University of Calgary

Transcriptional regulation of the *elt-2* gene in the nematode *Caenorhabditis elegans*

by

Janette Berg

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

CALGARY, ALBERTA
JULY, 2006

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Transcriptional regulation of the *elt-2* gene in the nematode *Caenorhabditis elegans*" submitted by Janette Berg in partial fulfillment of the requirements for the degree of Master of Science.

_____

Supervisor, Dr. J.D. McGhee
Department of Biochemistry and Molecular Biology

_____

Dr. Paul Mains
Department of Biochemistry and Molecular Biology

_____

Dr. William Brook
Department of Biochemistry and Molecular Biology

_____

Dr. Frans van der Hoorn
Department of Medical Biochemistry

_____

Dr. Dave Hansen
Department of Biology

_____

Date

ABSTRACT

Underlying the process of development is the selective use of genes, a phenomenon known as gene regulation. Genes are regulated at the level of transcription or post-transcriptional. At the level of transcription, genes need to be turned on at the correct time and in the appropriate cell. Once a gene is expressed, the level of transcription needs to be regulated in order to produce the correct amount of protein in the cell. I have chosen to study transcriptional regulation of the *elt-2* gene. *elt-2* codes for a single zinc finger GATA transcription factor responsible for the regulation of intestinal gene expression in the nematode *Caenorhabditis elegans*.

The main focus of this study was to determine experimentally whether the 5' region flanking *elt-2* contained important transcription factor binding sites necessary for *elt-2* regulation; more specifically, whether these sites corresponded to END-1 and END-3 binding sites. END-1 and END-3 are two redundant GATA type transcription factors expressed one cell cycle earlier than *elt-2* and are responsible for the specification of the intestine. Ectopic expression of END-1 throughout the embryo induces ectopic *elt-2* expression suggesting END-1 can initiate *elt-2* expression. Previous to this study, it was not known whether END-1,3 directly initiated *elt-2* expression by binding to the *elt-2* promoter or whether END-1,3 indirectly regulated *elt-2* through another transcription factor. I have located one potential END-1 binding site within the *elt-2* enhancer region. The binding site named GATA site -3820, is highly conserved within *Caenorhabditis* species and can bind END-1 *in vitro*. An expression vector construct containing a mutated GATA site -3820 within the full length *elt-2* promoter produces weaker embryonic *elt-2* expression than a control vector that contains the wild type site. I purpose that END-1,3 binds primarily to GATA site -3820 to initiate *elt-2* expression. I also purpose that in the absence of GATA site -3820, END-1 and END-3 are capable of binding to one other site within the *elt-2* promoter; albeit with less affinity, to regulate the expression of *elt-2* at early stages of development.

Another focus of the study was to determine whether any potential ELT-2 binding sites were important for the expression of *elt-2* at later stages of development. END-1 and END-3 are transiently expressed in the intestine and therefore are not responsible for maintaining *elt-2* expression in later stages of development (late gastrulation to adult

stage). Previous studies suggested that *elt-2* maintains its own expression through autoregulation; however, the specific site(s) ELT-2 bound to were unknown. We have experimentally determined that ELT-2 can bind *in vitro* to two sites with the *elt-2* enhancer, GATA site -3820 and GATA site -3394, and that both sites are necessary for *elt-2* expression.

The current model is that shortly after END-1,3 are expressed, either END-1 or END-3 binds to GATA site -3820 to initiate the expression of *elt-2*. Once sufficient levels of ELT-2 protein are present in the cell, ELT-2 binds both GATA site -3820 and -3394 to maintain its expression in embryogenesis.

Table of Contents

Page

Page

List of Tables

Page

List of Figures

Page

Page

Page

Page

List of Abbreviations

AA	amino acid(s)
bp	base pair
cDNA	Complementary deoxyribonucleic acid
DAPI	4,6-diamindino-2-phenyl-indole
DNA	deoxyribonucleic acid
dsRNA	double stranded RNA
GCS	γ-Glutamylcysteine synthetase
GSH	glutathione
GFP	green fluorescent protein
Gob	gut obstructed
HSE	heat chock element
NGM	nematode growth media
PCR	polymerase chain reaction
RNA	ribonucleic acid
RNAi	double stranded-RNA mediated interference
SDS	sodium dodecyl sulfate
UTR	untranslated region
2xYT	two times yeast tryptone media

CHAPTER ONE: INTRODUCTION

Introduction

The GATA family of transcription factors are critical in both vertebrates and invertebrates in specifying key organs in the body. Despite their important role in development not much is known about how these factors are regulated in higher organisms; due in part, to their expression in multiple organs as well as the complexity of their regulatory regions. Fortunately, the role of GATA factors in regulating early endoderm development has been conserved through evolution. In *Caenorhabditis elegans* GATA factors have been confirmed to have a role in regulating endoderm development. The GATA transcription factor *elt-2* is present in *C. elegans* and has a central role in the differentiation of the intestine. ELT-2 is the only non redundant GATA factor present in the endoderm and is critical for the worms survival. Due to the lack of knowledge concerning to the regulation of GATA factors, I have chosen to study the regulation of *elt-2,* in hopes that this study may in give insight into how GATA factors in general are regulated.

The *C. elegans* intestine is the ideal experimental system in which to decipher how GATA factors may be regulated. The entire intestine derives clonally from a single cell, called E, that is set aside early in development. At completion of development, the functional intestine contains only 20 cells (Sulston *et al.,* 1983). Many key regulators of *C. elegans* intestinal development are known and derive from two families of transcription factors, forkhead and GATA factors. Many of these families are present in higher organisms and function through a transcription factor cascade to develop the intestine.

In a transcription factor cascade each transcription factor is responsible for initiating its own set of genes as well as to initiate expression of the next transcription factor in the series. In *C. elegans*, ELT-2 is a vital player in the transcription factor cascade, since it lies at a central developmental transition point where the intestine has been specified and now needs to differentiate. My first goal was to determine whether the transcription factors END-1 and END-3 which are responsible for intestinal specification,

directly bind the *elt-2* promoter to initiate its expression or whether they indirectly regulate *elt-2* expression through another factor. My second goal was to determine the sites on the *elt-2* promoter used by ELT-2 for autoregulation. The introduction will include a review of our current knowledge of *elt-2* and the intestinal transcription factor cascade in *C. elegans*. I will also highlight the benefits of using *C. elegans* as a model system and give a simple overview on basic elements of transcriptional regulation.

Basic transcriptional regulation

Although genes code for an array of different proteins that can have numerous functions in the cell, the basic arrangement of a gene and its regulatory regions have been generally conserved in eukaryotes. The basic transcriptional unit includes the coding region and the regulatory DNA sequences that direct transcription. The regulatory DNA sequence usually includes the basal and regulatory promoter, enhancer and boundary elements (Figure 1). The basal promoter is closest to the transcription start site and functions to bind and assemble the pre-initiation complex containing the RNA polymerase II (Goldberg, 1979; Pelham, 1982; Chandler *et al.,* 1983). The main purpose of the basal promoter is to direct the RNA polymerase II enzyme to transcribe from the correct start site and in the correct orientation (Breathnach and Chambon, 1981). In general, the basal promoter does not function to initiate transcription but rather to respond to transcription factors recruited by the regulatory promoter or enhancer region.

Regulatory promoters are generally located closer to the basal promoter than enhancers which are located a larger distance away from the transcription start site (reviews see Hatzopoulos *et al.,* 1988; Muller *et al.,* 1988). The location of an enhancer can vary; it can be located upstream, downstream or internally to the gene (Garcia *et al.,* 1986). Enhancers, unlike regulatory promoters, can work in either orientation relative to the start site of transcription. Enhancers and regulatory promoters contain an assortment of binding sites for transcription factors. It is the assortment of binding sites within the enhancer or regulatory region that in turn determines the expression of gene (reviews see Merika and Thanos, 2001; Strahl, 2001).

There are two main models for how enhancers and regulatory regions function to initiate gene expression (Figure 2). The first model, known as the looping model,

suggests that DNA between the enhancer and the basal promoter loops out to allow for a direct interaction between the enhancer and basal promoter. It is thought that transcription factors located on enhancers or regulatory promoters interact with primed basal transcription machinery at the basal promoter to initiate transcription (Tolhuis *et al.,* 2002; Spilianakis and Flavell 2004).

The second model, known as the tracking model, suggests that the enhancer is used as a platform for loading transcription factors and cofactors. This complex then tracks towards the promoter where it encounters the basal transcriptional machinery. The complete transcriptional apparatus is formed and transcription of the gene is initiated (Blackwood and Kadonaga, 1998; Tuan *et al.,* 1992) DNA-tracking proteins.

To prevent enhancers or regulatory promoters from inadvertently interacting with adjacent genes, boundary or insulator elements flank individual genes within a regulatory region (Bell *et al.,* 2001; Labrador and Corces 2002; Kuhn and Geyer, 2003).

GATA factors and the regulation of intestinal development

Cells within an embryo will give rise to the three germ layers of the organism: the ectoderm, the mesoderm and the endoderm. The ectoderm will give rise to the skin and nervous system; the mesoderm will give rise to muscle, bone and blood; the endoderm will produce the respiratory and digestive tracts. Development of the endoderm is regulated principally by two families of transcription factors; namely forkhead and GATA transcription factor families. The GATA transcription factor family is composed of proteins with a characteristic two zinc-finger motif and bind to the DNA consensus sequence WGATAR (Evans and Felsenfeld, 1989; Orkin, 1992; Tsai *et al.,* 1989).

In general, three GATA factors, GATA-4/5/6, are expressed in the developing and adult gut of vertebrates (Arceci *et al.,* 1993; Laverriere, 1994). Overexpression and deficiency studies of these GATA factors illustrate their conserved role in endoderm development. For example in mice, overexpression of GATA-4 or GATA-6 can induce embryonic stem cells to differentiate into extra-embryonic endoderm (Fujikura *et al.,* 2002). Mice mutant for GATA-4 or 6 die in embryogenesis signifying the necessity of these factors early in development (Kuo *et al.,* 1997). In zebrafish, GATA-5 is expressed early in development, in mesendodermal cells and is required for endoderm formation

(Soudais *et al.,* 1995; Kuo *et al.,* 1997; Molkentin *et al.,* 1997; Reiter *et al.,* 1999). Overexpression of GATA-5 in zebrafish expands the endoderm at the expense of the mesoderm while morpholino knock-downs of GATA-5 block endoderm formation (Reiter *el al.,* 2001; Davidson *et al.,* 2002). In *Xenopus*, cells from the vegetal pole develop into the endoderm and express GATA-4,5,6 (Afouda *et al.,* 2005). The dominant role of GATA factors in *Xenopus* endoderm development was illustrated by the misexpression of GATA factors in the presumptive ectoderm induced early and late endoderm markers (Weber *et al.,* 2000; Afouda *et al.,* 2005). Overall, it is evident that GATA factors are present in the vertebrate endoderm and have a conserved role in endoderm development; however, the mechanism by which GATA factors regulate intestinal development are less understood than in invertebrates.

Invertebrates use numerous GATA transcription factors to regulate the development of the intestine. For example in *Drosophila* three GATA factors are expressed in the gut. (Reuter, 1994; Rehorn *et al.,* 1996). *Serpent*, the first GATA transcription factor to be expressed in the gut, is only transiently expressed in the midgut primordia and is downregulated before terminal differentiation occurs (Reuter 1994; Rehorn *et al.,* 1996). Overexpression of *serpent* induces endodermal gene expression, whereas loss of *serpent* activity causes prospective endodermal cells to transform into ectodermal foregut and hindgut cells (Okumura, 2005; Reuter, 1994). *dGATAc/grain* is also expressed in the *Drosophila* endoderm as well as other regions of the body (Lin *et al.,* 1995). However, in overexpression and deficiency studies, *dGATAc* does not affect morphology or endodermal gene expression suggesting that *dGATAc* does not play a critical role in intestinal development (Brown *et al.,* 2000). Instead, the GATA factor *dGATAe* may be responsible for intestinal differentiation. *dGATAe* expression is gut specific and is expressed both early (at a time when *serpent* expression is declining) and later in development (maintained throughout the life of the fly) (Murakani *et al.,* 2005). Overexpression of *dGATAe* can induce ectopic expression of midgut specific genes suggesting that it may be responsible for intestinal differentiation (Okumura *et al.,* 2005). Overexpression of *serpent* induces *dgatae*, suggesting that in *Drosophila*, these two genes regulate the specification and differentiation of the intestine (Murakani *et al.,* 2005). In *C. elegans,* a series of GATA factors are also used to specify the endoderm.

Two GATA factors, *med-1,2* are expressed in the mesendoderm progenitor cell. *med-1* has been shown to induce the expression of GATA factors *end-1* and *end-3*, which act to redundantly specify the endoderm (Maduro *et al.,* 2001). However, the majority of mutants for *med-1,2* have an intestine suggesting that other factors like *skn-1* or *pop-1* may directly initiate the expression of *end-1* and *end-3* (Goszczynski and McGhee, 2005). Ectopic expression of *end-1* induces *elt-2* expression outside of the presumptive intestine, indicating that the role of *end-1,3* are to initiate *elt-2* expression (Zhu *et al.,* 1998). *elt-2* is the final GATA factor in the series and is responsible for intestinal differentiation. In the intestine, GATA factors *elt-4* and *elt-7* are also expressed; however, mutations within these genes cause no visible defect (Fukushige *et al.,* 2003; Maduro and Rothman, 2002).

Invertebrate GATA factors can substitute for vertebrate GATA factors in both specification and differentiation of the intestine suggesting that some of their molecular mechanism has been conserved. For example, GATA factor END-1 responsible for the specification of the *C. elegans* intestine induces endodermal differentiation in *Xenopus* explants (Shoichet *et al.,* 2000). Also, *Drosophila* GATA factors *serpent* and *dGATAe* induce endodermal marker genes in the *Xenopus* ectoderm (Marakami *et al.,* 2005). The ability of invertebrate GATA factors to induce endodermal markers in higher organisms suggests that the role of GATA factors in intestinal development have been conserved. This conservation between GATA factors suggests that determining how GATA factors function in invertebrates may give insight into their roles in intestinal development in higher organisms.

Caenorhabditis elegans **as model system**

Caenorhabditis elegans is a free-living nematode species found in the soil (Wood, 1988). The organism has a relatively simple body plan with the adult worm containing fewer than 1000 cells (Sulston *et al.,* 1983). The worm is transparent, which allows for visualizing cells through-out development, including cell divisions, cell migrations and cell deaths (Sulston *et al.,* 1983). The lineage for each cell has been determined and is invariant. For example, the most anterior daughter cell of the intestinal precursor cell E is known always to give rise to the most anterior cell in the intestine.

A natural population of *C. elegans* contains two sexes hermaphrodites and males. A hermaphrodite makes both sperm and oocytes; with the brood size limited to the amount of sperm made. However, while a self-fertilized hermaphrodite produces about 280 progeny, a mated hermaphrodite can produce more than 1000 progeny.

Embryogenesis, from fertilization to hatching, is completed in 14 hours at 20°C and is composed of 8 landmark developmental stages (Figure 3). The first landmark is the generation of the founder cells, cells that will give rise to specific cell types in the worm. The second mark, gastrulation, occurs when intestinal precursor cells ingress into the interior of the embryo. The next four landmarks, comma, 1 ½-fold, movement, and 2-fold all occur in a total of about 50 minutes. The three-fold stage, also known as the pretzel stage, occurs 520 minutes after fertilization. Hatching is the final stage of embryogenesis occurring approximately 800 minutes post fertilization. After hatching, worms progress through four larval stages (L1-4) to become adults. However, under starving or overcrowding conditions, L3 worms may progress through an alternate dauer pathway that allows worms to survive for months without food (Cassada and Russell, 1975). Worms resume normal development with the re-introduction of food.

C. elegans has a relatively simple anatomy (reviewed White, 1988) (Figure 4). The outer layer is composed of an extracellular cuticle secreted by the hypodermis. Four muscle quadrants are anchored to the cuticle and are used for locomotion. The nervous system is composed of 302 neurons and consists of the nerve ring found close to the pharynx, and the dorsal and ventral nerve cord. The pharynx is located at the anterior of the animal and is responsible for the grinding and pumping of food through the pharyngeal-intestinal valve into the intestine (Figure 5). The intestine is located centrally and is responsible for nutrient uptake. The reproductive organ in the hermaphrodite is composed of two U-shaped arms that extend in opposite directions from the middle of the animal. Each arm is an independent reproductive system and consists of both female and male organs. In males, the male reproductive system is contained in one L-shaped arm that extends from the middle to the anterior and then posterior region of the animal.

Another benefit of using this model organism is the wealth of information collected and stored concerning; the genetics, genomics, biology of *C. elegans* and the accessibility of this information through Wormbase. The complete *C. elegans* genome

has been annotated and is accessible through Wormbase (http://www.wormbase.org/). Wormbase allows for easy access to cosmid libraries, expression libraries, worm strains, and RNAi (RNA interference) libraries. In the absence of an available mutant allele the RNAi library has been an effective resource to silence gene expression and investigate gene function. Also available is the Fire lab *C. elegans* expression kit that contains 288 expression vectors with GFP, YFP and *lacZ* that can be used to determine the expression pattern of a regulatory region or gene of interest. Transformation of these vectors in *C. elegans* is relatively simple compared to efficiency of transformations in other model systems.

C. elegans **embryogenesis**

In *C. elegans*, the fertilized egg (called Po) contains the necessary nutrients and complete gene expression machinery to produce viable offspring. The first cell division in embryogenesis is the division of Po to produce two daughter cells AB and P1 (Figure 6). The AB cell, or blastomere, undergoes numerous cell divisions to give rise to pharynx, hypodermal and nervous system. The P1 cell after two rounds of division produces the C, P3, MS and E cells, each of which will go on to produce distinct cell types. The C blastomere will give rise to muscle and hypodermal cells. P3 will give rise to the germ line and muscle cells. The MS blastomere will give rise to muscle and pharynx cells. The E blastomere will clonally produce the intestine.

The ability of the E blastomere to clonally derive the intestine is accomplished using both internal cues (signaling between intestinal cells) and external cues (outside of the intestinal lineage) as well as using both maternal and zygotic expressed genes. SKN-1 is a maternally provided transcription factor. SKN-1 is important for endoderm development since eighty percent of embryos lacking *skn-1* do not produce an endoderm (Bowerman *et al.,* 1993). While *skn-1* mRNA is present in both AB and P1, *skn-1* mRNA is only translated into protein in the P1 descendants (Bowerman *et al.,* 1993; Seydoux and Fire, 1994). The P1 descendants, EMS (mesendoderm) and P2 (germline) both contain SKN-1; however, SKN-1 is unable to induce transcription in P2 due to the presence of the germline transcriptional repressor PIE-1 (Batchelder *et al.,* 1999; Seydoux *et al.,* 1996). Therefore SKN-1 is only active in the EMS cell. Until recently,

SKN-1 was thought to initiate intestinal specification mainly by initiating transcription of two GATA binding transcription factors *med-1* and *med-2* (Maduro *et al.,* 2001). The MEDs activated endoderm specification genes *end-1* and *end-3* in E, while high levels of POP-1 (a member of the TCF/LEF class of HMG box proteins) in MS induced mesodermal fate (Brunner *et al.,* 1997; Korswagen and Clevers, 1999; Maduro, 2001). Although, overexpression experiments have indicated that the *meds* can induce *end* gene expression; the normal role of *med-1,2* in intestinal development is minor since 85% of *med-1,2* minus worms still produce an intestine (Goszczynski and McGhee, 2005). Although it remains to be tested, *end-1,3* may be direct targets of SKN-1, since numerous SKN-1 binding sites are located within their promoters.

The role of the GATA factors, END-1 and END-3, in specification of the endoderm was shown when a chromosomal deletion containing *end-1* and *end-3* abolished endoderm formation (Zhu *et al.,* 1997). Moreover, ectopic *end-1* or *end-3* expression throughout the embryo caused non-endoderm cells to become endodermal. However, END-1 and END-3 are only expressed early in development (tapers off at the 8E cell stage) suggesting their role is limited to specification. END-1 has been shown to induce the expression of the *elt-2* gene suggesting END-1 and END-3 may be responsible for initiating *elt-2* expression (Zhu *et al.,* 1998). ELT-2 is the final GATA transcription factor within the intestinal transcription factor cascade and is responsible for the proliferation and differentiation of the intestine (Hawkins and McGhee, 1995).

ELT-2 and intestinal differentiation

elt-2 codes for a single zinc finger transcription factor (Hawkins and McGhee, 1995). It is first expressed at the 2E cell stage and continues to be expressed in the intestine for the entire lifespan of the worm (Figure 7). Homozygous *elt-2* null mutants die shortly after hatching with a gut obstructed phenotype and a deformed intestine suggesting that the role of *elt-2* is critical in the maturation of the intestine (Figure 8).

SAGE (serial analysis of gene expression) has been performed on RNA extracted from the adult intestine. The method catalogued more than 4000 genes that are expressed in the intestine as well as their transcript frequency (McGhee *et al.,* in preparation). From the 74 genes with the highest expression in the intestine, 56 contain within their promoter

the sequence AX**TGATAA** suggesting that these genes may be regulated by GATA transcription factors. In the adult intestine there are a total of three expressed GATA factors, ELT-2, ELT-4 and ELT-*7*; however, it is doubtful that ELT-4 and ELT-7 regulate gene expression since *elt-4/elt-7* double mutants have a functional intestine. Although it remains to be tested whether these 56 genes are regulated by ELT-2, previous data showing ELT-2 regulating intestinal gene expression through a TGATAA binding site supports this conclusion (Figure 9) (Marshall and McGhee, 2001; Bossinger *et al.,* 2004; Fukushige *et al.,* 2005).

ELT-2 has been visualized *in vivo* localizing to its own promoter, suggesting *elt-2* autoregulates its expression. Homozygous transgenic worms containing the full length *elt-2* construct (*elt-2* promoter and gene fused to GFP) contain two foci per gut nucleus, suggesting the ELT-2::GFP protein binds directly to its promoter (Figure 10) (Fukushige *et al.,* 1999). However, since there are 12 TGATAA sites within the 5044 base pairs flanking the 5' end of the *elt-2* gene it has not been experimentally determined which site(s) have a role in *elt-2* autoregulation.

Intestinal organogenesis in *C. elegans*

The fully developed *C. elegans* intestine is composed of 20 cells that are arranged to form nine intestinal rings (Figure 11). Each of the nine intestinal rings (int I-IX) have two cells except for int I, which has 4 cells. The cells that make up int II,III,IV are rotated 90º relative to the other cells of the intestine giving these cells a twisted appearance. The gut lumen runs through the midline of the adjoining cells. The apical domain faces the lumen and is composed of a microvillar brush border that is anchored by actin filaments to the subapical terminal web structure (review Fath and Burgess 1999). The terminal web is a collection of intermediate filaments, spectrin, myosin, and actin-binding proteins. The basal surface faces outwards towards an extracellular basement membrane.

The intestine develops from one cell, named E, which goes through a series of divisions to clonally derive the intestine (Figure 12). Intestinal development occurs through a series of cell divisions, movements and polarization events and has been characterized thoroughly by Sulston *et al.,* 1983 and Leung *et al.,* 1999. The first division occurs along the anterior/posterior axis at the surface of the embryo. At gastrulation, the

two intestinal precursor cells migrate into the interior of the embryo; where they divide along the left/right axis. The next two divisions occur in a skewed anterior/posterior axis to form two layers, a ventral row comprised of six cells and a dorsal row of ten cells. At this stage of development, nuclei that are centrally located within each cell move towards the apical side (future lumen), while cytoplasmic components move towards the basal boundary. Cytoplasmic polarization is followed by the formation of gaps between cells, which leads to the formation of the lumen. As the lumen develops, the remaining cells from the dorsal level move into the ventral level; with 7 moving in between 3 and 4, 8 moving behind 5. The first two anterior cells of the intestine subsequently divide along the dorsal/ventral axis to produce the four cells that make up int I. The two last posterior cells will also divide but in an anterior/posterior axis. After the final two cell divisions, cells that form int II move in between int I and III. As the cells are intercalating, a coordinated movement of int II,III,IV create the 90º rotated positions of int II,III,IV compared to the left/right positions of posterior cells. Once morphogenesis is complete, intestinal nuclei are relocated centrally in the cell and cytoplasmic components have a normal distribution. The apical and basal surfaces continue to develop and the complete intestinal primordium elongates along the anterior/posterior axis of the embryo.

Interestingly, cells of the intestinal primordium have an intrinsic ability to produce an apical/basal polarity in the absence of external signals. For example, an isolated E cell produces E descendants that differentiate and have an apical/basal polarity around a central cavity (prospective lumen). The E descendants have the coordinated movement of nuclei towards the apical region. However, E descendants never form a linear symmetrical tube suggesting that external cues are responsible for the normal configuration of the intestine.

elt-2 conservation between _Caenorhabditis species_

There are 18 species of *Caenorhabditis*; of interest to this study are *C. elegans*, *C. briggsae* and *C. remanei* since the genomes of these species have been sequenced. Within the phylogenetic tree, *C. briggsae* and *C. remanei* are more closely related to each other than they are to *C. elegans. C. elegans* diverged from the last common ancestor approximately 80 to 110 million years ago (Cho *et al.,* 2004; Kiontke *et al.,* 2004;

Coghlan and Wolfe, 2002; and Stein *et al.,* 2003). Sequence comparisons between these three *Caenorhabditis* species have been used to locate genomic regions that have been conserved through evolution and reveals a correlation between conservation and function (Coroian *et al.,* 2006). The genomic region surrounding *elt-2* has been sequenced in these three *Caenorhabditis* species, and shows that *elt-2* is present and highly conserved.

In *C. elegans, elt-2* is located in the central region of the X chromosome (position 10481169-10483357) and is given the sequence name C33D3.1 (Figure 13). Between *elt-2* and the upstream gene *elt-4* (C39B10.6) there is 5144bp of non-coding region. *elt-4* like *elt*-2 is expressed in the intestine and may have arisen from an *elt-2* gene duplication event (Fukushige, 2003).

The *elt-2* orthologue in *C. briggsae* is CBG17257 and is also present on the X chromosome (Figure 14). dsRNA corresponding to the *C. briggsae* CBG17257 gene injected into *C. elegans* gives a gut obstructed phenotype resembling the *elt-2* RNAi phenotype suggesting that CBG17257 functions similar to *elt-2*. Protein alignments between CBG17257 and ELT-2 illustrate that the zinc finger motif is 100% identical between the two species while the full protein is 68% identical (Figure 15). The 5' flanking region of CBG17257 specifies intestinal expression similar to the analogous region of *elt-2*, as illustrated by the ability of CBG17257 promoter::GFP construct to direct intestinal specific expression in *C. elegans* (Figure 16). However, the 5' flanking region of CBG17257 is not completely comparable to the 5' flanking region of *elt-2.* The *C. briggsae* genome does not contain the *elt-4* gene. The divergence of *C. elegans* from the last common ancestor occurred before the *elt-2* duplication and therefore *elt-4* is not present in *C. briggsae* nor *C. remanei* (see below). Instead, in *C. briggsae* the gene CBG17255 is located 17Kb upstream and is most likely the orthologue of the G-protein coupled receptor gene located upstream of *elt-4* in *C. elegans*.

The *elt-2* orthologue in *C. remanei* is most likely cr01.sctg11.wum.290.1 and is also present on the X chromosome (Figure 17). The zinc finger motif is 94% identical between the two species while the full protein is 72% identical (Figure 15). Upstream of cr01.sctg11.wum.290.1 is the proposed gene cr01.sctg11.wum.291.1 containing sequence similarity to the *elt-2* enhancer. Further upstream is cr01.sctg11.wum.292.1 and is the

orthologue of the G-protein coupled receptor gene located upstream in both *C. elegans* and *C. briggsae*.

The plan

To locate sites important in the regulation of *elt-2*, numerous techniques will be used. Interspecies sequence comparison of the *elt-2* 5' flanking region will be used to locate conserved sequences. These conserved sequences may contain *elt-2* regulatory sites and will be analyzed for similarity to known transcription factor binding sites. These conserved binding sites will be experimentally tested to determine if indeed such transcription factors can bind *in vitro*. Ultimately, each site will be tested for its *in vivo* function in regulating *elt-2*; which, will be accomplished by analyzing GFP/β-galactosidase expression from transgenic worms with reporter constructs containing each potentially important site mutated within the full length *elt-2* promoter.

CHAPTER TWO: Materials and Methods

General *C. elegans* maintenance and storage

C. elegans strains were kept in conditions described by Brenner (1974). The Bristol N2 strain was used as the Wildtype strain. Worms were maintained on nematode growth medium (NGM) plates that were seeded with the E. coli strain OP50. Worms were kept at 16ºC. Strains were frozen indefinitely at -80ºC by rinsing recently cleared plates with M9 (42.2 mM Na_2HPO4, 22 mM KH_2PO4, 85.6 mM NaCl and 1.0mM $MgSO_4$) and adding equal volume of freezing solution (100 mM NaCl, 50 mM KH_2PO_4, 300 mM glycerol, 0.3 mM $MgSO_4$ (Sulston and Brenner, 1974). Worms were placed into Nunc cryogenic tubes and placed into a Styrofoam freezing box. The box was placed at -80ºC for two days. Vials were then removed from Stryrofoam box and placed into standard lab boxes previously cooled to -80ºC.

Small scale purification of plasmid DNA

JM109 competent cells were used for plasmis transformations unless otherwise noted. Ten mL 2XYT cultures were grown overnight with 10μL of 50 mg/mL Ampicillin. The plasmid was then extracted from pelleted cells using the Wizard Plus Minipreps DNA Purification system from Promega. Plasmids were normally eluted into water unless otherwise specified. Plasmid concentration was calculated using a Nanodrop ND-1000 Spectophotometer (Nanodrop technologies).

Microinjections

Expression constructs were injected at 50 μg/mL with the plasmid pRF4 (50 μg/mL) unless otherwise indicated. pRF4 (*rol-6*(su1006)) codes for a cuticle collagen mutant that acts as a dominant transformation marker (Mello *et al.*, 1991). Purified salmon sperm DNA was used when concentrations of the expression construct were less than 50 μg/mL. DNA was injected into the syncytial gonad arms of mid-staged adult worms. Injected hermaphrodites were then placed onto plates (two worms per plate) and allowed to recover at room temperature. Rolling F2 progeny were selected and placed onto individual plates to establish transgenic lines. For each injected construct, a

minimum of three lines were established. Each line was analyzed at three stages of embryogenesis; early (during gastrulation), mid stage (comma), and late (pretzel). Transgenic lines were also analyzed after embryogenesis; shortly after hatching (L1) and as older adults.

β-galactosidase staining

Embryos were cut out from adult hermaphrodites and transferred into a well slide that contained 180 µL of M9 buffer. Twenty µ of 6% hypochlorite was added to the M9 solution and mixed with a pipette. Embryos were then quickly moved to a new well that contained 200 µL of 1% bovine serum albumin (BSA). Embryos were then transferred to the middle of a gelatin-subbed slide that contained double-sided tape 22mm apart from each other. Excess M9 solution surrounding the embryos was removed (total volume of approximately 10 µL remained). A 22x50 mm cover slip was placed evenly over the double sided tape. Embryos were attached to slides by gently squashing the cover slip. Embryos were washed five times with M9. Embryos were then washed three times with *ges-1* fixative (5 mM NaOH, 0.125M phosphate buffer ((36 mM KH_2PO_4 and 89 mM Na_2HPO_4), 22.5 mg/mL paraformaldehyde, 0.1% glutaraldehyde)). Embryos were permeablized with gentle pressure applied to the top of the coverslip. Slides were kept on ice for three minutes. Embryos were rinsed thoroughly with M9 to remove any trace of fixative. β-galactosidase staining solution was applied and removed five times. β-galactosidase solution is composed of 895 µL of β-galactosidase buffer (36 mM NaH_2PO_4, 186 mM $NaHPO_4$), 0.004% SDS, 0.05% X-GAL (5-bromo-4-chloro-3-indoyl β-galactopyranoside) in DMF solution and 100 µl of Redox solution (5 mM $K_3Fe(CN)_6$ and 5 mM $K_4Fe(CN)_6$). Staining was done at 37ºC overnight. Once stained, three washes of M9 were used to remove excess solution. One wash of M9 containing 0.1 µg/mL DAPI (4,6-diamidino-2-phenylindole) was followed by a final M9 wash. Slides were then sealed with nail polish.

RNAi against open reading frame within the *elt-2* enhancer

Two primers were made to make double stranded RNA corresponding to the open reading frame (ORF) located -4180bp upstream of the *elt-2* gene. A sequence

corresponding to -4367bp to -3420bp upstream of the *elt-2* ATG site was amplified using primers oJBT7-683 and oJBT7-1630. Each primer contains the T7 polymerase initiation sequence at the 5' end of the primer. PCR annealing and extension temperatures were both 72ºC. Extension time was 1 minute. Twelve identical PCR reactions were made and pooled into four tubes and ethanol precipitated. The size of the PCR product was confirmed using a 0.7% agarose gel.. To make double stranded RNA from the PCR product, a modified protocol from Promega was followed. One hundred µL of RNA transcription solution ((water (46 µL), 1x transcription buffer (Promega), 10 mM dithiothreitol, 35 units of RNAguard (Amersham Pharmacia), 0.5 mM of rNTPs, and 40 units of T7 RNA polymerase (Promega)) was added to each of the four tubes. The RNA transcription solution was incubated at 37ºC overnight. After overnight incubation, two units of DNase I (Amersham Pharmacia) were added and the solution was incubated at 37ºC for 15 minutes. The solution was then centrifuged and 200 µL of transcription stop solution (sodium acetate) was added. To purify dsRNA, phenol:chloroform:isoamylalcohol extraction was used (see extraction protocol below). dsRNA was then resuspended into 20 µL of 10 mM phosphate EDTA solution. To verify the product was RNA and not DNA, Pancreatic Ribonuclease was incubated with approximately 600 ng of RNA product for one hour at 37ºC. Ribonuclease digested RNA was run beside undigested RNA on a 0.7% gel to confirm that the product was RNA and not DNA. RNA was injected (1227 ng/µL) into both the syncytial portion of the gonad arm and in the intestine of hermaphrodite N2 and JM73 worms. JM73 worms contain the integrated plasmid pJM86, which has the full length *elt-2* promoter and coding region (minus the last nine C-terminal amino acids of ELT-2) fused to GFP. Hermaphrodites were allowed to recover overnight and then transferred onto individual plates.

RNAi against genes found in RNAi feeding library

dsRNA corresponding to the genes *acn-1*(C42D8.5), *skn-1*(T19E7.2a, T19E7.2b, T19E7.2c), and *elt-2* (C33D3.1) were made using the L4440 RNAi feeding vectors assembled by Kamath and Ahringer (2003). Primers L4440F and L4440R were used in a PCR reaction to amplify the genes while the T7 RNA transcription protocol from Promega was used to make RNA from the PCR products (same as above)**.** dsRNA (1

mg/mL) was injected into hermaphrodites in both the syncytium portion of the gonad arm and in the intestine. Hermaphrodites were allowed to recover overnight and transferred to individual plates.

Purification of Salmon Sperm DNA

A highly polymerized form of DNA from salmon sperm (ICN Biomedical INC) was dissolved in TE solution and sheared using a 26 gauge needle. The 35 mL solution was then placed into a 50 mL Falcon tube along with 25 mg of proteinase K; 5 mL of 10% SDS, and 10 mL of 0.5 EDTA. After an one hour incubation at 55ºC, 5 mL of chloroform was added and the solution was then vortexed for one minute. After a five minute incubation at room temperature the solution was centrifuged for 15 minutes. The supernatant was transferred to a new tube and an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added. The solution was mixed for one minute and then centrifuged for five minutes at 10,000 RPM. Extraction was repeated. After the final extraction, 95% ethanol was added at 2x the original volume of supernatant. The solution was then mixed and stored at room temperature for five minutes. The solution was then centrifuged (10,000 RPM) for 15 minutes at 4ºC. After the supernatant was removed, the pellet was washed with 200 µL of 70% ethanol. After a final centrifugation step, ethanol was removed. The DNA was allowed to air dry and then was resuspended into water to a concentration of 1022 ng/µL.

Antibody staining of ELT-2

To determine ELT-2 levels in heat-shocked worms, N2 worms were placed in a 30ºC incubator overnight. Larval worms were fixed as described by Finney and Ruvkun (1990). Worms were placed in 250 µL of MRWB (160 mM KCl, 40 mM NaCl, 20 mM EGTA, 10 mM Spermidine, 0.3mM PIPES, 50% final volume Methanol). Heads were severed from bodies and 50 µL of 20% paraformaldehyde and 200 µL of water were added. After the mixture was incubated for 35 minutes at 4ºC, worms were washed twice with 500 µL of Tris-Triton buffer (0.1 mM Tris HCl, 0.1% Triton-X-100, 1mM EDTA). Worms were then transferred to a 500 µL solution containing 480 µL of TTB, 15 µL 20% paraformaldehyde, 5 µL of β-mercaptoethanol and kept at 37ºC overnight. Worms were

then immersed in 250 µL of Borate buffer (pH 9.2) (62mg/ml H_3BO_3, 0.5 NaoH) + 0.01% Triton solution for 15 minutes and then resuspended into a 250 µL Borate buffer solution containing 200 mM DTT. After 15 minutes at room temperature, worms were washed with Borate buffer solution and then resuspended into AbB buffer (AbA buffer + 0.2% BSA). Worms were resuspended with the ELT-2 antibody (diluted 1/10 in AbA buffer (1xPBS, 1% BSA, 0.5% Triton X-100, 0.05% sodium azide and 1 mM EDTA)) and incubated overnight at room temperature. Worms were then resuspended in AbB buffer overnight at four degrees Celsius. Worms were rinsed twice with AbA buffer and then labeled using the secondary antibody Cy3-labeled goat anti-rabbit antibody or FITC-labeled donkey anti-mouse IgG at a dilution of 1/1000 in 200 µL. The solution was shaken for two hours and kept from light. Worms were then washed for 15 minutes with AbB buffer containing DAPI and mounted with antifade solution (80-90% glycerol containing antifade solution) (Wood, 1988).

Verification of strain T-413

T-413 is an integrated line made by Tetsunari Fukushige. Although it was known that the strain contained a reporter construct with the *elt-2* enhancer insert, the vector backbone and enhancer sequence was not known.. PCR reactions determined that the construct contains the *elt-2* enhancer from -3876 to-3142 base pairs upstream of the *elt-2* ATG start site inserted into the heat-shock vector pJM76.

Primers oJBhindIII1388 and 96'SmaI were used to amplify by PCR a 682bp sequence. The ability of the primer 96'SmaI to amplify a PCR product revealed that the insert is in the vector pJM76. The amplified DNA product was inserted into pCR2.1-Topo (Invitrogen). Sequencing (used Reverse primer) revealed that the 3' end of the insert is -3145bp upstream of the *elt-2* ATG site. To amplify the 5' end of the insert, primers Reverse and Enhancer1741Pst1I were used. The PCR product was cloned into pGEM-T (Promega) and sequenced with the Primer Sp6. Sequencing indicated that the 5' end of insert is -3859 from the *elt-2* ATG start site.

Constructs for analyzing the *elt-2* promoter

pJM67

pJM67 is the parental plasmid used for most expression constructs made in this study. To make the plasmid the complete *elt-2* 5' flanking region (5048bp) and the first 92bp of the *elt-2* coding region was inserted into the multiple cloning site of pPD96.04 using restriction enzymes EcoRI and BamHI. pPD96.04 is a promoterless vector that contains a multiple cloning site upstream of a nuclear localization sequence (NLS), full length GFP/*lacZ* coding regions (with a total of 15 introns) and an *unc-54* 3' untranslated (UTR) sequence. A complete list of expression vectors including pPD96.04 can be found at,
(ww.ciwemb.edu/pub/FireLabInfo/FireLabVectors/1995_Vector_Kit/Vec95_Docs/Vec95Doc.rtf)

pJM258-pJM244 Constructs of the 3' deletion series

The constructs in the 3' deletion series were made by Lana Wong using pJM67 as the template. Each construct has approximately 500bp of the *elt-2* 5' flanking region shortened from the 3' end of the proceeding deletion construct in the series. Three main sites were used to assemble the nine constructs, the Reverse primer site and the SbfI and NheI restriction sites. The Reverse primer binding site is located 210bps upstream of the multiple cloning site (MCS). The MCS contains the SbfI restriction site. The NheI restriction site is located endogenously 493bp upstream of the *elt-2* ATG start site. The desired promoter sequence was amplified by PCR using the primer Reverse (vector specific) and a promoter specific primer (see end of Materials and Methods section for primer sequences). All promoter specific primers contained an engineered NheI restriction sequence. The promoter specific primer specified the 3' promoter position represented in each construct. The amplified DNA was then cloned into the intermediate subcloning vector pCR2.1-Topo (Invitrogen). Each insert was released from pCR2.1-Topo using SbfI and NheI restriction enzymes. The insert was then ligated into pJM67 also cut with SbfI and NheI. All constructs were injected at 50 μg/mL with 50 μg/mL of transformation marker pRF4(*rol-6*(su1006)). At least three independent transgenic lines

were maintained and analyzed. Expression in the intestine was visualized using GFP and then confirmed by β-galactosidase staining.

Below, Table 1 contains the name of each construct, as well as the 5' and 3' end of the insert. The table also contains the names of the primers used, but not their specific sequences; specific sequences for all primers used in the study are located at the end of the materials and methods section. It is important to note that each PCR product was cloned into pJM67 (cut with NheI and SbfI), and therefore each construct contains 493bp of the *elt-2* basal promoter from the NheI site to the ATG start site.

Table 1: 3' deletion series constructs

Construct Name	Left primer	Right primer	Sub-cloned into pCR2.1-Topo	5' end of *elt-2* promoter insert	3'end of *elt-2* promoter insert
pJM258	Reverse	olw08	pJM257	-5049	**-715**
pJM256	Reverse	olw07	pJM255	-5049	**-1422**
pJM254	Reverse	olw06	pJM253	-5049	**-1786**
pJM252	Reverse	olw05	pJM251	-5049	**-2568**
pJM250	Reverse	olw04	pJM249	-5049	**-3022**
pJM248	Reverse	olw03	pJM247	-5049	**-3575**
pJM246	Reverse	olwo2	pJM245	-5049	**-4323**
pJM244	Reverse	olw01	pJM243	-5049	**-4584**

pJM259

The final construct in the series pJM259 is the parental plasmid pJM67 cut with SbfI and NheI. The cohesive ends were endfilled to produce blunt ends that could then be used to recircularize the plasmid.

pJM284-pJM288 Constucts of the 5' enhancer-specific deletion series

The 5' deletion series is compiled of five constructs that have the enhancer sequentially shortened from the 5' end. The series removes from -4287 to -3205 bp upstream of the *elt-2* ATG site while the 3' enhancer position remains the same (-3022 bp upstream of *elt-2*).

pJM67 was used as the PCR template to amplify each enhancer deletion insert. An annealing temperature of 68ºC and an extension of 1 minute and 10 seconds was used for all PCR reactions. All five 5' primers in the series contain an engineered HindIII site while the 3' primer oJBPstI2025 contains an engineered PstI site. Each PCR product was cloned into the pCR2.1-Topo vector and then released using HindIII/PstI restriction enzymes. The released PCR products were inserted into the HindIII/PstI cleaved vector pJM76. pJM76 is an expression testing vector that contains, downstream of the HindIII/PstI insert site, a basal heatshock promoter, as well as the coding regions for GFP/*lacZ*. All constructs were verified with diagnostic restriction digests, and, in addition each insert was sequenced.

Table 2 contains the name of each construct, the left and right primers designed to PCR amplify each enhancer deletion insert and the size of the PCR product. The table also includes the primers 5' and 3' position with respect to the ATG start site of *elt-2*.

Table 2: 5' enhancer deletion constructs

Construct Name	Left primer	Right primer	PCR product size (bp)	5' end of the *elt-2* enhancer	3' end of the *elt-2* enhancer
pJM284	oJBhindIII760	oJBPstI2025	1285	**-4287**	-3022
pJM285	oJBhindIII1055	oJBPstI2025	995	**-3995**	-3022
pJM286	oJBhindIII1388	oJBPstI2025	659	**-3660**	-3022
pJM287	oJBhindIII1660	oJBPstI2025	391	**-3392**	-3022
pJM288	oJBhindIII1845	oJBPstI2025	205	**-3205**	-3022

pJM289

pJM289 contains 351bp of the enhancer cloned into pJM76. Primers oJBhindIII1388 and Enhancer1741PstI were designed to PCR amplify -3661 to -3310 bp upstream of *elt-2* ATG site. The PCR product was digested with restriction enzymes HindIII and PstI. The product was then inserted into the cleaved HindIII/PstI vector pJM76. Positive clones were selected using Primers Reverse and 96E Sma'. Insert was sequenced using the Reverse primer.

pJM290

pJM290 was designed to contain 605bp of the enhancer cloned into pJM76. Primers HindIII1175 and Enhancer1741PstI were used to amplify by PCR a portion of the enhancer region (-3915 to -3310 upstream of *elt-2* ATG site). The PCR product was digested with restriction enzymes HindIII and PstI. The product was to be inserted into pJM76 (also digested with HindIII and PstI); however, ligation into pJM76 was not completed.

pJM281-pJM283 Conserved Region I, II, III

Sequence comparisons of the 5' flanking region of *elt-2* with the *elt-2* orthologues in *C. remanei* and *C. briggsae* located three regions of high sequence similarity, and were named Conserved Region I, II, III. DNA sequences corresponding to Conserved Regions I, II, III were each placed into pJM76. pJM281 has an insert that corresponds to Conserved Region I (C.R. I). pJM282 has an insert that corresponds to Conserved Region II (C.R. II). pJM283 has an insert that corresponds to Conserved Region III (C.R. III). Sub-cloning vectors pGEM-T (Promega) and pCR2.1 Topo were used. After the inserts were cloned into pJM76 using the HindIII and PstI restriction sites, diagnostic restriction digests and, in addition each insert was sequenced to confirm for the correct insert. Table three contains the names of the left and right primers used to amplify the insert, the size of the PCR product, the sub-cloning vector and the 5' and 3' end of the insert.

Table 3: Conserved Region I, II, III constructs

Construct Name	Left primer	Right primer	PCR product size (bp)	sub-cloning vector	5' end of *elt-2* promoter	3' end of *elt-2* promoter
pJM281 C.R I	IVhindIII	IVPstI	280	pGEM-T	-437	-169
pJM282 C. R. II	IIIhindIII	IIIPstI	589	pGEM-T	-2092	-1531
pJM283 C.R. III	IIhindIII	IIPstI	1052	Topo	-4434	-3421

pJM324

pJM324 is an altered version of pJM67. pJM324 does not contain the first 32 amino acids of ELT-2. To make pJM324, pJM67 was cut with the restriction enzyme XmaI to release the full length *elt-2* promoter and the 98 base pairs that codes for the *elt-2* amino terminus. After ethanol precipitation, the 5'-phosphate groups were removed using Calf intestinal alkaline phosphatase (CIP). After one hour incubation at 37ºC, the *elt-2* insert and vector backbone were separated on a 0.7% agarose gel. The vector backbone (8088bp band) was isolated and then ligated to the PCR product below.

Primers Reverse and XmaIATG were designed to PCR amplify 1123bp of the *elt-2* basal promoter region from construct pJM244 (smallest construct from 3' deletion series). The PCR product was then sub-cloned into pCR2.1-Topo and named pJM312. pJM312 was then digested with restriction enzyme XmaI to release the insert. The insert was ligated to the pJM67 XmaI digested vector backbone from above. JM109 electrocompetent cells were used for transformations. PCR colony screening with primers Reverse and olw01 was used to locate a colony that contained a plasmid with an insert in the correct orientation. The intermediate clone was named pJM323. pJM323 has the *elt-2* coding region removed and contains the *elt-2* basal promoter fused to GFP/*LacZ*.

pJM323 was cut with restriction enzymes SbfI and NheI. After separation on a 0.7% agarose gel, the 8583bp fragment was isolated and ligated to a 4606bp fragment released from a pJM67 NheI/SbfI digestion. The vector was renamed pJM324. pJM324 contains the complete 5' flanking region of *elt-2* fused directly to GFP/*LacZ* coding region. A diagnostic HincII restriction enzyme digest was used to differentiate pJM324 from pJM67 since pJM324 has the 92 bp of *elt-2* coding region removed that contains the HincII restriction digest site.

After transformation and plasmid purification (see small scale purification of plasmid DNA section), Phenol:chloroform:isoamyl alcohol extraction was used to remove any remaining impurities. The enhancer region and basal promoter were sequenced. The plasmid was injected at 15 μg/mL with 35 μg/mL salmon sperm DNA and 50 μg/mL pRF4._Worms transformed with pJM324 were visualized to determine if any differences in expression were seen from worms containing the older version pJM67.

pJM309

pJM309 is the "enhancer construct" and contains the sequence from -4157 to -3310 bp of the *elt-2* 5' flanking region. The enhancer was PCR amplified using primers enhancer893 and enhancer1741. The 875 bp PCR product was separated on an agarose gel and isolated. The DNA was cut with SbfI and NheI and ligated to SbfI/NheI cleaved pJM67. SspI restriction enzyme was used for a diagnostic restriction digest. The insert was then sequenced using the primer Reverse.

pJM308 GATA Site -3820 Construct

Site directed mutagenesis was performed first by using an overlapping splicing technique. For site directed mutagenesis against GATA site -3820 primers; 3882B, Reverse, 3882A and olwo5 were used for the primary round of PCR. Primer 3882B and Reverse were used to amplify by PCR a 1483 bp product from pJM67. Primers 3882A and olwo5 were used to amplify by PCR a 1283 bp product from pJM67. These two primary products were mixed together and boiled for two minutes and placed on ice for two minutes. For the secondary round of PCR the primers Reverse and olwo5 were used. The PCR product was cloned into pGEM-T and designated as plasmid pJM307. pJM307 was then cut with SbfI and NheI to release the insert. The insert was then cloned into a cleaved SbfI/NheI pJM67vector, and named pJM308. For site directed mutagenesis against GATA site -3394 primers; 33394B, Reverse, olwo5, 3394B were used. The same protocol for site directed mutagenesis against GATA site -3820 was used. Site directed mutagenesis against GATA site -3820 produced unwanted base pair changes within the enhancer. Therefore site directed mutagenesis against GATA site -3820 was redone using Stratagene QuickChange XL Site-Directed Mutagenesis Kit (see section below).

pJM310-312 Site directed mutagenesis on pJM309

Stratagene QuickChange XL Site-Directed Mutagenesis Kit was used to alter; site 1360, site 1601, and GATA site -3394 with the recommended protocol being precisely followed. The enhancer construct pJM309 was used as the template for all three constructs. Various concentrations of pJM309 were used (5ng, 30ng and 50ng) with 50ng producing optimal results. For each construct, PCR was cycled 16 times with their

respective primers (see table below). A 4.5 hour DpnI digestion was used to remove template DNA. Five μL of DpnI treated DNA were used to transform XL1-Blue supercompetent cells. Positive clones were selected using primers En893 and olwo3. After transformation and plasmid purification (see small scale purification of plasmid DNA section), the restriction enzyme DdeI was used to ensure no vector rearrangements had occurred. The restriction enzyme SalI was used to confirm that GATA -3394 was altered to the engineered SalI restriction digest site. Four clones positive for the SalI restriction digest site were sequenced using primer 1128. All four clones contained the altered GATA-3394 site with no inadvertent point mutations in the enhancer. Since site directed mutagenesis against GATA -3394 had produced no aberrant clones, clones containing the potentially altered Site 1360 and Site 1601 were directly sequenced. Sequencing showed that Site 1360 and Site 1601 were altered with no aberrant mutations in the enhancer. The three altered enhancers were placed into an untouched SbfI/NheI cleaved pJM309 backbone to ensure that no aberrant mutations were on the vector backbone.

Table 4: Site directed mutagenesis on enhancer construct pJM309

Construct Name	Site name	Site location	Vector backbone	Altered to restriction site:	Left primer	Right primer
pJM310	GATA -3394	-3394	pJM309	SalI	gata-3394A	gata-3394B
pJM311	Site 1360	-3665	pJM309	ClaI	oJB1360	oJB1406
pJM312	Site 1601	-3425	pJM309	SalI	oJB1601	oJB1656

pJM319-pJM321 Site directed mutagenesis on pJM251

pJM251 is a pCR2.1-Topo clone containing an insert with -5050 to -2568 bp of the *elt-2* 5' flanking region. Stratagene QuickChange XL Site-Directed Mutagenesis kit was used with the optimized protocol designed for pJM310 (see above). Three sites, GATA site -3394, Site 1360 and Site 1601 were individually mutated in the pJM251

backbone using the previously designed primers for site directed mutagenesis on template pJM309 (see table below). Positive clones were selected using primers Reverse and olwo1. The enhancer portion was sequenced using primer 1128.

Table 5: Site directed mutagenesis on pJM251

Construct Name	Site name	Site location	Vector backbone	Altered to restriction site	Left primer	Right primer
pJM319	GATA -3394	-3394	pJM251	SalI	gata-3394A	gata-3394B
pJM320	Site 1360	-3665	pJM251	ClaI	oJB1360	oJB1406
pJM321	Site 1601	-3425	pJM251	SalI	oJB1601	oJB1656

pJM325, pJM327, pJM330 Constructs containing the distal region of the *elt-2* promoter with mutations against GATA site -3394, Site 1360, Site 1601

The inserts from the pCR2.1-Topo clones pJM319-pJM321 were placed into the pJM323 backbone using the restriction digest sites NheI and SbfI. The resulting vectors contain the *elt-2* promoter from -5050 to -2568 bp upstream of the ATG site inserted in front of the *elt-2* basal promoter fused to GFP/*LacZ* coding region. Table 6 (below) indicates the name of the constructs, the specific sites altered, the coordinates of the inserts with respect to the full length *elt-2* promoter, and the vector backbone.

Table 6: pJM325-pJM330

Construct Name	Altered Site	*elt-2* promoter insert	backbone
pJM325	GATA site-3394	-5050 to -2568	pJM323
pJM327	Site 1360	-5050 to -2568	pJM323
pJM330	Site 1601	-5050 to -2568	pJM323

pJM328, pJM329, pJM331, pJM326 Full length *elt-2* promoter constructs with mutated GATA -3394, Site 1360 and Site 1601

To make the full length *elt-2* promoter constructs, the constructs pJM325, pJM327 and pJM330 were digested with the restriction enzyme NcoI. CIP was used to inhibit self religation. Each vector was then ligated to a 3340bp (-3028 to +307) fragment released from pJM324 digested with NcoI. Primers gfp-3 and gfp-5 were used to screen for colonies with the correct insert orientation. The enhancer portion was sequenced with primer1128. GATA site -3394 in full length backbone has yet to be made. Table below indicates the name of the construct, the site altered, the length of the insert, and the vector backbone.

Table 7: Full length *elt-2* promoter constructs with altered GATA site -3820, Site 1360 and Site 1601

Construct Name	Altered Site	*elt-2* promoter insert	backbone
pJM329	GATA site -3820	-5050 to 0 (full length)	pJM324
pJM328	Site 1360	-5050 to 0 (full length)	pJM324
pJM331	Site 1601	-5050 to 0 (full length)	pJM324
pJM326	GATA site-3394		

Constructs with pD95.77 vector backbone

pJM313

pPD95.77 is a *yfp* containing vector that has a cryptic basal promoter 3' of the *yfp* ATG site. The Stratagene QuickChange XL Site-Directed Mutagenesis Kit was used with primers 95.75Nhe1forward and 95.75Nhe1reverse to create a NheI site in the 30th base pair position of the multiple cloning site (MCS). The site change was confirmed by digesting the vector with NheI. The vector was then cut with NcoI and PstI to release the altered MCS. The cleaved NcoI/PstI MCS site was then ligated to an untouched pPD95.77 vector (also cut with NcoI/PstI) to remove possible aberrant base pair changes that may have occurred elsewhere on the vector. The altered vector containing a NheI site within the MCS was named pJM313.

pJM314, pJM315, pJM316, pJM317, pJM318 Constructs containing the *elt-2* enhancer (with site directed mutations) inserted into the backbone pJM313

pJM309 (enhancer containing construct) as well as pJM310-pJM312 (enhancer containing constructs with site directed mutations) were digested with restriction enzymes NheI and SbfI. Each insert was ligated to a NheI/SbfI cleaved pJM313 vector. Primers Reverse and olwo3 were used to screen for positive clones. Diagnostic restriction digests were performed using DdeI and SalI. Primer 1135Sbf1 was used to sequence each insert. Table below indicate the name of each construct, the site altered, the coordinates of the insert with respect to the full length *elt-2* promoter, and the name of the vector backbone.

Table 8: Enhancer inserts (with site directed mutations) in the pJM313 backbone

Construct Name	Altered Site	*elt-2* promoter insert	backbone
pJM314	no altered site	-4157 to -3310	pJM313
pJM315	GATA site-3394	-4157 to -3310	pJM313
pJM316	Site 1360	-4157 to -3310	pJM313
pJM317	Site 1601	-4157 to -3310	pJM313
pJM318	GATA site -3820	-5050 to -2568	pJM313

pJM333.

pJM333 contains the complete *elt-2* 5' flanking untranslated region inserted into the *yfp* containing vector pPD95.77. To make this construct, pJM324 was digested with restriction enzyme XmaI. The released XmaI fragment (5062 bp) contains the complete *elt-2* 5' flanking promoter region. The fragment was then ligated to an XmaI digested, CIP dephosphorylated pPD95.77 vector. Primers Reverse and olwo1 were used to screen for colonies containing the insert in correct orientation. A diagnostic restriction digest was performed using the NdeI restriction enzyme.

pJM334

pJM334 contains most of the 5' *elt-2* promoter region (-5078 to -513 upstream of *elt-2* ATG site). To make the construct, pJM324 was digested with the restriction enzyme SalI. The 4565 bp fragment was inserted into a SalI digested, CIP dephosphorylated pPD95.77 vector. Primers Reverse and olwo1 were used to screen for colonies containing the insert in the correct orientation. A diagnostic restriction digest was performed using the NdeI restriction enzyme.

***end-1* and *end-3* reporter constructs**

end-1 was amplified from cosmid T26F2 using primers oJbEnd1Sph and oJbEnd1Bam. The 4199bp PCR product contains the complete *end-1* promoter (5' flanking region up to the upstream T26f2.2 gene) as well as the full *end-1* coding region and the 3' untranslated region. The PCR product was then sub-cloned into the pCR2.1-Topo vector. The restriction enzymes BamHI and SphI were used to release the insert pCR2.1-Topo. The insert was then ligated to a BamHI/SphI cleaved pJM277 vector. pJM277 is a stripped down version of the Fire Lab vector pPD95.11. pJM277 was made by digesting pPD95.11 with SmaI and StuI restriction enzymes and then religating the vector onto itself.

end-3 was amplified from cosmid F58E10 using primers oJbEnd3sph and oJBbEnd3Bam. The 2500bp PCR product includes the 5' untranslated region up to the *aip-1* gene, the complete *end-3* coding region and the 3' untranslated region down to the gene *srh-308*. The PCR product was sub-cloned into the pCR2.1-Topo vector. The insert was released from pCR2.1-Topo using restriction enzymes SphI and BamHI. The insert was then ligated to a SphI/BamHI cleaved pDp95.11 vector. The intermediate vector was then digested with SmaI and StuI restriction enzymes. The *end-3* containing fragment was the relegated onto itself.

The Stratagene QuickChange XL Site-Directed Mutagenesis Kit with primers End3-NheI-3' and End3-NheI-5' was used to create a NheI restriction digest site one base pair before the *end-3* TGA stop codon.. The protocol was directly followed with only one exception. The Dpn1 digestion time was extended from the suggested one hour to four hour incubation time. A diagnostic restriction digest with enzymes NheI and HindIII was

used to confirm the insertion of the NheI site. The construct was then cut with NheI and the 5' phosphate groups were removed using CIP. *mRFP* (see below) was subsequently ligated into the vector.

The *mRFP* gene encodes for a red fluorescent protein. The sequence was PCR amplified from the *mRFP1* plasmid (Campbell *et al.,* 2002) using primers mRFP-NheI-5' and mRFP-NheI-e3-3'. The PCR product was then cut with the restriction enzyme NheI and inserted into the engineered *end-3* NheI site. A diagnostic restriction digest (EcoRI /NcoI) was used to locate the *end-3*::RFP vector with RFP in the correct orientation.

Heatshock protocols

RNA isolation

Plates containing healthy N2 worms were incubated at 30ºC for either 5 hours or 20 hours. Worms were washed off plates with DEPC treated water, placed in RNAse free eppendorf tubes and centrifuged for one minute. Excess water was removed and one mL of Trizol reagent was added. Tubes were vortexed and left to stand at room temperature for 10 minutes followed by a ten minute centrifugation at 14,000 rpm at 4ºC. The supernatant was removed and 200 µL of chloroform was added into each eppendorf tube. The tubes were vortexed for 15 seconds and then left at room temperature for three minutes. Tubes were spun at 12,000 rpm for 15 minutes at 4ºC. The top layer was removed and placed into a new eppendorf tube with 500 µL of Isopropanol. After the solution was mixed and left to stand for 10 minutes, it was spun at 12000 rpm for ten minutes at 4ºC. The supernatant was removed and the pellet was washed with 100 µL of 75% ethanol. After the solution was spun for five minutes, the ethanol was removed and the pellet was left to air dry. The pellet was then dissolved into 25 µL of DEPC-treated water. The resulting solution was heated for ten minutes at 60ºC to dissolve the RNA pellet. RNA preparations had various concentrations: room temperature worms yielded 498.8 ng/µL, five hour heat treated worms yielded 669.7 ng/µL, while 20 hour heat treated worms yielded spectrometer readings ranging from 121.4 ng/µL to 540ng/µL. The Ambion kit, Cells to cDNA II, was used to make cDNA from RNA. The amount of RNA template added was 2490 ng of RNA from room temperature worms, 2678.8 ng of RNA from five hour heat treated worms, 968 ng to 4320 ng from 20 hour heat treated worms.

Quantitative Real Time PCR

Three genes were chosen as controls for quantitative RT-PCR experiments to determine whether *elt-2* expression is elevated in heat shock conditions. PCR reactions were done in a DNA engine Opticon Z system with the PTC-200 DNA Engine cycler and CFD-3220 Opticon 2 detector from MJ research. *hsp-70* and *cyp-2* were chosen based on their expression profiles from microarray hybridization experiments of 11,917 genes (GuhaThakurta *et al.,* 2002). *hsp-70* was chosen as a positive control since heat shocked worms were shown in the same study to have an increase in *hsp-70* expression. *cyp-2* was chosen as a negative control based on its unchanged expression profile (room temperature worms vs. heat shocked worms). *gcs-1* was also used as a positive control since transgenic worms containing the *gcs-1* promoter fused to GFP (pro*gcs-1*::GFP) have an increase in GFP expression in heat induced conditions (see oxidative stress results).

cDNA made from RNA of room temperature worms was used to PCR amplify the *cyp-2*, *gcs-1*, and *hsp-70* cDNA. The three PCR products were cloned into p-GEMT and sequenced. Primer annealing temperatures were optimized using each cloned cDNA product as a template. Annealing temperatures were then re-optimized for each set of primers for amplification from mixed cDNA of non heat shock worms. The sybergreen (Sybr) PCR mixture was composed of 10x PCR buffer, 5 mM dNTP, primer1 5μM, primer 2 5μM, taq (5U/μL) 0.5 μL, Sybr 2.5μL of 1/1000 diluted stock. Below is Table 9 that includes gene, gene function, and information on PCR primer optimization.

Table 9: Genes used as heatshock controls for Quantitative RT-PCR experiments

Gene name	*hsp-70* heat shock protein	*cyp-2* (A.K.A. *cyn-2*) Cyclophilin type peptidyl-prolyl cis-trans isomerase	*gcs-1* gamma-glutamine cysteine synthetase heavy chain
Function	Prevention of protein aggregation and assistance in protein folding	Molecular chaperone	The catalytic sub-unit of γ-Glutamylcysteine synthetase (GCS). GCS makes glutathione which acts as an antioxidant in the cell.

Gene	C30C11.4	Zk520.5	F37B12.2
Left primer	hsp-70left	cyp-2 left	gcs-1left
Right primer	hsp-70right	cyp-2 right	gcs-1right
PCR annealing temp	56ºC	56ºC	64ºC
PCR extension time	1 minute	1 minute	1 minute
PCR product size	544 bp	268 bp	294 bp
Optimized annealing temperature for cDNA in p-GEMT	66.5ºC	70.8ºC	70.8ºC
Optimized annealing temperature for cDNA	61.3ºC	61.3ºC	64.1ºC

Four primers; oJB1227, oJBsplice 986, oJBsplice500 and oJBsplice 688 were tested for the ability to amplify *elt-2* cDNA. Annealing temperatures for primers were first optimized using the template pJM68 (plasmid contains the complete *elt-2* cDNA). Specific binding occurred with primer pair oJBsplice500 and oJBsplice986 at 62.6ºC. The primer annealing temperature was re-optimized against a cDNA mixture. An annealing temperature of 62ºC produced the best results. However, the product size from the primer pair is 783bp and is considered too large for RT-PCR.

Worm Lysis for PCR

The worm lysis protocol was modified from Truett *et al.* (2000). A minimum of three transgenic worms were placed into a PCR tube with 2 µL of alkaline lysis solution (25 mM NaOH and 0.2 mM disodium ethylenediaminetetraacetic acid (EDTA)) and overlayed with 5 µL of paraffin oil. PCR tubes were then placed into a PCR machine for 20 minutes at 95ºC and then cooled to 4ºC. Two µL of neutralizing solution (40mM Tris-HCl) and two µL of 3X worm lysis solution (150 mM KCl, 7.5 mM $MgCl_2$, 30 mM Tris-HCl pH 8.3, 1.35% Tween 20, 1.35% NP40, 0.003% gelatin and 1.25 mg/mL Proteinase K) were added below the paraffin oil. Samples were then reloaded into the PCR machine

and heated for 90 minutes at 60ºC and then for 15 minutes at 95ºC. Six µL were then used as the DNA template for the PCR reaction.

Primers used in this study

Below in Table 10 are the primers used to amplify sequences from the 5' flanking region of *elt-2*. The primers are listed in order according to their corresponding position on the elt-2 promoter; from the most distal position to most proximal to the *elt-2* ATG start site. The table also includes the primer sequence and whether the primer contains a restriction digest site (R.D site).

Table 10: Primers corresponding to the *elt-2* 5' flanking region.

Name	Sequence	upstream of *elt-2*	Function
oJB1	tgtattcagctcgaccaacgc	-4987	screening
olw01	ggtggtgctagcaatgttctatccaagttctgatgc (contains NheI R.D site)	-4584	The 3'deletion series
oJB577	cacgccaatataccgaaaaacg (contains HindIII R.D site)	-4474	screening
IIhindIII	ggtggaagctttggtatatacactatcc (contains HindIII R.D site)	-4435	Conserved Region III
oJB655	tcatcacaaagcgaatcttcagc	-4393	screening
oJbT7-683	taatacgactcactatagggagaccacgaatttgt gtataggagg	-4368	Amplify ORF
olwo2	ggtggtgctagcaaccgatgtatatccgtgtgagg (contains NheI R.D site)	-4323	3'deletion Series
oJBhindIII760	ggtggaagcttcatcgtaaaatgcggcttc (contains HindIII R.D site)	-4287	5' enhancer deletion series
enhancer893	ggtggtcctgcaggttcccctaaagttgcaattcc (contains SbfI R.D site)	-4157	amplify enhancer
en-del-910	ataagaatgcggccgcaataggaattgcaacttta ggg	-4132	delete enhancer
oJBhindIII1055	ggtggaagcttggcaatgcaaatacggatg (contains HindIII R.D site)	-3995	5' deletions of enhancer

			enhancer
oJB1128	aggtaatatacacattgtccgcaac	-3924	sequencing
1135SbfI	ggtggtgcctgcaggacacattgtccgcaacaaaa (contains SbfI R.D site)	-3914	amplify enhancer
hindIII1175	ggtggaagcttctccttcccatgtgc (contains HindIII R.D site)	-3874	amplify enhancer
oJBhindIII1388	ggtggaagcttcgaacccgttcagcatctc (contains HindIII R.D site)	-3661	5' enhancer deletion series
Sbf11473	ggtggtcctgcaggattataatcctattacccacc (contains SbfI R.D site)	-3579	
olw03	ggtggtgctagctaataataatagttacagagtcag (has NheI R.D site)	-3575	3' deletion series
IIpstI	ggtggctgcagatggtacaatcgagagg (contains PstI R.D site)	-3422	Conserved region III
oJbT7-1630	taatacgactcactatagggagaccactagatggtacaatcgagagg	-3414	Amplify ORF for RNAi
oJBhindIII1660	ggtggaagcttaacagatcgaatgtgaaagacc (contains HindIII R.D site)	-3393	5' deletions of enhancer
1663SbfI	ggtggtgcctgcaggatcgaatgtgaaagaccaat	-3388	
Enhancer 1741PstI	ggtggctgcagcccacataagtctctaaatac (contains PstI R.D site)	-3310	Amplify enhancer
enhancer1741	ggtggtgctagccccacataagtctctaaatac	-3310	Amplify enhancer
en-del-1721	atggtggtgcggccgctatttagagacttatgtggg	-3330	delete enhancer
oJBhindIII1845	ggtggaagcttggaaattagtggcgagaaaac (contains HindIII R.D site)	-3205	5' deletions of enhancer

1866NheI	ggtggtgctagctgttttctcgccactaatttcc (contains NheI R.D site)	-3183	
oJBpstI2025	ggtggctgcagccatggaagtatgttagcaag (contains PstI R.D site)	-3022	5' deletions of enhancer
olw04	ggtggtgctagcccatggaagtatgttagcaagtac (contains NheI R.D site)	-3022	3'deletion series
oJB2089	agcggtgatttcttctggtgc	-2962	
oJB2198	ctgtcagaatacttggtgcgagc	-2852	
olw05	ggtggtgctagcggtccgagagctttttttaatacg (contains NheI R.D site)	-2568	3' deletion series
IIIHindIII	ggtggtaagcttatacagaattgataatgttatcttc (contains HindIII R.D site)	-2093	Conserved Region II
olw06	ggtggtgctagctgctctcatttctcaatgtttgcg (contains NheI R.D site)	-1786	3' deletion series
IIIpstI	ggtggctgcagacctgtagaaaaattgataa (contains PstI R.D site)	-1531	Conserved region II
olw07	ggtggtgctagcttataaacaaaacgtgatgcgacc (has NheI R.D site)	-1422	3'deletion series
olw08	ggtggtgctagcagtagataaaacaggcgatcagac (contains NheI R.D site)	-715	3' deletion series
IVHindIII	ggtggtaagcttccaatacgctttgtgccttca (contains HindIII R.D site)	-435	Conserved Region I
IVPstI	ggtggtctgcagaaattaaaactatttgaagac (contains PstI R.D site)	-168	Conserved Region I
XmaI-ATG	tagctcccccgggccattctataatc (contains XmaI R.D site)	+1	remove *elt-2* coding sequence

Table 11: Primers designed to alter sites located in the 5' flanking region of *elt-2*

Name	Sequence	Position on *elt-2* promoter	Function
gata-3822A	caaacgtcatgatgactcattcaacgtcgaccatacctc	-3845	GATA-3820
gata-3822B	acttccggcagccacggaggtatggtcgacgttgaa	-3809	GATA-3820
gata-3394A	agtgtcttgtcgtgtgcataactacgtcgaccagatcga	-3420	GATA-3394
gata-3394B	ttgggtctttcacattcgatctggtcgacgtag	-3367	GATA-3394
oJB1360	ttcagattcttctactccaagtaccgatcgatcacccgt	-3689	Site 1360
oJB1406	atgctgaacgggtgatcgatcggtactggagtagaaga	-3645	Site 1360
oJB1601	agtttattgcctctcgattgtaaaccagctggtcttgtcgtg	-3448	Site 1601
oJB1656	agtagttatgcacacgacaagaccagctggtttacaatcg agaggc	-3395	Site 1601
gata4007A	aacccaacggtcacgtcgactatttttggcaatgcaaatac g	-4022	GATA-4006
gata4007B	ttgccaaaaatagtcgacgtgaccgttgggtttctgtaata g	-3990	GATA-4006

Table 12: Primers designed for Quantitative RT-PCR.

Name	sequence	Function
oJBsplice986	aacttggatgttatcggcag	*elt-2* cDNA
oJB1227	aagcagaccgagaactgagc	*elt-2* cDNA
oJBsplice500	agtaaacggaggaatgatgtgc	*elt-2* cDNA
oJBsplice688	acgcagatccgtttgacc	*elt-2* cDNA
oJBGFP-1,	aacagttgcgaagcttgaatggc	GFP
oJBGFP-2,	aacgactgtcctggccgtaacc	GFP
gcs-1left	attatatcaagaagcacggaattgc	*gcs-1* cDNA
gcs-1right	aagcgatgagacctccgtaag	*gcs-1* cDNA
cyp-2left	ttatggagctctacaacgacatc	*cyp-2* cDNA
cyp-2right	acagtttgccatcgacaag	*cyp-2* cDNA
hsp70left	aagacaaaatacggaattgatgc	*hsp-70* cDNA
hsp70right	acgtcgttttctcctccg	*hsp-70* cDNA

Table 13: Primers designed for Constructs *end-1*::RFP and *end-3*::RFP

Name	Seqeunce	Gene	Function
oJBEnd1sph	acatacatgcatgcaaggagcgaactgtaagaaac g	*end-1*	amplify gene
oJBEND1Bam	ttgatgaatgatggatccttttaaatttaataatg	*end-1*	amplify gene
oJBEnd3sph	acatacatgcatgcgaactgttgttaatgggtgttctg	*end-3*	amplify gene
oJBEnd3Bam	tatcaagaggatccagttttttatttatttattatggaaa cc	*end-3*	amplify gene
End1-NheI-3'	aagataagataaatttagctagcagagaaatg ttgtgatggaatg	*end-1*	create NheI site in second last codon

End1-NheI-5'	catttctctgctagctaaatttatcttatctttttatgtgtg tatg	*end-1*	create NheI site in second last codon
End3-Nhe-I-5'	tgcttttaattttgctagctgatttaactgatcaactctgt gc	*end-3*	create NheI site in second last codon
End3-NheI-3'	ttgatcagttaaatcagctagcaaaattaaaagcatta g	*end-3*	create NheI site in second last codon
mRFP-NheI-5'	ctactagctagcggaggaatggcctcctccgagga cg	*mRFP*	amplify gene
mRFP-NheI-e1-3',	ctactagctagcttattaggcggcggtggagtggcg	*mRFP*	amplify *mRFP* with complementary sequence to *end-1* stop codon
mRFP-NheI-e3-3'	ctactagctagctgatgaggcgccggtggagtggc gg	*mRFP*	amplify *mRFP* with complementary sequence to *end-3* stop codon

Table 14: Primers used for fusion PCR protocol ((Biotechniques 32:728-730 (April2002))

Name	Sequence	Function	*elt-2* promoter
95.77MCS	aagcttgcatgcctgcaggtcgactc	amplify 95.77 first round	
UNC54UTR	aagggcccgtacggccgactagtagg	amplify 95.77 first round	
UNC54NEST	ggaaacagttatgtttggtatattggg	amplify 95.77 second round	
oJB2510	acctgcaggcatgcaagcttttaatacg tgaaacattccatcac	amplify *elt-2* promoter first round	-2588
oJB2510B	agcgattatggcgttagagcttttaatac gtgaaacattccatcac	amplify *elt-2* promoter	-2588
oJB4540	acactaacgccataatcgctagccagc	amplify *elt-2* promoter	-512
oJB90	aagcgttggtcgagctgaatacac	amplify *elt-2* promoter	-4585

Table 15: Primers that correspond to vector sequences

Name	Sequence	Function
L4440F	catgttctttcctgcgttatc	RNAi
L4440R	ctgcaaggcgattaagttg	RNAi
Reverse	gagcggataacaatttcacacaggaaa	
Sp6	agctatttaggtgacactatag	p-GEMT
96E' SmaI	tacctttgggtcctttggcc	pJM76
95.75nheIforward	aagcttgcatgcctgcaggtcgactctagagctagcggat ccccggg	Site directed against MCS
95.75nheIreverse	tcctttggccaatcccggggatccgctagctctagagtcga cctg	Site directed against MCS

Table 16: primers designed to amplify *C. briggsae* CBG17257 promoter

Name	sequence	Function
brigg893	ggtggtcctgcaggtctcccaaggtagctgttcc	*briggsae elt-2* promoter
brigg1741	ggtggtgctagcccatcctgtccgtatgttgtc	*briggsae elt-2* promoter

CHAPTER THREE: IDENTIFICATION OF THE ENHANCER REGION

Specific Aim I: To determine whether the 5' flanking region of *elt-2* regulates *elt-2* expression in embryonic and larval stages of development

It was important to determine whether the 5' flanking region of *elt-2* was independently responsible for the regulation of *elt-2* expression, or whether *elt-2* was also regulated through sites present within introns and downstream of the gene. With the expectation that regulatory regions would be conserved through evolution; the genomic region surrounding *elt-2* was compared to the *elt-2* orthologues in *C. remanei* and *C. briggsae* to locate regions with a high degree of sequence similarity. Two way alignments using the program EMBOSS Dotmatcher were done between *C. elegans* and *C. remanei*, and *C. elegans* and *C. briggsae* (Figure 18+19). The parameters were set to create a dot, wherever a minimum sequence conservation of 35 bp was located within a 40 bp sliding window. The DotMatrix illustrates that within a 7000 bp region upstream of the *elt-2* gene, there are four regions that contain a high degree of sequence similarity and are denoted I, II, III, and IV. Region IV flanks the 5' end of the *elt-4* gene and is presumed to be the basal promoter of *elt-4*. The three remaining regions are within the 5144 bp region that flanks the 5' end of the *elt-2* gene. Sequence comparisons between the three species did not locate regions of sequence similarity within the eight introns or downstream of *elt-2*; suggesting that the upstream region may contain most, if not all, of the *elt-2* regulatory sites (data not shown).

To determine experimentally whether the upstream region flanking *elt-2* was indeed responsible for regulating expression; I compared the expression from transgenics containing a reporter construct with the complete *elt-2* 5' flanking region fused to GFP/*lacZ* (pJM67) to the staining pattern of the ELT-2 antibody. The ELT-2 polyclonal antibody depicts that ELT-2 is present half way through the 2E cell stage of intestinal development and continues to be expressed uniformly in all intestinal cells for the complete life of the worm (Figure 7) (Fukushige *et al.,* 1998). Three independent transformed lines containing the reporter construct pJM67 were analyzed in three embryonic stages of development (gastrulation stage, comma stage, pretzel stage) as well as larval stages.

As illustrated in Figure 20, the staining pattern from β-galactosidase activity in embryos containing the transgene pJM67 is similar to embryos stained with the ELT-2 antibody. Temporally, β-galactosidase activity is reproducibly detected in such embryos at the 2E cell stage and in all embryonic stages of development. The above result indicates that the 5144 bp upstream flanking region contains the complete transcriptional regulatory region necessary to control the temporal and spatial pattern of *elt-2* in embryonic stages.

Like the uniform staining pattern of the ELT-2 antibody in the intestine of larval stage 1 (L1) worms, recently hatched transgenic L1 worms containing pJM67 have uniform expression of GFP/β-galactosidase throughout the intestine. However as worms mature, GFP/β-galactosidase expression is maintained only in the anterior and posterior of the intestine while expression in the mid intestine becomes weak and mosaic. The loss of GFP/β-galactosidase expression in the mid region of the intestine does not recapitulate the uniform intestinal staining pattern of the ELT-2 antibody in all larval stages of development. The discrepancy between the expression pattern from the reporter construct and ELT-2 antibody staining may indicate that *elt-2* is regulated in larval stages through site(s) outside of its 5' flanking region. The discrepancy may also be explained by the fact that intestinal specific reporter constructs containing *lacZ* have a tendency not to be expressed as strongly in older worms (McGhee, personnel communication). Consequently, the expression from the reporter construct pJM67 can not be used as a reliable marker for *elt-2* expression after the L1 stage.

Specific Aim II: To determine the minimum promoter region necessary for *elt-2* embryonic expression

Sequence comparisons between the three *Caenorhabditis* species indicated that, within the *elt-2* promoter (5' flanking region), there are three highly conserved regions and are denoted as I, II, III. It was first important to determine whether one conserved region independently regulates *elt-2* expression or whether *elt-2* is coordinately regulated by more than one conserved region. The first construct series, the 5' deletion series, was created to delete the conserved regions in a sequential 5' to 3' arrangement, thereby removing Conserved Region III, then II, and then I. The series would test whether Conserved Region I was sufficient for *elt-2* expression or whether Conserved Region II

and I coordinately regulated *elt-2* expression. The second series, named the 3' deletion series, sequentially removed Conserved Region II and then III, while maintaining Conserved Region I. The series was designed to test whether Conserved Region II was necessary for *elt-2* expression or whether Conserved Region III and I coordinately regulated *elt-2* expression (Figure 21).

The 3' deletion series contains ten constructs and was created using pJM67 as the parental vector. Conserved Region I was preserved in all ten constructs, since it was presumed that a basal promoter was necessary for expression and that this region contained the endogenous *elt-2* basal promoter (-498 to 0 base pairs from the *elt-2* ATG start site). The constructs in the 3' deletion series were CsCl purified and injected at 50 μg/mL with the transformation marker *rol-6* used as a visible marker for transformed worms.

As figure 21 illustrates, the constructs pJM258 to pJM252, sequentially remove the region between Conserved Region I and III, and in doing so remove Conserved Region II. All four constructs produced the same expression pattern; β-galactosidase activity remained uniform throughout the embryonic intestine similar to control embryos containing the reporter construct pJM67 (contains complete *elt-2* promoter) (Figure 22-25). These results indicate that Conserved Region II does not contain regulatory sites necessary for *elt-2* expression.

Transgenic embryos containing the reporter construct pJM250, which contains both Conserved Region I and III, had strong uniform β-galactosidase staining similar to control embryos containing pJM67 (Figure 26). However, embryos containing the next reporter construct in the series pJM248 (contains the distal half of Conserved Region III but not the proximal half) had a dramatic loss in expression with only weak staining of β-galactosidase activity occurring sporadically (Figure 27). Consequently, the 3' deletion series indicated that although Conserved Region II is not necessary for *elt-2* regulation, the proximal region of Conserved Region III is necessary and most likely contains regulatory sites responsible for the regulation of *elt-2*.

The next three constructs in the deletion series (pJM246, pJM244 and pJM259) sequentially remove Conserved Region III as well as the remaining distal region of the *elt-2* promoter. All three constructs produced sporadic weak staining in embryos similar

to pJM248 suggesting that once the proximal region of Conserved Region III is removed, no other distal sites are sufficient to direct *elt-2* expression. (Figure 28-30). pJM259, the final construct in the series contains only Conserved Region I and its lack of expression indicates that the *elt-2* basal promoter region is not sufficient to direct *elt-2* expression.

The 5' deletion series had been made by Mark Hawkins (a previous student in the lab) to identify the most distal end of the promoter necessary for *elt-2* regulation. The 5' deletion series is composed of nine constructs (Figure 31). The series sequentially removes the *elt-2* promoter from the distal to proximal region; the largest construct, named 0elt-2.lac-z, has 5044 base pairs of the *elt-2* promoter fused to *lacZ*, the final construct, named 5006elt-2.lac-z, contains the 38 base pairs preceding the *elt-2* ATG start site fused to *lacZ*. The first three constructs contain all three conserved regions and produced normal gut staining in embryogenesis. The next construct in the series, 1193elt-2.lac-z, has approximately half of the distal region of Conserved Region III removed with no change in the staining pattern of β-galactosidase activity. However, construct 1882elt-2.lac-z has all of Conserved Region III removed and produced weak sporadic staining of β-galactosidase activity indicating that within the proximal region of Conserved Region III there is at least one critical site necessary for *elt-2* expression.

In the 5' deletion series, the next three constructs in the series (2144elt-2.lac-z, 4094elt-2.lac-z, 4531elt-2.lac-z) produced weak sporadic staining of β-galactosidase activity similar to the reporter construct 1882elt-2.lac-z suggesting that Conserved Region II and I are not sufficient to direct *elt-2* expression. Interestingly, both deletion series indicate that the weak sporadic staining may derive from Conserved Region I. In the 5' deletion series, construct 4531elt-2.lac-z contains only Conserved Region I and produced sporadic staining while construct 5006elt-2.lac-z has Conserved Region I removed and produced no staining. In the 3' deletion series, the reporter construct pJM259 contains only Conserved Region I and also produced sporadic staining. Therefore, Conserved Region I is responsible for the sporadic staining pattern of β-galactosidase activity and may indicate that the *elt-2* basal promoter has an active role in *elt-2* regulation.

In conclusion, the 5' and 3' deletion series indicated that proximal region of Conserved Region III is necessary for *elt-2* expression. The 3' deletion series also

indicated that Conserved Region II could be removed without abolishing expression suggesting that Conserved Region II is not necessary for *elt-2* expression. The ectopic staining produced by Conserved Region I indicated that the *elt-2* basal promoter may be important for *elt-2* expression. Therefore it was important to further characterize: i) the minimum region of Conserved Region III necessary for *elt-2* expression; ii) whether the expression of *elt-2* is dependant on the presence of the endogenous *elt-2* promoter.

Specific Aim III: To determine the minimum enhancer region necessary for *elt-2* expression

The 5' and 3' deletion series revealed that the most proximal region of Conserved Region III (821 bps) is necessary for *elt-2* expression and most likely contains the *elt-2* enhancer. To further narrow down the region and responsible for *elt-2* expression, a more thorough enhancer-specific 5' deletion series was constructed. This series contained five constructs (pJM284 to pJM288) and sequentially deleted Conserved Region III from the distal to the proximal end (-4287bp to -3205bp from *elt-2* ATG) while the 3' end remained fixed (-3022 bp from *elt-2* ATG) (Figure 32). The five deletion constructs have pJM76 as their reporter vector backbone. pJM76 was chosen as the backbone for the series for two reasons: first, it was the standard GFP/*lacZ* reporter vector in the lab; secondly, the basal promoter is a heat shock basal promoter and not the endogenous *elt-2* basal promoter (Conserved Region I). It was important to isolate Conserved Region III from Conserved Region I in order to characterize the role of the enhancer, independent of Conserved Region I. In the enhancer-specific 5' deletion series, pJM284 and pJM285 were designed as positive controls for expression since they both contain more than the minimum enhancer sequence necessary for expression; as determined by the 5' and 3' deletion series. The next three constructs in the series were designed to sequentially remove sequences within the determined enhancer region. Interestingly, GFP expression was not present in transgenics containing any of the five constructs; furthermore, staining for β-galactosidase activity which can be used to locate weak or sporadic expression did not locate any aberrant staining. Transgenic embryos containing pJM284 and pJM285 were expected to have expression, since the constructs contain more than minimum enhancer region. Transgenic lines for each construct were verified by PCR to contain their respective construct. Each construct had been injected at the same concentration as

the previous deletion series (50 μg/mL) indicating that the lack of expression was not due to low vector copy numbers in the worms. Taken together, the results indicate that the lack of GFP and β-galactosidase expression in the enhancer-specific 5' deletion series is due to the inability of transgenics to express the construct.

The lack of expression from the enhancer-specific 5' deletion series indicates that the expression of *elt-2* is dependant on the presence of the endogenous *elt-2* promoter. In the two previous deletion series, expression was visualized from any reporter construct that contained both the proximal region of Conserved Region III and the *elt-2* basal promoter (Conserved Region I). In the enhancer-specific 5' deletion series, the proximal region of Conserved Region III was fused to a heat shock basal promoter instead of the endogenous *elt-2* basal promoter. The discrepancy between the lack of expression from the enhancer-specific 5' deletion series and expression from previous deletion constructs suggests that the *elt-2* needs its own endogenous basal promoter for expression and that this basal promoter is within Conserved Region I.

To circumvent problems with basal promoters, pJM309 was made which contains the proximal region of Conserved Region III (-4157 to -3311 bp) placed into the same reporter vector backbone used in the 3' deletion series. The backbone contains the *elt-2* basal promoter plus 98 base pairs of the *elt-2* coding region fused to *gfp* and *lacZ*. pJM309 was initially injected at the standard concentration of 50 μg/mL with the transformation marker *rol-6*. However, at this concentration, F1 rolling progeny occurred at a very low frequency and stable lines could not be established. By decreasing the injected plasmid concentration to two μg/mL, the frequency of F1 rolling progeny increased and only two independent stable lines could be generated. The first line recapitulated the expression of transgenic control embryos containing the reporter construct pJM67 while the second line had undetectable expression. The two transgenic lines were verified by PCR to contain the pJM309 construct. The discrepancy between the two lines can be explained by a toxic effect produced by the transgenes. It is reasonable to assume that the plasmid would only be present at a very low copy number level; at a level too low to be detected or at a level detectable only by staining for β-galactosidase activity. The lethal effect from pJM309 is discussed in greater detail below in the section, Specific Aim VII.

The GFP/β-galactosidase expression from transgenic embryos recapitulated the staining pattern of the ELT-2 antibody in embryogenesis; expression initiated at the 2E cell stage and remained throughout the intestine, suggesting that pJM309 contains a sufficient region of the *elt-2* enhancer to direct *elt-2* expression (Figure 33). Although the 841bp sequence is most likely not the minimum enhancer region, the sequence was of a reasonable size for the next set of approaches. As discussed in the next section, Specific Aim IV, a bioinformatics approach was used to locate potentially important sites within this enhancer sequence. As described in Specific Aim V, band shift assays were then used to experimentally determine if known intestinal transcriptional regulators could bind *in vitro* to the these potentially important sites. As explained in Specific Aim VI, some of these potential transcription factor binding sites were ultimately tested for their *in vivo* role in the regulation of *elt-2*.

CHAPTER FOUR: ANALYSIS OF THE ENHANCER REGION

Specific Aim IV: Interspecies sequence comparisons of the *elt-2* enhancer and identification of potential transcription factor binding sites.

Caenorhabditis species sequence alignments of the complete 5' promoter region flanking *elt-2* were successful in locating important *elt-2* regulatory regions. Therefore, it was reasonable to assume that a more detailed sequence alignment of the proximal region of Conserved Region III would also be successful in locating potentially important sites within the enhancer. Figure 34A+B contains a three way alignment of the more proximal region of Conserved Region III and depicts numerous sites that are highly conserved between these three species. It is important to note, and is illustrated in figure 34A, there are sequences that are highly conserved between these three species that are located outside of the experimentally determined 5' end of the enhancer. Figure 34B illustrates that the 3' end of the enhancer, as determined by the enhancer construct pJM309, is within the 3' end of Conserved Region III. To determine if any of the conserved sequences were of significance, the computer program Transfac 6.0 was used to locate potential protein binding sites within the *elt-2* enhancer. Transfac 6.0 is a database of eukaryotic transcription factors and their genomic binding sites. Transfac 6.0 identified numerous potential transcription factor binding sites within the proximal region of Conserved Region III; of immediate interest were WGATAR and RTCAT binding motifs since WGATAR motifs have been previously characterized as an END-1 and ELT-2 binding site and the RTCAT motif as a SKN-1 binding site (Zhu *et al.,* 1997; Marshall and McGhee 2001; Blackwell *et al.,* 1994).

As figure 34A illustrates, within the proximal region of Conserved Region III are three potential SKN-1 binding sites. SKN-1 site -3945, although present and highly conserved between species is outside of the experimentally determined enhancer region, cautioning that a conserved potential binding site does not signify that the site is necessary for *elt-2* expression. Two overlapping SKN-1 binding sites are present within the experimentally determined enhancer region of *C. elegans* (-3805 bp upstream of *elt-2* ATG site); however, the SKN-1 binding sites are not present in either *C. briggsae* or *C. remanei*. Therefore according to the sequence alignment and the location of SKN-1

binding sites, the three potential SKN-1 binding sites are most likely not necessary for regulating *elt-2* transcription.

Transfac 6.0 located four potential GATA transcription factor binding sites within the proximal region of Conserved Region III in *C. elegans*. As illustrated by figure 34A, the first GATA site -4055 is conserved between *C. elegans*, *C. briggsae* and *C. remanei*; however, the site is located outside of the experimentally determined enhancer region. The second GATA site, -3820, is present within the experimentally determined enhancer region and is conserved in all three species. The third *C. elegans* GATA site, -3611, is not present in *C. briggsae* or *C. remanei*. However, *C. briggsae* and *C. remanei* contain a GATA site 37 base pairs downstream of the *C. elegans* GATA site -3611. The sequence flanking the *C. remanei* and *C. briggsae* GATA site is not similar to the *C. elegans* GATA site -3611 suggesting that the two sites are not derived from the same ancestral site. The most proximal GATA site, -3394, is completely conserved between the three species, with moderate conservation flanking the GATA site. Therefore, sequence comparisons between the three species highlighted two conserved and potentially functional GATA sites within the enhancer region.

The alignment of the proximal region of Conserved Region III highlighted two sites, named 1360 (Figure 34A) and 1601 (Figure 34B), as being the most highly conserved sites within the enhancer. Site 1360 in *C. elegans*, is present and highly conserved in both *C. remanei* (31/32 base pairs conserved) and *C. briggsae* (30/32 base pairs conserved). Site 1601 in *C. elegans*, is conserved in *C. briggsae* (24/25 base pairs conserved) and in *C. remanei* (23/25 base pairs conserved). Although Transfac did not highlight Site 1601 as a potential transcription binding site; it did highlight that Site 1360 contained a CCAAT motif, previously reported as an important regulatory site in the mouse alpha2(I) collagen gene (Karsenty *et al.,* 1998).

Specific Aim V: To determine whether SKN-1, END-1 and ELT-2 can bind *in vitro* to conserved binding sites located within the *elt-2* enhancer

In the previous section, Transfac 6.0 located five sites within the *elt-2* enhancer that were potential binding sites for known regulators of intestinal development. Although interspecies alignments of the *elt-2* enhancer sequence highlighted that only two GATA factor transcription binding sites were conserved between all three species; it

was important to determine if key regulators of intestinal development could bind *in vitro* to any of their respective sites within the enhancer region.

Band shift assays were done by Barbara Goszczynski to experimentally determine if the transcription factors; ELT-2, END-1 and SKN-1, could bind *in vitro* to potential binding sites located within the experimentally determined enhancer. The enhancer sequence was subdivided into three 150 to 200 bp probes (Figure 34A+B). Probe A (-3886 to -3740 from ATG) contains GATA site -3820 (G**TGATAA**G) and the two overlapping SKN-1 binding sites (GTCATCAT). ELT-2, END-1 and SKN-1 were all able to bind the probe to form Protein::DNA complexes, indicating that these transcription factors are capable of binding the sequence *in vitro*. The second (-3739 to -3540 from ATG) and the third probe (-3539 to -3388 from ATG) each contain a WGATAR sequence motif; TCT**TGATAA**C (-3611) and TAC**TGATAA**C (-3394) respectively. Band shifts using ELT-2 and END-1 protein indicated that while ELT-2 could bind to both probes, END-1 was unable to bind to either.

The band shift assay indicated that there are four potential binding sites within the enhancer region. ELT-2 could bind to all the three GATA sites (-3820, -3659, -3394) while END-1 could only bind to GATA site -3820. The observation that END-1 was able to bind to only one GATA site suggests that END-1 has a more stringent binding profile than ELT-2. The fourth potential binding site, SKN-1 site -3805, was shown to be bound by SKN-1 *in vitro*. Therefore according strictly to binding assays, all four transcription factor binding sites have the potential to interact with key transcription factors, indicating that these sites may participate in the regulation of *elt-2* expression.

In summary, the methods previously described in Specific Aim IV and V were used to locate numerous sites that could be potential *elt-2* regulatory sites. However, a potential site had to first pass all three of the following requirements in order to be tested for an *in vivo* function: i) site must be within the experimentally determined enhancer region; ii) site must be conserved between all three species; iii) a known regulator of intestinal development must bind the site *in vitro*. As table 17 depicts, GATA site -3820 and -3394 were the only two sites that passed all three requirements. It should be highlighted that only GATA site -3820 was shown to interact with END-1 *in vitro* suggesting that if indeed END-1 or END-3 initiates *elt-2* expression, then this site may be

necessary. Since ELT-2 can bind *in vitro* to both GATA site -3820 and GATA site -3394, ELT-2 may interact with one or both sites to autoregulate expression. As described in the next section, Specific Aim VI, a series of site directed mutagenesis experiments were used to determine if GATA site -3820, -3394 and Site 1360 and 1601 do in fact regulate *elt-2* expression *in vivo*. Although Site 1360 and 1601 do not correspond to binding sites of known regulators of intestinal development, it was important to test these sites *in vivo* in the hopes that their role in the regulation of *elt-2* would be determined.

Table 17: potential sites of *elt-2* regulation within the proximal region of Conserved Region III

Name	Within experimentally determined enhancer region	Conserved between all three species	Binds a known regulator of intestinal development
SKN-1 -3945	No	Yes	not tested
SKN-1 -3805	Yes	No	SKN-1
GATA-4055	No	Yes	not tested
GATA -3820	**Yes**	**Yes**	**END-1 and ELT-2**
GATA -3611	Yes	No	ELT-2, not END-1
GATA-3394	**Yes**	**Yes**	**ELT-2,** not END-1
Site 1360	Yes	Yes	unknown
Site 1601	Yes	Yes	unknown

Specific Aim VI: To determine the *in vivo* function of GATA site -3820, -3394, Site 1360 and 1601 using site directed mutagenesis

The first set of site directed mutagenesis experiments were designed to knock out GATA site -3820 and GATA site -3394 independently within the pJM252 reporter construct backbone. pJM252 contains the distal portion of the *C. elegans elt-2* promoter (-5048 to -2568bp upstream of *elt-2*) fused to the *elt-2* basal promoter and *elt-2* coding region corresponding to the first 32 amino acids, and GFP/*lacZ* coding sequence (Figure 21). Initially, a two step site directed mutagenesis PCR protocol was used to achieve the desired mutations (see materials and methods). Although GATA site -3820 and GATA site -3394 were correctly mutated, sequencing uncovered unwanted base pair changes

within the enhancer region of the GATA site -3394 construct. For site directed mutagenesis against GATA site -3820, no unwanted base pair changes were found within the enhancer region. To circumvent problems with unwanted base pair changes, the two step site directed mutagenesis PCR protocol was replaced with Stratagene QuickChange XL Site-Directed Mutagenesis Kit, given that the kit contains a high fidelity polymerase.

GATA site -3820 may act redundantly with at least one other site in the *elt-2* promoter

The first construct in the site directed mutagenesis series (pJM308) has a mutated GATA -3820 site within the pJM252 backbone (see above). pJM308 was injected at 50 μg/mL and produced no F1 rolling progeny. The concentration of injected pJM308 was decreased until F1 progeny were detected. When pJM308 was injected at a concentration of two μg/mL a total of 347 F1 progeny were collected; however, most were sick and died before reproducing. Nine independent transgenic lines were established and PCR was used to verify that only two of the lines contained the transgene. pJM308 transgenic embryos were stained for β-galactosidase activity parallel to control embryos (pJM252). In contrast to pJM252 embryos, which contained strong GFP/β-galactosidase expression, embryos containing the pJM308 transgene showed no detectable β-galactosidase activity.

Although the results indicated that GATA site -3820 was necessary for *elt-2* expression, I was hesitant to conclude that the results were definitive. First, pJM308 was injected at a concentration of two μg/mL, a concentration drastically lower than the standard injected concentration of 50 μg/mL. To test whether the concentration of the plasmid may be responsible for the lack of expression in transgenics, pJM67 was injected at two μg/mL. pJM67 contains the full length *elt-2* promoter and is larger than pJM308. Robust expression was seen in all transgenic lines produced from pJM67 injections indicating that two μg/mL of plasmid was sufficient for expression. The frequency of establishing stable transgenic lines from pJM67 injections was similar to the frequency obtained from previous injections of pJM252 (15-20%), both of which were in sharp contrast to injections of pJM308 (0.5%), suggesting that the low frequency of recovering stable lines from pJM308 injections was probably due to a lethal effect of the plasmid.

Since full length *elt-2* promoter constructs had never produced a lethal effect, the reporter construct pJM329 was made and injected at five μg/mL (contains the full length

elt-2 promoter with GATA site -3820 mutated). No lethal effect was seen and transgenics containing pJM329 had GFP expression and β-galactosidase activity in embryos and larval worms. However, staining of β-galactosidase activity indicated that expression was slightly different between transgenic embryos of pJM329 and pJM324 (contains the full length *elt-2* promoter fused to GFP/*lacZ*) (Figure 35). Worms containing the transgene pJM329 had weaker embryonic β-galactosidase activity than worms containing pJM324. However, at the time of hatching transgenic, worms containing pJM329 had the same intensity of staining as worms containing pJM324, suggesting that expression from pJM329 may represent a less robust onset of *elt-2* expression.

It was curious that pJM308 (GATA -3820 mutated within the partial *elt-2* promoter) produced no detectable β-galactosidase activity while pJM329 (GATA -3820 mutated within the full *elt-2* promoter) did; albeit, slightly weaker than control embryos. One possible explanation for the discrepancy in expression was that worms were injected with the full length construct pJM329 (5 μg/mL) at a higher concentration than the partial length promoter construct pJM308 (2 μg/mL). However, the difference in concentration of injected plasmid is most likely not the cause for the discrepancy given that the difference in concentration is relatively small. Another possibility for the inconsistency in expression is that pJM308 is selected against since it is toxic to the worms; hence, transgenics for pJM329 have a higher copy number of plasmid than pJM308 transgenics. Another more probable explanation is that the full length promoter contains sites that facilitate *elt-2* expression that are not present in the shorter version construct pJM308. Therefore the reporter constructs pJM308 and pJM329 indicate that GATA -3820 may act redundantly with at least one other site to regulate *elt-2* expression, and that the site is located in the proximal region of the *elt-2* promoter (-2568 to -498 bp from ATG).

GATA site -3394 is necessary for *elt-2* expression

The reporter construct pJM325 has GATA site -3394 mutated within the distal region of the *elt-2* promoter (-5050 to -2568 bp upstream of ATG) fused to the *elt-2* basal promoter and GFP/*lacZ*. The construct was injected at a concentration of five μg/mL and produced numerous stable healthy lines. PCR verified 13 of the 14 lines contained the plasmid. Both GFP and β-galactosidase staining was used to observe expression from

transgenic lines. No expression of the transgene was seen at any stage of development, with the exception of one adult worm that contained ectopic cuticle staining (Figure 36A). The lack of expression from transgenics agrees with the previous results from the 5' and 3' deletion series, which indicated that removal of any region containing GATA site -3394 eliminated expression. The 5' deletion series also demonstrated that no sites in the proximal promoter region were sufficient to drive *elt-2* expression, indicating that a full length promoter construct containing a mutated GATA site -3394 was not necessary.

SITE 1360 may bind a negative regulator of *elt-2* transcription

pJM328 contains the full length *elt-2* promoter with Site 1360 mutated (11bps altered -3717 bp upstream of the *elt-2* ATG site). Both pJM328 and the control plasmid pJM324 were injected at a concentration of five μg/mL. Transgenic worms containing pJM328 had an estimated two fold increase in GFP and β-galactosidase expression at all stages in development compared to pJM324 control worms (Figure 37). I analyzed the altered site using Transfac 6.0 to determine if the site was unintentionally switched to a potential binding site; however, the site was not similar to any known transcription factor binding site. Therefore as illustrated by figure 37E, under normal conditions Site 1360 may be a potential repressor binding site responsible for the down regulation of *elt-2* transcription.

It is important to review some previous experiments on *elt-2* autoregulation before describing the entire results for site directed mutagenesis against Site 1360. As previously discussed in the introduction, ELT-2 has been visualized localizing *in vivo* to its own promoter, suggesting that ELT-2 autoregulates expression (Fukushige *et al.,* 1999). Although ELT-2 has not been experimentally shown to down regulate its expression; it is reasonable to assume that an excess of ELT-2 protein in the cell would trigger the down regulation of *elt-2* transcription. Experimentally ELT-2 has been shown to positively regulate its expression; ELT-2 expressed though out the embryo drives ectopic *elt-2* expression outside of the intestinal primordium (Fukushige *et al.,* 1998). However, *elt-2* RNAi experiments indicate that *elt-2* remains expressed throughout embryogenesis and the L1 stage in the absence of the ELT-2 protein suggesting that at least one other positive regulator exists.

To confirm that *elt-2* was positively regulated by at least one other factor and not solely by ELT-2, *elt-2* RNAi was performed on pJM328 worms. I reasoned that, in the absence of ELT-2 (positive regulator) and Site 1360 (negative regulator binding site) *elt-2* expression would increase; if in fact, one other factor positively regulated *elt-2* expression. Mothers containing pJM328 (Site 1360 mutated) were injected with dsRNA corresponding to the *elt-2* or *acn-1* gene (control). RNAi against the *acn-1* gene produce a lethal L1 arrest similar to *elt-2* RNAi experiments (minus the Gob phenotype). F1 progeny from transgenic mothers injected with dsRNA corresponding to *elt-2* had an increase in GFP expression compared to F1 progeny from transgenic mothers injected with dsRNA corresponding to *acn-1*. Furthermore, the transgenic mothers injected with dsRNA corresponding to *elt-2* also had an increase in intestinal GFP expression compared transgenic mothers injected with dsRNA corresponding to *acn-1*. Interestingly, GFP expression increased exponentially as the injected mothers aged suggesting no upper limit in expression. As illustrated in figure 37F, I interpret the above results as implying that: i) Site 1360 is responsible for negatively regulating *elt-2* expression; ii) a positive regulator other than ELT-2 can regulate *elt-2* expression; iii) the positive regulator is not END-1 or END-3 since these factors are not expressed after embryogenesis.

Site 1601 has no identifiable function

Next I wanted to determine if altering Site 1601 had any effect on *elt-2* expression. pJM331 contains the full length *elt-2* promoter with a mutated Site 1601. pJM331 transgenic worms had similar GFP/β-galactosidase expression as transgenic worms containing the pJM324 (full length promoter construct). Injections of dsRNA corresponding to *elt-2* into such transgenics produced no change in expression. Even though transgenics containing pJM331 have GFP/β-galactosidase expression similar to control worms, I do not exclude the possibility that Site 1601 may have a yet undetected function.

In summary, I determined by site directed mutagenesis that GATA site -3820 and GATA site -3394 have a role in the regulation of *elt-2*. GATA site -3820 may be responsible for the initiation of *elt-2* expression since; the site binds END-1 *in vitro*, and

transgenic embryos containing the reporter construct pJM308 (GATA site -3820 mutated within the distal region of *elt-2* promoter) have no expression. Transgenic embryos containing the reporter construct pJM329 (GATA site -3820 mutated within the full length *elt-2* promoter) have a weak onset of expression suggesting that at least one other site in the proximal promoter region can initiate *elt-2* expression in the absence of GATA site -3820; albeit, less robustly. I also determined that, GATA site -3394 binds ELT-2 *in vitro*, and that the site is necessary for *elt-2* expression. Although altering Site 1601 produced no change in expression, altering Site 1360 produced an increase in *elt-2* expression suggesting that Site 1360 may bind a negative regulator of *elt-2* transcription.

Specific Aim VII: To determine the source of lethality produced by reporter constructs

Although not described in detail in the previous section, numerous constructs in this study were lethal to worms such that transgenic lines were difficult to generate and were usually unstable. One source of lethality was due simply to impurities in plasmid preparations using the Wizard Plus Minipreps DNA Purification system from Promega. Once plasmids were purified using a phenol chloroform system, some lethality was removed.

Another source of lethality was that many of the constructs contained the coding region for the first 32 amino acids of *elt-2*. For example, numerous rounds of injections were necessary to establish two transgenic lines containing the enhancer construct pJM309. When transgenic worms containing pJM309 were analyzed at high magnification, L1 stage worms contained a gut obstructed phenotype (Gob) similar to *elt-2* nulls (Figure 38). At this time, it is difficult to understand how this sequence may cause the lethal effect; however, it is possible that less than optimal promoters create partial *elt-2* transcripts from both template strands that then go on to be primary products for RNAi.

Unfortunately, the *elt-2* sequence was not the final culprit for a lethal effect. The 3' deletion series pointed to the distal region of the *elt-2* promoter also producing a lethal effect. In the 3' deletion series, the first four constructs (pJM258 to pJM252) produced numerous stable lines while numerous rounds of injections were needed to produce stable lines for constructs pJM250 to pJM244. The final construct pJM259 (contains only the *elt-2* basal promoter) produced ten lines after the first round of injection suggesting that

Conserved Region I was not responsible for the lethal effect. All of the constructs contained the problematic 32 amino acids of *elt-2*. Although the *elt-2* coding region undoubtedly produced some lethal effect; embryos were shown to be severely malformed (figure 26C and 28B), a phenotype not seen in *elt-2* RNAi. Therefore, in the 3' deletion series the lethal effect of the smaller constructs pointed to the distal region of the *elt-2* promoter as producing the lethal effect. The lethal effect of the distal region of the *elt-2* promoter was supported by the observation that stable lines could not be obtained from constructs pJM314 to pJM318. Constructs pJM314 to pJM318 contain the distal region of the *elt-2* promoter in pJM313; a *yfp* reporter backbone that does not contain the *elt-2* coding region (Figure 39). Each construct was injected and no stable lines were obtained suggesting, once again, that the distal end of the promoter was responsible for the lethality.

Within the distal end of the promoter is Conserved Region III. Conserved Region III is 998 base pairs in length and is the largest of the three conserved regions in the *elt-2* promoter. As depicted in figure 40 and 41, a Dot plot of Conserved Region III indicates that this region is highly conserved and may be sectioned into two parts: Section A (distal) and B (proximal). Since constructs containing the distal portion of the *elt-2* promoter were lethal and that Conserved Region III was within this region, it was presumed that Conserved Region III was in fact responsible for this lethality.

Previously Genefinder analysis had indicated that within the distal region of Conserved Region III in *C. elegans* there was a putative protein coding region that started -4180 base pairs upstream of the *elt-2* ATG start site (Figure 42). Genefinder analysis also indicated that the protein was present in all three species; although, the predicated splicing was different after the first exon that codes for the first 36 amino acids. I analyzed the first 36 amino acids of the protein between *C. elegans* and *C. remanei* and discovered that 24/36 amino acids were identical, while 28/36 amino acids were identical between *C. elegans* and *C. briggsae*. Interestingly, the third codon position of many of the conserved amino acids contained a high degree of synonymous changes (40% in *briggsae* and 56% in *remanei*) suggesting an evolutionary constraint at the protein level. BlastP searches were performed; however, the protein did not correspond to any known protein.

Since many reporter constructs contained the putative coding region, it was possible that overexpression of the protein was responsible for the lethal effect. I presumed that if the protein had an overexpression phenotype, it could also confer a lethal effect when removed. To determine if the putative protein region had a lethal effect when removed, I knocked out this region by both RNAi and a deletion allele. dsRNA corresponding to the ORF was injected into wild type worms. Worms were viable and showed no indication of any intestinal defects. The deletion allele *elt-2(gk153)* removes -4790 to -3860 bp upstream of the *elt-2* ATG site and thus removes the putative coding region (Figure 13). No defects were observed in mutant worms. Therefore, if the ORF encodes a protein, it is either redundant with another gene or is not necessary for worm survival. The above results indicate that the putative coding region is most likely not responsible for the lethal effects produced by the reporter constructs.

Section B of Conserved Region III contains the *elt-2* enhancer. It is possible that the enhancer is, to a certain extent, responsible for the lethality of the distal region. One explanation is that in the worm, each cell contains a limited amount of the proteins used for transcription. When the enhancer is present at high copy numbers a critical amount of these proteins are sequestered, and are no longer available to bind their endogenous targets in the cell. The cell therefore is no longer able to function correctly which leads to cell death and eventual death of the worm. However the question remains: Why are larger constructs that contain the enhancer not lethal to the worm? It remains possible that larger constructs contain sequences adjacent to the enhancer that may insulate the enhancer or actively inhibit protein binding.

In summary, it was difficult to determine where the lethality was arising from, since it was caused by more than one factor. Plasmid preparations were in part to blame for a lethal effect. The Gob phenotype indicated that constructs containing the *elt-2* gene were also causing lethality. Ultimately, the removal of the *elt-2* coding sequence uncovered the lethal effect from the enhancer region of *elt-2*.

CHAPTER FIVE: ANALYSIS OF CONSERVED REGION I AND II

Specific Aim VIII: To determine the function of Conserved Region I and II

Two way alignments between *C. elegans* and *C. remanei* as well as *C. elegans* and *C. briggsae* using the program EMBOSS Dotmatcher revealed that there are four regions of conservation upstream of *elt-2* (Figure 18+19). Region IV corresponds to the promoter region of the upstream *elt-4* gene. The three remaining regions are within the 5kb flanking region upstream of the *elt-2* gene and were numbered I, II, and III. Conserved Region III was previously discussed and contains the *elt-2* enhancer. Although Conserved Region I was presumed to contain the *elt-2* basal promoter, it was important to experimentally determine the function of Conserved Region II and I.

Conserved Region I confers *elt-2* transcriptional specificity

Conserved Region I (C. R. I) is 230 bp in length and is the closest conserved region to *elt-2* (-220 bp from *elt-2* ATG start site). In the 5' enhancer-specific deletion series, C. R. III was shown not to be sufficient to drive *elt-2* expression. In the 5' and 3' deletion series, expression was only seen from constructs that contained both C. R. III and I, suggesting that Conserved Region I was necessary for expression. Furthermore, the 5' and 3' deletion series indicated that a small level of ectopic staining derived from Conserved Region I. I wanted to examine the function of C.R. I independent of the 220 bp sequence located between C. R. I and the *elt-2* ATG start site. Construct pJM281 was made and contains C. R. I placed into the expression testing vector pJM76. pJM76 contains a heat shock basal promoter fused to GFP and *lacZ*. Transgenic worms containing pJM281 produced ectopic staining similar to pJM259. After hatching, worms expressed GFP in the first 4 anterior and in the last 2 posterior intestinal cells. Taken together, these results indicate that although Conserved Region I is not sufficient for *elt-2* expression, it cooperates with Conserved Region III to regulate *elt-2* expression.

I aligned Conserved Region I between all three species to locate conserved sites that may be potential sites of interest. Sequence alignments of C. R. I indicated that there are two highly conserved GATA sites (Figure 43). GATA site -315 (**ACTGATA**T) is completely conserved between all three species and is similar to an ELT-2 binding site.

GATA site -400 is the highest conserved sequence within C.R. I with 36/40 bp conserved between *C. elegans* and *C. remanei* and 33/35 bp conserved between *C. elegans* and *C. briggsae*.

What is the function of Conserved Region I? C.R. I most likely contains basal promoter binding sites necessary for the loading of general transcription machinery. However, the *elt-2* basal promoter also contains sites required for the specificity of *elt-2* transcription since *elt-2* expression was previously shown to be dependant on the presence of its own basal promoter. I therefore consider it likely that the *elt-2* basal promoter contains specific sites required for *elt-2* expression and that this specificity is due to the highly conserved GATA sites.

Conserved Region II is functionally redundant with Conserved Region I

Conserved Region II is 548 base pairs in length and extends from -2090 to -1542 bp upstream of the ATG site. In the 3' deletion series, Conserved Region II was removed with no change in expression, indicating that Conserved Region II was not necessary for *elt-2* expression. This result leads to the question: if the region is not necessary for *elt-2* expression why is this region still present in all three species?

As depicted in figure 44, the distal region of Conserved Region II is highly similar to the sequence of the Conserved Region IV. Since Conserved Region IV most likely corresponds to the *elt-4* basal promoter, it was important to determine if Conserved Region II also had the capacity of functioning as a basal promoter. To examine this possibility, I injected worms with the reporter construct pJM334 which contains Conserved Region III and II but not Conserved Region I (*elt-2* promoter -5098 to -516 bp inserted into *yfp* reporter construct), and compared their expression pattern to transgenics carrying the control reporter pJM333 (contains the complete *elt-2* promoter) (Figure 39). Both constructs produced strong YFP expression throughout the developing intestine suggesting that Conserved Region II is capable of functioning as a basal promoter.

I also wanted to determine whether Conserved Region II had a function independent of Conserved Region I and III. Transgenic worms containing the reporter construct pJM282 (C. R. II inserted into pJM76 expression testing vector) were examined for expression in embryogenesis and larval stages. While no expression was seen in

embryogenesis, adults at room temperature contained the same anterior/posterior expression seen in transgenics containing the reporter construct for Conserved Region I (Figure 45A). Taken together the results indicate that Conserved Region II: i) alone is not sufficient for *elt-2* expression; ii) has the capacity to function as an *elt-2* basal promoter; iii) is functionally redundant with Conserved Region I iv) can cooperate with Conserved Region III to drive *elt-2* expression.

Conserved Region II up-regulates *elt-2* expression in heat stress conditions

Interestingly, pJM282 transgenics placed at 30ºC for five hours have an increase in GFP expression compared to transgenic worms containing pJM282 kept at room temperature. Transgenic worms at room temperature contained GFP expression only in the anterior and posterior cells of the intestine while heat shocked counterparts had an expansion of GFP expression in posterior section of the intestine (Figure 45B). The increase in GFP expression from transgenics containing Conserved Region II was the first indication that Conserved Region II may up-regulate *elt-2* expression in response to heat stress conditions. As a control, transgenic worms containing Conserved Region I or III in the same reporter backbone as Conserved Region II were heat shocked but produced no such increase. Therefore, the increase in GFP expression only in transgenics containing Conserved Region II indicates that the increase in expression is not an artifact of the pJM76 vector backbone and that Conserved Region II is sufficient for the up-regulation of *elt-2* during heat stress conditions.

I next wanted to determine whether Conserved Region II was not only sufficient but necessary for the up-regulation of *elt-2* expression under heat shock conditions. Transgenics containing the reporter construct pJM256 (contains C. R. III, II, I) were heat shocked and had an increase in GFP expression. Transgenics containing the reporter construct pJM254 (contains C. R. III, the distal region of II, I) were also heat shocked and had a similar increase in expression. However, transgenics containing the construct pJM252 (contains only C. R. III and I) did not have an increase in GFP expression under heat shock conditions suggesting that the distal region of Conserved Region II is necessary for the up-regulation of *elt-2* in heat stress conditions (Figure 46).

I wanted to confirm that the increase in expression was due to an increase in transcription and not an artifact produced by the vector. More specifically, I was

concerned that the vector supplied 3' UTR (untranslated region) was stabilizing the mRNA transcript; which, in turn would increase the amount of protein in the cell. To directly test the effect of the 3' UTR, I would have had to alter pJM282 by replacing the vector supplied UTR with the *elt-2* UTR. Instead, I made use of the previously made reporter construct pJM188. pJM188 was chosen on three criteria: i) contains Conserved Region IV upstream of *elt-4*; ii) GFP inserted immediately upstream of the *elt-4* stop codon; iii) contains the endogenous *elt-4* 3' UTR. It was reasoned that since Conserved Region IV and II are highly similar and pJM188 contained the endogenous *elt-4* UTR an increase in expression in heatshocked worms would indicate that, the vector supplied 3' UTR in pJM282 was not responsible for the increase in expression. Transgenics containing pJM188 were placed at 30ºC for five hours and were found to have an increase in GFP expression. It is important to note that before this study, *elt-4* was presumed to have arisen from a gene duplication of *elt-2* but the function of the GATA factor was unknown. The up-regulation of *elt-4* transcription may indicate that *elt-4* has an active role in the transcriptional regulation of intestinal genes when the worm is under a heat stress state.

The distal region of Conserved Region II was analyzed for known binding sites like the heat response elements TTCTAGAA and GGGTGTC; however, no such elements were found (Guhathakurta *et al.,* 2002). Conserved Region II was then compared between all three species to locate potentially important sites (Figure 47). In the distal region of Conserved Region II, sequence alignments located five conserved GATA sites (Figure 47A). GATA site -2079 was determined to have the highest sequence conservation with 66/70 bp conserved between *C. elegans* and *C. briggsae* and 63/72 bp between *C. elegans* and *C. remanei*. To elucidate whether ELT-2 was responsible for the up-regulation of *elt-2* through Conserved Region II, dsRNA corresponding to *elt-2* or *acn-1* was injected into pJM282 worms. F1 worms were left at room temperature for two days and then placed overnight at 30ºC. F1 worms from mothers injected with dsRNA corresponding to the *elt-2* gene had the same expression pattern as compared to F1 worms from mothers injected with the control (dsRNA corresponding to *acn-1*). Therefore, the up-regulation of *elt-2* through Conserved Region II is not regulated by ELT-2 and is regulated by an unknown factor.

CHAPTER SIX: OXIDATIVE STRESS

Increased GFP expression in heat treated transgenic worms containing pJM282 indicated that *elt-2* may be up-regulated in stress conditions. This section gives a brief overview of what is currently known on oxidative stress and describes a series of experiments done to determine whether *elt-2* is necessary for the activation of a stress response in the intestine. However it should be noted that the project was eventually terminated due to the variability of expression from transgenes.

Introduction

Organisms are affected by their surroundings and have developed mechanisms to respond to changes in their environment. The stress response has evolved to protect organisms against thermal changes, food shortage, overpopulation, heavy metal stress, UV irradiation, and oxidative stress. While high stress levels are lethal to an organism, moderate environmental stresses can increase the lifespan of an organism (Johnson and Hartman, 1988; Lithgowet *et al.,* 1995). In this section, I will review the basic mechanism of removing reactive oxygen species from the cell as well as recent findings on an oxidative stress response in the *C. elegans* intestine.

Increased resistance to oxidative damage has been shown to increase lifespan. For example, transgenic over-expression of antioxidant genes like superoxide dismutase (SOD) and catalase in *Drosophila* has been shown to increase lifespan (Sun and Tower, 1999; Orr and Shoal, 1994**).** In worms, the addition of EUK-134 and EUK-8 (synthetic SOD/catalase mimetics) increases mean and maximum lifespan (Melov *et al.,* 2000). Genetically, mutants with increased lifespan have an up-regulation of antioxidant genes that code for peroxisomal catalase (*ctl-2*), cytosolic catalase (*ctl-1*), and manganese superoxide dismutase (*sod-3*) (Murphy *et al.,* 2003). Worms lacking stress response genes have a decreased lifespan, supporting the theory that oxidative damage caused by reactive oxygen species in the cell is the underlying mechanism of aging (Liao and Yu, 2005; Harman, 1956; Harman, 1992).

Reactive oxygen species (ROS) like superoxide radicals, hydrogen peroxide and hydroxyl radicals are generated during normal cellular metabolism. ROS are produced during aerobic respiration when electrons are transferred along the respiratory chain to

generate ATP. Usually oxygen is the final electron acceptor, yielding water. However, ROS can be produced as a result of inappropriate electron donation and are available to react and damage DNA and proteins (Fridovich, 1995; Stadtman, 1992).

The cell counteracts ROS with protective antioxidants. Antioxidants are mainly enzymatic and include superoxide dismutase, glutathione peroxidase and catalase (Klauniung, 2001). Superoxide dismutase removes superoxide radicals by catalyzing their conversion into hydrogen peroxide (Fridovich, 1995). Catalase then breaks down hydrogen peroxide into oxygen and water.

Hydrogen peroxide is also removed from the cell by oxidizing glutathione (GSH). The oxidation of GSH to GSSG is done mostly by Glutathione peroxidase (Betteridge, 2000; Cotgreave *et al.,* 1988). Glutathione reductase regenerates GSH from GSSG, with NADPH as the source of reducing power. GSH is also used as a cofactor for GSH peroxidase in the reduction of peroxides (Flohe, 1978). Therefore glutathione is a critical component in removing reactive oxygen species from the cell.

Glutathione is made by γ-Glutamylcysteine synthetase (GCS) and glutathione synthetase (Meister and Anderson, 1983). GCS catalyzes the first rate limiting step in glutathione biosynthesis and is composed of a catalytic heavy chain subunit and a regulatory subunit (Richman and Miester, 1975). The catalytic subunit in C. *elegans* is encoded by the *gcs-1* gene while the regulatory subunit is encoded by E01A2.1 (Soltaninassab *et al.,* 2000).

Most antioxidant genes are regulated through an anti-oxidant response element (GCNNNG/ATCAT/C) found within their promoters (Rushmore, 1990; Nguyen, 2000; Wasserman, 1997; Frilling *et al.,* 1992; Jaiswal, 1991; Favreau *et al.,* 1991). In humans, the *gcs-1* gene is regulated through an anti-oxidant response element (ARE) sequence found in its promoter (Mulcahy, 1997; Moinova, 1998). In *C. elegans,* intestinal expression of *gcs-1* is regulated through a modified ARE in its promoter, indicating transcriptional regulation of antioxidant genes through an ARE element has been evolutionarily conserved.

C. elegans kept at standard room temperature conditions, express *gcs-1* in the pharynx, ASI chemosensory neurons and in the anterior and posterior region of the intestine (An and Blackwell, 2003). The expression of *gcs-1* in each tissue (intestine,

pharynx and ASI neurons) is regulated through different sites within the 1840bp promoter (Figure 48). Pharyngeal expression is regulated through distal sites on the promoter while intestinal expression is regulated through an ARE motif found more proximal to the gene. Interestingly, it is this ARE motif that confers an expansion of *gcs-1* expression throughout the intestine when worms are exposed to heat or to the herbicide paraquat. Furthermore, the ARE motif (ACTTTATCATCAT) contains an overlapping SKN-1/ELT-2 binding site (Figure 49). Electrophoretic mobility shift assays indicate that SKN-1 binds to the sequence *in vitro* (An and Blackwell, 2003). *In vivo*, a *skn-1* deletion eliminates *gcs-1* transgene expression in the gut. Site directed mutations against the overlapping SKN-1/ELT-2 binding site abolishes intestinal expression *in vivo.* Although the study determined that SKN-1 was necessary for *gcs-1* expression, it did not investigate whether ELT-2 was also necessary for *gcs-1* expression. Therefore ELT-2 may also regulate *gcs-1* intestinal expression and may also be responsible for the regulation of other oxidative stress genes in the intestine.

Although *skn-1* is transcribed in the cells of the pharynx, hypodermis, intestine and ASI neurons, the SKN-1 protein is only present in the ASI neurons and at modest levels in intestinal cells of worms kept at standard room temperature conditions. However, when adult worms are heat stressed or treated with 50mM sodium azide, SKN-1 is present in intestinal nuclei five minutes after treatment. The investigators suggest that under stress conditions, the *skn-1* transcripts are rapidly made into protein in order to activate directly oxidative stress genes like *gcs-1*.

A collaborative study with the Blackwell lab set out to determine whether ELT-2 regulated transcription of *gcs-1* in the intestine. More specifically, if ELT-2 regulated *gcs-1* expression, I wanted to determine whether ELT-2 directly regulated *gcs-1* expression by binding to the ARE sequence or whether ELT-2 regulated *gcs-1* expression indirectly by regulating *skn-1* expression.

Results

Specific Aim IX: To determine whether ELT-2 regulates the transcription of stress response genes in the intestine

Before determining whether ELT-2 regulated *gcs-1* expression in the intestine, I first wanted to determine whether gcs-1 was the only stress response gene to contain a potential ELT-2 binding site. I analyzed the 5' untranslated regions of 11 stress response genes for potential ELT-2 binding sites. Nine of the eleven stress response genes analyzed, contain at least one potential ELT-2 binding site (Table 20). When calculated, the frequency of an ELT-2 binding site (WGATAR) occurring randomly in the genome is every 1024 bp. The number of potential ELT-2 binding sites located within the promoters of stress response genes was found at a greater frequency than the frequency of site occurring randomly suggesting that these sites have undergone positive selection (Table 20). Interestingly, ELT-2 binding sites were also located on the *skn-1* promoter. The above findings suggest that ELT-2 could act independently, or cooperatively with SKN-1 to bind the ARE sequence of oxidative stress genes, to regulate their expression in the intestine. Another possibility is that ELT-2 may regulate oxidative stress genes indirectly through the regulation of *skn-1*.

As indicated in Table 20, the stress response gene *gcs-1* contains 6 potential ELT-2 binding sites within the promoter. More importantly, *gcs-1* intestinal expression was previously determined to be regulated through an overlapping SKN-1/ELT-2 binding site (An and Blackwell, 2003). To determine whether *elt-2* was necessary for *gcs-1* transgene expression, transgenic worms containing the construct *gcs*Δ4::GFP were ordered from Hyung An in the Blackwell lab. The *gcs*Δ4::GFP transgene contains the most proximal 163 bp of the total 1840 bp *gcs-1* promoter, which includes the overlapping SKN-1/ELT-2 site responsible for intestinal expression (Figure 48).

The Blackwell lab published the intestinal expression pattern of the transgene *gcs*Δ4::GFP (An and Blackwell, 2003). The expression pattern from transgenic worms containing *gcs*Δ4::GFP was noted as similar between worms incubated at 29ºC for 20 hours or incubated at 34ºC for two to four hours. To determine whether I could replicate the results, I exposed transgenic worms containing *gcs*Δ4::GFP to similar conditions. Unfortunately expression was not similar between worms, worms with longer incubation

times at 30ºC always had higher levels of GFP expression compared to worms incubated at 34ºC for 5 hours. Although not described by An and Blackwell in the published article, GFP expression was variable within cohorts, such that some worms had strong expression in all cells of the intestine while some had weaker, more variable expression. Longer incubation times decreased the variability of GFP expression between worms; however, expression remained variable.

Although expression from heat treated worms containing the *gcs*Δ4::GFP transgene was variable, it remained possible to determine whether *elt-2* was necessary for *gcs-1* expression in the intestine. dsRNA corresponding to *elt-2* and *acn-1* was injected into *gcs*Δ4::GFP worms. F1 worms were kept at room temperature for two days and then heat shocked overnight at 30ºC. GFP expression in F1 transgenic worms remained, indicating that *elt-2* is not necessary for *gcs-1* expression in the gut. However, F1 worms from *elt-2* RNAi injections appeared to have a decrease in GFP expression compared to *acn-1* control worms suggesting that *elt-2* may have a role in regulating *gcs-1* expression. It is important to note that the decrease in expression from *elt-2* RNAi worms is not an artifact of the lethal effect of *elt-2* RNAi since both *elt-2* and *acn-1* RNAi are lethal to the worm.

To further assess whether *gcs-1* expression in the intestine was regulated by *elt-2*, I ordered the *gcs-1* construct containing the full length promoter (Figure 48). Transgenic worms containing the full length reporter construct have GFP expressed in both the pharynx and intestine. Pharyngeal expression served as an internal control in *elt-2* RNAi experiments since worms always had equal to stronger expression in the intestine than the pharynx. dsRNA corresponding to *acn-1*, *elt-2* and *skn-1* were injected into transgenics containing the full length promoter construct. Although expression remained variable between worms, *elt-2* RNAi treated worms had weak to undetectable GFP expression in the gut compared to pharyngeal expression (Figure 50). These results indicated that although ELT-2 is not necessary for *gcs-1* expression, it does have an impact on *gcs-1* expression.

Intestinal expression of *gcs-1* is regulated through an overlapping SKN-1/ELT-2 binding site (An and Blackwell, 2003). I next asked whether RNAi performed against both *skn-1* and *elt-2* had a different effect on *gcs-1* expression than each alone. However,

skn-1 RNAi was found to be ineffective in removing *skn-1* and therefore *elt-2*/*skn-1* RNAi was not performed. As a result, it was never determined whether *elt-2* and *skn-1* worked cooperatively in regulating *gcs-1* expression.

It was previously shown that SKN-1 was necessary for *gcs-1* expression (An and Blackwell, 2003). It was important to ask whether ELT-2 indirectly regulated *gcs-1* by regulating *skn-1* in the intestine. To test this hypothesis we performed *elt-2* RNAi on *skn*PRO::*gfp* transgenics. The *skn*PRO::*gfp* construct contains the 5' untranslated region of *skn-1* fused to GFP and is expressed in the intestine, hypodermis and pharynx (Figure 53). dsRNA corresponding to *elt-2* was injected into transgenics containing the *skn*PRO::*gfp* construct. Unfortunately, F1 worms had variable GFP expression, which included worms without a change in GFP expression to worms with GFP expression everywhere except the intestine (Figure 51). The result indicated one of two possibilities; *skn-1* expression in the intestine is regulated by ELT-2, or lack of expression in the intestine is due to transgene mosaic expression. Therefore, I obtained the *skn-1* antibody to look at SKN-1 protein levels in *elt-2* RNAi worms. Although the antibody has been previously shown to work in embryos, I was unable to see SKN-1 staining in adults despite optimization of the antibody staining protocol. Therefore all of these confounding factors resulted in the inability to conclude whether *elt-2* directly regulates *skn-1* expression in the intestine.

In summary, although the stress response section of the project was problematic, preliminary information implied that *elt-2* may play a role in an oxidative stress response in the intestine. First, many of the stress response genes have potential ELT-2 binding sites in their promoters. Second, the stress response gene *gcs-1* has weakened expression in *elt-2* RNAi experiments, suggesting that *elt-2* may regulate its expression.

CHAPTER SEVEN: DISCUSSION

Although we currently understand the general transcription factor framework that determines intestine development, the entire pathway is by no means completely understood. For example, even though we know the main players in the regulatory cascade from SKN-1 to MED-1,2 to END-1,3 to ELT-2, we do not presently understand how the transition from one transcription factor to the next takes place. To truly understand the regulatory cascade, not only is it necessary to know the players but also to understand their interactions. The promoters of the *end-1,3* and *med-1,2* genes have been analyzed in order to understand the interactions in the cascade up to the point of gut specification (Coroian *et al.,* 2005; Morris *et al.,* 2005). In turn, promoter analysis of gut specific genes regulated by ELT-2 have started to elucidate how gut differentiation occurs (Marshall and McGhee, 2001; Fukushige *et al.,* 2005). However until now, no studies have tried to understand the transitional part of the cascade, the point where the gut has been specified and is primed to differentiate. More specifically, it was unclear whether the gut specification transcription factors END-1 and END-3 directly initiated expression of the gut differentiation gene *elt-2*, and whether ELT-2 directly regulated its expression by binding to sites within its promoter.

Promoter size and regulatory complexity

The 5' flanking region upstream of *elt-2* contains over 5000 bp. Comparisons of 5' flanking regions of intestinal specification genes like *end-1* and *end-3*, as well as genes like *ges-1* and *pho-1* that are expressed in the intestine and are regulated by *elt-2* indicates that the regulatory region of *elt-2* is substantially larger. The size of the *elt-2* 5' flanking region is more comparable to genes that are key regulators to not only the intestine but to other organs as well. For example *skn-1*, a key regulator of intestinal development (80% of embryos mutant for *skn-1* do not produce an intestine), is part of an operon and has a large intronic four kb region. *pha-4*, which regulates the formation of the foregut –or pharynx; has a 5' flanking region of approximately eight kb.

The question arises: why is the *elt-2* promoter more similar in size to the promoters of genes like *pha-4* and *skn-1*? One explanation is that *elt-2,* like *pha-4* and *skn-1,* governs the expression of numerous genes and must also interpret its own

expression level in the cell. The idea that regulators that orchestrate multiple developmental processes tend to have larger intergenic regions surrounding the respective gene is supported by an analysis that was done on the genomes of *C. elegans* and *Drosophila*. The study found a positive correlation between the inferred regulatory complexity of a gene and the size of the noncoding DNA region flanking the gene (Nelson *et al.,* 2004). Another explanation for the larger regulatory region of *elt-2* compared to other *C. elegans* GATA factors is that ELT-2 is expressed early in development and is maintained throughout the life of the worm, unlike both *med* and *end* genes that are expressed transiently. However, until the *elt-2* 5' flanking region was thoroughly dissected and analyzed, it was pure speculation as to the complexity of the region.

Before determining whether the complete 5' flanking region was necessary for *elt-2* expression, I compared the upstream region of *elt-2* to the same region in the *elt-2* orthologues in *C. remanei* and *C. briggsae.* Four regions were found to be highly conserved between the three species, and were denoted Conserved Region IV, III, II, I. Conserved Region III, II, I are located within the 5144 bp upstream of *elt-2*. Conserved Region IV is located upstream of the *elt-4* gene and has a high degree of sequence similarity to the distal region of Conserved Region II. The sequence conservation located throughout the promoter supported the proposed complexity of *elt-2* regulation.

The *elt-2* 5' flanking region is responsible for *elt-2* embryonic expression

The recently described analysis of the *C. elegans* and *D. melanogaster* genomes indicated that *C. elegans* usually partitions its regulatory information upstream of the promoter, whereas no strong bias was apparent in flies (Nelson *et al.,* 2004). Comparison of the ELT-2 antibody staining to the expression from the transgene pJM67 (contains the complete *elt-2* 5' flanking region fused to GFP/*lacZ*) indicated that the *elt-2* 5' flanking region is responsible for *elt-2* embryonic expression. Interestingly, the 5' flanking region is not responsible for *elt-2* expression in larval stages suggesting that *elt-2* expression in larval stages may be dependant on sites outside of the 5' flanking region. Although sequence alignments between *elt-2* and its orthologues did not reveal any highly

conserved regions within the introns or downstream of *elt-2*, it remains possible that these regions may contain sites important for the regulation of *elt-2*.

elt-2 is cooperatively regulated by Conserved Region III and I.

In the present study, a series of reporter constructs were used to locate the minimum promoter region necessary for *elt-2* embryonic expression. It was determined that Conserved Region I and the proximal region of Conserved Region III are necessary for *elt-2* expression. The location of an enhancer within Conserved Region III suggests that *elt-2* is not regulated in same manner as the GATA factors *end-1,3* and *med-1,2*. GATA factors *end-1,3* and *med-1,2* do not contain enhancers and are solely regulated by regions that are proximally located to each gene (Coroian *et al.,* 2005; Morris *et al.,* 2005). The location of an enhancer over 3000 bp away from the *elt-2* gene suggests that *elt-2* may have a regulatory complexity similar to GATA factors in higher organisms. For example, in vertebrates, expression of cGATA-5 has been shown to be regulated by a bipartite enhancer. The distal region the enhancer is located -9.6 to -6.3 kb upstream of the gene and regulates early embryonic gut expression; the proximal region is located -5 to -4.5 kb upstream of the gene and is responsible for late embryonic gut expression. However, unlike the vertebrate study, our study determined a more defined enhancer region (maximum of 543 bp). Our study also determined that *elt-2* expression was dependant on its own basal promoter (Conserved Region I). Previous expression studies in invertebrates/vertebrates have not tested whether the endogenous basal promoter of GATA genes are needed for expression. This is first study that has tested and shown that expression of a GATA factor is dependant on its own endogenous promoter.

Conserved Region III contains the *elt-2* enhancer

The critical enhancer region is located within the proximal region of Conserved Region III. In this study, two sites were identified within the *elt-2* enhancer that were: i) also present within the enhancer region of the *elt-2* orthologues in *C. briggsae* and *C. remanei*; ii) corresponded to a known transcription factor binding site and; iii) bound such transcription factors *in vitro*. The two sites, GATA site -3820 and GATA site -3394,

were mutated independently within *elt-2* promoter::*gfp/lacZ* constructs and produced both a decrease and lack of GFP expression respectively.

END-1 and END-3 may bind to GATA site -3820 to initiate *elt-2* expression

GATA site -3820 was present in all three species, and was bound by both END-1 and ELT-2 *in vitro*. GATA site -3820 was tested for its *in vivo* role in the regulation of *elt-2* by mutating the site within two *elt-2* promoter::*gfp/lacZ*.constructs. The first construct contained only the distal region of the *elt-2* promoter and produced no GFP/β-galactosidase expression. The second construct contained GATA site -3820 mutated within the full length *elt-2* promoter and produced expression; albeit, weaker in embryogenesis than control embryos that contained the full length *elt-2* promoter with no altered GATA -3820. One interpretation is that GATA site -3820 is the primary site for initiating *elt-2* expression and that the onset of a weaker expression is due to a secondary, more proximal, site present in only the full length promoter construct. A previously unpublished result indicates that ectopic expression of END-1 throughout the embryo can induce, by way of the *elt-2* enhancer region, ectopic expression of ELT-2 through out the embryo (Fukushige, personnel communication). However, it was unknown whether END-1 or END-3 (functionally redundant) initiated *elt-2* expression by directly binding the enhancer. Our study has determined that END-1 (and therefore END-3) can bind to a site within the enhancer and that this site, GATA -3820, is important for *elt-2* expression.

GATA site -3394 is necessary for *elt-2* expression

GATA site -3394 binds ELT-2 *in vitro* and is the only site within the *elt-2* promoter necessary for expression. The 5' and 3' deletion construct series and site directed mutagenesis experiments indicated that removing this site abolished *elt-2* expression. One explanation for the lack of expression is that ELT-2 binds GATA site -3394, and that this interaction is necessary for *elt-2* expression. GATA site -3394 contains the consensus ELT-2 binding motif that has been previously identified as a binding/regulatory site of key intestinal genes. Our study also confirmed that this motif does bind ELT-2 *in vitro* and is important for gene expression.

Could GATA site -3394 be the site responsible for *elt-2* autoregulation? At first it seemed unlikely since a previously unpublished experiment had shown that ectopic expression of ELT-2 throughout the embryo did not direct ectopic expression of an *elt-2*enhancer::*lacZ* construct (Fukushige, personnel communication). However, I have shown that the *elt-2* enhancer is not sufficient to direct expression and that at least one other site is necessary for *elt-2* embryonic expression (within Conserved Region II or I). I purpose that the lack of ectopic expression from the *elt-2*enhancer::*lacZ* construct is due to the absence of this second necessary site. I therefore predict that ELT-2 does indeed bind to GATA site -3394 to autoregulate its expression and that ectopic expression of ELT-2 throughout the embryo would drive ectopic expression of a construct containing the *elt-2* enhancer and Conserved Region I (or II).

Site 1360 may be an important site for the down-regulation of *elt-2* expression

I also located two highly conserved sequences that had no sequence similarity to any known transcription factor binding sites (Site 1601 and 1360). Although mutations within Site 1601 did not effect *elt-2* expression, altering Site 1360 increased *elt-2* expression. The increase in expression from transgenics containing the altered Site 1360 suggests that, under endogenous conditions, a negative regulator uses Site 1360 (unaltered) to decrease *elt-2* expression. Furthermore, in the absence of ELT-2 protein, expression increased exponentially in 1360 transgenics confirming the ELT-2 is not the only positive regulator of *elt-2* expression.

Conserved Region I contains the *elt-2* basal promoter

Conserved Region I is located close to the *elt-2* ATG start site and most likely contains a basal promoter used to bind general transcription machinery (i.e. RNA polymerase). However, this study reveals that Conserved Region I is specialized to cooperate with the *elt-2* enhancer to direct *elt-2* expression. Analysis of Conserved Region I indicates that there are three highly conserved sites within this region. While site -355 has no sequence similarity to any known transcription factor binding site, sites -400 and -315 have sequence similarity to GATA transcription factor binding sites. One interpretation is that ELT-2 binds to these two sites within the enhancer to direct proper

elt-2 expression. In all the constructs used in this study which contained the *elt-2* basal promoter but not the enhancer, ectopic sporadic expression was seen. It remains possible that this sporadic expression is due to endogenous *elt-2* binding to these sites and sporadically starting transcription. Interestingly, *end-1* and *end-3* both contain potential GATA transcription factor binding sites within the proximal region of their promoters. Furthermore, removal of a single GATA site located -414 bp upstream of the *end-1* gene revealed that the site contributes positively to *end-1* activation (Maduro *et al.,* 2005). Although it remains to be tested whether the GATA sites within Conserved Region I are important for *elt-2* expression, it will be interesting to see if a reporter construct containing Conserved Region III and I, and mutations in either GATA -400 and -315 will have a wild-type expression pattern.

Conserved Region II and the heat shock response

Conserved Region II, although highly conserved, is not necessary for *elt-2* expression. However, Conserved Region II can substitute as the *elt-2* basal promoter in the absence of Conserved Region I. Sequence analysis of Conserved Region II indicates that there are seven conserved potential *elt-2* binding sites within the 562 bp region. Due to the above findings, it remains possible that GATA sites within the basal promoter are responsible for its specificity. It was also demonstrated that the distal region of Conserved Region II up-regulated *elt-2* expression in heat stress conditions. Interestingly, Conserved Region IV (*elt-4* basal promoter) is highly conserved to the distal portion of Conserved Region II and up-regulates *elt-4* expression in heat induced conditions, suggesting a potential role of *elt-4* in heat stress conditions.

MODEL

As previously discussed in the introduction, there are two main current models of transcriptional regulation. The first model, the looping model, suggests that factors present on the enhancer may reel in the chromatin in search of the promoter (West and Fraser, 2005). The reeling of the DNA produces a looping of the DNA between the enhancer and promoter. Enhancer bound factors interact with factors present on the basal promoter; but do not themselves directly bind to the basal promoter region. The tracking model suggests that proteins present on the enhancer form an activation complex that tracks along the promoter until reaching the basal promoter (Tuan *et al.,* 1992; Blackwood *et al.,* 1998). Once on the basal promoter, the complete transcriptional apparatus is assembled and the gene is transcribed. One proposed scenario is that *elt-2* expression is regulated by a hybrid of these two models. First, END-1 initiates *elt-2* expression by binding to GATA site -3820 (Figure 52). Since GATA factors are thought to initiate transcription by first binding and opening compacted chromatin and then stabilizing loop conformation (Cirillo *et al.,* 2002; Johnson *et al.,* 2001; Letting *et al.,* 2001); END-1 may cause a change in chromatin conformation and arrange the enhancer close to the basal promoter permitting the RNA polymerase unit to initiate *elt-2* transcription. The ability of END-1 to initiate *elt-2* expression in the absence of the endogenous basal promoter indicates that END-1 only needs to bind the *elt-2* enhancer to initiate expression. *elt-2* expression is maintained by END-1 until ELT-2 protein is present. ELT-2 then initiates autoregulation by binding GATA site -3394 and possibly GATA site -3820 as well as site(s) within Region I or II. Our study indicates that *elt-2* has conserved GATA binding sites present on the enhancer as well as the basal promoter suggesting that ELT-2 directly binds to both regions. Also the inability of the enhancer to independently autoregulate expression supports the idea that the conserved ELT-2 binding sites within the basal promoter are necessary. It remains possible that ELT-2 maintains the open chromatin structure by binding both the enhancer and the basal promoter. Although this model suggests that embryonic expression of *elt-2* is regulated by the enhancer and its basal promoter, it is important to note that *elt-2* expression in larval stages is regulated by at least one other site not present in the 5' flanking region.

CHAPTER EIGHT: FUTURE EXPERIMENTS

Some key general questions remain to be answered. Although I have determined that two sites GATA -3820 and GATA -3394 are important for *elt-2* expression and bind END-1 (GATA site -3820) and ELT-2 (GATA site -3820 and -3394) *in vitro*, I have not shown whether these sites bind these factors *in vivo*. I have also not characterized GATA site -400 and -315 that are located within Conserved Region I. Fortunately, in worms due to the transparent nature of the animal, direct binding of proteins to their respective sites can be visualized *in vivo*. As discussed in the introduction, ELT-2 has been visualized co-localizing to transgenic arrays containing the full length *elt-2* promoter (Fukushige *et al.,* 1999). Below, is a series of experiments that have been designed to test the ability of END-1, END-3, ELT-2 to bind to the *elt-2* promoter *in vivo*. Although the section is somewhat descriptive, it is important to show that this study is primed for a very interesting ending.

GATA site -3820 is bound by END-1 *in vitro*. To test whether END-1 binds GATA -3820 *in vivo*, I have made an END-1::RFP construct. The construct contains the full length END-1 promoter and gene fused to the red fluorescent protein coding sequence. The END-1::RFP construct will be injected into pJM252 transgenics (contains distal region of the *elt-2* promoter fused to *gfp::lac-z*) and pJM308 transgenics (pJM252 with altered GATA site -3820). I expect that END-1::RFP protein will bind to GATA site -3820 *in vivo.* This will be visualized in pJM252 transgenics which will contain two red foci within each nucleus of intestinal cells. I also fully expect that transgenics containing pJM308 (altered GATA site -3820) will not contain red foci since the END-1 binding site is not present.

An ELT-2::GFP construct containing the *elt-2* promoter and coding region has been previously constructed and used to determine that ELT-2 was able to bind its own promoter *in vivo* (Fukushige *et al.,* 1999). The ELT-2::GFP construct will be used to further our studies with GATA site -3394 and Conserved Region I and II. First, the ELT-2::GFP construct will be injected into pJM251 transgenics (distal region of *elt-2* promoter inserted into pCR2.1-Topo vector) and pJM319 transgenics (pJM251 with altered GATA site -3394). I expect results similar to the END-1 nuclear spot assay; in that, ELT-2::GFP

will produce two foci per gut nucleus in the pJM251 transgenics but not in pJM319 transgenics. This experiment will confirm that ELT-2 binds GATA site -3394 *in vivo*.

One of the key remaining questions of whether ELT-2 binds directly to its basal promoter can be answered using the ELT::GFP construct. The ELT-2::GFP construct will be injected into worms containing pJM281 (Conserved Region I). I expect that ELT::GFP will co-localize with the transgene pJM281 and produce two foci per gut nucleus, suggesting that ELT-2 binds to the conserved GATA sites within the basal promoter. In a final experiment, ELT-2::GFP and END-1::RFP constructs will both be injected into transgenics containing the full length *elt-2* promoter. I believe that a temporal relationship will be seen. RFP will be expressed early in development with a very short window in which red foci will be present. I expect that shortly after *elt-2* is expressed (2E cell stage) green foci will dominate. Since END-1 is currently thought to express until the 8E cell stage, this experiment will determine whether END-1 maintains *elt-2* expression until the 8E cell stage or whether ELT-2 actively removes END-1 from the *elt-2* promoter to activate autoregulation.

REFERENCES

Afouda, B. A., Ciau-Uitz, A., and Patient, R. (2005). GATA4, 5 and 6 mediate TGFbeta maintenance of endodermal gene expression in Xenopus embryos. *Development* **132,** 763-74.

An, J. H., and Blackwell, T. K. (2003). SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. *Genes & Development* **17,** 1882-93.

An, J. H., Vranas, K., Lucke, M., Inoue, H., Hisamoto, N., Matsumoto, K., and Blackwell, T. K. (2005). Regulation of the Caenorhabditis elegans oxidative stress defense protein SKN-1 by glycogen synthase kinase-3. *Proc Natl Acad Sci U S A* **102,** 16275-80.

Arceci, R. J., King, A. A., Simon, M. C., Orkin, S. H., and Wilson, D. B. (1993). Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol Cell Biol* **13,** 2235-46.

Batchelder, C., Dunn, M. A., Choy, B., Suh, Y., Cassie, C., Shim, E. Y., Shin, T. H., Mello, C., Seydoux, G., and Blackwell, T. K. (1999). Transcriptional repression by the Caenorhabditis elegans germ-line protein PIE-1. *Genes & Development* **13,** 202-12.

Bell, C. E., Barry, J., Matthews, K. S., and Lewis, M. (2001). Structure of a variant of lac repressor with increased thermostability and decreased affinity for operator. *J Mol Biol* **313,** 99-109.

Benian, G. M., L'Hernault, S. W., and Morris, M. E. (1993). Additional sequence complexity in the muscle gene, unc-22, and its encoded protein, twitchin, of Caenorhabditis elegans. *Genetics* **134,** 1097-104.

Betteridge, D. J. (2000). What is oxidative stress? *Metabolism* **49,** 3-8.

Blackwell, T. K. (2004). Germ cells: finding programs of mass repression. *Curr Biol* **14,** R229-30.

Blackwell, T. K., Bowerman, B., Priess, J. R., and Weintraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. *Science* **266,** 621-8.

Blackwell, T. K., and Walker, A. K. (2002). Getting the right dose of repression. *Genes Dev* **16,** 769-72.

Blackwell, T. K., and Walker, A. K. (2003). Transcription elongation: TLKing to chromatin? *Curr Biol* **13,** R915-6.

Blackwood, E.M,, Kadonaga J.T.(1998). Going the distance: Current view of enhancer action. *Science*. **281,** 61-63.

Boag, P. R., Nakamura, A., and Blackwell, T. K. (2005). A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in C. elegans. *Development* **132,** 4975-86.

Boskey, A. L., Gadaleta, S., Gundberg, C., Doty, S. B., Ducy, P., and Karsenty, G. (1998). Fourier transform infrared microspectroscopic analysis of bones of osteocalcin-deficient mice provides insight into the function of osteocalcin. *Bone* **23,** 187-96.

Bossinger, O., Fukushige, T., Claeys, M., Borgonie, G., and McGhee, J. D. (2004). The apical disposition of the Caenorhabditis elegans intestinal terminal web is maintained by LET-413. *Dev Biol* **268,** 448-56.

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,** 443-52.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* **50,** 349-83.

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

Broitman-Maduro, G., Maduro, M. F., and Rothman, J. H. (2005). The noncanonical binding site of the MED-1 GATA factor defines differentially regulated target genes in the C. elegans mesendoderm. *Dev Cell* **8,** 427-33.

Bronckers, A. L., Price, P. A., Schrijvers, A., Bervoets, T. J., and Karsenty, G. (1998). Studies of osteocalcin function in dentin formation in rodent teeth. *Eur J Oral Sci* **106,** 795-807.

Brown, S., and Castelli-Gair Hombria, J. (2000). Drosophila grain encodes a GATA transcription factor required for cell rearrangement during morphogenesis. *Development* **127,** 4867-76.

Brunner, E., Peter, O., Schweizer, L., and Basler, K. (1997). pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. *Nature* **385,** 829-33.

Busson, M., Djoulah, S., Karsenty, E., Bleux, H., Bouteiller, A. M., and Charron, D. (1998). Proposal for a new classification of HLA-DR alleles based on electric charges of pockets of amino acid residues. *Transplant Proc* **30,** 2855-6.

Carmi, I., Kopczynski, J. B., and Meyer, B. J. (1998). The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex.[see comment]. *Nature* **396,** 168-73.

Carroll, A. S., Gilbert, D. E., Liu, X., Cheung, J. W., Michnowicz, J. E., Wagner, G., Ellenberger, T. E., and Blackwell, T. K. (1997). SKN-1 domain folding and basic region monomer stabilization upon DNA binding. *Genes Dev* **11,** 2227-38.

Chandler, V. L., Maler, B. A., and Yamamoto, K. R. (1983). DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. *Cell* **33,** 489-99.

Chen, D., Ji, X., Harris, M. A., Feng, J. Q., Karsenty, G., Celeste, A. J., Rosen, V., Mundy, G. R., and Harris, S. E. (1998). Differential roles for bone morphogenetic protein (BMP) receptor type IB and IA in differentiation and specification of mesenchymal precursor cells to osteoblast and adipocyte lineages. *J Cell Biol* **142,** 295-305.

Cho, S., Jin, S. W., Cohen, A., and Ellis, R. E. (2004). A phylogeny of caenorhabditis reveals frequent loss of introns during nematode evolution. *Genome Res* **14,** 1207-20.

Coghlan, A., and Wolfe, K. H. (2002). Fourfold faster rate of genome rearrangement in nematodes than in Drosophila. *Genome Res* **12,** 857-67.

Coroian, C., Broitman-Maduro, G., and Maduro, M. F. (2006). Med-type GATA factors and the evolution of mesendoderm specification in nematodes. *Dev Biol* **289,** 444-55.

Corral, D. A., Amling, M., Priemel, M., Loyer, E., Fuchs, S., Ducy, P., Baron, R., and Karsenty, G. (1998). Dissociation between bone resorption and bone formation in osteopenic transgenic mice. *Proc Natl Acad Sci U S A* **95,** 13835-40.

Cotgreave, I. A., Moldeus, P., and Orrenius, S. (1988). Host biochemical defense mechanisms against prooxidants. *Annu Rev Pharmacol Toxicol* **28,** 189-212.

Cirillo, L. A., Lin, F. R., Cuesta, I.., Friedman, D., Jarnik. M., Zaret K.S. (2002) Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell.* **9,** 279-278.

Ducy, P., and Karsenty, G. (1998). Genetic control of cell differentiation in the skeleton. *Curr Opin Cell Biol* **10,** 614-9.

Evans, T., and Felsenfeld, G. (1989). The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* **58,** 877-85.

Fath, K. R., and Burgess, D. R. (1995). Microvillus assembly. Not actin alone. *Curr Biol* **5,** 591-3.

Favreau, L. V., and Pickett, C. B. (1991). Transcriptional regulation of the rat NAD(P)H:quinone reductase gene. Identification of regulatory elements controlling basal level expression and inducible expression by planar aromatic compounds and phenolic antioxidants. *J Biol Chem* **266,** 4556-61.

Flohe, L. (1978). Glutathione peroxidase: fact and fiction. *Ciba Found Symp***,** 95-122.
Frendo, J. L., Xiao, G., Fuchs, S., Franceschi, R. T., Karsenty, G., and Ducy, P. (1998). Functional hierarchy between two OSE2 elements in the control of osteocalcin gene expression in vivo. *J Biol Chem* **273,** 30509-16.

Fridovich, I. (1995). Superoxide radical and superoxide dismutases. *Annu Rev Biochem* **64,** 97-112.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J., and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* **16,** 784-9.

Fukushige, T., Goszczynski, B., Yan, J., and McGhee, J. D. (2005). Transcriptional control and patterning of the pho-1 gene, an essential acid phosphatase expressed in the C. elegans intestine. *Dev Biol* **279,** 446-61.

Fukushige, T., Hawkins, M. G., and McGhee, J. D. (1998). The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. *Dev Biol* **198,** 286-302.

Fukushige, T., Hendzel, M. J., Bazett-Jones, D. P., and McGhee, J. D. (1999). Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 11883-8.

Garcia, J. V., Bich-Thuy, L. T., Stafford, J., and Queen, C. (1986). Synergism between immunoglobulin enhancers and promoters. *Nature* **322,** 383-5.

Geoffroy, V., Corral, D. A., Zhou, L., Lee, B., and Karsenty, G. (1998). Genomic organization, expression of the human CBFA1 gene, and evidence for an alternative splicing event affecting protein function. *Mamm Genome* **9,** 54-7
.
Goldberg, G., Caldwell, P., Weissbach, H., and Brot, N. (1979). In vitro regulation of DNA-dependent synthesis of Escherichia coli ribosomal protein L12. *Proc Natl Acad Sci U S A* **76,** 1716-20.

Goszczynski, B., and McGhee, J. D. (2005). Reevaluation of the role of the med-1 and med-2 genes in specifying the Caenorhabditis elegans endoderm. *Genetics* **171,** 545-55.

GuhaThakurta, D., Palomar, L., Stormo, G. D., Tedesco, P., Johnson, T. E., Walker, D. W., Lithgow, G., Kim, S., and Link, C. D. (2002). Identification of a novel cis-regulatory

element involved in the heat shock response in Caenorhabditis elegans using microarray gene expression and computational methods. *Genome Res* **12,** 701-12.

Harman, D. (1956). Aging: a theory based on free radical and radiation chemistry. *J Gerontol* **11,** 298-300.

Harman, D. (1992). Free radical theory of aging. *Mutat Res* **275,** 257-66.
Hatzopoulos, P., and Kambysellis, M. P. (1988). Comparative biochemical and immunological analysis of the three vitellogenins from Drosophila grimshawi. *Comp Biochem Physiol B* **89,** 557-64.

Hawkins, M. G., and McGhee, J. D. (1995). elt-2, a second GATA factor from the nematode Caenorhabditis elegans. *Journal of Biological Chemistry* **270,** 14666-71.

Helder, M. N., Karg, H., Bervoets, T. J., Vukicevic, S., Burger, E. H., D'Souza, R. N., Woltgens, J. H., Karsenty, G., and Bronckers, A. L. (1998). Bone morphogenetic protein-7 (osteogenic protein-1, OP-1) and tooth development. *J Dent Res* **77,** 545-54.

Inoue, H., Hisamoto, N., An, J. H., Oliveira, R. P., Nishida, E., Blackwell, T. K., and Matsumoto, K. (2005). The C. elegans p38 MAPK pathway regulates nuclear localization of the transcription factor SKN-1 in oxidative stress response. *Genes Dev* **19,** 2278-83.

Jaiswal, A. K. (1991). Human NAD(P)H:quinone oxidoreductase (NQO1) gene structure and induction by dioxin. *Biochemistry* **30,** 10647-53.

Johnson, T. E., and Hartman, P. S. (1988). Radiation effects on life span in Caenorhabditis elegans. *J Gerontol* **43,** B137-41.

Johnson, K.D., et al., (2001). Distinct mechanisms control RNA polymerase II recruitment to a tissue- specific locus control region and a downstream promoter. *Mol. Cell* **8**, 465-471.

Kamath, R. S., and Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. *Methods* **30,** 313-21.

Karsenty, G. (1998). Genetics of skeletogenesis. *Dev Genet* **22,** 301-13.

Karsenty, G. (1998). Transcriptional regulation of osteoblast differentiation during development. *Front Biosci* **3,** d834-7.

Katagiri, T., Boorla, S., Frendo, J. L., Hogan, B. L., and Karsenty, G. (1998). Skeletal abnormalities in doubly heterozygous Bmp4 and Bmp7 mice. *Dev Genet* **22,** 340-8.

Kiontke, K., Gavin, N. P., Raynes, Y., Roehrig, C., Piano, F., and Fitch, D. H. (2004). Caenorhabditis phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proc Natl Acad Sci U S A* **101,** 9003-8.

Kophengnavong, T., Carroll, A. S., and Blackwell, T. K. (1999). The SKN-1 amino-terminal arm is a DNA specificity segment. *Mol Cell Biol* **19,** 3039-50.

Korswagen, H. C., and Clevers, H. C. (1999). Activation and repression of wingless/Wnt target genes by the TCF/LEF-1 family of transcription factors. *Cold Spring Harbor Symposia on Quantitative Biology* **64,** 141-7.

Kuhn, E. J., and Geyer, P. K. (2003). Genomic insulators: connecting properties to mechanism. *Curr Opin Cell Biol* **15,** 259-65.

Kuo, C. T., Morrisey, E. E., Anandappa, R., Sigrist, K., Lu, M. M., Parmacek, M. S., Soudais, C., and Leiden, J. M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev* **11,** 1048-60.

Kuo, H., Chen, J., Ruiz-Lozano, P., Zou, Y., Nemer, M., and Chien, K. R. (1999). Control of segmental expression of the cardiac-restricted ankyrin repeat protein gene by distinct regulatory pathways in murine cardiogenesis. *Development* **126,** 4223-34.

Labrador, M., and Corces, V. G. (2002). Setting the boundaries of chromatin domains and nuclear organization. *Cell* **111,** 151-4.

Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R. E., Burch, J. B., and Evans, T. (1994). GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J Biol Chem* **269,** 23177-84.
Letting, D.L., et al., (2003). Formation of a tissue specific histone acetylation pattern by the hematopoietic transcription factor GATA-1. *Mol. Cell. Biol.* **23**, 1334-1340.

Leung, B., Hermann, G. J., and Priess, J. R. (1999). Organogenesis of the Caenorhabditis elegans intestine. *Dev Biol* **216,** 114-34.

Liao, V. H., and Yu, C. W. (2005). Caenorhabditis elegans gcs-1 confers resistance to arsenic-induced oxidative stress. *Biometals* **18,** 519-28.

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,** 599-609.

Lin, W. H., Huang, L. H., Yeh, J. Y., Hoheisel, J., Lehrach, H., Sun, Y. H., and Tsai, S. F. (1995). Expression of a Drosophila GATA transcription factor in multiple tissues in the developing embryos. Identification of homozygous lethal mutants with P-element insertion at the promoter region. *J Biol Chem* **270,** 25150-8.

Lithgow, G. J., White, T. M., Melov, S., and Johnson, T. E. (1995). Thermotolerance and extended life-span conferred by single-gene mutations and induced by thermal stress. *Proc Natl Acad Sci U S A* **92,** 7540-4.

Maduro, M. F., Hill, R. J., Heid, P. J., Newman-Smith, E. D., Zhu, J., Priess, J. R., and Rothman, J. H. (2005). Genetic redundancy in endoderm specification within the genus Caenorhabditis. *Dev Biol* **284,** 509-22.

Maduro, M. F., Kasmir, J. J., Zhu, J., and Rothman, J. H. (2005). The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. *Dev Biol* **285,** 510-23.

Maduro, M. F., Meneghini, M. D., Bowerman, B., Broitman-Maduro, G., and Rothman, J. H. (2001). Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3beta homolog is mediated by MED-1 and -2 in C. elegans. *Molecular Cell* **7,** 475-85.

Maduro, M. F., and Rothman, J. H. (2002). Making worm guts: the gene regulatory network of the Caenorhabditis elegans endoderm. *Dev Biol* **246,** 68-85.

Marshall, S. D., and McGhee, J. D. (2001). Coordination of ges-1 expression between the Caenorhabditis pharynx and intestine. *Dev Biol* **239,** 350-63.

McMahon, M., Itoh, K., Yamamoto, M., Chanas, S. A., Henderson, C. J., McLellan, L. I., Wolf, C. R., Cavin, C., and Hayes, J. D. (2001). The Cap'n'Collar basic leucine zipper transcription factor Nrf2 (NF-E2 p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes. *Cancer Research* **61,** 3299-307.

Meister, A., and Anderson, M. E. (1983). Glutathione. *Annu Rev Biochem* **52,** 711-60.

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,** 3959-70.

Mello, C. C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J. R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. *Nature* **382,** 710-2.

Melov, S., Ravenscroft, J., Malik, S., Gill, M. S., Walker, D. W., Clayton, P. E., Wallace, D. C., Malfroy, B., Doctrow, S. R., and Lithgow, G. J. (2000). Extension of life-span with superoxide dismutase/catalase mimetics. *Science* **289,** 1567-9.

Merika, M., and Thanos, D. (2001). Enhanceosomes. *Curr Opin Genet Dev* **11,** 205-8.

Moinova, H. R., and Mulcahy, R. T. (1999). Up-regulation of the human gamma-glutamylcysteine synthetase regulatory subunit gene involves binding of Nrf-2 to an electrophile responsive element. *Biochem Biophys Res Commun* **261,** 661-8.

Molkentin, J. D., and Olson, E. N. (1997). GATA4: a novel transcriptional regulator of cardiac hypertrophy? *Circulation* **96,** 3833-5.

Mulcahy, R. T., Wartman, M. A., Bailey, H. H., and Gipp, J. J. (1997). Constitutive and beta-naphthoflavone-induced expression of the human gamma-glutamylcysteine synthetase heavy subunit gene is regulated by a distal antioxidant response element/TRE sequence. *J Biol Chem* **272,** 7445-54.

Muller, M. M., Gerster, T., and Schaffner, W. (1988). Enhancer sequences and the regulation of gene transcription. *Eur J Biochem* **176,** 485-95.

Murakami, R., Okumura, T., and Uchiyama, H. (2005). GATA factors as key regulatory molecules in the development of Drosophila endoderm. *Dev Growth Differ* **47,** 581-9.

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature* **424,** 277-83.

Navarro, R. E., and Blackwell, T. K. (2005). Requirement for P granules and meiosis for accumulation of the germline RNA helicase CGH-1. *Genesis* **42,** 172-80.

Navarro, R. E., Shim, E. Y., Kohara, Y., Singson, A., and Blackwell, T. K. (2001). cgh-1, a conserved predicted RNA helicase required for gametogenesis and protection from physiological germline apoptosis in C. elegans. *Development* **128,** 3221-32.

Nguyen, T., Huang, H. C., and Pickett, C. B. (2000). Transcriptional regulation of the antioxidant response element. Activation by Nrf2 and repression by MafK. *J Biol Chem* **275,** 15466-73.

Okumura, T., Matsumoto, A., Tanimura, T., and Murakami, R. (2005). An endoderm-specific GATA factor gene, dGATAe, is required for the terminal differentiation of the Drosophila endoderm. *Dev Biol* **278,** 576-86.

Orkin, S. H. (1992). GATA-binding transcription factors in hematopoietic cells. *Blood* **80,** 575-81.

Orr, W. C., and Sohal, R. S. (1994). Extension of life-span by overexpression of superoxide dismutase and catalase in Drosophila melanogaster. *Science* **263,** 1128-30.

Owen, M. J., and Karsenty, G. (1998). New developments in bone formation. *Curr Opin Nephrol Hypertens* **7,** 363-6.

Pelham, H. R. (1982). A regulatory upstream promoter element in the Drosophila hsp 70 heat-shock gene. *Cell* **30,** 517-28.

Pramila, T., Miles, S., GuhaThakurta, D., Jemiolo, D., and Breeden, L. L. (2002). Conserved homeodomain proteins interact with MADS box protein Mcm1 to restrict ECB-dependent transcription to the M/G1 phase of the cell cycle. *Genes Dev* **16,** 3034-45.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* **386,** 569-77.

Rehorn, K. P., Thelen, H., Michelson, A. M., and Reuter, R. (1996). A molecular aspect of hematopoiesis and endoderm development common to vertebrates and Drosophila. *Development* **122,** 4023-31.

Reiter, J. F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N., and Stainier, D. Y. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev* **13,** 2983-95.

Reiter, J. F., Kikuchi, Y., and Stainier, D. Y. (2001). Multiple roles for Gata5 in zebrafish endoderm formation. *Development* **128,** 125-35.

Reuter, R. (1994). The gene serpent has homeotic properties and specifies endoderm versus ectoderm within the Drosophila gut. *Development* **120,** 1123-35.

Reuter, R., and Leptin, M. (1994). Interacting functions of snail, twist and huckebein during the early development of germ layers in Drosophila. *Development* **120,** 1137-50.

Richman, P. G., and Meister, A. (1975). Regulation of gamma-glutamyl-cysteine synthetase by nonallosteric feedback inhibition by glutathione. *J Biol Chem* **250,** 1422-6.

Rushmore, T. H., King, R. G., Paulson, K. E., and Pickett, C. B. (1990). Regulation of glutathione S-transferase Ya subunit gene expression: identification of a unique xenobiotic-responsive element controlling inducible expression by planar aromatic compounds. *Proc Natl Acad Sci U S A* **87,** 3826-30.

Sasmal, D., Guhathakurta, B., Bhattacharya, S. K., Pal, C. R., and Datta, A. (2002). N-acetyl-D-glucosamine specific hemagglutinin receptor of Vibrio cholerae O1 in chicken erythrocyte membranes. *FEMS Immunol Med Microbiol* **32,** 187-9.

Schinke, T., McKee, M. D., Kiviranta, R., and Karsenty, G. (1998). Molecular determinants of arterial calcification. *Ann Med* **30,** 538-41.

Seydoux, G., and Fire, A. (1994). Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans. *Development* **120,** 2823-34.

Seydoux, G., Mello, C. C., Pettitt, J., Wood, W. B., Priess, J. R., and Fire, A. (1996). Repression of gene expression in the embryonic germ lineage of C. elegans. *Nature* **382,** 713-6.

Shi, Y., and Blackwell, T. K. (2003). A two-tiered transcription regulation mechanism that protects germ cell identity. *Mol Cell* **12,** 1062-4.

Shim, E. Y., Walker, A. K., and Blackwell, T. K. (2002). Broad requirement for the mediator subunit RGR-1 for transcription in the Caenorhabditis elegans embryo. *J Biol Chem* **277,** 30413-6.

Shim, E. Y., Walker, A. K., Shi, Y., and Blackwell, T. K. (2002). CDK-9/cyclin T (P-TEFb) is required in two postinitiation pathways for transcription in the C. elegans embryo. *Genes Dev* **16,** 2135-46.

Shoichet, S. A., Malik, T. H., Rothman, J. H., and Shivdasani, R. A. (2000). Action of the Caenorhabditis elegans GATA factor END-1 in Xenopus suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. *Proc Natl Acad Sci U S A* **97,** 4076-81.

Sies, H. (1993). Strategies of antioxidant defense. *European Journal of Biochemistry* **215,** 213-9.

Soltaninassab, S. R., Sekhar, K. R., Meredith, M. J., and Freeman, M. L. (2000). Multi-faceted regulation of gamma-glutamylcysteine synthetase. *Journal of Cellular Physiology* **182,** 163-70.

Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C. A., Narita, N., Saffitz, J. E., Simon, M. C., Leiden, J. M., and Wilson, D. B. (1995). Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. *Development* **121,** 3877-88.

Spilianakis, C. G., and Flavell, R. A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* **5,** 1017-27.

Stadtman, E. R. (1992). Protein oxidation and aging. *Science* **257,** 1220-4.

Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C., Willey, D., Wilson, R. K., Durbin, R., and Waterston, R. H. (2003). The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. *PLoS Biol* **1,** E45.

Sulston, J. E., and Brenner, S. (1974). The DNA of Caenorhabditis elegans. *Genetics* **77,** 95-104.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology* **100,** 64-119.

Sun, J., and Tower, J. (1999). FLP recombinase-mediated induction of Cu/Zn-superoxide dismutase transgene expression can extend the life span of adult Drosophila melanogaster flies. *Mol Cell Biol* **19,** 216-28.

Takagi, T., Walker, A. K., Sawa, C., Diehn, F., Takase, Y., Blackwell, T. K., and Buratowski, S. (2003). The Caenorhabditis elegans mRNA 5'-capping enzyme. In vitro and in vivo characterization. *J Biol Chem* **278,** 14174-84.

Thirunavukkarasu, K., Mahajan, M., McLarren, K. W., Stifani, S., and Karsenty, G. (1998). Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfbeta. *Mol Cell Biol* **18,** 4197-208.

Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10,** 1453-65.

Tsai, F. Y., Browne, C. P., and Orkin, S. H. (1998). Knock-in mutation of transcription factor GATA-3 into the GATA-1 locus: partial rescue of GATA-1 loss of function in erythroid cells. *Dev Biol* **196,** 218-27.

Tsai, S. F., Martin, D. I., Zon, L. I., D'Andrea, A. D., Wong, G. G., and Orkin, S. H. (1989). Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* **339,** 446-51.

Tuan, D., Kong, S., Hu, K. (1992). Transcription of the hypersensitive site HS2 enhancer in erthroid cells. *Proc. Natl Acad Sci* U.S.A. **89**, 11219-11223.

Walker, A. K., and Blackwell, T. K. (2003). A broad but restricted requirement for TAF-5 (human TAFII100) for embryonic transcription in Caenorhabditis elegans. *J Biol Chem* **278,** 6181-6.

Walker, A. K., Rothman, J. H., Shi, Y., and Blackwell, T. K. (2001). Distinct requirements for C.elegans TAF(II)s in early embryonic transcription. *Embo J* **20,** 5269-79.

Walker, A. K., See, R., Batchelder, C., Kophengnavong, T., Gronniger, J. T., Shi, Y., and Blackwell, T. K. (2000). A conserved transcription motif suggesting functional parallels between Caenorhabditis elegans SKN-1 and Cap'n'Collar-related basic leucine zipper proteins. *J Biol Chem* **275,** 22166-71.

Walker, A. K., Shi, Y., and Blackwell, T. K. (2004). An extensive requirement for transcription factor IID-specific TAF-1 in Caenorhabditis elegans embryonic transcription. *J Biol Chem* **279,** 15339-47.

Wang, J. C., Walker, A., Blackwell, T. K., and Yamamoto, K. R. (2004). The Caenorhabditis elegans ortholog of TRAP240, CeTRAP240/let-19, selectively modulates gene expression and is essential for embryogenesis. *J Biol Chem* **279,** 29270-7.

Wasserman, W. W., and Fahl, W. E. (1997). Functional antioxidant responsive elements. *Proc Natl Acad Sci U S A* **94,** 5361-6.

Weber, H., Symes, C. E., Walmsley, M. E., Rodaway, A. R., and Patient, R. K. (2000). A role for GATA5 in Xenopus endoderm specification. *Development* **127,** 4345-60.

West, A.G and Fraser, P. (2005). Remoter control of gene transcription. *Human Molecular Genetics* **14**, R101-111.

Wood, W. B. (1988). Determination of pattern and fate in early embryos of Caenorhabditis elegans. *Dev Biol (N Y 1985)* **5,** 57-78.

Xiao, G., Wang, D., Benson, M. D., Karsenty, G., and Franceschi, R. T. (1998). Role of the alpha2-integrin in osteoblast-specific gene expression and activation of the Osf2 transcription factor. *J Biol Chem* **273,** 32988-94.

Zhang, F., Barboric, M., Blackwell, T. K., and Peterlin, B. M. (2003). A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. *Genes Dev* **17,** 748-58.

Zhu, J., Fukushige, T., McGhee, J. D., and Rothman, J. H. (1998). Reprogramming of early embryonic blastomeres into endodermal progenitors by a Caenorhabditis elegans GATA factor. *Genes Dev* **12,** 3809-14.

Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., and Rothman, J. H. (1997). end-1 encodes an apparent GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos. *Genes Dev* **11,** 2883-96.

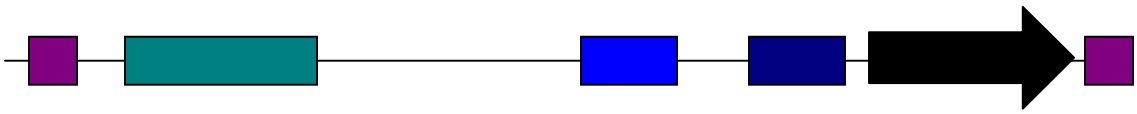


Figure 1. Basic transcriptional unit; includes the coding region (black arrow) and the regulatory DNA sequences that direct transcription. The regulatory DNA sequence usually includes the basal (in dark blue) and regulatory promoter (light blue), enhancer (green) and boundary elements (purple).

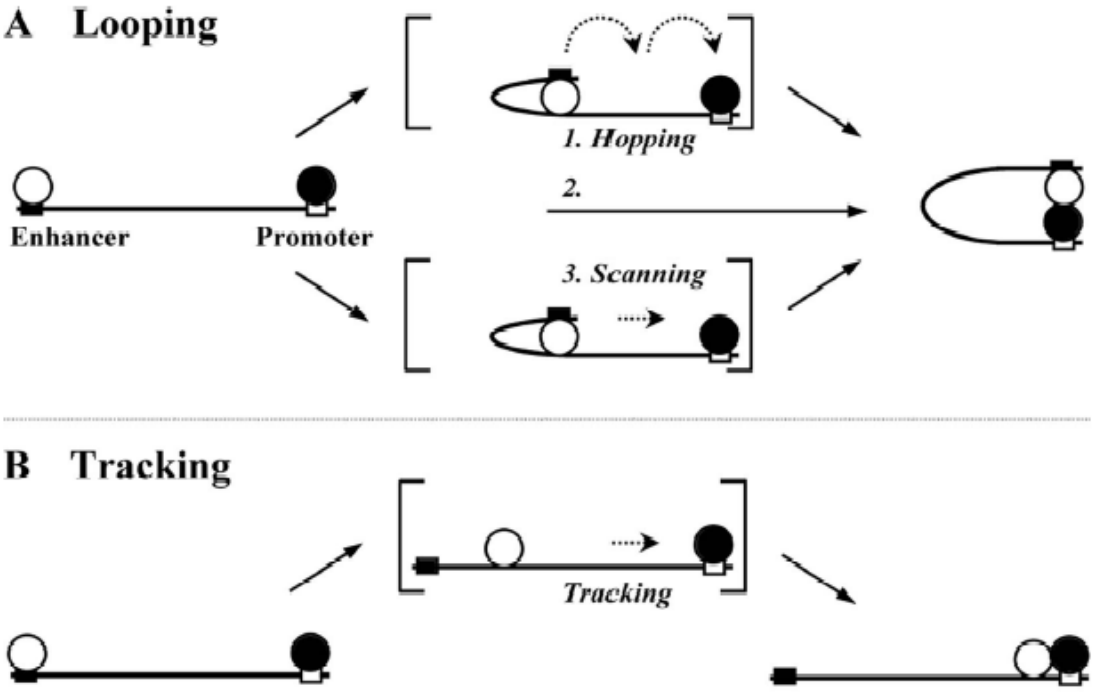


Figure 2. Transcriptional regulation. Mechanisms proposed to explain enhancer action over a distance are shown. (A) Looping Model. Enhancer-bound protein could either directly interact with a protein at the promoter (2), or search the spacer DNA using a hopping (1) or scanning (3) mechanism. (B) Tracking Model. The enhancer and promoter bound proteins are indicated by circles (white and black, respectively) and their DNA-binding sites are indicated by rectangles. The intermediate complexes formed during enhancer–promoter communication are shown in brackets; the dashed arrows indicate direction of productive communication. Figure and legend taken from Bondarenko et al., 2003.

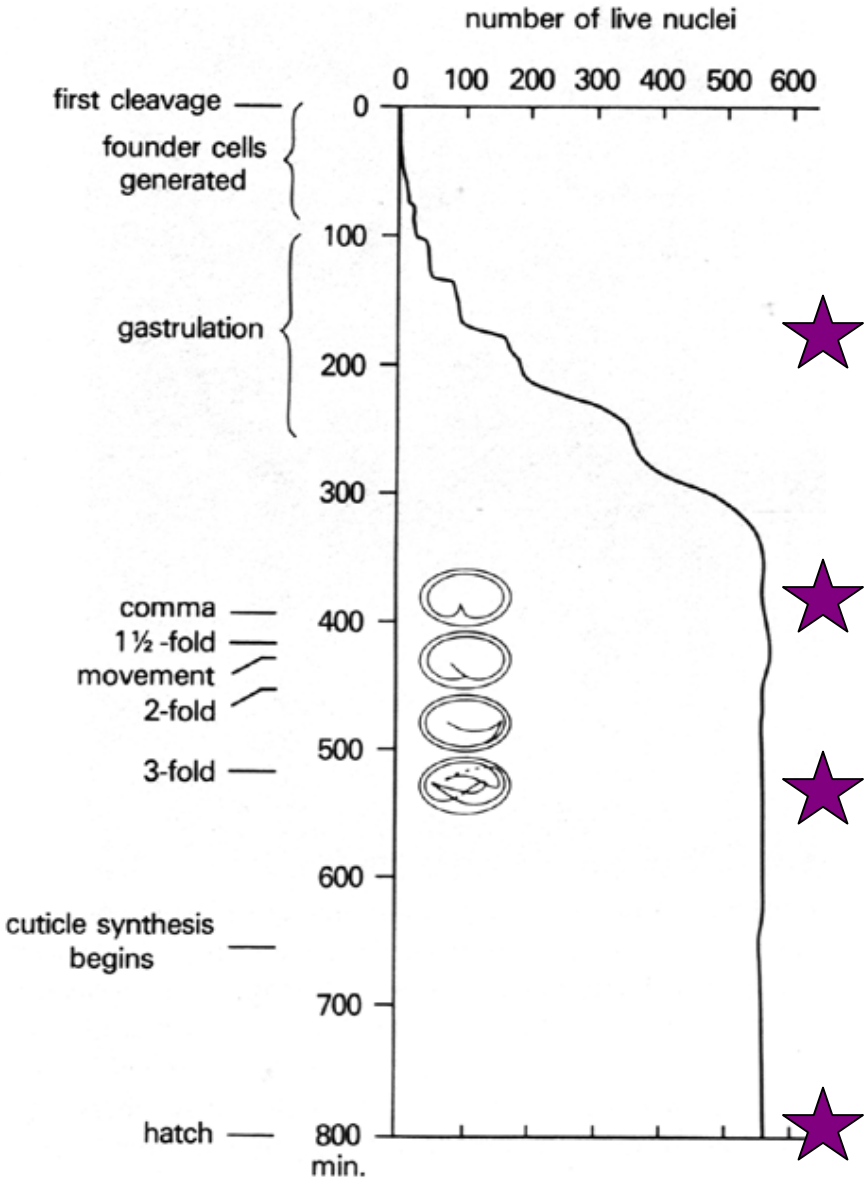


Figure 3. Summary of events in embryogenesis. The timing of key events and stages of elongation (at 20°C) are shown on the left. The number of living nuclei in different stage embryos is plotted on the right. Purple stars represent key developmental stages at which *elt-2* reporter constructs were analyzed. Fertilization is normally at –50 minutes. Figure and legend adapted from Strome [1989] and Sulston et al. [1983].

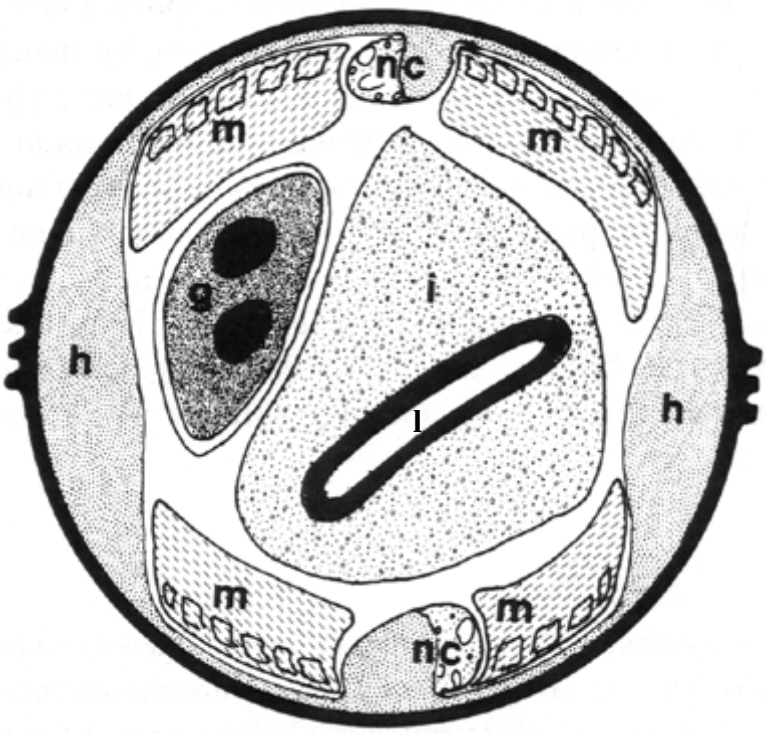


Figure 4. Diagram of a posterior cross section through the adult hermaphrodite. (g) Gonad; (h) hypodermal ridge; (i) intestine; (m) muscle; (nc) nerve cord; (l) lumen. Figure and legend adapted from Edwards and Wood 1983.

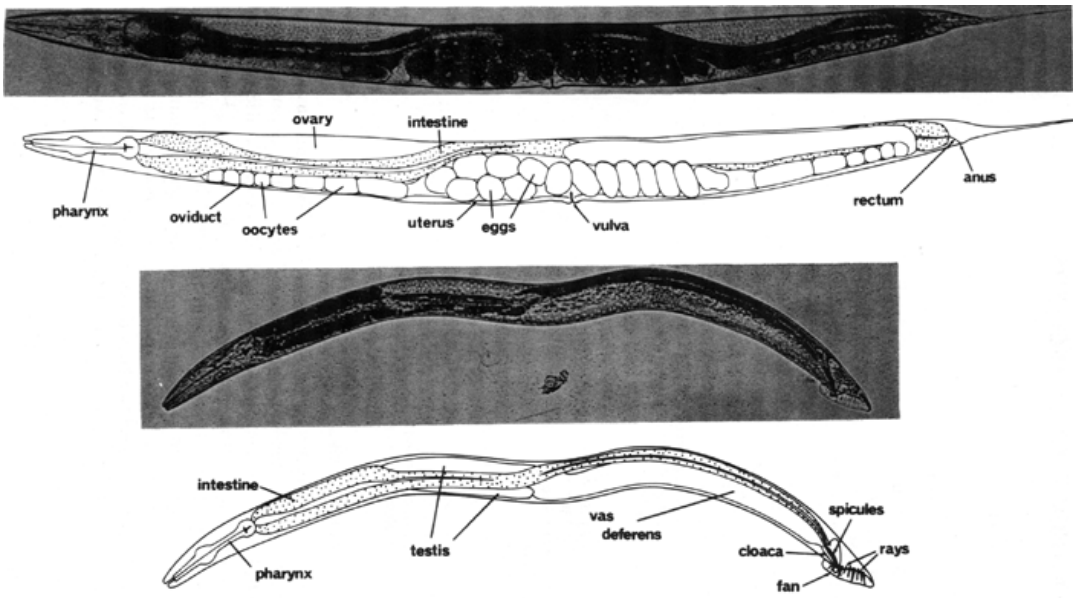


Figure 5. Photomicrographs showing major anatomical features of the *C. elegans* adult hermaphrodite (*top*) and male (*bottom*). Shown are lateral views under bright-field illumination. Figure and legend adapted from Sulston and Horvitz 1977.

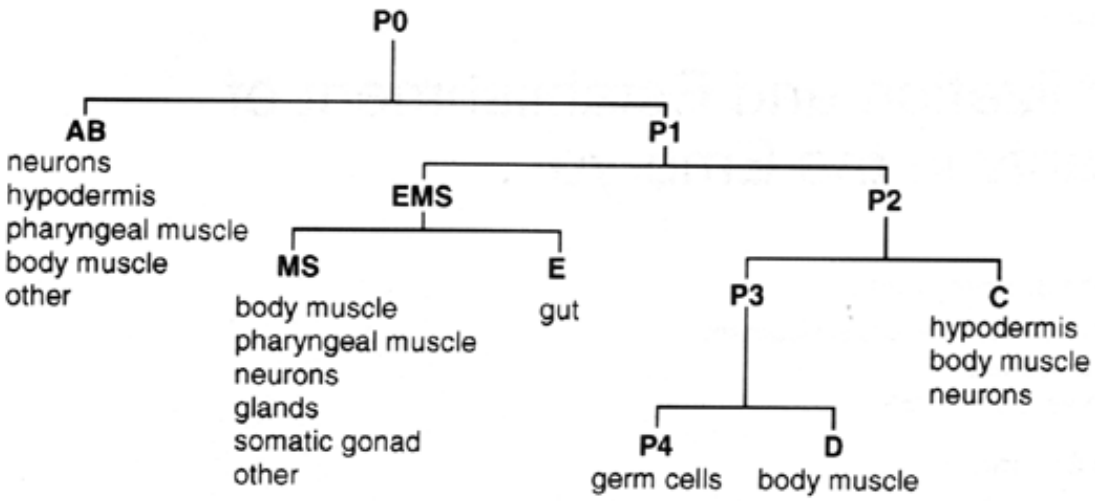


Figure 6. Early cell lineage tree. Cell divisions are indicated by horizontal lines; the anterior daughter of each division is placed on the left. A series of unequal divisions of the germ-line or P cells results in formation of five somatic founder cells also called blastomeres (AB, MS, E, C, and D) and the primordial germ cell (P_4). The tissues generated by each founder cell are indicated. Figure and legend adapted from Sulston et al. [1983] and Schierenberg [1987].

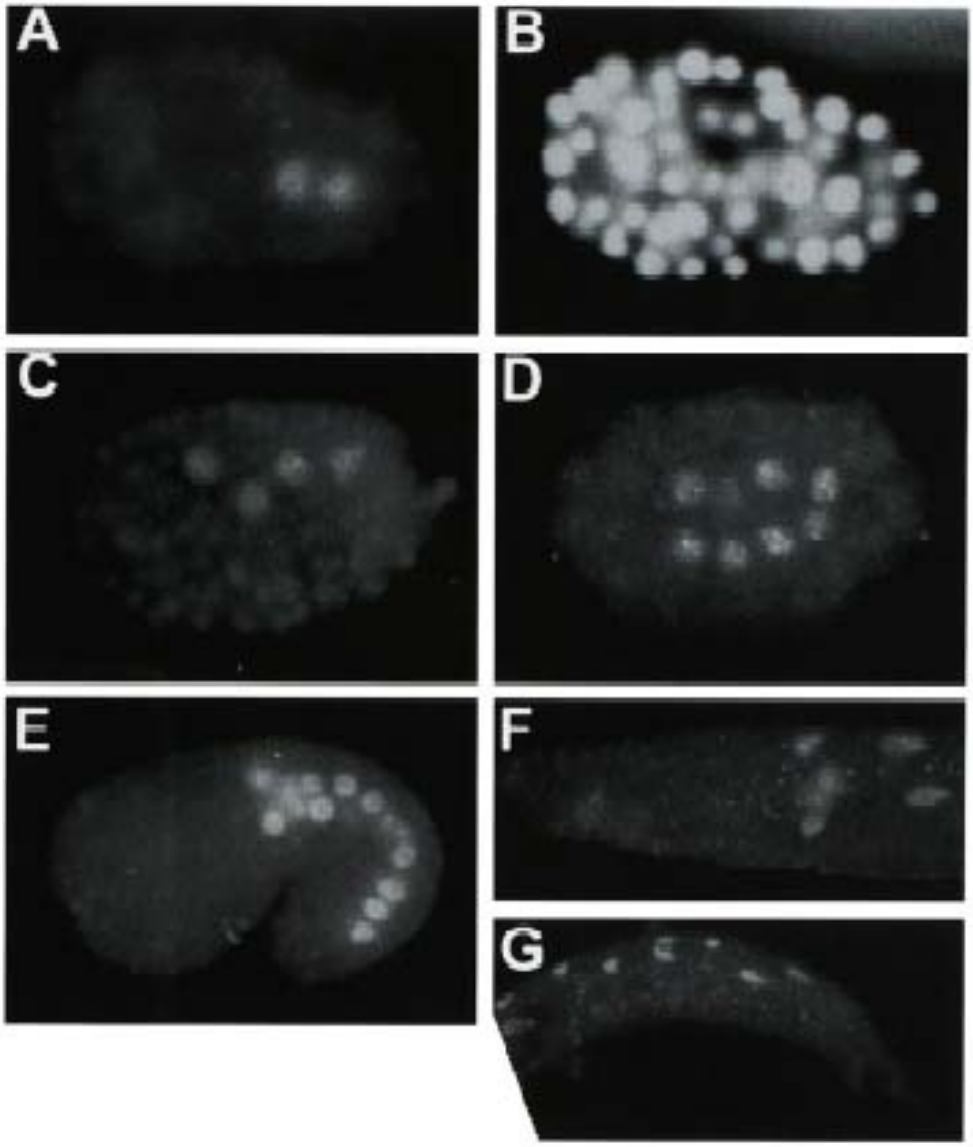


Figure 7. Immunological detection of ELT-2 protein in wild-type embryos and adults. (A) Anti-ELT-2 staining of N2 embryo at the gastrulation stage, showing the presence of ELT-2 protein; (B) DAPI staining of the same embryo as in A; (C-D}anti-ELT-2 staining of N2 embryos at the mid and late gastrulation stage; (E) anti-ELT-2 staining of N2 embryos at comma stage; (F and G) anti-ELT-2 staining shown in the anterior intestine of an adult N2 hermaphrodite and in the posterior intestine of an adult N2 male, Respectively. In all figures, anterior is to the left. Figure and legend adapted from Fukushige et al., 1998.

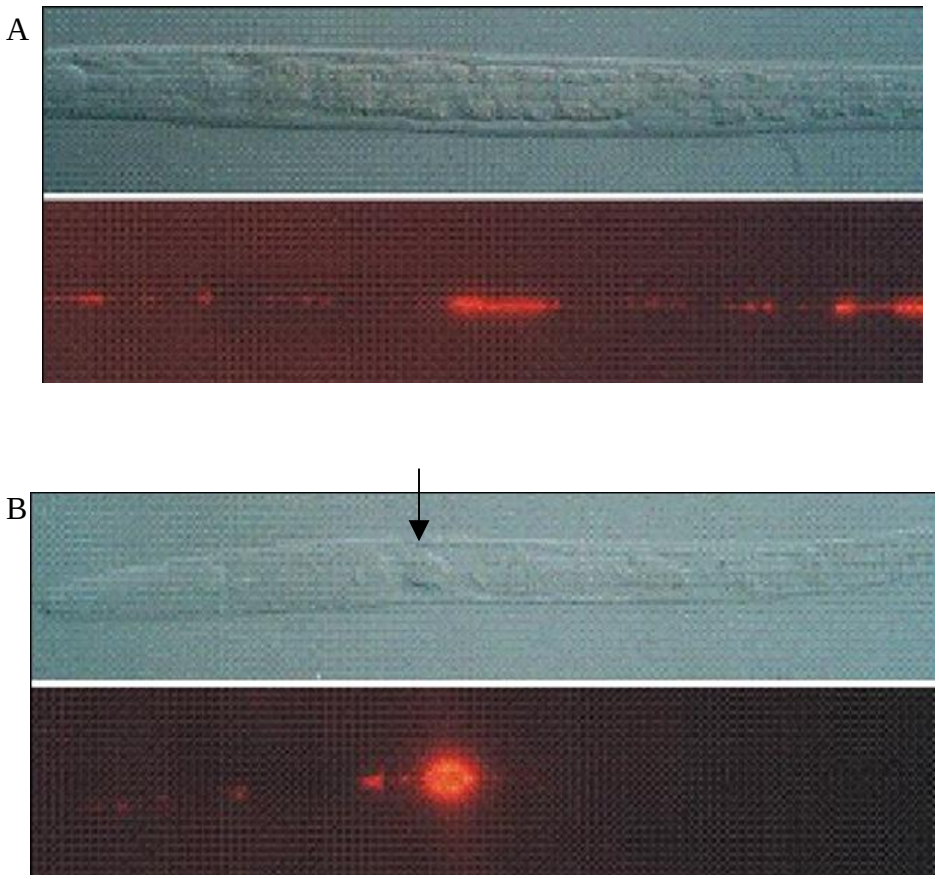


Figure 8. *elt-2* RNAi phenotype (A) N2 worm fed fluorescent beads. Top panel is a nomarski image and bottom panel is fluorescence image (B) *elt-2* RNAi worm fed fluorescent beads have a gut obstructed (Gob) phenotype. Arrow indicates Gob phenotype in the anterior intestine. Top panel is a nomarski image and bottom panel is fluorescence image.

ACTGATAGCA	*C. elegans ges-1* upstream GATA
ACTGATAAGG	*C. elegans ges-1* downstream GATA
TGTGATAAGA	*C. briggsae ges-1* upstream GATA
ACTGATAAGA	*C. briggsae ges-1* downstream GATA
ACTGATAAAA	*C. elegans pho-1*
ACTGATAAAA	*C. briggsae pho-1*
ACTGATAAGA	*C. elegans cpr-1*
ACTGATAAAA	*C. elegans mtl-1* GATA site 1
GCTGATAACA	*C. briggsae mtl-1* GATA site 1
ACTGATAGCG	*mab-3* intestinal enhancer
AATGTTGCAATTTGTTTCTGATAAGG	*vit-2* upstream GATA with adjacent **MAB-3** site
ACTGATAAGN	vitellogenin genes consensus GATA sequence

Figure 9. Alignment of critical GATA sites regulating expression of intestinal specific genes. Core GATA site is conserved, while sequence surrounding the site varies between genes. Bold sequence is a MAB-3 binding sequence.

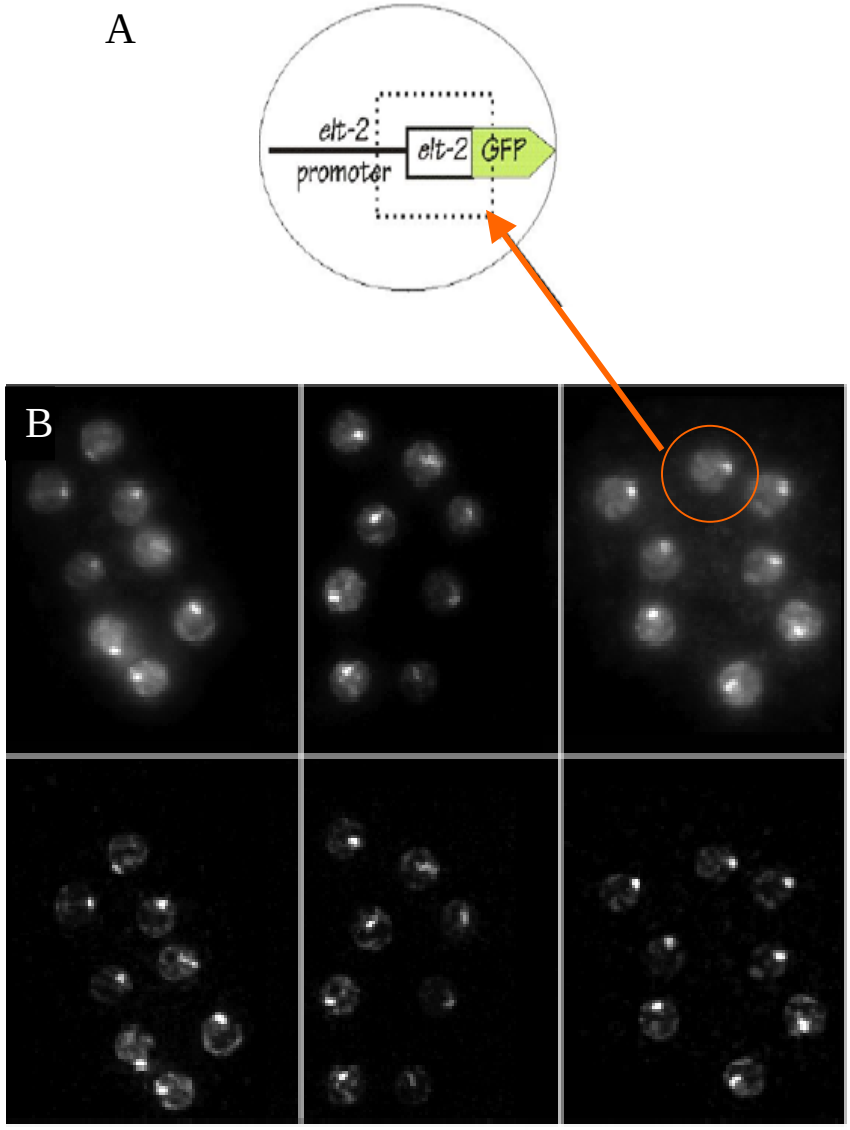


Figure 10. *elt-2* autoregulation is direct. (a) cartoon depicting *elt-2* promoter and *elt-2* gene fused to GFP reporter construct backbone. (b) Gut primordium at eight cell stage. One focus per gut nucleus in embryos heterozygous for the transgenic array containing the full rescuing *elt-2*::GFP construct pJM86. top row are photos that have not been deconvolved while bottom row are photos that have been digitally deconvolved. Orange arrow connects cartoon depiction to one nucleua. Figure and legend adapted from Fukushige et al., 1999**.**

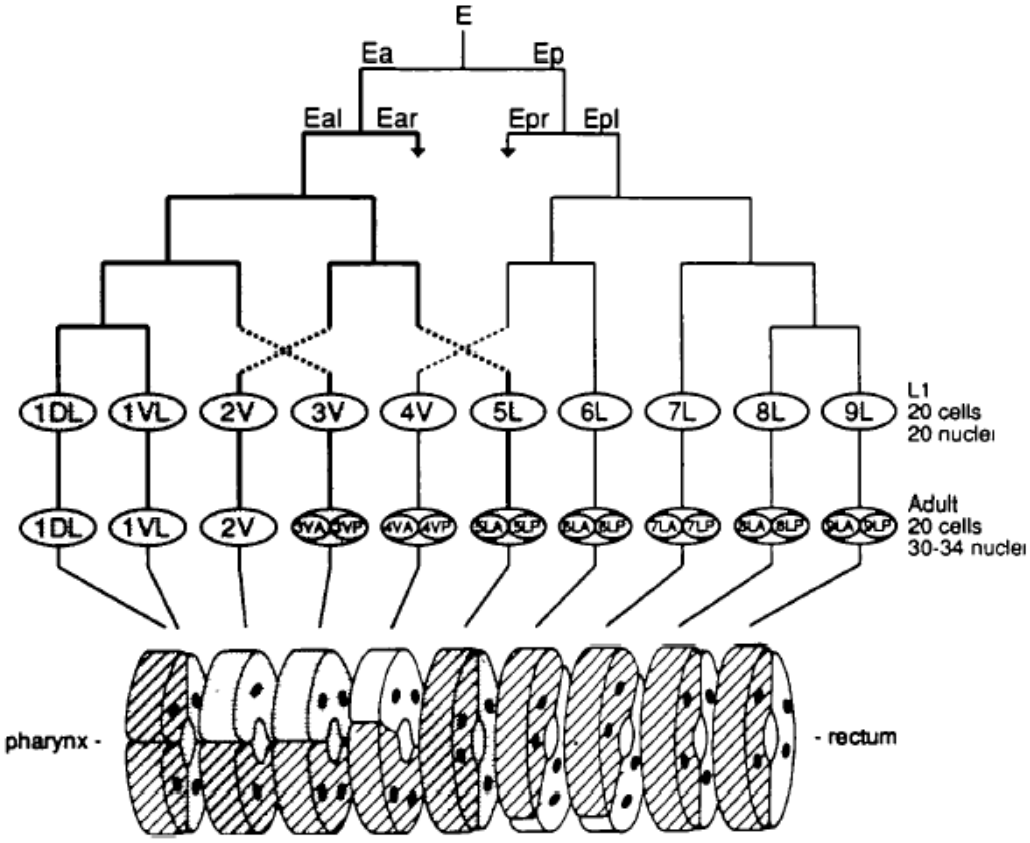


Figure 11. Diagrammatic representation of the *C. elegans* intestinal lineage and morphology. The upper diagram shows the lineage of the intestinal cells. The lower portion of figure shows a schematic of the structure of the adult intestine, which consists of 20 cells arranged in 9 structural units (designated int1-9 as shown). Figure and legend adapted from Schedin et al., 1991.

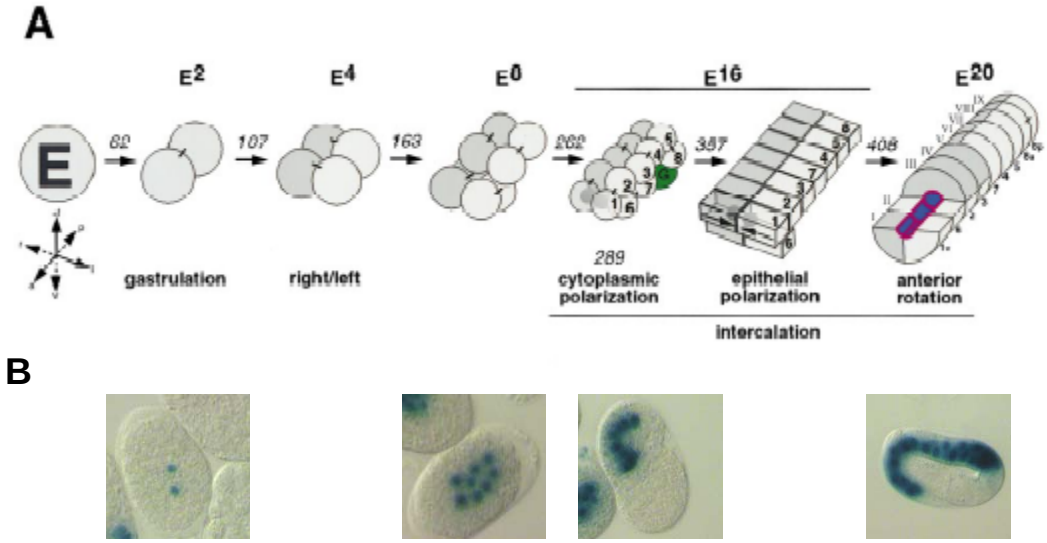


Figure 12. Cellular events in organogenesis. (A) Schematic diagram showing the cell divisions and major events of intestinal organogenesis. Times are indicated in italics and represent minutes past the 2-cell stage of embryogenesis; on this scale the E blastomere is born at 35 min. For the E16 primordium, the anterior most left/right pair of cells is shown transparent to indicate nuclear positions. The two germ cells (G) are on the ventral side of the E16 intestine, and one of these is shown in green. The intestinal rings (int rings) at the E20 stage are labeled with roman numerals following the convention of Sulston *et al.* (1983). The dorsal cells of the int I and II rings are removed to show the lumenal surface (blue) and the surrounding adherens junctions (magenta). The diagram is not to scale; the volume of the E blastomere is about the same as the volume of the entire intestine. B) Embryos stained with β-galactosidase. Stage of embryos 2E, 8E, 16E, 20E and correspond to time points in intestinal development from the schematic in A. Figure and legend adapted from Leung et al., 1999.

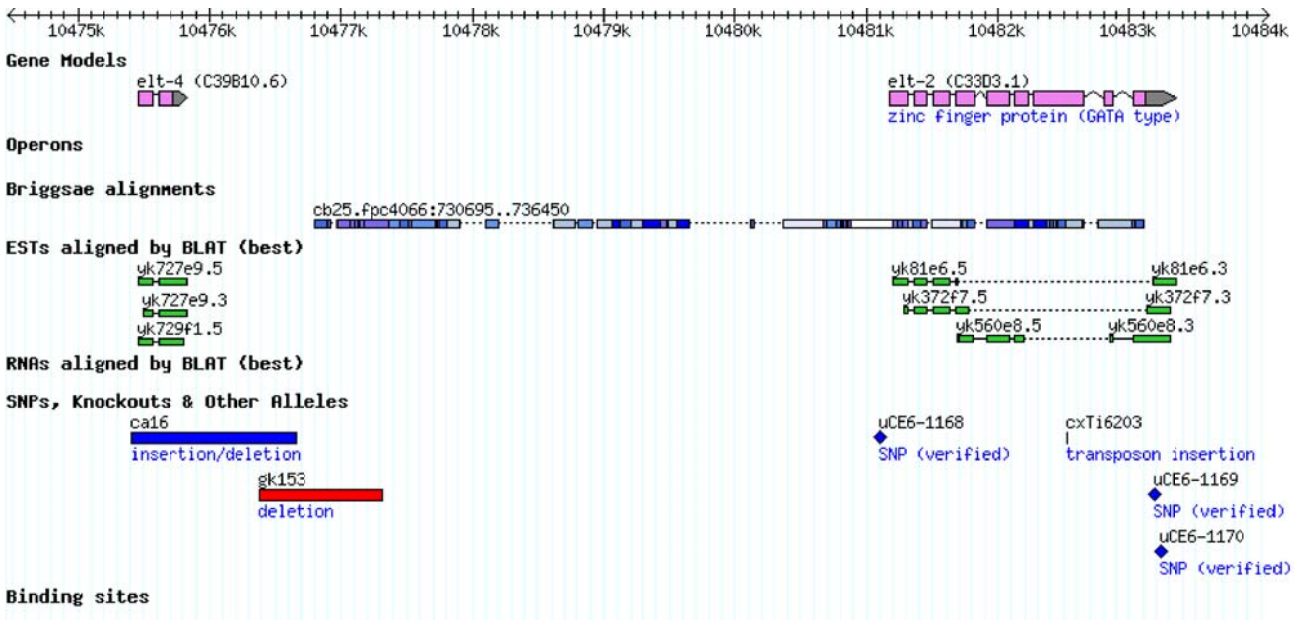


Figure 13. *elt-2* gene plus surrounding region. The *elt-4* gene is upstream of *elt-2*. In blue are genome alignments between *C. briggsae* and *C. elegans* using the computer program WABA; low similarity regions (light blue), high similarity regions (dark blue). In green are ESTs (Expressed sequence Tags). Red bar represents location of the deletion removing the distal region of the enhancer. Blue diamonds are SNP (single nucleotide polymorphisms). Diagram taken from Wormbase; http://www.wormbase.org/db/seq/gbrowse/wormbase/

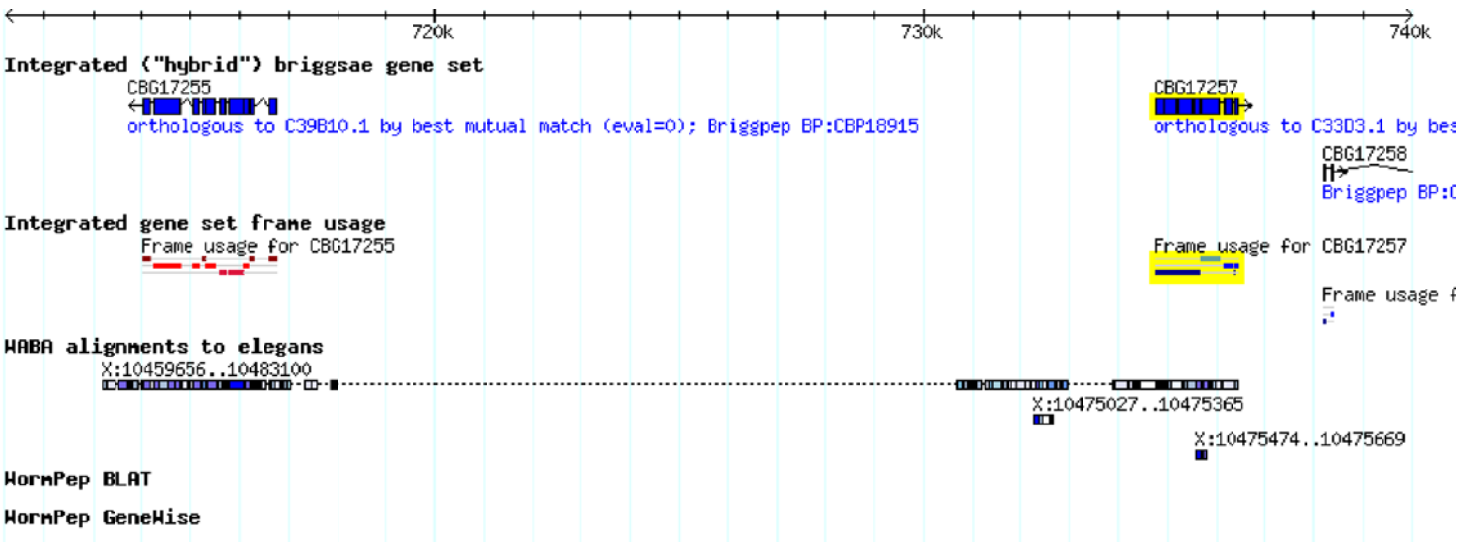


Figure 14. CBG17257 gene and surrounding region. Note *elt-4* orthologue is not present. In blue are genome alignments between *C. briggsae* and *C. elegans* using the computer program WABA; low similarity regions (light blue), high-similarity regions (dark blue). Red bars indicate the reading frame used for CBG17257. Diagram taken from Wormbase; http://www.wormbase.org/db/seq/gbrowse/briggsae/?name=CBG17255

Figure 15. Comparison of the ELT-2 protein with its orthologues in *C. remanei* and *C. briggsae*. Black letters represent 100% conservation. *C. elegans* amino acids highlighted in pink are conserved between *C. elegans* and *C. remanei*, but not *C. briggsae*. *C. elegans* amino acids highlighted in orange are conserved between *C. elegans* and *C. briggsae*, but not *C. remanei*. *C. elegans* amino acids highlighted in purple are not conserved with *C. remanei* or *C. briggsae*. Gaps are introduced to allow for maximum alignment. The DNA binding domain consisting of the GATA type zinc finger and basic region are underlined with the GATA zinc finger double underlined. Cysteine residues of the $CX_2CX_{17}CX_2C$ GATA zinc finger are marked "*".

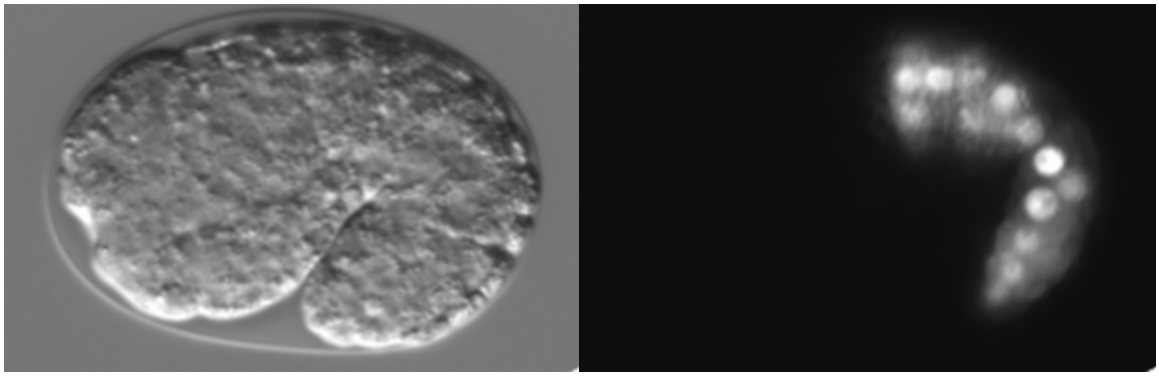


Figure 16. The 5’ flanking region of the *C. briggsae* gene CBG17257 specifies intestinal expression similar to *elt-2* as illustrated by the ability of an CBG17257 promoter::GFP construct to drive intestinal specific expression in *C. elegans*. Left image is a Nomarski image and the right is a fluorescence image of the same comma staged embryo.

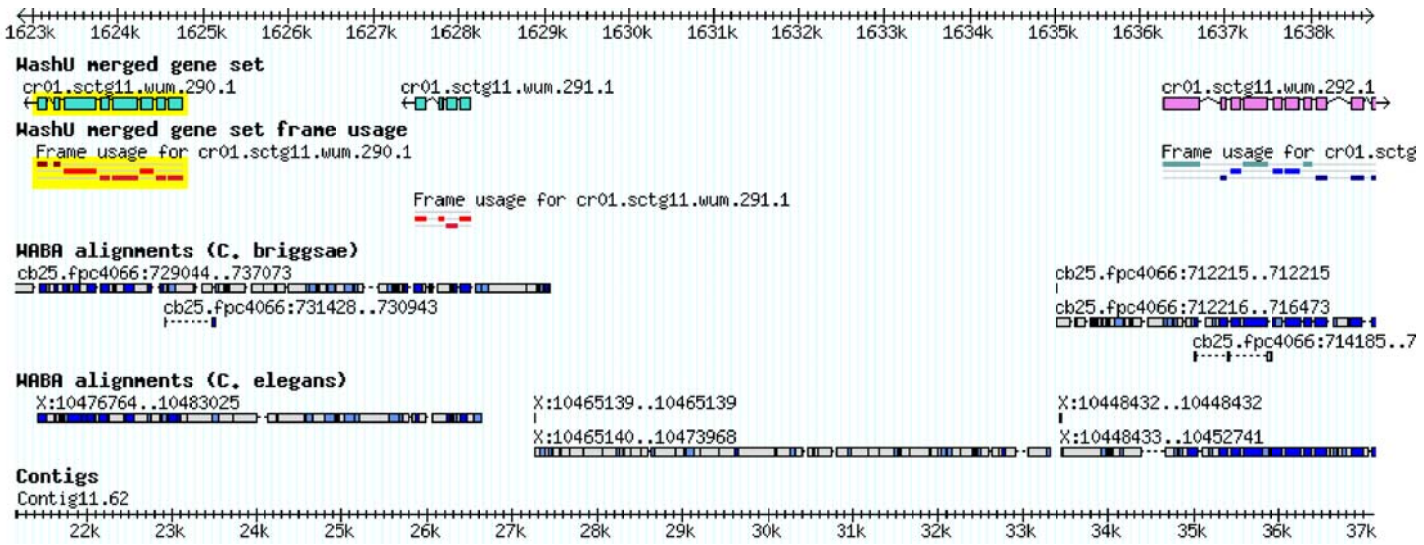


Figure 17. The predicted gene cr. 7 cr01.sctg11.wum.290.1 in *C. remanei* and its surrounding region. The gene is located on the left side of the schematic. The predicted upstream gene cr01.sctg11.wum.291.1 is most similar to the *elt-2* enhancer region. No *elt-4* orthologue is present. In blue are genome alignments of *C. briggsae* and *C. elegans* to *C. remanei.* Red bars indicate the reading frame used for each potential gene. Diagram taken from Wormbase; http://dev.wormbase.org/db/seq/gbrowse/remanei/?name=cr01.sctg11.wum.290.1

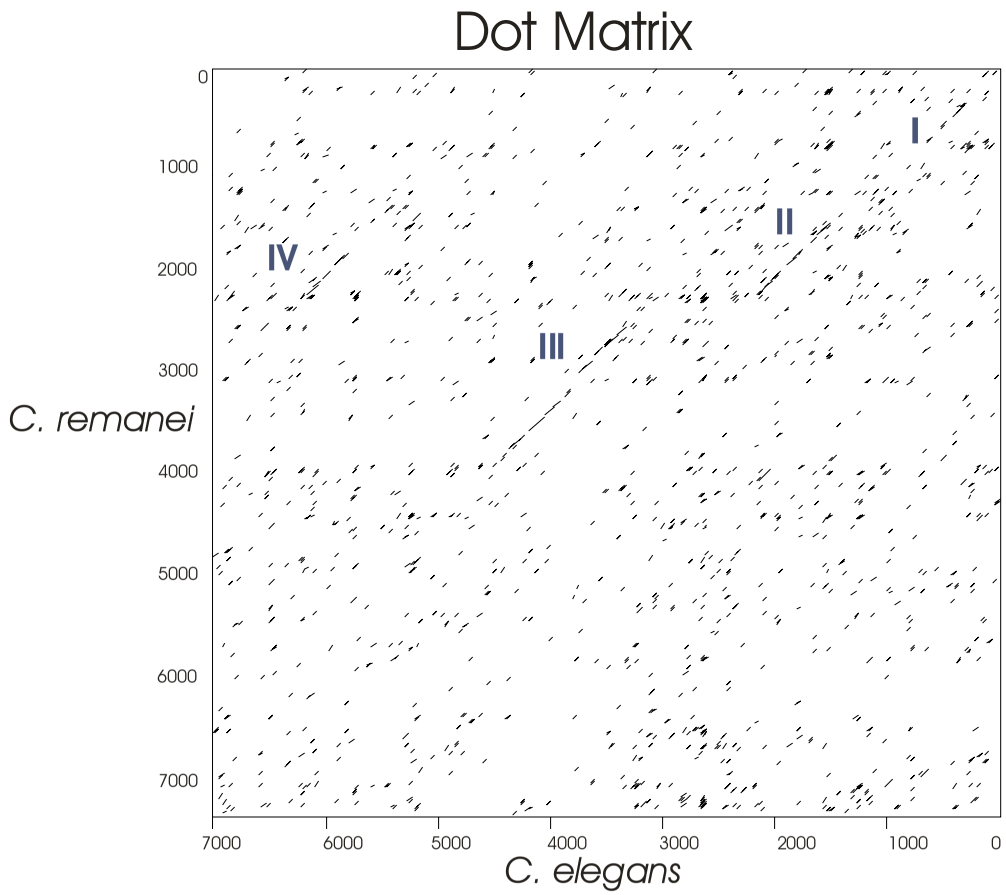


Figure 18. Dot Matrix of the 7000 bp sequence flanking the 5' end of the *C. remanei* gene cr01.sctg11.wum.290.1 and the *C. elegans* gene *elt-2*.
Dot matrix made using the program Dotmatcher. The program was specified to find a minimum sequence conservation of 35 bp within a 40 bp sliding window. Each dot on the matrix pertains to finding a 35 bp sequence conservation between the two species.

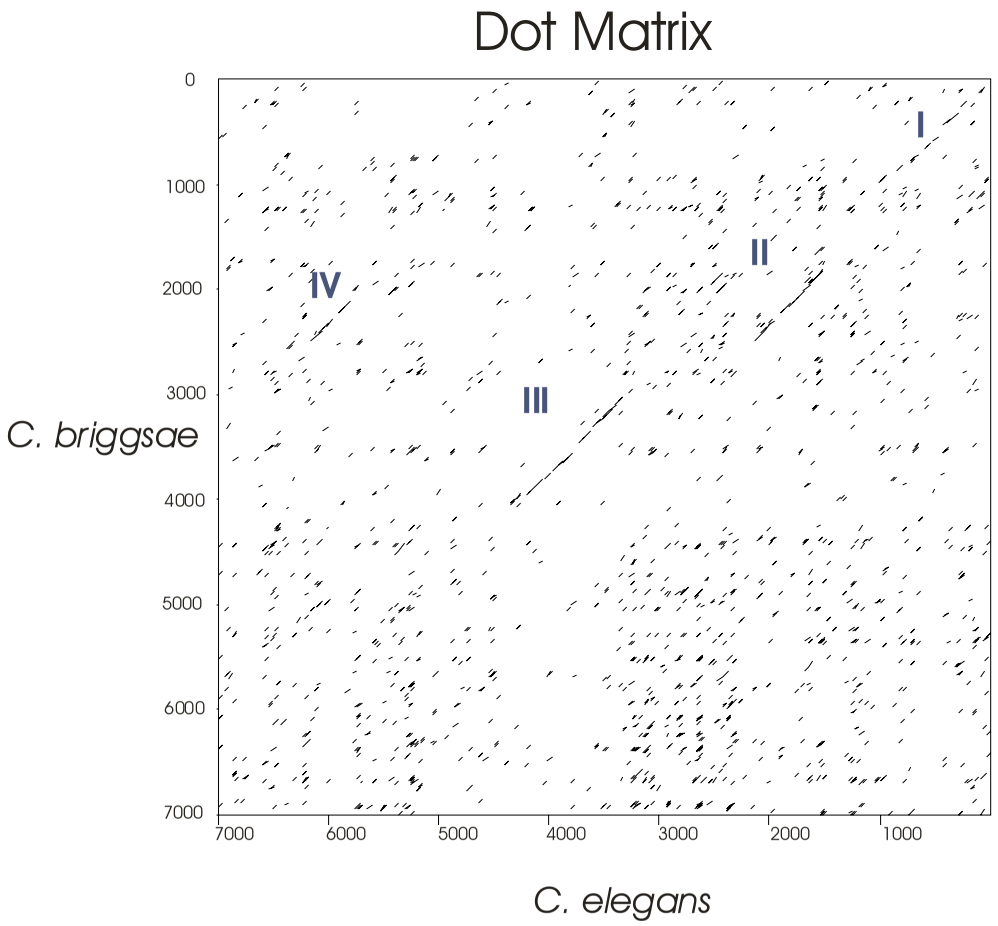


Figure 19. Dot Matrix of the 7000 bp sequence flanking the 5' end of the *C. briggsae* gene CBG17255 and the *C. elegans* gene *elt-2*.
Dot matrix made using the program Dotmatcher. The program was specified to find a minimum sequence conservation of 35 bp within a 40 bp sliding window. Each dot on the matrix pertains to finding a 35 bp sequence conservation between the two species.

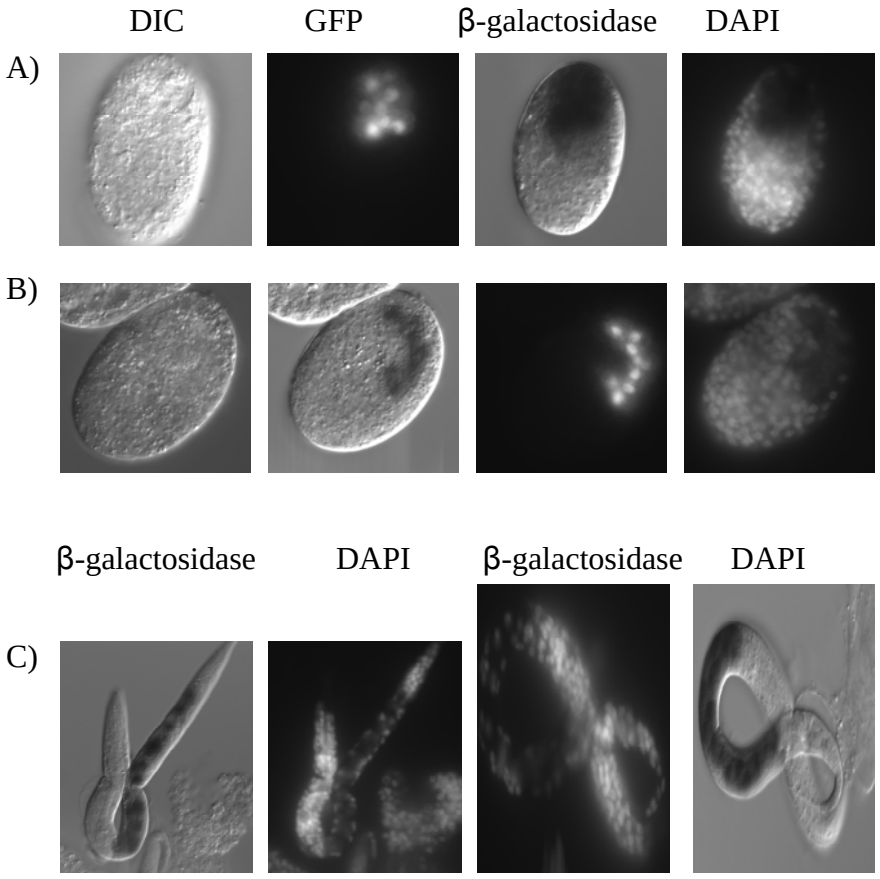


Figure 20. Construct pJM67 transgenic lines. pJM67 contains the complete 5140bp of the 5' flanking region of *elt-2*. A) + B) Nomarski image, GFP/fluorescence, β-galactosidase staining, DAPI staining of late gastrulation embryo and comma stage embryo, respectively. C) Worms recently hatched, β-galactosidase and DAPI stained respectively.

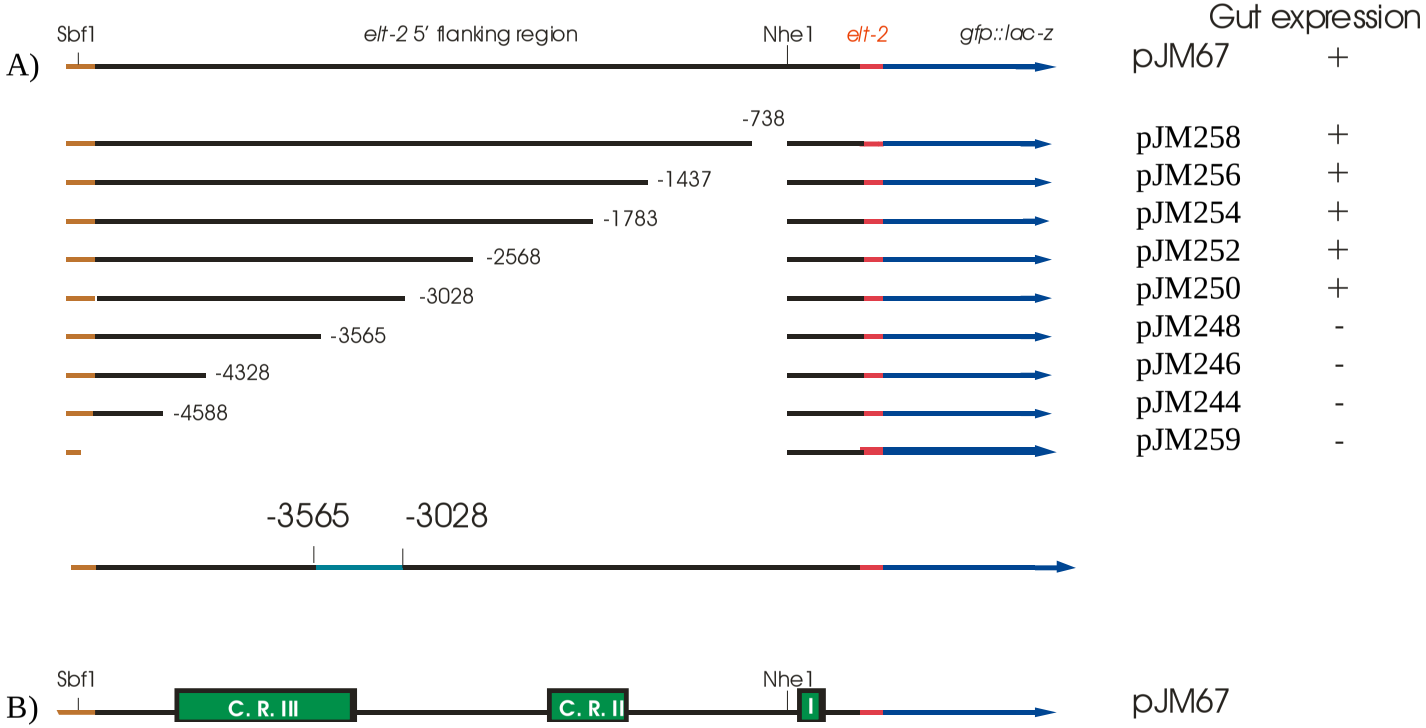


Figure 21. A) Schematic representation of the 3' deletion series. pJM67 contains the complete 5' flanking region of *elt-2* and recapitulates ELT-2 antibody staining in the embryo. All constructs contain the *elt-2* basal promoter located 3' of the NheI restriction digest site. Red bars represent the 98 base pairs of the *elt-2* gene; blue bars represent the coding region for *lacZ/gfp*. Brown bars represent the multiple coding region and contains the restriction digest site SbfI. GFP expression is seen in the developing intestine from transgenics containing plasmids pJM67 to pJM250, but not from transgenics containing the plasmids pJM248 to pJM259. B) Position of Conserved Regions (C. R.) with respect to the full length promoter. Diagram is to scale.

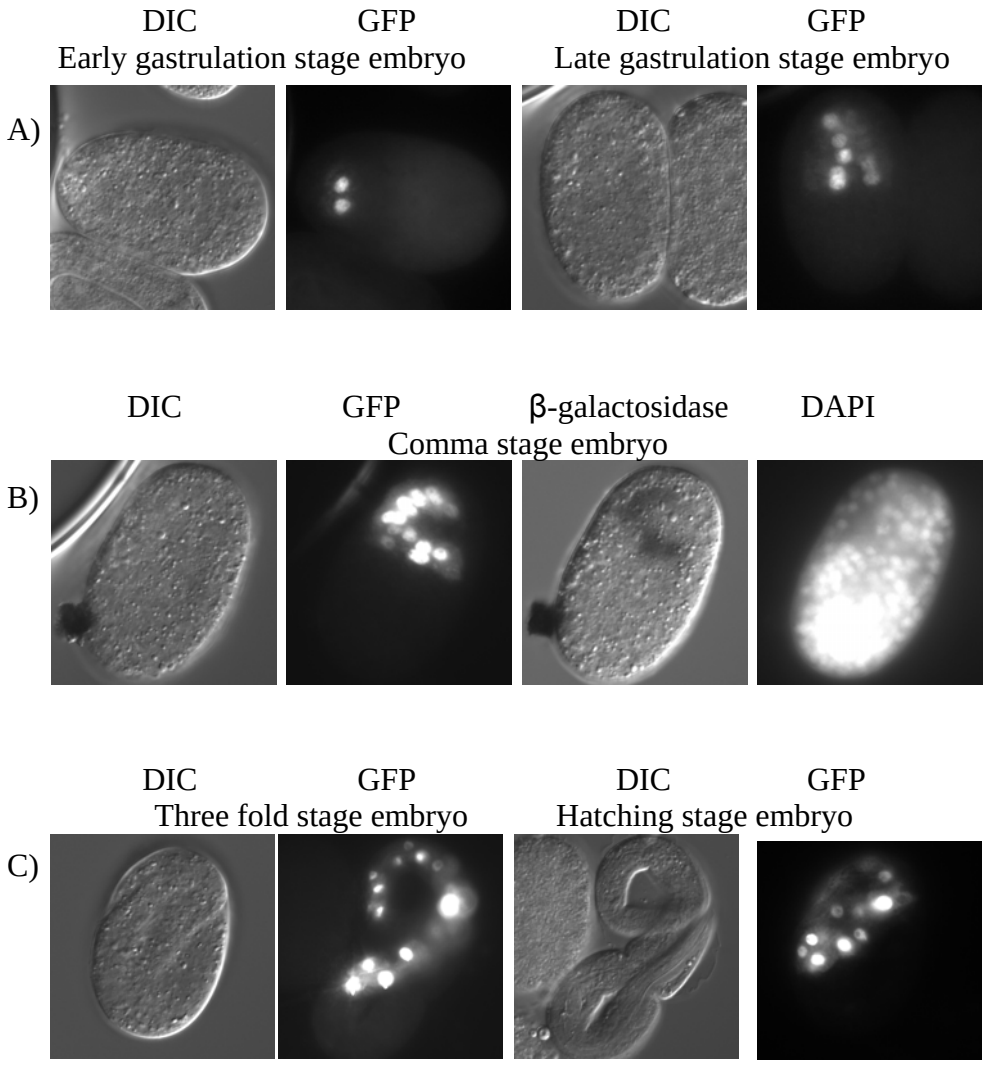


Figure 22. Construct pJM258 transgenic lines. The reporter construct, pJM258, contains the *elt-2* promoter from -5140 bp to -738 bp from the *elt-2* ATG start site fused to gfp/lacZ. A) Nomarski and GFP/fluorescence image of early gastrulation stage embryo and late gastrulation stage embryo, respectively B) Comma stage embryo image; Nomarski, GFP/ fluorescence, β-galactosidase and DAPI stained, respectively C) Late stage of embryogenesis Nomarski and GFP/ fluorescence image; Three fold and shortly after hatching.

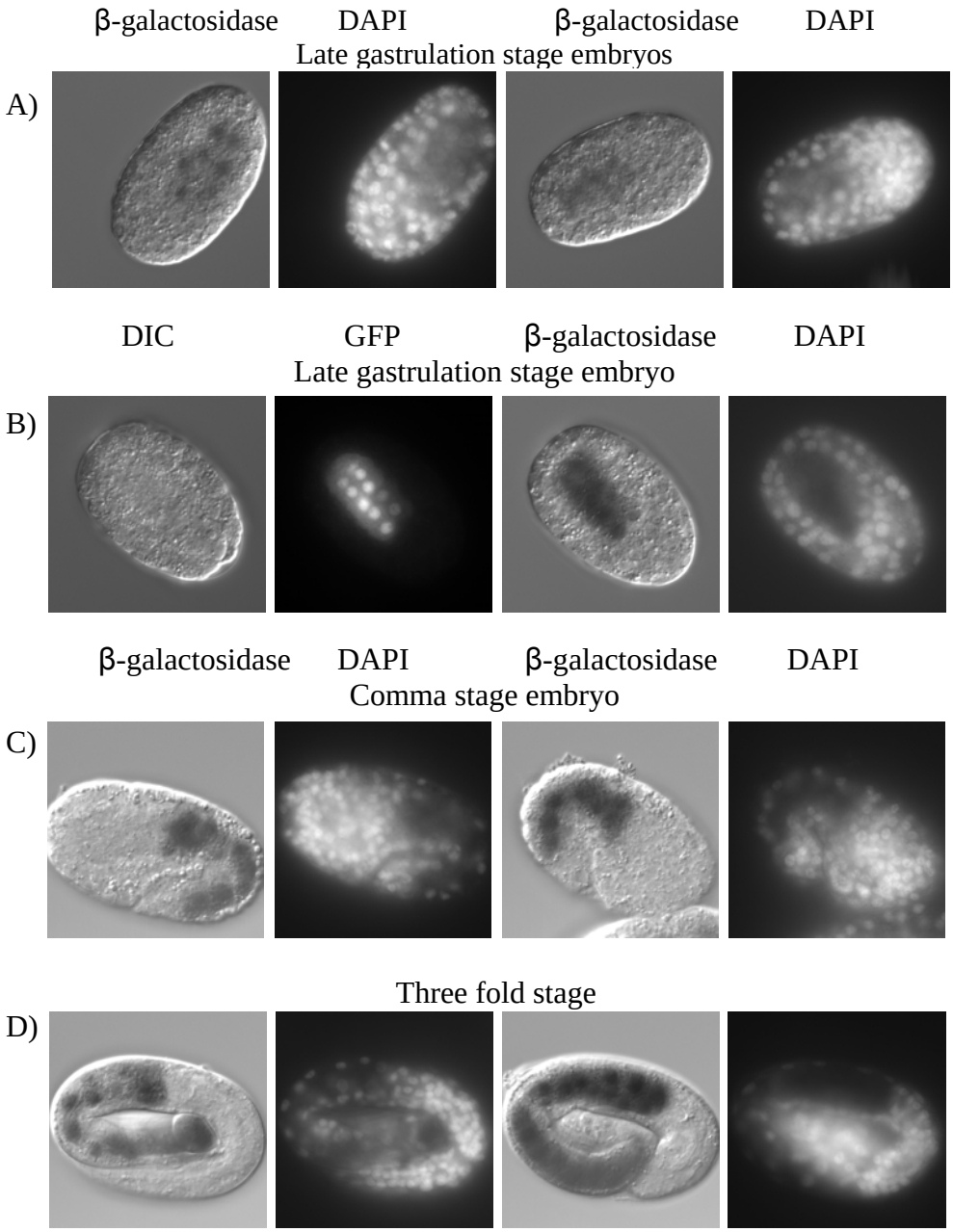


Figure 23. Construct pJM256 transgenic lines. The reporter construct, pJM256, contains the *elt-2* promoter from -5140 bp to -1426 bp from the *elt-2* ATG start site fused to gfp/lacZ. A) β-galactosidase and DAPI stained embryos at late gastrulation stage B) Nomarski, GFP/ fluorescence, β-galactosidase and DAPI stained; image of late gastrulation embryo C) β-galactosidase and DAPI stained embryos at comma stage D) β-galactosidase and DAPI stained embryos at the three fold stage.

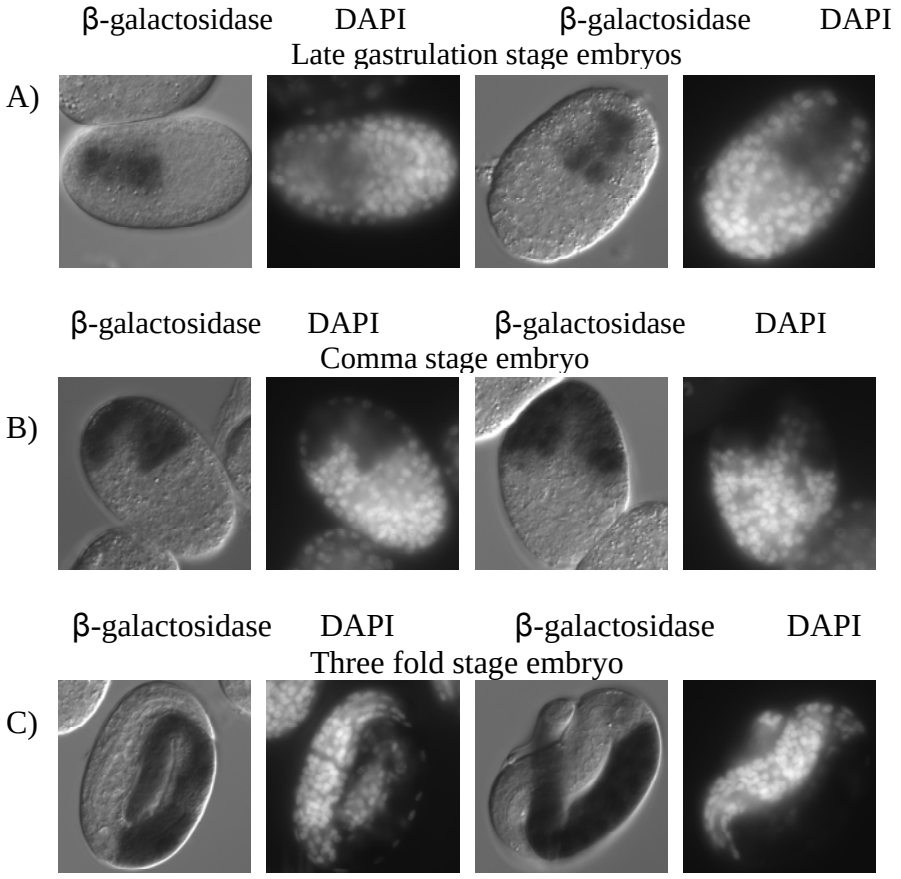


Figure 24. Construct pJM254 transgenic lines. The reporter construct, pJM254, contains the *elt-2* promoter from -5140 bp to -1783 bp from the *elt-2* ATG start site fused to gfp/lacZ. A) β-galactosidase and DAPI stained embryos at late gastrulation stage B) β-galactosidase and DAPI stained embryos at the comma stage C) β-galactosidase and DAPI stained embryos at the three fold stage.

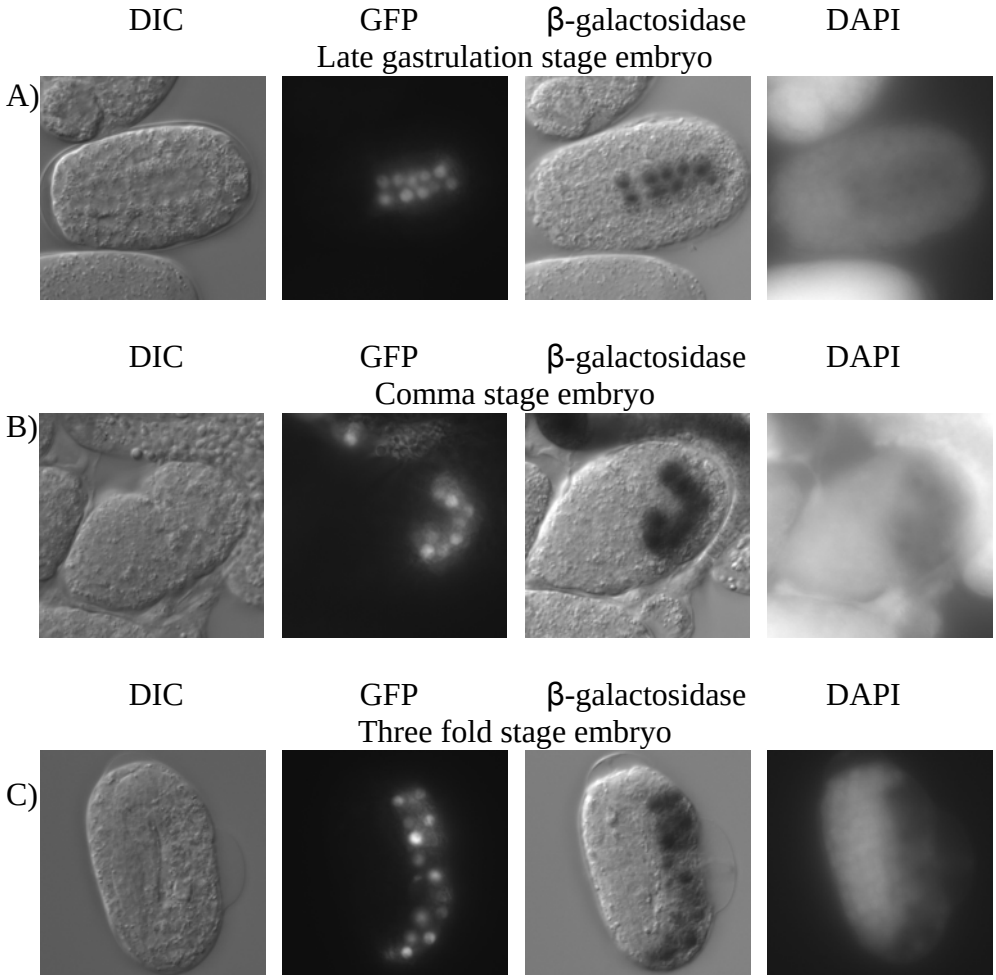


Figure 25. Construct pJM252 transgenic lines. The reporter construct, pJM252, contains the *elt-2* promoter from -5140 bp to -2568 bp from the *elt-2* ATG start site fused to gfp/lacZ. Nomarski image, GFP/ fluorescence image, β-galactosidase and DAPI stained embryos at A) late gastrulation B) comma stage C) three fold stage.

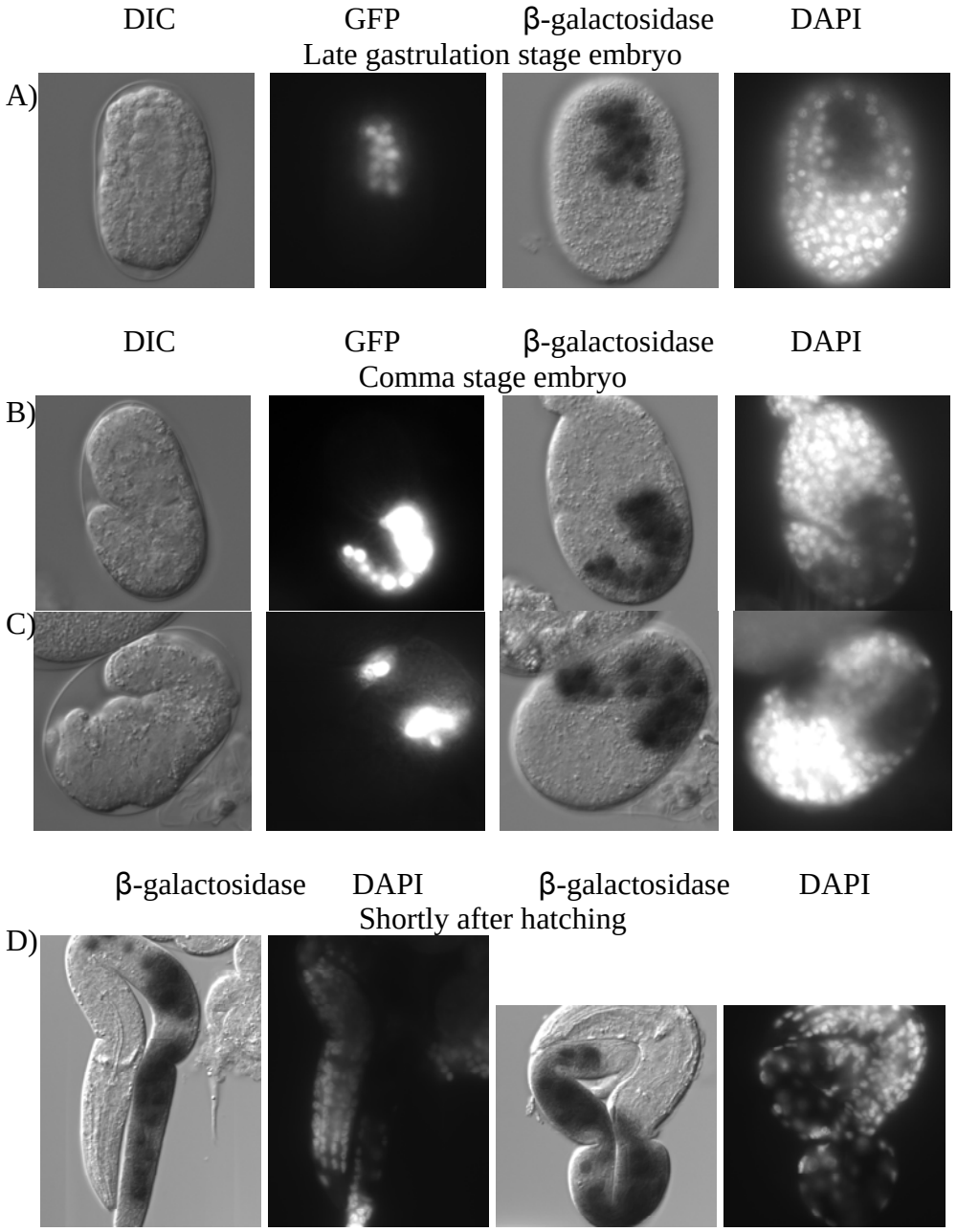


Figure 26. Construct pJM250 transgenic lines. The reporter construct, pJM250, contains the *elt-2* promoter from -5140 bp to -3028 bp from the *elt-2* ATG start site fused to gfp/lacZ. Nomarski image, GFP/ fluorescence image, β-galactosidase and DAPI stained embryos at A) Late gastrulation stage B) comma stage C) comma stage. Note embryo in C. is malformed compared to embryo in B. D) larvae shortly after hatching; β-galactosidase and DAPI stained, respectively.

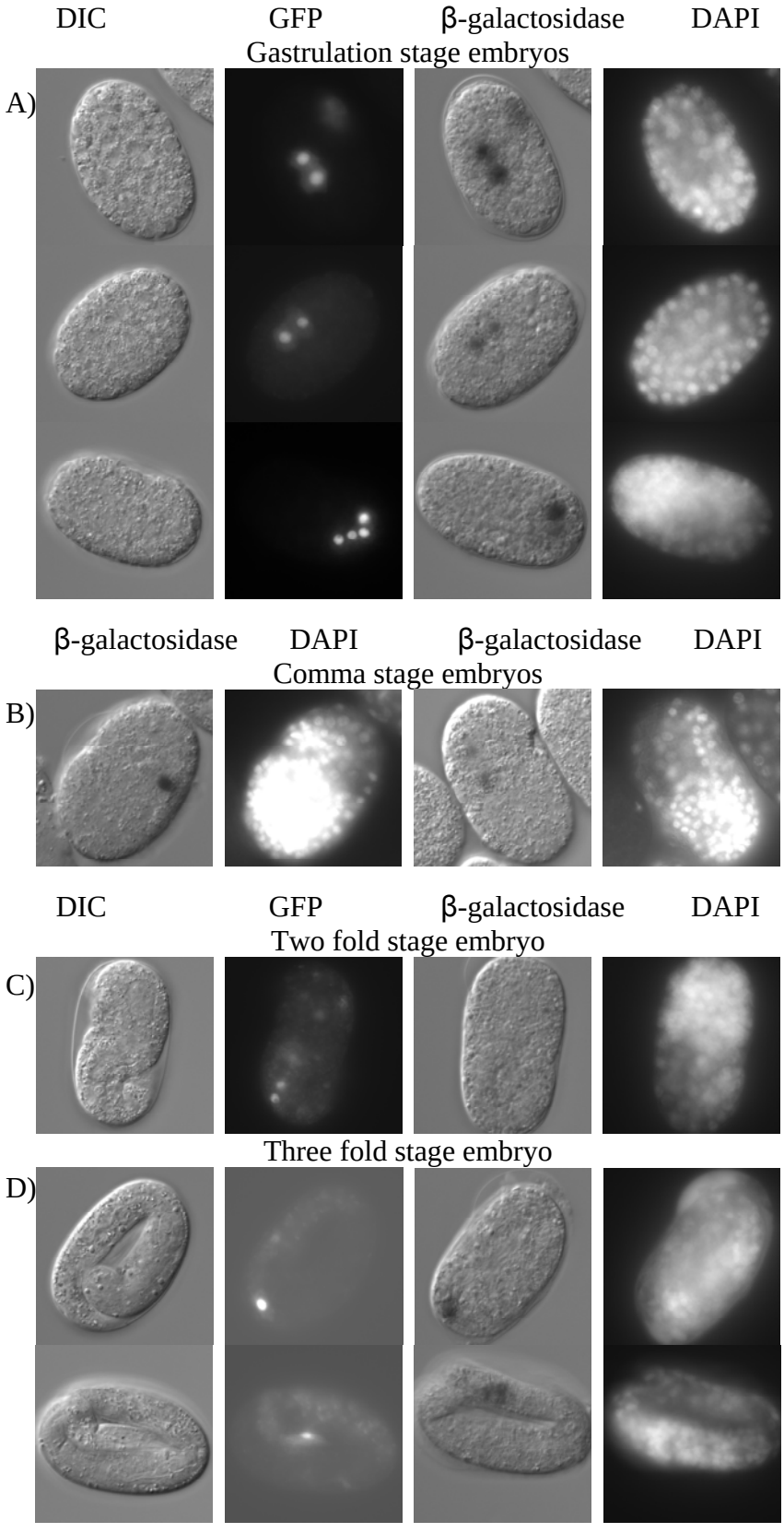

DIC
GFP
β-galactosidase
DAPI
Gastrulation stage embryos
A)
β-galactosidase
DAPI
β-galactosidase
DAPI
Comma stage embryos
B)
DIC
GFP
β-galactosidase
DAPI
Two fold stage embryo
C)
Three fold stage embryo
D)

Figure 27. Construct pJM248 transgenic lines. The reporter construct, pJM248, contains the *elt-2* promoter from -5140 bp to -3565 bp from the *elt-2* ATG start site fused to gfp/lacZ. A) Nomarski image, GFP/ fluorescence image, β-galactosidase and DAPI stained embryos at gastrulation stage B) β-galactosidase and DAPI stained embryos at comma stage. Nomarski image, GFP/fluorescence image, β-galactosidase and DAPI stained embryos at C) two-fold stage D) three-fold stage.

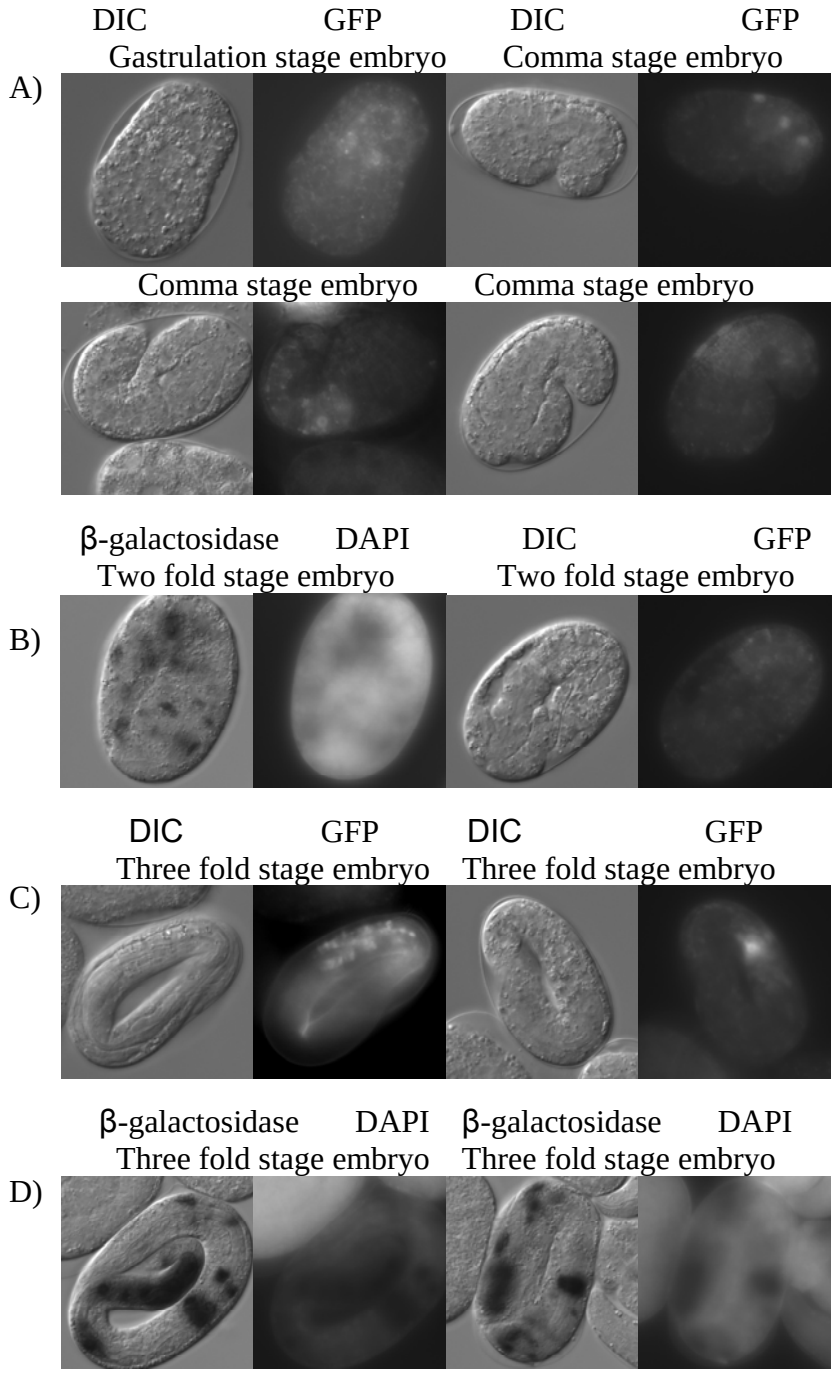


Figure 28. Construct pJM246 transgenic lines. The reporter construct, pJM246, contains the *elt-2* promoter from -5140 bp to -4328 bp from the *elt-2* ATG start site fused to gfp/lacZ. A) Nomarski and GFP/ fluorescence images of gastrulation and comma stage embryos. B) Embryos at comma stage. Left panel shows embryo with ectopic β-galactosidase staining outside of the intestine but morphology is normal. Right panel shows aberrant embryo. C) Nomarski and GFP/fluorescence image of three-fold stage embryos D) three-fold stage embryos; β-galactosidase and DAPI stained.

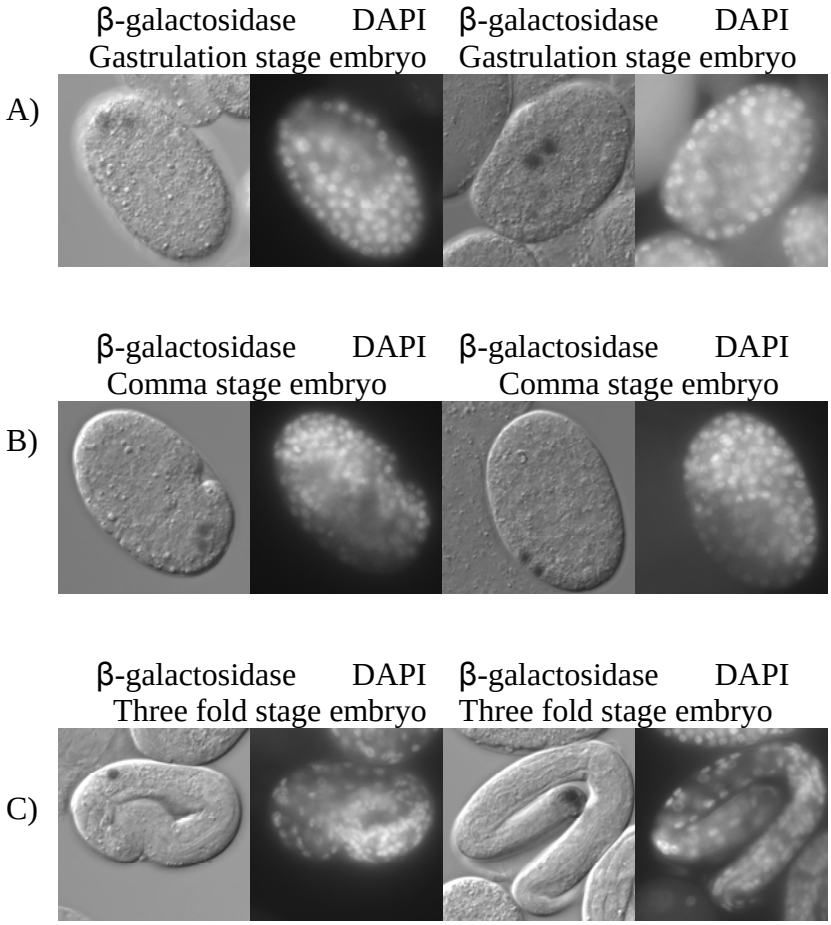


Figure 29. Construct pJM244 transgenic lines. The reporter construct, pJM244, contains the *elt-2* promoter from -5140 bp to -4588 bp from the *elt-2* ATG start site fused to gfp/lacZ. β-galactosidase and DAPI stained embryos at A) Gastrulation stage B) Comma stage C) Three-fold stage.

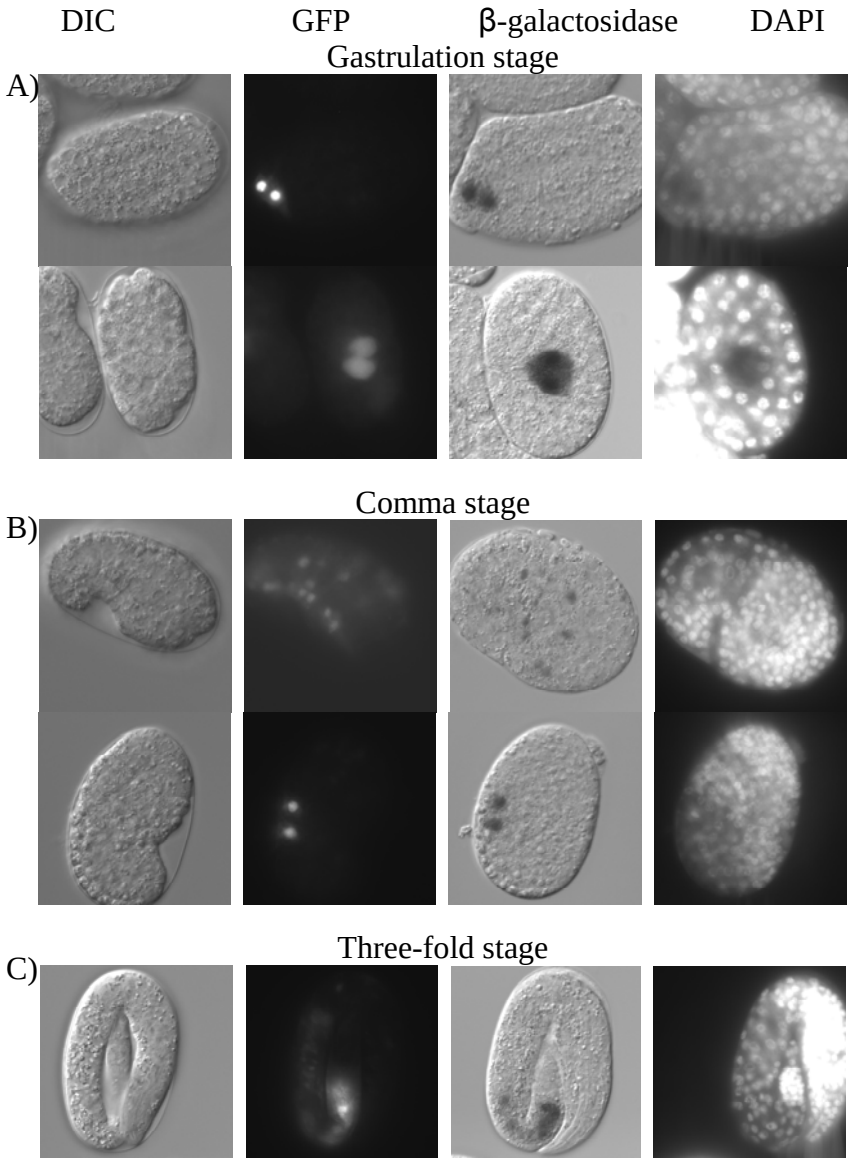


Figure 30. Construct pJM259 transgenic lines. The reporter construct, pJM259, contains only the *elt-2* basal promoter fused to gfp/lacZ. Nomarski, GFP/fluorescence, β-galactosidase and DAPI stained embryos at A) Gastrulation stage B) Comma stage C) Three-fold stage.

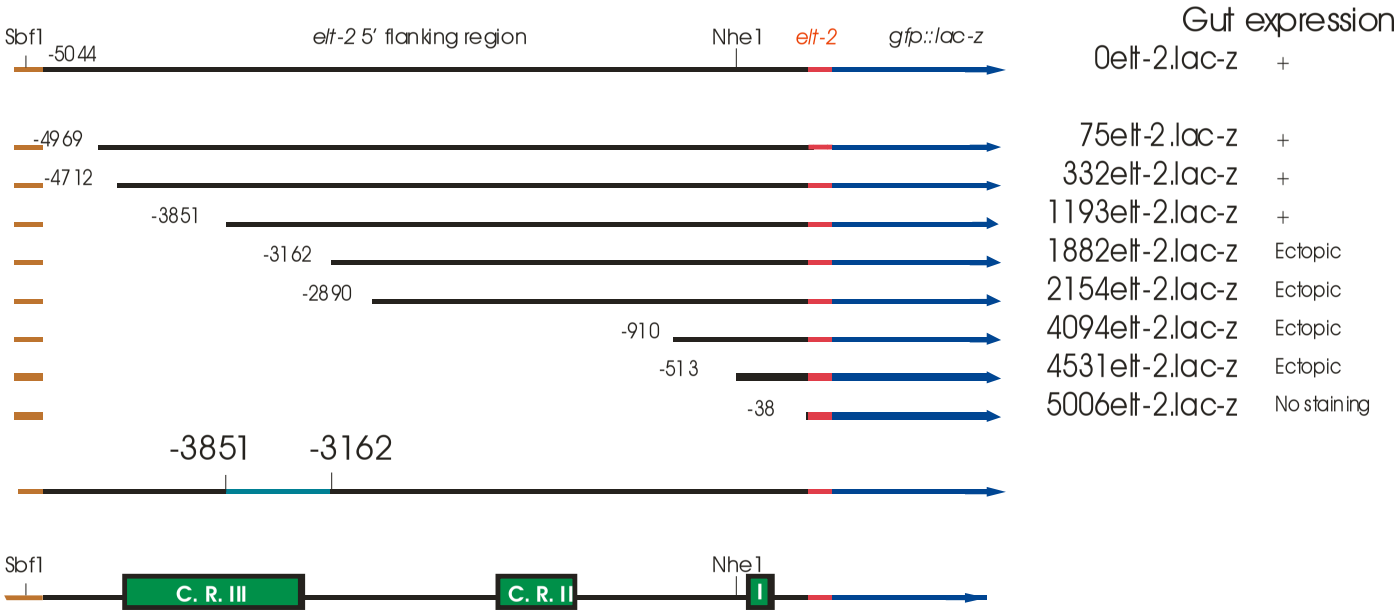


Figure 31. A) Schematic representation of the 5' deletion series. Red bars represent the 98 base pairs of the *elt-2* gene; blue bars represent the coding region for *lacZ.* Brown bar represents the multiple coding region. B) Position of Conserved Regions (C. R.) with respect to the full length promoter. Diagram is to scale.

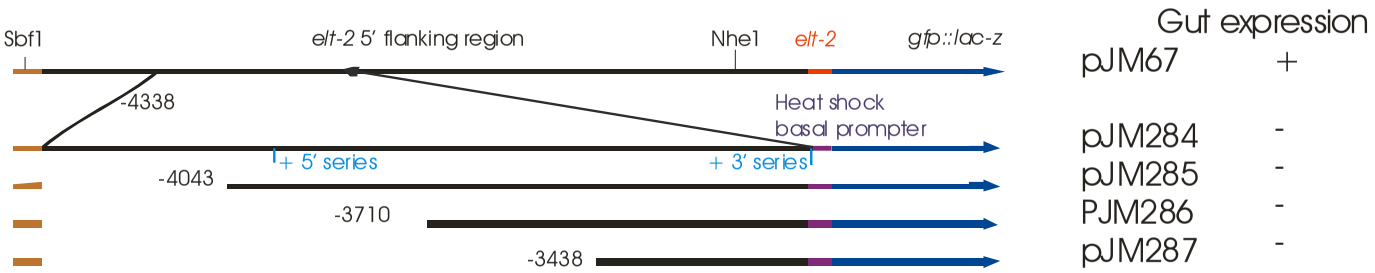


Figure 32. 5' enhancer deletion series contains four reporter constructs. pJM284 has from -4338 to -3025 base pairs of the *elt-2* promoter inserted into the expression testing vector pJM76. pJM284 and pJM285 contain the minimum enhancer region as determined by the 5' and 3' deletion series. Transgenics containing pJM284 to pJM287 did not have GFP/β-galactosidase expression.
The basal promoter used for the series is a heat shock basal promoter and not the *elt-2* basal promoter.

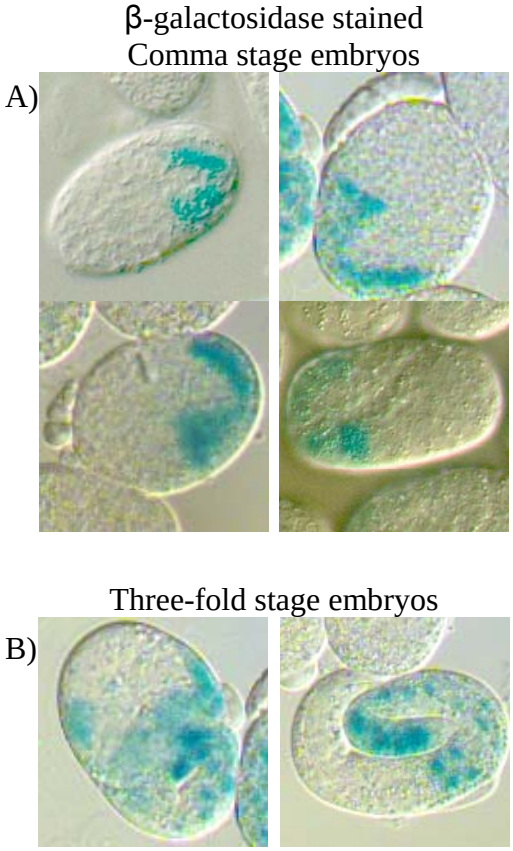


Figure 33. A) Transgenics containing the reporter construct pJM309 have expression in embryogenesis. pJM309 contains -4157 to -3309 base pairs of the *elt-2* promoter (from the ATG start site) inserted upstream of the *elt-2* basal promoter and first 98 bp of coding region fused to gfp/lacZ. A) Comma stage embryos stained with β-galactosidase B)Three-fold stage embryos stained with β-galactosidase. Transgenic lines are noted as containing a high percentage of malformed and dead embryos.

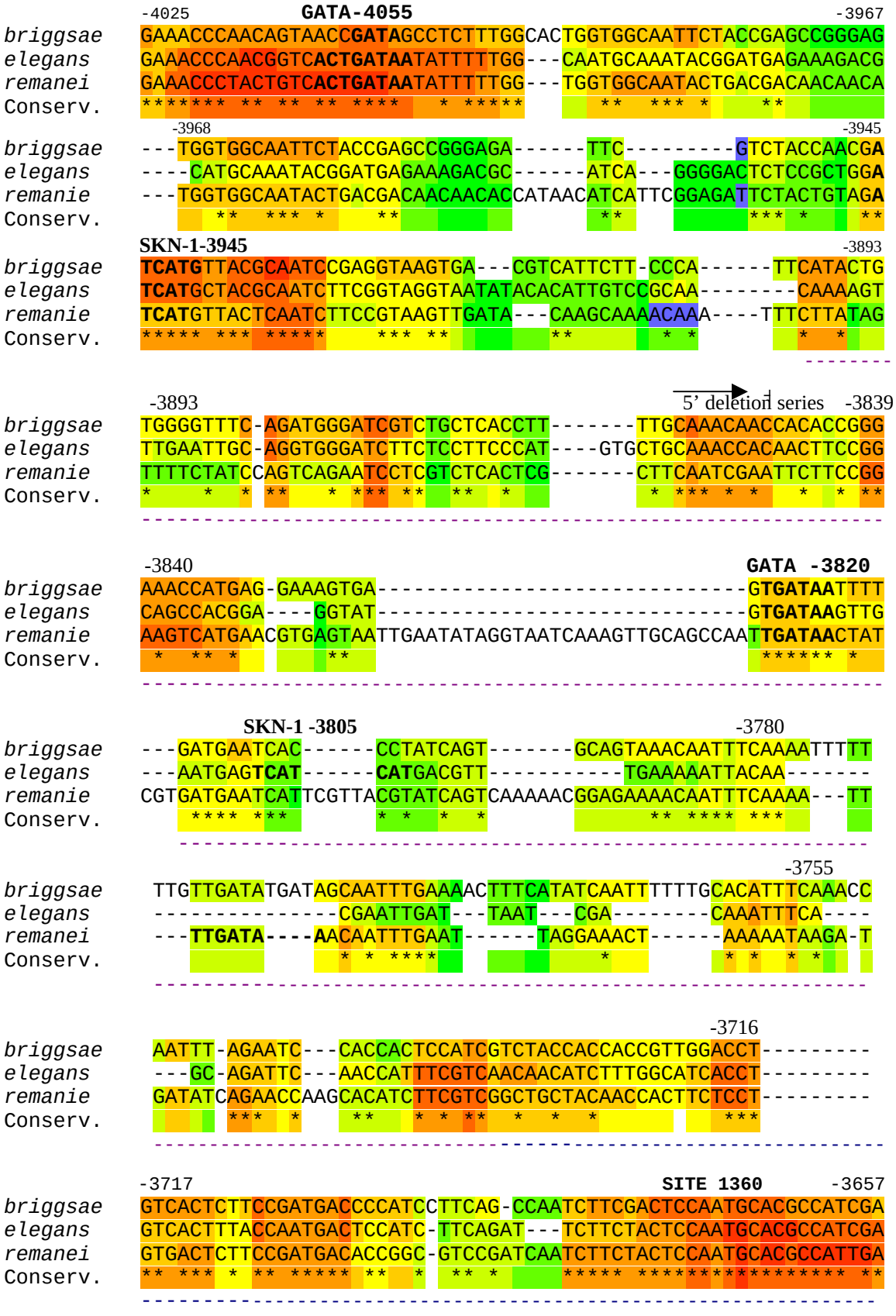

-4025
GATA-4055
-3967
briggsae
elegans
remanei
Conserv.
-3968
-3945
SKN-1-3945
-3893
5' deletion series
-3839
-3840
GATA -3820
SKN-1 -3805
-3780
-3755
-3716
-3717
SITE 1360
-3657

Figure 34A. Proximal region of Conserved Region III is divided into two pages. This page contains the sequence from -4025 to -3657 bp upstream of the *elt-2* ATG start site and contains important sites in the regulation of *elt-2*; 5’ deletion series transgenics with the reporter construct 1882elt-2.lacz have GFP expression in the intestine (5’ end of the insert marked with arrow); GATA-3820, Site 1360. Purple dashed lines are under sequences used for probe A (-3886 to -3740), Blue dashed lines are under sequences used for probe B (-3739 to 3540).

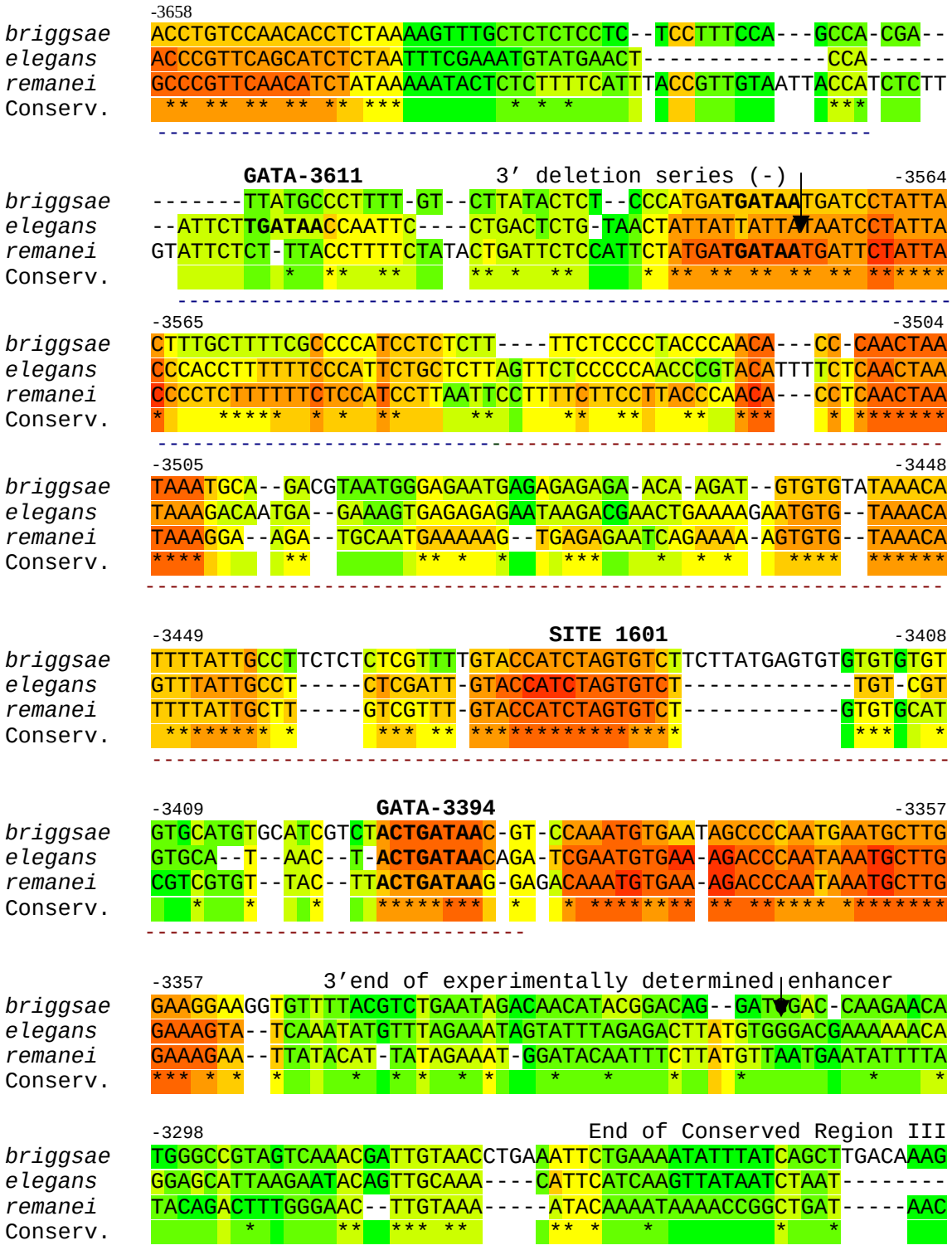

-3658
briggsae ACCTGTCCAACACCTCTAAAAGTTTGCTCTCTCCTC--TCCTTTCCA---GCCA-CGA--
elegans ACCCGTTCAGCATCTCTAATTTCGAAATGTATGAACT--------------CCA------
remanei GCCCGTTCAACATCTATAAAAATACTCTCTTTTCATTTACCGTTGTAATTACCATCTCTT
Conserv.
GATA-3611
3′ deletion series (-)
-3564
briggsae -------TTATGCCCTTTT-GT--CTTATACTCT--CCCATGATGATAATGATCCTATTA
elegans --ATTCTTGATAACCAATTC----CTGACTCTG-TAACTATTATTATTATAATCCTATTA
remanei GTATTCTCT-TTACCTTTTCTATACTGATTCTCCATTCTATGATGATAATGATTCTATTA
Conserv.
-3565
-3504
briggsae CTTTGCTTTTCGCCCCATCCTCTCTT----TTCTCCCCTACCCAACA---CC-CAACTAA
elegans CCCACCTTTTTTCCCCATTCTGCTCTTAGTTCTCCCCCAACCCGTACATTTTCTCAACTAA
remanei CCCCTCTTTTTTCTCCATCCTTAATTCCTTTTCTTCCTTACCCAACA---CCTCAACTAA
Conserv.
-3505
-3448
briggsae TAAATGCA--GACGTAATGGGAGAATGAGAGAGAGA-ACA-AGAT--GTGTGTATAAACA
elegans TAAAGACAATGA--GAAAGTGAGAGAGAATAAGACGAACTGAAAAGAATGTG--TAAACA
remanei TAAAGGA--AGA--TGCAATGAAAAAG--TGAGAGAATCAGAAAA-AGTGTG--TAAACA
Conserv.
-3449
SITE 1601
-3408
briggsae TTTTATTGCCTTCTCTCTCGTTTTGTACCATCTAGTGTCTTCTTATGAGTGTGTGTGTGT
elegans GTTTATTGCCT-----CTCGATT-GTACCATCTAGTGTCT-------------TGT-CGT
remanei TTTTATTGCTT-----GTCGTTT-GTACCATCTAGTGTCT------------GTGTGCAT
Conserv.
-3409
GATA-3394
-3357
briggsae GTGCATGTGCATCGTCTACTGATAAC-GT-CCAAATGTGAATAGCCCCAATGAATGCTTG
elegans GTGCA--T--AAC--T-ACTGATAACAGA-TCGAATGTGAA-AGACCCAATAAATGCTTG
remanei CGTCGTGT--TAC--TTACTGATAAG-GAGACAAATGTGAA-AGACCCAATAAATGCTTG
Conserv.
-3357
3′end of experimentally determined enhancer
briggsae GAAGGAAGGTGTTTTACGTCTGAATAGACAACATACGGACAG--GATGGAC-CAAGAACA
elegans GAAAGTA--TCAAATATGTTTAGAAATAGTATTTAGAGACTTATGTGGGACGAAAAAACA
remanei GAAAGAA--TTATACAT-TATAGAAAT-GGATACAATTTCTTATGTTAATGAATATTTTA
Conserv.
-3298
End of Conserved Region III
briggsae TGGGCCGTAGTCAAACGATTGTAACCTGAAATTCTGAAAATATTTATCAGCTTGACAAAG
elegans GGAGCATTAAGAATACAGTTGCAAA----CATTCATCAAGTTATAATCTAAT--------
remanei TACAGACTTTGGGAAC--TTGTAAA-----ATACAAAATAAAACCGGCTGAT-----AAC
Conserv.

Figure 34B. Proximal region of Conserved Region III is separated into two pages. This page contains from (-3658 to -3252) and contains important sites in the regulation of *elt-2*. In the 3' deletion series transgenics for pJM248 do not have robust expression in the intestine (3' end of insert is -3575 bp from *elt-2* ATG site and is marked with an arrow); Site 1601; GATA-3394; transgenics for the enhancer construct pJM309 have GFP expression (3'end of insert is -3309 bp from *elt-2* ATG); end of Conserved Region III. Blue dashed lines are under sequences used for probe B (-3739 to 3540). Brown dashed lines are under sequences used for probe C (-3539 to -3388).

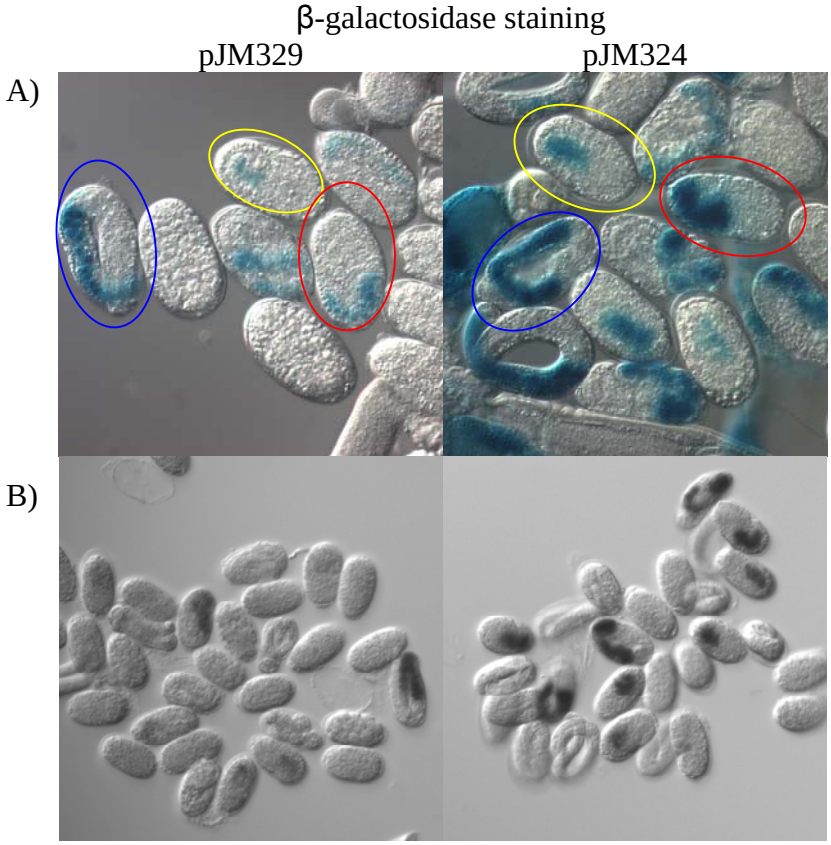


Figure 35. The reporter construct pJM329 has GATA site -3820 mutated in the full length *elt-2* promoter fused to GFP/LacZ. Embryos and larvae were visualized for GFP and β-galactosidase. A) Transgenics for pJM329 were stained in parallel to transgenics containing pJM324 (full length *elt-2* promoter fused to GFP/LacZ). Color image shows pJM329 embryos are weaker than pJM324 embryos; gastrulation embryo (yellow circle); comma stage embryos (red circle); and three-fold stage embryo (blue circle) B) Black and white image of pJM329 transgenics stained in parallel to pJM324 transgenics.

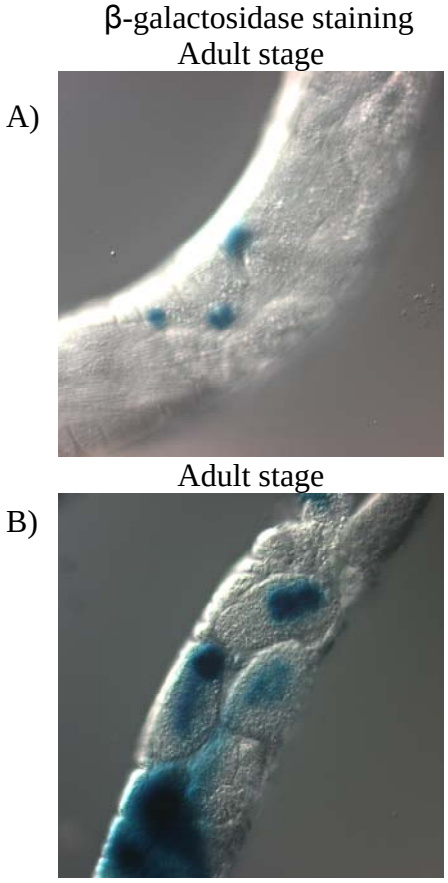


Figure 36. The reporter construct pJM325 has GATA site -3394 mutated within the distal region of the *elt-2* promoter (-5050 to -2568 bp upstream of ATG) fused to the *elt-2* basal promoter and GFP/LacZ. The construct was injected at (5 µg/mL) and produced numerous stable healthy lines. PCR verified 13 of the 14 lines contained the plasmid. Embryos and larvae were visualized for GFP and β-galactosidase. A) Adult worm containing reporter construct pJM325 has ectopic staining in the cuticle B) pJM324 transgenics stained in parallel to worm in A. β-galactosidase staining is seen in embryos as well as in adult.

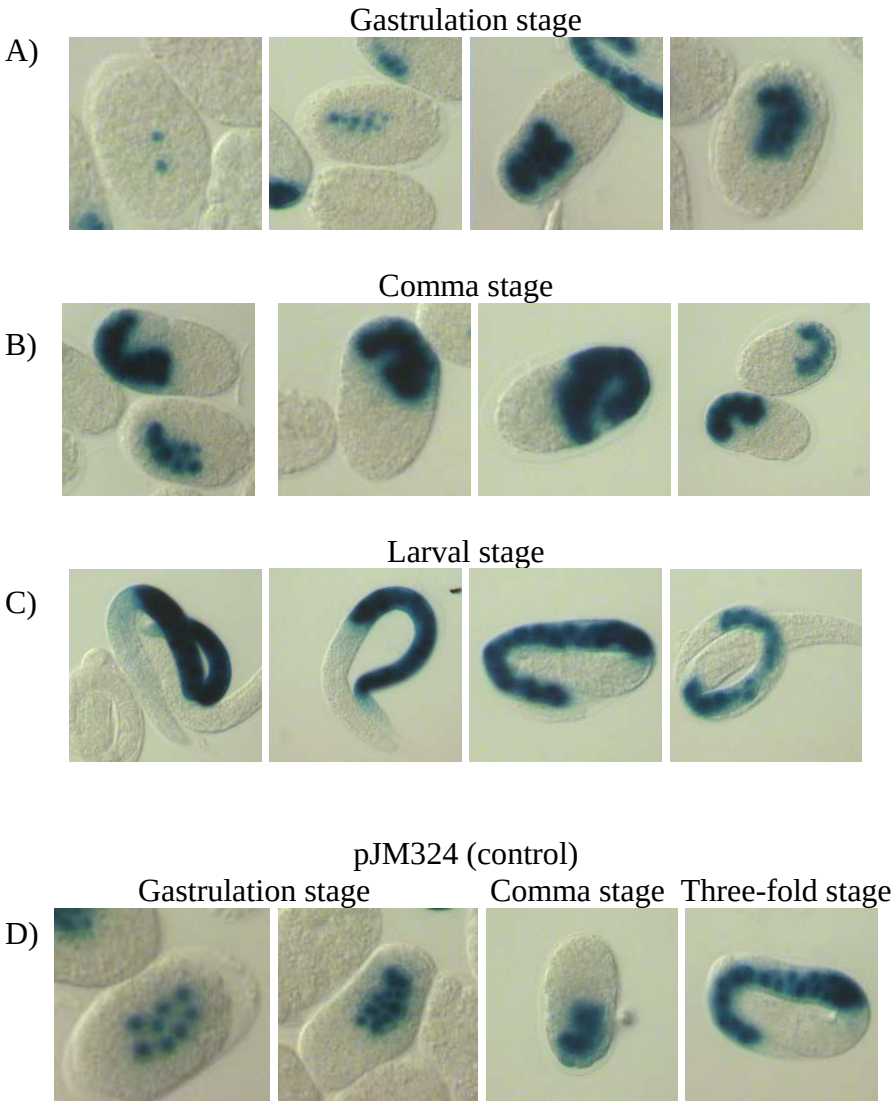


Figure 37. pJM328 has the full length *elt-2* promoter with Site 1360 mutated (11bps altered -3717 bps upstream of the *elt-2* ATG site). Transgenic worms containing pJM328 have predominantly a two fold increase in GFP and β-galactosidase expression at all stages in development. A) Gastrulation stage embryos stained with β-galactosidase. Embryos in the left panels have weaker staining than embryos in the right panels indicating variability in intensity of β-galactosidase staining. B) Comma stage embryos. Note right panel contains a strong and weak staining embryo. C) Hatched worms have strong staining compared to pJM324 control worms. D) pJM324 control worms at the gastrulation, comma and three-fold stage, respectively.

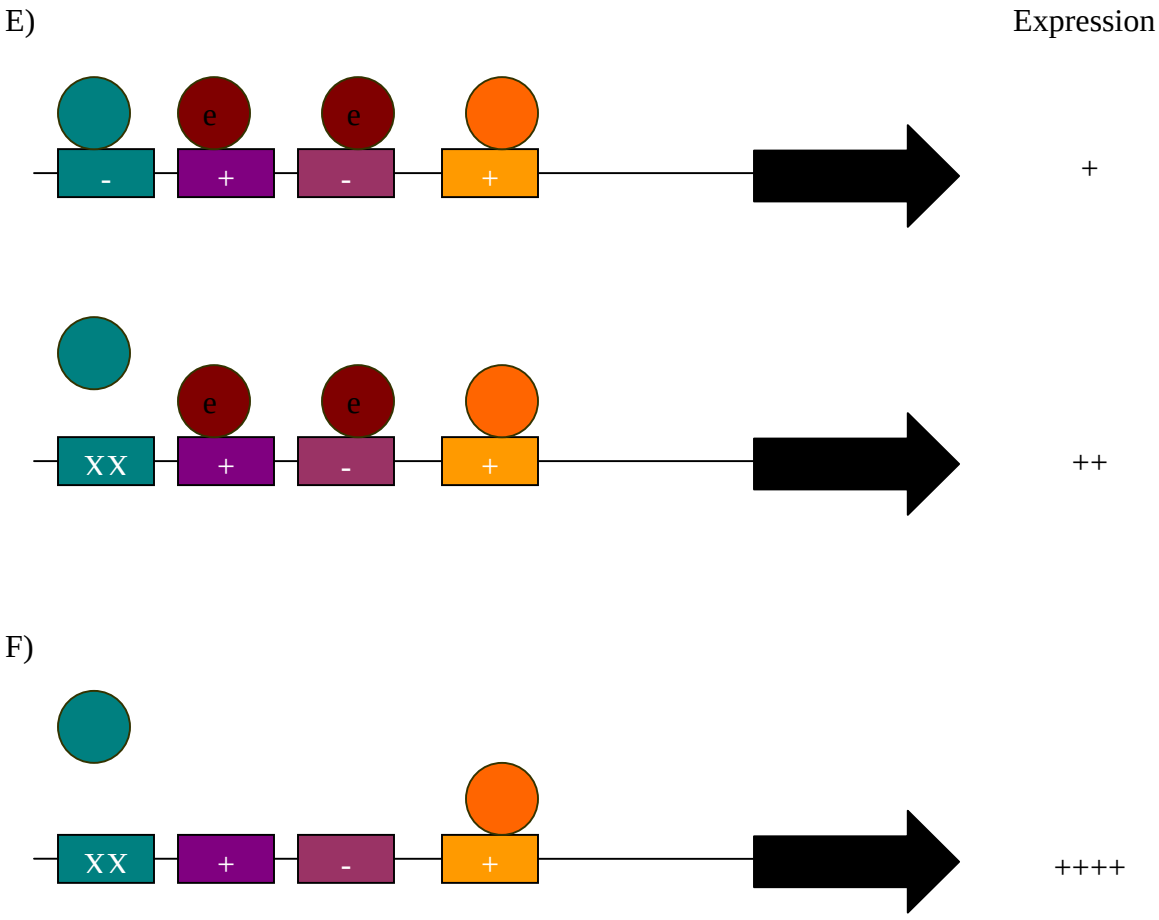


Figure 37. Green circle represents negative regulator and its corresponding binding site (Site 1360). Red circle represents ELT-2 binding to both its negative and positive regulation sites. Orange circle represents positive regulator binding to its corresponding orange site.
E) Under normal conditions *elt-2* is regulated by both positive and negative regulators. In the absence of Site 1360, expression increases. F) In the absence of Site 1360 and the ELT-2 protein, the positive regulator still can bind and increase *elt-2* transcription.

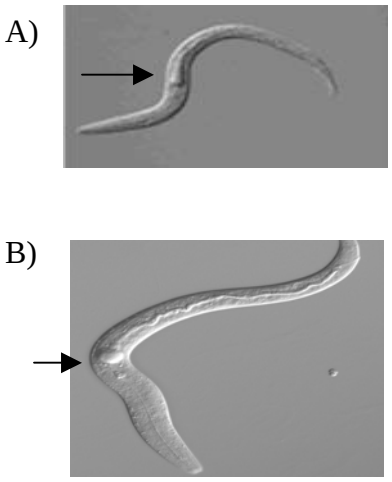


Figure 38. A) Transgenics containing reporter constructs with 98 bp of *elt-2* coding region have a gut obstructed (Gob) phenotype B) dsRNA corresponding to the *elt-2* gene injected into Wild-type worms have progeny with a Gob phenotype. Arrow points to the obstructed area in anterior portion of the intestine.

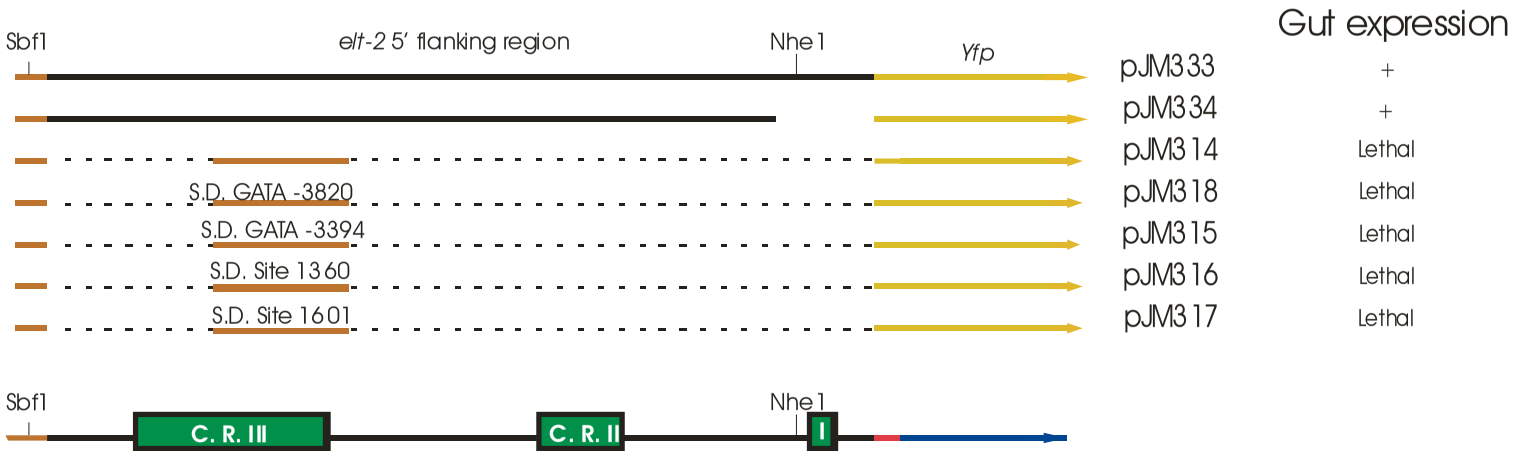


Figure 39. Reporter constructs pJM314 to pJM317 are lethal to worms. pJM314 to pJM317 contain the enhancer region of the *elt-2* promoter (-4147 to -3309 base pairs from *elt-2* ATG start site) fused to a *yfp* reporter backbone. pJM315 to pJM318 contain site directed (S.D) mutations within the enhancer; pJM318 has GATA -3820 altered from gtgataagtt to g**GTCGAC**gtt; pJM315 has GATA -3394 altered from tactgataacag to ta**GAGTCGACC**ag; pJM316 has site 1360 altered from aa**tgca**cg**ccatcga**acc to aa**GTAC**cg**ATCGATC**acc; pJM317 has site 1601 altered from taccatctagtgt to ta**AACCAGCTG**gt.

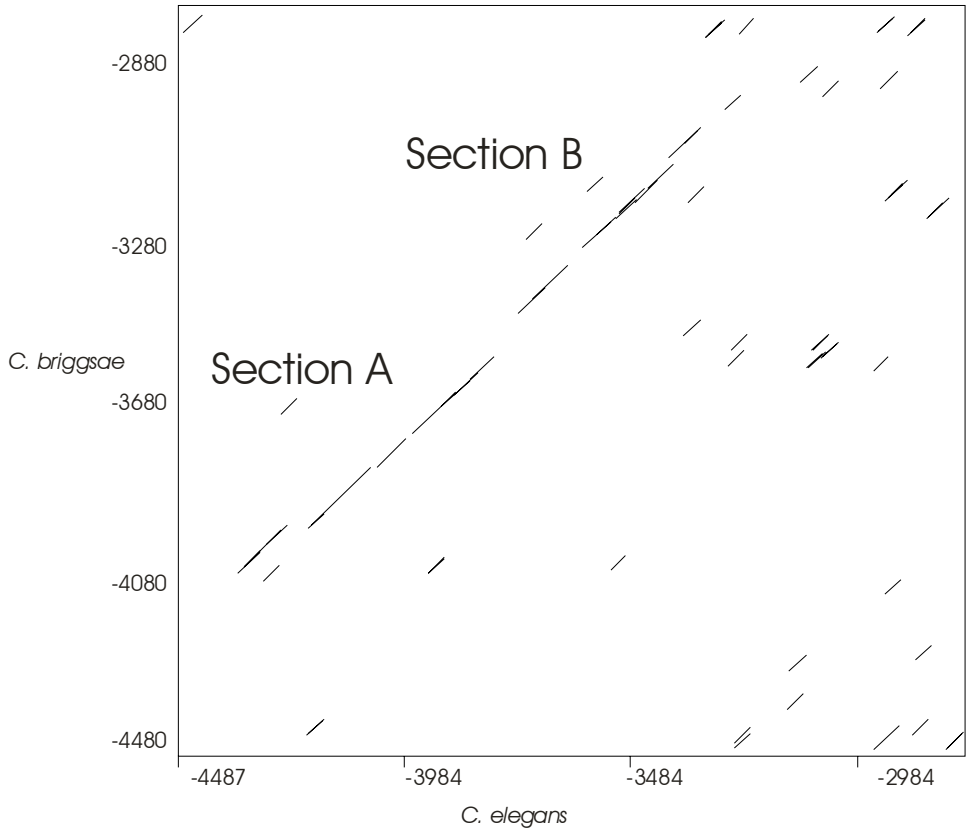


Figure 40. Within the distal end of the promoter is Conserved Region III. Conserved Region III is 998 base pairs in length. The computer program Dotmatcher was used to compare Conserved Region III between *C. briggsae* and *C. elegans*. Each dot on the matrix indicates a 35 bp sequence conservation within a sliding window of 40 base pairs. Dot Matrix of Conserved Region III indicates the region is highly conserved and may be sectioned into two parts Section A (distal) and B (proximal).

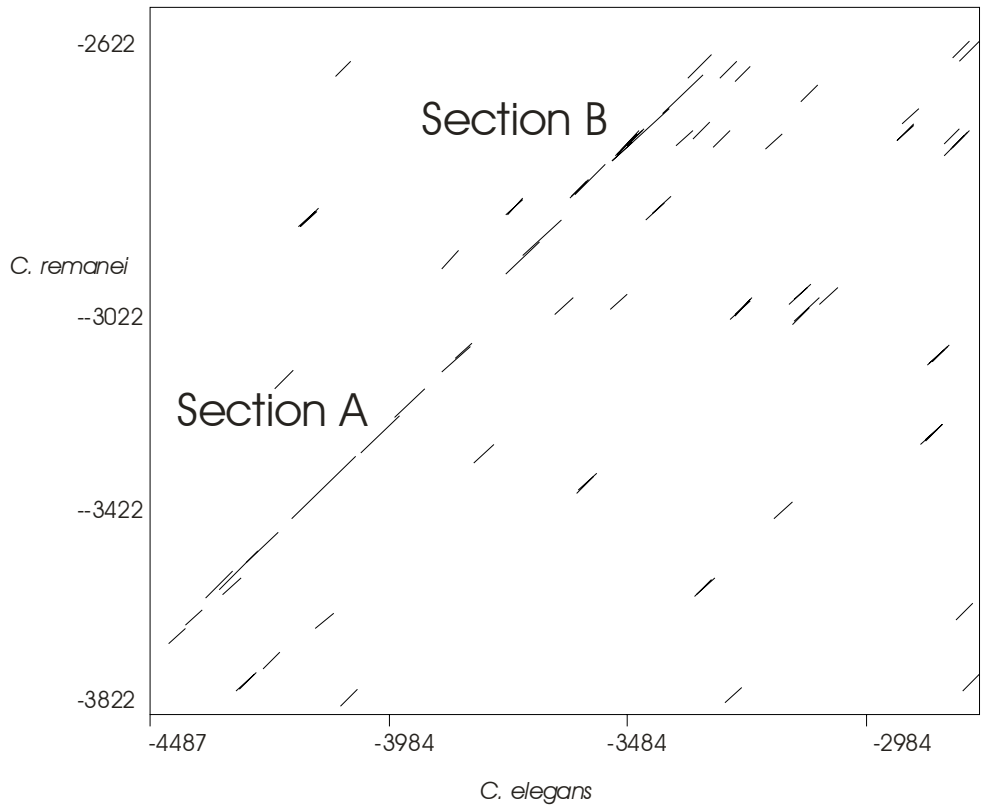


Figure 41. Within the distal end of the promoter is Conserved Region III. Conserved Region III is 998 base pairs in length. The computer program Dotmatcher was used to compare Conserved Region III between *C. remanei* and *C. elegans*. Each dot on the matrix indicates a 35 bp sequence conservation within a sliding window of 40 base pairs. Dot Matrix of Conserved Region III indicates the region is highly conserved and may be sectioned into two parts Section A (distal) and B (proximal).

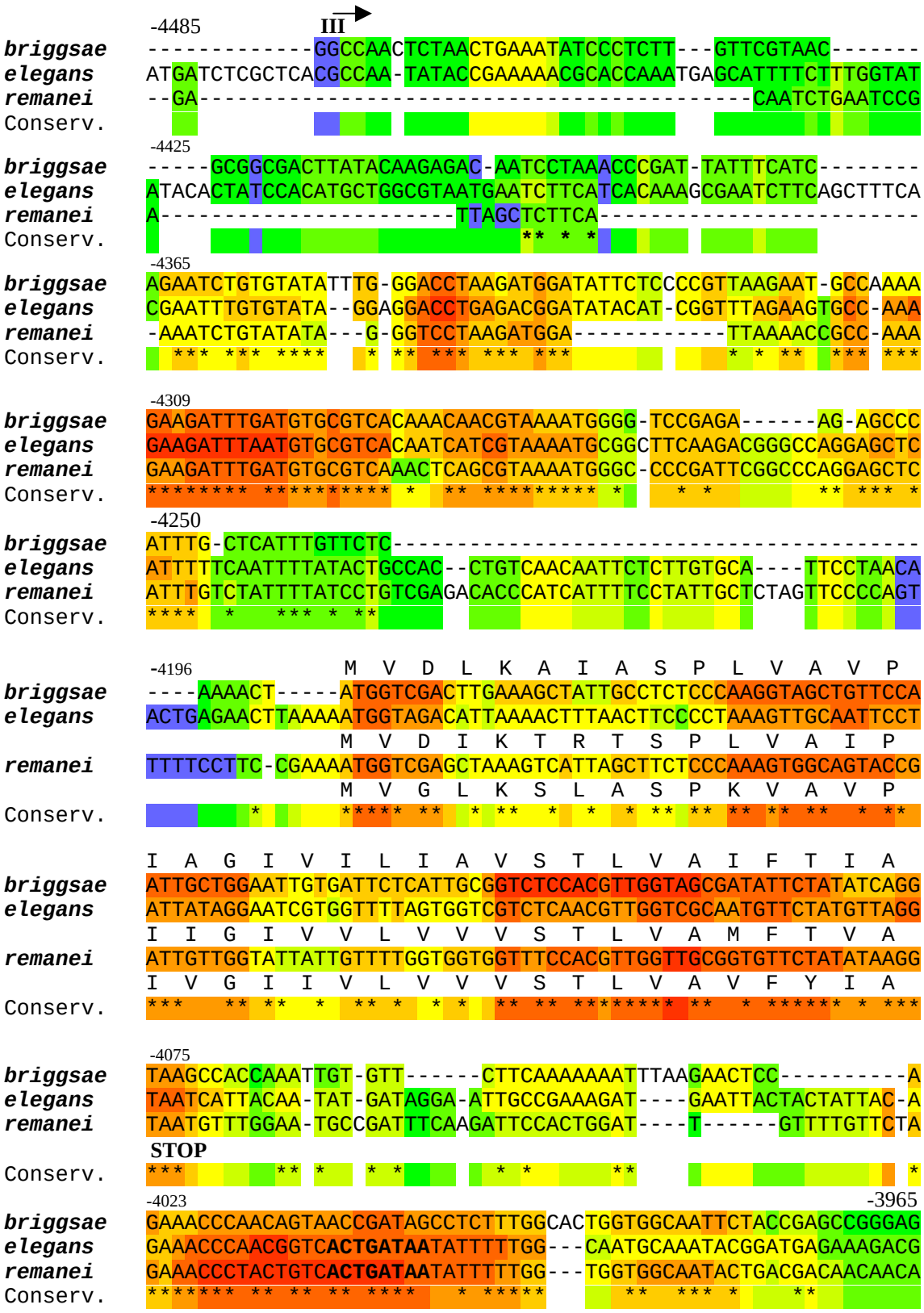

-4485
III
briggsae ----------GGCCAACTCTAACTGAAATATCCCTCTT---GTTCGTAAC-------
elegans ATGATCTCGCTCACGCCAA-TATACCGAAAAACGCACCAAATGAGCATTTTCTTTGGTAT
remanei --GA----------CAATCTGAATCCG
Conserv.
-4425
briggsae -----GCGGCGACTTATACAAGAGAC-AATCCTAAACCCGAT-TATTTCATC--------
elegans ATACACTATCCACATGCTGGCGTAATGAATCTTCATCACAAAGCGAATCTTCAGCTTTCA
remanei A----------TTAGCTCTTCA----------
Conserv.
-4365
briggsae AGAATCTGTGTATATTTG-GGACCTAAGATGGATATTCTCCCCGTTAAGAAT-GCCAAAA
elegans CGAATTTGTGTATA--GGAGGACCTGAGACGGATATACAT-CGGTTTAGAAGTGCC-AAA
remanei -AAATCTGTATATA---G-GGTCCTAAGATGGA----------TTAAAACCGCC-AAA
Conserv.
-4309
briggsae GAAGATTTGATGTGCGTCACAAACAACGTAAAATGGGG-TCCGAGA------AG-AGCCC
elegans GAAGATTTAATGTGCGTCACAATCATCGTAAAATGCGGCTTCAAGACGGGCCAGGAGCTC
remanei GAAGATTTGATGTGCGTCAAACTCAGCGTAAAATGGGC-CCCGATTCGGCCCAGGAGCTC
Conserv.
-4250
briggsae ATTTG-CTCATTTGTTCTC----------
elegans ATTTTTCAATTTTATACTGCCAC--CTGTCAACAATTCTCTTGTGCA----TTCCTAACA
remanei ATTTGTCTATTTTATCCTGTCGAGACACCCATCATTTTCCTATTGCTCTAGTTCCCCAGT
Conserv.
-4196
M V D L K A I A S P L V A V P
briggsae ----AAAACT-----ATGGTCGACTTGAAAGCTATTGCCTCTCCCAAGGTAGCTGTTCCA
elegans ACTGAGAACTTAAAAATGGTAGACATTAAAACTTTAACTTCCCCTAAAGTTGCAATTCCT
M V D I K T R T S P L V A I P
remanei TTTTCCTTC-CGAAAATGGTCGAGCTAAAGTCATTAGCTTCTCCCAAAGTGGCAGTACCG
M V G L K S L A S P K V A V P
Conserv.
I A G I V I L I A V S T L V A I F T I A
briggsae ATTGCTGGAATTGTGATTCTCATTGCGGTCTCCACGTTGGTAGCGATATTCTATATCAGG
elegans ATTATAGGAATCGTGGTTTTAGTGGTCGTCTCAACGTTGGTCGCAATGTTCTATGTTAGG
I I G I V V L V V V S T L V A M F T V A
remanei ATTGTTGGTATTATTGTTTTGGTGGTGGTTTCCACGTTGGTTGCGGTGTTCTATATAAGG
I V G I I V L V V V S T L V A V F Y I A
Conserv.
-4075
briggsae TAAGCCACCAAATTGT-GTT------CTTCAAAAAAATTTAAGAACTCC----------A
elegans TAATCATTACAA-TAT-GATAGGA-ATTGCCGAAAGAT----GAATTACTACTATTAC-A
remanei TAATGTTTGGAA-TGCCGATTTCAAGATTCCACTGGAT----T------GTTTTGTTCTA
STOP
Conserv.
-4023
-3965
briggsae GAAACCCAACAGTAACCGATAGCCTCTTTGGCACTGGTGGCAATTCTACCGAGCCGGGAG
elegans GAAACCCAACGGTCACTGATAATATTTTTGG---CAATGCAAATACGGATGAGAAAGACG
remanei GAAACCCTACTGTCACTGATAATATTTTTGG---TGGTGGCAATACTGACGACAACAACA
Conserv.

Figure 42. Distal region of Conserved Region III (-4485 to -3965 bp upstream of the *elt-2* ATG start site) contains a putative protein coding region. The protein is conserved in all three species and starts -4180 base pairs upstream of the *elt-2* ATG start site. Wormbase predicts the protein is either 36 amino acids total length or spliced differently in all three species into larger proteins. Between *C. elegans* and *C. remanei* 24/36 amino acids are identical, while 28/36 amino acids are identical between *C. elegans* and *C. briggsae*. At the third position of each codon there is a high degree of sequence similarity (40% in *briggsae* and 56% in *remanei*).

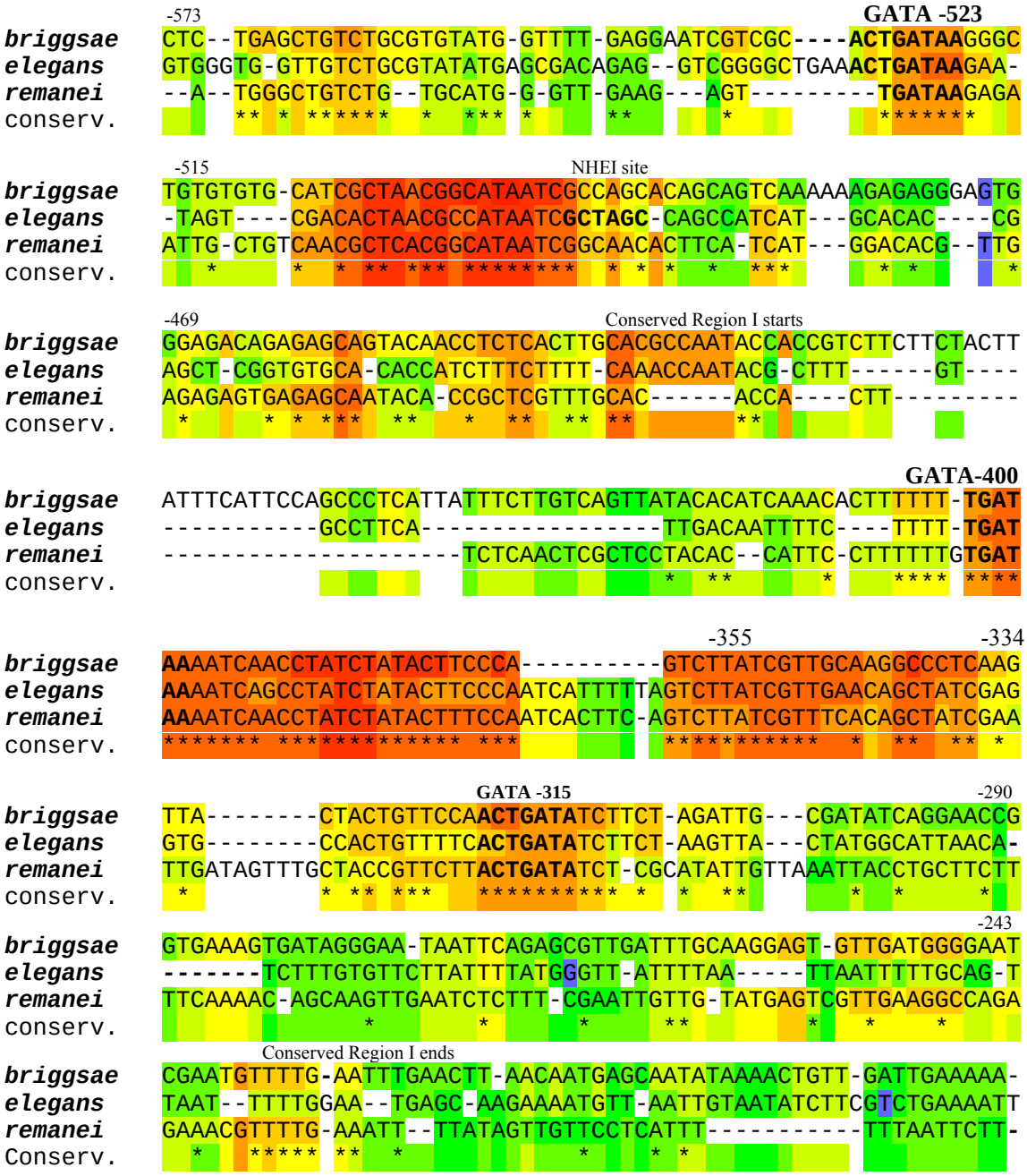


Figure 43. Conserved Region I sequence compared between *C. elegans*, *C. briggsae* and *C. remanei*. GATA site -315 is completely conserved between all three species. GATA site -400 is conserved in all three species and is part of a highly conserved domain with 36/40 bp conserved between *C. elegans* and *C. remanei* and 33/35 conserved between *C. elegans* and *C. briggsae*.

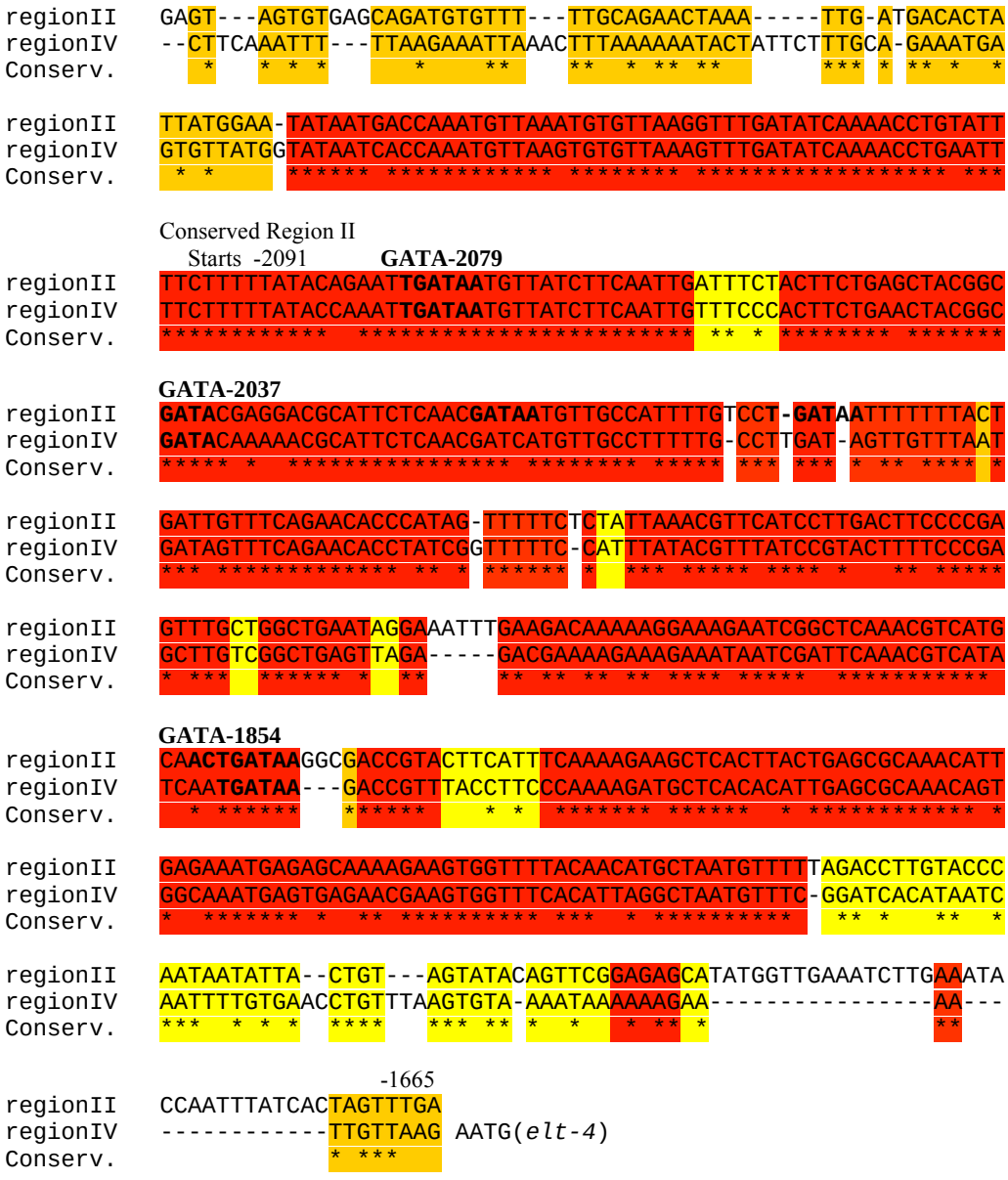


Figure 44. Sequence comparison between Conserved Region IV (*elt-4* basal promoter) and the distal region of Conserved Region II. (2091 to -1665). Three GATA sites are conserved; GATA -2079,GATA-2037 and GATA -1854.

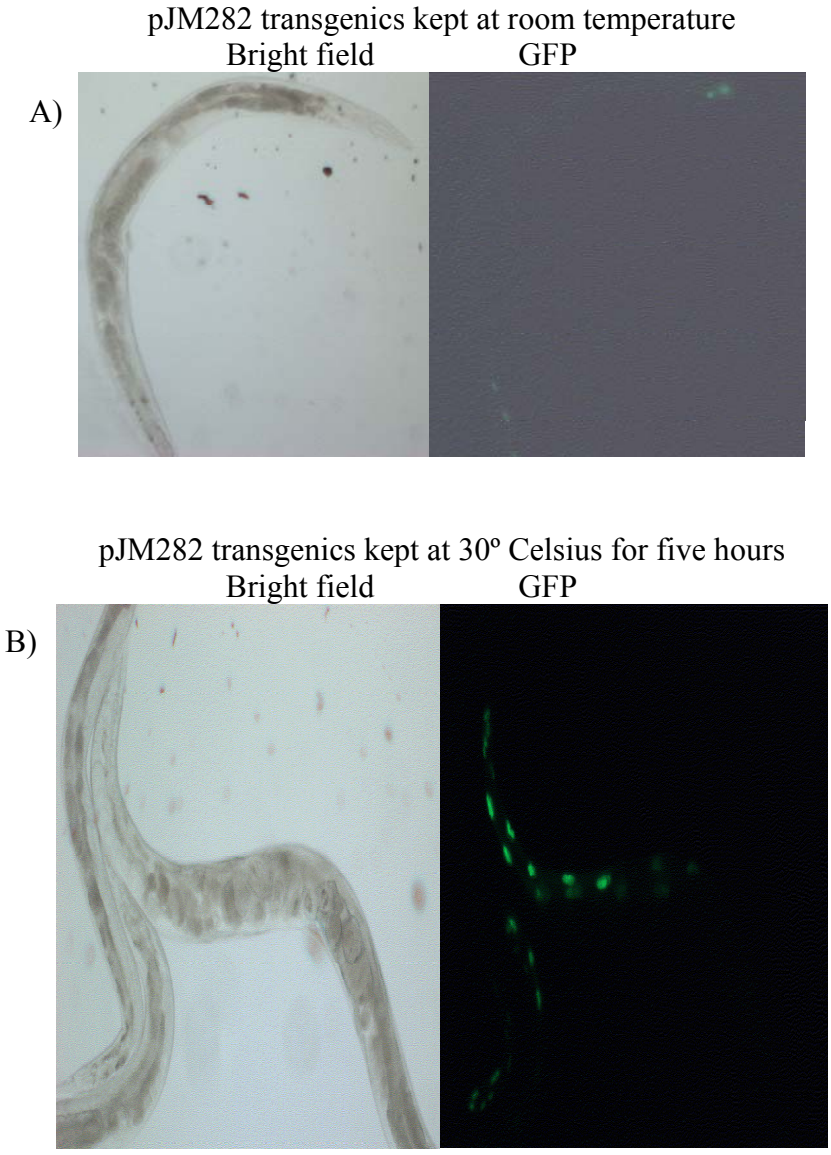


Figure 45. Adult worms containing the reporter construct pJM282 have under normal room temperature conditions an anterior posterior GFP expression in the intestine. A) Bright field and GFP/fluorescence image of adults kept at room temperature B) Bright field and GFP/fluorescence image of adults placed at 30º Celsius for five hours. Adult worms have an increase in posterior expression of GFP.

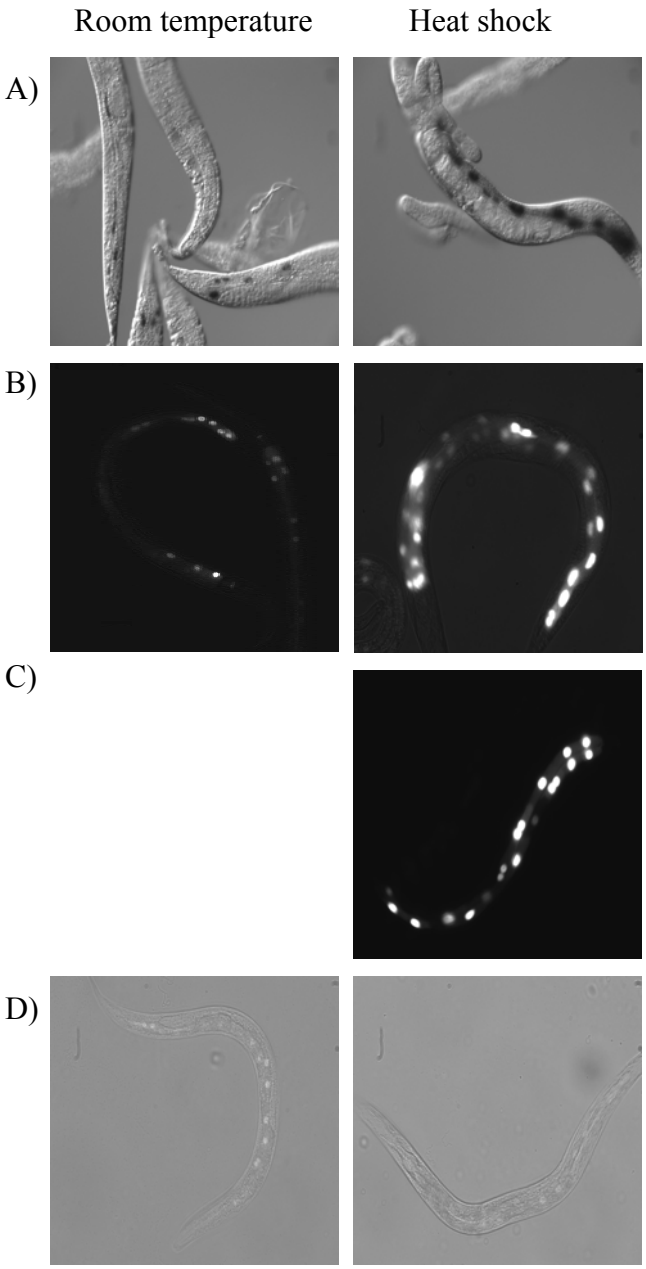


Figure 46. A) Adult worms containing the reporter construct pJM67 under room temperature and heat shocked conditions. The reporter construct contains conserved Region III, II, I B) Transgenics containing the reporter construct pJM256 have an increase in GFP expression in the intestine when worms are subjected to heat. pJM256 contains Conserved Regions III,II,I C) Transgenics containing the reporter construct pJM254 have an increase in GFP expression in the intestine when subjected to heat. pJM254 contains Conserved Region III, distal region of II, I. Image for room temperature expression is not shown D) Transgenics containing the reporter construct pJM252 do not have an increase in GFP expression when worms are subjected to heat. Both images have been altered with respect to brightness and contrast so that the worms are visible. pJM252 contains Conserved Region III and I but not II.

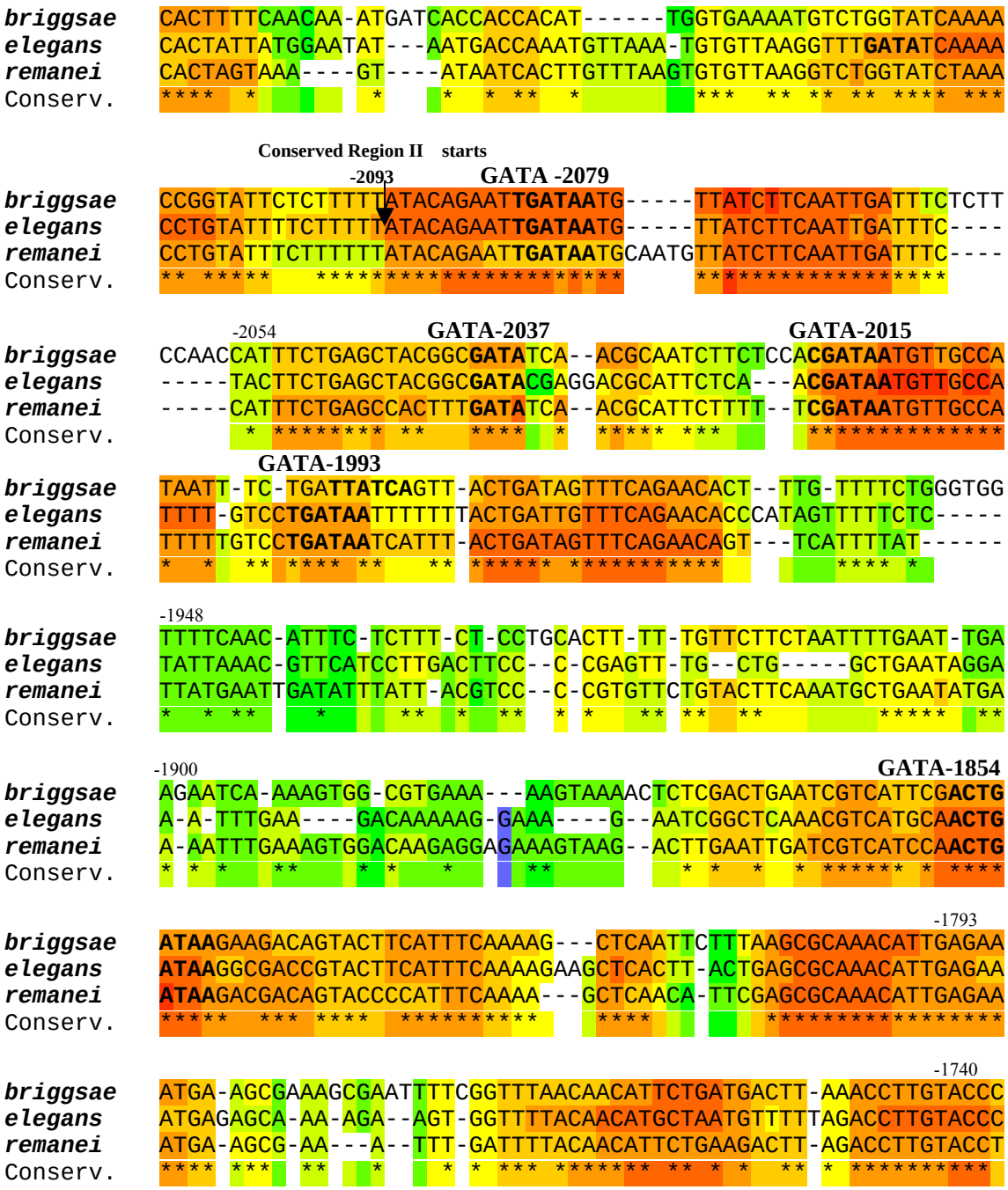


Figure 47A. The distal region of Conserved Region II (-2091 to -1740 bp from *elt-2* ATG start site). There are five conserved GATA sites with GATA site -2079 having the highest conservation between species.

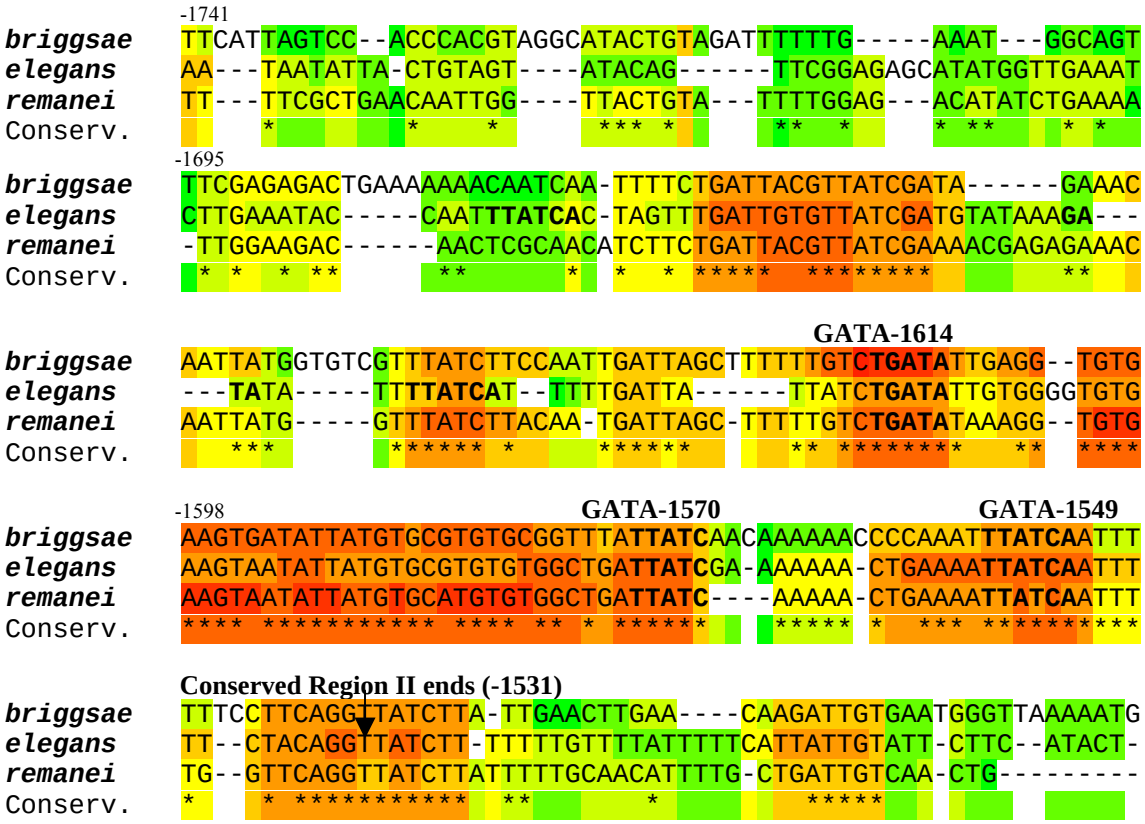


Figure 47B. The proximal Region of Conserved Region II (-1741to -1531 bp from *elt-2* ATG start site). Within the proximal region of Conserved Region II are three highly conserved GATA sites

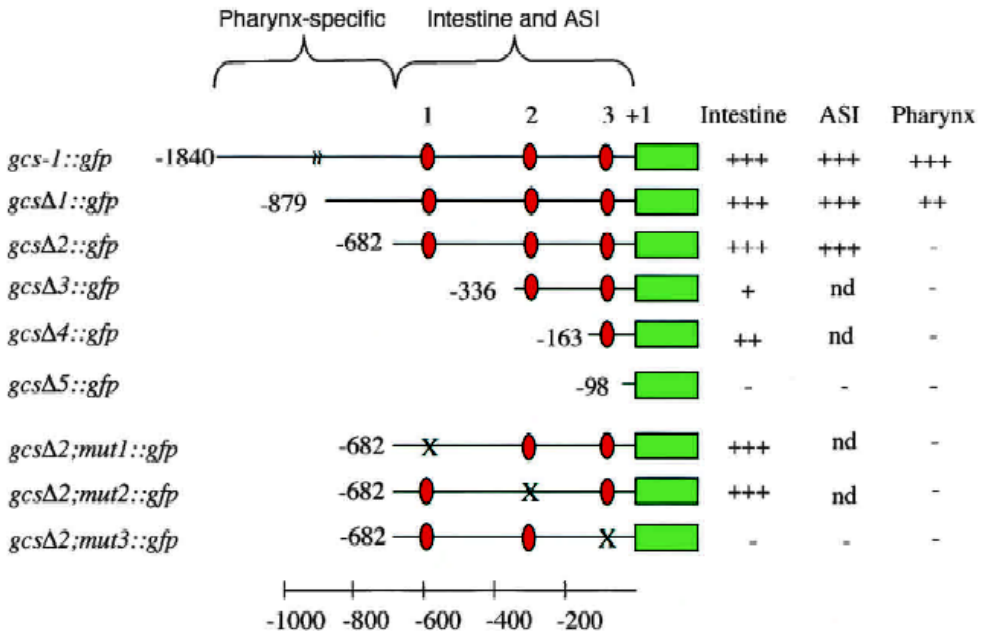


Figure 48. The expression of *gcs-1* in each tissue (intestine, pharynx and ASI neurons) is regulated through different sites within the 1840bp promoter. Expression from the indicated constructs from transgenics are indicated by plus signs, with ++ indicating a reproducible reduction and + indicating barely detectable expression. Red ovals are predicted SKN-1 binding sites. Green bar indicates the 5' end of the gcs::gfp coding region. Construct gcsΔ;mut3::gfp does not have expression in the intestine.

gcs-1 –124 CA- C**TTTATCAT**CATGA-GATTTAATGTTTCCTTTTGAT - -TTTCT -83
med-1 –127 CACCTCTGTCATCATGATGATTTTTGGAG-CATTATCATCATTTCT –83
med-1 – 127 CACCTCTGTCATCATGATGATTTTTAGAG-CATTATCATCATTTCT -83

Figure 49. SKN-1 binding sites on three separate promoters. Red sequence represents SKN-1 binding sites. Bold lettering indicates that a plausible overlapping ELT-2 binding site occurs in the same location as the SKN-1 binding site. The sequence is compared to known SKN-1 binding sites on the *med-1* gene in *C. elegans* and *C. briggsae*. Construct gcsΔ;mut3::gfp has the SKN-1/ELT-2 binding site altered to TTTCTGCAG. Worms containing the gcsΔ;mut3::gfp transgene do not contain GFP expression in the intestine.

ENZYME	GENE	SITES	Promoter length / 1024(WGATAR)
Glutamyl –cysteine syntheatase	*gcs-1*	6	1.5
NADH quinine oxidoreductase	F39B2.3	3	0.7
Glutathione S-transferase	R03D7.6	3	1.2
Glutathione S-transferase	F35E8.8	3	0.4
Glutathione S-transferase	F11G11.2	1	0.6
Glutathione S-transferase	K08F4.7	4	0.7
Superoxide dismutase	*sod-1* *sod-2* *sod-3*	6 0 4	2.7 1.5 1.1
Glutathione synthetase	M176.2	2	0.6
Catalase	*ctl-1*	9	1.5
SKN-1	*skn-1*	6	1.5

Table 20. Oxidative stress genes contain an ELT-2 consensus binding site within their promoters. The table contains the number of sites present in each promoter. The frequency of the site occurring randomly within the promoter length was also calculated.

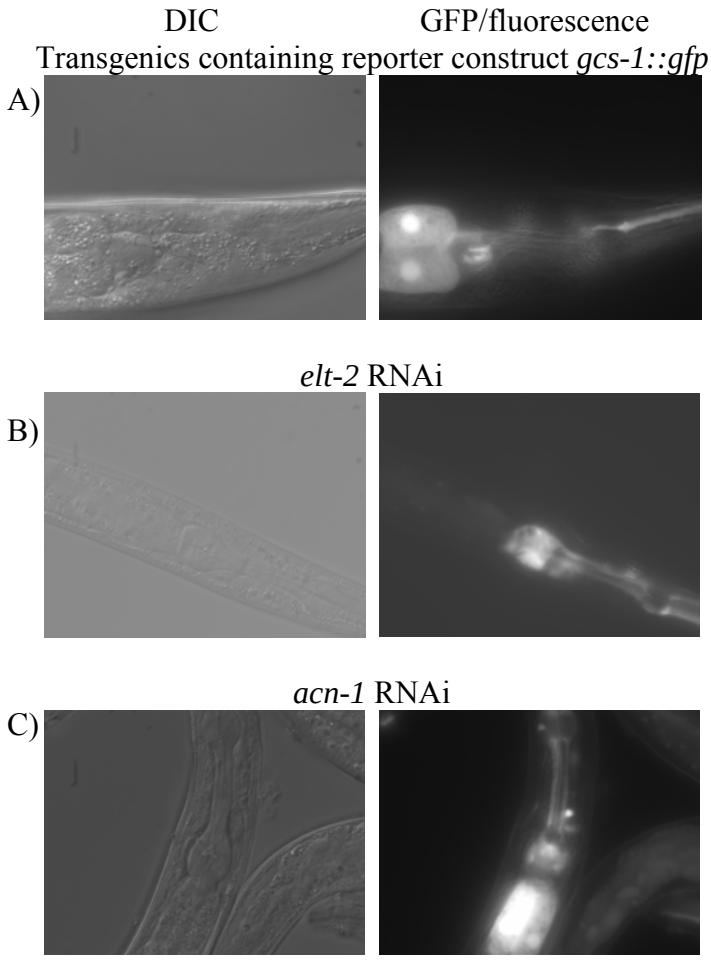
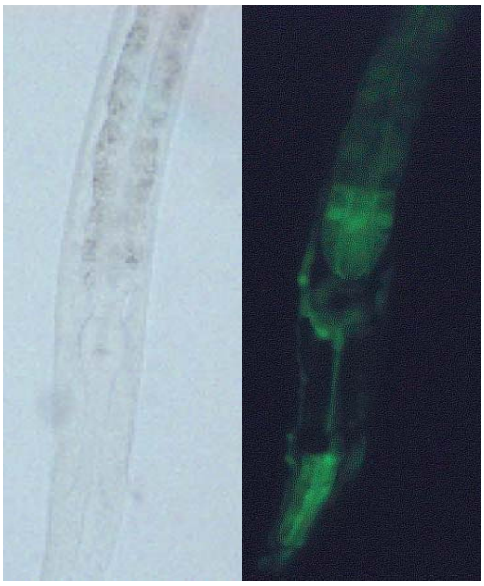


Figure 50. A) Transgenic worms containing the reporter construct *gcs-1::gfp* (the full length promoter) have GFP expression in both the pharynx and intestine. After worms are heat shocked for 20 hrs at 30° Celsius, GFP expression in the pharynx is equal to or greater in strength than GFP expression in the intestine. B) Transgenic adults containing the reporter construct *gcs-1::gfp* were injected with dsRNA corresponding to *elt-2*. Progeny are heat shocked and have a decrease in GFP expression in the intestine compared to GFP expression in the pharynx. C) dsRNA corresponding to *acn-1* was injected into transgenics containing *gcs-1::gfp.* Progeny were heat shocked in parallel to *elt-2* RNAi experiment. GFP is expressed similar to uninjected transgenics in A.

DIC GFP/Fluorescence

Transgenics containing the reporter construct *skn-1*PRO::*gfp*

A)



elt-2 RNAi

B)

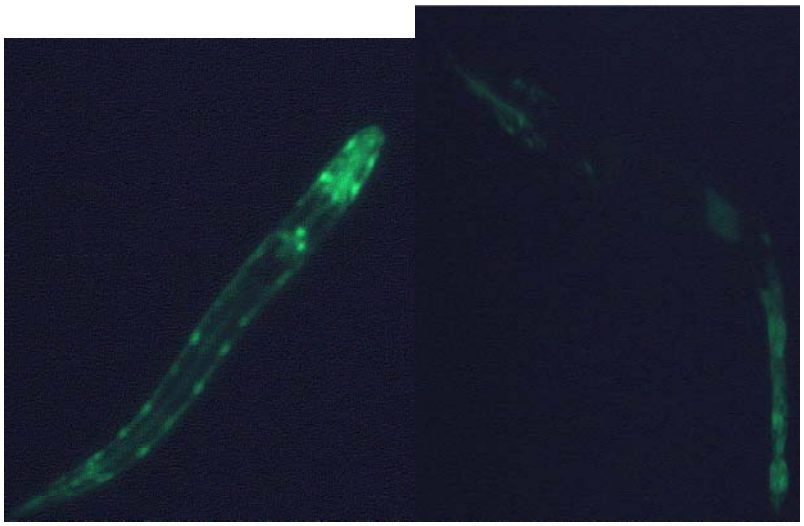
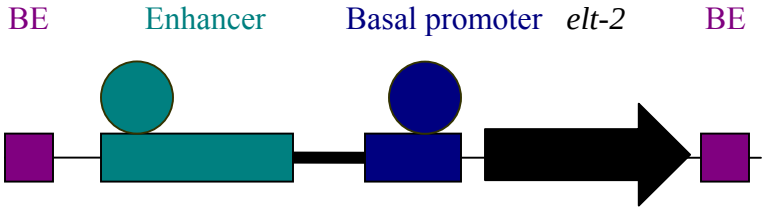


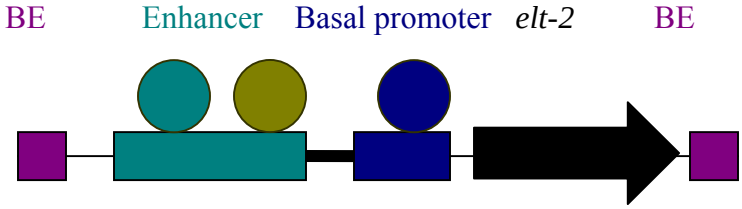
Figure 51. A) Transgenics containing the reporter construct sknPRO::*gfp* (*skn-1* full length promoter fused to *gfp*) have GFP expression in the intestine, hypodermis and pharynx of the worm. B) Progeny from mothers injected with dsRNA corresponding to *elt-2* have variable GFP expression. Left panel shows a worm with expression in the hypodermis and pharynx but not in the intestine. Right panel shows a worm with GFP expression in the intestine.

A) END-1 binds to GATA site -3820

B) Movement



C) Intermediate



D) Maintaining embryonic expression

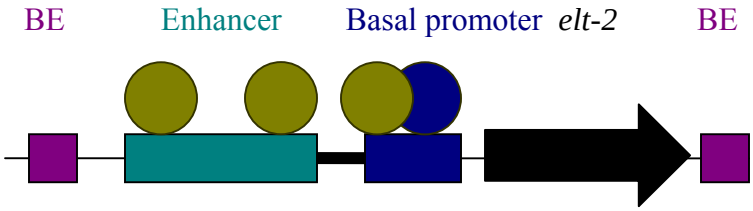


Figure 52. Model for the regulation the *elt-2* gene (black arrow). END-1 (Dark green circle) binds the enhancer (enhancer-Dark green box) and opens chromatin as well as positioning the enhancer close to the basal promoter (Navy box) which facilitates the RNA polymerase unit (Navy circle) to initiate *elt-2* transcription. A) Binding of END-1 to GATA site -3820 in the enhancer. B) END-1 moves enhancer and basal promoter together. C) Before the 8E cell stage intestinal cells contain both END-1 and ELT-2.
D) ELT-2 protein is responsible for maintaining *elt-2* expression after the 8E cell stage. ELT-2 (Olive color circle) may bind to GATA -3394 (enhancer Dark green box) alone or to GATA site -3394 and GATA site -3820. ELT-2 also binds to the basal promoter to facilitate an open chromatin structure.