test if EGRH-1 antibody staining was specific, we stained *egrh-1(tm1736)* mutants, which are predicted to encode a truncated protein that would be recognized by our antibody and *egrh-1(tm1736)* with *egrh-1* expression further reduced by RNAi. Staining was decreased in both cases, indicating this antibody stain is specific for EGRH-1 (Table IV).

EGRH-1 function does not appear to be required in sperm. Wild-type and *egrh-1(tm1736)* males produce similar numbers of cross progeny when mated with *fog-1(q253)* females [+/+ males, average cross progeny 206 (n=9); *egrh-1(tm1736)* males, average cross progeny 143 (n=9)] , and these cross progeny exhibit similar low levels of embryonic lethality [wild-type, 9.0% Emb (n=556); *egrh-1(tm1736)*, 11.9% Emb (n=616)].
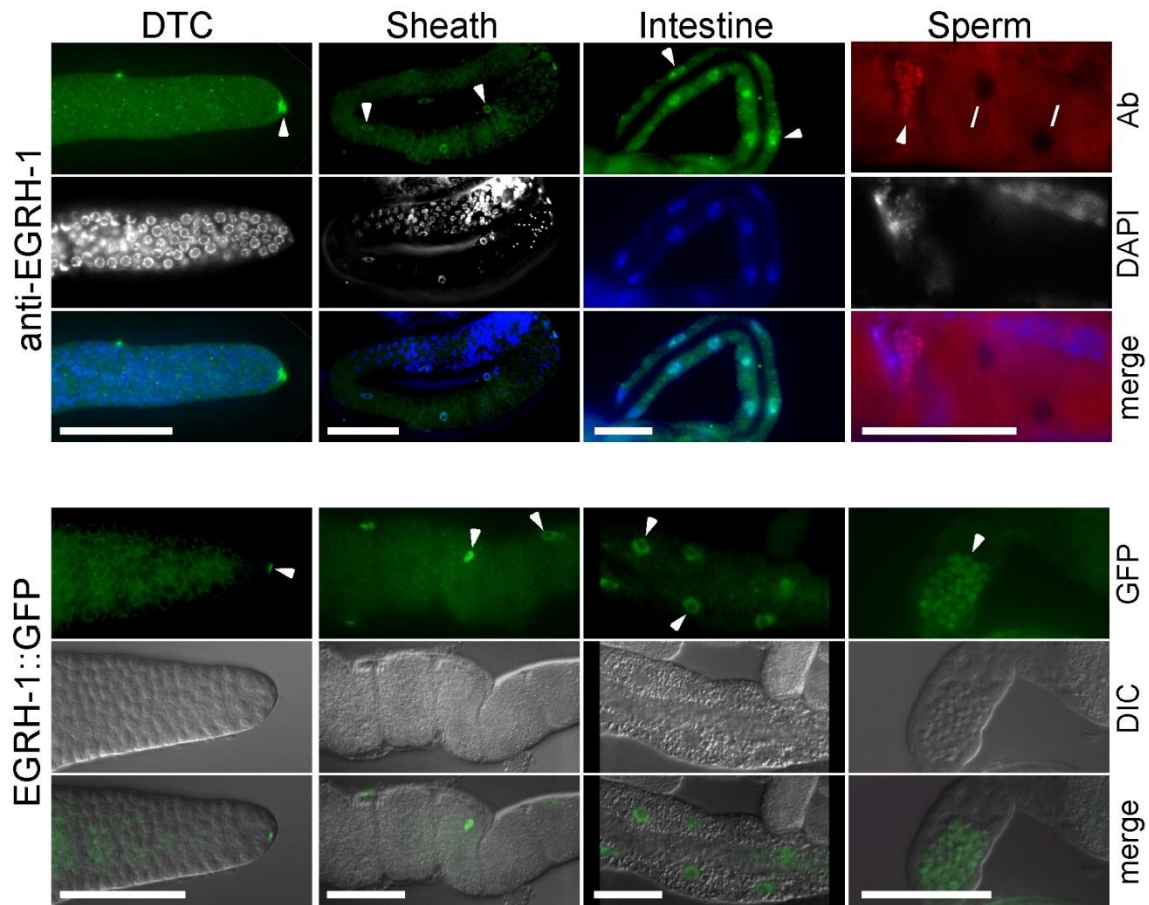
Figure 10: EGRH-1 antibody staining and EGRH-1::GFP expression in the somatic gonad and intestine.

EGRH-1 antibody staining (top panel) and EGRH-1::GFP expression (bottom panel) in nuclei of the distal tip cell (DTC) and sheath cells of the somatic gonad and intestinal cells. Arrowheads mark representative EGRH-1 containing nuclei. White lines point out oocyte nuclei lacking detectable EGRH-1 staining. For anti-EGRH-1, antibody stained and DAPI stained images are shown individually and merged. For EGRH-1::GFP, GFP fluorescence and DIC images are shown individually and merged. Distal tip cell staining with EGRH-1 antibody co-localized with *lag-2::gfp* expression which marks this cell type (data not shown) (Siegfried and Kimble, 2002). White bars indicate 40 μm.

TABLE IV: FREQUENCY OF EGRH-1 ANTIBODY STAINING IN DISSECTED SOMATIC GONADS

genotype	% distal tip cell staining (n)	% sheath cell staining (n)
+/+	93% (110)	91% (93)
egrh-1(tm1736) [a]	69% (42)	60% (40)
egrh-1(tm1736); egrh-1(RNAi)	28% (43)	7% (70)

[a] *egrh-1(tm1736)* is predicted to encode a truncated protein that would be detected by this targeted antibody (Figure 5A)

3.4.6   *fog-1(q253ts); egrh-1(tm1736)* oocytes are defective in sperm recruitment

During mating males deposit sperm in the uterus of hermaphrodites and females. Sperm will then crawl through the uterus to the spermatheca where they will be available to fertilize the next oocyte. We used MitoTracker to label wild type sperm and *egrh-1(tm1736)* sperm to analyze *egrh-1(tm1736)* male sperm motility and to analyze sperm recruitment in *fog-1(q253); egrh-1(tm1736)* females (Figure 11). When crossed with wild type females all of *egrh-1(tm1736)* labeled male sperm are located within the spermatheca within one hour. Indicating that *egrh-1(tm1736)* male sperm are capable of responding to oocyte signals. When *fog-1(q253); egrh-1(tm1736)* females are crossed with wild type male sperm, we found that after one hour some of the labeled sperm are located in the spermatheca however much of the sperm is still dispersed throughout the uterus (Figure 11E,F). This result indicates that *fog-1(q253ts); egrh-1(tm1736)* oocytes are defective in sperm recruitment.
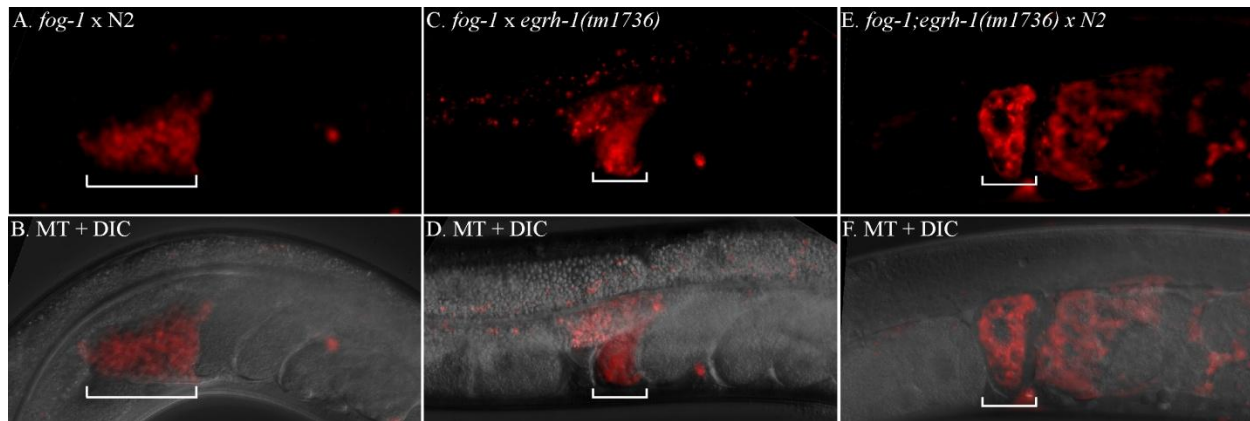
Figure 11: *egrh-1* **sperm function like wild type**.

MitoTracker Red CMXRos (Invitrogen) was used to label wild type sperm and *egrh-1(tm1736)* sperm. Males were then mated to *fog-1(q253); egrh-1(tm1736)* females to analyze *egrh-1(tm1736)* male sperm motility and to analyze sperm recruitment. In all images the proximal gonad is to the left and the uterus is to the right. Sperm is identified by red fluorescence and the spermatheca is marked by the white bracket. (A,B) 1 hr after mating with a wild type *fog-1* female, N2 sperm is located in the spermatheca. (C,D) 1 hr after mating with a wild type *fog-1* female, *egrh-1(tm1736)* sperm is located in the spermatheca. (E,F) 1 hr after mating with a *fog-1;egrh-1(tm1736)* female, N2 sperm is located in the spermatheca and dispersed throughout the uterus.

3.4.7   EGRH-1 function is required in the gut and somatic gonad
     Based upon the *egrh-1(tm1736)* mutant phenotype and EGRH-1 expression pattern, we

hypothesized EGRH-1 functions in the soma to inhibit maturation and ovulation. To test this

hypothesis, we first examined the effect of *egrh-1(RNAi)* in an *rrf-1(pk1417)* mutant background.

rrf-1(pk1417) mutants are specifically resistant to RNAi in the soma but remain sensitive to

RNAi in the germline (Sijen et al., 2001). We found *egrh-1(RNAi)* was less effective in *rrf-*

1(pk1417) than in a wild-type background (Table I). Indeed, *rrf-1(pk1417); egrh-1(RNAi)*

animals exhibited only a slight increase in the frequency of oocyte defects when compared to

untreated *rrf-1(pk1417)* mutants. Thus we conclude *egrh-1* functions in the soma.

     We next examined the phenotype of *egrh-1(tm1736)* animals that were mosaic for a

rescuing transgene [(Mello and Fire, 1995); see Materials and Methods]. Transformed DNAs in

C. elegans recombine to form extrachromosomal arrays that are partially unstable in mitosis. We

generated the strain *egrh-1(tm1736); cuEx531* containing the *egrh-1(+)* cosmid C27C12 and the

cell autonomous marker for somatic cells *sur-5::gfp* (Yochem and Herman, 2003). This

transgene rescued *egrh-1(tm1736)* mutants (Table I) but was lost at low frequency during early

mitotic cell divisions. We first identified animals that specifically lost this transgene in the EMS

blastomere, which is the precursor of the somatic gonad and gut, but not the germline, and 95%

of these animals exhibited abnormal and distal oocytes (Figure 12). Thus *egrh-1* function is

required in somatic tissues derived from EMS. While we were not able to find any animals that

lost the transgene only in the germline precursor P2, we did find one animal that lost the

transgene in P2 but had a second loss later in the E lineage. This animal exhibited a wild-type

phenotype suggesting *egrh-1* function is not required in the germline. We next identified mosaic

animals that lost the transgene only in MS or E, the descendants of EMS that give rise to the

somatic gonad or gut, respectively. Loss in either MS or E resulted in the *egrh-1* phenotype, strongly suggesting *egrh-1* function is required in both the somatic gonad and the gut for normal oocyte development.

We next expressed *egrh-1* in the E lineage using the *ges-1* promoter and asked if gut expression is sufficient to rescue *egrh-1* mutants. The *ges-1* promoter is active exclusively in the E lineage from early embryogenesis throughout the life of the worm (Kennedy et al., 1993). We found that *egrh-1* expressed using a *ges-1::egrh-1* fusion gene did not rescue sperm independent ovulation and MAPK activation (Table II). We suggest *egrh-1* in the sheath cells of the somatic gonad negatively regulates oocyte MAPK activation and ovulation. However, we do not have a well-characterized sheath cell promoter to drive *egrh-1* expression in these cells to test this hypothesis. In contrast, *egrh-1* expression in the E lineage partially rescued the formation of abnormal and distal oocytes in both *egrh-1* mutant hermaphrodites and *egrh-1* mutant females (Table I), although this result may reflect a role for *egrh-1* in other aspects of oocyte development.
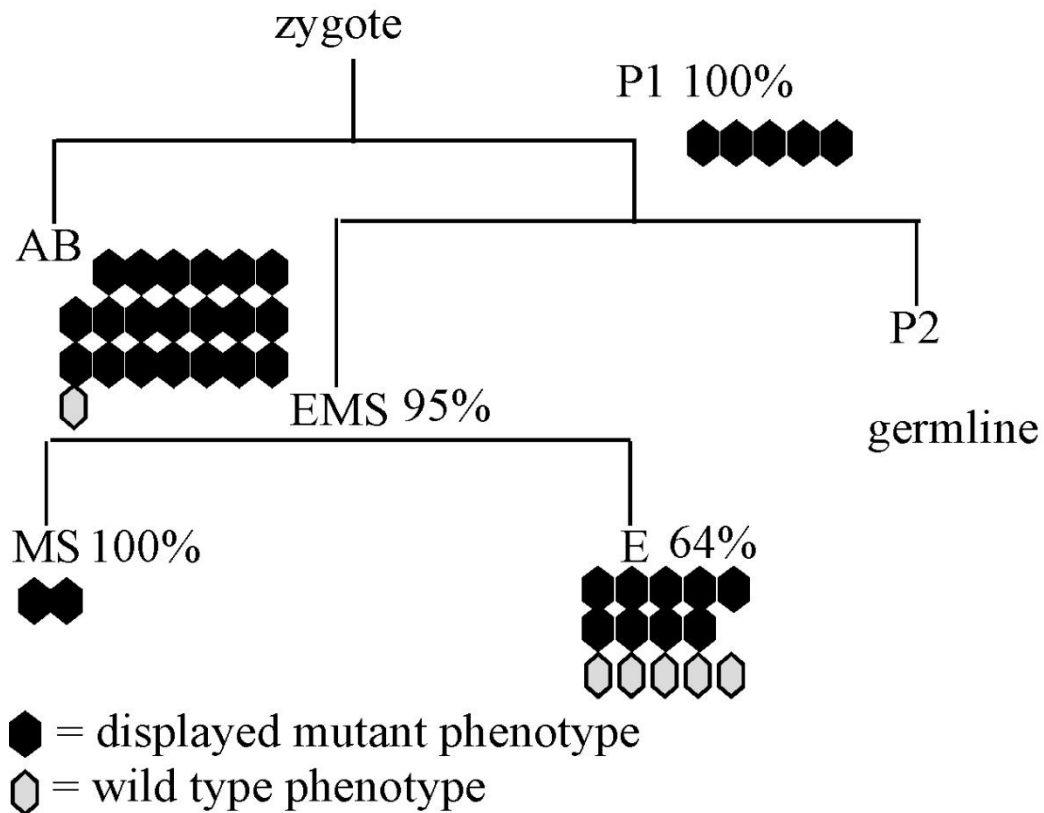
Figure 12: **Analysis of *egrh-1* genetic mosaics**

Diagram of the early *C. elegans* cell lineage indicating the phenotypes of *egrh-1* genetic mosaics that lost a rescuing transgene in specific sub-lineages. Black indicates individual transgenic animals that lost GFP expression in the indicated lineages and displayed abnormal and degraded oocytes in the proximal and distal gonadal arms. Light gray indicates transgenic animals which lost GFP expression in these lineages and exhibited wild type oocyte development. Numbers indicate the percent animals that lost the transgene in this sub-lineage that exhibited the mutant phenotype.

3.5   Discussion

In this paper, we characterize the role of the *C. elegans* EGR-family factor EGRH-1

during oogenesis. We find that *egrh-1* mutants exhibit ectopic oocyte development in the distal

gonad and sperm independent oocyte maturation and ovulation. EGRH-1 is broadly expressed

both temporally and spatially, and tissue specific RNAi analysis indicates it functions in the

soma to regulate oogenesis. Mosaic analysis indicates that EGRH-1 function is likely required in

the somatic gonad, which is descended from the MS blastomere, and more surprisingly in the

gut, which is descended from the E blastomere. Expression of *egrh-1* in the gut rescues the

formation of ectopic oocytes in *egrh-1* mutants, but does not rescue sperm independent oocyte

maturation and ovulation. Together, these results suggest that EGRH-1 regulates cell non-

autonomous mechanisms controlling multiple aspects of oocyte development, and that EGRH-1

has distinct functions in the sheath cells and gut.

3.5.1   EGRH-1 negatively regulates oocyte maturation and ovulation

In wild-type *C. elegans* hermaphrodites, oocyte maturation and ovulation depend on

signals between the oocytes, the sheath cells of the somatic gonad, and sperm (Greenstein, 2005).

In the absence of sperm, sheath cell-derived signals inhibit oocyte maturation and ovulation to

preserve oocytes for fertilization. This inhibition depends on gap junctions between the sheath

cells and the oocytes (Govindan et al., 2006; Whitten and Miller, 2007). When sperm are

present, sheath cell inhibition is overridden by major sperm proteins (MSPs), which are secreted

from sperm and bind the VAB-1/Eph receptor in oocytes and unidentified receptors in sheath

cells, allowing oocyte maturation and ovulation (Miller et al., 2001; Miller et al., 2003;

Govindan et al., 2006). Oocyte maturation coincides with the phosphorylation and activation of

the MPK-1 MAPK protein in proximal oocytes and MAPK activation is necessary for maturation

and ovulation. MAPK is typically highly activated in the most proximal oocyte, but this activation is rapidly lost as this oocyte completes maturation and is ovulated (Lee et al., 2007). Maturation leads to degradation of specific oocyte proteins allowing the transition from differentiated oocyte to totipotent zygote (Nishi and Lin, 2005; Shirayama et al., 2006; Stitzel et al., 2006).

egrh-1 mutants exhibit MPK-1/MAPK activation in proximal oocytes and ovulation independently of sperm signals, indicating that EGRH-1 negatively regulates oocyte maturation and ovulation. As in wild-type hermaphrodites, MPK-1/MAPK activation appeared transient and was occasionally observed in more distal oocytes. We suggest this transient MAPK activation in *egrh-1* mutant oocytes inappropriately initiates the transition from oocyte to zygote, and this contributes to the partially penetrant embryonic lethal phenotype.

A number of other *C. elegans* mutants have been described that exhibit sperm independent MAPK activation and ovulation. In particular, mutants affecting the CEH-18 POU-homeodomain transcription factor, which is expressed similarly to EGRH-1 in sheath cells, represses oocyte maturation (Greenstein et al., 1994; Rose et al., 1997). *ceh-18* is necessary for normal sheath cell differentiation and mutants exhibit disorganized myofilaments in sheath cells. We have not observed myofilament disorganization in the *egrh-1* mutant sheath cells, suggesting EGRH-1 and CEH-18 play distinct roles in inhibiting oocyte maturation.

3.5.2 EGRH-1 mutants accumulate differentiated oocytes in the distal gonad

Our analyses indicate that *egrh-1* function is necessary and sufficient in the soma for normal oocyte development, and that *egrh-1* function is necessary in the descendents of both the MS blastomere, which include the somatic gonad, and the E blastomere that gives rise to the gut.

These conclusions are consistent with the expression of the rescuing *egrh-1::gfp* transgene and accumulation of EGRH-1 protein in somatic tissues but not in oocytes.

egrh-1 mutants accumulate differentiated oocytes in the distal gonadal arm, and this phenotype can result from loss of *egrh-1* either in the somatic gonad or the gut. Loss of *egrh-1* in the somatic gonad leads to oocyte degradation and blockage of the proximal arm that interferes with ovulation, and distal oocytes may form as a secondary consequence of defective ovulation. A similar phenotype can be observed in other Emo mutants at late times (D. Greenstein personal communication). We do not understand how *egrh-1* loss in the gut leads to formation of distal oocytes. The gut produces complexes containing yolk proteins and precursors for signaling molecules involved in sperm recruitment that are transported into the psuedocoelom and are ultimately endocytosed by oocytes (Kimble and Sharrock, 1983; Hall et al., 1999; Kubagawa et al., 2006a). *rme-2* mutants defective in endocytosis of yolk complexes exhibit ovulation defects (Grant and Hirsh, 1999). We have observed yolk accumulation in *egrh-1* mutants, but perhaps EGRH-1 is necessary for synthesizing other components of yolk complexes necessary for normal ovulation, and loss of these components affects ovulation and leads indirectly to formation of distal oocytes.

3.5.3   Oocytes in EGRH-1 mutants are defective in sperm recruitment

During mating males deposit sperm in the uterus of hermaphrodites and females. Sperm will then crawl through the uterus to the spermatheca where they will be available to fertilize the next oocyte (Ward and Carrel, 1979). Our analysis shows that *fog-1(q253ts); egrh-1(tm1736)* oocytes are defective in sperm recruitment. Polyunsaturated fatty acids (PUFAs) function in oocytes to direct sperm towards the spermatheca, and loss of these PUFAs causes sperm to disperse throughout the uterus (Kubagawa et al., 2006a). PUFAs are transported from the

intestine to the oocytes yolk. We found that EGRH-1 is required in the gut for normal oocyte development. Therefore it is possible that EGRH-1 mutant oocytes may have defects in the development or transport of these PUFAs to the oocytes. However it is also possible that the defects in sperm recruitment may also be a secondary effect from oocyte degradation found in EGRH-1 mutants.

3.5.4   Relationship to EGR family in other species

In mammalian cells, EGR-family expression is induced by a variety of stimuli, including growth factor and cytokine signaling, and expression is regulated in response to MAPK pathway signaling (O'Donovan et al., 1999). EGR proteins in turn regulate expression of a variety of transcriptional regulatory proteins and signaling molecules (Fu et al., 2003; Virolle et al., 2003). Thus EGR family proteins function at the convergence of signaling pathways. In *C. elegans* MPK-1 is activated in the sheath cells of the somatic gonad (Lee et al., 2007), and *egrh-1* expression could be regulated in this tissue by activated MPK-1.

Mammalian EGR-family genes are broadly expressed, and mutants defective in EGR-family members exhibit a variety of defects (O'Donovan et al., 1999).   Egr1 mutant mice have fertility defects and fail to ovulate due to loss of luteinizing hormone-ß (LHß) expression in the pituitary (Lee et al., 1996; Topilko et al., 1998), and its function is partially redundant with Egr4 (Tourtellotte et al., 2000). In comparison, Egr3 mutant mice do not have fertility defects but exhibit defects in muscle spindle stretch receptors (Tourtellotte and Milbrandt, 1998; Tourtellotte et al., 2001). While *C. elegans* does not have an LHß ortholog, EGRH-1 may function like mouse Egr1 by regulating expression of secreted factors that have non-autonomous effects on oocyte development.

IV. IDENTIFYING TARGETS OF CAENORHABDITIS ELEGANS T-BOX TRANSCRIPTION FACTOR TBX-2 THROUGH MICROARRAY ANALYSIS

4.1   Abstract

T-box transcription factors are important developmental regulators, and they have been more recently implicated in a variety of human diseases and cancers. Despite their importance relatively few direct targets of T-box transcription factors have been identified. *C. elegans* TBX-2 is required for the development of ABa-derived pharyngeal muscles and is the sole *C. elegans* member of the conserved Tbx2 sub-family of T-box factors, which in other organisms includes both transcriptional activators and repressors. We are interested in identifying TBX-2 targets that function in pharyngeal muscle development, and in determining if TBX-2 is a transcriptional activator or repressor. To identify targets of TBX-2, we have compared mRNA expression levels in wild-type and *tbx-2(bx59)* mutant embryos using Affymetrix microarrays. Of 19,885 probe sets examined, we found 980 mRNAs that were significantly up-regulated in *tbx-2(bx59)* relative to wild-type and 175 mRNAs that were significantly down-regulated. Using clustering analysis, comparisons to existing data sets on pharyngeal gene expression, and phylogenetic foot-printing to identify promoters with consensus T-box factor binding sites, we have analyzed a subset of genes and identified D2096.6 as a direct target of TBX-2. Results from our microarray and semi-quantitative PCR analysis indicate D2096.6 is repressed by TBX-2. GFP promoter fusions indicate the spatial and temporal expression pattern of D2096.6 is directly regulated by TBX-2 at a variant T-box binding site in the D2096.6 promoter.

4.2.   Introduction to the *C. elegans* pharynx

4.2.1   Anatomy of the *C. elegans* pharynx

The pharynx is a bilobed feeding organ located at the anterior end of the digestive system (Figure 13A). The anterior region of the pharynx is connected to the mouth by the buccal cavity and the posterior end of the pharynx is connected to the intestine by the pharyngeal-intestinal valve. The pharynx can be subdivided into sections, the anterior structure is the procorpus, followed by metacorpus (anterior bulb), isthmus and at the posterior the terminal bulb. The pharynx is composed of 80 cells that can be grouped into five types: muscles, epithelia, neurons, glands and marginal cells (Figure 13B) (Albertson and Thomson, 1976). One of the advantages of studying the pharynx is that the morphology and position of each cell has been characterized, which allows for cell identification in DIC without the use of additional markers (Albertson and Thomson, 1976).
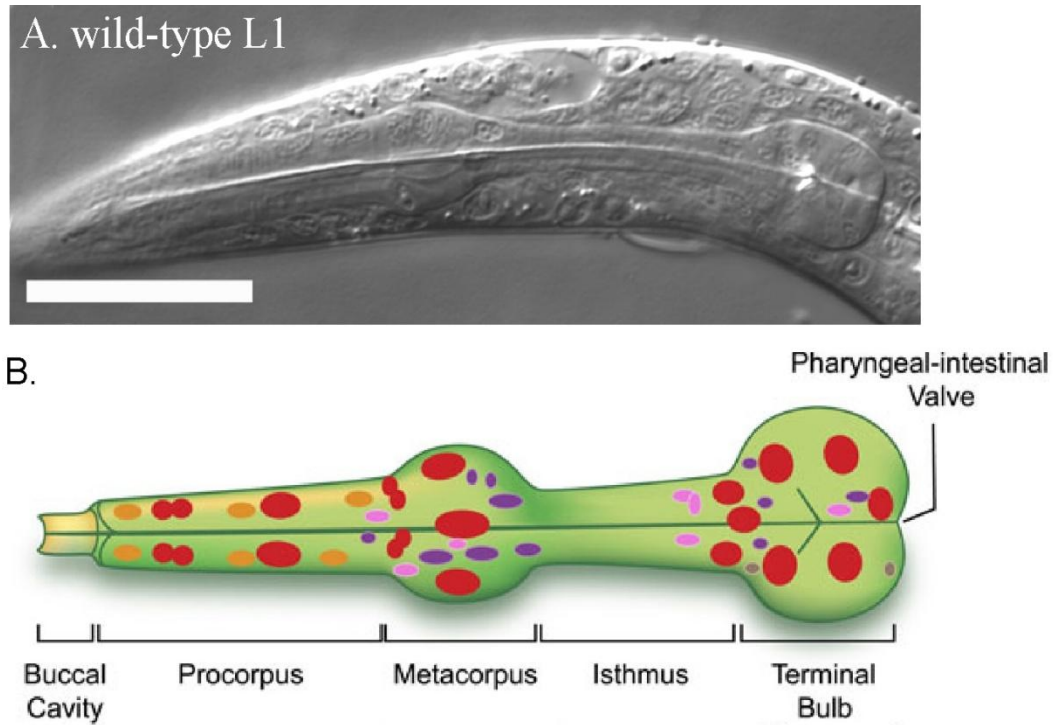
Figure 13: The structure of the *C. elegans* pharynx

(A) DIC image of a wild type L1 hermaphrodite pharynx. (B) Illustration of a wild type pharynx. The pharynx is divided into 5 sections, the buccal cavity, procorpus, metacorpus (anterior bulb), isthmus and terminal bulb. The positions of 4 types of pharyngeal nuclei are shown; muscles (red), neurons (purple), epithelia (orange) and marginal cells (pink). Illustration modified from http://www.wormbook.org.

4.2.2   Pharyngeal development

Precursor cells that give rise to the pharynx accumulate into a cell primordium (Sulston et al., 1983; Mango et al., 1994). The pharyngeal primordium consists of a group of undifferentiated cells which forms during mid-embryogenesis. The primordium is surrounded by a basement membrane which separates the pharynx from the rest of the worm (Albertson and Thomson, 1976). After formation of the pharyngeal primordium cells no longer divide, but the pharynx continues to elongate and develop in size as the worm develops (Sulston et al., 1983)

4.2.3   Anterior and posterior pharynx are specified by two distinct pathways

The pharynx is produced from two blastomeres at the four cell stage, ABa and EMS (Figure 14) (Sulston et al., 1983). ABa primarily gives rise to cells in the anterior pharynx, while EMS gives rise to cells in the posterior pharynx. In the mature pharynx several cells produced by EMS appear identical to those produced by ABa and have been found to express some of the same pharyngeal specific genes making the distinction between the two sometimes difficult (Bowerman et al., 1992). Pharyngeal precursor cells become committed to pharyngeal fate through regulation by multiple transcription factors which include PHA-4 and the T-box transcription factors TBX-2, TBX-35, TBX-37 and TBX-38 (Mango et al., 1994; Good et al., 2004; Broitman-Maduro et al., 2006; Roy Chowdhuri et al., 2006; Smith and Mango, 2007). PHA-4 is a Forkhead box A transcription factor that has been shown to be essential for development of pharynx derived from both ABa and EMS, and is thought to regulate most pharyngeal genes (Mango et al., 1994; Horner et al., 1998; Kalb et al., 1998; Gaudet and Mango, 2002). In *pha-4* mutants pharyngeal cells are transformed into an ectodermal cell type (Mango et al., 1994; Horner et al., 1998). Different T-box proteins have been implicated in both ABa and EMS derived pharyngeal development.

ABa and EMS will give rise to anterior and posterior pharyngeal cells through two distinct pathways. Pharyngeal fate in the ABa lineage is specified non-autonomously. At the 4-cell stage pharyngeal cells that develop from the ABa pathway depend on intercellular signaling between blastomeres and on *glp-1*, a Notch receptor ortholog. At the 24-cell stage two T-box proteins, TBX-37 and TBX-38, comprise a pair of redundant transcription factors that are activated and required for a subset of ABa descendents to develop into pharynx (Good et al., 2004). Both TBX-37/38 function and Notch signaling are needed to activate *pha-4*, which initiates pharyngeal development (Mango et al., 1994; Kalb et al., 1998; Good et al., 2004).

Pharyngeal development via the EMS pathway is specified autonomously and depends on the activity of two maternal genes *skn-1* and *pop-1* (Bowerman et al., 1992; Lin et al., 1995). *skn-1* encodes an bZIP-related transcription factor that functions autonomously to specify the EMS blastomere (Bowerman et al., 1992; Bowerman et al., 1993). *pop-1*, a member of the TCF/LEF family of transcription factors, is found in the nucleus of a daughter of EMS, called MS, and specifies MS derived pharyngeal cells (Lin et al., 1995; Lo et al., 2004). In the MS descendants the T-box protein, TBX-35, is required for the production of posterior pharyngeal cells (Broitman-Maduro et al., 2006). Loss of *tbx-35* function in worms results in arrested embryos or arrested L1 larvae. Mutants were missing MS derived pharynx and body muscle and were found to produce ectopic muscle and hypodermis normally produced by the C blastomeres (Broitman-Maduro et al., 2006).
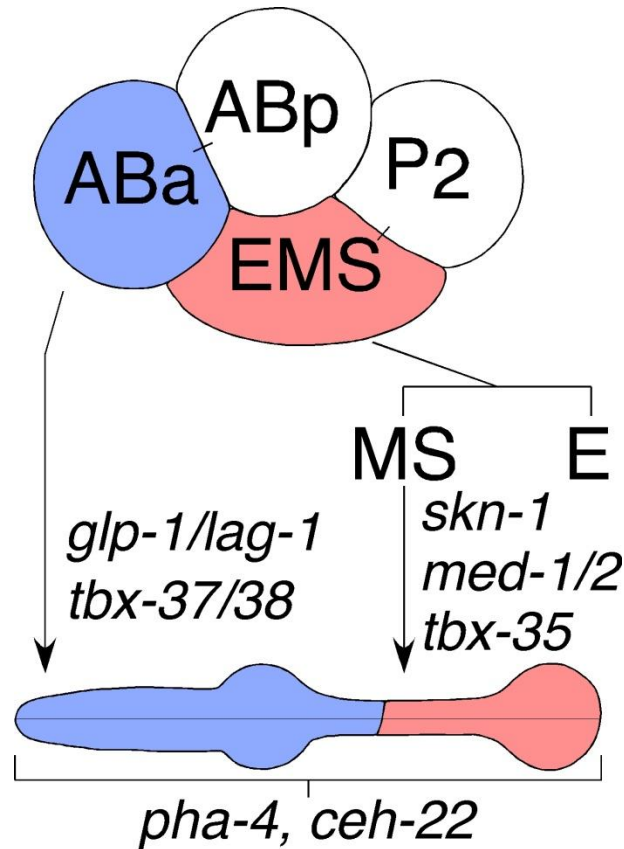
Figure 14: **The pharynx is produced from two blastomeres at the four cell stage.**

At the 4-cell stage two blastomeres ABa and EMS give rise to the anterior and posterior pharynx through two distinct pathways. (blue) ABa specifies the anterior pharynx non-autonomously and requires the activities of *glp-1* signaling and the T-box transcription factors *tbx-37/38*. (pink) EMS specifies the posterior pharynx autonomously through the activities of the maternal gene *skn-1* and the transcription factors *med-1/2* and *tbx-35*.

4.2.4   TBX-2 specifies anterior pharyngeal muscle

T-box transcription factors are important developmental regulators and are found in all multicellular organisms (Showell et al., 2004). T-box proteins are identified by a highly conserved 180-200 amino acid DNA binding domain, known as the T-domain, and are grouped into subfamilies based on their T-box protein sequence (Papaioannou, 2001). *C. elegans* TBX-2 is the sole member of the Tbx2 subfamily which consists of Tbx2, Tbx3, Tbx4 and Tbx5 in mammals (Papaioannou, 2001). The TBX-2 T-box domain shares 73% amino acid identity with human Tbx2 and Tbx3; 61% with Tbx4 and 60% with Tbx5 protein (Roy Chowdhuri et al., 2006).

It has been previously shown that TBX-2 is specifically required for the development of ABa derived pharyngeal muscle cells (Roy Chowdhuri et al., 2006; Smith and Mango, 2007). *tbx-2(ok529)* homozygotes and reduction of *tbx-2* by RNAi in a RNAi sensitive background results in larval lethality. Mutant worms display a long buccal cavity or unattached pharynx due to loss of anterior pharyngeal muscles (Roy Chowdhuri et al., 2006; Smith and Mango, 2007). *ceh-22* is a homeobox gene and is the earliest marker of pharyngeal muscle development. *tbx-2(ok529)* and *tbx-2(RNAi)* animals both have a reduction in the number of *ceh-22::gfp* expressing cells, indicating the presence of fewer pharyngeal muscle cells (Roy Chowdhuri et al., 2006). *ceh-22* is expressed in 7 ABa derived pharyngeal muscles and in 14 MS derived pharyngeal muscles (Sulston et al., 1983). The *tbx-2* mutant phenotype as well as reduction of *ceh-22::gfp* expression in *tbx-2* mutants is consistent with a loss specifically in the ABa derived pharyngeal muscles. In contrast mutant worms retained ABa marginal cells indicating that *tbx-2* is not required for all ABa pharyngeal cells (Roy Chowdhuri et al., 2006; Smith and Mango, 2007).

Pharyngeal cells in *tbx-2* mutants initiate anterior muscle development normally but then differentiation stops in the absence of *tbx-2* activity (Smith and Mango, 2007). This is based on the observation that *pha-4::gfp* was expressed normally in ABa descendents in TBX-2 mutant embryos (Smith and Mango, 2007). This result indicates that *glp-1* signaling and *tbx-37/tbx-38* were functional in *tbx-2* mutants, because both of these factors are needed for PHA-4 activation in ABa descendants (Kalb et al., 1998; Good et al., 2004; Smith and Mango, 2007). Later in development *pha-4* expression and *ceh-22* expression is reduced; suggesting *tbx-2* functions to maintain *pha-4* and activates *ceh-22*, which results in ABa muscle development (Smith and Mango, 2007).

A full length *tbx-2::gfp* reporter consisting of the entire *tbx-2* gene and 4 kb of 5' flanking DNA is expressed in a dynamic pattern. Expression initiated in 2 anterior cells in a 100 cell embryo and increased to 12 cells by the 200 cell stage. Expression was found in a subset of pharyngeal precursors prior to the bean stage of development only. Expression in body wall muscles and pharyngeal neurons continued into larval stages (Roy Chowdhuri et al., 2006). This construct does not rescue the *tbx-2(ok529)* mutant phenotype. Results in our laboratory suggest that *tbx-2* is expressed in more cells than what is seen with this GFP reporter, including hypodermal cells and intestine. Suggesting that this construct is missing additional regulatory elements and these elements may be needed for rescue (Clary and Okkema unpublished and Milton and Okkema unpublished).

A *tbx-2::gfp* translational fusion containing 5.2 kb upstream sequence has been constructed and analyzed (Smith and Mango, 2007). This reporter showed expression beginning at ~100 cell stage of embryogenesis similar to what was seen in our GFP fusion described above.

At the 1 ½ fold stage GFP was seen in the pharyngeal muscle (pm) cells 3-5. Expression was also seen outside the ABa lineage after the 1 ½ fold stage in pm8, pharyngeal neurons and body wall muscles. Pharyngeal muscle expression persisted into later stages which is unlike the translational fusion described above which showed no expression in pharyngeal precursors at the bean stage of development. In addition the 5.2 kb translational fusion reportedly was capable of rescuing the pharyngeal defects in *tbx-2(ok529)* mutants but worms did not reach adulthood (Smith and Mango, 2007). These results suggest that there is an additional element located within 4 to 5.2 kb of upstream sequence needed for expression, however, there are still additional elements missing since full rescue of *tbx-2(ok529)* was not seen.

4.2.5   Additional functions of TBX-2 in *C. elegans*

Genetic evidence suggests that TBX-2 function in *C. elegans* is dependent on SUMOylation (Roy Chowdhuri et al., 2006). SUMO is a small ubiquitin-related modifier that is covalently attached to proteins to control their function (Meulmeester and Melchior, 2008). SUMO is expressed in all eukaryotes and in *C. elegans* there is a single gene which encodes the SUMO peptide SMO-1. The SUMOylation pathway is essential and reduction of SMO-1 results in embryonic and larval lethality. SUMO attachment occurs at a lysine residue on its target protein and requires an E1 activating enzyme (AOS-1/UBA-2), an E2 conjugating enzyme (UBC-9) and an E3 ligase (GEI-17) (Meulmeester and Melchior, 2008). Reduction of UBC-9 or GEI-17 by RNAi results in pharyngeal phenotypes that are similar to TBX-2 mutants. In addition, TBX-2 was found to interact with the UBC-9 and GEI-17 in yeast-two hybrid assays (Roy Chowdhuri et al., 2006). TBX-2 contains two SUMOylation sites and we hypothesize that SUMOylation of TBX-2 is important for its function.

Outside of the pharynx TBX-2 has a role in neural fate specification in the HSN and PHB

lineage (Singhvi et al., 2008). The HSNs are a pair of serotonergic motor neurons that innervate

the vulval muscles and stimulate egg laying (J. G. White, 1986; White et al., 1986; White and

Brenner, 1986; Desai et al., 1988). The HSNs are produced in the tail by two HSN/PHB

precursors that will divide and generate one HSN and one PHB phasmid chemosensory neuron.

The HSNs will eventually migrate to the center of the embryo. The *tbx-2* mutant characterized

was isolated in an EMS (ethyl methanesulfonate) screen for HSN development defects. This

mutation is predicted to cause an Ala238Val change in the protein and is believed to affect

binding ability. TBX-2 mutants had displaced HSN cells and were missing PHB cells. It was

shown that TBX-2 functions as a negative regulator of cell death for the PHB neuron and

functions in HSN migration (Singhvi et al., 2008).

TBX-2 functions in olfactory adaptation and dauer regulation and identifies additional

roles for TBX-2 outside the pharynx (Miyahara et al., 2004). Olfactory adaption allows the

worm to detect volatile chemical cues and adapt accordingly. *tbx-2* is expressed in pharyngeal

and non-pharyngeal neurons and it was found that expression in the AWC neuron is sufficient

for normal adaptation. Two missense mutations in *tbx-2* impair adaptation, but not chemotaxis,

to odorants sensed by AWC neurons (Miyahara et al., 2004). These mutations were isolated in an

EMS screen for dauer constitutive phenotypes and subsequently identified as *tbx-2*. Mutations in

the T-box domain from lysine (Lys164) to glutamic acid or arginine were identified.

4.2.6   Targets of T-box proteins

Despite their importance, few targets of T-box transcription factors have been

characterized (Wardle and Papaioannou, 2008). Some targets of T-box proteins have been

identified. *Xenopus Brachyury (Xbra)* is expressed throughout the mesoderm of embryos at an early stage in gastrulation (Smith et al., 1991). Misexpression of this gene results in ectopic formation of mesoderm (Cunliffe and Smith, 1992). *Xbra* directly activates expression of *embryonic FGF* (*eFGF*) by binding to a T-box site about 1 kb upstream of the *eFGF* start site (Casey et al., 1998). Gel shift assays showed that *Xbra* specifically binds this site as a monomer and is needed for reporter gene expression (Casey et al., 1998). *C. elegans*, TBX-35 directly activates *ceh-51*. The MS blastomere, produces mesodermal cell types, including pharyngeal cells and body muscles. A *tbx-35* null mutation results in a decrease in MS-derived tissues. Loss of *ceh-51* causes weak defects in the muscle and pharynx and results in larval lethality, while loss of *tbx-35* and *ceh-51* results in a penetrant loss of MS tissues and embryonic arrest (Broitman-Maduro et al., 2009). Four T-box sites were identified in the upstream sequence of *ceh-51* and based on EMSA all are thought to be important for TBX-35 binding (Broitman-Maduro et al., 2009). As discussed previously, mutations in *TBX5* cause heart and limb malformation in Holt-Oram syndrome. Cell culture and in-vitro assays have shown that Tbx5 and Nkx2-5, a homeobox protein essential in cardiac development, interact and directly regulate the cardiac-specific natriuretic peptide precursor type A (*Nppa*) gene (Hiroi et al., 2001). For those T-box proteins where no direct targets have been identified, a lot is known about expression patterns as well as loss of function phenotypes in different species such as mouse and *Xenopus*. However, there is still little information known about the regulatory mechanisms through which these transcription factors function.

T-box proteins contain a conserved DNA binding domain called the T-domain that is 180-200 amino acids long. This domain is used to divide members into 5 subfamilies based on sequence similarities (Papaioannou, 2001). *Brachyury* (*T*) gene was the first T-box identified in

mice and has the best characterized binding site which is used as a consensus sequence for the binding sites of other T-box proteins. A PCR-based binding-site selection procedure showed that the Brachyury protein binds to a 20 nucleotide partially palindromic sequence T[G/C]ACACCTAGGTGTGAAATT as a dimer (Kispert and Herrmann, 1993; Kispert et al., 1995). It has also been shown that the Brachyury protein binds and a half-palindromic (T/C)TTCACACCT as a monomer (Casey et al., 1998; Tada et al., 1998) Examination of downstream targets and binding-site selection experiments for a number of T-box proteins have shown that members of the T-box family that have been examined are capable of binding to the DNA sequence AGGTGTGA (Wilson and Conlon, 2002).

There are currently no known direct targets of *C. elegans* TBX-2 and therefore we do not yet know if this protein acts as an activator, repressor or both. We believe that TBX-2 is a transcriptional repressor based on data that shows TBX-2 function depends on SUMOylation, which is usually associated with transcriptional repressors (Shiio and Eisenman, 2003) and it was shown in yeast 2-hybridstudies that TBX-2 interacts with Groucho corepressor UNC-37 (Roy Chowdhuri et al., 2006). Previous studies have shown that the closely related TBX2 and TBX3 are the only mammalian T-box factors that are known to function as transcriptional repressors (Carreira et al., 1998; He et al., 1999), while TBX4 and TBX5 are transcriptional activators (Gibson-Brown et al., 1998; Takeuchi et al., 2003). Transcriptional assays using the GAL4-TBX5 fusion protein in a yeast-one hybrid system showed that TBX5 contains a transactivating domain in its C-terminus (Hiroi et al., 2001; Zaragoza et al., 2004). Through deletions the transactivating domain was narrowed to amino acids 349-351 (Zaragoza et al., 2004). In mice Tbx4 has recently been shown to have transcriptional activator and repressor activities (Ouimette et al., 2010). In Tbx4 both activator and repressor domains were identified in the C-terminus

using GAL4 DNA binding domain analysis (Ouimette et al., 2010). The activator domain in Tbx4 contains 49% sequence conservation with Tbx5 and the data suggests that the repressor domain identified is unique to Tbx4 and not found in Tbx5 (Ouimette et al., 2010). *ET* is a *Xenopus* T-box gene that represses transcription. Deletion constructs of the ET protein fused to the DNA-binding domain of GAL4 were expressed in 293T cells and used to identify the repressor domain in this protein (He et al., 1999). The repressor domain was identified in the C-terminus (residues 557-647) of ET and was found to be conserved in human TBX3 and TBX2 (He et al., 1999).

4.2.7 Identifying targets of *C. elegans* TBX-2

 Comparing mRNA expression levels in wild-type and hypomorphic *tbx-2(bx59)* mutant embryos using Affymetrix microarrays resulted in identification of a direct target of TBX-2. Of 19,885 probe sets examined, we found 980 mRNAs that were significantly up-regulated in *tbx-2(bx59)* relative to wild-type and 175 mRNAs that were significantly down-regulated. Using clustering analysis, comparisons to existing data sets on pharyngeal gene expression, and phylogenetic foot-printing to identify promoters with consensus T-box factor binding sites, we have analyzed a subset of genes and identified D2096.6 as a direct target of TBX-2. Results from our microarray and semi-quantitative PCR analysis indicate D2096.6 is repressed by TBX-2. GFP promoter fusions indicate the spatial and temporal expression pattern of D2096.6 is directly regulated by TBX-2 at a variant T-box binding site in the D2096.6 promoter.

Identification of additional targets will provide information on T-box protein transcriptional regulation in general and specifically aid in the understanding of how *C. elegans* TBX-2 regulates pharyngeal muscle development. TBX-2 is also expressed in neurons and body

wall muscles and data in our laboratory indicates TBX-2 may also be expressed in the intestine and seam cells (Milton and Okkema unpublished). Thus it is likely we will also identify targets of TBX-2 that function outside of the pharynx, which will provide information on the overall role of TBX-2 in *C. elegans* development.

4.3   Materials and Methods

4.3.1   <u>RNAi analyses</u>

In this study *tbx-2(RNAi)* analyses were performed essentially as previously described (Roy Chowdhuri et al., 2006). *tbx-2* dsRNA was produced by in-vitro transcription (Ambion Megascript T7 and T3) from PCR amplified pOK165.07 using the primer pair PO566/567. The PCR program used was as follows; (1) 94° 30 sec. (2) 94° 30 sec. (3) 53° 30 sec. (4) 72° 1 min. (5) got to step 2 4x, (6) 92° 30 sec., (7) 59° 30sec., (8) 72° 1 min., (9) go to step 6 29x, (10) 72° 5 min.

In addition *ubc-9(RNAi)* analysis was performed by feeding *D2096.6::gfp* (OK0666) worms bacteria expressing double stranded *ubc-9.* The *ubc-9* clone was obtained from the *C. elegans* RNAi feeding library and grown on LB-agar plates with 50 µg/ml ampicillin and 15 µg/µl tetracycline. Genomic fragments in the *C. elegans* RNAi library were cloned into L4440 and transformed into HT115 (DE3), an RNase III-deficient *E. coli* strain with IPTG-inducible T7 polymerase activity (Sijen et al., 2001). To seed the feeding plates, a single colony of *ubc-9* grown on the ampicillin and tetracycline plate was picked and inoculated into a 1 ml liquid culture (1 ml 2xTY + 50 µg/ml Ampicillin). The 1 ml liquid culture was then incubated with shaking at 37ºC for 6-8 hour and used to seed 6 inch feeding plates. Feeding plates contained NGM (2g NaCl, 3g Bacto Tryptone, 3g KH_2PO_4 , 0.5g K_2HPO_4 , 18g agar, 1.66 ml Cholesterol) with 25 µg/ml Carbenicillin and 1 mM IPTG. The feeding plates were allowed to dry overnight at room temperature to allow bacterial growth and induction of *ubc-9*.

After the plates are seeded and dry, ten L4 hermaphrodites were placed on one RNAi feeding plate and pre-fed for 24 hours at 20ºC. After 24 hours of growth 3 resulting adult hermaphrodites were placed on each of 3 individual RNAi feeding plates. These plates were

incubated for 24 hours at 20ºC and the adult hermaphrodites were then transferred to 3 new RNAi plates and returned to 20ºC. The embryos and L1 larvae remaining on each RNAi plate were mounted on slides and GFP was observed. The adult hermaphrodites were transferred for 2 additional 24 hr periods and embryos and larvae were observed after each transfer.

4.3.2 Isolation of N2 and *tbx-2(bx59)* mutant embryos

To isolate populations of N2 and *tbx-2(bx59)* mutant embryos, hermaphrodites from each strain were grown on ten 10 cm NGM agar plates containing a lawn of *E. coli* strain OP50. Each of ten plates for N2 and *tbx-2(bx59)* was started with 15 hermaphrodites. Worms were allowed to grow for 4 days at 25ºC for *tbx-2(bx59)* but only 3 days at 25ºC for wild type hermaphrodites. After the three or four day growth period, a rubber policeman was used to scrape worms and embryos from all ten plates into a single 15 ml conical tube containing 500 µl dH20. Once embryos and worms from all ten plates were collected, 2 µl samples were taken from each embryo preparation and mounted on agarose slides for microscopy. The percent of embryos at each stage was then compared between preparations to verify comparable distributions. The 15 ml conical tube containing the embryo samples was centrifuged at 1,700 rpm for 1 minute to pellet the worms and embryos (centrifuge 5804 R, rotor F45-30-11). The supernatant was then removed from the pellet, and the pellets were resuspended in 10 ml of hypochlorite solution (0.5M NaOH, 20% bleach in M9). The resuspended pellets were incubated at room temperature for 3 minutes with agitation to allow digestion of larvae. The samples were again centrifuged at 1,700 rpm for 1 minute. The supernatant was removed and the pellets were resuspended in fresh hypochlorite solution for an additional 3 minutes with agitation to allow digestion of remaining larvae. A dissecting scope was used to check that larva and adult worms were not present in the

sample. The remaining embryos in the sample were washed twice with 15 ml of PBS (137 mM

NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4, pH 7.4) to remove the hypochlorite

solution (Seydoux and Fire, 1995). The samples were then centrifuged again at 1,700 rpm for 1

minute to pellet the embryos and RNA was isolated as described in section 4.3.3.

4.3.3   Isolation of RNA

RNA was isolated from embryos using TRIzol® Reagent (Invitrogen) according to the

protocol provided by the manufacturer. TRIZOL was added in a 10:1 ratio (1 ml of TRIZOL

added to 100 µl of pelleted embryos) and samples were vortexed for 15 minutes at room

temperature. The samples were then centrifuged at 14,000 rpm for 10 minutes at 4ºC (centrifuge

5804 R, rotor F45-30-11). The supernatant from each sample was collected into fresh microfuge

tubes and the volume was recorded (~1,500 µl). Next, 0.2 ml of chloroform per 1 ml of TRIZOL

was added to each sample. The samples were shaken vigorously by hand for 15 seconds and

incubated at room temperature for 3 minutes. Samples were then centrifuged at 14,000 rpm at

4ºC for 15 minutes. The colorless aqueous phase present in each sample, which contains the

RNA, was removed and placed in a new centrifuge tube (~1000 µl). Next, RNA was

precipitated from the aqueous phase using 0.5 ml of isopropyl alcohol per 1ml of TRIZOL used

in initial homogenization step. Samples were incubated at room temperature for 10 minutes and

then centrifuged at 14,000 rpm for 10 minutes at 4ºC. After centrifugation a visible pellet

appeared in each tube. Using a pipette, the supernatant was removed from each tube containing

a pellet and the pellet was then washed once with 75% ethanol (in DEPC treated water). To wash

the pellet 1 ml of 75% ethanol was added for every 1 ml of TRIZOL used in homogenization

step. The pellet was resuspended by vortexing and then centrifuged at 7,500 rpm for 5 minutes at

4ºC. The resulting pellet was air dried for 20 minutes and resuspended in 30 µl of DEPC treated

water by incubating at 60ºC for 10 minutes. The RNA samples were then purified with a Qiagen

RNeasy mini kit and concentrations were quantified using a nanodrop spectrophotometer (Wang

Laboratory UIC). The concentrations of the RNA samples used in the microarray analysis are

shown in Table V.

TABLE V: RNA CONCENTRATIONS FOR INDIVIDUAL N2 AND *tbx-2(bx59)* EMBRYO
PREPARATIONS

Sample	OD 260/280	µg/ µl
N2 #1	2.08	2.439
N2 #2	1.88	3.601
tbx-2(bx59) #2	2.13	1.443
tbx-2(bx59) #3	2.09	2.925
tbx-2(bx59) #4	1.96	3.228

N2 and *tbx-2(bx59)* embryo populations grown at 25°C were collected and RNA was
isolated as described (sections 6.3.1 and 6.3.2). RNA concentrations for each sample
used in the microarray experiment were measured using a nanodrop spectrophotometer.

4.3.4   <u>Microarray and Data analysis</u>

Labeling and hybridization to the Affymetrix '*C. Elegans* Genome' GeneChips was performed by the University of Illinois at Chicago Core Genomic Facility (CGF). A pair-wise comparison was performed between two wild type (A, B) and three mutant samples (A, B, C).

Samples were labeled according to standard Affymetrix recommended protocols. Affymetrix GeneChip arrays were hybridized and scanned according to standard Affymetrix protocols. Each array was analyzed for the following quality metrics: total background, raw noise (Q), average signal present, signal intensity of species-specific house-keeping genes, 3'/5' signal ratio of house-keeping genes, relative signal intensities of labeling controls, absolute signal intensities of hybridization controls, and GCOS scale factors.

Microarray data was analyzed by the CGF and differential expression values were calculated according to the following criteria. Data was analyzed in 'S-Plus' 6.2 statistical package with the 'S+ArrayAnalyzer' v2.0.1 from Insightful. Data was normalized by quantiles and summarized using the Robust Multi-array Average (RMA) (Irizarry et al., 2003). Analysis was performed including and excluding samples 7-9. Local Pooled Error (LPE) test (Jain et al., 2003). Statistical test p-values were corrected for False Discovery Rate (FDR) by Benjamini-Hochberg (BH) procedure. At a significance of BH-adjusted pValue <0.05, 1186 probe sets were identified as statistically significant, differentially expressed. Of the significant differentials, 195 were down regulated in mutant relative to wild type and 991 were up regulated. Differentially expressed transcripts from comparisons were annotated according to data available in Affymetrix's 'NetAffx Analysis Center' (Release R25, March 1, 2008).

4.3.5   Semi-quantitative RT-PCR

Semi-quantitative reverse transcription polymerase chain reaction (semi-qPCR) was used

to compare the expression levels of genes between N2 and *tbx-2(bx59)*. RNA samples collected

as described in section 4.3.3 were treated with Ambion® TURBO DNA-*free*™ DNase

Treatment and Removal Reagents to remove DNA from the RNA samples according to the

manufacturer's instructions. Briefly, 0.1 volume of 10X Turbo DNase buffer and 1µl of Turbo

DNase was added to each RNA sample and gently mixed. The samples were then incubated at

37ºC for 30 minutes. Next, 0.1 volume of DNase Inactivation Reagent was added to each sample

and incubated for two minutes at room temperature with occasional mixing. Then samples were

then centrifuged at 14,000 rpm for 1.5 minutes and the supernatant was transferred to fresh

eppendorf tubes (centrifuge 5804 R, rotor F45-30-11).

For the reverse transcription of the RNA samples the following was added to individual

eppendorf tubes, 2 µg of total RNA, 1µl of 200 ng/µl oligo(dT), 0.5 µl RNase Inhibitor

(RNaseOUT Invitrogen) and DEPC-H_2O to a final volume of 12 µl. The samples were

incubated at 65ºC for 10 minutes. Next, the samples were placed on ice and 4µl of 5X First

Strand buffer (Invitrogen), 2µl 0.1M DTT and 1 µl 40mM dNTP mix was added to each tube and

placed at 37ºC for 2 minutes. After the incubation, 1 µl SS III (Invitrogen) was added to each

tube and placed at 50ºC for 1 hour. After 1 hour, 1µl RNase H (Fermentas) was added to each

tube and placed at 37ºC for 20 minutes. Finally, the samples were incubated at 95ºC for 10

minutes to heat inactivate enzymes and a 1:10 dilution of each sample was used for the PCR

reactions.

PCR primers were designed for each gene and when possible primers spanned two exons

to avoid amplification of genomic DNA that may be present (Table VI). The PCR conditions

used for all primers were as follows: (1) 94ºC 2 minutes, (2) 94ºC 30 seconds, (3) 56ºC 30 seconds, (4) 72ºC for 30 seconds , (5) go to step 2 39 times, (6) 72ºC 5 minutes. After the primer pairs were designed for each gene they were tested to ensure amplification of the correct product size was produced. Varying amounts of cDNA template was used initially to determine the proper amount needed to produce comparable *ama-1* transcript from wild type and *tbx-2(bx59)* samples. To obtain PCR products in the linear range five PCR reactions were set up for each gene and a sample was pulled out at cycle 30, 33, 36, 39 and 42. Samples for each PCR cycle for all 13 genes were then run on 2% ethidium bromide gels. ImageJ software was used to quantify the expression levels on the ethidium bromide gels and then normalized to *ama-1* using the equation Relative expression = [*tbx-2*(exp)/*tbx-2*(*ama-1*)]/[N2 (exp)/N2 (*ama-1*)].

TABLE VI: GENES ANALYZED BY SEMI-QUANTITIVE PCR

Gene	Primer Pair	Product size
T10E10.4	PO878/879	421 bp
D2096.6	PO880/881	309 bp
myo-5	PO882/883	436 bp
flp-1c	PO884/885	175 bp
flp-2b	PO886/887	151 bp
flp-9	PO888/889	230 bp
flp-11c	PO890/891	247 bp
flp-15	PO892/893	175 bp
flp-16	PO894/895	179 bp
flp-19	PO896/897	202 bp
T25E4.1	PO898/899	411 bp
phg-1	PO900/901	432 bp
pqn-71	PO902/903	514 bp
ama-1	PO906/907	354 bp

14 genes were used in a semi-qPCR experiment to validate the microarray results. The gene names as well as the primer pairs and product sizes amplified are listed.

4.3.6   Site-directed mutagenesis

Site directed mutagenesis was performed with Stratagene's QuikChange II site-directed

Mutagenesis kit according to the manufacturer's instructions. Briefly, primers were designed

using the following criteria: primers for mutagenesis were 25-45 bases in length with a melting

temperature $\geq$78ºC and the deletion positioned in the middle of the primer sequence. In a 50 µl

reaction 5 µl reaction buffer was added (provided in kit), 10 ng plasmid, 125 ng of each primer designed

for the desired mutation and 1 µl of dNTP mix (provided in kit). 1 µl pfu Ultra High-Fidelity DNA

polymerase was added (provided in kit). The following PCR program was used with the

appropriate time in step 4 based on the size of the plasmid used: step (1) 95ºC 30 seconds, (2)

95ºC 30 seconds, (3) 55 ºC 1 minute, (4) 68ºC 1 minute/kb of plasmid length, (5) go to step 2 17

times. The amplified products were digested with 1 µl Dpn1 at 37 ºC for 1 hour. 1 µl of the

digested DNA was transformed with 50 µl XL1-Blue supercompetent cells (provided in kit). 0.5

ml of NZY$^+$ broth was added to each tube and incubated 1 hour at 37 ºC with shaking. The cells

were then spread on 2xTY + ampicillin plates and incubated at 37 ºC overnight. Plasmid

pOK253.05 was used as a template for the D2096.6 promoter mutations. Primer pair

PO952/PO953 were used to mutate site 1 and PO1002/PO1003 were used to mutate site 2.

4.3.7   GST::TBX-2 expression

The DNA binding domain of TBX-2 was PCR amplified from pOK246.03 with primer

pair PO1055/PO1056 (made by Tom Ronan). This 578 bp PCR product contains the DNA

binding domain with a BamH1 and Xho1 linker and was cloned into a pGEX-4-T2 vector

containing GST. To transform GST::TBX-2 into Rosetta competent cells (Novagen), 1 µl of

DNA from pOK269.07 (GST::TBX-2 containing the TBX-2 binding domain) was added to 100

µl Rosetta competent cells in 15 ml conical tubes and placed on ice for 30 minutes. The cells

were then heat shocked for 60 seconds in a 42ºC water bath. After heat shock the cells were placed directly on ice long enough to add 100 µl 2xTY. The cells were then incubated at 37ºC for 30 minutes with shaking. At this point the cells were then spread on 2xTY + ampicillin + chloramphenicol plates using a ETOH sterilized glass spreader and allowed to dry. The plates were incubated overnight at 37ºC.

A single colony from the transformation of pOK269.07 into Rosetta competent cells was picked and used to inoculate a 3 ml culture of (3 ml 2xTY + 0.1 mg/ml ampicillin + 0.17 mg/ml chloramphenicol) in a 125 ml flask. The culture was incubated overnight at 20ºC with shaking. The following day 3 ml of the overnight culture was transferred to a fresh 50 ml culture of M9 minimal medium (1L M9 salts: 15g KH_2PO_4, 64g Na_2HPO_4 $7H_2O$, 2.5g NaCl, 5g NH_4 Cl pH adjusted to 7.2 with NaOH, autoclave)[1L minimal growth medium: 200ml 5X M9 salts, 20ml D-glucose (20g/100ml), 10ml Basal vitamin eagle (Fisher), 1M $MgSO_4$ (autoclaved), 1M $CaCl_2$ (autoclaved) (Marley et al., 2001)] with 0.1 mg/ml ampicillin and 0.17 mg/ml chloramphenicol. The culture was again incubated overnight at 20ºC with shaking. The following day 20 ml of the overnight culture was transferred to 480 ml M9 minimal medium with 0.1 mg/ml ampicillin and 0.17 mg/ml chloramphenicol. The culture was incubated at 20ºC for 15-17 hours with shaking or until the OD reached 0.6-0.8. Once the appropriate OD was reached, 0.5 mM IPTG was added and the culture was induced for 24 hours at 20ºC with shaking. After 24hrs the induced culture was centrifuged at 5,000 rpm for 30 minutes (centrifuge 5804 R, rotor F-34-6-38) and the resulting pellet was stored at -20ºC until purified.

4.3.8   GST::TBX-2   purification

GST::TBX-2 protein was expressed as described in section 4.3.7 and purified on a glutathione affinity column. To prepare the column 3 ml of 50% glutathione agarose (Pierce)

was measured and added to the affinity column. The column was then centrifuged at 3000 rpm

for 30 seconds (centrifuge 5804 R, rotor F-34-6-38). The supernatant was removed and the

agarose was washed with 4 ml PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM

KH_2PO_4, pH 7.4) for 3 minutes and centrifuged at 3000 rpm for 30 seconds. The PBS wash was

then repeated one time. Next, the agarose was washed with 4 ml glutathione buffer [50mM Tris

base, 10mM reduced glutathione (Sigma), pH adjusted to 8 with HCl. The solution was prepared

fresh daily and stored at 4ºC] for 3 minutes and centrifuged at 3000 rpm for 30 seconds. The

agarose was then washed with 7 ml PBS/EDTA (1X PBS, 5mM EDTA, pH 7.4 with 1M

NAOH, store up to 1 month at 4ºC) for 3 minutes and centrifuged at 3000 rpm for 30 seconds.

The PBS/EDTA wash was repeated one time. After the column is prepped, the pellet from the

500 ml induced culture was resuspended in 6 ml ice cold lysis buffer (50mM NaCl, 50mM Tris

base, 5mM EDTA, pH 8.0 with HCL and prepare fresh daily) and 50 µl of Protease Inhibitor

Cocktail (Sigma P8465) was added. Next, DTT was added to a final concentration of 10mM and

lysozyme was added to a final concentration of 0.2 mg/ml. The sample was then incubated on

ice for 30 minutes. To break apart the cells the sample was sonicated 10 times for 10 seconds on

ice with a 1 minute break in between each sonication. A 100 µl sample was saved as total

protein. After sonication is complete, 1% triton X-100 from a 20% stock was added and the

sample was rocked for 20 minutes at 4ºC. Next, the sample was centrifuged at 5,000 rpm for 1

hour at 4ºC to pellet (centrifuge 5804 R, rotor F-34-6-38). The supernatant (soluble fraction)

was transferred to the prepared glutathione column and rocked for 4 hours at 4ºC to bind the

protein to the column. A 100 µl sample of the supernatant was taken as total soluble protein.

The left over pellet was resuspended in 6 ml of lysis buffer and a 100 µl sample was taken as

total insoluble protein. After rocking for 4 hours the column containing the supernatant was

centrifuged at 5000 rpm for 1 minute at 4ºC. The column was then washed with 3 ml

PBS/EDTA cold buffer for 3 minutes and then centrifuged at 5000 rpm for 1 minute at 4ºC. The

supernatant was removed and the wash was repeated 2 times. After the final wash the agarose

resin was resuspended in 2 ml PBS/EDTA. The column was allowed to settle and the buffer was

allowed to flow out of column until it reached the top of the resin. The bound protein was eluted

with 10 ml glutathione buffer at 4ºC and 1 ml fractions were collected in eppendorf tubes. To

identify which fraction contained the eluted protein samples were run [10 µl protein + 10 µl 2X-

SDS (for 25 ml solution: 6.25 ml 4x Tris-Cl/SDS pH6.8, 20% glycerol, 1g SDS, 0.25mg

bromophenol blue, 5% b-mercaptoethanol)] on a 10-12% SDS-PAGE gel. Protocol and recipes

modified from Current Protocols in Protein Science Chapter 6.

4.3.9   Radiolabeling probes for gel mobility shift assays

Probes were designed for two possible TBX-2 binding sites in the D2096.6 promoter, the

Brachyury consensus binding site 1 which is conserved amongst other nematode species (bp

282-292 in pOK253.05, mutated site 1 pOK256.01) and the new position weighted matrix

designated site 2 (bp 608-618 in pOK253.05, mutated newpwm pOK258.03). Probes were

designed for the wild type sequence as well as a mutated sequence at site 1 and site 2 in D2096.6

(Table VII). DNA oligonucleotides were ordered through IDT (Integrated DNA Technologies)

and were purified under standard desalting conditions. Single stranded oligonucleotide pairs for

each site were annealed to produce double stranded probes. To anneal probes 20µl reactions

were prepared with 4 µl of each oligonucleotide pair (1 µg/µl), 2 µl of 10X TE and 10 µl of

dH20 for a final concentration of 0.4 µg/µl. Probes were heated at 100ºC for 3 minutes and

slowly cooled to room temperature by removing the heating block. Annealed probes were then

stored at -20ºC. Double stranded oligonucleotides were labeled via end-filling reactions with the

Klenow fragment (3'-5' exo- NEB) of DNA Polymerase I in the presence of [32P]-dCTP (800

Ci/mmol). Labeling reactions were mixed on ice and contained: 0.5 µl ds oligo (0.4µg/µl stock),

2 µl 10X REact 2 buffer, 10 µl (0.5mM dG, dA, dT mix), 5 µl 32p dCTP (800 ci/mmol), 2.5 µl

dH$_2$0 for a final volume of 20 µl. 0.5 µl Klenow Fragment was added to each reaction and

incubated for 1 hour in a 37ºC water bath. After incubation reactions were stopped by adding 40

µl 2X TE. To determine the radioactive counts incorporated, 1 µl of each labeled probe was

spotted on 3 separate DE81 Whatman filters and allowed to dry. 2 out of the 3 filters was

washed for each probe with 75 ml 0.5M Na2HPO4 for 5 minutes 3 times each. Next the filters

were rinsed briefly with 100% EtOH and allowed to dry. Each dry filter was placed in a separate

scintillation vials and radioactive counts were analyzed with a scintillation counter. The

remainder of the labeled probes was purified with a QIAquick Nucleotide Removal Kit (Qiagen)

and eluted in 100 µl.

TABLE VII: D2096.6 OLIGONUCLEOTIDE PROBES USED IN EMSA

Oligonucleotide	Site	Sequence 5'-3'[a]
PO1157/PO1158	site 1	CAAGCGCTCACTTT<u>TTCTCACCT</u>TAATGTGTACCC
PO1209/PO1210	mutated site 1	CAAGCGCTCACTTT<u>TTCTACAAT</u>TAATGTGTACCC
PO1159/PO1160	site 2	CAAGATTTTGATATTT<u>GGTTTCAA</u>CATTTGATAT
PO1211/PO1212	mutated site 2	CAAGATTTTGATATTT<u>AAATTCAA</u>CATTTGATAT

[a] Wild type and mutated T-box binding sites are underlined

4.3.10 Electrophoretic mobility shift assay (EMSA)

The following EMSA assay was modified from (Thatcher et al., 1999). For probe 1 50 ng GST::TBX-2 protein was incubated in 20 μl binding reactions containing buffer 2 [10 mM Tris pH 7.8, 50 mM NaCl, 5 mM EDTA pH8, 10% glycerol modified from (Carlson et al., 2001)]; and 40 ng poly dI/dC. For competition assays 200 to 800 times unlabeled probe was also added at this time. The 20 μl binding reactions are then incubated at room temperature for 15 minutes. 500,000 cpm of probe 1 was then added and the binding reaction was incubated at room temperature for an additional 10 minutes. For probe 2 250ng GST::TBX-2 protein was incubated in 20 μl binding reactions containing buffer 2 [10 mM Tris pH 7.8, 50 mM NaCl, 5 mM EDTA pH8, 15% glycerol (modified from Carlson et al., 2001)]; and 40 ng poly dI/dC. For competition assays 2 to 18 times unlabeled probe was also added at this time. The 20 μl binding reactions were then incubated at room temperature for 15 minutes. One million cpm of probe 2 was added and the binding reaction was incubated at room temperature for 1 hour. Binding reactions were then electrophoresed on a 5% non-denaturing acrylamide gel at 4ºC. The binding reactions were run about half way down the gel and visualized by the addition of blue juice in the first lane. Gels were removed and dried on a gel dryer and then imaged on the phosphorimager and exposed to film.

4.4   Results

4.4.1   <u>*tbx-2(bx59)* mutants exhibit defects in the anterior pharynx</u>

tbx-2(bx59) is a temperature sensitive allele containing a missense mutation affecting the

T-box DNA binding domain (King Chow pers comm). This strain can be maintained as

homozygotes and at permissive temperatures (16ºC) these worms have a wild type pharynx. At

the non-permissive temperature (25ºC), *tbx-2(bx59)* mutants exhibit partially penetrant L1 arrest

and anterior pharyngeal defects (Figure 15D). These defects are similar, but less severe, to

defects seen in *tbx-2(ok529)* null mutants and *tbx-2(RNAi)* animals which were found to lack

ABa derived anterior pharyngeal muscles (Figure 15B,C) (Roy Chowdhuri et al., 2006).
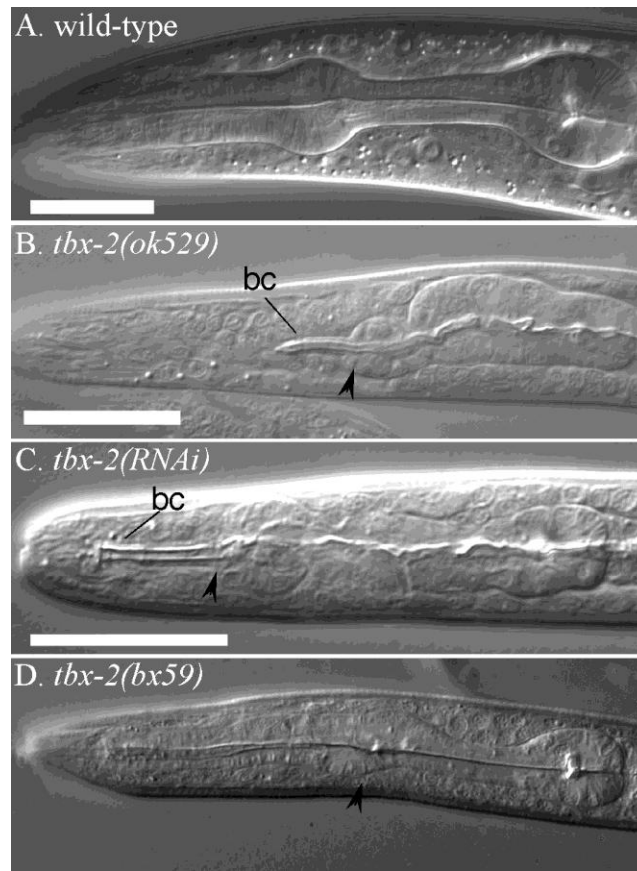
Figure 15: *tbx-2(bx59)* **mutants exhibit defects in the anterior pharynx**

(A) DIC image of a wild type pharynx with the anterior and posterior pharynx extending the length of the head. (B) DIC image of a *tbx-2(ok529)* mutant with defects in the anterior pharynx. Arrowhead indicates the position of the anterior pharynx. An elongated and detached buccal cavity is indicated by (bc). (C) DIC image of a *tbx-2(RNAi)* animal with defects in the anterior pharynx. Arrowhead indicates the position of the anterior pharynx. An elongated and detached buccal cavity is indicated by (bc) (D) DIC image of a *tbx-2(bx59)* mutant (OK0660) at 25ºC. Arrowhead indicates misshapen anterior bulb.

4.4.2   <u>Identifying differentially expressed genes in *tbx-2(bx59)*</u>

C. elegans TBX-2 is required for the development of ABa-derived pharyngeal muscles and is the sole *C. elegans* member of the conserved Tbx2 sub-family of T-box factors, which in other organisms includes both transcriptional activators and repressors. Currently there are no known targets of TBX-2. To help identify TBX-2 targets we performed a microarray to compare RNA expression in wild type and *tbx-2(bx59)* mutant embryos. The *tbx-2(bx59)* temperature sensitive allele was used in the microarray because this strain is homozygous viable and when shifted to the nonpermissive temperature 100% of these animals have reduced levels of TBX-2. The *tbx-2(ok529)* null allele is lethal and maintained as a heterozygous strain segregating only 25% TBX-2 homozygous mutants and was therefore not used for the microarray. We are specifically interested in identifying TBX-2 targets involved in pharyngeal muscle development.

To identify genes that are differentially regulated by TBX-2 in a microarray analysis, we isolated total RNA from wild-type and *tbx-2(bx59)* mixed embryo populations grown at the non-permissive temperature (25˚C). We compared mRNA expression levels in wild-type and *tbx-2(bx59)* embryos grown at the non-permissive temperature using the Affymetrix GeneChip *C. elegans* Genome Array.

The GeneChip *C. elegans* Genome Array was designed using the December 2000 genome sequence, predicted transcripts and EST sequences from the Sanger Center. In addition mRNA sequences from GenBank (release 121) were used. For the GeneChip *C. elegans* Genome Array, probe sets were selected against the 3' ends to generate greater than 22,500 probe sets against 22,150 unique transcripts. The oligo length was 25-mer and the individual gene sequences for *C. elegans* was represented by approximately 11 different probe sequences arranged in probe pair sets. Control sequences for hybridization controls included; bioB, bioC

and bioD from E.coli, and cre from P1 bacteriophage. Poly-A controls were; *dap*, *lys*, *phe*, *thr* and *trp* from *B. subtilis.* Finally controls consisted of the *C. elegans* maintenance genes; actin, catalase,GAPDH, gly14 and ubiquitin (GeneChip *C. elegans* Genome Array (#900383) www.affymetrix.com)

The Affymetrix GeneChip *C. elegans* Genome Array monitors expression of 19,885 genes, the RNA labeling, hybridization and microarray analysis was performed by the University of Illinois RRC. We compared RNA levels in two wild type and three *tbx-2(bx59)* mutant embryo preparations. The microarray identified 1155 / 19,885 genes that were statistically significant and differentially expressed between wild type and *tbx-2(bx59).* Of those genes that were found to be differentially expressed, 980 genes are over expressed in *tbx-2(bx59)* relative to wild type (26.6x to 1.1x) and 175 genes are under expressed in *tbx-2(bx59)* relative to wild type (-13.4x to -1.1x).

4.4.3   Validating microarray results using semi-quantitative RT-PCR

To test the validity of our microarray results, we used semi-quantitative PCR to examine the changes in gene expression of 13 genes in wild type and *tbx-2(bx59)* embryos (Table VIII). Each transcript was normalized to *ama-1*. *ama-1* encodes the large subunit of RNA polymerase II and its expression levels remain relatively constant throughout development, and it has been previously used as an internal control for comparing expression levels of a target gene during different stages of *C. elegans* development (Bird and Riddle, 1989; Larminie and Johnstone, 1996; Nelms and Hanna-Rose, 2006). We performed semi-quantitative PCR (semi-qPCR) as previously described in material and methods (section 6.3.5) on 13 selected genes. Twelve of these genes were over expressed in the *tbx-2(bx59)* microarray (*T10E10.4, D2096.6, flp-1, flp-2, flp-9, flp-11, flp-15, flp-16, flp-19, T25E4.1, phg-1, and pqn-71*) while one of these genes (*myo-*

5) was under expressed. ImageJ software was used to quantify expression levels on ethidium

bromide stained gels and we then normalized to *ama-1* using the equation Relative expression =

[*tbx-2*(exp)/*tbx-2*(*ama-1*)]/[N2(exp)/N2(*ama-1*)]. We found 11 of these 13 genes (84%)

exhibited similar changes in expression in *tbx-2(bx59)* mutants using semi-qPCR to what we

observed using the microarray. Two genes (*flp-2* and *flp-9*) that exhibited ectopic expression in

the *tbx-2(bx59)* microarray showed a moderate decrease in mRNA expression in our semi-qPCR

assay (Table VIII).

TABLE VIII: SEMI-qPCR AND MICROARRAY DATA SHOW SIMILAR TRENDS IN
RNA EXPRESSION

Gene	Fold change semi-qPCR *tbx-2(bx59)*/N2	Fold change microarray *tbx-2(bx59)*/N2
T10E10.4	1.29	2.70
D2096.6	1.38	1.39
flp-1	1.15	2.17
flp-2	0.66	2.02
flp-9	0.52	1.85
flp-11	1.58	2.29
flp-15	3.17	1.66
flp-16	1.76	2.89
flp-19	1.70	1.93
T25E4.1	1.77	1.86
phg-1	1.74	1.57
pqn-71	4.42	2.42
myo-5	0.38	-1.37

Semi-quantitative PCR to examine the changes in gene expression of 13 genes in wild type and *tbx-2(bx59)* embryos and each transcript was normalized to *ama-1*. ImageJ software was used to quantify expression levels on ethidium bromide stained gels and we then normalized to *ama-1* using the equation Relative expression = [*tbx-2*(exp)/*tbx-2*(*ama-1*)]/[N2(exp)/N2(*ama-1*)].

4.4.4   Filtering of differentially expressed genes

To identify pharyngeal genes that are regulated by TBX-2 we compared our microarray data to existing data sets on pharyngeal gene expression. A previous microarray was used to examine expression of 62% of genes (11,917/19,099) and identified 240 genes that are likely expressed in the pharynx (Gaudet and Mango, 2002). We compared our data set to those genes identified in that microarray that are likely expressed in the pharynx and found 48 (~20%) of those pharyngeal genes are up-regulated in *tbx-2(bx59)* and 3 (~1%) of those pharyngeal genes are down-regulated in *tbx-2(bx59)*. We also looked at published literature and Yugi Kohara's Nematode Expression database and identified 24 genes that have been shown to be expressed in the pharynx (Table IX).

Next, we identified sites matching the Brachyury T-box consensus binding site (Figure 21) in 5' intergenic regions upstream of microarray positive genes using JASPAR Core database and ConSite database. Utilizing the JASPAR Core and ConSite databases, we looked in the 5'-intergenic region from the genes that overlapped with the data from the pharyngeal specific microarray (Gaudet and Mango, 2002) and found 34/51 genes contain Brachyury consensus binding sites. We then performed phylogenetic footprinting to ask if these T-box binding sites are conserved in *C. elegans, C. remanei,* and *C. briggsae.* We found 3 genes that contained at least one conserved Brachyury consensus binding site in all three species (Table IX).

TABLE IX: SELECTED GENES IDENTIFIED IN THE *TBX-2(BX59)* AND PHARYNGEAL SPECIFIC MICROARRAY AS DIFFERENTIALLY EXPRESSED[a]

Gene	Gene title[b]	Fold (*tbx-2*/N2)[c]	Fold (*par-1skin-1*)[d]	# T-box sites in 5'-intergenic[e]	# sites conserved Ce vs C[f]	# sites conserved Ce vs	exp data in pharynx[h]
abu-10	Activated in Blocked Unfolded protein response	2.47	2.69	2	0	0	
abu-6	Activated in Blocked Unfolded protein response	1.54	2.35	0	0	0	
abu-7 // abu-8	Activated in Blocked Unfolded protein response	1.59	1.92	0	0	0	
abu-8	Activated in Blocked Unfolded protein response	1.83	2.26	3	0	0	
B0238.12	protein coding	3.35	1.54	1	0	0	NEXTDB
C14C6.5	protein coding	3.72	1.03	0	0	0	
C18A11.1	protein coding	2.57	1.32	0	0	0	
C18H9.6	protein coding	1.89	1.14	0	0	0	
C24A3.2	protein coding	1.25	1.12	0	0	0	
C53C9.2	protein coding	2.29	1.51	1	0	0	NEXTDB
cdr-4	CaDmium Responsive	2.23	1.9	0	0	0	(Dong et al., 2008)
cpn-4	CalPoNin	2.51	2.56	0	0	0	NEXTDB
cyp-33C8	CYtochrome P450 family	2.22	1.25	0	0	0	
D2096.6	protein coding	1.39	2.82	4	1	1	(Gaudet and Mango, 2002)

TABLE IX (continued)

Gene	Gene title[b]	Fold (*tbx-2*/N2)[c]	Fold (*par-1skin-1*)[d]	# T-box sites in 5'-intergenic[e]	# sites conserved Ce vs	# sites conserved Ce vs	exp data in pharynx[h]
F07C4.11	protein coding	1.71	1.48	0	0	0	
F08B12.4	protein coding	2.4	1.29	1	0	0	NEXTDB
F08G2.5	protein coding	3.38	1.43	1	0	0	
F09F3.6 // ttr-21	Transthyretin-like family	9.52	1.58	0	0	0	
F13H10.1	protein coding	1.57	1.69	0	0	0	Wormbase
F13H8.4	protein coding	1.7	1.7	2	0	0	
F20B10.3	protein coding	6.11	4.68	1	0	0	
F35E12.5	protein coding	9.71	1.18	2	0	0	
fat-5	FATty acid desaturase	2.18	1.08	1	0	0	(Brock et al., 2006)
fbxa-3	F-box A protein	2.53	1.13	2	0	0	
flp-1	FMRF-Like Peptide	2.17	-1.84	0	0	0	(Li et al., 1999)
flp-11	FMRF-Like Peptide	2.29	-1.88	2	0	0	
flp-12	FMRF-Like Peptide	2.65	1.77	2	0	0	
flp-16	FMRF-Like Peptide	2.89	0.16	4	0	0	
flp-19	FMRF-Like Peptide	1.93	-0.81	2	0	0	
flp-3	FMRF-Like Peptide	3.32	0.29	2	0	0	
lec-8	gaLECtin // galaptin domain	1.67	1.91	1	0	0	

TABLE IX (continued)

Gene	Gene title[b]	Fold (tbx-2/N2)[c]	Fold (par-1skin-1)[d]	# T-box sites in 5'-intergenic [e]	# sites conserved Ce vs Cr[f]	# sites conserved Ce vs Cb[g]	exp data in pharynx[h]
myo-5	protein coding // myosin	1.37	2.76	1	0	0	NEXTDB,Wormbase
nspb-12	"Nematode Specific Peptide family	3.89	1.9	0	0	0	
phat-2	PHAryngeal gland Toxin-related	1.41	3.11	0	0	0	(Smit et al., 2008)
phat-3	PHAryngeal gland Toxin-related	2.39	1.48	0	0	0	(Smit et al., 2008)
phat-5	PHAryngeal gland Toxin-related	1.28	3.79	1	0	0	(Smit et al., 2008)
pqn-5	Prion-like-(Q/N-rich)-domain-bearing protein	1.75	2.17	0	0	0	
pqn-71	Prion-like-(Q/N-rich)-domain-bearing protein	2.42	2.11	1	0	0	NEXTDB
rnh-1.3	RNase H	8.64	1.47	0	0	0	
T10E10.4	protein coding	2.7	1.83	1	1	1	NEXTDB
T25E4.1	protein coding	1.86	3.41	4	2	1	NEXTDB
T28B4.4	protein coding	1.81	1.54	2	0	0	
tbx-2	T BoX family	1.69	1.23	1	0	0	(Roy Chowdhuri et al., 2006; Smith and Mango, 2007)
tnc-2	TropoNin C	1.5	2.37	0	0	0	NEXTDB,Wormbase
ttr-26	Transthyretin-like family	17.8	1.07	1	0	0	

TABLE IX (continued)

Gene	Gene title[b]	Fold (*tbx-2*/N2)[c]	Fold (*par-1*/*skin-1*)[d]	# T-box sites in 5'-intergenic [e]	# sites conserved Ce vs Cr[f]	# sites conserved Ce vs Cb[g]	expression data in pharynx[h]
ttr-54	Transthyretin-like family	1.62	1	0	0	0	
W03F8.6	protein coding	2.11	1.57	1	0	0	Wormbase
Y73F4A.2	protein coding	2.23	1.37	0	0	0	NEXTDB,Wormbase
ZC116.3	protein coding // bone morphogenetic protein 1 like	-1.9	1.32	2	0	0	
ZK1025.7	protein coding	2.13	1.01	1	0	0	
ZK662.2	protein coding	1.65	2.7	0	0	0	NEXTDB

[a] Table contains the list of differentially regulated genes that are common between the *tbx-2(bx59)* microarray described in this chapter and the pharyngeal specific microarray described in (Gaudet and Mango, 2002).

[b] Gene titles are assigned by Wormbase (http://www.wormbase.org/) and UniProt (http://www.uniprot.org/).

[c] The signal ratio for samples *tbx-2(bx59)* to samples N2 reported in the *tbx-2(bx59)* microarray.

[d] The signal ratio for samples *par-1* to samples *skin-1* reported in (Gaudet and Mango, 2002)*. par-1* mutants contain excess pharyngeal cells and *skn-1* mutants contain no pharyngeal cells

[e] The number of T-box sites identified within the 5'-intergenic region of *C. elegans* identified using the JASPAR (http://jaspar.genereg.net/) and ConSite (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/) databases*.*

[f] The number of T-box sites identified within the 5'intergenic region of *C.elegans* that are conserved within the 5'intergenic region of *C. remanei.*

[g] The number of T-box sites identified within the 5'intergenic region of *C.elegans* that are conserved within the 5'intergenic region of *C. briggsae.*

[h] Cited literature, NEXTDB The Nematode Expression Pattern DataBase (http://nematode.lab.nig.ac.jp/) and Wormbase (http://www.wormbase.org/) which contain information for genes that are expressed in the pharynx.

4.4.5   TBX-2 represses D2096.6 expression in embryos and larvae

We hypothesize that TBX-2 directly binds the promoters of its targets to regulate their expression and therefore, we utilized GFP promoter fusions to identify genes that are regulated by TBX-2. Based on known expression patterns, clustering analysis and phylogenetic foot-printing, we chose eight genes for further analysis and obtained or constructed GFP promoter fusions for 8 genes (*D2096.6, T25E4.1, pqn-71, myo-5, flp-1, flp-2, flp-11, flp-15*) on which to begin our analysis and test if these candidates are targets of TBX-2. To test whether expression of these reporters is affected by TBX-2, we compared their expression in wild type and *tbx-2(RNAi)* animals. We found no alteration of expression for six of these seven reporters, and these genes are discussed in Appendices B, C and D. We found that expression of the *D2096.6::gfp* reporter was brighter in *tbx-2(RNAi)* animals than in wild type, and we focused on asking if D2096.6 is regulated by TBX-2.

D2096.6 produces an 879 base pair transcript that encodes a 171 amino acid protein. This is a small gene located in the intron of another uncharacterized gene transcribed from the opposite strand (Figure 16). Nothing is known about the function of D2096.6; however some information has been reported about the regulation of this gene. A *D2096.6::gfp* transcriptional fusion containing 673 base pair 5'flanking DNA was expressed in pharyngeal cells and this expression was regulated by PHA-4 (Gaudet and Mango, 2002). PHA-4 is a Forkhead box A transcription factor that is essential for development of the pharynx and is believed to regulate most pharyngeal genes. Two PHA-4 binding sites were identified in the D2096.6 promoter and when these sites were deleted or mutated D2096.6 expression was lost (Gaudet and Mango, 2002).
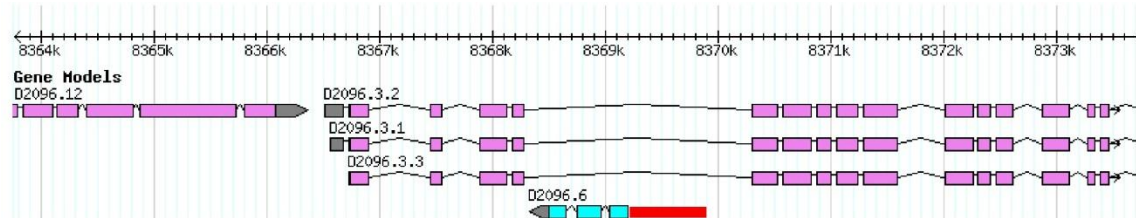
Figure 16: D2096.6 is a small gene located in the intron of another uncharacterized gene on opposite strand

D2096.6 is an 879 base pair transcript that encodes a 171 amino acid protein. A *D2096.6::gfp* transcriptional fusion containing 673 base pair 5'flanked DNA was constructed and expressed in the pharyngeal cells (red bar) provided by Susan Mango (Gaudet and Mango, 2002). Image downloaded and modified from http://www.wormbase.org, release WS224, date June 6,2011

Based on our microarray and semi-qPCR results we hypothesize that TBX-2 represses

D2096.6 expression. To test this hypothesis we examined expression of *D2096.6::gfp* in wild

type animals and *tbx-2* mutants. We obtained this GFP construct (kindly provided by Susan

Mango) and made transgenic lines expressing *D2096.6::gfp* as an extrachromosomal array. In

wild type worms *D2096.6::gfp* expression initiates at approximately the 200 cell stage of

embryogenesis, and is observed in one to two cells. Later in embryogenesis at the 1 ½ fold stage

D2096.6::gfp expression is typically observed in one cell in the developing pharynx (Figure

17A-C). The number of pharyngeal cells expressing *D2096.6::gfp* increases and by hatching

D2096.6::gfp is expressed in many of the pharyngeal muscles (Figure 19A, B). Expression

continues in the pharynx through the remainder of the life cycle, and in adults *D2096.6::gfp*

expression is observed in pharyngeal muscle and marginal cells.

We crossed *D2096.6::gfp* transgenic worms into *tbx-2(bx59)* and *tbx-2(ok529)* null

mutant backgrounds and found *D2096.6::gfp* was expressed in more cells in *tbx-2* mutants that

in wild type during embryogenesis and L1 larvae. Among *tbx-2(bx59)* animals that expressed

GFP we found 71% showed over expression of *D2096.6::gfp* and among *tbx-2(ok529)/dpy-*

17(e164)unc-32(e189) progeny we found 20% showed over expression of *D2096.6::gfp* (Table

X). We believe that this 20% represents the majority of *tbx-2(ok529)* homozygous animals. In

tbx-2(bx59) and *tbx-2(ok529)* embryos, we see additional cells expressing *D2096.6::gfp* in the

pharyngeal primordium as well as outside the pharynx and in body wall muscle cells (Figure 17;

Figure 18). In *tbx-2(bx59)* and *tbx-2(ok529)* larvae we see additional cells expressing

D2096.6::gfp in the intestine, body wall muscles and hypodermal cells (Figure 19). We also see

early onset of expression in pharyngeal marginal cells which is normally not seen until the adult

stage. Thus we conclude *D2096.6::gfp* expression increases both inside and outside the pharynx

in *tbx-2* mutants. This is consistent with our microarray and semi q-PCR results which indicate that the endogenous D2096.6 gene is over expressed in *tbx-2(bx59)* compared to wild type.

TABLE X: LOSS OF TBX-2, UBC-9 OR MUTATING BINDING SITE 2 RESULTS IN ECTOPIC *D2096.6::gfp* EXPRESSION

genotype	% animals with ectopic *D2096.6::gfp* (n)[1]
+/+; *D2096.6::gfp*[2]	5% (136)
tbx-2(bx59); D2096.6::gfp[3]	71% (55)
tbx-2(ok529)/tbx-2(ok529); D2096.6::gfp[4]	20% (35)
+/+; *D2096.6::gfp*[5] site 2 mutant	100% (61)
tbx-2(bx59); D2096.6::gfp site 2 mutant[6]	87% (38)
+/+; *D2096.6::gfp; ubc-9(RNAi)*[7]	75% (48)

All embryonic stages and L1 larvae that expressed GFP were scored for wild type or ectopic expression pattern based on the wild type expression pattern in OK0666.

[1] n= the total number for all embryonic stages and L1 counted
[2] wild type strain OK0666 carrying the *D2096.6::gfp*;*cuEx553*
[3] *tbx-2(bx59)* strain OK0692 carrying the *D2096.6::gfp*;*cuEx553*
[4] *tbx-2(ok529)/tbx-2(ok529)* progeny from *tbx-2(ok529)/dpy-17(e164)unc-32(e189)* strain OK0741 carrying the *D2096.6::gfp;cuEx553*
[5] wild type strain OK0801 carrying the *D2096.6::gfp*;*cuEx644*
[6] *tbx-2(bx59)* strain OK0802 carrying the *D2096.6::gfp*;*cuEx644*
[7] *ubc-9(RNAi)* was performed in strain OK0666 carrying the *D2096.6::gfp cuEx553*
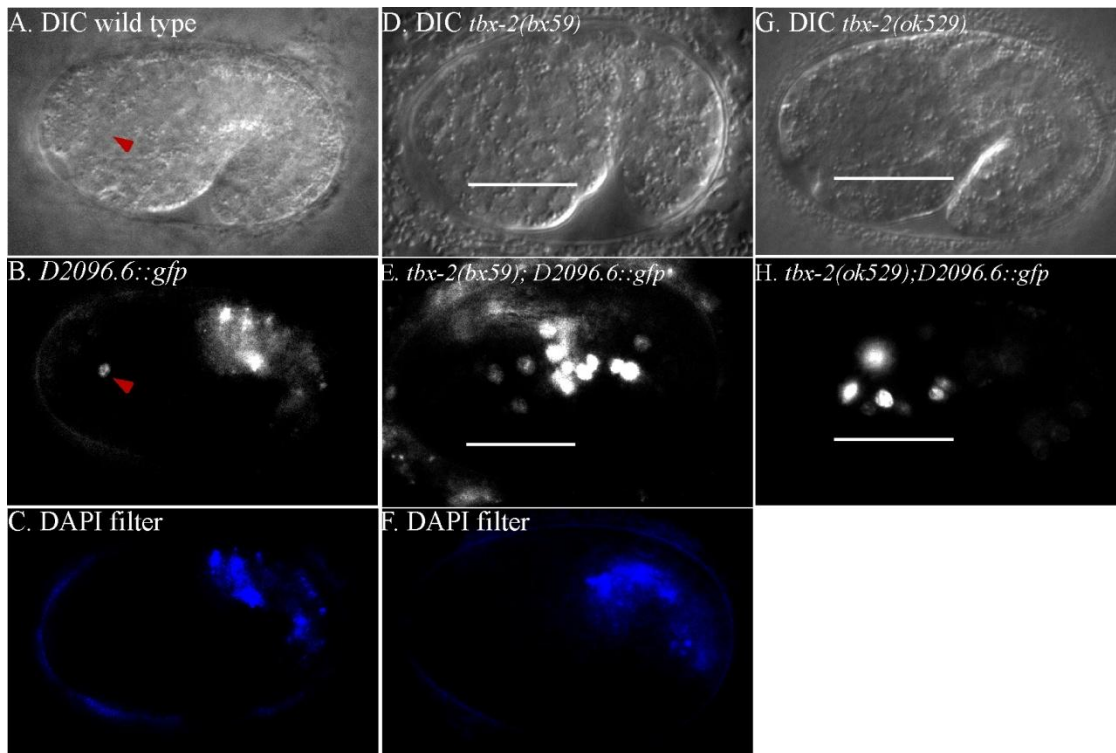
Figure 17: *D2096.6::gfp* is ectopically expressed in *tbx-2* mutant backgrounds during embryogenesis.

(A-C) 1 ½ fold stage wild type embryo expressing *D2096.6::gfp;cuEx553* (OK0666). (A) DIC image. (B) *D2096.6::gfp* is expressed in 1 cell in the head of the embryo, arrow head. (C) Autofluorescence in DAPI filter. (D-F) 1 ½ fold stage *tbx-2(bx59);cuEx553* mutant has ectopic *D2096.6::gfp* expression at the non-permissive temperature 25ºC (OK0692). (D) DIC image. (E) Ectopic *D2096.6::gfp* expression in the head, underlined (white bar). (F) Autofluorescence in DAPI filter.. (G,H) 1 ½ fold stage homozygous *tbx-2(ok529)* progeny of *tbx-2(ok529)/dpy-17(e164)unc-32(e189);cuEx553* (OK0741) has ectopic *D2096.6::gfp* expression in the head. (G) DIC image. (H) Ectopic *D2096.6::gfp* expression in the head, underlined (white bar).
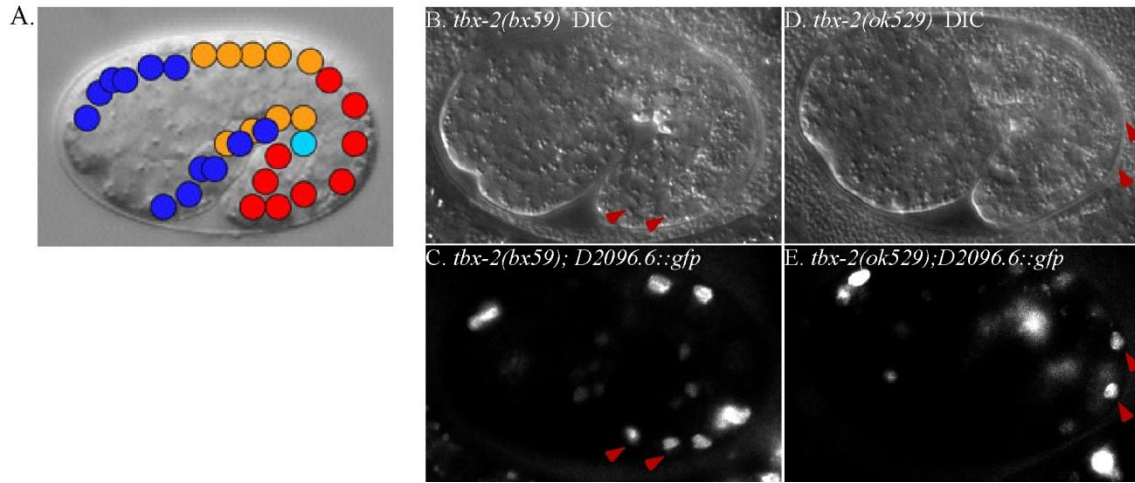
Figure 18: *D2096.6::gfp* is ectopically expressed in body wall muscle cells in *tbx-2* mutants during embryogenesis.

(A) An image of a 1 ½ fold stage embryo with body wall muscle nuclei super-imposed on the embryo. The four different colored nuclei represent body wall muscle nuclei from four different lineages: MS (blue), D (orange), C (red), AB (turquoise) (modified from (Okkema and Krause, 2005). (B, C) 1 ½ fold stage *tbx-2(bx59);cuEx553* (OK0692) mutant at the nonpermissive temperature (25ºC). (B) DIC image. (C) *D2096.6::gfp* expression in body wall muscle cells (arrow heads). (D, E) 1 ½ fold stage homozygous *tbx-2(ok529)* progeny of *tbx-2(ok529)/dpy-17(e164) unc-32(e189);cuEx553* (OK0741) mutants. (D) DIC image. (C) *tbx-2(ok529)* homozygote with *D2096.6::gfp* expression in body wall muscle cells (arrow heads).
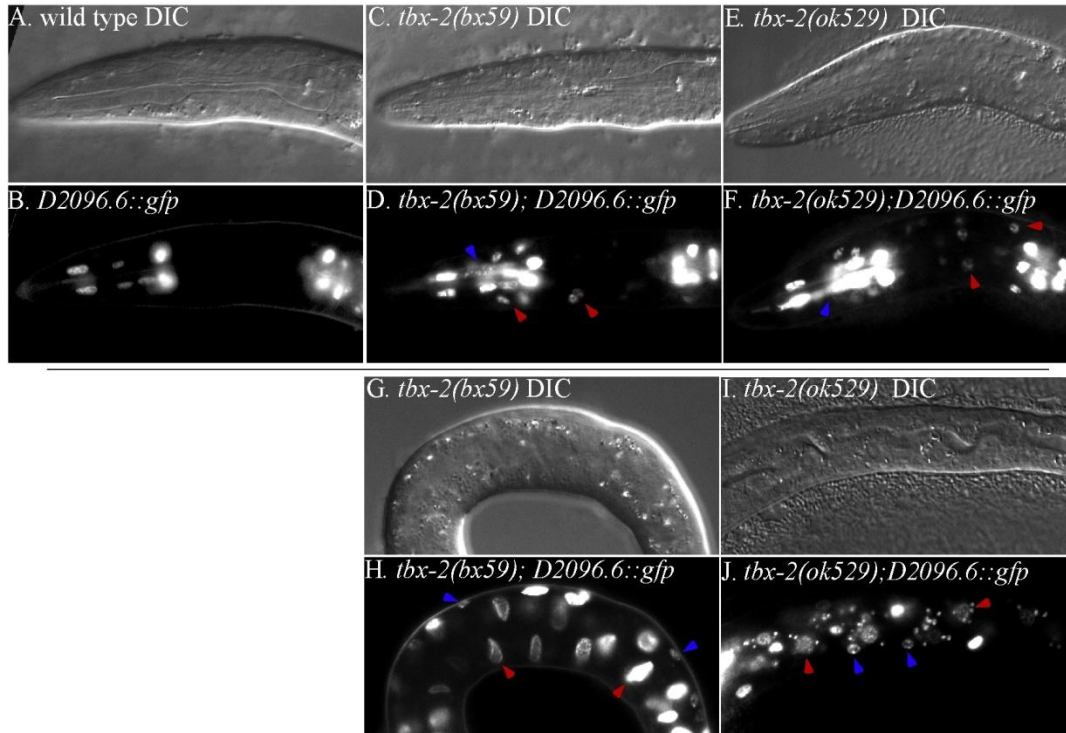
Figure 19: *D2096.6::gfp* is ectopically expressed in *tbx-2* mutant backgrounds during larval stages.

(A, B) Wild type L1 larva expressing *D2096.6::gfp;cuEx553*. (A) DIC image. (B) *D2096.6::gfp* expression in pharyngeal muscle cells. (C, D) *tbx-2(bx59);cuEx553* mutant larva at the nonpermissive temperature (25°C). (C) DIC image. (D) Ectopic *D2096.6::gfp* expression in marginal cells within the pharynx (blue arrow head) and outside the pharynx in hypodermal cells and neurons (red arrow heads). (E, F) Homozygous *tbx-2(ok529)* progeny from *tbx-2(ok529)/dpy-17(e164)unc-32(e189);cuEx553*. (E) DIC image. (F) Ectopic *D2096.6::gfp* expression in marginal cells within the pharynx (blue arrow head) and outside the pharynx in hypodermal cells and neurons (red arrow heads). (G, H) *tbx-2(bx59);cuEx553* mutant larva at the nonpermissive temperature (25°C). (C) DIC image. (D) Ectopic *D2096.6::gfp* expression in gut nuclei (red arrow head) and body wall muscle nuclei (blue arrow heads). (I,J) Homozygous *tbx-2(ok529)* progeny from *tbx-2(ok529)/dpy-17(e164)unc-32(e189);cuEx553*. (I) DIC image. (J) Homozygous *tbx-2(ok529)* progeny show ectopic expression in gut nuclei (red arrow head) and body wall muscle nuclei (blue arrow heads). Wild type larva do not express *D2096.6::gfp* in gut and body wall muscle nuclei.

4.4.6 *ubc-9(RNAi)* results in overexpression of D2096.6 similar to that seen in *tbx-2* mutants

We hypothesize TBX-2 is a SUMOylation dependent transcriptional repressor. Our previous genetic evidence suggests TBX-2 function is SUMOylation dependent, as reduction of the SUMO-conjugating enzyme UBC-9 produces pharyngeal phenotypes similar to loss of TBX-2. In addition, TBX-2 interacts with the SUMOylation E2 conjugating enzyme UBC-9 and the E3 SUMO ligase GEI-17 in yeast 2-hybrid assays (Roy Chowdhuri et al., 2006).

To test this hypothesis, we asked if reducing SUMOylation by performing *ubc-9(RNAi)* using the feeding method (Kamath et al., 2003) would affect *D2096.6::gfp* expression. *ubc-9(RNAi)* results in embryonic and larval lethality, and the most severely affected *ubc-9(RNAi)* animals have a highly disorganized morphology that makes it difficult to identify specific tissues (Roy Chowdhuri et al., 2006). Therefore we characterized *D2096.6::gfp* expression in older embryos that had undergone morphogenesis and the few surviving L1 larvae. We see that reduction of *ubc-9* by RNAi results in *D2096.6::gfp* expression in posterior body wall muscles in embryos in a pattern similar to that which we have observed in *tbx-2(bx59)* embryos (Figure 20 A-D). In larvae we see expression in body wall muscles (data not shown) and hypodermal cells in the posterior of the worm similar to the expression pattern we see in *tbx-2(bx59)* mutants (Figure 20 E-H). We also see additional GFP expressing cells outside the pharynx in the anterior of embryos and larva when compared to wild type (data not shown). Our results suggest TBX-2 functions as a transcriptional repressor whose function is dependent on SUMOylation.
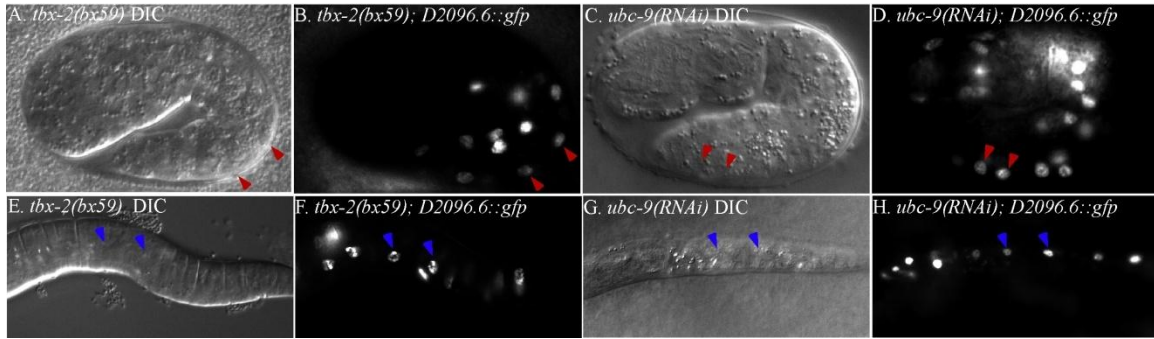
Figure 20: *ubc-9(RNAi)* results in over expression of D2096.6 in embryos and L1 larvae.

(A,B) 2 fold stage *tbx-2(bx59)* embryo at the non-permissive temperature 25ºC. (A) DIC image. (B) *tbx-2(bx59)* with ectopic *D2096.6::gfp* expression in the posterior body wall muscles (red arrowheads). (C,D) *ubc-9(RNAi)* embryo. (C) DIC image. (D) *ubc-9(RNAi)* embryo showing ectopic *D2096.6::gfp* expression in the posterior body wall muscles (red arrowheads). (E,F) *tbx-2(bx59)* L1 larvae at the non-permissive temperature 25ºC. (E) DIC image. (F) *tbx-2(bx59)* larva showing ectopic *D2096.6::gfp* expression in the posterior hypodermal cells (blue arrowheads). (G,H) *ubc-9(RNAi)* L1 larvae. (G) DIC image. (H) *ubc-9(RNAi)* larva showing ectopic *D2096.6::gfp* expression in the posterior hypodermal cells (blue arrowheads).

4.4.7 <u>TBX-2 specifically binds two T-box binding sites in the D2096.6 promoter</u>

To ask if D2096.6 is a direct target of TBX-2 we first analyzed its promoter sequence for T-box binding sites. *D2096.6::gfp* fusion contains 673 bp of sequence upstream of D2096.6 start site fused to GFP (Gaudet and Mango, 2002). Using JASPAR and ConSite we identified site 1 (Figure 21) which was a high scoring match to the Brachyury consensus sequence as well as a lower scoring matches. Site 1 is located 400 bp upstream of the start site in our promoter fusion and is conserved between *C. elegans*, *C. briggsae* and *C. remanei*.

In parallel, an analysis of all of the differentially regulated genes reported in our microarray has been performed and a modified position weighted matrix (newpwm) sequence for potential TBX-2 binding sites was constructed. This sequence was constructed by collecting all differentially expressed genes that contained Brachyury consensus binding sites within 500 bp of the start codon. These Brachyury consensus sites were then compiled and the newpwm was created based on the variations among these sites (Tom Ronan unpublished). A search with the newpwm sequence corresponded to a lower scoring Brachyury site identified in our previous analysis using JASPAR. This newpwm site 2 is located 100 bp upstream of the start site in our promoter fusion and is not conserved amongst other nematode species (Figure 21).
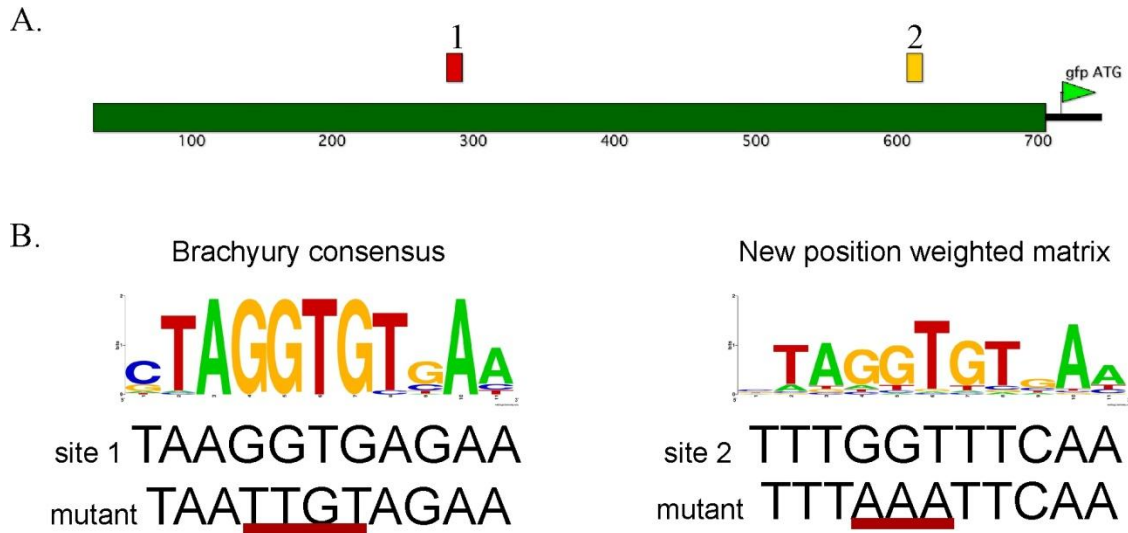
Figure 21: Diagram of D2096.6 promoter and T-box binding sites.

(A) *D2096.6::gfp* transcriptional fusion containing 673 base pair 5'flanked DNA, provided by Susan Mango. T-box binding site 1 (red box) is located approximately 400 bp upstream of the start site in our promoter fusion and is conserved between *C. elegans*, *C. briggsae* and *C. remanei*. T-box bind newpwm site 2 (yellow box) is located 100 bp upstream of the start site in our promoter fusion and is not conserved amongst other nematode species. (B) Sequences for position weighted matrices from JASPAR and ConSite and the matrix constructed by our lab are shown. Binding sites 1 and 2 identified using these matrices as well as mutations made for analysis are underlined in red.

Based on our promoter analysis we have identified two potential TBX-2 binding sites. To ask if TBX-2 binds either the consensus site 1 or the newpwm site 2, we performed gel shifts with purified TBX-2 protein. Our protein consists of the TBX-2 DNA binding domain (amino acids 53 to 193 of TBX-2/F21H11.3) fused to GST and was purified on a GST affinity column (Figure 22C). A 35 bp probe to site 1 and site 2 was designed which consisted of the 10 bp binding sequence with 12-13 bp of flanking sequence on either side.

We first asked if *C. elegans* TBX-2 is capable of binding the conserved Brachyury consensus binding site. This site is similar to those which most T-box proteins in other species have been shown to bind in vitro. To do this we used ^{32}P labeled probe for site 1, which is the conserved consensus T-box binding site in the D2096.6 promoter, and performed gel shift assays. We found that *C. elegans* TBX-2 binds an oligonucleotide containing site 1 in vitro but binding is significantly reduced to an oligonucleotide in which site 1 is mutated (Figure 23A). To show that this binding is specific, competition assays were performed with unlabeled probe 1, mutated probe 1 and probe 2 (Figure 23B). Excess probe 1 competes for binding, but we found that mutated probe 1 and probe 2 did not compete with binding to wild type probe 1. This data indicates TBX-2 specifically binds a consensus T-box binding site in vitro. This result is not surprising because examination of downstream targets and binding-site selection experiments have shown that members of the T-box family are capable of binding to the DNA consensus sequence AGGTGTGA (Wilson and Conlon, 2002), and this is very similar to the sequence of binding site 1.

Next, we performed gel shift assays with probe 2 to ask if TBX-2 can bind the newpwm site. We found that TBX-2 also binds probe 2 but it does not bind an oligonucleotide in which

site 2 is mutated (Figure 24A). We performed competition assays with unlabeled probe 2,

mutated probe 2 and probe 1 to ask if this binding is specific (Figure 24B). We found that wild

type probe 2 competed for binding but mutated probe 2 competed for binding less than wild type

probe 2, and probe 1 did not compete with binding. This binding is surprising because site 2 is

less similar to a typical T-box binding site. In addition we sometimes see two shifted complexes

with site 2 probe indicating TBX-2 may form monomer and dimer complexes. We located a

second partial binding site in the site 2 probe sequence which was identified by ConSite (Figure

24, underlined in blue) which may explain the appearance of multiple bands. This partial site is

similar to site 2 and further examination will need to be done to ask if TBX-2 binds this site. It is

possible that in *C. elegans* TBX-2 functions at a variant site and not a Brachyury consensus site;

however more targets are needed to answer this question. Our results suggest TBX-2 directly

binds a conserved Brachyury consensus binding site and a newpwm site in the D2096.6 promoter

in vitro.
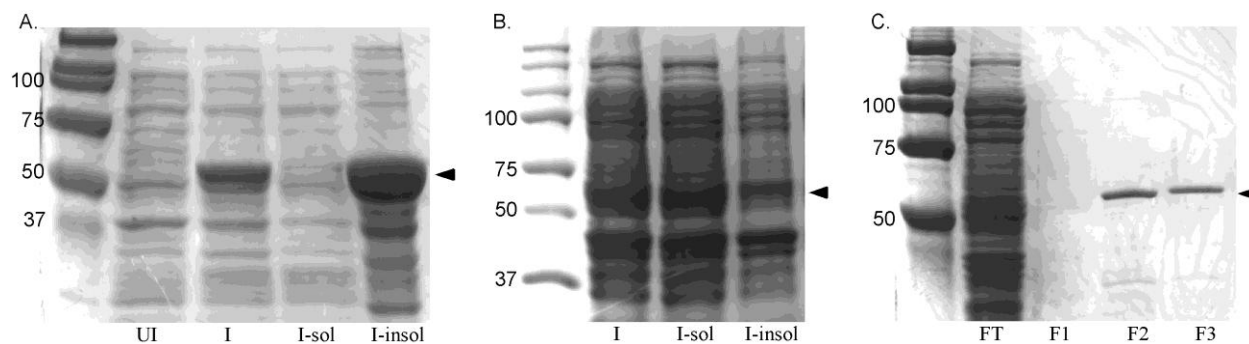
Figure 22: GST::TBX-2 expression and purification.

Arrowhead indicates GST::TBX-2 at 51 kd. (A) Coomassie staining shows GST::TBX-2 is induced (lane I) and it is induced in the insoluble fraction (lane I-insol) at 37ºC for 2 hours with 1 mM IPTG. (B) Gel shows induced protein and increased amounts in the soluble fraction (lane I-sol) when protein is grown in M9 minimal medium at 22ºC and induced with 0.5 mM IPTG for 24hours (protocol describes in methods of this chapter). (C) Gel shows purification of GST::TBX-2 on GST affinity column in fraction 2 and 3 (F2 and F3). FT indicates flow through not bound to column.
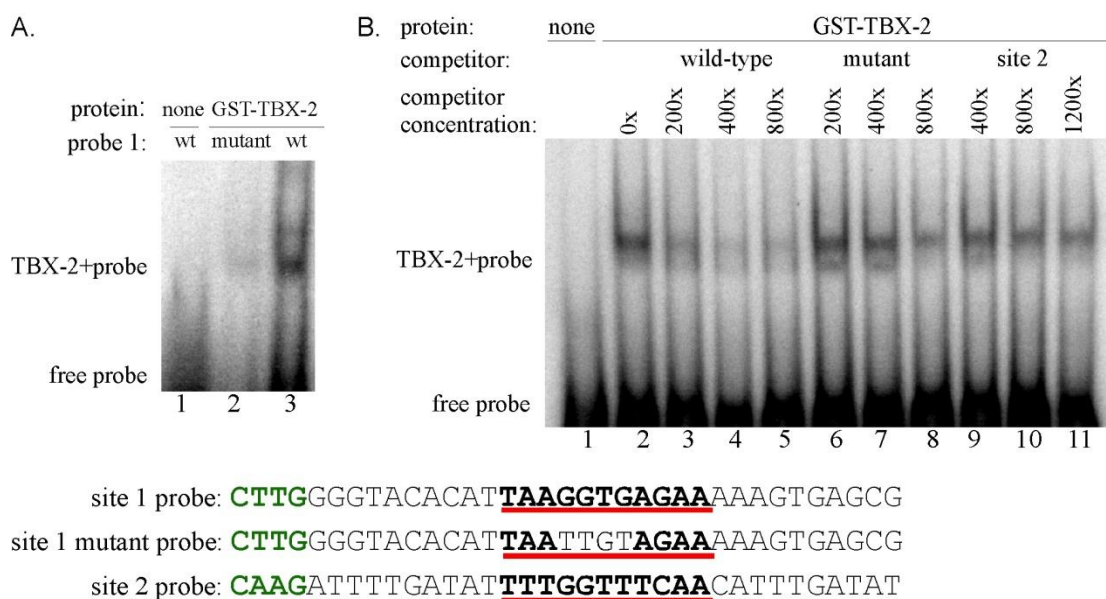
Figure 23: TBX-2 specifically binds the conserved Brachyury consensus site 1 in vitro.

Purified GST::TBX-2 containing (amino acids 53-193 of TBX-2/F21H11.3) the T-box domain was used in the following gel shift assays. Probes shown below are 35-36 nucleotides long and highlighted green were added to sequences for labeling with [32]P dCTP. The wild type or mutated binding site sequence is underlined in red. Nucleotides not mutated are in bold font (A) TBX-2 forms a shifted complex with site 1 probe. Lane 1 contains no TBX-2 protein, lanes 2 and 3 contain 200 ng of TBX-2 protein and mutated site 1 probe or site 1 probe respectively. (B) TBX-2 specifically binds site 1. Lane 1 contains no TBX-2 protein; all subsequent lanes contain 50 ng of TBX-2 protein. Lanes 3-5, binding is competed with site 1. Lanes 6-8, binding does not compete with 200x and 400x mutated site 1, slight competition may be seen with 800x mutated site 1. Lanes 9-11, binding is not competed with site 2.

A.

protein: none | GST-TBX-2
probe 2: wt | mutant | wt

TBX-2+probe

free probe

1　2　3

B.

protein: none | GST-TBX-2
competitor: | wild-type | mutant | site 1
competitor concentration: 0x 2x 6x 18x | 2x 6x 18x | 2x 6x 18x

TBX-2+probe

free probe

1　2　3　4　5　6　7　8　9　10　11

site 2 probe: **CAAG**ATTTTGATAT**TTTGGTTTCAA**CATTTGATAT

site 2 mutant probe: **CAAG**ATTTTGATAT**TTT**AAA**TTCAA**CATTTGATAT

site 1 probe: **CTTG**GGGTACACAT**TAAGGTGAGAA**AAAGTGAGCG

Figure 24: TBX-2 specifically binds the variant site 2 in vitro.

Purified GST::TBX-2 containing (amino acids 53-193 of TBX-2/F21H11.3) the T-box domain was used in the following gel shift assays. Probes shown below are 35-36 nucleotides long and highlighted green were added to sequences for labeling with ^{32}P dCTP. The wild type or mutated binding site sequence is underlined in red. Nucleotides not mutated are in bold font. A partial second potential binding site is underlined in blue in site 2 probe sequences. (A) TBX-2 forms a shifted complex with site 2 probe. Lane 1 contains no TBX-2 protein, lanes 2 and 3 contain 200 ng of TBX-2 protein and mutated site 2 probe or wild type site 2 probe respectively. (B) TBX-2 specifically binds site 2. Lane 1 contains no TBX-2 protein; all subsequent lanes contain 250 ng of TBX-2 protein. Lanes 3-5, binding is competed with increasing amounts of site 2. Lanes 6-8, binding is competed but to a lesser extent with increasing amounts of mutated site 2. Lanes 9-11, binding is not competed with site 1. Open and filled arrowheads indicate possible monomer and dimer binding complexes.

4.4.8   D2096.6 expression is regulated by activator and repressor T-box binding sites

We have shown that the TBX-2 binding domain protein is able to bind two sites in the D2096.6 promoter. Based on our microarray and GFP expression data, we believe that TBX-2 is directly repressing D2096.6 expression through one of these sites. To test this hypothesis we mutated the Brachyury consensus site 1 and the newpwm site 2 and asked how these mutations affect D2096.6 expression.

We first mutated the conserved Brachyury consensus site 1 and found that this reduced *D2096.6::gfp* expression, suggesting site 1 functions as an activator site. We mutated 4 base pairs in the binding site from GGTG to TTGT using site directed mutagenesis (Figure 21B). This is the same mutation that reduced TBX-2 binding in vitro. We examined transgenic worms expressing the *D2096.6::gfp* promoter fusion with site 1 mutated and found *D2096.6::gfp* expressed later in embryogenesis and in fewer cells in larvae (Figure 25, Table XI). In site 1 mutants we see expression initiate at the 1 ½ fold stage in embryos where as in wild type and *tbx-2(bx59)* embryos we see D2096.6 expression at approximately the 200 cell stage (Table XI). Of the animals that expressed GFP we found a reduction in *D2096.6::gfp* expression in 68% percent of embryos and L1 animals (Table XII).

We believe that TBX-2 binds a site in the D2096.6 promoter to repress expression. Based on these results site 1 appears to function as an activator site and it is possible that a T-box protein may bind this site to activate D2096.6 expression. However, we do not believe that this site is targeted by TBX-2, our previous results examining both *D2096.6::gfp* and endogenous D2096.6 mRNA indicates D2096.6 expression is increased in the *tbx-2(bx59)* and *tbx-2(ok529)* null backgrounds. In addition our microarray and semi-qPCR indicate that *D2096.6* mRNA is

over expressed in *tbx-2(bx59)* compared to wild type. We crossed the binding site 1 mutant into

the *tbx-2(bx59)* background and found no change in the expression pattern. We found that 67%

of animals showed a reduction in *D2096.6::gfp* expression, which is the same expression pattern

we see when site 1 is mutated in a wild type background (Table XII).

TABLE XI: MUTATIONS IN BINDING SITE 1 RESULT IN DELAYED ONSET OF
D2096.6::gfp EXPRESSION

Embryonic stage	% animals with GFP expressing cells (n)			
	+/+;D2096.6::gfp[1]	*tbx2(bx59); D2096.6::gfp* [2]	*+/+; D2096.6::gfp* site 1 mutant[3]	*+/+; D2096.6::gfp* site 2 mutant [4]
200 cell	2 % (95)	5% (72)	0% (22)	40% (5)
bean	12% (32)	17% (18)	0% (21)	40% (10)
comma	11% (42)	31% (13)	0% (21)	50% (12)
1 ½	5% (28)	44% (50)	11% (19)	NA
2 fold	33% (31)	35% (26)	5% (19)	60% (5)

[1] Wild type strain OK0666 carrying the *D2096.6::gfp*;cuEx553
[2] *tbx-2(bx59)* strain OK0692 carrying the *D2096.6::gfp;cuEx553*
[3] Wild type strain OK0694 carrying the *D2096.6::gfp;cuEx566* with site 1 mutated
[4] Wild type strain OK0801 carrying the *D2096.6::gfp;cuEx644* with site 2 mutated

TABLE XII: MUTATION OF THE CONSERVED BRACHYURY BINDING SITE 1
RESULTS IN REDUCED *D2096.6::gfp* EXPRESSION

genotype	% animals with reduced *D2096.6::gfp* (n)[1]
+/+; D2096.6::gfp[2]	4% (136)
+/+; D2096.6::gfp site 1 mutant[3]	68% (129)
tbx-2(bx59); D2096.6::gfp site 1 mutant [4]	67% (60)

All embryonic stages and L1 larvae that expressed GFP were scored for wild type or reduced
expression pattern based on the wild type expression pattern in OK0666.

[1] n= the total number for all embryonic stages and L1 counted
[2] Wild type strain OK0666 carrying the *D2096.6::gfp;cuEx553*
[3] Wild type strain OK0694 carrying the *D2096.6::gfp;cuEx566* with site 1 mutated
[4] *tbx-2(bx59)* strain OK0696 carrying the *D2096.6::gfp; cuEx566* with site 1 mutated
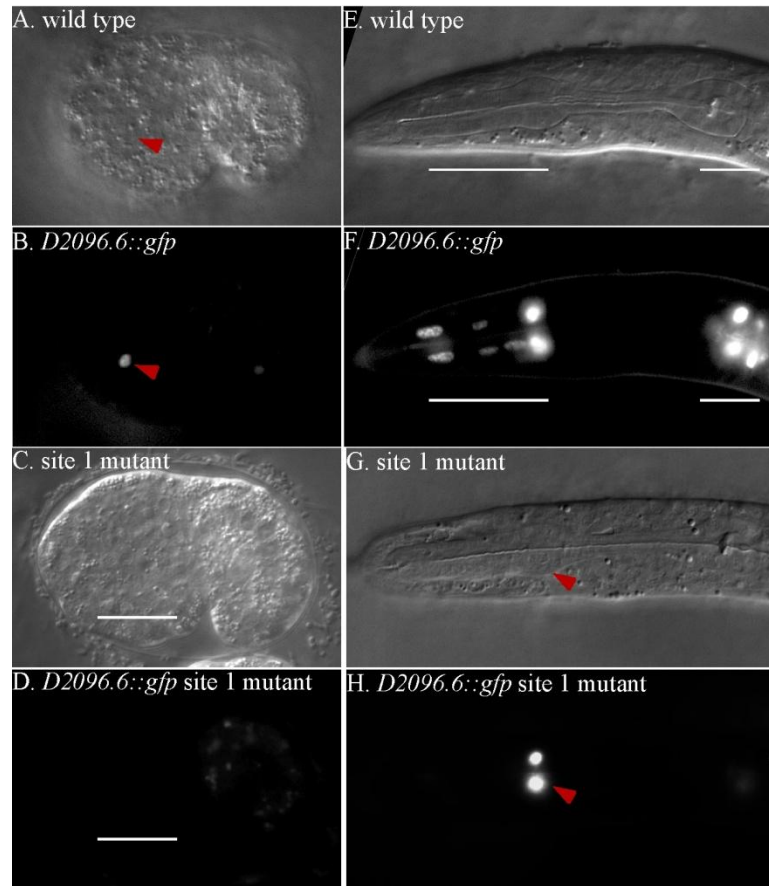
Figure 25: Mutations in the conserved Brachyury binding site 1 result in decreased expression of *D2096.6::gfp*.

(A, B) Comma stage wild type embryo expressing of *D2096.6::gfp* (OK0666). (A) DIC image. (B) Wild type embryo expressing *D2096.6::gfp* in one cell in the head (red arrowhead). (C, D) Comma stage embryo carrying *D2096.6::gfp* with site 1 mutated (OK0694). (C) DIC image. (D) Comma stage embryo showing no *D2096.6::gfp* expression in the head (white bar). (E, F) Wild type larva shown in Figure 19, expressing of *D2096.6::gfp*. (E) DIC image. (F) *D2096.6::gfp* expression in pharyngeal muscle cells (underlined by white bar). (G, H) Larva carrying *D2096.6::gfp* with site 1 mutated (OK0694). (G) DIC image. (H) L1 larvae showing a reduced number of pharyngeal muscle cells expressing *D2096.6::gfp* (red arrowhead).

Mutations in binding site 2 result in ectopic expression indicating a repressor binds this site. Binding site 2 was identified by JASPAR and the newpwm model and is not conserved among other nematode species. We mutated 3 base pairs in the binding site from ACC to TTT using site directed mutagenesis (Figure 21B). We examined transgenic worms expressing the *D2096.6::gfp* promoter fusion with site 2 mutated and found the onset of *D2096.6::gfp* is comparable to what we found in the wild type strain but we see the spatial pattern has changed (Table XI). Throughout embryogenesis and in L1 larvae we see that 100% of animals show ectopic expression of *D2096.6::gfp* in the anterior and/or posterior regions (Table X). In embryos we see additional cells expressing GFP in the pharyngeal primordium as well as outside the pharynx and in the posterior region (Figure 26). In larvae we see additional expression inside and outside the pharynx as well as the intestine and muscle cells (Figure 27). This over expression resembles the over expression we see in *tbx-2(bx59)* and *tbx-2(ok529)* background, and is consistent with our microarray and semi-qPCR which indicates D2096.6 is over expressed. In addition, we crossed the binding site 2 mutant into the *tbx-2(bx59)* background and found ectopic expression in 87% of embryos and L1 larvae, similar to what we see when site 2 is mutated in a wild type background (Table X). This suggests that TBX-2 is not necessary for the activation of this promoter. Based on these results, we hypothesize that TBX-2 binds site 2 to repress D2096.6 expression.
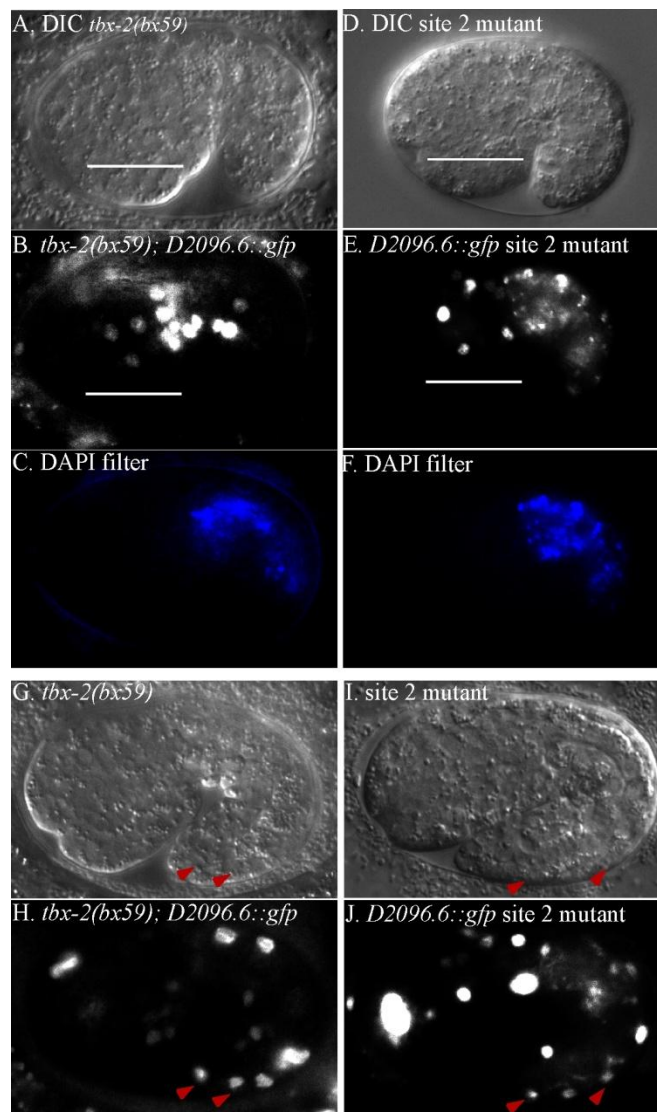
Figure 26: Mutations in binding site 2 result in ectopic expression of *D2096.6::gfp* similar to ectopic expression in *tbx-2* mutant embryos.

(A-C) *tbx-2(bx59)* 1 ½ fold embryo at non-permissive temperature (25ºC) carrying the *D2096.6::gfp* (OK0692) shown previously in Figure17. (A) DIC image. (B) Ectopic *D2096.6::gfp* expression in the head (white bar). (C) DAPI filter showing autofluorescence in gut granules. (D-F) Wild type 1 ½ fold embryo expressing *D2096.6::gfp* with site 2 mutated (OK0801). (D) DIC image. (E) Ectopic *D2096.6::gfp* expression in the head (white bar). (F) DAPI filter showing autofluorescence in gut granules (F). (G,H) *tbx-2(bx59)* 1 ½ fold stage embryo at the nonpermissive temperature (25ºC) previously shown in Figure 18. (G) DIC image. (H) Ectopic *D2096.6::gfp* expression in body wall muscle cells (red arrowheads). (I,J) Wild type 1 ½ fold embryo expressing *D2096.6::gfp* with site 2 mutated (OK0801). (I) DIC image. (J) Ectopic *D2096.6::gfp* expression in body wall muscle cells (red arrowheads)**.**
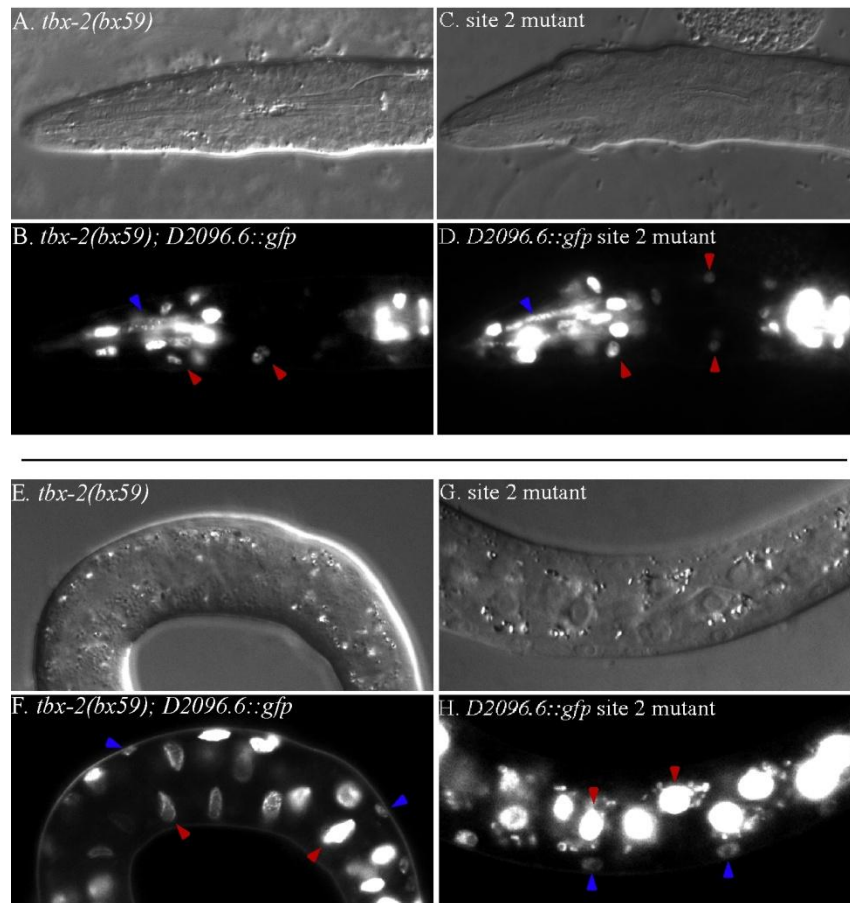
Figure 27: Mutations in binding site 2 result in ectopic expression of *D2096.6::gfp* similar to ectopic expression in *tbx-2* mutant larvae.

(A-B) *tbx-2(bx59)*;*D2096.6::gfp* larva at the nonpermissive temperature (25ºC) previously shown in Figure 19. (A) DIC image. (B) Ectopic *D2096.6::gfp* expression in marginal cells within the pharynx (blue arrow head) and outside the pharynx in hypodermal cells and neurons (red arrow heads). (C,D) Wild type larva expressing *D2096.6::gfp* with binding site 2 mutated (OK0801). (C) DIC image. (D) Ectopic *D2096.6::gfp* expression in marginal cells within the pharynx (blue arrow head) and outside the pharynx in hypodermal cells and neurons (red arrow heads). (E,F) *tbx-2(bx59)*;*D2096.6::gfp* larva at the nonpermissive temperature (25ºC) previously shown in Figure 19. (E) DIC image. (F) Ectopic *D2096.6::gfp* expression in gut nuclei (red arrow head) and body wall muscle nuclei (blue arrow heads). (G, H) Wild type larva expressing *D2096.6::gfp* with binding site 2 mutated (OK0801). (G) DIC image. (H) Ectopic expression in gut nuclei (red arrow heads) and body wall muscle nuclei (blue arrow heads).

4.4.9   <u>A balance between activator and repressor sites is needed for proper regulation</u>

      Our results indicate that site 1 is an activator site and site 2 is a repressor site. We wanted to ask if we would lose expression or gain expression of *D2096.6::gfp* if both site 1 and 2 were mutated. We found that mutation of both activator and repressor T-box sites results in a wild type expression pattern in 92% of worms (Table XIII). *D2096.6::gfp* promoter fusions containing mutations for site 1 and 2 respectively, were cloned using restriction enzymes to obtain a GFP promoter fusion containing both mutated site 1 and 2 together. We examined transgenic worms expressing the *D2096.6::gfp* promoter fusion with site 1 and 2 mutated and found *D2096.6::gfp* expressed in a wild type expression pattern (Figure 28). In addition we crossed this binding site mutant into the *tbx-2(bx59)* background and found a wild type expression pattern in 79% of worms. This suggests that TBX-2 is not functioning at additional unidentified sites in this promoter region.

      We know that PHA-4 activates D2096.6 expression at the promoter therefore it was not unexpected to see expression in this mutant. However a wild type expression pattern suggests that either PHA-4 regulation is sufficient for wild type expression or there are additional activator and repressor sites regulating D2096.6. We hypothesize that there may be a balance needed between activating and repressing sites to maintain expression at proper levels.

TABLE XIII: A BALANCE BETWEEN ACTIVATOR AND REPRESSOR SITES IS NEEDED FOR WILD TYPE *D2096.6::gfp* EXPRESSION

genotype	% animals with wild type *D2096.6::gfp* expression (n)[1]
+/+; D2096.6::gfp [2]	91% (136)
+/+; D2096.6::gfp site 1 and 2 mutant [3]	92% (50)
tbx-2(bx59); D2096.6::gfp site 1 and 2 mutant[4]	79% (50)

All embryonic stages and L1 larvae that expressed GFP were scored for wild type expression pattern based on the wild type expression pattern in OK0666.

[1] n= the total number for all embryonic stages and L1 counted
[2] Wild type strain OK0666 carrying the *D2096.6::gfp*;*cuEx553*
[3] Wild type strain OK0737 carrying the *D2096.6::gfp*;*cuEx589* with site 1 and 2 mutated .
[4] *tbx-2(bx59)* strain OK0799 carrying the *D2096.6::gfp*;*cuEx589* with site 1 and 2 mutated**.**
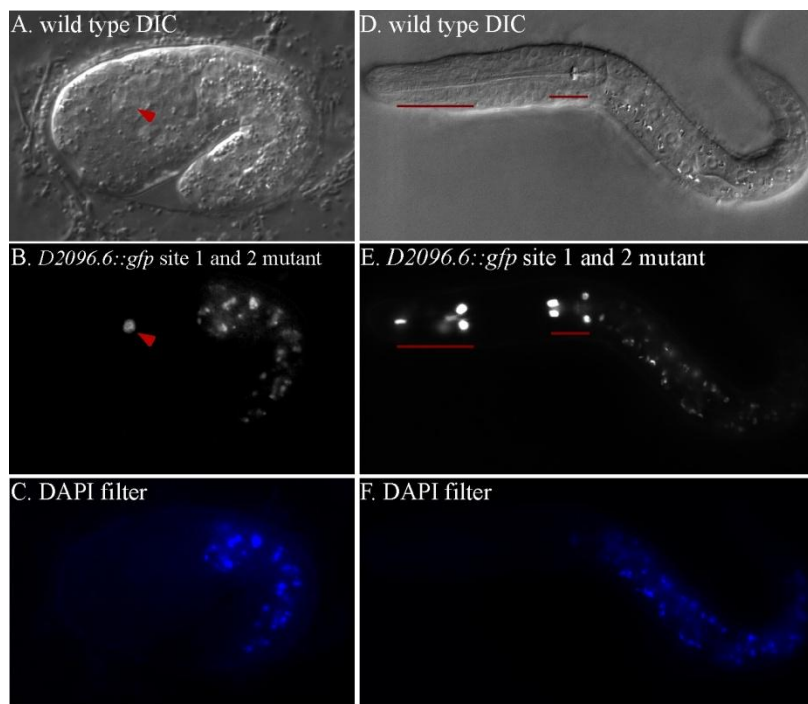
Figure 28: Mutation of both activator and repressor T-box sites results in wild type
D2096.6::gfp **expression.**

(A-F) OK0737 wild type strain carrying *D2096.6::gfp* with site 1 and 2 mutated. (A-C) 1 ½ fold stage embryo. (A) DIC image. (B) *D2096.6::gfp* expression in 1 cell in the head (red arrowhead). Wild type *D2096.6::gfp* expression pattern shown in Figure 17. (C) DAPI filter showing autofluorescence in gut granules. (D-F) L1 larva. (D) DIC image. (E) *D2096.6::gfp* expression in pharyngeal muscle cells (underlined by red bars). Wild type *D2096.6::gfp* expression pattern shown in Figure 19. (F) DAPI filter showing autofluorescence in gut granules.

4.5   Discussion

We have identified the first direct target of the *C. elegans* T-box transcription factor

TBX-2. To identify targets of TBX-2, we compared mRNA expression levels in wild-type and

tbx-2(bx59) mutant embryos using Affymetrix microarrays. Our initial goal was to identify

targets that are expressed in the pharynx and that may have a role in pharyngeal development. To

begin our analysis of the microarray data, we compared our list of differentially regulated genes

to existing data sets on pharyngeal gene expression and phylogenetic footprinting to identify

promoters with consensus T-box factor binding sites. We analyzed a subset of genes and

identified D2096.6 as a direct target of TBX-2. D2096.6 is expressed in pharyngeal cells and

results from our microarray and semi-qPCR indicate D2096.6 is directly repressed by TBX-2.

4.4.10 <u>TBX-2 specifically binds a variant T-box binding site to regulate D2096.6</u>

A direct target of TBX-2 has not been previously identified but we believe that TBX-2

functions as a repressor. *C. elegans* TBX-2 is most closely related to mammalian TBX2 and

TBX3, which function as repressors. Previous evidence also suggests that TBX-2 function

requires SUMOylation (Roy Chowdhuri et al., 2006) a post-translational modification usually

associated with repressor activity (Gill, 2005).   TBX-2 has also been shown to interact with the

Groucho corepressor UNC-37 in yeast two-hybrid experiments (Roy Chowdhuri et al., 2006).

We hypothesize that TBX-2 represses D2096.6 by binding a variant T-box binding site in

the promoter. Transgenic worms expressing the *D2096.6::gfp* promoter fusion with site 2

mutated show *D2096.6::gfp* is over expressed similar to the over expression seen in TBX-2

mutant strains. Consistent with these findings, we found that purified TBX-2 DNA binding

domain protein specifically binds site 2 in vitro.

The gene, Brachyury, is the founder of the T-box family and in vitro site selection experiments identified a 24 base pair imperfect palindromic sequence as the preferred binding site for the Brachyury protein. A half site was identified as AG<u>GTG</u>TGAA and extensive studies have shown that most T-box proteins can bind as monomers to this half site. In addition it was found that the core GTG bases were conserved in all strong binding sites that were selected (Herrmann et al., 1990; Kispert and Herrmann, 1993). The crystal structures for the T-domain of human TBX3 bound to DNA and for Brachyury bound to DNA have been solved. Based on this crystal structure data, a more stringent consensus sequence, xxGTGxxAx, which identifies the critical residues where TBX3 contacts DNA (Coll et al., 2002). This site allows for more variation in the T-box consensus sequence and since this study was published, additional targets have been identified that support target sequence variation (Di Gregorio and Levine, 1999). Mammalian TBX2 and TBX3 were found to bind a variant palindromic T-site, CACCNNNGGTG, in the human p14[ARF] promoter. This site matched 13 out of 20 nucleotides of the consensus T-box site, and contains the GTG core bases (Lingbeek et al., 2002).

C. elegans TBX-2 binds two sites in the D2096.6 promoter in vitro. Site 1 is a close match to the Brachyury consensus binding sequence and is conserved in other nematode species while site 2 is a variant T-box site that is not conserved. Site 2 lacks the GTG core bases found in most T-box binding sites but contains 3 out of 4 bases thought to be important based on the TBX3 crystal structure, and these bases may be sufficient for binding (Coll et al 2002). In addition, results from our GFP promoter fusions suggest that this is a functional site.

TBX-2 binds site 1 and site 2 differently. Site 1 has a higher binding activity but needs a lot of excess oligonucleotide to compete binding. Site 2 has less binding activity and needs very

little excess oligonucleotide to compete for binding. This suggests that there may be two conformations of TBX-2 protein in our purified sample. One form binds site 1 and is in abundance and is not all bound by our probe so it takes much more oligonucleotide to compete for binding. The second form of protein which binds site 2 is not in abundance and is all bound by our probe, therefore less oligonucleotide is needed to compete with binding. In addition our site 1 oligonucleotide does not appear to compete for binding to site 2 and our site 2 oligonucleotide does not compete for binding to site 1. Indicating there is a distinct difference in TBX-2 binding to each of these sites.

4.4.11 D2096.6 expression is regulated by multiple transcription factors

Analysis of the D2096.6 promoter GFP fusion suggests multiple transcription factors are involved in regulation D2096.6 expression. We analyzed the D2096.6 promoter and identified four potential T-box binding sites through JASPAR. We mutated site 1 and 2 and found these two sites are functional. Mutating site 1 reduced *D2096.6::gfp* expression resulting in a delayed onset of D2096.6 expression and a reduction in the number of cells that express GFP. Mutating site 2 resulted in an increase in the number of cells expressing D2096.6 GFP which was similar to what we observed when TBX-2 was reduced in mutant backgrounds. Based on our results we believe that TBX-2 directly binds site 2 to repress D2096.6 expression, while an unknown T-box protein is binding the Brachyury consensus site 1 to activate expression. When we mutate both sites 1 and site 2 together we see a wild type expression pattern for D2096.6. This data suggests that a balance of activating and repressing T-box functions are necessary for normal D2096.6 expression. JASPAR and ConSite databases identify lower scoring Brachyury sites present in the D2096.6 promoter that we have not yet examined and therefore do not know if these sites affect D2096.6 expression.

Multiple T-box proteins have been shown to bind a single promoter and regulate expression. *Nppa* (<u>n</u>atriuetic <u>p</u>recursor <u>p</u>eptide <u>A</u>) is a cardiac gene and was initially identified as a target for Tbx5 and contains three T-box bonding sites (Bruneau et al., 2001; Hiroi et al., 2001). It was later discovered that Tbx5 activates *Nppa* while Tbx2 and Tbx3 repress Nppa (Habets et al., 2002; Hoogaars et al.,2004). In addition the cardiac expressed gene *Connexin 40* was also found to be activated by Tbx5 and repressed by Tbx2 and Tbx3 (Hiroi et al., 2001; Bruneau etal., 2001; Christoffels et al., 2004;Hoogaars et al., 2004).

Other transcription factors also target the D2096.6 promoter. PHA-4 activates D2096.6 expression at two possible PHA-4 binding sites in the promoter region (Gaudet and Mango 2002). PHA-4 is a Forkhead box A transcription factor that is essential for development of the pharynx and is believed to regulate most pharyngeal genes. These sites are located approximately 30 bp upstream of site 1 in the D2096.6 promoter. TBX-2 and PHA-4 are dependent on each other to maintain expression and suggest a direct or indirect positive regulatory loop (Smith and Mango 2007). If PHA-4 and TBX-2 function in a positive regulatory loop, then it is possible that a balance is needed between PHA-4 and TBX-2 for proper regulation of D2096.6 expression. D2096.6 also contains potential binding sites for CEH-22, a pharyngeal muscle specific homeodomain factor related to the vertebrate heart specific factor NKx2-5. In the mouse NKx2-5 regulates cardiac gene expression in combination with both T-box activators and repressors, and a similar mechanism may be used to regulate D2096.6 (Hoogars 2004; Habets 2002).

4.4.12 <u>Factors that contribute to target specificity in T-box family proteins</u>

We found that our TBX-2 protein was capable of binding the consensus Brachyury sequence, site 1. We do not think that TBX-2 binds and functions at the consensus site 1 in vivo

to regulate D2096.6 expression. Interestingly, unlabeled site 1 probe was not able to compete for

TBX-2 binding to site 2 and unlabeled site 2 probe did not compete for binding to site 1;

suggesting that TBX-2 protein may bind these two sequences in different conformations. It is

also possible that the TBX-2 protein may have two different binding sites within the T-domain.

We currently have not distinguished between these two possibilities. Many T-box proteins have

been shown to have overlapping expression patterns but show differences in function and target

specificity (reviewed in Naiche et al., 2005). It is not clear what determines specificity of most

T-box proteins.

It has been suggested that T-box protein specificity may be due to differences in binding

affinities to target sequences or interactions with other transcription factors may play a role in

determining specificity (reviewed in Naiche et al., 2005). DNA binding experiments on T-box

proteins Tbx20 and Tbx5 suggest that the concentrations of different T-box members as well as

T half site affinities may affect if activator or repressor complexes are targeted to promoters

(Macindoe 2009). Specificity of T-domain proteins may be determined by interactions with

other transcription factors. Analysis in mouse and cell culture show cardiac T-box proteins

interact with the transcription factors GATA-4 and Nkx2-5, and these interactions are important

for heart development (Maitra et al., 2009; Hiroi et al. 2001; Ferin et al., 2007). It is possible that

the *C. elegans* TBX-2 interacts with PHA-4 or CEH-22 or both to target and regulate D2096.6.

4.4.13 TBX-2 repression of D2096.6 may be SUMOylation dependent
We hypothesize TBX-2 functions as a SUMOylation dependent repressor. It has been

previously shown that UBC-9 knockdown produces pharyngeal defects similar to *tbx-2* mutants,

and TBX-2 interacts with the SUMOylation E2 conjugating enzyme UBC-9 and the E3 SUMO

ligase GEI-17 in yeast 2-hybrid assays (Roy Chowdhuri 2006). Additional results from our lab are consistent with SUMOylation dependent function of TBX-2. TBX-2 can be SUMOylated in mammalian cell culture assays (Paul Huber unpublished), and RNAi knockdown of SUMOylation components enhance lethality of pharyngeal defects in the hypomorphic allele *tbx-2(bx59)* (Tanya Crum unpublished).

Here we show that *ubc-9(RNAi)* results in derepressed D2096.6 expression in a nearly identical pattern to that seen in *tbx-2* mutants or when site 2 is mutated in the D2096.6 promoter. Therefore we suggest SUMOylated TBX-2 directly represses D2096.6 expression through site 2. However, because *ubc-9(RNAi)* globally reduces SUMOylation we cannot rule out that other SUMOylation dependent repressors target D2096.6.

We have identified the first TBX-2 target in *C. elegans,* D2096.6. Further analysis of the D2096.6 promoter is needed to understand if this gene is regulated by multiple T-box proteins functioning as activator and repressors, and if TBX-2 regulation depends on interactions with other transcription factors. In addition functional analysis will need to be done on D2096.6 to understand its role in *C. elegans* development. Our microarray data along with the identification of TBX-2 binding site sequences will allow us to identify additional targets in *C. elegans* and further our understanding of T-box family regulation.

V. GENERAL DISCUSSION OF EGRH-1 AND TBX-2

5.1   EGR and T-box family genes are important developmental regulators

Mammalian EGR-family genes and T-box family genes are broadly expressed and mutants result in a variety of defects and diseases (O'Donovan et al., 1999). For example, Egr1 mutant mice have fertility defects and fail to ovulate (Lee et al., 1996; Topilko et al., 1998). Egr3 mutant mice exhibit defects in muscle spindle stretch receptors (Tourtellotte and Milbrandt, 1998; Tourtellotte et al., 2001). In recent years human diseases have been identified in which multiple T-box factors are altered. Haploinsufficiency of TBX5, TBX3 and TBX1 causes Holt-Oram syndrome, Ulnar-mammary syndrome and DiGeorge Syndrome respectively (Basson et al., 1997). This thesis describes the characterization of two conserved transcription factors from the EGR and T-box families. The invariant cell lineage, simple anatomy and wide range of genetic tools available in *C. elegans* have allowed us to characterize the function of the EGR family gene EGRH-1 and to identify the first direct target of T-box family member TBX-2. The information we have gained as well as our ongoing projects will continue to advance our understanding of these transcription factor families and how they control development.

5.2   Understanding EGRH-1 function in the sheath cells and gut through further mutant characterization and identification of downstream targets

We have characterized the role of the *C. elegans* EGR-family factor EGRH-1 during oogenesis. We find that *egrh-1* mutants exhibit ectopic oocyte development in the distal gonad and sperm independent oocyte maturation and ovulation. Mosaic analysis indicates that EGRH-1 function is likely required in the somatic gonad, which is descended from the MS blastomere, and more surprisingly in the gut, which is descended from the E blastomere. Expression of *egrh-1* in the gut rescues the formation of ectopic oocytes in *egrh-1* mutants, but does not rescue

sperm independent oocyte maturation and ovulation. This result suggests that EGRH-1 has distinct functions in the sheath cells and gut.

5.2.1   Further characterization of the EGRH-1 mutant phenotype

We do not understand how *egrh-1* loss in the gut leads to formation of distal oocytes. The gut produces complexes containing yolk proteins and precursors for signaling molecules involved in sperm recruitment that are transported into the psuedocoelom and are ultimately endocytosed by oocytes (Kimble and Sharrock, 1983; Hall et al., 1999; Kubagawa et al., 2006a). *rme-2* encodes the yolk receptor and *rme-2* mutants that are defective in endocytosis of yolk complexes exhibit ovulation defects (Grant and Hirsh, 1999). Receptor-mediated endocytosis is an essential process in all eukaryotes and is required for the uptake of nutrients and recycling of membranes and membrane proteins (Mukherjee *et al.*, 1997). The basic components and pathways of endocytic trafficking are well conserved between *C. elegans* and vertebrates. An assay has been developed using a YP170::GFP reporter to visualize yolk endocytosis in vivo in the *C. elegans* oocyte. This fusion protein is transported like endogenous yolk, from intestine to oocyte and can be visualized by fluorescence microscopy (Grant and Hirsh, 1999) and will allow us to ask if endocytosis of yolk complexes is defective in EGRH-1 mutants.

5.2.2   Identification of EGRH-1 target genes

To understand how EGRH-1 functions in oocyte development we would like to identify downstream targets. With the experience gained from performing a microarray with *tbx-2(bx59)* we are in an ideal position to apply the same methods used with the *tbx-2(bx59)* microarray on an *egrh-1(tm1736)* microarray.

Some advantages with *egrh-1(tm1736)* compared to *tbx-2(bx59)* is that *egrh-1(tm1736)* is a null allele and these mutants are viable. In addition we understand the expression pattern of

EGRH-1 and were able to rescue *egrh-1(tm1736)* mutant phenotypes. EGRH-1 protein shares

high sequence identity with EGR-family factors in the zinc-finger region, including residues that

contact DNA (Wolfe et al., 2000) which suggests EGRH-1 binds the same GC-rich sequence

motif GCGKGGGCG recognized by mammalian EGR-family factors. This motif is found at 469

distinct sites in the *C. elegans* genome (Markstein et al., 2002). Having a well characterized

binding site that is not abundant in the *C. elegans* genome will aid in our microarray data

analysis. Unlike TBX-2 we were able to produce an antibody for the EGRH-1 protein and this

could be used in ChIP (chromatin immunoprecipitation) analyses to test potential targets

identified from the microarray data.

**5.3	Understanding TBX-2 function in pharyngeal development by characterizing
D2096.6 function and identifying additional targets**
We have identified the first direct target of the *C. elegans* T-box transcription factor

TBX-2. To identify targets of TBX-2, we compared mRNA expression levels in wild-type and

tbx-2(bx59) mutant embryos using Affymetrix microarrays. To begin our analysis of the

microarray data, we compared our list of differentially regulated genes to existing data sets on

pharyngeal gene expression and phylogenetic foot-printing to identify promoters with consensus

T-box factor binding sites. We analyzed a subset of genes and identified D2096.6 as a direct

target of TBX-2. D2096.6 is expressed in pharyngeal cells and results from our microarray and

semi-qPCR indicate D2096.6 is directly repressed by TBX-2.

5.3.1   What is the role of D2096.6 in *C. elegans* development?
Our initial goal was to identify targets that are expressed in the pharynx and that may

have a role in pharyngeal development. We identified D2096.6 as a direct target of TBX-2 and

we believe TBX-2 represses D2096.6 by binding a variant T-box binding site in its promoter.

In wild type animals *D2096.6::gfp* is expressed in pharyngeal cells and based on our promoter analysis it seems likely that TBX-2 acts to restrict D2096.6 expression to these cells. When TBX-2 is reduced we see expanded expression in additional pharyngeal cells as well as outside the pharynx in the hypodermis, body wall muscles and gut. Based on this broad expression pattern it is unlikely that D2096.6 has a role specifically in ABa derived pharyngeal muscle development. However, overexpression of D2096.6 may reveal weaker phenotypes in pharyngeal development as well as phenotypes outside the pharynx. During early development the hypodermis has multiple functions including depositing basement membrane components, regulating cell fate specification of neighboring cells and guiding cell and axon migrations (Johnstone and Barry, 1996; Greenwald, 1997; Michaux et al., 2001). It will be interesting to see if overexpression of D2096.6 in the hypodermis results in any observable defects. In wild type worms *D2096.6::gfp* is expressed in pharyngeal muscle cells, therefore in addition to overexpression we would also ask if loss of D2096.6 results in pharyngeal defects. This experiment may not provide information on TBX-2 regulation of ABa pharyngeal muscle specifically but it could provide information on other aspects of pharyngeal muscle development.

5.3.2   How to identify additional targets of TBX-2

To understand how TBX-2 regulates ABa derived pharyngeal muscle cells we must identify direct targets of TBX-2 that have a role in the development of these cells. At this time we do not know the function of D2096.6; further analysis is needed to ask if this gene functions in pharyngeal development. However we have established that this gene is a direct target of TBX-2 and we can use information we have gained from this target to identify additional targets.

T-box proteins have been shown to bind sequences resembling a Brachyury consensus binding site. Our gel shift assays have shown that TBX-2 is capable of binding two sequences, a

Brachyury consensus binding site 1 and a more variant binding site 2. We believe TBX-2 represses D2096.6 by binding this variant T-box binding site 2 and currently it is not known if TBX-2 functions through a Brachyury consensus binding site in other targets.

Results from our gel shift assay show that mutation of 3 nucleotides in the binding site of probe 2 is still able to compete for binding to TBX-2, but to a lesser extent when compared to the wild type probe 2. This suggests that other nucleotides may be important for TBX-2 binding to probe 2. Additional mutations should be made to the binding site sequence in probe 2 and used in gel shifts to identify which nucleotides are important for TBX-2 to bind this sequence. Once these nucleotides are identified we can use this information to analyze binding sites in other potential targets of TBX-2 identified through analysis of the *tbx-2(bx59)* microarray.

Concurrent with this work, an analysis of differentially regulated genes from the *tbx-2(bx59)* microarray is being performed by Tom Ronan (Ronan and Okkema unpublished). Based on his analysis Tom has constructed multiple lists of potential TBX-2 targets based on various criteria such as the presence of binding sites within 1 kb of the transcription start site, conservation among nematode species and anti-correlation of expression levels. It will be interesting to see if any of these genes contain binding sites similar to the variant binding site 2 where we have shown TBX-2 is capable of functioning, or if a direct target is identified that shows TBX-2 functioning through a Brachyury consensus binding site.

APPENDICES

APPENDIX A: Pharyngeal expressed genes *egrh-1(tm1736), egrh-2(Y55F3AM.7), ZC328.2* and *ZK337.2(tm1134)* do not function in pharyngeal development

Introduction

      C. elegans pharyngeal muscle development involves *ceh-22*, an NK-2 family homeobox gene. *ceh-22* is the earliest known gene expressed in pharyngeal muscle, and its expression is initiated by a transcriptional enhancer termed the distal enhancer (Kuchenthal et al., 2001). Within the distal enhancer a short enhancer sequence (*de199*) was identified, and found to be active in multiple pharyngeal cell types (Vilimas et al., 2004).

      A yeast one-hybrid identified candidate proteins which were found to bind the de199 distal enhancer (Vilimas et al., 2004)(Vilimas and Okkema unpublished). Among those identified were the proteins (EGRH-1, ZC328.2, B0336.7, Y55F3AM.7 and ZK337.2). EGRH-1 was renamed EGRH-1 based on our characterization described in chapter IV. The expression patterns of these genes were examined using GFP reporter fusions. B0336.7, EGRH-1 and ZK337.2 are expressed inside and outside the pharynx and ZC328.2 is expressed in a specific subset of head neurons and Y55F3AM.7 expression was not determined (Vilimas and Okkema unpublished). In addition, four of these proteins (EGRH-1, Y55F3AM.7, ZC328.2, and ZK337.2) have similar zinc finger regions, and recognize similar sequences in *de199* (Vilimas unpublished).

      To ask if these genes shown to be expressed in the pharynx have a role in pharyngeal development, we obtained deletion mutants for 3 genes *egrh-1*, *B0336.7* and *ZK337.2* from the National BioResource Project (NBRP Japan). Observation of these mutant strains revealed no

abnormalities in pharyngeal development. Four of the zinc finger proteins EGRH-1, Y55F3AM.7 [EGRH-2], ZC328.2, and ZK337.2 have similar zinc finger regions, and recognize similar sequences in *de199.* Therefore, we used RNAi to perform combinations of double knockdowns to test redundancy between these proteins. Despite sequence similarities, we found no visible synthetic interactions between these proteins. Our characterization of these zinc-finger proteins identified no visible role in pharyngeal development.

Materials and Methods

Single worm PCR

Worm lysis and genotyping for the following genes was performed as described in chapter 2.

APPENDIX A (continued)

TABLE XIV: GENOTYPING *ZK337.2(TM706), ZK337.2(TM1134), B0336.7(TM1642), EGRH-1(TM1736)*

Gene	Primers for genotyping	PCR program
ZK337.2(tm706)	PO755/756/757	(1) 94°C, 2 min (2) 94°C, 30 sec (3) 54°C, 30 sec (4) 72°C, 30 sec (5) got to step 2 39x (6) 72°C, 5 min
ZK337.2(tm1134)	PO758/759/760	(1) 94°C, 2 min (2) 94°C, 30 sec (3) 56°C, 30 sec (4) 72°C, 30 sec (5) got to step 2 39x (6) 72°C, 5 min
B0336.7(tm1642)	PO764/765/766	(1) 94°C, 2 min (2) 94°C, 30 sec (3) 52.3°C, 30 sec (4) 72°C, 50 sec (5) got to step 2 39x (6) 72°C, 5 min
EGRH-1(tm1736)	PO761/762/763	(1) 94°C, 2 min (2) 94°C, 30 sec (3) 55.7°C, 30 sec (4) 72°C, 1 min (5) got to step 2 39x (6) 72°C, 5 min.

Lethality test

Lethality tests were conducted as follows for wild type and mutant hermaphrodites. Two L4 hermaphrodites were placed on each of 3 individual plates and placed at 25ºC. After 24 hours adult hermaphrodites were transferred to new plates and embryos were counted and transferred to separate plates for 3 days. After 24 hours un-hatched embryos were counted and after 48 hours adults were counted.

RNAi analyses

RNAi analyses were performed essentially as previously described (Fire et al., 1998). dsRNAs were synthesized by in vitro transcription (Ambion). For RNAi experiments *egrh-1* was PCR amplified from pOK165.12 which contains 2149 bp of *egrh-1* cDNA.

APPENDIX A (continued)

egrh-2(Y55F3AM.7) was PCR amplified from pOK173.19 which contains 1229 bp of

Y55F3AM.7 cDNA and ZC328.2 was PCR amplified from pOK172.13 which contains 1765 bp

of ZC328.2 cDNA. The cDNA for each gene was cloned into a pBluescript cloning vector and

PCR amplified with primers PO3 and PO340 which recognize the T3 and T7 primer binding

sites present in the Bluescript backbone. PCR program for *egrh-1*, *egrh-2* and ZC328.2 was the

following: (1) 94°C 45 sec., (2) 56°C 45 sec, (3) 72°C 2 min, (4) got to step 1 35x.

Results

Gene descriptions and mutant alleles

EGRH-1 has been previously described in chapter IV of this thesis.

ZK337.2 consists of 9 exons, covering 5038 bp of genomic sequence on Chromosome I.

This gene is predicted to encode a 543 amino acid protein which contains four C_2H_2 zinc fingers

(www.wormbase.org; Vilimas and Okkema unpublished). Based on previous analysis in our lab

a ZK337.2 translational GFP reporter fusion was found to be expressed in neurons and

pharyngeal muscles (Vilimas and Okkema unpublished). To characterize ZK337.2 function *in*

vivo, we obtained two deletion mutants from the National BioResource Project. *ZK337.2(tm706)*

contains a 503 bp deletion, removes exons 4 and 5 upstream of the zinc finger coding region.

The second allele *ZK337.2(tm1134),* contains 626 bp deletion +1 bp insertion and removes exons

7 and 8 which removes the zinc finger coding region.

B0336.7 consists of 6 exons, covering 2565 bp on chromosome III, and is predicted to

encode a 501 amino acid protein (www.wormbase.org; Vilimas and Okkema unpublished).

Based on previous analysis B0336.7 a translational GFP reporter fusion was found to be

expressed in the pharyngeal muscles and head neurons (Vilimas and Okkema unpublished). To

APPENDIX A (continued)

characterize B0336.7 function *in* vivo, we obtained deletion mutants from the National

BioResource Project. *B0336.7(tm1642)* 430 bp deletion removes the exons 1 and 2 of B0336.7

as well as part of an upstream gene *lgg-3*. *B0336*.7*(tm1642)* is a homozygous lethal strain and

displays partially penetrant maternal affect lethality. This strain can be maintained as viable

heterozygote. This gene appears to be in an operon with *lgg-3* (Vilimas and Okkema

unpublished; Blumenthal et. al., 2002). Since this deletion removes part of both genes, we do not

know which gene is responsible for the observed phenotype and did not continue with the

analysis of this allele.

Embryonic and larval lethality in the *egrh-1(tm1736), ZK337.2(tm706)* and *ZK337.2(tm1134)*

Pharyngeal defects can result in lethality due an inability to feed, therefore, we first

examined if any of these mutant alleles results in embryonic or larval lethality. Lethality tests

were conducted on *egrh-1(tm1736), ZK337.2(tm706), ZK337.2(tm1134)* and *B0336.7(tm1642)*

mutants and compared to wild type strain. L4 hermaphrodites were incubated at 25ºC. After 24

hours moms were transferred to new plates and embryos were counted and transferred to

separate plates for 3 days. After 24 hours unhatched embryos were counted and after 48hours

hatched embryos that have reached adult stage were counted.

We did not see a significant increase in the percent of dead embryos for either

ZK337.2(tm706) or *ZK337.2(tm1134)* mutants when compared to wild type worms (Table XV).

We did see a 25% increase in embryonic lethality in *egrh-1(tm1736)* mutants when compared to

wild type worms (Table XV). The embryonic lethality found *egrh-1(tm1736)* mutants may be

APPENDIX A (continued)

due to the loss of EGRH-1, however, it may also be an indirect effect from defective ovulation. As discussed in Chapter IV, we found that *egrh-1(tm1736)* mutants are defective in oocyte maturation and ovulation. We have observed endomitotic oocytes (Emo) in these mutants and it is known that an Emo phenotype can arise from a number of possible causes: defective dilation of the distal spermatheca; defective sheath cell contractions needed to progress the oocyte through the proximal gonad; defective sperm that are unable to provide the proper signals to arrested oocytes; and oocytes defective in meiotic cell cycle regulation (Iwasaki et al., 1996; Yin et al., 2004).

TABLE XV: PERCENT DEAD EMBRYOS AND ARRESTED LARVAE IN *egrh-1* AND ZK337.2 MUTANT STRAINS

Strain	% Dead embryos (n)	% Arrested larvae (n)
N2	2.7 (331)	3.9 (331)
egrh-1(tm1736)	36.4 (162)	10.5 (162)
ZK337.2(tm706)	3.4 (294)	3.7 (295)
ZK337.2(tm1134)	3.2 (284)	3.2 (284)

For each strain two L4 hermaphrodites were placed on each of 3 individual plates and placed at 25ºC. After 24 hours adult hermaphrodites were transferred to new plates and embryos were counted and transferred to separate plates for 3 days. After 24 hours un-hatched embryos were counted and after 48 hours adults were counted.

APPENDIX A (continued)

egrh-1(tm1736), egrh-2(Y55F3AM.7), ZC328.2 and *ZK337.2(tm1134)* do not function in
pharyngeal development

egrh-1(tm1736), egrh-2(Y55F3AM.7), ZC328.2 and *ZK337.2(tm1134)* have similar zinc

finger regions, and recognize similar sequences in the *ceh-22* distal enhancer (Vilimas and

Okkema unpublished). We hypothesized that these genes may have redundant roles in

pharyngeal development and therefore examined double knockdowns of these genes and looked

for defects in the pharynx. We examined the double mutant phenotypes of *egrh-1(tm1736),*

egrh-2(Y55F3AM.7), ZC328.2 and ZK337.2(tm1134) by utilizing deletions mutants in

combination with RNAi to see if there are synthetic interactions between these zinc-finger

proteins. We performed *egrh-1(tm1736), egrh-2(Y55F3AM.7)* and *ZC328.2(RNAi)* in

ZK337.2(tm1134) mutants, *egrh-2(Y55F3AM.7), ZC328.2* and *ZK337.2* RNAi in *egrh-1(tm1736)*

mutants and and *egrh-2(Y55F3AM.7)* and *ZC328.2(RNAi)* in N2 worms. We examined these

double knockdown worms for embryonic lethality (Table XVI) and pharyngeal defects and saw

no increase in either category when compared to the single mutants or single knockdown by

RNAi. Based on our phenotypic characterization and lethality tests, we do not believe that *egrh-*

1, *egrh-2*, ZK337.2 and ZC328.2 function in pharyngeal development

TABLE XVI: PERCENT DEAD EMBRYOS IN MUTANTS WITH DOUBLE KNOCKDOWNS OF PHARYNGEAL PROTEINS

Strain	% Dead Embryos (n)
egrh-1(RNAi); ZK337.2(tm1134)	1.2 (345)
egrh-1(RNAi); N2	2.5 (357)
egrh-2(RNAi); ZK337.2(tm1134)	3.1 (326)
egrh-2(RNAi); N2	1.7 (346)
egrh-2(RNAi); egrh-1(tm1736)	4 (327)
ZC328.2(RNAi); egrh-1(tm1736)	2.6 (266)
ZC328.2(RNAi); ZK337.2(tm1134)	1.8 (281)
ZK337.2(RNAi); egrh-1(tm1736)	5 (240)
egrh-2(RNAi); ZC328.2(RNAi); N2	8 (239)
uninjected N2	6 (121)

12 young adult hermaphrodites were injected per strain and placed at 20ºC overnight to recover. Moms were transferred to new plates twice a day for 2 days at 25ºC. Embryos were counted and transferred to new plates each time moms were transferred. After 24 hours unhatched embryos were counted from each transfer.

APPENDIX A (continued)

Conclusion

We set out to ask if EGRH-1, ZC328.2, B0336.7, Y55F3AM.7 and ZK337.2 played a role in pharyngeal development. Four of these proteins EGRH-1, Y55F3AM.7, ZC328.2, and ZK337.2 have similar zinc finger regions, and recognize similar sequences in the distal enhancer of the pharyngeal muscle specific gene *ceh-22* (Vilimas and Okkema unpublished). In addition GFP reporter fusions showed B0336.7, EGRH-1 and ZK337.2 are expressed inside and outside the pharynx (Vilimas and Okkema unpublished).

To ask if these genes shown to be expressed in the pharynx have a role in pharyngeal development, we obtained deletion mutants for 3 genes EGRH-1, B0336.7 and ZK337.2 from the National BioResource Project (NBRP Japan). Observation of these mutant strains revealed no abnormalities in pharyngeal development. Next, we used these mutants as well as RNAi to ask if the proteins that contain similar zinc finger regions, and bind similar sequence in the *ceh-22* enhancer function redundantly in pharyngeal development. Despite these similarities, we found no visible synthetic interactions between these proteins. Our characterization of these zinc-finger proteins identified no visible role in pharyngeal development.

APPENDIX B: T25E4.1 is an indirect target of TBX-2

T25E4.1 is ectopically expressed in *tbx-2(bx59)*

Our analysis of the T25E4.1 promoter region indicates this gene is downstream of TBX-2, but it is likely an indirect target. T25E4.1 was identified in our microarray and semi-qPCR as a gene that was over expressed in *tbx-2(bx59)* compared to wild type. We performed an analysis of the T25E4.1 promoter to ask if this gene is directly regulated by TBX-2. T25E4.1 is uncharacterized and encodes a predicted 231 amino acid protein. BLAST results identify T25E4.1 as a paralog of D2096.6, a TBX-2 target gene discussed in chapter VI (www.wormabase.org, WS224, 20 April 2011).

APPENDIX B (continued)
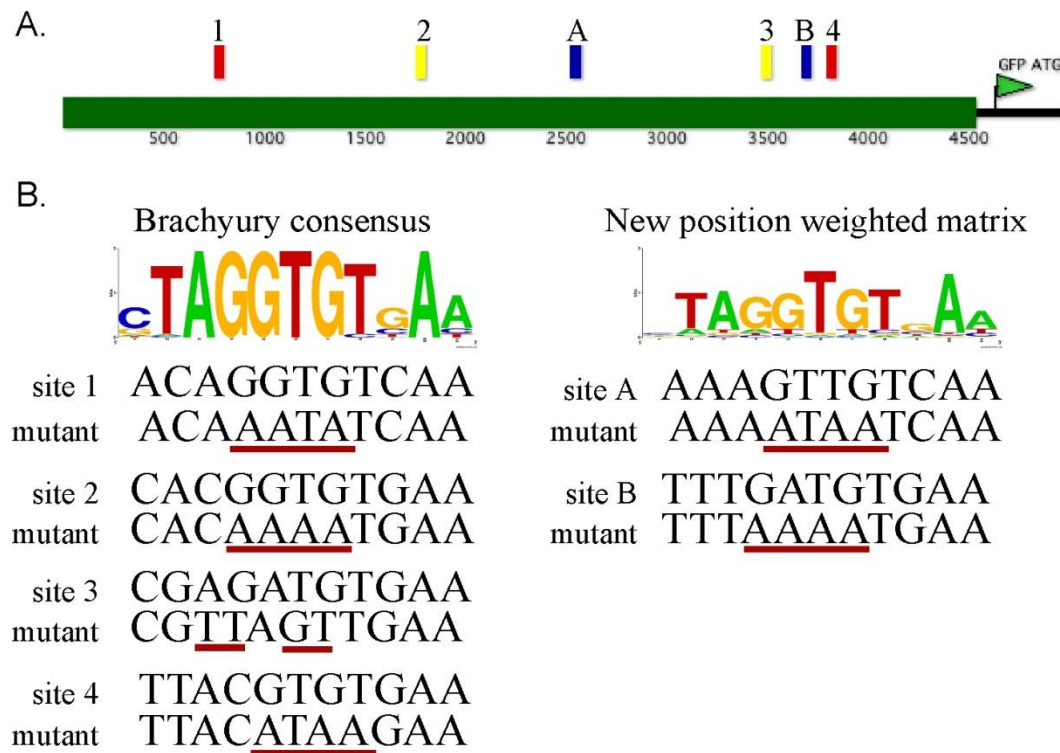
Figure 29: The T25E4.1 promoter contains 6 potential TBX-2 binding sites.

(A) Diagram of the T25E4.1 transcriptional GFP fusion containing 4.5 kb of upstream sequence cloned into pPD 95.77 GFP vector. Sites 1 and 4 shown in red are conserved. Site 1 (777 bp) is conserved in *C. remanei* and site 2 (3815 bp) is conserved in *C. briggsae* and *C. remanei.* Sites 2 (1774 bp) and 3 (3493 bp) are shown in yellow and these are not conserved in other nematode species. Sites A and B are newpwm sites shown in blue. Site A (2546 bp) is conserved in *C. briggsae* and site B (3694 bp) is not conserved in other nematode species. (B) Position weighted matrix of the Brachyury consensus sequence form JASPAR Database. Sequences of sites 1 through 4 indentified using the JASPAR and ConSite Database are shown along with the corresponding mutation made at each site underlined in red (left). The new position weighted matrix (newpwm) sequence constructed based on upstream sequences of genes reported in the *tbx-2(bx59)* microarray is shown. Sequences for sites A and B, identified by the newpwm sequence are shown along with the corresponding mutation made at each site underlined in red (right).

APPENDIX B (continued)

To ask if TBX-2 regulates T25E4.1 expression, we constructed a transcriptional GFP fusion, pOK258.01, containing 4.5 kb of sequence upstream of T25E4.1 (Figure 29A). The 4.5 kb promoter of T25E4.1 was PCR amplified from genomic DNA with primers PO979 and PO980 and cloned into pPD 95.77 at Sph-1 and BamH1 sites. We found *T25E4.1::gfp* is expressed in the pm6 pharyngeal muscle cell beginning at the pretzel stage of embryogenesis and continuing through the adult stage (Figure 30A,B). Our microarray and semi-qPCR data indicate that T25E4.1 is over expressed in *tbx-2(bx59)* compared to wild type. We used RNAi analysis to ask if reduction of TBX-2 function alters T25E4.1 expression. *tbx-2* dsRNA was synthesized in vitro as previously described and injected into a wild type strain expressing *T25E4.1::gfp* promoter fusion. Reduction of TBX-2 by RNAi resulted in a wild type spatial expression pattern. *T25E4.1::gfp* was expressed in pm6 in *tbx-2(RNAi)* worms, however the GFP in these worms appeared brighter. Based on this result we then crossed our promoter construct into *tbx-2(bx59)* mutants to further characterize T25E4.1 expression. *tbx-2(bx59); T25E4.1::gfp* (pOK0798) showed expression in pm6 as well as additional anterior pharyngeal cells in 25% of animals starting at pretzel stage. (Figure 30C-F). This change in expression varied from what was observed in *tbx-2(RNAi)* animals, however it is possible that the RNAi as not as effective at reducing TBX-2 function as our temperature sensitive mutant. The expanded expression pattern in *tbx-2(bx59)* mutants is consistent with our microarray and semi-qPCR results which indicate T25E4.1 is over expressed in *tbx-2(bx59)* compared to wild type.

APPENDIX B (continued)
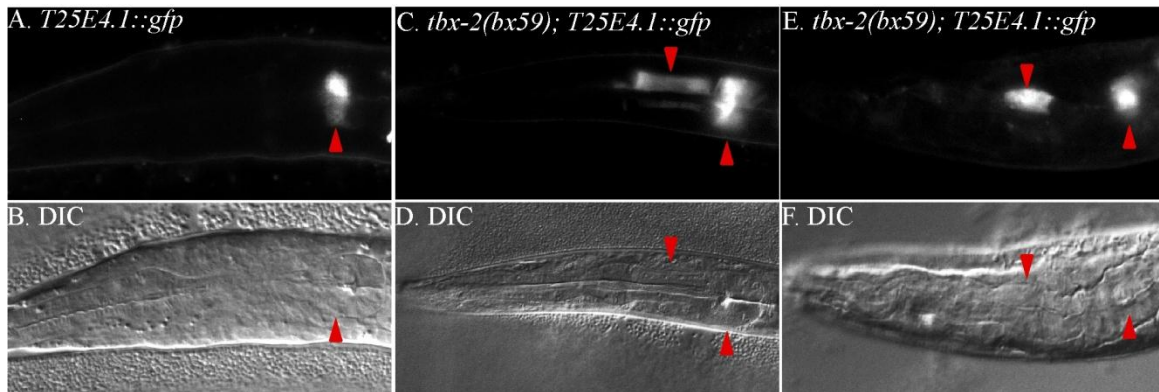
Figure 30: *T25E4.1::gfp* expression is increased in *tbx-2(bx59)*.

(A) Larva showing wild type *T25E4.1::gfp* expression (in the pharyngeal muscel cell pm6 in the terminal bulb (OK0806). (B) DIC image. (C) *T25E4.1::gfp* exrpression in *tbx-2(bx59)* larva (OK0798). Arrow heads point to expression in pm6 as well as expanded expression in the isthmus. (D) DIC image. (E) *T25E4.1::gfp* exrpression in *tbx-2(bx59)* larva (OK0798). Arrow heads point to expression in pm6 as well as expanded expression in the anterior bulb. (F) DIC image.

APPENDIX B (continued)

<u>T25E4.1 contains 6 potential TBX-2 binding sites</u>

Analysis of the T25E4.1 4.5 kb promoter sequence using JASPAR and ConSite databases revealed 4 potential T-box binding sites (Figure 29A). Phylogenetic analysis of the T25E4.1 promoter sequence showed site 1 to be conserved in *C. remanei* and site for 4 to be conserved in both *C. Briggsae* and *C. remanei*, while sites 2 and 3 are not conserved.

An analysis of all differentially regulated genes reported in our microarray was performed and a modified new position weighted matrix sequence for a potential TBX-2 binding site was identified (Tom Ronan unpublished). Two of these sites (newpwm A and newpwm B) were identified in the T25E4.1 promoter, for a total of 6 potential T-box binding sites (figure 1A). Phylogenetic analysis showed that site A conserved in *C. briggsae* and site B was not conserved in other nematode species.

In *tbx-2(bx59)* mutants, *T25E4.1::gfp* is over expressed compared to wild type which is consistent with our microarray and semi-quantitative PCR results. We hypothesized TBX-2 may directly repress T25E4.1 expression at one of the 6 candidate binding sites identified. Therefore, we individually mutated each binding site and asked if these mutations affected T25E4.1 expression. We first mutated the conserved Brachyury consensus sites 1 and 4 (Figure 29B). Mutating site 1 resulted in a loss of GFP at stages L1-adult (n= 40) (Figure 31A, B). Mutating site 2 also resulted in a loss of GFP at stages L1-adult (n=40) (Figure 31G, H). These results suggest that sites 1 and 4 function as activator sites. We next mutated sites 2 and 3 (Figure 29B). Mutating site 2 resulted in a loss of GFP at stages L1-adult (n=50) (Figure 31C, D), and mutating site 3 resulted in a loss of GFP at stages L1-adult (n=66) (Figure 31E, F). These results suggest

APPENDIX B (continued)

sites 2 and 3 also function as activator sites. Based on our microarray, semi-qPCR and

expression data in *tbx-2(bx59)* mutants we do not believe that TBX-2 activates T25E4.1

expression. We see a loss of expression when sites 1 through 4 are individually mutated, and

therefore do not believe that TBX-2 binds these sites to regulate T25E4.1.

APPENDIX B (continued)
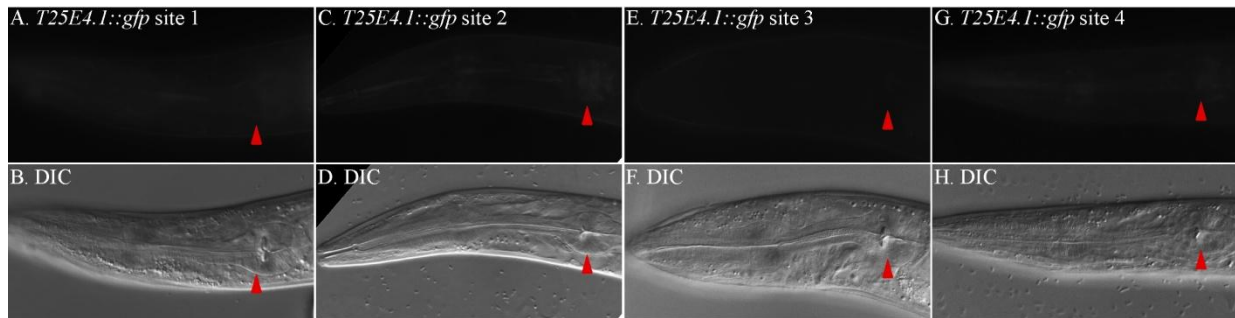
Figure 31: Mutations of Brachyury consensus sites result in a loss of *T25E4.1::gfp* expression.

(A) *T25E4.1::gfp* expression was lost in larva when the conserved Brachyury consensus site 1 was mutated (OK0804). (B) DIC image. (C) *T25E4.1::gfp* expression was lost in larva when the Brachyury consensus site 2 was mutated (OK0803). (D) DIC image. (E) *T25E4.1::gfp* expression was lost in larva when the Brachyury consensus site 3 was mutated. (F) DIC image. (G) *T25E4.1::gfp* expression was lost in larva when the conserved Brachyury consensus site 4 was mutated (OK0805). (H) DIC image.

APPENDIX B (continued)

A modified position weighted matrix sequence for potential TBX-2 binding was

constructed and we identified 2 sites, newpwm site A and B, which match this sequence in the

T25E4.1 promoter (Tom Ronan unpublished) (Figure 29A). To ask if TBX-2 regulates

expression at these sites we mutated each one individually (Figure 29B). Mutating newpwm site

A resulted in worms expressing GFP in the pm6 muscle from pretzel through adult stages

(pretzel 7% n=57; adult 31% n=13) in a wild type pattern. In addition, we saw 10% (n=281) of

worms L1 through adult stage that had faint GFP expression throughout the pharynx which was

not seen in either wild type or *tbx-2(bx59)* mutants (Figure 32A, B). The expanded expression

pattern we see when site A is mutated indicates that it functions as a site for repression. However

the over expression we see in this mutant is not similar to the over expression we see in *tbx-*

2(bx59) mutants, indicating TBX-2 is not the protein repressing T25E4.1 expression at this site.

Mutating newpwm site B showed expression in pm6 from pretzel through L4 larval stage

(pretzel 40% n=55; L4 6% n=49). However, we found expression in adults was completely lost

(n=30), suggesting site B may have a role in maintaining expression in adult worms (Figure 32C,

D). Mutation of either site did not result in the expanded pharyngeal expression pattern observed

in *tbx-2(bx59);* therefore, we do not believe that TBX-2 regulates T25E4.1 expression at these
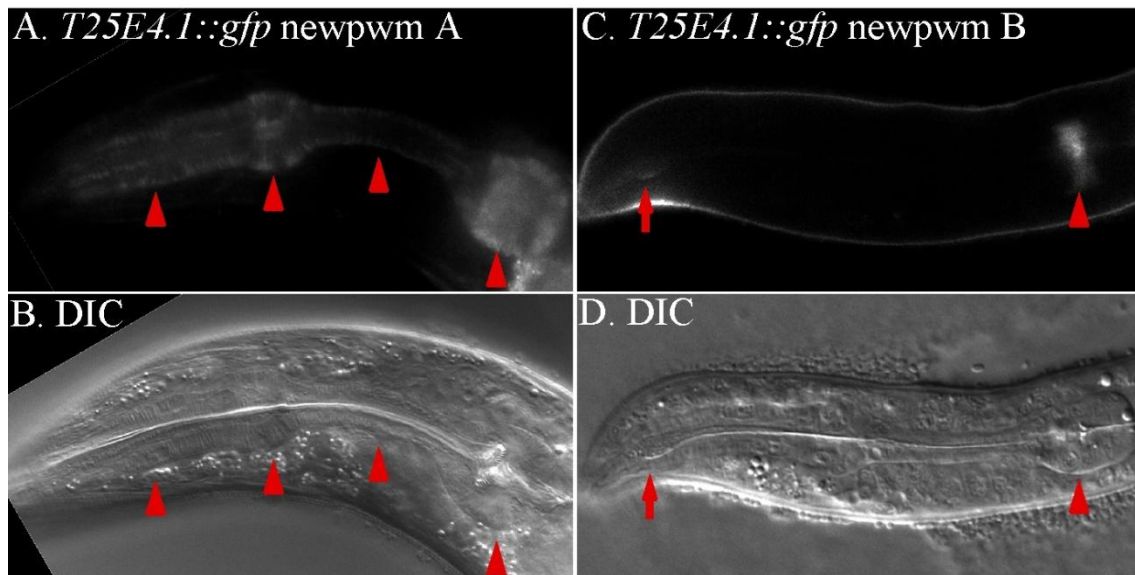
sites.

Figure 32: *T25E4.1::gfp* **expression is altered when newpwm sites A and B are mutated.**

(A) *T25E4.1::gfp* is expressed in the pharyngeal muscle pm6, as well as faintly throughout the anterior pharynx in an adult worm when site A is mutated (OK0808). Arrow heads indicate expanded expression. (B) DIC image. (C) L4 larva showing wild type *T25E4.1::gfp* expression in pm6 when site B is mutated (OK0811). Arrow head indicates expression in pm6 and arrow indicates auto-fluorescence in pharyngeal lumen. (D) DIC image.

APPENDIX B (continued)

Discussion

T25E4.1 is an indirect target of TBX-2

T25E4.1 is over expressed in *tbx-2(bx59)* according to our microarray, semi-qPCR and expression analysis, indicating T25E4.1 is downstream of TBX-2. Based on this data, we hypothesized TBX-2 directly or indirectly represses T25E4.1 expression. Our binding site analysis showed that mutating 5 out of 6 possible T-box sites results in decreased *T25E4.1::gfp* expression at different stages in development. Suggesting these sites are important for the activation of T25E4.1 expression. We see 10% of animals show faint expression throughout the worm when newpwm site A is mutated. However this does not resemble the over expression we see in *tbx-2(bx59)* mutants suggesting that some other protein is responsible for this expression pattern. Our analysis did not identify a site in the T25E4.1 promoter where TBX-2 directly represses T25E4.1 expression. We believe that T25E4.1 may be an indirect target of TBX-2. It is also possible that T25E4.1 is directly regulated by TBX-2 but we have not identified the site where TBX-2 binds.

T25E4.1 promoter contains multiple predicted T-box binding sites

Our analysis revealed that conserved binding sites 1 and 4 closely match the monomer sequence for the Brachyury consensus binding site. Based on the crystal structure of human TBX3, both sites contain all 4 nucleotides that are important for T-box proteins binding to their targets (Coll et. al., 2002). Therefore, it is possible that a T-box protein may bind sites 1 and 4 to activate expression. Binding sites 2 and 3 as well as the modified position weighted matrix sequences also look like good T-box binding sites when compared to the Brachyury consensus

APPENDIX B (continued)

site, differing by only one or two nucleotides in the monomer sequence. Site 2 contains all 4

nucleotides important for binding based on human TBX3 crystal structure. Site 3 and both new

postion weighted matrix sites contain 3 out of 4 nucleotides important for binding. Therefore, it

is possible that one or multiple T-box proteins are binding these sites and regulating expression.

It is also possible that still unknown transcription factors may be regulating T25E4.1 expression

at these or additional regulatory sites.

APPENDIX C: FMRFamide (Phe-Met-Arg-Phe-NH2)-like neuropeptide (FaRPs) family members analyzed are not regulated by TBX-2

FMRFamide (Phe-Met-Arg-Phe-NH2)-like neuropeptide (FaRPs) family members identified in our microarray

Utilizing the DAVID Bioinformatics database we identified a group of 13 genes from the FMRFamide-like neuropeptide family as overrepresented in our microarray data. We hypothesized that these genes may be targets of TBX-2. DAVID Bioinformatics database allows you to identify genes of interest in large datasets by grouping enriched functional-related gene groups and annotation clustering. To ask if we could identify related gene groups among our microarray data, we analyzed the full list of differentially expressed genes from the microarray using DAVID Bioinformatics Resources 2008. This database identified the FMRFamide-like peptide (FLP) protein family as an overrepresented gene group in our microarray data, however no other groups were identified using this method.

Several of these genes in the FMRFamide (Phe-Met-Arg-Phe-NH2)-like neuropeptides (FaRPs) family are known to be expressed in the pharynx and we found that several contained Brachyury consensus binding sites in the 5' intergenic region. The family of FMRFamide (Phe-Met-Arg-Phe-NH2)-like neuropeptides (FaRPs) all share an RFamide sequence at their C-termini and have been shown to have a variety of functions in the central and peripheral nervous systems. In the nematode *C. elegans*, FMRFamide-like peptides are expressed in multiple neurons, including motor, sensory, and interneurons which have been shown to affect movement, feeding, defecation, and reproduction respectively. In *C. elegans* there are 23 genes that have been identified and encode FaRPs. The gene family has been named *flp* (for *F*MRFamide-*l*ike *p*eptides) and numbered *flp-1* through *flp-23*. (Li et al., 1999; Kim and Li, 2004). The complete

expression pattern of TBX-2 is not known, however, it is known that TBX-2 is expressed in head neurons in adult worms. Therefore it is possible that TBX-2 may target neurons at earlier stages in development. It has been reported that most of the *flp* genes are found to be expressed throughout the life cycle of *C. elegans* (Li et al., 1999). We obtained GFP expressing strains for 4 *flp* genes identified in our microarray *flp-1, flp-2, flp-15* and *flp-11*; and asked if a reduction in TBX-2 function affects GFP expression in these lines (provided by Chris Li at The City College of New York).

flp-1::gfp expression is not affected by *tbx-2(RNAi)*

flp-1 was identified in our microarray as well as semi-qPCR experiments as over expressed in *tbx-2(bx59)* when compared to wild type . *flp-1* is expressed in the AVK neuron (Kim and Li, 2004). We obtained the integrated *Pflp-1::gfp* (NY2097 ynIs97) expressing strain and verified expression in the AVK neuron. We found *flp-1* expression beginning at embryonic stage 1 ½ fold in a single cell and continuing into the adult stage in 100% (n=17) of worms (Figure 33). To ask if TBX-2 regulates the expression of *flp-1* we performed *tbx-2(RNAi)* in the *flp-1::gfp* expressing strain. *tbx-2* dsRNA was synthesized in vitro as previously described and injected into adult hermaphrodites. We examined all stages of development and found the *flp-1* wild type expression pattern in 100% (n=23) of worms. ConSite database was used to identify Brachyury consensus binding sites in the upstream region of *flp-1*. Using ConSite we identified 3 Brachyury consensus sites within 2 kb of *flp-1* upstream sequence (Table VXII). Based on our microarray results, *flp-1* is over expressed in *tbx-2(bx59)* mutants compared to wild type indicating that TBX-2 represses this gene. Therefore, we would expect an increase

in *flp-1* expression when TBX-2 is reduced by RNAi. Knockdown of TBX-2 protein did not

affect *flp-1::gfp* expression.

APPENDIX C (continued)

Figure 33: *flp-1::gfp* **expression is not altered by** *tbx-2(RNAi).*

(A-D) *flp-1::gfp* expression at the 1 ½ fold stage in the head and was identical between wild type and *tbx-2(RNAi);flp-1::gfp* worms. (E-H) *flp-1::gfp* expression in pretzel stage embryos in the AVK interneuron located next to the terminal bulb of the pharynx.. Arrow heads point to expression in the AVK neuron.

APPENDIX C (continued)

TABLE XVII: BRACHYURY BINDING SITES IDENTIFIED THOROUGH CONSITE DATABASE

Gene	Binding site sequence[1]	Conservation[2]
flp-1	GTAGGTTTCAT	no
	CTCAGTGTGAT	no
	GTAGTTGTGCA	*C. briggsae*
flp-2	CTAGATTTGAA	no
flp-11	ATGGGTGTGAA	no
flp-15	GTAGGTTTCAA	no

[1]Binding sites were identified using the Brachyury consensus sequence found in the ConSite database. The upstream sequence analyzed was based on the GFP constructs used for each gene. For *flp* genes 2 kb of upstream sequence was analyzed for Brachyury binding sites.
[2] Upstream sequences from *C. briggsae* and *C. remanei* was analyzed to ask if these sites were conserved in either species.

APPENDIX C (continued)

<u>*flp-2::gfp* expression is not affected by *tbx-2(RNAi)*</u>

flp-2 was identified in our microarray as well as semi-qPCR experiments as over expressed in *tbx-2(bx59)* compared to wild type embryos. *flp-2::gfp* is expressed in several neurons as well as head muscles (AIA, RID, PVW, I5, MC,ASI, M4) (Kim and Li, 2004). We obtained an integrated *Pflp-2::gfp* (NY2057 ynIs57) expressing strain (a gift from Chris Li at The City College of New York). We found *flp-2* expression in the AIA interneuron beginning at embryonic stage 1 ½ fold and continuing into larval stages in 89% (n=38) of worms (Figure 34). Using ConSite database we identified 1 Brachyury consensus site in 2kb of upstream sequence (Table XVII). Based on our microarray results, *flp-2* is over expressed in *tbx-2(bx59)* mutants compared to wild type indicating that TBX-2 represses this gene. Therefore, we would expect an increase in *flp-2* expression when TBX-2 is reduced by RNAi. We performed *tbx-2(RNAi)* in the *flp-2::gfp* expressing strain. We examined all stages of development and found the *flp-2* wild type expression pattern in 100% (n=18) of worms.

APPENDIX C (continued)
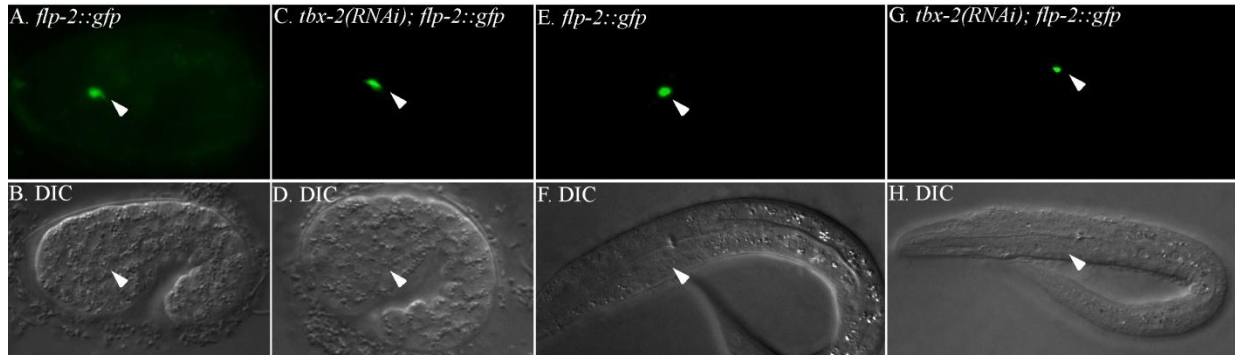
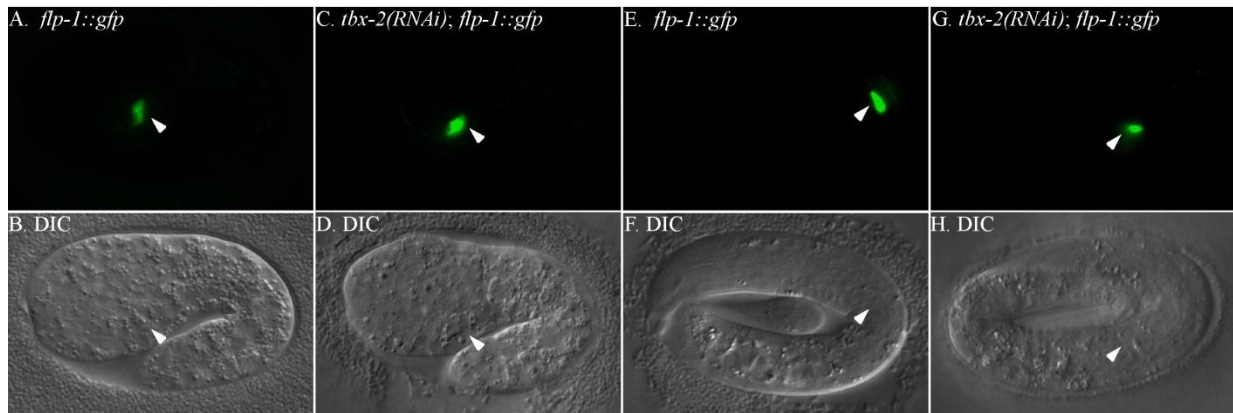**Figure 34:** *flp-2::gfp* **expression is not altered by** *tbx-2(RNAi)***.**

(A-D) *flp-2::gfp* expression at the 1 ½ fold stage in the head (E-H) *flp-2::gfp* expression in larva in the AIA interneuron located next to the terminal bulb of the pharynx. Arrow heads point to expression in the AIA interneuron.

APPENDIX C (continued)

flp-11::gfp expression is not affected by *tbx-2(RNAi)*

   flp-11 was identified our microarray as well as semi-quantitative PCR experiments as

over expressed in *tbx-2(bx59)* when compared to wild type . We obtained the integrated *flp-*

11::gfp strain ynIs40*[Pflp-11::gfp]* from the Caenorhabditis Genetics Center. *flp-11* is expressed

in neurons, head muscles, sheath cells and uterine cells. Expression is observed in the AVA

neuron which regulates feeding behavior (Coates and de Bono, 2002; Kim and Li, 2004). We

observed *flp-11* expression in the head within and outside the pharyngeal primordium beginning

in embryos and in the AVA neuron in larva (Figure 35). Using ConSite database we identified 1

Brachyury consensus site in 2kb of upstream sequence (Table XVII). Based on our microarray

results, *flp-11* is over expressed in *tbx-2(bx59)* mutants compared to wild type indicating that

TBX-2 represses this gene. Therefore, we would expect an increase in *flp-11* expression when

TBX-2 is reduced by RNAi. We performed *tbx-2(RNAi)* in the *flp-11::gfp* expressing strain. *tbx-*

2 dsRNA was synthesized in vitro and injected into adult hermaphrodites. We examined all

stages of development and found the *flp-11* wild type expression pattern in 100% (n=110) of

worms. We also examined expression in the *tbx-2(bx59)* background and saw no change

compared to the wild type expression pattern.

APPENDIX C (continued)
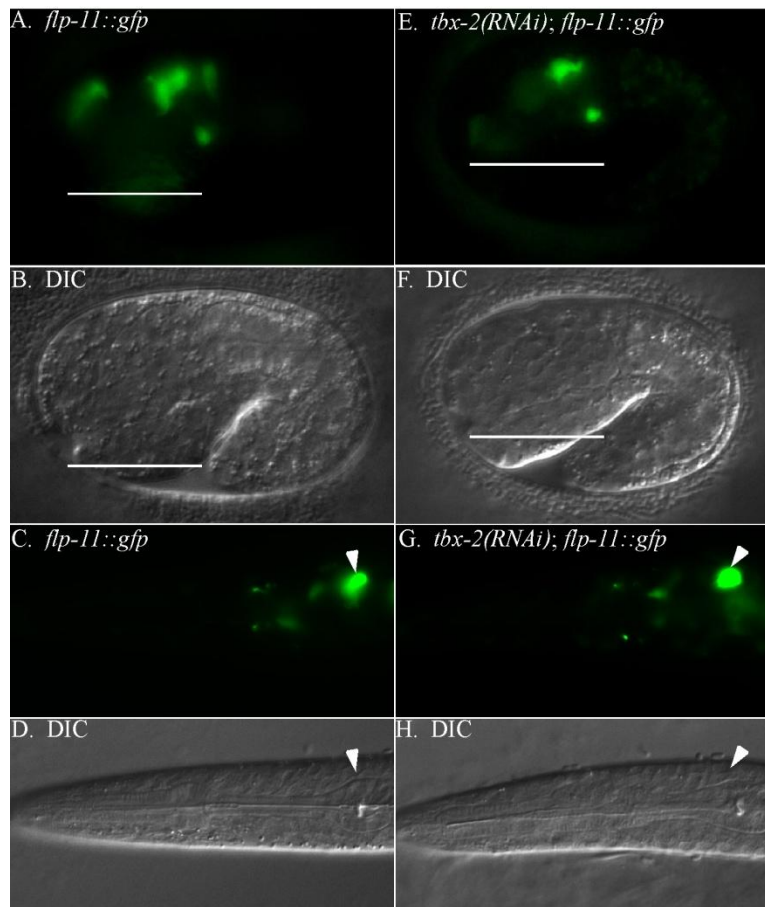
Figure 35: *flp-11::gfp* expression is not altered by *tbx-2(RNAi)*.

(A-B, E-F) In 1 ½ fold embryos *flp-11::gfp* is expressed in the head within and outside the pharyngeal primordium. (C-D, G-H) In larva *flp-11::gfp* is expressed in the AVA neuron. Bracket spans expression in the head and the arrow heads point to expression in the AVA neuron.

APPENDIX C (continued)

flp-15:gfp expression is not affected by *tbx-2(RNAi)*

      flp-15 was identified in our microarray and semi-quantitative PCR experiments as over expressed in *tbx-2(bx59)* when compared to wild type . We obtained the integrated *flp-15::gfp* strain ynIs45[*Pflp-15::gfp*] from the Caenorhabditis Genetics Center. *flp-15* is expressed in the pharyngeal neuron I2 as well as additional neurons in the head (Kim and Li, 2004). We observed *flp-15* expression in the anterior head cells of comma stage embryos and in the I2 and I5 pharyngeal interneurons in larva (Figure 36). The pharyngeal neuron I2 is locates in the anterior bulb of the pharynx. The pharyngeal neuron I5 is located in the terminal bulb of the pharynx and participates in regulating pharyngeal relaxation (Avery, 1993). Using ConSite database we identified 1 Brachyury consensus site in 2kb of upstream sequence (Table XVII). Based on our microarray results, *flp-15* is over expressed in *tbx-2(bx59)* mutants compared to wild type indicating that TBX-2 represses this gene. Therefore, we would expect an increase in *flp-15* expression when TBX-2 is reduced by RNAi. We performed *tbx-2(RNAi)* in the *flp-15::gfp* expressing strain. We examined all stages of development and found the *flp-15* wild type expression pattern in 100% (n=44) of worms.

APPENDIX C (continued)
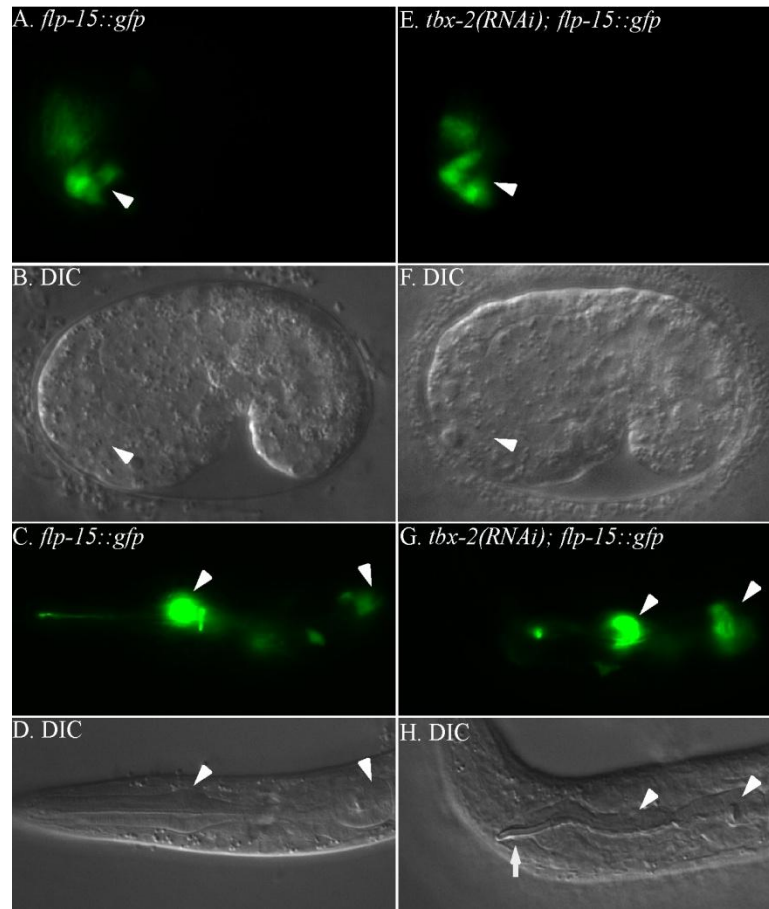
Figure 36: *flp-15::gfp* **expression is not altered by** *tbx-2(RNAi).*

(A-B, E-F) In comma stage embryos *flp-15::gfp* is expressed in anterior head cells. (C-D, G-H) In larva *flp-15::gfp* is expressed in pharyngeal neurons I2 and I5.. (H) The arrow points to the unattached buccal cavity

APPENDIX C (continued)

Discussion

Knockdown of TBX-2 protein did not affect *flp-1::gfp*, *flp-2::gfp, flp-11::gfp*, and *flp-15::gfp* expression. Based on our microarray results as well as semi-qPCR, mRNA levels of these *flp* family genes are over expressed in *tbx-2(bx59)* compared to wild type; suggesting that these genes are downstream of TBX-2. It is possible that the promoter GFP fusions we obtained do not contain the complete regulatory sequences that are needed for TBX-2 regulation and therefore *tbx-2(RNAi)* had no affect on expression. Transcriptional reporter genes lack introns and 3'UTRs which may contain elements important for regulation (Okkema et al., 1993). Therefore, translational reporter fusions or *in situ* hybridization experiments may need to be performed to observe changes in the expression.

APPENDIX D: *pqn-71* and *myo-5* are not regulated by TBX-2

pqn-71::gfp expression is not affected by *tbx-2(RNAi)*

       pqn-71 was identified our microarray as well as semi-quantitative PCR experiments as over expressed in *tbx-2(bx59)* when compared to wild type. *pqn-71 - (Prion-like-(Q/N-rich)-domain-bearing protein)* is an uncharacterized protein (http://www.wormbase.org). We constructed a *pqn-71::gfp* fusion consisting of 843 bp of sequence spanning a large intron between the first and second exon cloned into pPD95.75 GFP vector (pOK252.09) . Using ConSite database we identified 2 Brachyury consensus sites within this 843 bp region (Table XVIII). We observed *pqn-71* expression in the pharynx as well as the hypodermis (Figure 37). Expression initiated at the 2 fold stage in embryos and remained through adult stages. Based on our microarray results, *pqn-71* is over expressed in *tbx-2(bx59)* mutants compared to wild type indicating that TBX-2 represses this gene. Therefore, we would expect an increase in *pqn-71* expression when TBX-2 is reduced by RNAi. *tbx-2(RNAi)* in animals expressing *pqn-71::gfp* resulted in no change in expression 100% (n=101). We also examined expression in the *tbx-2(bx59)* background and saw no change compared to the wild type expression pattern (n=117).

APPENDIX D (continued)

TABLE XVIII: BRACHYURY BINDING SITES FOR *pqn-71* AND *myo-5* PROMOTERS
IDENTIFIED THOROUGH CONSITE DATABASE

Gene	Binding site sequence[1]	Conservation[2]
pqn-71	GTTGGTGTTAG	no
	ATTGGTGTGGA	no
myo-5	TTAGGTGTCGA	no

[1]Binding sites were identified using the Brachyury consensus sequence found in the ConSite
database. The upstream sequence analyzed was based on the GFP constructs used for each gene.
For *pqn-71* 834 bp of upstream sequence (pOK252.09) and for *myo-5* 2 kb of upstream sequence
was analyzed for Brachyury binding sites.
[2] Upstream sequences from *C. briggsae* and *C. remanei* was analyzed to ask if these sites were
conserved in either species.

APPENDIX D (continued)
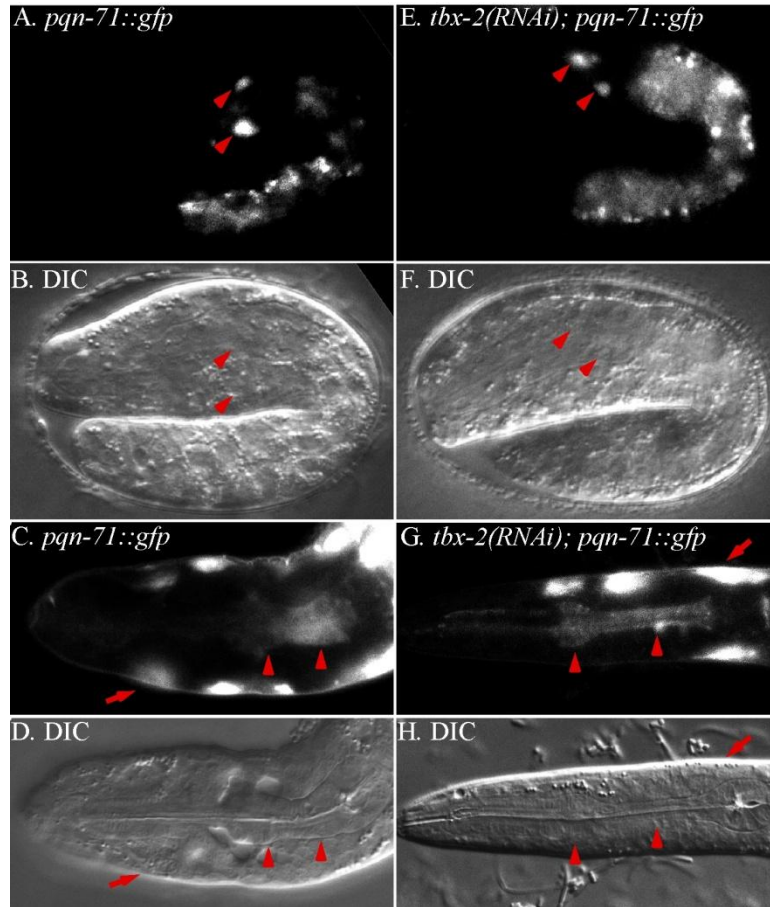
Figure 37: *pqn-71::gfp* **is expression is not altered by** *tbx-2(RNAi).*

(A-B, E-F) In 2-fold stage embryos *pqn-71::gfp* (pOK252.09) is expressed in 2 cells located in pharyngeal primordium as well as outside (arrow heads). Additional expression is due to autofluorescence from gut granules. (C-D, G-H) In larva *pqn-71::gfp* is expressed in the anterior bulb (arrow head) and isthmus (arrow head) of the pharynx and in the hypodermis (arrow).. The arrows point to expression in the hypodermis and the arrow heads point to expression in the pharynx.

APPENDIX D (continued)

<u>*myo-5::gfp* expression is not affected by *tbx-2(RNAi)*</u>

In addition to genes that were over expressed in we also examined one gene that was under expressed *tbx-2(bx59)* compared to wild type. *myo-5* was identified in our microarray as well as semi-quantitative PCR experiments as under expressed in *tbx-2(bx59)* when compared to wild type. *myo-5* is designated as part of the MYOsin heavy chain structural gene family and is expressed in pharyngeal muscle (www.wormbase.org). We obtained the strain BC11598 [*dpy-5(e907); sex11598*] from the Caenorhabditis Genetics Center which contains 2 kb of upstream sequence fused to GFP. Expression initiated in one to two cells in the head at the comma stage in embryos and in pretzel and larva *myo-5::gfp* was expressed in pharyngeal muscle (Figure 38). Using ConSite database we identified 1 Brachyury consensus sites within this 2 kb region (Table XVIII). Based on our microarray results, *myo-5* is under expressed in *tbx-2(bx59)* mutants compared to wild type indicating that TBX-2 activates this gene. Therefore, we would expect a decrease in *myo-5* expression when TBX-2 is reduced by RNAi. We performed *tbx-2(RNAi)* in the *myo-5::GFP* expressing strain. We examined all stages of development and found the *myo-5* wild type expression pattern in 100% (n=228) of worms.

APPENDIX D (continued)
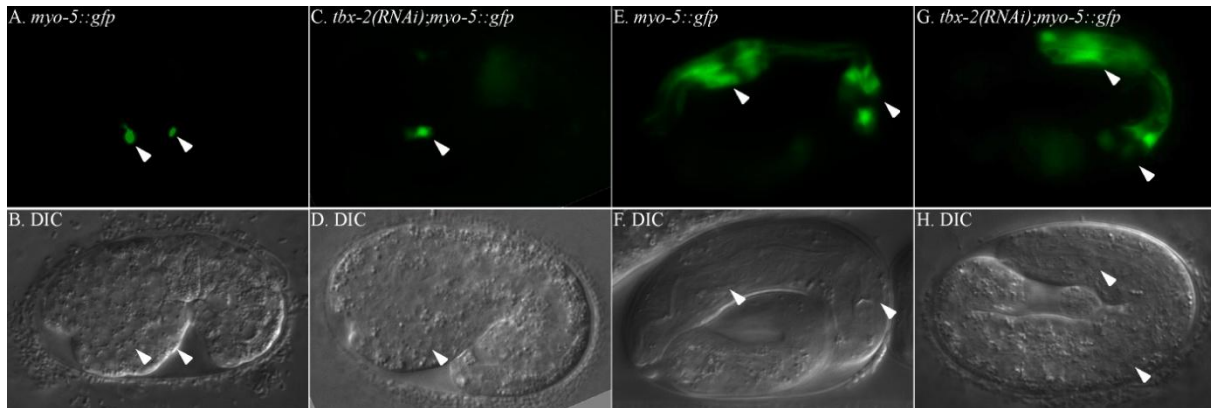
Figure 38: *myo-5::gfp* **expression is not altered by** *tbx-2(RNAi).*

(A-D) In 1 ½ fold embryos *myo-5::gfp* is expressed in unidentified cells in the head. (E-H) In pretzel stage embryos *myo-5::gfp* is expressed in pharyngeal muscle cells.. The arrow heads point to expression in the head and pharynx.

APPENDIX D (continued)

Discussion

Knockdown of TBX-2 protein did not affect *pqn-71::gfp* or *myo-5::gfp* expression. Based on our microarray results as well as semi-qPCR, mRNA levels of *pqn-71* is over expressed in *tbx-2(bx59)*; and *myo-5* is under expressed in *tbx-2(bx59)* compared to wild type; suggesting that these genes are downstream of TBX-2. It is possible that the promoter GFP fusions we obtained do not contain the complete regulatory sequences that are needed for TBX-2 regulation and therefore *tbx-2(RNAi)* had no affect on expression. Transcriptional reporter genes lack introns and 3'UTRs which may contain elements important for regulation (Okkema et al., 1993). Therefore, translational reporter fusions or *in situ* hybridization experiments may need to be performed to observe changes in the expression.

APPENDIX E: Oligonucleotides

TABLE XIX: OLIGONUCLEOTIDES

Name	Sequence 5'-3'	Description
PO3	AGCGGATAACAATTTCACACAGGA	M13 reverse sequencing primer
PO290	CTGCAAGGCGATTAAGTTGG	sequencing primer for pHISi
PO308	TCGAGGTGCCGTAAAGCACTAA	reverse primer that binds to the f1 origin
PO340	CGATTAAGTTGGGTAACGCC	sequencing primer for Bluescript inserts
PO566	TAATACGACTCACTATAGGGAGACGTACAATCACTGGGGACG	primer for T7 RNA polymerase promoter
PO567	AATTAACCCTCACTAAAGGGAGAAGCGTGAGAAAAAGCGAC	primer for T3 RNA polymerase promoter
PO604	TTTTGTGCGAACATCGACC	primer for genotyping *tbx-2(ok529)*
PO605	ACACACACACCTACAGTAATGCG	primer for genotyping *tbx-2(ok529)*
PO606	TCAGGCGAGCATTGGATG	primer for PCR genotyping *tbx-2(ok529)*

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO646	TGCATCACCTTCACCCTC	sequencing primer for promoter inserts in pOK173.06, 173.07, and 173.09 gfp vectors
PO695	GGAATTCATGGCACTTCACGAGCCAC	primer for amplifying *egrh-1* orf.
PO696	TTCTGCAGTTAATCCGAAGAGCTAGACGGC	primer for amplifying *egrh-1* orf.
PO755	CTCAATCAACGAAACAACACG	primer for genotyping *ZK337.2(tm706)*
PO756	TGTTGTTGAGACCTTCGTCG	primer for genotyping *ZK337.2(tm706)*
PO757	TGTCATCTTTGCTTCTGGGG	primer for genotyping *ZK337.2(tm706)*
PO758	GACCAATTCCTTTGTGAGCAGC	primer for genotyping *ZK337.2(tm1134)*
PO759	CGATCATGTACCGCAAACGC	primer for genotyping *ZK337.2(tm1134)*
PO760	GGTGGTAGTGGTGAAGTAGTTGCAC	primer for genotyping *ZK337.2(tm1134)*
PO761	GGCACCATCATCGTCATACAAGG	primer for genotyping *EGRH-1(tm1736)*
PO762	GGCAGCAGATAAGCCTGAAAATTC	primer for genotyping *egrh-1(tm1736)*
PO763	GCTGTAGTCATCCATTGGCTCGG	primer for genotyping *egrh-1(tm1736)*
PO764	TCGTGAGCATTGTCTACTTGAACG	primer for genotyping *B0337.6(tm1642)*
PO765	TGGATGATGATGGGGAAGAAGC	primer for genotyping *B0337.6(tm1642)*

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO766	ACATTTGCCCCATCTCCTGAC	primer for genotyping *B0337.6(tm1642)*
PO816	GTGATTCCCCTTTTTTTGGTTGTAGG	primer for amplifying *egrh-1* genomic coding sequence
PO817	CCAATTTCCCAATTTTTTTGCGTCG	primer for amplifying *egrh-1* genomic coding sequence
PO818	ATGGGCCCTTTCTTATTGCTGCGGCG	primer for amplifying *egrh-1* genomic coding sequence
PO878	ATCTTGCCCATCTGGAACAC	primer for *T10E10.4* RT-PCR
PO879	ACGGGCTGTTACTCCATTTG	primer for *T10E10.4* RT-PCR
PO880	CACAATGTCCACCACAATCC	primer for *D2096.6* RT-PCR
PO881	ACAACTTTTGGCGTTGAACC	primer for *D2096.6* RT-PCR
PO882	TTGCAAAACTTGTGAGCGAC	primer for *F58G4.1 (myo-5)* RT-PCR
PO883	GGAAGAGCTTCAACAGGCAC	primer for *F58G4.1 (myo-5)* RT-PCR
PO884	ATCTTTTGCAGTGCCACCTC	primer for *flp-1c* RT-PCR
PO885	GGTTATTACTCCTTGTGGCAGC	primer for *flp-1c* RT-PCR

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO886	GGAAGTCGTAATCTGGCAGC	primer for *flp-2b* RT-PCR
PO887	CCAGCTGTCAACGACAACAC	primer for *flp-2b* RT-PCR
PO888	ATGCACTATTCCTTGTCGCC	primer for *flp-9* RT-PCR
PO889	TACCAGCGGGTATCCTGAAC	primer for *flp-9* RT-PCR
PO890	CTTACCAAATGGTTGTGGGG	primer for *flp-11c* RT-PCR
PO891	TCAATTCTCTGCATTGGCAC	primer for *flp-11c* RT-PCR
PO892	AGATGGTCCACGTCGTTTTC	primer for *flp-15* RT-PCR
PO893	CACGCTTATTCGAGTTGCAG	primer for *flp-15* RT-PCR
PO894	CTGCGCTCTCTTTCCAAATC	primer for *flp-16* RT-PCR
PO895	CGATTGTTGCCTTCTTCCTC	primer for *flp-16* RT-PCR
PO896	ACCGAACCTCACTTGATTCG	primer for *flp-19* RT-PCR
PO897	GCTTTTCCTGTTAATTGCCG	primer for *flp-19* RT-PCR
PO898	CATAGGTGGACCAGATTGGG	primer for *T25E4.1* RT-PCR
PO899	CCACTGAAGAAGGCAAAAGC	primer for *T25E4.1* RT-PCR
PO900	GTGACGAGCTCAGAACATGC	primer for *phg-1* RT-PCR

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO901	ACACGAGCACAGTGAGCAAG	primer for *phg-1* RT-PCR
PO902	TGTTTGCAGTTGCTTTGGC	primer for *pqn-71* RT-PCR
PO903	CCTTGCTGATTGTACCCTCC	primer for *pqn-71* RT-PCR
PO906	TTCCAAGCGCCGCTGCGCATTGTCTC	primer for *ama-1* RT-PCR
PO907	CAGAATTTCCAGCACTCGAGGAGCGGA	primer for *ama-1* RT-PCR
PO910	GGAAACAGTTATGTTTGGTATA	primer D* for PCR gene fusions in pPD95.75
PO924	TGGCTGATCGCTTACAACAG	*T25E4.1* primer A for PCR gene fusions
PO925	AGTCGACCTGCAGGCATGCAAGCTTGCGAAGAACA GAGGGTTACG	*T25E4.1* primer B for PCR gene fusions
PO926	ATTCATTCGGTAATTTCTTCC	*T25E4.1* primer A' for PCR gene fusions
PO931	AGTTTGACACCGATTTTCTCG	primer for amplifying *tbx-2(bx59)* mutation with PO932
PO932	GTGATGATGGATCTTGTTCCG	primer for amplifying *tbx-2(bx59)* mutation with PO931
PO952	CCCCCCATCGCTCACTTTTTCTACAATTAATGTGTA CCCGAGGTGGGT	primer for mutating Brachyury site 1 in *D2096.6::gfp* pOK253.05

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO953	ACCCACCTCGGGTACACATTAATTGTAGAAAAGTGAGCGATGGGGGG	primer for mutating Brachyury site 1 in *D2096.6::gfp* pOK253.05
PO979	CAGCATGCTGGCTGATCGCTTACAACAG	PCR primer for T25E4.1 promoter with SphI site
PO980	CAGGATCCTGCGAAGAACAGAGGGTTACG	PCR primer for T25E4.1 promoter with BamHI linker
PO1002	GGAATGCGGGAGGAATATCAAATGTTGAATTTAAAATATCAAAATGTGAAACTTTTTAAATGTATAGATAAAATGG	primer for mutating t-box site (4) 2 in *D2096.6::GFP::his2b*
PO1003	CCATTTTATCTATACATTTAAAAAGTTTCACATTTTGATATTTAAATTCAACATTTGATATTCCTCCCGCATTCC	primer for mutating t-box site (4) 2 in *D2096.6::GFP::his2b*
PO1004	GGGCACGGGAGGATGAAAAGTTACATAAGAAGTGTCAAATTCAGGTTTTTTG	primer for mutating site 4 (closest to ATG) for *T25E4.1* in pPD95.77 GTGT to ATAA
PO1005	CAAAAAACCTGAATTTGACACTTCTTATGTAACTTTTCATCCTCCCGTGCCC	primer for mutating site 4 (closest to ATG) for *T25E4.1* in pPD95.77 GTGT to ATAA
PO1006	TATCAGTAAGCACCGACTCTTCACAAATATCAAGTTGCTCTTTTCTCCG	primer for mutating site 1 (farthest from ATG) for *T25E4.1* in pPD95.77 GGTG to AATA

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO1007	CGGAGAAAAGAGCAACTTGATATTTGTGAAGAGTC GGTGCTTACTGATA	primer for mutating site 1 (farthest from ATG) for *T25E4.1* in pPD95.77 GGTG to AATA
PO1025	GAAGAAGGCGATGCTTATGTG	primer to sequence *T25E4.1* site 1 mutation
PO1026	CCTGCTAACAAGACAAGGCAG	primer to sequence *T25E4.1* site 1 mutation
PO1027	GTCATCACAATCTACGGCTCC	primer to sequence *T25E4.1* site 4 mutation
PO1028	GCGATGTCGTCAGTATTTTCC	primer to sequence *T25E4.1* site 4 mutation
PO1036	ACGGTCAAAATACTATCATCTTCGCCTAAAAATTT AAAGGTGCAAAGAGTGTATCGCTTC	mutagenesis primer for *T25E4.1* newpwm site at 3320 bp in pOK258.01
PO1037	GAAGCGATACACTCTTTGCACCTTTAAATTTTTAGG CGAAGATGATAGTATTTTGACCGT	mutagenesis primer for *T25E4.1* newpwm site at 3320 bp in pOK258.01
PO1038	ACTAGCTTCACATCAAATGACTCATGCAACCTTTG ACATTAAAATCAATTCAATACATTATAGACAG	mutagenesis primer for *T25E4.1* newpmw at 3725 bp in pOK258.01
PO1039	CTGTCTATAATGTATTGAATTGATTTTAATGTCAAA GGTTGCATGAGTCATTTGATGTGAAGCTAGT	mutagenesis primer for *T25E4.1* newpmw at 3725 bp in pOK258.01
PO1040	AACATGCTTGCACTTGAGAGCACAAAATGAATGGG GGTGAGAACGATCTT	mutagenesis primer for *T25E4.1* brachyury site 2

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description
PO1041	AAGATCGTTCTCACCCCCATTCATTTTGTGCTCTCAAGTGCAAGCATGTT	mutagenesis primer for *T25E4.1* brachyury site 2
PO1042	TCTTTAAAAATAGGAAAGTATTGGTAGTTTTCAGTAAATTCTTCAACTAACGGCGAAACGTGAAATGAATGAC	mutagenesis primer for *T25E4.1* brachyury site 3
PO1043	GTCATTCATTTCACGTTTCGCCGTTAGTTGAAGAATTTACTGAAAACTACCAATACTTTCCTATTTTTAAAGA	mutagenesis primer for *T25E4.1* brachyury site 3
PO1055	CATGGATCCGAAGATGATGGGGTAACTGATGA	primer for amplifying Ce TBX-2 T-box with BamHI linker
PO1056	CATCTCGAGTTATCCAGCATCCCGAAATCCTT	primer for amplifying Ce TBX-2 T-box with xho1 linker
PO1157	GCAAGTTGCAAAGCTTGCAAAGAAACG	*D2096.6* site 1 probe
PO1158	GGCTCGAGCCTTATTTACATTCATATTGC	*D2096.6* site 1 probe
PO1159	CTCGTCGACTCGTGTTACTTGAATATC	*D2096.6* site 2 (4) probe
PO1160	AGCGTCGACGAATATTCAACTGCATAG	*D2096.6* site 2 (4) probe

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description 5'-3'
PO1063	CCGCATGTCCATAGCTTGT	primer for sequencing *T25E4.1* brachyury site 2 mutation
PO1064	AATGGGTTCTTTGGCAGAGT	primer for sequencing *T25E4.1* brachyury site 2 mutation
PO1083	CTTTGGCAGCTCATATCTCAGTAGTCGTTTGTTCTATAGA AAAATAATCAACTGATAAAATATTTTTA	mutagenesis primer for *T25E4.1* newPWM site 2546-2556
PO1084	TAAAAAATATTTTATCAGTTGATTATTTTTCTATAGAACA AACGACTACTGAGATATGAGCTGCCAAAG	mutagenesis primer for *T25E4.1* newPWM site 2546-2556
PO1085	GAAAGGAATGTCAAAGGTTGCATGAGTCATTTAAAATGA AGCTAGTTTGAATTAACAATGTA	mutagenesis primer for *T25E4.1* newPWM site 3694-3704
PO1086	TACATTGTTAATTCAAACTAGCTTCATTTTAAATGACTCA TGCAACCTTTGACATTCCTTTC	mutagenesis primer for *T25E4.1* newPWM site 3694-3704
PO1094	ACAACAGCTGAGGGTTACGC	primer to sequence mutation in *T25E4.1* newPWM at 2546-2556 of POK258.01
PO1095	TTTTGCACCCCTTTCTCAAC	primer to sequence mutation in *T25E4.1* newPWM at 2546-2556 of POK258.01
PO1096	CATCACAATCTACGGCTCCA	primer to sequence mutation in *T25E4.1* newPWM at 3694-3704 POK258.01

APPENDIX E TABLE XIX: OLIGONUCLEOTIDES (continued)

Name	Sequence 5'-3'	Description 5'-3'
PO1097	TGACAGTTCTTCCCATGTGC	primer to sequence mutation in *T25E4.1* newPWM at 3694-3704 POK258.01
PO1209	CAAGCGCTCACTTTTTCTACAATTAATGTGTACCC	*D2096.6* site 1 probe mutated
PO1210	CTTGGGGTACACATTAATTGTAGAAAAAGTGAGCG	*D2096.6* site 1 probe mutated
PO1211	CAAGATTTTGATATTTTAAATTCAACATTTGATAT	*D2096.6* site 2 (4) probe mutated
PO1212	CTTGATATCAAATGTTGAATTTAAAATATCAAAAT	*D2096.6* site 2 (4) probe mutated

APPENIDX F: Plasmids

APPENDIX F TABLE XX: PLASMIDS

Name	Description
pOK165.07	yk112c4 (*tbx-2*/F21H11.3) cDNA in pbluscript sk-. from Y. Kohara
pOK165.11	yk386a11 (B0336.7) cDNA cloned into EcoRi/XhoI sites of pBS SK-.
pOK165.12	yk484c12 (*egrh-1*) cDNA cloned into EcoRi/XhoI sites of pBS SK-. Sequence file may not be completely accurate since it was made using cDNA sequence predicted on Wormbase.
pOK172.13	ZC328.2 in KS+
pOK173.19	Y55F3AM.7 cDNA in pBS KS+. XhoI fragment from 31.1.F67 (- orientation) cloned into pBS KS+ linearized with XhoI.
pOK179.41	*egrh-1::gfp* translational fusion to pOK173.06. Sph-BglII fragment of C27C12 (bp 10647-17295) ligated to pOK173.06 linearized with SphI and BglII
pOK196.02	MBP::EGRH-1 fusion. EGRH-1 C-terminus amplified with PO695/696 and cloned into EcoRI-PstI sites of pMal C2
pOK196.08	B0336.7 cDNA XhoI-BglII fragment from pOK165.11 cloned into SmaI-BamHI digested pGEX-1

APPENDIX F TABLE XX: PLASMIDS (continued)

Name	Description
pOK210.03	B0336.7 BglII to Kpn fragment from pOK165.11 cloned into BamHI to Kpn of pRSETB
pOK236.01, pOK236.01@	full length *egrh-1* plus 3'UTR. PCR amplified *egrh-1* from C27C12 cosmid using PO817/818 cut with ApaI-MluI ligated to 179.41 cut with ApaI-MluI
pOK237.04	*tbx-2* cDNA (1396bp frag of pOK186.55/Sac1, Kpn1) cloned into pGEM-4Z (2740bp frag of pOK237.10/ Sac1, Kpn1) downstream from SP6 promoter.
pOK240.01	pPD103.87 xba1-Bgl1 GFP fragmented inserted into pOK236.01 cut with Nhe1. Full length *egrh-1::GFP* construct
pOK246.03	*tbx-2* cDNA amplifed from pOK237.04 with PO873 & PO874, digested with XbaI and XhoI, and cloned into XbaI-SalI digested pMalC2. Insert sequenced by RL (8/13/08) and it's perfect. 246.04 and 246.05 were not sequenced.
pOK252.09	*pqn-71::gfp* constructed by Ashley
pOK253.05	*D2096.6::gfp::his2B* received from Susan Mango constructed from pAP.10 contains about 1 kb of upstream sequence
pOK256.01	Site directed mutagenesis used to mutate T-box site 1 (farthest from ATG) in *D2096.6* insert. Mutated CACC to ACAA
pOK258.01, pOK258.02	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector

APPENDIX F TABLE XX: PLASMIDS (continued)

Name	Description
pOK258.03, pOK258.04, pOK258.05	Site directed mutagenesis used to mutate T-box site #2 (1=farthest from ATG, 5=closest to ATG) in *D2096.6* insert. Mutated GGT to AAA,
pOK261.05	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector newpwm at site 3320-3330 was mutated from CCTC to AAAA using primers PO1036/PO1037
pOK261.06	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector. newpwm at site 3725-3735 was mutated from CCTT to AAAA using primers PO1038/PO1039
pOK262.01, pOK262.02	*ges-1* plasmid pJM16
pOK262.03, pOK262.04	*ges-1* insert is reverse from original sequence from McGhee. based on sequencing with M13R and PO290 & PO360 *egrh-1* cDNA downstream of *ges-1* promoter
pOK262.05, pOK262.06	*elt-2* promoter
pOK262.07, pOK262.08	site directed mutagenesis used to mutate T-box site 1 and 2 (4) (1=farthest from ATG, 5=closest to ATG) in *D2096.6* insert. cloned from single mutant plasmids

APPENDIX F TABLE XX: PLASMIDS (continued)

Name	Description
pOK262.09, pOK262.10	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector site 1 at 777-787 bp mutated by site directed mutagenesis
pOK262.11, pOK262.12	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector site 2 at 3815-3825 mutated by site directed mutagensis
pOK267.04	pOK258.01 with newPWM site 2546-2556 mutated from AGTTGT-AATAAT with PO1083/1084 by site directed muatgenesis
pOK267.05	pOK258.01 with newPWM site 3694-3704 with sequence mutated from GATG-AAAA with primers PO1085/1086 by site directed mutagenesis
pOK267.07	pOK258.01 site 3 at 3493-3503 bp mutated by site directed mutagenesis with primers PO1042/1043
pOK267.08	pOK258.01 site 2 at 1774-1784 bp mutated by site directed mutagenesis with primers PO1040/1041
pOK269.07	pGEX-4T2 cloning vector, containing the DNA binding domain of tbx-2
pPD95.77	Fire expression vector
pMAL-2C	Expression vector for maltose binding protein fusions
pRSETB	pRSETB with fusion ATG at +1

APPENIDX G: *C. elegans* strains

TABLE XXI: *C. ELEGANS* STRAINS

Strain	Genotype	Description
OK0003	*dpy-11(e224) unc-76(e911) V*	Dpy Unc
OK0507	*cuIs21 III*	*ceh-22::gfp* integrated strain expressing nuclear localized gfp.
OK0553	*ZK337.2(tm1134)*	homozygous viable strain, outcrossed 4 times
OK0558	*ZK337.2 (tm706)*	homozygous viable strain, outcrossed 4 times
OK0559	*egrh-1(tm1736)*	homozygous viable strain, outcrossed 2 times
OK0570	*B0336.7 (tm1642)/+*	heterozygous strain not outcrossed
OK0579	*dpy-17(e164) B0336.7(tm1642); sDp3(III;f)*	transfer WT to maintain. segregates Dpy dpy-17(e164) B0336.7(tm1642) and WT dpy-17(e164) B0336.7(tm1642); sDp3
OK0580	*dpy-17(e164) B0336.7(tm1642)/ qC1[dpy-19(e1259) glp-1(q339)]*	transfer WT to maintain. segregates WT heterozygotes, Dpy-17 dpy-17(e164) B0336.7(tm1642) homozygotes and Dpy-19 Ste qC1 homozygotes
OK0600	*fog-1(q253); egrh-1(tm1736)*	temperature sensitive strain. maintain at 16 or 20 degrees. shift to 25 degrees for *egrh-1(tm1736)* females.
OK0601	*egrh-1(tm1736);cuEx503*	carrying *C27C12* (at 2ug/ml) and pRF4 (at 100 ug/ml) as an extrachromosomal array. maintain by passing rollers segregates 40% rollers. 87% rescued

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Genotype	Description
OK0602	*egrh-1(tm1736); cuEx504*	carrying *C27C12* (at 2ug/ml) and pRF4 (at 100 ug/ml) as an extrachromosomal array. maintain by passing rollers segregates 25% rollers. 83% rescued
OK0603	*egrh-1(tm1736); cuEx505*	carrying *C27C12* (at 2ug/ml) and pRF4 (at 100 ug/ml) as an extrachromosomal array. maintain by passing rollers segregates 30% rollers
OK0604	*egrh-1(tm1736); cuEx506*	carrying *C27C12* (at 2ug/ml) and pRF4 (at 100 ug/ml) as an extrachromosomal array. maintain by passing rollers segregates 20% rollers
OK0636	*egrh-1(tm1736); cuEx531*	carrying *C27C12* (at 2ug/ml), pRF4 (at 100 ug/ml) and pOK236.02 (pTG96 *sur-5* GFP(NLS) (at 100ug/ml) as an extrachromosomal array. maintain by passing rollers
OK0637	*unc-51(e369) rol-9(scl48) V; egrh-1(tm1736)*	Unc rollers
OK0638	*dpy-11(e224) unc-76 (e911) V; egrh-1(tm1736)*	Dpy Unc
OK0639	*fog-2(q71); egrh-1(tm1736)*	male/female strain maintain by crossing

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Genotype	Description
OK0653	*egrh-1(tm1736); cuEx542*	carrying pOK236.01 (full length *egrh-1* plus 3' UTR 2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array. passage rollers 82% rescue 50% rollers
OK0654	*egrh-1(tm1736); cuEx543*	carrying pOK236.01 (full length *egrh-1* plus 3' UTR 2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array. passage rollers 69% rescue 31% rollers
OK0655	*egrh-1(tm1736); cuEx544*	carrying pOK236.01 (full length *egrh-1* plus 3' UTR 2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array. passage rollers 63% rescue 40% rollers
OK0656	*egrh-1(tm1736); cuEx545*	carrying pOK236.01 (full length *egrh-1* plus 3' UTR 2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array
OK0658	*egrh-1(tm1736); cuEx547*	carrying pOK240.01 (full length *egrh-1* with GFP inserted in frame in 5th exon plus 3' UTR 2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array passage rollers. 75% rescue
OK0660	*mab-22(bx59)III*	*tbx-2* mutant outcrossed from him-5(e1490) Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult stage. pick wild type hermaphrodites to maintain

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Genotype	Description
OK0661	*cu13 II; tbx-2(bx59)III*	*tbx-2* mutant with enhancer of ts larval arrest outcrossed from *him-5(e1490)* Wild type at 16C. Strong temp sensitive at 25C, no progeny grow to adult stage
OK0666	*cuEx553*	WT carrying *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array (plasmid pOK253.05) segregates 50% rollers
OK0667	*cuEx554*	WT carrying *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array (plasmid pOK253.05) segregates 30% rollers
OK0669	*cuEx556*	WT carrying *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array
OK0670	*cuEx557*	WT carrying *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array segregates low % rollers

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Genotype	Description
OK0692	*tbx-2(bx59)III; cuEx553*	Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult stage. Carrying D2096.6::GFP::his2B (2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array(plasmid pOK253.05)
OK0693	*tbx-2(bx59)III; cuEx554*	Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult stage.Carrying *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul) as an extrachromosomal array.(plasmid pOK253.05). ln S in my notes
OK0694	*cuEx566*	carrying pOK256.01 site directed mutagenesis used to mutate T-box site 1. *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul). Mutated CACC to ACAA, sequenced
OK0695	*cuEx567*	carrying pOK256.01 Site directed mutagenesis used to mutate T-box site 1. *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul). Mutated CACC to ACAA, sequenced
OK0696	*tbx-2(bx59)III; cuEx566*	Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult stage. Site directed mutagenesis used to mutate T-box site 1. *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul). Mutated CACC to ACAA, sequenced
OK0697	*tbx-2(bx59)III; cuEx567*	Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult stage carrying pOK256.01 Site directed mutagenesis used to mutate T-box site 1. *D2096.6::GFP::his2B* (2ng/ul) and pRF4 (100ng/ul). Mutated CACC to ACAA, sequenced line 32

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Gentotype	Description
OK0702	*cuEx560*	carrying pOK261.05. Site directed mutagenesis used to mutate newpwm at site 3320 bp *T25E4.1::gfp* (pOK258.01) (2ng/ul) and pRF4 (100ng/ul). Mutated CCTC to AAAA, sequenced
OK0703	*cuEx561*	carrying pOK261.06 site directed mutagenesis used to mutate newpwm at site 3725 bp of *T25E4.1::gfp* (pOK258.01) (2ng/μl) and pRF4 (100ng/μl). Mutated CCTT to AAAA, sequenced
OK0704	*egrh-1(tm1736); cuEx562*	*egrh-1(tm1736)* mutants carrying an array with *ges-1* promoter driving *egrh-1* cDNA this array shows rescue of *egrh-1* mutant defects at ~50% (1.4ng/μl DNA) (100ng/μl pRF4) injected pOK262.03
OK0705	*egrh-1(tm1736); cuEx563*	*egrh-1(tm1736)* mutants carrying an array with *ges-1* promter driving egrh-1 cDNA (1.4ng/μl DNA) (100ng/μl pRF4)
OK0706	*cuEx564*	carrying pOK261.05 Site directed mutagenesis used to mutate newpwm at site 3320 bp *T25E4.1::gfp* (pOK258.01) (2ng/μl) and pRF4 (100ng/μl). Mutated CCTC to AAAA, sequenced
OK0707	*cuEx565*	carrying pOK261.06 Site directed mutagenesis used to mutate newpwm at site 3725 bp of *T25E4.1::gfp* (pOK258.01) (2ng/μl) and pRF4 (100ng/μl). Mutated CCTT to AAAA, sequenced
OK0733	*fog-2(q71); egrh-1(tm1736) cuEx562*	*fog-2(q71); egrh-1(tm1736)* mutants carrying an array with *ges-1* promter driving *egrh-1* cDNA. OK0704 was crossed into the *fog-2(q71); egrh-1(tm1736)* OK0639 rescue ~55% male female rolling strain maintain by crossing

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Gentotype	Description
OK0734	*egrh-1(tm1736); cuEx586*	*egrh-1(tm1736)* mutants carrying an array with *ges-1* promter driving egrh-1 cDNA this array shows rescue of *egrh-1* mutant defects at ~22% (1.4ng/μl pOK262.03) (100ng/μl pRF4) 22% rescue
OK0735	*egrh-1(tm1736); cuEx587*	*egrh-1(tm1736)* mutants carrying an array with *ges-1* promter driving egrh-1 cDNA this array shows rescue of EGRH-1 mutant defects at ~19% (1.4ng/μl pOK262.03) (100ng/μl pRF4) 19% rescue
OK0736	*cuEx588*	*D2096.6* t-box binding site 1 and 2 mutated and injected at 2ng/μl (pOK262.07) +100ng/μl pRF4
OK0737	*cuEx589*	*D2096.6* t-box binding site 1 and 2 mutated and injected at 2ng/μl (pOK262.07) +100ng/μl pRF4 pick rollers
OK0738	*cuEx590*	*T25E4.1* plasmid pOK262.11 containing 4.5 kb promoter region of T25E4.1 cloned into pPD 95.77 GFP vector site 4 at 3815-3825 mutated by site directed mutagensis. injected at (1.6 ng/μl DNA) and (100ng/μl pRF4)
OK0739	*cuEx591*	*T25E4.1* plasmid pOK262.11 containing 4.5 kb promoter region of T25E4.1 cloned into pPD 95.77 GFP vector site 4 at 3815-3825 mutated by site directed mutagenesis. injected at (1.6 ng/μl DNA) and (100ng/μl pRF4)

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Gentotype	Description
OK0740	*cuEx592*	*T25E4.1* plasmid pOK262.12 containing 4.5 kb promoter region of T25E4.1 cloned into pPD 95.77 GFP vector site 4 at 3815-3825 mutated by site directed mutagenesis. injected at (2 ng/µl DNA) and (100ng/µl pRF4)
OK0741	*tbx-2(ok529)/dpy-17(e164) unc-32(e189); cuEx553*	crossed into with OK0666 carrying *D2096.6::GFP::his2B* (2ng/µl) and pRF4 (100ng/µl) as an extrachromosomal array (plasmid pOK253.05) Balanced hermaphrodite segregating WT heterozygotes, dpy unc, and ok529 homozygotes.passage WT rollers
OK0798	*tbx-2(bx59)III; cuEx642*	T25E4.1 4.5 kb promoter GFP PCR fusion
OK0799	*tbx-2(bx59)III; cuEx589*	Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult stage *D2096.6* t-box binding site 1 and 2 mutated and injected at 2ng/µl (pOK262.07) +100ng/µl pRF4
OK0800	*cuEx643*	pOK258.03 carrying *D2096.6::GFP::his2B* with mutated T-box site 2. (2ng/µl) and pRF4 (100ng/µl). Mutated GGT to AAA
OK0802	*tbx-2(bx59)III; cuEx644*	Wild type at 16C. Weak temp sensitive at 25C, some larval arrest several grow to adult pOK258.03 carrying *D2096.6::GFP::his2B* with mutated T-box site 2. (2ng/µl) and pRF4 (100ng/µl). Mutated GGT to AAA
OK0803	*cuEx645*	*T25E4.1* plasmid pOK267.08 containing 4.5 kb promoter region of T25E4.1 cloned into pPD 95.77 GFP vector site 2 at 1774-1784 bp mutated by site directed mutagenesis. injected at (2 ng/µl DNA) and (100ng/µl pRF4)

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Gentotype	Description
OK0804	*cuEx646*	*T25E4.1* plasmid pOK262.09 containing 4.5 kb promoter region of T25E4.1 cloned into pPD 95.77 GFP vector site 1 at 777-787 bp mutated by site directed mutagenesis injected at (2 ng/µl DNA) and (100ng/µl pRF4)
OK0805	*cuEx647*	*T25E4.1* plasmid pOK262.11 containing 4.5 kb promoter region of T25E4.1 cloned into pPD 95.77 GFP vector site 4 at 3815-3825 mutated by site directed mutagensis. injected at (1.6 ng/µl DNA) and (100ng/µl pRF4)
OK0806	*cuEx648*	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector
OK0807	*cuEx649*	4.5 kb promoter region of *T25E4.1* cloned into pPD 95.77 GFP vector
OK0808	*cuEx650*	*T25E4.1* plasmid with newPWM site 2546-2556 mutated from AGTTGT-AATAAT with PO1083/1084 by site directed muatgenesis
OK0809	*cuEx651*	*T25E4.1* plasmid with newPWM site 2546-2556 mutated from AGTTGT-AATAAT with PO1083/1084 by site directed muatgenesis
OK0810	*cuEx652*	*T25E4.1* plasmid newPWM site 2546-2556 mutated from AGTTGT-AATAAT with PO1083/1084 by site directed muatgenesis
OK0811	*cuEx653*	*T25E4.1* plasmid with newPWM site 3694-3704 with sequence mutated from GATG-AAAA with primers PO1085/1086 by site directed mutagenesis

APPENDIX G TABLE XXI: *C. ELEGANS* STRAINS (continued)

Strain	Gentotype	Description
OK0812	*cuEx654*	*T25E4.1* plasmid with newPWM site 3694-3704 with sequence mutated from GATG-AAAA with primers PO1085/1086 by site directed mutagenesis
OK0813	*cuEx655*	*T25E4.1* plasmid with newPWM site 3694-3704 with sequence mutated from GATG-AAAA with primers PO1085/1086 by site directed mutagenesis
BC4697	*sDf121(s2098) unc-32(e189) III; sDp3 (III;f)*	sDf121.Unc strain.
BW1935	*unc-119(ed3) III; ctIs43 him-5(e1490) V.*	worms carrying *dbl-1::gfp.* passage unc worrns
CB2167	*dpy-5 (e61) unc-13 (e1091) I*	Dpy. Unc.strain
CB 3297	*vab-9(e1744)II; him-5(e1490)V.*	Him. M-MATING+POOR <1%WT. Tail abnormalities->knob-like swelling in larvae and adults.
CB4147	*fog-1(e2121) unc-11 (e47)l; sDp2(l;f)*	animals with Dup are WT. Animals which have lost the Dup are Unc and female. maintain by picking WT
DR 103	*dpy-10(e128) unc-4(e120) II*	DpyUnc
DR466	*him-5(e1490)V*	pick wild type hermaphrodites to maintain. approximately 33% males. received from CGC
EM207	*mab-22(bx59)III; him-5(e1490)V*	*tbx-2* mutant. Conditional lethal; inviable when grown at 25C. Wild type at 16C.
JK560	*fog-1(q253) l*	temperature sensitive allele grow at 15C or 20C

APPENDIX G TABL XXI: *C. ELEGANS* STRAINS (continued)

Strain	Gentotype	Description
JK574	*fog-2(q71)V*	male-female strain maintain by mating
JK2868	*[unc-119(ed3); qIs56 (unc-119 (+); lag-2::GFP]*	*lag-2::gfp*
NY2045	*ynIs45;him-5(e1490)*	*Pflp-15::GFP*
NY2040	*ynIs40*	*Pflp-11::GFP*
NY2097	*ynIs97*	*Pflp-1::GFP*
NL2099	*rrf-3(pk1426) II*	homozygous *rrf-3* deletion allele. Increased sensitivity to RNAi when compared to WT animals.
NL2098	*rrf-1(pk1417)*	RNAi interference for genes expressed in somatic tissue is lost in *rrf-1* deletion mutants
SP471	*dpy-17(e164) unc-32(e189) III*	Dpy Unc strain.

APPENDIX H: Copyright

GNU Free Documentation License

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this

Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled \"History\", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled \"History\" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the \"History\" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled \"Acknowledgements\" or \"Dedications\", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled \"Endorsements\". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled \"Endorsements\" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the

previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/
.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

Copyright (c) YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this documentunder
the terms of the GNU Free Documentation License, Version 1.2or any later version published by the Free
Software Foundation;with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.A copy of the
license is included in the section entitled "GNUFree Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with...Texts." line with this:

with the Invariant Sections being LIST THEIR TITLES, with theFront-Cover Texts being LIST, and with the
Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

CITED LITERATURE

Alberini, C. M. (2009) 'Transcription factors in long-term memory and synaptic plasticity', *Physiol Rev* 89(1): 121-45.

Albertson and Thomson (1976) The pharynx of Caenorhabditis elegans. in Thomson (ed.), vol. 275. Philos Trans R Soc Lond B Biol Sci.

Ausubel, F. M. (1990) *Current protocols in molecular biology*, New York: Greene Pub. Associates and Wiley-Interscience : J. Wiley.

Avery, L. (1993) 'The genetics of feeding in Caenorhabditis elegans.', *Genetics* 133(4): 897-917.

Barton, M. K. and Kimble, J. (1990) 'fog-1, a regulatory gene required for specification of spermatogenesis in the germ line of Caenorhabditis elegans', *Genetics* 125(1): 29-39.

Basson, C. T., Bachinsky, D. R., Lin, R. C., Levi, T., Elkins, J. A., Soults, J., Grayzel, D., Kroumpouzou, E., Traill, T. A., Leblanc-Straceski, J. et al. (1997) 'Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome.', *Nat Genet* 15(1): 30-5.

Beaster-Jones, L. and Okkema, P. G. (2004) 'DNA binding and in vivo function of C. elegans PEB-1 require a conserved FLYWCH motif', *J Mol Biol* 339(4): 695-706.

Bird, D. M. and Riddle, D. L. (1989) 'Molecular cloning and sequencing of ama-1, the gene encoding the largest subunit of Caenorhabditis elegans RNA polymerase II.', *Mol Cell Biol* 9(10): 4119-30.

Bowerman, B., Draper, B. W., Mello, C. C. and Priess, J. R. (1993) 'The maternal gene skn-1 encodes a protein that is distributed unequally in early C. elegans embryos.', *Cell* 74(3): 443-52.

Bowerman, B., Tax, F. E., Thomas, J. H. and Priess, J. R. (1992) 'Cell interactions involved in development of the bilaterally symmetrical intestinal valve cells during embryogenesis in Caenorhabditis elegans.', *Development* 116(4): 1113-22.

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

Brock, T. J., Browse, J. and Watts, J. L. (2006) 'Genetic regulation of unsaturated fatty acid composition in C. elegans.', *PLoS Genet* 2(7): e108.

CITED LITERATURE (continued)

Broitman-Maduro, G., Lin, K. T., Hung, W. W. and Maduro, M. F. (2006) 'Specification of the C. elegans MS blastomere by the T-box factor TBX-35.', *Development* 133(16): 3097-106.

Broitman-Maduro, G., Owraghi, M., Hung, W. W., Kuntz, S., Sternberg, P. W. and Maduro, M. F. (2009) 'The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in C. elegans mesoderm development.', *Development* 136(16): 2735-46.

Byerly, L., Cassada, R. C. and Russell, R. L. (1976) 'The life cycle of the nematode Caenorhabditis elegans. I. Wild-type growth and reproduction.', *Dev Biol* 51(1): 23-33.

Carlson, H., Ota, S., Campbell, C. E. and Hurlin, P. J. (2001) 'A dominant repression domain in Tbx3 mediates transcriptional repression and cell immortalization: relevance to mutations in Tbx3 that cause ulnar-mammary syndrome.', *Hum Mol Genet* 10(21): 2403-13.

Carreira, S., Dexter, T. J., Yavuzer, U., Easty, D. J. and Goding, C. R. (1998) 'Brachyury-related transcription factor Tbx2 and repression of the melanocyte-specific TRP-1 promoter.', *Mol Cell Biol* 18(9): 5099-108.

Casey, E. S., O'Reilly, M. A., Conlon, F. L. and Smith, J. C. (1998) 'The T-box transcription factor Brachyury regulates expression of eFGF through binding to a non-palindromic response element.', *Development* 125(19): 3887-94.

Coates, J. C. and de Bono, M. (2002) 'Antagonistic pathways in neurons exposed to body fluid regulate social feeding in Caenorhabditis elegans.', *Nature* 419(6910): 925-9.

Coll, M., Seidman, J. G. and Müller, C. W. (2002) 'Structure of the DNA-bound T-box domain of human TBX3, a transcription factor responsible for ulnar-mammary syndrome.', *Structure* 10(3): 343-56.

Consortium, C. e. S. (1998) 'Genome sequence of the nematode C. elegans: a platform for investigating biology.', *Science* 282(5396): 2012-8.

Cunliffe, V. and Smith, J. C. (1992) 'Ectopic mesoderm formation in Xenopus embryos caused by widespread expression of a Brachyury homologue.', *Nature* 358(6385): 427-30.

Desai, C., Garriga, G., McIntire, S. L. and Horvitz, H. R. (1988) 'A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons.', *Nature* 336(6200): 638-46.

CITED LITERATURE (continued)

Di Gregorio, A. and Levine, M. (1999) 'Regulation of Ci-tropomyosin-like, a Brachyury target gene in the ascidian, Ciona intestinalis.', *Development* 126(24): 5599-609.

Dong, J., Boyd, W. A. and Freedman, J. H. (2008) 'Molecular characterization of two homologs of the Caenorhabditis elegans cadmium-responsive gene cdr-1: cdr-4 and cdr-6.', *J Mol Biol* 376(3): 621-33.

Duerr, J. S. (2006) 'Immunohistochemistry.', *WormBook*: 1-61.

Eckmann, C. R., Crittenden, S. L., Suh, N. and Kimble, J. (2004) 'GLD-3 and control of the mitosis/meiosis decision in the germline of Caenorhabditis elegans', *Genetics* 168(1): 147-60.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998) 'Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.', *Nature* 391(6669): 806-11.

Francis, R., Barton, M. K., Kimble, J. and Schedl, T. (1995) 'gld-1, a tumor suppressor gene required for oocyte development in Caenorhabditis elegans', *Genetics* 139(2): 579-606.

Fu, M., Zhu, X., Zhang, J., Liang, J., Lin, Y., Zhao, L., Ehrengruber, M. U. and Chen, Y. E. (2003) 'Egr-1 target genes in human endothelial cells identified by microarray analysis', *Gene* 315: 33-41.

Gaudet, J. and Mango, S. E. (2002) 'Regulation of organogenesis by the Caenorhabditis elegans FoxA protein PHA-4.', *Science* 295(5556): 821-5.

Gibson-Brown, J. J., Agulnik, S. I., Silver, L. M., Niswander, L. and Papaioannou, V. E. (1998) 'Involvement of T-box genes Tbx2-Tbx5 in vertebrate limb specification and development.', *Development* 125(13): 2499-509.

Gill, G. (2005) 'Something about SUMO inhibits transcription.', *Curr Opin Genet Dev* 15(5): 536-41.

Good, K., Ciosk, R., Nance, J., Neves, A., Hill, R. J. and Priess, J. R. (2004) 'The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C. elegans embryos.', *Development* 131(9): 1967-78.

CITED LITERATURE (continued)

Govindan, J. A., Cheng, H., Harris, J. E. and Greenstein, D. (2006) 'Galphao/i and Galphas signaling function in parallel with the MSP/Eph receptor to control meiotic diapause in C. elegans', *Curr Biol* 16(13): 1257-68.

Grant, B. and Hirsh, D. (1999) 'Receptor-mediated endocytosis in the Caenorhabditis elegans oocyte', *Mol Biol Cell* 10(12): 4311-26.

Greenstein, D. (2005) 'Control of oocyte meiotic maturation and fertilization', *WormBook*: 1-12.

Greenstein, D., Hird, S., Plasterk, R. H., Andachi, Y., Kohara, Y., Wang, B., Finney, M. and Ruvkun, G. (1994) 'Targeted mutations in the Caenorhabditis elegans POU homeo box gene ceh-18 cause defects in oocyte cell cycle arrest, gonad migration, and epidermal differentiation', *Genes Dev* 8(16): 1935-48.

Hall, D. H., Winfrey, V. P., Blaeuer, G., Hoffman, L. H., Furuta, T., Rose, K. L., Hobert, O. and Greenstein, D. (1999) 'Ultrastructural features of the adult hermaphrodite gonad of Caenorhabditis elegans: relations between the germ line and soma', *Dev Biol* 212(1): 101-23.

Hansen, D. and Schedl, T. (2006) 'The regulatory network controlling the proliferation-meiotic entry decision in the Caenorhabditis elegans germ line', *Curr Top Dev Biol* 76: 185-215.

He, M., Wen, L., Campbell, C. E., Wu, J. Y. and Rao, Y. (1999) 'Transcription repression by Xenopus ET and its human ortholog TBX3, a gene involved in ulnar-mammary syndrome.', *Proc Natl Acad Sci U S A* 96(18): 10212-7.

Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. and Lehrach, H. (1990) 'Cloning of the T gene required in mesoderm formation in the mouse.', *Nature* 343(6259): 617-22.

Hillier, L. W., Miller, R. D., Baird, S. E., Chinwalla, A., Fulton, L. A., Koboldt, D. C. and Waterston, R. H. (2007) 'Comparison of C. elegans and C. briggsae genome sequences reveals extensive conservation of chromosome organization and synteny.', *PLoS Biol* 5(7): e167.

Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R. and Komuro, I. (2001) 'Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation.', *Nat Genet* 28(3): 276-80.

CITED LITERATURE (continued)

Hirsh, D., Oppenheim, D. and Klass, M. (1976) 'Development of the reproductive system of Caenorhabditis elegans', *Dev Biol* 49(1): 200-19.

Hodgkin, J., Horvitz, H. R. and Brenner, S. (1979) 'Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS.', *Genetics* 91(1): 67-94.

Horner, M. A., Quintin, S., Domeier, M. E., Kimble, J., Labouesse, M. and Mango, S. E. (1998) 'pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in Caenorhabditis elegans.', *Genes Dev* 12(13): 1947-52.

Houlden, H. and Reilly, M. M. (2006) 'Molecular genetics of autosomal-dominant demyelinating Charcot-Marie-Tooth disease', *Neuromolecular Med* 8(1-2): 43-62.

Hubbard, E. J. and Greenstein, D. (2000) 'The Caenorhabditis elegans gonad: a test tube for cell and developmental biology', *Dev Dyn* 218(1): 2-22.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. and Speed, T. P. (2003) 'Exploration, normalization, and summaries of high density oligonucleotide array probe level data.', *Biostatistics* 4(2): 249-64.

Iwasaki, K., McCarter, J., Francis, R. and Schedl, T. (1996) 'emo-1, a Caenorhabditis elegans Sec61p gamma homologue, is required for oocyte development and ovulation', *J Cell Biol* 134(3): 699-714.

J. G. White, E. S., J. N. Thomson and S. Brenner (1986) The Structure of the Nervous System of the Nematode Caenorhabditis elegans, vol. 314: Phil. Trans. R. Soc. Lond. B.

Jain, N., Thatte, J., Braciale, T., Ley, K., O'Connell, M. and Lee, J. K. (2003) 'Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays.', *Bioinformatics* 19(15): 1945-51.

Joslin, J. M., Fernald, A. A., Tennant, T. R., Davis, E. M., Kogan, S. C., Anastasi, J., Crispino, J. D. and Le Beau, M. M. (2007) 'Haploinsufficiency of EGR1, a candidate gene in the del(5q), leads to the development of myeloid disorders', *Blood* 110(2): 719-26.

Kalb, J. M., Lau, K. K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P. G. and McGhee, J. D. (1998) 'pha-4 is Ce-fkh-1, a fork head/HNF-3alpha,beta,gamma homolog that functions in organogenesis of the C. elegans pharynx.', *Development* 125(12): 2171-80.

CITED LITERATURE (continued)

Kaletta, T. and Hengartner, M. O. (2006) 'Finding function in novel targets: C. elegans as a model organism.', *Nat Rev Drug Discov* 5(5): 387-98.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M. et al. (2003) 'Systematic functional analysis of the Caenorhabditis elegans genome using RNAi.', *Nature* 421(6920): 231-7.

Kennedy, B. P., Aamodt, E. J., Allen, F. L., Chung, M. A., Heschl, M. F. and McGhee, J. D. (1993) 'The gut esterase gene (ges-1) from the nematodes Caenorhabditis elegans and Caenorhabditis briggsae', *J Mol Biol* 229(4): 890-908.

Kim, K. and Li, C. (2004) 'Expression and regulation of an FMRFamide-related neuropeptide gene family in Caenorhabditis elegans.', *J Comp Neurol* 475(4): 540-50.

Kimble, J. and Crittenden, S. L. (2005) 'Germline proliferation and its control', *WormBook*: 1-14.

Kimble, J. and Sharrock, W. J. (1983) 'Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans', *Dev Biol* 96(1): 189-96.

Kimble, J. E. and White, J. G. (1981) 'On the control of germ cell development in Caenorhabditis elegans', *Dev Biol* 81(2): 208-19.

Kispert, A. and Herrmann, B. G. (1993) 'The Brachyury gene encodes a novel DNA binding protein.', *EMBO J* 12(8): 3211-20.

Kispert, A., Koschorz, B. and Herrmann, B. G. (1995) 'The T protein encoded by Brachyury is a tissue-specific transcription factor.', *EMBO J* 14(19): 4763-72.

Klass, M., Wolf, N. and Hirsh, D. (1976) 'Development of the male reproductive system and sexual transformation in the nematode Caenorhabditis elegans.', *Dev Biol* 52(1): 1-18.

Kohara, Y. (2001) '[Systematic analysis of gene expression of the C. elegans genome]', *Tanpakushitsu Kakusan Koso* 46(16 Suppl): 2425-31.

Kramer, J. M., French, R. P., Park, E. C. and Johnson, J. J. (1990) 'The Caenorhabditis elegans rol-6 gene, which interacts with the sqt-1 collagen gene to determine organismal morphology, encodes a collagen.', *Mol Cell Biol* 10(5): 2081-9.

Kubagawa, H. M., Watts, J. L., Corrigan, C., Edmonds, J. W., Sztul, E., Browse, J. and Miller, M. A. (2006a) 'Oocyte signals derived from polyunsaturated fatty acids control sperm recruitment in vivo', *Nat Cell Biol* 8(10): 1143-8.

CITED LITERATURE (continued)

Kubagawa, H. M., Watts, J. L., Corrigan, C., Edmonds, J. W., Sztul, E., Browse, J. and Miller, M. A. (2006b) 'Oocyte signals derived from polyunsaturated fatty acids control sperm recruitment in vivo.', *Nat Cell Biol* 8(10): 1143-8.

Kuchenthal, C. A., Chen, W. and Okkema, P. G. (2001) 'Multiple enhancers contribute to expression of the NK-2 homeobox gene ceh-22 in C. elegans pharyngeal muscle.', *Genesis* 31(4): 156-66.

Kuwabara, P. E. (2003) 'The multifaceted C. elegans major sperm protein: an ephrin signaling antagonist in oocyte maturation', *Genes Dev* 17(2): 155-61.

Larminie, C. G. and Johnstone, I. L. (1996) 'Isolation and characterization of four developmentally regulated cathepsin B-like cysteine protease genes from the nematode Caenorhabditis elegans.', *DNA Cell Biol* 15(1): 75-82.

Lee, J. C., VijayRaghavan, K., Celniker, S. E. and Tanouye, M. A. (1995) 'Identification of a Drosophila muscle development gene with structural homology to mammalian early growth response transcription factors', *Proc Natl Acad Sci U S A* 92(22): 10344-8.

Lee, M. H., Ohmachi, M., Arur, S., Nayak, S., Francis, R., Church, D., Lambie, E. and Schedl, T. (2007) 'Multiple functions and dynamic activation of MPK-1 extracellular signal-regulated kinase signaling in Caenorhabditis elegans germline development', *Genetics* 177(4): 2039-62.

Lee, S. L., Sadovsky, Y., Swirnoff, A. H., Polish, J. A., Goda, P., Gavrilina, G. and Milbrandt, J. (1996) 'Luteinizing hormone deficiency and female infertility in mice lacking the transcription factor NGFI-A (Egr-1)', *Science* 273(5279): 1219-21.

Lewis, J. A. and Fleming, J. T. (1995) Basic Culture Methods. in M. i. C. Biology (ed.) *Caenorhabditis elegans: Modern Biological Analysis of an Organism*, vol. 48. San Diego, CA: Academic Press.

Li, C., Kim, K. and Nelson, L. S. (1999) 'FMRFamide-related neuropeptide gene family in Caenorhabditis elegans.', *Brain Res* 848(1-2): 26-34.

Lin, R., Thompson, S. and Priess, J. R. (1995) 'pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos.', *Cell* 83(4): 599-609.

CITED LITERATURE (continued)

Lingbeek, M. E., Jacobs, J. J. and van Lohuizen, M. (2002) 'The T-box repressors TBX2 and TBX3 specifically regulate the tumor suppressor gene p14ARF via a variant T-site in the initiator.', *J Biol Chem* 277(29): 26120-7.

Lo, M. C., Gay, F., Odom, R., Shi, Y. and Lin, R. (2004) 'Phosphorylation by the beta-catenin/MAPK complex promotes 14-3-3-mediated nuclear export of TCF/POP-1 in signal-responsive cells in C. elegans.', *Cell* 117(1): 95-106.

MacQueen, A. J. and Villeneuve, A. M. (2001) 'Nuclear reorganization and homologous chromosome pairing during meiotic prophase require C. elegans chk-2', *Genes Dev* 15(13): 1674-87.

Mango, S. E., Lambie, E. J. and Kimble, J. (1994) 'The pha-4 gene is required to generate the pharyngeal primordium of Caenorhabditis elegans.', *Development* 120(10): 3019-31.

Markstein, M., Markstein, P., Markstein, V. and Levine, M. S. (2002) 'Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the Drosophila embryo', *Proc Natl Acad Sci U S A* 99(2): 763-8.

Marley, J., Lu, M. and Bracken, C. (2001) 'A method for efficient isotopic labeling of recombinant proteins.', *J Biomol NMR* 20(1): 71-5.

McCarter, J., Bartlett, B., Dang, T. and Schedl, T. (1997) 'Soma-germ cell interactions in Caenorhabditis elegans: multiple events of hermaphrodite germline development require the somatic sheath and spermathecal lineages', *Dev Biol* 181(2): 121-43.

McCarter, J., Bartlett, B., Dang, T. and Schedl, T. (1999) 'On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans', *Dev Biol* 205(1): 111-28.

McKay, S. J., Johnsen, R., Khattra, J., Asano, J., Baillie, D. L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E. et al. (2003) 'Gene expression profiling of cells, tissues, and developmental stages of the nematode C. elegans', *Cold Spring Harb Symp Quant Biol* 68: 159-69.

McMahon, A. P., Champion, J. E., McMahon, J. A. and Sukhatme, V. P. (1990) 'Developmental expression of the putative transcription factor Egr-1 suggests that Egr-1 and c-fos are coregulated in some tissues', *Development* 108(2): 281-7.

CITED LITERATURE (continued)

Mello, C. and Fire, A. (1995) DNA Transformation. in H. F. Epstein and D. C. Shakes (eds.) *Caenorhabditis elegans: Modern Biological Analysis of an Organism*, vol. 48. San Diego, CA: Academic Press.

Mello, C. C., Kramer, J. M., Stinchcomb, D. and Ambros, V. (1991) 'Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences.', *EMBO J* 10(12): 3959-70.

Meulmeester, E. and Melchior, F. (2008) 'Cell biology: SUMO.', *Nature* 452(7188): 709-11.

Miller, M. A., Nguyen, V. Q., Lee, M. H., Kosinski, M., Schedl, T., Caprioli, R. M. and Greenstein, D. (2001) 'A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation', *Science* 291(5511): 2144-7.

Miller, M. A., Ruest, P. J., Kosinski, M., Hanks, S. K. and Greenstein, D. (2003) 'An Eph receptor sperm-sensing control mechanism for oocyte meiotic maturation in Caenorhabditis elegans', *Genes Dev* 17(2): 187-200.

Minasaki, R., Puoti, A. and Streit, A. (2009) 'The DEAD-box protein MEL-46 is required in the germ line of the nematode Caenorhabditis elegans.', *BMC Dev Biol* 9: 35.

Miyahara, K., Suzuki, N., Ishihara, T., Tsuchiya, E. and Katsura, I. (2004) 'TBX2/TBX3 transcriptional factor homologue controls olfactory adaptation in Caenorhabditis elegans.', *J Neurobiol* 58(3): 392-402.

Nelms, B. L. and Hanna-Rose, W. (2006) 'C. elegans HIM-8 functions outside of meiosis to antagonize EGL-13 Sox protein function.', *Dev Biol* 293(2): 392-402.

Nishi, Y. and Lin, R. (2005) 'DYRK2 and GSK-3 phosphorylate and promote the timely degradation of OMA-1, a key regulator of the oocyte-to-embryo transition in C. elegans', *Dev Biol* 288(1): 139-49.

O'Donovan, K. J., Tourtellotte, W. G., Millbrandt, J. and Baraban, J. M. (1999) 'The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience', *Trends Neurosci* 22(4): 167-73.

Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A. and Fire, A. (1993) 'Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans.', *Genetics* 135(2): 385-404.

CITED LITERATURE (continued)

Okkema, P. G. and Krause, M. (2005) 'Transcriptional regulation.', *WormBook*: 1-40.

Ouimette, J. F., Jolin, M. L., L'honoré, A., Gifuni, A. and Drouin, J. (2010) 'Divergent transcriptional activities determine limb identity.', *Nat Commun* 1: 35.

Page, B. D., Guedes, S., Waring, D. and Priess, J. R. (2001) 'The C. elegans E2F- and DP-related proteins are required for embryonic asymmetry and negatively regulate Ras/MAPK signaling', *Mol Cell* 7(3): 451-60.

Papaioannou, V. E. (2001) 'T-box genes in development: from hydra to humans.', *Int Rev Cytol* 207: 1-70.

Rose, K. L., Winfrey, V. P., Hoffman, L. H., Hall, D. H., Furuta, T. and Greenstein, D. (1997) 'The POU gene ceh-18 promotes gonadal sheath cell differentiation and function required for meiotic maturation and ovulation in Caenorhabditis elegans', *Dev Biol* 192(1): 59-77.

Roy Chowdhuri, S., Crum, T., Woollard, A., Aslam, S. and Okkema, P. G. (2006) 'The T-box factor TBX-2 and the SUMO conjugating enzyme UBC-9 are required for ABa-derived pharyngeal muscle in C. elegans.', *Dev Biol* 295(2): 664-77.

Schedl, T. and Kimble, J. (1988) 'fog-2, a germ-line-specific sex determination gene required for hermaphrodite spermatogenesis in Caenorhabditis elegans', *Genetics* 119(1): 43-61.

Schneider-Maunoury, S., Topilko, P., Seitandou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C. and Charnay, P. (1993) 'Disruption of Krox-20 results in alteration of rhombomeres 3 and 5 in the developing hindbrain', *Cell* 75(6): 1199-214.

Seydoux, G. and Fire, A. (1995) 'Whole-mount in situ hybridization for the detection of RNA in Caenorhabditis elegans embryos.', *Methods Cell Biol* 48: 323-37.

Shiio, Y. and Eisenman, R. N. (2003) 'Histone sumoylation is associated with transcriptional repression.', *Proc Natl Acad Sci U S A* 100(23): 13225-30.

Shirayama, M., Soto, M. C., Ishidate, T., Kim, S., Nakamura, K., Bei, Y., van den Heuvel, S. and Mello, C. C. (2006) 'The Conserved Kinases CDK-1, GSK-3, KIN-19, and MBK-2 Promote OMA-1 Destruction to Regulate the Oocyte-to-Embryo Transition in C. elegans', *Curr Biol* 16(1): 47-55.

Showell, C., Binder, O. and Conlon, F. L. (2004) 'T-box genes in early embryogenesis.', *Dev Dyn* 229(1): 201-18.

CITED LITERATURE (continued)

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H. and Fire, A. (2001) 'On the role of RNA amplification in dsRNA-triggered gene silencing', *Cell* 107(4): 465-76.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J. and Plasterk, R. H. (2002) 'Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi', *Curr Biol* 12(15): 1317-9.

Singhvi, A., Frank, C. A. and Garriga, G. (2008) 'The T-box gene tbx-2, the homeobox gene egl-5 and the asymmetric cell division gene ham-1 specify neural fate in the HSN/PHB lineage.', *Genetics* 179(2): 887-98.

Singson, A. (2001) 'Every sperm is sacred: fertilization in Caenorhabditis elegans', *Dev Biol* 230(2): 101-9.

Smit, R. B., Schnabel, R. and Gaudet, J. (2008) 'The HLH-6 transcription factor regulates C. elegans pharyngeal gland development and function.', *PLoS Genet* 4(10): e1000222.

Smith, J. C., Price, B. M., Green, J. B., Weigel, D. and Herrmann, B. G. (1991) 'Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response to mesoderm induction.', *Cell* 67(1): 79-87.

Smith, P. A. and Mango, S. E. (2007) 'Role of T-box gene tbx-2 for anterior foregut muscle development in C. elegans.', *Dev Biol* 302(1): 25-39.

Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A. et al. (2003) 'The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics.', *PLoS Biol* 1(2): E45.

Stitzel, M. L., Pellettieri, J. and Seydoux, G. (2006) 'The C. elegans DYRK Kinase MBK-2 Marks Oocyte Proteins for Degradation in Response to Meiotic Maturation', *Curr Biol* 16(1): 56-62.

Strome, S. (1986) 'Fluorescence visualization of the distribution of microfilaments in gonads and early embryos of the nematode Caenorhabditis elegans.', *J Cell Biol* 103(6 Pt 1): 2241-52.

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

CITED LITERATURE (continued)

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

Swirnoff, A. H. and Milbrandt, J. (1995) 'DNA-binding specificity of NGFI-A and related zinc finger transcription factors', *Mol Cell Biol* 15(4): 2275-87.

Tabara, H., Grishok, A. and Mello, C. C. (1998) 'RNAi in C. elegans: soaking in the genome sequence.', *Science* 282(5388): 430-1.

Tada, M., Casey, E. S., Fairclough, L. and Smith, J. C. (1998) 'Bix1, a direct target of Xenopus T-box genes, causes formation of ventral mesoderm and endoderm.', *Development* 125(20): 3997-4006.

Takeuchi, J. K., Koshiba-Takeuchi, K., Suzuki, T., Kamimura, M., Ogura, K. and Ogura, T. (2003) 'Tbx5 and Tbx4 trigger limb initiation through activation of the Wnt/Fgf signaling cascade.', *Development* 130(12): 2729-39.

Thatcher, J. D., Haun, C. and Okkema, P. G. (1999) 'The DAF-3 Smad binds DNA and represses gene expression in the Caenorhabditis elegans pharynx.', *Development* 126(1): 97-107.

Timmons, L. and Fire, A. (1998) 'Specific interference by ingested dsRNA.', *Nature* 395(6705): 854.

Topilko, P., Schneider-Maunoury, S., Levi, G., Trembleau, A., Gourdji, D., Driancourt, M. A., Rao, C. V. and Charnay, P. (1998) 'Multiple pituitary and ovarian defects in Krox-24 (NGFI-A, Egr-1)-targeted mice', *Mol Endocrinol* 12(1): 107-22.

Tourtellotte, W. G., Keller-Peck, C., Milbrandt, J. and Kucera, J. (2001) 'The transcription factor Egr3 modulates sensory axon-myotube interactions during muscle spindle morphogenesis', *Dev Biol* 232(2): 388-99.

Tourtellotte, W. G. and Milbrandt, J. (1998) 'Sensory ataxia and muscle spindle agenesis in mice lacking the transcription factor Egr3', *Nat Genet* 20(1): 87-91.

Tourtellotte, W. G., Nagarajan, R., Auyeung, A., Mueller, C. and Milbrandt, J. (1999) 'Infertility associated with incomplete spermatogenic arrest and oligozoospermia in Egr4-deficient mice', *Development* 126(22): 5061-71.

CITED LITERATURE (continued)

Tourtellotte, W. G., Nagarajan, R., Bartke, A. and Milbrandt, J. (2000) 'Functional compensation by Egr4 in Egr1-dependent luteinizing hormone regulation and Leydig cell steroidogenesis', *Mol Cell Biol* 20(14): 5261-8.

Unoki, M. and Nakamura, Y. (2001) 'Growth-suppressive effects of BPOZ and EGR2, two genes involved in the PTEN signaling pathway', *Oncogene* 20(33): 4457-65.

Vilimas, T., Abraham, A. and Okkema, P. G. (2004) 'An early pharyngeal muscle enhancer from the Caenorhabditis elegans ceh-22 gene is targeted by the Forkhead factor PHA-4.', *Dev Biol* 266(2): 388-98.

Virolle, T., Krones-Herzig, A., Baron, V., De Gregorio, G., Adamson, E. D. and Mercola, D. (2003) 'Egr1 promotes growth and survival of prostate cancer cells. Identification of novel Egr1 target genes', *J Biol Chem* 278(14): 11802-10.

Ward, S. and Carrel, J. S. (1979) 'Fertilization and sperm competition in the nematode Caenorhabditis elegans.', *Dev Biol* 73(2): 304-21.

Wardle, F. C. and Papaioannou, V. E. (2008) 'Teasing out T-box targets in early mesoderm.', *Curr Opin Genet Dev* 18(5): 418-25.

Warner, L. E., Mancias, P., Butler, I. J., McDonald, C. M., Keppen, L., Koob, K. G. and Lupski, J. R. (1998) 'Mutations in the early growth response 2 (EGR2) gene are associated with hereditary myelinopathies', *Nat Genet* 18(4): 382-4.

Watson, M. A. and Milbrandt, J. (1990) 'Expression of the nerve growth factor-regulated NGFI-A and NGFI-B genes in the developing rat', *Development* 110(1): 173-83.

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

White, Southgate, Thomson and Brenner (1986) The Structure of the Nervous System of the Nematode Caenorhabditis elegans, vol. 314: Phil. Trans. R. Soc. Lond. B.

White, S., Thomson and Brenner (1986) The Structure of the Nervous System of the Nematode Caenorhabditis elegans, vol. 314: Phil. Trans. R. Soc. Lond. B.

Whitten, S. J. and Miller, M. A. (2007) 'The role of gap junctions in Caenorhabditis elegans oocyte maturation and fertilization', *Dev Biol* 301(2): 432-46.

CITED LITERATURE (continued)

Wightman, B., Ha, I. and Ruvkun, G. (1993) 'Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.', *Cell* 75(5): 855-62.

Wilson, V. and Conlon, F. L. (2002) 'The T-box family.', *Genome Biol* 3(6): REVIEWS3008.

Wolfe, S. A., Nekludova, L. and Pabo, C. O. (2000) 'DNA recognition by Cys2His2 zinc finger proteins', *Annu Rev Biophys Biomol Struct* 29: 183-212.

Yin, X., Gower, N. J., Baylis, H. A. and Strange, K. (2004) 'Inositol 1,4,5-trisphosphate signaling regulates rhythmic contractile activity of myoepithelial sheath cells in Caenorhabditis elegans', *Mol Biol Cell* 15(8): 3938-49.

Yochem, J. and Herman, R. K. (2003) 'Investigating C. elegans development through mosaic analysis', *Development* 130(20): 4761-8.

Yochem, J., Sundaram, M. and Bucher, E. A. (2000) 'Mosaic analysis in Caenorhabditis elegans.', *Methods Mol Biol* 135: 447-62.

Zaragoza, M. V., Lewis, L. E., Sun, G., Wang, E., Li, L., Said-Salman, I., Feucht, L. and Huang, T. (2004) 'Identification of the TBX5 transactivating domain and the nuclear localization signal.', *Gene* 330: 9-18.

VITA

NAME: Lynn Clary

EDUCATION: PhD, Department of Biological Sciences, University of Illinois, Chicago, 2011

B.S., Animal Science, University of Illinois, Urbana-Champaign, 2001

TEACHING: Teaching Assistant, Microbiology Laboratory, Department of Biological Sciences, University of Illinois at Chicago (2004-2006)

Teaching Assistant, Genetics Laboratory, Department of Biological Sciences, University of Illinois at Chicago (2009)

HONORS: Award for Research Achievement, 2011

Travel Fellowship, Cincinnati Midwest Regional Developmental Biology Meeting, 2008

Travel Fellowship, Santa Cruz Developmental Biology Meeting, 2008

Teaching award Microbiology Laboratory, University of Illinois at Chicago, 2006

PRESENTATIONS:

"The EGR-family gene *egrh-1* functions to inhibit oocyte production, maturation and ovulation in *C. elegans*" (selected talk) Cincinnati Midwest Regional Developmental Biology Meeting, 2010

"Identifying targets of the C. elegans T-box transcription factor TBX-2" (poster), Cincinnati Midwest Regional Developmental Biology Meeting, 2010 "Hunting for targets of TBX-2 in the developing pharynx" (poster), 17th International *C. elegans* Meeting, 2009

"C. elegans oocyte maturation, ovulation and sperm recruitment are regulated by the EGR family protein EGRH-1" (poster), Santa Cruz Developmental Biology Meeting, 2008

"The C. elegans EGR family protein EGRH-1 regulates ovulation, oocyte maturation and sperm recruitment" (poster), 46th Annual Midwest

VITA (continued)

Developmental Biology Meeting and (poster)16th International *C. elegans* Meeting, 2008

"A fistful of zinc-fingers" (poster), *C. elegans* Development and Evolutions Topic Meeting, 2006

PUBLICATIONS:

Clary, L. M. and Okkema, P. G. (2010) The EGR family gene egrh-1 functions non-autonomously in the control of oocyte meiotic maturation and ovulation in *C. elegans*. *Development* 137(18): 3129-3137.

Clary, L. M., Ronan T.J., Milton A.C., Okkema, P. G. Identifying targets of the *C. elegans* T-box transcription factor TBX-2. (in preparation)

EMPLOYMENT:

Research Specialist/Technician, Department of Physiology, Loyola University Medical Center, Maywood, IL. 2003-2004

Associate Toxicologist, Experimur, Michael Reese Hospital, Chicago, IL. 2001-2003